PhD THESIS

Identification of genes related to seed longevity in *Arabidopsis thaliana* using genomic molecular techniques

PhD student

Joan Renard Meseguer

Program

PhD in Biotechnology

Thesis directors

Eduardo Bueso Ródenas

José Gadea Vacas

Emeritus director

Ramón Serrano Salom

Valencia, March 2021

Abstract

Seed longevity, or period that seeds remain viable, is important for biodiversity conservation, agriculture and economy. In addition, the study of this parameter could ease the knowledge about molecular mechanisms common to all organisms to prevent aging. One of the main strategies of seeds to reduce their aging consists to stop their metabolism, through drying. Other molecular mechanisms to avoid damages are the isolation from the environment with the seed coat, and the production of antioxidants and other molecules to avoid oxidative damage, one of the main seed aging causes. Repair mechanisms relieve part of the accumulated damage. The model plant *Arabidopsis thaliana* provides the opportunity to carry out genomic studies for the research of, in this case, seed longevity to discover determinant genetic factors and molecular mechanisms. This will serve to better understand seed deterioration processes and it will be key to increase seed longevity.

Using natural genotyped varieties of *Arabidopsis thaliana* and a genome-wide association study (GWAS) followed by reverse genetic studies, 14 new genes related to seed longevity have been identified. They are related to embryo protection, oxidative damage control, and seed coat permeability. Seed coat development is determined by transcription factors. Mutant plants in some transcription factors involved in the seed coat development present altered seed longevity. The over-expression of the transcription factors AtHB25 and COG1 resulted in seeds with increased longevity due to an increased lipid polyester deposition. These lipid polyesters barriers are the cuticle, formed by cutin, and the suberin layer. Both participate positively in the embryo protection from the external environment. Genomic studies of both transcription factors have revealed that AtHB25 directly regulates biosynthetic enzymes of suberin and cutin monomers, and COG1 regulates the expression of enzymes related to the polymerization of lipid polyesters and lignin.

The regulation involving AtHB25 is crucial due to the high conservation of genomic sequences and functions of AtHB25 in angiosperms, and it seems to be involved in the response to low temperatures. On the other hand, COG1, which is involved in light perception, regulates part of the development of the external integument through its regulation by AP2, a key factor in establishing the tissue identity of this seed coat integument, where suberin is located. AtHB25 and COG1 are involved in seed longevity adaptation through environmental signals such as temperature and light, respectively, regulating lipid polyesters deposition.

Resumen

La longevidad de las semillas, o el tiempo durante el cual permanecen las semillas viables, es de gran importancia para la conservación de la biodiversidad, la agricultura y la economía. Además, el estudio de este parámetro puede contribuir a conocer mejor los mecanismos moleculares comunes a todos los organismos para prevenir el envejecimiento. Una de las principales estrategias de las semillas para ralentizar su envejecimiento consiste en detener su metabolismo, a través de su deshidratación. Otros mecanismos moleculares para evitar daños son el aislamiento frente al entorno a través de la cubierta de la semilla, y la producción de antioxidantes y otras moléculas para evitar el daño oxidativo, uno de los principales causantes del envejecimiento de las semillas. Los mecanismos de reparación mitigan parte del daño acumulado. El organismo modelo de plantas *Arabidopsis thaliana* brinda la oportunidad de la realización de estudios genómicos para el estudio de, en este caso, la longevidad de las semillas para descubrir nuevos factores genéticos y mecanismos moleculares determinantes. Este conocimiento servirá para entender mejor los procesos de deterioro de las semillas y que también será clave para aumentar la longevidad de estas.

Mediante el uso de variedades naturales genotipadas de *Arabidopsis thaliana* y un estudio de asociación del genoma conocido como GWAS, seguido de estudios de genética reversa, se han identificado 14 nuevos genes relacionados con la longevidad de las semillas, relacionados con la protección del embrión, el control del daño oxidativo, y la permeabilidad de la cubierta de la semilla. El desarrollo de la cubierta de la semilla está determinado por factores de transcripción. Plantas mutantes en diversos factores de transcripción involucrados en el desarrollo de la cubierta de la semilla presentan una longevidad alterada. La sobreexpresión de los factores de transcripción AtHB25 y COG1 provoca que las semillas presenten una mayor longevidad debido a una incrementada deposición de poliésteres lipídicos. Estas barreras de poliésteres lipídicos son la cutícula, formada por cutina, y la suberina. Ambas participan positivamente en la protección del embrión frente al ambiente exterior. Estudios genómicos de ambos factores de transcripción han demostrado que AtHB25 regula directamente a enzimas biosintéticos de los monómeros de suberina y cutina, y COG1 regula la expresión de enzimas relacionados con la polimerización de poliésteres lipídicos y lignina.

La regulación en la que participa AtHB25 es muy importante debido a la alta conservación de las secuencias genómicas y funciones de AtHB25 en angiospermas, y parece involucrado en la respuesta a bajas temperaturas. Por otra parte, COG1, que está involucrado en la percepción de luz, regula parte del desarrollo del tegumento externo a través de la regulación de AP2, un factor clave en el establecimiento de la identidad de tejido de este tegumento de la cubierta de la semilla, donde se localiza la suberina. AtHB25 y COG1 están involucrados en la adaptación de la longevidad de la semilla a través de señales ambientales como la temperatura y la luz, respectivamente, regulando la deposición de poliésteres lipídicos.

Resum

La longevitat de les llavors, o el temps que romanen les llavors viables, es de gran importància per la conservació de la biodiversitat, l'agricultura i l'economia. A més a més, l'estudi d'aquest paràmetre pot contribuir a conèixer millor els mecanismes moleculars comuns a tots els organismes per prevenir l'envelliment. Una de les principals estratègies de les llavors per retardar el seu envelliment consisteix detenir el seu metabolisme, mitjançant la seua deshidratació. Altres mecanismes moleculars per evitar danys son el seu aïllament de l'entorn per mitjan de la coberta de la llavor, i la producció d'antioxidants i altres molècules per evitar el dany oxidatiu, un dels principal causants del envelliment de les llavors. Els mecanismes de reparació mitiguen part del dany acumulat. L'organisme model *Arabidopsis thaliana* brinda la oportunitat de la realització d'estudis genòmics per a l'estudi de, en aquest cas, la longevitat de les llavors per descobrir nous factors genètics y mecanismes moleculars determinants. Aquest coneixement servirà per entendre millor els processos de deteriorament de les llavor i serà clau per augmentar la longevitat d'aquestes.

Mitjançant l'ús de varietats naturals genotipades d'*Arabidopsis thaliana* i un estudi d'associació del genoma conegut com GWAS, seguits d'estudis de genètica inversa, s'han identificat 14 nous gens relacionats amb la longevitat de les llavors, relacionats amb la protecció de l'embrió, el control del dany oxidatiu, i la permeabilitat de la coberta de la llavor. El desenvolupament de la coberta de la llavor està determinada per factors de transcripció. Plantes mutants a diversos factors de transcripció involucrats al desenvolupament de la coberta de la llavor presenten una longevitat alterada. La sobreexpressió dels factors de transcripció AtHB25 i COG1 provoca que les llavors presenten una major longevitat degut a una deposició de polièsters lipídics incrementada. Aquestes barreres de polièsters lipídics son la cutícula, formada per cutina, i la suberina. Ambdues participen positivament la protecció de l'embrió enfront de l'entorn exterior. Estudis genòmics d'ambdós factors de transcripció han demostrat que AtHB25 directament regula a enzims biosintètics dels monòmers de suberina i cutina, i COG1 regula enzims relacionats amb la polimerització de polièsters lipídics i lignines.

La regulació en la que participa AtHB25 es molt important degut a l'alta conservació de les seqüències genòmiques i funcions de AtHB25 en angiospermes, i parteix estar involucrat en la resposta a baixes temperatures. Per altra banda, COG1, que està involucrat en la percepció de la llum, regula part del desenvolupament del tegument extern mitjançant la regulació de AP2, un factor clau en l'establiment de la identitat de teixit de aquest tegument de la coberta de la llavor, on es localitza la suberina. AtHB25 i COG1 estan involucrats en l'adaptació de la longevitat de la llavor per mitjan de senyals ambientals com la temperatura i la llum, respectivament, regulant la deposició de polièsters lipídics.

Abbreviations

2-MAG - 2-monoacylglycerol

3-AT - 3-Amino-1,2,4-triazole

A_{660} - absorbance at 660 nm

AAT - accelerated aging treatment

ABA - abscisic acid

ABSL - acetyl bromide soluble lignin

AHCs - alkyl hydroxycinnamates

APS - ammonium persulfate

BER - base excision repair

bp - base pair

BSA - bovine serum albumin

cDNA - RNA-complementary DNA

CDS - Coding sequence

CDT - controlled deterioration treatment

ChIP - Chromatin Immunoprecipitation

ChIP-seq - Chromatin Immunoprecipitation followed by new generation sequencing

CoA - Coenzyme A

CTAB - hexadecyltrimethylammonium bromide

CYP - cytochrome P450

DAP - Days-after-pollination

DC - direct current

DCAs - α,ω-dicarboxylic fatty acids

DEGs - differentially expressed genes

DHARs - dehydroascorbate reductases

DMSO - Dimethyl sulfoxide

DW - dry weight

EDTA - Ethylenediaminetetraacetic acid (disodium salt dihydrate)

EPPO - elevated partial pressure of oxygen treatment

ER - endoplasmic reticulum

FA - fatty acids

FAE - fatty acid elongase

FAR - fatty acyl reductases

G3P - glycerol 3-phosphate

GA - gibberellins

GC-MS - Gas chromatography-mass spectrometry

GO - Gene Ontology

GPAT - glycerol-3-phosphate acyl-transferases

GWAS - Genome-Wide Association Study

HCAs - hydroxycinnamic acids

HSPs - heat-sock proteins

IAA - auxins

JA - jasmonic acid

Kb - kilobase

KCS - β-Ketoacyl-Coenzyme A synthases enzymes

KO - knock-out

LACS - Long-chain acyl-CoA synthetases

LB - Luria-Bertanil

LEA - late embryogenesis abundant proteins

LiTE - Litium acetate - Tris-HCl-EDTA

LPA - lysophosphatidic acid

MES - 2-(N-morpholino)ethanesulfonic acid

MS - Murashige and Skoog

NADPH - nicotinamide adenine dinucleotide phosphate

NAT - natural aging treatment

NGS - new generation sequencing

OHFA - ω-hydroxy fatty acids

PA - primary alcohols (fatty alcohols)

PAs - proanthocyanidins

PCD - program cell death

PCR - Polymerase chain reaction

PIPES - piperazine-N,N′-bis

PTM - Photomultiplier Tube

PVDF - polyvinylidene difluoride

RFOs - raffinose family oligosaccharides

RH - Relative humidity

RNA-seq - Whole transcriptome analysis by new generation sequencing

ROS - Reactive oxygen species

rpm - revolutions per minute

SD - synthetic media - Dextrose (Glucose)

SDS - Sodium dodecyl sulfate

SEM - Scanning Electron Microscopy

SNPs - single-nucleotide polymorphisms

ssDNA 1% - salmon sperm DNA 1%

SSPs - Seed Storage Proteins

TB - transformation buffer

TBE - Tris Base-Boric acid-EDTA

TEM - Transmission Electron Microscopy

TEMED - Tetramethyl ethylenediamine

TF - transcription factor

Tris base - tris (hydroxymethyl) aminomethane

TSS - transcription start site

TTS - transcription termination site

VLCFA - very long chain fatty acids

Y1H - yeast-one-hybrid

Y2H - yeast-two-hybrid

YNB - yeast nitrogen buffer

YPD - Yest extract - Peptone - Dextrose

YPDA - Yest extract - Peptone - Dextrose - Adenine

ZF-HD - Zinc Finger-Homeodomain

Table of contents

1. Introduction and objectives

Aging in biological systems

Life is glorious, but not eternal. All organisms, from unicellular to more complex, as plants and animals, age to death. The aging term refers to all changes produced in organisms during their life span. This term includes all deleterious aging events that drive organism to death in absence of other external factors such as disease, stress, or physical damage. This process is named senescence (Dollemore, 2002; Kirkwood, 2005).

Biological immortality, defined as the absence of aging, does not exist. All cells and organisms age. However, aging velocity differs significantly between different organism species and even some organisms are improperly considered as immortal. Commonly, unicellular organisms are considered eternal organisms, as they divide by mitosis and resulting cells can be considered as identical. However, studies in *Escherichia coli,* a symmetrical-dividing bacterium, have reported that cell divisions are not equal and aging damage accumulates mainly in the old cell, finally affecting to their reproductive efficiency (Stewart *et al.*, 2006). This fact demonstrates that aging also occurs in bacteria. On the other hand, cell aging damage can be reduced in multicellular organisms, as they can renewal cells and regenerate tissues. In this way, tissue regeneration can prolong individual organism live. In some organisms, as *Hydra vulgaris*, apparently this prolongation is indefinite (Martinez, 1998). Other animal strategy to prolong life indefinitely consists into trans-differentiate form adult to juveniles in solitary stage observed in some jellyfishes, as *Turritopsis dohrnii* (Piraino *et al.*, 1996). In contrast to animals, plants, which have a more versatile tissue regeneration, present postembryonic development and develop new tissues and organs during most of their live (Perez-Garcia and Moreno-Risueno, 2018). Some plant species can apparently grow indefinitely. This results in millenarian plant individuals, as estimated in some bristlecone pines (Schulman, 1958; Currey, 1965).

The biological reason of aging is still under discussion (da Costa *et al.*, 2016). Program theories suggest that individuals age due to an evolutionary benefit: life decays to avoid competition with offspring. In this way, life would be limited in a genetically programmed way. This effect has been observed in bacteria, yeast, plants and animals (Long *et al.*, 2005; Allocati *et al.*, 2015: Woo *et al.*, 2018). Genes affecting aging have been described and common mechanisms are found among different organisms (Johnson, 1990; Mattson, 2003). Programmed aging could be considered as development growth continuation (Blagosklonny, 2013). On the other hand, damage theories remark the absence of evolutionary selection after reproduction. Thereby, aging is

caused due to the accumulation of imperfections in the metabolism and the failure on maintenance systems. The oxidative damage is thought to be the major damaging reaction (Harman, 1956). Reactive oxygen species (ROS) causes damages in biological macromolecules such as DNA, proteins, and lipids (Dizdaroglu, 1992; Shah *et al.*, 2001; Mishra *et al.*, 2011). These damages directly deteriorate organisms. The three ROS main sources in living cells are the electron transport chains, the nicotinamide adenine dinucleotide phosphate oxidases (NADPH oxidases) and the 5-lypoxigenase (Novo and Parola, 2008). According to the damage theory, reducing the metabolism and ROS damage may prolong organism life.

Aging and the different ways to cope aging are of special interest in research as a selfish attempt to prolong human life eternally. Thereby, aging research has been focused on animal model organisms such as *Caenorhabditis elegans*, *Drosophila melanogaster* and *Mus musculus*. These studies provide genetic evidence for both, damage and programmed theories (Johnson, 1990; Orr and Sohal, 1994; Lithgow *et al.*, 1995), but same strategies, as overexpression of antioxidant enzymes, do not extend life in all organisms (Pérez *et al.*, 2009). Nevertheless, both theories point to metabolism as the main aging cause. Aging is the cost of living and metabolism reduction is an essential strategy to enhance longevity.

Physical factors affecting longevity

The exposure of organisms to different physical factors alters metabolism rates and causes direct physical damage. For example, internal organism temperature correlates with the metabolism rate. Low temperatures generally prolong organism life-span. In many poikilothermic organisms, moderate environmental changes in temperatures have important effects in their longevity (Conti, 2008). Thermodynamically, low temperature slows diffusion and thus, metabolism rate. Thereby, cell cycle division rates, development and thus, aging, get slow down at lower temperatures (Keil *et al.*, 2015). In the model organisms *D. melanogaster* and *C. elegans* it is well established that their life span inversely correlates with temperatures (Miquel *et al.*, 1976; Leiser *et al.*, 2011). This temperature effect has been also observed in mammals, although they barely present internal temperature changes. However, mammalian hibernators can drop their body temperature up to a 10%, without hypothermic damage (Carey *et al.*, 2003; Storey, 2010). This low temperature in addition to the prolonged fasting during hibernation significantly increases their life-span through a reduction of their metabolic rate, which leads to an important reduction in protein synthesis (Wu and Storey, 2016). In non-hibernator mammals, this temperature effect has been studied through caloric restriction. Limiting calorie intake without causing malnutrition reduces slightly but significantly the organism temperature in mammals (Duffy *et al.*, 1989; Col-

man *et al.*, 2009; Soare *et al.*, 2011). Caloric restriction is associated with metabolism reduction. It decreases ROS production through the reduction of the mitochondrial electron-chain reactions (Pamplona and Barja, 2006). Thereby, caloric restriction extend life in numerous organisms (Fontana and Partridge, 2015). It is one of the most promising approaches to extend human life.

However, aging is not the only effect of temperature: very low temperatures are associated with severe enzyme inhibition preventing organism development, while extremely high temperatures produces protein denaturation and membrane dysfunction due to an increased fluidity (Pörtner, 2002). If organism temperature decays below zero degrees, water freezes, molecules stop diffusing, and thereby metabolism stops. During the freezing process, cell cytoplasm gets more viscous and solutes concentrate progressively damaging cells. This process is harmful to cells even though there are some organisms that can tolerate it. Cryo-conservation has been widely used as a process to preserve cells alive for long terms as aging events stops with metabolism. Optimal freezing and unfreezing kinetics and cryo-potector molecules are needed to rapidly vitrify water avoiding water-crystal formation damage in cellular components and tissue structures (Glenister and Thornton, 2000). There are well-established protocols to cryoconserve microorganisms, such as bacteria and yeast, single cells like human stem cells and gametes, and even some tissues (Baust *et al.*, 2009). However, it does not work in complex organisms, as damage is irreversible. In nature, continuously frozen soils, also called permafrost, preserve ancient life forms. Permafrost sediments are a repository of ancient organisms and ancient genetic resources. For example, the maternal placental tissue of the plant *Silene stenophylla* recovered from permafrost sediments dated between 30,000 and 32,000 years was able to regenerate a viable plant (Yashina *et al.*, 2012), demonstrating the value of permafrost in the preservation of ancient live forms, even for multicellular organisms.

Metabolism is also greatly reduced upon desiccation. Cellular media gets high viscous and molecules diffuse very slow (Wood and Jenks, 2007). As in freezing processes, during desiccation, water availability decays and solute concentration increases. However, water depletion is harmful and can be lethal. Intracellular water maintains the structure of molecules and membranes. Upon desiccation macromolecules aggregate and organelles disintegrates (Alpert, 2006). Life emerged and evolved initially in water. Land areas presents lower and variable water availability, including dry periods. Many life forms developed desiccation tolerance, defined as the ability to dry to equilibrium with the ambient atmosphere and then recover to its normal moisture content (Proctor and Pence, 2002). This tolerance permits tolerant organisms to survive in absence of water. In addition, desiccation slows down and stops metabolism reactions, permitting desiccation tolerant organisms to endure for long time with reduced aging events. Several desiccation-tolerant-species can survive from years to decades (Alpert

and Oliver, 2002; Guidetti and Jonsson, 2002). Desiccation tolerance of organisms was crucial for land colonization as it allows them to survive until conditions are favorable. For example, cyanobacteria, one of the most ancient organisms occupy extreme niches that are inaccessible to other organisms as they can survive in adverse physiological growth conditions due to their great adaptive abilities like desiccation tolerance (Singh, 2018). Insects are also spread in a wide range of ecologically diverse niches due to their variety of strategies to avoid desiccation damage (Chown and Nicolson, 2004). Plants had to acquire desiccation tolerance necessary for land colonization (Oliver *et al.*, 2000), but latter this tolerance was lost. Desiccation tolerant plants, called resurrection plants, maintained this ability. Nevertheless, several plants develop desiccation tolerant seeds, also called orthodox seeds. Similar genetic pathways control plant desiccation tolerance and seed maturation of orthodox seeds (Costa *et al.*, 2017). Contrarily to desiccation sensitive seeds (also called recalcitrant seeds), orthodox seeds can dry until low moisture contents (5%) and even can tolerate freezing as low water content avoids crystal formation and cellular damage. As other desiccation tolerance organisms, orthodox seeds are small (Singh, 2018) and can survive longer than recalcitrant seeds. The high-water content in recalcitrant seeds promotes microbial contamination and a faster seed deterioration reducing their life span to months (Berjak and Pammenter, 2008). On the other hand, orthodox-seed metabolism is highly decreased, and they can last for years. Seed longevity, defined as period that seeds remain viable, varies enormously depending the plant species (Walter *et al.*, 2005) and in some species they can last for hundreds or even more than thousands of years, as described in *Nelumbo nucifera* and *Phoenix dactylifera* seeds (Shen-Miller *et al.*, 1995; Sallon *et al.*, 2008).

Avoiding aging goes through avoiding oxygen damage. Oxygen is highly reactive and as discussed above, is the main aging cause through ROS-induced damages in biomolecules. Nowadays, atmosphere contains about 21% of oxygen. However, initially the atmosphere and oceans were largely or entirely anoxic (Holland, 2006). Photosynthesis emergence is responsible of this change (Pierson, 1994), and it shifted out anaerobic life from the earth surface, as the oxygen molecule resulted toxic for it. In response, oxygen-tolerant life emerged. One important advantage of oxygen is that allows respiration in eukaryotic life forms, which is highly energy-efficient compared with fermentation. On the other side, oxygen is highly reactive causing organism damage and provoking aging, as discussed above. This controversial effect of oxygen, essential but damaging, conforms the Oxygen Paradox (Ursini and Davies, 1995). The oxygen damaging effect is widely known, and for example, oxygen-lacking modified atmospheres are generally used in food industry to preserve aliments. This practice avoids oxidation events such as lipid peroxidation. In addition, it prevents aerobic microorganism growth and thus, food deterioration (Parry, 1993). In living organisms, aging rates also increase with the oxygen concentration in the atmosphere. *C. elegans*

life span decreases in a high-oxygen atmosphere and augments in low-oxygen atmosphere, while aging goes oppositely (Honda *et al.,* 1993).

Other physical factor that accelerates aging is the radiation. The harmful effect of UV light is widely known. Highly energetic UV light directly damages DNA, creating genetic mutations, while low energetic UV light influences ROS production (Shen and Tower, 2019). However, UV light in moderate doses results in beneficial effects. Not only UV lights affects aging, visible light has been also described to reduce life-span in *C. elegans* and *D. melanogaster* (De Magalhaes Filho *et al.*, 2018; Shen *et al.*, 2019). Visible light controls circadian rhythms, and thus metabolism. In addition, it creates photooxidative stress along with protein unfolding not related with circadian clock (De Magalhaes Filho *et al.*, 2018). Nevertheless, light effects on life are not always deleterious: near far-red light provokes beneficial effects and delay aging in some animal models (Shen and Tower, 2019). Moreover, phototrophs organisms, as plants, need light for their development, and they develop photoprotective mechanisms to reduce collateral light-induced damages (Pinnola and Bassi, 2018). Life depends on light as the main source of energy through carbon fixation in photosynthesis.

Natural radioactivity, coming from radionuclides in the earth and cosmic radiation, is an often-forgotten physical factor that affects aging. There are not many studies that describe radioactive effect in longevity, but they are emerging (Al Attar *et al.*, 2016; Nyhan *et al.*, 2018). Radioactivity has increased lastly through human activity and resource exploitation. In addition, other physical factors have been altered by human activity affecting organisms aging in different ways. One example is global warming. In addition, the emergence of widespread plastics and chemicals, water and air contamination, and natural ecosystem modifications among others, affects different life processes and it could increase aging events (Li *et al.*, 2018; Burraco *et al.*, 2020).

Seed longevity

Dry seeds are an interesting source for longevity research as their aging events occur in low water-content conditions. Essential molecular mechanisms are conserved among eukaryotes. Thus, seed longevity research can lead to the discovery of general conserved-metabolic pathways that can provide general aging knowledge and new insights for aging theories. As discussed above, seed longevity varies drastically among different plant species, form months to thousands of years. Genetics and molecular pathways determining seed development drive these seed longevity differences. Moreover, seed storage conditions affect seed deterioration. Dry seeds can resist a wide range of conditions such as low and high temperatures and desiccation (Sano *et al.*, 2016). In this way, seeds not only are part of the reproductive mechanism of most

plants, but they also present an important adaptive strategy. Seeds permit to interrupt plant life cycle in adverse environmental conditions as seeds maintain life in a quiescent state until favorable conditions (Bewley, 1997; Bentsink and Koornneef, 2008). This aspect has been crucial for land plant adaptation, as it permits plants species to survive in each area even when plants cannot survive concrete environmental conditions. Seeds are time-vectors of plant life. The evolutionary importance of seeds is remarkable, as seed-producing plants constitute a major part of plant species in earth. However, as it occurs in organisms, the exposure to damaging physical factors such as high temperatures, high humidity, high oxygen pressure, UV light and radioactivity increases aging events and decreases their longevity. Upon desiccation, embryo cytoplasm reduces its fluidity into a glass-viscosity state with a low moisture content (less than 20%). The mobility of cellular components gets limited, and thus metabolic rate gets reduced (Buitink and Leprince, 2008). Thereby, as metabolism aging-events are slowed down, desiccation is crucial for seed longevity acquisition. High humidity and high temperature are deleterious for seed storage (Figure 1) as they increase fluidity of embryo cytoplasm reactivating aging events, such as ROS production (Walters, 1998). Oxygen, light, and radioactivity directly produces damages in cellular structures and ROS formation (Harrington, 1970; Groot *et al.*, 2012; Al Attar *et al.*, 2016).

Seeds present a valuable resource for human alimentation. In several cases they are a direct food resource, and in many other cases seed are crucial for plant propagation. Highly germinating seeds, also called vigorous seeds, are necessary for agronomic propagation. Obtaining vigorous seeds and preserving them properly is crucial for agriculture and the associated economy. First symptoms of aged seeds are delayed and poor seedling establishment. Finally, seed germination ability is lost. Seed aging and germination loss do not occur simultaneously in all seeds from same seed lot (Figure 1). In each seed batch, seed aging follows a probabilistic distribution due to the random and progressive accumulation of aging events in the different seeds of the batch. Germination ratios are used to test how seed lots have aged, and it is useful to determine maximum crop germination to adjust crop yield. Other tests, as seed vigor tests or artificial aging tests, are established for numerous species. Damaging physical factors such as humidity and temperature are used to establish vigor and thus predict aging behavior on different seed lots (ISTA, 2018).

Agronomically, seeds are also important for plant biodiversity conservation. Plant biodiversity is the genetic source for plant breeding. *In situ* biodiversity conservation represents the best method to preserve life as it develops in its natural ecosystems. However, extensive agriculture and human environmental impact often damage and destroy different ecosystems, leading to the loss of biodiversity. Other alternative for biodiversity conservation is *ex situ* conservation. For *ex situ* conservation seeds represent the best way to preserve plant ancestral cultivars and other wild plants species,

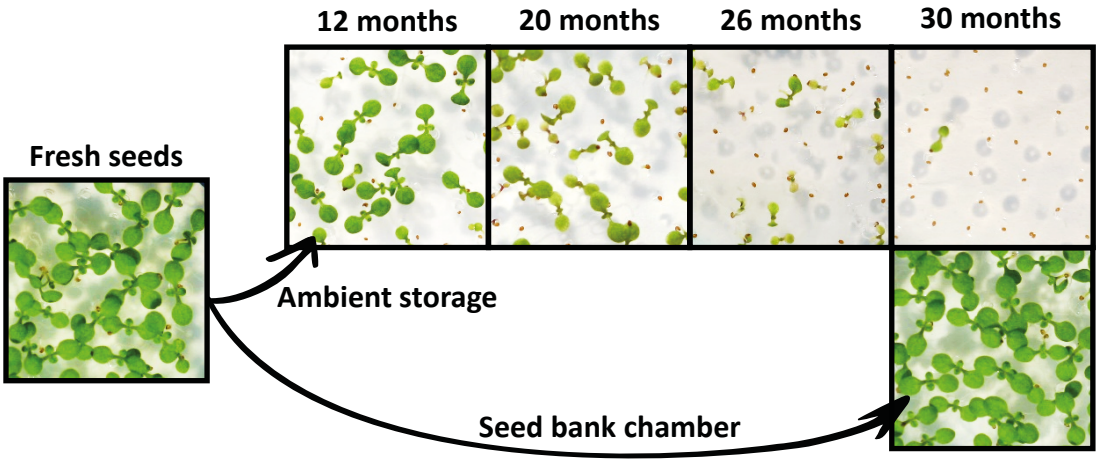


Figure 1: The germination rate decays on time but it can be prolonged with optimal seed storage conditions. Seed aging is progressive, and it can be measured through germination ratio. Small seedlings indicate a delay of germination, present in aged seeds. Low temperature and low humidity during seed storage allow seeds to endure longer time comparing with ambient seed storage. The same lot of *Arabidopsis thaliana* (Col-0) seeds was stored at room temperature (20 - 25 ºC) and ambient humidity (40% - 80%) for 12, 20, 26 and 30 months (up), and in a seed storage chamber at 10 ºC and 16% relative humidity for 30 months (below). Stratified seeds were shown and grown in MS media for seven days.

as seeds require a reduced space for storage, reduced maintenance and, depending on the species, they could last for years. A dry and cold ambient permits to maintain seeds for years and even decades in seed banks (Vertucci and Roos, 1990, Figure 1). One of the most famous seed banks is in Norway. Svalbard Global Seed Vault is the largest seed bank in earth with more than a million different seed lots, maintained by different national and international institutions (Westengen *et al.*, 2013; Asdal and Guarino, 2018). Its location allows maintaining low temperatures in a highly efficient way, even if the artificial cooling fails. It aims to ensure food safety through seed biodiversity conservation of agronomic cultivars in case of global catastrophes. However, even if seeds are stored at best conditions in seed banks, eventually seed-intrinsic senescence events provoke seed aging and their germination ability decay. Plant growing for seed regeneration is necessary every few years, depending on the plant species, to ensure their bioconservation. In this way, other seed banks distributed around the world are also necessary for food safety. They take part on seed conservation, regeneration, and distribution (Westengen *et al.*, 2013).

The Arabidopsis seed - Structure and development

Seeds (Figure 2a) are a complex specialized plant organ, with the function to maintain the embryo alive and protected from detrimental environments until conditions are favorable. The new individual develops from a single cell into an embryo that should remain alive during seed storage to generate a new plant. Seed is not just the embryo itself. In Angiosperms, two other different-origin tissues form the seed: the endosperm and the surrounding maternal tissues (Sreenivasulu and Wobus, 2013) (Figure 2b). These different seed compartments conforming the seed present different genomes. Much of the seed complexity aims to prevent aging and deleterious events in seed embryos. Seed development differs among different plant species and following remarks are based in the model plant *Arabidopsis thaliana*, which is widely studied in plant molecular biology.

After egg cell fertilization from, the diploid zygote cell starts dividing with a concrete developmental pattern program and resulting cells undergo progressively differentiation forming the different tissues. There are four described embryo developmental stages: globular, heart, mid-torpedo, and bent cotyledon (Mayer *et al.*, 1991). This last stage can be already visualized at 7 days after pollination (DAP). Later, from 8 to 16 DAP, major storage reserves, as sugars, oils, storage proteins and RNA pools, accumulates and embryo fills the seed sac (Focks and Benning, 1998). This period is called seed filling and these reserves are crucial for seedling establishment and seed vigor acquisition, and therefore seed longevity. During seed filling, embryo produces chlorophyll (Puthur *et al.*, 2013). The photosynthesis activity in embryo permits the synthesis of the reserve biomolecules, and it is necessary for seed vigor and seed longevity (Allorent *et al.*, 2015). In addition, the photosynthesis activity of silique walls also participates in the seed filling process (Zhu *et al.*, 2018). Finally, the embryo development stops, and the maturation phase starts (up to 21 DAP). Seed maturation consists of the acquisition of desiccation tolerance, the degradation of chlorophylls and the transformation of chloroplast in storage plastids. In angiosperms, embryos of different plant species do no degrade the chlorophylls and remain green (chloroembryophytes) while others, like Arabidopsis, do and become white (leucoembryophytes). In the maturation period, dormancy and seed longevity are acquired (Goldberg *et al.*, 1994; Smolikova, *et al.*, 2017). Seed dormancy avoids premature germination of seeds and depends on the plant species and on its environment adaptation. Dormancy is controlled by environmental factors such as light, temperature and time of seed dry storage (Bentsink and Koornneef, 2008). It is inversely correlated with seed longevity in diverse Arabidopsis natural accessions or ecotypes and mutants (Nguyen *et al.*, 2012).

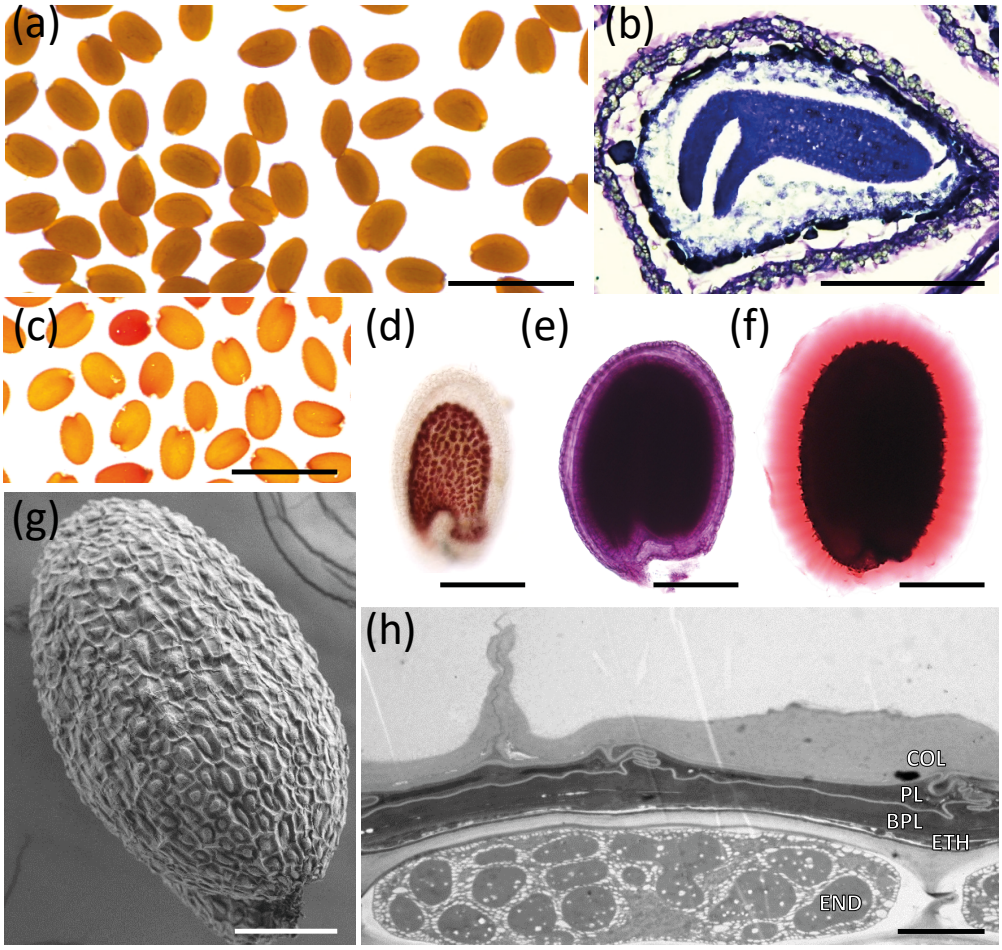


Figure 2: The Arabidopsis seed. Col-0 wild-type seed analysis. (a) Dry seeds. Scale bar, 1 mm. (b) Toluidine Blue staining of a developing seed section between mid-torpedo and bent cotyledon. Notice that endosperm degrades around the embryo, endothelium already has proanthocyanidins (dark blue), and outer integument cells, separated from inner integument by a thick cell wall, accumulate amyloplasts (grey) and mucilage (purple). Scale bar, 200 µm. (c) Tetrazolium salt reduction assay (for 48h at 28 ºC) test seed coat permeability and/or viability. Notice that same seed-batch present slightly differences in tetrazolium reduction rates, visualized through a reddish seed staining. Scale bar, 1 mm. (d): Developing seed at day-after-pollination (DAP) 5 stained with vanillin to visualize proanthocyanidins (PAs) in the endothelium. Scale bar, 200 µm. (e) Lipid polyester dry seed staining with Sudan Red. Notice that two lipidic layers can be observed: the suberin layer close to the seed coat epidermis; and the cutin layer, close to the embryo. Scale bar, 200 µm. (f) Mucilage halo visualization in 30 minutes imbedded dry seed with Ruthenium Red. Scale bar, 200 µm. (g) Scanning Electron Microscopy (SEM) of the seed surface. Notice the characteristic seed surface formed by columella structures. Scale bar, 100 µm. (h) Transmission Electron Microscopy (TEM) of a dry seed section. In the image, the endosperm (END, below) and the seed coat (above) are shown, but not the embryo. Notice that endosperm cell has not collapsed as seed coat cells, the endothelium (ETH) has breaks, below the endosperm there is the cutin layer (below the last dark layer). The brown pigment layer (BPL) supposes big part of seed coat thick. The palisade layer (PL) is composed by smaller cells with a lighter electron-transmission cell wall and they collapse and fold with the adjacent palisade cell forming a sealed structure, and the columella cells (COL) present a characteristic volcano shape. Finally, is remarkable that the darker areas (BPL) are within the two impermeable barriers, the cutin and the suberin, and thus PA liberation stains only inner integument layers. Scale bar, 4 µm.

Fertilization in Arabidopsis, as in other angiosperm plants, is in fact, a double fertilization. The two sperm cells present in the pollen grain fertilize the egg cell to develop the new individual, and the central cell (a diploid cell) leading to the endosperm, a triploid tissue that envelops the seed embryo (Dresselhaus *et al.*, 2016). The endosperm is a nourishing tissue as it participates in the nutrient transfer from the mother plant into the developing embryo (Bleckmann *et al.*, 2014). Endosperm development and cell proliferation are different from any other plant tissue. First, successions of nuclear divisions without cytokinesis take place (Li and Berger, 2012). After a certain nuclear division number, endosperm cellularizes (Boisnard-Lorig *et al.*, 2001). Finally, endosperm cells progressively undergo program cell death (PCD) while embryo absorbs them for its growth (Sreenivasulu and Wobus, 2013) (Figure 2b). Only one cell layer surrounds the embryo in the mature seed. Mature endosperm imposes a mechanical resistance to radicle extrusion during germination; this resistance acts as a dormancy barrier. Upon-germination the endosperm cell-wall softs in a genetically programed way (Endo *et al.*, 2012).

The outermost seed-tissue surrounding embryo and endosperm is called seed coat in dicot plants, as Arabidopsis. Seed coat, with a protective role, origins from dipliod ovule-integuments of the mother plant, so it is not a direct product of the polen grain fertilization. Upon the fertilization hormonal signal, they start to differentiate coordinately with the embryo (Bencivenga *et al.*, 2011; Figueiredo *et al.*, 2016) following a complex and specialized genetic program. Five integument-cell layers form the seed coat: three originate from the inner integument and two from the outer integument. Each cell layer undergoes a concrete developmental program to get specialized and synthesize their specific associated biomolecules (Haughn and Chaudhury, 2005). The endothelium, the innermost seed coat layer, synthesize a cutin-based layer: the cuticle. Cutin is a lipid polyester barrier formed by very long chain fatty acids. Cutin monomers are deposited at outer side of the cell wall. There, cutin monomers crosslink forming the impermeable cuticle layer bounded into the cell wall (Figure 2e). The cuticle layer also presents polysaccharides and associated waxes. This layer is also present in other plant organs such as leaves (Yeats and Rose, 2013). The cuticle avoids organ fusion with the endosperm (Voisin *et al.*, 2009) and allows seed coat and endosperm slip during embryo growth (Coen *et al.*, 2019). After development, cuticle is involved in seed longevity and dormancy because it prevents water embryo intake (De Giogi *et al.*, 2015). Moreover, cutin functions include mechanical reinforcement and UV light protection. In addition to cutin, endothelial cells also produce condensed tannins or proanthocyanidins (PAs) which derive from the flavonoid biosynthetic pathway. They provide protection against fungal and insects (Dixon *et al.*, 2005), protection against oxidative stress through oxygen scavenging as they act as antioxidants (Bagchi *et al.*, 1997) and prevent light induced damage (Li *et al.*, 1993). Colorless PAs are accumulated in the immature seed during seed development (Figure 2d) and turn brown

during seed desiccation and seed storage due to their oxidation (Debeaujon *et al.*, 2001), providing the seed characteristic brown color. Endothelium breaks during development by endosperm growth, releasing synthetized PAs (Ondzighi *et al.*, 2008). In mature seeds, PAs affect positively to dormancy probably through hormone regulation (Debeaujon *et al.*, 2000; Jia *et al.*, 2012). From the endothelium grows the subepidermal layer, which may fill empty spaced between integuments (Coen *et al.*, 2017). The outer layer of the inner integument together with the subepidermal layer undergo PCD and crush forming the brown pigment layer seen by transmission electron microscopy (Figure 2h, Beeckman *et al.*, 2000), which seems important for seed thickness (Nakaune *et al.*, 2005). The outer integument is composed by two layers different in shape and function; however, both initially accumulate starch granules in amyloplasts that later are degraded (Windsor *et al.*, 2000). The inner layer of the outer integument present smaller cell areas and increased cell number compared to the outermost or epidermal seed-coat cell layer (Truernit and Haseloff, 2008). Thus, we call palisade layer to this layer. The palisade layer cells present a cell wall thickening produced by suberin deposition (Beeckman *et al.*, 2000: Molina *et al.*, 2008; Gou *et al.*, 2017). Aliphatic monomers of suberin probably are produced from starch degradation, as they are synthesized in plastids and mutants in starch degradation present deficient suberin synthesis (Vishwanath *et al.*, 2015; unpublished data). However, it is also possible that starch degradation is also involved in the cell wall reinforcement of outer integument cells (Windsor *et al.*, 2000). Like cutin, suberin is a lipid polyester barrier, but deposited between the cell wall and the plasma membrane. It presents a characteristic lamellae disposition pattern visualized under electron microscopy, probably caused by aliphatic-aromatic domains (Bernards, 1998; Gadini *et al.*, 2006). Suberin produces impermeable barriers from water and air (Figure 2c,e). When palisade cells collapse, lateral cell walls fold by the compression creating the impermeable structure (Figure 2h). Gou *et al.* (2017) claimed to visualize the suberin lamellae, but the lamellae pattern does not fit with the suberin model. Suberin function in dry seeds is directly linked with seed permeability (Beisson *et al.*, 2007; Yadav *et al.*, 2014), and thus, dormancy (Fedi *et al.*, 2017) and seed longevity (Bueso *et al.*, 2016). Finally, seed coat epidermis cells present a unique structure called the columella, a volcano-shaped structure (Figure 2g). These cells present reinforced radial and columella, but not outer, cell walls and produce large quantities of a pectic polysaccharide, called the mucilage, from accumulated starch granules (Windsor *et al.*, 2000; Western *et al.*, 2000). Mucilage is highly hydrophilic, and, upon water seed-imbibition, it expands drastically breaking the outer cell wall and forms a big halo around the seed (Figure 2f). The mucilage halo results an important evolutionary trait, as it permits water retention for germination avoiding drought stress (Yang *et al.*, 2011). In addition, it might help seed dispersion due to its sticky nature (Western, 2011). During seed maturation, all integuments progressively end up with a cell death program and collapse leading to a compacted and reinforced cell tissue with all its associated biomolecules (Figure 2h).

Embryo molecular mechanisms promoting seed longevity

Seeds possess molecular mechanisms to prevent embryo aging that act during seed development, seed storage and seed germination. Embryo developmental mechanisms allow acquisition of seed longevity during seed maturation through seed desiccation tolerance. Seed storage mechanisms avoid ROS-induced aging events and during germination molecular mechanisms repair embryo damages (Figure 3). The seed coat isolates the embryo form external damaging physical factors.

Maturation and seed desiccation tolerance

Several Arabidopsis mutants affected in seed longevity are developmental mutants (Nguyen *et al.*, 2012). Embryo development and maturation are essential for acquisition of seed longevity. Hormones, such as abscisic acid (ABA), gibberellins (GAs) and auxins (IAA) play essentials roles in seed longevity establishment. ABA is responsible of seed desiccation tolerance, seed longevity acquisition and seed dormancy. The ABA-insensitive *abi3* mutant seed does not maturate properly: seeds do not degrade chlorophyll, present a poor desiccation tolerance and seed longevity and show no dormancy (Ooms *et al.*, 1993). ABI3 is the master regulator of seed maturation (Sano *et al.*, 2016). The transcription factor (TF) ABI3 controls important key aspects for the embryo seed longevity such as chlorophyll degradation (Nakajima *et al.*, 2012; Delmas *et al.*, 2013) and desiccation tolerance through aquaporins, raffinose family oligosaccharides (RFOs), heat-sock proteins (HSPs) and late embryogenesis abundant proteins (LEA) (Kotak, *et al.*, 2007; Mönke *et al.*, 2012; Mao and Sun, 2015, Sengupta *et al.*, 2015; Leprince *et al.*, 2017). Desiccation tolerance is one of the key aspects of the high longevity of orthodox seeds, as dehydrated life forms endure longer. However, desiccation is harmful for biological molecular structures and components. The embryo of orthodox seeds can tolerate desiccation thanks to their glassy-state cytoplasm, a highly viscosity state that reduces cellular molecular-mobility and metabolisms. Sugar and RFOs accumulation play a protective role avoiding water-crystal formation through crystal matrix disruption (Koster and Leopold, 1988). The hydrophilic properties of LEA proteins and their progressive folding during desiccation suggest that they participate in the glassy state formation (Leprince *et al.*, 2017). They are also induced upon plant desiccation and other stresses (Hong-Bo *et al.*, 2005). Indeed, in seed maturation the main abundance of most LEA coincides with the final desiccation phase (Verdier *et al.*, 2013). The role of LEA proteins in seed longevity seems crucial through their contribution to the desiccation tolerance. Nevertheless, direct confirmations are needed to establish the relation of LEA and seed longevity. It has been shown that downregula-

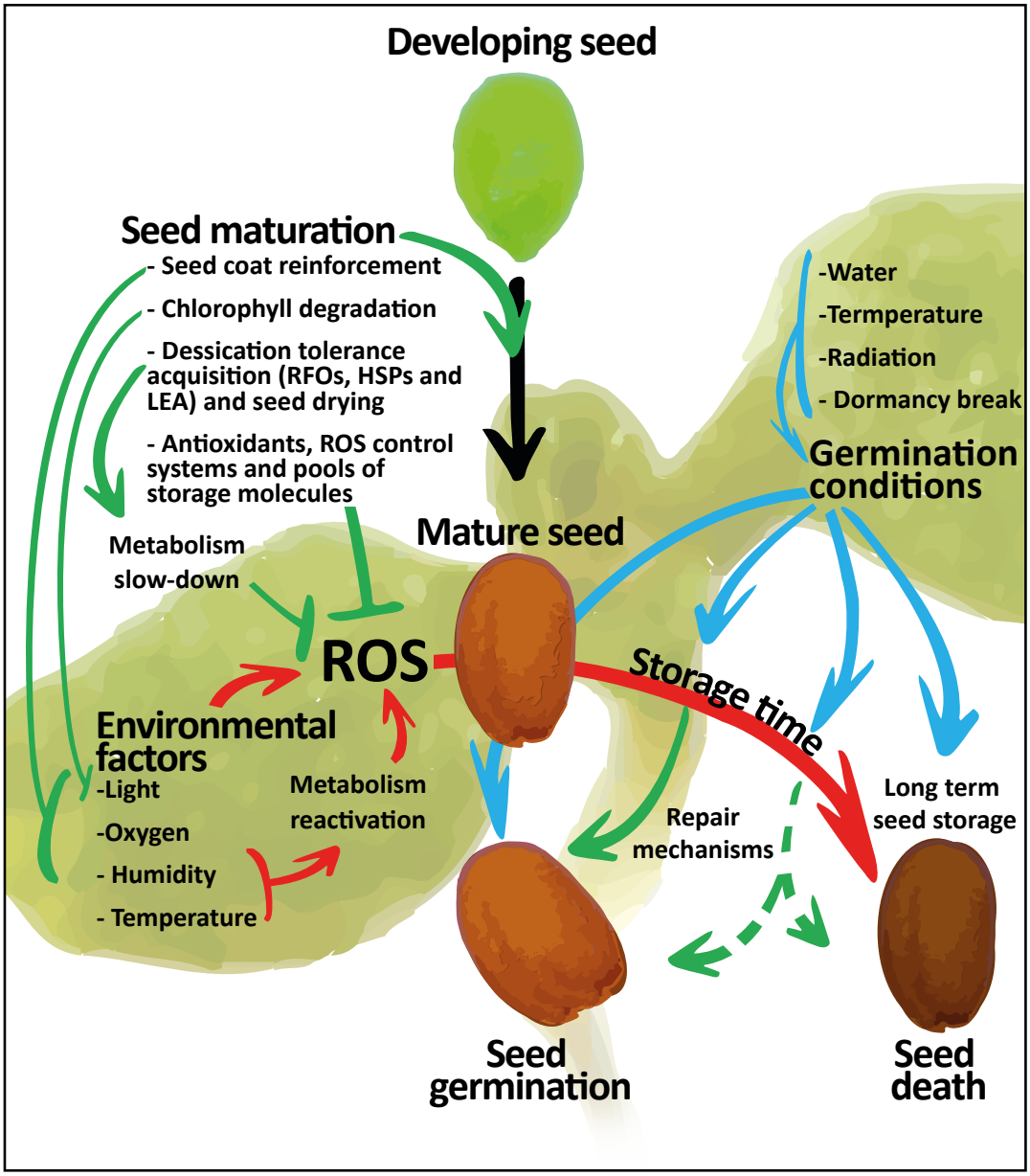


Figure 3: Schematic diagram of seed aging. Seed longevity is acquired during seed maturation through different mechanisms that reduce ROS damage during dry seed storage. ROS damage is the result of ROS accumulation and storage time. The progressive ROS accumulation during the storage depends on environmental factors, and an extended ROS damage eventually causes seed death. However, part of this damage can be repaired by embryo repair systems during seed rehydration and germination, but repair systems may be insufficient upon prolonged ROS damage. Environmental factors that are negative for seed longevity are linked with red arrows, seed molecular mechanism to avoid embryo damage, positive for seed longevity are linked with green arrows and flat head arrows, and germination processes are indicated with blue arrows. ROS: reactive oxygen species. RFOs: raffinose family oligosaccharides. HSPs: heat shock proteins. LEA: late embryogenesis abundant proteins.

tion of family members of group 2 LEA or dehydrins transcripts through interference RNA diminish seed longevity (Hundermarkt *et al.*, 2011). Members belonging to other LEA classes, in total 71 members in Arabidopsis classified in twelve groups (Jaspard *et al.*, 2012), have not demonstrated role in seed longevity. However, insights of their contribution in the seed longevity trait have been found in other plant species (Sano *et al.*, 2016). Other important component for seed longevity are small HSPs as they help to protect and fold proteins during seed desiccation and storage (Prieto-Dapena *et al.*, 2006). The sunflower transcription factor HaHSFA9, controlled by ABI3, drives the expression of various small HSPs during seed development through the interaction with the HSFs repressor -HaIAA27, also repressed by auxins (Kotak *et al.*, 2007; Carranco *et al.*, 2010). The HaHSFA9 has a positive role for seed longevity in tobacco (Tejedor-Cano *et al.*, 2010). Chlorophyll degradation, also controlled by ABI3, is an important maturation process for acquisition of seed longevity to avoid light photoreception and, thus light-induced ROS damage from the chloroplast electron transport chain (Nakajima *et al.*, 2012).

ROS-damage control

The first mechanism to avoid embryo oxidative stress caused by ROS accumulation is embryo desiccation. The seed final water content is reduced in relationship with the environmental humidity. The optimum moisture level for reducing metabolism in the glassy state in seed embryos coincides to the minimum water content permitted in seeds (around 5%). To reach this optimum moisture content, environmental relative humidity between 19 and 27% is required (Vertucci and Roos, 1990). Environments with higher humidity and temperatures break the glassy state and increase cytoplasm fluidity in seed embryos (rubbery cytoplasm). That is the case in most of land areas in earth (70%-80% relative humidity (RH), except in dry land areas or high terrains) (Dai, 2006). Thereby, seed aging in most natural environments occurs through activation of metabolism and ROS accumulation. Moreover, fluctuation in RH causes desiccation-dehydration cycles and enhance ROS accumulation affecting seed longevity. This oxidative stress damage occurs also in hypoxic environments (Considine and Foyer, 2014).

Embryo storage molecules, such as DNA, RNA pools, Seed Storage Proteins (SSPs), oils and sugars are needed for plant establishment, but ROS accumulation causes damage. Damages in DNA and genome integrity is directly related with seed aging (Waterworth *et al.*, 2015). Accumulation of seed DNA damages occurs even in the absence of external damaging factors due to the inherent DNA instability (Lindahl, 1993). Indeed, genome integrity is a maker of seed quality (Navashin and Shkvarnikov, 1933). RNA degradation also correlates with seed aging (Fleming *et al.*, 2019).

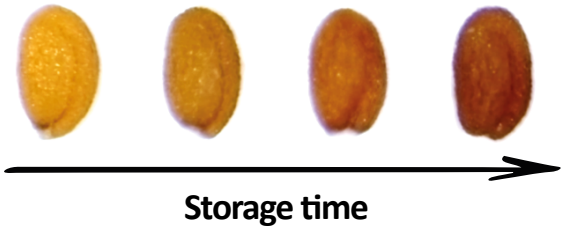


Figure 4: Seed browning occurs during seed storage. Seeds of *Arabidopsis thaliana* (ecotype Col-0) collected at different times (from left to right: freshly collected, two months old, two years old and four years old representative seeds) were imaged. Seed browning is associated with progressive oxidation of proanthocyanidins (PAs), and seeds of mutant plants in the PAs biosynthetic pathway do not present this effect.

Although RNA can be synthesized from DNA, stored RNA pools are important for germination, seed vigor and seed longevity as they drive first metabolic germination processes (Sano *et al.*, 2020). RFOs are important for the desiccation tolerance, as commented above, but they also serve as energy source during germination thanks to a rapid mobilization (Blöchl *et al.*, 2008). Decline of RFOs correlates with seed aging (Bernal-Lugo and Leopold, 1992). Oxidative damage progressively produce protein carbonylation during natural and artificial seed aging and these patterns can be used as viability markers (Job *et al.*, 2005; Rajjou *et al.*, 2008; Cabiscol *et al.*, 2014). Seed storage proteins (SSPs) are important for germination as a source of free amino acids and energy necessary for seed germination. Proteases are synthesized during seed maturation for a rapid SSPs degradation (Wang *et al.*, 2014). The aspartyl-protease ASPG1 has an important role in early proteolytic activity with an important impact in seed longevity (Shen *et al.*, 2018). Lipid peroxidation also occurs during seed storage and correlates with seed aging (Sung, 1996).

Embryos over-accumulate macromolecules to buffer ROS-induced damages. This allows having non-damaged essential macromolecules. SSPs and RFOs are examples of this passive ROS protection: they have a high affinity for ROS, they are very abundant and mutant seeds in their metabolic genes present affected germination and higher oxidative damage. For example, SSPs mutant seeds in 12S globulin genes are affected in germination (Job *et al.*, 2005; Nguyen *et al.*, 2015). For RFOs, mutant seeds in the gene *GOLS2*, a galactinol synthase, present a diminished germination and higher oxidative stress (Nishizawa *et al.*, 2015; de Souza Vidigal *et al.*, 2016).

Polyphenols, such as PAs, located in the seed coat, but also in embryo and endosperm, are important for seed longevity as shown in diverse *transparent testa (tt)* mutants (Debeaujon *et al.*, 2000). The oxidation of seed coat PAs is clearly visible during early seed storage in Arabidopsis seeds, as they gradually take their characteristic brown color due to PAs oxidation. During long seed storage, seed browning continues, and it can be used as a qualitative indicative of seed aging and oxidative damage (Figure 4). Presumably, PAs oxidation reduces ROS damage in the seed embryo. In addition,

PAs also can act as a filter to UV light, avoid pathogen growth and herbivore damage (Griffen *et al.*, 2004; Barbehenn and Peter Constabel, 2011; Wang *et al.*, 2017*b*). Plants lacking biosynthetic enzymes involved in PAs synthesis present seeds whose seed coats range from yellow to light brown colors, and in some cases, present a reduced seed longevity (Debeaujon *et al.*, 2000). This visible phenotype allowed the early characterization of mutant plants in this metabolic pathway and their genetic regulators (Koornneef, 1990; Debeaujon *et al.*, 2000; Debeaujon *et al.*, 2001; Clerkx *et al.*, 2004; Pourcel *et al.*, 2005). Other mutants in the biosynthesis and regulators of seed coat PAs are banyuls (*ban*) transparent testa glabra mutants (*ttgs*), and tannin deficient seed (*tds*) (Koornneef, 1990; Winkel-Shirley, 2001; Abrahams *et al.*, 2002; Dixon *et al.*, 2005). Some proteins may actively oxidize PAs such as the laccase TT10 (Pourcel *et al.*, 2007), other polyphenol oxidases, peroxidases and chitinases (Moïse *et al.,* 2005; Pourcel *et al.*, 2005). Low molecular weight antioxidants are part of the passive protective systems. Tocopherol (Vitamin E) has a demonstrated direct effect on seed longevity through the biosynthetic mutant *vte1* (Sattler *et al.*, 2004). Tocopherols together with lipocalins protect lipids from peroxidation and thus, enhance seed longevity (Havaux *et al.*, 2005; Boca *et al.*, 2014). Other important antioxidants for seed longevity are ascorbic acid (Vitamin C) and glutathione. Glutathione is highly accumulated in seeds constituting the main passive antioxidant in seeds: its redox potential can be used as a viability marker, whereas ascorbic acid is barely accumulated in the seed embryo (Kranner *et al.*, 2006; Nagel *et al.*, 2015). Both, ascorbic acid, and glutathione, can be directly oxidized by ROS, but there is also a tightly regulated system to oxidize and reduce them for a precise redox control to reduce ROS damage

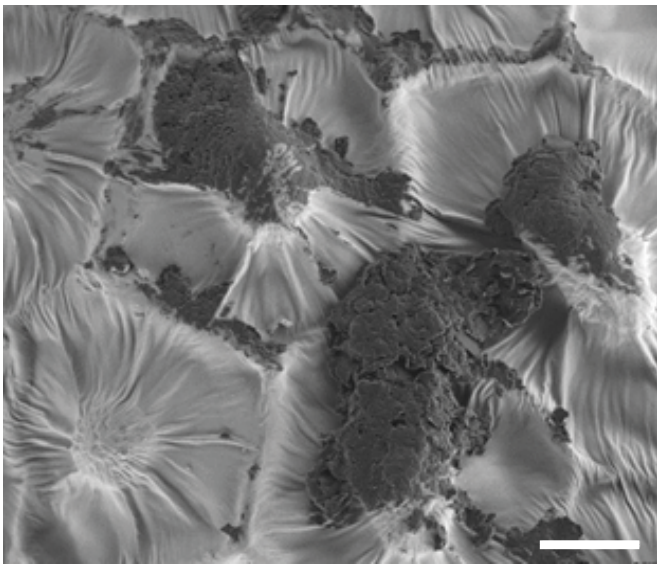


Figure 5: Fungi is growing in the seed coat surface of a three-year-old seed. Scanning electron microscope (SEM) visualization of a three-year-old seed of *Arabidopsis thaliana* with a fungus growing in the seed coat surface. Scale bar, 10 μm.

and seed deterioration: the ascorbate-glutathione-NADPH system (Foyer and Noctor, 2011). The ascorbate peroxidase enzyme oxides ascorbic acid with H_2O_2 reduction and detoxification and its overexpression increases oxidative stress tolerance, enhancing seed longevity and other stress resistance tolerance (Wang *et al.*, 1999; Chen *et al.*, 2015; Chiang *et al.*, 2017; Chen *et al.*, 2019). The ascorbate can be regenerated by monodehydroascorbate reductases, which regenerates ascorbate from NADPH and by dehydroascorbate reductases (DHARs) reducing dehydroascorbate into ascorbate with reduced glutathione, which is stored in seeds. Glutathione can be used by glutathione peroxidases and glutathione S-transferase to eliminate H_2O_2, while glutathione reductase regenerates glutathione from NADPH (Foyer and Noctor, 2011; Smirnoff, 2011; Rahantaniaina *et al.*, 2017). Other enzymes involved in avoiding ROS damage are superoxide dismutases (transform into O_2 and H_2O_2), catalases (detoxify H_2O_2 into H_2O and O_2), methionine sulfoxide reductase (repair methionine oxidized residues in proteins), thioredoxins (reduce cysteine and cleavages disulfide bonds), peroxiredoxins (detoxify ROS using electrons from thioredoxins) and glutaredoxins (reduce dehydroascorbate, peroxiredoxins and methionine sulfoxide reductase) (Nakamura *et al.*, 1997; Bailly, 2004, Weissbach *et al.*, 2005: Rouhier *et al.*, 2008; Jeevan Kumar *et al.*, 2015). The effect of all these active ROS detoxification systems in seed longevity is not completely characterized and further research is needed (Sano *et al.*, 2016). In addition, the aquaporins TIP3;1 and TIP3;2, regulated by ABI3, probably help to ROS detoxification permitting H_2O_2 transport (Mao and Sun, 2015). Although ROS-damage has a negative impact on seed longevity, ROS constitute an important signaling system for plant development and stress signaling. ROS signaling regulates the hypersensitive response upon biotic stress, abiotic stress responses, hormone levels and developmental processes including programmed cell death (Miller *et al.*, 2008; Suzuki *et al.,* 2011; Marino *et al.*, 2012). In seeds, ROS signaling controls embryogenesis, seed dormancy and germination modulating hormone levels, drives seed coat and endosperm PCD and act as protection against pathogens (Murphy *et al.*, 1998; Bailly *et al.*, 2008; El-Maarouf-Bouteau and Bailly, 2008; Jeevan Kumar *et al.*, 2015). During seed storage, high humidity can lead to pathogen, such as fungi, growth in seed tissues damaging the seed (Figure 5). Besides the putative biotic protection from pathogens by PAs (Wang *et al.*, 2017*b*), ROS production in the embryo can also inhibit pathogen growth, and eventually might locally kill infected cells to control pathogen infection (Jeevan Kumar *et al.*, 2015). The importance of ROS signaling and ROS pathogen control remarks that ROS-producing enzymes expressed during seed development and seed storage, as different NADPH oxidases (also named RBOHs), are important players in seed longevity mediating seed developmental and seed defense processes. Nevertheless, ROS production directly damages the embryo and a tight balance between seed longevity and biotic defense controls the seed survival.

Embryo repair mechanisms

Different ROS-damages might be repaired enzymatically. Embryo repair mechanisms mainly act during seed imbibition while metabolism reactivates. The observed delay in germination in aged seeds probably is related to repair processes (Powell and Matthews, 2012). Nevertheless, an extended and continued embryo ROS damage in DNA, RNA, protein, and other macromolecules leads to embryo death (Sano *et al.*, 2016; Waterworth *et al.*, 2019). Repair mechanisms are specific for each type of macromolecules. Regarding DNA, guanine hydroxylation represents the major damage. The base excision repair (BER) system can repair it. Two enzymes of the BER system whose mutant plants exhibit reduced seed longevity are OGG1 and PARP3 (Bray and West, 2005; Biedermann *et al.*, 2011; Chen *et al.*, 2012). The second characteristic DNA damage is double strand breaks, which can be repaired by ligases, and LIG4 and LIG6 have been demonstrated to have an impact on seed longevity (Waterworth *et al.*, 2010). Cell cycle control permits DNA reparation for a genome integrity maintenance, controlled by the kinases ATM and ATR, activated by double strand breaks. Both mutant seeds, *atm* and *atr*, present increased seed longevity and an early germination after seed aging treatments, but *atm* seeds present also important chromosomal abnormalities (Waterworth *et al.*, 2016).

Methionine oxidation represents the major ROS-induced protein damage. This damage can be repaired by the methionine sulfoxide reductase (Stadtman, 2006; Châtelain *et al.*, 2013). Other characteristic protein damage is the spontaneous covalent modification of asparaginyl residues, which causes protein denaturation. This damage can be repaired by an L-Isoaspartyl O-methyltransferase (PIMT1), which has an evident direct effect on seed longevity (Lowenson and Clarke, 1992; Mudgett *et al.*, 1997; Ogé *et al.* 2008). Finally, part of the protein repair systems involves heat-shock proteins and chaperones to avoid protein miss-folding and permit protein renaturation. The transcription factor HSFA9 regulating HSPs and the prolyl-isomerases ROF1 and ROF2 have been described to participate in protein protection and folding system contributing to seed longevity (Tejedor-Cano *et al.*, 2010; Bissoli *et al.*, 2012).

Regarding to embryo stored RNA it has not been reported any repair mechanisms with a clear impact on seed longevity (Sano *et al.*, 2016). However, it is likely that stored RNA pools attenuate the RNA damage, as the transcript abundance permits to maintain part of the RNA molecules undamaged for the rapid transcription needed during germination. Arabidopsis accumulates more than 12000 different storage RNA transcripts types during seed formation (Nakabayashi *et al.*, 2005). Nevertheless, progressive RNA damage reduces seed vigor until seed death (Fleming *et al.*, 2017). Longer stored RNA molecules are more likely to be damaged than smaller ones (Sano *et al.*, 2020).

Seed coat protects the seed embryo

All previous molecular mechanisms controlling seed longevity occur mainly in the embryo. However, as commented above, seeds are formed by two other tissues, the nourishing endosperm, and the protective seed coat. Indeed, the seed coat, originated from mother ovule integuments, constitute the first physical barrier, which aims, among other functions, the protection of the embryo. The seed coat has an important role in seed longevity as it protects the embryo and from physical and oxidative damage preventing the entry of water and oxygen. The different seed coat components, the cuticle, PAs, lignin-reinforced cell walls, and suberin, contribute to this protection (De Giorgi *et al.*, 2015).

Transcription factors determine seed coat

The genetic program of seed coat differentiation from ovule integuments is complex and tightly regulated. The five layers forming the final seed coat must undergo through different and precise genetic developmental programs to perform their function and synthesize their specific components. This genetic program starts with floral identity TFs that regulate ovule integuments differentiation. Although the complete regulatory cascade of TFs and their interaction regulation are far to be known, several TFs have been described to be involved in seed coat and seed longevity.

In Arabidopsis, BEL1 is required for ovule identity and interacts with AGAMOUS (AG) and SEPALLATA3 (SEP3) in a ternary protein complex, required for proper integument identity (Brambilla *et al.*, 2007). These two TF also form ternary complexes with SEEDSTICK (STK), SHATTERPROOF1 (SHP1) and SHATTERPROOF2 (SHP2) (Brambilla *et al.*, 2007) that determine redundantly the integument tissue differentiation. AINTEGUMENTA (ANT) controls integument cellular proliferation (Klucher *et al.*, 1996). ABERRANT TESTA SHAPE (ATS) is involved in the integument separation and growth (McAbee *et al.*, 2006) and INNER NO OUTER (INO) together with SUPERMAN (SUP) are involved in the outer integument growth (Gaiser *et al.*, 1995). Later, upon fertilization, integuments follow different differentiation stages regulated by several TFs (Haughn and Chaudhury, 2005; Arsovski *et al.*, 2010). The inner integument differentiation is regulated by TRANSPARENT TESTA16 (TT16/ABS) and STK (Coen *et al.*, 2017). The outer integument differentiation regulation probably goes thought APETALLA2 (AP2) and GORDITA (GOA) (Prasad *et al.*, 2010).

Mutants in these TFs show direct effects in seed development. For example, the *ap2* mutant, which also fails in petal differentiation, present an affected outer integument differentiation. This effect produces an aberrant seed shape and seed coat, with no mu-

cilage, no suberin, and a drastically reduced seed longevity (Leon-Kloosterziel *et al.*, 1994; Debeaujon *et al.*, 2000). Mutant seeds *tt16* fails in inner integument PA synthesis and present altered seed size (Nesi *et al.*, 2002; Ehlers *et al.*, 2016). Seeds of *goa* mutant present a rounder seed shape (Prasad *et al.*, 2010). Seeds of strong *ino* mutants fails in integuments growth, while *sup* seeds present an overdevelopment of the outer integuments, as SUP is a negative regulator of INO (Meister *et al.,* 2002; Gallagher and Gasser, 2008). As the gene name indicates, *ats* seeds present aberrant seed coat formation (McAbee *et al.*, 2006). In some cases, single mutants do not show defects in seed coat development, as there are other redundant TFs with overlapping functions on integument development regulation. The effect of SHP1, SHP2 and STK in integument cell-identity regulation was observed only in the triple mutant *shp1, shp2, stk* (Pinyopich *et al.*, 2003), although individual mutants have alterations defects in fruit opening (*shp1* and *shp2*) and seed abscission (*stk*).

Later, other TFs and protein complexes are involved in the regulation of the specific metabolic pathways of the different single-cell layer. MYB–bHLH–WDR protein ternary complexes have been described to precisely regulate part of these metabolic pathways. TT2–TT8/GL3/EGL3–TTG1 and MYB5–TT8–TTG1 complexes together with TT1, TT16, TTG2 and STK control the expression of DFR, LDOX, TT19, TT12, AHA10 and BAN, which are involved in seed PA biosynthesis (Xu *et al.*, 2015). Oppositely, MYB–bHLH–TTG1 complexes negative regulate cuticle biosynthetic genes LACS3, FAR6, WSD1, LTP5, LTP10, ABCG23 and LTP7, through the repressors MYBL2, CPC, OFP18 and OFP18L (Li *et al.*, 2020). In addition, TTG1 complexes regulate jasmonic acid (JA), hormone involved in biotic and abiotic stress signaling, seed dormancy and PAs synthesis; activating the JA biosynthetic genes LOX3, LOX4 and AOC3s (Wasternack and Feussner, 2018; Li *et al.*, 2020).

Similar MYB–bHLH–WDR complexes participate in the epithelial layer differentiation of the outer integument and mucilage formation. MYB5/MYB23/TT2-TT8/EGL3-TTG1, together with AP2, NARS1 and NARS2, initially regulate columella cell differentiation (Golz *et al.*, 2017). They control TTG2 and GL2 expression, whose directly regulate mucilage biosynthetic genes such as GATL5, PRX36, MUM4, PMEI6, and other TFs such as KNAT7, involved in negative regulation of cell wall lignification, directly affecting mucilage extrusion, together with MYB75 (Bhargava *et al.*, 2013). In addition, MYB61 independently also regulates KNAT7 expression in epidermal seed coat cells (Romano *et al.*, 2012). Nevertheless, knowledge of transcriptional control of seed suberin is not as complete. MYB107 is responsible of direct activation of the suberin biosynthetic genes GPAT5, FAR1, HHT and FACT (Gou *et al.,* 2017).

In our laboratory, two seed coat TFs have been isolated through an Arabidopsis activation-tagging mutant collection screening subjected to an artificial seed aging treat-

ment (Bueso *et al.*, 2014*b*; Bueso *et al.*, 2016). Both isolated mutant lines, *athb25-1D* and *cog1-2D*, overexpress the TFs AtHB25 and COG1 respectively and present similar phenotypes: enhanced seed longevity, reduced tetrazolium salt uptake and/or reduction and maternal heritage. These findings suggested that the enhanced seed longevity of both mutants was due to a seed coat effect, probably through a reduced seed coat permeability (Bueso *et al.*, 2016). In addition, lipid polyester staining with Sudan Red was more intense in mutant seeds compared to control seeds. The enhanced lipid polyester deposition must be responsible for the increased seed longevity observed in these mutants, limiting the water and air penetration and thus, the oxygen penetration, reducing the oxidative damage during seed storage and artificial aging treatments (Bueso *et al.*, 2016). These mutants also presented increased GAs levels in seedlings and siliques, and increased expression levels of the GA biosynthetic gene GA3OX3. This hormonal effect might be involved in the regulation of their reinforced seed coat, as the exogenous GAs application and the constitutive GA-signaling of the quintuple DELLA mutant also enhances seed longevity (Bueso *et al.*, 2014*b*). Nevertheless, transcriptomic analysis showed that both mutants, *athb25-1D* and *cog1-2D*, overexpressed two biosynthetic enzymes, LACS9 and PRX25 which might be directly involved in biosynthesis of lipid polyesters, suberin and cutin, responsible of the enhanced seed longevity (Bueso *et al.*, 2016). In addition, SEP3 was also induced in both activation-tagging mutants.

The endogenous role of AtHB25 and COG1 and their location in the TF regulatory pathway of the seed coat is still unclear. The transcription factor regulatory pathway regulating seed coat components is complex, and the list of TFs is still growing. However, the comprehension of these genetic determinants on seed coat together with the regulation of their components is crucial to understand the regulation of seed coat components during seed coat performance and, thus, the seed longevity associated processes. Several works have grouped seed coat TFs in different levels of regulation (depending if they regulate other TFs, directly biosynthetic enzymes, or both). For a more compressive understanding see Gou *et al.,* 2017 and Li *et al.*, 2020. Nevertheless, the TF regulation network is more complex as there are diverse TFs that participate in different pathways. In addition, the hierarchical order is not always evident, there is gene redundancy, and the regulatory cascade is far to be completely revealed.

Seed coat lipid polyester barriers

The importance of seed coat lipid polyester barriers in seed longevity will be remarked along this PhD thesis. Both lipid polyester barriers, the cuticle layer and the suberin layer, controls the flux of gases, water, and solutes. As commented above, these barriers limit atmospheric oxygen diffusion avoiding oxygen-derived ROS damage. This effect is likely to extend seed viability.

Both cutin and suberin are complex matrices composed by aliphatic monomers (fatty acids, fatty alcohols, ω-hydroxy fatty acids, α,ω-dicarboxylic acids), aromatic monomers (hydroxycinnamic acids (HCAs), such as ferulate) and glycerol, linked by ester-bonds (Philippe *et al.*, 2020). Generally, suberin has higher proportion of fatty alcohols, hydroxycinnamic acids and glycerol (Pollard *et al.*, 2008). In addition, suberin and cutin monomer composition also variates among different plants and tissues.

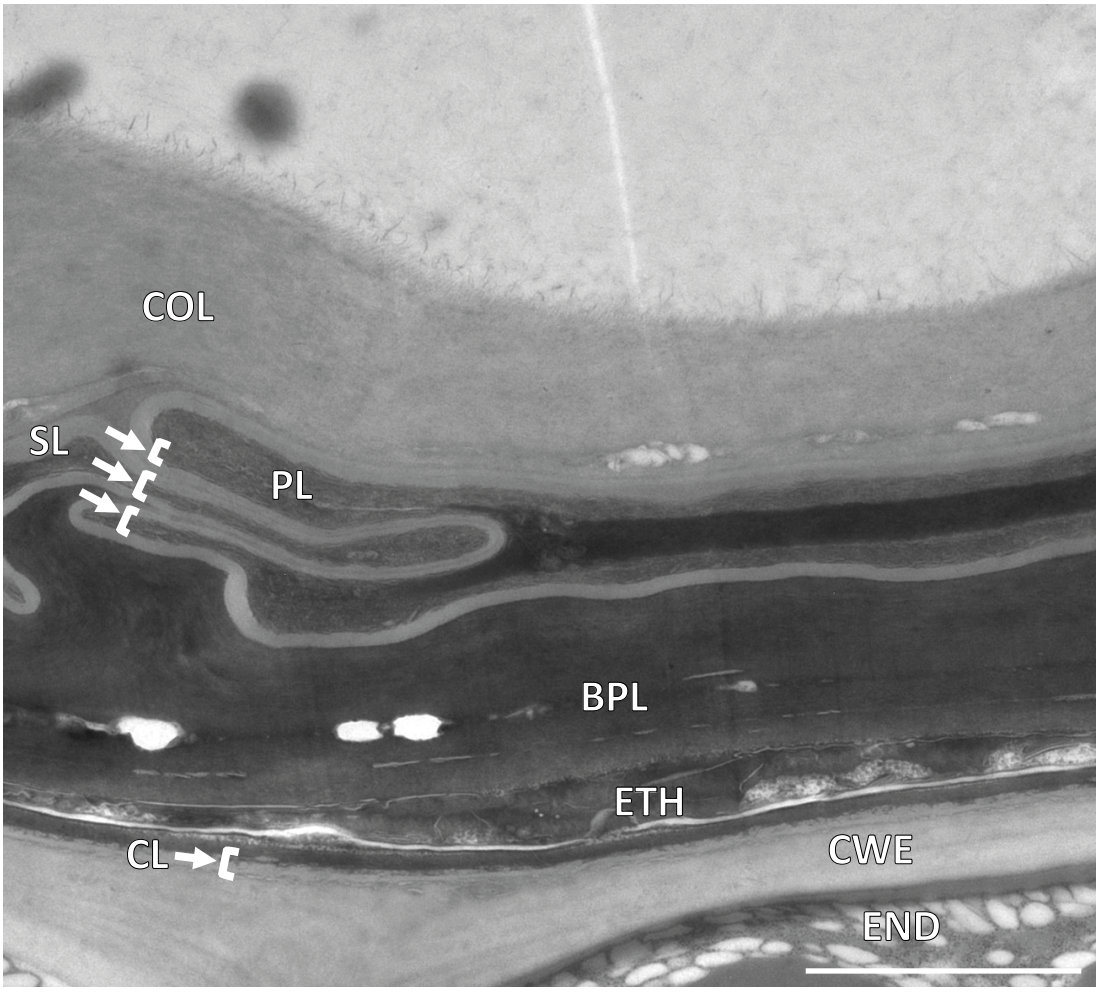


Figure 6: Arabidopsis Col-0 seed coat visualized with the Transmission Electron Microscope (TEM). The image is focused on the seed coat. The endosperm can be visualized below. Upper arrows and brackets indicate the location of the suberin lamellae, formed by suberin deposited in the inner side of the palisade cell wall. The lamellae pattern is barely visible (dark area). Nevertheless, in the intersection between two collapsed palisade cells (where arrows and brackets indicate), both cell walls fold and permit a better suberin lamellae visualization compared to other zones. Brackets show three areas where the lamellae can be visualized. However, the suberin also must impermeabilize the whole cell wall as it is not stained by proanthocyanidins (PAs). The arrow below points the cuticle layer. The cuticle layer is located below the endothelium cell wall, visualized as a dark layer as it is impregnated by PAs. SL: Suberin layer, CL: Cuticle layer, COL: columella cell, PL: palisade layer cell, BPL: brown pigment layer, CWE: endosperm cell wall, ETH: endothelium, END: endosperm cell. Scale bar, 2 μm.

Concretely, in seeds, cutin monomer composition is similar to suberin composition, but suberin present higher loads of ω-hydroxy fatty acids (C18:1, C18:2, C22 and C24) α,ω-dioic acids (C24 and C22), and ferulate. These differences were obtained through an *ap2* mutant seeds analysis, which fails in outer integument differentiation and exhibits reduced suberin (Molina *et al.,* 2008). Nevertheless, the main difference between cuticle and suberin layers is not monomer composition, but their location and their characteristic disposition pattern. Both layers are covalent bound to the plasma membrane, but while suberin is in the plasma membrane face and forms a characteristic lamellae pattern; the cuticle layer localizes in the outer cell wall face and can present different deposition patterns (lamellae, reticulate, or amorphous) (Jeffree, 2007). The cuticle is not only formed by cutin, but also by waxes and polysaccharides, forming a chemical gradient where the internal part contains cutin, waxes and polysaccharides and the external part is formed only by cutin and waxes (Jeffree, 2007, Bakan and Marion, 2017). In Arabidopsis, cutin composition differs from other cuticles from organisms such as tomato, which has standard cuticles. Arabidopsis cuticles mainly contain dicarboxylic fatty acids and glycerol 3-phosphate, but not dihidroxy monomers, resulting in fragile cuticle layers with a suberin-like monomer composition. The Arabidopsis seed coat cuticle layer present amorphous deposition (Loubéry *et al.*, 2018), and it can be considered a protocuticle as it is not completely differentiated as other cuticle layers.

The biosynthesis of main lipid polyester monomers of seed coat is schematized in the Figure 7, and their abundance can be visualized in Figure 8. Fatty suberin and cutin monomers derive from C16:0, C18:0 and C18:1 fatty acids. They are synthetized in plastids and transported into the endoplasmic reticulum (ER) where undergo further modifications (Li-Beisson *et al.*, 2013). The fatty acid elongase (FAE) complex involving the β-Ketoacyl-Coenzyme A (CoA) synthases enzymes (KCS) elongate them into very long chain fatty acids (VLCFA). The KCS2 and KCS20 enzymes are involved in suberin monomer elongation. KCS2, also named DAISY, elongates suberin monomers at the seed coat chalaza-micropyle region (Franke *et al.*, 2009; Lee *et al.*, 2009*b*). KCS10, also named FIDDLEHEAD, is involved in leaf cuticle formation (Pruitt *et al.*, 2000). Further studies are needed to establish the importance of VLCFA with seed coat sealing and seed longevity.

Fatty acids are transformed into fatty alcohols by FATTY ACYL REDUCTASE (FAR) genes. Suberin fatty alcohols are synthesized by FAR1, FAR4 and FAR5 enzymes (Domergue *et al.*, 2010). FAR5 is involved in C18:0-OH synthesis, FAR4 in C20:0-OH and FAR1 in C22:0-OH (Vishwanath *et al.*, 2013). In addition, double mutants, but mainly triple mutant seeds *far1 far4 far5*, present higher permeability to tetrazolium salts, indicating an important role of fatty alcohols in seed coat sealing and embryo impermeabilization. The tetrazolium salts assay assesses qualitatively the seed coat

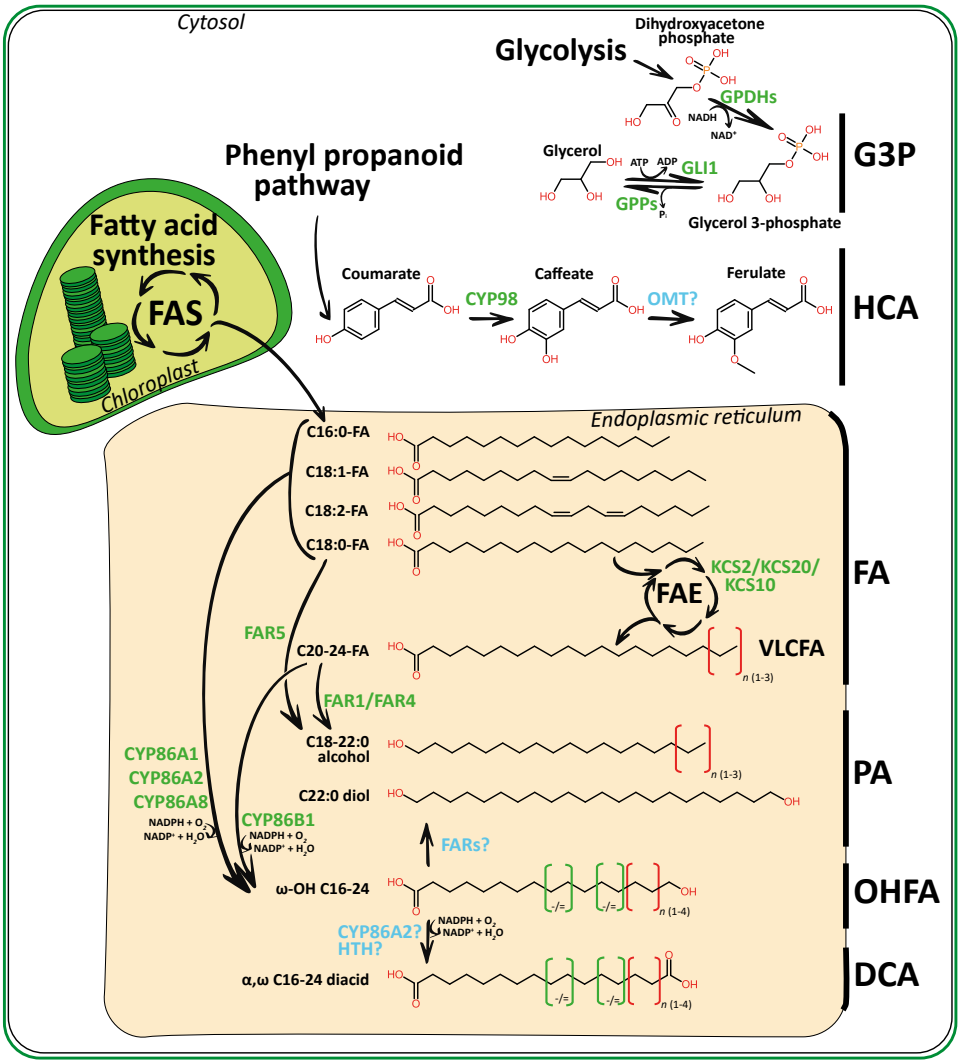


Figure 7: Lipid polyester monomer synthesis. Schematic representation of the synthesis of abundant lipid polyester monomers detected in seeds in our analysis. The aliphatic monomers start synthesizing in plastid through the Fatty Acid Synthesis complex (FAS) and then transported into the endoplasmic reticulum. They may elongate by the Fatty Acid Elongate complex (FAE) for very long chain fatty acid (VLCFA) synthesis. Modification of fatty acids (FA) by different enzymes transform them into primary alcohol (PA) or fatty alcohols, ω-hydroxy fatty acids (OHFA) and of α,ω-dicarboxylic fatty acids (DCA). By contrast, aromatic monomers of lipid polyesters, called hydroxycinnamic acids (HCA), are synthesized in the cytosol. They come from the phenylpropanoid pathway. Finally, the glycerol 3-phosphate (G3P) is synthesized from both, glycolysis intermediates and glycerol. The G3P is necessary for monomer polymerization, but it is removed. This figure is partially based in Vishwanath *et al.* (2015) and Philippe *et al.* (2020) lipid polyester schemes. Molecules names are in black. Enzymes demonstrated to participate in this biosynthetic process are marked in green. When different enzymes from the same family are described for a particular reaction, those with a demonstrated role in lipid polyester synthesis in the seed coat are represented. Enzymes that present a direct effect on the synthesis, but they are not biochemically characterized, or it is unclear if they participate in the process are in colored blue with a question mark. Red brackets indicate similar molecules different in the length of the carbon chain, and green brackets indicate C18 similar molecules with or without unsaturation.

permeability of fresh seed lots. If the tetrazolium salt pass through the seed coat it is reduced by the embryo to formazans, a coloured molecule easily visible (Figure 2c, Debeaujon *et al.*, 2000). However, a non-viable embryo cannot reduce tetrazolium salts. Thus, this assay can be used to evaluate the viability of seed lots after storage.

The cytochrome P450 family of 86 members (CYP86) catalyses the ω-hydroxylation of fatty acids to form ω-hydroxy fatty acids (OHFA), the precursor of α,ω-dicarboxylic fatty acids (DCAs) (Bak *et al*., 2011). Both are part of seed-coat lipid polyesters (Figure 8). Members of the family have been related with cutin and suberin monomer synthesis. In vitro and mutant plant analysis has shown that the CYP86A clade, formed by 5 members, generally transform 16:0, 18:0 and 18:1 fatty acids into ω-hydroxy fatty and DCAs, while CYP86B (two members) clade would use VLCFA C22 and C24 (Wellesen *et al.*, 2001; Höfer *et al.*, 2008; Molina *et al.*, 2008; Compagnon *et al.*, 2009). The CYP86C clade, composed by four members, needs further investigation, but they might act like the CYP86A clade (Kai *et al.*, 2009; Bak *et al.*, 2011). Different members of the CYP86 family have been studied in cutin and suberin monomer synthesis. CYP86B1 and CYP86A2 are involved in seed coat ω-hydroxy fatty acids and α,ω-dicarboxylic fatty acids synthesis (Molina *et al.*, 2008; Compagnon *et al.*, 2009). However, neither *att1* (*cyp86a2*) and *ralph1* (*cyp86b1*) mutant seeds

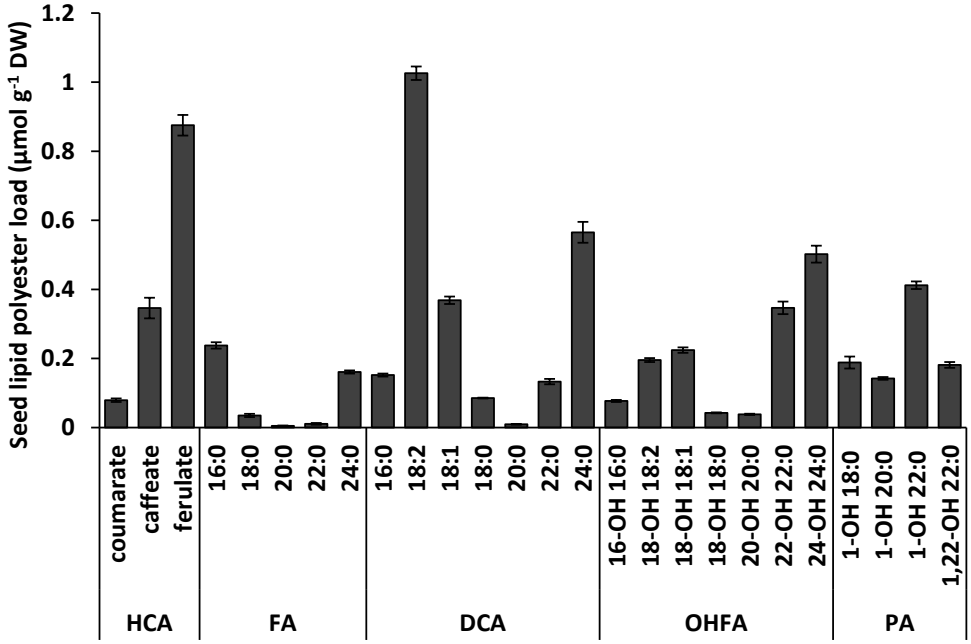


Figure 8: Lipid polyester monomer composition of Arabidopsis Col-0 seeds. Mean values are shown in µM g^{-1} delipidated dry weight (DW) ± SE of three biological replicates. *Significantly differing from controls (wild type) at $p < 0.05$ (Student's *t* test). HCA: hydroxycinnamic acids; FA: fatty acids; DCA: α,ω-dicarboxylic acids; OHFA: ω-hydroxy fatty acids; PA: primary alcohols (fatty alcohols). Caffeate values are obtained from Kosma *et al.* (2012) analysis. The rest of seed lipid-polyester monomer analysis were conducted in collaboration with Isabel Molina.

where notably more permeable to tetrazolium salts. These studies did no assessed the seed longevity trait. Other CYP86 family members have not been described to determine seed coat properties: CYP86A1 is related to root suberin (Höfer *et al.*, 2008), CYP86A4 and CYP86A7 with petal cutin (Kannangara *et al.*, 2007), and CYP86A8 is important for epidermis cuticle, as mutant plants *lcr* (*cyp86a8*) present organ fusions and pleiotropic defects (Wellesen *et al.*, 2001). However, the effect of these other CYP86 family members have not been studied in seeds.

The HOTHEAD (HTH) is involved in the synthesis of cutin DCAs. The cuticle of *hth* mutants present reduced DCAs and increased ω-hydroxy fatty acids (Kurdyukov *et al.*, 2006*b*). Although its molecular reaction is not characterized, it might be involved in DCA biosynthesis from ω-hydroxy fatty acids. However, the reduction of ω-hydroxy fatty acids into DCAs is an expected secondary reaction (Bak *et al.*, 2011). Its expression in seeds might uncover a role for the embryo cuticle formation (Kurdyukov *et al.*, 2006*b*).

CYP77A clade members CYP77A4 and CYP77A6 oxidize cutin monomers in the midchain. CYP77A4 uses unsaturated fatty acids to add epoxide groups (Sauveplane *et al.*, 2009) and controls embryo patterning. However, it is not expressed in the seed coat during seed development and thus is unlikely to have an important role in seed coat cutin biosynthesis (Kawade *et al.*, 2018). CYP77A6 performs midchain hydroxylation ω-hydroxy fatty acids of cutin monomers after CYP86 reactions and has a determinant role in flower cutin monomers (Li-Beisson *et al.*, 2009). CYP77A6 products can be used by the EPOXIDE HYDROLASE1 (EH1), a cytosolic epoxide hydrolase involved in the synthesis of poly-hydroxylated cutin monomers (Fich, *et al.*, 2016). Mutant seeds *eh1* present an increased permeability (Pineau *et al.*, 2017), suggesting an importance of poly-hydroxylated cutin in seed coat sealing, and the involvement of CYP77A6 or another CYP77 family member in the EH1 seed coat cutin precursor synthesis.

Long-chain acyl-CoA synthetases (LACS) catalyze the synthesis of acyl-CoA intermediates. Although several members are important for oil biosynthesis in seed embryos, few of them have reported roles in lipid polyester biosynthesis. LACS1, LACS2 and LACS4 participate in cutin monomer biosynthesis (Lü *et al.,* 2009; Zhao *et al.,* 2019). They add CoA to a broad range of long-chain lipid-polyester precursors, such as fatty acids, ω-hydroxy fatty acids and DCAs, activating all these suberin and cutin monomers (Schnurr *et al.*, 2004). Then, the CoA is removed in successive reactions (Figure 9). In addition, LACS activity is also needed for monomer trafficking (Philippe *et al.*, 2020). The LACS2 enzyme has an important role in the biosynthesis of lipid polyester barriers of the seed coat, and mutant *lacs2* seeds present an affected seed germination upon an artificial seed aging treatment (De Giorgi *et al.*, 2015). In leaves, *lacs2* mutant plants present a strong reduction of DCAs (Bessire *et al.*, 2007). This indicates that

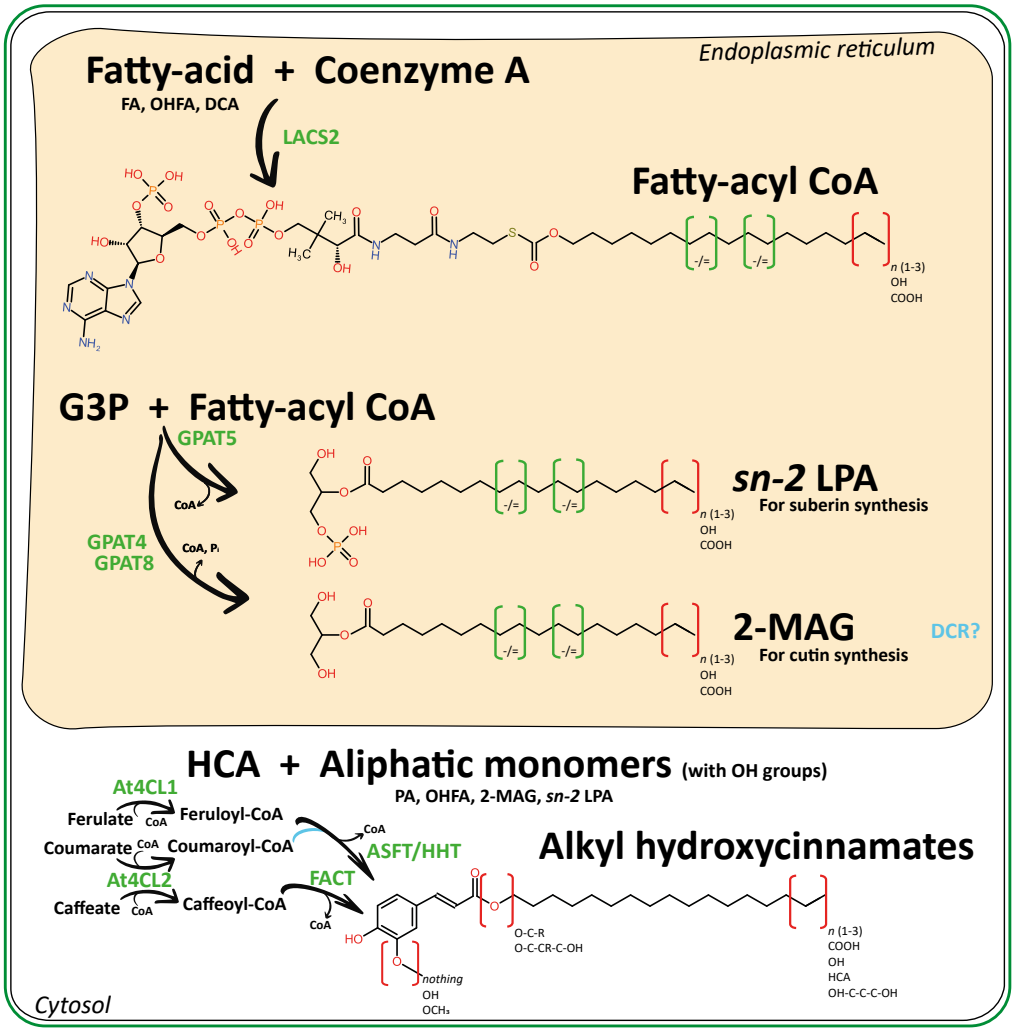


Figure 9: Lipid polyester intermediates synthesis. Schematic representation of the synthesis of lipid polyester intermediates which are not detected in out lipid polyester analysis. In the cytosol the hydroxycinnamic acids (HCA) are ester bound to OH groups of fatty alcohols or primary alcohols (PA) or hydroxy fatty acids (OHFA) and their derivates as 2-MAG or *sn-2* LPA (see below). In the endoplasmic reticulum fatty acids (FA), hydroxy fatty acids (OHFA) and dicarboxylic acids (DCA) are activated with the Coenzyme A (CoA) forming fatty-acyl CoA which is then conjugated with glycerol 3-phosphate (G3P) producing *sn-2* LPA and 2-MAG intermediates, used in suberin and cutin synthesis, respectively. The addition of Coenzyme A remains unclear if it is produced during monomer synthesis, or during intermediate synthesis, or even both. CoA addition has been demonstrated to be important during fatty acid synthesis, but also after in cutin and suberin monomer synthesis. It may be also important to monomer trafficking. This figure is partially based in Vishwanath *et al.* (2015) and Philippe *et al.* (2020) lipid polyester schemes. Molecules names are in black. Enzymes demonstrated to participate in this biosynthetic process are marked in green. When different enzymes from the same family are described for a particular reaction, those with a demonstrated role in lipid polyester synthesis in the seed coat are represented. Enzymes that present a direct effect on the synthesis, but they are not biochemically characterized, or it is unclear if they participate in the process are in colored blue with a question mark. Red brackets indicate similar molecules different in the length of the carbon chain, terminal group and/or glycerol backbone presence, and green brackets indicate C18 similar molecules with or without unsaturation.

LACS2 is important for DCA synthesis, probably acting after CYP86 ω-hydroxylation reaction, but the concrete reaction order has not been completely elucidated (Pollard *et al.*, 2008). The role of LACS enzymes in monomer trafficking might involve also other members of the family, as LACS2 is localized in the endoplasmic reticulum (Weng *et al.*, 2010, Philippe *et al.*, 2020). In the embryo, LACS activity also must be important, as some simple, double and triple *lacs* mutant seeds present an aberrant testa shape phenotype and this effect may be caused by defects in storage of oil accumulation in seed embryos (Schnurr *et al.*, 2004; Zhao *et al.,* 2019), probably affecting also seed longevity.

The glycerol 3-phosphate (G3P), one of the main components of lipid polyesters, is added to aliphatic monomers by GLYCEROL-3-PHOSPHATE *SN-2*-ACYLTRANSFERASES (GPAT) enzymes. The addition to the acyl group in the carbon 2 of the glycerol molecule (*sn-2* position) is characteristic of suberin and cutin precursors, in contrast to *de novo* synthesis of membrane and storage lipids, where the acyl group is added in the carbon 1 (*sn-1* position) (Murata and Tasaka, 1997; Wendel *et al.*, 2009; Yang *et al.*, 2010). The biosynthetic enzymes GPAT1, GPAT4, GPAT5, GPAT6, GPAT7 and GPAT8 have mainly *sn-2* acyltransferase activity, although they also present little *sn-1* activity (Yang *et al.*, 2010; Yang *et al.*, 2012). The *in vitro* activities of GPAT2 and GPAT3 remain uncharacterized, but they are predicted to be localized in the mitochondria together with GPAT1 (Zheng *et al.*, 2003; Beisson *et al.*, 2007). The GPAT4, GPAT6 and GPAT8 add the G3P in cutin monomers of leaves (GPAT4 and GPAT8) and flowers (GPAT6) (Li *et al.*, 2007; Li-Beisson *et al.*, 2009) and present phosphatase activity producing monoacylglycerol (2-MAG) products (Yang *et al.*, 2010; Yang *et al.*, 2012). The double mutant *gpat4 gpat8* have reduced seed longevity upon accelerated aging treatments (De Giorgi *et al.*, 2015) indicating an important role in seed cuticle biosynthesis. By contrast, GPAT5 and GPAT7 lack the phosphatase activity and produce 2-MAG-3-phosphate, also named *sn-2* lysophosphatidic acid (*sn-2* LPA), monomers (Yang *et al.*, 2010; Yang *et al.*, 2012), which are considered the suberin building blocks. Probably, one important difference between suberin and cutin monomer biosynthesis regards this phosphate extra in the suberin monomers (Figure 9, Philippe *et al.*, 2020). Perhaps the different deposition pattern depends on this extra phosphate. GPAT7 has important functions in wounding-induced suberin (Yang *et al.*, 2012). GPAT5 is the only GPAT enzyme directly involved in seed coat suberin synthesis, with a drastic effect in the seed coat impermeability (Beisson *et al.*, 2007). Mutant *gpat5* seeds present one of the strongest seed coat permeability defects in tetrazolium salt uptake, comparable to the *myb107* (Beisson *et al.*, 2007, Gou *et al.*, 2017), and it is used as negative control of seed coat suberin in seed coat permeability studies (Compagnon *et al.*, 2009; Vishwanath *et al.*, 2013; Fedi *et al.*, 2017). This fact remarks the importance of glycerol 3-phosphate in seed suberin impermeabilization. However, direct studies of the effect of glycerol and its metabolism regarding lipid

polyester synthesis are still missing.

Contrary to aliphatic monomers of cutin and suberin that are synthesized in the ER, the aromatic monomers are synthesized in the cytosol. They are formed form hydroxycinnamic acid, which derives from the phenylpropanoid pathway (Bernards *et al.*, 1995). Few members of the BAHD acyltransferase family ester-bond hydroxycinnamic acids and aliphatic monomers (and intermediates) that present OH groups, producing alkyl hydroxycinnamates (AHCs) (Figure 9) (D'Auria, 2006; Vishwanath *et al.*, 2015;

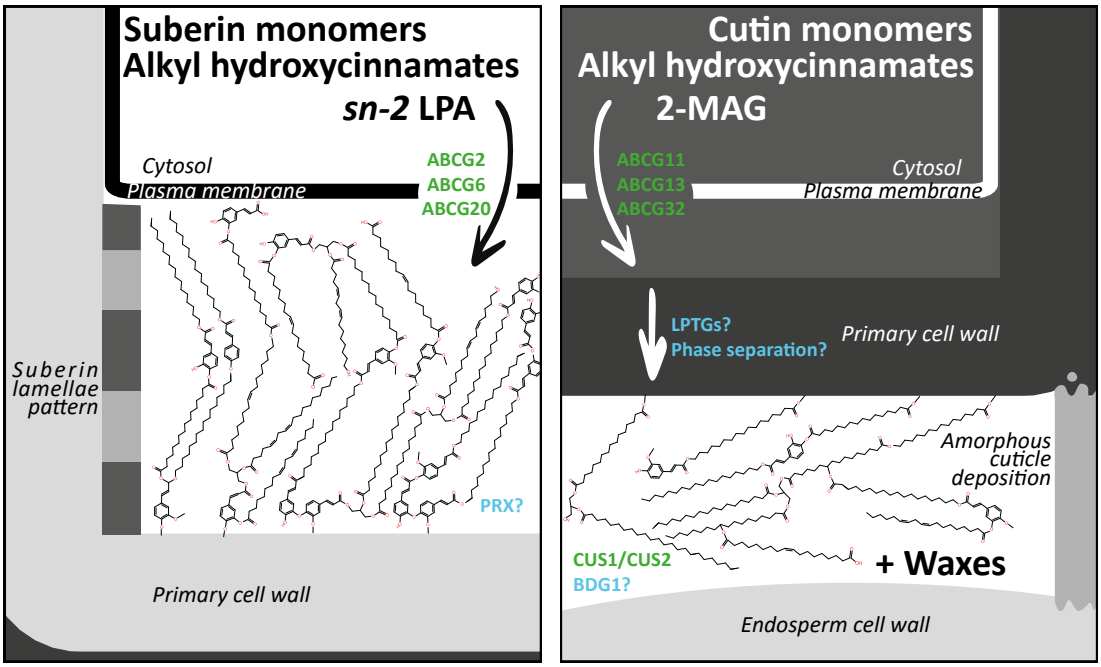


Figure 10: Hypothetical (a) suberin and (b) cutin polymerization synthesis schemes in the seed coat. Partially based in Fich *et al.* (2016) and Domergue and Kosma (2017) schemes and current views of lipid polyester structures. Suberin and cutin monomers and intermediates are transported through ABCG membrane transporters. While suberin deposition in the seed coat occurs in the inner side of the cell wall of palisade cells, the cutin monomers are transported through the hydrophilic cell wall of endothelium cells. This transport may occur by phase separation and/or by LPTG transport. Suberin is polymerized by ester bonds by unknown enzymes, probably involving peroxidases (PRX). It is proposed that the suberin lamellae pattern is caused by the distribution of aromatic and aliphatic components. The cutin polymerization, also by ester bonds, is produced by CUS1 and CUS2 and probably BDG1. Both impermeable layers limit proanthocyanidin (PA) diffusion. PA-stained areas constitute the brown pigment layer, visualized as electron dense zones with the transmission electron microscope. PAs liberation occurs during seed coat collapse, after lipid polyester deposition, however it is represented in the scheme with grey colors for an easier structure identification. The cuticle layer is attached to the endosperm cell wall. Molecule group names are in black. Enzymes demonstrated to participate in this biosynthetic process are marked in green. When different enzymes from the same family are described for a particular reaction, those with a demonstrated role in lipid polyester synthesis in the seed coat are represented. Enzymes that present a direct effect on the synthesis, but they are not biochemically characterized, or it is unclear if they participate in the process are in colored blue with a question mark. Grey colors represent the transmission electron microscope visualization. Grey bars represent both lipid polyester barriers: the suberin lamellae and the amorphous cuticle deposition.

Domergue and Kosma, 2017). Ferulate, the major hydroxycinnamic acid, is ester-bond into suberin monomers (ω-hydroxy acids and fatty alcohols) by the ALIPHATIC SUBERIN FERULOYL TRANSFERASE (ASFT) forming alkyl ferulates. Mutant *asft* seeds completely lack ferulate in their lipid polyester, but the rest of monomer components remain unaltered (Molina *et al.*, 2009). Mutant seeds of *asft* plants present increased seed coat permeability (Gou *et al.*, 2017). ASFT, also named HHT, can use *p*-coumaroyl-CoA as substrate, but with lower specificity compared to feruloyl-CoA (Gou *et al.*, 2009). This may explain the coumarates monomers seen in the seed coat lipid polyester analysis (Figure 8). Nevertheless, it is plausible that another enzyme catalyzes the biosynthesis of alkyl coumarates (Vishwanath *et al.*, 2015). The FATTY ALCOHOL:CAFFEOYL-CoA CAFFEOYL TRANSFERASE (FACT) is expressed in the seed coat palisade layer and it is responsible of the esterification caffeotyl-CoA into suberin aliphatic monomers producing alkyl caffeates (Kosma *et al.*, 2012). Mutant fact seeds present high reduction in seed-suberin caffeate content. The higher hydroxycinnamic acids present in suberin monomers has been suggested as responsible of the characteristic lamellae deposition pattern of suberin by the succession of phenolic domains and aliphatic components (Bernards, 1998; Gadini *et al*., 2006). However, a direct reduction of phenolic compounds does not alter the suberin lamellae structure in roots, as in mutant *asft* plants (Molina *et al.*, 2009). Therefore, the cause of suberin lamellae structure remains an open question. In cutin biosynthesis, DEFICIENT IN CUTIN FERULATE (DCF) conducts the feruloylation of ω-hydroxy fatty acids (Rautengarten *et al.*, 2012), like ASFT in suberin. The DEFECTIVE IN CUTICULAR RIDGES (DCR) may use 2-MAG to form cutin precursors prior to their extrusion, as it is a BAHD acyltransferase, and their mutant plants present altered cuticle formation and defect in seed coat epidermis and mucilage (Panikashvili *et al.*, 2009). This defect leads to a reduced seed longevity in *drc* seeds (De Giorgi *et al.*, 2015). However, the DCR molecular mechanism is still not characterized. Other important enzymes involved in the synthesis of this lipid polyester aromatic compounds and their involvement in seed coat lipid polyester proprieties have yet to be characterized (Philippe *et al.*, 2020).

ATP-BINDING CASSETTE (ABC) are transmembrane transporters directly energized by ATP that transport complex organic molecules (Do *et al.*, 2018). Different members of the subfamily G (ABCG) transport the different suberin and cutin monomers across the plasma membrane (Figure 10). The ABCG2, ABCG6 and ABCG20 transport suberin monomers, and the triple mutant *abcg2 abcg6 abcg20* present a drastic reduction of suberin monomers in seeds, but not in roots, probably due to an increase-seed suberin monomer synthesis compensatory effect in roots but not in seeds (Yadav *et al.*, 2014). The triple mutant, but also the double *abcg2 abcg20* and the strong *abcg20* mutant *awake1*, show a drastic seed permeability increase similar to *gpat5* mutant seeds, remarking the importance of these ABCG transporters in

seed coat suberin biosynthesis (Yadav *et al.*, 2014; Fedi *et al.*, 2017). The ABCG11, ABCG13 and ABCG32 are involved in 2-MAG cutin monomer transport into the apoplast and they are well-characterized in flower and leaves cuticle formation (Pighin *et al.*, 2004; Bird *et al.*, 2007; Takeda *et al.*, 2014; Fabre *et al.*, 2016; Do *et al.*, 2018), but their role is not molecular characterized in seed cuticle formation. In addition to plasma membrane extrusion, cutin monomers (and associated waxes) pass through the cell wall for their deposition. Cell wall hydrophilic properties difficult this transport. Although simple diffusion may occur through aliphatic aggregation and phase separation, GLYCOSYLPHOSPHATIDYLINOSITOL-ANCHORED LIPID TRANSFER PROTEINS (LPTG) are proposed to be involved in cell wall cuticular wax transport. LPTG1, LPTG2 might be involved in this speculative transport (Fich *et al.*, 2016), as mutant *lptg1* and *lptg2* plants present a reduced cutin deposition and altered cuticular structure (Lee *et al.*, 2009*a*; Kim *et al.*, 2012). They are expressed in the seed coats, as in other cuticle forming tissues. Hereby, they could have important roles in seed coat cuticle sealing. Recently, LPTG15 has been characterized in VLCFA suberin monomer transport for seed coat suberin formation. Although authors claim that this gene is important for seed coat suberin sealing, mutant *lptg15* seeds are barely affected in permeability compared to wild type seeds (Lee and Shu, 2018). On the other hand, suberin monomers do not pass through the cell wall, as suberin depositions occurs in the inner part of the polysaccharide cell wall next to the plasma membrane.

Finally, the last lipid-polyester biosynthesis step is the *in situ* monomer assembly. Some enzymes have been suggested to be involved in cutin polymerization. The CUTIN SYNTHASE 1 (CUS1/LTL1) polymerize 2-MAG cutin precursors through transesterification of end-chain and mild-chain hydroxyl groups (Yeats *et al.*, 2012; Philippe *et al.*, 2016). In Arabidopsis, CUS2 is involved in petal cuticle formation. Nevertheless, both CUS1 and CUS2 are expressed in siliques and might have a role in seed coat cuticle synthesis (Hong *et al.*, 2017). However, other family members may have important roles. In addition, BODYGUARD1 (BDG1) might thus be involved in the extracellular polymerization of Arabidopsis C18 unsaturated cutin precursors (Jakobson *et al.*, 2016). BDG1 belongs to the hydrolase gene family, but its molecular function in cutin monomer polymerization and/or organizing the cuticle formation is yet unknown (Kurdyukov *et al.*, 2006*a*; De Giorgi *et al.*, 2015). BDG1 has an important role in cutin biosynthesis as mutant *bdg1* plants are dwarf and present fused organs and cuticle defects (Kurdyukov *et al.*, 2006*a*; Chassot *et al.*, 2007). Mutant *bdg1* seeds present endospermic cuticle alterations and alterations in seed coat permeability (De Giorgi *et al.*, 2015). In addition, other enzymes might be involved in cutin and suberin polymerization. For example, peroxidases are postulated to have a role in cell wall polyphenolic assembly of suberin (Bernards *et al.*, 2004). Peroxidases can generate phenolic radicals that form lignin polymers spontaneously (Lewis and Yamamoto, 1990), and similarly might act in suberin phenolic monomer cell-wall

polymerization. Little evidence supports this idea, a tomato peroxidase directly affects suberin and lignin (Quiroga *et al.*, 2000). However, further investigation is needed to support this idea.

In summary, the lipid polyester biosynthetic pathway is characterized partially. In addition, seed coat polyesters are not the principal tissue of research of these biosynthetic pathways. Epithelial and petal cuticles and root suberin are in the main research focus. Nevertheless, different lipid polyester biosynthetic enzymes exhibit seed coat phenotypes indicating an important role of them in seed coat sealing and impermeabilization, a characteristic that seems important to seed longevity. Biosynthetic cuticle mutant seeds *lacs2*, *dcr*, *bdg1* and *gpat4 gpat8* present reduced seed deterioration upon accelerated aging tests (De Giorgi *et al.*, 2015). On the other hand, biosynthetic suberin mutant seeds *gpat5*, *abcg2 abcg6 abcg20*, *far1 far4 far5, eh1* and with minor impact *lptg15,* present increased seed coat permeability (Beisson *et al.*, 2007; Vishwanath *et al.*, 2013; Yadav *et al.*, 2014; Pineau *et al.*, 2017), but their effect in seed longevity has not been yet determined.

As commented above, the transcription factor MYB107 is the only TF described to regulate seed coat suberin biosynthesis, and their mutant seed present a similar seed permeability phenotype like suberin biosynthetic mutants (Gou *et al.*, 2017). Nevertheless, cutin regulatory TFs in seed coat have not been described in Arabidopsis. The only TF described to regulate seed coat cuticles is the *Medicago* KNOX4 TF (Chai *et al*., 2016). In addition, other MYB transcription factors, MYB16 and MYB106, have been described to regulate cuticle formation, affecting directly different cuticle biosynthetic genes such as *LACS2*, *CYP86A4*, *CYP77A6*, *KCS10/FDH* (Oshima and Mitsuda, 2013; Oshima *et al.*, 2013), and thus, they might have a role in seed coat cuticle development in a similar regulatory pathway than MYB107. Other Arabidopsis cuticle regulatory TFs are described: SHN1, SHN2, SHN3, CD2 and ANL2 are positive regulators, while CFL1 and NFXL2 are negative regulators (Fich *et al.*, 2016; Philippe *et al.*, 2020). It is only a matter of time before more TF regulating seed coat lipid polyesters are described.

Environmental adaptation of seed longevity

Seed longevity varies among different species and within the same species. Plant adaptation to ecological niches is, in part, due to genetic changes, acquired through mutation and selection. This constitutes the molecular bases of ecotypes, which are different individual groups of the species adapted to different environments. Seed longevity is an important trait for plant adaptation and survival, but there are other important

seed traits, such as seed dormancy, seed production, seed dispersal and biotic defense. Seed performance is the balance of these traits. Part of plant adaptation to their niches regards in the seed performance to ensure their reproduction and survival (Clerkx *et al.*, 2004*a*; Bentsink *et al.*, 2010). Thus, seed longevity is not in all cases the most determinant seed trait and plants might adapt with reduced seed longevity. Variation in seed longevity among different ecotypes is present in Arabidopsis, and some reduced seed longevity ecotypes present and increased seed dormancy (Nguyen *et al.*, 2012; Mondoni *et al.*, 2014; Nagel *et al.*, 2015).

However, seed longevity is not only genetically, but also environmentally determined. Optimal mother-plant growth conditions are crucial for the seed longevity acquisition and pronounced biotic and abiotic stresses are detrimental. However, in some cases they might be beneficial. For example, a pronounced drought stress is deleterious for seed development if seeds stop developing, but it might be beneficial for seed maturation, as seed may reduce their moisture content at lower levels (Zinsmeister *et al.*, 2020). Nutrient availability is also important for seed longevity and lower nutrient supply negatively impacts seed longevity (He *et al.*, 2014; Zinsmeister *et al.*, 2020). However, plants grown with lower iron availability produced seeds with an increased seed longevity (Murgia *et al.*, 2015). Virus infection might affect seed longevity, either enhancing or reducing it (Bueso *et al.*, 2017), probably by hormonal alterations, and genetic pathways controlling both, seed longevity and biotic defense pathways, which share common molecular components (Righetti *et al.*, 2015), probably providing a balance between biotic defense and seed longevity.

Environmental cues, such as temperature and light, are perceived by the mother plant and the embryo and determine the seed performance. Both, warm temperature, and high light intensity are positive for seed longevity as demonstrated in different studies (He *et al.*, 2014; Mondoni *et al.*, 2014; Nagel *et al.*, 2015). This adaptive effect on seed longevity might be determinant in plant adaptation to dry climates. Plants naturally from cold wet forests have lower seed longevity than plants living in hotter and dryer areas, which usually present an increased seed longevity (Probert *et al.*, 2009; Mondoni *et al.*, 2014). In contrast, seed dormancy determination behaves oppositely to seed longevity plasticity (He *et al.*, 2014) and this permits plant species adaptation in specific niches (Bewley, 1997; Donohue *et al.*, 2005). This effect might be explained with warmer and cooler geographic areas: in cold wet areas, seed deterioration is more unlikely to occur than seed premature germination while in hot dry areas, seed longevity is a more important trait for seed survival than avoiding premature germination. Light incidence and intensity are also different in each geographical area: higher light intensity usually correlates with tropical areas, while low light intensity with polar-close areas. For instance, the ratio of red/far red light determine seed longevity (Contreras *et al.*, 2009), and mutants affected in light perception showed altered seed longevity,

caused by differences in seed coat lipid polyesters (Bueso *et al.*, 2016). In addition, mutants affected in seed coat lipid polyesters present altered seed dormancy (Chai *et al.*, 2016; Fedi *et al.*, 2017), establishing these lipid polyesters in the environmental control of seed aging and seed dormancy, probably by limiting the embryo interaction with the environment.

Hormones are important signaling molecules that control the balance between plant stress defense and plant growth through environmental cues for the sessile plant adaptation. They have been crucial for plant land adaptation, as they emerged in this period (Wang *et al.*, 2015). In addition, hormones have also an important adaptive role in seeds. One important, well-studied role is their involvement in the tight control between seed longevity and seed dormancy controlled by GAs and ABA. The antagonistic role of GAs and ABA has been studied in different plant developmental processes and environmental responses (Liu and Hou, 2018). In seeds, the hormonal balance between GAs and ABA controls seed development, dormancy, and germination (North *et al.*, 2010). During seed development and embryogenesis GAs levels are high while ABA levels are low. GAs are required for the formation of the seed coat (Kim *et al.*, 2005), and GAs content correlates positively with seed longevity (Bueso *et al.*, 2014*b*). Later, during seed maturation, GAs levels decay and ABA accumulates (Yamaguchi *et al.*, 2007). ABA accumulation in seeds leads to seed maturation events driven by ABI3 necessary for seed longevity acquisition and prevent seed germination, as discussed above (Nonogaki, 2014). During germination, ABA levels drop, and GA levels rise permitting germination events and seedling establishment. However, if conditions are not favorable, ABA levels are maintained by ABI5, a ABI3-downstream TF that regulates the inhibition of germination controlling the abundance ABA receptors in a feed-back regulation (Lopez-Molina *et al.*, 2002; Zhao *et al.*, 2020). In addition, in humid seeds ABI5 can reactivate late embryo maturation events, to permit *de novo* seed desiccation, and further seed storage. Seed stratification and seed storage promote GAs, dormancy break and seed germination (Finch-Savage and Leubner-Metzger, 2006). These perceived cues avoid seed germination before cold periods and ensure a sufficient storage time before seed germination. Light perception, an important germination cue, also contribute to regulation of GAs levels (Hedden and Thomas, 2012).

On the other hand, adverse environmental conditions repress GAs and thus germination. ABA, known as the stress hormone, avoids germination until optimal conditions. ABA levels determine seed maturation events and prepare seeds for water-loss. As discussed above, ABA determines seed longevity acquisition through the accumulation of sugars, RFOs, LEA proteins and HSPs (Kotak *et al.*, 2007; Mönke *et al.*, 2012; Sengupta *et al.*, 2015; Leprince *et al.*, 2017). In addition, ABA controls chlorophyll degradation to avoid light induced ROS damage upon storage (Nakajima *et al.*, 2012;

Delmas *et al.*, 2013) and embryo dehydration and ROS movement through aquaporins (Mao and Sun, 2015). Moreover, part of the plant ABA-induced developmental program concerns apoplastic lipid production (Leide *et al.*, 2012), which protects plant tissues from external physical, chemical and biological factors (Kolattukudy, 2001). It protects from freezing, osmotic, salt, oxidative, wounding, and biotic stresses through the permeability regulation of both water and mineral uptake (Franich *et al.*, 1983; Griffith *et al.*, 1985; Kolattukudy and Espelie, 1989; Baxter *et al.*, 2009). During seed maturation, ABA also controls seed coat lipid polyester production (Beisson *et al.*, 2007; Lashbrooke *et al.*, 2016; Gou *et al.*, 2017). Abiotic stress could enhance lipid polyester deposition. Salt induces root suberin deposition (Vishal *et al.*, 2019) and cold induces lipid polyester deposition in the seed coat (Fedi *et al.*, 2017). On the other hand, exogenous GAs and constitutive GAs signaling of quintuple DELLA mutant enhances seed longevity (Bueso *et al.*, 2014*b*). Maybe, the GAs effect in seed longevity regards to the seed development phase and not in the seed maturation phase, due to the antagonistic roles of ABA and GAs. The overexpressing mutants *athb25-1D* and *cog1-2D* present a higher GAs content, produced by the GAs biosynthetic gene GA3OX3, and an enhanced seed deterioration resistance through the development of a reinforced seed coat (Bueso *et al.*, 2016). GAs are essential seed coat development (Kim *et al.*, 2005) and GAs overaccumulation must improve seed coat. This GAs regulation might be determined by environmental factors. In addition, the *rsl1-1D* mutant is deficient in ABA response and presents extended longevity. This mutant overexpresses *RSL1*, which target to degradation the ABA receptors PYL4 and PYR1. Probably the extended longevity of this mutant is due to an increase of GAs response (Bueso *et al.*, 2014*a*; Bueso *et al.*, 2014*c*). Thus, ABA and GAs signaling pathways and their interplay are important during seed development for seed longevity acquisition.

Auxin signaling has also an important role in seed longevity during seed maturation, and exogenous auxin treatment enhances seed longevity (Pellizzaro *et al.*, 2020). However, auxin biosynthetic mutants present enhanced seed longevity and auxin downregulates the seed longevity positive TF HSFA9 (Carranco *et al.*, 2010). The involvement of other hormones in seed longevity must be further investigated. For example, mutants in jasmonic acid and ethylene signaling (*jar1* and *etr1* mutants) presented an enhanced seed longevity (Clerkx *et al.*, 2004*b*). Probably, these traditionally associated biotic stress hormones negatively affect seed longevity to enhance biotic defenses, for example through ROS production (Jeevan Kumar *et al.*, 2015). The absence of their signaling pathways should reinforce seed longevity, perhaps through a reduced ROS production. Common regulators of seed longevity and biotic defense have been described (Righetti *et al.*, 2015) which may regulate this aspect. In the seed coat, the TF TTG1, involved in PAs synthesis and seed dormancy, also regulates JA. Probably this regulation takes part of the environmental determining performance of seed coat (Wasternack and Feussner, 2018; Li *et al.*, 2020).

How to study the seed longevity trait

Aging events during seed dry storage occurs slowly, and seed deterioration, depending on the species, lasts from months to decades. To avoid long aging periods and go on with the research, artificial aging methods have been developed. Those use high temperatures, high humidity and/or high oxygen pressure to damage seed lots and simulate long storage periods. The correlation between artificial methods and dry storage or natural aging has been widely validated (Demir and Mavi, 2008; Rajjou *et al.*, 2008; Bueso *et al.*, 2014*b*; Fenollosa *et al.*, 2020). However, the different environments of seed aging treatments affect differently the seed machinery to prevent aging. Therefore, the choice of the aging treatment may determine the genetic findings of seed longevity studies (Hay *et al.*, 2019). In this thesis, we use three different artificial aging methods to speed up the seed deterioration: the accelerated aging treatment (AAT), the controlled deterioration treatment (CDT) and the elevated partial pressure of oxygen treatment (EPPO). AAT consists in warm water imbibition (39-41 ºC) for 24-48 hours (Bueso *et al.*, 2014*b*; Prieto-Dapena *et al.*, 2006). The CDT, avoids water imbibition, but uses a highly humid atmosphere (75% RH) and high temperature (37 ºC) aging seeds during for 15 days (Basak *et al.*, 2006, International Seed Testing Association, 2018; Hay *et al.*, 2019). EPPO, however, occurs at ambient temperature (20-25 ºC) where seed age for five months in a five-bar oxygen chamber with 40% RH (Groot *et al.*, 2012).

In addition, seed dry storage under ambient storage, or natural aging treatment (NAT), corroborates the results when it has been possible and confirms the seed longevity differences observed. In our laboratory conditions dry seed storage at ambient conditions (20-25 ºC) occurs at 40-60% RH. As commented, for an optimal seed storage the optimal RH is between 19 and 27% (Vertucci and Roos, 1990). Thus, in our NAT conditions aging events are more likely occurring in a rubbery cytoplasm rather than in glassy cytoplasm (Walters *et al.*, 2010). However, we use optimal storage conditions to maintain seeds alive for longer periods (10 ºC and 10% RH). It is important to grow plants simultaneously to avoid growth dependent seed longevity effects, and to include biological replicas. In addition, independently grown seed lots should be used to ensure the consistence of the observed seed-longevity effects. All seed lots need to develop and dry under optimal conditions, but short storage periods may affect the behavior among the different aging treatments. Thus, perfect reproducibility of germination ratios in concrete aging conditions is not always possible (Hay *et al.*, 2019). Due to this effect is mandatory to have control seeds in the assay, preferably from a representative wild-type plant. In Arabidopsis, the reference ecotype Col-0 is widely used for genetic and mutant analysis. NAT is more reproducible due to its slower aging events. Within two years, Col-0 seed germination in our NAT conditions germinate around 50%. In addition, as aging events are progressive, seed germination can be as-

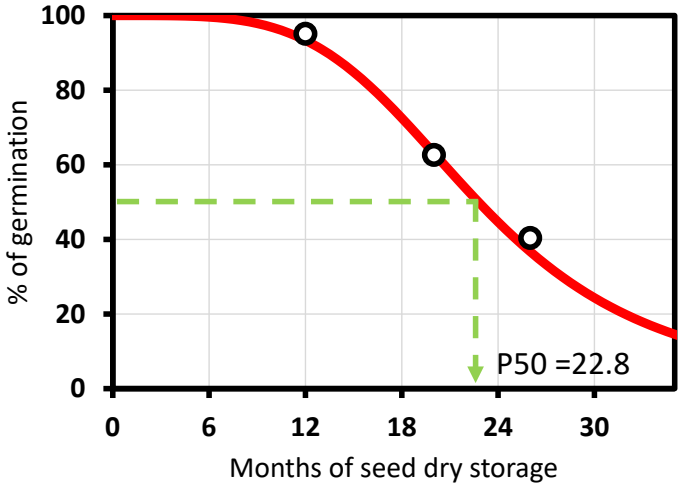


Figure 11: Natural seed aging treatment (NAT) for 0, 12, 20 and 26 months of *Arabidopsis thaliana* Col-0 seeds and the estimated probit curve with the calculated P50 value. The probit curve was calculated with the dcr package (Tausch *et al.*, 2019).

sessed at different time-points. These data usually fit a probit curve (Ellis and Roberts, 1980). From this analysis, the P50 value (time at which seeds should germinate 50%) can be estimated as standard indicative of the seed longevity inherent of the studied seed lot (Figure 11). Nevertheless, differences between cultivars and/or mutant variants present a maximum-difference time on which differences are clearly visible. For variants with decreased seed longevity, shorter storage times and/or weaker seed aging treatments might light up the seed longevity differences with control seeds, while in lines with enhanced seed longevity longer times and/or strong aging treatments are needed.

Genomic molecular techniques

The emergence of new generation sequencing (NGS) techniques has permitted the genome sequencing of thousands of organisms, including plants and different cultivars or ecotypes within a particular specie. Full-sequence information of organisms permits comparative genomic studies, such as evolutionary studies or sequence conservation. Genes are usually conserved due to their importance in different molecular processes. However, gene duplication allows gene differentiation and neofunctionalization and, thus sequence divergence (Zhang, 2003), forming gene families. Genes with the higher sequence similarity within species are called orthologous genes (Fitch, 1970), which should have similar functions among different species. As main molecular processes are similar in different species, the use of model organisms is widely extended in all molecular biology fields. In molecular plant biology, the model organism is *Arabidopsis thaliana*, which presents several advantages such as low genome complexity,

short life period and little growth requirements and it is a self-pollination specie with an important seed production. In addition to these advantages, the extended study on the model organism permits a better characterization of common molecular processes. Then, thanks to these genomic resources, this knowledge can be corroborated and transferred onto other species with agronomical relevance.

In addition, whole-sequence analysis allows genotype-phenotype association studies. The Genome-Wide Association Study (GWAS) is based on the statistical association of single-nucleotide polymorphisms (SNPs) with a particular phenotype among different individuals (cultivars or ecotypes in plants). GWAS is a powerful assay to analyze polygenic traits (Atwell *et al.*, 2010; Korte and Farlow, 2013). Arabidopsis is a worldwide-distributed species, adapted to different niches, and presents enough genomic variance for the GWAS study. More than a thousand different ecotypes have been microarray genotyped and/or full sequenced (Horton *et al.*, 2012; 1001 Genomes Consortium, 2016). GWAS genotype-phenotype association point to causal loci that may be involved in the trait of study, which may be part of the cause of the observed natural variation between accessions, even if they contribute with a little phenotypic effect (Curlis *et al.*, 2017). However, this statistical study based in large-scale datasets need empirical validation due to the high rate of false positives obtained (Ioannidis *et al.*, 2009; Korte and Farlow, 2013). *Arabidopsis thaliana* is an excellent specie for Genome Wide Association Studies as it is a plant with a simple well-annotated genome through numerous molecular studies and present a wide range of natural variation. Moreover, Arabidopsis is a self-pollination specie, permitting the diverse ecotypes to be maintained as inbred lines (Atwell *et al.*, 2010; Horton *et al.*, 2012; Korte and Farlow, 2013). The availability of different large collections of Arabidopsis mutant lines, such as T-DNA insertion lines (O'Malley and Ecker, 2010), permits a rapid empirical validation of GWAS candidates. In addition, mutant collections, such as activation tagging collections with randomly genomic-distributed insertions, have permitted to screen different mutant phenotypes and to determine efficiently gene functions (Gou and Li, 2012). For instance, the mutants *athb25-1D* and *cog1-2D* were obtained through an activation tag mutant screening from W. Sheible and C. Sommerville collection (Bueso *et al.*, 2014*b*; Bueso *et al.*, 2016).

NGS also allows to determine gene expression profiles among different tissues, developmental stages, stress conditions and different mutant individuals through whole transcriptome analysis (RNA-seq), a more precise method comparing to microarray studies (Wang *et al.*, 2009). The differences in gene expression determine a big part of the molecular regulation as it is a crucial step in protein synthesis, and it is tightly regulated by diverse processes. The chromatin structure and DNA packing is regulated by histones, DNA methylation and non-coding RNAs (Holoch and Moazed, 2015). TF regulation controls the expression of genes through the binding to close regulatory se-

quences (Latchman, 1993). RNA degradation is determined by RNA sequence, RNA structure and interaction with proteins. This allows the control of RNA renewal for a more precise gene expression regulation (Nakaminami *et al.*, 2012). As explained above, TF regulation is determinant in seed development, as they are key factors involved in tissue identity of embryo, endosperm, and seed coat, and they promote seed longevity acquisition through different mechanisms. Transcriptome analysis shows differential expressed genes that might be directly or indirectly regulated by TFs when using mutant lines in the TF of study, such as overexpression lines or knock-out (KO) lines. TFs work in cascades of regulation: TFs regulate the expression of other TFs in a complex regulatory network (Jothi *et al.*, 2009). In addition, protein-protein interaction between TFs is important in gene expression regulation (Latchman, 1993). This complex TF regulation results in expression changes of hundreds and even thousands of genes, which may lead to big phenotypic effects. RNA-seq provide direct information on these changes. However, downstream analysis of large differentially expressed gene lists generated by RNA-seq analysis is complicated. Categorical enrichment analysis of differentially expressed genes and other gene relations might help to determine the effect of the TF of study.

Nevertheless, with RNA-seq studies is not possible to know the direct targets of the TF of study. Other techniques are used for this purpose. The Chromatin Immunoprecipitation (ChIP) followed by NGS (ChIP-seq) is a complex technique that allows to determine the direct target of TFs, but also other protein-DNA interaction such as histone-DNA interaction. ChIP-seq consists of covalent-crosslinking the DNA and interacting proteins, breaking the DNA into short length fragments, purification of the protein (and the bound DNA) with a specific antibody, reverse-crosslinking the DNA-protein interaction and purifying the DNA (Kuo and Allis, 1999). Purified DNA is sequenced and compared to a non-enriched sample, which might be identical processed sample with no antibody. This allows to obtain enriched DNA fragments that correspond to genomic areas where the protein interacts with the DNA (Kaufmann *et al.*, 2010). These regions, in the case of TFs, are normally regulatory regions that usually are close to the transcription start site (TSS) of the TF-regulated gene. However, other regions (downstream gene regions or even intergenic regions) are also enriched. Motif enrichment analysis can be assessed to determine the DNA motif recognized by the TF. Other motifs, not recognized by the TF might light up with this approach, because of the presence of an interacting TF co-immunoprecipitated with the TF of study (Li *et al.*, 2016). Analysis of gene expression is crucial to confirm the regulation of the putative target gene by the TF, and to determine is the TF is a positive or a negative regulator of that gene. This analysis can be assessed by RNA-seq or qRT-PCR to dissect the transcriptional regulatory networks (Muhammad *et al.*, 2020).

The study of up-stream regulators of TFs is different. Techniques such as yeast-one-

hybrid (Y1H) can be used to determine the activation of putative regulatory sequences by TFs in the heterologous system of the yeast *Saccharomyces cerevisiae*. For a genomic approach, this assay can be performed with a library of TFs to screen the ones activating the selected regulatory sequence. The study of protein-protein interactions can be assessed with complementary assays, which consist of covalent linking by genetic engineering both proteins to complementary fragments of a reporter protein. The yeast-two-hybrid (Y2H) is the most popular assay, and it can be performed as a library screen like the Y1H. Other techniques are also used for protein-protein interaction studies, as Split-Tpr (similar to Y2H but more suitable for TFs) or BiFC (reconstitute the Green Fluorescent Protein and permit to detect the interaction in plants).

Objectives

The main objective of this PhD thesis is the study and characterization of new molecular mechanisms affecting seed longevity using genomic molecular techniques in the model plant *Arabidopsis thaliana*. For this purpose, two main approaches have been used. The first approach is to find new factors involved in seed longevity through natural variation of this trait among Arabidopsis ecotypes. Natural variation is an important source of variability and might be used to identify molecular mechanisms and genes involved in the trait of interest in a way that might not be possible with other techniques, such as screening of mutant collections. Different aging treatments would be used to identify SNPs associated with seed longevity, and nearby genes that might have a role in this trait. Mutant analysis in those genes might reveal important roles of some of these genes in the seed longevity trait. Confirmed genes would be further characterized as far as possible.

The second main approach to characterize new molecular mechanisms of seed longevity consists on further investigate the regulatory role of AtHB25 and COG1, two TFs that confers extended seed longevity in the overexpression lines *athb25-1D* and *cog1-2D* (Bueso *et al.*, 2014*b*; Bueso *et al.* 2016). Both mutant lines were isolated in the laboratory of Ramón Serrano by Eduardo Bueso and colleagues through an AAT screening using an activating-tag collection. The increased seed longevity observed in both mutants depends on the seed coat, as demonstrated by the maternal heritance, and the lipid polyesters barriers, suberin and cutin, might have an important role on this extended seed longevity. Further characterization of these TFs goes through the genomic techniques ChIP-seq and RNA-seq. Lipid polyester biosynthetic enzymes might be regulated by them, explaining the observed phenotypes. In addition, promoter expression analysis with reporter genes and seed analysis techniques will be used to determine the endogenous role of AtHB25 and COG1 in seed coat development and seed longevity.

2. Chapters

This PhD thesis is a recapitulation of three research articles and unpublished results. Chapters are the following:

Chapter 1

Plant, Cell & Environment (2020) **43**:2523–2539 DOI: 10.1111/pce.13822

Identification of novel seed longevity genes related to oxidative stress and seed coat by genome-wide association studies and reverse genetics

Joan Renard, Regina Niñoles, Irene Martínez-Almonacid, Beatriz Gayubas, Rubén Mateos-Fernández, Gaetano Bissoli, Eduardo Bueso, Ramón Serrano* and José Gadea

Chapter 2

New Phytologist (2021) **231**:679–694 DOI: 10.1111/nph.17399

Apoplastic lipid barriers regulated by conserved homeobox transcription factors extend seed longevity in multiple plant species

Joan Renard, Irene Martínez-Almonacid, Indira Queralta Castillo, Annika Sonntag, Aseel Hashim, Gaetano Bissoli, Laura Campos, Jesús Muñoz-Bertomeu, Regina Niñoles, Thomas Roach, Susana Sánchez-León, Carmen V. Ozuna, José Gadea, Purificación Lisón, Ilse Kranner, Francisco Barro, Ramón Serrano, Isabel Molina and Eduardo Bueso

Chapter 3

Plant, Cell & Environment (2020) **43**:315–326 DOI: 10.1111/pce.13656

PRX2 and PRX25, peroxidases regulated by COG1, are involved in seed longevity in Arabidopsis

Joan Renard, Irene Martínez-Almonacid, Annika Sonntag, Isabel Molina, José Moya-Cuevas, Gaetano Bissoli, Jesús Muñoz-Bertomeu, Isabel Faus, Regina Niñoles, Jun Shigeto, Yuji Tsutsumi, José Gadea, Ramón Serrano and Eduardo Bueso

Chapter 4

Work in progress and unpublished data

Environmental regulation of seed longevity is mediated by AtHB25 and COG1 through temperature and light cues

Joan Renard, Gaetano Bissoli, Regina Niñoles, Irene Martínez-Almonacid, Luis Oñate-Sánchez, José Gadea, Ramón Serrano and Eduardo Bueso

Chapter 1

Plant, Cell & Environment (2020) **43**:2523–2539 DOI: 10.1111/pce.13822

Identification of novel seed longevity genes related to oxidative stress and seed coat by genome-wide association studies and reverse genetics

Joan Renard, Regina Niñoles, Irene Martínez-Almonacid, Beatriz Gayubas, Rubén Mateos-Fernández, Gaetano Bissoli, Eduardo Bueso, Ramón Serrano* and José Gadea*

Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universitat Politècnica de València-C.S.I.C., Camino de Vera, 46022, València, Spain

*corresponding authors (e-mail: jgadeav@ibmcp.upv.es and rserrano@ibmcp.upv.es)

Abstract

Seed longevity is a polygenic trait of relevance for agriculture and for understanding the effect of environment on the aging of biological systems. In order to identify novel longevity genes, we have phenotyped the natural variation of 270 ecotypes of the model plant *Arabidopsis thaliana* for natural aging and for three accelerated aging methods. Genome-wide analysis, using publicly available single-nucleotide polymorphisms (SNPs) data sets, identified multiple genomic regions associated with variation in seed longevity. Reverse genetics of twenty candidate genes in Columbia ecotype resulted in seven genes positive for seed longevity (*PSAD1, SSLEA, SSTPR, DHAR1, CYP86A8, MYB47* and *SPCH*) and five negative ones (*RBOHD, RBOHE, RBOHF, KNAT7* and *SEP3*). In this uniform genetic background, natural and accelerated aging methods provided similar results for seed-longevity in knock-out mutants. The NADPH oxidases (RBOHs), the dehydroascorbate reductase (DHAR1) and the photosystem I subunit (PSAD1) highlight the important role of oxidative stress on seed aging. The cytochrome P-450 hydroxylase CYP86A8 and the transcription factors MYB47, KNAT7 and SEP3 support the protecting role of the seed coat during seed aging.

Introduction

Imperfections of metabolism result in accumulation of deleterious and damaged molecules that age cells and organisms (Gladyshev, 2014; López-Otín *et al.*, 2016). According to the oxidative stress theory of aging, a major metabolic reaction contributing to aging is the production of reactive oxygen species (ROS) (Harman, 1956). Current views (López-Otín *et al.*, 2016) consider that aging is controlled by metabolism in a more general way than just ROS production. However, it has been argued that oxidative stress contributes to all the metabolic and genomic alterations associated with aging (Schmidlin *et al.*, 2019).

One model for the study of biological aging is the seed of vascular plants. Aging is slowed down in desiccated forms of organisms, where metabolism is reduced because of high viscosity and slow diffusion inside cells. Accordingly, desiccation is the mechanism developed by many life forms to decrease metabolism and endure for long times (Wood and Jenks, 2007). Most organisms die upon desiccation and therefore acquisition of desiccation tolerance is a prerequisite for this longevity mechanism. In the case of plants, it has been suggested that the evolution of desiccation tolerance was a crucial step in the colonization of land by primitive plants living in fresh water (Oliver *et al.*, 2000). Nowadays, very few plants (the so called resurrection plants) are desiccation tolerant in their vegetative form. However, orthodox seeds are tolerant to desiccation, a feature acquired during their late maturation phase. This tolerance results from the accumulation of osmolytes (sucrose and raffinose family oligosaccharides, RFO), late embryogenesis abundant proteins (LEA) and small heat shock proteins (Sano *et al.*, 2016; Leprince *et al.*, 2017). Under these conditions, cytoplasm enters an immobilized glassy state that protects macromolecules and cellular membranes during desiccation (Hoekstra *et al.*, 2001; Ballesteros and Walters, 2011; Sano *et al.*, 2016; Leprince *et al.*, 2017). This desiccated state of seeds permits to keep the plant alive for long periods of time, until environmental conditions are favourable. However, dry seeds also experience deterioration and aging and it has been proposed that oxidative damage is a major cause reducing longevity in this system (Harman and Mattick, 1976; Bailly, 2004; Sano *et al.*, 2016; Nagel *et al.*, 2019).

High temperature and humidity are deleterious for seeds because they break the glassy state and promote metabolism, and therefore ROS and cellular damage. Equilibrating mature seeds between 19 and 27% relative environmental humidity (0.04-0.06 g H_2O / g dry weight in seeds) provides the optimum moisture level for reducing metabolism and maintaining seed longevity during long-term storage (Vertucci and Roos, 1990). Seed deterioration occurs under ambient storage or natural conditions (about 20-25 ºC, 40-60 % relative humidity (RH), 0.08-0.10 g H_2O / g dry weight), where dried seeds still exhibit some metabolism.

Analysis of Quantitative Trait Loci (QTLs) has demonstrated the polygenic nature of seed longevity in different plants such as rice (Jiang *et al.*, 2011), wheat (Landjeva *et al.*, 2010), lettuce (Schwember *et al.*, 2010), tobacco (Agacka-Mołdoch *et al.*, 2015) and barley (Nagel *et al.*, 2016). Several genes important for seed longevity in the model plant *Arabidopsis thaliana* have been identified by reverse genetics and compiled by Sano *et al.* (2016). They refer to three biological functions: anti-oxidative defence, repair of damaged proteins and nucleic acids, and development of seed coat. Some important aspects that need to be clarified include the nature of the ROS producing systems and of antioxidant defences during seed storage, the protective mechanisms of the seed coat and the complete gene networks controlling seed longevity. Many antioxidant defences such as enzymes and chemicals have been proposed to participate in seed longevity but only vitamin E (tocopherol) has been demonstrated to be important by mutant studies (Sattler *et al.*, 2004). Also, the seed coat has been proposed to be important, specially the endothelium flavonoids (proanthocyanidins; Clerkx *et al.* 2004*b*) and the palisade suberin (Bueso *et al.*, 2014*b*; Bueso *et al.*, 2016, Renard *et al.*, 2020*a*) but the transcription factors and enzymes involved in the development of this structure (Haughn and Chaudhury, 2005) are not completely known.

In the present work, we took advantage of natural variation in seed longevity between accessions (ecotypes) of *Arabidopsis thaliana* with well-characterized genomes to perform a Genome Wide Association Study (GWAS) of this trait. GWAS analysis is a powerful tool to dissect polygenic traits (Atwell *et al.*, 2010; Korte and Farlow, 2013) and *Arabidopsis* is an excellent organism for this purpose (Horton *et al.*, 2012; Weigel, 2012). GWAS consists on the association of the molecular genetic information (Single Nucleotide Polymorphisms, SNPs) with the observed variation in a given trait through statistical computation. These associations point to causal loci that may be involved in the trait and may explain part of the natural variation between accessions (Curtin *et al.*, 2017). However, for each candidate gene in these loci empirical validation is required due to the high number of false positives with this genomic approach (Ioannidis *et al.*, 2009; Korte and Farlow, 2013). Arabidopsis is the most studied plant species at the molecular level and a large collection of sequence-indexed T-DNA insertion mutants is available (O'Malley and Ecker, 2010). This valuable resource allows validation of GWAS candidate genes by testing the phenotype of the corresponding loss-of-function mutants in the same cultivar background.

As the study of seed longevity under natural, ambient storage conditions requires long times for aging, several accelerated aging procedures have been utilized in this field. The quickest is the accelerated aging test (AAT), which consists in a high temperature treatment of water imbibed seeds during days (Prieto-Dapena *et al.*, 2006; Bueso *et al.*, 2014*b*). In the controlled deterioration test (CDT) aging of dry seeds occurs in weeks with a high humidity and high temperature atmosphere. Finally, to target the

oxygen-produced damage, there is the elevated partial pressure of oxygen test (EPPO), in which dry seeds are stored for months in a low humidity and high-pressured oxygen atmosphere (ISTA, 2018; Hay *et al.*, 2019; Groot *et al.*, 2012). Although in some studies with few genotypes there is a similar behaviour in both natural and artificial aging (Rajjou *et al.*, 2008; Bueso *et al.*, 2014*b*), each aging procedure might affect different physiological and molecular aspects of the seed deterioration-resistance trait. Accordingly, to increase the list of candidate genes, we have performed natural aging and different artificial aging treatments.

Our results suggest that the different tests used to evaluate seed longevity are highlighting different aspects of the machinery deployed by the seed to avoid deterioration. However, when seed longevity was determined in mutants from the same genetic background, natural and artificial aging procedures gave similar results. Validation of candidate genes resulted in several novel participants in seed longevity such as ROS-producing NADPH oxidases (RBOHs), antioxidant dehydroascorbate reductase and, related to seed coat development, a P450 cytochrome hydroxylase and several transcription factors (TFs).

Results

CDT correlates best with natural aging

To elucidate new genetic components involved in seed longevity, a core-set of 360 Arabidopsis ecotypes was obtained from NASC to perform GWAS. All accessions were grown on soil under long-day conditions, and mature seeds where collected from all accessions that were able to flower and develop seeds in a time-window not wider than 160 days. 270 natural accessions were selected for aging treatments after the control (non-aged) germination test. Four different aging methods were performed: Natural Aging Treatment (NAT) and three artificial aging treatments (AAT, CDT and EPPO). Aged-seeds germination data were corrected using control germination data (Supplemental data 1). A wide range of variability was observed among all the accessions in each of the different assays (Figure 12a-d). The accession Col-0 (Columbia), used as reference to set up all treatments, varied from 40% to 60% of germination in all treatments.

To estimate how each artificial aging treatment resembles natural aging, Spearman correlations were calculated (Figure 13). CDT is the artificial aging treatment correlating best with NAT ($r_s = 0.68$), followed by EPPO ($r_s = 0.47$). AAT does not correlate with NAT in this study ($r_s = 0.11$). AAT is the quickest (days) artificial aging method used and at variance with other treatments, it involves immersion of seeds into water.

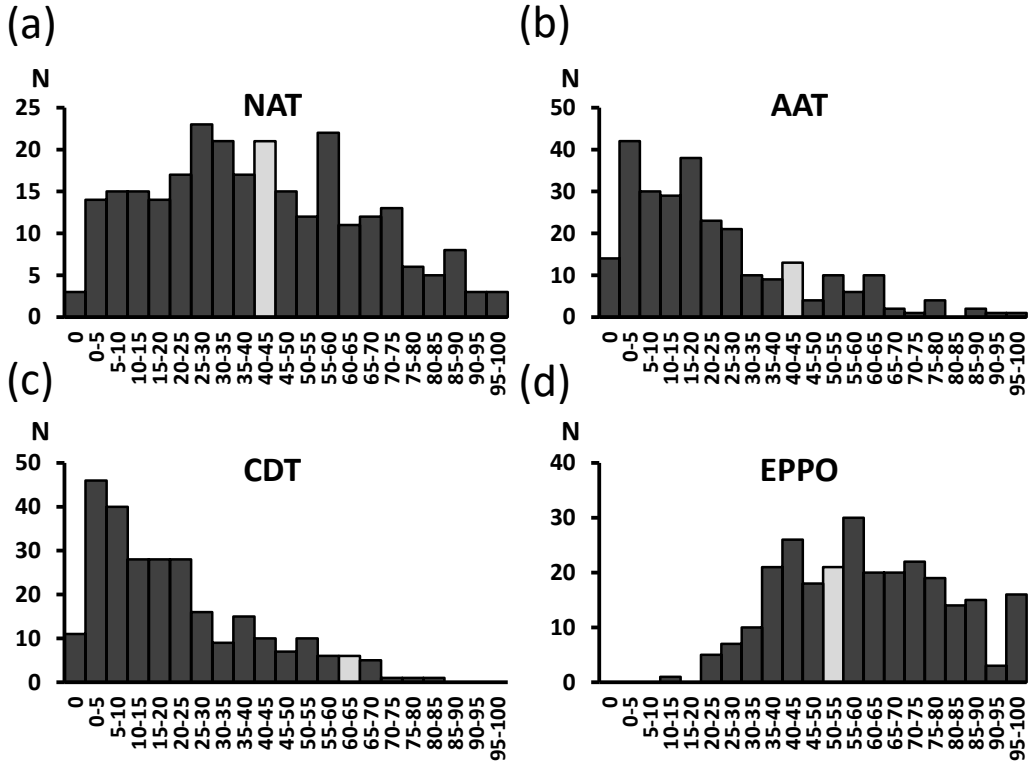


Figure 12: Arabidopsis ecotypes present a wide range of variability in seed deterioration resistance among all four aging treatments. (a) Histogram distribution of 270 ecotypes after NAT. (b) AAT histogram. (c):CDT histogram. (d) EPPO histogram. Germination data of each treatment is corrected by control germination of each ecotype. Histograms represent the number of ecotypes (y-axis) in the indicated range of germination (x-axis). The bar in light grey corresponds to the range including the reference accession Col-0.experiments with 100 seeds per line. Non-treated seed from all lines germinated more than 99% after 3 days. The error bars denote standard errors. *, $p < 0.05$.

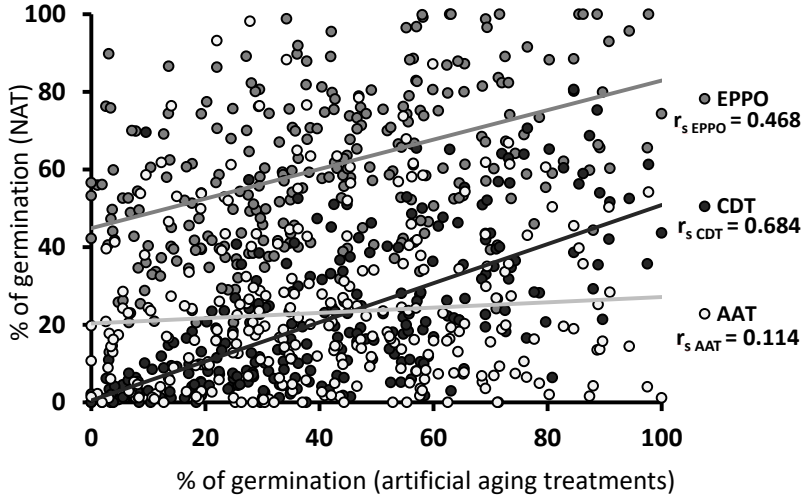


Figure 13: CDT correlates better to NAT. Dispersion plots of EPPO, CDT and AAT (y-axis) and NAT (x-axis). Regression lines are plotted and Spearman correlations are indicated for each artificial aging correlation with NAT.

Table 1: Correlation indexes of different phenotypic traits with seed longevity determined with different aging treatments (NAT, AAT, CDT and EPPO).

	NAT	*AAT*	*CDT*	*EPPO*	*Ecotypes*	*Reference*
Plant life span	**-0.35**	**-0.33**	**-0.30**	-0.22	270	Seed collection time
Days to flower	-0.28	**-0.35**	-0.29	**-0.31**	71	Grimm *et al.* (2017)
Rossete leaves number	**-0.40**	**-0.37**	**-0.36**	**-0.32**	71	Grimm *et al.* (2017)
Fe content in leaves	**-0.32**	0.14	**-0.44**	**-0.30**	51	Atwell *et al.* (2010)
Latitude at collection point	-0.11	0.00	0.03	0.10	256	Horton *et al.* (2012)
Seed dormancy	0.09	-0.28	0.00	-0.09	48	Atwell *et al.* (2010)
Anthocyanin content	0.13	-0.01	-0.07	-0.09	83	Atwell *et al.* (2010)

Indicated correlation indexes correspond to Spearman correlations. Ecotypes column refers to the number of coinciding ecotypes in the correlations. Blood letters indicate correlation indexes above 0.3 and below -0.3. Correlated phenotypes are listed in the Supplemental data 2. Upper part: correlated traits. Down part: non-correlated traits.

We next calculated correlations with public datasets and ecotype collection data. Although the coincident ecotypes with our collection was small in most cases (48-83 ecotypes), an inverse correlation between seed longevity and flowering time, rosette leave number and plant life span could be observed (Table **1**). These three traits are linked and suggest the importance of early flowering for seed longevity. There is also an inverse correlation of iron content in leaves with NAT, CDT and EPPO. No correlation was found with seed dormancy, latitude of collection point or anthocyanin content, among others. The complete list of correlated phenotypes with at least one seed aging treatment is in Supplemental data 2.

The different seed aging treatments highlight multiple genomic areas and candidate genes.

The corrected germination data obtained after each aging treatment (Supplemental data 1) was used to perform four different GWAS analysis, one for each seed aging treatment. Results can be visualized at https://gwas.gmi.oeaw.ac.at/#/study/5406/overview. We established a relaxed *p* value ($< 10^{-5}$) threshold to consider putative longevity-related genomic areas (Figure 14). Genes enclosed within the 1.5 Kb-region spanning significant SNPs were listed (Supplemental data 3). Natural aging (NAT) highlighted 99 genes, the AAT treatment, 280 genes, the CDT, 130 genes and the EPPO, 337 genes. We further filtered for candidate genes following three criteria: (a) biological function based on the state-of-knowledge (Sano *et al.*, 2016); (b) gene expression during seed development (Le *et al.*, 2010); and (c) distance to highly scored SNPs. Occasionally, a few genes with *p* value slightly above 10^{-5} were considered for reverse genetic analysis. In some cases, gene family members from candidate genes were also selected. For validation and further experiments, double and triple mutants

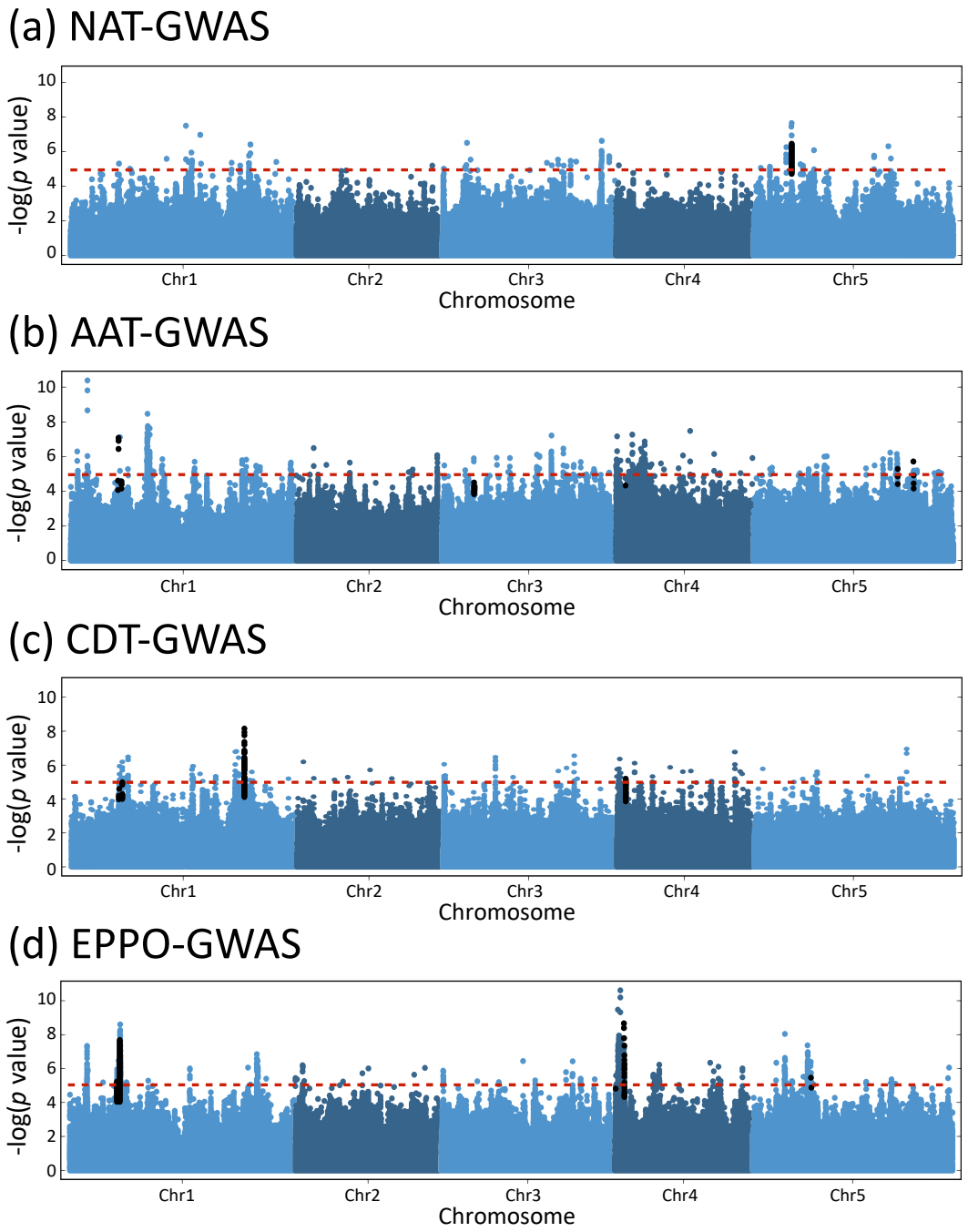


Figure 14: Different seed longevity GWAS highlight different and similar genomic areas. (a) NAT-GWAS Manhattan plot. (b) AAT-GWAS Manhattan plot. (c) CDT-GWAS Manhattan plot. (d) EPPO-GWAS Manhattan plot. Dashed lines indicate the threshold significance ($p < 10^{-5}$). In black, SNPs showing $p < 10^{-4}$ that are at less than 1.5 Kb distance of genes used for reverse genetics.

were obtained. Table 2 shows the list of candidate genes selected for reverse genetic analysis in this study.

Table 2: Highlighted genes used for reverse genetics analysis.

GWAS	*Score*	*AGI*	*Gene alias*
EPPO, CDT	8.51	*AT4G02750*	*SSTPR*
EPPO, AAT, CDT	8.24	*AT4G02770*	*PSAD1*
CDT	8.14	*AT1G62990*	*KNAT7*
EPPO, CDT	7.49	*AT1G19570*	*DHAR1*
AAT, EPPO, CDT	6.99	*AT1G18710*	*MYB47*
NAT	6.26	*AT5G15800*	*SEP1*
AAT	5.63	*AT5G53210*	*SPCH*
EPPO	5.54	*AT5G23190*	*CYP86B1*
AAT	5.28	*AT5G47910*	*RBOHD*
EPPO, AAT	*4.85*	*AT1G19230*	*RBOHE*
EPPO	*4.75*	*AT4G00360*	*CYP86A2*
AAT	*4.41*	*AT3G17520*	*SSLEA*

GWAS column: aging treatments where the indicated genes appeared at less than 1.5 Kb distance of SNPs with *p* value $\leq 10^{-4}$; Score column or –log(*p* value): score of most significant SNP close to the indicated gene in first-listed GWAS; AGI column: Arabidopsis gene identifier; Gene alias: gene name used in this paper. Lower part of table correspond to genes whose SNPs were not significant in GWAS (score between 4 and 5) but were used for reverse genetics.

T-DNA insertion mutants identify novel genes involved in seed longevity

Three different seed lots of two different plant generations from every mutant line were aged with two artificial treatments (AAT and CDT; Table 3) and with natural conditions (NAT). Figure 16 describes three seed aging sensitive mutants: a Late Embryogenesis Abundant (LEA) protein (AT3G17520), a subunit of Photosystem I (PSAD1), and a tetratricopeptide repeat (TPR) protein (AT4G02750). The *LEA* and *TPR* genes were named *SSLEA* and *SSTPR*, with *SS* meaning Seed Storability. *PSAD1* and *SSTPR* were highlighted by two and three artificial aging treatments, respectively (Table 2). After 18 months of natural aging, all three mutant seeds present a drastic reduction of germination. *PSAD1* and *LEA* are expressed in mature embryos, while *SSTPR* is expressed in developing embryos (Figure 15). As indicated in Table 3, mutants in these three genes also exhibited reduced seed longevity with two artificial aging treatments (AAT and CDT). Mutant *sstpr* seeds present an increased tetrazolium reduction (Figure 17) indicating greater seed coat permeability.

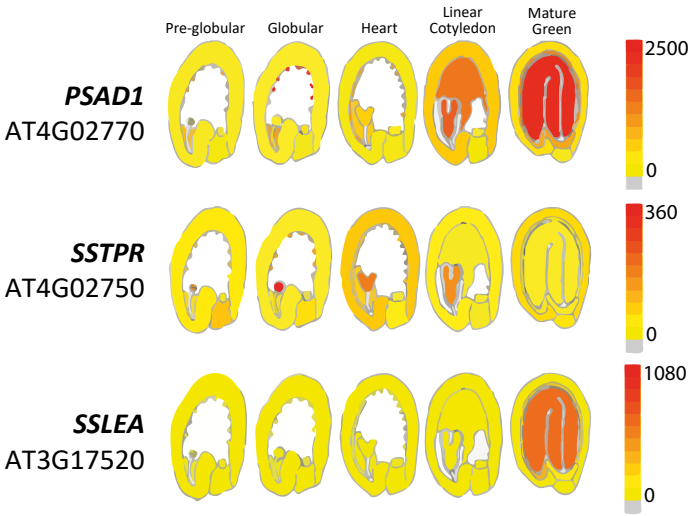


Figure 15: *In silico* expression analysis of *PSAD1*, *SSTPR* and *SSLEA* during seed development based in Le *et al.* (2010) data. Images were obtained with the eFP browser (Winter *et al.*, 2007).

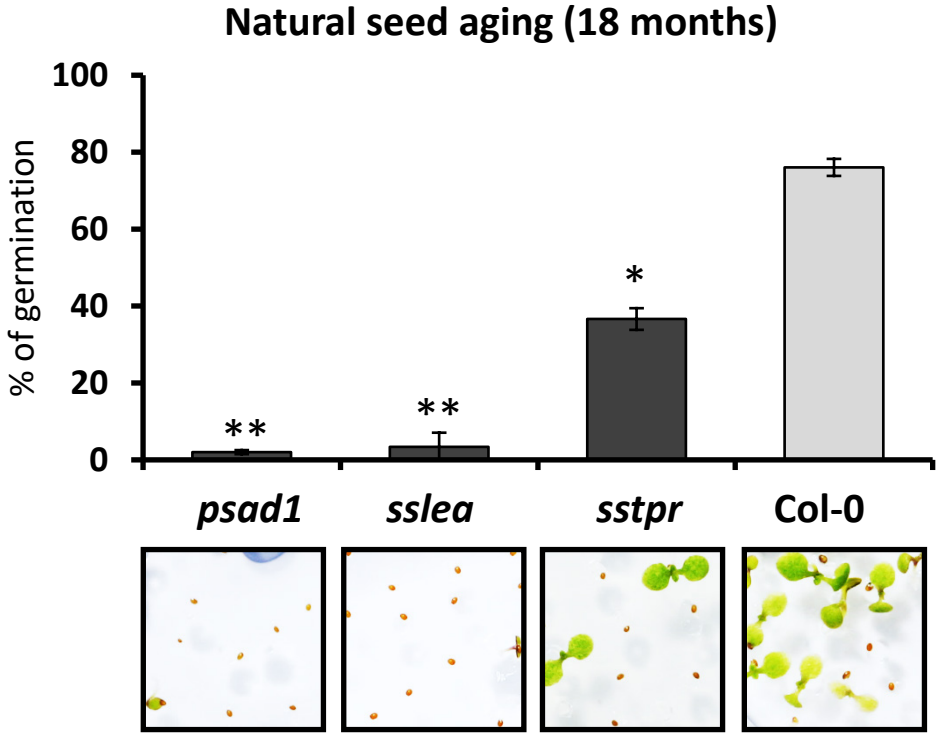


Figure 16: Seeds of *psad1*, *sslea* and *sstrp* mutant exhibit reduced seed longevity. Seed lots were stored for 18 months under natural conditions (see Materials and Methods) and sown on MS plates. Above: The percentage of germination was recorded after one week. The results are the average of three experiments with 50 seeds per line. Bars indicate standard errors. Not-aged seeds from all lines germinated more than 95%. *Significantly differing from wild type seeds at $p < 0.05$ (Student´s *t* test). ** Significantly differing from wild type seeds at $p < 0.01$ (Student´s *t* test). Below: representative images.

Table 3: Schematic results of aging treatments (NAT, AAT and CDT) of loss-of-function mutants *psad1*, *sslea*, *sstpr* in Col-0 background

Mutant line	*NAT*	*AAT*	*CDT*
psad1	Susceptible**	Susceptible*	Susceptible**
sslea	Susceptible**	Susceptible*	Susceptible*
sstpr	Susceptible**	Susceptible*	Susceptible*

Reduced seed germination (Susceptible), in comparison to Col-0 seeds. *Significantly differing from wild type seeds at $p < 0.05$ (Student´s *t* test). **Significantly differing from wild type seeds at $p < 0.01$ (Student´s *t* test).

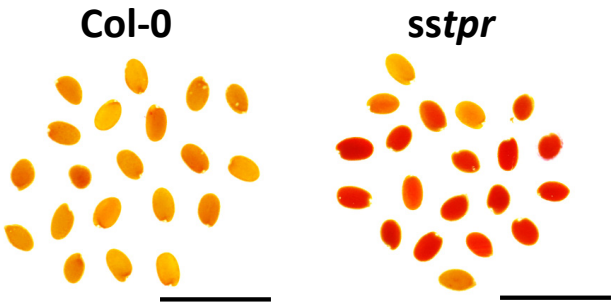


Figure 17: Mutant seeds of *sstpr* present an increased tetrazolium reduction rate, and thus putatively an increased seed coat permeability, compared to Col-0 seeds. Tetrazolium salt assay performed for 24 h at 28°C. Scale bar: 2 mm.

ROS accumulation by NADPH oxidases (RBOHs) is detrimental for seed longevity

Interestingly, two RBOHs genes, *RBOHD* (AT5G47910) and *RBOHE* (AT1G19230) were highlighted after AAT and EPPO treatments, respectively. RBOHs are transmembrane NADPH oxidases important for ROS production. They produce superoxide anion (O_2^-), rapidly transformed to H_2O_2, and play important roles in biotic and abiotic stresses (Chang *et al.*, 2016; Qu *et al.*, 2017). We decided to study them given the importance of ROS in seed aging (Bailly, 2004; Sano *et al.*, 2016). There are ten *RBOH* genes in Arabidopsis named from A to J. Only three of them are abundantly expressed during seed development according to Le *et al.* (2010): *RBOHE* and *RBOHF*, mainly in the seed coat, and *ROBHD*, in the endosperm (Figure 18). Seeds of *rbohd, rbohe* and *rbohf* mutant plants exhibited increased longevity and the double mutant *rbohd,f* has even more seed longevity (Figure 19). RBOH-mutant seeds showed no significant seed coat alterations (Figure 20). As indicated in Table 4, these four mutants also exhibited increased seed longevity with two artificial aging treatments (AAT and CDT).

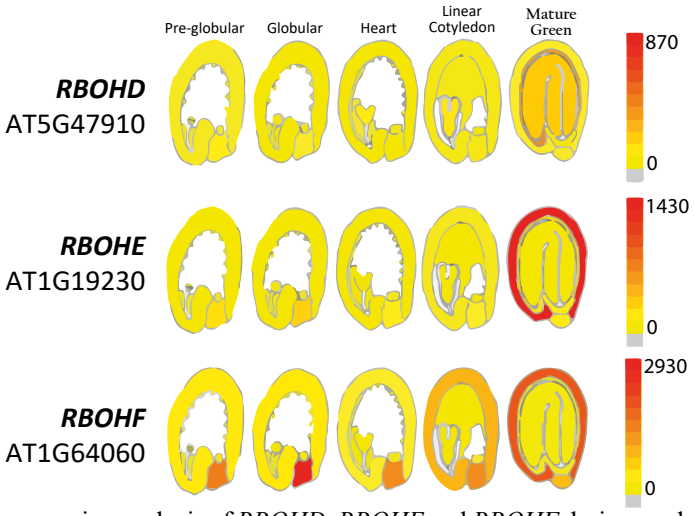


Figure 18: *In silico* expression analysis of *RBOHD*, *RBOHE* and *RBOHF* during seed development based in Le *et al.* (2010) data. Images were obtained with the eFP browser (Winter *et al.*, 2007).

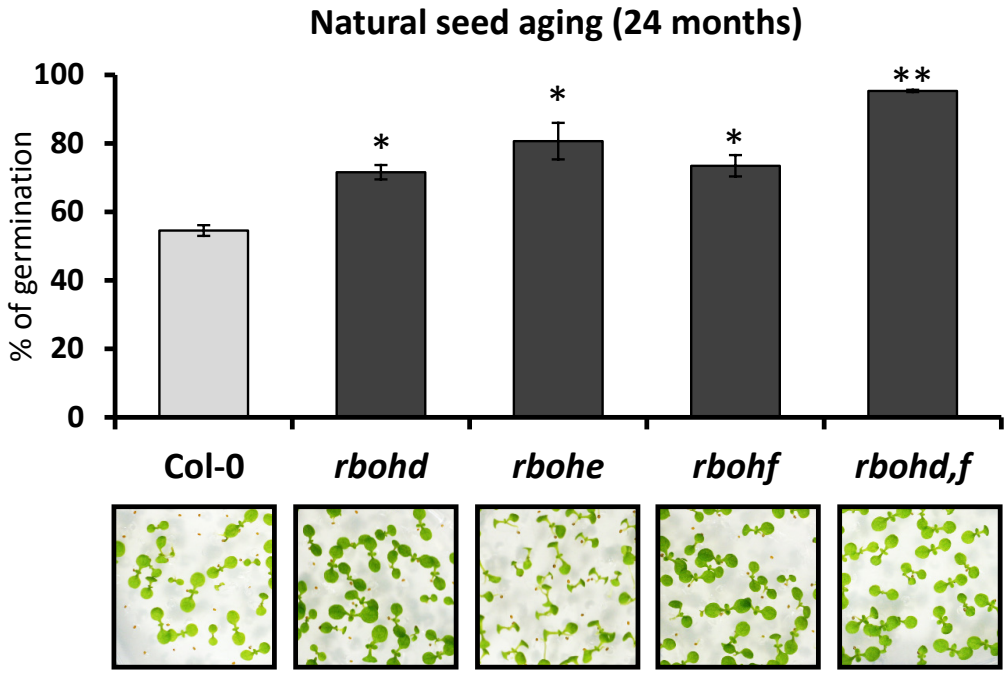


Figure 19: Seeds of *rbohd*, *rbohe*, *rbohf* and double mutant *rbohd,f* exhibit increased seed longevity. Seed lots were stored for 24 months and sown on MS plates. Above: The percentage of germination was recorded after one week. The results are the average of three experiments with 50 seeds per line. Bars indicate standard errors. Not-aged seeds from all lines germinated more than 97%. *Significantly differing from wild type seeds at $p < 0.05$ (Student´s *t* test). **Significantly differing from wild type seeds at $p < 0.01$ (Student´s *t* test). Below: representative images.

Table 4: Schematic results of aging treatments (NAT, AAT and CDT) of loss-of-function mutants *rbohd, rbohe, rbohf* and double mutant *rbohd,f* in Col-0 background.

Mutant line	*NAT*	*AAT*	*CDT*
rbohd	Resistant*	Resistant*	Resistant*
rbohe	Resistant*	Resistant*	Resistant**
rbohf	Resistant*	Resistant*	Resistant*
rbohd,f	Resistant**	n.s.	Resistant**

Increased seed germination (Resistant), in comparison to Col-0 seeds. *Significantly differing from wild type seeds at $p < 0.05$ (Student´s *t* test). **Significantly differing from wild type seeds at $p < 0.01$ (Student´s *t* test). n.s.: non significant.

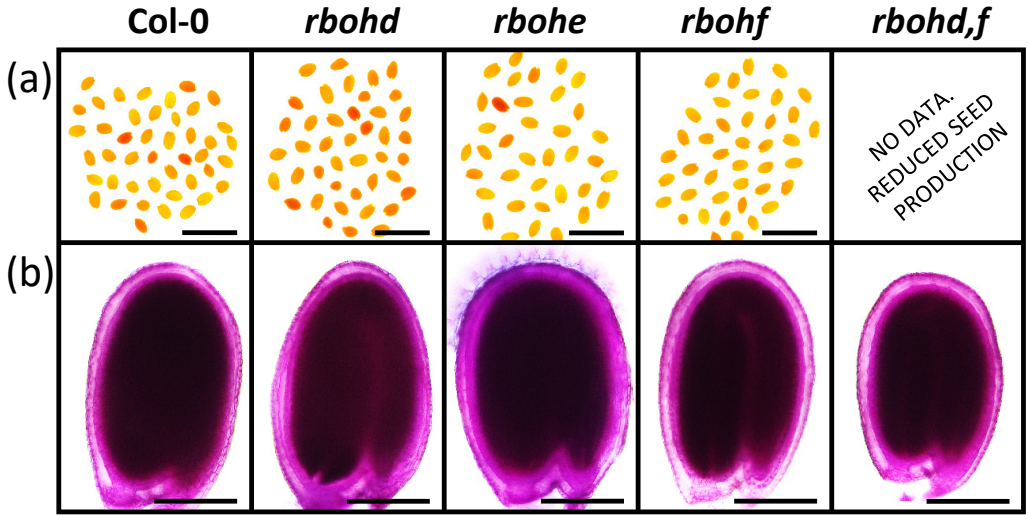


Figure 20: Seed permeability and lipid polyester staining showed no significant differences in *rbohd*, *rbohe*, *rbohf* and double mutant *rbohd,f* mutant seeds. (a) Tetrazolium salt reduction assay for 24 h. Seeds of *rbohd* mutant are lightly more permeable than Col-0 seeds, while rbohf seed are lightly less permeable. However, there is not observed an overall behavior in all rboh mutant seeds. Scale bars: 2 mm. (b) Sudan Red staining did not show differences in lipid polyesters barriers that could explain their enhanced seed longevity. Scale bars: 200 µm.

DHAR1 ROS-detoxification system is important for seed longevity

The *DHAR1* (AT1G19570) gene was highlighted after the EPPO and CDT treatments. DHARs are glutathione-dependent dehydroascorbate reductases involved in ROS detoxification. They catalyse the regeneration of ascorbate oxidized during detoxification of H_2O_2 by ascorbate peroxidase (Foyer and Noctor, 2011; Smirnoff, 2011). Given the importance of ROS in seed longevity we decided to study the implication of this gene. Three functional isoforms are described in Arabidopsis (Dixon *et al.*, 2002; Dixon and Edwards, 2010). We tested all three single mutants and the triple mutant *dhar1,2,3*

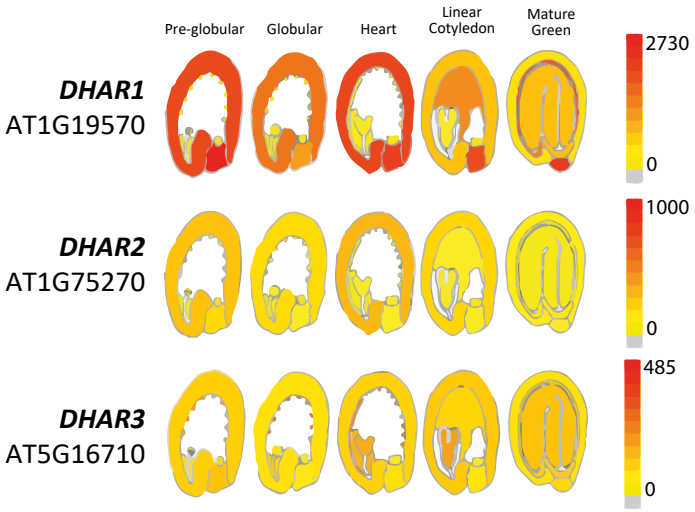


Figure 21: *In silico* expression analysis of *DHAR1*, *DHAR2* and *DHAR3* during seed development based in Le *et al.* (2010) data. Images were obtained with the eFP browser (Winter *et al.*, 2007).

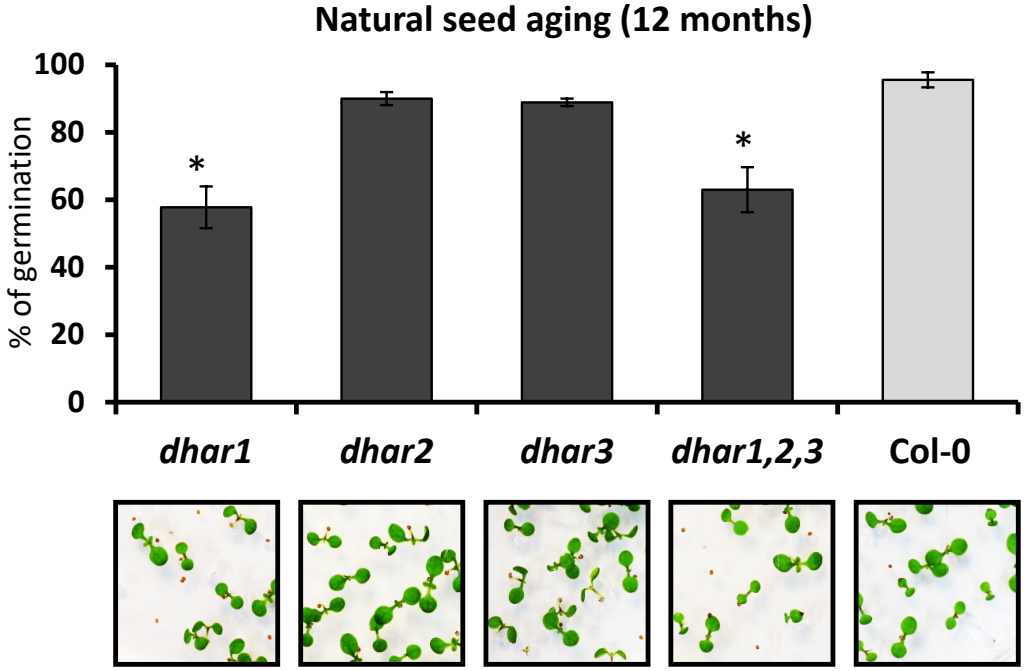


Figure 22: Seeds of *dhar1* and triple mutant *dhar1,2,3* present a reduced seed longevity, but *dhar2* and *dhar3* seeds no phenotype. Seed lots were stored for 12 months and sown on MS plates. Above: The percentage of germination was recorded after one week. The results are the average of three experiments with 50 seeds per line. Bars indicate standard errors. Not-aged seeds from all lines germinated more than 99%. *Significantly differing from wild type seeds at $p < 0.05$ (Student´s *t* test). Below: representative images.

Table 5: Schematic results of aging treatments (NAT, AAT and CDT) of loss-of-function mutants *dhar1, dahr2, dhar3* and the triple mutant *dhar1,2,*3 in Col-0 background.

Mutant line	*NAT*	*AAT*	*CDT*
dhar1	Susceptible**	Susceptible*	Susceptible**
dhar2	n.s.	n.s.	n.s.
dhar3	n.s.	n.s.	Susceptible**
dhar1,2,3	Susceptible*	Susceptible*	Susceptible*

Reduced seed germination (Susceptible), in comparison to Col-0 seeds. *Significantly differing from wild type seeds at $p < 0.05$ (Student´s *t* test). **Significantly differing from wild type seeds at $p < 0.01$ (Student´s *t* test). n.s.: non significant.

to clarify if there is a major player contributing to seed longevity.After one year of NAT treatment, *dhar1* seeds already lost part of their germination ability, while *dhar2* and *dhar3* mutant seeds did not. The triple mutant *dhar1,2,3* exhibited the same seed longevity reduction as the single *dhar1* mutant, indicating that other isoforms are not needed for seed longevity (Figure 22). According to Le *et al.* (2010) dataset, *DHAR1* is the most abundantly expressed *DHAR* isoform during seed development and in mature seed (Figure 21). This expression pattern is consistent with our natural aging results. As indicated in Table 5, *dhar1* and *dhar1,2,3* mutants also exhibited reduced seed longevity with both artificial aging treatments (AAT and CDT).

CYP86A8 is involved in synthesis of seed lipid-polyesters needed for seed longevity

GWAS highlights genomic areas causing the observed natural variation of traits (Curtin *et al.*, 2017). Natural variation, however, may affect less to major player genes due to their important physiological role. Here we present the example of a gene family, whose two members where highlighted after the aging treatments, although a third member resulted to be more important for seed longevity in a mutant background. After the EPPO treatment, *CYP86A2* (AT4G00360) and *CYP86B1* (AT5G23190) were highlighted. They are related to suberin and cutin biosynthesis (Bak *et al.*, 2011), lipid polymers involved in seed permeability and recent results suggest an important role in seed longevity (Renard, 2020*a*). We obtained T-DNA mutant lines affecting members of the *CYP86* family abundantly expressed in seeds accordingly to Le *et al.* (2010) data (Figure 23): *CYP86A1* (AT5G58860), *CYP86A2* (AT4G00360), *CYP86A8* (AT2G45970), *CYP86B1* (AT5G23190), *CYP86B2* (*AT5G08250*) and *CYP86C1* (*AT1G24540*). A T-DNA insertion line close to *CYP86C3* was not available. After one year of seed dry storage, only seeds of *cyp86a8* showed a drastic reduction in their germination (Figure 24a), also observed during AAT and CDT treatments (Table 6). Interestingly, *cyp86a8* seeds are notably different to wild type: they were smaller and

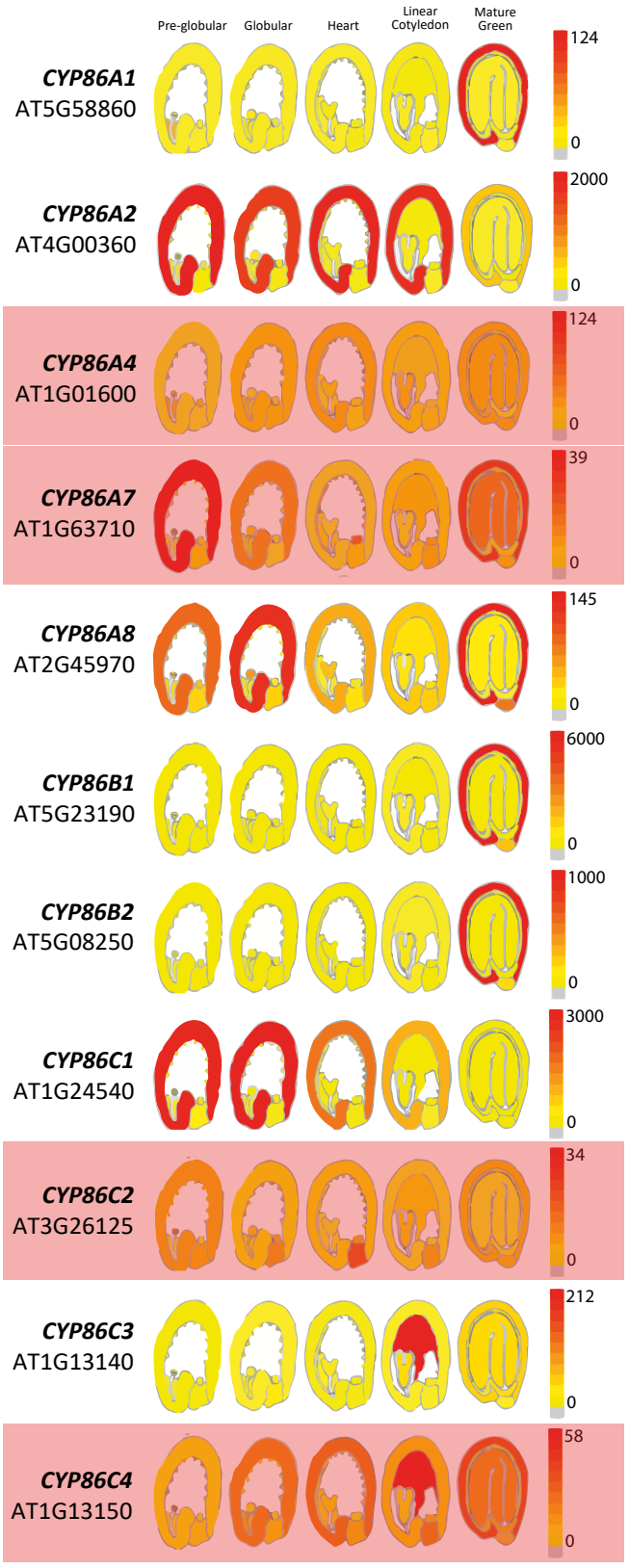


Figure 23: *In silico* expression analysis of the CYP86 gene family during seed development based in Le *et al.* (2010) data. Images were obtained with the eFP browser (Winter *et al.*, 2007).

rounder, they present a reduced lipid polyester staining (Figure 24b) and they are more permeable due to the higher rate of tetrazolium reduction (Figure 24c).

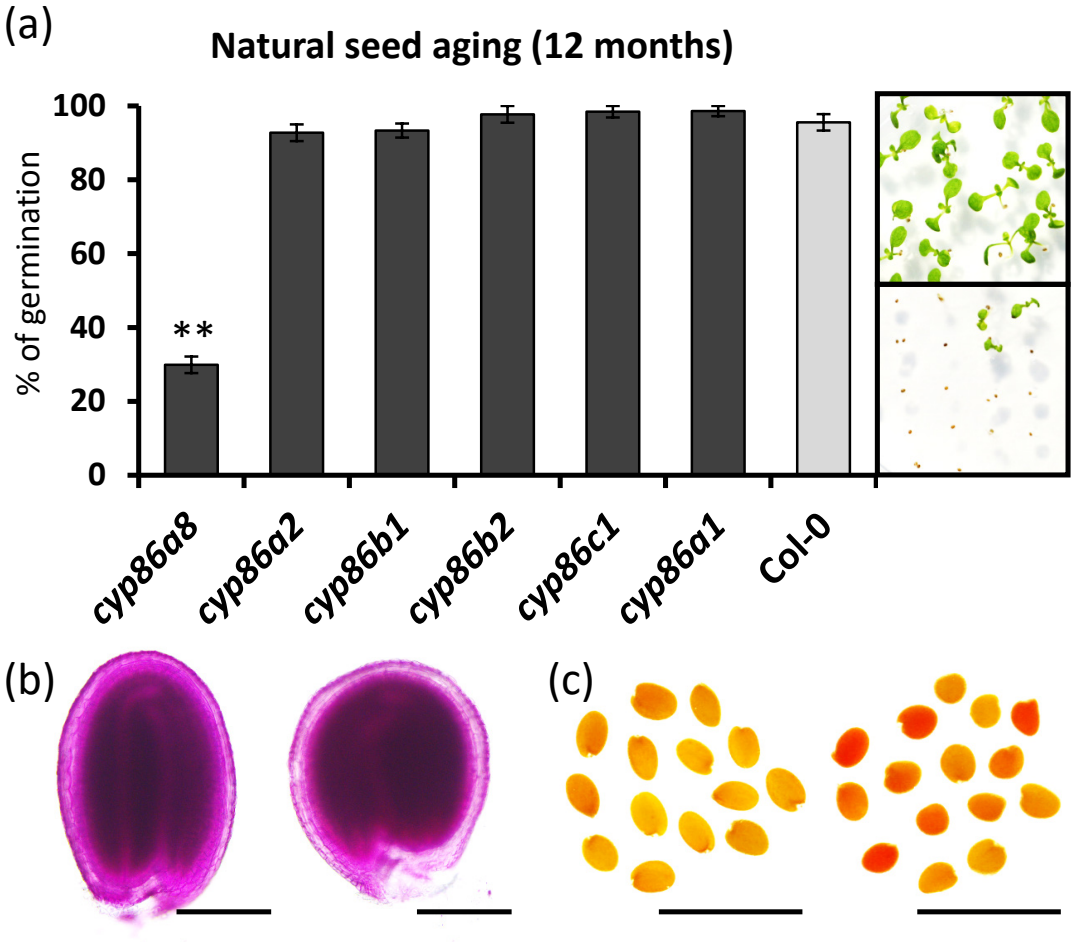


Figure 24: Reduced seed longevity of mutant *cyp86a8* correlates with reduced lipid polyester barriers. (a) Seeds of *cyp86a8* present a drastically reduced seed longevity, while seeds of mutant plants in other family members have no phenotype. Seed lots were stored for 12 months and sown on MS plates. Left: The percentage of germination was recorded after one week. Results are the average of three experiments with 50 seeds per line. Bars indicate standard errors. Not-aged seeds from all lines germinated more than 95%. **Significantly differing from wild type seeds at $p < 0.01$ (Student´s *t* test). Right: representative images of Col-0 (up) and *cyp86a8* (down) germination. (b) Mutant *cyp86a8* seeds present a reduced suberin layer and a rounder shape. Representative image of Sudan Red staining of Col-0 (left) and *cyp86a8* (right) seeds. Scale bars: 200 µm. (c) Mutant seeds of *cyp86a8* are more permeable than Col-0 seeds. Tetrazolium reduction assay (24 h) in Col-0 (left) and *cyp86a8* (right) seeds. Scale bars: 2 mm.

Table 6: Schematic results of aging treatments (NAT, AAT and CDT) of loss-of-function mutants of different *cyp86* family members in Col-0 background.

Mutant line	*NAT*	*AAT*	*CDT*
cyp86a1	n.s.	n.s.	n.s.
cyp86a2	n.s.	n.s.	n.s.
cyp86a8	Susceptible**	Susceptible**	Susceptible*
cyp86b1	n.s.	n.s.	n.s.
cyp86b2	n.s.	n.s.	n.s.
cyp86c1	n.s.	Susceptible*	n.s.

Reduced seed germination (Susceptible), in comparison to Col-0 seeds. *Significantly differing from wild type seeds at $p < 0.05$ (Student´s *t* test). **Significantly differing from wild type seeds at $p < 0.01$ (Student´s *t* test). n.s.: non significant.

TFs expressed in seed coat and endosperm are important for seed longevity

Transcription factors (TFs) are essential for seed coat differentiation. We found some transcription factors expressed in seed coat and endosperm (Figure 25, Le *et al.*, 2010) among our candidate gene list. *SEP1* (AT5G15800), highlighted after natural aging, *KNAT7* (AT1G62990) highlighted in the CDT treatment, *SPCH* (AT5G53210), highlighted in the AAT, and *MYB47* (AT1G18710), highlighted in all three artificial aging treatments.

Seeds of *myb47* and *spch* showed an important reduction of seed longevity, while seed longevity of *knat7* and *sep3* (but not of *sep1,2,4*) was increased (Figure 26a). Similar behaviour was observed in artificial aging tests (Table 7). Here, a candidate gene (*SEP1*) was again not the member of the gene family with the strongest seed-phenotype (this was *sep3*). As those TFs were expressed in the seed coat (Figure 25), we performed a tetrazolium assay to assess their seed coat permeability (Figure 26b). Seeds of *myb47* and *sep1,2,4* do no present significant differences to Col-0 seeds. Interestingly, *spch* and *knat7* seeds present enhanced and reduced tetrazolium salt reduction respectively, in correlation with their seed longevity phenotype (Figure 27a for *knat7* seeds). Surprisingly, *sep3* seeds present a high reduction to tetrazolium salts. Seed coat assays to visualize proanthocyanidins (PAs), the mucilage halo, the suberin layer, and seed size were performed. Differences were observed only in *sep3* and *knat7* seeds. As described, *knat7* mutant seeds present less mucilage extrusion (Romano *et al.*, 2012), and *sep3* seeds exhibit a similar feature (Figure 28). Seed size is reduced in *knat7* seeds (Figure 27b, c) and *sep3* seeds presented a rounder shape (Figure 26b, Figure 28).

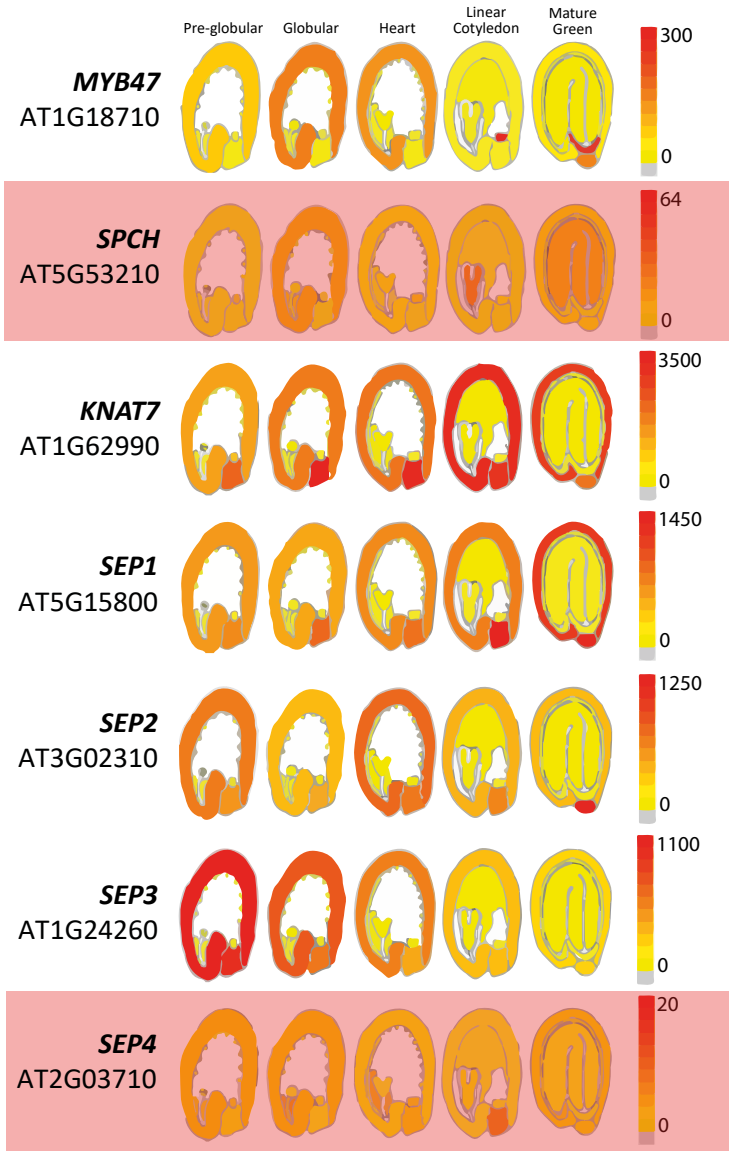


Figure 25: *In silico* expression analysis of *MYB47*, *SPCH*, *KNAT7* and the *SEPALLATA* gene family during seed development based in Le *et al.* (2010) data. Images were obtained with the eFP browser (Winter *et al.*, 2007). Pale red boxes indicate those genes with reduced seed-development expression

MYB47 is expressed in the seed coat and it is a positive regulator of seed longevity

To further investigate the MYB47 mechanism, we developed Arabidopsis transgenic lines. Confocal imaging of *proMYB47::GFP* plants locates the expression of *MYB47* in the seed coat during seed development. More precisely, GFP expression was visualized during first days of seed development in the seed coat, and in latter days, it was

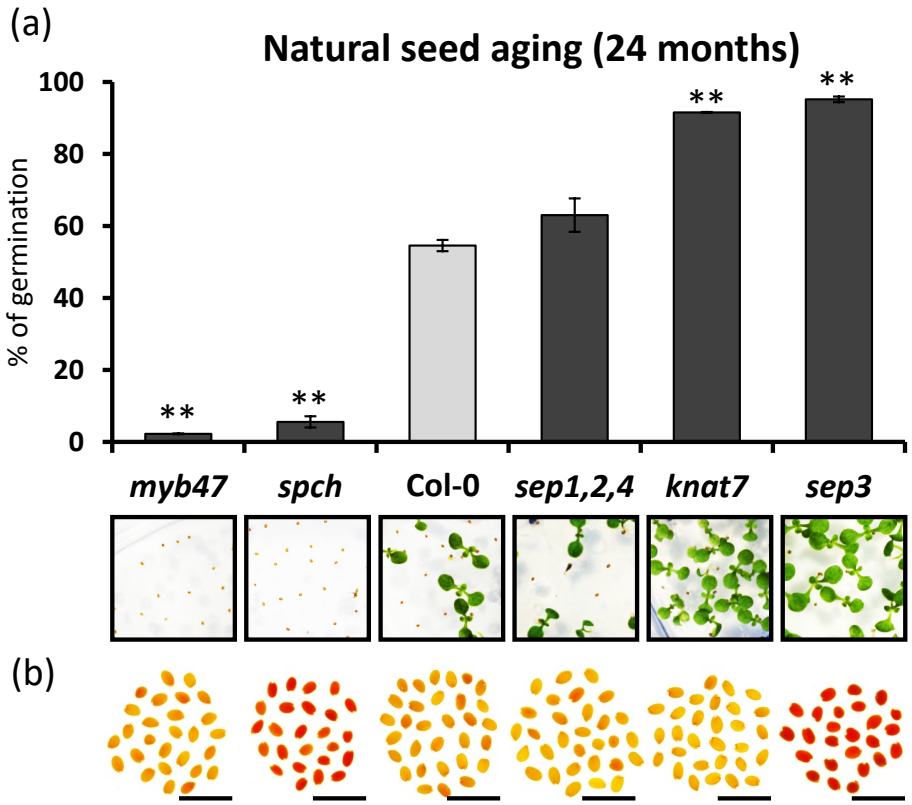


Figure 26: Mutants of different TFs exhibit differences in seed longevity and seed coat permeability. (a) Seeds of *spch* and *myb47* have reduced seed longevity; seeds of *knat7* and *sep3* have increased seed longevity, but triple mutant *sep1,2,4* has no phenotype. Seed lots were stored for 24 months and sown on MS plates. Above: The percentage of germination was recorded after one week. The results are the average of three experiments with 50 seeds per line. Bars indicate standard errors. Not-aged seeds from all lines germinated more than 97%. **Significantly differing from wild type seeds at $p < 0.01$ (Student´s *t* test). Below: representative images. (b) Mutant seeds of *spch* and *sep3* are more permeable and *knat7* seeds are less permeable than Col-0 seeds. Tetrazolium reduction assay (24 h) in *myb47*, *spch*, Col-0, *knat7* and *sep3* seeds. Scale bars: 2 mm.

Table 7: Schematic results of aging treatments (NAT, AAT and CDT) of loss-of-function *myb47, spch, knat7, sep3* mutants and the triple mutnat *sep1,2,4* in Col-0

Mutant line	*NAT*	*AAT*	*CDT*
myb47	Susceptible**	Susceptible*	Susceptible*
spch	Susceptible**	Susceptible*	Susceptible*
knat7	Resistant**	Resistant*	Resistant*
sep3	Resistant**	Resistant**	Resistant**
sep1,2,4	n.s.	Resistant**	n.s.

Reduced seed germination (Susceptible) and increased seed germination (Resistant), in comparison to Col-0 seeds. *Significantly differing from wild type seeds at $p < 0.05$ (Student´s *t* test). **Significantly differing from wild type seeds at $p < 0.01$ (Student´s *t* test). n.s.: non significant.

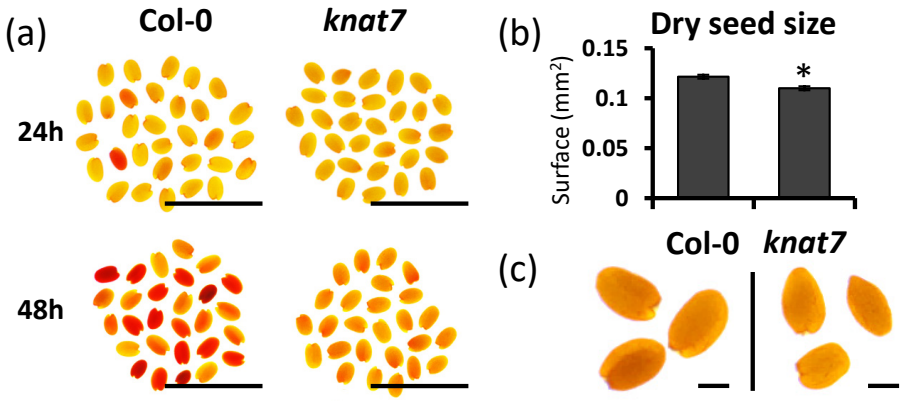


Figure 27: Seeds of *knat7* seeds are less permeable and smaller than Col-0 seeds. (a) Tetrazolium salt reduction assay for 24 h and 48 h at 28°C. Scale bars: 2 mm. (b) Seed size analysis. Average measure of projected seed surface of 100 dry seeds. Size analysis was measured with Fiji program. Bars indicate standard errors. *Significantly differing from average wild type seed size at $p<0.05$ (Student´s *t*-test). (c) Representative images. Scale bars: 200 µm.

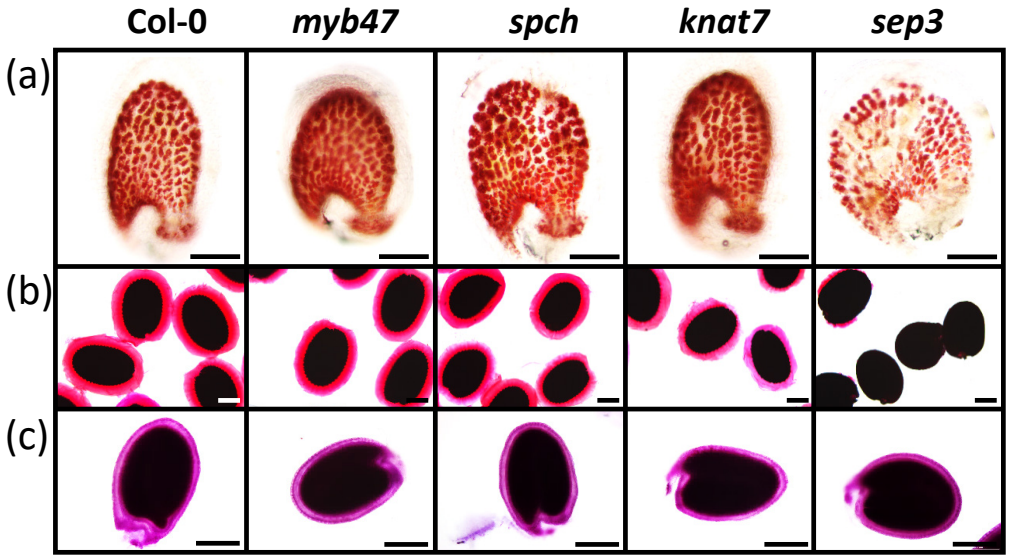


Figure 28: Seed coat analysis in seeds of TF mutants of PAs, mucilage halo and lipid polyesters barriers. (a) Vanillin staining in 5 days-after-pollination developing seeds showed no differences in PAs. Little differences observed are related to little differences in pollination times, observed also in same silique seeds. Representative mutant *sep3* seed show a rounder shape and a fissure, but there are not highly significant differences in PAs in mutant seeds. Scale bars: 100 µm. (b) Ruthenium red mucilage staining to visualize the mucilage halo. Differences are observed in *knat7* and *sep3* seeds. As published, *knat7* present a reduced halo. Here we can observe a bigger halo compared to previous reports, probably due to an increased incubation time, permitting more mucilage extrusion. Seeds of *sep3* mutant present a drastically reduction of mucilage. Scale bars: 200 µm. (c) Sudan Red staining did not show differences in lipid polyesters barriers that could explain their different seed longevity. Scale bars: 200 µm.

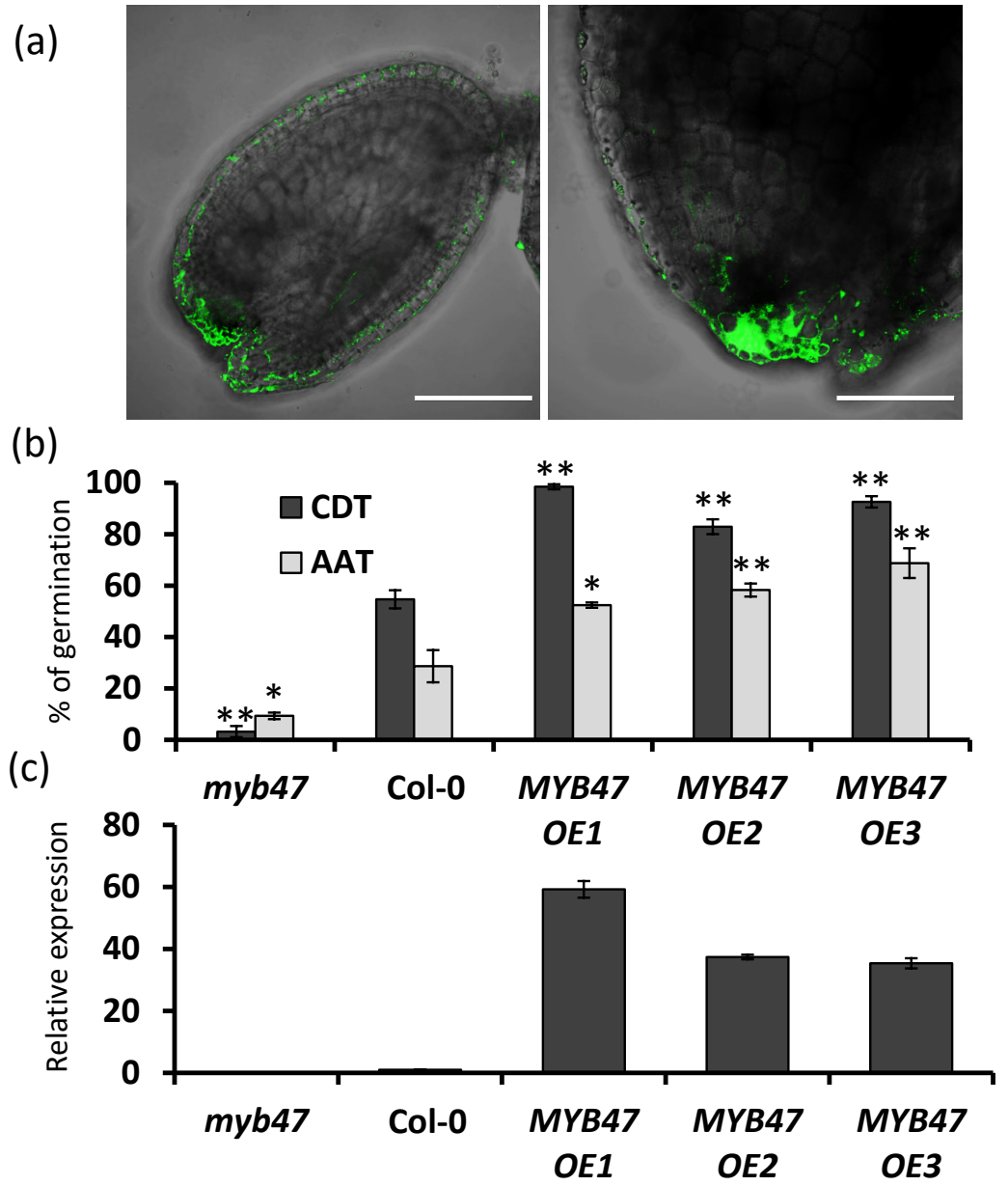


Figure 29: MYB47 is a positive TF gene for seed longevity and it is expressed in the chalaza during seed development. (a) *MYB47* expression during seed development is localized initially at the seed coat (left, DAP3) and later in the chalaza (right, DAP7). Confocal imageing of developing seeds of *proMYB47::GFP* plants. Scale bar: 100 µm. (b) Seeds of three independent lines over-expressing *MYB47* are more resistant to accelerated aging treatments CDT (dark bars) and AAT (light bars). The percentage of germination was recorded after one week. The results are the average of three experiments with 50 seeds per line. Bars indicate standard errors. Not-aged seeds from all lines germinated more than 99%. *Significant differences from wild type seeds at $p < 0.05$ (Student´s *t* test). **Significantly differing from wild type seeds at $p < 0.01$ (Student´s *t* test). (c) qRT-PCR gene expression analysis of *MYB47* in 7 days-old seedlings from *myb47* mutant and three *MYB47* over-expressing lines. Expression values are relative to housekeeping gene *PP2AA3* and the resulting ratios are normalized to wild type, taken as 1. Results are the average of three determinations with bars corresponding to standard errors.

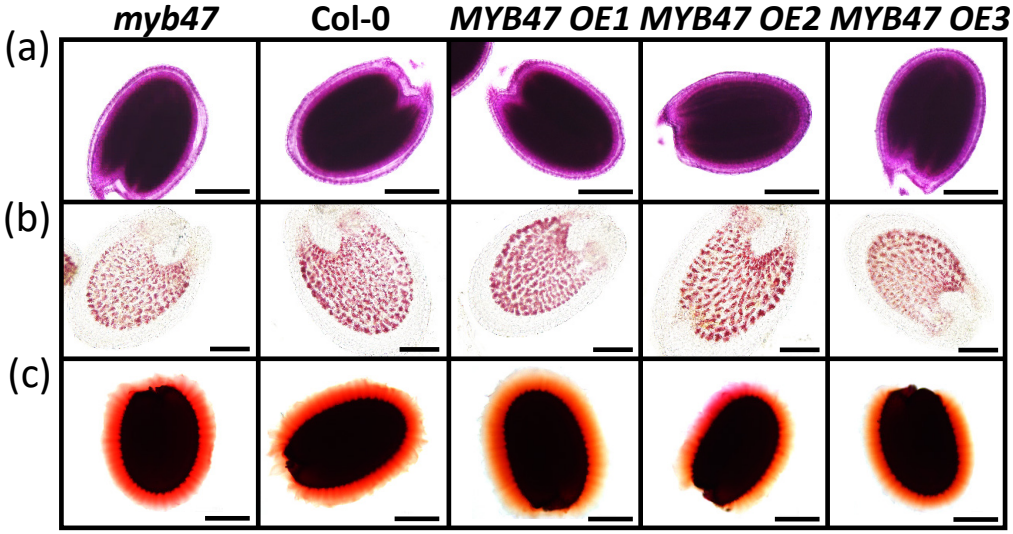


Figure 30: Different dosses of *MYB47* do no produced changes in seed coat lipid polyester barriers, PAs or mucilage halo that could explain the different seed longevity observed among different *MYB47* mutants. (a): Sudan Red staining did not show differences in lipid polyesters barriers that could explain their different seed longevity. Scale bars: 200 µm. (b): Vanillin staining in 5 days-after-pollination developing seeds showed no differences in PAs. Scale bars: 100 µm. (c): Ruthenium red mucilage staining to visualize the mucilage halo and there are not significant differences. Scale bars: 200 µm.

localized at the chalaza (Figure 29a). This expression pattern fits with Le *et al.* (2010) data (Figure 25).

Over-expression of *MYB47* driven by the UBQ10 promoter conferred resistance to AAT and CDT treatments compared to *wild-type* plants in three independent lines, highlighting the importance of *MYB47* in seed longevity as a positive regulator (Figure 29b). Expression analysis in seedlings demonstrates differences in MYB47 gene expression among different over-expressing lines. The *myb47* mutant has 30-fold reduction of *MYB47* gene expression, while over-expression lines increased *MYB47* expression 35 and 60-fold compared to wild-type plants (Figure 29c). Although *MYB47* expression is located in seed coat during seed development, we did not observe significant differences in seed coat component abundance (mucilage, PAs, or lipid polyester barriers) among over-expressing lines, compared to Col-0 (Figure 30).

Discussion

Many of the genes required for seed longevity are still unknown. In this work we have utilized a combination of GWAS, rational filtering and reverse genetics in *Arabidopsis thaliana* Col-0 to identify twelve novel genes involved in this trait. Seven are positive for longevity (*PSAD1, SSLEA, SSTPR, DHAR1, CYP86A8, MYB47* and *SPCH*) as

their knock-out mutants have less seed longevity than their corresponding wild type. Five are negative ones (*RBOHD, RBOHE, RBOHF, KNAT7* and *SEP3*) as their knock-out mutants have more seed longevity than their corresponding wild type. These results contribute to the understanding of this polygenic trait.

Our data demonstrates a wide variation in seed longevity among Arabidopsis ecotypes (Figure 12). Correlations between artificial and natural aging treatments indicate that, with exception of the AAT case, CDT and EPPO are valuable treatments to test seed longevity (Figure 13). Bueso *et al.* (2014) reported a good correlation between AAT and NAT. However, their plant material consisted of mutant lines in the same *Arabidopsis* background. In our case, each ecotype presents a different background genome, affecting the expression of multiple genetic components, probably causing this miss-correlation. In our T-DNA mutant approach, lines have the same genetic background (Col-0) and behave similarly with natural, AAT and CDT aging procedures (Table 3, Table 4, Table 5, Table 6 and Table 7), as in Bueso *et al.* (2014) experiments.

One important aspect observed in natural variation of seed longevity is that it inversely correlates with plant life span, flowering time and rosette leave number (Table 1). These phenotypes are linked, as rosette leave number increases as flowering delays, and plant life span is dependent on flowering time to ensure plant reproduction. Flowering initiates the developmental program of tissues that will develop into seeds. In some way, the rapid flowering initiation signals must be important for seed development and seed longevity acquisition, although the connection is not straightforward. There are no previous reports concerning this trait association. No correlation with ecotype latitude-collection point discards an effect of temperature or photoperiod on adaptation of seed-longevity. The negative correlation found between seed longevity and iron content has been already suggested (Murgia *et al.*, 2015). It is plausible that high amounts of iron drives oxidative stress during seed storage. Seed dormancy, described to be inversely correlated with seed longevity (Nguyen *et al.*, 2012), was neither correlated with our longevity data, according to our data. Anthocyanin content influences PAs deposition in seed coats, a component putatively implicated in seed longevity (Debeaujon *et al.*, 2000) but was not correlated.

New mechanisms regulating seed longevity

The AAT treatment highlighted a LEA protein (AT3G17520) that we named SSLEA after the observed reduced seed storability of the mutant. There are 71 putative LEA proteins in Arabidopsis and their functions are usually associated with desiccation tolerance. A link between seed longevity and LEA proteins has been proposed based on the RNA interference targeting of LEA14 of Arabidopsis resulting in decreased seed longevity (Hundertmarkt *et al.*, 2011). This LEA protein belongs to group 2 (de-

hydrins) and SSLEA belongs to group 6 according to Jaspard *et al.* (2012). Dry seeds of the *sslea* mutant are able to germinate completely after ripening, but germination ability decays rapidly in months. This fact discards that *sslea* mutant seeds cannot cope with desiccation.

The EPPO treatment highlighted a tetratricopeptide-repeat (TPR) protein (AT4G02750) that we named SSTPR after the reduced seed storability of the mutant. These mutant seeds exhibit a high permeability by the tetrazolium test. SSTPR belongs to the TPR-PTR (tetratricopeptide-pentatricopeptide) family of repeat domain proteins (Sharma and Pandey, 2016). Not much is known about this TPR protein, or other family members. PTRs are described to play constitutive and essential roles in mitochondria and chloroplasts, probably binding to RNA (Lurin *et al.*, 2004). Indeed, SSTPR protein is sublocalized in the mitochondria according to Hooper *et al.* (2017).

ROS generation and detoxification are important in seed longevity

All three artificial aging treatments highlighted genes involved in the ROS metabolism and detoxification. It has been widely accepted that this oxidative damage is one of the major causes of seed aging. This ROS negative effect has been remarked by the ROS-producing enzymes RBOHD, RBOHE and RBHOF, the ROS detoxification system involving DHAR1 and the photosystem I subunit PSAD1. ROS signalling is essential for diverse cellular processes and developmental programs. It is involved in the fine-tuning of the hypersensitive response (HR) upon pathogen infection, abiotic stress signalling and in hormone and developmental signalling, including programmed cell death (PCD) (Miller *et al.*, 2008; Suzuki *et al.*, 2011; Marino *et al.*, 2012). In seeds, ROS are required for embryogenesis, programmed cell death of the seed coat and endosperm and during seed germination, and act as protection against pathogens (Murphy *et al.*, 1998; Bailly *et al.*, 2008; Jeevan Kumar *et al.*, 2015). Here we present evidence that RBOH-produced ROS are a major cause of deterioration during aging. The enhanced seed longevity of RBOH mutants could indicate RBOH-ROS activity through seed storage. RBOHE and RBOHF are expressed in the seed coat and RBOHD in the endosperm and embryo (Figure 18), pointing to a role during cell death of these tissues. Plants express these proteins in mature seeds because ROS are needed for maturation of the seed coat and for defence against pathogen attack (Jeevan Kumar *et al.*, 2015).

A ROS-detoxification system has been highlighted by means of the *DHAR1* gene. DHARs are glutathione-dependent dehydroascorbate reductases. They catalyse the regeneration of ascorbate oxidized during detoxification of H_2O_2 by ascorbate perox-

idase (Foyer and Noctor, 2011; Smirnoff, 2011). This model is supported by the high oxidation of glutathione in the triple mutant *dhar1,2,3* (Rahantaniaina *et al.*, 2017) and the use of glutathione redox equilibrium as a seed viability marker (Kranner *et al.*, 2006). Over-expression of ascorbate peroxidases leads to resistance to oxidative stress (Wang *et al.*, 1999), highlighting the importance of this ROS-detoxification system. We observed that plants lacking *DHAR1,* but not *DHAR2* or *DHAR3*, presented less seed longevity, probably because less ROS detoxification. This is consistent with their expression profiles, with *DHAR1* being the most expressed isoform in seed embryo (Le *et al.*, 2010).

Another highlighted ROS-related protein is the Photosystem I subunit D1 or PSAD1 (AT4G02770). It participates in the stability of this photosystem (Ihnatowicz *et al.*, 2004). This hydrophilic protein is exposed in the stroma and interacts directly with ferredoxin in the electron transport chain (Andersen *et al.*, 1992; Merati and Zanetti, 1987; Zilber and Malkin, 1988). Although there is a homologue gene in Arabidopsis (*PSAD2*), *PASD1* loss of function leads to growth defect due to a deficient photosynthesis, indicating that *PSAD2* is not completely redundant to *PSAD1*. Mutant *psad1* plants are pale-green and dwarf, and they have an increased photosensitivity and altered redox state of the stroma (Haldrup *et al.*, 2003). The double mutant *psad1 psad2* is not viable (Ihnatowicz *et al.*, 2004). The importance of PSAD1 in seed longevity points to its role in photosynthesis and the imbalance of photosystems in the mutant plant. An imbalance of photosystems leads to increased ROS production (Pinnola and Bassi, 2018) and this may explain the reduced seed longevity phenotype. *PSAD1* transcripts are highly expressed at the embryo in latter stages of seed development according to Le *et al.* (2010) data (Figure 15). Chlorophyll develops during embryo development for proper seed filling (Ruuska *et al.*, 2002; Goffman *et al.*, 2005) and latter disappears for seed storage avoiding light-induced ROS accumulation (Nakajima *et al.*, 2012). However, we cannot discard that no direct correlation may exist and that the seed longevity phenotype of *psad1* seeds is due to seed developmental problems due to the strong pale and dwarf phenotype observed in the mutant plant.

Lipid polyester barriers prevents embryo aging

The physical protection by seed coat, through lipids polyester barriers, has been highlighted in our study by two members of the CYP86 family. The CYP86 family belongs to the Cytochrome P450 superfamily, with 244 members in Arabidopsis and is involved in numerous metabolic pathways in all organisms (Mansuy, 1998; Nelson, 1999). Members of CYP86 gene family catalyse the ω-hydroxylation of fatty acids (Benveniste *et al.*, 1998; Wellesen *et al.*, 2001; Duan and Schuler, 2005), and mutant plants in different CYP86 members show a reduction in ω-hydroxy fatty acids and

α,ω-dicarboxylic fatty acids (Höfer *et al.*, 2008; Compagnon *et al.*, 2009; Kai *et al.*, 2009). CYP86 enzymes participate in the synthesis of lipid-polyester barriers, such as cutin and suberin (Watson *et al.*, 2001; Compagnon *et al.*, 2009; Kannangara *et al.*, 2007). Recent studies point to an important role of these lipid polyester barriers in seed longevity and tetrazolium impermeability (Beisson *et al.*, 2007; Yadav *et al.*, 2014; Bueso *et al.*, 2016; Renard *et al.*, 2020*a*).

The drastic reduction of seed longevity, the high tetrazolium reduction rate and the lighter Sudan Red staining suggest that *cyp86a8* seeds are more permeable due to a reduced suberin and/or cutin layer. With this analysis, we can conclude that CYP86A8 is probably the major cytochrome P-450 of the CYP86 family producing the ω-hydroxylation for seed lipid-polyester synthesis necessary for seed longevity. However, CYP86A8 is also important for diverse developmental and signalling processes as described by Wellesen *et al.* (2001), and we cannot discard that these processes are affecting also seed development.

TFs modulate seed longevity in different ways

The first physical barrier to protect the seed embryo is the seed coat, which requires a precise developmental program of its cell layers. Floral identity TFs regulate the ovule integuments that will develop into the seed coat. Many TFs and different protein-protein interactions important for seed coat development have been described (Sano *et al.*, 2016; Golz *et al.*, 2018). However, the complete TF cascade is far to be completely known. Here we demonstrate the implication of four TFs in the seed development transcriptional programme whose mutant lines present seeds with altered seed longevity: SEP3, SPCH, KNAT7 and MYB47.

SEP3 is part of the *SEPALLATA* (*SEP*) gene family, formed by four MAD-box TFs (Pelaz *et al.*, 2000). They interact with other MAD-box proteins to determine flowering and ovule development (Favaro *et al.*, 2003; Hugouvieux *et al.*, 2018). They are not completely redundant as they differ in DNA-binding patterns (Jetha *et al.*, 2014; Soza *et al.*, 2016). SEP1 was highlighted in our analysis, but we found SEP3 to be determinant, due to the enhanced seed longevity and seed coat phenotype of *sep3* mutant seeds. The different seed shape and size and the reduced mucilage halo of *sep3* seeds suggest that SEP3 has a role in seed coat development. The SEP3 seed-coat expression pattern during seed development (Le *et al.*, 2010) supports this idea. The results of the tetrazolium reduction by *sep3* seeds is intriguing. Normally tetrazolium salts reduction is related with a higher seed coat permeability, which is inversely correlated to seed longevity (Debeaujon *et al.*, 2000). Mutant seeds of two SEP3 interactors present similar trait effects: *stk* seeds present a round shape as *sep3* seeds do (Mizzotti *et al.*, 2014) and *tt16* seeds present high tetrazolium reduction rate (Debeaujon *et al.*, 2000). Both

are involved in inner integument regulation and seed shape (Coen *et al.*, 2017). An explanation could be that *sep3* seed coat is physically different, perhaps softer. Internal embryo pressure would be responsible for this round shape. Pressure may be stronger during seed development and during water imbibition (producing seed coat fissures), but no significant in the dry stage. This would explain the miss-correlation between tetrazolium reduction and their enhanced seed longevity in the dry state.

SPCH is a well-described basic helix-loop-helix (bHLH) TF involved in stomatal lineage determination during the asymmetrical division of future guard cells (MacAlister *et al.*, 2007). Mutant plants have no stomata at all. They have been widely studied for stomata formation but not in relation with seed development and longevity. We found that seed longevity of *spch* mutant seeds was reduced. Tetrazolium assay in *spch* seeds showed enhanced permeability (Figure 26b), indicating a putative function in seed development.

KNAT7 is a Homeobox-TF and seeds of the *knat7* mutant are more tolerant to seed aging, are notably smaller and present decreased tetrazolium reduction (Figure 27). *knat7* mutant seeds have been described to have affected epidermis cells responsible for mucilage secretion (Romano *et al.*, 2012). The expression pattern is at seed coat during last maturation stages (Figure 25). Recent studies propose that KNAT7 inhibits cell wall formation by regulating lignin synthesis through interaction with BHL6 and MYB75 (Bhargava *et al.*, 2013; Liu *et al.*, 2014). Cell wall deposition is an important aspect of seed coat development and seed longevity and the absence of this TF must improve it.

MYB47 is the only TF not previously described in other works. It belongs to R2-R3 MYB transcription factor family formed by 126 members involved in the regulation of diverse metabolic pathways, including PAs, suberin and cuticle biosynthesis (Dubos *et al.*, 2010). *MYB47* was highlighted after the three artificial treatments (AAT, CDT and EPPO) and mutant seeds present reduced seed longevity. Over-expression plants had enhanced seed longevity, confirmed with artificial aging treatments (Figure 29b). This corroborates its positive effect in seed longevity. GFP promoter-driven expression locates *MYB47* expression in the seed coat during seed development, coinciding with published expression dataset (Figure 25). Nevertheless, we did not find differences in mucilage, lipid polyester staining or PAs deposition in the knock out mutant or in the over-expression mutants (Figure 30). Thus, MYB47 participates in an important seed coat mechanism contributing to seed longevity, not related with seed coat components such as PAs, mucilage or lipid polyester barriers, pointing to a new unknown protecting mechanism of the seed coat which remains unclear.

We have provided insights of the participation of four TFs in the seed longevity trait. Observed seed phenotypes suggest a role of this four seed TFs in seed coat develop-

ment. However, further study is necessary to complete the understanding of these TF and their role in the TF cascade regulating seed development. Seed RNA-seq analysis of these mutants might provide us new players and new molecular mechanisms involved in seed longevity.

Concluding remarks

Our work demonstrates the power of combining GWAS and reverse genetics to identify novel seed longevity genes. We have described 12 novel genes involved in seed longevity. They were found by GWAS and validated by T-DNA insertion mutants. These genes are coding for seed coat TFs, ROS-producing enzymes, a ROS-detoxification enzyme, a mature seed protein, a photosystem component, a cytochrome P-450 and a scaffolding protein. More candidate genes remain to be validated in future studies, increasing the list of genes involved in this highly polygenic trait.

Chapter 2

New Phytologist (2021) **231**:679–694 DOI: 10.1111/nph.17399

Apoplastic lipid barriers regulated by conserved homeobox transcription factors extend seed longevity in multiple plant species

Joan Renard[1], Irene Martínez-Almonacid[1], Indira Queralta Castillo[2], Annika Sonntag[2], Aseel Hashim[2], Gaetano Bissoli[1], Laura Campos[1], Jesús Muñoz-Bertomeu[1], Regina Niñoles[1], Thomas Roach[3], Susana Sánchez-León[4], Carmen V. Ozuna, José Gadea[1], Purificación Lisón[1], Ilse Kranner[3], Francisco Barro[4], Ramón Serrano[1], Isabel Molina[2] and Eduardo Bueso[1*].

[1]***Instituto de Biología Molecular y Celular de Plantas, Universitat Politècnica de València-Consejo Superior de Investigaciones Científicas, Camino de Vera, 46022 Valencia, Spain***

[2] ***Department of Biology, Algoma University, 1520 Queen Street East, Sault Ste Marie, ON, P6A 2G4, Canada***

[3]***Institute of Botany, Functional Plant Biology, University of Innsbruck, Innsbruck, A-6020 Austria***

[4]***Department of Plant Breeding, Institute for Sustainable Agriculture (IAS-CSIC), 14004 Córdoba, Spain***

*corresponding author (e-mail: edbuero@ibmcp.upv.es)

Abstract

- Cutin and suberin are lipid polyesters deposited in specific apoplastic compartments. Their fundamental roles in plant biology include controlling the movement of gases, water and solutes, and conferring pathogen resistance. Both cutin and suberin have been shown to be present in the Arabidopsis seed coat where they regulate seed dormancy and longevity.
- In this study, we use accelerated and natural aging seed assays, glutathione redox potential measures, optical and transmission electron microscopy and gas chromatography-mass spectrometry to demonstrate that increasing the accumulation of lipid polyesters in the seed coat is the mechanism by which the AtHB25 transcription factor regulates seed permeability and longevity.
- Chromatin immunoprecipitation during seed maturation revealed that the lipid polyester biosynthetic gene *LACS2* (long-chain acyl-CoA synthetase 2) is a direct AtHB25 binding target. Gene transfer of this transcription factor to wheat and tomato demonstrates the importance of apoplastic lipid polyesters for the maintenance of seed viability.
- Our work establishes AtHB25 as a trans-species regulator of seed longevity and has identified the deposition of apoplastic lipid barriers as a key parameter to improve seed longevity in multiple plant species.

Introduction

The main role of plant lipid polyesters is to contribute to the formation of a protective barrier against external physical, chemical and biological factors (Kolattukudy, 2001). This enclosure could be used by the plant to prevent water loss or gas diffusion. For instance, cuticles of the pericarp of *Solanum lycopersicum* (tomato) and *Capsicum annuum* comprise the main diffusion barrier for oxygen (Lendzian, 1982). In the organs of many plants, cell wall suberization is a common strategy to restore diffusion resistance after wounding. Thus, during wound-healing, suberin can be deposited in tubers, but also on fruits and leaves (Kolattukudy and Espelie, 1989). The protective role of suberin from freezing damage has been tested in different species, including *Secale cereale* (Griffith *et al.*, 1985), and many reports have described the importance of cutin and suberin to prevent biotic stress (Franich *et al.*, 1983). An important function of these polymers is to seal off specific plant tissues; for example, suberization of the root endodermis participates in the regulation of both water and mineral uptake (Baxter *et al.*, 2009). Finally, ultrastructural studies indicate that suberin and cutin are deposited in seeds of many species (Kolattukudy and Espelie, 1989). In *Citrus paradise*, detailed examination showed that an amorphous cuticle layer encircles the entire seed except in the chalazal region (Espelie *et al.*, 1980). It has also been reported that 98% of the total amount of monomers found in the inner bran of white *Triticum aestivum* (wheat) caryopses were indicative of the presence of cutin, representing almost 0.5% of the dry weight of caryopses (Matzke and Riederer, 1990). Finally, seed permeability in *Glycine max* depends on the composition of the outermost cuticle (Shao *et al.*, 2007).

In *Arabidopsis thaliana* seeds, both barriers are found; suberin monomers are preferentially associated with the outer integument (Molina *et al.*, 2008), while a cuticle layer synthesized by the seed coat inner integument is found covering the outer face of the endosperm cells in mature seeds (De Giorgi *et al.*, 2015). Apoplastic lipid polymers have been proposed to modulate seed dormancy and viability. For instance, Arabidopsis mutant seeds deficient in cutin biosynthesis display decreased seed dormancy and viability levels (De Giorgi *et al.*, 2015). In *Medicago truncatula*, KNOX4, a transcription factor that controls CYP86A, a gene associated with cutin biosynthesis, regulates seed dormancy (Chai *et al.*, 2016). Furthermore, analysis of the *abcg20* mutant provided evidence that suberin composition affects seed dormancy (Fedi *et al.*, 2017) and mutations in key genes required for suberin and lignin polyphenolic accumulation in seeds demonstrated that these barriers are crucial to modulate seed longevity (Renard *et al.*, 2020*a*). Lastly, the importance of the suberin accumulation in palisade layer to cope with osmotic, salt and oxidative stress is also well described (Beisson *et al.*, 2007; Lashbrooke *et al.*, 2016; Gou *et al.*, 2017).

Although cutin and suberin present differences in their location and deposition pat-

tern, the monomers and enzymes participating in their biosynthesis are similar and it has been clearly established that both polyesters contain unsubstituted fatty acids, ω-hydroxy fatty acids, α,ω-dicarboxylic acids, glycerol and phenolic compounds. It has also been reported that fatty acid oxidases and acyltransferases play central roles in the biosynthesis of both types of polymers (Pollard *et al.*, 2008). However, we do not yet have a full understanding of their biosynthesis, structure and functions because of the complexity of these polyesters, their intractability in organic solvents and their variability between species and even within organs of the same species.

Relatively few enzymes have been described to play an important role in the biosynthesis of cutin in seeds. Among them are LACS2, a long-chain acyl-CoA synthetase, that catalyzes the synthesis of acyl-CoA intermediates (Schnurr *et al.*, 2004) and BDG1, an extracellular enzyme thought to be involved in polymerizing or organizing the cuticle at the cell wall (Kurdyukov *et al.*, 2006; De Giorgi *et al.*, 2015). In addition, the acyltransferase DCR is required for incorporation of 9(10),16-dihydroxy-hexadecanoic acid into seed coat cutin (Panikashvili *et al.*, 2009). Finally, ATT1/CYP86A2 is a crucial ω-hydroxylase required for the production of C18 unsaturated α,ω-dicarboxylic fatty acids (Molina *et al.*, 2008).

Several seed suberin biosynthetic enzymes have been described. GPAT5, glycerol-3-phosphate acyltransferase 5, is involved in the synthesis of acylglycerol and *gpat5* mutants display a strong reduction in 22:0/24:0 fatty acid and their derivatives (Beisson *et al.*, 2007), whereas aliphatic suberin feruloyl transferase (ASFT) /hydroxycinnamoyl-CoA:ω-hydroxyacid O-hydroxycinnamoyltransferase (HHT) catalyzes the incorporation of ferulic acid into the polyesters (Molina *et al.*, 2009; Gou *et al.*, 2009). Analysis of the double and triple mutants lacking the FATTY ACYL REDUCTAS genes, *FAR1*, *FAR4* and *FAR5* demonstrated the important role of these enzymes in seed suberin fatty alcohol accumulation and seed coat permeability (Vishwanath *et al.*, 2013). 3-Ketoacyl-CoA synthase 2 (KCS2/DAISY) is involved in the biosynthesis of VLCFA (very long chain fatty acids) suberin monomers at the chalaza-micropyle region (Franke *et al.*, 2009). And finally, the role of the very long chain fatty hydroxylase CYP86B1 has also been described. This gene is highly expressed in the seed coat and is critical for the production of C22- and C24-hydroxyacids and α,ω-dicarboxylic acids (Compagnon *et al.*, 2009; Molina *et al.*, 2009).

The expression of specific transcription factors that mediate the timely expression of these biosynthetic genes is essential to properly develop the seed coat. MYB107 and MYB9 control the expression of several acyltransferases, hydroxylases, reductases and elongases (Lashbrooke *et al.*, 2016; Gou *et al.*, 2017). On the other hand, COG1, a transcription factor whose expression is seed coat specific, regulates the polymerization of polyphenolic compounds through the activation of peroxidase genes (Renard *et al.*, 2020*a*).

Arabidopsis HOMEOBOX25 (AtHB25) has been described as a transcription factor that regulates seed longevity by mediating structural modifications in the Arabidopsis seed coat (Bueso *et al.*, 2014*b*). Open questions remained about the compounds that confer this resistance to seed aging in plants overexpressing AtHB25 and which genes are directly regulated by this TF. Plants harbouring a gain-of-function mutation in AtHB25 (*athb25-1D*) present improved seed viability that correlates with an increase in GAs and ABA levels during silique formation (Bueso *et al.*, 2016). GAs are required for normal formation of the Arabidopsis seed coat (Kim *et al.*, 2005) and ABA is known to regulate apoplastic lipid production in several species as *Solanum lycopersicum* (Leide *et al.*, 2012). In the present study, we demonstrate that the AtHB25 transcription factor regulates specific genes involved in lipid polyester synthesis. In addition, we show that this conserved transcription factor regulates the accumulation of these biopolymers not only in Arabidopsis, but also in tomato and wheat seeds. Thus, our work defines a molecular mechanism involving AtHB25-mediated apoplastic lipid barrier deposition that directly impacts seed longevity in Arabidopsis, wheat and tomato.

Results

AtHB25 is expressed during the seed development and regulates seed oxidation, permeability and longevity

Gain of function mutants of TF AtHB25 (*athb25-1D*) are resistant to natural and accelerated seed aging treatments. However, loss-of-function mutants of this TF are not more sensitive than wild type to accelerated treatments. The role of AtHB25 in coping with seed deterioration is partially redundant with AtHB22, as demonstrated by the sensitivity of *athb22 athb25* double mutants to artificial seed aging treatments (Bueso *et al.*, 2014*b*). In order to confirm that the biological function of AtHB22 and AtHB25 is to regulate seed longevity under natural conditions, we carried out an experiment in which seeds were stored for 12, 20 and 26 months at 20-25 ºC and 40-60% RH and sown on MS medium. Germination data were fitted to sigmoidal curves to estimate P50 (Zinsmeister *et al.*, 2016), which indicates the time at which the loss of seed viability during storage is reduced to 50%. For *athb25-1D* plants, germination after 26 months of storage was higher than 60%, while the viability of wild type and the loss-of-function double mutant was significantly lower (P50 of 22.8 ± 4.6 and 15.9 ± 3.8 months, respectively) (Figure 31a).

Oxidative stress has been proposed to be the main cause of seed deterioration during aging (Bailly *et al.*, 2008). In order to study whether the regulation of seed aging by AtHB25 depends on the permeability or/and the detrimental effects of oxygen expo-

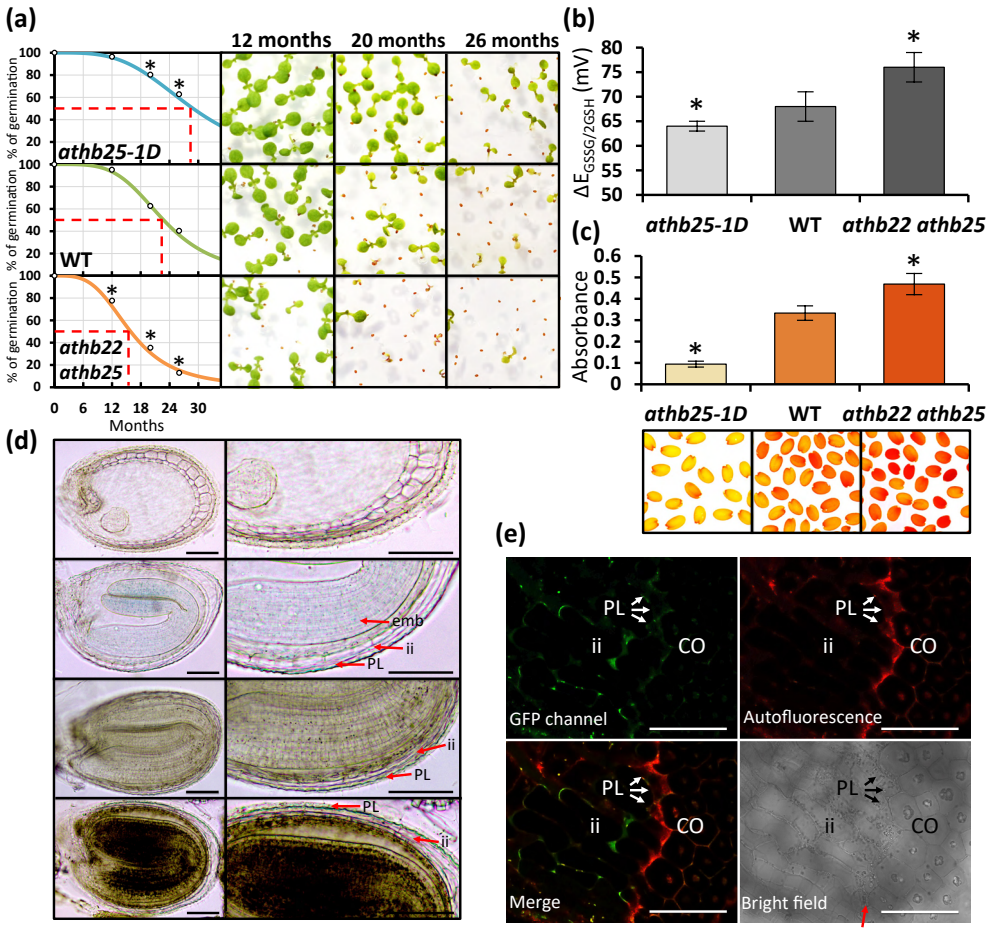


Figure 31: *AtHB25* is expressed in seed coat and regulates longevity, oxidation and permeability in *Arabidopsis thaliana* seeds. (a) Seeds from the *athb25-1D* mutant, wild type (WT) and the *athb22 athb25* double mutant were subjected to natural aging for 12, 20 and 26 months and sown on MS plates. The percentage of germination was recorded after seven days. The results are the average of three experiments with 100 seeds per line. Fitted probit and P50 value curve was calculated with the drc R package. Red dashed lines indicate estimated P50 values. Representative pictures of the germination are at the right of each graph. (b) Effects of accelerated aging on glutathione half-cell reduction potential increase ($\Delta E_{GSSG\ :\ 2GSH}$) in seeds of *athb25-1D*, WT and *athb22 athb25*. The results are the average of four experiments with 20 mg per replicate. Darker bar colours represent higher seed oxidation. (c) Seed coat permeability test of *athb25-1D,* WT and *athb22 athb25* double mutant seeds, and representative stained seed images. The error bars denote standard errors. *, $p < 0.05$. (d) GUS staining of *proAtHB25::GUS:GFP* in developing seeds at different stages (from up to down: globular cotyledon, bent cotyledon, early maturation and middle maturation). Magnifications are shown at the right side. Red arrows indicate the GUS signal and the identified tissue. Bars, 100µm. (e) Representative image of GFP-fluorescence in 10 DAP developing seeds of different independent *proAtHB25::GUS:GFP* Arabidopsis lines. The image was obtained with confocal laser scanning microscopy using the lambda scan mode and linear unmixing mode to separate the GFP channel from seed coat autofluorescence in the red channel. In the bright field image, it can be appreciated that the seed coat is broken, marked with a red arrow. This fracture allows the visualization of the GFP signal in inner cell layers, although it is also present in the outer integument with lower intensity. The outer integument presents a strong autofluorescence, presumably due to cell differentiation and increased lignification. Bars, 50 µm. emb: embryo ii: inner integument, PL: palisade layer, CO: epidermis columella cells. Arrows indicate signal in PL.

sure, an alternative artificial assay to accelerate deterioration was performed. Seeds were subjected to high-pressure oxygen storage (0.5 MPa) for five months and sown on MS plates (Renard *et al.*, 2020*a*). Again, *athb25-1D* presented resistance displaying a germination rate 50% higher than the WT control, while the *athb22 athb25* mutant was sensitive to this treatment and no seeds germinated (Figure 32). In addition, we used the redox state of glutathione as a marker for seed viability after a controlled deterioration treatment (Kocsy *et al.*, 2015). After this accelerated aging procedure, *athb22 athb25* seeds presented the most pronounced oxidative shift of the glutathione redox potential ($E_{GSSG\ :\ 2GSH}$), indicating that this mutant coped a higher level of oxidative stress during aging. In contrast, the glutathione redox potential of *athb25-1D* seeds shifted less than that of wild-type controls, showing a reduction of cellular oxidation during aging (Figure 31b).

In previous studies, we determined that the mechanism by which AtHB25 regulates seed longevity has a maternal origin, as expected if it depends on changes in the seed coat (Bueso *et al.*, 2014*b*). One important function of the seed coat is to limit the diffusion between the embryo and the surrounding environment. To quantify the seed coat permeability of these mutants, we measured tetrazolium salt uptake using the triphenyltetrazolium reduction method (Debeaujon *et al.*, 2000; Molina *et al.*, 2008). As shown in Figure 31c, the *AtHB25* overexpression line accumulated lower levels of red formazans as compared to the wild type control. In contrast, the *athb22 athb25* seed extract presented an intense red color, suggesting increased seed coat permeability.

To determine the spatial-temporal expression of *AtHB25* during seed coat development, we generated transgenic plants expressing the green fluorescent (GFP) and the

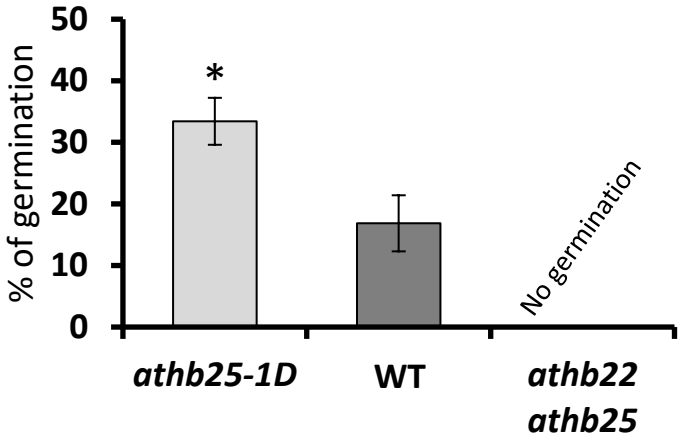


Figure 32: AtHB25 regulates germination after high-pressure oxygen storage. Arabidopsis seeds from *athb25-1D* mutant, wild type (WT) and *athb22 athb25* double mutant were subjected to high-pressure oxygen storage (0.5 MPa) for five months and sown on MS plates. The percentage of germination was recorded after seven days. The results are the average of four experiments with 100 seeds per line. Non-treated seed from all lines germinated more than 99% after 3 days. The error bars denote standard errors. *, $p < 0.05$.

beta-glucoronidase (GUS) proteins under the control of the AtHB25 promoter. The expression of AtHB25 was detected in the inner integument during the embryogenesis coinciding in space and time with the endosperm-associated cuticle formation (Figure 31d). During the maturation stage, GUS staining was also observed in palisade layer where the suberin is accumulated (Figure 31d). Finally, a specific GFP signal was observed at the beginning of the maturation in the inner integument and in the palisade layer (Figure 31e).

AtHB25 regulates lipid polyester accumulation in the Arabidopsis testa

AtHB25 regulates mucilage synthesis (Bueso *et al.*, 2014*b*), however *mum2,* a loss-of-function mutant in the biosynthesis of this pectinaceous compound did not show alterations in germination after accelerated aging treatments (Figure 33). Therefore, the mechanism by which AtHB25 confers tolerance to seed aging should involve additional processes. It has been previously shown that lipid polyesters and cell wall polyphenolics (suberin and lignin) are crucial to extend seed longevity (De Giorgi *et al.*, 2015; Renard *et al.*, 2020*a*). We studied lipid polyester accumulation in the seed coat of *athb25-1D* and *athb22 athb25* mutants. Seeds were delipidated and stained with Sudan Red 7B, a lipophilic dye that is used for staining polyesters of Arabidopsis seed coats after removing free fatty acids (Beisson *et al*., 2007). Seed coat staining was

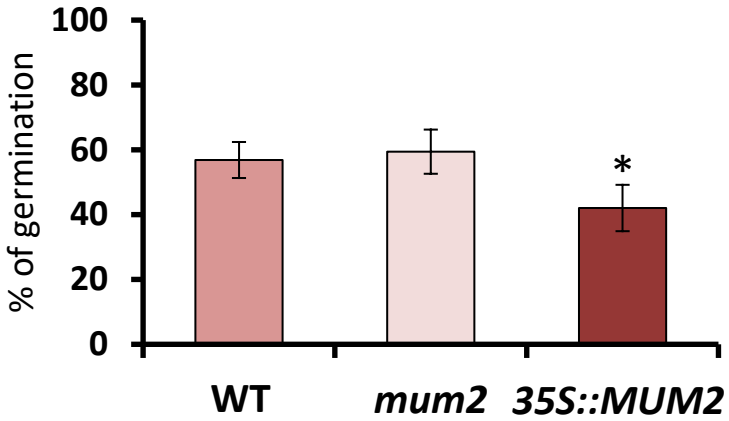


Figure 33: Mucilage production in seed coat epidermal layer is not regulating seed longevity. Arabidopsis seeds from wild type (WT), *MUCILAGE MODIFIED 2* mutant (*mum2*) and overexpression of *MUCILAGE MODIFIED 2* line (*35S::MUM2*) were subjected to accelerated aging treatment for 24 h and sown on MS plates. The percentage of germination was recorded after seven days. The results are the average of four experiments with 100 seeds per line. Non-treated seeds from all lines germinated more 99% after three days. The error bars denote standard errors. *, $p < 0.05$. Bars are red colored with different intensities qualitatively according to mucilage abundance of each line.

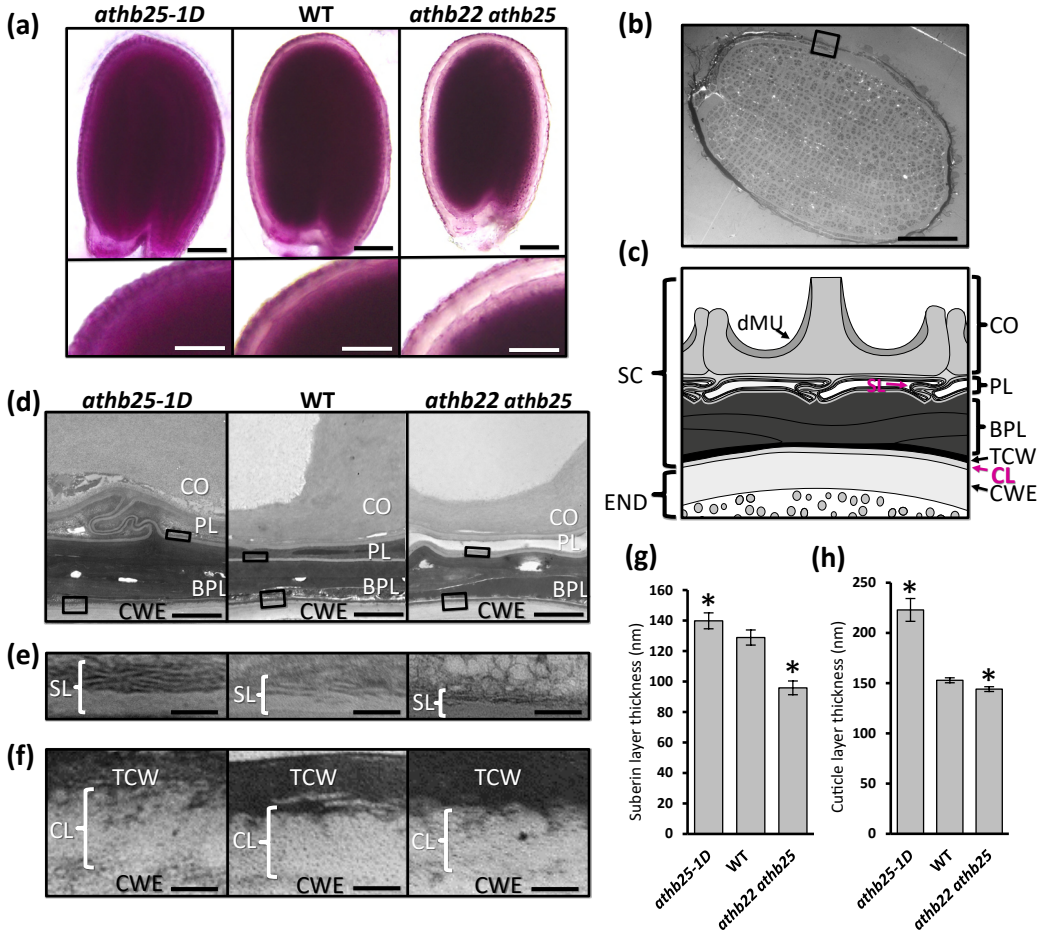


Figure 34: AtHB25 regulates lipid polyester deposition in *Arabidopsis thaliana* seed coat. (a) Sudan Red 7B staining of a representative *athb25-1D*, wild type (WT) and *athb22 athb25* seed (above). Bars, 100 µm. Detail of Sudan Red 7B staining in seed coat of the corresponding seeds (below). Bars, 50 µm. (b) Transmission Electron Microscope (TEM) of a WT mature seed ultrathin section. The black box indicates the seed coat and endosperm location, where next images are focused. Bar, 100 µm. (c) Schematic representation of the seed coat visualization under the TEM. On the left side, the seed coat (SC) and the endosperm (END) are identified. On the right side the different cell layers and components of the seed coat and endosperm are indicated. CO: columella cell; PL: palisade layer, BPL: brown pigment layer, TCW: tannic cell wall or endothelium cell wall, CL; cuticle layer and CWE: cell wall of the endosperm; dMU: dehydrated mucilage, and SL: suberin layer. The SL and CL are in purple to indicate that they constitute the two lipid polyester layers, stained by Sudan Red. (d) TEM observation of a representative seed coat of the *athb25-1D*, WT and *athb22 athb25*. Black boxes indicate the location area of next figures focusing in the suberin layer (above) and cuticle layer (below). The different components from the schematic seed coat representation can be easily identified in these pictures. Mucilage, however, is not present as it was released during seed rehydration prior to the ultrathin sectioning. Bars, 2 µm. (e) Detail of suberin layer associated to palisade layer. Bars, 100 nm. Brackets indicate the suberin layer thickness. (f) Detail of cuticle layer. Bars, 100 nm. Brackets indicate the cuticle layer thickness. (g) Thickness measurement of the suberin layer and (h) cuticle layer of *athb25-1D*, WT and *athb22 athb25* seed coat sections. Different sections were used for the analysis and the measurement was performed with the Image J program. The error bars denote standard errors. *, $p < 0.05$.

reduced in *athb22 athb25* seeds, whereas the seed coat of the aging-tolerant *athb25-1D* mutant presented more intense staining than wild type seeds (Figure 34a). To better characterize the seed coat of these mutants, we generated histological sections of siliques in the maturation stage. After Sudan Black staining, *athb25-1D* histological sections showed a thicker seed coat that accumulated more stain (Figure 35).

Next, we observed histological sections of dry seeds (Figure 34b) using optical and transmission electron microscopy (TEM). Seeds were fixed and included in resin, then were ultrasectioned and examined with a JEM-1010. As expected, we could observe the structures previously reported in the literature. The epidermal cells of the mature seed coat are represented by the mucilage and the columella, a thick cellulose-rich cell wall with a shape similar to a volcano. Beneath the epidermis lies the palisade layer where the suberin is deposited. The more internal layer is formed by three fused inner integuments (Haughn and Chaudhury, 2005). A cuticle, synthesized by the inner integument endothelium during development, covers the endosperm outer surface at maturity (De Giorgi *et al.*, 2015; Figure 34c). TEM analysis revealed epidermal cells with larger columellas (Figure 34d) and the typical lamellation of the suberin layer associated with the palisade layer was more evident in AtHB25 gain-of-function mutants (Figure 34e and g). In addition, the thickness of the cuticle depended on the AtHB25 expression level (Figure 34f and h), supporting the role of this TF in the lipid polyester deposition of both barriers in the seed coat.

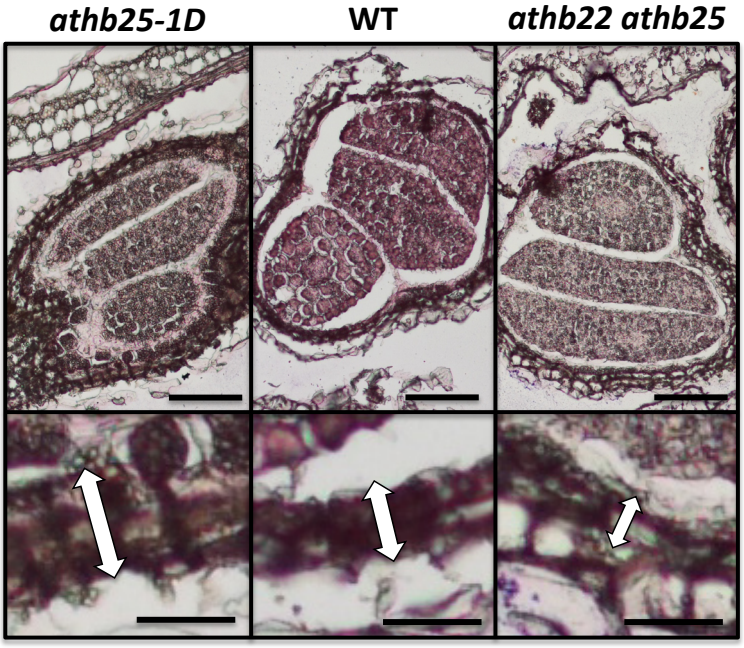


Figure 35: Central transversal sections in maturing developing seeds of *athb25-1D* mutant, WT and *athb22 athb25* double mutant stained with Black Sudan (above). Bars, 100 µm. Detail of Black Sudan staining in seed coat of the corresponding seeds (below). Bars, 25 µm. Arrows indicate thickness of stained seed coat.

To determine if changes in lipid staining intensity and the morphologies observed by electron microscopy correlate with changes in composition, we performed a chemical analysis of seed coat lipid polyesters. Lipid polyester monomers were released by chemical depolymerization of the cell wall-enriched residues remaining after thorough solvent extraction of seeds (Molina *et al.*, 2006). Gas chromatography-mass spectrometry (GC-MS) analysis of the extracted monomers revealed significant dif-

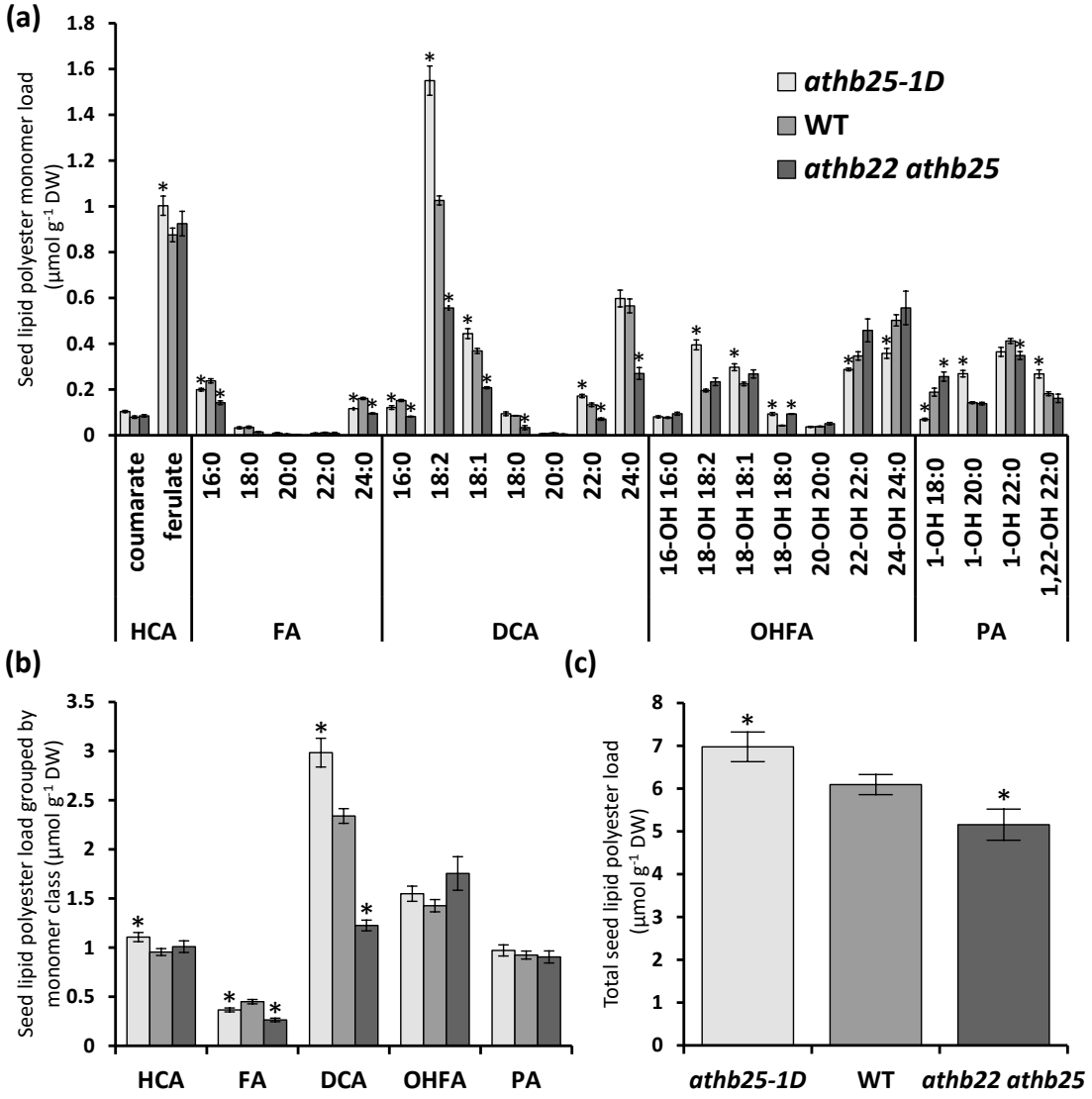


Figure 36: Lipid polyester composition in *athb25-1D*, WT and *athb22 athb25 Arabidopsis thaliana* seeds. (a) Seed lipid polyester monomer composition of the *athb25-1D*, wild type (WT) and *athb22 athb25* seeds. (b) Total lipid polyester distribution grouped by monomer class. (c) Total lipid polyester monomer content. Mean values of three biological replicates are shown in µmols g^{-1} delipidated dry residue (DW). The error bars denote standard errors. *, $p < 0.05$. HCA: hydroxycinnamic acids; FA: fatty acids; DCA: a,ω-dicarboxylic acids; OHFA: ω-hydroxy fatty acids; PA: primary fatty alcohols.

ferences in the loads of specific monomers in the mutants as compared to wild type seeds (Figure 36a). Seeds overexpressing *AtHB25* showed increased loads of ferulate, 18:2, 18:1 and 22:0 α,ω-dicarboxylic acids (DCAs), all the 18C ω-hydroxy fatty acids (OHFA), 20:0 primary fatty alcohol (PA) and 22:0 1,ω-diol. Conversely, the *athb22 athb25* double mutant presented significant reductions in the loads of all DCA monomers (Figure 36a). When the total amount of identified monomers were grouped and summed according to monomer class, it was evident that main group regulated by AtHB25 is DCAs, as *athb25-1D* seeds present a 20% increase and *athb22 athb25* seeds show a 40% reduction, as compared with wild type seeds (Figure 36b). Seed coats of the gain-of-function mutant had an approximately one μmol g^{-1} higher content of these compounds than wild type seed coats, while the loss-of-function double mutant presented a reduction of about one μmol g^{-1} of these lipids than control seeds (Figure 36c). Collectively, these observations suggest that AtHB25 regulates the accumulation of lipid polyesters in the seed coat**.**

Genome-wide identification of in vivo AtHB25 binding sites

To elucidate the molecular mechanism of AtHB25-mediated seed longevity improvement, we determined which genes are direct targets of this transcription factor. To define the global *in vivo* binding sites of AtHB25 during seed maturation stage, we performed a ChIP-seq analysis in *athb22 athb25* seeds transformed with the *proUBQ10::AtHB25:3xHA* construct. Resulting sequences are deposited in GEO (GSE154886) (Edgar *et al.*, 2002). Libraries of two ChIP replicas had 2.7 and 3.7 million Arabidopsis reads, respectively and the control with no antibody had 2.9 million reads. The Model-based Analysis of ChIP-seq (MACS2) program (Zhang *et al.*, 2008) was used to identify 175 enriched regions using a false discovery cut-off of 0.05. CLC Genomics Workbench (CLC Bio, Aarhus, Denmark) was used to check normal and equitable distribution of reads and 146 peaks were identified as AtHB25 binding sites during seed maturation (Supplemental data 4). Peak Annotation and Visualization (PAVIS) detected 119 AtHB25 binding sites (81.5%) in upstream regions (2000 bp from transcription start site (TSS) plus 5´UTR), 14 biding sites (9.6%) in downstream regions (1000 bp from transcription termination site (TTS) plus 3´UTR) and only 13 biding sites (8.9%) in different regions (Figure 37a and b). Significantly, 90% of the AtHB25 binding sites found in upstream regions were located in 5´UTRs or within 1000 bp from transcriptional start sites (Figure 37a).

The genes putatively associated with AtHB25 binding sites (located within 3000 bp from binding sites) are shown in Supplemental data 5. A Gene Ontology (GO) analysis of these genes was performed using agriGO (Tian *et al.*, 2017) and were classi-

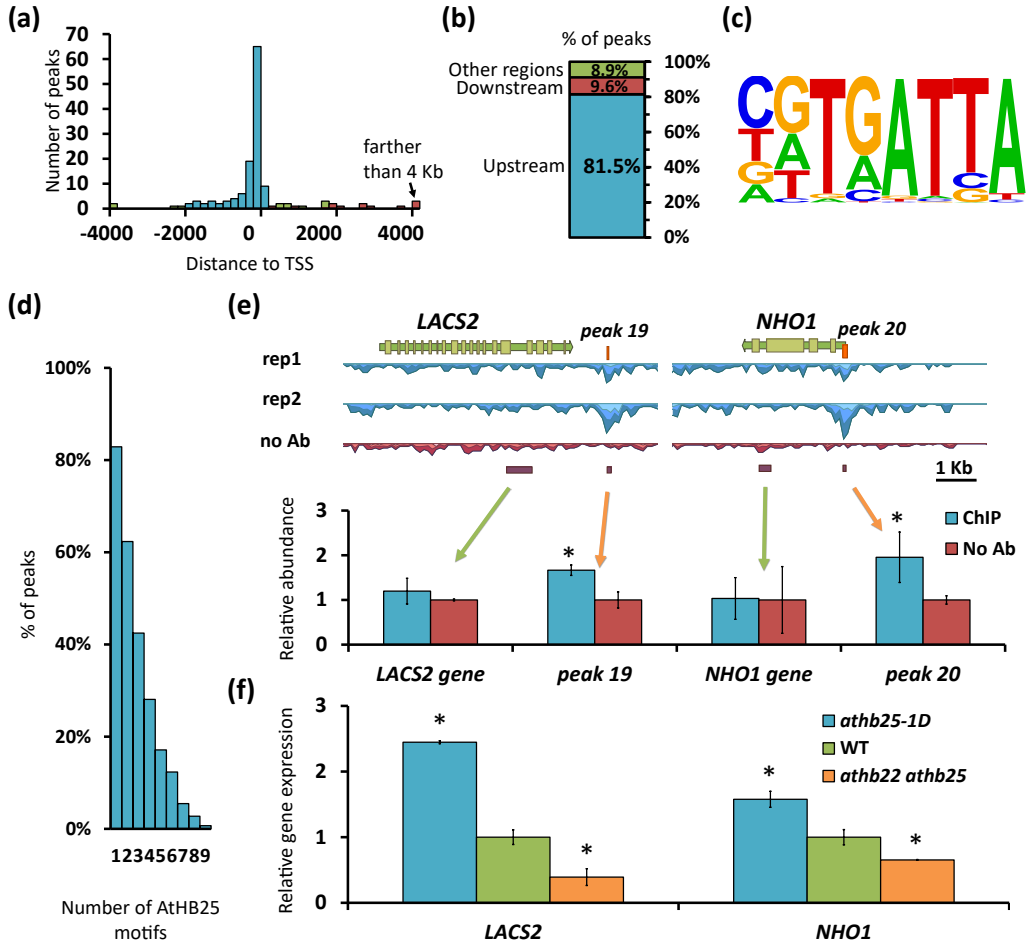


Figure 37: AtHB25 chromatin immunoprecipitation analysis in *Arabidopsis thaliana* maturing seeds. (a) Histogram of Transcription Start Site (TSS) distances from AtHB25 binding regions. TSS distances were calculated with the PAVIS suggested target gene. (b) Genomic location of AtHB25 binding regions respectively to PAVIS suggested gene. Upstream (2000 bp from TSS plus 5´UTR); Downstream (1000 bp from Transcription Termination Site (TTS) plus 3´UTR) and other (rest of regions in genome including introns, exons and intergenic regions). Colours in the histogram correspond to regions in the graph. (c) Most enriched motif in AtHB25 binding region displayed after *de novo* motif finding analysis. This analysis was performed in 200 bp from AtHB25 genomic binding sites. Best match was AtHB25-DAP-Seq motif, present in 82.88% of AtHB25 binding sites. (d) Percentage of peaks containing the indicated number of the AtHB25-DAP-Seq motives. Peaks containing more than one motif are also included in bars of minor numbers. (e) Above, ChIP AtHB25 read coverage in *LACS2* and *NHO1*. The figure was obtained with CLC Genomics Workbench program, TAIR10 genes and BAM files resulting from both Antibody samples (rep1 and rep2) and the No-Antibody control (no Ab). The representative gene model is illustrated with the arrow tip at 3' end of the gene and boxes to represent exons. Orange boxes indicate the peak area obtained in the MACS2 analysis. Bar, 1 Kilobase (Kb). Below, quantitative PCR in AtHB25 form an independent ChIP experiment performed in maturing siliques. Purple boxes indicate the regions amplified by PCR. Peak regions are linked with orange arrows and gene regions, used as negative control, are lined with green arrows. The enrichment of other genomic regions is analyzed in the Figure 38. (f) Gene expression analysis of *LACS2* and *NHO1* in maturing seeds. Expression values are relative to the housekeeping gene *PP2A3* and normalized to wild type. Results are the average of three determinations. The error bars denote standard errors. *, $p < 0.05$.

fied for biological process, molecular function and cellular component, as shown in Table 8. Post-embryonic development (GO:0009791) and response to auxin stimulus (GO:0009733) are enriched biological processes among these putative AtHB25-target genes and they could be relevant to clarify the molecular mechanism underlying the resistance of *athb25-1D* seeds to aging (Carranco *et al.*, 2010; Pellizzaro *et al*., 2020) (Table 8). The analysis revealed 16 genes related to the post-embryonic development category, including the light signalling suppressors *PIF3* and *SPA1*. The *ICE1* gene, a positive regulator of cold and ABA signalling, and *LACS2,* a gene that encodes an enzyme that catalyses the activation of fatty acids to acyl-CoA intermediates, were the most remarkable (Table 9). Regarding the response to auxin stimulus, *IAA11*, *IAA33*,

Table 8: GO enrichment analysis in AtHB25 putative target genes

Biological process	***Description***	***N list***	***N BG/Ref***	***p value***	***FDR***
GO:0042221	response to chemical stimulus	37	2085	$2.2 \cdot 10^{-05}$	0.017
GO:0010033	response to organic substance	25	1342	0.00025	0.049
GO:0009791	post-embryonic development	16	705	0.00044	0.049
GO:0009719	response to endogenous stimulus	21	1068	0.00039	0.049
GO:0009725	response to hormone stimulus	20	982	0.00035	0.049
GO:0009733	response to auxin stimulus	11	360	0.00033	0.049
GO:0005975	carbohydrate metabolic process	19	866	0.0002	0.049
Molecular function	***Description***	***N list***	***N BG/Ref***	***p value***	***FDR***
GO:0003824	catalytic activity	112	9638	$9.7 \cdot 10^{-05}$	0.028
GO:0016874	ligase activity	14	564	0.00042	0.03
GO:0016881	acid-amino acid ligase activity	11	364	0.00036	0.03
GO:0016879	ligase activity, forming carbon-nitrogen bonds	12	423	0.00034	0.03
GO:0004842	ubiquitin-protein ligase activity	10	324	0.00057	0.032
GO:0019787	small conjugating protein ligase activity	10	336	0.00074	0.035
Cellular component	***Description***	***N list***	***N BG/Ref***	***p value***	***FDR***
GO:0043227	membrane-bounded organelle	96	7622	$1.9 \cdot 10^{-05}$	0.0017
GO:0043231	intracellular membrane-bounded organelle	96	7615	$1.8 \cdot 10^{-05}$	0.0017
GO:0043229	intracellular organelle	99	8149	$5.7 \cdot 10^{-05}$	0.0027
GO:0043226	organelle	99	8155	$5.9 \cdot 10^{-05}$	0.0027
GO:0044424	intracellular part	109	9302	$9.1 \cdot 10^{-05}$	0.0033
GO:0005622	intracellular	112	9671	0.00011	0.0034
GO:0044444	cytoplasmic part	79	6289	0.00015	0.004
GO:0005886	plasma membrane	26	1456	0.00035	0.008
GO:0005737	cytoplasm	80	6822	0.0011	0.023

GO enrichment analysis performed in genes located within 3 Kb of AtHB25 binding sites, listed in Supplemental data 5. The online tool agriGO was used for this analysis (Tian *et al.*, 2017).

Table 9: Gene lists from two enriched GO terms related to embryo development and auxin response

postembrionic development

AT1G04110	SDD1	Subtilase family protein
AT1G09530	PAP3,PIF3,POC1	phytochrome interacting factor 3
AT1G12770	EMB1586,ISE1	P-loop nucleoside triphosphate hydrolases superfamily
AT1G49430	LACS2,LRD2	long-chain acyl-CoA synthetase 2
At1G69490	ANAC029,ATNAP,NAP	NAC-like, activated by AP3/PI
AT1G80410	EMB2753	tetratricopeptide repeat (TPR)-containing protein
AT2G46340	SPA1	SPA (suppressor of phyA-105) protein family
AT3G26744	ATICE1,ICE1,SCRM	basic helix-loop-helix (bHLH) DNA-binding superfamily
AT3G50790		esterase/lipase/thioesterase family protein
AT3G61630	CRF6	cytokinin response factor 6
AT3G61650	TUBG1	gamma-tubulin
AT4G03190	AFB1,ATGRH1,GRH1	GRR1-like protein 1
AT5G37770	CML24,TCH2	EF hand calcium-binding protein family
AT5G64050	ATERS,ERS,OVA3	glutamate tRNA synthetase
AT5G64930	CPR5,HYS1	CPR5 protein, putative
AT5G65700	BAM1	Leucine-rich receptor-like protein kinase family protein

response to auxin stimulus

AT2G42580	TTL3,VIT	tetratricopetide-repeat thioredoxin-like 3
AT4G28640	IAA11	indole-3-acetic acid inducible 11
AT5G67300	MYB44,MYBR1	myb domain protein r1
AT5G57420	IAA33	indole-3-acetic acid inducible 33
AT3G05630	PDLZ2,PLDP2,PLDZETA2	phospholipase D P2
AT4G27260	GH3.5,WES1	Auxin-responsive GH3 family protein
AT1G80390	IAA15	indole-3-acetic acid inducible 15
AT4G03190	AFB1,ATGRH1,GRH1	GRR1-like protein 1
AT5G37770	CML24,TCH2	EF hand calcium-binding protein family
AT5G65670	IAA9	indole-3-acetic acid inducible 9
AT3G60690		SAUR-like auxin-responsive protein family

Genes form GO enriched terms *GO:0009791* (post-embryonic development) and *GO:0009733* (response to auxin stimulus) from Table 8. Genes were annotated with The Arabidopsis Genome Integrative Explorer - PlantGenIE.org (Sundell *et al.*, 2015).

IAA15, *IAA9* and *SAUR59* are proposed as putative direct AtHB25 targets based on our analyses (Table 8). These results suggest an important role of AtHB25 in regulating the expression of key genes with a function in seed maturation and this assumed AtHB25-mediated gene regulation would be carried out mainly through binding to the promoter of the regulated gene.

AtHB25 binds to the canonical homeodomain-binding site

Using the HOMER Plants software (Heinz *et al.*, 2010), we performed a motif discovery analysis on the regions surrounding the 146 AtHB25 binding-peak centre (± 100 bp). *De novo* motif analysis revealed that the TRATTA sequence, which corresponds to colamp-ATHB25-DAP-Seq (GSE60143) motif (*p* value 10^{-34}), was present in 121 of 146 peaks (Figure 37c). It is important to highlight that this motif was present at least twice in the 60% of the binding-peaks and four repetitions of the motif were found in 25% of them (Figure 37d). Homer Known Motif Enrichment results showed four additional motifs in more than 70% of the sequences (Table 10) that contained the ZF-HD core Consensus-Binding Sequence ATTA (Tan and Irish, 2006). These motifs are named as ATHB34, ATHB24, ATHB23 and ATHB33 because they were identified in DNA affinity sequencing experiments (DAP-seq) with the corresponding in-vitro-expressed homeobox TFs (O'Malley *et al.*, 2016). In addition, most of these proteins do not contain an intrinsic activation domain, suggesting that interactions with transcription factors from different families could be required for transcriptional activation (Tan and Irish, 2006). Consistent with this idea, our analysis revealed 79 enriched motifs statically overrepresented that are known to be binding sites for TF that are not members of the Homeobox family (Supplemental data 6). Interestingly, 25 of these motifs are characteristic of the basic leucine zippers (bZIPs) TF. Nine of these correspond to bHLH class TF binding sites and eight of them have sequences that could be transcriptionally regulated by TEOSINTE BRANCHED / CYCLOIDEA / PROLIFERATING cell factor (TCP). Interestingly, bZIPs, bHLH and TCP TFs play key roles in plant development, environmental signaling and stress response (Duek and Fankhauser, 2005; Nicolas and Cubas, 2016; Dröge-Laser *et al.*, 2018). These results support the possibility that the ZF-HD proteins act in a combinatorial manner to modulate transcription, functioning as homo or heterodimers.

Table 10: Homeobox binding sites are enriched in Arabidopsis AtHB25 ChIP-seq peak regions.

Known TF motif	***Motif sequence***	***% in peak sequences***	***% in background sequences***	***p value***
AtHB25	TRATTAVB	82.99 %	37.64 %	10^{-28}
AtHB33	NGTRATTAAK	79.59 %	38.30 %	10^{-23}
AtHB23	HTAATTARNN	76.87 %	31.45 %	10^{-28}
AtHB34	TRATTARS	74.83 %	28.36 %	10^{-30}
AtHB24	TAATTAAS	72.11 %	27.04 %	10^{-29}

The table shows the first five results from motif enrichment analysis. Motif enrichment analysis was performed with HOMER program in 200 bp around AtHB25 ChIP-seq biding region centres. Background genome was randomly sampled from the Arabidopsis genome. TF: Trasncription factor.

Expression analysis of genes encoding lipid polyester biosynthesis enzymes

We validated the ChIP-Seq results with additional ChIP-qPCR experiments using specific primers around selected AtHB25 binding-peaks (Figure 37e and Figure 38), including those closer to *LACS2* and *NHO1*. *LACS2* is involved in the biosynthesis of leaf cutin and *lacs2* seeds show lower amounts of total lipids (Schnurr *et al*., 2004). *NHO1* encodes a glycerol kinase, which converts glycerol to glycerol 3-phosphate (G3P) (Eastmond *et al.*, 2004); its potential function in cutin or suberin biosynthesis has not been investigated. G3P is a substrate of glycerol-3-phosphate acyltransferases (GPAT), which is essential for cutin and suberin deposition (Li *et al.*, 2007; Yang *et al.*, 2010). As shown in Figure 37e and Figure 38, amplifications around peaks were more efficient in ChIP samples than in samples where no antibody was utilized. However, in random genomic areas, amplifications were similar in both cases (Figure 38).

Whereas our analysis clearly indicated that AtHB25 could directly regulate *LACS2* and *NHO1* transcription, it would be unusual that the high lipid polyester accumulation observed in the seed coat of AtHB25 gain-of-function mutants is solely explained by the regulation of these two genes. Therefore, we carried out a broad gene expression study of all polyester biosynthesis genes described and those involved in the formation of seed lipid barriers (Gou *et al.*, 2017). qRT-PCR was performed on three independent maturing-seed samples.

We could detect expression changes in 10 genes (Figure 37f and Figure 39).*LACS2* and *NHO1* were induced in the *AtHB25* gain-of-function mutant and repressed in *athb22 athb25* double mutant (Figure 37f). In addition, we found a positive correlation

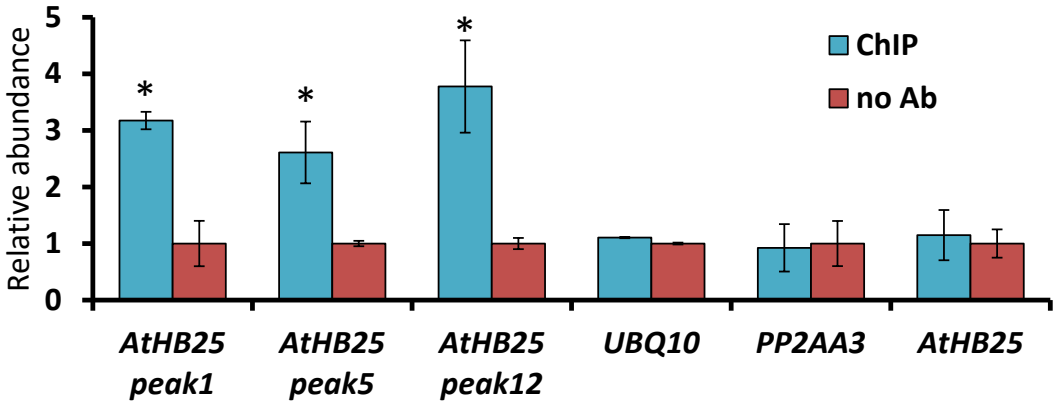


Figure 38: ChIP-qRT in siliques confirms the enrichment of different AtHB25 binding regions. Relative abundance of different peak regions and other areas (*UBQ10*, *PP2AA3* and *AtHB25* gene sequences). Expression values are normalized to the no Antibody (no Ab) sample as a negative control. Results are the average of three determinations on two independent chromatin immunoprecipitations. The error bars denote standard errors. *, $p < 0.05$.

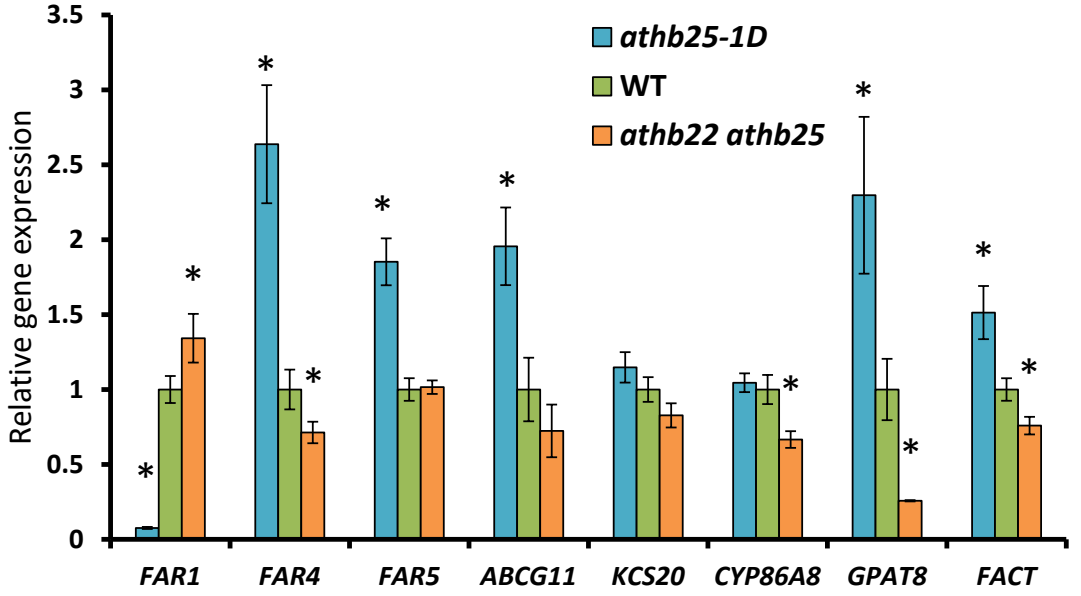


Figure 39: Gene expression analysis of lipid polyester biosynthetic enzymes in developing seeds of *athb25-1D*, WT and *athb22 athb25* at the maturation stage. Expression values are relative to the housekeeping gene *PP2A3* and normalized to wild type. Results are the average of three determinations. The error bars denote standard errors. *, $p < 0.05$

between *AtHB25* expression and *FAR4*, *GPAT8* and *FACT* (Figure 39). These results implicate AtHB25 in the regulation of the expression of these lipid polyester biosynthesis genes during seed maturation.

LACS2 and NHO1 are essential to AtHB25-mediated seed longevity

In our laboratory, we previously demonstrated that seed coat cell wall polyphenolics derived from suberin and/or lignin are key factors for tolerance to seed deterioration (Renard *et al.*, 2020*a*). In addition, a genome-wide association study (GWAS) of seed longevity identified CYP86A8 as being crucial for the maintenance of seed longevity and permeability (Renard *et al.*, 2020*b*).

To determine whether LACS2 and NHO1 are important for seed viability given their role in seed coat lipid polyester deposition, *lacs2* and *nho1* mutant seeds were subjected to a controlled deterioration treatment (CDT), a standard method to measure seed longevity (Renard *et al.*, 2020*a*). We used freshly collected fully germinating seeds to rule out problems in dormancy or desiccation tolerance and ensuring correct seed maturation. The experiment was performed after 7, 12, 15, 18, 22 and 28 days of treatment to accurately determine the putative sensitivity of both mutants. Again, germination

data were fitted to sigmoidal curves to estimate P50 (Figure 40a). *lacs2* showed a high sensitivity to artificial aging, since at all timepoints tested the mutant presented lower germination than the WT control. However, *nho1* mutant seeds presented sensitivity only after the 18-day treatment. Consistently, the calculated P50 in *lacs2* was half (10.1 ± 2.4) than that of WT (18.6 ± 3.14). P50 in *nho1* (15.5 ± .9) was not different than that observed in the WT (Figure 40b). In addition, as *LACS2* and *NHO1* are direct targets of AtHB25, it would be very interesting to know the impact of these mutations in the *athb25-1D* background. Thus, we next performed genetic crosses and CDT was performed on those lines. Both mutations (*lacs2* and *nho1*) affected the higher resistance of *athb25-1D* to seed aging, becoming more sensitive than WT. In fact, the P50 of *athb25-1D lacs2* and *athb25-1D nho1* were 9.8 ± 2.8 and 12.7 ± 2.4 days respectively (Figure 40a and b). Remarkably, this reduction in germination was associated with an increase in seed coat permeability in the double mutants, as shown in Figure 40c**.** We next delipidated the seeds and stained them with Sudan red. Whereas *lacs2* presented an observable reduction in staining intensity comparing with WT, this was not evident in *nho1* seeds (Figure 41). However, the loss-of function of both *lacs2* and *nho1* in the *athb25-1D* gain-of-function background caused a considerable reduction in the Sudan Red staining of the seeds (Figure 40d**),** suggesting a reduction in the lipid polyester content in the testa of the double mutants compared to *athb25-1D*.

Finally, double mutants of *athb25-1D* and other important genes involved in seed coat lipid polyester accumulation as *GPAT5* and *MYB107* were generated and seed longevity, permeability and lipid polyester staining were analyzed. *athb25-1D gpat5* and *athb25-1D myb107* displayed a decreased in seed longevity and staining as well as increase in seed coat permeability (Figure 42 and Figure 43). This result suggests that AtHB25 and MYB107 could be involved in the same regulatory pathway but additional experiments should be performed to test this hypothesis.

To investigate if changes in lipid staining correlate with changes in composition, we performed a chemical analysis of seed coat lipid polyesters. GC-MS analysis of the extracted monomers revealed significant differences in the total loads of *lacs2* (but not *nho1*) and WT. However, the increase of lipid polyesters normally observed in *athb25-1D* seed coats was reduced to wild type levels when the *nho1* mutation was added to the *athb25-1D* background (Figure 44). Analysis of individual monomer loads in *lacs2* seeds revealed significant reduction in the ferulate content, in 16:0 and 18:2 DCAs, the main monomers regulated by AtHB25 (Figure 36a), and in 24-OH 24:0 FA (Figure 45). Our GC-MS analysis also showed that the increased loads of ferulate, 18:2 and 22:0 DCAs and all 18 OHFA observed in *athb25-1D* were supressed in the double mutant *athb25-1D nho1* (Figure 46).

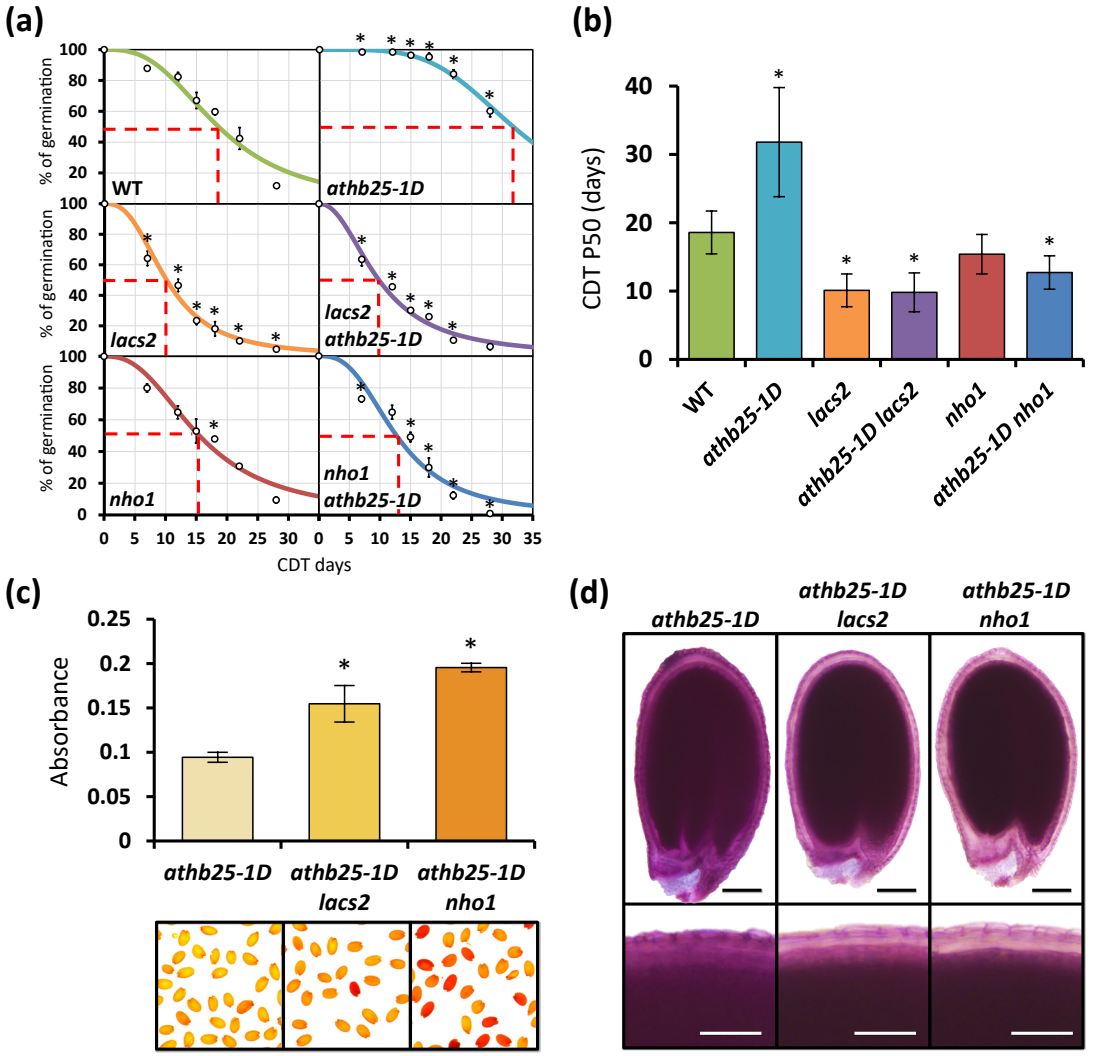


Figure 40: *LACS2* and *NHO1* are required for the increased seed longevity and the enhanced lipid deposition of *athb25-1D*. (a) Seed germination curves of the *Arabidopsis thaliana* wild type (WT) *athb25-1D, lacs2-3*, *athb25-1D lacs2-3* double mutant, *nho1* and *athb25-1D nho1* double mutant obtained aftercontrolled deterioration treatment (CDT) for 0, 7, 12, 15, 18, 22 and 28 days. The percentage of germination was recorded after seven days. The results are the average of three experiments with 100 seeds per line. Fitted probit and P50 value curve was calculated with the drc R package. Red dashed lines indicate estimated P50 values. (b) P50 values of each mutant genotype obtained from (a) data with the corresponding standard errors. (c) Seed coat permeability test of *athb25-1D*, *athb25-1D lacs2-3* and *athb25-1D nho1* double mutants seeds, and representative stained seed images. Colour of bars is representative of the formazan extraction colour. (d) Representative image of the staining of the lipid polyester layers of the seed coat with Sudan red 7B in *athb25-1D*, *athb25-1D lacs2-3* and *athb25-1D nho1* double mutant seeds. Bars, 100 µm. Details below. Bars, 50 µm. The error bars denote standard errors. *, $p < 0.05$.

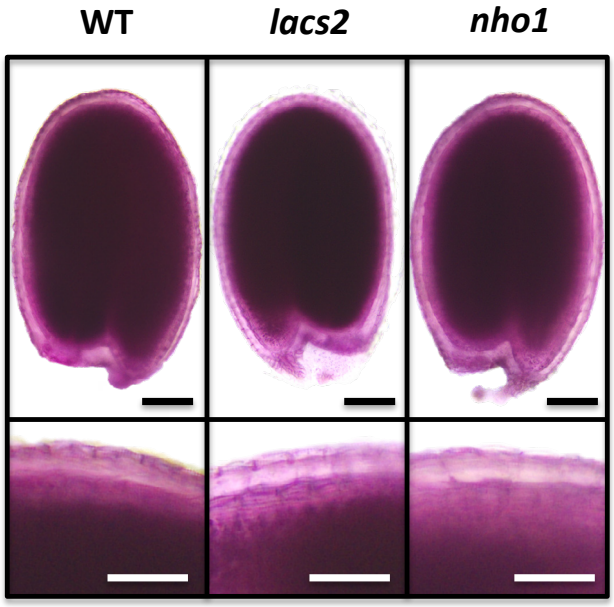


Figure 41: Lipid polyester staining in *lacs2* and *nho1* seeds. Representative image of the staining of the lipid polyester layers of the seed coat with Sudan red 7B in wild type (WT), *lacs2-3*, and *nho1* double mutant seeds. Bars, 100 µm. Detail (below). Bars, 50 µm.

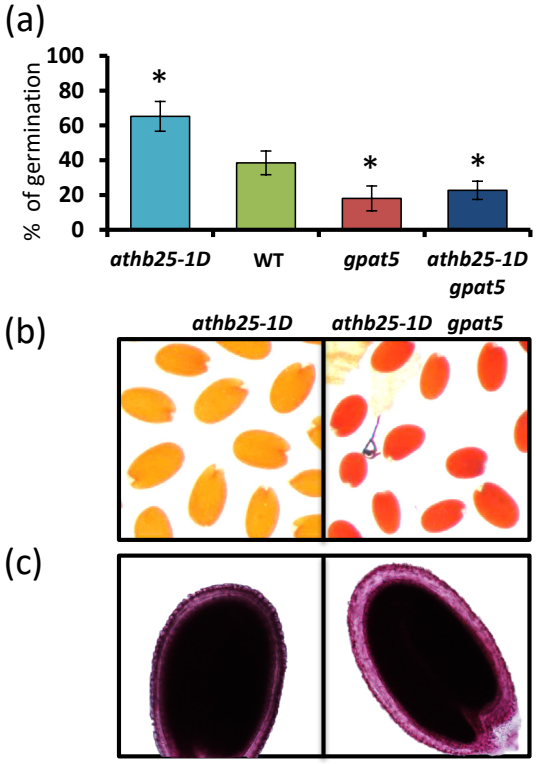


Figure 42: Seed longevity, permeability and Sudan Red 7B staining in *gpat5* and *athb25-1D gpat5* double mutant. (a) Arabidopsis seeds from *athb25-1D* mutant, wild type (WT), mutant seeds of *GLYCEROL-3-PHOSPHATE SN-2-ACYLTRANSFERASE 5* (*gpat5*) *and athb25-1D gpat5 double* mutant were subjected to controlled deterioration treatment (CDT) for 15 days. The percentage of germination was recorded after seven days. The results are the average of three experiments with 100 seeds per line. The error bars denote standard errors. *, $p < 0.05$. (b) Representative staining pattern of the seed coat permeability test in seeds of *athb25-1D* and *athb25-1D gpat5 double* mutant. (c) Representative image of the staining of the lipid polyester layers of the seed coat with Sudan red 7B in *athb25-1D* and *athb25-1D gpat5* double mutant seeds.

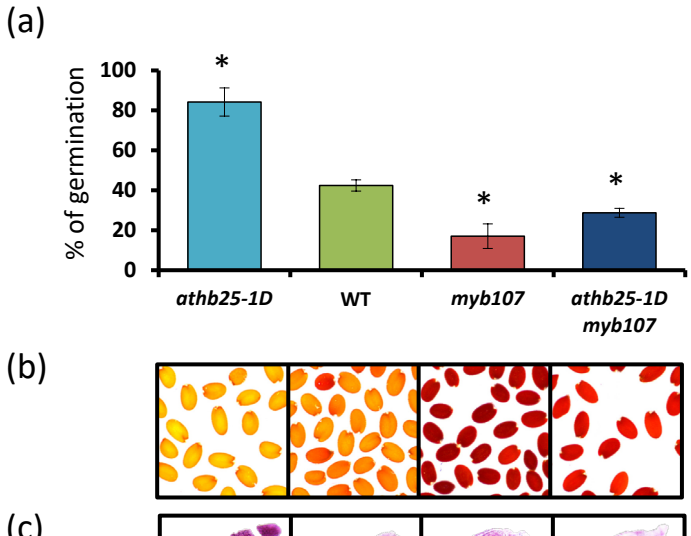


Figure 43: Seed longevity, permeability, and Sudan red 7B staining in *athb25-1D* and *athb25-1D myb107* double mutant. (a) Arabidopsis seeds from *athb25-1D* mutant, wild type (WT), mutant seeds of *MYB107* (*myb107*) and *athb25-1D myb107 double* mutant were subjected to controlled deterioration treatment (CDT) for 15 days. The percentage of germination was recorded after seven days. The results are the average of three experiments with 100 seeds per line. The error bars denote standard errors. *, $p < 0.05$. (b) Representative staining pattern of the seed coat permeability test in seeds of *athb25-1D*, WT, *myb107* and *athb25-1D myb107 double* mutant. (c) Representative image of the staining of the lipid polyester layers of the seed coat with Sudan red 7B in *athb25-1D*, wild type (WT), *myb107* and *athb25-1D myb107* double mutant seeds.

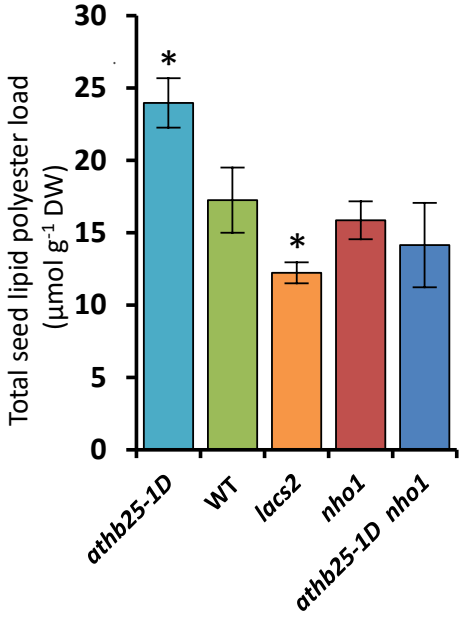


Figure 44: Lipid polyesters monomer loads in *lacs2*, *nho1*, *athb251-D lacs2* and *athb251-D nho1* mutant seeds. Total lipid polyester monomer content of *athb25-1D*, wild type (WT), *lacs2-1*, *nho1* and *athb25-1D nho1* double mutant. Bars represent the mean content of three biological replicates of each identified monomer and error bars represent standard deviations. *, $p < 0.05$.

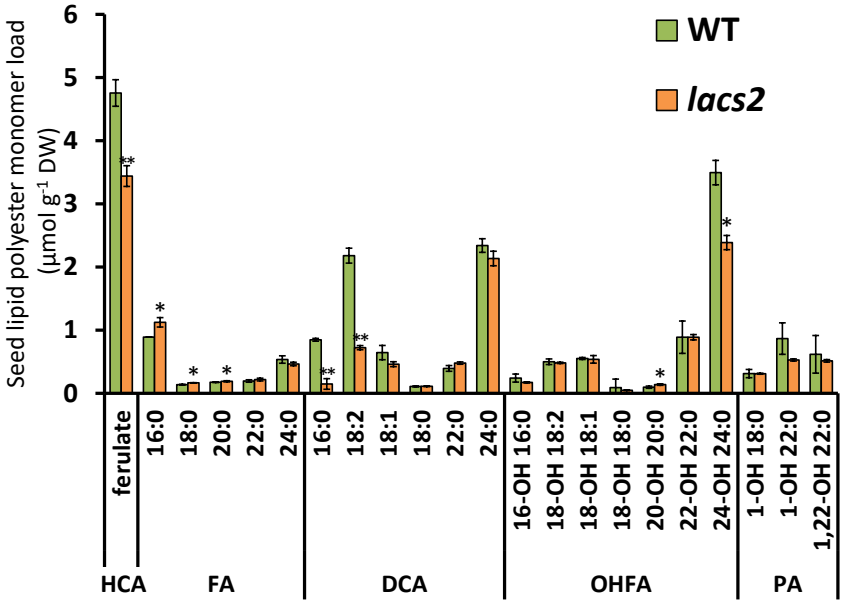


Figure 45: Comparison of lipid polyester monomer composition between WT and *lacs2-1* seeds. Bars represent the mean content of three biological replicates of each identified monomer and error bars represent standard deviations. Asterisks indicate significant differences with WT (*, $p < 0.05$; **, $p < 0.01$) between samples in unpaired *t* tests. HCA; Hydroxycinnamic acids; FA: fatty acids; DCA: a,ω-dicarboxylic acids; OHFA: ω-hydroxy fatty acids; PA: primary fatty alcohols.

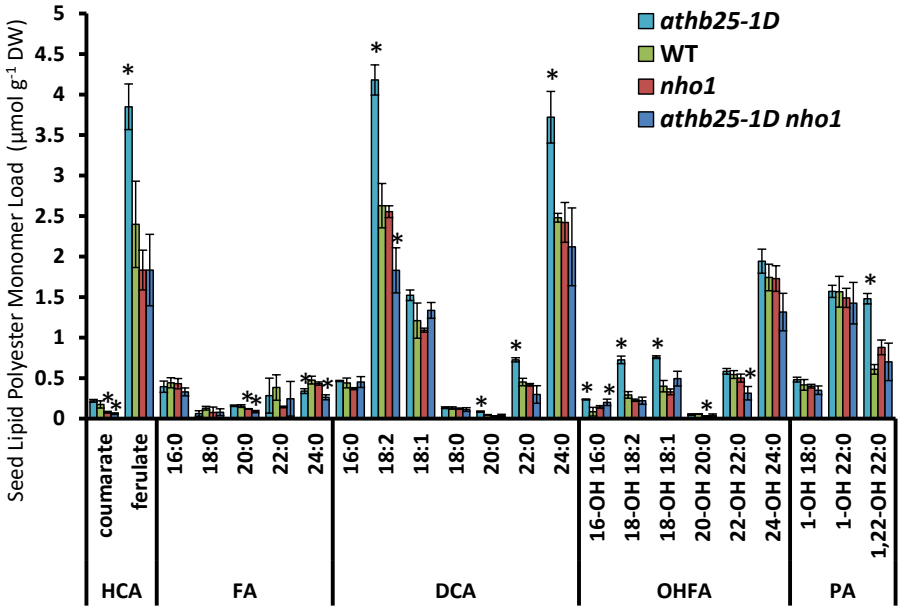


Figure 46: NHO1 is necessary for the *athb25-1D* increased lipid polyester deposition. Seed lipid polyester monomer composition in Arabidopsis seeds of *athb25-1D*, wild type (WT), *nho1* and *athb25-1D nho1* double mutant. Mean values of three biological replicates are shown in µmol g^{-1} delipidated dry residue (DW). The error bars denote standard deviations. *, $p < 0.05$ comparing to WT. HCA: hydroxycinnamic acids; FA: fatty acids; DCA: a,ω-dicarboxylic acids; OHFA: ω-hydroxy fatty acids; PA: primary fatty alcohols.

Cutin deposition in wheat caryopses enhances embryo longevity

To explore whether AtHB25 is a trans-species regulator of seed longevity through the control of the lipid polyester biosynthesis gene expression, we transformed a bread wheat cultivar, denoted as THA53, by particle bombardment with this Arabidopsis TF under the control of the maize *UBI1* promoter. Plants were selected after confirming the presence of the *BAR* gene by PCR and *AtHB25* expression was demonstrated by qRT-PCR. Plants from the T3 non-segregating transgenic line, the parental THA53, and BW208, another bread wheat cultivar, were grown under the same conditions and seeds were subjected to accelerated aging assays. Freshly collected seeds from all lines fully germinated. However, after the accelerated aging treatment, we observed differences in seed survival. Germination of the THA53 cultivar was below 20%, indicating that this line is very sensitive to this treatment. In addition, approximately 50% of the BW208 seeds maintained their viability, whereas THA53 transgenic lines expressing AtHB25 achieved close to 80% germination, showing marked resistance to the accelerated aging treatment (Figure 47a). Sudan Red staining on histological sections from caryopses showed a thin cuticle in THA53 grains, located in the outer part of the inner bran as described in Matzke and Rieder (1990), whereas BW208 presented a thicker cuticle than THA53 and the transgenic lines had a remarkably thicker cuticle than any of the control lines (Figure 47b). In order to quantify these visual differences and to determine what compounds are altered, we performed a chemical analysis of lipid polyesters. GC-MS analysis of the extracted monomers after chemical depolymerization revealed significant differences in the loads of specific monomers between the different lines. As expected, THA53 presented a reduction in some compounds. The most remarkable difference between this seed aging sensitive cultivar and the other two is that caryopses of this cultivar contained an approximately 30% reduction in one of the main compounds, 18-hydroxyoctadecenoic acid. Conversely, transgenic lines showed significantly increased amounts of coumarate, ferulate, 16:0 unsubstituted fatty acids, and 9-epoxy 18-hydroxyoctadecanoic acid (Figure 47c). The total amount of lipid polyesters, calculated by adding all the identified monomer quantities, revealed that caryopses overexpressing AtHB25 accumulated almost twice as much cutin as the THA53 cultivar. Total cutin in the transgenic line was also significantly higher than that of BW208 caryopses (Figure 47d). Taken together, we observed a positive correlation between resistance to aging, cuticle thickness and cutin amount, further supporting the results observed in Arabidopsis and extending these observations to an economically important crop species.

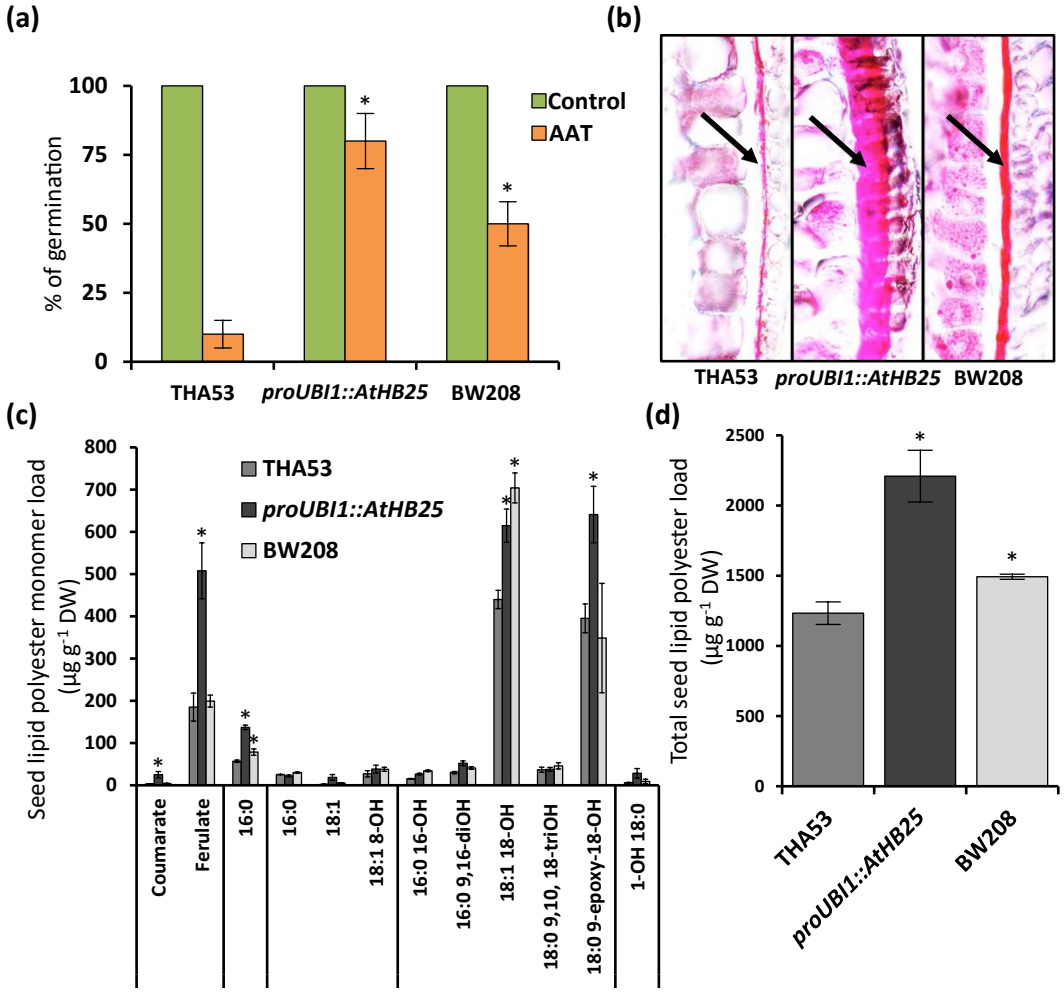


Figure 47: Ectopic expression of AtHB25 in wheat enhances seed longevity and lipid polyester barriers. (a) Wheat (*Triticum aestivum*) caryopses from THA53 cultivar, *proUBI1::AtHB25* transgenic lines (data is the mean of four independent *AtHB25* overexpressing lines under the maize *UBI1* promotor in THA53 background) and BW208 cultivar were germinated in control conditions and after an accelerated aging treatment (AAT, 41ºC for 24 hours at 100% relative humidity). The results are the average of three experiments with 20 seeds per line, and bars indicate standard errors. (b) Sudan red 7B staining of seed sections of THA53, *proUBI1::AtHB25* transgenic lines (image is representative of four independent *AtHB25* overexpressing lines in THA53 background) and BW208 to visualize the seed cuticle. Arrows indicate the lipid polyester barrier. (c) Seed lipid-polyester monomer composition of THA53, *proUBI1::AtHB25* and BW208 caryopses. (d) Total seed lipid polyester monomer content. Mean values of three biological replicates are shown in µg g^{-1} delipidated dry residue (DW). The error bars denote standard errors *, $p < 0.05$. HCA: hydroxycinnamic acids; FA: fatty acids; DCA: dicarboxylic acids; OHFA: hydroxyl fatty acids; PA: primary fatty alcohols.

An internal tomato seed coat cuticle enhances embryo longevity

To determine whether the function of AtHB25 in seed longevity is conserved in a dicot crop species, we examined AtHB25 function in tomato. Cotyledons from tomato plants (*Solanum lycopersicum* cv Moneymaker, *Tm2²*) were transformed with a construct designed to overexpress *AtHB25*. T3 overexpression homozygous lines were used to check seed viability. Around 90% of the non-treated seeds germinated. The radicle emergence test after accelerated aging treatment and seed bank storage (SBS) showed that *AtHB25* overexpression also confers extended longevity to tomato seeds (Figure 48a). Histological sections and Sudan red staining of tomato seeds revealed

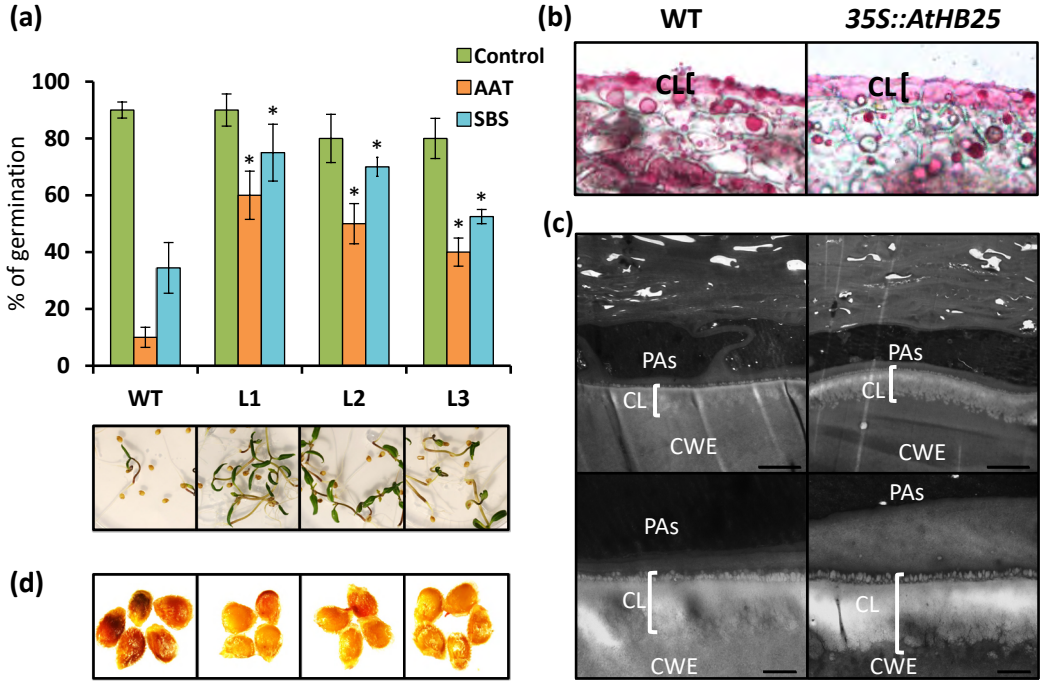


Figure 48: Ectopic expression of *AtHB25* in tomato enhances seed longevity, cuticle formation and reduces seed permeability. (a) Tomato (*Solanum lycopersicum* cv Moneymaker) seeds from three independent *35S::AtHB25* transgenic lines (L1, L2 and L3) and the control cultivar line were germinated in control conditions, after an accelerated aging treatment (AAT, 41ºC for five daysat 100% relative humidity) and after 36 months of seed bank storage (SBS). The results are the average of three experiments with 20 seeds per line. Representative pictures of the germination after SBS are at the bottom of each graph. The error bars denote standard errors. *, $p < 0.05$. (b) Sudan red 7B staining of a representative section of control and *35S::AtHB25* tomato seeds (outer integument was removed). Pictures are representative of analyses of 15 different seeds. CL: cuticle layer. (c) Transmission electron microscopy observation of representative control and *35S::AtHB25* tomato seed coat. Mature seeds were ultrathin sectioned, and image was focused to show seed coat cuticle layer. The pictures are representative of analyses of 10 different seeds. Brackets indicate the thickness of the cuticle layer. Top panels bars, 4 µm. Bottom panels bars, 1 µm. PAs: proanthocyanidins, CL: cuticle layer, CWE: cell wall of the endosperm. (d) Tetrazolium permeability test seeds WT and the three independent transgenic lines *35S::AtHB25* lines incubated in 1% tetrazolium red at 30 °C for 48 hours.

a significantly thicker stained layer in transgenic lines, as compared to control seeds, presumably corresponding to the cuticle (Figure 48b, note that outer seed coat tissues were not preserved in histological sections). To confirm that the stained area corresponds to the cuticle, we performed ultrafine sections and examined the tomato seed coat by TEM. The TEM image analysis revealed that the cuticle was thicker and more compact in *AtHB25* overexpressing lines (Figure 48c). Lastly, to determine whether this thicker barrier of cutin impedes seed-environment interactions, we measured tetrazolium salt uptake using the triphenyltetrazolium reduction method (Debeaujon *et al.*, 2000). Control seeds showed red coloration as a result of the production of red formazans, indicating a higher permeability of control seed coats to the substrate (Figure 48d). Conversely, transgenic lines remained yellow, suggesting that these seeds had lower permeability to tetrazolium salts. In summary, these results indicate that AtHB25 is able to improve seed longevity in tomato by reinforcing lipid barriers, similarly to our observations in Arabidopsis and wheat.

Discussion

Multiple reports have shown that seed longevity is a crucial biological parameter required to maintain plant biodiversity. Furthermore, understanding the genetic and environmental parameters that could modify this complex trait is essential for vegetable seed companies. Farmers, increasingly concerned about climate change, demand seeds with maximum vigor to maximize crop yield (Waterworth *et al.*, 2019).

The concept of seed longevity is highly complex and this is confounded by the variety of terminologies, criteria and different experimental methodologies that can be found in the literature. Some authors defend that for an artificial method to be comparable to natural aging, it must not break the glassy state in which the seeds are found after maturation, and thus the seed must not exceed a certain degree of humidity during the treatment (which depends on the species). In most cases, these precise determinations have not yet been performed. But the special conditions used in artificial treatments to determine seed longevity are typically seed bank conditions. In nature, seeds often have to cope with humid conditions that are nevertheless unfavorable for germination. Therefore, seed aging can also arise in wet environments. In any case, in the present study, we have employed artificial seed aging protocols, but more importantly, we studied the effect of AtHB25 in Arabidopsis seeds stored for up to 26 months, which would represent natural seed aging.

The plasticity of the seed longevity trait is another difficulty to deal with, as it is very strongly influenced by the environmental conditions in which the seed develops. Both abiotic and biotic stresses that affect the mother plant during seed development

are determinant for the seed lifespan (Nagel *et al.*, 2015). However, the number of studies specifically addressing this important point is very limited. For example, in Arabidopsis low temperature during seed development significantly decreased seed longevity through unknown mechanisms (He *et al.*, 2014). In terms of biotic stress, it has been demonstrated that virus infection in the mother plant does not always have a detrimental effect on seed viability (Bueso *et al.*, 2017). The main strategy that most seeds use to stay alive as long as possible when facing environmental stresses is to reach a desiccated state with the purpose of limiting the metabolic activity, which necessarily leads to oxidative stress and thus aging (Leprince *et al.*, 2016). Nevertheless, dry seeds also experience deterioration by oxidation processes, such as Amadori and Maillard reactions, lipid peroxidation and protein carbonylation, as a consequence of an oxygen-enriched atmosphere (Sano *et al.*, 2016). Seeds are set with a large battery of antioxidants including tocopherol, glutathione and LEA proteins to prevent these non-enzymatic oxidations; however, these compounds are finite and it is not possible to regenerate them. To avoid misusing these resources, the mother plant generates the seed coat from ovule integuments, a barrier formed by several layers that protects the embryo from changing environmental conditions and reduces oxygen diffusion.

Being part of the seed cover, lignin, condensed tannins and the lipid polyester barriers, cutin and suberin, have important roles in seed longevity. These compounds are highly regulated by environmental parameters and hormones synthesized in response to stress. For instance, low temperature induces the production of proanthocyanidins and suberin (MacGregor *et al.*, 2015; Fedi *et al.*, 2017). Abscisic acid (ABA), a hormone related to many abiotic stress conditions, is a well-established inductor of suberization in different species and organs (Soliday *et al.*, 1978; Cottle and Kolattukudy, 1982; Leide *et al.*, 2012). It is also known to regulate leaf cuticle formation upon drought stress and during leaf expansion (Martin *et al.*, 2017). Precisely ABA, and not the AtHB25 hormone-regulated gibberellic acid (Bueso *et al.*, 2014*b*), controls the acquisition of tolerance to desiccation during the maturation phase (Braybrook and Harada, 2008). However, it is also during this stage when AtHB25 expression is more evident in the seed coat inner integument (Figure 31d and e). The accumulation of antioxidants and the establishment of physical barriers to protect the embryo are two important events during the seed maturation phase. In this work, we demonstrate that AtHB25 controls the formation of lipid polyester barriers in the seed coat and this is likely to be one mechanism by which augmented AtHB25 activity confers less permeability to seed coats (Figure 31c) and reduced oxidation (Figure 31b).

In Arabidopsis, the expression of *AtHB25* is mostly localized to the inner integument, and mutants in this transcription factor display differences in the thickness of the cuticle (Figure 34f and g), suggesting that this transcription factor may control the formation of this apoplastic layer. This possibility is further supported by the lipid polyester

monomer profiles of *AtHB25* gain- and loss-of-function mutants, which reveal drastic changes in the levels of polyester monomers (in particular 18:2 α,ω-dicarboxylic acid) (Figure 36a). However, changes in DCA and fatty alcohol monomers were also observed, suggesting AtHB25 directly or indirectly controls seed coat suberization.

From the ChIP-sequencing results (Supplemental data 4 and Supplemental data 5), many candidates genes were identified that may explain the increased longevity in *AtHB25* overexpressing seeds, but none of these are related to GAs biosynthesis, probably because the seeds were collected in the maturation phase, and this hormone does not play an important role at this stage (Braybrook and Harada, 2008). Our analysis revealed increased expression of antioxidant genes, such as catalases CAT1 and CAT3 and one LEA protein; enzymes that could be involved in cell wall reinforcement, such as CESA6 or a xylose isomerase, since xylans are known to enhance seed coat impermeability (Mullin and Xu, 2001); and an aspartic protease with a demonstrated role in seed longevity, ASPG1 (Shen *et al.*, 2018). However, genetic crosses with loss-of-function *lacs2* or *gpat5* mutants showed that lipid polyester accumulation in seed coat is the main mechanism by which AtHB25 regulates seed longevity (Figure 40b and Figure 42).

Our study demonstrates that AtHB25 may regulate the biosynthesis of both polyesters (Figure 49), a possibility that is further substantiated by our gene expression analysis. This TF binds to regulatory sequences of *LACS2* (Figure 37e) during seed maturation, positively regulating its transcription. But it is also demonstrated that AtHB25 regulates the expression of *FAR4*, *GPAT8* and *FACT* by undetermined mechanisms (Figure 39).

Interestingly, LACS2 is necessary for cutin synthesis (Schnurr *et al.*, 2004; Zhao *et al.*, 2019), and the lipid polyester monomer profile of the loss-of-function mutant is in part consistent with an activating function of the LACS2 gene by AtHB25 in *athb25-1D*. More specifically, *athb25-1D* shows increased levels of ferulate and 18:2 DCA (Figure 36a), whereas the *lacs2* mutant shows decreased levels of these compounds (Figure 45), thus indicating that *LACS2* is a possible target of AtHB25.

On the other hand, NHO1, encodes a protein similar to glycerol kinase, which converts glycerol to glycerol 3-phosphate (G3P). G3P is also a substrate of the GPAT enzymes that produce monoacylglycerol intermediates in the synthesis of cutin and suberin. However, *nho1* mutants did not show significant changes in their seed lipid polyester composition compared to WT (Figure 44), and thus it is not possible to speculate on the specific role of NHO1 in the formation of cell wall-associated lipid polymers during normal seed development. Nevertheless, the *nho1* mutation suppresses the phenotype displayed by *athb25-1D* overexpression mutants, suggesting a role for NHO1 in the pathway activated by AtHB25 that leads to lipid polyesters accumulation in the seed coat (Figure 44).

Homeodomain proteins are conserved among species and regulate their specificity by interacting with other proteins that control cell and tissue-specific gene expression (Bürglin and Affolter, 2016). Many homeobox genes encode transcription factors with regulatory roles in animal and plant development (Holland, 2013). Because the seed coat is the main structure protecting the embryo from the hostile environment in spermatophytes, it is plausible that there should be a conserved TF regulating the most abundant compounds found in this cover; lignin, tannins and the lipid barriers, cutin and suberin. Many TFs with relevant functions in Arabidopsis have been described,

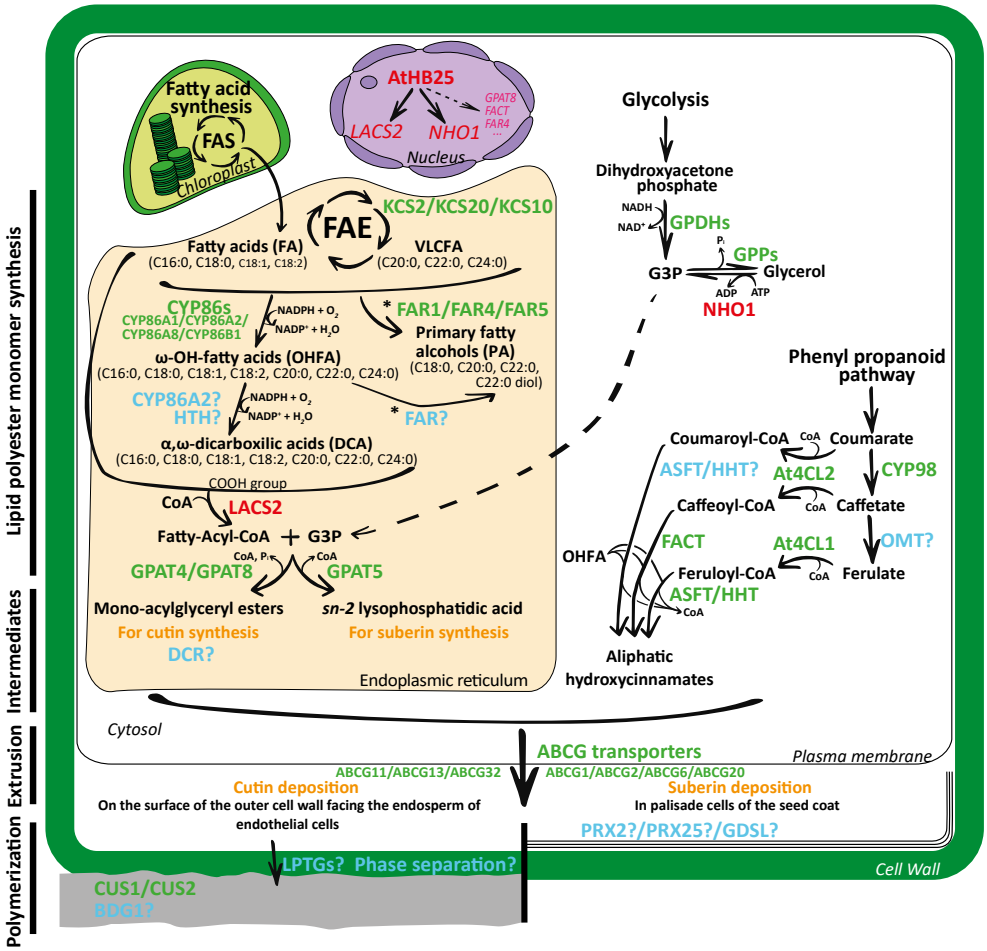


Figure 49: Scheme showing the biosynthesis of suberin and cutin and the effect of AtHB25 regulation. Simplified biosynthetic pathways of suberin and cutin monomers according to current views. Enzymes described to be involved are illustrated in green. Enzymes that are not still biochemically characterized are shown in blue with a question mark. The effect of AtHB25 directing regulating LACS2 and NHO1 is illustrated in red. Monomers and intermediates are extruded by ABCG transporters. While suberin deposition locates in the inner side of the cell wall and shows a lamellae pattern, the amorphous cutin deposition occurs in the outer side of the cell wall. Enzymes involved in the polymerization are not completely dilucidated. FAS: Fatty Acid Synthesis complex, FAE: Fatty Acid Elongation complex, VLCFA: Very long chain fatty acids, CoA: Coenzyme A; G3P: Glycerol 3-phosphate.

but the study of their orthologues in different species is limited. Our work constitutes one of the first studies that identify an Arabidopsis TF with conserved functions in distant species, namely tomato and wheat, suggesting that orthologous genes could have a similar role in seed coat development. The best AtHB25 orthologue in wheat is TraesCS5A02G246500 with an identity of 38% and a similarity of 56% at the protein level. Similarly, Solyc04g0142260.1, scoring 40% identity and 60% similarity at the amino acid sequence level is the best candidate we found in tomato (Figure 50). Interestingly, the amino acid similarity within the Zinc Finger region is 96% for the wheat homologue and 98% for the tomato homologue, whereas in the Homeobox domains (Tan and Irish, 2006) the similarities are 92.2% and 90.2%, respectively. Thus, given this high level of homology in the functional domains of these transcription factors and our data showing the cross-species activity of AtHB25, it is very likely that the function of these transcription factors is conserved across species. It will be a fascinating task to confirm whether these proteins really regulate the lipid polyester deposition in these species. The present study establishes the role of AtHB25 as a trans-species regulator of seed longevity and substantially adds to our knowledge regarding the molecular mechanism through which this TF achieves this regulation. These results may be directly applied to improve this important trait in economically-relevant crop species, such as wheat and tomato.

```
TraesCS5A02G246500      MDFDDHDDGDEEMRMSMP-----MPVSSSYEPPPPLSTGFGGAGVAPSKPPGGGGGGGGG
Solyc04g014260.1        MEFEDQQKQRDEE----------------IAAAPRHNDSLDNSDLTTKMPPSP--QLELE
At5g65410               MEFEDNNNNNDEEQEEDMNLHEEEEDDDAVYDSPPLSRVLPKASTESHETTGTTSTGGGG
                        *:*:*::.  :*                    .*  .  :  :.  .  ...

TraesCS5A02G246500      ELGGRAKAPGGGARYRECLKNHAVGIGGYAVDGCGEFIAGGEEGSIDALRCAACSCHRNF
Solyc04g014260.1        PLTAVQLWTNNKPKYKECLKNHAVGIGGHAVDGCGEFLPAGEDGSIDSLKCAACNCHRNF
At5g65410               GFMVVHGGGGSRFRFRECLKNQAVNIGGHAVDGCGEFMPAGIEGTIDALKCAACGCHRNF
                         :       ..  :::*****:**.***:********:..* :*:**:*:****.*****

TraesCS5A02G246500      HRRESDFPAGGEGSPFSPTAMVPYGGVPPHHQFSPYYRTPAGYLHHHQHHMAAAAAAAAA
Solyc04g014260.1        HRKIAPPPIAAAAAGVGGEPVP--FVYHSHNQLPTYYRTLPPPCGYLQYHVAP-------
At5g65410               HRKELPYFHHAPPQHQPPPPPP------------GFYRLPAPVSYRPPPSQAP-------
                        **:       .        .               :**  .          *.

TraesCS5A02G246500      GHPRPLALPSTSHSGRDDADELSGMAAGPMSALVPLSSMSLGAGPSGYGSGSGSGKKRFR
Solyc04g014260.1        -NQRPLALPSTSGGYREDQEDIS------------------------HPNYSGGSKKRFR
At5g65410               --PLQLALPPPQRERSEDPMETS------------------------SAEAGGGIRKRHR
                             ****...    :*  : *                          . .*. :**.*

TraesCS5A02G246500      TKFTQEQKDKMLAFAERVGWRIQKHDEAAVLQFCDEVGVKRHVLKVWMHNNKHTLGKKPP
Solyc04g014260.1        TKFSQVQKDKMQELADKLGWRIQREDEELVQQLCNETGITRQVFKVWMHNNKHTLGKKP-
At5g65410               TKFTAEQKERMLALAERIGWRIQRQDDEVIQRFCQETGVPRQVLKVWLHNNKHTLGKSPS
                        ***:  **::*  :*:::*****:.*:  : ::*:*.*: *:*:***:*********.*

TraesCS5A02G246500      SI----------------------
Solyc04g014260.1        ------------------------
At5g65410               PLHHHQAPPPPPPQSSFHHEQDQP
```

Figure 50: Alignment of AtHB25 protein sequence with orthologues proteins in wheat and tomato. The protein sequence of the Arabidopsis protein AtHB25 (At5g65410) is aligned with Clustal W to the tomato orthologue Solyc04g0142260.1 and the wheat orthologue TraesCS5A02G246500 previously identified with Plaza monocot 4.5.

Chapter 3

Plant, Cell & Environment (2020) **43**:315–326 DOI: 10.1111/pce.13656

PRX2 and PRX25, peroxidases regulated by COG1, are involved in seed longevity in Arabidopsis

Joan Renard[1], Irene Martínez-Almonacid[1], Annika Sonntag[2], Isabel Molina[2], José Moya-Cuevas[1], Gaetano Bissoli[1], Jesús Muñoz-Bertomeu[1], Isabel Faus[1], Regina Niñoles[1], Jun Shigeto[3], Yuji Tsutsumi[4], José Gadea[1], Ramón Serrano[1] and Eduardo Bueso[1*]

[1]***Instituto de Biología Molecular y Celular de Plantas, Universitat Politècnica de València-Consejo Superior de Investigaciones Científicas, 46022 València, Spain***

[2]***Department of Biology, Algoma University, Sault Ste Marie, ON, Canada, P6A 2G4***

[3]***Incubation Center for Advanced Medical Science, Kyushu University, Fukuoka 819-0395, Japan***

[4]***Faculty of Agriculture, Kyushu University, Fukuoka 812-8581, Japan***

*corresponding author (e-mail: edbuero@ibmcp.upv.es)

Abstract

Permeability is a crucial trait that affects seed longevity and is regulated by different polymers including proanthocyanidins, suberin, cutin and lignin located in the seed coat. By testing mutants in suberin transport and biosynthesis, we demonstrate the importance of this biopolymer to cope with seed deterioration. Transcriptomic analysis of *cog1-2D*, a gain-of-function mutant with increased seed longevity, revealed the upregulation of several peroxidase genes. Reverse genetics analyzing seed longevity uncovered redundancy within the seed coat peroxidase gene family; however, after controlled deterioration treatment, seeds from the *prx2 prx25* double and *prx2 prx25 prx71* triple mutant plants presented lower germination than wild-type plants. Transmission electron microscopy analysis of the seed coat of these mutants showed a thinner palisade layer, but no changes were observed in proanthocyanidin accumulation or in the cuticle layer. Spectrophotometric quantification of acetyl bromide- soluble lignin components indicated changes in the amount of total polyphenolics derived from suberin and/or lignin in the mutant seeds. Finally, the increased seed coat permeability to tetrazolium salts observed in the *prx2 prx25* and *prx2 prx25 prx71* mutant lines suggested that the lower permeability of the seed coats caused by altered polyphenolics is likely to be the main reason explaining their reduced seed longevity.

Introduction

The progressive oxidation of nucleic acids, proteins and lipids of the embryo by external environmental sources accelerates seed deterioration (Bailly *et al.*, 2008). One of the most important mechanisms in plants to prevent the impact of ROS on embryo integrity is the formation of the seed coat, a structure that functions as a barrier between the embryo and its external environment, thus conferring chemical and physical protection (Sano *et al.*, 2016).

In Arabidopsis, the outermost epidermal cells of seed coat secrete mucilage, a pectic polysaccharide, that has high water-retention properties, but does not appear to play a role in seed deterioration (Haughn and Chaudhury, 2005). Below the epidermis is the palisade layer followed by the seed coat endothelium, which is the innermost layer and the site of proanthocyanidin synthesis. These flavonoid compounds are antioxidants that act as ROS scavengers. On the other hand, the structure and biosynthesis of lignin have received little attention with regard to seed coat location. Some testa cells develop lignified secondary cell walls that are thought to reinforce the cell and make the seed coat impermeable to water and gas (Tobimatsu *et al.*, 2013, Liang *et al.*, 2006).

In addition, cell walls of the palisade layer of Arabidopsis seed coats contain suberin and the cell walls of the inner integument contain cutin (Molina *et al.*, 2008), These polymers constitute a lipophilic cell wall barrier that controls the flux of gases, water and solutes. In particular, suberin is a complex biopolymer consisting of a polyaliphatic domain (polyester) and a polyphenolic domain (Bernards *et al.*, 2002). Although a role for suberin has been described in relation to different biotic and abiotic stresses (Franke *et al.*, 2005; Franke and Schreiber 2007; Vishwanath *et al.*, 2015), no studies have linked this complex molecule with seed longevity. Since suberin in the seed coat could limit the diffusion of atmospheric oxygen into the seed, it might play an important role in preventing oxidation and associated deteriorative processes that contribute to seed aging.

The order of the reactions in suberin biosynthesis has yet to be definitively elucidated, but several enzyme families are known to be involved, including β-ketoacyl-CoA synthases (KCS), fatty acid oxidases of the cytochrome P450 monooxygenases (CYP), fatty acyl reductases (FAR), glycerol-3-phosphate acyl-transferases (GPAT), and HXXXD-motif family hydroxycinnamoyl-CoA transferases (ASFT or FHT) (Pollard *et al.*, 2008; Li-Beisson *et al.*, 2013; Vishwanath *et al.*, 2015). *GPAT5* acquires relevance, since it encodes a glycerol-3-phosphate acyltransferase specifically involved in the synthesis of acylglycerol precursors of the suberin polymer. Seed coat lipid polyester analysis of the *gpat5* mutant revealed a strong reduction in 22:0/24:0 fatty acids and their derivatives, in agreement with a reduction in seed coat suberin content observed after Sudan Red staining (Beisson *et al.*, 2007). Several of these sub-

erin biosynthesis genes are positively regulated in seeds by MYB107, a transcription factor (TF) that directly binds directly to their promoters (Gou *et al.*, 2017). On the other hand, suberin transport to the cell wall may occur through the secretory pathway and, in part, also mediated by ATP-binding cassette (ABC) transporters and/or lipid transfer proteins (LTPs). In *Arabidopsis*, an *abcg2 abcg6 abcg20* triple mutant was shown to have alterations in the structure, composition and properties of root and seed coat suberin (Yadav *et al.* 2014). In addition, an important role for peroxidases in the macromolecular assembly of the suberin polyphenolic domain at the cell wall has also been postulated (Bernards *et al.*, 2004). Peroxidases are heme-containing proteins that catalyze the reduction of H_2O_2 by transferring electrons from various donor molecules and, in the case of plant peroxidases (class III), these acceptor molecules can be very different, for example lignin precursors or the growth hormone auxin (Passardi *et al.*, 2004). This type of peroxidase gene family appeared with the colonization of land by plants and consists of a large (73 genes in *Arabidopsis thaliana*) and varied multigene family in land plants (Duroux and Welinder, 2003). Class III peroxidases participate in a broad range of physiological processes such as lignin formation, synthesis of phytoalexins or metabolism of ROS (Almagro *et al.*, 2009). Our results suggest a new function for these peroxidases in the maintenance of seed viability likely due to their role in the polymerization of suberin polyphenolics and lignin in the seed coat.

Results

Suberin accumulation in the seed coat subepidermal layer is a key factor for tolerance to seed deterioration

Seed storability can be measured by many different approaches, however the International Seed Testing Association (ISTA) has defined very specific protocols for accelerated aging treatment (AAT), which is thought to reflect seed vigor, and for controlled deterioration treatment (CDT), the standard method to measure seed longevity (ISTA, 2018). COG1 (At1g29160) is a TF that positively regulates the tolerance of *Arabidopsis* seeds to deterioration. The most likely mechanism of the gain of function mutant *cog1-2D* is the high accumulation of suberin in the palisade layer of the seed coat (Bueso *et al.* 2016). To confirm this hypothesis, we carried out a deterioration test of full germination capacity of seeds from mutants affected in suberin content. We used *cog1-2D* seeds as the resistant mutant control in an accelerated aging treatment (AAT). As expected, we observed an increase in germination for this mutant, as compared to WT plants (80 versus 50 %, respectively; Figure 44). In addition, *gpat5, abcg2 abcg6 abcg20* and *myb107* mutants, all showing altered seed coat aliphatic suberin composition, were also tested. All three mutants with defects in suberin composition presented

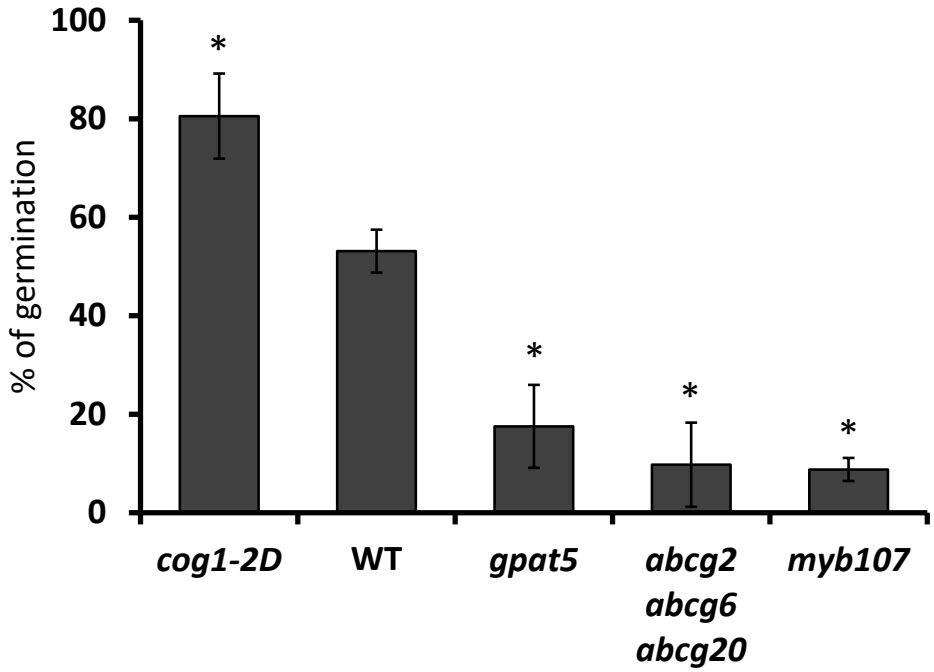


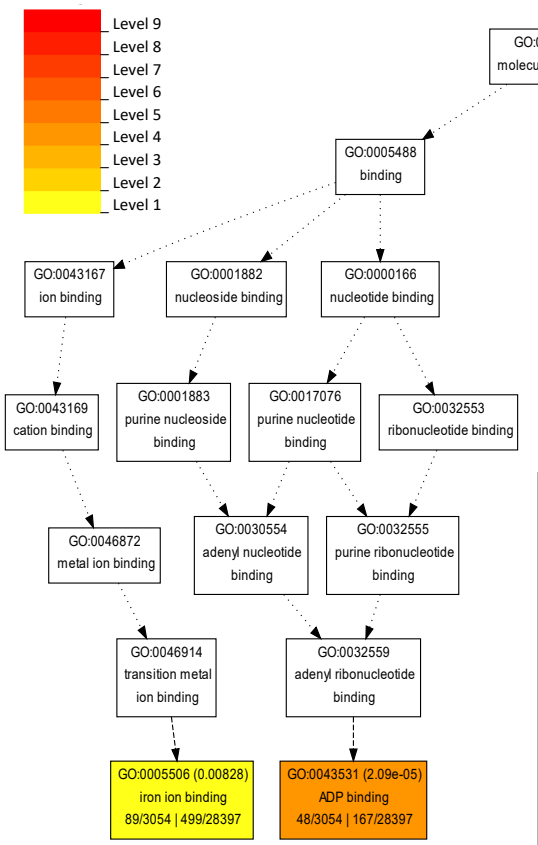
Figure 51: Mutants with less suberin in the palisade layer of seed coat are sensitive to seed deterioration treatment. Seeds from wild type, activation tagging *cog1-2D*, glycerol-3-phosphate acyltransferase (*gpat5*), ABCG half-transporters *(abcg2 abcg6 abcg20*) and MYB107 transcription factor (*myb107*) mutants were subjected to seed deterioration treatment for 24 h at 41 °C and sown on Murashige and Skoog plates. The percentage of germination was recorded after one week. The results are the average of four experiments with 100 seeds per line, and the bars indicate standard errors. Not-aged seeds from all lines germinated more than 99% after 3 days. *Significantly differing from controls (wild type) at $p < 0.05$ (Student's *t* test)

100% germination in non-aged seeds, but were highly sensitive to AAT with less than 25% of the seeds able to germinate one week after sowing (Figure 51). These results suggest that suberin accumulation in the palisade layer of the outer integument is important for tolerance to seed deterioration during aging.

COG1 regulates peroxidase activity

It is unclear how the COG1 TF could regulate suberin accumulation in the palisade layer of the seed coat, however the participation of growth and stress hormones is expected (Bueso *et al.*, 2016). To shed light on the global effect of the overexpression of this TF in the whole plant, we sequenced the mRNA in both WT and COG1-overexpressing seedlings. The analysis of this data (FDR < 0.05) revealed 3124 differentially expressed genes (DEGs) in *cog1-2D*. Of these, 1360 genes were upregulated and 1764 genes were downregulated (Supplemental data 7). In order to determine the processes affected in this gain of function mutant, we searched for over-represented

gene ontology (GO) terms within differentially expressed genes and performed Singular Enrichment Analysis (SEA) as implemented in the agriGO website. Regarding "biological process", we detected 90 GO terms significantly over-represented in the sets of DEGs in *cog1-2D* using a significance level of 0.01 (Supplemental data 8). Interestingly, some enriched GO terms could be related to suberin biosynthesis. For instance, related to the polyaliphatic fraction, we identified the "lipid metabolic process" (GO:0006629), and "carboxylic acid biosynthetic process" (GO:0046394) categories. In addition, our analysis revealed 71 DEGs classified as "phenylpropanoid biosynthetic process" (GO:0009699), which are genes that could be directly involved in suberin and lignin synthesis (Supplemental data 9). Phenylpropanoids are a diverse family of organic compounds, which are all initially synthesized via the shikimate pathway, with 4-coumaryl CoA as a central metabolite that provides the basis for all subsequent branches and resulting compounds (Vogt, 2010). In addition, our analysis identified a glycerol-3-phosphate acyltransferase, two cinnamyl alcohol dehydrogenases, several



Genomic locus	*Protein name*
AT1G05260	PRX3
AT1G30870	PRX7
AT2G34060	PRX19
AT2G35380	PRX20
AT3G21770	PRX30
AT3G28200	PRX31
AT3G49110	PRX33
AT3G49120	PRX34
AT4G11290	PRX39
AT4G26010	PRX44
AT4G37520	PRX50
AT4G37530	PRX51
AT5G15180	PRX56

Figure 52: COG1 regulates peroxidase activity in seedlings. Molecular function chart resulted from mRNA sequencing comparison analysis using agriGO between seedlings wild type and *cog1-2D* mutant. Inserted chart represents 13 peroxidases differentially expressed between both genotypes.

fatty acyl-CoA reductases, and putative and described laccases as DEGs. Although class III family peroxidases are not included in GO:0009699, these enzymes catalyse the oxidation of phenylpropanoids to their phenoxyl radicals, and the subsequent non-enzymatic coupling controls the pattern and extent of polymerization (Russell *et al.*, 2006). For instance, using the KEGG PATHWAY, peroxidases are included in the "phenylpropanoid biosynthesis" group (ath00940) and we observed that both *PRX33* and *PRX39* were induced at least two-fold (Supplemental data 10). In accordance with this observation and regarding the "molecular function" the most significant group was the GO:0016491: "oxidoreductase activity", where 13 class III peroxidases were identified as DEGs (Figure 52). This result points to peroxidases as a COG1-regulated gene family.

prx2 prx25 and *prx2 prx25 prx71* are sensitive to accelerated aging treatments

The RNA sequencing analysis from *cog1-2D* seedlings displayed differentially expressed peroxidases genes that are highly expressed in this vegetative stage, but not in seeds. Thus, these isoforms are unlikely to play an important role in the seed coat formation. However, we hypothesized that peroxidases expressed in the seed coat may have a physiological role in determining seed longevity. Therefore, we performed a reverse screening of loss-of-function mutants of peroxidases with high expression in seed coat based on data provided by Belmonte *et al.* (2013) and represented in eFP browser (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi). We isolated homozygous mu-

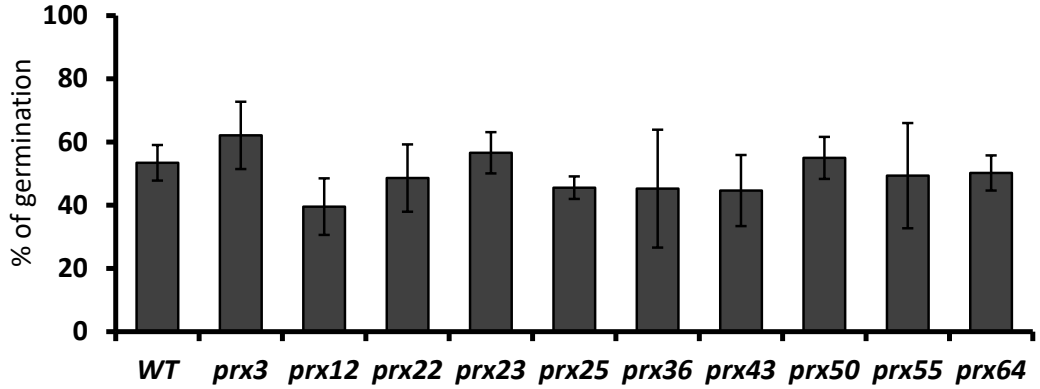


Figure 53: Simple peroxidases mutants do not show differences in seed deterioration resistance. Seeds from wild type (WT) and seed coat peroxidases *prx3*, *prx11*, *prx12*, *prx21*, *prx22*, *prx23*, *prx25*, *prx36*, *prx43*, *prx50*, *prx5*, *prx55* and *prx64* mutants were subjected to accelerated aging treatment for 24 h and sown on MS plates. The percentage of germination was recorded after 1 week. The results are the average of four experiments with 100 seeds per line, and bars indicate standard errors. Not-aged seeds from all lines germinated more than 99% after 3 days.

tant lines of peroxidases 3, 12, 22, 23, 25, 36, 43, 50, 55 and 64 and sensitivity to deterioration was checked using an AAT treatment. The viability of the seeds of all these single mutants was the same as the wild type (Figure 53).

We next conducted experiments to establish which of these seed coat peroxidases are regulated by COG1 during seed formation. According to databases, COG1 is highly

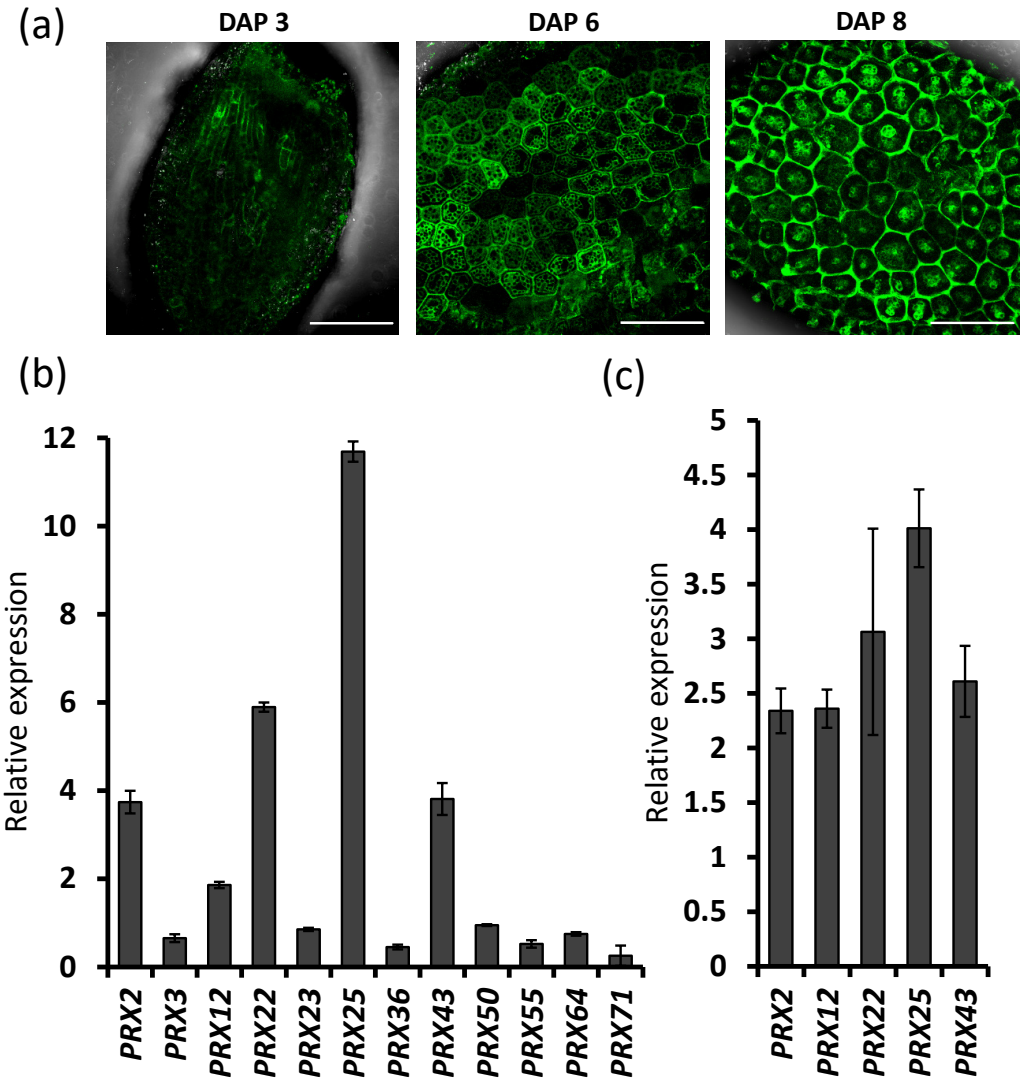


Figure 54: Gene expression analysis of COG1 in developing seeds and seed coat peroxidases in *cog1-2D* and GA3 treated plants. (a) Fluorescence images of 3, 6 and 8 days after pollination Arabidopsis seeds transformed with *proCOG1::GUS:GFP* using confocal laser scanning microscopy. Scale bar, 100 µm. (b) Quantitative real-time polymerase chain reaction analysis of expression of *PRX2, PRX3, PRX12, PRX22, PRX23, PRX25, PRX36, PRX43, PRX50, PRX55, PRX64* and *PRX71* from *cog1-2D* mutant *one*-week-old siliques. (c) Quantitative real-time polymerase chain reaction analysis of expression of *PRX2, PRX12, PRX22, PRX25* and *PRX43* from GA treated wild type *one*-week-old siliques. Expression values are relative to housekeeping genes, and the resulting ratios are normalized to wild type, taken as one. Results are the average of three determinations with bars corresponding to standard errors

expressed in the seed coat at the end of embryogenesis. To determine with more precision the spatial-temporal expression of *COG1* during seed development, we generated transgenic plants expressing green fluorescent protein (GFP) under the control of the COG1 promoter. On the third day, COG1 expression starts to be detected in seeds and at the end of embryo differentiation the expression of this TF is higher in the columella and in the radial walls of cells that form the seed epidermis (Figure 54a). Thus, 6-8 days after pollination (DAP), wild type and *cog1-2D* siliques were collected and seed coat peroxidase expression was analysed by qRT-PCR. The COG1 TF positively regulated seed coat peroxidases *PRX12*, *PRX22*, *PRX25* and *PRX43* (Figure 54b).

On the other hand, overexpression of COG1 increases gibberellin levels during seed maturation (Bueso *et al.*, 2016) and some peroxidases are induced by exogenous application of GA3 (Zieslin and Ben-Zaken 1992; Wang *et al.*, 2017*a*). To ascertain whether COG1-regulated seed coat peroxidases are induced by this growth hormone, siliques from GA3-treated plants were collected and peroxidase gene expression was analysed. All these peroxidases were induced after this treatment (Figure 54c).

PRX25 was the most expressed in *cog1-2D*, but seeds from this loss-of-function mutant were not affected in seed viability (Figure 46). Shigeto *et al.*, 2013 reported that *PRX25* and two homologues genes, *PRX2* and *PRX71,* are involved in stem lignification. *PRX2* was also induced in *cog1-2D* and in GA3-treated siliques more than 3-fold,

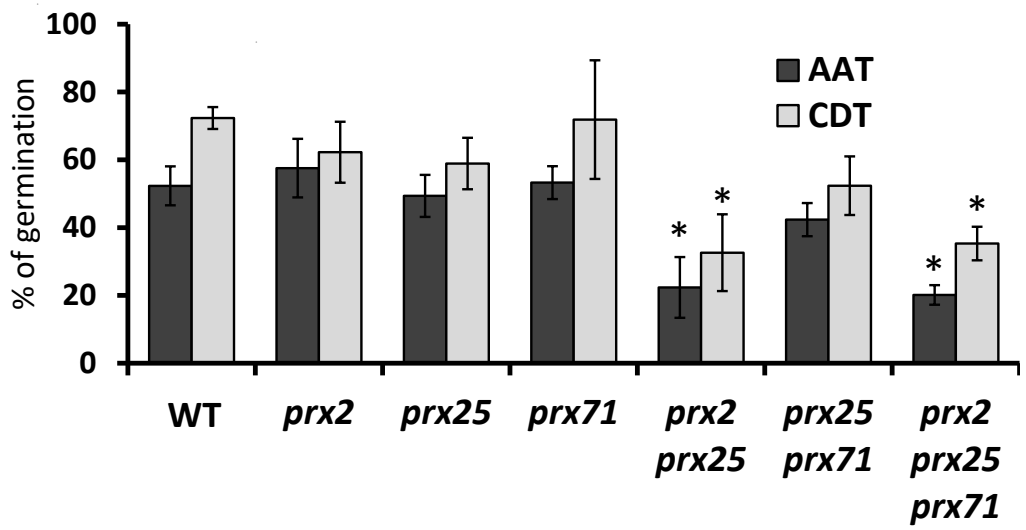


Figure 55: Double mutant *prx2 prx25* and triple mutant *prx2 prx25 prx71* are sensitive to *artificial* aging treatments. Seeds from wild type (WT), *prx2, prx25, prx71*, *prx2 prx25*, *prx25 prx71* and *prx2 prx25 prx71* mutants were subjected to *accelerated aging* treatment *(AAT)* for 24 h at 41 °C (black bars) and to controlled deterioration treatment *(CDT)* for 14 days (grey bars) and sown on Murashige and Skoog plates. The percentage of germination was recorded after *one* week. The results are the average of four experiments with 100 seeds per line, and the bars indicate standard errors. Not-aged seeds from all lines germinated more than 99% after3 days. *Significantly differing from controls (wild type) at $p < 0.05$ (Student's *t* test*)*.

while *PRX71* was clearly repressed in this activation-tagging mutant (Figure 54b and Figure 54c).

In order to find out whether these peroxidases are involved in seed viability, single loss-of-function mutants *prx2*, *prx25, prx71,* double mutants *prx2 prx25* and *prx25 prx71* and the *prx2 prx25 prx71* triple mutant were subjected to both accelerated aging treatments (AAT and CDT). All mutants presented 100% germination in non-aged seeds. As mentioned above, the AAT treatment, used to evaluate seed vigour, revealed that the double mutant *prx2 prx25* and the *prx2 prx25 prx71* triple mutant were sensitive to the deterioration treatment to the same extent, indicating that the loss-of-function of *PRX71* was not additive (Figure 55). In addition, these mutants were subjected to a controlled deterioration treatment, an alternative test to check seed longevity where seeds are exposed to high humidity instead of being imbibed in water. Both the *prx2 prx25* and the *prx2 prx25 prx71* triple mutant exhibited same sensitivity after 14 days of accelerated aging, with less than 50% of the sowed seeds germinating (Figure 55).

prx2 prx25 and *prx2 prx25 prx71* show seed coat alterations

To characterize potential defects in the seed coat of these peroxidase double and triple mutants, we performed histological sections of dry seeds and observed them using transmission electron microscopy (TEM). Analyzing the different layers starting from the inside of the seed (Figure 56), the cutin layer (CL) was located in the inner integument and associated to the cell wall of the endosperm. The *prx2 prx25* and *prx2 prx25 prx71* mutant lines presented a structure comparable to the wild type. Tannins are deposited in the endothelium (EN), which is observed as a broad brown pigment layer (Gou *et al.*, 2017). This layer was also very similar between the wild type and mutants. The next identified layer is called the palisade (PL), where suberin is deposited adjacent to cell walls. The thickness of this subepidermal layer was highly reduced in both mutants. Finally, no change was detected in the columella cells (COL), responsible of mucilage extrusion, in any of the samples analyzed. This result suggests the specific participation of these peroxidases in the formation of the palisade layer in the seed coat. To completely discard the role of these peroxidases in both mucilage and proanthocyanidins synthesis, ruthenium red and vanillin staining were performed respectively. As shown in Figure 57 no differences were observed between the wild type and mutant lines. Finally, the fact that the double and triple mutants do not display a transparent testa phenotype also discards a role for these peroxidases in tannin accumulation (Debeaujon *et al.*, 2000).

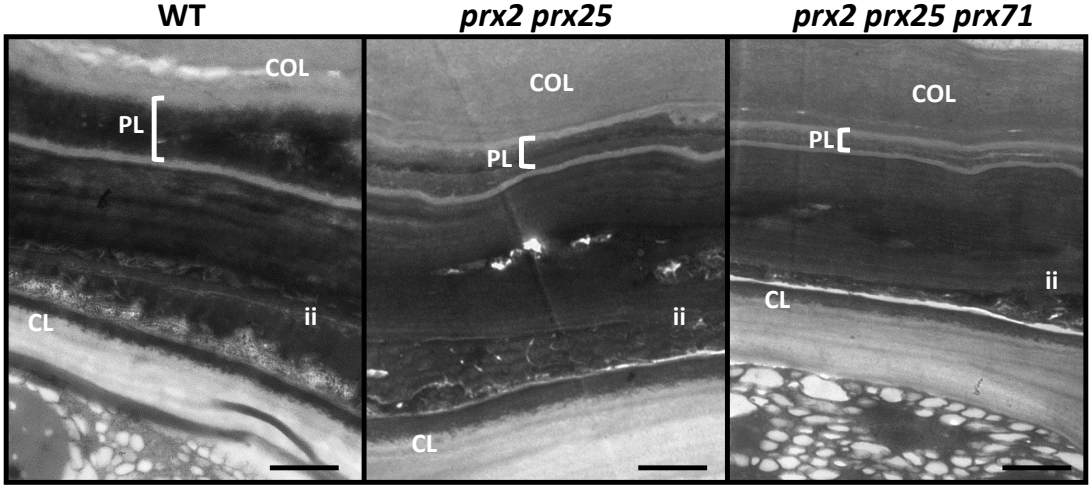


Figure 56: Double mutant *prx2 prx25* and triple mutant *prx2 prx25 prx71* are altered in the formation of palisade layer of seed coat. Transmission electron microscopy observation of representative seed coat of the wild type, *prx2 prx25* and *prx2 prx25 prx71*. Mature seeds were ultrathin sectioned, and image was focused to show seed coat and. ii, inner integument; CL, cutin layer; COL, columella cell; PL, palisade layer. The pictures are representative of 20 different seeds. Arrows indicate the thickness of palisade layer. Scale bar, 2 µm.

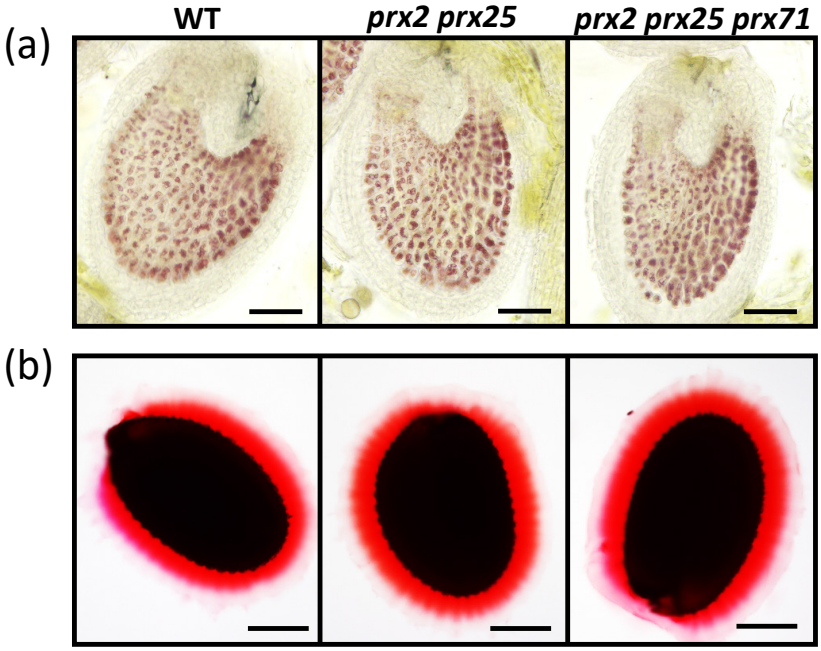


Figure 57: *prx2 prx25* and *prx2 prx25 prx71* do not present alterations in proanthocyanidins accumulation or mucilage extrusion. (a) Vanillin staining of representative wild type, *prx2 prx25* and *prx2 prx25 prx71* seeds after 5 days of pollination. Scale bar, 100 µm. (b) Ruthenium red staining of representative wild type, *prx2 prx25* and *prx2 prx25 prx71* seeds. Scale bar, 200 µm.

PRX2 and PRX25 regulate cell wall polyphenolics in seed coat

Suberin is a complex heteropolymer, characterized by both polyphenolic and poly-aliphatic domains (Bernards *et al.*, 2002). There is no single method by which both fractions can be effectively isolated and studied together (Vishwanath *et al.*, 2015), and there is no method to chemically characterize suberin polyphenolics and lignin separately. Although the exact structure of the suberin-associated polyaromatics remains an open question, it is known that polyaromatics are a substantial component of suberized cells (Graça *et al.*, 2015). The analysis of the insoluble suberin-associated polyaromatics includes degradative lignin techniques like the acetyl bromide protocol, based on the formation of acetyl derivatives in non-substituted OH groups and bromide replacement of the α-carbon OH groups to produce a complete solubilization under acidic conditions (Rains *et al.*, 2018). In order to confirm whether PRX2, PRX25 and PRX71 have an important role in the macromolecular assembly of lignin and/or the suberin polyphenolic domain in seed coats, we performed a cell wall preparation from seeds that was solubilized with acetyl bromide. Spectrophotometry measurements indicated that the *prx2 prx25* mutant had a 35% reduction of polyphenolic compounds, and the triple mutant seed preparations showed similar values (Figure 58). Aliphatic components of suberin were also measured by GC-MS in these mutants. Total suberin amounts were comparable to WT in the in *prx2 prx25* and in *prx2 prx25 prx71* mutants (Data not shown).

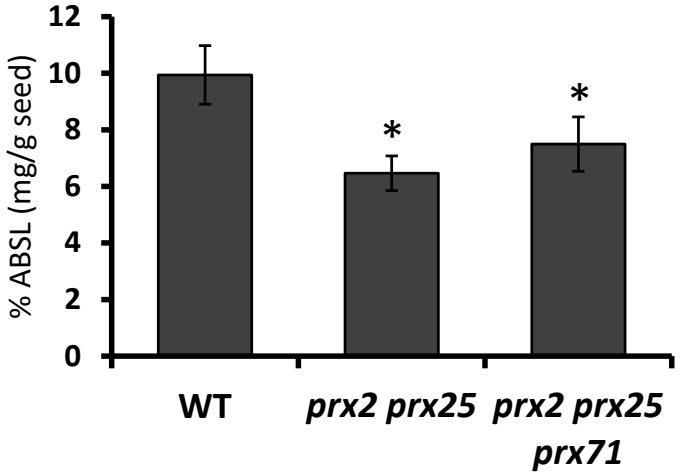


Figure 58: PRX2 and PRX25 regulate cell wall polyphenolics in seed coat. Total acetyl bromide soluble polyphenolic from suberin and lignin from wild type, *prx2 prx25* and *prx2 prx25 prx71* seeds. The results are the average of 10 samples. *Significantly differing from controls (wild type) at $p < 0.05$ (Student's *t* test). %ABSL: acetyl bromide soluble lignin percentage

prx2 prx25 and *prx2 prx25 prx71* seeds are more permeable

The correlation between seed permeability and the lack of some components, such as proanthocyanidins, suberin and cutin, in the seed coat has been widely studied (Debeaujon *et al.*, 2000; Vishwanath *et al.*, 2013; Beisson *et al.*, 2007 and Yadav *et al.*, 2014). A common method to assess seed permeability consists of incubating seeds in an aqueous solution of colourless tetrazolium salts. When seeds are permeable, tetrazolium is metabolized by NADH-dependent reductases present in embryo cells, producing red-colored insoluble precipitates made up of formazans (Debeaujon *et al.*, 2000). Our results indicate that the mutants sensitive to accelerated aging, *prx2 prx25* and *prx2 prx25 prx71*, have significant compositional alterations in seeds and as explained above, these seeds should present higher permeability. To test this, we incubated seeds from both mutants in a tetrazolium salt solution and after 48 h we observed the amount of red staining (Molina *et al.*, 2008). All seeds from the mutant peroxidase lines displayed increased accumulation of formazan precipitates compared to the WT, as determined by microscopy (Figure 59). To quantify seed permeability, we extracted the red-colored formazans after 72 h of incubation and measured the absorbance of the solution at 490 nm. Extracts from mutant seeds presented higher absorbance than WT (Figure 59) confirming our microscopical observations.

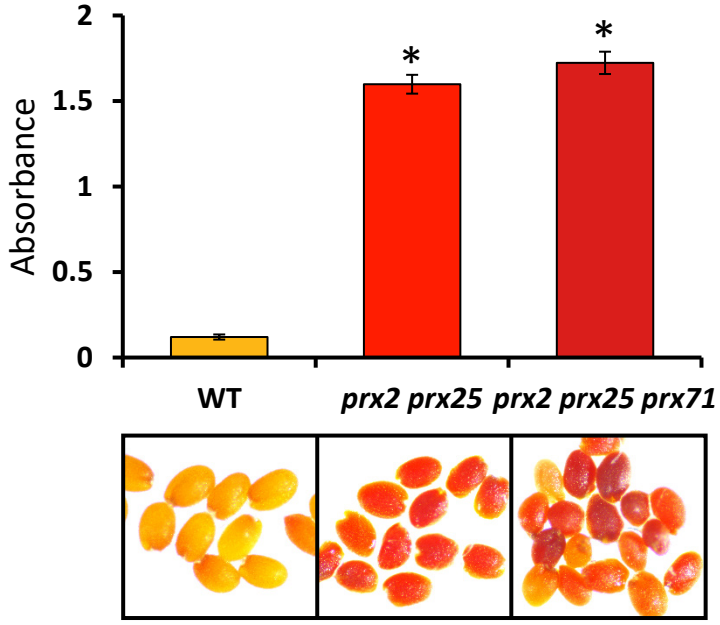


Figure 59: Seed permeability is increased in seed deterioration treatment-sensitive seed coat peroxidases mutants. Above: Quantification of formazan extracted from seeds of different genotypes incubated in 1% tetrazolium red at 30 °C at the end of 72 hours. The data (absorbance at 490 nm) are mean values of three replicates, and the bars indicate standard errors. The bar color is representative of the formazan extraction color. *Significantly differing from controls (wild type) at $p < 0.05$ (Student's *t* test). Below: Representative images of the staining pattern in seeds of wild type (WT) and mutants incubated in 1% tetrazolium red at 30 °C for 48 hours.

Discussion

The seed coat is a maternal tissue developed from the integuments of the ovule whose main function is to protect the embryo from external factors. Class III peroxidases are enzymes with a role in the polymerization of different seed coat components, as well as their hydrolysis (Francoz *et al.*, 2015). For instance, PEROXIDASE36 (PRX36) is required for seed coat mucilage extrusion (Kunieda *et al.*, 2013). Proanthocyanidins are oligomeric and polymeric end products of the flavonoid biosynthetic pathway and peroxidases could be involved in the polymerization of these phenolic compounds, although no gene encoding a protein with this function has been reported so far (Dixon *et al.*, 2004). The accumulation of these flavonoids provides a barrier, as demonstrated by two main studies. Debeaujon and collaborators compared mutants affected in testa pigmentation, *transparent testa (tt)* and *transparent testa glabra (ttg)* for dormancy, tetrazolium uptake and storability and they established an inverse correlation between permeability and the presence of catechins and proanthocyanidins (Debeaujon *et al.*, 2000). In addition, Clerkx *et al.*, 2004*b* showed that condensed tannins are crucial to maintain seed viability because of their antioxidant properties. Finally, lipid polyesters are compounds commonly found in seed coats of different species. It was described that both suberin and cutin are found in *Arabidopsis* seed coats (Molina *et al.*, 2008). In addition, peroxidases also have a role in assembling monolignols during lignin synthesis. Although seed coat lignin has been less explored than vascular lignin, there is evidence of its accumulation in cell walls of seed coat integuments, possibly making seeds less permeable to water and gases (Tobimatsu *et al.*, 2013; Liang *et al.*, 2006). Finally, the biosynthesis of the suberin poly(phenolic) domain is hypothesized to follow a peroxidase-mediated oxidative coupling process (Bernards *et al.*, 2004). This apoplastic lipid barrier is located on the palisade cell walls of the outer integument of the seed coat (Molina *et al.*, 2008). It has been postulated to have a function in dormancy (Fedi *et al.*, 2017) and in abiotic stress protection during germination (Gou *et al.*, 2017). Our work demonstrates that suberin deposition in the palisade layer of testa also increases seed tolerance to deterioration and it constitutes the first barrier protecting the embryo from this stress. Mutants sensitive to AAT such as *gpat5, myb107* and *abcg2 abcg6 abcg20* display an important reduction in the suberin layer in seeds that generates very permeable seeds (Beisson *et al.*, 2007, Gou *et al.*, 2017 and Yadav *et al.*, 2014).

On the other hand, class III peroxidases are a gene family with a large number of members involved in numerous processes, suggesting a particular specialization for each isoform. Genetically, this hypothesis is also supported by the nucleotide variation observed in the peroxidase promoters. Another feature of these oxidoreductases is the high redundancy, which substantially complicates the observation of a specific phenotype upon mutation of a single peroxidase gene (Cosio and Dunand, 2009). Our

data reinforce this concept, as no single mutant line in seed coat peroxidases showed lower germination after AAT, as compared to control seeds. The most studied function of peroxidases is in cell wall stiffening mostly through lignin polymerization. The first and most well-studied is the ATP A2 peroxidase, named PRX53 using the accepted peroxidase classification nomenclature. A mutant with high levels of this peroxidase presents an increase in lignin accumulation. Moreover, the substrate specificity for p-coumaryl and coniferyl alcohols was determined by X-ray crystallography (Ostergaard *et al.*, 2000). In *Arabidopsis*, the role of PRX64 in the Casparian Strip formation has also been reported (Lee *et al.*, 2013), as well as the importance of PRX2, PRX25 and PRX71 in controlling the content of lignin in stem (Shigeto *et al.*, 2015). Pedreira *et al.*, 2010 described the function of PRX37 as a peroxidase associated with vascular bundles functioning mainly in mature leaves and the flower stem. Finally, the *prx72* mutant was reported to present a low content in syringyl units and a decrease in the amount of lignin (Herrero *et al.*, 2013).

Suberin is a biopolymer involved in cell wall permeability and stiffening and the polyphenolic domain shares some similarities with lignin, because they are partly formed by monolignols, such as coumaryl, coniferyl and sinapyl alcohols. In addition, peroxidases are proposed to participate in the oxidative coupling process of this domain (Bernards *et al.*, 2004). This role of peroxidases has been confirmed in the case of *TPX1*, a tomato peroxidase gene whose expression was detected in the endodermis and exodermis where suberin accumulates (Quiroga *et al.*, 2000). Our study shows the important function of PRX2 and PRX25 in the normal development of the palisade layer, where suberin is accumulated in the cell wall. In addition, *prx2 prx25* double mutant seeds present a marked reduction in the % of acetyl bromide soluble lignin (ABSL) that measures the amount of suberin and lignin polyphenolic compounds (Rains *et al.*, 2018). Thus, our results suggest that the high permeability of these mutants is because of a deficient phenolic-based polymer. However, our analyses cannot establish if the affected polymer represents a part of suberin or lignin or both.

During the last years, we have been trying to define the molecular mechanism underpinning COG1 and ATHB25-mediated resistance to seed aging. Overexpression of these TFs produces seeds with an increase in the accumulation of suberin in the palisade layer of the seed coat, likely reducing the embryo oxidation and extending seed viability (Bueso *et al.*, 2014*b*; Bueso *et al.*, 2016). The specific target genes and processes involved in mediating this reinforcement of the seed coat is the next issue to clarify. *cog1-2D* and *athb25-1D* accumulate more ABA. Interestingly, suberization in potato tuber by exogenous ABA application has been known for a long time (Soliday *et al.*, 1978, Cottle and Kolattukudy, 1982). More recently, MYB41 has been described as an ABA-inducible TF that activates steps necessary for aliphatic suberin synthesis and deposition (Kosma *et al.*, 2014). Another explanation is that COG1 and ATHB25

could directly regulate suberin biosynthetic genes, including glycerol-3-phosphate acyltransferases, cinnamyl alcohol dehydrogenases, fatty acyl-CoA reductases and laccases. At this time, we do not favour a mechanism in which the peroxidase genes are directly regulated by these TFs because we are unable to detect binding of this transcription factor to the *PRX25* promoter in chromatin immunoprecipitation experiments. Thus, the most plausible mechanism is that the COG1 TF regulates the expression of GA-inducible seed coat peroxidases *PRX2* and *PRX25* through direct regulation of gibberellic oxidase biosynthesis genes.

Chapter 4

Work in progress and unpublished data

Environmental regulation of seed longevity is mediated by AtHB25 and COG1 through temperature and light cues

Joan Renard[1], Gaetano Bissoli[1], Regina Niñoles[1], Irene Martínez-Almonacid[1], Luis Oñate-Sánchez[2], José Gadea[1], Ramón Serrano[1] and Eduardo Bueso[1]

[1]Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universitat Politècnica de València-C.S.I.C., Camino de Vera, 46022, València, Spain

[2]Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid - Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Campus de Montegancedo, Pozuelo de Alarcón, Madrid 28223, Spain

Abstract

Transcription factor regulation has a pivotal role in the study of the developmental and molecular regulation mechanism. Plant are sessile organisms, and their adaptation relies in their tightly regulated developmental program. Seeds, and its highly specialized mother-originated tissue (the seed coat) are determined by transcription factor regulation. The overexpression of the transcription factors (TFs) AtHB25 and COG1 produce reinforced seed coats that extend seed longevity. Neither AtHB25 nor COG1 interacted with other TFs described in seed coat differentiation in protein-protein interaction assays pointing to a different regulation. Further analysis demonstrates that AtHB25 was regulated by low temperatures and confers cold resistance, and COG1, described in the regulation of light signaling, contributes to the outer integument differentiation. Both AtHB25 and COG1 play essential roles in the adaptation of seed longevity to environmental cues, such as temperature and light, respectively. The importance of the cold regulation of AtHB25 is remarked in the highly conserved regulatory region, that it is putatively regulated by a cold induced TF belonging to a stress response subfamily of TFs, the DREB2 TF subfamily. The enhanced cold resistance involving AtHB25 probably goes through ICE1 regulation. On the other hand, COG1 is regulated by AP2, a floral TF determining the outer integument differentiation, and thus the mucilage and the suberin synthesis.

Introduction

Life emerged and adapted to different ecosystems. Evolution implies directly changes in the DNA, which directly determines the protein sequence and thus the function. However, diverse mechanisms emerged to control the abundance of proteins for a regulation of different metabolic pathways. Gene expression is one of the main regulatory steps, as it occurs at first instance in the protein synthesis from the gene coding sequence. It consists of the regulation of the number of copies of RNA for a given gen. The number of mRNA from a gen correlates with the protein synthesis rate. Other regulation steps control protein abundance and function, such as RNA processing (determine the splicing variant and RNA half-life), and the post-translational modifications (determine protein structure, cofactors, protein switching and protein degradation) (Nakaminami *et al.*, 2012).

Differences in gene expression are determined by the chromatin structure and DNA packing regulated by histones, DNA methylation, non-coding RNAs and transcription factor (TF) regulation (Holoch and Moazed, 2015). TF regulation is essential for the presence of a concrete mRNA set and their derived proteins in different cell types. TFs play essential roles through gene regulation, as organism adaptation to environment, tissue differentiation and organism development. They can control gene expression through the biding to DNA regulatory sequences (Latchman, 1993), which are usually localized in up-stream (or 5') regions of the target gene (Adcock and Caramori, 2009). These regulatory regions, which may act as enhancers or repressors, are not always located close of the transcription start site (TSS) of the gene (Spitz and Furlong, 2012). Nevertheless, DNA is not a straight molecule and it folds and can form loops. This structure permits distanced TF to regulate gene expression, as the TSS and TF regulatory sequences might be close. Transcription factors bind to specific sequences, and these sequences can be determined by the probability of union by the TF, named the motif (Stormo, 2000). Transcriptome analysis allows to find differential expressed genes that might be directly or indirectly regulated by TFs. Mutant lines are used for the TF of study, such as overexpression lines or knock-out (KO) lines. Normally TFs work in cascade of regulation: TFs regulate the expression of other TFs in a complex regulatory network (Jothi *et al.*, 2009). In addition, protein-protein interaction between TFs is important in gene expression regulation (Latchman, 1993). This complex TF regulation results in expression changes of hundreds and even thousands of genes, which may lead to phenotypic modifications.

In Arabidopsis there are 1,851 TFs, grouped in 50 families, according to the *Arabidopsis thaliana* Transcription Factor Database (Palaniswamy *et al.*, 2006), which suppose a 5.5% of total genes according to the last Arabidopsis genome version TAIR10. This high TFs abundance remarks the importance of TF regulation for plant regulatory

processes. For seed development the TF regulation network is determinant, as they are key factors involved in tissue identity of embryo, endosperm and seed coat, determining seed performance and seed longevity. Several TFs factors are described in seed coat tissue differentiation and development regulation and determine integument identity and biomolecules associated (Xu *et al.*, 2015; Golz *et al.*, 2017).

Two TFs, *COG1* and *AtHB25* were found overexpressed in two activation tag lines resistant to the accelerated aging treatment, named *cog1-2D* and *athb25-1D,* respectively (Bueso *et al.*, 2014*b*; Bueso *et al.*, 2016). Seeds of both lines presented reinforce seed coats with an increase lipid polyester deposition (Renard *et al.*, 2020*a*; Renard *et al.*, 2021) and an expanded seed longevity (Bueso *et al.*, 2014*b*; Bueso *et al.*, 2016).

COG1 belongs to the DOF transcription factor family. Diverse DOF TFs are involved in seed development, dormancy and germination through ABA and GAs hormone regulation (Ruta *et al.*, 2020). DOF6 increases ABA contents and DAG1 repress GAs perception. Thereby, both prevent germination increasing seed dormancy (Papi *et al.*, 2000; Rueda-Romero *et al.*, 2012). Contrarily, DAG2 promotes light induced seed germination via the repression of DAG1 (Santopolo *et al.*, 2015). In addition, CDF1, CDF2, CDF3 and CDF4 redundantly control the photoperiodic control of flowering in Arabidopsis through the CONSTANS TF degradation (Imaizumi *et al.*, 2005; Ruta *et al.*, 2020). The *cdf4-1D*, overexpressing *CDF4* a close homolog of *COG1*, present increased seed longevity such as *cog1-2D* (Bueso *et al.*, 2016). The COG1 TFs is involved in the different regulations described in DOF TFs. COG1 regulates ABA and GAs metabolism, as the line *cog1-2D* present increased ABA and GA contents (Bueso *et al.*, 2016). It negatively regulates phytochrome signaling, and mutant expression lines, *cog1-1D* and *cog1-2D,* present deficient a light sensing phenotype including increased hypocotyl elongation and reduced light-induced anthocyanidin accumulation (Park *et al.*, 2005; Bueso *et al.*, 2016). The expression of *COG1* during seed development is mainly located in the seed coat, specifically during seed maturation in the outer integument (Renard *et al.*, 2020*a*). All these data suggest a role of COG1 in light-mediated seed longevity.

AtHB25 is a homeobox TF belonging to the Zinc Finger-Homeodomain (ZF-HD) TF family. Members of this TF family are expressed in flowering and they form homodimers and heterodimers with other members of the family (Tan and Irish, 2006). Moreover, other homeodomain TF families, such as WOX, KNAT, BEL and HD-Zip are also involved in flowering development and environmental adaptation of plants (Hay and Tsiantis, 2010; Chew *et al.*, 2013; Costanzo *et al.*, 2014; Jha *et al.*, 2020). Indeed, homeobox TFs are conserved among species, and they participate in tissue identity regulation in plants (and animals), necessary for flowering development. Interaction proteins determine their regulatory specificity (Bürglin and Affolter, 2016). AtHB25 TF regulates, not only flowering, but also seed longevity through the seed coat rein-

forcement (Bueso *et al.*, 2014*b*). Indeed, *AtHB25* is expressed in the maturing seed coat and it directly targets the lipid polyester biosynthetic genes LACS2, CYP86A8, KCS20 and the putative NHO1 expression, determining the lipid polyester deposition of seed coat (Renard *et al.*, 2021). AtHB22, the closest AtHB25 homologue (Tan and Irish, 2006), act redundantly to AtHB25 in the regulation of seed longevity. The double mutant *athb22 athb25* line presents a reduced seed coat longevity phenotype, like the antisense line targeting *AtHB25*, which target to degradation transcripts of *AtHB25*, *AtHB22* and *AtHB33* (Bueso *et al.*, 2014*b*). AtHB25 seed coat lipid polyester deposition pathways are conserved among different angiosperm (tomato and wheat), and respective orthologue genes are highly conserved (Renard *et al.*, 2021) remarking the importance of this AtHB25 regulation.

Thereby, molecular mechanisms of the reinforced seed longevity caused by COG1 and AtHB25 in the overexpression mutants *cog1-2D* and *athb25-1D* have been revealed. However, the regulation of these seed coat transcription factors during seed development is still unknown. We found that AtHB25 is up regulated by cold, probably by a DREB2 TF, which directly binds to a conserved AtHB25 regulatory region. In addition, AtHB25 positively contributed to cold resistance in plants. Cold resistance regulated by AHB25 might go through ICE1, a direct target of AtHB25. However, direct seed coat fissures of *ice1* seeds explain their reduced seed longevity rather than seed lipid polyester deposition in the seed coat, responsible of the seed longevity phenotypes of *AtHB25* mutants. On the other hand, COG1 was found to be a downstream target of AP2 in outer integument differentiation, participating in columella cell formation and suberin deposition, which explains the increased seed longevity of *cog1-2D* seeds. COG1 integrates light signaling which might influence seed coat development. Cold signaling pathways and light-dependent outer integument differentiation regulate seed coat lipid polyester deposition, probably through AtHB25 and COG1, respectively. COG1 and AtHB25 did not interact with known seed development TF. Our results suggest that AtHB25 and COG1 regulate the seed longevity adaptation to environmental cues through the seed coat regulation.

Results

AtHB25 forms homodimers

To determine the endogenous role of AtHB25 and COG1 in seed longevity, we first determined if they interact with other TFs in the determination of seed coat development. TF often interact with other TFs modulating their effect. Moreover, the TF AtHB25 presumably does no present an activation domain to recruit the RNA polymerase II (Tan and Irish, 2006). Known seed coat transcription factors involved in the tissue

identity and biomolecule regulation were assayed as putative interactors of AtHB25 or COG1, used as bait proteins in the protein-protein interaction Split-Tpr assay. This assay consists of covalent link complementary fragments of TPR1 protein to both proteins. The reconstituted activity of the TPR1 protein is used to assess the interaction among both tested proteins. In comparison to the yeast-two-hybrid (Y2H) assay, the Split-Tpr present the advantage that TFs cannot autoactivate the reporter gene, resulting in reduced false positives. Nevertheless, this technique presents a reduced sensitivity compared to Y2H. The TF MYB47, recently associated to seed longevity with a GWAS analysis (Renard *et al.*, 2020*b*), was analyzed together with AtHB25 and COG1. The transcription factors AP2, TT2, TT8, TT16, GL2, TTG1, TTG2, EGL3, MYB5, MYB61, SKT, SEP3, KNAT7, FUS3, LEC1, LEC2, WRKY46, SPCH and FLY1 were cloned as prey proteins, in addition to AtHB25, COG1 and MYB47. Although, the Split-Tpr assay did not provide any positive interaction among the different combinations of transcription factors tested, AtHB25 was found to interact with itself forming homodimers. This interaction has been previously described (Tan and Irish, 2006).

AtHB25 is regulated by a conserved regulatory sequence

AtHB25 is a Homeobox protein conserved among angiosperms and ectopic *AtHB25* expression enhances seed coat lipid-polyester deposition in different plant species (Renard *et al.*, 2021). In order to reveal the *AtHB25* regulation, we seek for conserved regions in the promoter of the gene. The best candidate orthologous of *AtHB25* were identified by PLAZA 4.5 DICOTS in closely related species to *Arabidopsis thaliana*. The plant species *Brassica rapa, Schrenkiella parvula, Capsella rubella* and *Theobroma cacao* were selected for promoter multiple sequence alignment of their respective genes *Brara.I00817*, *Tp2g28490*, *Carubv10026891m.g* and *TCA.TCM_034470*, together with *AtHB25* promoter. A region of 71 pb ranging from -982 bp to -911 bp

```
ACAAAA--TTCAGACCTTGAAATTATTAAGGGCATTACT---GATTTTGA-TGTGATT--AGACTGCCCATCACTTTATTGCAGAAAA-ATATTAATGCA
ACAAAAA-TTCAGACCTTGAAATTATTAAGGGCATTACT---GATTTCGA-TGTGATT--AGACTGCCCATCACTTTATTGCAGAAA---TATTAATGCA
ACAAAAA-TTCAGACCTTGAAATTATTAAGGGCATTACT---GATTTTGA-TGTGATT--AGACTGCCCATCACTTTATCGCAGAAG---TATTAATGCA
AAGAAAAGTTCAGACCTTGAAATTATTAAGGGCATTACT---GATTTTGA-TGTGATT--AGACTGCCCATCACTTTATTGCAGAAAATATATTAATGCA
GCAATTCTCTCCCCTTTTGGATTAAGCAAACTCCTCTCTCCCAATCTTATCTATAATTCTATATCACGCACATCTTAATTCCTTCAGTCTTATTTGTAGA
   *     **     *** * * *  **   * *  **    ** *    * * ***  * *   * **   *** **  *   *    **** *  *
```

Figure 60: Sequence alignment on the conserve promoter region of *AtHB25* and best orthologous candidates genes from *Brassica rapa, Schrenkiella parvula, Capsella rubella* and *Theobroma cacao*, respectively. The alignment was performed with Clustal W with the upstream and gene regions of the genes *At5g65410, Brara.I00817, Tp2g28490, Carubv10026891m.g and TCA.TCM_034470* from up to down. The selected fragment form the *AtHB25* promoter region is highlighted in yellow, with may be important for its endogenous regulation due to the highly conservation. Notice that in the cacao sequence the conserve regulatory sequence starts to degenerate. * indicate completely conserved DNA base among all four regulatory regions aligned.

relative to the Arabidopsis *AtHB25* translation start site was found highly conserved in these five plant species (Figure 60). This highly conservation suggest an important regulatory role, putatively in the regulation of *AtHB25* expression.

The conserved sequence has transcriptional activity

To test the transcriptional activity of the selected conserved fragment of the promoter of *AtHB25* four repetitions were cloned in the plasmid *pYRO*, which contains the luciferase gene and the minimal promoter, a short sequence necessary for the transcription activation of the luciferase gene (Castrillo *et al.*, 2011). The luciferase activity assay showed a 10-fold transcriptional activation of luciferase driven by the four conserved-fragment repetitions in transiently transformed leaves of *Nicotiana benthamiana* in comparison with the empty *pYRO* plasmid (Figure 61). This assay confirmed the regulatory activity of the selected conserved *AtHB25* promoter fragment in *Nicotiana benthamiana,* and thus presumably in Arabidopsis. Other promoter regions might also have a regulatory role in the AtHB25 expression. However, this conserved region might unravel important transcription regulation important in plant development or adaptation.

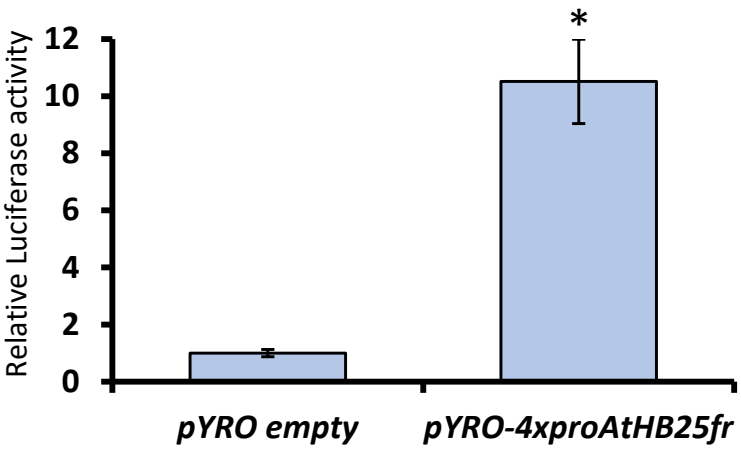
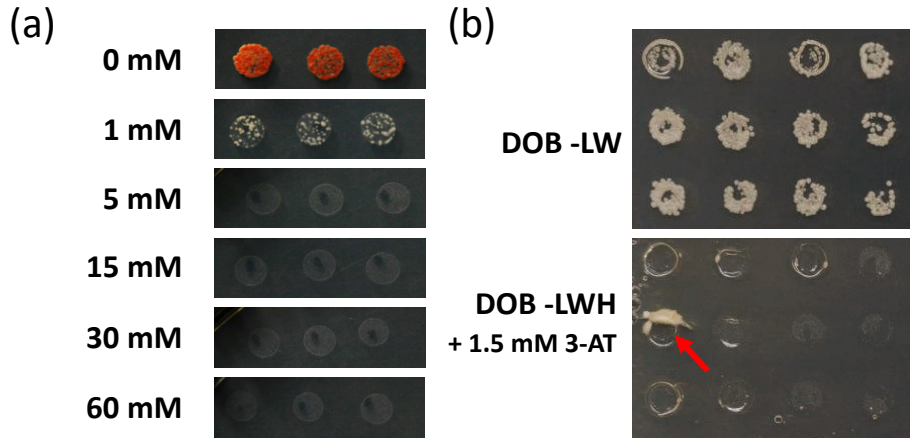


Figure 61: Luciferase activity assay in *Nicotiana benthamiana* transiently transfected leaves with the *pYRO-4xproAtHB25fr::LUC* construction, relative to the empty *pYRO* plasmid luciferase activity, taken as 1. Three independent transfection leaves and three independent reads per sample were analyzed. Bars indicate standard errors. *, $p < 0.05$.

Y1H screening points to *AtHB25* cold regulation

We use the selected conserved *AtHB25* promoter fragment to perform a DNA-protein screening assay to determine regulatory TFs of AtHB25. In collaboration with Luis Oñate-Sánchez (CNB, Madrid), we determined using a yeast-one-hybrid assay (Y1H), what TFs from an arrayed TFs library interact with the selected conserved *AtHB25* promoter fragment. To assess the promoter autoactivation in yeast we analyzed growth of yeast with the *pTUY1H-proAtHB25fr* construction and the empty *pDEST22* plasmid in media without the amino acid used for the auxotrophic recovery through the DNA-protein interaction assay. Protein interactions with the *AtHB25* promoter fragment, located next to the *HIS3* gene, would lead to *HIS3* expression, and permit the yeast to survive in synthetic media without histidine. However, the intrinsic *HIS3* expression, not depending on TF interaction, might increase false positives. Different concentrations of the histidine competitive inhibitor 3-Amino-1,2,4-triazole (3-AT) were used to inhibit the yeast growth in interaction selective media (Figure 62a). A second assay was used to set precisely the minimal 3-AT concentration that inhibited



(c)

Position	*AGI*	*TF name*
006-C01	AT5G53980	Homeobox (HB52)
008-H10	AT1G06850	bZIP (bZIP52)
015-C12	AT1G24540	MADS (ALG24)
015-G03	AT2G40350	AP2/EREBP (DREB2)

Figure 62: The Y1H screening with the selected conserved *AtHB25* promoter fragment provided four positives interactions. (a) The effect of HIS3 autoactivation in yeast containing *pTUY1H-proAtHB25fr* and the empty *pDEST22* plasmids was inhibited within 1 and 5 mM 3-AT concentrations. Yeast were grown under DOB-LWH media with different concentrations of 3-AT for seven days. (b) Representative image of the Y1H screening. Upper part: Control diploid growth (two days). Down part: Selective media for the protein-DNA interaction (seven days). The red arrow indicates a positive interaction. (c) Schematic results of the obtained protein-DNA interactions through the Y1H screening. The position in the arrayed library indicated the TF that permit the yeast growth. Four interactions were obtained through the Y1H screening.

the yeast growth, which was set in 1.5 mM. The Y1H screening was performed in the Oñate-Sanchez TFs arrayed collection as Sánchez-Montesino and Oñate-Sánchez (2017). Yeast (Mat*α*) containing the *pTUY1H-proAtHB25* were matted with the different yeast (Mat*a*) containing the different TFs in the *pDEST22* plasmid. A bunch of 1,358 different TFs was assayed in the Y1H screening. After one week of growth in protein-DNA interaction selective media, colonies grown were considered as positive screening results (Figure 62b). Four positive interactions were obtained with the yeast containing the transcription factors HB52, bZIP52, ALG24 and a DREB2 family member, named DREB2H by Nakano *et al.* (2006) (Figure 62c). Positive interactions were reconfirmed in an independent Y1H directed assay. To assess which of the TFs might have a regulatory role in seed coat development and seed longevity, *in silico* expression analysis was performed with the Le *et al.* (2010) data. Only *DREB2H* was significantly expressed in the seed coat, with an expression peak in the maturation stage (Figure 63a), similar to the observed seed coat *AtHB25* expression in *proHB25::GUS:GFP* transgenic plants (Renard *et al.*, 2021). DREB2 transcription factors are a TFs class that binds to dehydration-responsive cis-elements (DRE) and

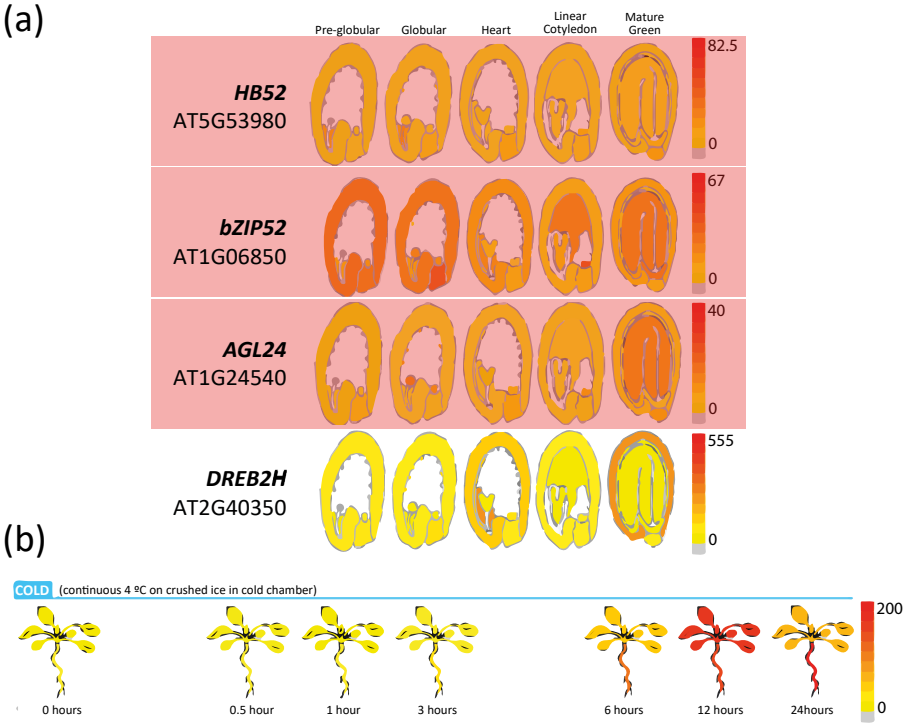


Figure 63: *In silico* expression analysis of the positive Y1H screening results. (a) Expression during seed development based in Le *et al.* (2010) data. Pale red boxes indicate those genes with reduced seed-development expression. (b) Expression of the *DREB2* family member during cold shock based in Kilian *et al.* (2007) data. Images were obtained with the eFP browser (Winter *et al.*, 2007).

C-repeats cis-elements (CRT) in response to various abiotic stresses (Sakuma *et al.*, 2002; Lata and Prasad, 2011). The selected *AtHB25* promoter fragment does not contain a DRE (ACCGAC) or CRT (GCCGAC) sequence (Sakuma *et al.*, 2002), and the most similar sequence contained in this promoter was GCCCAT. To identify abiotic stresses that regulate *DREB2H* expression, we analyzed the Kilian *et al.* (2007) data with the browser online tool (Winter *et al.*, 2007). Cold stress is apparently the main abiotic stress regulation of *DREB2H*. Transcripts levels of *DREB2H* are accumulated after 12 hours of continuous plant chilling (Figure 63b). Taken together, the Y1H screening pointed to a *AtHB25* regulation in the seed coat through DREB2H, among others, which might be influenced by cold.

AtHB25 is up regulated by chilling and it is involved in cold resistance

In order to determine the cold-dependent regulation of *AtHB25*, putatively due to the Y1H-positive DREB TF, we chilled for 24 hours at 9 ºC *proAtHB25::GUS:GFP* transgenic 25-days old Arabidopsis plants. GUS expression was analyzed in whole control

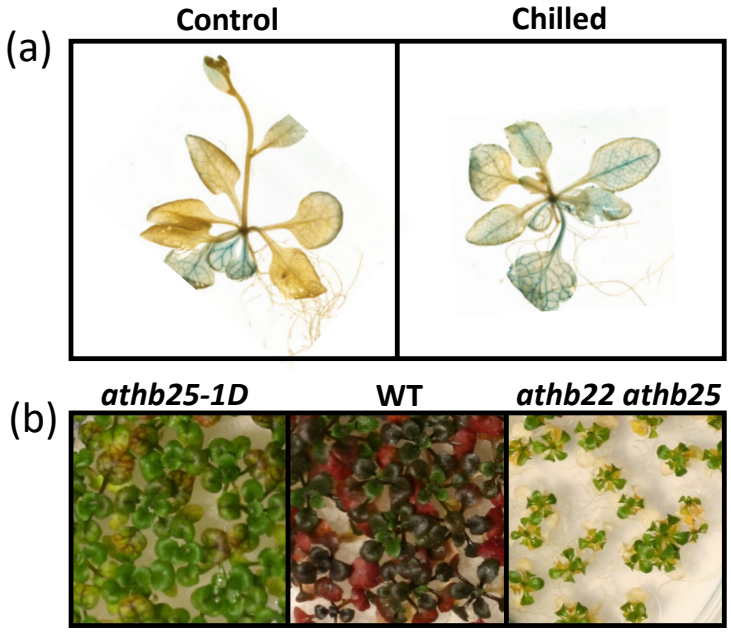


Figure 64: AtHB25 is up regulated by cold and confers cold resistance. (a) Representative image of GUS staining of 25-days old *proAtHB25::GUS:GFP* plants grown in MS were chilled for 24 hours at 9 ºC, while control plants not. Chilled plants presented vascular leaf GUS staining that was not present in control plants. (b) *AtHB25* mutants present cold resistant alterations. Plants were grown in MS media for 85 days at 9ºC and continuous light. The overexpression mutant barely present anthocyanin pigmentation visualized in wild type (WT) plants, probably indicating a reduced cold stress. The double mutant *athb22 athb25* present a serious growth defect in cold *in vitro* cultivation.

and chilled plants. Beside the cotyledonal GUS signal observed in both conditions, chilled plants present a strong GUS induction in leaves, mainly vascularly located (Figure 64a).

AtHB25 mutants were grown under low temperature conditions to clarify their role in cold-stress response. Seven-day old plants were grown in MS media at 9 ºC for 85 days. The different *AtHB25* mutants showed important differences in cold adaptation

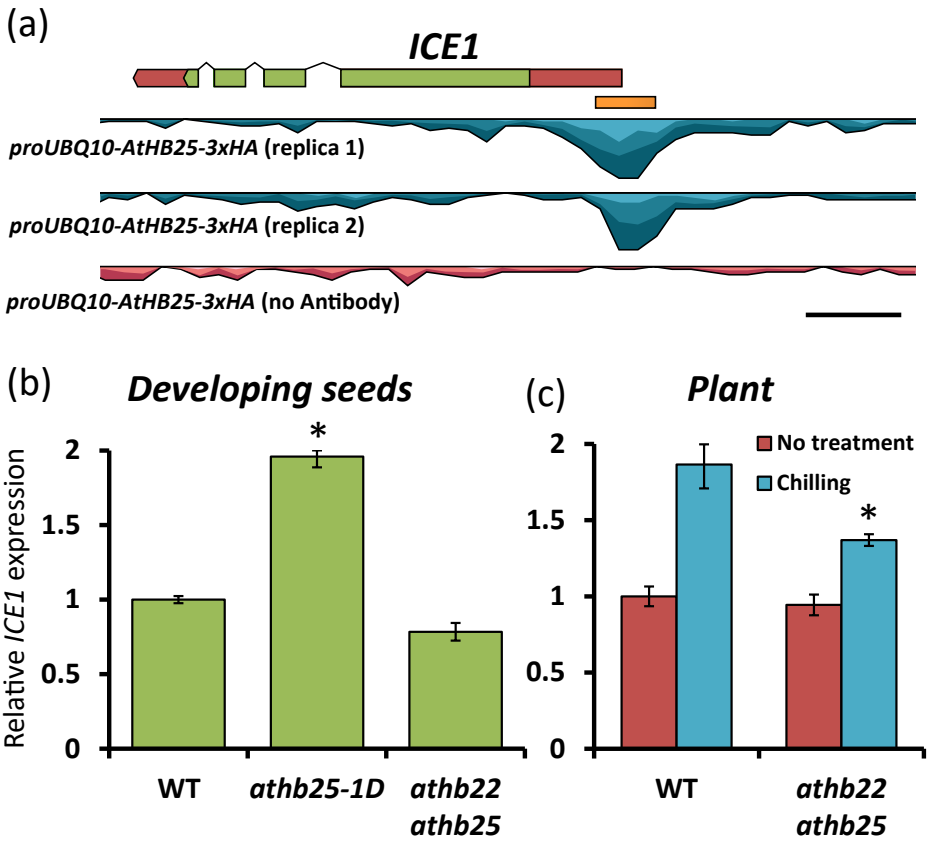


Figure 65: AtHB25 regulates *ICE1* cold response. (a) Read coverage of AtHB25 ChIP-seq in *ICE1* genomic locus. The figure was obtained with CLC Genomics Workbench program, TAIR10 genes and BAM files resulting from both Antibody samples and the no antibody control sample from the ChIP-seq experiment of AtHB25 in developing seeds (Renard *et al.*, 2021). The *ICE1* gene is represented by its exons with the untranslated regions (UTR) in red and the coding sequence (CDS) in green. The arrow tip indicates the 3'-end of the gene. Only the representative gene model transcript is illustrated. The orange bar indicates the peak area obtained with MACS2. Scale bar, 500 bp. (b) The overexpression *athb25-1D* mutant presents an induction of *ICE1* in developing seeds. Gene expression analysis of *ICE1* in wild type (WT), *athb25-1D* and double mutant *athb22 athb25* developing seeds at maturation stage. (c) The induction of *ICE1* by cold is diminished in *athb22 athb25* double mutant plants. Gene expression analysis in whole plants of *ICE1* in 25-day old WT plants and double mutant *athb22 athb25* plants chilled for 24 hours at 6 ºC. Non chilled plants were also analyzed (No treatment). Expression values are relative to the housekeeping gene *PP2A3* and normalized to untreated wild type. Results are the average of three determinations. The error bars denote standard errors. *, $p < 0.05$ in comparison to WT.

Figure 66: Representative plants of *athb25-1D*, Col-0 wild type (WT), double mutant *athb22 athb25*, and *ice1* plants grown at growth chamber conditions (16 h light/8 h dark, at 23 °C and 70-75% relative humidity). Mutant *athb25-1D* plants present early flowering, reduced leave number, ticker flower stem and rounder leaves compared to wild type plants, while the double mutant plants *athb22 hb25* are smaller. Mutant plants of *ice1* present a similar plant phenotype to the double mutant *athb22 athb25* mutant plants.

(Figure 64b). The double mutant *athb22 athb25* presented a delayed growth that produced a dwarf plant phenotype comparing to wild type plants. In addition, cotyledon yellowing was observed, which might imply an affected nutrient uptake. On the other hand, the overexpressing mutant *athb25-1D* presented bigger leaves and reduced anthocyanin production comparing to wild type plants that might be related to a reduced cold stress. AtHB25 directly binds and to the *ICE1* upstream promoter region (Figure 65a, Renard *et al.*, 2021), and this regulation might explain part of the AtHB25-dependent cold resistance. The overexpression mutant directly regulates *ICE1* expression, confirming *ICE1* as a direct target of AtHB25 (Figure 65b), but no in the double mutant *athb22 athb25*. However, we found that *ICE1* gene induction by chilling was reduced in the double mutant plants in comparison to wild-type plants (Figure 65c).

In addition, we observed that mutant plants of *ice1* grown in growth chamber conditions were similar to *athb22 athb25* plants, presenting a smaller rosette and slightly retarded flowering (Figure 66). In contrast, mutant plants of *athb25-1D* present early flowering reduced leave number, thicker flower stem and rounder leaves.

ICE1 role in endosperm breakdown avoids seed coat fissures formation

To clarify the role of *ICE1* in seed longevity we preformed AAT and CDT aging assays that showed that seed of *ice1* present a reduced seed longevity (Figure 67a). It was easily observable that seed browning was intense in *ice1* two months seeds (Figure 67d) suggesting a higher seed oxidation and quick seed aging, probably due to a higher seed coat permeability. In addition, seed present amorphous shape. The high seed coat permeability of *ice1* seeds in comparison to wild type seeds was demonstrated by the tetrazolium salt assay (Figure 67e), and this high permeability explains the rapid

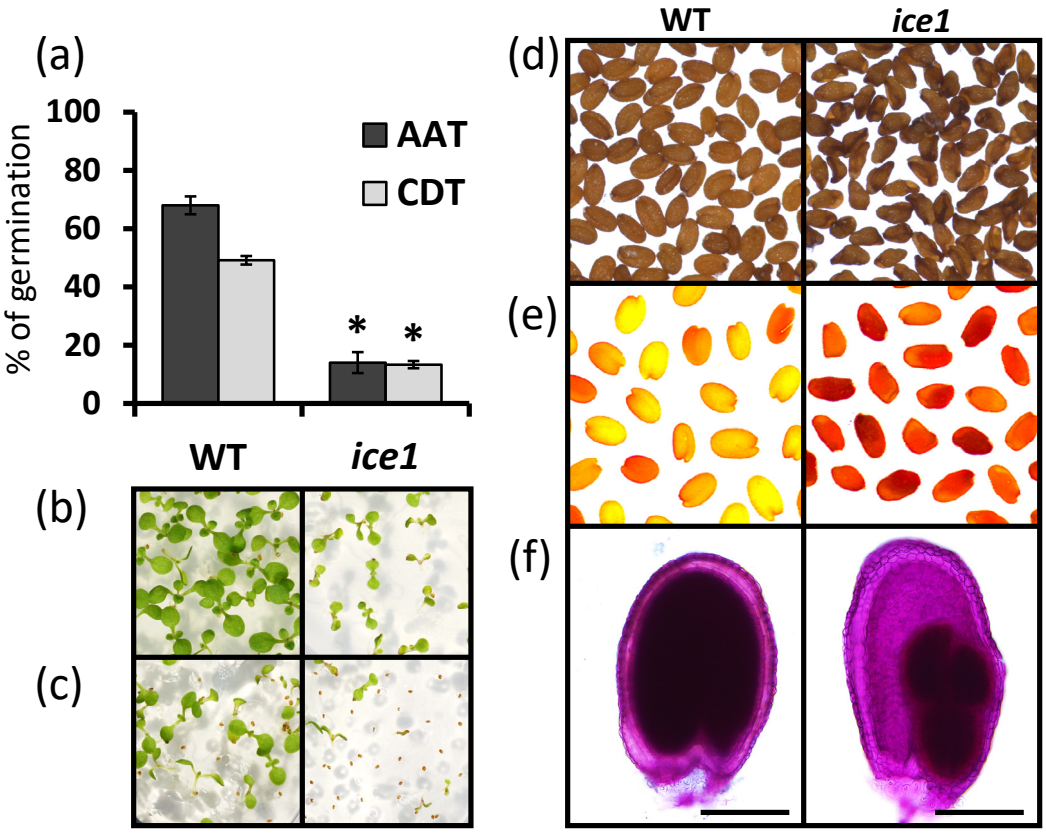


Figure 67: Mutant seeds *ice1* present reduced seed longevity and higher permeability, but similar lipid polyester deposition the seed coat. (a) Arabidopsis seeds from wild type (WT), mutant seeds of ICE1 (*ice1*) and were subjected to accelerated aging treatment (AAT, 40 ºC for 24 hours in water) and controlled deterioration treatment (CDT, 37 ºC for 15 days at 75% relative humidity). The percentage of germination was recorded after seven days. The results are the average of three experiments with 100 seeds per line. The error bars denote standard errors. *, $p < 0.05$. (b) Representative images of WT and *ice1* seeds in control germination conditions (notice that *ice1* seedlings are smaller) (c) and CDT germination (d) Seeds of WT and *ice1* plants stored at room temperature for two months. (e) Representative staining pattern in seeds of WT and *ice1* incubated in 1% tetrazolium red at 30 °C for 72 hours. (f) Sudan red 7B staining of a representative seed of WT and *ice1* for seed coat lipid polyester visualization. The embryo constriction by the endosperm can be visualized. This can lead to embryo developmental defects. Scale bar: 200 µm.

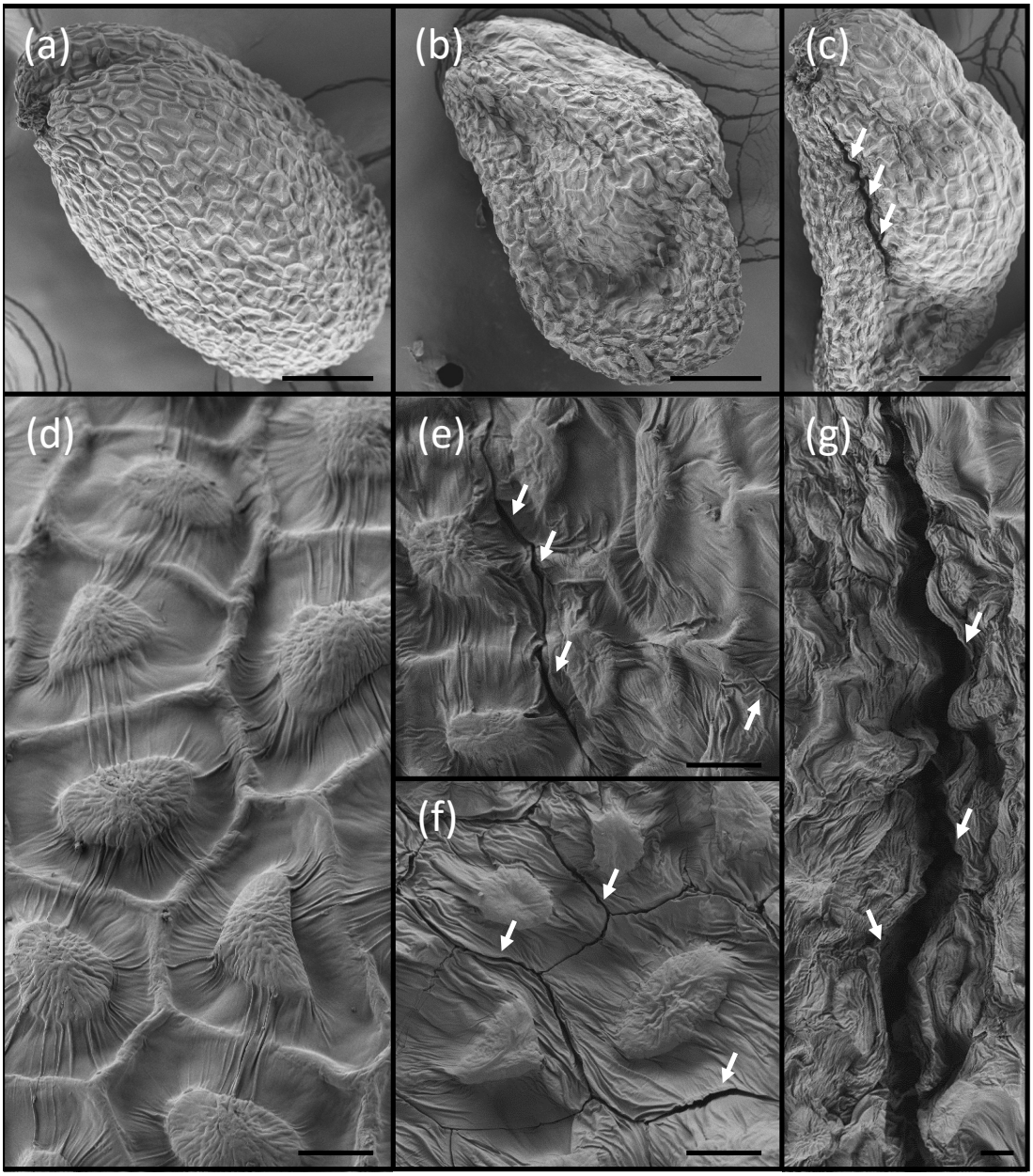


Figure 68: Mutant *ice1* seeds present seed coat fissures. Scanning Electron Microscopy (SEM) of the seed surface of wild type (WT) seeds (a) and *ice1* seeds (b and c). Scale bars, 100 µm. An important seed coat fissure can be observed in the *ice1* seed of the image (c). Magnification of WT seed (d), and *ice1* seeds (e, f and g). Seed fissures can be observed in the seed coat surface of *ice1* seeds, such as slits (e), columella cell fragmentation (f) and big seed coat fissures (g), not observed in WT seed surface (d). Arrows indicate seed coat fissures. Scale bars, 10 µm

PA oxidation of *ice1* seeds. This effect might be caused by a decreased lipid polyester deposition, as in suberin mutants. However, we found that the lipid polyester deposition was apparently similar to wild type seeds when stained with Sudan red (Figure 67f). Seed coat visualization with Scanning Electron Microscopy (SEM) showed that *ice1* seeds presented fissures in the seed coat surface. Different kind of fissures, from slits and columella cell fragmentation to big seed coat fissures, were visualized in different *ice1* seeds, but not in any observed wild-type seed (Figure 68). This seed coat defect in *ice1* seeds directly explains the high seed coat permeability, the rapid PAs oxidation, and the reduced seed longevity of *ice1* seeds.

COG1 expression is downregulated in *ap2* mutants

In an attempt to known COG1 direct targets, we constructed *proUBQ10::COG1:3xHA* overexpressing plants. However, low number of transformed plants were obtained and most of them died during flowering, probably due to the high transcription deregulation produced with higher expression of *COG1*. Only two lines were able to develop successfully seeds and reproduce. These lines presented an important flowering delay and increased leave number as *cog1-2D* plants. However, they presented smaller leaves, contrarily to the *cog1-2D* mutant, which present bigger leaves, comparing to wild-type plants (Figure 69).

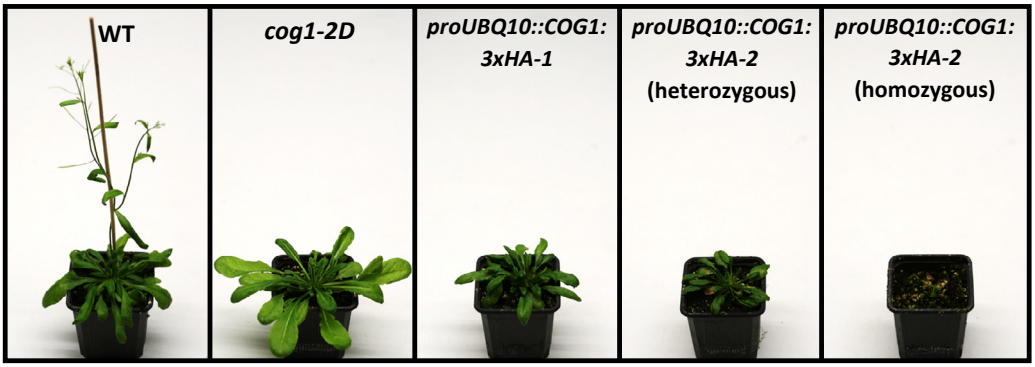


Figure 69: *COG1* overexpression leads to a flowering delay and high COG1 levels are deleterious. Plants were grown simultaneously for five weeks at green house conditions. *COG1* overexpression mutants are presented from left to right according to the expected *COG1* relative abundance. The *cog1-2D* mutant, from the activation tagging collection, present a tandem 35S promoter insertion distanced five kilobases from the 5'-end of the *COG1* gene (Bueso *et al.*, 2016). The *proUBQ10::COG1:3xHA* presented presumably higher levels of *COG1* as the UBQ10 promoter drive directly the *COG1:3xHA* expression. Only two transgenic lines were obtained as other isolated lines died when they flower. Only, one line, *proUBQ10::COG1:3xHA-1*, survived in homozygosis, while the other line, *proUBQ10::COG1:3xHA-2*, only completed its life cycle when presented the insertion in heterozygosis, as died during flowering at homozygosis. Levels of *COG1* produce a flowering delay according with the COG1 relative abundance. In addition, bigger leaves are present in the *cog1-2D* mutant, while *proUBQ10::COG1:3xHA* present smaller leaves, in comparison to wild-type plants (WT)

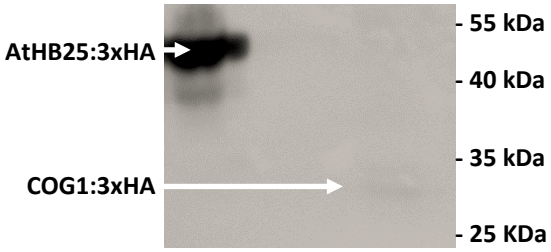


Figure 70: Western analysis of nuclei protein extraction from developing seeds of *proUBQ10::AtHB25:3xHA* and *proUBQ10::COG1:3xHA-1* transgenic plants. The antibody HA.11 (clone 16B12) was used to detect the presence of HA-tagged recombinant proteins. Similar protein loads were used. Arrows indicate the respective bands.

ChIP-seq analysis in developing seeds of *proUBQ10::COG1:3xHA-1* plants, performed similarly to the AtHB25 ChIP-seq (Renard *et al.*, 2021), did not provide significantly enriched genomic regions, presumably to the lower protein abundance observed in western blots (Figure 70). Contrarily to AtHB25 endogenous regulation, we did not perform a Y1H screening in a conserved COG1 fragment.

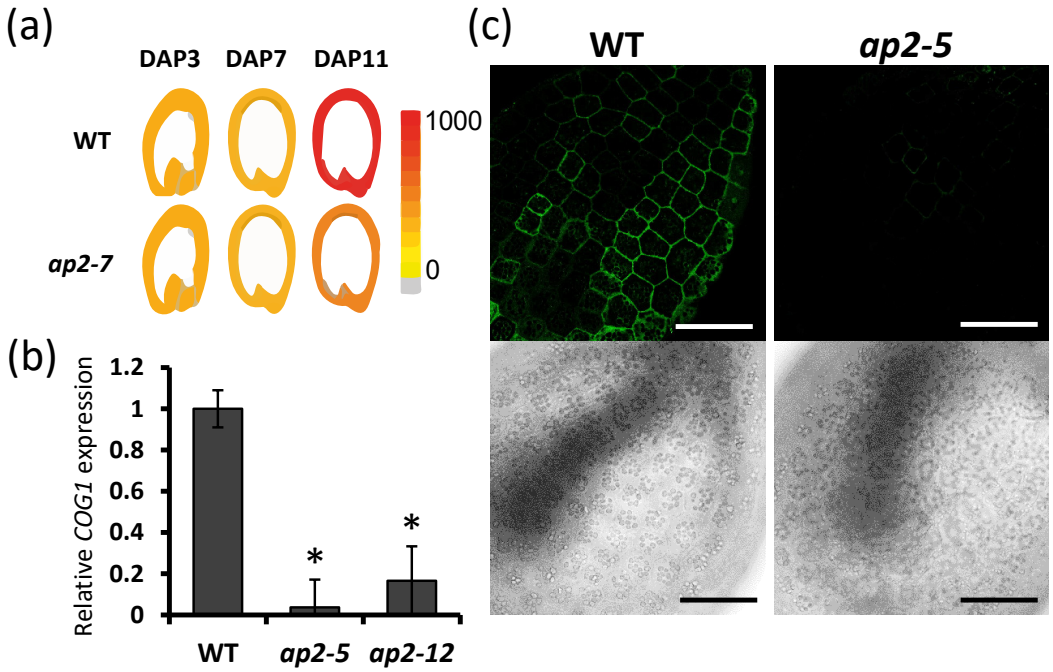


Figure 71: *COG1* is downregulated in different *ap2* mutants. (a) *In silico* expression analysis of *COG1* in the developing seed coat of wild type (WT) and mutant *ap2-7* plants. Images were obtained with the seed coat eFP browser (Dean *et al.*, 2011). (b) Gene expression analysis of *COG1* in *ap2-5* and *ap2-12* in leaves in comparison to wild type plants (WT). Expression values are relative to the housekeeping gene *PP2A3* and normalized to wild type. Results are the average of three determinations. The error bars denote standard errors. *, $p < 0.05$. (c) Fluorescence images 6 days after pollination Arabidopsis seeds transformed with *proCOG1::GUS:GFP* in the WT and *ap2-5* backgrounds using confocal laser scanning microscopy. Scale bar, 100 µm.

Analysis of the *COG1* expression through *proCOG1::GUS:GFP* reporter lines located the expression of *COG1* in the outer integument (Figure 54a, Renard *et al.*, 2020*a*). Thus, we investigated its expression in seed coat development. The Arabidopsis seed coat eFP browser shows that the *ap2-7* mutant does not present the high *COG1* seed coat expression during the seed maturation that is present in the corresponding wild-type plant (Figure 71a). *COG1* expression was assessed in leaves in two *ap2* mutants: a weak mutated *APTETALLA2* allele *ap2-5* (Ohto *et al.*, 2005) and a t-DNA insertion knock-out mutant *ap2-12* (Yant *et al.*, 2010) and both showed *COG1* downregulation (Figure 71b). This demonstrated that the downregulation of COG1 observed in the seed coat eFP browser is directly influenced by AP2, and it is not a secondary effect of the seed coat outer integument COG1 expression, absent in *ap2* strong mutant seeds. Finally, to confirm the outer-integument seed coat *COG1* in *ap2* mutants, genetical crosses of the *proCOG1::GUS:GFP* plants with *ap2*-5 mutants were obtained. Seed coats of 6 day-after-pollination developing seed were visualized with the confocal microscope. Epithelial outer integument cells were similarly developed, while the GFP expression was clearly reduced in the *ap2-5* background seed, in comparison to the wild-type background seed (Figure 71c).

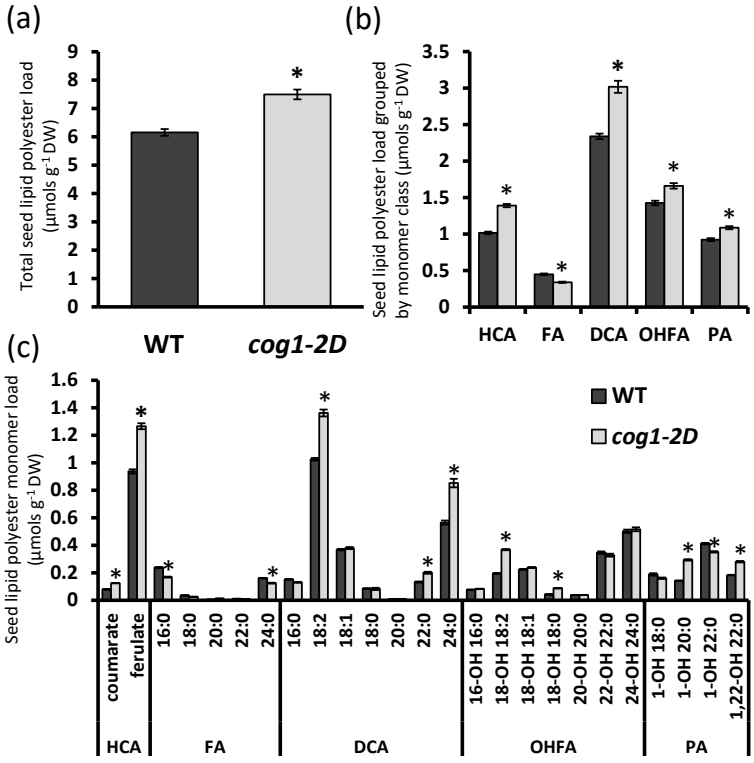


Figure 72: Lipid polyester composition in wild-type (WT) and *cog1-2D* Arabidopsis seeds. (a) Seed lipid polyester monomer composition of the WT and *cog1-2D* seeds. (b) Total lipid polyester distribution grouped by monomer class. (c) Total lipid polyester monomer content. Mean values of three biological replicates are shown in µmols g^{-1} delipidated dry residue (DW). The error bars denote SE. *, $p < 0.05$. HCA: hydroxycinnamic acids; FA: fatty acids; DCA: a,ω-dicarboxylic acids; OHFA: ω-hydroxy fatty acids; PA: primary fatty alcohols.

COG1 recovers seed suberin accumulation and seed longevity in the *ap2-5* mutant

The reduced longevity in *ap2* mutants might be related with the seed coat outer integument differentiation, which consists in two layers: the inner layer is named palisade layer, and the outer layer is named epithelial layer, which produces suberin and mucilage, respectively. The mucilage does not determine seed longevity, as mutants in mucilage production do not present changes in seed longevity (Figure 33, Renard *et al.*, 2021). However, suberin deficient mutants present reduced seed longevity (Figure 44, Renard *et al.*, 2020*a*). Gas-chromatography analysis of seed lipid polyesters showed an increase of total monomers in the *cog1-2D* mutant (Figure 72a). Generally, except fatty acids, which were reduced, the other monomers groups were increased in *cog1-2D* mutant seeds (Figure 72b). The ferulates, such as the coumarate, were increased in *cog1-2D* mutant (Figure 72c), corroborating the contribution of peroxidases to phenolic monomer polymerization in this mutant described (Renard *et al.*, 2020*a*).

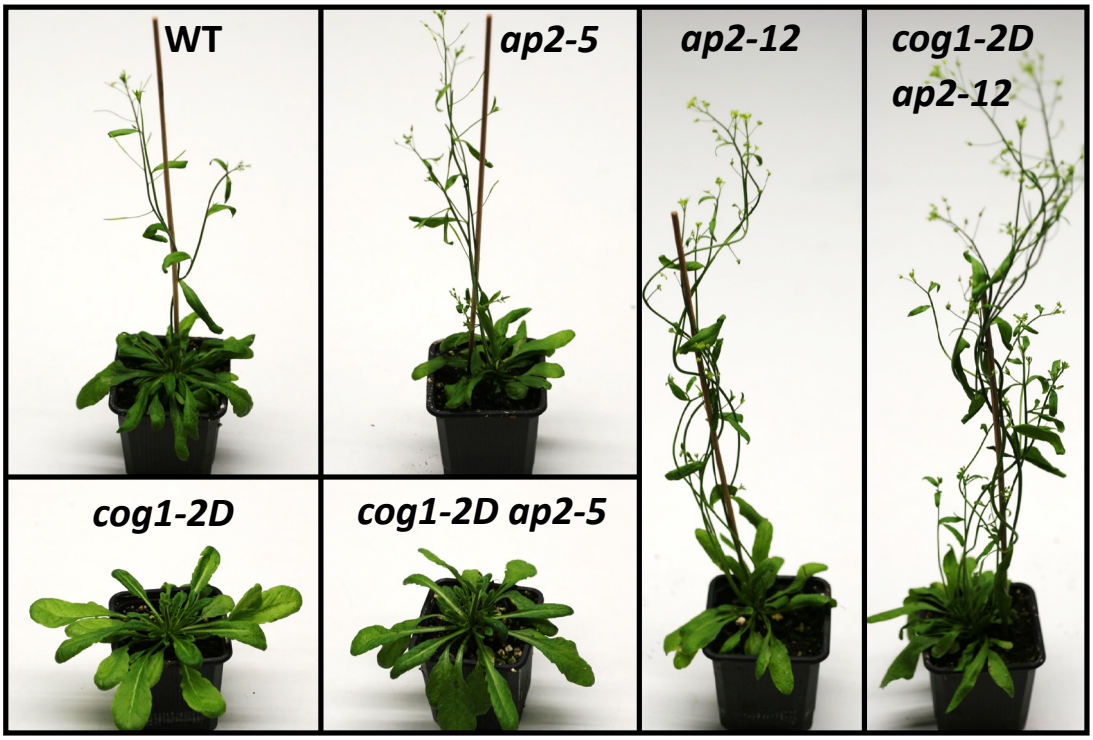


Figure 73: Flowering delay resulting of *COG1* overexpression is dependent of AP2. Plants were grown simultaneously for five weeks at green house conditions. Plants of *ap2*-5, and specially *ap2-12* present early flowering in comparison to wild type (WT) plants. The mutant phenotype of *cog1-2D* is reverted in the *cog1-2D ap2-12* double mutant, but not in the *cog1-2D ap2-5*. However, leaves are smaller in *cog1-2D ap2-5* in comparison to *cog1-2D*.

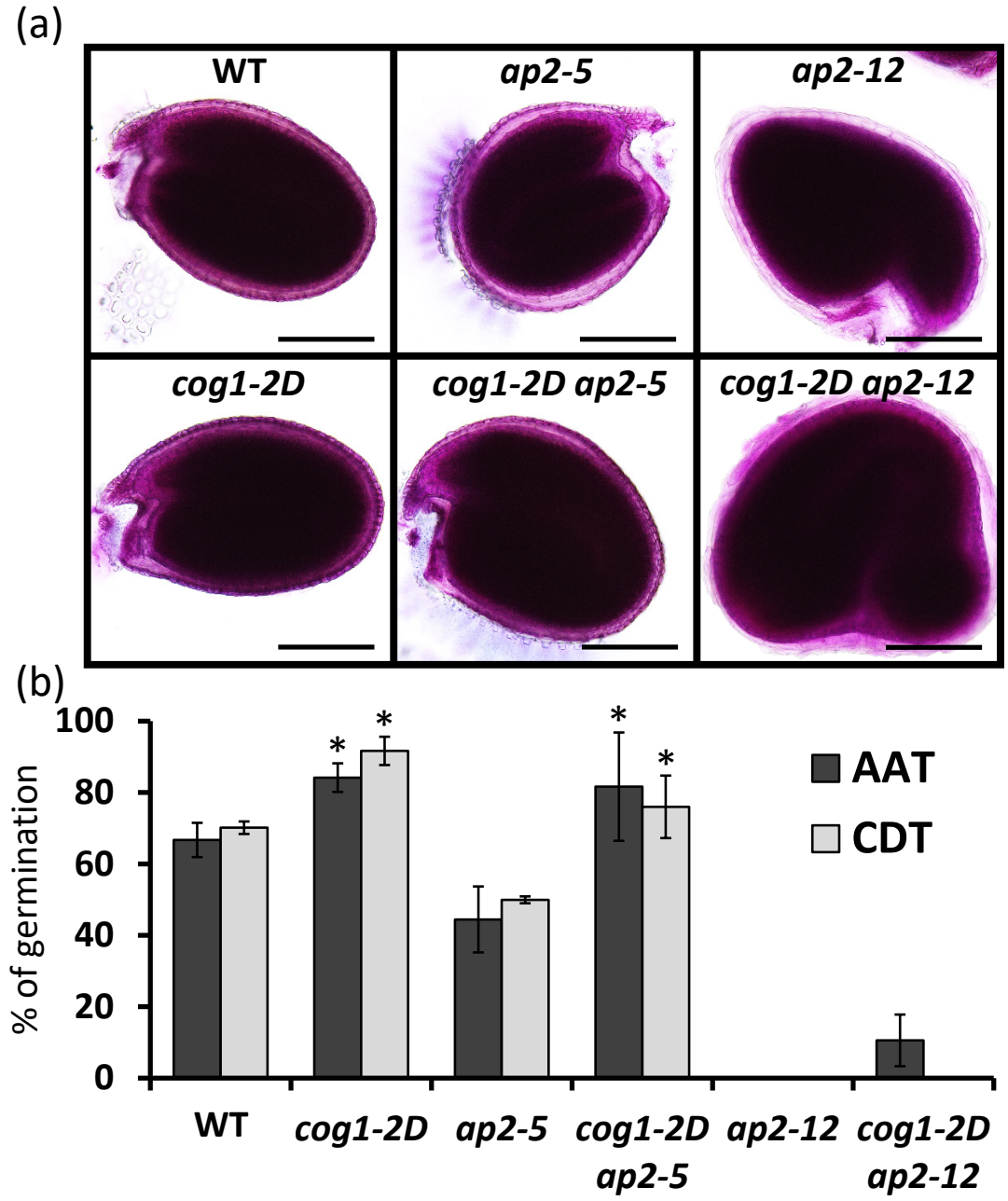


Figure 74: COG1 partially recovers the seed suberin accumulation and the seed longevity reduced in *ap2* mutants. (a) Representative image of Sudan red 7B staining of a representavitve seeds of *ap2* mutants, *cog1-2D* and their respective geneical crosses. *COG1* is downregulated in different *ap2* mutants. Scale bars, 200 µm. (b) Artificial aging assays AAT (41 ºC for 24 hours, dark grey bars) and CDT (15 days at 37 ºC and 75% relative humidity, light grey bars) in *ap2* mutants, *cog1-2D* and their respective geneical crosses. Seeds of *ap2* and *cog1-2D ap2-12* mutant plants were not bleach sterilized, but ethanol 70% washsed. The results are the average of three experiments with 50 seeds per line. The error bars denote standard errors. Statistical tests were perfomed to asses th effect of COG1, in the different *ap2* mutant and wild type backgrounds. *, $p < 0.05$.

Double mutants *cog1-2D ap2-5* and *cog1-2D ap2-12* were obtained to determine if the effect of COG1 recovered the reduced seed longevity of the *ap2* mutants. Interestingly, double mutant plants *cog1-2D ap2-12* did not present the characteristic flowering delay of *cog1-2D* plants (Figure 73), suggesting that COG1 control of flowering requires AP2. This effect was not observed in the genetic cross of the weak allele *ap2-5*. In seeds, suberin production is localized in the inner layer of the outer integument, and it is affected in *ap2* mutants. Lipid polyester staining showed that the double *cog1-2D ap2-5* mutant recovered wild type suberin levels through COG1 overexpression, but not *cog1-2D* levels. However, the double *cog1-2D ap2-12* mutant did not recover the suberin accumulation, as seed coat epithelial cells are completely undifferentiated (Figure 74a). It is clearly visible that *ap2-12* mutant seeds do not present suberin in its undifferentiated palisade layer. However, the cuticle layer is perfectly formed in this mutant. In addition, seeds are amorphous, rounder, and bigger. These developmental effects are exacerbated in the double mutant *cog1-2D ap2-12*.

A similar effect was seen in the different artificial aging experiment. Double *cog1-2D ap2-5* mutant seeds recovered wild type germination levels while double *cog1-2D ap2-12* mutant seeds not, in both artificial aging treatments, AAT and CDT (Figure 74b). Mutant seeds *ap2-12* and *cog1-2D ap2-12* present a highly reduced germination and they cannot be bleach sterilized as their germination ratio rapidly decays and were sterilized only with a quick ethanol 70% wash in the different seed aging assays.

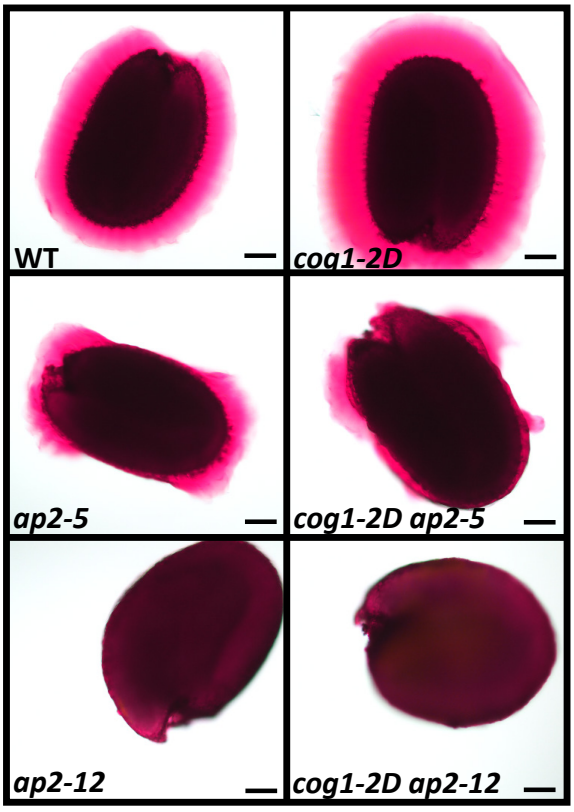


Figure 75: *COG1* overexpression does not recover the affected mucilage halo observed in *ap2* mutants. Ruthenium red mucilage staining was used to visualize the mucilage halo. Scale bars, 100 µm.

COG1 is involved in columella cell differentiation

The AP2 TF is involved in the outer integument differentiation. As COG1 overexpression recovered partially the palisade layer suberin deposition we asked if COG1 also recovered part of the seed coat epithelial cell differentiation lost in strong *ap2* mutants. The mucilage area projection, increased in *cog1-2D* mutant seeds compared to wild type, was not recovered in the double crosses *cog1-2D ap2-5* and *cog1-2D ap2-12*,

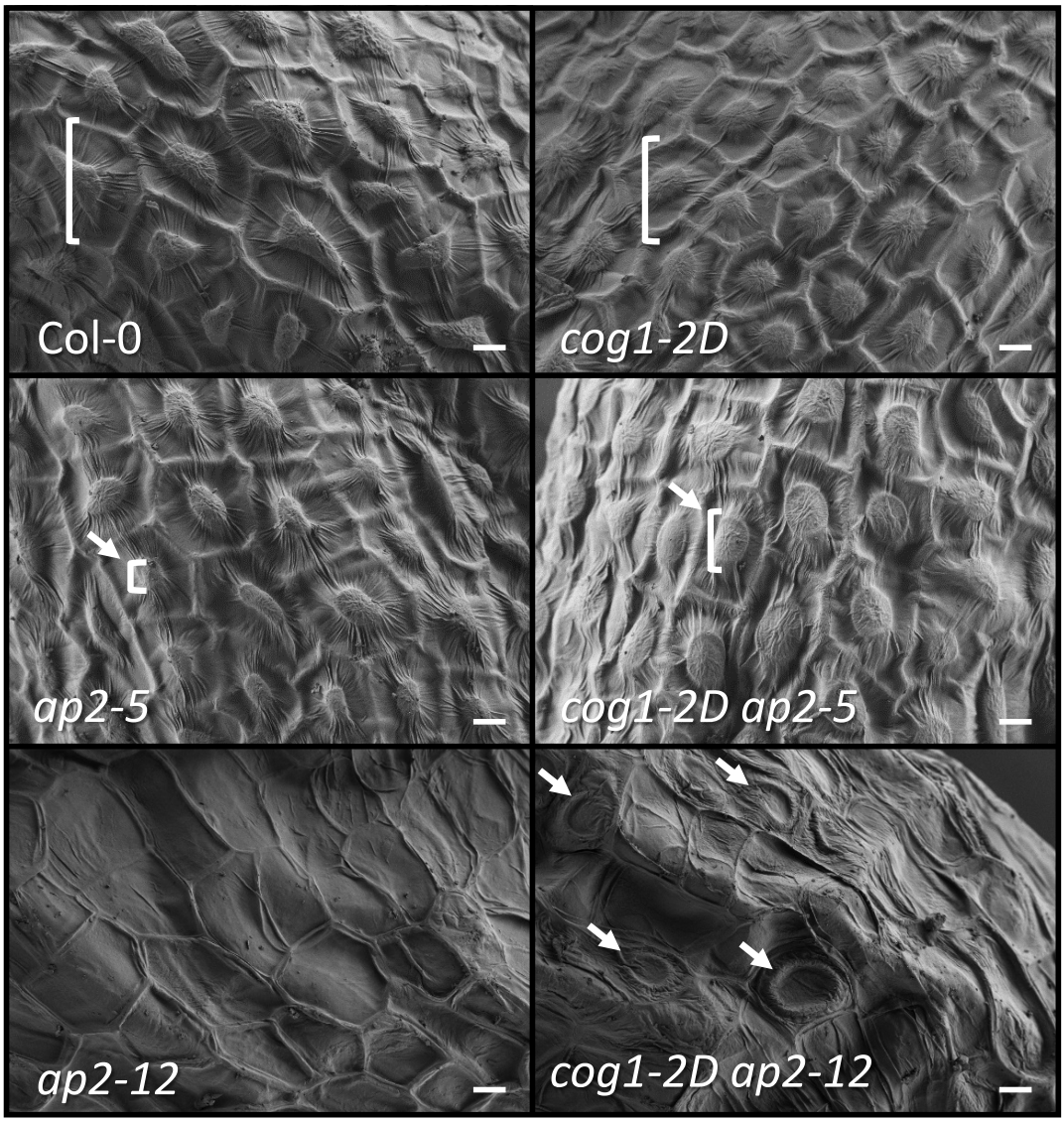


Figure 76: COG1 participates in columella cell differentiation. Scanning Electron Microscopy (SEM) of the seed surface. Brackets in upper images indicate the size of columella cells, smaller in the *cog1-2D* mutant. Brackets and the arrow in middle images indicate the volcano structure size, bigger in the double mutant *cog1-2D ap2*-5. Arrows in the lower images indicate the columella-like structures observed in the double mutant *cog1-2D ap2-12*, but not in the *ap2-12* mutant. Scale bars, 10 μm.

comparing to the respective *ap2* mutants (Figure 75). However, seed surface visualization with Scanning Electron Microscopy (SEM) showed differences in the *ap2* mutants and the respective double crosses with *cog1-2D* (Figure 76). Columella cells in *cog1-2D* seed were smaller than in wild type seeds. In addition, a thicker adjacent-cell wall reinforcement can be observed, as in the double mutant *cog1-2D ap2-5*. This double mutant, in comparison with *ap2-5* present a bigger volcano structure. Finally, the genetic cross with *ap2-12* mutant, which completely fails in outer integument differentiation and there is not volcano structure in their seed coat epithelial cells, slightly recovered columella cell differentiation as columella-like structures were visualized in the seed coat surface of the double mutant *cog1-2D ap2-12* (Figure 76). This effect directly links COG1 to columella cell differentiation.

Discussion

The seed coat is a specialized mother-origin tissue that protects the seed to extend their life span. Transcription factor regulation is crucial for its complex differentiation and specialization from ovule integuments. The transcription factors AtHB25 and COG1 extend seed longevity in the respective overexpressing lines *athb25-1D* and *cog1-2D* (Bueso *et al.*, 2014*b*; Bueso *et al.*, 2016). Both TFs are expressed during seed development (Renard *et al.*, 2020*a;* Renard *et al.*, 2021) observed through promoter-reporter fusion lines. The idea to situate AtHB25 and COG1 in the seed coat TF regulation network leads us to investigate their regulation and to test protein-protein interaction with different transcription factors described to participate in the seed coat development. However, no interactions were found between AtHB25, COG1 and the described TFs through the Split-Tpr system used. The only positive result obtained was that AtHB25 interacting with itself forming homodimers. This interaction was already described (Tan and Irish, 2006). Despite the negative interactions found, we cannot discard that proteins might interact *in planta.* The Split-Tpr system used is less robust that the yeast-two-hybrid system, which leads to a higher false negative rate. In addition, it is possible that a third protein is needed to form ternary complexes, common in seed coat differentiation (Li *et al.*, 2020). The yeast-three-hybrid system may be used to assess ternary protein complexes. Nevertheless, the more plausible explanation is that COG1 and HB25 do not act at the same regulatory level of tested TFs. In fact, most of the transcription factors tested present important embryo and seed coat development alterations, while knock out mutant *athb22 athb25* and *cog1* apparently do not present such strong structural alterations (data not shown). Thereby, probably COG1 and AtHB25 are involved in seed coat development but not determining tissue identity of the seed coat. To identify protein interactors of AtHB25 and COG1, a screening approach should be used. However, the TFs library from Oñate-Sanchez (CNB, Madrid) could not be used for this porpoise as it was used in the Y1H screening, because it was designed for the Y2H and Y1H systems, and not for Split-Tpr.

AtHB25 and cold regulation

The regulation of *AtHB25* was studied with a selected conserved promoter fragment among close Arabidopsis plant species. This promoter fragment, which presented transcriptional activity in *Nicotiana benthamiana* was used for a Yeas-one-hybrid screening against 1,358 different TFs. AtHB52, bZIP52, ALG24 and a DREB2 family member were found as positive regulators of this conserved fragment, but only the DREB2H presented abundant expression during seed development, which may explain the seed longevity regulation of AtHB25. This TF belong to the TFs DREB2 subfamily of the of ERF/AP2 transcription factor family, composed by eight members in Arabidopsis. They specifically bind to DRE/CRT cis elements that are responsive to drought and low-temperature stresses, regulating gene expression of different water- stress inducible genes. Curiously, the selected *AtHB25* promoter fragment did not contain a DRE/CRT element (A/GCCGAC) (Sakuma *et al.*, 2002; Lata and Prasad, 2011), but contained a GCCCAT sequence, candidate for the DREB2 TF recognition. Further study is needed to determine if the DREB2 TF binds to this sequence to regulate the AtHB25 expression. DREB2A, the most studied member of this subfamily, and other subfamily members respond to cold, salt, ABA and drought in different plant species (Nakashima *et al.*, 2000, Sakuma *et al.*, 2002, Lim *et al.*, 2007, Agarwal *et al.*, 2017). The expression of this *DREB2 TF* is up regulated with chilling, As DREB2 TFs respond to drought, it is plausible that seed coat *DREB2H* expression is related to seed coat dehydration. Nevertheless, the up regulation of *DREB2H* by chilling might affect to seed coat performance in response to low temperatures, involving *AtHB25* regulation. *AtHB25* was confirmed to be induced after chilling by promoter-driven GUS expression, matching with *DREB2H* expression. The GUS protein is stable for long periods (Jefferson *et al.*, 1987), so the measurement at 24 hours after chilling can detect the induction of AtHB25 by chilling even if it occurred up to 24 hours earlier. Further analyses, such as *AtHB25* expression analysis in different *dreb2* mutants, are necessary to clarify if the cold induction of *AtHB25* is dependent of this DREB2 TF.

Even if this is not the case, we found that AtHB25 is involved in the cold resistance as the double mutant *athb22 athb25* was highly affected in continuous cold comparing to wild type plants, while the overexpression mutant presented reduced cold-induced anthocyanidins, and bigger leaves, probably indicating a reduced cold stress (Figure 63). Overexpression of DREB2 TFs lead to abiotic stress resistance, including in some cases cold stress, in different plant species (Agarwal *et al.*, 2017). Perhaps part of this DREB2 stress tolerance depends on AtHB25 regulation by DREB2 TFs. In addition, the high conservation of the AtHB25 amino acid sequence in angiosperms (Renard *et al.*, 2021) and its conserved promoter fragment in close Arabidopsis species points to an important role of AtHB25. It is plausible that this role is related with cold response for plant adaptation. Furthermore, AtHB25 directly regulates ICE1 (Figure 59,

Renard *et al.*, 2021), a TF involved in the induction of around 40% of cold-response genes (Lee *et al.*, 2005; Chinnusamy *et al.*, 2007). ICE1 SUMOylation controls CBF3/DREB1A transcript levels in response to cold, for the regulation of cold tolerance (Miura *et al.*, 2007; Zhou *et al.*, 2011). Thus, the induction of ICE1 transcript levels by AtHB25 may explain part of the cold resistance regulation dependent of *AtHB25* observed in the mutant *athb25-1D*. The induction of ICE1 by chilling was reduced in the *athb22 athb25* double mutant. This effect can partially explain the cold susceptibility of *athb22 athb25* plants. Thereby, AtHB25 regulates *ICE1* cold regulation binding to the *ICE1* upstream promoter region. Probably *ICE1* regulation by chilling is controlled partially by AtH25, activating part of the genes necessary for low-temperature adaptation and resistance. To confirm the role of *ICE1* in *athb25-1D* cold resistance, genetic crosses to generate a double mutant *athb25-1D ice1* must be performed to determine if the cold resistance phenotype of *athb25-1D* is lost.

ICE1 TF not only controls freezing tolerance, but also regulates different developmental aspects such as the endosperm breakdown, primary seed dormancy, stomatal cell differentiation and ABA signaling (Kanaoka *et al.*, 2008; Denay *et al.*, 2014; MacGregor *et al.*, 2019; Hu *et al.*, 2019). We found that ICE1 is also involved in seed longevity. Mutant seeds of *ice1* plants present a reduced seed germination through a higher seed coat permeability, but not altered lipid polyester deposition (Figure 67). These seeds are amorphous and present seed coat fissures that permit a rapid air entry and quick seed aging. This effect is probably due to endosperm breakdown role of ICE1 (Denay *et al.*, 2014). Mutants seeds of *ice1* fail in the endosperm breakdown and must present higher internal pressure, as *zou* mutant seeds (Creff *et al.*, 2015; Fourquin *et al.*, 2016). The high internal pressure produces the observed seed coat fissures that allows the air entry and the rapid seed oxidation (Figure 68). This effect must accelerate aging processes and reduce seed longevity. In addition, this effect produces embryo constriction, and it leads to smaller seed embryos (Figure 61f) and thus, smaller seedlings (Figure 67b). Probably, resulting small embryos present developmental defects which are also negative for seed longevity. Thus, the main seed longevity effects of *ice1* mutant seeds are not related with lipid polyester deposition in the seed coat, a key aspect of the increased seed longevity of the mutant *athb25-1D*. However, other effects, uncovered by these seed coat fissures, might be related to the regulation of seed longevity by AtHB25, such as the hormone level control.

The plant phenotype of the overexpression mutant *athb25-1D* involves long hypocotyl, early flowering, bigger and rounder leaves, thicker stem, and wider siliques, and it is probably caused in part by the increase of GAs levels of this mutant (Figure 68; Bueso *et al.*, 2014*b*). In addition, the *athhb25-1D* also present increased ABA levels (Bueso *et al.*, 2016). Both different hormone levels might be caused by the *ICE1* regulation. Moreover, the similar phenotype of *athb22 athb25* and *ice1* plants may be

also related to ICE1 hormone regulation. ICE1 interacts with ABI5 and DELLA proteins, negatively regulating ABA signaling in seed germination (Hu *et al.*, 2019), and probably also during seed development. In addition, ICE1 binds to the ABI3 promoter and positively regulates it (MacGregor *et al.*, 2019). ABI3 is an important regulator of seed maturation and seed longevity (Sano *et al.*, 2016; Leprince *et al.*, 2017). Genetic cross of *ice1* with the overexpression mutant *athb25-1D* must be performed to determine if the GAs content and the different associated phenotypes are due to *ICE1* overexpression. GAs are important for seed coat development and the increased GAs levels contribute to the enhanced seed longevity, as endogenous GAs treatment and constitutive GA-signaling mutant (DELLA mutants) lead to seeds with enhanced seed longevity (Bueso *et al.*, 2014*b*).

Suberin deposition is a key aspect of the seed longevity of *AtHB25* mutants. Suberin deposition is regulated by ABA levels (Soliday *et al.*, 1978, Cottle and Kolattukudy, 1982), through a transcriptional program involving MYB41 (Kosma *et al.*, 2014). In seeds, the close homologue MYB107, which regulates seed coat suberin and is essential for seeds longevity and seed coat sealing (Gou *et al.*, 2017; Renard *et al.*, 2020*a*), might also respond to the different ABA levels to regulate suberin deposition, up-regulating different suberin biosynthetic enzymes, such as GPAT5 (Gou *et al.*, 2017). Levels of ABA are increased with low temperatures (Kendall *et al.*, 2011), including seeds maturated at low temperatures (MacGregor *et al.*, 2019). It is plausible that this regulation involves AtHB25. In the seed coat, cold temperatures increase the expression of diverse suberin biosynthetic enzymes, such as *LACS2*, and results in a different lipid polyester composition of the seed coat and reduces seed dormancy, probably due to a higher permeable seed coat that oxidates seed coat PAs (Fedi *et al.*, 2017). Seed dormancy dependent of low temperatures is also regulated by PA deposition as shown in diverse Arabidopsis ecotypes (MacGregor *et al.*, 2015). AtHB25 directly regulates suberin biosynthetic genes, such as *LACS2* and *NHO1,* controlling suberin and cutin deposition in seed coat (Renard *et al.*, 2021). Hereby, it is plausible that part of this cold regulation of suberin deposition in seeds involves AtHB25. The effect of low temperatures in lipid polyester composition must be assessed in *AtHB25* mutants to confirm the putative role of AtHB25 in the temperature regulation of lipid polyester composition.

Further research is necessary to reveal the key aspects of the AtHB25 cold regulation and how affect seed longevity. Both, hormones and direct regulation of lipid polyester biosynthetic enzymes are apparently key aspects of this seed longevity adaptation. The role of the AtHB25-direct target NHO1 might be crucial in cold adaptation as its substrate, glycerol is an important cryoprotectant and lipid precursor.

COG1 is regulated by AP2 and contributes to outer integument differentiation

To characterize the endogenous role of COG1, we followed a different strategy. COG1 direct targets were not obtained through a ChIP-seq, probably because low protein dosses produce a high toxicity in plants (Figure 69 and Figure 70). The developmental defects produced by the COG1 expression might have been avoided with the endogenous COG1 promoter. The Y1H screening in a COG1 conserved promoter fragment was not performed, in contrast to AtHB25. The low COG1 amounts and the high protein toxicity in plant might affect negatively to the assay. Nevertheless, the high expression on COG1 in the outer integument and published transcriptional analysis in *ap2* seed coat suggested a positive role of AP2 in the COG1 regulation (Dean *et al.*, 2011; Renard *et al.*, 2020*a*; Figure 71a). To corroborate the expression analysis, *proCOG1::GUS:GFP* lines in the weak allele *ap2-5* background showed a reduction of the GFP signal in the outer integument of developing seeds (Figure 71c). The *COG1* expression was confirmed to be AP2-dependent through expression analysis in leaves of *COG1* in *ap2* mutants (Figure 71b).

AP2 is a TF important in flowering, for petal differentiation and for seed coat development controlling the outer integument differentiation (Kunst *et al.*, 1987; Leon-Kloosterziel *et al.*, 1994). Mutant seeds of *ap2* mutants present a reduced seed longevity, and no outer integument differentiation, lacking columella cell differentiation, mucilage and suberin production (Western *et al.*, 2001, Molina *et al.*, 2008), resulting in reduced seed longevity (Debeaujon *et al.*, 2000). These effects can be visualized in the knock out *ap2-12* in the Figure 74, Figure 75 and Figure 76. In addition, AP2 controls seed mass (Ohto *et al.*, 2005). COG1 acts downstream of AP2 and the overexpression of COG1 can revert partially different aspects of the outer integument differentiation. Although the differentiation of outer integument cell was not completely recovered, and the associated biomolecules, the suberin and the mucilage, were still absent in *cog1-2D ap2-12* double mutant seeds, we could observe columella-like structures in the seed surface of these double mutant seeds (Figure 76). In addition, *cog1-2D* mutant presented smaller columella cells with thicker adjacent cell walls. Thereby, COG1 is involved in columella cell differentiation, acting downstream of AP2 regulation. Effects on suberin synthesis and seed longevity were visible with the *cog1-2D ap2-5* double mutant. The weak mutant allele *ap2-5* presented a diminished seed longevity and reduced suberin deposition, compared to wild-type plants, which was recovered with the COG1 overexpression (Figure 74). However, this effect was not seen in the *ap2-12* probably due to the undifferentiated outer integument tissue. Taken together, we found that COG1 by itself cannot complement the AP2-dependent differentiation of the outer integument, but its transcriptional activity is involved in this tissue differentiation. Negative results of COG1 protein-protein interaction with known seed

coat TFs suggest that COG1 is not interacting with the outer integument regulators AP2, GL2, TT8, TTG1, TTG2, EGL3, MYB5, MYB61, KNAT7 or STK (Golz *et al.*, 2017). Thereby, the outer integument regulation of COG1 may be independently to the transcriptional regulatory program. COG1 is involved in light perception through the negative regulation of the phytochrome signaling (Park *et al.*, 2005). Light signaling pathway is important for seed developmental processes determining seed performance (He *et al.*, 2014), and it determines seed lipid polyester deposition, seed permeability and seed longevity (Bueso *et al.*, 2016). Thereby, it is plausible that COG1 directly influences in the outer integument differentiation in response to light signals. This effect must be important for the seed plasticity and the balance of seed dormancy and seed longevity. Indeed, the overexpression of COG1 produces an increase of GAs and ABA, directly affecting both adaptative traits. The red/far red-light ratio signaling determines seed longevity (Contreras *et al.*, 2009), and thus COG1 might be responsible of this regulation important for geographical seed performance adaptation, probably through hormone regulation. Indeed, both hormones, ABA and GAs, are required for the seed coat development. The AP2 regulation of COG1 must be important for COG1 expression in the seed coat.

The overexpression of *COG1* in the mutants *cog1-2D* and *proUBQ10::COG1:3xHA* produced flowering delay flowering delay, probably due to light-dependent flowering regulation. This flowering delay was completely lost in the doble mutant *cog1-2D ap2-12* (Figure 73) suggesting that the COG1 flowering-control depends on AP2. Further investigation is needed to stablish the role of AP2 in flowering COG1-dependent. Perhaps COG1 and AP2 constitute a feedback loop to regulate flowering and seed coat in a light dependent manner.

Concluding remarks

In conclusion, the regulation of AtHB25 and COG1 stablish direct links in cold, light and flowering in seed coat development regulating seed longevity. The cold regulation of *AtHB25* is probably regulated by DREB2H and this regulation might explain the cold-induced lipid polyester regulation found in seed coat (Fedi *et al.*, 2017). Temperature also affects to potato suberin composition (Dean, 1989). Changes in suberin composition might help to control permeability to water and air, probably important for cold acclimatization. *COG1* is regulated by AP2 in the seed coat, and probably regulates the outer integument composition in a light dependent way, regulating lipid polyester deposition. In addition, COG1 controls flowering through the action of AP2, which may suppose a feedback loop controlling flowering signals in a light dependent way.

3. General discussion

This PhD thesis aimed to identify new seed-longevity genes in the model plant *Arabidopsis thaliana* using two approaches using different genomic molecular techniques. First approach took advantage of the wide Arabidopsis natural variation to find new seed longevity genes using GWAS. The second main approach consisted in further characterize the positive effect in seed longevity of the TFs HB25 and COG1, described previously in our lab by the overexpression mutants *athb25-1D* and *cog1-2D*, which presented enhanced seed longevity (Bueso *et al.*, 2014*b*; Bueso *et al.*, 2016). We used the genomic techniques ChIP-seq for AtHB25 (not possible for COG1) and RNA-seq for COG1 characterization. In addition, Y1H, protein-protein interaction experiments and gene expression assays were used to determine part of the expression regulation of the TFs AtHB25 and COG1.

The GWAS analysis identified different molecular mechanisms controlling seed longevity. In addition, the seed coat protective effects were remarked in both, in the GWAS study and in AtHB25 and COG1 characterization. TFs factor regulating seed coat development appeared as important seed longevity regulators which may act as positive or negative seed longevity determinants. Lipid polyester barriers present in the seed coat, the suberin layer and the cuticle, were found crucial for seed longevity protecting the embryo from the environment. Diverse biosynthetic enzymes involved in the development of these protective layers with a direct effect in seed longevity were described in this PhD thesis, from the seed longevity GWAS (CYP86A8), the AtHB25 ChIP-seq (LACS2 and NHO1) and the COG1 RNA-seq (PRX2 and PRX25).

In addition, the regulation of seed longevity is not only genetically, but also environmentally controlled. AtHB25 might control seed lipid polyester deposition by temperature cues, and COG1, involved in light perception (Park *et al.*, 2003; Bueso *et al.,* 2016), controls flowering and seed coat outer integument formation. Thereby, it is plausible that COG1 participates in light-seed longevity determination.

Validation of seed aging methods with Arabidopsis

In our work, we found that Arabidopsis ecotypes presented a wide variation in seed longevity among in all four seed aging treatments (the artificial aging treatments: AAT, CDT and EPPO, and the dry seed storage treatment or NAT) (Figure 12). Genes obtained through GWAS were mainly from artificial treatments, as NAT data was obtained later and NAT-GWAS genes were more difficulty validated for the reduced remaining time for the research (Figure 14, Table 2). In addition, the significance of the associated SNPs was lower (Figure 14, Supplemental data 3). Although artificial aging

treatments are used to predict how seeds would age during dry storage, we did not observe a perfect correlation between aging treatments in Arabidopsis ecotypes. The wide range of ecotypes used (270) must have reduced the correlation coefficients (Figure 13). The treatment that most correlated with NAT was CDT (r_s = 0.68), followed by EPPO (r_s =0.47). The AAT presented a low correlation index (r_s = 0.11). Indeed, AAT is the most different aging treatment compared to NAT, as occurs for short time in hot water imbibition. This might cause membrane rupture through a rapid water uptake, causing physical damages. Nevertheless, the no-perfect correlation between artificial treatments must occur due to the different widespread genomic changes in the seed longevity molecular machinery due to the different Arabidopsis genomic backgrounds. In this thesis, as in other studies, same-background Arabidopsis lines presented a similar behaviour among the different seed aging treatments: AAT and CDT (Figure 48, Figure 55 and Figure 74); AAT, CDT and NAT (Table 3, Table 4, Table 5, Table 6 and Table 7); and AAT, CDT, EPPO and NAT (Bueso *et al.*, 2014*b*, Figure 31, Figure 32 and Figure 40). Nevertheless, the differences among aging treatments in ecotypes suggest that each aging treatment affect similar and different aspects of the machinery deployed by the seed to avoid deterioration. These differences permitted to obtain different valuable information with the different seed longevity GWAS. Highlighted genes among the different GWAS were confirmed to directly affect the seed longevity trait by knock-out mutant analysis. It is curious that although AAT was the lowest correlated aging treatment with the others, NAT was the treatment with lowest coinciding genes close to significant SNPs (Table 2, Supplemental data 3). The corroborated effectiveness of artificial aging treatments to emulate natural seed aging observed in Col-0 background mutant lines would be used among the discussion of this PhD thesis as indicative of increased or decreased seed longevity of the different seed lots compating to the simultaneously grown wild type seed lots. Validation with natural seed aging treatment was used when possible, but seed storage for long periods was not possible for all seed lots.

New embryo genes important for seed longevity

The GWAS analysis identified different genes involved in known and unknown seed longevity mechanisms. However, none of them was apparently related with repair mechanisms. It is possible that repair mechanisms contribute slightly to seed aging recovery and they were not detected due to our methods and significance tests. Nevertheless, maturation desiccation tolerance genes and ROS-related embryo molecular mechanisms controlling seed longevity were found with the GWAS analysis and corroborated by reverse genetics.

The maturation protein SSLEA protein constitute the second Late embryogenesis

abundant protein directly described in seed longevity, after LEA14 (Hundertmarkt *et al.*, 2011). LEA proteins accumulates during seed maturation and desiccation (Verdier *et al.*, 2013), and they participate in seed dehydration protection for seed desiccation (Leprince *et al.*, 2017) and other stresses (Hong-Bo *et al.*, 2005). However, fresh seeds of *sslea* mutant plants did not presented an affected germination, suggesting that they can tolerate seed dehydration. The protective effect of SSLEA during seed aging should be further studied to determine its molecular mechanism. Other members of the group 6 of LEA proteins, according to Jaspard *et al.* (2012) classification, can also present an important role on seed longevity.

Seed aging damages are mainly produced by ROS (Kurek *et al.*, 2019). The three ROS main sources in living cells are the electron transport chain, the nicotinamide adenine dinucleotide phosphate oxidases (NADPH oxidase) and the 5-lypoxigenase (Novo and Parola, 2008). However, ROS constitute and important stress and developmental signals (Miller *et al.*, 2008; Suzuki *et al.,* 2011; Marino *et al.*, 2012). We found that three NADPH oxidases or RBOH, expressed in seeds during seed development, were negative for seed longevity, as the mutant seeds of *rbohd, rbohe, rbohf* and double mutant seeds *rbohd,f* presented increased seed longevity. They are expressed in seeds probably due to the ROS developmental role in embryogenesis, seed dormancy and germination (Murphy *et al.*, 1998; Bailly *et al.*, 2008; El-Maarouf-Bouteau and Bailly, 2008). However, the residual activity in dry seed is deleterious for seed longevity. In addition, it is plausible that seed ROS production help to protect seed against pathogens, and RBOH activity might contribute to this protection (Jeevan Kumar *et al.*, 2015). Indeed, genetic pathways of seed storability genes and biotic defense-related pathways controlled by common regulators (Righetti *et al.*, 2015). It is plausible that RBOH regulation participate in the balance between pathogen resistance and seed longevity.

The metabolism reduction by desiccation is the major strategy for seed time endurance in orthodox seeds. ROS derived from metabolism, as mitochondrial and chloroplast electron transport chains, are reduced. Mutant seeds of *psad1*, lacking a Photosystem I subunit that participates in the Photosystem I stability (Ihnatowicz *et al.*, 2004) demonstrates the deleterious effect of electron transport chains during seed storage. The instability of Photosystem I assembly leads to a rupture in the electron transport chain, an imbalance of photosynthesis and increased ROS production (Haldrup *et al.*, 2003; Pinnola and Bassi, 2018). It is plausible that light directly influences this reduced seed aging as mutant plants *psad1* present increased photosensitivity. Indeed, even in wild-type seeds with functional photosystems, photosynthesis is deleterious for seeds during storage. Chlorophyll disappears during seed maturation for seed storage to avoid light-induced ROS accumulation (Nakajima *et al.*, 2012). However, we cannot discard that no direct correlation may exist as is possible that the seed longevity

phenotype of *psad1* seeds is related to seed developmental problems as the mutant plant present strong pale and dwarf phenotypes. Seeds would have reduced seed longevity due to a diminished seed filling, as photosynthesis is necessary for a proper seed filling (Ruuska *et al.*, 2002; Goffman *et al.*, 2005). Both effects might directly contribute to the reduced seed longevity of *psad1*.

ROS detoxification systems are crucial to reduce ROS-induced damages. Macromolecules pools buffer ROS damages (Nguyen *et al.*, 2015; de Souza Vidigal *et al.*, 2016; Sano *et al.*, 2020) and antioxidants, such as vitamins, can reduce ROS to avoid their damaging effect in dry seeds (Kranner *et al.*, 2006; Sattler *et al.*, 2004). Pools of glutathione, with a high reduction power, are stored in seed, and the amount of reduced glutathione decays during seed storage. It constitutes a marker for seed viability and oxidation (Kranner *et al.*, 2006; Nagel *et al.*, 2015). Glutathione, as ascorbic acid, can be directly oxidized by ROS. However, the ascorbate-glutathione-NADPH system tightly controls and detoxify ROS (Foyer and Noctor, 2011). We demonstrate that the dehydroascorbate reductase DHAR1, but not DHAR2 and DHAR3, involved ascorbate regeneration by glutathione reduction, is crucial for embryo ROS detoxification and seed longevity. Presumably, DHAR1 will use the reducing power accumulated in glutathione pools to reduced actively ascorbate, as part of the ascorbate-glutathione-NADPH system.

Seed coat permeability determines seed longevity

The seed coat constitutes the first physical barrier between the external media and the seed embryo. It protects the seed embryo from detrimental physical factors such as water entry, oxygen penetration and light, through their different components: the cuticle layer, PAs accumulation, lignin-reinforced cell walls, and the suberin layer. Moreover, it also contributes to seed adaptation to the environment during development by the environmental cues perceived by the mother plant, and during seed storage, by embryo impermeabilization and ROS-oxidation (Jia *et al.*, 2012; De Giorgi *et al.*, 2015). In this PhD thesis, different TFs and other genes have been demonstrated to contribute to seed coat dependent seed longevity.

The seed coat imposes a physical barrier that restricts the entry to water and air to protect the embryo. The relative permeability can be assessed by the tetrazolium salt reduction assay (Debeaujon *et al.*, 2000). This assay needs to be performed in fresh seeds as it also depends on embryo viability for the tetrazolium salt reduction into formazan, a red compound easily visible. Different mutants presented enhanced tetrazolium permeability, coinciding with a reduced seed longevity. The mutants seed of *sstpr*, *spch*, *ice1*, double mutant *hb22 hb25*, the suberin mutants *gpat5* and *myb107*,

the cutin biosynthetic mutant *cyp86a8* and the peroxidases double and triple mutants *prx2 prx25* and *prx2 prx25 prx71* presented increased seed permeability and reduced seed longevity. On the other hand, the overexpression mutant *athb25-1D* and *cog1-2D* (tetrazolium assay not shown) and the knock out mutant line *knat7*, present reduced tetrazolium salt uptake and reduction and increased seed longevity. All these mutant lines show that seed coat impermeabilization contributes positively to seed longevity. Limitation of water and oxygen flow must limit the oxidative damage. Indeed, we demonstrated this with *athb25-1D* and double *athb22 athb25* mutants. The overexpression mutant *athb25-1D* was resistant to all four aging treatments, including the EPPO in which aging damages are related to oxygen penetration, and present a reduction of cellular oxidation after aging in accordance with the glutathione redox potential ratio comparted to wild type seeds (Figure 31b). The double mutant behaved oppositely.

Some mechanisms involving these mutants remained unclear, as the SSTPR molecular mechanism and the SPCH regulation. Not much is known about members of the TPR-PTR (tetratricopeptide-pentatricopeptide) family of repeat domain proteins (Sharma and Pandey, 2016). SSTPR is sublocalized in the mitochondria according to Hooper *et al.* (2017) where it may have an important role as other described PTR proteins (Lurin *et al.*, 2004).

The role of SPCH in seed coat permeability remains unclear. It is not abundantly expressed during seed development in the seed coat or in other seed tissues according to Lee *et al.* (2010) data. SPCH has been widely studied in stomatal cell differentiation and their mutant plants present no stomata (MacAlister *et al.*, 2007). ICE and SPCH interact in the Y2H system (Kanaoka *et al.*, 2008). In a speculative way, it might be important in the endosperm development or hormone control as ICE1, which is involved in stomata lineage determination, but also in endosperm breakdown, primary seed dormancy, ABA signaling, and freezing tolerance (Kanaoka *et al.*, 2008; MacGregor *et al.*, 2018; Hu *et al.*, 2019). ICE1 is a direct target of AtHB25 according to ChIP seq data (Figure 65, Supplemental data 5). The involvement of both, ICE1 and SPCH, must be further investigated to stablish the link with seed longevity. Stomata formation regulation is dependent of environmental conditions and ABA (Casson and Hetherington, 2010), and thus, ICE1 and/or SPCH might be similarly regulated in seeds, controlling seed longevity

KNAT7, curiously, is negative for seed longevity, as seeds of the mutant plant *knat7* present enhanced seed longevity. KNAT7 is described as a negative regulator of secondary cell wall lignin formation (Bhargava *et al.*, 2013; Liu *et al.*, 2014). Columella cells of *knat7* seeds present problems in mucilage extrusion (Romano *et al.*, 2012), which may be influenced by reinforced cell walls in columella cells. This effect might also explain their increased seed longevity and reduced seed coat permeability, through

cell wall reinforcement in other seed coat tissues. KNAT7 transcripts are accumulated in seed coat during seed maturation according to Lee *et al.* (2010) data (Figure 25). The expression of KNAT7 during seed development may limit or delay seed coat reinforcement, which might be produced earlier in mutant *knat7* seeds and thereby, limiting the seed growth, which results in a reduced seed size of *knat7* seeds (Figure 27b,c).

Two characterized seed coat TFs regulating seed longevity were not related with seed coat permeability, SEP3 and MYB47. While the abundance of the MYB47 TF was directly linked with an increased seed longevity, mutant seeds did not present an evident seed coat permeability change. The close homology to master regulators to lipid polyester deposition such as MYB107, MYB41 and MYB9 suggests a role in lipid biosynthesis, but changes in seed coat lipid polyester biosynthesis were not clearly visualized. *MYB47* expression in the chalaza strongly suggest a role in late maturation, which remains unclear. Perhaps it is involved in seed excision and chalazal sealing. Further study is needed to reveal the role of this interesting seed coat TF, highlighted through all three different artificial aging GWAS.

Intriguingly, mutant seeds of *sep3* mutant plants behaved oppositely as other seed coat permeability mutants: they presented an increased seed longevity but also a highly increased tetrazolium salt permeability. The contribution of *SEP3* to seed longevity is negative. SEP3 is a floral TFs that interacts with a wide range of TFs acting as scaffolding protein. Mutant *sep3* seeds presented a rounder shape as *stk* mutant seeds (Mizzotti *et al.*, 2014); and *sep3* seeds barely presented mucilage extrusion, suggesting a role in integument differentiation. This SEP3 role may involve TT16 and STK, two TFs, are involved in inner integument differentiation that interact with SEP3 (Favaro *et al.*, 2003; de Folter *et al.*, 2005; Kaufmann *et al.*, 2005; Coen *et al.*, 2017). The absence of STK-SEP3 interaction might explain the seed round shape, and the absence of TT16 interaction might explain the increased tetrazolium salt uptake, as in *tt16* mutant seed (Prasad *et al.*, 2010). In addition, outer integument differentiation is also affected, as seen with the reduced mucilage extrusion in *sep3* seeds. Amounts of SEP3 may regulate the seed coat properties in response to flowering and environment. A putative explanation of the surprisingly longevity-permeability miscorrelation is that the mutant *sep3* seed coat it is softer than wild type seed coat. During embryo development the growth might cause this round shape, and later upon seed rehydration, embryo expands and create fissures in the seed coat, explaining the high tetrazolium salt uptake. In the dry state this soft-seed coat would not be determinant, and even advantageous. Seed coat fissures in dry seeds were not visualized with the Scanning Electron Microscope (data not shown). Perhaps, molecular resources involving mucilage formation, which is diminished in *sep3* mutants, may be used for the biosynthesis of other molecular components during seed coat development, and this effect may help to increase seed longevity. Mucilage, as demonstrated in the Figure 34, is

not detrimental for seed longevity, and even the slight reduction of seed longevity of *35S::MUM2* seeds observed may be due to an increased mucilage production. The role of seed coat mucilage has been evolutionary maintained probably for its role in water retention in seed germination and its contribution for seed dispersal (Yang *et al.*, 2011), but molecular resources for its synthesis suppose a big part of total molecular resources necessary for seed development. Balance of seed resources must affect seed longevity.

Seed coat lipid polyesters and seed longevity

In this PhD thesis we found and demonstrate the direct effect of lipid polyesters deposited in the seed coat directly affect seed longevity through the seed coat impermeabilization. Traditionally, seed coat lipid polyester biosynthetic and regulation mutants have been characterized through the seed permeability assay using tetrazolium salts. Mutant seeds on the suberin biosynthetic mutants of *gpat5* (Beisson *et al.*, 2007), *abcg20* (Fedi *et al.*, 2017), *abcg2 abcg6, abcg2 abcg6 abcg20* (Yadav *et al.*, 2014), *far1 far4 far5* (Vishwanath *et al.*, 2013), and *eh1* (Pineau *et al.*, 2017); and the seed coat suberin regulator *myb107* (Gou *et al.*, 2017), presented increased tetrazolium salt uptake and reduction. In addition, the cutin biosynthetic mutants *lacs2* and the double *gpat4 gpat8* mutant present reduced seed longevity upon accelerated aging treatments (De Giorgi *et al.*, 2015). However, the complete link between seed longevity and seed coat permeability dependent of seed lipid polyester deposition has not been completely established. Different suberin mutants present reduced seed longevity (Figure 51). We demonstrate that the increased longevity of the overexpressing lines *athb25-1D* and *cog1-2D* is directly linked with lipid polyester deposition, through the regulation of diverse lipid biosynthetic mutants. The molecular characterization of the TFs AtHB25 and COG1 show a direct link between seed coat permeability and seed longevity involving the seed coat lipid polyester barriers, the suberin layer and the cuticle layer. The hydrophobic properties of these sealing layers protect the embryo from air and water entry, reducing oxygen penetration and associated damages, and thus positively contributing to seed longevity. In addition, the regulation of AtHB25 and COG1 points to environmental adaptation of seed longevity, through temperatures cues and light-dependent flowering cues, respectively.

Both lines, *athb25-1D* and *cog1-2D*, present increased GA and ABA content (Bueso *et al.*, 2016). Gibberellins are important in seed coat development (Kim *et al.*, 2005), and constitutive GA-signalling mutants present increased seed longevity (Bueso *et al.*, 2014*b*). On the other hand, ABA signalling is necessary for seed maturation (Leprince *et al.*, 2017). The exogenous application of GAs increases the expression of diverse peroxidases in siliques, including the *PRX2* and *PRX25* (Figure 54c), also overexpressed

in *cog1-2D* siliques (Figure 54b). In addition, *PRX25* was found overexpressed also in the mutant *athb25-1D* (Bueso *et al.*, 2014*b*). Both peroxidases, *PRX2* and *PRX25*, are essential for seed coat lipid polyester biosynthesis, as double mutant seeds present increased permeability and decreased longevity. Peroxidases may have a role in suberin polymerization HCAs, as well as in lignin polymerization (Lewis and Yamamoto, 1990; Bernards *et al.*, 2004) as seen in tomato (Quiroga *et al.*, 2000). Mutant seeds *prx2 prx25* presented a reduction of the percentage of acetyl bromide soluble lignin, either from lignin or suberin. In the stem, PRX2, PRX25 and PRX71 control the content of lignin (Shigeto *et al.*, 2015). Lignin, as suberin, might also affect water and gases permeability (Tobimatsu *et al.*, 2013; Liang *et al.*, 2006). However, other indicatives suggested that suberin may be reduced in these mutants. The seed coat palisade layer was thinner in prx*2 prx25 prx71* and *prx2 prx25* mutant seeds (Figure 56). In addition, ABA promotes suberin accumulation through TF regulation (Kosma *et al.*, 2014). Indeed, mutant *cog1-2D* seeds present increased seed coat lipid polyesters, with significative changes in HCAs, the lipid polyester aromatic compounds (Figure 72), which may be explained by the increased peroxidase activity. However, both biomolecules, suberin and lignin may be related to the peroxidase. As in Like *cog1-2D*, peroxidases should contribute to *athb25-1D* reinforced seed longevity.

In addition, COG1 determines other seed coat developmental processes. AP2 regulates COG1 expression, probably to determine part of the outer integument differentiation (Figure 71c). We found that COG1 overexpression partially recovers the undifferentiated seed coat outer integument of *ap2-12* seeds. The genetical cross *cog1-2D ap2-12*, present volcano-like structures in the epithelial seed coat layer, that are completely undifferentiated in the *ap2-12* mutant (Figure 76). The outer integument differentiation in necessary for seed suberin synthesis, completely absent in *ap2-12* seeds, but reduced in *ap2-5* seeds. The reduced suberin accumulation and the seed longevity of *ap2-5* seeds were recovered in the *cog1-2D ap2-5* double mutant seeds (Figure 74). The absence of suberin in the outer integument of *ap2-12* was not recovered by *COG1* overexpression. Other transcriptional players should act downstream of AP2, together with COG1, for outer integument regulation.

Downstream target analysis was possible for AtHB25, but no for COG1 due to the low protein amounts of viable mutants in the obtained *proUBQ10::COG1:3xHA-1* transgenic line (Figure 70). Four lipid polyester biosynthetic genes were found as AtHB25 direct targets, but only two, *LACS2* and *NHO1*, were overexpressed in *athb25-1D* dry seeds. We established dried seeds to measure the relative expression of the AtHB25 targets to avoid differences of expression related with seed developmental time points processes, controlled by the hormones ABA and GAs, as they may be altered in *athb25-1D* mutant. However, *KCS20* and *CYP86A8* might be overexpressed during *athb25-1D* seed development but not in dry seeds. Indeed, transcript levels of both

KCS20 and *CYP86A8* were reduced in dry seeds of the double mutant *athb22 athb25*. Mutant seeds of *lascs2* and *nho1* presented reduced seed longevity (Figure 40) and lipid polyester deposition (Figure 41). Crosses with *athb25-1D* reverted completely the *athb25-1D* increased seed longevity and lipid polyester increased deposition (Figure 40). In addition, genetic crosses of *athb25-1D* with mutant plants lacking the suberin biosynthetic gene *GPAT5* and the seed coat suberin regulator *MYB107*, both with reduced seed longevity and increased seed permeability (Figure 42 and Figure 43) reverted *athb25-1D* seed phenotypes and confirmed the involvement of lipid polyesters in the enhanced seed longevity of *athb25-1D.*

LACS2 is described to participate in the cutin monomer biosynthesis (Lü *et al.,* 2009; Zhao *et al.,* 2019). LACS enzymes catalyze the addition CoA to lipid polyester precursors, necessary for their polymerization. The *lacs2* reduced seed longevity was already described (De Giorgi *et al.*, 2015). In leaves DCAs are reduced in *lacs2* mutants (Bessire *et al.*, 2007). We found that 18:2 DCAs levels, together with ferulate levels, are a reduced in *lacs2-1* seeds. Seeds of *athb25-1D* and *athb22 athb25* presented more and less lipid polyester monomers, respectively (Figure 36) and 18:2 DCAs was the monomer presenting higher differences in abundance among these mutants. Thus, this effect can be explained by the positive direct regulation of *LACS2* by AtHB25.

On the other hand, analysis of the CYP86 gene family, involved in the ω-hydroxylation of fatty acids, to form OHFA and DCAs (Bak *et al.*, 2011), revealed that CYP86A8 is the member with a higher impact in lipid polyester biosynthesis necessary for seed longevity. Although other members such as, CYP86B1 and CYP86A2, with higher expression in the seed coat (Figure 23), are described to participate in the seed coat suberin biosynthesis (Molina *et al.*, 2008; Compagnon *et al.*, 2009), mutant seeds of *cyp86a8* presented a drastic reduced seed longevity accompanied by a higher tetrazolium salt uptake and reduced lipid polyester deposition (Figure 24). Curiously, CYP86B1 and CYP86A2, were highlighted in the GWAS analysis. However, their mutant lines did not present remarkable seed longevity reduction as *cyp86a8*. It is possible that CYP86B1 and CYP86A2 act redundantly in seed coat suberin deposition as described, and thus we do not find a seed longevity phenotype, whereas CYP86A8 acts in cutin monomer biosynthesis with no gene redundancy, and the mutant *cyp86a8* present defects in seed coat cuticle. Indeed, mutant plants *cyp86a8* present organ fusions and pleiotropic defects (Wellesen *et al.*, 2001) due to defect in the leaf cuticle. Thereby, seed longevity defect in *cyp86a8* mutant seeds might be also related with these pleotropic effects. The study of its expression in the seed coat during seed development, in the endothelium or in the palisade layer, will clarify its importance in suberin, cuticle or both layers. Cutin and suberin biosynthesis share a common biosynthetic process (Philippe *et al.*, 2020), and same players might be involved in both polymer composition. Main differences rely in the polymerization place and structure.

Although we found that AtHB25 binds to a close genomic region of *CYP86A8*, we could not detect an upregulation of this cytochrome in developing seeds of *athb25-1D*. Nevertheless, we found a downregulation in *athb22 athb25* developing seeds (Figure 39). DCAs synthesis are also dependent of a CYP86 member, which might be the CYP86A8, together with LACS2, which regulate the DCAs levels in the *AtHB25* mutants.

Glycerol is a principal component of suberin and cutin and participates in lipid polyester polymerization (Pollard *et al.*, 2008; Philippe *et al.*, 2020). However, glycerol-3-phospate is the molecule directly added to suberin and cutin monomers by GPAT enzymes (Murata and Tasaka, 1997; Wendel *et al.*, 2009; Xue *et al.*, 2019). This reaction is essential for the impermeabilization role in suberin and cuticle. Mutant *gpat5* seeds, present a decreased seed coat suberin deposition, an increased seed coat impermeabilization and reduced seed longevity (Yang *et al.*, 2010; Yang *et al.*, 2012; Figure 44). Double mutant *gpat4 gpat8* seeds present deficient seed longevity caused by a defective cuticle layer (De Giorgi *et al.*, 2015). AtHB25 directly regulates the expression of GLI, also named NHO1, which synthetizes a glycerol kinase. Glycerol kinase phosphorylates glycerol to produce G3P. In addition, the G3P, which is also essential for glycolysis and glycolipid biosynthesis, can be obtained from dihydroxyacetone phosphate through the G3P dehydrogenase (Driver *et al.*, 2017). Both metabolic G3P biosynthesis pathways are important to biotic stress resistance (Li *et al.*, 2005; Venugopal *et al.*, 2009; Maeda *et al.*, 2010). Indeed, systemic acquired resistance requires a correctly formed cuticle layer (Xia *et al.*, 2009). In seeds, the glycerol phosphorylation through NHO1 is not so essential for seed longevity, as mutant seeds do not present an affected seed longevity and an evident reduced lipid polyester deposition (Figure 40 and Figure 41). However, the double mutant *athb25-1D nho1* reverts the increased seed longevity, the reduced seed coat permeability, and the increased lipid polyester barrier deposition phenotypes of the *athb25-1D* mutant. (Figure 40, Figure 44 and Figure 45). There is only this glycerol kinase in Arabidopsis, and it performs a rate-limiting step in glycerol metabolism (Venugopal *et al.*, 2009). However, glycerol can be also obtained from glycolysis. Although normal lipid polyester synthesis might be using G3P from glycolysis, it is plausible that NHO1 regulates the lipid polyester deposition increasing the amounts of G3P. This regulation might be part of the environmental adaptation of seed longevity and seed coat lipid polyester barriers, which might be controlled by AtHB25. Indeed, the increased lipid polyester deposition in *athb25-1D* seed coat is dependent of NHO1, and AtHB25 is regulated by cold environmental cues. Glycerol is an important cryoprotectant, and NHO1 regulation by *AtHB25* might contribute to cold resistant. Levels of Glycerol 3-phosphate need to be assayed in AtHB25 mutants in different temperatures to further investigate this. Other lipid polyester biosynthetic enzymes are regulated by cold, as LACS2 (also regulated by AtHB25) and cold cues regulate lipid polyester barriers composition and deposition (Fedi *et al.*, 2017). Presumably LACS2, and other lipid polyester biosynthetic enzymes may be regulated by

AtHB25 in response to low temperature determining suberin and cutin composition.

We found also that other lipid polyester biosynthetic genes were regulated by AtHB25 (Figure 39). FAR genes were regulated by AtHB25. Interestingly, while *FAR1* was negatively regulated by AtHB25, FAR4 and FAR5 were upregulated in *athb25-1D* developing seeds. FAR1 is described to regulate mainly the C22:0-OH, FAR4 the C20:0-OH and FAR5 the C18:0-OH fatty alcohols (Vishwanath *et al.*, 2013). Accordingly, levels of C20:0-OH were augmented in seed coat lipid polyesters of *athb25-1D.* Controversially, levels of C18:0-OH were reduced, and they do not correlate with *FAR5* regulation by AtHB25. Levels of C22:0 were similar in the AtHB25 mutants, but the C22:0 diol were increased in the *athb25-1D* mutant (Figure 36a), in discordance to the FAR1 expression, reduced in this mutant. However, we understand that the different FAR enzymes can catalyze different reactions, and they also might act during other developing phases of the seed coat development, and not only in the maturation phase analyzed by qRT-PCR in the AtHB25 mutants.

The positive *FACT* regulation (Figure 39) by AtHB25 explains part of the increased HCA deposition (Figure 36b). The positive regulation of ABCG11 and GPAT8 might explain the increased lipid polyester deposition through an increased monomer trafficking and increased monomer glycerol esterification, respectively. It is also possible that they can also affect to the different monomer relative abundance.

The KCS20 regulation was slightly detected in developing seeds but, very long chain DCAs and PA, but not OHFA were reduced in seed of the double mutant *athb22 athb25* and augmented in *athb25-1D* seeds. This points to a slight effect of KCS20 in AtHB25 monomer composition of the seed coat lipid polyester. Mutants seeds of *kcs20* do not show a reduced seed longevity (data not shown). Perhaps KCS20 is redundant to KCS2, as both are described to elongate suberin monomers in the seed coat (Franke *et al.*, 2009; Lee *et al.*, 2009*b*), and only the double mutant presents a drastic effect on seed longevity.

Finally, during the gene expression analysis in developing seeds we notice that seed developing time was increased in the mutant plant *athb25-1D* and reduced in the double mutant *athb22 athb25* in comparison to wild type plants. This effect can also affect to the gene expression of the lipid polyester biosynthetic genes, which might be regulated for longer and shorter times respectively, even if there are not differences in their abundance in the maturation stage, the time point when the developing seeds were analyzed. We tried to analyze other time points, but they were difficult to establish in the different lines due to the differences in their developing times. Other lipid polyester biosynthetic enzymes might be important for the seed longevity regulation dependent of AtHB25 affecting to other seed developmental stages. For example, we found that PRX25 was upregulated in developing seeds of *athb25-1D* at earlier stages

(data not shown). This regulation is putatively related with the increased GAs levels in the *athb25-1D* mutant, such as the *cog1-2D* mutant. Indeed, the GAs levels in seeds are important for the seed coat development prior to maturation stages (Kim *et al.*, 2005). It is plausible that also the increased developing time of seeds is due to the GAs increased levels.

The suberin lamellae in the seed coat

Previously, only Gou *et al.* (2017) showed the suberin lamellae in the primary cell wall of the palisade layer. TEM Analysis of *AtHB25* mutants were performed with a high amplification (Figure 34) to visualize the suberin lamellae pattern in seed coats. Contrarily to Gou *et al.* (2017), but accordingly to Molina *et al.* (2009) root suberin visualization, we found the suberin lamellae within the primary cell wall and the cytoplasm. Nevertheless, suberin should also affect the primary cell wall structure as it is not stained by PAs, indicating that it is not permeable, differently to endothelium cell wall, which is stained by PAs (Demonsais *et al.*, 2020). In the Figure 34e, the seed coat suberin lamellae is clearly visible, moreover in the *athb25-1D* mutant seed, which present an increased seed suberin deposition. With this image, we show at first instance the suberin lamellae in Arabidopsis seed coat palisade cells. Probably it was not correctly described early due to the difficultness for the lamellae pattern visualization. Not all palisade cell-wall areas present a visible suberin lamellae pattern, as it depends on the histological cut, among others. In addition, as cells of the seed coat collapse, the suberin lamellae is not as regular as root suberin lamellae. We found that intersection areas within different palisade cells presented better visible suberin lamellae patterns. Lateral cell walls of palisade cells are subjected to a higher pressure due to palisade cell collapse. Cell walls folds and create an impermeable structure (Figure 34c). Probably this cell compression permits that cell structures are better maintained in these areas after the ultrathin cut due to a higher pressure, making easier the suberin lamellae visualization.

Regulation of *AtHB25* and *COG1*

The endogenous roles of AtHB25 and COG1 remained unclear with the overexpression mutants, *athb25-1D* and *cog1-2D*, analysis. We found expression of both TFs during the seed coat development. AtHB25 expression was situated in the maturation seed coat, mainly in the inner integument, but also less abundantly in the palisade layer (Figure 31c and d and Figure 33). In contrast, COG1 was found abundantly expressed in the outer integument of the seed coat, during early stages of seed coat differentiation (Figure 54a).

Direct regulators of *AtHB25* were screen through a Y1H using a 71bp conserved *AtHB25* promoter fragment. Four TF were found to interact with this DNA fragment: HB52, bZIP52. AGL24 and DREB2H (Figure 62a). However, only *DREB2H* was found abundantly expressed during the seed coat development. The DREB2 TF is subfamily of the ERF/AP2 TF family that responds to diverse abiotic stress such as drought, salt and high temperatures, binding to dehydration responsive elements (Lata and Prasad, 2011). Seed coat expression of AtHB25 coincides with seed maturation and seed coat drying. Thus, DREB2H might be the responsible for this AtHB25 expression maybe through desiccation cues, as DREB2 TFs respond to abiotic stresses such as drought (Nakashima *et al.*, 2000; Agarwal *et al.*, 2017). In addition, *DREB2H* was induced by chilling in plants (Figure 63b), suggesting that AtHB25 might be also induced by cold. We found that the AtHB25 promoter driven GUS expression was upregulated in leaves by chilling, and that overexpression *athb25-1D* plants presented increased cold resistance while the double mutant *athb22 athb25* were susceptible in low temperature growth (Figure 64). Hereby, AtHB25 participates in the cold adaptation of plants, probably regulated by *DREB2H*. In seeds, it is plausible that DREBH regulates AtHB25 in response to temperatures cues for developmental seed longevity regulation.

The role of AtHB25 in cold resistance probably goes, in part, through ICE1 regulation. ICE1 is a well-studied TF involved in stomatal cell differentiation, endosperm breakdown and freezing tolerance (Kanaoka *et al.*, 2008; MacGregor *et al.*, 2019) and it is a direct AtHB25 target found by ChIP-seq analysis. In developing seeds *athb25-1D* presented higher *ICE1* transcript levels, in comparison to wild type plants, while the double mutant *athb22 athb25* presented a reduced ICE1 up regulation by chilling (Figure 65). ICE1 controls the induction of 40% of cold-response genes, including CBFs TFs, whose play a central role in signal transduction in response to low temperature (Lee *et al.*, 2005; Chinnusamy *et al.*, 2007; Zhou *et al.*, 2011). Thereby, regulation of *ICE1* must be important for cold resistance. In addition, ICE1 regulates ABA signaling through ABI5 interaction (Hu *et al.*, 2019). As ABI5 is involved in seed longevity, it is plausible that part of the seed longevity of the *athb25-1D* is due to *ICE1* regulation. Seed longevity of *ice1* mutant plants was reduced, but lipid polyester levels not. They present a highly seed coat permeability due to seed coat fissures produced by the high internal pressure resulting from endosperm stiffness and embryo growth constriction, as in *zou* mutants (Figure 67; Figure 68; Denay *et al.*, 2014; Creff *et al.*, 2015; Fourquin *et al.*, 2016). Thereby, direct air entry rapidly oxidises and damages *ice1* embryos. Other effects of ICE1 in seed longevity remain uncovered by the seed coat fissures. These effects might be related with hormone levels, as ICE1 participate in hormone regulation and signalling of ABA and GAs, and the overexpression *athb25-1D* mutant present higher levels of both hormones, ABA and GAs (Bueso *et al.*, 2016). For instance, plants of *athb22 athb25* are similar to *ice1* plants, and this may imply that

ICE1 deregulation affect similarly the double mutant *athb22 athb25*. The different hormone levels, putatively produced by ICE1, can contribute to part of the enhance seed longevity during acquired during seed development in the *athb25-1D* mutant. Thus, it is plausible that the increased seed longevity of *athb25-1D* seeds is not only controlled by the increased levels of biosynthetic genes such as LACS2 and NHO1, that promote lipid polyester deposition in the seed coat, but also by other processes regulated by hormones, that might affect to the seed development time among others. Indeed, constitutive GAs signaling of DELLA mutants leads to increased seed longevity, such as external GAs treatment during seed development (Bueso *et al.*, 2014*b*).

Cold regulation of seed longevity during seed development can be the endogenous role of AtHB25. In plants, low temperatures increase ABA (Kendall *et al.*, 2011), and ABA promotes suberinization (Soliday *et al.*, 1978, Cottle and Kolattukudy, 1982) by TFs regulation such as MYB41 (Kosma *et al.*, 2014). In seeds the close homolog, MYB107 controls suberin deposition, and it might also respond to ABA as it is expressed in maturing seed coats (Gou *et al.*, 2017). Lipid polyester composition of the seed is modified by low temperatures, by the regulation of suberin biosynthetic enzymes such as LACS2 (Fedi *et al.*, 2017), which is regulated by AtHB25. Taken together, it is possible that AtHB25 is upregulated by low temperatures, also in the developing seed coat, by DREB2H, to control lipid polyester biosynthetic genes and other cold response genes regulated by ICE1 and hormone levels, determining seed longevity and other seed traits for seed adaptation to environmental cues. The importance of the AtHB25 role, probably in cold response in plants and seeds, is remarked through it highly conserved in angiosperm plants, such as tomato and wheat, and the conserved regulatory regions in close species. In addition, the ectopic expression AtHB25 also produces increased lipid polyester deposition in the seed coat of wheat and tomato cultivars. Thereby, the regulatory mechanism of AtHB25 is conserved among angiosperm and must constitute an important adaptive mechanism as it has been maintained during the evolution of the different plant species. AtHB25 might play a central role in cold adaptation of plant and seeds.

On the other hand, COG1 was found to be regulated by the floral TF AP2. The role of AP2 in the regulation of the outer integument differentiation might explain the seed coat *COG1* downregulation in *ap2* mutants. However, this *COG1* downregulation was found also in *ap2* seedlings demonstrating a direct regulation of *COG1* by AP2. We found that COG1 was able to recover partially phenotypes of *ap2* seeds. The reduced suberin accumulation and reduced seed longevity of the *ap2-5* mutant was recovered in the *cog1-2D ap2-5* double mutant. The absent columella cell differentiation in the *ap2-12* mutant was partially recovered by COG1. Thereby, COG1 is a regulator of outer integument differentiation, acting downstream of AP2. No interaction was found with other TFs participating in outer integument differentiation. This fact suggests that

COG1 do no interact with then and probably they act at a different level of regulation, although it is also possible that our Split-Tpr assay was no robust enough, or multimeric complexes are needed. However, the previously described AtHB25-AtHB25 positive interaction, corroborated the adequate assay conditions. The involvement of COG1 in light perception regulation (Park *et al.*, 2005) which determined an increased seed longevity through seed coat impermeabilization by an increased lip polyester deposition (Bueso *et al.*, 2016) might reveal a role of COG1 in light-perception in seed longevity adaptation through the regulation of outer integument differentiation.

Seed longevity is an adaptive trait

The different Arabidopsis ecotypes used in the seed longevity assays showed a wide range of variation in the resistance for each aging treatment. This high variance was advantageous for the seed longevity GWAS study. The reason of this variance relies in the adaptation of this world-wide distributed plant in different niches and climates. The natural selection of Arabidopsis in each climate diversified its genomic information. Seed longevity is an important trait for plant adaptation, as it is involved in plant reproduction aiming to ensure plant reproduction even under unfavorable environments. However, other traits such as dormancy, seed yield or seed propagation are essential in seed adaptation. Sometimes seed traits are inversely correlated, such as seed dormancy and seed longevity found in diverse ecotypes and mutant lines (Nguyen *et al.*, 2012). Not in all cases seed longevity is the most determining trait in plant survival within different generations. This effect can explain part of the variability found among all tested ecotypes. We found that an early flowering was correlated with a higher seed longevity (Table 1, Supplemental data 2), which has not been previously described. Nevertheless, this correlation is, in some way, expectable. The seed developmental program initiates with flowering, and flowering is regulated by diverse environmental cues, such as temperature, light, photoperiod, nutrient availability perceived by the mother and the embryo, and these perceived cues also determine seed longevity (He *et al.*, 2014). In our study we stablish same conditions for Arabidopsis plant growth and germination, which are optimal for the Col-0 ecotype. However, other ecotypes might have other optimal plant growing conditions, and the environmental cues perceived, such as light and temperature, might affect differently to each ecotype. Indeed, form the 360 ecotypes initially used only 287 flowered in within 180 days, because the growing conditions were not adequate to that ecotypes. In addition, diverse ecotypes included in the study presented an important flowering delay and extended life span and produced low amounts of seeds barely viable and with extremely reduced seed longevity. This effect explains part of this correlation. Nevertheless, the correlation between rapid flowering and enhanced seed longevity must be further

studied. He *et al.* (2014) study was demonstrated that warm temperature and sufficient light conditions were positive for seed longevity. These conditions also produce early flowering, confirming this correlation. The mutant *athb25-1D* present early flowering and increased seed longevity, but the mutant *cog1-2D* present a retarded flowering and increased seed longevity.

Nutrient availability is necessary for a proper seed filling and seed longevity acquisition. However, we found a negative correlation between seed longevity and iron content in Arabidopsis ecotypes. The iron content was measured by Atwell *et al.* (2010) in leaves; however, we expect that seed iron content correlates with leave iron content. Iron is used by electron transport chains participating in respiration and photosynthesis, as iron can accept electrons acceptor by redox reactions. A higher abundance of iron might lead to increased electron acceptance and ROS accumulation, derived from electron transport chains. This effect would explain the positive effect of low iron content and enhanced seed longevity. This correlation has been previously suggested by Murgia *et al.* (2015).

Nguyen *et al.* (2012) found a negative correlation between seed dormancy and seed longevity in a reduced group of ecotypes. This inverse correlation was found also in adaptative seed performance in response to environmental cues (He *et al.*, 2014). Embryo and PAs oxidation affect directly to seed longevity and seed dormancy. Lipid polyester barriers of the seed coat isolate the seed from the environment, retarding seed aging and the seed dormancy release (Fedi *et al.*, 2017). We do not find a correlation among ecotypes in our seed aging assays and public dormancy data. In addition, we neither found correlation of seed longevity with the latitude at collection point of the different ecotypes. Latitude determines the annual temperature, the hours of light and the ratio red/far red light, environmental cues determining adaptive seed longevity (He *et al.*, 2014). In our experiments, plants were grown in similar conditions. Hereby, the seed longevity plasticity determined by environmental was removed (He *et al.*, 2014). We can conclude that the genomic adaptation of Arabidopsis was not influenced by the latitude of the niche or the dormancy.

Contribution of this thesis to global aging theories

The seed coat, a plant specific structure, represent a major part of molecular mechanisms regulating seed longevity characterized in this PhD thesis. However, we also found general aspects involving aging that could affect human aging, mainly related with ROS damage. The different results of this thesis remark the importance of ROS-induced aging. Thereby, damaging theories of aging, which postulate that aging is a consequence of ROS produced by imperfect metabolisms reactions, seem to affect

mainly seed aging. Seeds dry to reduce the metabolism and higher temperatures and humidity negatively affect seed longevity due to a partial reactivation of the metabolisms and increased oxidative damage.

The three RBOH genes that negatively affect seed longevity encode homologous genes to NOX enzymes that are NADPH oxidases (Torres *et al.*, 2002). NADPH oxidases are a ROS source together with electron transport chains and the 5'-lypoxigenase (Novo and Parola, 2008). NOX enzymes, as RBOHs in plants, produce ROS for fine-tuning cellular signalling to maintain normal physiology and ROS signalling misfunction leads to different pathologies (Krause, 2007). Aging is correlated with chronic activation of these NAPH oxidases which would lead to an increased ROS damage (Ewald, 2018). Our investigation in seed longevity supports the damaging role of NADPH oxidases in aging and confirms the aging theories based in ROS damage. Perhaps, drug therapies targeting to specific NOX isoforms might help to reduced diverse tissue aging to prolong human life (Ewald, 2018).

ROS are not only produced for signalling, but also because of metabolism imperfections. The disruption of electron transport chain is directly provoking accelerated aging, as corroborated with the mutant *psad1*. The *dhar1* mutant demonstrated again the protective role of ROS detoxification systems in longevity, described in plants and animals (Johnson, 1990; Orr and Sohal, 1994; Lithgow *et al.*, 1995). In addition, we demonstrate the lipid polyester deposition in the seed coat protects the embryo through the impermeabilization to water an air, avoiding oxidative damages. The direct effect of oxygen was observed in *C. elegans* (Honda *et al.,* 1993). Reducing atmospheric oxygen and other oxidant molecules might be beneficial to extend human life.

The negative correlation between seed aging and iron might reveal an interesting approach to extend life. Iron excess is a potent source of oxidative damage, and dysregulated metabolism of iron is associated with oxidative and cellular stress involved in diverse neurodegenerative disorders related to aging (Belaidi and Bush, 2016). In addition, aging is associated with anemia which causes an increased risk of mortality (Balducci, 2010). It is plausible that anemia caused by age is a consequence of a molecular mechanism that aims to protect the individual from iron associated damages. Caloric restriction, without causing malnutrition, extends animal life in diverse animal organisms, including humans (Fontana and Partridge, 2015), through the reduction of mitochondrial ROS production (Pamplona and Barja, 2006). It is plausible that a big part of this ROS reduction regards to a reduction of iron uptake. Probably similar extending-life effects can be obtained through iron restriction in diet. Perhaps, food content with lower iron content, such as vegetables, but not red meat, might help to extend human life due to the reduction of iron provoked oxidative events. Red meat consumption is associated with cancer (Kouvari *et al.*, 2016; Carr *et al.*, 2017), and iron produced ROS damage might be responsible of this association. However, in old-

er people might be beneficial to compensate anemia age-induced effects.

Nevertheless, seeds also present programed induced aging. The TFs SEP3 and KNAT7 are negative for seed longevity. Negative traits often are lost during natural selection and plant adaptation. SEP3 and KNAT7 have important roles in flowering and lignification, respectively, and these regulations might be responsible of their maintenance. However, its expression in during seed development are deleterious, but curiously maintained, such as the RBOHs. Seed longevity is an adaptative trait, and other traits such as seed dormancy are important in seed performance and environment adaptation. Thereby, seed longevity is adjusted in a programmed way.

4. Conclusions

1 - The variation on seed longevity found among Arabidopsis ecotypes permitted to identify and validate 12 new seed longevity genes regarding to the desiccation tolerance acquisition (*SSLEA*), ROS damage (*RBOHD*, *RBOHE*, *RBOHF*, PSAD1 and *DHAR1*) and seed coat development (*CYP86A8*, *SSTPR* and the TFs *SEP3*, *SCPH*, *KNAT7* and *MYB47*, which was expressed in the chalaza, but it did not show an observable seed coat phenotype). However, their molecular mechanisms regulating seed longevity remain unclear.

2 - AtHB25 and COG1 regulate seed coat lipid polyester deposition that explains the phenotypes of enhanced seed longevity and reduced seed coat permeability of *athb25-1D* and *cog1-2D* mutants. The restriction to water, air and oxygen must reduce the seed oxidative damage, and thus aging. By contrast, the mucilage does not apparently participate in seed longevity.

3 - It has been properly visualized the suberin lamellae in Arabidopsis seed coat. We found that it is better visualized in folded cell wall areas within two palisade cell layers. The suberin lamellae is thicker in the *athb25-1D* mutant, and thinner in the *athb22 athb25* double mutant. The cuticle layer presented was increased in *athb25-1D* mutant seed coats.

4 - The ectopic AtHB25 expression produce similar effects in tomato and wheat, increasing the cuticle layer deposition and seed longevity. The conserved function and genetic sequence of AtHB25 in angiosperm remarks the importance of AtHB25 regulation in Angiosperms.

5 - AtHB25 directly regulates the lipid polyester biosynthetic genes. *NHO1* and *LACS2* are key enzymes in the *athb25-1D* increased seed longevity. NHO1 has been proposed to regulate lipid polyester deposition by increasing G3P levels.

6 - *AtHB25* is up regulated by cold and it is involved in cold resistance. *ICE1* regulation by AtHB25 must explain part of the cold resistance regulation dependent of AtHB25 and the different hormone levels present in the *athb25-1D* mutant.

7 - The reduced seed longevity observed in *ice1* mutant seeds is produced by seed coat fissures produced by the high internal pressure during seed development. This effect supposes a new molecular mechanism affecting seed longevity. A high internal pressure is damaging for seed coat integrity and seed longevity.

8 - COG1 regulates lipid polyester deposition affecting specifically in the outer integument regulation. The mutant *cog1-2D* present higher GA and ABA levels, inducing the expression of peroxidases *PRX2* and *PRX25*. The role of PRX2 and PRX25, controlling suberin and/or lignin polymerization, is crucial in seed coat sealing.

9 - The expression of COG1 in the outer integument is regulated by AP2 expression controlling the suberin deposition and columella cell differentiation.

10 - Seed longevity is an adaptative trait. Plants adapt seed longevity determinants to ensure their survival. Light and temperatures are environmental cues for the adaptation of seed longevity, and AtHB25 and COG1 participates, respectively, in this regulation.

5. Materials and Methods

Plant material

The Arabidopsis ecotype Col-0 was used as the wild-type reference for *Arabidopsis thaliana* assays. We used the *Solanum lycopersicum* cv Moneymaker (*Tm2*^{2}) for tomato plant assays. For wheat studies, we utilized *Triticum aestivum* cv BW208 and cv THA53 cultivars from the germplasm collection of the Institute for Sustainable Agriculture (CSIC, Cordoba, Spain).

To assess the Arabidopsis natural variation in seed longevity and perform GWAS a core set of 360 Arabidopsis thaliana ecotypes was obtained from the Nottingham Arabidopsis Stock Centre (NASC) (reference N76309). These accessions were genotyped with a 250k-SNP/high density tiling array technology by the Justin Borevitz laboratory (University of Chicago, USA). Recently, full sequence from thousands of Arabidopsis ecotypes data was published (1001 Genomes Consortium, 2016). However, just a fraction (about 1/3) of this Arabidopsis-ecotype core set was full-sequenced.

Homozygous T-DNA and transposon insertion Arabidopsis lines used in the GWAS study are: *psad1* (SALK_036224), *sslea* (SAIL_681_A07), *sstpr* (SAIL_208_C05), *rbohd* (SALK_109396), *rbohe-1* (SALK_146126, Xie *et al.*, 2014), *rbohf* (SAIL_1242_C07), *rbohd,f* (double mutant from Torres *et al.* 2002), *dhar1* (SALK_005238, Rahantaniaina *et al.,* 2017), *dhar2* (SALK_026089, Rahantaniaina *et al.,* 2017), *dhar3* (SAIL_435_A09, Rahantaniaina *et al.,* 2017), *dhar1,2,3* (triple mutant *dhar1 dhar2 dhar3,* Rahantaniaina *et al.,* 2017), *cyp86a1* (GK-055C08), *cyp86a2* (SALK_005826), *cyp86a8* (GK-719C03), *cyp86b1* (SALK_203846), *cyp86b2* (SALK_070150), *cyp86c1* (SALK_050565), *myb47* (SALK_200360), *spch-3* (SAIL_36_B06, MacAlister *et al.*, 2007), *knat7-1* (SALK_002098, Romano *et al.*, 2012), *sep3* (transposon insertion line *sep3-2* from Pelaz *et al.*, 2000), and *sep1,2,4* (triple mutant transposon insertion line from Ditta *et al.*, 2004).

Homozygous Arabidopsis Col-0 background mutant lines used in the AtHB25 study are: *athb25-1D* (activation tagging mutant overexpressing *AtHB25,* Bueso *et al.*, 2014*b*), double mutant *athb22 athb25* (SALK_017963 SALK_014023, Bueso *et al.*, 2014*b*), *lacs2-3* (GABI_368C02, Bessire *et al.*, 2007), *lacs2-1* (Schnurr *et al.*, 2004), *nho1* (SALK_067205), *kcs20* (SALK_019727, Lee *et al.*, 2009*b*), *gpat5-2* (SALK_142456, Beisson *et al.*, 2007), *myb107-1* (SAIL_242_B04), *mum2-1* (Western *et al.*, 2001) and *35S::MUM2* (Dean *et al.,* 2007). Double homozygous F3 seeds of *athb25-1D* crosses with *lacs2-3*, *nho1*, *gpat5* and *myb107* were obtained.

Homozygous Arabidopsis Col-0 background mutant lines used in the COG1 study are: *cog1-2D* (activation tagging mutant overexpressing *COG1,* Bueso *et al.*, 2016),

prx2 (SAIL_355_A01), *prx3* (SALK_140204), *prx12* (SALK_043965), *prx22* (SALK_144487), *prx23* (SALK_061249), *prx25* (GK-716G12), *prx36* (SAIL_194_G03), *prx43* (SAIL_849_H03), *prx50* (SALK_063662), *prx55* (SALK_102284), *prx64* (SALK_203548), *prx71* (SALK_123643), *prx2 prx25* (double mutant from Shigheto *et al.*, 2015), *prx2 prx25 prx71* (triple mutant from Shigheto *et al.*, 2015), and the *abcg2 abcg6 abcg20* (triple mutant from Yadav *et al.*, 2014). Mutant seeds form *gpat5-2* and *myb107-1* used in the AtHB25 study were also used in this study.

Homozygous Arabidopsis Col-0 background mutant lines used in the AtHB25 and COG1 regulation study are: *ice1-2* (SALK_003155, Kanaoka *et al.*, 2008), *ap2-5* (Ohto *et al.*, 2005) and *ap2-12* (SALK_071140, Yant *et al.*, 2010).

Plant growth conditions

Mutant and Col-0 Arabidopsis lines and wheat plants were grown under growth chamber conditions (16 h light/8 h dark, at 23 °C and 70-75% relative humidity). Arabidopsis ecotypes and tomato plants were grown similarly but at 18 ºC at night. Plants were grown in pots containing a 1:2 vermiculite:soil mixture. Plants were watered with nutrient solution (5 mM KNO_3, 2.5 mM K_3PO_4, 2 mM $MgSO_4$, 2 mM $Ca(NO_3)_2$, 0.036% EDTA–$FeSO_4$, 70 µM H_3BO_3, 14 µM $MnCl_2$, 0.5 µM $CuSO_4$, 1 µM $ZnSO_4$, 0.2 µM Na_2MoO_4, 10 µM NaCl, 0.01 µM $CoCl_2$). Seeds lots were simultaneously grown to avoid growth-dependent phenotypes. Different seed lots grown independently were used to confirm the different described mutant phenotypes. Seven different Arabidopsis plants of each ecotype were grown, and seed pooled. Arabidopsis plants treated with GAs were sprayed twice per week with a 50 µM GA3 solution.

For *in vitro* culture, Arabidopsis and tomato seeds were surface sterilized, stratified for 3 days (7 days for ecotypes) at 4 °C and sown in Murashige and Skoog salts (MS) agar plates. Plates were incubated in a controlled environment growth room at 23 °C under long-day conditions (16 h light/8 h dark) at 80 µmol m^{-2} s^{-1}.

Seed sterilization

Arabidopsis and tomato seeds were surface sterilized by soaking in 70% ethanol 0.01% Triton X-100 solution for 10 min, followed by incubation in 50% commercial bleach (2.5% NaClO) 0.05% Triton X-100 solution for 5 min. Seeds were washed three times by rinsing with sterile water.

Ecotype seed sterilization consisted on eight hours chlorine gas treatment. Chlorine gas was obtained through the addition of 3 mL of HCl to 100mL of commercial bleach

(2.5% NaClO). Seeds were enclosed in a chamber and later, ventilated in a laminar flow cabin. This seed-sterilization method permitted the simultaneous and equal sterilization of the hundreds of ecotypes. However, a higher contamination ratio was obtained.

The last sterilization method used during this PhD thesis is suitable for transgenic selection after Arabidopsis transformation, by floral deep, as it permits the sterilization of big amounts of seeds. Consists on the soaking in 70% ethanol 0.05% Triton X-100 solution for 10 min, followed by 70% ethanol wash. Later, seeds are disposed in sterilized pieces of paper to dry in laminar flow cabin. Doing this, seeds are dry sterilized and can be sown easily.

MS media preparation

Solid MS media for plant *in vitro* culture was prepared with 4.3 g/L Murashige and Skoog medium (Duchefa), 1 g of 2-(N-morpholino)ethanesulfonic acid (MES) monohydrated (Duchefa), Tris base to adjust pH to 5.7 with, 10 g of sucrose and 8 g Phyto agar (Duchefa). Sterilized media was then distributed in 100 mm petri dishes or plates.

Antibiotics, if needed, were added in warm media (~50ºC) before solidification. For easy management they are prepared in freeze 1000X stocks. The stocks are the following: Hygromycin (50 mg/ mL), Glufosinate-ammonium (BASTA) (20 mg/mL), Kanamycin (50 mg/ mL) and Sulfadiazine (25 mg/mL).

Seed aging assays

In this thesis, we use three different artificial aging methods to speed up the seed deterioration: the accelerated aging treatment (AAT), the controlled deterioration treatment (CDT) and the elevated partial pressure of oxygen treatment (EPPO). In addition, seed dry storage at ambient temperature or natural seed aging treatment (NAT) was performed when possible. Seed bank conditions were also use to long-term seed storage and mimic tomato seed aging in optimal conditions.

The quickest aging treatment is the AAT and consists in seed hot-water imbibition for one or two days (Prieto-Dapena *et al.*, 2006; Bueso *et al.*, 2014*b*). In order to avoid denaturalization deterioration processes it is important to not overcome the temperature of 42 ºC. For the ecotype phenotyping the AAT was adjusted at 39ºC for 48h. However, for Col-0 background seeds 40º-41ºC different AAT conditions may be applied for 24-48h and 24h, respectively. 41 ºC for 24h implies a strong Arabidopsis seed aging treatment, as 40 ºC for 48h. However, 40 ºC for 24h sometimes is a soft treatment.

Col-0 seeds germinate 50% with 40 ºC for 36h. Nevertheless, it is remarkable that not always seeds of different treatments in same conditions age similarly. For tomato seeds Accelerated Aging Treatment (AAT) was set at 41 ºC for 72 h and for wheat, AAT was set at 41 ºC for 24h.

Following, the second quickest aging treatment is the CDT. This seed aging treatment avoids water imbibition but uses a highly humid atmosphere (75% RH) and high temperature (37-38 ºC) for seed aging. (Basak *et al.*, 2006; Châtelain *et al.*, 2013; Righetti *et al.*, 2015; International Seed Testing Association, 2018; Hay *et al.*, 2019). This high humid atmosphere is achieved with close chamber containing a NaCl-saturated solution at 37-38 ºC. Different days might be applied, but normally 14-15 days are used. This time window usually makes that Col-0 seeds germinate about a 50%. For ecotype aging, CDT was set with 14 days at 38 ºC in a 75%, similarly to other Arabidopsis assays.

The last artificial seed aging used is the EPPO. This treatment consists in aging seeds during months in a high-pressure oxygen atmosphere (Groot *et al.*, 2012; Nagel *et al.*, 2016; ISTA, 2018; Hay *et al.*, 2019). In is way, the main cause of aging in EPPO is the Oxygen-produced damage. In our laboratory, we stablished EPPO for five months in a five-bar oxygen chamber with 40% RH at 20-25 ºC. These conditions were used in both, Arabidopsis ecotype and mutant aging assays. Assays with reduced RH barely produced aging damage, suggesting the importance of humidity in the oxygen damage in EPPO.

All these treatments speed seed aging and permit to work in short times in seed longevity research. Artificial aging treatments age seeds similarly to NAT (Rajjou *et al.*, 2008; Bueso *et al.*, 2014*b*), however, each aging procedure might affect different physiological and molecular aspects of the seed deterioration-resistance trait. Therefore, seed dry storage at ambient temperature is the best assay to test seed longevity. In 18 months, Col-0 seeds germinate around 50%, thus the NAT for Arabidopsis ecotypes was set at 18 months (20-25 ºC, 40-60% RH). However, shorter times (~12 months) light-up differences in lines with affected seed longevity, while longer times (~24 months) shows differences with resistant lines. For tomato seeds, dry storage (10% RH) for 36 months at 4º C was used to mimic seed bank conditions, where tomato seeds are usually stored for long periods.

As seed longevity is an adaptive trait, seeds used in seed aging treatment were harvested at the same time, dried, and stored under the same conditions for at least 2 weeks prior to the experiment. More than 100 seeds were tested for each treatment. Mutant lines were tested three times in two different seed generations in all different assays. The germination ratio was measured at 7, 10 and 14 days after sowing for Arabidopsis, tomato and wheat respectively. Radicle extrusion and seedling establishment

with green cotyledons were used as germination indication. The number of established seedlings under our experimental conditions did not increase after the measure. The corrected germination ratio was calculated dividing the germination ratio in each treatment by the germination ratio in control conditions. This correction was more important in Arabidopsis ecotype treatments, as not all ecotypes germinated 100% in the control germination test, and some ecotypes with affected germination were discarded. Significant differences with control plants were assessed by Student's *t* test ($p < 0.05$). Calculation of probit curve and P50 value was performed with the R package dcr (Tausch *et al.*, 2019).

Proanthocyanidins visualization

Proanthocyanidins (PA) staining in seeds was performed 5 days-after-pollination (DAP) developing seeds when PAs are synthetized and prior to seed gets opaque. Seed coat PAs were stained with a 1% vanillin 6M HCl solution for 15 minutes (Debeaujon *et al.*, 2000). Later, seeds were washed two times and visualized under the microscope.

Mucilage staining

For mucilage halo visualization, seeds were first hydrated in water, and then mucilage-stained in a 0.2% Ruthenium Red solution. After 15 minutes, seeds were water-washed three times and visualized in the microscope.

Seed histological cuts

Siliques and seeds were incubated in freshly prepared FAE fixation buffer (10% formaldehyde (37%), 5% Acetic acid (glacial) 100% and 50% ethanol). To facilitate FAE tissue-penetration, samples were putted under three 5-minues vacuum cycles. Then, FAE buffer was renewed and incubated for 30 minutes. After, samples were dehydrated by serial ethanol incubations (50% and 70%) for 30 minutes each. In a Leica TP1020 Tissue Processor, samples were treated with ethanol, Histoclear® and eosin, and then they were included in paraffin wax. Paraffin blocks were cut in sections under the rotatory microtome Leica RM2025. Sections were then included in poly-L-lysine coated glass slides through a 40ºC water bath. Slides were dried at 40 ºC for 24 hours.

Toluidine blue and Sudan Black staining

Seed histological cuts were deparaffinated and hydrated through serial incubations in Histoclear® (10 min), ethanol 100% (10 min), ethanol 90% (5 min), ethanol 70% (5 min) and water (5 min, only for Toluidine blue staining). Toluidine Blue staining was performed through slide submergence in 0.02% toluidine blue solution for 2 minutes. On the other hand, Sudan Black staining was performed with a 1% Sudan Black B and 75% ethanol solution for 1 hour. After the desired staining, slides were water-washed and air-dried. Finally, water and coverslip were used for microscope visualization.

GUS staining

GUS staining was performed as described by Jefferson *et al.* (1987) with some modifications. For GUS staining, plant tissue expressing the glucuronidase protein might be fixed with 2% formaldehyde. This procedure fixes the glucuronidase protein in the tissue to permit a better localized signal. However, the glucuronidase activity is partially reduced. The plant tissues were washed up to three times in 100 mM sodium phosphate buffer pH 7 (577 mM Na_2HPO_4 and 423 mM NaH_2PO_4). Then, tissues were incubated overnight at 37 ºC imbibed in the GUS staining solution (1 mg/ml X-gluc also named 5-bromo-4-chloro-3-indolyl β-D-glucuronide, 2.5 mM potassium ferrocyanide and 2.5 mM potassium ferricyanide diluted in 100 mM sodium phosphate buffer pH 7). GUS reagents are degraded by light, so they must be stored in darkness. Vacuum helps the solution to penetrate the tissue for a better staining. After the GUS solution incubation (time might be adjusted for an adequate GUS staining), tissues were progressively hydrated with ethanol washes of 20 min (20%, 35%, 50%, and 70%). For quick protocol, samples might be directly washed in ethanol 70%. Incubation in ethanol 70% decolorate the samples by removing chlorophyll. Later, for a completely tissue decoloring, incubation in chloral hydrate for 24 hours at 4 ºC might be used. For seed coat GUS staining pieces of developing siliques were incubated for 2 - 4 hours in commercial bleach (2.5%) solution after chloral hydrate incubation to remove PAs facilitating GUS signal visualization in the developing seed coat. Finally, samples were visualized with water and coverslip in the microscope.

Sudan Red lipid polyester barrier staining

Lipid polyester barriers staining was performed as described in Brundrett *et al.* (1991) and Beisson *et al.* (2007) in dry seeds. Seed were bleached in a 100% commercial bleach (5% NaClO) containing 0.01% Triton X-100. Seeds were shacked until seeds became white. The bleaching solution may be renewed each 15 minutes for a quicker

bleach. After, seeds were washed with water, and then with ethanol. Later seed were delipidated through a chlorophorm:methanol (2:1) incubation for 30 minutes. Seeds were cleaned with ethanol and air-dried. Finally, the previously prepared Sudan Red staining solution was added to stain the lipid polyester barriers for 3-4 hours. The Sudan Red staining solution consists in 2.2 g/L Sudan Red 7B (MP Biomedicals 15803) diluted in PEG-300 of Sudan Red. The solution is heated at 90ºC in bain-marie to dissolve the Sudan Red 7B. For microscope visualization, Sudan Red solution was removed through multiple water-washings until the remaining solution was transparent.

Size calculation

Size calculation was done with Fiji program from dry seeds images (Schindelin *et al.*, 2012). The scale was set with a known distance, such as the scale bar of the image, drawing a straight line. Then, in the Analyse menu, the Set scale tools was used to configure the scale. Global scale was used for same size images. Images were transformed in white and black images for a better threshold analysis. In the Adjust menu, of the image menu, the Threshold tool was set and fixed for seed area selection. Particles were measured (in the analyse menu) and the area results were filtered discarding big and small values that corresponded to other image particles or seed aggregates. Diverse images were used to confirm the measurement and perform the statistical analysis.

Seed coat permeability test

Tetrazolium red assays were used for seed coat permeability tests. Dried *Arabidopsis* seeds were incubated in the dark in an aqueous solution of 1% tetrazolium red (2,3,5-triphenyltetrazolium chloride, Sigma-Aldrich) at 28°C for 48 h. The seeds were observed for change in color and imaged using a macroscope. If the tetrazolium salts penetrate into the seed, they are reduced by the embryo to a red precipitate called formazans.

The quantitative formazan assay was carried out as described Vishwanath *et al.*, (2013). 50 mg of seeds were incubated in 500 µL of 1% tetrazolium red solution at 30 ºC for 72 in darkness. After incubation, the samples were washed twice with water, resuspended in 1 mL 95% ethanol, and finely ground with a mortar and pestle to extract formazans. The final volume was adjusted to 1 mL with 95% ethanol. Samples were centrifuged for 3 min at maximum speed and the supernatant was recovered. The absorbance at 490 nm was measured by triplicate with spectrophotometer in each biological replica. Three biological replicas were assessed.

Acetyl bromide method

Acetyl bromide method was used to measure polyphenolic compounds from suberin and lignin in seed (Moreira-Vilar *et al.*¸ 2014). Approximately 50 mg of seeds were bleached in a 100% commercial bleach (5% NaClO) containing 0.01% Triton X-100, and homogenized. Then they were resuspended in 2 mL of pH7 50 mM potassium phosphate buffer. Samples were pellet at 1400*g* for 5 minutes, and 1 mL-washed two times with phosphate buffer (pH7), three times with 2% Triton X-100 in phosphate buffer (pH7), two times with 1M NaCl in phosphate buffer (pH7), two times with water, and two times with acetone. Then, the pellet was dried at 60ºC for 24 hours. This protein-free cell wall samples, were included in the tap of a microtube containing 400 µL of 25% acetyl bromide diluted in glacial acetic acid. Samples were incubated at 70 ºC for 30 minutes. After, samples were placed on ice. Once cooled, 700 µL of 2 M NaOH and 100 µL of 5M hydroxylamine-HCl and 400 µL of glacial acetic acid were added and mixed. Samples were centrifugated for 5 min at 2000*g* and the absorbance of the supernatant was measured at 280 nm. The percentage of acetyl bromide soluble lignin (ABSL) was calculated according to Foster *et al.,* (2010) and Rains *et al.*, (2018), through the following formula:

Seed lipid polyester analysis

Seed samples (approximately 0.1 g per replicate, 4 replicates per genotype) were incubated in hot isopropanol at 85°C for 15 minutes, allowed to cool down, then further ground with a Polytron, and delipidated by subsequent solvent extractions (i.e., isopropanol, CH_3OH:$ClCH_3$ [1:2], CH_3OH:$ClCH_3$ [2:1]) (Jenkin and Molina, 2015). The cell wall-enriched dry residues were depolymerized by NaOMe-catalyzed transmethylation, with the addition of 1 mg g^{-1} DW each of pentadecanol methyl heptadecanoate (17:0 ME) and pentadecalactone as internal standards., following the protocol described by Molina *et al.* (2006). The depolymerization products were extracted with CH_2Cl_2 and derivatized by treatment with pyridine/*N,O-bis*-(trimethylsilyl)-trifluoroacetamide (BSTFA) (wheat and tomato seeds) or with pyridine/acetic anhydride (Arabidopsis seeds). Derivatized samples were evaporated to dryness, resuspended in 500 µL hexane and subsequently analyzed by GC-MS on a TRACE 1300 Thermo Scientific gas chromatograph with a ISQ Single Quadrupole mass spectrometer detector. Split injection (5:1 ratio, 310°C) was used with TG-5MS column (Thermo Scientific; 30-m length, 0.25-mm i.d., and 0.25-mm film thickness). The oven temperature was programmed from 140 °C to 310 °C at 3 °C min^{-1} and increased to a 310 °C at a rate of 10 °C min^{-1}, with a final 10 min-hold at 310 °C. The helium flow rate was set at 1.5 mL min^{-1}. The mass spectrometer was run in scan mode over 40-600 amu (electron impact ionization).

HPLC analysis of LMW thiols and disulphides

Seeds (15-20 mg/replicate) before and after CDT were freeze-dried for 5 days before grinding with two 3-mm quartz beads at 30 Hz for 2 min at -80 °C. 1 mL of ice-cold 0.1 M HCl and polyvinylpolypyrrolidone equal to seed weight was added before extracting at 30 Hz for 2 min. The extract was centrifuged at 20,000 *g* for 10 min at 4 °C and 700 µL of the supernatant was transferred, avoiding the lipid phase, for re-centrifugation as before. LMW thiols and disulphides were analysed by HPLC as described (Kranner, 1998). Briefly, for the determination of total LMW thiols and disulphides, 120 µL of the supernatant was pH-adjusted to 8.0 with 180 µL of 200 mM bicine buffer, reduced with dithiothreitol (DTT) for 1 h, before labelling thiols with monobromobimane (mBBr) for 15 min and stopping the reaction with methanesulfonic acid. To measure disulphides only, thiols in 400 µL of supernatant were blocked for 15 min with *N*-ethylmaleimide (NEM), which was added just before pH-adjustment with 600 **µL** bicine buffer. Excess NEM was removed with an equal volume of toluene four times. Disulphides in 300 **µL** of NEM-treated extract were reduced with DTT and labelled with mBBr, as for total LMW thiols and disulphides. Labelled LMW thiols were separated by reversed-phase HPLC using an Agilent 1100 HPLC system (Agilent Technologies, Santa Clara, CA, USA) on a ChromBudget 120-5-C18 column (5.0 µm, BISCHOFF GmbH, Leonberg, Germany), and detected by a fluorescence detector (excitation: 380 nm, emission: 480 nm). The concentration of each LMW thiol was calculated by subtracting the amount of the corresponding LMW disulphide from the total amount of each LMW disulphide and thiol.

Calculation of the cellular thiol-disulphide based redox environment

The half-cell reduction potential (E_{hc}) for each thiol-disulphide couple was calculated according to the Nernst equation, and mathematically combined into the LMW thiol-disulphide redox environment ($E_{thiol\text{-}disulphide}$) (Kranner *et al.*, 2006; Schafer and Buettner, 2001):

$$E_i = E^{0'} - \frac{R\,T}{n\,\mathrm{F}}\, ln \frac{[LMW\ thiol]^2}{[LMW\ disulphide]}$$

where *R* is the gas constant (8.314 J K^{-1} mol^{-1}); *T*, temperature in K; *n*, number of transferred electrons (2 GSH → GSSG + 2H^+ + 2 e^-); F, Faraday constant (9.6485 x 10^4 C mol^{-1}); $E^{0'}$, standard half-cell reduction potential of a thiol-disulphide redox couple at an assumed cellular pH of 7.3 ($E^{0'}_{GSSG/2GSH}$ = -258 mV). E_i is the half-cell reduction

potential of an individual redox couple i, and [reduced species]$_i$ is the concentration of the reduced species in that redox pair. The molar concentrations of LMW thiols and disulphides for each redox couple were calculated based on seed water content.

Plant genomic DNA extraction

Plant genomic DNA extraction was performed either in leaves or in seedlings following the CTAB extraction method (Weigel and Glazebrook, 2002). Plant material was smashed directly in the microtube (N_2 freezing helps) and rapidly 400 µL of CTAB solution was added. CTAB solution consist on 2% hexadecyltrimethylammonium bromide (CTAB), 1.4 M NaCl, 0.1M Tris-HCl pH8 and 20 mM EDTA. Mixed solutions were incubated at 65ºC for 20 min. Later, chloroform extraction (400 µL) was used to remove organic components, keeping the aqueous phase obtained after a soft centrifugation (3 min, 10000 rpm). DNA was precipitated with one volume of isopropanol (~350 µL) (1 h at -20 ºC, then spin for 30 min at maximum speed). One mL of cold ethanol 70% washed the DNA pellet, which was detached form the tube walls for a better washing. Finally, ethanol was removed, pellet was air-dried and DNA was resuspended in 50 µL of MilliQ H_2O.

RNA extraction

Total RNA from 7-days Arabidopsis seedlings, wheat leaves and tomato leaves was extracted either with E.Z.N.A. Plant RNA Kit (Omega Bio-tek) or with NucleoSpin RNA II (Macherey-Nagel). One-week old siliques and dry seed Arabidopsis RNA was extracted from seeds following the protocol of Oñate-Sánchez and Vicente-Carbajosa (2008). The complete protocol is attached at the end of the manuscript, at the extended protocol section. RNA was then purified using silica spin columns from previous RNA purification kits. DNase treatment was performed as indicated by the manufacturer.

Gene expression analysis

From one to three µg RNA was reverse transcribed using the Maxima first-strand cDNA synthesis kit for RT-qPCR (Thermo Scientific, Waltham, Massachusetts, USA) according to the manufacturer's instructions. Quantitative (real-time) PCR (qRT-PCR) was performed in triplicate using an Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher Scientific) with 5x PyroTaq EvaGreen qPCR Mix Plus (ROX) (Cultek S.L.U., Madrid, Spain). Three technical replicas were analysed. PCR amplification specificity was confirmed with a heat-dissociation curve (from 60 to 95 °C).

Relative mRNA abundance to *PP2AA3* or *At5g55840*, depending the higher or lower relative abundance of the transcript of interest respectively (Czechowski *et al.*, 2005), was calculated using the comparative ΔCt method (Pfaffl, 2001). Primers used in qRT analysis are listed in Supplemental data 11.

Ecotype-phenotype correlation analysis

The easyGWAS online tool (Grimm *et al.*, 2017) was used to calculate Spearman correlations (r_s) between our data and different public data. Strong direct correlations were considered if $r_s \geq 0.3$, and strong inverse correlation if $r_s \leq -0.3$.

GWAS analysis

All four GWAS were performed using the online application GWAPP (Seren *et al.*, 2012). Corrected germination ratios were used as GWAPP input. Only NAT data were transformed logarithmically, for Shapiro-wilk score and pseudo-heritability calculation. The dataset used for the GWAS was the Imputed Fullsequence Dataset (Horton *et al.*, 2012; Long *et al.*, 2013; Cao *et al.*, 2011; Gan *et al.*, 2011). GWAS results were filtered by Minor Allele Frequency (MAF) ≥ 0.05. Genes 1.5 kilobases (Kb) up and downstream of significant SNPs were mapped using TAIR10 (The Arabidopsis Information Resource version 10). The significance threshold in this study was set at *p* value $\leq 10^{-5}$.

RNA-Seq and bioinformatics analysis

Total RNA from 7-day-old-seedling was used for library construction, performed using the TruSeq Stranded mRNA Library Preparation Kit (Illumina). Resulting fragments were sequenced in the lllumina Hiseq 2000 system at Sistemas Genómicos (Paterna, Spain). The sequence alignment and the quantification of gene expression levels were performed as previously described in Mandel *et al.*, 2016 with some modifications. The reads (an average of 25 M per sample at 101 nucleotides long) were quality filtered and trimmed using Trimmomatic version 0.36 (Bolger *et al.*, 2014). The resulting reads were then aligned to the TAIR10 version of the *Arabidopsis thaliana* genome sequence (https://www.arabidopsis.org/) using Hisat2 version 2.1.0 (Kim *et al.*, 2015). The resulting read alignments (in BAM format) were used for transcript quantification with cuffdiff program of the Cufflinks version 2.2.1 package (Trapnell *et al.*, 2013). Three biological replicates were used for each genotype. The resulting read alignments were visualized and explored using Tablet software (Milne *et al.*,

2013) and CummeRbund R package version 2.23.0 (Goff *et al.*, 2014). Differentially expressed genes (DEGs) were subjected to Singular Enrichment Analysis (SEA) for the identification of overrepresented Gene Ontology (GO) terms using agriGO (http://bioinfo.cau.edu.cn/agriGO/) (Berardini *et al.*, 2004) with the default options (statistical test: hypergeometric, multi-test adjustment method: Yekutieli, significance level: 0.01). Using the criteria of two-fold up- or down-regulation, DEGs

In silico gene expression analysis

Gene expression analysis was performed with the online tool eFP browser (Winter *et al.*, 2007). Developing seed gene-expression was obtained from Le *et al.* (2010) data. Subcellular location information from Hooper *et al.* (2017) data. *In situ* hybridization could be used to confirm the gene expression of different genes, but was not used due to the complexity of the technique.

Chromatin immunoprecipitation and sequencing

Double mutant *athb22 athb25* plants were transformed with the *proUBQ10::AtHB25:3xHA* construction. Western blot was used to confirm the expression of AtHB25:3xHA in developing seeds. The line with higher AtHB25:3xHA abundance and similar *athb25-1D* phenotype was selected for ChIP-seq experiment.

Transgenic *proUBQ10::AtHB25:3xHA* siliques in maturation stage were collected and gently ground. They were filtered using liquid nitrogen to obtain high seed enrichment (5 grams approximately). The ChIP protocol was performed as described by Gendrel *et al.* (2005). The protocol is detailed below in attachments (Extended protocols section). Mouse monoclonal antibody HA.11 (clone 16B12) COVANCE and Dynabeads® Protein G were used for the immunoprecipitation. An equivalent sample with no antibody was used as a negative control. The presence of AtHB25:3xHA after the immunoprecipitation with antiHA:HPR antibody was confirmed by immunodetection. The Input, two ChIP replicas and the sample without antibody were sequenced with the 1x50 bp HiSeq2500 by the Genomic Unit of CRG (Barcelona, Spain).

ChIP-seq downstream analysis

For ChIP-seq data analysis, from sequencing reads adaptor trimming was performed using cutadapt (Martin, 2011) and mapped using the BWA tool (Li and Durbin, 2009) and TAIR10 Arabidopsis genome reference. Unique mapped sequences were filtered

and sorted with SAMTOOLS (Li *et al.*, 2009). Duplicated sequences were marked with Picard MarkDuplicates (Picard Toolkit, 2019). MACS2 (Zhang *et al.*, 2008) was used to call peaks. Threshold significance was stabilised with a False Discovery Rate (FDR) of 0.05. We used the CLC genomics program (QIAGEN) for the graphical corroboration to confirm peaks as AtHB25 binding sites. The HOMER motif analysis tool (Heinz *et al.*, 2010) was used for motif discovery and motif enrichment analysis in 200 bp around the peak centre. PAVIS (Huang *et al.*, 2013) was used for TSS distance analysis. Genes within 3000 pb around the peak centre were annotated with TAIR10 and PlantGenIE (Sundell *et al.*, 2015) and were considered as putative AtHB25-target genes. GO enrichment analysis was performed with agriGO (Tian *et al.*, 2017).

AtHB25 orthologue identification

AtHB25 best candidates orthologous in wheat and tomato were candidates identified by PLAZA monocot 4.5 (Van Bel *et al.*, 2018). AtHB25 aminoacidic sequence was compared with the tomato orthologous Solyc04g0142260.1 and the wheat orthologous TraesCS5A02G246500. Multiple sequence alignment was performed with Clustal W and LALIGN was used for the identity-similarity analysis. Both tools are available on www.expasy.org.

Conserved promoter fragment identification

AtHB25 best candidate orthologous were identified by PLAZA 4.5 DICOTS in *Arabidopsis thaliana* closely related species. Then, four species with a highly conserved *AtHB25* orthologue up-stream region were selected for promoter alignment. The best-conserved promoter regions of corresponding genes is expected to have an important regulatory role as has been maintained among species. Clustal W alignment was used to obtain this region.

LB media preparation

Luria-Bertanil (LB) media was prepared liquid and solid for *Escherichia coli* and *Agrobaterium tumefanciens* propagation with 10 g/L of Tryptone (Conda), 5 g/L of yeast extract (Conda), and 10 g/L of NaCl. pH was adjusted to 7. For solid media preparation 15 g/L of American bacteriological agar (Conda) was added. Sterilized media was then distributed in 100 mm petri dishes or plates for solid media preparation, or 5 mL for liquid bacteria growth.

Antibiotics, if needed, were added in warm media (~50ºC) before solidification. For easy management they are prepared in freeze 1000X stocks. The stocks are the following: Kanamycin (50 mg/mL), Spectinomycin (50 mg/mL), Ampicillin or the stable homologue Carbenicillin (100 mg/mL), Gentamicin (25 mg/mL), Cefotaxime (200 mg/mL), Rifampicin (10 mg/mL) and Tetracycline (10 mg/ml diluted in ethanol).

SOC media

Liquid SOC media was prepared with 20 g/L of Tryptone (Conda), 5 g/L of yeast extract (Conda), 10 mM of NaCl, 2.5 mM KCl, 10 mM $MgSO_4$ and 20 mM glucose. The $MgSO_4$ and the glucose were added after media sterilization from prior-sterilized stock solution (1M and 20% respectively).

Escherichia coli growth

E. coli was grown overnight at 37 ºC in darkness with the pertinent antibiotic. For liquid media culture, tubes were shacked at 200 rpm. Cell cultures might be stored at -4ºC in agar plates for two months. For long-term storage, liquid LB (with antibiotic) saturated *E. coli* cell culture from unique confirmed cell colonies were stored with glycerol (200 µL of glycerol 80% for each 800 µL of cell culture) at -80ºC.

Escherichia coli competent cell preparation

E. coli strain DH5α was used for competent cell preparation with the Inoue *et al.* (1990) method. A 200 µL aliquot from a saturated *E. coli* cell liquid culture was inoculated in 400 mL of SOB media (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 3 mM KCl, 10 mM $MgCl_2$ and 10 mM $MgSO_4$) and grown overnight with gentle agitation at room temperature. It is recommended to parallel two cell cultures inoculated at different time points (usually 14:00 and 18:00) as the room temperature might vary among the year, and cells might grow differently. Next day, the absorbance at 660 nm (A_{660}) should be within 0.4 and 0.8. The absorbance is important for the final bacteria transformation efficiency. For a more precise absorbance calculation is recommended to dilute three times the cell suspension and then multiply by 3 the result. Then the cell culture is placed on ice for 10 minutes. Cells are pellet in by centrifugation in four 50 mL tubes for 5 minutes at 5000 rpm at 4 ºC. Then, cells are resuspended in 10 mL of cold transformation buffer (TB) (20 mM piperazine-N,N′-bis (PIPES) pH 6.7 (KOH adjusted), 20 mM $CaCl_2$, 210 mM KCl and 50 mM $MnCl_2$). Cell suspension were placed in two tubes and again centrifuged in same conditions. Then, cells were

resuspended in 4 mL of cold TB and placed in a 15 mL tube, and the suspension was rinsed upon 10 mL with cold TB. Cold 750 µL DMSO was added in a continued way, little by little and with softly shaking. After 10 minutes in ice, the cell suspension was distributed in 100 µL aliquots that were N_2 frozen and stored under -80 ºC. To check the presence of contamination, a solid LB with ampicillin (or with other antibiotics) culture might be used. There should no be colonies next day after the culture at 37ºC. The transformation efficiency should be checked next days to confirm the adequate competence of *E. coli* cells.

Escherichia coli transformation

E. coli cells were transformed following the Heat Shock method. Frozen *E. coli* competent cells aliquots were ice placed for a slow unfreezing. From 1-10 µg of plasmid was added and mixed. Cells were incubated in ice for 30 minutes. Heat shock was set at 42 ºC for 30 s - 1 min. Later, in sterile conditions, 300 µL of SOC was added, and cells were incubated for 1 hour at 37 ºC with shaking. Finally, bacteria were plated and distributed with crystal balls in LB solid media with the pertinent antibiotic. Reduced times worked also efficiently. After overnight growth, bacteria colonies were isolated for plasmid insertion corroboration and further analysis.

Polymerase chain reaction (PCR)

Two different kind of polymerase chain reactions (PCR) were used. For analytical purposes, Taq PCR was performed. The Taq-PCR reaction mix was prepared in ice as following (per reaction): 13.5 µL of MilliQ H_2O, 2 µL of 10X Taq buffer (Biotools), 0.5 µL 10 mM dNTPs (dATPs, dTTP, dGTP and dCTP; 2.5 mM each), 1 µL of forward primer (10 µM), 1 µL of reverse primer (10 µM) and 1 µL Taq polymerase (Biotools) corresponding to 1U. A negative control, and if possible, a positive control was used for each PCR reaction. In addition, extra volume was needed to avoid pipetting errors and to have enough mix to distribute. Around 10% extra samples were considered for mix calculation. The reaction mix is gently vortex and 19 µL were distributed in different wells and the DNA (1 µL) was added. For colony PCR, explained below, 1 µL extra of water was added during the mix preparation and 20 µL distributed. The PCR program depended on the PCR reaction primers and product, but generally followed similar steps. The program starts with 95 ºC for 5 min for an initial strong denaturation, and then cycles start. We usually wait for thermal cycler block overcoming 80 ºC to insert samples in the machine, to avoid possible unspecific and primer-dimer amplification, whose might occur at low temperatures. Each cycle is repeated 30-35 times and consists of 30 sec at 95ºC for DNA denaturation, 30 sec at the specified

primer annealing temperature (which usually varies among 55 to 60 ºC) and 1 min per kilobase to amplify at 72 ºC, which depends on the expected PCR product length. Finally, a final step of 5 min at 72 ºC was used for the final extension of DNA fragments. The PCR program ends by holding the samples at low temperature (4-10 ºC) for DNA protection until collection.

The other PCR reaction used was the Phusion PCR. This reaction was used for preparative purpose as gene cloning. The Phusion polymerase presents proofreading activity (3'-5' exonuclease activity) that permit the correction of replication errors, while the Taq polymerase do not present this activity. The Taq polymerase has an error rate of 2.28×10^{-5}, while the Phusion polymerase error rate is 4.4×10^{-7}. In 30 cycles for a one kilobase PCR product, the estimated percentage of DNA molecules with an error which the Taq is around the 68.4%, while it is only a 1.32% for the Phusion enzyme. As the Phusion polymerase is considerably more expensive than Taq polymerase, Taq PCR is used for analytical preparations and the Phusion PCR for preparative purposes.

The Phusion-PCR reaction mix was prepared in ice as following: 30.5 µL of MilliQ H_2O, 10 µL 5X of Phusion® HF Buffer (BioLabs), 1 µL of 10 mM dNTPs, 2.5 µL of forward primer (10 µM), 2.5 µL of reverse primer (10 µM), 1.5 µL of DMSO, 1.5 µL of template DNA (~100 ng/µL, usually cDNA) and 0.5 µL Phusion® polymerase (BioLabs) corresponding to 1U. The PCR program used followed a similar structure than Taq PCR program. The program starts with 98 ºC for 3 min, and then cycles start. Each cycle is repeated 35 times and consists of 10 sec at 98 ºC, 15 sec at 55-60 ºC and 72 ºC for 30 sec per kilobase of the PCR product. Finally, a final step of 7 min at 72 ºC was used for the final extension of DNA fragments. The PCR program ends by holding the samples at low temperature (4-10 ºC) for DNA protection until collection.

There is another optional Phusion protocol if this normal Phusion amplification does not work (even if adjusting annealing conditions and extension times are optimal). The Phusion® GC buffer (BioLabs) permits difficult PCR DNA amplification. Mix is prepared in ice as following: 35.75 µL of MilliQ H_2O, 10 µL 5X of Phusion® GC Buffer (BioLabs), 1 µL of 10 mM dNTPs, 1 µL of forward primer (10 µM), 1 µL of reverse primer (10 µM), 1 µL of template DNA (~100 ng/µL, usually cDNA) and 0.25 µL Phusion® polymerase (BioLabs) corresponding to 0.5 U. The program starts with 98 ºC for 30 sec, and then cycles start. Each cycle is repeated 35 times and consists of 5 sec at 98 ºC, 10 sec at 55-60 ºC and 72 ºC for 30 sec per kilobase of the PCR product. Finally, a final step of 7 min at 72 ºC was used for the final extension of DNA fragments. The PCR program ends by holding the samples at low temperature (4-10 ºC) for DNA protection until collection.

Colony PCR

For rapid correct plasmid insertion corroboration, direct PCR on *E. coli* and Agrobacterium colonies was performed. Selected colonies were orderly striped in a new plate and placed in a 20-µL PCR reaction mix. High bacteria amounts might inhibit the PCR reaction. To ensure that bacteria membranes break, and plasmid gets liberated in the PCR mix, cells are boiled for 5 minutes prior to PCR cycles of the PCR program.

For yeast-colony PCR, a generous loop of yeast colony was resuspended in 30 µL of fresh 20 mM NaOH solution and boiled at 95 ºC for 5 minutes. Then the solution was rapidly placed in ice and used as DNA template to let debris get to the bottom of the tube. One µL of the supernatant was used in the PCR mix (preferably in a 30-µL Taq PCR mix). The PCR mix should turn a bit cloudy.

Plasmid annealing primers used in colony PCR are listed in Supplemental data 11.

DNA electrophoresis

DNA electrophoresis was performed with TBE buffer and TBE agarose gels. TBE 10X stock buffer is prepared for later dilution. It is formed by 54 g/L of Tris-Base, 27.5 g/L of Boric Acid and 20 mL of EDTA (pH 8) 0.5 M. For gel preparation, 7 g/L of Agarose D1 Low EEO (Conda) was added to the TBE buffer at working solution. Microwave heating until boiling and gently mixing for two times was used to dissolve the agarose. Ethidium bromide (50 µL from 1% stock) was added to warm TBE-agarose (~50 ºC) and TBE buffer. Solid gels were covered (1mm) with the TBE buffer. DNA for electrophoresis was prepared by adding a 1/5 volume loading buffer 6X (40% Sucrose, 0.25% Bromophenol blue and EDTA 0.1M) and loaded. DNA ladder was also loaded for fragment size identification. Electrophoresis was set at 80 V (10 V/cm) of direct current (DC). DNA was visualized under UV light in a transilluminator.

E. coli plasmid DNA extraction (miniprep)

Normally, plasmid DNA from 4 mL saturated culture of *E. coli* was extracted with the use of commercial plasmid miniprep kits, which use silica spin columns for a rapid and clean plasmid extraction. The protocol followed was as manufacturers indicate. However, different buffers of from different kits might be combined and applied as the extraction principle is the same: resuspend cells, lysate cells with NaOH, neutralize the suspension with acetic acid, centrifuge to discard solid components, bind DNA to the silica spin column, clean DNA through alcohol and elute DNA with water (or TE).

However, for higher amounts we performed a traditional plasmid DNA extraction. Cells where grown overnight in LB with the plasmid selective antibiotic, and then they were pellet. The pellet was resuspended in 200 µL of GTE buffer (50 mM Glucose, 25 mM Tris-HCl pH8 and 10 mM EDTA). Then 300 µL lysis buffer (0.2M NaOH and 1% SDS, freshly made) was added, and the suspension was softly mixed by tube inversion (5-6 times). In no more than 5 minutes, 300 µL of neutralization buffer (3M Potassium acetate pH 4.8) was added and the solution was softly mixed similarly. A 5 min ice incubation was used prior to a 10 min centrifuge-spin (maximum velocity). Supernatant (~750 µL) was transferred to a 2 mL tube and 800 µL of chloroform was added for organic component extraction (two extractions of 400 µL of chloroform might be better). Tubes were centrifuged for 1 min at 10000 rpm, and the aqueous phase was transfer to a clean tube. DNA was precipitated with 1 volume of isopropanol and tubes were incubate on ice for 10 minutes. Then, the DNA was pellet for 15 minutes centrifuging at maximum speed. The pellet was cleaned with 1 mL of ethanol 70% and air dried. Finally, the DNA was resuspended in 50 µL of water (or TE) with 0.2 µL of RNase 1%.

Two independent plasmid DNA digestion with restriction DNA endonucleases were used to confirm that the extracted plasmid was the expected plasmid of interest, through the agarose DNA electrophoresis created-fragment size visualization. In some cases, plasmids were sequenced to confirm the correct insertion and the absence of mutation. In addition, plasmids were used for molecular cloning and for *E. coli*, Agrobacteria, and yeast transformation. Plasmid DNA was stored at -20 ºC.

DNA restriction with DNA endonucleases

DNA was restricted or cut for plasmid corroboration and for molecular cloning with the desired enzyme, with the indicated manufacturer's buffer, which usually included bovine serum albumin (BSA). Normally, 1-3 µg of plasmid DNA was used for the 20-µL digestion (with 2 µL of the specific 10X buffer, and 0.5 µL of the digestion enzyme (~1 U)). Digestion was set for 1 hour at the specific enzyme temperature (usually 37 ºC) but might be longer if the enzyme do not present star activity. For plasmid corroboration, two or three cuts are necessary to easy interpret the fragmentation pattern. It is useful to search one cut inside the specific insertion and another in the plasmid skeleton. In addition, the creation of DNA fragments ranging from 300 bp to 3000 bp makes the fragment identification easier, as shorter and longer usually are more difficult to stablish their concrete size through the electrophoresis in 0.7% agarose gels. If possible one enzyme was used, if not double digestions were used.

In case of double digestions, if possible, they were performed in one-step in the same buffer. This compatibility can be rapidly assessed with the DoubleDigestionCalucula-

tor online tool of Thermo Scientific. If it was not possible to perform a one-step double digestion, first the reaction with the buffer with less salt concentration was used, and then the other adjusting the final buffer concentration to a similar original concentration. If this did not work properly, chloroform-extraction, followed by ethanol DNA precipitation was used, but this leaded to a reduced plasmid yield.

In molecular cloning, the design should include unique restriction enzyme cuts that permit efficiently cut and ligated desired fragment in the desired place and direction in the plasmid. Compatible enzymes, which create compatible DNA ends or same overhang extremes (sticky ends) can be used. 5' overhang ends can be filled, while 3' ends removed with the T4 DNA Polymerase or the Klenow fragment of DNA Polymerase I. In addition, overhang extremes can be fill partially or totally (blunt end) to permit ligation compatibility by adding only the required dNTPs (dATP, dCTP, dTTP and/or dGTP). Required buffer, dNTPs and 1 µL of enzyme was used for this. Samples were incubated at 37ºC for 1 hour to overnight. In double digestions, the end to blunt or fill should be cut in first place, then blunted or filled, the enzyme inactivated (by a 75ºC heat for 20 minutes) and then cut at with the second restriction enzyme.

DNA ligation

DNA ligation was performed with the T4 ligase. Similar amounts of both previously cut and agarose-gel purified DNA molecules were used for DNA ligation. They should have compatible ends, preferably sticky. It is not recommended more than one blunt end per ligation, as the efficiency is highly reduced. The reaction mix was in 15 µL, containing 1.5 µL of ligase buffer (which contains dATP, however other buffers as used in DNA restriction might be used if adding dATP), ~2 µL of plasmid backbone and ~5 µL of DNA insert and 1 µL of T4 ligase. After overnight incubation at 4-10ºC, the ligation reaction was transformed in *E. coli*. For an easy DNA ligation recovery, only the recombined ligation product should form a replicable plasmid with the antibiotic resistance gene.

Plasmids cut with one single restriction enzymes usually present compatible ends, and may autoligate, a process which is statistically more probable that the double ligation with the fragment of interest. To avoid plasmid autoligation, an alkaline phosphatase can be used in the plasmid linearized molecule. With this treatment, the overhanging 5'-phosphate, required for the ligation through the formation of phosphodiester bond with the 3'-OH of the other DNA molecule, is removed and autoligation is restricted. However, any molecular process works 100% efficiently and some autoligated molecules might be formed. To form the circularized desired plasmid only the phosphate of the 5'-end of the insert can contribute with the bond, and a DNA nick is present in each strand (in different locations, corresponding with each restriction reconstructed

site). Nevertheless, during bacterial plasmid replication, these nicks are repaired, and plasmid is completely formed with all nucleotide-nucleotide bounds. Alkaline phosphatase was used for the phosphatase treatment (rAPid Alkaline Phosphatase, Roche).

Agarose gel DNA band extraction

DNA band were extracted for preparative DNA purposes, as in molecular cloning. It was necessary after the Phusion PCR, prior to pCR8 recombination to eliminate primer-dimers and small and undesirable DNA fragments to foment the desired PCR product-plasmid recombination. In addition, for traditional cloning, after DNA digestion, and prior to DNA ligation, the desired fragment was extracted from an agarose gel to avoid the ligation to undesired fragments.

Agarose gel DNA band extraction was performed with commercial DNA band extraction kits as indicated by manufacturers. These kits consisted in to liquate the agarose with a buffer rich in chaotropic salts, which in addition, permit to bind the DNA with a silica spin column. Then, the DNA in the column is wash with ethanol and eluted with TE.

pCR8 recombination

Normally, cloned products with Phusion PCR, were inserted in the plasmid pCR8 from the pCR™8/GW/TOPO® TA Cloning® Kit from Thermo Scientific (Waltham, Massachusetts, USA). This small plasmid permits the easy DNA incorporation and corroboration. When the correct insert is corroborated, through a miniprep, it is possible to obtain low-cost high amounts of the DNA fragment of interest, to subclone it in the final and bigger plasmid in a more efficient way.

The linearized *pCR8* (with the TOPO isomerase) has 3'-T overhanging ends, and thus requires a 3'-A overhanging PCR product for the recombination and plasmid recircularization. These 3'-A overhanging ends are commonly created by Taq PCR amplification, but not in the Phusion PCR as the 3'-5' exonuclease activity remove the 3'-A overhanging ends, and thus DNA Phusion products present blunt ends. For this reason, the agarose-band purified Phusion PCR product should have 3'-A overhanging addition by the Taq polymerase. The Phusion PCR band is purified by agarose-band extraction to ensure the expected fragment size, select the correct one (in the case of unspecific amplification), and to eliminate other DNA fragments such as primer-dimers or other undesirable DNA fragments, as the TOPO isomerase, responsible of the recombination, can recombine *pCR8* with any DNA fragment with 3'-A overhanging ends. Moreover, smaller fragments would be inserted in *pCR8* at higher rates due to

the higher probability that both DNA-fragment ends recombine with the same plasmid molecule. The 3'-A overhanging ends are added with a 15 min Taq extension at 72 ºC. In this mix, dATPs (or dNTPs), 10X Taq buffer and Taq polymerase are added at similar amounts than in the Taq PCR reaction (0.5 µL, 2 µL and 1 µL respectively, to 16.5 µL of purified DNA band).

The *pCR8* recombination reaction was performed by the addition of 0.5 µL the provided salt solution and 0.1 µL of *pCR8*/ TOPO to 2 µL of the 3'-A overhanging-end Phusion PCR production. After 1 hour (or overnight incubation for better results), *E. coli* cells were transformed with the reaction product and selected with Spectinomycin.

However, even if the inserted fragment product is the correct, it might be inserted in the undesired direction. This *pCR8* plasmid present Gateway *attL1* and *attL2* recombination sites, flanking the DNA insertion. Normally, Gateway destination vectors (with *attR* sites) are designed to have *attR1* whereas the start codon or initial sequence should recombine and the *attR2* where the end-part of the insert sequence should recombine. Therefore, the insertion in *pCR8* of the DNA fragment should have recombined with the *attL1* region next to the beginning of the sequences and the *attL2* next to the end (as *attL1* recombine with *attR1*, and *attL2* with *attR2* in the LR reaction). A colony PCR with the forward primer used for insert amplification and *T7* primer (located close to the *attR2*) can be use to assess this insertion direction prior to plasmid DNA extraction. The plasmid *pCR8* also present EcoRI restriction sites flanking the insert to rapidly isolate the insert. If EcoRI is used to extract the DNA insert, this direction corroboration is not necessary, but it probably is necessary to check it after the following DNA insertion.

LR reaction

In some constructs (*pMDC* and *pUB-DEST* plasmids), *pCR8* vectors containing the DNA insertion were recombined into the final destination plasmid through the Gateway LR reaction, were *attL* sites recombine with *attR* and form *attB* sites in the final plasmid with the recombined insert. This reaction was performed as following: 3 **µL** of the *pCR8* plasmid with the insert, 1 **µL** of the destination plasmid, 0.2 **µL** of LR clonase II enzyme mix (Thermo Scientific). After 1 hour (or overnight incubation for better results), *E. coli* cells were transformed with the reaction product and selected with the destination-plasmid selection antibiotic. Only, the correctly recombined (re-circularized) plasmid should growth, as the unrecombined plasmid contains the *ccdB* lethal gene for *E. coli* DH5**α**. For a correct Gateway system work, the destination plasmid and the entry plasmid (*pCR8* in our case) should have a different antibiotic resistance gene. In this way, bacteria containing the entry plasmid should not be selected. This was not the case in *pUB-DEST* LR, as the *pUB-DEST* plasmid contains

the Spectinomycin resistance gene, as *pCR8*. To avoid *pCR8* selection, the *pCR8* plasmid containing *MYB47* insertion was backbone-cut (outside the insertion) for plasmid linearization, and thus to avoid *pCR8* bacterial-selection (which would be detrimental for the reaction, as initial plasmids are enormously more abundant than LR products, which occurs eventually. However, with the appropriate selection, it is enough to obtain efficiently the desired final vector.

Western blot

For recombinant protein AtHB25:3xHA corroboration western analysis was performed. SDS-PAGE electrophoresis was performed in Mini-PROTEAN 3 cells (Biorad). Acrylamide gel preparation consisted in two parts: the resolving gel preparation and the stacking gel preparation. The preparation of 1mm resolving gel consisted in 3 mL of 2X Resolving buffer (0.75M Tris, 0.2% SDS, pH8.8 with HCl), 1.5 mL of Acrylamide (38% Acrylamide, 2% Bis-acrylamide, filtered with 0.45-µm pore nitrocellulose) and 1.5 mL of water for a 10% acrylamide gel. For 12%, 1.8 mL of Acrylamide and 1.2 mL of water was used, and for 8% 1.2 mL and 1.8 mL, respectively. It is important wear globes due to the acrylamide toxicity. To start the acrylamide polymerization, 40 µl of 10% ammonium persulfate (APS) and 4 µl of Tetramethyl ethylenediamine (TEMED) were added and mixed. The resolving mix was poured carefully in the gel cassette, about one centimeter below the wells. Isopropanol was used to eliminate bubbles formed in the gel surface. Polymerization last 1 hour approximately and it can be corroborated in the residual gel mix. Once polymerized the resolving gel, the isopropanol was removed, and the stacking gel poured. Stacking gel mix for 1mm gel consists in: 1.2 mL of 2X stacking buffer (0.25M Tris, 0.2% SDS, pH6.8 with HCl), 0.2 mL of Acrylamide, 1 mL of water, 20 µl of 10% APS and 2µ of TEMED. Plastic wells were added. After the stacking gel polymerization, the electrophoretic gel was inserted in the cell with sufficient electrophoresis buffer (14.4 g/L glycine, 0.1% SDS, pH8.3 with Tris).

Protein samples were treated as following: samples were grinded in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) (or other solution) and then, the different extractions were quantified with the method to normalize total protein extraction. Biorad protein assay (with Coomassie Brilliant Blue G-250) was diluted five times, and 200 µL of the reagent was distributed a spectrophotometer 96-well microplate. The calibration curve line was adjusted with known protein concentration (0, 200 ng, 400 ng, 600 ng, 800 ng and 1 µg of BSA). From 2-5 µL was added (and visually corroborated that is within the calibration standards, if not it was adjusted). Absorbance was measured at 530 nm. The calculated concentration was then used to dilute sample to a determine concentration. Then, 1/4 volume of 5X lamely buffered

(250 mM Tris-HCl pH 6.8, 10% SDS, 5% bromophenol blue, 50% glycerol and 500 mM of β-mercaptoethanol) was added, and samples were boiled for 10 minutes at 95ºC.

A maximum of 40 µL was loaded in the PAGE-SDS gel. The protein molecular marker PageRuler™ Prestained Protein Ladder (Thermo Scientific) was also loaded (3 µL) to have a protein size reference. Samples ran at 50 volts until they entered in the resolving gel. Then the voltage was increased to 100-120 volts, and electrophoresis ran until the electrophoresis front arrived at the bottom of the gel.

Then, proteins were transfer onto a polyvinylidene difluoride (PVDF) membrane. The PVDF membrane was previously activated with methanol. In addition, the TOWBIN buffer was prepared (14.4 g/L glycine, 25 mM Tris-base, 0.01% SDS, 20% methanol). The membrane and the electrophoresis gels were included in a "sandwich" with TOWBIN wet Whatman papers (three each side). Then, the sandwich submerged in TOWBIN was set to a directional voltage to transfer the negatively charged proteins to the membrane (membrane was located in the positive side). Reusable dry ice packs should be used to maintain the low temperature. For a rapid transfer, electrotransfer was set at 100 volts at 4 ºC for 90 minutes. For slow transfer, electrotransfer was set at 15 volts at room temperature overnight.

After the electrotransfer, the membrane was put in a tray containing 10 mL of TBS (150 mM NaCl, 20 mM Tris-CIH pH 7.6) with 0.05% Tween 20 and 2% non-fat milk, and was incubated for 1 hour on a rocker to block all unspecific protein-binding sites. The primary antibody was 1:10000 diluted (the dilution factor needs to be determined empirically for each antibody) in 10 mL of TBS-Tween 20-Milk solution. The membrane was then incubated with the primary antibody overnight at 4 ºC. Next day, the membrane was washed five times for 10 minutes with 10 mL of TBS-Tween 20. Then, similarly the secondary antibody (HRP labeled) was 1:10000 diluted in 10 mL of TBS-Tween 20-Milk solution. The membrane was in the secondary antibody solution for more than one hour. After, the membrane was washed for 3 times with 10 mL of TBS-Tween 20 for 10 minutes, as before.

To detect the protein the ECL Select™ Wester Blottinh Dectection Reagen (GE Health care) was used. Reagents were room warmed. A 1:1 solution was used for HPR detection from both ECL reagents. The membrane was TBS-drained, covered by the reaction mix and incubated in darkness for 5 minutes. Finally, the quimoluminescence was measured with a LAS-3000 Imager (Fujifilm). Exposure time was adjusted in accordance with the signal. A digitalized EPI image was taken for the protein ladder visualization for protein size determination.

In Arabidopsis, for the *proAtHB25::GUS:GFP* construction, a 1943 bp fragment corresponding to the native *AtHB25* promoter was selected and amplified with *proAtHB25-F* and *proAtHB25-R* primers from Arabidopsis genomic DNA. In the case of *proCOG1::GUS:GFP* a 1957 bp fragment corresponding to the native *COG1* promoter was selected and amplified with *proCOG1-F* and *proCOG1-R* primers. These regions comprise from -1943 and -1957 to 0 bp relative to the *AtHB25* and *COG1*, translation start site, respectively. Each promoter fragment was inserted in *pCAMBIA1303* (Muñoz-Bertomeu *et al.*, 2009) with *XbaI* and *NcoI* to drive the expression of the GUS-GFP construct.

For the *proMYB47::GFP* construction, a different strategy was followed. 1975 bp fragment corresponding to the native *MYB47* promoter (-1957 to 0 relative to the translation start site) was selected and amplified with *proMYB47-F* and *proMYB47-R* primers from Arabidopsis genomic DNA, as in the others promoter-reporter fusion constructions. However, the promoter fragment was inserted into *pCR8* plasmid by using the pCR™8/GW/TOPO® TA Cloning® Kit from Thermo Fisher Scientific (Waltham, Massachusetts, USA). LR reaction recombined our construct with the *pMDC107* plasmid (Curtis and Grossniklaus, 2003).

Similarly, for the construction of *MYB47* over-expression lines, the *MYB47* coding sequence was amplified with *MYB47-F* and *MYB47-R* primers from cDNA to clone it into pCR8. The resulting plasmid was recombined with *pUB-DEST* plasmid (Grefen *et al.*, 2010) by LR reaction.

For the *proUBQ10::AtHB25:3xHA* construction, PCR from cDNA with *AtHB25-pCM262-F* and *AtHB25-pCM262-R* primers was performed. To construct the *AtHB25:3xHA* fusion, the PCR product was combined with the *NotI*-digested *pCM262* plasmid providing 3xHA and 6xHis tags at the *AtHB25* 3'-CDS end through yeast homologous recombination. The *pCM262* vector is derived from *pCM190* (Garí *et al.*, 1997). The resulting *AtHB25:3xHA* fragment was inserted in *pGPTVII–Hyg–pUBQ10* (Krebs *et al.*, 2012) after *Apa*I and *Asc*I digestion. The *proUBQ10::COG1:3xHA* was similarly constructed.

The construction *35S::AtHB25*, for tomato transformation, was conducted as following. *AtHB25* CDS was amplified from Arabidopsis cDNA with *AtHB25-F* and *AtHB25-R* primers and it was cloned into the *pCR8* plasmid. The LR reaction was used to recombine the *AtHB25* CDS and the *pGWB2* plasmid (Nakagawa *et al.*, 2007).

For the *proUBI1::AtHB25* wheat construction, the Arabidopsis *AtHB25* CDS from *pBluescript-AtHB25* (Bueso *et al.*, 2014*b*) was amplified using *AtHB25-BAMH1-F* and *AtHB25-BAMH1-R* primers and subcloned in the *pAHC17* plasmid containing the

UBI1 maize promoter (Christensen and Quail, 1996) after digestion with *BamHI*.

For *pYRO::4xproAtHB25fr* for *Nicotiana benthamiana* transient expression, four repetitions of the conserved *AtHB25* promoter fragment were inserted into the *pYRO* with the minimal promoter (-58F8) as Castrillo *et al.* (2011). The cloned fragment was cut with BglII and HindIII, and the pYRO plasmid with BamHI and HindIII. BamHI and BglII create compatible ends, and after the ligation the restriction site disappear. In addition, the insert provides and extra BamHI restriction site for consecutive cloning. In four cycles of plasmid cut, fragment ligation, bacteria transformation, and plasmid DNA extraction the *proAtHB25fr* cloned fragment was inserted four times in *pYRO*.

All genetic constructs derived from PCR products were sequenced to ensure the correct sequence insertion and to avoid possible process-derived product mutations. Primers used for cloning and sequencing are listed in Supplemental data 11.

Molecular cloning for yeast interaction assays

For the Split-Tpr protein-protein interaction assay, TFs of interest were cloned in pCR8 were cloned with the *TF-SfiI-F* and *TF-bait-SfiI-R* primers for bait constructions (for *pIC-Trp*) and the *TF-SfiI-F* and *TF-prey-SfiI-R* primers for prey constructions (for *pN-Trp*). Notice that F primer is identical in both constructions. The difference of R primers relies in the *STOP* codon, present in prey primers to delimit the end of the recombinant protein. However, bait final construction present the split-Tpr protein in the C-terminal part of the protein and should not present this *STOP* codon. Both, forward and reverse primers provide SfiI restriction sites, but with different overhanging ends, compatible to the SfiI produced overhanging ends in *pIC-Trp* and *pN-Tryp*. This permit the direction insertion of with one enzyme cut, due to the properties of SfiI, which recognizes GGCCNNNNNGGCC, were N are any DNA letters. Thus, the 5-N central sequence can be designed for a concrete compatibility and incompatibility among different SfiI-cut sites. TFs were inserted into pIC-Trp (for bait constructions, HB25 and COG1) and pN-Tpr (for prey constructions). Centromeric plasmids pIC-Trp and pN-Tpr were used, but 2-µ version for a higher plasmid replication, and thus higher recombinant proteins levels was available. Nevertheless, it does not improve the results in our case.

The *proAtHB25fr::HIS3* construction was obtained through the insertion of the cloned genomic DNA fragment with the primers *proAtHB25fr-F* and *proAtHB25fr-R* in *pTUY1H* as described in Sánchez-Montesino and Oñate-Sánchez (2017). Primers provided the necessary restriction sites. The restriction enzymes XmaI and XbaI were used to cut both, plasmid and insert. Corresponding purified bands were ligated. Primers used are listed in Supplemental data 11.

Primer design

Primers were designed following different procedures according to the final purpose. For cloning, primers were designed to flank the DNA region of interest: for gene cloning primers, the forward primer contained 18-25 bp first coding sequence (CDS) bases and the reverse contained 18-25 bp last CDS bases. Primer length was set up to adjust a similar melting temperature in both primers (~57-60 ºC). The NEB Tm Calculator (Biolabs) online tool can used to calculate easily this melting temperature. For promoter regions, similar process was used, from the genomic region upstream the gene. Sequences were obtained from TAIR (CDS) and PLAZA DICOTS (genomic upstream regions) web sites. Restriction enzyme sites were added if needed in accordance with the plasmid molecular cloning design. In the case of recombinant proteins, primers were carefully designed to not disturb the open reading frame for a proper protein formation.

Primers for qRT-PCR were designed with Primer BLAST (NCBI) online tool following this protocol. The TAIR CDS gene sequence was pasted in the Primer-BLAST tool. PCR product length was set up between 70 to 250 bp. Max Tm was set up at 62 ºC and min Tm at 59 ºC. *Arabidopsis thaliana* was selected as organism, and the Primer-BLAST tool was launched. The primer pair from with less self-complementary score was selected from first five results. If gene present introns, it is interesting to design primers in different consecutive exons, to be able to discern between genomic DNA and cDNA amplification. Is necessary to check primer specificity in the melting curve obtained after the qRT-PCR, corroborating the presence of only one peak.

Primers for Arabidopsis T-DNA insertion lines were obtained through the iSet Primers tools from the T-DNA express web name with the default parameters, by just typing the T-DNA line. Obtained LP and RP primers were used for wild-type band reaction (which should be negative for the homozygous mutant plant isolation) and the RP and the LB primer (*SALK-LB*, *SAIL-LB* or *GABI-LB*, depending on the type of T-DNA mutant line) was used for mutant line calculation. Usually 57 ºC for Tm worked properly, and 1 min and 20 seconds of extension ensured correctly wild-type band elongation (which is the larger one, as calculated by the program).

In addition, sequencing primers were also designed to confirm the correct cloning of long DNA sequences that cannot be read with the plasmid common sequences (in *pCR8* the *M13-F* for forward sequence reading, and *M13-R* and *T7* for reverse sequence reading) or due to the absence of these sequences as in *proUBQ10::AtHB25:3xHA*. For a proper Sanger-sequencing lecture the primers should be placed in a ~30 bp distance, and for long inserts, primers should be distanced within 800 bp (or 1600 bp if one is a forward primer and the other is a reverse one)

Published primers were preferably used. Primers are listed in Supplemental data 11.

Agrobacterium tumefaciens growth

Agrobacterium was grown for three days in darkness at 28ºC in LB with the pertinent antibiotics for colony selection, and two days for liquid Agrobacterium propagation. For liquid media culture, tubes were shacked at 180 rpm. Cell cultures might be stored at -4ºC in agar plates for two weeks. For long-term storage, liquid LB (with antibiotics) saturated Agrobacteria cell culture from unique confirmed cell colonies were stored with glycerol (200 µL of glycerol 80% for each 800 µL of cell culture) at -80ºC.

Agrobacterium tumefaciens competent cell preparation

The *A. tumefaciens* strain GV3101 (C58C1) carrying *pSOUP* (Koncz and Schell, 1986) was used for competent cell preparation. A preculture of 50 mL LB with Gentamycin (for *pSOUP* selection) and Tetracycline or Rifampicin (for *E. coli* negative selection) was grown at 28 ºC until A_{660} of 1.0. The cell suspension was then centrifuged at 4500 rpm for 15 min. Pellet was resuspended in 1 mL of 20 mM $CaCl_2$. Aliquots of 100 µL were liquid nitrogen frozen and stored at -80 ºC.

Agrobacterium tumefaciens transformation

The freeze-thaw method was used for Agrobacterium transformation (Holsters *et al.* 1978). The 100 µL competent cell aliquot was thawed and 1 µg (~10 µL) of the binary DNA plasmid was added and mixed. Then, the mixture was rapidly N_2 frozen and incubate at 37 ºC for 15 minutes. Later, 900 µL of LB with Gentamycin and Tetracycline or Rifampicin was added, and the sample was incubated for 4 hours at 28 ºC with slow agitation. Finally, the cell suspension was plated in LB containing Gentamycin, Tetracycline or Rifampicin and the insertion-plasmid selection antibiotic (Kanamycin in this case). Plates were incubated in darkness at 28 ºC. Agrobacterium colonies appeared in 2-3 days.

Nicotiana benthamiana transient transformation

Agrobacteria were grown in 50 mL of LB with appropriate antibiotics for 24 hours at 28ºC with shacking for aeration. Cultures were pellet through a centrifugation at 4000 rpm for 30 minutes. Then, the cells were resuspended in 10 mL of infiltration solution (10 mM $MgCl_2$, 10 mM MES pH5.6 and 200 µM acetosyringone (from ethanol 100 mM stock)). Then, the cell suspension was incubated for 90 minutes at 28ºC and darkness with shacking. Then, Agrobacteria cell suspension should be diluted with

the infiltration solution until an A_{660} of ~0.4 (1:20 dilutions may be used for accurated results). Again, cell cultures were incubated for 90 minutes at same conditions, but with a softer shaking (50 rpm). For Agrobacteria coinfiltration, cells suspensions were mixed following this recipe for the luciferase assay: 200 µL of the Agrobacterium with the *promoter::luciferase* construction, 1 mL of Agrobacterium with the *35S::TF* construction (different Agrobacteria containing different TF may be used for TF-complex assays), and 200 µL of Agrobacterium containing the *p19* plasmid, which suppress gen silencing. Negative controls with empty plasmid (with no TFs) are necessary to be included. Then the Agrobacteria mix solution were rinsed to 10 mL with infiltration solution. Back sides *Nicotiana benthamiana* of leaves from 3-week old plants were carefully infiltrated with a plastic syringe. The cell suspension should be visualized expanding in the interior of the leave. Leaves correctly infiltrated without excessive damage should be labeled with the concrete cell mix suspensions. Different leaves (4-6) should be infiltrated with each Agrobacteria mix. Leaves were analyzed after three days.

Plant transformation

Arabidopsis plant transformation was performed by floral deep. To maximize the transformation efficiency 6-week old Col-0 plants were flower stem cut, to promote a high increase in the number of flowers in two weeks. Transformed *Agrobacterium tumefaciens* were grown overnight at 28 ºC (shaking at 225 rpm) in a 50 mL tube containing 1 mL of LB without antibiotic. The following day 10 mL of LB were added for overnight Agrobacterium propagation. The A_{660} should be around 1.7 next morning. 40 mL of water containing sucrose (5 g) and Silwet (25 µL) to have a final concentration of 10% and 0.05% respectively. Immediately, the suspension was used for dip flowers for 10 seconds. Finally, plants were 24h covered in plastic to maintain the humidity. Dry seeds formed were sterilized only with ethanol as explained in the Seed sterilization section and sown in MS media with the appropriated antibiotic.

For wheat transformation, two plasmids, *proUBI1::AtHB25* and *pAHC25*, with *BAR* and *GUSA* reporter genes (Christensen and Quail, 1996), were precipitated onto 0.6 µm gold particles. Immature scutella harvested approximately 16 days after anthesis of THA53 cultivar was transformed by particle bombardment as described previously (Pistón *et al.*, 2009). Phosphinothricin was used for selecting transgenic explants (Barro *et al.*, 1998). Putative transgenic plants were then transferred to soil and grown in the greenhouse. Transgene insertion was confirmed by detecting the presence of *pAHC25* using genomic PCR (*GUS-F* and *GUS-F*; *BAR-F* and *BAR-R*). *AtHB25* expression was confirmed through qRT-PCR.

Transformed *Agrobacterium tumefaciens* LBA4404 was used for tomato plant trans-

formation. Cotyledons from tomato plants were used for the co-culture. Explant preparation, selection and regeneration were performed as described by Ellul *et al.* (2003). Kanamycin-resistant transformants were selected and propagated in soil to obtain seeds. *AtHB25* expression was confirmed through qRT-PCR. Seeds from non-transgenic *in vitro*-regenerated tomato plants, obtained in parallel with the transgenic plants, were used as control plants.

YPD media preparation

YPD media, used for yeast strain propagation and competent cell preparation, consisted on 10 g/L of yeast extract (Conda), 20 g/L of peptone (Conda) and 20 g/L of glucose. For solid media preparation of 20 g/L of American bacteriological agar (Conda) was added. For YPDA preparation, Adenine was added from a sterile stock to a final concentration of 100 mg/L. Sterilized media for solid YPD or YPDA was distributed in 80 mm petri dishes or plates for solid media preparation.

SD media preparation

SD media, used for yeast selection, was performed by mixing pre-sterilized solutions. Glucose 20%, yeast nitrogen buffer (YNB) 7% and 0.5 M succinic acid pH 5.5 were added in a 10% of the final volume each. The rest of the volume correspond to water and added amino acids and vitamins. For solid media preparation, warm sterilized water contained American bacteriological agar for a 20 g/L final concentration. Added amino acids corresponded to the yeast amino acid expected autotrophy.

The 100X common amino acids stocks (stored at 4 ºC) are the following: Adenine (3 mg/mL), Histidine (3 mg/mL), Leucine (10 mg/mL), Lysine (10 mg/mL), Methionine (10 mg/mL), Tryptophan (10 mg/mL) and Uracil (3 mg/mL; stored at room temperature). Other amino acids should be prepared at 10 mg/mL, except Serine and Threonine that should be prepared at 30 mg/mL.

Yeast growth

Saccharomyces cerevisiae was grown for three days in darkness at 28ºC. For rapid and vigorous yeast growth yeast were grown in YPD/A. For yeast transformation and selection, yeast was grown in SD media with the necessary amino acids for auxotroph yeast strain grown, but without the autotrophy-selection amino acids. However, for Y1H screening commercially prepared DOB media was used instead of SD media.

The different amino acid preparation was also prepared and added as manufacturers indicated. For liquid media culture, the culture was shacked at 180 rpm. Cell cultures might be stored at -4ºC in agar plates for two months. For long-term storage, liquid selective SD media saturated yeast cell culture from unique confirmed cell colonies were stored with glycerol (200 µL of glycerol 80% for each 800 µL of cell culture) at -80ºC.

Yeast competent cell preparation

At the end of the day (~19:00), a 100 mL YPDA was inoculated with 100 µL of a saturated YPD yeast culture in a big flask for sufficient aeration. For homologous recombination in AtHB25 construction, the strain BY4741 (-H, -L, -M, -U) was used. For Split-Tpr, the strain CRY1 Mat*a* (-A, -H, -L, -U, -W) was used. For Y1H the strains Y187α (-A, -H, -L, -M, -W) was used for the *proAtHB25fr::HIS3* transformation (Bait) and YM4271*a* (-K -H, -L, -U, -W) was used for empty *pDEST22* transformation (negative control of Preys). Next morning, the A_{660} should be between 0.6 and 1. Cells were centrifuged (in two 50 mL sterile tubes) for five minutes at 2000 rpm and resuspended in 5 mL of LiTE (0.1 M lithium acetate, 10 mM Tris-HCl pH7.6, 1 mM EDTA) and both tubes were transfer to a single tube. Again, cells were centrifuged for five minutes at 2000 rpm but resuspended in 1 mL of LiTE. Cell suspension was incubated at 28ºC for 30 minutes. Sonicated and preheated at 95ºC 1% salmon sperm DNA (ssDNA 1%) was added as DNA carrier for the transformation (100 µL). The suspension was distributed in 100 µL aliquots. Cells are now ready to transform with a maximum efficiency. However, competent yeast can be stored at -80 ºC for later transformation. In this case, 23 µL of glycerol 80% was added in each aliquot, and the final suspension was mixed with vortex and stored at -80 ºC.

Yeast transformation

To transform yeast, the PEG-LiTE yeast transformation protocol was used. For this protocol, fresh PEG solution in needed for PEG-LiTE. It is prepared by mixing 9/10 volumes of 45% PEG-4000 and 1/10 volumes of LiTE 10X (1 M lithium acetate, 100 mM Tris-HCl pH7.6, 10 mM EDTA), for a final 40% PEG-LiTE. It is necessary to heat at 50-60 ºC and mix the solute each few minutes to completely dissolve PEG in water.

For Yeast transformation, 1 µg (~10 µL) of plasmid was added to the competent yeast aliquot. For two-plasmid transformation both plasmids were added, preferably in similar amounts. The mixture was placed in ice for 10 minutes. Then, 600 µL of the PEG-LiTE solution is added to the cells and the mixture should be gently mixed. This

mixture is incubated at 28ºC for 30 minutes and then, at 42 ºC for 10 minutes. Though a quick spin (15 seconds) cells are pelleted, and PEG-LiTE solution removed. Finally, cells are resuspended in 100 µL of sterile water and plated in a SD plate without the amino acid for the plasmid-rescue autotrophy selection. Plates were incubated at 28 ºC. Yeast colonies appeared within 2-3 days.

Yeast homologous recombination

For yeast homologous recombination, linearized plasmid (preferably without a fragment to avoid plasmid recircularization) was co-transformed with the DNA fragment with homologous flanking sequences. It is necessary 20-25 bp of identical sequence for the DNA recombination. Yeast was cultivated without the amino acid of the auxotrophy-rescue gen in the linearized plasmid. This force the yeast to recombine both DNA fragments to recircularize the plasmid and to be able to growth. A short rich-media incubation prior the selective media cultivation may improve the recombination.

Yeast matting and Y1H screening

Yeast matting was used to combine both necessary plasmids required for the Y1H screening as described by Sánchez-Montesino and Oñate-Sánchez (2017). YPAD saturated yeast cultures of Mat*a* (YM4271 yeast strain containing the *pDEST22* with the corresponding TF from the arrayed Oñate-Sánchez TF library, Madrid) and Matα (Y187 yeast strain containing the *pTUY1H* with *proAtHB25fr::HIS3*) were mixed in same proportion (100 µL each). Then, cells were cultivated for 2-3 days at 28 ºC without shacking for matting. Then, with a replicator, 200 µL of diploid selective media (DOB-L-W) was inoculated from the YPDA matting cell suspension. After two days of diploid enrichment, at 28ºC with shacking, yeast colonies were plated in diploid selective media (matting control, DOB-L-W) and the diploid and DNA-protein interaction selective media (Y1H assay, DOB-L-W-H with the pertinent 3-Amino-1,2,4-triazole (3-AT) concentration to avoid basal autoactivation, 1.5 mM in our case). The 3-AT molecule is a competitive inhibitor of the product of *HIS3*, and it serve to reduce the HIS3 effect and thus, it permits to avoid yeast growth due to *HIS3* autoactivation. The 3-AT concentration was determined by the diploid growth of yeast containing *pTUY1H* with *proAtHB25fr::HIS3* and the empty *pDEST22* plasmid in DOB-L-W-H and different 3-AT concentration (0, 1 mM, 5 mM, 15 mM, 30 mM and 40 mM). Then, the 3-AT concentration was more tightly adjusted (1 mM, 1.5 mM, 2 mM, 3mM and 5 mM), and set in the smaller 3-AT concentration that presented no-yeast growth in one week. After one week, colonies from the screening that presented grow in the DOB-L-W-H media were considered as positive screening results. The TF contained

in the *pDEST22* plasmid was thus, putatively regulating *AtHB25* expression through the interaction with the selected fragment. Positive interactions were reconfirmed in a directed assay.

Yeast plasmid DNA extraction (miniprep)

Yeast were grown overnight in 3 mL of selective SD media. After cells were pellet in a 1.5 mL microtube. Cells were resuspended in 200 µL of yeast lysis buffer (10 mM Tris-HCl pH8, 1 mM EDTA, 100 mM NaCl, 1% SDS, 2% Triton X-100). Then, 200 µL of phenol:chlorophrom:isoamilic (PCI) and around 500 µL of glass beads (0.5 mm dimeter) were added. Samples were vortex for 10 minutes. Then, 200 µL of TE buffer (10 mM Tris-HCl pH8, 1 mM EDTA) was added. Again, samples were vortex for 10 minutes. Then, samples were centrifuged for 5 minutes at 10000 rpm and the aqueous phase (~350 µL) was transferred to a new tube. Two volumes of ethanol was added. Samples were mixed and stored at -20 ºC for 15 minutes. After, samples were centrifuged at maximum speed for 5 minutes. The pellet was washed with ethanol 70% and then air dried. Finally the DNA pellet was resuspended in 20 µL of TE (with RNase). It is possible to transform *E. coli* with the extracted plasmid and perform an *E. coli* plasmid DNA extraction for a cleaner and higher quantity plasmid extraction.

Luciferase activity assay

To measure the promoter activity the Dual-Glo® Luciferase Assay System (Promega), the 5X Passive Lysis Buffer (Promega) and white 96-well Microplate LUMITRAC 200 – Ref 655075 (Greiner bio-one) are required.

Circles of ~2cm of diameter from three *Nicotiana benthamiana* transiently transfected leaves from Agrobacteria mix were extracted. Samples were frozen in liquid N_2 immediately. Then, samples were grinded and 350 µL of diluted 1X Passive Lysis Buffer was added rapidly. Samples were centrifuged for 5 minutes at 4 ºC and maximum speed. A 50 µL of supernatant was transferred to the 96-well microplate, trying to form any bubbles. To measure the Luciferase activity, 40 µL of Luciferase Assay reagent II (diluted 1:100 in 1X Passive Lysis Buffer, previously freshly prepared) was added. The plate was roughly moved in cercles to facilitate solution mixing. The luciferase activity was measured after a 10-minute incubation (for maximum peak activity), with a GloMax®-Multi Detection System (Promega). The Photomultiplier Tube (PTM) was activated 5 minutes prior to the measurement. The measurement consisted in three reads of 10 seconds recorded with a 2-minute period of read time-space. If a renilla luciferase control is available a second measurement can be used to assess the transformation efficiency and relativize measurements in a more precise way (luciferase/

renilla activity ratio). For *pYRO-4xproAtHB25fr* was not the case. To perform this measurement, after the reading 40 µL of Stop and Glo Reagent (diluted 1:100 in 1X Passive Lysis Buffer, previously freshly prepared) was added. Similarly, the plate was roughly moved in cercles to facilitate solution mixing and incubated 10 minutes prior to the renilla activity measurement. If 96-well microplate needs to be reused, it can be cleaned with dextran firstly, and then with distilled water.

Microscopy and macroscopy

Visible light seed images were acquired with Leica DM5000 microscope and with a Leica Leica DMS1000 macroscope.

Confocal microscopy

Transgenic driven-promoter GFP expression confocal imaging was performed using a Zeiss LSM 780 (www.zeiss.com) AxioObserver.Z1 laser-scanning microscope with C-Apochromat 940/1.20 W corrective water immersion objective. Argon 480 nm laser was used for GFP excitation. Lambda scan mode was used to discern between lignin autofluorescence signal and GFP in *proHB25::GUS:GFP* due to the low signal observed in during the seed maturation stage. Spectra analysis corroborated GFP signal. For *proCOG1::GUS:GFP* and *proMYB47::GFP* an emission spectrum of 500-550 nm was registered. Images were processed with the microscope software ZEN and Fiji program (Schindelin *et al.*, 2012).

SEM analysis

Gold-coated seeds were surface scanned with the ZEISS Ultra-55 Scanning Electron Microscope.

TEM analysis

Seeds were hydrated and fixed with Karnovsky's fixative solution (2.5% paraformaldehyde, 0.5% glutaraldehyde) and then were included in Lr-White resin. For the resin inclusion, samples were previously post-fixed with osmium 2% and dehydrated with ethanol washes (30%, 50%, 70% and 90%). The resin was infiltrated progressively, first with a 2:1 90% ethanol:resin mix, then with a 1:2 90% ethanol:resin, and finally in 100% resin. Samples were polymerized at 55 ºC for 24 h. Samples were sectioned

(60–90 nm) using an Ultracut Leica UC6 with a diamond blade (DIATOME). Sections for each sample were examined and imaged with a JEM-1010 (JEOL) transmission electron microscope. Suberin and cuticle layer thickness was measured with the ImageJ/Fiji program (Schindelin *et al.*, 2012).

6. References

1001 Genomes Consortium (2016) 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell* **166**, 481–491. https://doi.org/10.1016/j.cell.2016.05.063

Abrahams S., Tanner G.J., Larkin P.J. and Ashton A.R. (2002) Identification and Biochemical Characterization of Mutants in the Proanthocyanidin Pathway in Arabidopsis. *Plant Physiology* **130**, 561–576. https://doi.org/10.1104/pp.006189

Adcock I.M. and Caramori G. (2009) Chapter 31 - Transcription Factors. In *Asthma and COPD (Second Edition)*. (eds P.J. Barnes, J.M. Drazen, S.I. Rennard and N.C. Thomson), pp. 373–380. Academic Press, Oxford. https://doi.org/10.1016/B978-0-12-374001-4.00031-6

Agacka-Mołdoch M., Nagel M., Doroszewska T., Lewis R.S. and Börner A. (2015) Mapping quantitative trait loci determining seed longevity in tobacco (*Nicotiana tabacum* L.). *Euphytica* **202**, 479–486. https://doi.org/10.1007/s10681-015-1355-x

Agarwal P.K., Gupta K., Lopato S. and Agarwal P. (2017) Dehydration responsive element binding transcription factors and their applications for the engineering of stress tolerance. *Journal of Experimental Botany* **68**, 2135–2148. https://doi.org/10.1093/jxb/erx118

Al Attar L., Al-Oudat M., Safia B. and Abdul Ghani B. (2016) Aging impact on the transfer factor of 137Cs and 90Sr to lettuce and winter wheat. *Journal of Environmental Radioactivity* **164**, 19–25. https://doi.org/10.1016/j.jenvrad.2016.06.019

Alejandro, S., Rodríguez, P. L., Bellés, J. M., Yenush, L., García-Sanchez, M. J., Fernández, J. A., and Serrano, R. (2007) An Arabidopsis quiescin-sulfhydryl oxidase regulates cation homeostasis at the root symplast-xylem interface. *The EMBO Journal* **26**, 3203–3215. https://doi.org/10.1038/sj.emboj.7601757

Allocati N., Masulli M., Di Ilio C. and De Laurenzi V. (2015) Die for the community: an overview of programmed cell death in bacteria. *Cell Death & Disease* **6**, e1609. https://doi.org/10.1038/cddis.2014.570

Allorent G., Osorio S., Vu J.L., Falconet D., Jouhet J., Kuntz M., … Finazzi G. (2015) Adjustments of embryonic photosynthetic activity modulate seed fitness in Arabidopsis thaliana. *New Phytologist* **205**, 707–719. https://doi.org/10.1111/nph.13044

Almagro L., Gómez Ros L.V., Belchi-Navarro S., Bru R., Ros Barceló A. and Pedreño M.A. (2009) Class III peroxidases in plant defence reactions. *Journal of Experimental Botany* **60**, 377–390. https://doi.org/10.1093/jxb/ern277

Alpert P. (2006) Constraints of tolerance: why are desiccation-tolerant organisms so small or rare? *Journal of Experimental Biology* **209**, 1575–1584. https://doi.org/10.1242/jeb.02179

Alpert, P. and Oliver, M. J. (2002) Drying without dying. *In Desiccation and Survival in Plants: Drying Without Dying* (ed. M. Black and H. W. Prichard), pp.3 -43. Wallingford, UK: CAB International.

Andersen B., Koch B. and Scheller H.V. (1992) Structural and functional analysis of the reducing side of photosystem I. *Physiologia Plantarum,* **84**, 154–161. https://doi.org/10.1111/j.1399-3054.1992.tb08778.x

Aoki-Kinoshita K.F. and Kanehisa M. (2007) Gene Annotation and Pathway Mapping in KEGG. In *Comparative Genomics*. Methods In Molecular Biology™, (ed N.H. Bergman), pp. 71–91. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-59745-515-2_6

Arsovski A.A., Haughn G.W. and Western T.L. (2010) Seed coat mucilage cells of Arabidopsis thaliana as a model for plant cell wall research. *Plant Signaling and Behavior* **5**, 796–801. https://doi.org/10.4161/psb.5.7.11773

Asdal Å. and Guarino L. (2018) The Svalbard Global Seed Vault: 10 Years—1 Million Samples. *Biopreservation and Biobanking* **16**, 391–392. https://doi.org/10.1089/bio.2018.0025

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., … Nordborg, M. (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627–631. https://doi.org/10.1038/nature08800

Bagchi D., Garg A., Krohn R.L., Bagchi M., Tran M.X. and Stohs S.J. (1997) Oxygen free radical scavenging abilities of vitamins C and E, and a grape seed proanthocyanidin extract in vitro. *Research Communications in Molecular Pathology and Pharmacology* **95**, 179–189.

Bailly C. (2004) Active oxygen species and antioxidants in seed biology. *Seed Science Researc*, **14**, 93–107. https://doi.org/10.1079/SSR2004159

Bailly C., El-Maarouf-Bouteau H. and Corbineau F. (2008) From intracellular signaling networks to cell death: the dual role of reactive oxygen species in seed physiology. *Comptes Rendus Biologies* **331**, 806–814. https://doi.org/10.1016/j.crvi.2008.07.022

Bak, S., Beisson, F., Bishop, G., Hamberger, B., Höfer, R., Paquette, S., and Werck-Reichhart, D. (2011) Cytochromes P450. *The Arabidopsis book* **9**, e0144. https://doi.org/10.1199/tab.0144

Bakan B. and Marion D. (2017) Assembly of the Cutin Polyester: From Cells to Extracellular Cell Walls. *Plants* **6**, 57. https://doi.org/10.3390/plants6040057

Balducci L. (2010) Anemia, fatigue and aging. *Transfusion Clinique Et Biologique: Journal De La Societe Francaise De Transfusion Sanguine* **17**, 375–381. https://doi.org/10.1016/j.tracli.2010.09.169

Ballesteros D. and Walters C. (2011) Detailed characterization of mechanical properties and molecular mobility within dry seed glasses: relevance to the physiology of dry biological systems. *The Plant Journal* **68**, 607–619. https://doi.org/10.1111/j.1365-313X.2011.04711.x

Barbehenn R.V. and Peter Constabel C. (2011) Tannins in plant-herbivore interactions. *Phytochemistry* **72**, 1551–1565. https://doi.org/10.1016/j.phytochem.2011.01.040

Barro F., Cannell M.E., Lazzeri P.A. and Barcelo P. (1998) The influence of auxins on transformation of wheat and tritordeum and analysis of transgene integration patterns in transformants. *Theoretical and Applied Genetics* **97**, 684–695. https://doi.org/10.1007/s001220050944

Basak O., Demir I., Mavi K. and Matthews S. (2006) Controlled deterioration test for predicting seedling emergence and longevity of pepper (Capsicum annuum L.) seed lot. *Seed Science and Technology* **34**, 701–712. https://doi.org/10.15258/sst.2006.34.3.167

Baust J.G., Gao D. and Baust J.M. (2009) Cryopreservation. *Organogenesis* **5**, 90–96. https://doi.org/10.4161/org.5.3.10021

Baxter I., Hosmani P.S., Rus A., Lahner B., Borevitz J.O., Muthukumar B., … Salt D.E. (2009) Root Suberin Forms an Extracellular Barrier That Affects Water Relations and Mineral Nutrition in Arabidopsis. *PLOS Genetics* **5**, e1000492. https://doi.org/10.1371/journal.pgen.1000492

Beeckman T., De Rycke R., Viane R. and Inzé D. (2000) Histological Study of Seed Coat Development in Arabidopsis thaliana. *Journal of Plant Research* **113**, 139–148. https://doi.org/10.1007/PL00013924

Beisson F., Li Y., Bonaventure G., Pollard M. and Ohlrogge J.B. (2007) The Acyltransferase GPAT5 Is Required for the Synthesis of Suberin in Seed Coat and Root of Arabidopsis. *The Plant Cell* **19**, 351–368. https://doi.org/10.1105/tpc.106.048033

Belaidi A.A. and Bush A.I. (2016) Iron neurochemistry in Alzheimer's disease and Parkinson's disease: targets for therapeutics. *Journal of Neurochemistry* **139**, 179–197. https://doi.org/10.1111/jnc.13425

Belmonte M.F., Kirkbride R.C., Stone S.L., Pelletier J.M., Bui A.Q., Yeung E.C., … Harada J.J. (2013) Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E435–E444. https://doi.org/10.1073/pnas.1222061110

Bencivenga S., Colombo L. and Masiero S. (2011) Cross talk between the sporophyte and the megagametophyte during ovule development. *Sexual Plant Reproduction* **24**, 113–121. https://doi.org/10.1007/s00497-011-0162-3

Bentsink L. and Koornneef M. (2008) Seed dormancy and germination. *The Arabidopsis Book* **6**, e0119. https://doi.org/10.1199/tab.0119

Bentsink L., Hanson J., Hanhart C.J., Vries H.B., Coltrane C., Keizer P., … Koornneef M. (2010) Natural variation for seed dormancy in Arabidopsis is regulated by additive genetic and molecular pathways. *Proceedings of the National Academy of Sciences* **107**, 4264–4269. https://doi.org/10.1073/pnas.1000410107

Benveniste I., Tijet N., Adas F., Philipps G., Salaün J.-P. and Durst F. (1998) CYP86A1 from *Arabidopsis thaliana* encodes a Cytochrome P450-dependent fatty acid omega-hydroxylase. *Biochemical and Biophysical Research Communications* **243**, 688–693. https://doi.org/10.1006/bbrc.1998.8156

Berardini T.Z., Mundodi S., Reiser L., Huala E., Garcia-Hernandez M., Zhang P., … Rhee S.Y. (2004) Functional Annotation of the Arabidopsis Genome Using Controlled Vocabularies. *Plant Physiology* **135**, 745–755. https://doi.org/10.1104/pp.104.040071

Berjak P. and Pammenter N.W. (2008) From Avicennia to Zizania: Seed Recalcitrance in Perspective. Annals of Botany 101, 213–228. https://doi.org/10.1093/aob/mcm168

Bernal-Lugo I. and Leopold A.C. (1992) Changes in Soluble Carbohydrates during Seed Storage. *Plant Physiology* **98**, 1207–1210. https://doi.org/10.1104/pp.98.3.1207

Bernards M.A. (2002) Demystifying suberin. *Canadian Journal of Botany* **80**, 227–240. https://doi.org/10.1139/b02-017

Bernards M.A. and Lewis N.G. (1998) The macromolecular aromatic domain in suberized tissue: A changing paradigm. *Phytochemistry* **47**, 915–933. https://doi.org/10.1016/S0031-9422(98)80052-6

Bernards M.A., Lopez M.L., Zajicek J. and Lewis N.G. (1995) Hydroxycinnamic Acid-derived Polymers Constitute the Polyaromatic Domain of Suberin. *Journal of Biological Chemistry* **270**, 7382–7386. https://doi.org/10.1074/jbc.270.13.7382

Bernards M.A., Summerhurst D.K. and Razem F.A. (2004) Oxidases, peroxidases and hydrogen peroxide: The suberin connection. *Phytochemistry Reviews* **3**, 113–126. https://doi.org/10.1023/B:-PHYT.0000047810.10706.46

Bessire M., Chassot C., Jacquat A.-C., Humphry M., Borel S., Petétot J.M.-C., … Nawrath C. (2007) A permeable cuticle in Arabidopsis leads to a strong resistance to Botrytis cinerea. *The EMBO Journal* **26**, 2158–2168. https://doi.org/10.1038/sj.emboj.7601658

Bewley J.D. (1997) Seed Germination and Dormancy. *The Plant Cell* **9**, 1055–1066. https://doi.org/10.1105/tpc.9.7.1055

Bhargava A., Ahad A., Wang S., Mansfield S.D., Haughn G.W., Douglas C.J. and Ellis B.E. (2013) The interacting MYB75 and KNAT7 transcription factors modulate secondary cell wall deposition both in stems and seed coat in Arabidopsis. *Planta* **237**, 1199–1211. https://doi.org/10.1007/s00425-012-1821-9

Biedermann S., Mooney S. and Hellmann H. (2011) Recognition and Repair Pathways of Damaged DNA in Higher Plants. *Selected Topics in DNA Repair*. https://doi.org/10.5772/21380

Bird D., Beisson F., Brigham A., Shin J., Greer S., Jetter R., … Samuels L. (2007) Characterization of Arabidopsis ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion†. *The Plant Journal* **52**, 485–498. https://doi.org/10.1111/j.1365-313X.2007.03252.x

Bissoli G., Niñoles R., Fresquet S., Palombieri S., Bueso E., Rubio L., … Serrano R. (2012) Peptidyl-prolyl cis-trans isomerase ROF2 modulates intracellular pH homeostasis in Arabidopsis. *The Plant Journal* **70**, 704–716. https://doi.org/10.1111/j.1365-313X.2012.04921.x

Blagosklonny M.V. (2013) Aging is not programmed. *Cell Cycle* **12**, 3736–3742. https://doi.org/10.4161/cc.27188

Bleckmann A., Alter S. and Dresselhaus T. (2014) The beginning of a seed: regulatory mechanisms of double fertilization. *Frontiers in Plant Science* **5**. https://doi.org/10.3389/fpls.2014.00452

Blöchl A., Peterbauer T., Hofmann J. and Richter A. (2008) Enzymatic breakdown of raffinose oligosaccharides in pea seeds. *Planta* **228**, 99–110. https://doi.org/10.1007/s00425-008-0722-4

Boca S., Koestler F., Ksas B., Chevalier A., Leymarie J., Fekete A., … Havaux M. (2014) Arabidopsis lipocalins AtCHL and AtTIL have distinct but overlapping functions essential for lipid protection and seed longevity. *Plant, Cell & Environment* **37**, 368–381. https://doi.org/10.1111/pce.12159

Boisnard-Lorig C., Colon-Carmona A., Bauch M., Hodge S., Doerner P., Bancharel E., … Berger F. (2001) Dynamic Analyses of the Expression of the HISTONE::YFP Fusion Protein in Arabidopsis Show That Syncytial Endosperm Is Divided in Mitotic Domains. *The Plant Cell* **13**, 495–509. https://doi.org/10.1105/tpc.13.3.495

Bolger A.M., Lohse M. and Usadel B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brambilla V., Battaglia R., Colombo M., Masiero S., Bencivenga S., Kater M.M. and Colombo L. (2007) Genetic and Molecular Interactions between BELL1 and MADS Box Factors Support Ovule Development in Arabidopsis. *The Plant Cell* **19**, 2544–2556. https://doi.org/10.1105/tpc.107.051797

Bray C.M. and West C.E. (2005) DNA repair mechanisms in plants: crucial sensors and effectors for the maintenance of genome integrity. *New Phytologist* **168**, 511–528. https://doi.org/10.1111/j.1469-8137.2005.01548.x

Braybrook S.A. and Harada J.J. (2008) LECs go crazy in embryo development. *Trends in Plant Science* **13**, 624–630. https://doi.org/10.1016/j.tplants.2008.09.008

Brundrett M.C., Kendrick B. and Peterson C.A. (1991) Efficient Lipid Staining in Plant Material with Sudan Red 7B or Fluoral Yellow 088 in Polyethylene Glycol-Glycerol. *Biotechnic and Histochemistry* **66**, 111–116. https://doi.org/10.3109/10520299109110562

Bueso E., Ibañez C., Sayas E., Muñoz-Bertomeu J., Gonzalez-Guzmán M., Rodriguez P.L. and Serrano R. (2014*a*) A forward genetic approach in Arabidopsis thaliana identifies a RING-type ubiquitin ligase as a novel determinant of seed longevity. *Plant Science* **215–216**, 110–116. https://doi.org/10.1016/j.plantsci.2013.11.004

Bueso E., Muñoz-Bertomeu J., Campos F., Brunaud V., Martínez L., Sayas E., … Serrano R. (2014*b*) ARABIDOPSIS THALIANA HOMEOBOX25 Uncovers a Role for Gibberellins in Seed Longevity. *Plant Physiology* **164**, 999–1010. https://doi.org/10.1104/pp.113.232223

Bueso E., Muñoz-Bertomeu J., Campos F., Martínez C., Tello C., Martínez-Almonacid I., … Serrano R. (2016) Arabidopsis COGWHEEL1 links light perception and gibberellins with seed tolerance to deterioration. *The Plant Journal* **87**, 583–596. https://doi.org/10.1111/tpj.13220

Bueso E., Rodriguez L., Lorenzo-Orts L., Gonzalez-Guzman M., Sayas E., Muñoz-Bertomeu J., … Rodriguez P.L. (2014*c*) The single-subunit RING-type E3 ubiquitin ligase RSL1 targets PYL4 and PYR1 ABA receptors in plasma membrane to modulate abscisic acid signaling. *The Plant Journal* **80**, 1057–1071. https://doi.org/10.1111/tpj.12708

Bueso E., Serrano R., Pallás V. and Sánchez-Navarro J.A. (2017) Seed tolerance to deterioration in Arabidopsis is affected by virus infection. *Plant Physiology and Biochemistry* **116**, 1–8. https://doi.org/10.1016/j.plaphy.2017.04.020

Buitink J. and Leprince O. (2008) Intracellular glasses and seed survival in the dry state. *Comptes Rendus Biologies* **331**, 788–795. https://doi.org/10.1016/j.crvi.2008.08.002

Bürglin T.R. and Affolter M. (2016) Homeodomain proteins: an update. *Chromosoma* **125**, 497–521. https://doi.org/10.1007/s00412-015-0543-8

Burraco P., Orizaola G., Monaghan P. and Metcalfe N.B. (2020) Climate change and aging in ectotherms. *Global Change Biology*. https://doi.org/10.1111/gcb.15305

Cabiscol E., Tamarit J. and Ros J. (2014) Protein carbonylation: Proteomics, specificity and relevance to aging. *Mass Spectrometry Reviews* **33**, 21–48. https://doi.org/10.1002/mas.21375

Cao J., Schneeberger K., Ossowski S., Günther T., Bender S., Fitz J., … Weigel D. (2011) Whole-genome sequencing of multiple Arabidopsis thaliana populations. *Nature Genetics* **43**, 956–963. https://doi.org/10.1038/ng.911

Carey H.V., Andrews M.T. and Martin S.L. (2003) Mammalian Hibernation: Cellular and Molecular Responses to Depressed Metabolism and Low Temperature. *Physiological Reviews* **83**, 1153–1181. https://doi.org/10.1152/physrev.00008.2003

Carr P.R., Jansen L., Bienert S., Roth W., Herpel E., Kloor M., … Hoffmeister M. (2017) Associations of red and processed meat intake with major molecular pathological features of colorectal cancer. *European Journal of Epidemiology* **32**, 409–418. https://doi.org/10.1007/s10654-017-0275-6

Carranco R., Espinosa J.M., Prieto-Dapena P., Almoguera C. and Jordano J. (2010) Repression by an auxin/indole acetic acid protein connects auxin signaling with heat shock factor-mediated seed longevity. Proceedings of the National Academy of Sciences 107, 21908–21913. https://doi.org/10.1073/pnas.1014856107

Casson S.A. and Hetherington A.M. (2010) Environmental regulation of stomatal development. *Current Opinion in Plant Biology* **13**, 90–95. https://doi.org/10.1016/j.pbi.2015.08.005

Castrillo G., Turck F., Leveugle M., Lecharny A., Carbonero P., Coupland G., … Oñate-Sánchez L. (2011) Speeding Cis-Trans Regulation Discovery by Phylogenomic Analyses Coupled with Screenings of an Arrayed Library of Arabidopsis Transcription Factors. *PLoS ONE* **6**. https://doi.org/10.1371/journal.pone.0021524

Chai M., Zhou C., Molina I., Fu C., Nakashima J., Li G., … Wang Z.-Y. (2016) A class II KNOX gene, KNOX4, controls seed physical dormancy. *Proceedings of the National Academy of Sciences* **113**, 6997–7002. https://doi.org/10.1073/pnas.1601256113

Chang, Y. L., Li, W. Y., Miao, H., Yang, S. Q., Li, R., Wang, X., Li, W. Q., and Chen, K. M. (2016) Comprehensive genomic analysis and expression profiling of the NOX gene families under abiotic stresses and hormones in plants. *Genome Biology and Evolution* **8**, 791–810. https://doi.org/10.1093/gbe/evw035

Chassot C., Nawrath C. and Métraux J.-P. (2007) Cuticular defects lead to full immunity to a major plant pathogen. *The Plant Journal* **49**, 972–980. https://doi.org/10.1111/j.1365-313X.2006.03017.x

Châtelain E., Satour P., Laugier E., Vu B.L., Payet N., Rey P. and Montrichard F. (2013) Evidence for participation of the methionine sulfoxide reductase repair system in plant seed longevity. *Proceedings of the National Academy of Sciences* **110**, 3633–3638. https://doi.org/10.1073/pnas.1220589110

Chen H., Chu P., Zhou Y., Li Y., Liu J., Ding Y., … Huang S. (2012) Overexpression of AtOGG1, a DNA glycosylase/AP lyase, enhances seed longevity and abiotic stress tolerance in Arabidopsis. *Journal of Experimental Botany* **63**, 4107–4121. https://doi.org/10.1093/jxb/ers093

Chen Y., Cai J., Yang F.X., Zhou B. and Zhou L.R. (2015) Ascorbate peroxidase from Jatropha curcas enhances salt tolerance in transgenic Arabidopsis. *Genetics and molecular research: GMR* **14**, 4879–4889. https://doi.org/10.4238/2015.May.11.20

Chen Z., Lu H.-H., Hua S., Lin K.-H., Chen N., Zhang Y., … Chen S.-P. (2019) Cloning and overexpression of the ascorbate peroxidase gene from the yam (Dioscorea alata) enhances chilling and flood tolerance in transgenic Arabidopsis. *Journal of Plant Research* **132**, 857–866. https://doi.org/10.1007/s10265-019-01136-4

Chew W., Hrmova M. and Lopato S. (2013) Role of Homeodomain leucine zipper (HD-Zip) IV transcription factors in plant development and plant protection from deleterious environmental factors. *International Journal of Molecular Sciences* **14**, 8122–8147. https://doi.org/10.3390/ijms14048122

Chiang C.-M., Chen C.-C., Chen S.-P., Lin K.-H., Chen L.-R., Su Y.-H. and Yen H.-C. (2017) Overexpression of the ascorbate peroxidase gene from eggplant and sponge gourd enhances flood tolerance in transgenic Arabidopsis. *Journal of Plant Research* **130**, 373–386. https://doi.org/10.1007/s10265-016-0902-4

Chinnusamy V., Zhu J. and Zhu J.-K. (2007) Cold stress regulation of gene expression in plants. *Trends in Plant Science* **12**, 444–451. https://doi.org/10.1016/j.tplants.2007.07.002

Chown S.L. and Nicolson S. (2004) *Insect Physiological Ecology: Mechanisms and Patterns*. Oxford University Press, Oxford, New York.

Christensen A.H. and Quail P.H. (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Research* **5**, 213–218. https://doi.org/10.1007/BF01969712

Clerkx E.J.M., El-Lithy M.E., Vierling E., Ruys G.J., Vries H.B.-D., Groot S.P.C., … Koornneef M. (2004*a*) Analysis of Natural Allelic Variation of Arabidopsis Seed Germination and Seed Longevity Traits between the Accessions Landsberg erecta and Shakdara, Using a New Recombinant Inbred Line Population. *Plant Physiology* **135**, 432–443. https://doi.org/10.1104/pp.103.036814

Clerkx E.J.M., Vries H.B.-D., Ruys G.J., Groot S.P.C. and Koornneef M. (2004*b*) Genetic differences in seed longevity of various Arabidopsis mutants. *Physiologia Plantarum* **121**, 448–461. https://doi.org/10.1111/j.0031-9317.2004.00339.x

Coen O., Fiume E., Xu W., Vos D.D., Lu J., Pechoux C., … Magnani E. (2017) Developmental patterning of the sub-epidermal integument cell layer in Arabidopsis seeds. *Development* **144**, 1490–1497. https://doi.org/10.1242/dev.146274

Coen O., Lu J., Xu W., De Vos D., Péchoux C., Domergue F., … Magnani E. (2019) Deposition of a cutin apoplastic barrier separating seed maternal and zygotic tissues. BMC Plant Biology 19. https://doi.org/10.1186/s12870-019-1877-9

Colman R.J., Anderson R.M., Johnson S.C., Kastman E.K., Kosmatka K.J., Beasley T.M., … Weindruch R. (2009) Caloric restriction delays disease onset and mortality in rhesus monkeys. *Science (New York, N.Y.)* **325**, 201–204. https://doi.org/10.1126/science.1173635

Compagnon V., Diehl P., Benveniste I., Meyer D., Schaller H., Schreiber L., … Pinot F. (2009) CYP86B1 Is Required for Very Long Chain ω-Hydroxyacid and α,ω-Dicarboxylic Acid Synthesis in Root and Seed Suberin Polyester. *Plant Physiology* **150**, 1831–1843. https://doi.org/10.1104/pp.109.141408

Considine M.J. and Foyer C.H. (2014) Redox regulation of plant development. *Antioxidants and Redox Signaling* **21**, 1305–1326. https://doi.org/10.1089/ars.2013.5665

Conti B. (2008) Considerations on Temperature, Longevity and Aging. *Cellular and molecular life sciences : CMLS* **65**, 1626–1630. https://doi.org/10.1007/s00018-008-7536-1

Contreras S., Bennett M.A., Metzger J.D., Tay D. and Nerson H. (2009) Red to Far-red Ratio During Seed Development Affects Lettuce Seed Germinability and Longevity. HortScience 44, 130–134. https://doi.org/10.21273/HORTSCI.44.1.130

Cosio C. and Dunand C. (2009) Specific functions of individual class III peroxidase genes. *Journal of Experimental Botany* **60**, 391–408. https://doi.org/10.1093/jxb/ern318

Costa M.-C.D., Artur M.A.S., Maia J., Jonkheer E., Derks M.F.L., Nijveen H., … Hilhorst H.W.M. (2017) A footprint of desiccation tolerance in the genome of Xerophyta viscosa. *Nature Plants* **3**, 1–10. https://doi.org/10.1038/nplants.2017.38

Costanzo E., Trehin C. and Vandenbussche M. (2014) The role of WOX genes in flower development. *Annals of Botany* **114**, 1545–1553. https://doi.org/10.1242/dev.030049

Cottle W. and Kolattukudy P.E. (1982) Abscisic Acid Stimulation of Suberization: Induction of Enzymes and Deposition of Polymeric Components and Associated Waxes in Tissue Cultures of Potato Tuber. *Plant Physiology* **70**, 775–780. https://doi.org/10.1104/pp.70.3.775

Creff A., Brocard L. and Ingram G. (2015) A mechanically sensitive cell layer regulates the physical properties of the Arabidopsis seed coat. *Nature Communications* **6**, 6382. https://doi.org/10.1038/ncomms7382

Currey D.R. (1965) An Ancient Bristlecone Pine Stand in Eastern Nevada. *Ecology* **46**, 564–566. https://doi.org/10.2307/1934900

Curtin, S. J., Tiffin, P., Guhlin, J., Trujillo, D. I., Burghart, L. T., Atkins, P., Baltes, N. J., Denny, R., Voytas, D. F., Stupar, R. M., and Young, N. D. (2017) Validating genome-wide association candidates controlling quantitative variation in nodulation. *Plant Physiology* **173**, 921–931. https://doi.org/10.1104/pp.16.01923

Curtis, M. D., and Grossniklaus, U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiology* **133**, 462–469. https://doi.org/10.1104/pp.103.027979

Czechowski T., Stitt M., Altmann T., Udvardi M.K. and Scheible W.-R. (2005) Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis. *Plant Physiology* **139**, 5–17. https://doi.org/10.1104/pp.105.063743

D'Auria J.C. (2006) Acyltransferases in plants: a good time to be BAHD. *Current Opinion in Plant Biology* **9**, 331–340. https://doi.org/10.1016/j.pbi.2006.03.016

da Costa J.P., Vitorino R., Silva G.M., Vogel C., Duarte A.C. and Rocha-Santos T. (2016) A synopsis on aging—Theories, mechanisms and future prospects. Aging Research Reviews 29, 90–112. https://doi.org/10.1016/j.arr.2016.06.005

Dai A. (2006) Recent Climatology, Variability, and Trends in Global Surface Humidity. *Recent Climatology, Variability, and Trends in Global Surface Humidity*. **19**, 3589–3606. https://doi.org/10.1175/JCLI3816.1

de Folter S., Immink R.G.H., Kieffer M., Pařenicová L., Henz S.R., Weigel D., … Angenent G.C. (2005) Comprehensive Interaction Map of the Arabidopsis MADS Box Transcription Factors. The

Plant Cell 17, 1424–1433. https://doi.org/10.1105/tpc.105.031831

De Giorgi J., Piskurewicz U., Loubery S., Utz-Pugin A., Bailly C., Mène-Saffrané L. and Lopez-Molina L. (2015) An Endosperm-Associated Cuticle Is Required for Arabidopsis Seed Viability, Dormancy and Early Control of Germination. *PLoS genetics* **11**, e1005708. https://doi.org/10.1371/journal.pgen.1005708

De Magalhaes Filho C.D, Henriquez B., Seah N.E., Evans R.M., Lapierre L.R. and Dillin A. (2018) Visible light reduces C. elegans longevity. Nature Communications 9, 1–13. https://doi.org/10.1038/s41467-018-02934-5

de Souza Vidigal D., Willems L., van Arkel J., Dekkers B.J.W., Hilhorst H.W.M. and Bentsink L. (2016) Galactinol as marker for seed longevity. *Plant Science* **246**, 112–118. https://doi.org/10.1016/j.plantsci.2016.02.015

Dean B.B. (1989) Deposition of Aliphatic Suberin Monomers and Associated Alkanes during Aging of Solanum tuberosum L. Tuber Tissue at Different Temperatures 1. *Plant Physiology* **89**, 1021–1023.

Dean G., Cao Y., Xiang D., Provart N.J., Ramsay L., Ahad A., … Haughn G. (2011) Analysis of Gene Expression Patterns during Seed Coat Development in Arabidopsis. *Molecular Plant* **4**, 1074–1091. https://doi.org/10.1093/mp/ssr040

Dean G.H., Zheng H., Tewari J., Huang J., Young D.S., Hwang Y.T., … Haughn G.W. (2007) The Arabidopsis MUM2 Gene Encodes a β-Galactosidase Required for the Production of Seed Coat Mucilage with Correct Hydration Properties. *The Plant Cell* **19**, 4007–4021. https://doi.org/10.1105/tpc.107.050609

Debeaujon I., Léon-Kloosterziel K.M. and Koornneef M. (2000) Influence of the Testa on Seed Dormancy, Germination, and Longevity in Arabidopsis. *Plant Physiology* **122**, 403–414. https://doi.org/10.1104/pp.122.2.403

Debeaujon I., Peeters A.J., Léon-Kloosterziel K.M. and Koornneef M. (2001) The TRANSPARENT TESTA12 gene of Arabidopsis encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. *The Plant Cell* **13**, 853–871. https://doi.org/10.1105/tpc.13.4.853

Delmas F., Sankaranarayanan S., Deb S., Widdup E., Bournonville C., Bollier N., … Samuel M.A. (2013) ABI3 controls embryo degreening through Mendel's I locus. Proceedings of the National Academy of Sciences of the United States of America 110, E3888-3894. https://doi.org/10.1073/pnas.1308114110

Demir I. and Mavi K. (2008) Controlled Deterioration and Accelerated Aging Tests to Estimate the Relative Storage Potential of Cucurbit Seed Lots. *HortScience* **43**, 1544–1548. https://doi.org/10.21273/HORTSCI.43.5.1544

Demonsais L., Utz-Pugin A., Loubéry S. and Lopez-Molina L. (2020) Identification of tannic cell walls at the outer surface of the endosperm upon Arabidopsis seed coat rupture. *The Plant Journal* **104**, 567–580. https://doi.org/10.1111/tpj.14994

Denay G., Creff A., Moussu S., Wagnon P., Thévenin J., Gérentes M.-F., … Ingram G. (2014) Endosperm breakdown in Arabidopsis requires heterodimers of the basic helix-loop-helix proteins

ZHOUPI and INDUCER OF CBP EXPRESSION 1. *Development* **141**, 1222–1227. https://doi.org/10.1242/dev.103531

Ditta G., Pinyopich A., Robles P., Pelaz S. and Yanofsky M.F. (2004) The SEP4 Gene of *Arabidopsis thaliana* functions in floral organ and meristem identity. *Current Biology* **14**, 1935–1940. https://doi.org/10.1016/j.cub.2004.10.028

Dixon D.P. and Edwards R. (2010) Glutathione transferases. *The Arabidopsis Book* **2010**. https://doi.org/10.1199/tab.0131

Dixon D.P., Davis B.G. and Edwards R. (2002) Functional divergence in the glutathione transferase superfamily in plants. Identification of two classes with putative functions in redox homeostasis in *Arabidopsis thaliana*. *Journal of Biological Chemistry* **277**, 30859–30869. https://doi.org/10.1074/jbc.M202919200

Dixon R.A., Xie D.-Y. and Sharma S.B. (2005) Proanthocyanidins – a final frontier in flavonoid research? *New Phytologist* **165**, 9–28. https://doi.org/10.1111/j.1469-8137.2004.01217.x

Dizdaroglu M. (1992) Oxidative damage to DNA in mammalian chromatin. *Mutation Research/DNAging* **275**, 331–342. https://doi.org/10.1016/0921-8734(92)90036-O

Do T.H.T., Martinoia E. and Lee Y. (2018) Functions of ABC transporters in plant growth and development. *Current Opinion in Plant Biology* **41**, 32–38. https://doi.org/10.1016/j.pbi.2017.08.003

Dollemore, D. (2002) Aging Under the Microscope: A Biological Quest. National Institutes of Health, National Institute on Aging, Office of Communications and Public Liaison.

Domergue F. and Kosma D.K. (2017) Occurrence and Biosynthesis of Alkyl Hydroxycinnamates in Plant Lipid Barriers. *Plants* **6**, 25. https://doi.org/10.3390/plants6030025

Domergue F., Vishwanath S.J., Joubès J., Ono J., Lee J.A., Bourdon M., … Rowland O. (2010) Three Arabidopsis Fatty Acyl-Coenzyme A Reductases, FAR1, FAR4, and FAR5, Generate Primary Fatty Alcohols Associated with Suberin Deposition. *Plant Physiology* **153**, 1539–1554. https://doi.org/10.1104/pp.110.158238

Donohue K., Dorn L., Griffith C., Kim E., Aguilera A., Polisetty C.R. and Schmitt J. (2005) The evolutionary ecology of seed germination of Arabidopsis thaliana: variable natural selection on germination timing. *Evolution; International Journal of Organic Evolution* **59**, 758–770. https://doi.org/10.1111/j.0014-3820.2005.tb01751.x

Dresselhaus T., Sprunck S. and Wessel G.M. (2016) Fertilization Mechanisms in Flowering Plants. *Current Biology* **26**, R125–R139. https://doi.org/10.1016/j.cub.2015.12.032

Driver T., Trivedi D.K., McIntosh O.A., Dean A.P., Goodacre R. and Pittman J.K. (2017) Two Glycerol-3-Phosphate Dehydrogenases from Chlamydomonas Have Distinct Roles in Lipid Metabolism. *Plant Physiology* **174**, 2083–2097. https://doi.org/10.1104/pp.17.00491

Dröge-Laser W., Snoek B.L., Snel B. and Weiste C. (2018) The Arabidopsis bZIP transcription factor family—an update. *Current Opinion in Plant Biology* **45**, 36–49. https://doi.org/10.1016/j.pbi.2018.05.001

Duan H. and Schuler M.A. (2005) Differential Expression and Evolution of the Arabidopsis CYP86A

Subfamily. *Plant Physiology* **137**, 1067–1081. https://doi.org/10.1104/pp.104.055715

Dubos C., Stracke R., Grotewold E., Weisshaar B., Martin C. and Lepiniec L. (2010) MYB transcription factors in *Arabidopsis*. *Trends in Plant Science* **15**, 573–581. https://doi.org/10.1016/j.tplants.2010.06.005

Duek P.D. and Fankhauser C. (2005) bHLH class transcription factors take centre stage in phytochrome signalling. *Trends in Plant Science* **10**, 51–54. https://doi.org/10.1016/j.tplants.2004.12.005

Duffy P.H., Feuers R.J., Leakey J.A., Nakamura K.D., Turturro A. and Hart R.W. (1989) Effect of chronic caloric restriction on physiological variables related to energy metabolism in the male Fischer 344 rat. *Mechanisms of Aging and Development* **48**, 117–133. https://doi.org/10.1016/0047-6374(89)90044-4

Duroux L. and Welinder K.G. (2003) The Peroxidase Gene Family in Plants: A Phylogenetic Overview. *Journal of Molecular Evolution* **57**, 397–407. https://doi.org/10.1007/s00239-003-2489-3

Eastmond P.J. (2004) Glycerol-insensitive Arabidopsis mutants: gli1 seedlings lack glycerol kinase, accumulate glycerol and are more resistant to abiotic stress. *The Plant Journal* **37**, 617–625. https://doi.org/10.1111/j.1365-313X.2003.01989.x

Ehlers K., Bhide A.S., Tekleyohans D.G., Wittkop B., Snowdon R.J. and Becker A. (2016) The MADS Box Genes ABS, SHP1, and SHP2 Are Essential for the Coordination of Cell Divisions in Ovule and Seed Coat Development and for Endosperm Formation in Arabidopsis thaliana. *PLOS ONE* **11**, e0165075. https://doi.org/10.1371/journal.pone.0165075

Ellis R.H. and Roberts E.H. (1980) Improved Equations for the Prediction of Seed Longevity. *Annals of Botany* **45**, 13–30. https://doi.org/10.1093/oxfordjournals.aob.a085797

Ellul P., Garcia-Sogo B., Pineda B., Ríos G., Roig L. and Moreno V. (2003) The ploidy level of transgenic plants in Agrobacterium-mediated transformation of tomato cotyledons (Lycopersicon esculentum L.Mill.) is genotype and procedure dependent. *Theoretical and Applied Genetics* **106**, 231–238. https://doi.org/10.1007/s00122-002-0928-y

El-Maarouf-Bouteau H. and Bailly C. (2008) Oxidative signaling in seed germination and dormancy. *Plant Signaling and Behavior* **3**, 175–182. https://doi.org/10.4161/psb.3.3.5539

Endo A., Tatematsu K., Hanada K., Duermeyer L., Okamoto M., Yonekura-Sakakibara K., … Nambara E. (2012) Tissue-Specific Transcriptome Analysis Reveals Cell Wall Metabolism, Flavonol Biosynthesis and Defense Responses are Activated in the Endosperm of Germinating Arabidopsis thaliana Seeds. *Plant and Cell Physiology* **53**, 16–27. https://doi.org/10.1093/pcp/pcr171

Espelie K.E., Davis R.W. and Kolattukudy P.E. (1980) Composition, ultrastructure and function of the cutin- and suberin-containing layers in the leaf, fruit peel, juice-sac and inner seed coat of grapefruit (Citrus paradisi Macfed.). *Planta* **149**, 498–511. https://doi.org/10.1007/BF00385755

Ewald C.Y. (2018) Redox Signaling of NADPH Oxidases Regulates Oxidative Stress Responses, Immunity and Aging. *Antioxidants* **7**, 130. https://doi.org/10.3390/antiox7100130

Fabre G., Garroum I., Mazurek S., Daraspe J., Mucciolo A., Sankar M., … Nawrath C. (2016) The

ABCG transporter PEC1/ABCG32 is required for the formation of the developing leaf cuticle in Arabidopsis. *New Phytologist* **209**, 192–201. https://doi.org/10.1111/nph.13608

Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., Borghi, L., Ditta, G., Yanofsky, M. F., Kater, M. M., and Colombo, L. (2003) MADS-box protein complexes control carpel and ovule development in *Arabidopsis*. *The Plant Cell* **15**, 2603–2611. https://doi.org/10.1105/tpc.015123

Fedi F., O'Neill C.M., Menard G., Trick M., Dechirico S., Corbineau F., … Penfield S. (2017) Awake1, an ABC-Type Transporter, Reveals an Essential Role for Suberin in the Control of Seed Dormancy. *Plant Physiology* **174**, 276–283. https://doi.org/10.1104/pp.16.01556

Fenollosa E., Jené L. and Munné-Bosch S. (2020) A rapid and sensitive method to assess seed longevity through accelerated aging in an invasive plant species. *Plant Methods* **16**, 64. https://doi.org/10.1186/s13007-020-00607-3

Fich E.A., Segerson N.A. and Rose J.K.C. (2016) The Plant Polyester Cutin: Biosynthesis, Structure, and Biological Roles. *Annual Review of Plant Biology* **67**, 207–233. https://doi.org/10.1146/annurev-arplant-043015-111929

Figueiredo D.D., Batista R.A., Roszak P.J., Hennig L. and Köhler C. Auxin production in the endosperm drives seed coat development in Arabidopsis. *eLife* **5**. https://doi.org/10.7554/eLife.20542

Finch-Savage W.E. and Leubner-Metzger G. (2006) Seed dormancy and the control of germination. *New Phytologist* **171**, 501–523. https://doi.org/10.1111/j.1469-8137.2006.01787.x

Fitch W.M. (1970) Distinguishing homologous from analogous proteins. Systematic Zoology 19, 99–113.

Fleming M.B., Hill L.M. and Walters C. (2019) The kinetics of aging in dry-stored seeds: a comparison of viability loss and RNA degradation in unique legacy seed collections. *Annals of Botany* **123**, 1133–1146. https://doi.org/10.1093/aob/mcy217

Fleming M.B., Richards C.M. and Walters C. (2017) Decline in RNA integrity of dry-stored soybean seeds correlates with loss of germination potential. *Journal of Experimental Botany* **68**, 2219–2230. https://doi.org/10.1093/jxb/erx100

Focks N. and Benning C. (1998) wrinkled1: A Novel, Low-Seed-Oil Mutant of Arabidopsis with a Deficiency in the Seed-Specific Regulation of Carbohydrate Metabolism. *Plant Physiology* **118**, 91–101. https://doi.org/10.1104/pp.118.1.91

Fontana L. and Partridge L. (2015) Promoting Health and Longevity through Diet: From Model Organisms to Humans. *Cell* **161**, 106–118. https://doi.org/10.1016/j.cell.2015.02.020

Foster C.E., Martin T.M. and Pauly M. (2010) Comprehensive Compositional Analysis of Plant Cell Walls (Lignocellulosic biomass) Part I: Lignin. *Journal of Visualized Experiments : JoVE*. https://doi.org/10.3791/1745

Fourquin C., Beauzamy L., Chamot S., Creff A., Goodrich J., Boudaoud A. and Ingram G. (2016) Mechanical stress mediated by both endosperm softening and embryo growth underlies endo-

sperm elimination in Arabidopsis seeds. *Development* **143**, 3300–3305. https://doi.org/10.1242/dev.137224

Foyer C.H. and Noctor G. (2011) Ascorbate and glutathione: the heart of the redox hub. *Plant Physiology* **155**, 2–18. https://doi.org/10.1104/pp.110.16756

Francoz E., Ranocha P., Nguyen-Kim H., Jamet E., Burlat V. and Dunand C. (2015) Roles of cell wall peroxidases in plant development. *Phytochemistry* **112**, 15–21. https://doi.org/10.1016/j.phytochem.2014.07.020

Franich R.A., Gadgil P.D. and Shain L. (1983) Fungistatic effects of Pinus radiata needle epicuticular fatty and resin acids on Dothistroma pini. *Physiological Plant Pathology* **23**, 183–195. https://www.cabdirect.org/cabdirect/abstract/19841396494

Franke R. and Schreiber L. (2007) Suberin — a biopolyester forming apoplastic plant interfaces. *Current Opinion in Plant Biology* **10**, 252–259. https://doi.org/10.1016/j.pbi.2007.04.004

Franke R., Briesen I., Wojciechowski T., Faust A., Yephremov A., Nawrath C. and Schreiber L. (2005) Apoplastic polyesters in Arabidopsis surface tissues – A typical suberin and a particular cutin. *Phytochemistry* **66**, 2643–2658. https://doi.org/10.1016/j.phytochem.2005.09.027

Franke R., Höfer R., Briesen I., Emsermann M., Efremova N., Yephremov A. and Schreiber L. (2009) The DAISY gene from Arabidopsis encodes a fatty acid elongase condensing enzyme involved in the biosynthesis of aliphatic suberin in roots and the chalaza-micropyle region of seeds. *The Plant Journal* **57**, 80–95. https://doi.org/10.1111/j.1365-313X.2008.03674.x

Gaiser J.C., Robinson-Beers K. and Gasser C.S. (1995) The Arabidopsis SUPERMAN Gene Mediates Asymmetric Growth of the Outer Integument of Ovules. *The Plant Cell* **7**, 333–345. https://doi.org/10.1105/tpc.7.3.333

Gallagher T.L. and Gasser C.S. (2008) Independence and Interaction of Regions of the INNER NO OUTER Protein in Growth Control during Ovule Development. *Plant Physiology* **147**, 306–315. https://doi.org/10.1104/pp.107.11460

Gan, X., Stegle, O., Behr, J., Steffen, J. G., Drewe, P., Hildebrand, K. L., … Mott, R. (2011) Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. *Nature,* **477**, 419–423.

Garí E., Piedrafita L., Aldea M. and Herrero E. (1997) A Set of Vectors with a Tetracycline-Regulatable Promoter System for Modulated Gene Expression in Saccharomyces cerevisiae. *Yeast* **13**, 837–848. https://doi.org/10.1002/(SICI)1097-0061(199707)13:9<837::AID-YEA145>3.0.CO;2-T

Gendrel A.-V., Lippman Z., Martienssen R. and Colot V. (2005) Profiling histone modification patterns in plants using genomic tiling microarrays. *Nature Methods* **2**, 213–218. https://doi.org/10.1038/nmeth0305-213

Gladyshev V. N. (2014) The free radical theory of aging is dead. Long live the damage theory!. *Antioxidants and Redox Signaling* **20**, 727–731. https://doi.org/10.1089/ars.2013.5228

Glenister P.H. and Thornton C.E. (2000) Cryoconservation—archiving for the future. *Mammalian Genome* **11**, 565–571. https://doi.org/10.4161/org.5.3.10021

Goff L., Trapnell C. and Kelley D. (2020) *cummeRbund: Analysis, exploration, manipulation, and vi-*

sualization of Cufflinks high-throughput sequencing data. Bioconductor version: Release (3.11). https://doi.org/10.18129/B9.bioc.cummeRbund

Goffman, F. D., Alonso, A. P., Schwender, J., Shachar-Hill, Y., and Ohlrogge, J. B. (2005) Light enables a very high efficiency of carbon storage in developing embryos of rapeseed. *Plant physiology* **138**, 2269–2279. https://doi.org/10.1104/pp.105.063628

Goldberg R.B., Paiva G. de and Yadegari R. (1994) Plant Embryogenesis: Zygote to Seed. Science 266, 605–614. https://doi.org/10.1126/science.266.5185.605

Golz J.F., Allen P.J., Li S.F., Parish R.W., Jayawardana N.U., Bacic A. and Doblin M.S. (2018) Layers of regulation – Insights into the role of transcription factors controlling mucilage production in the Arabidopsis seed coat. *Plant Science* **272**, 179–192. https://doi.org/10.1016/j.plantsci.2018.04.021

Gou J.-Y., Yu X.-H. and Liu C.-J. (2009) A hydroxycinnamoyltransferase responsible for synthesizing suberin aromatics in Arabidopsis. *Proceedings of the National Academy of Sciences* **106**, 18855–18860. https://doi.org/10.1073/pnas.0905555106

Gou M., Hou G., Yang H., Zhang X., Cai Y., Kai G. and Liu C.-J. (2017) The MYB107 Transcription Factor Positively Regulates Suberin Biosynthesis. *Plant Physiology* **173**, 1045–1058. https://doi.org/10.1104/pp.16.01614

Gou X. and Li J. (2012) Activation Tagging. In *Plant Signalling Networks: Methods and Protocols*. Methods in Molecular Biology, (eds Z.-Y. Wang and Z. Yang), pp. 117–133. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-61779-809-2_9

Graça J. (2015) Suberin: the biopolyester at the frontier of plants. *Frontiers in Chemistry* **3**. https://doi.org/10.3389/fchem.2015.00062

Grefen C., Donald N., Hashimoto K., Kudla J., Schumacher K. and Blatt M.R. (2010) A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. *The Plant Journal* **64**, 355–365. https://doi.org/10.1111/j.1365-313X.2010.04322.x

Griffen L.R., Wilczek A.M. and Bazzaz F.A. (2004) UV-B affects within-seed biomass allocation and chemical provisioning. *New Phytologist* **162**, 167–171. https://doi.org/10.1111/j.1469-8137.2004.01013.x

Griffith M., Huner N.P.A., Espelie K.E. and Kolattukudy P.E. (1985) Lipid polymers accumulate in the epidermis and mestome sheath cell walls during low temperature development of winter rye leaves. *Protoplasma* **125**, 53–64. https://doi.org/10.1007/BF01297350

Grimm, D. G., Roqueiro, D., Salomé, P. A., Kleeberger, S., Greshake, B., Zhu, W., … Borgwardt, K. M. (2017) easyGWAS: A Cloud-Based Platform for Comparing the Results of Genome-Wide Association Studies. *The Plant Cell* **29**(1), 5–19. https://doi.org/10.1105/tpc.16.00551

Groot, S. P., Surki, A. A., de Vos, R. C., and Kodde, J. (2012) Seed storage at elevated partial pressure of oxygen, a fast method for analysing seed aging under dry conditions. *Annals of botany* **110**(6), 1149–1159. https://doi.org/10.1093/aob/mcs198

Guidetti R. and Jönsson K.I. (2002) Long-term anhydrobiotic survival in semi-terrestrial micrometazoans. *Journal of Zoology* **257**, 181–187. https://doi.org/10.1017/S095283690200078X

Haldrup A., Lunde C. and Scheller H.V. (2003) Arabidopsis thaliana Plants Lacking the PSI-D Subunit of Photosystem I Suffer Severe Photoinhibition, Have Unstable Photosystem I Complexes, and Altered Redox Homeostasis in the Chloroplast Stroma. *Journal of Biological Chemistry* **278**, 33276–33283. https://doi.org/10.1074/jbc.M305106200

Hand S.C., Menze M.A., Toner M., Boswell L. and Moore D. (2011) LEA proteins during water stress: not just for plants anymore. *Annual Review of Physiology* **73**, 115–134. https://doi.org/10.1146/annurev-physiol-012110-142203

Harman D. (1956) Aging: A Theory Based on Free Radical and Radiation Chemistry. *Journal of Gerontology,* **11**, 298–300. https://doi.org/10.1093/geronj/11.3.298

Harman, G.E. and Mattick, L.R. (1976) Association of lipid oxidation with seed aging and death. *Nature,* **260,** 323–324. https://doi.org/10.1038/260323a0

Harrington J.F. (1970) Seed and pollen storage for conservation of plant gene resources. *In Genetic Resources in Plants—Their Exploration andConservation.* pp. 501–521.Blackwell Scientific, Oxford.

Haughn G. and Chaudhury A. (2005) Genetic analysis of seed coat development in Arabidopsis. *Trends in Plant Science* **10**, 472–477. https://doi.org/10.1016/j.tplants.2005.08.005

Havaux M., Eymery F., Porfirova S., Rey P. and Dörmann P. (2005) Vitamin E Protects against Photoinhibition and Photooxidative Stress in Arabidopsis thaliana. *The Plant Cell* **17**, 3451–3469. https://doi.org/10.1105/tpc.105.037036

Hay A. and Tsiantis M. (2010) KNOX genes: versatile regulators of plant development and diversity. *Development (Cambridge, England)* **137**, 3153–3165. https://doi.org/10.1242/dev.030049

Hay F.R., Valdez R., Lee J.-S. and Sta. Cruz P.C. (2019) Seed longevity phenotyping: recommendations on research methodology. *Journal of Experimental Botany* **70**, 425–434. https://doi.org/10.1093/jxb/ery358

He H., de Souza Vidigal D., Snoek L.B., Schnabel S., Nijveen H., Hilhorst H. and Bentsink L. (2014) Interaction between parental environment and genotype affects plant and seed performance in Arabidopsis. *Journal of Experimental Botany* **65**, 6603–6615. https://doi.org/10.1093/jxb/eru378

Hedden P. and Thomas S.G. (2012) Gibberellin biosynthesis and its regulation. *The Biochemical Journal* **444**, 11–25. https://doi.org/10.1042/BJ20120245

Heinz S., Benner C., Spann N., Bertolino E., Lin Y.C., Laslo P., … Glass C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular Cell* **38**, 576–589. https://doi.org/10.1016/j.molcel.2010.05.004

Herrero J., Fernández-Pérez F., Yebra T., Novo-Uzal E., Pomar F., Pedreño M.Á., … Zapata J.M. (2013) Bioinformatic and functional characterization of the basic peroxidase 72 from Arabidopsis thaliana involved in lignin biosynthesis. *Planta* **237**, 1599–1612. https://doi.org/10.1007/s00425-013-1865-5

Hoekstra F.A., Golovina E.A. and Buitink J. (2001) Mechanisms of plant desiccation tolerance. *Trends in Plant Science* **6**, 431–438. https://doi.org/10.1016/S1360-1385(01)02052-0

Höfer, R., Briesen, I., Beck, M., Pinot, F., Schreiber, L., and Franke, R. (2008) The Arabidopsis cytochrome P450 CYP86A1 encodes a fatty acid omega-hydroxylase involved in suberin monomer biosynthesis. *Journal of experimental botany* **59**(9), 2347–2360. https://doi.org/10.1093/jxb/ern101

Holland H.D. (2006) The oxygenation of the atmosphere and oceans. *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**, 903–915. https://doi.org/10.1098/rstb.2006.1838

Holland P.W.H. (2013) Evolution of homeobox genes. *WIREs Developmental Biology* **2**, 31–45. https://doi.org/10.1002/wdev.78

Holoch D. and Moazed D. (2015) RNA-mediated epigenetic regulation of gene expression. *Nature Reviews Genetics* **16**, 71–84. https://doi.org/10.1038/nrg3863

Holsters M., de Waele D., Depicker A., Messens E., van Montagu M. and Schell J. (1978) Transfection and transformation of Agrobacterium tumefaciens. *Molecular and General Genetics MGG* **163**, 181–187. https://doi.org/10.1007/BF00267408

Honda S., Ishii N., Suzuki K. and Matsuo M. (1993) Oxygen-dependent perturbation of life span and aging rate in the nematode. *Journal of Gerontology* **48**, B57-61. https://doi.org/10.1093/geronj/48.2.b57

Hong L., Brown J., Segerson N.A., Rose J.K.C. and Roeder A.H.K. (2017) CUTIN SYNTHASE 2 Maintains Progressively Developing Cuticular Ridges in Arabidopsis Sepals. *Molecular Plant* **10**, 560–574. https://doi.org/10.1016/j.molp.2017.01.002

Hong-Bo S., Zong-Suo L. and Ming-An S. (2005) LEA proteins in higher plants: Structure, function, gene expression and regulation. *Colloids and Surfaces B: Biointerfaces* **45**, 131–135. https://doi.org/10.1016/j.colsurfb.2005.07.017

Hooper C.M., Castleden I.R., Tanz S.K., Aryamanesh N. and Millar A.H. (2017) SUBA4: the interactive data analysis centre for Arabidopsis subcellular protein locations. *Nucleic Acids Research* **45**, D1064–D1074. https://doi.org/10.1093/nar/gkw1041

Horton, M. W., Hancock, A. M., Huang, Y. S., Toomajian, C., Atwell, S., Auton, A., … Bergelson, J. (2012) Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel. *Nature Genetics* **44**(2), 212–216. https://doi.org/10.1038/ng.1042

Hu Y., Han X., Yang M., Zhang M., Pan J. and Yu D. (2019) The Transcription Factor INDUCER OF CBF EXPRESSION1 Interacts with ABSCISIC ACID INSENSITIVE5 and DELLA Proteins to Fine-Tune Abscisic Acid Signaling during Seed Germination in Arabidopsis. *The Plant Cell* **31**, 1520–1538. https://doi.org/10.1105/tpc.18.00825

Huang W., Loganantharaj R., Schroeder B., Fargo D. and Li L. (2013) PAVIS: a tool for Peak Annotation and Visualization. *Bioinformatics* **29**, 3097–3099. https://doi.org/10.1093/bioinformatics/btt520

Hugouvieux V., Silva C.S., Jourdain A., Stigliani A., Charras Q., Conn V., … Zubieta C. (2018) Tetramerization of MADS family transcription factors SEPALLATA3 and AGAMOUS is required

for floral meristem determinacy in Arabidopsis. *Nucleic Acids Research* **46**, 4966–4977. https://doi.org/10.1093/nar/gky205

Hundertmark M., Buitink J., Leprince O. and Hincha D.K. (2011) The reduction of seed-specific dehydrins reduces seed longevity in Arabidopsis thaliana. *Seed Science Research* **21**, 165–173. https://doi.org/10.1017/S0960258511000079

Ihnatowicz A., Pesaresi P., Varotto C., Richly E., Schneider A., Jahns P., … Leister D. (2004) Mutants for photosystem I subunit D of Arabidopsis thaliana: effects on photosynthesis, photosystem I stability and expression of nuclear genes for chloroplast functions. *The Plant Journal* **37**, 839–852. https://doi.org/10.1111/j.1365-313X.2004.02011.x

Imaizumi T., Schultz T.F., Harmon F.G., Ho L.A. and Kay S.A. (2005) FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. *Science (New York, N.Y.)* **309**, 293–297. https://doi.org/10.1126/science.1110586

Inoue H., Nojima H. and Okayama H. (1990) High efficiency transformation of Escherichia coli with plasmids. *Gene* **96**, 23–28. https://doi.org/10.1016/0378-1119(90)90336-p

Ioannidis, J., Thomas, G. and Daly, M. (2009) Validating, augmenting and refining genome-wide association signals. *Nat Rev Genet* **10,** 318–329. https://doi.org/10.1038/nrg2544

ISTA (2018) Seed vigour testing. In *International rules for seed testing 2018*. International Seed Testing Association, Bassersdorf.

Jakobson L., Lindgren L.O., Verdier G., Laanemets K., Brosché M., Beisson F. and Kollist H. (2016) BODYGUARD is required for the biosynthesis of cutin in Arabidopsis. *New Phytologist* **211**, 614–626. https://doi.org/10.1111/nph.13924

Jaspard, E., Macherel, D., and Hunault, G. (2012) Computational and statistical analyses of amino acid usage and physico-chemical properties of the twelve late embryogenesis abundant protein classes. *PloS One* **7**, e36968. https://doi.org/10.1371/journal.pone.0036968

Jeevan Kumar S.P., Rajendra Prasad S., Banerjee R. and Thammineni C. (2015) Seed birth to death: dual functions of reactive oxygen species in seed physiology. *Annals of Botany* **116**, 663–668. https://doi.org/10.1093/aob/mcv098

Jefferson R.A., Kavanagh T.A. and Bevan M.W. (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO journal* **6**, 3901–3907.

Jeffree C.E. (2007) The Fine Structure of the Plant Cuticle. In *Annual Plant Reviews Volume 23: Biology of the Plant Cuticle*. pp. 11–125. John Wiley and Sons, Ltd. https://doi.org/10.1002/9780470988718.ch2

Jetha, K., Theißen, G., and Melzer, R. (2014) Arabidopsis SEPALLATA proteins differ in cooperative DNA-binding during the formation of floral quartet-like complexes. *Nucleic acids research* **42**(17), 10927–10942. https://doi.org/10.1093/nar/gku755

Jha P., Ochatt S.J. and Kumar V. (2020) WUSCHEL: a master regulator in plant growth signaling. *Plant Cell Reports* **39**, 431–444. https://doi.org/10.1007/s00299-020-02511-5

Jia L., Wu Q., Ye N., Liu R., Shi L., Xu W., … Zhang J. (2012) Proanthocyanidins Inhibit Seed Germination by Maintaining a High Level of Abscisic Acid in Arabidopsis thalianaF. *Journal of Integrative Plant Biology* **54**, 663–673. https://doi.org/10.1111/j.1744-7909.2012.01142.x

Jiang W., Lee J., Jin Y.-M., Qiao Y., Piao R., Jang S.M., … Koh H.-J. (2011) Identification of QTLs for seed germination capability after various storage periods using two RIL populations in rice. *Molecules and Cells* **31**, 385–392. https://doi.org/10.1007/s10059-011-0049-z

Job C., Rajjou L., Lovigny Y., Belghazi M. and Job D. (2005) Patterns of Protein Oxidation in Arabidopsis Seeds and during Germination. *Plant Physiology* **138**, 790–802. https://doi.org/10.1104/pp.105.062778

Johnson T.E. (1990) Increased life-span of age-1 mutants in Caenorhabditis elegans and lower Gompertz rate of aging. *Science* **249**, 908–912. https://doi.org/10.1126/science.2392681

Jothi R., Balaji S., Wuster A., Grochow J.A., Gsponer J., Przytycka T.M., … Babu M.M. (2009) Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. *Molecular Systems Biology* **5**, 294. https://doi.org/10.1038/msb.2009.52

Kai K., Hashidzume H., Yoshimura K., Suzuki H., Sakurai N., Shibata D. and Ohta D. (2009) Metabolomics for the characterization of cytochromes P450-dependent fatty acid hydroxylation reactions in *Arabidopsis*. *Plant Biotechnology* **26**, 175–182. https://doi.org/10.5511/plantbiotechnology.26.175

Kanaoka M.M., Pillitteri L.J., Fujii H., Yoshida Y., Bogenschutz N.L., Takabayashi J., … Torii K.U. (2008) SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to arabidopsis stomatal differentiation. *The Plant Cell* **20**, 1775–1785. https://doi.org/10.1105/tpc.108.060848

Kaufmann K., Anfang N., Saedler H. and Theissen G. (2005) Mutant analysis, protein–protein interactions and subcellular localization of the Arabidopsis Bsister (ABS) protein. *Molecular Genetics and Genomics* **274**, 103–118. https://doi.org/10.1007/s00438-005-0010-y

Kaufmann K., Muiño J.M., Østerås M., Farinelli L., Krajewski P. and Angenent G.C. (2010) Chromatin immunoprecipitation (ChIP) of plant transcription factors followed by sequencing (ChIP-SEQ) or hybridization to whole genome arrays (ChIP-CHIP). *Nature Protocols* **5**, 457–472. https://doi.org/10.1038/nprot.2009.244

Kawade K., Li Y., Koga H., Sawada Y., Okamoto M., Kuwahara A., … Hirai M.Y. (2018) The cytochrome P450 CYP77A4 is involved in auxin-mediated patterning of the Arabidopsis thaliana embryo. *Development* **145**. https://doi.org/10.1242/dev.168369

Keil G., Cummings E. and de Magalhães J.P. (2015) Being cool: how body temperature influences aging and longevity. *Biogerontology* **16**, 383–397. https://doi.org/10.1007/s10522-015-9571-2

Kendall S.L., Hellwege A., Marriot P., Whalley C., Graham I.A. and Penfield S. (2011) Induction of Dormancy in Arabidopsis Summer Annuals Requires Parallel Regulation of DOG1 and Hormone Metabolism by Low Temperature and CBF Transcription Factors. *The Plant Cell* **23**, 2568–2580. https://doi.org/10.1105/tpc.111.087643

Kilian J., Whitehead D., Horak J., Wanke D., Weinl S., Batistic O., … Harter K. (2007) The AtGen-

Express global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *The Plant Journal* **50**, 347–363. https://doi.org/10.1111/j.1365-313X.2007.03052.x

Kim D., Langmead B. and Salzberg S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* **12**, 357–360. https://doi.org/10.1038/nmeth.3317

Kim H., Lee S.B., Kim H.J., Min M.K., Hwang I. and Suh M.C. (2012) Characterization of Glycosylphosphatidylinositol-Anchored Lipid Transfer Protein 2 (LTPG2) and Overlapping Function between LTPG/LTPG1 and LTPG2 in Cuticular Wax Export or Accumulation in Arabidopsis thaliana. *Plant and Cell Physiology* **53**, 1391–1403. https://doi.org/10.1093/pcp/pcs083

Kim Y.-C., Nakajima M., Nakayama A. and Yamaguchi I. (2005) Contribution of Gibberellins to the Formation of Arabidopsis Seed Coat Through Starch Degradation. *Plant and Cell Physiology* **46**, 1317–1325. https://doi.org/10.1093/pcp/pci141

Kirkwood T.B.L. (2005) Understanding the Odd Science of Aging. *Cell* **120**, 437–447. https://doi.org/10.1016/j.cell.2005.01.027

Klucher K.M., Chow H., Reiser L. and Fischer R.L. (1996) The AINTEGUMENTA gene of Arabidopsis required for ovule and female gametophyte development is related to the floral homeotic gene APETALA2. *The Plant Cell* **8**, 137–153. https://doi.org/10.1105/tpc.8.2.137

Kocsy G. (2015) Die or survive? Redox changes as seed viability markers. *Plant, Cell & Environment* **38**, 1008–1010. https://doi.org/10.1111/pce.12515

Kolattukudy P.E. (2001) Polyesters in Higher Plants. In *Biopolyesters*. Advances in Biochemical Engineering/Biotechnology, (eds W. Babel and A. Steinbüchel), pp. 1–49. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-74075-6_11

Kolattukudy P.E. and Espelie K.E. (1989) Chemistry, Biochemistry, and Function of Suberin and Associated Waxes. In *Natural Products of Woody Plants: Chemicals Extraneous to the Lignocellulosic Cell Wall*. Springer Series in Wood Science, (ed J.W. Rowe), pp. 304–367. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-40021-4_1

Koncz C. and Schell J. (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. *Molecular and General Genetics MGG* **204**, 383–396. https://doi.org/10.1007/BF00331014

Koornneef M. (1990) Mutations affecting the testa colour in Arabidopsis. *Arabidopsis Information Service* **27**, 1–4.

Korte, A., and Farlow, A. (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant methods* **9**, 29. https://doi.org/10.1186/1746-4811-9-29

Kosma D.K., Molina I., Ohlrogge J.B. and Pollard M. (2012) Identification of an Arabidopsis Fatty Alcohol:Caffeoyl-Coenzyme A Acyltransferase Required for the Synthesis of Alkyl Hydroxycinnamates in Root Waxes. *Plant Physiology* **160**, 237–248. https://doi.org/10.1104/pp.112.201822

Kosma D.K., Murmu J., Razeq F.M., Santos P., Bourgault R., Molina I. and Rowland O. (2014) At-

MYB41 activates ectopic suberin synthesis and assembly in multiple plant species and cell types. *The Plant Journal* **80**, 216–229. https://doi.org/10.1111/tpj.12624

Koster K.L. and Leopold A.C. (1988) Sugars and desiccation tolerance in seeds. *Plant Physiology* **88**, 829–832. https://doi.org/10.1104/pp.88.3.829

Kotak S., Vierling E., Bäumlein H. and Koskull-Döring P. von (2007) A Novel Transcriptional Cascade Regulating Expression of Heat Stress Proteins during Seed Development of Arabidopsis. *The Plant Cell* **19**, 182–195. https://doi.org/10.1105/tpc.106.048165

Kouvari M., Tyrovolas S. and Panagiotakos D.B. (2016) Red meat consumption and healthy aging: A review. *Maturitas* **84**, 17–24. https://doi.org/10.1016/j.maturitas.2015.11.006

Kranner I. (1998) Determination of Glutathione, Glutathione Disulphide and Two Related Enzymes, Glutathione Reductase and Glucose-6-Phosphate Dehydrogenase, in Fungal and Plant Cells. In *Mycorrhiza Manual*. Springer Lab Manual, (ed A. Varma), pp. 227–241. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-60268-9_15

Kranner I., Birtić S., Anderson K.M. and Pritchard H.W. (2006) Glutathione half-cell reduction potential: A universal stress marker and modulator of programmed cell death? *Free Radical Biology and Medicine* **40**, 2155–2165. https://doi.org/10.1016/j.freeradbiomed.2006.02.013

Krause K.-H. (2007) Aging: A revisited theory based on free radicals generated by NOX family NADPH oxidases. *Experimental Gerontology* **42**, 256–262. https://doi.org/10.1016/j.exger.2006.10.011

Krebs M., Held K., Binder A., Hashimoto K., Herder G.D., Parniske M., … Schumacher K. (2012) FRET-based genetically encoded sensors allow high-resolution live cell imaging of Ca2+ dynamics. *The Plant Journal* **69**, 181–192. https://doi.org/10.1111/j.1365-313X.2011.04780.x

Kunieda T., Shimada T., Kondo M., Nishimura M., Nishitani K. and Hara-Nishimura I. (2013) Spatiotemporal Secretion of PEROXIDASE36 Is Required for Seed Coat Mucilage Extrusion in Arabidopsis. *The Plant Cell* **25**, 1355–1367. https://doi.org/10.1105/tpc.113.110072

Kunst L., Klenz J.E., Martinez-Zapater J. and Haughn G.W. (1989) AP2 Gene Determines the Identity of Perianth Organs in Flowers of Arabidopsis thaliana. *The Plant Cell* **1**, 1195–1208. https://doi.org/10.1105/tpc.1.12.1195

Kuo M.-H. and Allis C.D. (1999) In Vivo Cross-Linking and Immunoprecipitation for Studying Dynamic Protein:DNA Associations in a Chromatin Environment. *Methods* **19**, 425–433. https://doi.org/10.1006/meth.1999.0879

Kurdyukov S., Faust A., Nawrath C., Bär S., Voisin D., Efremova N., … Yephremov A. (2006*a*) The Epidermis-Specific Extracellular BODYGUARD Controls Cuticle Development and Morphogenesis in Arabidopsis. *The Plant Cell* **18**, 321–339. https://doi.org/10.1105/tpc.105.036079

Kurdyukov S., Faust A., Trenkamp S., Bär S., Franke R., Efremova N., … Yephremov A. (2006*b*) Genetic and biochemical evidence for involvement of HOTHEAD in the biosynthesis of long-chain α-,ω-dicarboxylic fatty acids and formation of extracellular matrix. *Planta* **224**, 315–329. https://doi.org/10.1007/s00425-005-0215-7

Kurek K., Plitta-Michalak B. and Ratajczak E. (2019) Reactive Oxygen Species as Potential Drivers of the Seed Aging Process. *Plants* **8**, 174. https://doi.org/10.3390/plants8060174

Landjeva S., Lohwasser U. and Börner A. (2010) Genetic mapping within the wheat D genome reveals QTL for germination, seed vigour and longevity, and early seedling growth. *Euphytica* **171**, 129–143. https://doi.org/10.1007/s10681-009-0016-3

Lashbrooke J., Cohen H., Levy-Samocha D., Tzfadia O., Panizel I., Zeisler V., … Aharoni A. (2016) MYB107 and MYB9 Homologs Regulate Suberin Deposition in Angiosperms. *The Plant Cell* **28**, 2097–2116. https://doi.org/10.1105/tpc.16.00490

Lata C. and Prasad M. (2011) Role of DREBs in regulation of abiotic stress responses in plants. *Journal of Experimental Botany* **62**, 4731–4748. https://doi.org/10.1093/jxb/err210

Latchman D.S. (1993) Transcription factors: an overview. *International Journal of Experimental Pathology* **74**, 417–422.

Le B.H., Cheng C., Bui A.Q., Wagmaister J.A., Henry K.F., Pelletier J., … Goldberg R.B. (2010) Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 8063–8070. https://doi.org/10.1073/pnas.1003530107

Lee B., Henderson D.A. and Zhu J.-K. (2005) The Arabidopsis cold-responsive transcriptome and its regulation by ICE1. *The Plant Cell* **17**, 3155–3175. https://doi.org/10.1105/tpc.105.035568

Lee S.B. and Suh M.-C. (2018) Disruption of glycosylphosphatidylinositol-anchored lipid transfer protein 15 affects seed coat permeability in Arabidopsis. *The Plant Journal* **96**, 1206–1217. https://doi.org/10.1111/tpj.14101

Lee S.B., Go Y.S., Bae H.-J., Park J.H., Cho S.H., Cho H.J., … Suh M.C. (2009*a*) Disruption of Glycosylphosphatidylinositol-Anchored Lipid Transfer Protein Gene Altered Cuticular Lipid Composition, Increased Plastoglobules, and Enhanced Susceptibility to Infection by the Fungal Pathogen Alternaria brassicicola. *Plant Physiology* **150**, 42–54. https://doi.org/10.1104/pp.109.137745

Lee S.-B., Jung S.-J., Go Y.-S., Kim H.-U., Kim J.-K., Cho H.-J., … Suh M.-C. (2009*b*) Two Arabidopsis 3-ketoacyl CoA synthase genes, KCS20 and KCS2/DAISY, are functionally redundant in cuticular wax and root suberin biosynthesis, but differentially controlled by osmotic stress. *The Plant Journal* **60**, 462–475. https://doi.org/10.1111/j.1365-313X.2009.03973.x

Lee Y., Rubio M.C., Alassimone J. and Geldner N. (2013) A Mechanism for Localized Lignin Deposition in the Endodermis. *Cell* **153**, 402–412. https://doi.org/10.1016/j.cell.2013.02.045

Leide J., Hildebrandt U., Hartung W., Riederer M. and Vogg G. (2012) Abscisic acid mediates the formation of a suberized stem scar tissue in tomato fruits. *New Phytologist* **194**, 402–415. https://doi.org/10.1111/j.1469-8137.2011.04047.x

Leiser S.F., Begun A. and Kaeberlein M. (2011) HIF-1 modulates longevity and healthspan in a temperature-dependent manner. *Aging Cell* **10**, 318–326. https://doi.org/10.1111/j.1474-9726.2011.00672.x

Lendzian K.J. (1982) Gas permeability of plant cuticles. *Planta* **155**, 310–315. https://doi.org/10.1007/BF00429457

Leon-Kloosterziel K.M., Keijzer C.J. and Koornneef M. (1994) A Seed Shape Mutant of Arabidopsis That Is Affected in Integument Development. *The Plant Cell* **6**, 385–392. https://doi.org/10.1105/tpc.6.3.385

Leprince O., Pellizzaro A., Berriri S. and Buitink J. (2017) Late seed maturation: drying without dying. *Journal of Experimental Botany* **68**, 827–841. https://doi.org/10.1093/jxb/erw363

Lewis N.G. and Yamamoto E. (1990) Lignin: Occurrence, Biogenesis and Biodegradation. *Annual Review of Plant Physiology and Plant Molecular Biology* **41**, 455–496. https://doi.org/10.1146/annurev.pp.41.060190.002323

Li D., Fu X., Guo L., Huang Z., Li Y., Liu Y., … Liu X. (2016) FAR-RED ELONGATED HYPOCOTYL3 activates SEPALLATA2 but inhibits CLAVATA3 to regulate meristem determinacy and maintenance in Arabidopsis. *Proceedings of the National Academy of Sciences* **113**, 9375–9380. https://doi.org/10.1073/pnas.1602960113

Li H. and Durbin R. (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li J. and Berger F. (2012) Endosperm: food for humankind and fodder for scientific discoveries. *New Phytologist* **195**, 290–305. https://doi.org/10.1111/j.1469-8137.2012.04182.x

Li J., Ou-Lee T.M., Raba R., Amundson R.G. and Last R.L. (1993) Arabidopsis Flavonoid Mutants Are Hypersensitive to UV-B Irradiation. *The Plant Cell* **5**, 171–179. https://doi.org/10.1105/tpc.5.2.171

Li S.F., Allen P.J., Napoli R.S., Browne R.G., Pham H. and Parish R.W. (2020) MYB–bHLH–TTG1 Regulates Arabidopsis Seed Coat Biosynthesis Pathways Directly and Indirectly via Multiple Tiers of Transcription Factors. *Plant and Cell Physiology* **61**, 1005–1018. https://doi.org/10.1093/pcp/pcaa027

Li T., Zhang Y., Wang J., Xu D., Yin Z., Chen H., … Shi X. (2018) All-cause mortality risk associated with long-term exposure to ambient PM2·5 in China: a cohort study. *The Lancet. Public Health* **3**, e470–e477. https://doi.org/10.1016/S2468-2667(18)30144-0

Li X., Lin H., Zhang W., Zou Y., Zhang J., Tang X. and Zhou J.-M. (2005) Flagellin induces innate immunity in nonhost interactions that is suppressed by Pseudomonas syringae effectors. *Proceedings of the National Academy of Sciences* **102**, 12990–12995. https://doi.org/10.1073/pnas.0502425102

Li Y., Beisson F., Koo A.J.K., Molina I., Pollard M. and Ohlrogge J. (2007) Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proceedings of the National Academy of Sciences* **104**, 18339–18344. https://doi.org/10.1073/pnas.0706984104

Liang M., Davis E., Gardner D., Cai X. and Wu Y. (2006) Involvement of AtLAC15 in lignin synthesis in seeds and in root elongation of Arabidopsis. *Planta* **224**, 1185. https://doi.org/10.1007/s00425-006-0300-6

Li-Beisson Y., Pollard M., Sauveplane V., Pinot F., Ohlrogge J. and Beisson F. (2009) Nanoridges that characterize the surface morphology of flowers require the synthesis of cutin polyester. *Proceedings of the National Academy of Sciences* **106**, 22008–22013. https://doi.org/10.1073/

pnas.0909090106

Li-Beisson Y., Shorrosh B., Beisson F., Andersson M.X., Arondel V., Bates P.D., … Ohlrogge J. (2013) Acyl-Lipid Metabolism. *The Arabidopsis Book* **2013**. https://doi.org/10.1199/tab.0161

Lim C.J., Hwang J.E., Chen H., Hong J.K., Yang K.A., Choi M.S., … Lim C.O. (2007) Over-expression of the Arabidopsis DRE/CRT-binding transcription factor DREB2C enhances thermotolerance. Biochemical and Biophysical Research Communications 362, 431–436. https://doi.org/10.1016/j.bbrc.2007.08.007

Lindahl T. (1993) Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715. https://doi.org/10.1038/362709a0

Liu X. and Hou X. (2018) Antagonistic Regulation of ABA and GA in Metabolism and Signaling Pathways. *Frontiers in Plant Science* **9**. https://doi.org/10.3389/fpls.2018.00251

Liu, Y., You, S., Taylor-Teeples, M., Li, W. L., Schuetz, M., Brady, S. M., and Douglas, C. J. (2014) BEL1-LIKE HOMEODOMAIN6 and KNOTTED ARABIDOPSIS THALIANA7 interact and regulate secondary cell wall formation via repression of REVOLUTA. *The Plant cell* **26**(12), 4843–4861. https://doi.org/10.1105/tpc.114.128322

Long, Q., Rabanal, F. A., Meng, D., Huber, C. D., Farlow, A., Platzer, A., … Nordborg, M. (2013) Massive genomic variation and strong selection in Arabidopsis thaliana lines from Sweden. *Nature genetics* **45**(8), 884–890. https://doi.org/10.1038/ng.2678

Longo V.D., Mitteldorf J. and Skulachev V.P. (2005) Programmed and altruistic aging. *Nature Reviews Genetics* **6**, 866–872. https://doi.org/10.1038/nrg1706

Lopez-Molina L., Mongrand S., McLachlin D.T., Chait B.T. and Chua N.-H. (2002) ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. *The Plant Journal: For Cell and Molecular Biology* **32**, 317–328. https://doi.org/10.1046/j.1365-313x.2002.01430.x

López-Otín, C., Galluzzi L., Freije J.M.P., Madeo F. and Kroemer G. (2016) Metabolic Control of Longevity. *Cell* **166**, 802–821. https://doi.org/10.1016/j.cell.2016.07.031

Loubéry S., Giorgi J.D., Utz-Pugin A., Demonsais L. and Lopez-Molina L. (2018) A Maternally Deposited Endosperm Cuticle Contributes to the Physiological Defects of transparent testa Seeds. *Plant Physiology* **177**, 1218–1233. https://doi.org/10.1104/pp.18.00416

Lowenson J.D. and Clarke S. (1992) Recognition of D-aspartyl residues in polypeptides by the erythrocyte L-isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis. *The Journal of Biological Chemistry* **267**, 5985–5995.

Lü S., Song T., Kosma D.K., Parsons E.P., Rowland O. and Jenks M.A. (2009) Arabidopsis CER8 encodes LONG-CHAIN ACYL-COA SYNTHETASE 1 (LACS1) that has overlapping functions with LACS2 in plant wax and cutin synthesis. *The Plant Journal* **59**, 553–564. https://doi.org/10.1111/j.1365-313X.2009.03892.x

Lurin, C., Andrés, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyère, C., … Small, I. (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role

in organelle biogenesis. *The Plant cell* **16**(8), 2089–2103. https://doi.org/10.1105/tpc.104.022236

MacAlister, C., Ohashi-Ito, K. and Bergmann, D. (2007) Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. *Nature* **445,** 537–540 https://doi.org/10.1038/nature05491

MacGregor D.R., Kendall S.L., Florance H., Fedi F., Moore K., Paszkiewicz K., … Penfield S. (2015) Seed production temperature regulation of primary dormancy occurs through control of seed coat phenylpropanoid metabolism. *New Phytologist* **205**, 642–652. https://doi.org/10.1111/nph.13090

MacGregor D.R., Zhang N., Iwasaki M., Chen M., Dave A., Lopez-Molina L. and Penfield S. (2019) ICE1 and ZOU determine the depth of primary seed dormancy in Arabidopsis independently of their role in endosperm development. *The Plant Journal* **98**, 277–290. https://doi.org/10.1111/tpj.14211

Maeda K., Houjyou Y., Komatsu T., Hori H., Kodaira T. and Ishikawa A. (2010) Nonhost resistance to Magnaporthe oryzae in Arabidopsis thaliana. *Plant Signaling & Behavior* **5**, 755–756. https://doi.org/10.4161/psb.5.6.11770

Mandel T., Candela H., Landau U., Asis L., Zelinger E., Carles C.C. and Williams L.E. (2016) Differential regulation of meristem size, morphology and organization by the ERECTA, CLAVATA and class III HD-ZIP pathways. *Development* **143**, 1612–1622. https://doi.org/10.1242/dev.129973

Mansuy D. (1998) The great diversity of reactions catalyzed by cytochromes P450. *Comparative Biochemistry and Physiology. Part C, Pharmacology, Toxicology and Endocrinology* **121**, 5–14. https://doi.org/10.1016/s0742-8413(98)10026-9

Mao Z. and Sun W. (2015) Arabidopsis seed-specific vacuolar aquaporins are involved in maintaining seed longevity under the control of ABSCISIC ACID INSENSITIVE 3. *Journal of Experimental Botany* **66**, 4781–4794. https://doi.org/10.1093/jxb/erv244

Marino D., Dunand C., Puppo A. and Pauly N. (2012. A burst of plant NADPH oxidases. *Trends in Plant Science* **17**, 9–15. https://doi.org/10.1016/j.tplants.2011.10.001

Martin L.B.B., Romero P., Fich E.A., Domozych D.S. and Rose J.K.C. (2017) Cuticle Biosynthesis in Tomato Leaves Is Developmentally Regulated by Abscisic Acid. *Plant Physiology* **174**, 1384–1398. https://doi.org/10.1104/pp.17.00387

Martin M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12. https://doi.org/10.14806/ej.17.1.200

Martinez D.E. (1998) Mortality Patterns Suggest Lack of Senescence in Hydra. *Experimental Gerontology* **33**, 217–225. https://doi.org/10.1016/S0531-5565(97)00113-7

Mattson M.P. (2003) *Energy Metabolism and Lifespan Determination*. Amsterdam ; Boston.

Matzke K. and Riederer M. (1990) The composition of the cutin of the caryopses and leaves ofTriticum aestivum L. *Planta* **182**, 461. https://doi.org/10.1007/BF02411400

Mayer U., Ruiz R.A.T., Berleth T., Miséra S. and Jürgens G. (1991) Mutations affecting body organization in the Arabidopsis embryo. *Nature* **353**, 402–407. https://doi.org/10.1038/353402a0

McAbee J.M., Hill T.A., Skinner D.J., Izhaki A., Hauser B.A., Meister R.J., … Gasser C.S. (2006) ABERRANT TESTA SHAPE encodes a KANADI family member, linking polarity determination to separation and growth of Arabidopsis ovule integuments. *The Plant Journal* **46**, 522–531. 10.1111/j.1365-313X.2006.02717.x

Meister R.J., Kotow L.M. and Gasser C.S. (2002) SUPERMAN attenuates positive INNER NO OUTER autoregulation to maintain polar development of Arabidopsis ovule outer integuments. *Development (Cambridge, England)* **129**, 4281–4289. https://doi.org/10.1242/dev.129.18.4281

Merati G. and Zanetti G. (1987) Chemical cross-linking of ferredoxin to spinach thylakoids: Evidence for two independent binding sites of ferredoxin to the membrane. *FEBS Letters,* **215**, 37–40. https://doi.org/10.1016/0014-5793(87)80109-6

Miller G., Shulaev V. and Mittler R. (2008) Reactive oxygen signaling and abiotic stress. *Physiologia Plantarum,* **133**, 481–489. https://doi.org/10.1111/j.1399-3054.2008.01090.x

Milne I., Stephen G., Bayer M., Cock P.J.A., Pritchard L., Cardle L., … Marshall D. (2013) Using Tablet for visual exploration of second-generation sequencing data. *Briefings in Bioinformatics* **14**, 193–202. https://doi.org/10.1093/bib/bbs012

Miquel J., Lundgren P.R., Bensch K.G. and Atlan H. (1976) Effects of temperature on the life span, vitality and fine structure of Drosophila melanogaster. *Mechanisms of Aging and Development* **5**, 347–370. https://doi.org/10.1016/0047-6374(76)90034-8

Mishra S., Jha A.B. and Dubey R.S. (2011) Arsenite treatment induces oxidative stress, upregulates antioxidant system, and causes phytochelatin synthesis in rice seedlings. *Protoplasma* **248**, 565–577. https://doi.org/10.1007/s00709-010-0210-0

Miura K., Jin J.B., Lee J., Yoo C.Y., Stirm V., Miura T., … Hasegawa P.M. (2007) SIZ1-Mediated Sumoylation of ICE1 Controls CBF3/DREB1A Expression and Freezing Tolerance in Arabidopsis. *The Plant Cell* **19**, 1403–1414. https://doi.org/10.1105/tpc.106.048397

Mizzotti, C., Ezquer, I., Paolo, D., Rueda-Romero, P., Guerra, R. F., Battaglia, R., Rogachev, I., Aharoni, A., Kater, M. M., Caporali, E., and Colombo, L. (2014) SEEDSTICK is a master regulator of development and metabolism in the Arabidopsis seed coat. *PLoS genetics* **10**(12), e1004856. https://doi.org/10.1371/journal.pgen.1004856

Moïse J.A., Han S., Gudynaitę-Savitch L., Johnson D.A. and Miki B.L.A. (2005) Seed coats: Structure, development, composition, and biotechnology. *In Vitro Cellular and Developmental Biology - Plant* **41**, 620–644. https://doi.org/10.1079/IVP2005686

Molina I., Bonaventure G., Ohlrogge J. and Pollard M. (2006) The lipid polyester composition of Arabidopsis thaliana and Brassica napus seeds. *Phytochemistry* **67**, 2597–2610. https://doi.org/10.1016/j.phytochem.2006.09.011

Molina I., Li-Beisson Y., Beisson F., Ohlrogge J.B. and Pollard M. (2009) Identification of an Arabidopsis Feruloyl-Coenzyme A Transferase Required for Suberin Synthesis. *Plant Physiology* **151**, 1317–1328. https://doi.org/10.1104/pp.109.144907

Molina I., Ohlrogge J.B. and Pollard M. (2008) Deposition and localization of lipid polyester in devel-

oping seeds of Brassica napus and Arabidopsis thaliana. *The Plant Journal* **53**, 437–449. https://doi.org/10.1111/j.1365-313X.2007.03348.x

Mondoni A., Orsenigo S., Donà M., Balestrazzi A., Probert R.J., Hay F.R., … Abeli T. (2014) Environmentally induced transgenerational changes in seed longevity: maternal and genetic influence. *Annals of Botany* **113**, 1257–1263. https://doi.org/10.1093/aob/mcu046

Mönke G., Seifert M., Keilwagen J., Mohr M., Grosse I., Hähnel U., … Altschmied L. (2012) Toward the identification and regulation of the Arabidopsis thaliana ABI3 regulon. *Nucleic Acids Research* **40**, 8240–8254. https://doi.org/10.1093/nar/gks594

Moreira-Vilar F.C., Siqueira-Soares R. de C., Finger-Teixeira A., Oliveira D.M. de, Ferro A.P., Rocha G.J. da, … Ferrarese-Filho O. (2014) The Acetyl Bromide Method Is Faster, Simpler and Presents Best Recovery of Lignin in Different Herbaceous Tissues than Klason and Thioglycolic Acid Methods. *PLOS ONE* **9**, e110000. https://doi.org/10.1371/journal.pone.0110000

Mudgett M.B., Lowenson J.D. and Clarke S. (1997) Protein repair L-isoaspartyl methyltransferase in plants. Phylogenetic distribution and the accumulation of substrate proteins in aged barley seeds. *Plant Physiology* **115**, 1481–1489. https://doi.org/10.1104/pp.115.4.1481

Muhammad I.I., Kong S.L., Akmar Abdullah S.N. and Munusamy U. (2020) RNA-seq and ChIP-seq as Complementary Approaches for Comprehension of Plant Transcriptional Regulatory Mechanism. *International Journal of Molecular Sciences* **21**, 167. https://doi.org/10.3390/ijms21010167

Mullin W.J. and Xu W. (2001) Study of Soybean Seed Coat Components and Their Relationship to Water Absorption. *Journal of Agricultural and Food Chemistry* **49**, 5331–5335. https://doi.org/10.1021/jf010303s

Murata N. and Tasaka Y. (1997) Glycerol-3-phosphate acyltransferase in plants. *Biochimica Et Biophysica Acta* **1348**, 10–16. https://doi.org/10.1016/s0005-2760(97)00115-x

Murgia I., Giacometti S., Balestrazzi A., Paparella S., Pagliano C. and Morandini P. (2015) Analysis of the transgenerational iron deficiency stress memory in Arabidopsis thaliana plants. Frontiers in Plant Science 6, 745. https://doi.org/10.3389/fpls.2015.00745

Murphy T.M., Asard H. and Cross A.R. (1998) Possible Sources of Reactive Oxygen during the Oxidative Burst in Plants. *In Plasma Membrane Redox Systems and their Role in Biological Stress and Disease* (eds H. Asard, A. Bérczi and R.J. Caubergs), pp. 215–246. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-94-017-2695-5_9

Nagel M., Kranner I., Neumann K., Rolletschek H., Seal C.E., Colville L., … Börner A. (2015) Genome-wide association mapping and biochemical markers reveal that seed aging and longevity are intricately affected by genetic background and developmental and environmental conditions in barley. *Plant, Cell & Environment* **38**, 1011–1022. https://doi.org/10.1111/pce.12474

Nagel M., Seal C.E., Colville L., Rodenstein A., Un S., Richter J., … Kranner I. (2019) Wheat seed aging viewed through the cellular redox environment and changes in pH. *Free Radical Research* **53**, 641–654. https://doi.org/10.1080/10715762.2019.1620226

Nagel, M., Kodde, J., Pistrick, S., Mascher, M., Börner, A., and Groot, S. P. (2016) Barley Seed Aging: Genetics behind the Dry Elevated Pressure of Oxygen Aging and Moist Controlled Deterioration.

Frontiers in plant science **7**, 388. https://doi.org/10.3389/fpls.2016.00388

Nakabayashi K., Okamoto M., Koshiba T., Kamiya Y. and Nambara E. (2005) Genome-wide profiling of stored mRNA in Arabidopsis thaliana seed germination: epigenetic and genetic regulation of transcription in seed. *The Plant Journal* **41**, 697–709. https://doi.org/10.1111/j.1365-313X.2005.02337.x

Nakagawa T., Suzuki T., Murata S., Nakamura S., Hino T., Maeo K., … Ishiguro S. (2007) Improved Gateway Binary Vectors: High-Performance Vectors for Creation of Fusion Constructs in Transgenic Analysis of Plants. *Bioscience, Biotechnology, and Biochemistry* **71**, 2095–2100. https://doi.org/10.1271/bbb.70216

Nakajima S., Ito H., Tanaka R. and Tanaka A. (2012) Chlorophyll b Reductase Plays an Essential Role in Maturation and Storability of Arabidopsis Seeds. *Plant Physiology* **160**, 261. https://doi.org/10.1104/pp.112.196881

Nakaminami K., Matsui A., Shinozaki K. and Seki M. (2012) RNA regulation in plant abiotic stress responses. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1819**, 149–153. https://doi.org/10.1016/j.bbagrm.2011.07.015

Nakamura H., Nakamura K. and Yodoi J. (1997) Redox regulation of cellular activation. *Annual Review of Immunology* **15**, 351–369. https://doi.org/10.1146/annurev.immunol.15.1.351

Nakano T., Suzuki K., Fujimura T. and Shinshi H. (2006) Genome-Wide Analysis of the ERF Gene Family in Arabidopsis and Rice. *Plant Physiology* **140**, 411–432. https://doi.org/10.1104/pp.105.073783

Nakashima K., Shinwari Z.K., Sakuma Y., Seki M., Miura S., Shinozaki K. and Yamaguchi-Shinozaki K. (2000) Organization and expression of two Arabidopsis DREB2 genes encoding DRE-binding proteins involved in dehydration- and high-salinity-responsive gene expression. *Plant Molecular Biology* **42**, 657–665. https://doi.org/10.1023/A:1006321900483

Nakaune S., Yamada K., Kondo M., Kato T., Tabata S., Nishimura M. and Hara-Nishimura I. (2005) A Vacuolar Processing Enzyme, δVPE, Is Involved in Seed Coat Formation at the Early Stage of Seed Development. *The Plant Cell* **17**, 876–887. https://doi.org/10.1105/tpc.104.026872

Navashin M. and Shkvarnikov P. (1933) Process of Mutation in Resting Seeds accelerated by Increased Temperature. *Nature* **132**, 482–483. https://doi.org/10.1038/132482c0

Nelson D.R. (1999) Cytochrome P450 and the Individuality of Species. *Archives of Biochemistry and Biophysics* **369**, 1–10. https://doi.org/10.1006/abbi.1999.1352

Nesi N., Debeaujon I., Jond C., Stewart A.J., Jenkins G.I., Caboche M. and Lepiniec L. (2002) The TRANSPARENT TESTA16 Locus Encodes the ARABIDOPSIS BSISTER MADS Domain Protein and Is Required for Proper Development and Pigmentation of the Seed Coat. *The Plant Cell* **14**, 2463–2479. https://doi.org/10.1105/tpc.004127

Nguyen T.-P., Cueff G., Hegedus D.D., Rajjou L. and Bentsink L. (2015) A role for seed storage proteins in Arabidopsis seed longevity. *Journal of Experimental Botany* **66**, 6399–6413. https://doi.org/10.1093/jxb/erv348

Nguyen T.-P., Keizer P., Eeuwijk F. van, Smeekens S. and Bentsink L. (2012) Natural Variation for Seed

Longevity and Seed Dormancy Are Negatively Correlated in Arabidopsis. *Plant Physiology* **160**, 2083–2092. https://doi.org/10.1104/pp.112.206649

Nicolas M. and Cubas P. (2016) TCP factors: new kids on the signaling block. *Current Opinion in Plant Biology* **33**, 33–41. https://doi.org/10.1016/j.pbi.2016.05.006

Nishizawa A., Yabuta Y. and Shigeoka S. (2008) Galactinol and Raffinose Constitute a Novel Function to Protect Plants from Oxidative Damage. *Plant Physiology* **147**, 1251–1263. https://doi.org/10.1104/pp.108.122465

Nonogaki H. (2014) Seed dormancy and germination-emerging mechanisms and new hypotheses. *Frontiers in Plant Science* **5**, 233. https://doi.org/10.3389/fpls.2014.00233

North H., Baud S., Debeaujon I., Dubos C., Dubreucq B., Grappin P., … Caboche M. (2010) Arabidopsis seed secrets unravelled after a decade of genetic and omics-driven research. *The Plant Journal: For Cell and Molecular Biology* **61**, 971–981. https://doi.org/10.1111/j.1365-313X.2009.04095.x

Novo E. and Parola M. (2008) Redox mechanisms in hepatic chronic wound healing and fibrogenesis. *Fibrogenesis and Tissue Repair* **1**, 5. https://doi.org/10.1186/1755-1536-1-5

Nyhan M.M., Coull B.A., Blomberg A.J., Vieira C.L.Z., Garshick E., Aba A., … Koutrakis P. (2018) Associations Between Ambient Particle Radioactivity and Blood Pressure: The NAS (Normative Aging Study). *Journal of the American Heart Association* **7**. https://doi.org/10.1161/JAHA.117.008245

O'Malley R.C. and Ecker J.R. (2010) Linking genotype to phenotype using the Arabidopsis unimutant collection. *The Plant Journal* **61**, 928–940. https://doi.org/10.1111/j.1365-313X.2010.04119.x

O'Malley R.C., Huang S.C., Song L., Lewsey M.G., Bartlett A., Nery J.R., … Ecker J.R. (2016) Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**, 1280–1292. https://doi.org/10.1016/j.cell.2016.04.038

Ogé L., Bourdais G., Bove J., Collet B., Godin B., Granier F., … Grappin P. (2008) Protein Repair l-Isoaspartyl Methyltransferase1 Is Involved in Both Seed Longevity and Germination Vigor in Arabidopsis. *The Plant Cell* **20**, 3022–3037. https://doi.org/10.1105/tpc.108.058479

Ohto M., Fischer R.L., Goldberg R.B., Nakamura K. and Harada J.J. (2005) Control of seed mass by APETALA2. *Proceedings of the National Academy of Sciences* **102**, 3123–3128. https://doi.org/10.1073/pnas.0409858102

Oliver M.J., Tuba Z. and Mishler B.D. (2000) The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology* **151**, 85–100. https://doi.org/10.1023/A:1026550808557

Oñate-Sánchez L. and Vicente-Carbajosa J. (2008) DNA-free RNA isolation protocols for Arabidopsis thaliana, including seeds and siliques. *BMC Research Notes* **1**, 93. https://doi.org/10.1186/1756-0500-1-93

Ondzighi C.A., Christopher D.A., Cho E.J., Chang S.-C. and Staehelin L.A. (2008) Arabidopsis Protein Disulfide Isomerase-5 Inhibits Cysteine Proteases during Trafficking to Vacuoles before Programmed Cell Death of the Endothelium in Developing Seeds. *The Plant Cell* **20**, 2205–2220. https://doi.org/10.1105/tpc.108.058339

Ooms J., Leon-Kloosterziel K.M., Bartels D., Koornneef M. and Karssen C.M. (1993) Acquisition of Desiccation Tolerance and Longevity in Seeds of Arabidopsis thaliana (A Comparative Study Using Abscisic Acid-Insensitive abi3 Mutants). *Plant Physiology* **102**, 1185–1191. https://doi.org/10.1104/pp.102.4.1185

Oshima Y. and Mitsuda N. (2013) The MIXTA-like Transcription factor MYB16 is a major regulator of cuticle formation in vegetative organs. *Plant Signaling and Behavior* **8**, e26826. https://doi.org/10.4161/psb.26826

Oshima Y., Shikata M., Koyama T., Ohtsubo N., Mitsuda N. and Ohme-Takagi M. (2013) MIXTA-Like Transcription Factors and WAX INDUCER1/SHINE1 Coordinately Regulate Cuticle Development in Arabidopsis and Torenia fournieri. *The Plant Cell* **25**, 1609–1624. https://doi.org/10.1105/tpc.113.110783

Østergaard L., Teilum K., Mirza O., Mattsson O., Petersen M., Welinder K.G., … Henriksen A. (2000) Arabidopsis ATP A2 peroxidase. Expression and high-resolution structure of a plant peroxidase with implications for lignification. *Plant Molecular Biology* **44**, 231–243. https://doi.org/10.1023/A:1006442618860

Palaniswamy S.K., James S., Sun H., Lamb R.S., Davuluri R.V. and Grotewold E. (2006) AGRIS and AtRegNet. A Platform to Link cis-Regulatory Elements and Transcription Factors into Regulatory Networks. *Plant Physiology* **140**, 818–829. https://doi.org/10.1104/pp.105.072280

Pamplona R. and Barja G. (2006) Mitochondrial oxidative stress, aging and caloric restriction: The protein and methionine connection. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1757**, 496–508. https://doi.org/10.1016/j.bbabio.2006.01.009

Panikashvili D., Shi J.X., Schreiber L. and Aharoni A. (2009) The Arabidopsis DCR Encoding a Soluble BAHD Acyltransferase Is Required for Cutin Polyester Formation and Seed Hydration Properties. *Plant Physiology* **151**, 1773–1789. https://doi.org/10.1104/pp.109.143388

Papi M., Sabatini S., Bouchez D., Camilleri C., Costantino P. and Vittorioso P. (2000) Identification and disruption of an Arabidopsis zinc finger gene controlling seed germination. *Genes and Development* **14**, 28–33. https://doi.org/10.1101/gad.14.1.28

Park D.H., Lim P.O., Kim J.S., Cho D.S., Hong S.H. and Nam H.G. (2003) The Arabidopsis COG1 gene encodes a Dof domain transcription factor and negatively regulates phytochrome signaling. *The Plant Journal* **34**, 161–171. https://doi.org/10.1046/j.1365-313X.2003.01710.x

Parry R.T. (1993) *Principles and Applications of Modified Atmosphere Packaging of Foods*. Springer US, Boston, MA. https://doi.org/10.1007/978-1-4615-2137-2

Passardi F., Longet D., Penel C. and Dunand C. (2004) The class III peroxidase multigenic family in rice and its evolution in land plants. *Phytochemistry* **65**, 1879–1893. https://doi.org/10.1016/j.phytochem.2004.06.023

Pedreira J., Herrera M.T., Zarra I. and Revilla G. (2011) The overexpression of AtPrx37, an apoplastic peroxidase, reduces growth in Arabidopsis. *Physiologia Plantarum* **141**, 177–187. https://doi.org/10.1111/j.1399-3054.2010.01427.x

Pelaz S., Ditta G.S., Baumann E., Wisman E. and Yanofsky M.F. (2000) B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* **405**, 200–203. https://doi.

org/10.1038/35012103

Pellizzaro A., Neveu M., Lalanne D., Vu B.L., Kanno Y., Seo M., … Buitink J. (2020) A role for auxin signaling in the acquisition of longevity during seed maturation. *New Phytologist* **225**, 284–296. https://doi.org/10.1111/nph.16150

Perez-Garcia P. and Moreno-Risueno M.A. (2018) Stem cells and plant regeneration. *Developmental Biology* **442**, 3–12. https://doi.org/10.1016/j.ydbio.2018.06.021

Pfaffl M.W. (2001) A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research* **29**, e45–e45. https://doi.org/10.1093/nar/29.9.e45

Philippe G., Gaillard C., Petit J., Geneix N., Dalgalarrondo M., Bres C., … Bakan B. (2016) Ester Cross-Link Profiling of the Cutin Polymer of Wild-Type and Cutin Synthase Tomato Mutants Highlights Different Mechanisms of Polymerization. *Plant Physiology* **170**, 807–820. https://doi.org/10.1104/pp.15.01620

Philippe G., Sørensen I., Jiao C., Sun X., Fei Z., Domozych D.S. and Rose J.K. (2020) Cutin and suberin: assembly and origins of specialized lipidic cell wall scaffolds. *Current Opinion in Plant Biology* **55**, 11–20. https://doi.org/10.1016/j.pbi.2020.01.008

Piedrafita G., Keller M.A. and Ralser M. (2015) The Impact of Non-Enzymatic Reactions and Enzyme Promiscuity on Cellular Metabolism during (Oxidative) Stress Conditions. *Biomolecules* **5**, 2101–2122. https://doi.org/10.3390/biom5032101

Pierson B.K (1994) The emergence, diversification, and role of photosynthetic bacteria. *In Early life on earth*. Nobel Symposium and Bengtson S. eds. Columbia University Press, New York. pp 161–180.

Pighin J.A., Zheng H., Balakshin L.J., Goodman I.P., Western T.L., Jetter R., … Samuels A.L. (2004) Plant Cuticular Lipid Export Requires an ABC Transporter. *Science* **306**, 702–704. https://doi.org/10.1126/science.1102331

Pineau E., Xu L., Renault H., Trolet A., Navrot N., Ullmann P., … Pinot F. (2017) Arabidopsis thaliana EPOXIDE HYDROLASE1 (AtEH1) is a cytosolic epoxide hydrolase involved in the synthesis of poly-hydroxylated cutin monomers. *New Phytologist* **215**, 173–186. https://doi.org/10.1111/nph.14590

Pinnola A. and Bassi R. (2018) Molecular mechanisms involved in plant photoprotection. *Biochemical Society Transactions* **46**, 467–482. https://doi.org/10.1042/BST20170307

Pinyopich A., Ditta G.S., Savidge B., Liljegren S.J., Baumann E., Wisman E. and Yanofsky M.F. (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* **424**, 85–88. https://doi.org/10.1038/nature01741

Piraino S., Boero F., Aeschbach B. and Schmid V. (1996) Reversing the Life Cycle: Medusae Transforming into Polyps and Cell Transdifferentiation in Turritopsis nutricula (Cnidaria, Hydrozoa). *The Biological Bulletin* **190**, 302–312. https://doi.org/10.2307/1543022

Pistón F., Marín S., Hernando A. and Barro F. (2009) Analysis of the activity of a γ-gliadin promoter in

transgenic wheat and characterization of gliadin synthesis in wheat by MALDI-TOF during grain development. *Molecular Breeding* **23**, 655–667. https://doi.org/10.1007/s11032-009-9263-1

Pollard M., Beisson F., Li Y. and Ohlrogge J.B. (2008) Building lipid barriers: biosynthesis of cutin and suberin. *Trends in Plant Science* **13**, 236–246. https://doi.org/10.1016/j.tplants.2008.03.003

Pörtner H.O. (2002) Climate variations and the physiological basis of temperature dependent biogeography: systemic to molecular hierarchy of thermal tolerance in animals. *Comparative Biochemistry and Physiology. Part A, Molecular and Integrative Physiology* **132**, 739–761. https://doi.org/10.1016/S1095-6433(02)00045-4

Pourcel L., Routaboul J.-M., Cheynier V., Lepiniec L. and Debeaujon I. (2007) Flavonoid oxidation in plants: from biochemical properties to physiological functions. *Trends in Plant Science* **12**, 29–36. https://doi.org/10.1016/j.tplants.2006.11.006

Pourcel L., Routaboul J.-M., Kerhoas L., Caboche M., Lepiniec L. and Debeaujon I. (2005) TRANSPARENT TESTA10 Encodes a Laccase-Like Enzyme Involved in Oxidative Polymerization of Flavonoids in Arabidopsis Seed Coat. *The Plant Cell* **17**, 2966–2980. https://doi.org/10.1105/tpc.105.035154

Powell A. and Matthews S. (2012) Seed Aging/Repair Hypothesis Leads to New Testing Methods. *Seed Technology* **34**, 15–25. https://www.jstor.org/stable/23433631

Prasad K., Zhang X., Tobón E. and Ambrose B.A. (2010) The Arabidopsis B-sister MADS-box protein, GORDITA, represses fruit growth and contributes to integument development. *The Plant Journal* **62**, 203–214. https://doi.org/10.1111/j.1365-313X.2010.04139.x

Prieto-Dapena P., Castaño R., Almoguera C. and Jordano J. (2006) Improved Resistance to Controlled Deterioration in Transgenic Seeds. *Plant Physiology* **142**, 1102–1112. https://doi.org/10.1104/pp.106.087817

Probert R.J., Daws M.I. and Hay F.R. (2009) Ecological correlates of ex situ seed longevity: a comparative study on 195 species. *Annals of Botany* **104**, 57–69. https://doi.org/10.1093/aob/mcp082

Proctor, M. C. F. and Pence, V. C. (2002) Vegetative Tissues: Bryophytes, vascular resurrection plants, and vegetative propogules. *In* M. Black and H.W. Pritchard (eds.), *Desiccation and survival in plants: Drying without dying*, pp. 207–237. CABI Publishing, Wallingford, Oxon.

Pruitt R.E., Vielle-Calzada J.-P., Ploense S.E., Grossniklaus U. and Lolle S.J. (2000) FIDDLEHEAD, a gene required to suppress epidermal cell interactions in Arabidopsis, encodes a putative lipid biosynthetic enzyme. *Proceedings of the National Academy of Sciences* **97**, 1311–1316. https://doi.org/10.1073/pnas.97.3.1311

Puthur J.T., Shackira A.M., Saradhi P.P. and Bartels D. (2013) Chloroembryos: a unique photosynthesis system. *Journal of Plant Physiology* **170**, 1131–1138. https://doi.org/10.1016/j.jplph.2013.04.011

Qu, Y., Yan, M., and Zhang, Q. (2017) Functional regulation of plant NADPH oxidase and its role in signaling. *Plant Signaling andBehavior* **12**(8), e1356970. https://doi.org/10.1080/15592324.2017.1356970

Quiroga M., Guerrero C., Botella M.A., Barceló A., Amaya I., Medina M.I., … Valpuesta V. (2000) A Tomato Peroxidase Involved in the Synthesis of Lignin and Suberin. *Plant Physiology* **122**,

1119–1128. https://doi.org/10.1104/pp.122.4.1119

Rahantaniaina M.-S., Li S., Chatel-Innocenti G., Tuzet A., Issakidis-Bourguet E., Mhamdi A. and Noctor G. (2017) Cytosolic and Chloroplastic DHARs Cooperate in Oxidative Stress-Driven Activation of the Salicylic Acid Pathway. *Plant Physiology* **174**, 956–971. https://doi.org/10.1104/pp.17.00317

Rains M.K., Gardiyehewa de Silva N.D. and Molina I. (2018) Reconstructing the suberin pathway in poplar by chemical and transcriptomic analysis of bark tissues. *Tree Physiology* **38**, 340–361. https://doi.org/10.1093/treephys/tpx060

Rajjou L., Lovigny Y., Groot S.P.C., Belghazi M., Job C. and Job D. (2008) Proteome-wide characterization of seed aging in Arabidopsis: a comparison between artificial and natural aging protocols. *Plant Physiology* **148**, 620–641. https://doi.org/10.1104/pp.108.123141

Rautengarten C., Ebert B., Ouellet M., Nafisi M., Baidoo E.E.K., Benke P., … Scheller H.V. (2012) Arabidopsis Deficient in Cutin Ferulate Encodes a Transferase Required for Feruloylation of ω-Hydroxy Fatty Acids in Cutin Polyester. *Plant Physiology* **158**, 654–665. https://doi.org/10.1104/pp.111.187187

Renard J., Martínez-Almonacid I., Sonntag A., Molina I., Moya-Cuevas J., Bissoli G., … Bueso E. (2020*a*) PRX2 and PRX25, peroxidases regulated by COG1, are involved in seed longevity in Arabidopsis. *Plant, Cell & Environment* **43**, 315–326. https://doi.org/10.1111/pce.13656 ***This article is part of this PhD thesis.**

Renard J., Niñoles R., Martínez-Almonacid I., Gayubas B., Mateos R., Bissoli G., … Gadea J. (2020*b*) Identification of novel seed longevity genes related to oxidative stress and seed coat by genome-wide association studies and reverse genetics. *Plant, Cell & Environment* **43**, 2523–2539. https://doi.org/10.1111/pce.13822 ***This article is part of this PhD thesis.**

Renard J., Niñoles R., Martínez-Almonacid I., Queralta-Castillo I., Sonntag A., Bissoli G., … Bueso E. (2021) Apoplastic lipid barriers regulated by conserved homeobox transcription factors extend seed longevity in multiple plant species. *New Phytologist* **231**, 679–694. https://doi.org/10.1111/nph.17399 ***This article is part of this PhD thesis.**

Righetti K., Vu J.L., Pelletier S., Vu B.L., Glaab E., Lalanne D., … Buitink J. (2015) Inference of Longevity-Related Genes from a Robust Coexpression Network of Seed Maturation Identifies Regulators Linking Seed Storability to Biotic Defense-Related Pathways. *The Plant Cell* **27**, 2692–2708. https://doi.org/10.1105/tpc.15.00632

Romano J.M., Dubos C., Prouse M.B., Wilkins O., Hong H., Poole M., … Campbell M.M. (2012) AtMYB61, an R2R3-MYB transcription factor, functions as a pleiotropic regulator via a small gene network. *New Phytologist* **195**, 774–786. https://doi.org/10.1111/j.1469-8137.2012.04201.x

Rouhier N., Lemaire S.D. and Jacquot J.-P. (2008) The Role of Glutathione in Photosynthetic Organisms: Emerging Functions for Glutaredoxins and Glutathionylation. *Annual Review of Plant Biology* **59**, 143–166. https://doi.org/10.1146/annurev.arplant.59.032607.092811

Rueda-Romero P., Barrero-Sicilia C., Gómez-Cadenas A., Carbonero P. and Oñate-Sánchez L. (2012) Arabidopsis thaliana DOF6 negatively affects germination in non-after-ripened seeds and interacts with TCP14. *Journal of Experimental Botany* **63**, 1937–1949. https://doi.org/10.1093/jxb/err388

Russell W.R., Burkitt M.J., Scobbie L. and Chesson A. (2006) EPR Investigation into the Effects of Substrate Structure on Peroxidase-Catalyzed Phenylpropanoid Oxidation. *Biomacromolecules* **7**, 268–273. https://doi.org/10.1021/bm050636o

Ruta V., Longo C., Lepri A., De Angelis V., Occhigrossi S., Costantino P. and Vittorioso P. (2020) The DOF Transcription Factors in Seed and Seedling Development. *Plants* **9**, 218. https://doi.org/10.3390/plants9020218

Ruuska S.A., Girke T., Benning C. and Ohlrogge J.B. (2002) Contrapuntal Networks of Gene Expression during Arabidopsis Seed Filling. *The Plant Cell* **14**, 1191. https://doi.org/10.1105/tpc.000877

Sakuma Y., Liu Q., Dubouzet J.G., Abe H., Shinozaki K. and Yamaguchi-Shinozaki K. (2002) DNA-Binding Specificity of the ERF/AP2 Domain of Arabidopsis DREBs, Transcription Factors Involved in Dehydration- and Cold-Inducible Gene Expression. *Biochemical and Biophysical Research Communications* **290**, 998–1009. https://doi.org/10.1006/bbrc.2001.6299

Sallon S., Solowey E., Cohen Y., Korchinsky R., Egli M., Woodhatch I., … Kislev M. (2008) Germination, Genetics, and Growth of an Ancient Date Seed. *Science* **320**, 1464–1464. https://doi.org/10.1126/science.1153600

Sánchez-Montesino, R., and Oñate-Sánchez, L. (2017). Yeast One- and Two-Hybrid High-Throughput Screenings Using Arrayed Libraries. En K. Kaufmann and B. Mueller-Roeber (Eds.), *Plant Gene Regulatory Networks: Methods and Protocols* (pp. 47-65). Springer. https://doi.org/10.1007/978-1-4939-7125-1_5

Sano N., Rajjou L. and North H.M. (2020) Lost in Translation: Physiological Roles of Stored mRNAs in Seed Germination. *Plants* **9**, 347. https://doi.org/10.3390/plants9030347

Sano N., Rajjou L., North H.M., Debeaujon I., Marion-Poll A. and Seo M. (2016) Staying Alive: Molecular Aspects of Seed Longevity. *Plant and Cell Physiology* **57**, 660–674. https://doi.org/10.1093/pcp/pcv186

Santopolo S., Boccaccini A., Lorrai R., Ruta V., Capauto D., Minutello E., … Vittorioso P. (2015) DOF AFFECTING GERMINATION 2 is a positive regulator of light-mediated seed germination and is repressed by DOF AFFECTING GERMINATION 1. *BMC Plant Biology* **15**, 72. https://doi.org/10.1186/s12870-015-0453-1

Sattler S.E., Gilliland L.U., Magallanes-Lundback M., Pollard M. and DellaPenna D. (2004) Vitamin E Is Essential for Seed Longevity and for Preventing Lipid Peroxidation during Germination. *The Plant Cell* **16**, 1419–1432. https://doi.org/10.1105/tpc.02136

Sauveplane V., Kandel S., Kastner P.-E., Ehlting J., Compagnon V., Werck-Reichhart D. and Pinot F. (2009) Arabidopsis thaliana CYP77A4 is the first cytochrome P450 able to catalyze the epoxidation of free fatty acids in plants. *The FEBS journal* **276**, 719–735. https://doi.org/10.1111/j.1742-4658.2008.06819.x

Schafer F.Q. and Buettner G.R. (2001) Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radical Biology and Medicine* **30**, 1191–1212. https://doi.org/10.1016/s0891-5849(01)00480-4

Schindelin J., Arganda-Carreras I., Frise E., Kaynig V., Longair M., Pietzsch T., … Cardona A. (2012) Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676–682. https://

doi.org/10.1038/nmeth.2019

Schmidlin C.J., Dodson M.B., Madhavan L. and Zhang D.D. (2019) Redox regulation by NRF2 in aging and disease. *Free Radical Biology and Medicine* **134**, 702–707. https://doi.org/10.1016/j.freeradbiomed.2019.01.016

Schnurr J., Shockey J. and Browse J. (2004) The Acyl-CoA Synthetase Encoded by LACS2 Is Essential for Normal Cuticle Development in Arabidopsis. *The Plant Cell* **16**, 629–642. https://doi.org/10.1105/tpc.017608

Schulman, E. (1958) Bristlecone pine, oldest known living thing. *Nat. Geogr. Mag*. **113**: 355-372.

Sengupta S., Mukherjee S., Basak P. and Majumder A.L. (2015) Significance of galactinol and raffinose family oligosaccharide synthesis in plants. *Frontiers in Plant Science* **6**.https://doi.org/10.3389/fpls.2015.00656

Seren Ü., Grimm D., Fitz J., Weigel D., Nordborg M., Borgwardt K. and Korte A. (2017) AraPheno: a public database for Arabidopsis thaliana phenotypes. *Nucleic Acids Research* **45**, D1054–D1059. https://doi.org/10.1093/nar/gkw986

Seren, Ü., Vilhjálmsson, B. J., Horton, M. W., Meng, D., Forai, P., Huang, Y. S., Long, Q., Segura, V., and Nordborg, M. (2012) GWAPP: a web application for genome-wide association mapping in Arabidopsis. *The Plant cell* **24**(12), 4793–4805. https://doi.org/10.1105/tpc.112.108068

Shah K., Kumar R.G., Verma S. and Dubey R.S. (2001) Effect of cadmium on lipid peroxidation, superoxide anion generation and activities of antioxidant enzymes in growing rice seedlings. *Plant Science* **161**, 1135–1144. https://doi.org/10.1016/S0168-9452(01)00517-9

Shao S., Meyer C.J., Ma F., Peterson C.A. and Bernards M.A. (2007) The outermost cuticle of soybean seeds: chemical composition and function during imbibition. *Journal of Experimental Botany* **58**, 1071–1082. https://doi.org/10.1093/jxb/erl268

Sharma, M., and Pandey, G. K. (2016) Expansion and Function of Repeat Domain Proteins During Stress and Development in Plants. *Frontiers in plant science* **6**, 1218. https://doi.org/10.3389/fpls.2015.01218

Shen J. and Tower J. (2019) Effects of light on aging and longevity. Aging Research Reviews 53, 100913. https://doi.org/10.1016/j.arr.2019.100913

Shen J., Zhu X., Gu Y., Zhang C., Huang J. and Xiao Q. (2019) Toxic Effect of Visible Light on Drosophila Life Span Depending on Diet Protein Content. The Journals of Gerontology: Series A 74, 163–167. https://doi.org/10.1093/gerona/gly042

Shen W., Yao X., Ye T., Ma S., Liu X., Yin X. and Wu Y. (2018) Arabidopsis Aspartic Protease ASPG1 Affects Seed Dormancy, Seed Longevity and Seed Germination. *Plant and Cell Physiology* **59**, 1415–1431. https://doi.org/10.1093/pcp/pcy070

Shen-Miller J., Mudgett M.B., Schopf J.W., Clarke S. and Berger R. (1995) Exceptional seed longevity and robust growth: ancient Sacred Lotus from China. *American Journal of Botany* **82**, 1367–1380. https://doi.org/10.1002/j.1537-2197.1995.tb12673.x

Shigeto J., Itoh Y., Hirao S., Ohira K., Fujita K. and Tsutsumi Y. (2015) Simultaneously disrupting

AtPrx2, AtPrx25 and AtPrx71 alters lignin content and structure in Arabidopsis stem. *Journal of Integrative Plant Biology* **57**, 349–356. https://doi.org/10.1111/jipb.12334

Shigeto J., Kiyonaga Y., Fujita K., Kondo R. and Tsutsumi Y. (2013) Putative Cationic Cell-Wall-Bound Peroxidase Homologues in Arabidopsis, AtPrx2, AtPrx25, and AtPrx71, Are Involved in Lignification. *Journal of Agricultural and Food Chemistry* **61**, 3781–3788. https://doi.org/10.1021/jf400426g

Singh H. (2018) Desiccation and radiation stress tolerance in cyanobacteria. *Journal of Basic Microbiology* **58**, 813–826. https://doi.org/10.1002/jobm.201800216

Smirnoff N. (2011) Chapter 4 - Vitamin C: The Metabolism and Functions of Ascorbic Acid in Plants. In *Advances in Botanical Research*. Biosynthesis of Vitamins in Plants Part B, (eds F. Rébeillé and R. Douce), pp. 107–177. Academic Press. https://doi.org/10.1016/B978-0-12-385853-5.00003-9

Smolikova G., Dolgikh E., Vikhnina M., Frolov A. and Medvedev S. (2017) Genetic and Hormonal Regulation of Chlorophyll Degradation during Maturation of Seeds with Green Embryos. International Journal of Molecular Sciences 18. https://doi.org/10.3390/ijms18091993

Soare A., Cangemi R., Omodei D., Holloszy J.O. and Fontana L. (2011) Long-term calorie restriction, but not endurance exercise, lowers core body temperature in humans. *Aging (Albany NY)* **3**, 374–379. https://doi.org/10.18632/aging.100280

Soliday C.L., Dean B.B. and Kolattukudy P.E. (1978) Suberization: Inhibition by Washing and Stimulation by Abscisic Acid in Potato Disks and Tissue Culture. *Plant Physiology* **61**, 170–174. https://doi.org/10.1104/pp.61.2.170

Soza V.L., Snelson C.D., Hewett Hazelton K.D. and Di Stilio V.S. (2016) Partial redundancy and functional specialization of E-class SEPALLATA genes in an early-diverging eudicot. *Developmental Biology* **419**, 143–155. https://doi.org/10.1016/j.ydbio.2016.07.021

Spitz F. and Furlong E.E.M. (2012) Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics* **13**, 613–626. https://doi.org/10.1038/nrg3207

Sreenivasulu N. and Wobus U. (2013) Seed-development programs: a systems biology-based comparison between dicots and monocots. *Annual Review of Plant Biology* **64**, 189–217. https://doi.org/10.1146/annurev-arplant-050312-120215

Stadtman E.R. (2006) Protein oxidation and aging. *Free Radical Research* **40**, 1250–1258. https://doi.org/10.1080/10715760600918142

Stewart E.J., Madden R., Paul G. and Taddei F. (2005) Aging and Death in an Organism That Reproduces by Morphologically Symmetric Division. *PLOS Biology* **3**, e45. https://doi.org/10.1371/journal.pbio.0030045

Storey K.B. Out Cold: Biochemical Regulation of Mammalian Hibernation – A Mini-Review. *Gerontology* **56**, 220–230. https://doi.org/10.1159/000228829

Stormo G.D. (2000) DNA binding sites: representation and discovery. *Bioinformatics* **16**, 16–23. https://doi.org/10.1093/bioinformatics/16.1.16

Sung J.M. (1996) Lipid peroxidation and peroxide-scavenging in soybean seeds during aging. *Physiologia Plantarum* **97**, 85–89. https://doi.org/10.1111/j.1399-3054.1996.tb00482.x

Suzuki N., Miller G., Morales J., Shulaev V., Torres M.A. and Mittler R. (2011) Respiratory burst oxidases: the engines of ROS signaling. *Current Opinion in Plant Biology* **14**, 691–699. https://doi.org/10.1016/j.pbi.2011.07.014

Takeda S., Iwasaki A., Tatematsu K. and Okada K. (2014) The Half-Size ABC Transporter FOLDED PETALS 2/ABCG13 Is Involved in Petal Elongation through Narrow Spaces in Arabidopsis thaliana Floral Buds. *Plants* **3**, 348–358. https://doi.org/10.3390/plants3030348

Tan Q.K.-G. and Irish V.F. (2006) The Arabidopsis Zinc Finger-Homeodomain Genes Encode Proteins with Unique Biochemical Properties That Are Coordinately Expressed during Floral Development. *Plant Physiology* **140**, 1095–1108. https://doi.org/10.1104/pp.105.070565

Tausch S., Leipold M., Reisch C. and Poschlod P. (2019) Dormancy and endosperm presence influence the ex situ conservation potential in central European calcareous grassland plants. *AoB PLANTS* **11**. https://doi.org/10.1093/aobpla/plz035

Tejedor-Cano J., Prieto-Dapena P., Almoguera C., Carranco R., Hiratsu K., Ohme-Takagi M. and Jordano J. (2010) Loss of function of the HSFA9 seed longevity program. *Plant, Cell & Environment* **33**, 1408–1417. https://doi.org/10.1111/j.1365-3040.2010.02159.x

Tian T., Liu Y., Yan H., You Q., Yi X., Du Z., … Su Z. (2017) agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Research* **45**, W122–W129. https://doi.org/10.1093/nar/gkx382

Tobimatsu Y., Chen F., Nakashima J., Escamilla-Treviño L.L., Jackson L., Dixon R.A. and Ralph J. (2013) Coexistence but Independent Biosynthesis of Catechyl and Guaiacyl/Syringyl Lignin Polymers in Seed Coats. *The Plant Cell* **25**, 2587–2600. https://doi.org/10.1105/tpc.113.113142

Torres M.A., Dangl J.L. and Jones J.D.G. (2002) Arabidopsis gp91phox homologues AtrbohD and AtrbohF are required for accumulation of reactive oxygen intermediates in the plant defense response. *Proceedings of the National Academy of Sciences* **99**, 517–522. https://doi.org/10.1073/pnas.012452499

Trapnell C., Hendrickson D.G., Sauvageau M., Goff L., Rinn J.L. and Pachter L. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology* **31**, 46–53. https://doi.org/10.1038/nbt.2450

Truernit E. and Haseloff J. (2008) Arabidopsis thaliana outer ovule integument morphogenesis: Ectopic expression of KNAT1 reveals a compensation mechanism. *BMC Plant Biology* **8**, 35. https://doi.org/10.1186/1471-2229-8-35

Ursini F. and Davies K.J.A. (1995) *The oxygen paradox*. CLEUP University Press, Padova, Italy.

Van Bel M., Diels T., Vancaester E., Kreft L., Botzki A., Van de Peer Y., … Vandepoele K. (2018) PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Research* **46**, D1190–D1196. https://doi.org/10.1093/nar/gkx1002

Venugopal S.C., Chanda B., Vaillancourt L., Kachroo A. and Kachroo P. (2009) The common metabolite glycerol-3-phosphate is a novel regulator of plant defense signaling. *Plant Signaling & Behavior* **4**, 746–749. https://doi.org/10.4161/psb.4.8.9111

Verdier J., Lalanne D., Pelletier S., Torres-Jerez I., Righetti K., Bandyopadhyay K., … Buitink J. (2013) A Regulatory Network-Based Approach Dissects Late Maturation Processes Related to the Acquisition of Desiccation Tolerance and Longevity of Medicago truncatula Seeds. *Plant Physiology* **163**, 757–774. https://doi.org/10.1104/pp.113.222380

Vertucci, C. W., and Roos, E. E. (1990) Theoretical basis of protocols for seed storage. *Plant physiology* **94**(3), 1019–1023. https://doi.org/10.1104/pp.94.3.1019

Vishal B., Krishnamurthy P., Ramamoorthy R. and Kumar P.P. (2019) OsTPS8 controls yield-related traits and confers salt stress tolerance in rice by enhancing suberin deposition. *The New Phytologist* **221**, https://doi.org/1369–1386. 10.1111/nph.15464

Vishwanath S.J., Delude C., Domergue F. and Rowland O. (2015) Suberin: biosynthesis, regulation, and polymer assembly of a protective extracellular barrier. *Plant Cell Reports* **34**, 573–586. https://doi.org/10.1007/s00299-014-1727-z

Vishwanath S.J., Kosma D.K., Pulsifer I.P., Scandola S., Pascal S., Joubès J., … Domergue F. (2013) Suberin-Associated Fatty Alcohols in Arabidopsis: Distributions in Roots and Contributions to Seed Coat Barrier Properties. *Plant Physiology* **163**, 1118–1132. https://doi.org/10.1104/pp.113.224410

Vogt T. (2010) Phenylpropanoid Biosynthesis. *Molecular Plant* **3**, 2–20. https://doi.org/10.1093/mp/ssp106

Voisin D., Nawrath C., Kurdyukov S., Franke R.B., Reina-Pinto J.J., Efremova N., … Yephremov A. (2009) Dissection of the Complex Phenotype in Cuticular Mutants of Arabidopsis Reveals a Role of SERRATE as a Mediator. *PLoS Genetics* **5**. https://doi.org/10.1371/journal.pgen.1000703

Walters C. (1998) Understanding the mechanisms and kinetics of seed aging. *Seed Science Research* **8**, 223–244. https://doi.org/10.1017/S096025850000413X

Walters C., Ballesteros D. and Vertucci V.A. (2010) Structural mechanics of seed deterioration: Standing the test of time. *Plant Science* **179**, 565–573. https://doi.org/10.1016/j.plantsci.2010.06.016

Walters C., Wheeler L.M. and Grotenhuis J.M. (2005) Longevity of seeds stored in a genebank: species characteristics. *Seed Science Research* **15**, 1–20. https://doi.org/10.1079/SSR2004195

Wang C., Liu Y., Li S.-S. and Han G.-Z. (2015) Insights into the origin and evolution of the plant hormone signaling machinery. *Plant Physiology* **167**, 872–886. https://doi.org/10.1104/pp.114.247403

Wang G.-L., Que F., Xu Z.-S., Wang F. and Xiong A.-S. (2017*a*) Exogenous gibberellin enhances secondary xylem development and lignification in carrot taproot. *Protoplasma* **254**, 839–848. https://doi.org/10.1007/s00709-016-0995-6

Wang J., Zhang H. and Allen R.D. (1999) Overexpression of an Arabidopsis Peroxisomal Ascorbate Peroxidase Gene in Tobacco Increases Protection Against Oxidative Stress. *Plant and Cell Physiology* **40**, 725–732. https://doi.org/10.1093/oxfordjournals.pcp.a029599

Wang L., Ran L., Hou Y., Tian Q., Li C., Liu R., … Luo K. (2017*b*) The transcription factor MYB115 contributes to the regulation of proanthocyanidin biosynthesis and enhances fungal resistance in poplar. *New Phytologist* **215**, 351–367. https://doi.org/10.1111/nph.14569

Wang W.-Q., Ye J.-Q., Rogowska-Wrzesinska A., Wojdyla K.I., Jensen O.N., Møller I.M. and Song S.-Q. (2014) Proteomic Comparison between Maturation Drying and Prematurely Imposed Drying of Zea mays Seeds Reveals a Potential Role of Maturation Drying in Preparing Proteins for Seed Germination, Seedling Vigor, and Pathogen Resistance. *Journal of Proteome Research* **13**, 606–626. https://doi.org/10.1021/pr4007574

Wang Z., Gerstein M. and Snyder M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**, 57–63. https://doi.org/10.1038/nrg2484

Wasternack C. and Feussner I. (2018) The Oxylipin Pathways: Biochemistry and Function. *Annual Review of Plant Biology* **69**, 363–386. https://doi.org/10.1146/annurev-arplant-042817-040440

Waterworth, W. M., Footitt, S., Bray, C. M., Finch-Savage, W. E., and West, C. E. (2016). DNA damage checkpoint kinase ATM regulates germination and maintains genome stability in seeds. *Proceedings of the National Academy of Sciences*, *113*(34), 9647-9652. https://doi.org/10.1073/pnas.1608829113

Waterworth W.M., Bray C.M. and West C.E. (2015) The importance of safeguarding genome integrity in germination and seed longevity. *Journal of Experimental Botany* **66**, 3549–3558. https://doi.org/10.1093/jxb/erv080

Waterworth W.M., Bray C.M. and West C.E. (2019) Seeds and the Art of Genome Maintenance. *Frontiers in Plant Science* **10**. https://doi.org/10.3389/fpls.2019.00706

Waterworth W.M., Masnavi G., Bhardwaj R.M., Jiang Q., Bray C.M. and West C.E. (2010) A plant DNA ligase is an important determinant of seed longevity. *The Plant Journal* **63**, 848–860. https://doi.org/10.1111/j.1365-313X.2010.04285.x

Watson C.J.W., Froehlich J.E., Josefsson C.A., Chapple C., Durst F., Benveniste I. and Coolbaugh R.C. (2001) Localization of CYP86B1 in the Outer Envelope of Chloroplasts. *Plant and Cell Physiology* **42**, 873–878. https://doi.org/10.1093/pcp/pce110

Weigel D. (2012) Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant physiology* **158**(1), 2–22. https://doi.org/10.1104/pp.111.189845

Weigel D. and Glazebrook J. (2002) *Arabidopsis: A Laboratory Manual*. CSHL Press.

Weissbach H., Resnick L. and Brot N. (2005) Methionine sulfoxide reductases: history and cellular role in protecting against oxidative damage. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1703**, 203–212. https://doi.org/10.1016/j.bbapap.2004.10.004

Wellesen K., Durst F., Pinot F., Benveniste I., Nettesheim K., Wisman E., … Yephremov A. (2001) Functional analysis of the LACERATA gene of Arabidopsis provides evidence for different roles of fatty acid ω-hydroxylation in development. *Proceedings of the National Academy of Sciences* **98**, 9694–9699. https://doi.org/10.1073/pnas.171285998

Wendel A.A., Lewin T.M. and Coleman R.A. (2009) Glycerol-3-phosphate acyltransferases: rate limiting enzymes of triacylglycerol biosynthesis. *Biochimica Et Biophysica Acta* **1791**, 501–506. https://doi.org/10.1016/j.bbalip.2008.10.010

Weng H., Molina I., Shockey J. and Browse J. (2010) Organ fusion and defective cuticle function in a lacs1lacs2 double mutant of Arabidopsis. *Planta* **231**, 1089–1100. https://doi.org/10.1007/s00425-010-1110-4

Westengen O.T., Jeppson S. and Guarino L. (2013) Global Ex-Situ Crop Diversity Conservation and the Svalbard Global Seed Vault: Assessing the Current Status. *PLOS ONE* **8**, e64146. https://doi.org/10.1371/journal.pone.0064146

Western T.L. (2012) The sticky tale of seed coat mucilages: production, genetics, and role in seed germination and dispersal. *Seed Science Research* **22**, 1–25. https://doi.org/10.1017/S0960258511000249

Western T.L., Burn J., Tan W.L., Skinner D.J., Martin-McCaffrey L., Moffatt B.A. and Haughn G.W. (2001) Isolation and Characterization of Mutants Defective in Seed Coat Mucilage Secretory Cell Development in Arabidopsis. *Plant Physiology* **127**, 998–1011. https://doi.org/10.1104/pp.010410

Western T.L., Skinner D.J. and Haughn G.W. (2000) Differentiation of Mucilage Secretory Cells of the Arabidopsis Seed Coat. *Plant Physiology* **122**, 345–356. https://doi.org/10.1104/pp.122.2.345

Windsor J.B., Symonds V.V., Mendenhall J. and Lloyd A.M. (2000) Arabidopsis seed coat development: morphological differentiation of the outer integument. *The Plant Journal* **22**, 483–493. https://doi.org/10.1046/j.1365-313x.2000.00756.x

Winkel-Shirley B. (2001) Flavonoid Biosynthesis. A Colorful Model for Genetics, Biochemistry, Cell Biology, and Biotechnology. *Plant Physiology* **126**, 485–493. https://doi.org/10.1104/pp.126.2.485

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007) An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PloS one* **2**(8), e718. https://doi.org/10.1371/journal.pone.0000718

Woo H.R., Masclaux-Daubresse C. and Lim P.O. (2018) Plant senescence: how plants know when and how to die. *Journal of Experimental Botany* **69**, 715–718. https://doi.org/10.1093/jxb/ery011

Wood A.J. and Jenks M.A. (2008) Plant Desiccation Tolerance: Diversity, Distribution, and Real-World Application. In *Plant Desiccation Tolerance*. pp. 1–10. John Wiley and Sons, Ltd. https://doi.org/10.1002/9780470376881.ch1

Wu C.-W. and Storey K.B. (2016) Life in the cold: links between mammalian hibernation and longevity. *Biomolecular Concepts* **7**, 41–52. https://doi.org/10.1515/bmc-2015-0032

Xia Y., Gao Q.-M., Yu K., Lapchyk L., Navarre D., Hildebrand D., … Kachroo P. (2009) An intact cuticle in distal tissues is essential for the induction of systemic acquired resistance in plants. *Cell Host & Microbe* **5**, 151–165. https://doi.org/10.1016/j.chom.2009.01.001

Xie H.-T., Wan Z.-Y., Li S. and Zhang Y. (2014) Spatiotemporal Production of Reactive Oxygen Species by NADPH Oxidase Is Critical for Tapetal Programmed Cell Death and Pollen Development in Arabidopsis. The Plant Cell 26, 2007–2023. https://doi.org/10.1105/tpc.114.125427

Xu W., Dubos C. and Lepiniec L. (2015) Transcriptional control of flavonoid biosynthesis by MYB–

bHLH–WDR complexes. *Trends in Plant Science* **20**, 176–185. https://doi.org/10.1016/j.tplants.2014.12.001

Xue M., Guo T., Ren M., Wang Z., Tang K., Zhang W. and Wang M. (2019) Constitutive expression of chloroplast glycerol-3-phosphate acyltransferase from Ammopiptanthus mongolicus enhances unsaturation of chloroplast lipids and tolerance to chilling, freezing and oxidative stress in transgenic Arabidopsis. *Plant physiology and biochemistry: PPB* **143**, 375–387. https://doi.org/10.1016/j.plaphy.2019.07.019

Yadav V., Molina I., Ranathunge K., Castillo I.Q., Rothstein S.J. and Reed J.W. (2014) ABCG transporters are required for suberin and pollen wall extracellular barriers in Arabidopsis. *The Plant Cell* **26**, 3569–3588. https://doi.org/10.1105/tpc.114.129049

Yamaguchi S., Kamiya Y. and Nambara E. (2007) Regulation of ABA and GA Levels During Seed Development and Germination in Arabidopsis. In *Annual Plant Reviews Volume 27: Seed Development, Dormancy and Germination*. pp. 224–247. John Wiley and Sons, Ltd. https://doi.org/10.1002/9780470988848.ch9

Yang W., Pollard M., Li-Beisson Y., Beisson F., Feig M. and Ohlrogge J. (2010) A distinct type of glycerol-3-phosphate acyltransferase with sn-2 preference and phosphatase activity producing 2-monoacylglycerol. *Proceedings of the National Academy of Sciences* **107**, 12040–12045. https://doi.org/10.1073/pnas.0914149107

Yang W., Simpson J.P., Li-Beisson Y., Beisson F., Pollard M. and Ohlrogge J.B. (2012) A Land-Plant-Specific Glycerol-3-Phosphate Acyltransferase Family in Arabidopsis: Substrate Specificity, sn-2 Preference, and Evolution. *Plant Physiology* **160**, 638–652. https://doi.org/10.1104/pp.112.201996

Yang X., Zhang W., Dong M., Boubriak I. and Huang Z. (2011) The Achene Mucilage Hydrated in Desert Dew Assists Seed Cells in Maintaining DNA Integrity: Adaptive Strategy of Desert Plant Artemisia sphaerocephala. *PLOS ONE* **6**, e24346. https://doi.org/10.1371/journal.pone.0024346

Yant L., Mathieu J., Dinh T.T., Ott F., Lanz C., Wollmann H., … Schmid M. (2010) Orchestration of the Floral Transition and Floral Development in Arabidopsis by the Bifunctional Transcription Factor APETALA2. *The Plant Cell* **22**, 2156–2170. https://doi.org/10.1105/tpc.110.075606

Yashina S., Gubin S., Maksimovich S., Yashina A., Gakhova E. and Gilichinsky D. (2012) Regeneration of whole fertile plants from 30,000-y-old fruit tissue buried in Siberian permafrost. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 4008–4013. https://doi.org/10.1073/pnas.1118386109

Yeats T.H. and Rose J.K.C. (2013) The formation and function of plant cuticles. *Plant Physiology* **163**, 5–20. https://doi.org/10.1104/pp.113.222737

Yeats T.H., Martin L.B.B., Viart H.M.-F., Isaacson T., He Y., Zhao L., … Rose J.K.C. (2012) The identification of cutin synthase: formation of the plant polyester cutin. *Nature Chemical Biology* **8**, 609–611. https://doi.org/10.1038/nchembio.960

Zhang J. (2003) Evolution by gene duplication: an update. *Trends in Ecology and Evolution* **18**, 292–298. https://doi.org/10.1016/S0169-5347(03)00033-8

Zhang Y., Liu T., Meyer C.A., Eeckhoute J., Johnson D.S., Bernstein B.E., … Liu X.S. (2008) Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **9**, R137. https://doi.org/10.1186/gb-2008-9-9-r137

Zhao H., Nie K., Zhou H., Yan X., Zhan Q., Zheng Y. and Song C.-P. (2020) ABI5 modulates seed germination via feedback regulation of the expression of the PYR/PYL/RCAR ABA receptor genes. *New Phytologist* **228**, 596–608. https://doi.org/10.1111/nph.16713

Zhao L., Haslam T.M., Sonntag A., Molina I. and Kunst L. (2019) Functional Overlap of Long-Chain Acyl-CoA Synthetases in Arabidopsis. *Plant and Cell Physiology* **60**, 1041–1054. https://doi.org/10.1093/pcp/pcz019

Zheng Z., Xia Q., Dauk M., Shen W., Selvaraj G. and Zou J. (2003) Arabidopsis AtGPAT1, a Member of the Membrane-Bound Glycerol-3-Phosphate Acyltransferase Gene Family, Is Essential for Tapetum Differentiation and Male Fertility. *The Plant Cell* **15**, 1872–1887. https://doi.org/10.1105/tpc.012427

Zhou M.Q., Shen C., Wu L.H., Tang K.X. and Lin J. (2011) CBF-dependent signaling pathway: A key responder to low temperature stress in plants. *Critical Reviews in Biotechnology* **31**, 186–192. https://doi.org/10.3109/07388551.2010.505910

Zhu X., Zhang L., Kuang C., Guo Y., Huang C., Deng L., … Hua W. (2018) Important photosynthetic contribution of silique wall to seed yield-related traits in Arabidopsis thaliana. *Photosynthesis Research* **137**, 493–501. https://doi.org/10.1007/s11120-018-0532-x

Zieslin N. and Ben-Zaken R. (1992) Effects of applied auxin, gibberellin and cytokinin on the activity of peroxidases in the peduncles of rose flowers. *Plant Growth Regulation* **11**, 53–57. https://doi.org/10.1007/BF00024433

Zilber A.L. and Malkin R. (1988) Ferredoxin Cross-Links to a 22 kD Subunit of Photosystem I. *Plant Physiology* **88**, 810. https://doi.org/10.1104/pp.88.3.810

Zinsmeister J., Leprince O. and Buitink J. (2020) Molecular and environmental factors regulating seed longevity. *Biochemical Journal* **477**, 305–323. https://doi.org/10.1042/BCJ20190165

7. Attachments

Supplemental data

Supplemental data 1: Ecotype aging germination ratio obtained through the different seed aging treatments (NAT, AAT, CDT and EPPO). Germination ratios have been corrected by the control germination data to assess the seed germination reduction due to the seed-aging assay.

Data is publicly available at the araPheno repository (Seren *et al.*, 2017) under the study name Seed longevity (https://doi.org/10.21958/study:84). Data form the four seed aging treatments can be obtained from:

- NAT: https://doi.org/10.21958/phenotype:1307
- AAT: https://doi.org/10.21958/phenotype:1308
- CDT: https://doi.org/10.21958/phenotype:1309
- EPPO: https://doi.org/10.21958/phenotype:1310

Supplemental data 2*:* Correlation found between all four seed aging phenotypes (NAT, AAT, CDT, EPPO) and public phenotype data. The upper part corresponds to our data and correlations analysis, and the lower part correspond to correlations with public phenotypes, with an average of 70 coinciding plants. Only public correlations with Spearman correlation Rho (r_s) number above 0.3 (or below -0.3) in one seed aging treatment is listed. Public correlations only present an average of 70 coinciding plants. Blood letters indicate those correlations considered strong. Correlation analysis were performed with the easyGWAS comparison tool.

	NAT	***AAT***	***EPPO***	***CDT***	***Description***
NAT	**1.00**	0.11	**0.47**	**0.68**	Natural seed aging treatment
AAT	0.11	**1.00**	0.15	0.13	Artificial seed aging treatment
EPPO	0.47	0.15	**1.00**	**0.44**	Elevated partial pressure of oxygen seed aging treatment
CDT	**0.68**	0.13	**0.44**	1.00	Controlled deterioration treatment
Plant life span	**-0.35**	**-0.33**	-0.22	**-0.30**	Days from plant sowing until plant senesce
Longitude	0.04	0.21	0.16	0.11	Longitude at ecotype collection point
Latitude	-0.11	0.00	0.10	0.03	Latitude at ecotype collection point
MT GH	**0.32**	0.16	0.14	0.22	Number of days between last flower senescence and complete plant senescence
Secondary Dormancy	0.05	**0.35**	0.02	-0.01	Secondary dormancy was given by the slope between the germination percentages of non-dormant seeds after one and six weeks of cold treatment.

Trichome avg JA	0.07	-0.11	**0.3**	0.03	Measurements were collected from leaf disks removed from the 11th leaf of plants that had been treated with 0.6 ml of a 0.45mM solution of jasmonic acid (Sigma J-2500) in water (JA).
S34	-0.12	**0.39**	-0.04	-0.03	Sulfur concentrations in leaves, grown in soil. Elemental analysis was performed with an ICP-MS (PerkinElmer). Sample normalized to calculated weights as described in Baxter *et al.*, 2008
Emwa1	-0.03	**0.35**	-0.06	-0.07	All interactions were scored specifically on first true leaves as compatible, incompatible, or intermediate depending on the consistency of presence / absence of sporangiophores.
Leaf roll 16	-0.07	**0.35**	-0.04	-0.09	Results expressed as binary data, determined by the presence (1) or absence (0) of rolled leaves in all 4 plants / accession at 5 weeks growth
DW	0.22	**-0.32**	-0.04	0.03	Lesioning measured by dry weight
Germ in dark	-0.04	**-0.31**	0.04	-0.06	The ability to germinate in the dark at 4°C was measured as the percentage of non-dormant seeds that can germinate during 1-week long cold exposure, in the absence of light
Seed bank 133-91	**-0.33**	-0.09	0	-0.06	Genotypes consistently decreased their percentage of germination between 91 and 133 days of dry storage.
Ni60	**-0.33**	0.07	-0.09	-0.24	Nickel concentrations in leaves, grown in soil. Elemental analysis was performed with an ICP-MS (PerkinElmer). Sample normalized to calculated weights as described in Baxter *et al.*, 2008
8W GH LN	**-0.35**	-0.09	-0.08	-0.16	Leaf number was scored as the number of rosette leaves and the number of cauline leaves when the bolt reached 5cm
bio15	-0.01	**-0.33**	-0.23	-0.12	Precipitation Seasonality (Coefficient of Variation)
FT10	**-0.31**	-0.15	-0.09	-0.18	Plants were checked bi-weekly for presence of first buds, and the average flowering time of 4 plants of the same accession were collected
0W GH LN	**-0.38**	-0.25	0.13	-0.23	Leaf number was scored as the number of rosette leaves and the number of cauline leaves when the bolt reached 5cm
2W	-0.3	-0.16	-0.09	-0.21	Number of days required for the bolt height to reach 5cm
Fe56	**-0.32**	0.14	**-0.3**	**-0.44**	Iron concentrations in leaves, grown in soil. Elemental analysis was performed with an ICP-MS
0W	**-0.35**	-0.18	-0.14	-0.25	Number of days required for the bolt height to reach 5cm
0W GH FT	**-0.36**	-0.24	-0.14	-0.24	Flowering time was scored as the number of days for the bolt to reach 5cm
LN16	**-0.44**	-0.18	-0.17	-0.25	Plants were checked bi-weekly for presence of first buds, and the average leaf number at flowering time of 4 plants of the same accession were collected
LN22	**-0.33**	-0.27	-0.13	**-0.32**	Plants were checked bi-weekly for presence of first buds, and the average leaf number at flowering time of 4 plants of the same accession were collected
FT GH	**-0.4**	-0.21	-0.18	**-0.27**	Flowering time was scored as the number of days between germination date and appearance of the first flower
LC Duration GH	**-0.37**	-0.23	-0.23	**-0.25**	Number of days between germination and plant complete senescence
LFS GH	**-0.4**	-0.27	-0.21	**-0.26**	Number of days between germination and senescence of the last flower
FT22	**-0.35**	-0.28	-0.21	**-0.34**	Plants were checked bi-weekly for presence of first buds, and the average flowering time of 4 plants of the same accession were collected

DTF2	-0.28	**-0.37**	**-0.3**	-0.26	Flowering time was scored as days until the inflorescence stem elongated to 1 cm
DTF1	-0.28	**-0.35**	**-0.31**	-0.29	Flowering time was scored as days until the emergence of visible flowering buds in the center of the rosette from time of sowing
DTF3	**-0.3**	**-0.37**	**-0.3**	-0.27	Flowering time was scored as days until first open flower
FT16	**-0.49**	-0.24	-0.27	**-0.38**	Plants were checked bi-weekly for presence of first buds, and the average flowering time of 4 plants of the same accession were collected
RL	**-0.4**	**-0.37**	**-0.32**	**-0.36**	Rosette leaf number

Supplemental data 3: Genes distanced up to 1.5 Kb from significant SNPs (*p* value $\leq 10^{-5}$) among all four GWAS analysis: NAT-GWAS, AAT-GWAS, CDT-GWAS and EPPO-GWAS. Minor allele frequency was 0.05. The gene identifier is accompanied by the score of the SNP with max significance in brackets (-log(*p* value)).

NAT-GWAS: AT5G15810 (7.70), AT5G15815 (7.70), AT5G15820 (7.70), AT5G15740 (7.52), AT1G41820 (7.51), AT5G15860 (7.12), AT5G15870 (7.12), AT1G47620 (6.92), AT3G58330 (6.64), AT3G58340 (6.64), AT5G15830 (6.48), AT1G65330 (6.43), AT1G65340 (6.43), AT1G65342 (6.43), AT1G65346 (6.43), AT1G65350 (6.43), AT1G65360 (6.42), AT5G45120 (6.28), AT5G15800 (6.26), AT5G15802 (6.26), AT5G15805 (6.26), AT5G24160 (6.06), AT3G58130 (6.06), AT3G58140 (6.06), AT3G58165 (6.03), AT3G58170 (6.03), AT3G58180 (6.03), AT3G58310 (5.97), AT3G58320 (5.97), AT5G40645 (5.87), AT5G40650 (5.87), AT3G58160 (5.81), AT1G64880 (5.80), AT5G15840 (5.75), AT5G15845 (5.75), AT5G40660 (5.75), AT3G61010 (5.73), AT5G15750 (5.66), AT5G15760 (5.66), AT5G15770 (5.66), AT5G15780 (5.66), AT3G58260 (5.58), AT3G47170 (5.57), AT3G12130 (5.52), AT3G48840 (5.51), AT3G48850 (5.51), AT3G58193 (5.46), AT3G58196 (5.46), AT3G58200 (5.46), AT1G73820 (5.44), AT1G47810 (5.36), AT1G47813 (5.36), AT1G64780 (5.35), AT1G64790 (5.35), AT1G64810 (5.35), AT1G64820 (5.35), AT1G64840 (5.35), AT1G64850 (5.35), AT5G13800 (5.34), AT5G13810 (5.34), AT1G58602 (5.31), AT1G64800 (5.31), AT1G43320 (5.30), AT1G18900 (5.29), AT1G18910 (5.29), AT3G58080 (5.28), AT3G58090 (5.28), AT3G58100 (5.28), AT4G01010 (5.28), AT4G01020 (5.28), AT3G47160 (5.25), AT3G10790 (5.24), AT3G10800 (5.24), AT5G15900 (5.24), AT5G15910 (5.24), AT3G10600 (5.22), AT3G10610 (5.22), AT5G19850 (5.21), AT5G19855 (5.21), AT1G64830 (5.11), AT3G10160 (5.11), AT5G19780 (5.08), AT5G19790 (5.08), AT5G19800 (5.08), AT3G44100 (5.08), AT5G03795 (5.06), AT5G45740 (5.05), AT5G45745 (5.05), AT5G45750 (5.05), AT3G01670 (5.03), AT3G01680 (5.03), AT5G19860 (5.02), AT5G19870 (5.02), AT5G19875 (5.02), AT3G10300 (5.01), AT3G10310 (5.01), AT1G22710 (5.01), AT5G13730 (5.00) and AT5G13740 (5.00)

AAT-GWAS: AT1G07280 (10.33), AT1G29630 (8.45), AT1G29640 (8.45), AT1G29750 (7.75), AT1G29720 (7.61), AT1G29730 (7.61), AT1G29740 (7.61), AT1G29760 (7.61), AT1G30280 (7.60), AT1G30282 (7.60), AT4G18090 (7.48), AT4G18100 (7.48), AT4G18110 (7.48), AT1G29550 (7.34), AT1G30270 (7.23), AT3G42670 (7.21), AT4G00390 (7.12), AT4G00400 (7.12), AT1G29560 (7.07),

AT1G29570 (7.07), AT1G29690 (7.05), AT1G18710 (6.99), AT1G29590 (6.95), AT1G29600 (6.95), AT1G29620 (6.95), AT1G29540 (6.77), AT1G29430 (6.50), AT1G29440 (6.50), AT2G06050 (6.48), AT1G29400 (6.44), AT3G44900 (6.43), AT3G44910 (6.43), AT1G29410 (6.38), AT1G29418 (6.38), AT1G29420 (6.38), AT1G29660 (6.33), AT1G30200 (6.28), AT1G30210 (6.28), AT4G00380 (6.27), AT1G29370 (6.24), AT1G29380 (6.24), AT4G04510 (6.22), AT5G45530 (6.22), AT1G03600 (6.19), AT1G03610 (6.19), AT1G29700 (6.14), AT5G47740 (6.14), AT5G47750 (6.14), AT4G06676 (6.13), AT3G42660 (6.13), AT2G46440 (6.11), AT2G46400 (6.07), AT4G18120 (6.07), AT3G32150 (6.02), AT3G32160 (6.02), AT1G07350 (6.02), AT1G07360 (6.02), AT5G27870 (6.00), AT4G15980 (5.99), AT4G04450 (5.94), AT4G39700 (5.91), AT4G39710 (5.91), AT2G46410 (5.91), AT5G27150 (5.90), AT5G47850 (5.89), AT5G47860 (5.89), AT1G29530 (5.89), AT3G13280 (5.88), AT3G25120 (5.87), AT3G25130 (5.87), AT3G21070 (5.87), AT1G34000 (5.85), AT1G34010 (5.85), AT1G68130 (5.82), AT5G47890 (5.81), AT5G47900 (5.81), AT4G00231 (5.80), AT4G00232 (5.80), AT4G00234 (5.80), AT5G45510 (5.79), AT5G45520 (5.79), AT1G33880 (5.79), AT1G63840 (5.78), AT4G03728 (5.76), AT1G29450 (5.75), AT1G29460 (5.75), AT1G29465 (5.75), AT1G29470 (5.75), AT1G30160 (5.75), AT1G30170 (5.75), AT1G62510 (5.75), AT1G62500 (5.74), AT1G29350 (5.74), AT1G29355 (5.74), AT1G29357 (5.74), AT4G00335 (5.74), AT4G00340 (5.74), AT4G18130 (5.71), AT4G06701 (5.71), AT3G13275 (5.70), AT3G13277 (5.70), AT3G50420 (5.68), AT3G50430 (5.68), AT4G06599 (5.68), AT1G29580 (5.66), AT1G29395 (5.65), AT5G53210 (5.63), AT2G16720 (5.61), AT4G00430 (5.61), AT1G78880 (5.61), AT1G78882 (5.61), AT1G78890 (5.61), AT1G78955 (5.60), AT4G06634 (5.58), AT4G09647 (5.58), AT5G45540 (5.57), AT4G00310 (5.56), AT5G47960 (5.55), AT5G47970 (5.55), AT3G01930 (5.55), AT3G01940 (5.55), AT5G24790 (5.54), AT4G00450 (5.53), AT4G00460 (5.53), AT3G48110 (5.51), AT3G48115 (5.51), AT3G48120 (5.51), AT5G43450 (5.50), AT5G43455 (5.50), AT5G43460 (5.50), AT4G00270 (5.50), AT5G40760 (5.48), AT4G00238 (5.48), AT4G00240 (5.48), AT3G42728 (5.46), AT1G07370 (5.46), AT1G07380 (5.46), AT1G67040 (5.45), AT4G00070 (5.45), AT4G00080 (5.45), AT4G00085 (5.45), AT4G00090 (5.45), AT5G27230 (5.45), AT1G33850 (5.43), AT5G47820 (5.41), AT5G47830 (5.41), AT5G47840 (5.41), AT3G21090 (5.41), AT1G67035 (5.40), AT3G25110 (5.38), AT3G48050 (5.38), AT1G68120 (5.38), AT3G50440 (5.37), AT1G78870 (5.36), AT5G47950 (5.36), AT5G07810 (5.35), AT4G00120 (5.34), AT4G00124 (5.34), AT4G00130 (5.34), AT2G46390 (5.34), AT3G45810 (5.32), AT1G78895 (5.32), AT1G44050 (5.31), AT4G14746 (5.31), AT4G00330 (5.30), AT1G29480 (5.29), AT1G29490 (5.29), AT4G00230 (5.28), AT4G00315 (5.28), AT4G00320 (5.28), AT5G47910 (5.28), AT1G29950 (5.28), AT1G29951 (5.28), AT1G29952 (5.28), AT1G31080 (5.25), AT3G31430 (5.24), AT4G00280 (5.23), AT4G00290 (5.23), AT5G22470 (5.23), AT4G04460 (5.22), AT5G47940 (5.22), AT3G28860 (5.21), AT3G25099 (5.21), AT3G25100 (5.21), AT2G37435 (5.21), AT5G07830 (5.20), AT3G56900 (5.20), AT3G56910 (5.20), AT3G56920 (5.20), AT3G55260 (5.19), AT5G54320 (5.18), AT5G54330 (5.18), AT4G00300 (5.18), AT4G01360 (5.17), AT4G01370 (5.17), AT3G10150 (5.17), AT3G10160 (5.17), AT1G68140 (5.16), AT5G22510 (5.16), AT5G22520 (5.16), AT1G07310 (5.16), AT1G07320 (5.16), AT1G07330 (5.16), AT5G47455 (5.16), AT5G47460 (5.16), AT5G47480 (5.16), AT5G47490 (5.16), AT5G47500 (5.16), AT5G47510 (5.16), AT2G46420 (5.16), AT1G21980 (5.16), AT1G79150 (5.15), AT1G79160 (5.15), AT5G27080 (5.14), AT5G27090 (5.14), AT1G31250 (5.14), AT4G15370 (5.13), AT5G07180 (5.13), AT4G00305 (5.13), AT3G45800 (5.13), AT5G43590 (5.12), AT5G43600 (5.12), AT1G43910 (5.12), AT4G05612 (5.12),

AT5G61120 (5.09), AT5G61130 (5.09), AT2G36305 (5.09), AT4G04220 (5.09), AT4G04221 (5.09), AT4G04223 (5.09), AT1G63530 (5.09), AT1G63535 (5.09), AT1G63540 (5.09), AT1G03810 (5.09), AT1G03820 (5.09), AT1G03830 (5.09), AT4G06688 (5.08), AT5G07820 (5.08), AT5G62100 (5.08), AT5G62110 (5.08), AT1G07390 (5.07), AT1G65700 (5.06), AT1G22020 (5.06), AT1G22030 (5.06), AT1G67100 (5.06), AT1G07175 (5.05), AT1G07180 (5.05), AT1G07200 (5.05), AT1G07230 (5.05), AT1G07240 (5.05), AT3G24518 (5.05), AT3G24520 (5.05), AT1G07840 (5.04), AT1G07850 (5.04), AT1G34150 (5.04), AT3G55590 (5.03), AT3G55600 (5.03), AT5G60710 (5.03), AT5G63140 (5.03), AT5G63145 (5.03), AT5G63150 (5.03), AT4G06598 (5.02), AT5G53420 (5.02), AT5G47730 (5.01), AT5G43420 (5.01), AT5G43430 (5.01), AT5G43440 (5.01), AT5G18710 (5.01), AT5G18720 (5.01), AT1G07010 (5.00), AT1G07020 (5.00), AT4G35390 (5.00), AT5G27860 (5.00), AT1G07040 (5.00), AT1G07050 (5.00), AT1G07051 (5.00), AT1G07060 (5.00), AT1G07070 (5.00), AT3G54900 (5.00) and AT3G54910 (5.00)

CDT-GWAS: AT1G62990 (8.14), AT1G63000 (7.92), AT1G63005 (7.92), AT1G62981 (6.88), AT1G60490 (6.82), AT1G60030 (6.79), AT4G33070 (6.79), AT5G50450 (6.69), AT3G48050 (6.56), AT1G21810 (6.51), AT1G62980 (6.49), AT3G20410 (6.47), AT3G20420 (6.47), AT3G20430 (6.47), AT3G20440 (6.47), AT1G61180 (6.46), AT1G63010 (6.44), AT1G62975 (6.28), AT1G62978 (6.26), AT2G02955 (6.21), AT2G02960 (6.21), AT1G19840 (6.21), AT4G04750 (6.13), AT3G48040 (6.09), AT3G01490 (6.07), AT3G01500 (6.07), AT3G20395 (6.03), AT3G20400 (6.03), AT1G18840 (5.96), AT1G43580 (5.95), AT4G12080 (5.89), AT3G45555 (5.85), AT4G00752 (5.81), AT5G04160 (5.79), AT4G33080 (5.79), AT4G33090 (5.79), AT4G33100 (5.79), AT4G33110 (5.79), AT4G33120 (5.79), AT4G00190 (5.78), AT4G00200 (5.78), AT1G43190 (5.76), AT2G23230 (5.73), AT2G23240 (5.73), AT1G60025 (5.68), AT4G18050 (5.67), AT3G43710 (5.67), AT1G62970 (5.66), AT4G35070 (5.64), AT4G35080 (5.64), AT5G24710 (5.63), AT5G24735 (5.63), AT5G24740 (5.63), AT4G15680 (5.62), AT1G65480 (5.62), AT3G48030 (5.60), AT1G62950 (5.58), AT1G62960 (5.58), AT4G04760 (5.57), AT1G61190 (5.52), AT1G47310 (5.50), AT4G35740 (5.50), AT4G35750 (5.50), AT1G61260 (5.49), AT1G18850 (5.48), AT1G21830 (5.44), AT3G01930 (5.41), AT3G01940 (5.41), AT3G01510 (5.38), AT1G52290 (5.35), AT1G52300 (5.35), AT4G00165 (5.34), AT4G00170 (5.34), AT4G00180 (5.34), AT2G15770 (5.31), AT3G25940 (5.31), AT3G25950 (5.31), AT3G25960 (5.31), AT1G21760 (5.30), AT1G21770 (5.30), AT1G21780 (5.30), AT4G01270 (5.29), AT1G21050 (5.29), AT1G19580 (5.29), AT1G19600 (5.29), AT1G19480 (5.27), AT1G19485 (5.27), AT2G05940 (5.24), AT1G52280 (5.24), AT5G49310 (5.24), AT4G02750 (5.23), AT4G02760 (5.23), AT4G33050 (5.23), AT4G33060 (5.23), AT5G48410 (5.23), AT5G48412 (5.23), AT3G01950 (5.22), AT1G77920 (5.21), AT1G77930 (5.21), AT1G77932 (5.21), AT1G19490 (5.21), AT1G19500 (5.21), AT1G19510 (5.21), AT3G48940 (5.21), AT5G19370 (5.21), AT5G19380 (5.21), AT3G01513 (5.19), AT2G12475 (5.16), AT2G12480 (5.16), AT1G64690 (5.14), AT1G64700 (5.14), AT4G33160 (5.12), AT4G33170 (5.12), AT1G65730 (5.11), AT4G00700 (5.10), AT4G00710 (5.10), AT4G33130 (5.09), AT1G21790 (5.09), AT1G48030 (5.08), AT1G48040 (5.08), AT4G24840 (5.07), AT4G24860 (5.07), AT3G20450 (5.05), AT3G20350 (5.04), AT1G19570 (5.03), AT1G62940 (5.02), AT5G16390 (5.02), AT5G16400 (5.02), AT5G16410 (5.02) and AT1G61250 (5.00).

EPPO-GWAS: AT4G01703 (10.70), AT4G01710 (10.70), AT4G01700 (10.27), AT4G00950 (9.38), AT4G02750 (8.51), AT4G02760 (8.51), AT1G19600 (8.39), AT1G19610 (8.39), AT4G02770 (8.24), AT5G13550 (7.99), AT5G13560 (7.99), AT1G19580 (7.95), AT4G01330 (7.90), AT4G01335 (7.90), AT4G01340 (7.90), AT4G01350 (7.90), AT4G01360 (7.90), AT1G19620 (7.74), AT1G19630 (7.74), AT4G01130 (7.67), AT4G01260 (7.53), AT4G01240 (7.49), AT1G19570 (7.49), AT4G01720 (7.43), AT5G22350 (7.41), AT5G22355 (7.41), AT4G01100 (7.39), AT4G01110 (7.39), AT4G00955 (7.37), AT4G00960 (7.37), AT1G19800 (7.33), AT4G02730 (7.33), AT4G02733 (7.33), AT4G01220 (7.32), AT4G01230 (7.32), AT4G01250 (7.29), AT4G01170 (7.23), AT4G02740 (7.18), AT1G07550 (7.17), AT1G19790 (7.14), AT4G00980 (7.12), AT4G00985 (7.12), AT4G00990 (7.12), AT4G01690 (7.12), AT4G01120 (7.10), AT4G02780 (7.08), AT4G01180 (7.00), AT5G22400 (6.98), AT4G01210 (6.94), AT1G67350 (6.87), AT1G67360 (6.87), AT1G67365 (6.87), AT1G07560 (6.87), AT4G01290 (6.86), AT4G01370 (6.84), AT4G01380 (6.84), AT4G01390 (6.84), AT4G01090 (6.79), AT5G22410 (6.78), AT4G00975 (6.75), AT4G00970 (6.73), AT1G19470 (6.71), AT4G01000 (6.70), AT4G01010 (6.70), AT1G67370 (6.61), AT4G01245 (6.59), AT1G19640 (6.58), AT3G48050 (6.58), AT1G19490 (6.57), AT1G19500 (6.57), AT1G19510 (6.57), AT5G13530 (6.56), AT4G01070 (6.55), AT1G18980 (6.50), AT1G18990 (6.50), AT4G02050 (6.45), AT4G02055 (6.45), AT4G02060 (6.45), AT5G13590 (6.42), AT1G67620 (6.40), AT1G67623 (6.40), AT4G01140 (6.40), AT4G01150 (6.40), AT4G01160 (6.40), AT4G01355 (6.39), AT1G19750 (6.35), AT1G19770 (6.35), AT1G67430 (6.31), AT1G67440 (6.31), AT2G41930 (6.31), AT5G23280 (6.30), AT4G01400 (6.30), AT5G10140 (6.29), AT5G10150 (6.29), AT5G22315 (6.28), AT5G22320 (6.28), AT1G67340 (6.19), AT5G22330 (6.19), AT5G22340 (6.19), AT4G36390 (6.17), AT1G19240 (6.16), AT1G19250 (6.16), AT4G24972 (6.13), AT4G02860 (6.11), AT4G02870 (6.11), AT5G13380 (6.11), AT5G13390 (6.11), AT5G13400 (6.11), AT1G67650 (6.10), AT1G19480 (6.09), AT1G19485 (6.09), AT1G19520 (6.08), AT2G03270 (6.05), AT2G03280 (6.05), AT5G22420 (6.04), AT5G22430 (6.04), AT5G22440 (6.04), AT4G36450 (6.02), AT4G36460 (6.02), AT1G19740 (6.00), AT5G13410 (5.99), AT5G13420 (5.99), AT5G13570 (5.97), AT5G13580 (5.97), AT5G66360 (5.97), AT4G02700 (5.97), AT4G02710 (5.97), AT4G28080 (5.97), AT1G67500 (5.96), AT1G67740 (5.96), AT1G67750 (5.96), AT4G02070 (5.95), AT4G01740 (5.95), AT4G01750 (5.95), AT4G26370 (5.94), AT4G26375 (5.94), AT4G26380 (5.94), AT1G19464 (5.91), AT5G22490 (5.91), AT5G22360 (5.90), AT5G22370 (5.90), AT4G01080 (5.90), AT1G67390 (5.90), AT1G07850 (5.89), AT1G07860 (5.89), AT1G07870 (5.89), AT2G23230 (5.89), AT2G23240 (5.89), AT4G01450 (5.86), AT1G19460 (5.85), AT1G18530 (5.85), AT1G18540 (5.85), AT4G02830 (5.83), AT4G02840 (5.83), AT5G13370 (5.82), AT2G38310 (5.81), AT3G01790 (5.78), AT3G01800 (5.78), AT2G21140 (5.78), AT3G48060 (5.78), AT4G01440 (5.77), AT2G03470 (5.75), AT4G02100 (5.75), AT3G01810 (5.74), AT5G22470 (5.74), AT2G03290 (5.73), AT1G19830 (5.71), AT4G01990 (5.70), AT4G01995 (5.70), AT3G48040 (5.67), AT1G18870 (5.67), AT1G18871 (5.67), AT1G18570 (5.67), AT1G18660 (5.66), AT1G18670 (5.66), AT5G23220 (5.65), AT5G13490 (5.64), AT3G45850 (5.64), AT3G45851 (5.64), AT3G45860 (5.64), AT4G36380 (5.63), AT4G01730 (5.63), AT4G01735 (5.63), AT5G13610 (5.62), AT5G13620 (5.62), AT5G13630 (5.62), AT5G20240 (5.62), AT5G23200 (5.62), AT1G19540 (5.60), AT1G19550 (5.60), AT1G18680 (5.60), AT1G19710 (5.59), AT1G19715 (5.59), AT3G50682 (5.59),

AT3G50685 (5.59), AT3G50690 (5.59), AT1G18890 (5.57), AT5G23210 (5.57), AT5G22450 (5.57), AT1G19780 (5.53), AT1G67720 (5.53), AT1G67730 (5.53), AT4G02110 (5.53), AT4G02120 (5.53), AT4G02850 (5.52), AT5G22380 (5.49), AT4G01533 (5.48), AT4G08180 (5.47), AT4G08190 (5.47), AT5G22480 (5.47), AT4G02720 (5.46), AT4G02725 (5.46), AT5G23190 (5.46), AT4G08176 (5.45), AT5G65920 (5.44), AT4G00280 (5.43), AT4G00290 (5.43), AT4G01430 (5.43), AT2G01520 (5.42), AT1G07540 (5.41), AT3G50670 (5.40), AT3G50674 (5.40), AT1G19000 (5.40), AT1G67400 (5.40), AT1G67410 (5.40), AT5G23250 (5.40), AT3G50695 (5.39), AT5G23170 (5.38), AT5G23180 (5.38), AT5G22220 (5.37), AT5G22240 (5.37), AT5G22250 (5.37), AT5G23290 (5.35), AT1G19340 (5.34), AT1G19350 (5.34), AT4G02820 (5.33), AT4G02020 (5.32), AT4G02030 (5.32), AT4G02040 (5.32), AT3G22104 (5.31), AT5G13480 (5.30), AT1G19370 (5.30), AT1G19371 (5.30), AT4G28860 (5.30), AT4G28870 (5.30), AT4G28880 (5.30), AT5G46230 (5.29), AT5G46240 (5.29), AT4G28100 (5.29), AT4G28110 (5.29), AT5G13600 (5.29), AT4G02680 (5.28), AT4G02690 (5.28), AT1G18800 (5.27), AT4G09510 (5.27), AT4G09520 (5.27), AT1G07640 (5.27), AT1G19835 (5.27), AT5G38317 (5.26), AT5G38320 (5.26), AT1G07420 (5.26), AT1G07430 (5.26), AT1G07530 (5.26), AT5G14220 (5.24), AT3G14520 (5.23), AT3G55880 (5.23), AT4G08455 (5.23), AT1G18520 (5.22), AT1G13120 (5.21), AT1G13130 (5.21), AT5G10030 (5.21), AT5G10040 (5.21), AT1G30120 (5.20), AT1G30130 (5.20), AT5G13460 (5.19), AT5G13470 (5.19), AT4G02810 (5.19), AT4G02010 (5.18), AT4G02640 (5.18), AT4G02650 (5.18), AT4G02655 (5.18), AT5G45950 (5.18), AT1G67420 (5.18), AT1G19450 (5.17), AT4G26130 (5.16), AT4G26140 (5.16), AT1G19530 (5.16), AT1G65490 (5.16), AT1G19300 (5.16), AT4G02670 (5.15), AT5G23630 (5.14), AT5G23640 (5.14), AT5G16270 (5.12), AT5G16280 (5.12), AT5G16285 (5.12), AT5G16286 (5.12), AT5G16290 (5.12), AT5G24280 (5.11), AT5G24290 (5.11), AT4G36470 (5.11), AT4G01760 (5.11), AT1G67630 (5.11), AT1G67635 (5.11), AT1G21810 (5.11), AT1G18710 (5.10), AT5G13430 (5.09), AT1G19990 (5.09), AT1G20000 (5.09), AT1G20010 (5.09), AT4G01650 (5.09), AT4G01660 (5.09), AT1G21740 (5.08), AT5G23420 (5.08), AT4G28890 (5.08), AT1G19270 (5.08), AT5G47455 (5.06), AT5G47460 (5.06), AT5G47480 (5.06), AT5G47490 (5.06), AT5G47500 (5.06), AT5G47510 (5.06), AT4G09430 (5.06), AT1G18550 (5.05), AT1G18560 (5.05), AT4G08370 (5.04), AT5G16250 (5.04), AT5G16260 (5.04), AT4G15160 (5.04), AT4G15165 (5.04), AT1G18630 (5.04), AT1G18640 (5.04), AT4G01670 (5.03), AT4G01671 (5.03), AT5G16340 (5.03), AT5G16350 (5.03), AT4G02080 (5.02), AT1G18650 (5.01),

Supplemental data 4: AtHB25 ChIP-seq binding sites during Arabidopsis seed maturation. Results correspond to MACS2 peaks confirmed as AtHB25 ChIP-seq binding sites through graphical visualization. Last Colum indicated the number of AtHB25 motifs found in 200 bp from the peak centre obtained with the HOMER motif analysis tool (Heinz *et al.*, 2010)

Peak	*chr*	*start*	*end*	*length*	*-log(q value)*	*num motifs*
Peak1	Chr3	9835133	9835456	324	13.64	4
Peak2	Chr5	26078179	26078504	326	11.46	4
Peak3	Chr5	26822114	26822395	282	9.53	1
Peak4	Chr1	29734329	29734456	128	7.84	9
Peak5	Chr4	17050287	17050576	290	7.60	4
Peak6	Chr1	4348460	4348566	107	6.91	1
Peak7	Chr1	6062155	6062294	140	6.91	4
Peak8	Chr1	20210685	20210781	97	6.91	4
Peak9	Chr3	1632135	1632283	149	6.91	1
Peak10	Chr3	7096324	7096392	69	6.91	2
Peak11	Chr3	21680138	21680286	149	6.91	0
Peak12	Chr4	13652520	13653481	962	6.91	4
Peak13	Chr5	828780	828840	61	6.91	3
Peak14	Chr5	17591511	17591761	251	6.91	8
Peak15	Chr5	19308524	19308822	299	6.91	2
Peak16	Chr5	24904901	24905079	179	6.19	4
Peak17	Chr1	3278787	3278851	65	6.03	3
Peak18	Chr1	4610619	4610862	244	6.03	2
Peak19	Chr1	18296653	18296713	61	6.03	2
Peak20	Chr1	30249224	30249349	126	6.03	2
Peak21	Chr2	380123	380195	73	6.03	2
Peak22	Chr2	1050706	1050907	202	6.03	2
Peak23	Chr5	16069995	16070210	216	6.03	6
Peak24	Chr5	23346634	23346963	330	6.03	1
Peak25	Chr5	26853037	26853165	129	6.03	2
Peak26	Chr2	18258314	18258377	64	5.63	2
Peak27	Chr1	5195715	5196047	333	5.14	3
Peak28	Chr1	9879234	9879461	228	5.14	0
Peak29	Chr1	25514513	25514929	417	5.14	6
Peak30	Chr1	27646675	27646736	62	5.14	0
Peak31	Chr2	6392881	6393055	175	5.14	1
Peak32	Chr2	17728438	17728671	234	5.14	3
Peak33	Chr3	3725659	3725809	151	5.14	3
Peak34	Chr3	5081735	5081967	233	5.14	1

Peak35	Chr3	6346405	6346489	85	5.14	0
Peak36	Chr3	21065934	21066111	178	5.14	2
Peak37	Chr4	59283	59410	128	5.14	1
Peak38	Chr4	289659	289899	241	5.14	5
Peak39	Chr4	15798396	15798461	66	5.14	3
Peak40	Chr4	17709483	17709570	88	5.14	2
Peak41	Chr4	17774579	17774645	67	5.14	1
Peak42	Chr5	266305	266373	69	5.14	6
Peak43	Chr5	14888211	14888292	82	5.14	1
Peak44	Chr5	23166086	23166244	159	5.14	0
Peak45	Chr5	26653587	26653660	74	5.14	2
Peak46	Chr3	20571380	20571460	81	4.95	0
Peak47	Chr1	7265436	7265665	230	4.67	8
Peak48	Chr5	25948703	25948767	65	4.67	0
Peak49	Chr2	12173528	12173580	53	4.63	2
Peak50	Chr1	2082993	2083201	209	4.30	2
Peak51	Chr1	3081276	3081342	67	4.30	0
Peak52	Chr1	26240158	26240212	55	4.30	5
Peak53	Chr1	30156144	30156202	59	4.30	2
Peak54	Chr2	13085358	13085439	82	4.30	2
Peak55	Chr2	17643055	17643255	201	4.30	3
Peak56	Chr3	20941081	20941326	246	4.30	2
Peak57	Chr3	22815104	22815193	90	4.30	3
Peak58	Chr4	145842	145902	61	4.30	5
Peak59	Chr4	13083226	13083293	68	4.30	1
Peak60	Chr4	17018410	17018459	50	4.30	3
Peak61	Chr4	18406442	18406873	432	4.30	3
Peak62	Chr5	24635122	24635484	363	4.30	3
Peak63	Chr5	25633161	25633290	130	4.30	1
Peak64	Chr4	16431702	16431753	52	3.97	0
Peak65	Chr5	26239324	26239440	117	3.97	1
Peak66	Chr1	7146543	7146639	97	3.49	4
Peak67	Chr1	7909932	7909997	66	3.49	0
Peak68	Chr1	9451812	9451940	129	3.49	6
Peak69	Chr1	26439311	26439543	233	3.49	1
Peak70	Chr1	30225081	30225172	92	3.49	1
Peak71	Chr2	286476	286667	192	3.49	2
Peak72	Chr2	14999567	14999690	124	3.49	2
Peak73	Chr2	17268847	17269073	227	3.49	0
Peak74	Chr2	17778092	17778141	50	3.49	1

Peak75	Chr3	5018086	5018237	152	3.49	4
Peak76	Chr3	21715707	21715808	102	3.49	2
Peak77	Chr3	22436043	22436092	50	3.49	3
Peak78	Chr3	23324488	23324650	163	3.49	0
Peak79	Chr4	2620526	2620584	59	3.49	3
Peak80	Chr4	7962218	7962267	50	3.49	2
Peak81	Chr4	14141377	14141427	51	3.49	6
Peak82	Chr5	661910	661961	52	3.49	3
Peak83	Chr5	1420505	1420571	67	3.49	6
Peak84	Chr5	7257033	7257092	60	3.49	0
Peak85	Chr5	20786818	20786937	120	3.49	1
Peak86	Chr5	21968467	21968547	81	3.49	1
Peak87	Chr5	22156293	22156449	157	3.49	7
Peak88	Chr5	23267956	23268065	110	3.49	1
Peak89	Chr5	25872915	25873049	135	3.49	4
Peak90	Chr5	26171408	26171470	63	3.49	3
Peak91	Chr5	26251647	26251812	166	3.49	8
Peak92	Chr3	17316576	17316751	176	3.40	4
Peak93	Chr5	25477471	25477538	68	3.30	0
Peak94	Chr1	7140097	7140150	54	3.28	0
Peak95	Chr5	2710594	2710691	98	3.28	4
Peak96	Chr3	3304146	3304219	74	2.88	2
Peak97	Chr1	452936	453048	113	2.72	1
Peak98	Chr1	1877593	1877674	82	2.72	1
Peak99	Chr1	7420883	7421500	618	2.72	1
Peak100	Chr1	8558920	8559019	100	2.72	4
Peak101	Chr1	26137935	26138031	97	2.72	7
Peak102	Chr1	27241715	27241776	62	2.72	0
Peak103	Chr1	28423207	28423363	157	2.72	0
Peak104	Chr2	15608699	15608748	50	2.72	1
Peak105	Chr2	18910165	18910227	63	2.72	7
Peak106	Chr2	19027575	19027970	396	2.72	0
Peak107	Chr3	4672458	4672539	82	2.72	1
Peak108	Chr3	7089837	7089923	87	2.72	7
Peak109	Chr3	7934009	7934065	57	2.72	5
Peak110	Chr3	22805046	22805104	59	2.72	4
Peak111	Chr3	23089419	23089548	130	2.72	1
Peak112	Chr4	1407251	1407334	84	2.72	6
Peak113	Chr4	12502336	12502390	55	2.72	3
Peak114	Chr5	2821615	2821679	65	2.72	3

Peak115	Chr5	3473391	3473687	297	2.72	2
Peak116	Chr5	8732175	8732245	71	2.72	1
Peak117	Chr5	25880965	25881060	96	2.72	5
Peak118	Chr5	26281146	26281230	85	2.72	1
Peak119	Chr4	16696801	16696973	173	2.62	4
Peak120	Chr1	1064355	1064547	193	2.31	1
Peak121	Chr1	7781533	7781611	79	2.15	5
Peak122	Chr1	19299968	19300048	81	1.97	4
Peak123	Chr1	26097920	26097987	68	1.97	3
Peak124	Chr1	26120789	26120900	112	1.97	6
Peak125	Chr1	26400369	26400798	430	1.97	6
Peak126	Chr2	7899293	7899375	83	1.97	0
Peak127	Chr2	14489720	14489800	81	1.97	0
Peak128	Chr3	403279	403328	50	1.97	1
Peak129	Chr3	4595759	4595827	69	1.97	3
Peak130	Chr3	6797371	6797425	55	1.97	2
Peak131	Chr3	18883284	18883333	50	1.97	0
Peak132	Chr3	21410544	21410634	91	1.97	2
Peak133	Chr4	11897460	11897545	86	1.97	2
Peak134	Chr4	18119978	18120037	60	1.97	2
Peak135	Chr5	4874014	4874063	50	1.97	1
Peak136	Chr5	14998991	14999040	50	1.97	0
Peak137	Chr5	18972608	18972687	80	1.97	2
Peak138	Chr5	21698817	21698889	73	1.97	5
Peak139	Chr5	22307817	22307907	91	1.97	0
Peak140	Chr5	24887102	24887380	279	1.97	0
Peak141	Chr5	23577181	23577275	95	1.95	6
Peak142	Chr1	27541941	27541990	50	1.31	1
Peak143	Chr4	1143922	1144013	92	1.31	3
Peak144	Chr4	12616945	12617127	183	1.31	2
Peak145	Chr5	3457490	3457578	89	1.31	0
Peak146	Chr5	4633699	4634200	502	1.31	3

Supplemental data 5*:* Putative target genes of AtHB25. List of genes located within 3000 bp from AtHB25 binding sites in seeds according to TAIR10 genome annotation

AT3G26744, AT5G65240, AT5G65250, AT5G65260, AT5G67210, AT5G67220, AT5G67230, AT5G67240, AT1G79030, AT1G79040, AT1G79050, AT4G36040, AT4G36050, AT1G12750, AT1G12760, AT1G12770, AT1G17615, AT1G17620, AT1G17630, AT1G54130, AT3G05620, AT3G05625, AT3G05630, AT3G20340, AT3G20350, AT3G58620, AT4G27260, AT5G03370,

AT5G43760, AT5G43770, AT5G43780, AT5G43790, AT5G47630, AT5G47635, AT5G47640, AT5G47650, AT5G61990, AT1G10030, AT1G10040, AT1G10050, AT1G13440, AT1G13450, AT1G49430, AT1G49435, AT1G80460, AT2G03480, AT5G40170, AT5G40180, AT5G40190, AT5G40200, AT5G57640, AT5G57650, AT5G57655, AT5G67300, AT5G67310, AT2G44140, AT2G44150, AT1G15100, AT1G28260, AT1G68060, AT1G68070, AT1G68080, AT1G73530, AT1G73540, AT1G73550, AT2G14870, AT2G14880, AT2G42580, AT3G11773, AT3G11780, AT3G11800, AT3G11810, AT3G15090, AT3G15095, AT3G15110, AT3G18490, AT3G56890, AT3G56891, AT3G56900, AT4G00150, AT4G00700, AT4G00710, AT4G32730, AT4G32750, AT4G32760, AT4G37690, AT4G37700, AT4G37800, AT5G01700, AT5G01710, AT5G01712, AT5G01720, AT5G37480, AT5G37490, AT5G37500, AT5G57170, AT5G57180, AT5G66750, AT5G66760, AT3G55480, AT3G55490, AT1G20880, AT1G20890, AT5G64930, AT5G64940, AT2G28470, AT2G28480, AT1G06770, AT1G06780, AT1G09530, AT1G69750, AT1G69760, AT1G80170, AT1G80180, AT1G80190, AT2G30700, AT2G30710, AT2G42360, AT2G42370, AT3G56480, AT3G56490, AT3G56500, AT3G56510, AT3G61650, AT3G61660, AT4G00330, AT4G00335, AT4G00340, AT4G25650, AT4G25660, AT4G25670, AT4G25672, AT4G35930, AT4G39650, AT4G39660, AT5G61240, AT5G61250, AT5G61260, AT5G64050, AT5G64060, AT4G34350, AT4G34360, AT4G34370, AT5G65640, AT5G65650, AT1G20620, AT1G20630, AT1G22410, AT1G27200, AT1G70209, AT1G70210, AT1G80390, AT1G80400, AT1G80410, AT2G01640, AT2G01650, AT2G35670, AT2G35680, AT2G35690, AT2G41420, AT2G41430, AT2G41440, AT2G42680, AT2G42690, AT2G42700, AT3G14900, AT3G14910, AT3G14920, AT3G14930, AT3G58700, AT3G58710, AT3G58720, AT3G58730, AT3G60690, AT3G60700, AT3G63120, AT3G63130, AT3G63140, AT4G05100, AT4G13710, AT4G28630, AT4G28640, AT4G28650, AT5G02870, AT5G02880, AT5G04870, AT5G04885, AT5G21950, AT5G21960, AT5G51130, AT5G51140, AT5G51150, AT5G54140, AT5G54530, AT5G54531, AT5G54540, AT5G54550, AT5G57410, AT5G57420, AT5G64710, AT5G64720, AT5G64730, AT5G65460, AT5G65470, AT5G65670, AT3G46990, AT3G47000, AT3G47010, AT5G63630, AT5G63640, AT1G20610, AT5G08400, AT5G08410, AT5G08415, AT5G08420, AT3G10560, AT3G10570, AT3G10572, AT3G10580, AT1G02280, AT1G02290, AT1G02300, AT1G02305, AT1G06148, AT1G21170, AT1G21190, AT1G21200, AT1G24160, AT1G24170, AT1G24180, AT1G72350, AT1G72360, AT1G72370, AT1G75690, AT1G75700, AT2G37150, AT2G37160, AT2G45950, AT2G45960, AT2G45970, AT2G46340, AT2G46360, AT3G14090, AT3G14100, AT3G20320, AT3G20330, AT3G22400, AT3G22410, AT3G61620, AT3G61630, AT3G62390, AT3G62400, AT3G62410, AT3G62420, AT3G62422, AT4G03180, AT4G03190, AT4G03200, AT4G24050, AT4G24060, AT5G08670, AT5G08680, AT5G10980, AT5G64740, AT5G65700, AT4G35070, AT4G35080, AT1G04110, AT1G04120, AT1G22060, AT1G51940, AT1G69420, AT1G69430, AT1G69485, AT1G69490, AT1G70080, AT1G70090, AT1G70100, AT2G18160, AT2G18162, AT2G34330, AT2G34340, AT3G02170, AT3G02180, AT3G02190, AT3G13920, AT3G13930, AT3G19560, AT3G19570, AT3G50790, AT3G50800, AT3G57790, AT3G57800, AT4G22590, AT4G22592, AT4G22600, AT4G38810, AT4G38820, AT4G38825, AT4G38830, AT5G15050, AT5G15060, AT5G15070, AT5G37750, AT5G37760, AT5G37770, AT5G46750, AT5G46760, AT5G53460, AT5G54920, AT5G54930, AT5G54940, AT5G54950, AT5G54960, AT5G61960,

AT5G61970, AT5G58300, AT5G58310, AT5G58320, AT5G58330, AT1G73220, AT1G73230, AT1G73240, AT4G02600, AT4G24390, AT4G24400, AT5G10946, AT5G10950, AT5G14360, AT5G14370 and AT5G14380.

.

Supplemental data 6: Motif enrichment analysis of AtHB25-binding regions in 200 bp around the peak centre. Analysis was performed with HOMER (Heinz *et al.*, 2010). Background genome was randomly sampled from Arabidopsis. *p* value cut-off was set at 0.0001.

Motif Name	*Sequence*	*Target Sequences*	*Background Sequences*	*p value*
ATHB25(ZFHD)	TRATTAVB	82.99%	37.64%	10^{-28}
ATHB33(ZFHD)	NGTRATTAAK	79.59%	38.30%	10^{-23}
ATHB23(ZFHD)	HTAATTARNN	76.87%	31.45%	10^{-28}
ATHB34(ZFHD)	TAATTARS	74.83%	28.36%	10^{-30}
ATHB24(ZFHD)	TAATTAAS	72.11%	27.04%	10^{-29}
AT1G20910(ARID)	THAATTRAWN	66.67%	50.05%	10^{-04}
CRC(C2C2YABBY)	TWATSATA	51.70%	35.49%	10^{-04}
At1g76110(ARID)	ATTTAATG	36.73%	21.98%	10^{-04}
SPCH(bHLH)	WNBCACGTGA	29.93%	11.14%	10^{-09}
PHV(HB)	RTAATSATTA	29.93%	11.80%	10^{-08}
HY5(bZIP)	RRTSACGTSD	27.89%	12.56%	10^{-06}
IBL1(bHLH)	CACGTGCC	21.77%	9.42%	10^{05}
PCF	NNWWWTGGGCYTDDN	19.73%	6.87%	10^{-06}
bZIP3(bZIP)	DWKNHSACGTGGCAD	18.37%	6.10%	10^{-06}
TGA6(bZIP)	TGACGTCABC	18.37%	7.86%	10^{-04}
BIM2(bHLH)	NNNNCACGTGNN	18.37%	7.97%	10^{-04}
ABF1(bZIP)	CACGTGGC	17.69%	5.68%	10^{-06}
PIF5ox(bHLH)	BCACGTGVDN	17.69%	7.26%	10^{-04}
bZIP53(bZIP)	NDNHSACGTGKMNNN	15.65%	4.67%	10^{-06}
GBF3(bZIP)	TGCCACGTSAYC	14.97%	4.32%	10^{-06}
AREB3(bZIP)	NKGMCACGTGDCMNN	14.97%	4.37%	10^{-06}
CAMTA1(CAMTA)	WWAACGCGTT	14.97%	5.46%	10^{-04}
ZML2(C2C2gata)	CATCATCATC	14.29%	4.19%	10^{-05}
ABI5(bZIP)	GCCACGTG	14.29%	4.25%	10^{-05}
E-box	GCCACGTG	14.29%	4.36%	10^{-05}
bZIP68(bZIP)	WGCCACGTGK	14.29%	4.59%	10^{-05}
CAMTA5(CAMTA)	ACGCGTTTTANACRC	14.29%	4.77%	10^{-05}
GBF6(bZIP)	WWTGMCACGTCABCW	13.61%	3.36%	10^{-06}
bZIP28(bZIP)	TGCCACGTSABH	13.61%	3.59%	10^{-06}

bZIP16(bZIP)	TGCCACGTGD	13.61%	3.99%	10^{-05}
GBF5(bZIP)	WKNWSACGTGGCAWN	13.61%	4.09%	10^{-05}
ABF2(bZIP)	KGMCACGTGDCMHHH	12.93%	3.01%	10^{-06}
FHY3(FAR1)	HHCACGCGCBTN	9.52%	1.75%	10^{-06}
At1g78700(BZR)	NNNNCACGTGNNNNN	9.52%	2.84%	10^{-04}
ZML1(C2C2gata)	ATCWYRACCGTTSRW	7.48%	1.22%	10^{-05}
FAR1(FAR1)	TKNNNYYCACGCGCY	6.80%	0.72%	10^{-06}
ZIM(C2C2gata)	ATCSRACGGTYRAGA	6.80%	1.12%	10^{-05}
PIF7(bHLH)	CCACGTGGNH	6.80%	1.38%	10^{-04}
AT5G59990 C2C2COL	TCTCAACCGTTCATT	6.12%	1.04%	10^{-04}
bZIP44(bZIP)	NDTGCCACGTCAGCH	4.76%	0.55%	10^{-04}
AT4G27900 C2C2COL	TCTCVACCGTTSATT	4.08%	0.47%	10^{-04}
bZIP69(bZIP)	GACAGCTGKCAW	2.72%	0.15%	10^{-04}

Supplemental data 7: Differentially expressed genes in *cog1-2D*

Data is publicly available in the GEO repository **GSE128953** as a Supplementary file.

Supplemental data 8: GO terms significantly over-represented in the sets of DEGs in *cog1-2D* using a significance level of 0.01 FDR)

Biological Process	***Description***	***FDR***
GO:0050896	response to stimulus	$7.6 \cdot 10^{-27}$
GO:0006950	response to stress	$2.6 \cdot 10^{-15}$
GO:0042221	response to chemical stimulus	$8.7 \cdot 10^{-14}$
GO:0006952	defense response	$1.9 \cdot 10^{-10}$
GO:0010033	response to organic substance	$5 \cdot 10^{-10}$
GO:0051707	response to other organism	$1.7 \cdot 10^{-8}$
GO:0044283	small molecule biosynthetic process	$6.6 \cdot 10^{-8}$
GO:0019748	secondary metabolic process	$6.6 \cdot 10^{-8}$
GO:0065007	biological regulation	0.00000023
GO:0009607	response to biotic stimulus	0.00000024
GO:0009628	response to abiotic stimulus	0.00000079
GO:0006725	cellular aromatic compound metabolic process	0.000003
GO:0002376	immune system process	0.0000037
GO:0006955	immune response	0.0000037
GO:0045087	innate immune response	0.0000045
GO:0044281	small molecule metabolic process	0.0000064
GO:0006575	cellular amino acid derivative metabolic process	0.000017

GO:0008152	metabolic process	0.00002
GO:0009987	cellular process	0.00002
GO:0009719	response to endogenous stimulus	0.000022
GO:0051704	multi-organism process	0.00003
GO:0050789	regulation of biological process	0.000034
GO:0051716	cellular response to stimulus	0.000035
GO:0009698	phenylpropanoid metabolic process	0.000038
GO:0042398	cellular amino acid derivative biosynthetic process	0.000038
GO:0009605	response to external stimulus	0.00004
GO:0006519	cellular amino acid and derivative metabolic process	0.000059
GO:0009699	phenylpropanoid biosynthetic process	0.00012
GO:0055114	oxidation reduction	0.00012
GO:0019438	aromatic compound biosynthetic process	0.00012
GO:0050794	regulation of cellular process	0.00014
GO:0006790	sulfur metabolic process	0.00023
GO:0016143	S-glycoside metabolic process	0.00031
GO:0019757	glycosinolate metabolic process	0.00031
GO:0019760	glucosinolate metabolic process	0.00031
GO:0009725	response to hormone stimulus	0.00033
GO:0033554	cellular response to stress	0.00036
GO:0009617	response to bacterium	0.00036
GO:0009743	response to carbohydrate stimulus	0.00037
GO:0016137	glycoside metabolic process	0.00055
GO:0023052	signaling	0.00073
GO:0009814	defense response, incompatible interaction	0.00097
GO:0080134	regulation of response to stress	0.001
GO:0010200	response to chitin	0.0014
GO:0015698	norganic anion transport	0.0014
GO:0009266	response to temperature stimulus	0.0015
GO:0009611	response to wounding	0.0015
GO:0010817	regulation of hormone levels	0.0021
GO:0008219	cell death	0.0021
GO:0031347	regulation of defense response	0.0021
GO:0016265	death	0.0021
GO:0012501	programmed cell death	0.0022
GO:0009414	response to water deprivation	0.0022
GO:0046394	carboxylic acid biosynthetic process	0.0024
GO:0016053	organic acid biosynthetic process	0.0024
GO:0044237	cellular metabolic process	0.0025
GO:0042434	indole derivative metabolic process	0.0027

GO:0042430	indole and derivative metabolic process	0.0027
GO:0006820	anion transport	0.0027
GO:0042742	defense response to bacterium	0.0031
GO:0009415	response to water	0.0034
GO:0034641	cellular nitrogen compound metabolic process	0.0034
GO:0006629	lipid metabolic process	0.0036
GO:0042435	indole derivative biosynthetic process	0.0036
GO:0048583	regulation of response to stimulus	0.0036
GO:0016138	glycoside biosynthetic process	0.0039
GO:0044272	sulfur compound biosynthetic process	0.0042
GO:0019761	glucosinolate biosynthetic process	0.0044
GO:0019758	glycosinolate biosynthetic process	0.0044
GO:0016144	S-glycoside biosynthetic process	0.0044
GO:0009851	auxin biosynthetic process	0.0048
GO:0043067	regulation of programmed cell death	0.0053
GO:0002682	regulation of immune system process	0.0054
GO:0050776	regulation of immune response	0.0054
GO:0008610	lipid biosynthetic process	0.0055
GO:0009684	indoleacetic acid biosynthetic process	0.0056
GO:0048513	organ development	0.006
GO:0048731	system development	0.006
GO:0080135	regulation of cellular response to stress	0.0061
GO:0009683	indoleacetic acid metabolic process	0.0061
GO:0045088	regulation of innate immune response	0.0061
GO:0009626	plant-type hypersensitive response	0.0061
GO:0034050	host programmed cell death induced by symbiont	0.0065
GO:0010054	trichoblast differentiation	0.0067
GO:0010941	regulation of cell death	0.0078
GO:0006970	response to osmotic stress	0.0083
GO:0010363	regulation of plant-type hypersensitive response	0.0085
GO:0010053	root epidermal cell differentiation	0.0086
GO:0015706	nitrate transport	0.009
GO:0009627	systemic acquired resistance	0.0094

Mollecular Function	***Description***	***FDR***
GO:0016491	oxidoreductase activity	0.0000009
GO:0003824	catalytic activity	0.000001
GO:0043531	ADP binding	0.000021
GO:0005506	iron ion binding	0.0083

Cellular Component	*Description*	*FDR*
GO:0005886	plasma membrane	0.0000064
GO:0030312	external encapsulating structure	0.0000064
GO:0005618	cell wall	0.0000064
GO:0016020	membrane	0.000029
GO:0009505	plant-type cell wall	0.000088
GO:0005623	cell	0.000088
GO:0044464	cell part	0.000088
GO:0009536	plastid	0.0065

Supplemental data 9: Differentially Expressed Genes classified as "phenylpropanoid biosynthetic process" (GO:0009699) in *cog1-2D*.

AT5G05270, AT3G03990, AT4G14690, AT4G25310, AT5G04720, AT1G50740, AT3G61220, AT1G15950, AT4G05590, AT2G26170, AT4G19840, AT1G66800, AT1G80820, AT5G19730, AT1G19960, AT1G64160, AT1G17745, AT1G78570, AT5G25890, AT2G23910, AT1G16310, AT1G51990, AT2G21730, AT5G52250, AT1G77520, AT1G75040, AT3G51240, AT4G37980, AT5G65870, AT3G44550, AT3G44720, AT5G13930, AT3G18280, AT4G27430, AT4G38550, AT1G09480, AT5G60020, AT1G13930, AT4G09820, AT2G43800, AT5G22630, AT5G22500, AT4G00040, AT1G31710, AT5G08640, AT3G11000, AT3G55470, AT1G79270, AT2G30930, AT3G23000, AT4G25300, AT4G11190, AT4G11210, AT2G39430, AT5G15950, AT5G07990, AT1G62660, AT1G65840, AT1G65960, AT2G24540, AT2G28660, AT5G63580, AT5G66230, AT3G44540, AT5G17220, AT3G28220, AT1G09500, AT3G13610, AT3G11430, AT1G61810 and AT2G02990.

Supplemental data 10: Differentially Expressed Genes classified as "phenylpropanoid biosynthetic process" (ath00940) in Arabidopsis thaliana in KEGG analysis in *cog1-2D*.

AT1G24735, AT1G26560, AT1G45191, AT1G48130, AT2G21730, AT2G44460, AT3G49110, AT3G49120, AT3G6014 and AT4G11290.

Supplemental data 11: Primers used

Cloning primers for plant transformation

proAtHB25-F	TGCTCTAGAGTGTCTCCACGTAGGAAAGCG
proAtHB25-R	CATGCCATGGTCAAGAAGTCGAGAAATGTGAAAC
proCOG1-F	TGCTCTAGATCATCCAATCATGAAAAAGGTAC
proCOG1-R	CTTGGGTCCATGGTAAAATGAGAGAGAGATAGAG
proMYB47-F	ATTATATTTAATCTCCATGTGGTCG
proMYB47-R	TTTCTCCAGCCAACCAAAC
AtHB25-pCM262-F	CCGGATCAATTCGGGGGATCAGTTTAAACGCGGGCCCATG-GAGTTTGAAGACAACAAC
AtHB25-pCM262-R	CCTGGTGATCCGTCGACCTGCAGGCGGCCGCGTGGTTG-GTCTTGTTCATGATG
COG1-pCM262-F	CCGGATCAATTCGGGGGATCAGTTTAAACGCGGGCCCATGGC-GACCCAAGATTC
COG1-pCM262-R	CCTGGTGATCCGTCGACCTGCAGGCGGCCGCGACAAGATT-GACCATCGGTG
MBY47-F	ATGGGGAGGACGACAT
MYB47-R	TCAAAAGAGATGATCAAGTATGTC
AtHB25-F	ATGGAGTTTGAAGACAACAAC
AtHB25-R	TCATGGTTGGTCTTGTTCATG
AtHB25-BAMH1-F	GGATCCATGGAGTTTGAAGACAACAAC
AtHB25-BAMH1-R	TCATGGTTGGTCTTGTTCATGGGATCC
pCM262-R	GATAGCCTCCTGCATAGTCCG
proUBQ10-F	GTCGACGAGTCAGTAATAAAC
GPF-R	GTCGTGCCGCTTCATATGATCAGG
GUS-F	AGTGTACGTATCACCGTTTGTGTGAAC
GUS-R	ATCGCCGCTTTGGACATACCATCCGTA
BAR-F	GTCTGCACCATCGTCAACC
BAR-R	GAAGTCCAGCTGCCAGAAAC
proAtHB25fr-pYRO-F	GAAGATCTTTCAGACCTTGAAATTATTAAG
proAtHB25fr-pYRO-F	GGGAAGCTTAAAGGGATCCTTCTGCAATAAAGTGATGGGC

Cloning primers for yeast assays

proAtHB25fr-F	TCCCCCCGGGGGGATTCAGACCTTGAAATTATTAAG
proAtHB25fr-R	GCTCTAGAGCTCTGCAATAAAGTGATGGGC
pTUY1H-seq-F	CACGAGGCCCTTTCGTCTTC
HB25-SfiI-F	CTGGCCATTACGGCCATGGAGTTTGAAGACAACAAC
HB25-bait-SfiI-R	TTAGGCCGAGGCGGCCAGTGGTTGGTCTTGTTCATGATG
HB25-prey-SfiI-R	GGCCGAGGCGGCCTCATGGTTGGTCTTGTTCATG
COG1-SfiI-F	CTGGCCATTACGGCCATGGCGACCCAAGATTCTC
COG1-bait-SfiI-R	TTAGGCCGAGGCGGCCAGACAAGATTGACCATCGGTG

COG1-prey-SfiI-R	GGCCGAGGCGGCCTTAACAAGATTGACCATCGGTG
MYB47-SfiI-F	CTGGCCATTACGGCCATGGGGAGGACGACA
MYB47-bait-SfiI-R	TTAGGCCGAGGCGGCCAAAGAGATGATCAAGTATG
MYB47-prey-SfiI-R	GGCCGAGGCGGCCTCAAAAGAGATGATCAAGTATG
AP2-SfiI-F	CTGGCCATTACGGCCAATGTGGGATCTAAACGACGC
AP2-prey-SfiI-R	GGCCGAGGCGGCCTCAAGAAGGTCTCATGAGAGGAG
TT2-SfiI-F	CTGGCCATTACGGCCATGGGAAAGAGAGCAACTACTAG
TT2-prey-SfiI-R	GGCCGAGGCGGCCTCAACAAGTGAAGTCTCGGAG
TT8-SfiI-F	CTGGCCATTACGGCCATGGATGAATCAAGTATTATTCC
TT8-prey-SfiI-R	GGCCGAGGCGGCCCTATAGATTAGTATCATGTATTATGACTTG
TT16-SfiI-F	CTGGCCATTACGGCCATGGGTAGAGGGAAGATAGAGATAAAG
TT16-prey-SfiI-R	GGCCGAGGCGGCCCTTAATCATTCTGGGCCGTTG
GL2-SfiI-F	CTGGCCATTACGGCCATGAAGTCGATCGATGGCTG
GL2-prey-SfiI-R	GGCCGAGGCGGCCTCAGCAATCTTCGATTTGTAGAC
TTG1-SfiI-F	CTGGCCATTACGGCCATGGATAATTCAGCTCCAGA
TTG1-prey-SfiI-R	GGCCGAGGCGGCCTCAAACTCTAAGGAGCTGCA
TTG2-SfiI-F	CTGGCCATTACGGCCATGGAGGTGA ATGATGGTG
TTG2-prey-SfiI-R	GGCCGAGGCGGCCTCAAATTGTTTGCTTAGAAAGTTG
EGL3-SfiI-F	CTGGCCATTACGGCCATGGCAACCGGAGAAAAC
EGL3-prey-SfiI-R	GGCCGAGGCGGCCTTAACATATCCATGCAACCC
MYB5-SfiI-F	CTGGCCATTACGGCCATGATGTCAT GTGGTGGG
MYB5-prey-SfiI-R	GGCCGAGGCGGCCCTAGTCATGTCCTAAGCTAGAAGAC
MYB61- SfiI-F	CTGGCCATTACGGCCATGGGGAGAC ATTCTTGC
MYB61- SfiI-R	GGCCGAGGCGGCC CTAAAGGGACTGACCAAAAG
FUS3-SfiI-F	CTGGCCATTACGGCCATGATGGTTGATGAAAATG
FUS3-prey-SfiI-R	GGCCGAGGCGGCCCTAGTAGAAGTCATCGAGAGAG
LEC1-SfiI-F	CTGGCCATTACGGCCATGGAACGTGGAGCTC
LEC1-prey-SfiI-R	GGCCGAGGCGGCCTCACTTATACTGACCATAATG
LEC2-SfiI-F	CTGGCCATTACGGCCATGGATAACTTCTTACCCTTTC
LEC2-prey-SfiI-R	GGCCGAGGCGGCCTCACCACCACTCAAAGTC
WRKY42-SfiI-F	CTGGCCATTACGGCCATGTTTCGTTTTCCGG
WRKY42-prey-SfiI-R	GGCCGAGGCGGCCTTATTGCCTATTGTCAACG
WRKY46-SfiI-F	CTGGCCATTACGGCCATGATGATGGAAGAGAAACTTG
WRKY46-prey-SfiI-R	GGCCGAGGCGGCCCTACGACCACAACCAATCC
SPCH-SfiI-F	CTGGCCATTACGGCCATGCAGGAGATAATACCG
SPCH-prey-SfiI-R	GGCCGAGGCGGCCCTAGCAGAATGTTTGCTG
SEP3-SfiI-F	CTGGCCATTACGGCCATGGGAAGAGGGAGAGTAG
SEP3-prey-SfiI-R	GGCCGAGGCGGCCTCAAATAGAGTTGGTGTCATAAG
KNAT7-SfiI-F	CTGGCCATTACGGCCATGCAAGAAGCGGCACTAGG
KNAT7-prey-SfiI-R	GGCCGAGGCGGCCTTAGTGTTTGCGCTTGGACTTC

FLY1-SfiI-F	CTGGCCATTACGGCCATGAAGAAGCGGGAACATTTG
FLY1-prey-SfiI-R	GGCCGAGGCGGCCCTATGCTGGAGGAAGAGACCG
STK-SfiI-F	CTGGCCATTACGGCCATGCTCTTTCCCCATG
STK -prey-R(SfiI)	GGCCGAGGCGGCCTTATCCGAGATGAAGAATTTTC

Sequencing and PCR comprobation primers

T7	TAATACGACTCACTATAGGG
M13-F	TGTAAAACGACGGCCAGT
M13-R	CAGGAAACAGCTATGACC
pCM262-R	GATAGCCTCCTGCATAGTCCG
proUBQ10-F	GTCGACGAGTCAGTAATAAAC
proUBQ10-2-F	CTCAATTCTCTCTACCGTGATCAAG
HA-R	CCWACRTCRTANGGRTA
pYRO-seq-R	GGGGTCGACGGAATTGGATGAAGAAG
pTUY1H-seq-F	CACGAGGCCCTTTCGTCTTC
GPF-R	GTCGTGCCGCTTCATATGATCAGG
GUS-F	AGTGTACGTATCACCGTTTGTGTGAAC
GUS-R	ATCGCCGCTTTGGACATACCATCCGTA
BAR-F	GTCTGCACCATCGTCAACC
BAR-R	GAAGTCCAGCTGCCAGAAAC

Expression analysis primers

MYB47-qRT-F	TCTCGGAAACAGGTGGGCAG
MYB47-qRT-R	GATGATGGGCTCGTGGGTCA
LACS2-qRT-F	CGTGCCGAGAGGAGAGATTT
LACS2-qRT-R	TCTGAGCAATGAGGGGACATC
NHO1-qRT-F	TGTGCTGGAGAGCATGTGTT
NHO1-qRT-R	CTATGTCCACTGGCCTCACC
FAR1-qRT-F	GTCTCAACCGCGTATATA
FAR1-qRT-R	AAGCTTTGCCCTTGCCAT
FAR4-qRT-F	CCTTGATCTTATACCTGT
FAR4-qRT-R	GATCATGAATCTGTTCGA
FAR5-qRT-F	CGGAATGGAAAGGGCCAAG
FAR5-qRT-R	CGCTGTCGACAGTTCTCAA
ABCG11-qRT-F	GGAACAGCTGCATATGTG
ABCG11-qRT-R	CCCATCTCTATTATTGTC
KCS20-qRT-F	CATTTCCTCTCCGCCACACT
KCS20-qRT-R	GGGACGGGTGGTGAAGTAAG
CYP86A8-qRT-F	GAAACACCACGACGATCTCTTA
CYP86A8-qRT-R	GTCCAGCTAGGATGAAGTTGAG

GPAT8-qRT-F	CACCAAGTCGCCTTTCTCG
GPAT8-qRT-R	CCCCTGGAACTCTTGGCAA
FACT-qRT-F	GTGGGATAGAGAAGATAAAG
FACT-qRT-R	TCCATGGCTGCGATACCAT
AtHB25-qRT-R	TCACAAGCTCCTCCTCTTCA
AtHB25-qRT-F	TAATCCCTCCTCCTGCTTCA
ICE1-qRT-F	GGCCAGCAAGCTAGAGTTGA
ICE1-qRT-R	ATGGTAGCGAGCAACAGACC
COG1-qRT-F	CGAGCTAGATGCTTTGCTAGTG
COG1-qRT-R	GTGTAACAACGGAGCCTCTTC
PRX2-qRT-F	CGCCCTTGTGGTTCTTCTTA
PRX2-qRT-R	CTGTGACACCACGAACAATTTC
PRX3-qRT-F	GTGTATCTCACTGCTCGTCTTT
PRX3-qRT-R	CTTGAGATTGGCTGCGTACT
PRX12-qRT-F	CTATCAGGAGGGCCAGACTAT
PRX12-qRT-R	TTGAAGAACGGTGGTGGTAAG
PRX22-qRT-F	TTCGGCTCGAGGGTTTAATG
PRX22-qRT-R	GAGAGGCGATGGTGAGAATATC
PRX23-qRT-F	CAGGACTTGCCCACCTATTT
PRX23-qRT-R	GAACAAAGCAGTCATGGAAGTG
PRX25-qRT-F	GCGACGGAAACGCTATTCTA
PRX25-qRT-R	CTCTAAGACGACTCGCATACTTC
PRX36-qRT-F	AGCTCTGTAGCCTCTCTATCC
PRX36-qRT-R	TGCGAGGGTCATTGAAGTATG
PRX43-qRT-F	CCATACTTCTACGGCTCCATTT
PRX43-qRT-R	GCTTTGACGGCTTCTACTATCT
PRX50-qRT-F	GCCACTCTACGCCTCTATTT
PRX50-qRT-R	CATGATCCTTCTCCGCCTTATT
PRX55-qRT-F	CAACCCAGATGCTGTAGTAGAC
PRX55-qRT-R	GAGCTTGATCGGAGGTGAAA
PRX64-qRT-F	GGAGGTTGACCCAACACTAAA
PRX64-qRT-R	TCACGGTTCCATCCATGTTC
PRX71-qRT-F	GGGACGTAGAGATGGTAGAGTT
PRX71-qRT-R	GACGACGAGATCACGAGTATTG
PP2A3-qRT-F	ACCTGCGGTAATAACTGCATCTA
PP2A3-qRT-R	CCGAACATCAACATCTGGGTC
AT5G55840-qRT-F	GATGATATTGCAGTTTGTCACCGT
AT5G55840-qRT-R	CACTGTCTTGCTTGTCTTGTTCTG
UBQ10-qRT-F	TGTGTTTTGGGGCCTTGTAT
UBQ10-qRT-R	CAAGTTTCGCAGAACTGCAC

Mutant genotyping primers

psad1-LP	ACCAACACCTTCTCTTCCTGG
psad1-RP	CCAACCCAACATTCAAACAAC
sslea-LP	CATCTCCTGCTTTGCGTTTAG
sslea-RP	TTCCAAATCGTGTGGAAGATC
sstpr-LP	AACCTGCAATCATTGTGTTCC
sstpr-RP	AGTTGTTCGATGAAATGCCTG
rbohd-LP	TTTGATGCCAAACTCCAAGTC
rbohd-RP	GCCGCCTAAGACTTTCTAAGC
rbohe-LP	ATTTTGGCGGCTTAAACTTTG
rbohe-RP	ATTGACTTACCAACGCACTCG
rbohf-LP	ATTCGGTTCGACTGCTTAAGG
rbohf-RP	TTACCCGTTGGTCAAGTTCTG
cyp86a1-LP	AAGTCCTTTCCCAAGCAAGTC
cyp86a1-RP	GGATAATCCCTTCTCTGTCGC
cyp86a2-LP	ATCGAACACATGCTCAAGACC
cyp86a2-RP	GAATTCCAAGCAATCCTCTCC
cyp86a8-LP	TTGCTCCGGTACGTATCAGAC
cyp86a8-RP	TCCGGTAACACATCGTCTTTC
cyp86b1-LP	TTCTGGTCCAAACACAGGTTC
cyp86b1-RP	GTTTCTTGGGATGCTTCCTTC
cyp86b2-LP	TAATCCCCTCCAAAAGATTCG
cyp86b2-RP	CCTCGTCCGGATTACTTCTTC
cyp86c1-LP	ATGTTGGCGCTAAACAAACAC
cyp86c1-RP	AATCCTAGCGCCTTTCTTCAC
myb47-LP	CCTCTTTTGCTTGATCACTGC
myb47-RP	AAAATGATTTCCATCTTTTAACGG
spch-LP	TTTCCCTTTGCAATATGCAAC
spch-RP	GGACTTCGGCGTAGGTTTTAC
knat7-LP	GAGATTAGTGTTTGCGCTTGG
knat7-RP	TATGCGTAAGGGCATATCAGG
sep3-RP	GTTGGGAAAATCGTACGAGGCTTCACCTAGT
sep3-LP	CGTCACTTGCCTATTGATCTTGTT
lacs2-RP	GTACTCGAGACTTCCCCGAAG
lacs2-LP	GAGGCTGTGAAGAAATATCCG
nho1-RP	CATAGGCTGCTCCTAATGCTG
nho1-LP	AAGACGCTATCAAGAAAGGGG
kcs20-RP	TCATGGACAGATCTCAACGTG
kcs20-LP	AAACTTGCAAGCACGTGTTTC
gpat5-RP	CCAGCGAGAACCCTATACTTATCT

gpat5-LP	CTAAGGAGCATCTTAGAGCAGATGA
myb107-RP	GCTGAATTGGACTGGTGAGAC
myb107-LP	TTGATGTGTTTTCTGATTTTTGC
prx12 -LP	TTAATGCGACATGTCCTATTGC
prx12 -RP	AATAGGTGGACCACTGTCACG
prx21-LP	TTTTTGGTGGAACATGAAAATG
prx21-RP	CGCAGAAAGAGCGACAATATC
prx36-LP	AGATTAAAGAGAAGCTGCCGG
prx36-RP	CAAGGCAGACTTGATCTCGTC
prx55-LP	CGACATGTCCTAGCGTAGAGC
prx55-RP	CCTGAAAAATAGGTCAACCCC
prx71-LP	TCATCTATCCCCCACAAACAC
prx71-RP	TGTCCGGTGGTATTGAATAGC
ice1-2 RP	AATCTGATGGCTGAGAGGAGAA
ice1-2 LP	TTCATGGTAGCGAGCAACAGAC
SALK-LB	ATTTTGCCGATTTCGGAAC
SAIL-LB	GCTTCCTATTATATCTTCCCAAATTACCAATACA
GABI-LB	ATATTGACCATCATACTCATTGC
sep3-transposon	GAGCGTCGGTCCCCACACTTCTATAC

ChIP-qRT primers

AtHB25peak1 -F	TCAAAGTGTAATTAGCATCATTCAT
AtHB25peak1-R	ATATGAGTTAAAGAAATGCATGGAA
AtHB25peak5- F	GCGAAGAGGCGATTAGTCAT
AtHB25peak5- R	CCCGACCCGTAACTTAATCC
AtHB25peak12- F	GCACGAATAGTAAGCGTTCA
AtHB25peak12- R	GACACAATAAGGAAGATGAAGGG
AtHB25peak19- F	GGTGTCCGGAGTCCATAAAT
AtHB25peak19- R	CATCCGTTTTGTCGTCGAG
AtHB25peak20- F	CTTGACCATCCACCAATAAAAAG
AtHB25peak20- R	CGTTCTCTATCCATCTACGCA

Extended protocols

Seed RNA extraction

For siliques, developing seeds and dry seeds.

Collect and freeze the material

- 5-10 developing siliques
- 20-30 siliques to extract developing seeds (soft smash and filtering with liquid N_2)
- 25 mg of dry seeds

Buffers and reagents

- Seed extraction buffer: 0.4M LiCl, 0.2M Tris pH 8, 25mM EDTA, 1% SDS
- Chloroform : Isoamyl Alcohol (24:1)
- Phenol : Chloroform : Isoamyl Alcohol (PCI) (25:24:1)
- LiCl 6M
- Sodium acetate 3M, pH 5.2
- Ethanol (100% and 70%)
- DNAse
- H_2O miliQ RNAse free
- Liquid N_2

Recommendations

* Use gloves during the RNA extraction
* Sterilize previously all tips and microtubes
* Work with ice
* Separate properly organic solvents

Protocol

1 - Prepare 2 mL microtubes with 550 µL of seed extraction buffer and 550 µL of chloroform and keep them in ice

2 - Grind the tissue frozen in N_2, transfer the sample and vortex 10s

3 - Centrifuge at 10000 rpm for 3 min at 4 ºC. Transfer the aqueous phase to a new 2 mL microtube

4 - Add 700µL of PCI and vortex 10 seconds

5 - Centrifuge at 10000 rpm for 3 min at 4 ºC. Transfer the aqueous phase to a new 1.5 mL microtube

6 - Add 1/2 volume of LiCl 6 M and mix

7 - Precipitate at -20ºC for 1h and centrifuge at 4 ºC for 30 min

8 - Discard the supernatant and resuspend the pellet in 500 µL of H_2O

9 - Add 7 µL of sodium acetate 3 M pH 5.2 and 250 µL of ethanol 100% and vortex briefly

10 - Centrifuge at 4ºC for 10 min to precipitate carbohydrates

11 - Transfer the supernatant phase to a new 1.5 mL microtube

12 - Add 43 µL of sodium acetate 3 M pH 5.2 and 750µL of ethanol 100% and vortex briefly

13 - Precipitate at -20ºC for 1h and centrifuge at 4 ºC for 30 min

14 - Discard the supernatant and clean the pellet with 1 mL of ethanol 70%

15 - Discard the supernatant and dry pellet at room temperature

16 - Resuspend in H_2O and proceed to DNAse treatment as manufacturer

- ***a.*** Use Kit columns for the DNAse treatment (fix properly the RNA into the column with the appropriate buffer and follow manufacter's instructions)
- ***b.*** Liquid DNAse treatment and precipitate again (10% of NaAc 3 M pH 5.2 and 2 volumes ethanol 100%)

17 - Measure RNA concentration (above 50ng/µL may be good enough)

18 - Store RNA DNA-free in a -80ºC freezer

Chromatin Immunoprecipitation

For sequencing or qRT-PCR

Collect the material

- 3-5 grams of transgenic tissue of plants overexpression the recombinant-tagged protein of interest. The tissue of choice determines the chromatin state, which determine the accessibility of the DNA-binding protein. In addition, the coexpression with different cofactors might also determine diverse genomic protein-binding sites.

Buffers and reagents

- Extraction Buffer 1 (0.4 M sucrose, 10 mM Tris-HCl pH8, 10 mM $MgCl_2$, 5 mM β-mercaptoethanol, Protease Inhibitor (1 cOmplete, EDTA-free (Roche) tablet each 50 mL). Cold in ice.
- Extraction Buffer 2 (0.25 M sucrose, 10 mM Tris-HCl pH8, 10 mM $MgCl_2$, 1% Triton X-100, 5 mM β-mercaptoethanol, Protease Inhibitor tablet). Cold in ice.
- Extraction Buffer 3 (1.7 M sucrose, 10 mM Tris-HCl pH8, 10 mM $MgCl_2$, 1.5% Triton X-100, 5 mM β-mercaptoethanol, Protease Inhibitor tablet). Cold in ice.
- Nuclei lysis buffer (50mM Tris-HCl pH8, 10 mM EDTA, 1% SDS, Protease Inhibitor tablet). Make fresh. Do not place in ice.
- ChIP dilution Buffer (1.1% Triton X-100, 1.2 mM EDTA, 16.7 Tris-HCl pH8, 167 mM NaCl). Make fresh.
- Low salt wash buffer (150 mM NaCl, 0.1% SDS, 1% TritonX-100, 2mM EDTA, 20mM Tris-HCl pH8). Cold in ice when indicated.
- High salt wash buffer (500 mM NaCl, 0.1% SDS, 1 % TritonX-100, 2mM EDTA, 20mM Tris-HCl pH8). Cold in ice when indicated.
- LiCl wash buffer (0.25 M LiCl, 1 % NP40, 1% sodium deoxycholate, 1mM EDTA, 10mM Tris-HCl pH 8). Cold in ice when indicated.
- TE buffer (10 mM Tris-HCl pH 8, 1mM EDTA). Cold in ice when indicated.
- Elution buffer (1% SDS, 0.1 $NaHCO_3$). Make fresh when indicated. Cold in ice.
- Specific antibody for the recombinant-tag
- Dynabeads® protein G (or A, depend on the antibody)
- Magnet (for microtubes)
- Formaldehyde
- Glycine
- Chloroform : Isoamyl Alcohol (24:1)
- NaCl 5M
- Sodium acetate 3M, pH 5.2
- Ethanol (100% and 70%)
- Proteinase K
- Glycogen
- Liquid N_2

Recommendations

- Use gloves to avoid sample contamination (NGS are really sensitive)
- Use filter tips and sterile microtubes
- Use clean MilliQ water to prepare indicated solutions, preferably from commercial flask or sterilized water to avoid DNA contamination.

Protocol

1 - Transfer the plant material (~3-5 g) in 50 mL tube and wash it with 40 mL of Milli-Q water by gently shacking the tube. Repeat once

2 - After removing as much water as possible, add 37mls of 1 % formaldehyde (in water) (freshly prepared) covering completely the plant material. Crosslink in vacuum for 15 minutes. At this stage, seedlings should become water-soaked and translucent. Stop the crosslinking by adding Glycine to a final concentration of 0.125M (2.5 mL of 2M Glycine in 37 mL of (1% formaldehyde). Put under vacuum for additional 5 minutes

3 - Grind the plant material in liquid Nitrogen to a fine powder (1-2 grams of tissue)

4 - Add the powder to 30 mL Extraction Buffer in a 50 mL. Let sit on ice for 5 min

5 - Filter the solution through Miracloth into a fresh 50 mL tube. Repeat once

6 - Spin the filtered solution for 20 minutes at 4 ºC and 4000 rpm

7 - Gently remove supernatant and resuspend the pellet in 1 mL of Extraction Buffer 2. Transfer the solution to a 1.5 mL microtube

8 - Centrifuge at 12000 g for 10 minutes at 4 ºC

9 - Remove supernatant and resuspend pellet in 300ul of Extraction Buffer 3

10 - In a clean Eppendorf, add 300 µL of Extraction Buffer 3. Take the 300 µL solution (resuspended pellet) from step 16 carefully layer it on top of the clean 300 µL of Extraction Buffer 3.

11 - Spin for 1 hour at 16000g at 4 ºC. Make Nuclei Lysis Buffer and ChIP dilution buffer at this stage. You will need 10 mLs of ChIP dilution buffer for 3 IPs, and only 5 mL of Nuclei Lysis Buffer. Repeat steps 10 and 11 if the pellet is green. Normally 3-4 repetitions are needed.

12 - Resuspend the pellet by pipetting-up and down and vortexing (keep solution cold between vortexing). Keep a 10 µL aliquot to run on a gel (crude and deproteinized).

13 - Once resuspended, sonicate the chromatin solution (sonicator program: 30 s ON, 30 s OFF, 9 cycles). Keep 5 µL to run on a gel to check sonication efficiency (crude and deproteinized). Following this step, the chromatin solution can be frozen at -20 ºC.

14 - Spin at full speed the chromatin solution for 5 minutes at 4 degrees to pellet debris. Remove supernatant to a new tube.

15 - Remove 25 µL from for Total DNA INPUT Control

16 - Bring volume up to 3 mL with ChIP Dilution Buffer. The point here is to dilute the 1% SDS to 0.1 % SDS with ChIP dilution buffer.

17 - In three new tubes, rinse three times and resuspend 40 µL of Dynabeads® protein G.

18 - Split the chromatin solution for each genotype in those three tubes (1 mL each) to preclear the chromatin solution1 hour at 4 ºC with rotation to eliminate unspecific bead-DNA binding.

19 - Separate the chromatin solution plus beads with a magnet

20 - Label a new set of tubes with the appropriate antibody label and the no antibody control far the IP. Transfer the supernatant from previous step to a newly labeled tube. Be careful not to add any beads

21 - Add 5 µL of your antibody (concentrations may vary between antibodies and should be determined empirically) to two of the three tubes. Do not add antibody to the third tube as this the No Antibody control. Incubate for 4 hours or overnight at 4ºC with gentle agitation

22 - Collect immune complexes with 50 µL Dynabeads® protein G (rinsed first in ChIP dilution buffer as before) for one hour at 4 ºC gentle agitation. Again, prepare beads in a new set of labeled tubes and add IP solutions to these new tubes

23 - Prepare Elution Buffer during the incubation of the previous step and put cleaning and elution solutions in ice.

24 - Pellet beads with the magnet and wash eight times, for 5 minutes per wash, using the sequence of buffers listed below. Use 1 mL of each buffer per wash and wash at 4 degrees with gentle agitation. Magnet Dynabeads and remove supernatant.

- ***a.*** Low Salt Wash Buffer (Two washes: One quick, second for 5 minutes)
- ***b.*** High Salt Wash Buffer (Two washes: One quick, second for 5 minutes)
- ***c.*** LiCl Wash Buffer: (Two washes: One quick, second for 5 minutes)
- ***d.*** TE Buffer (Two washes: One quick, second for 5 minutes)

25 - Transfer the beads to a new set of tubes. Elute immunocomplexes by adding 250 µL of Elution buffer. Vortex briefly to mix and incubate at 65 ºC for 15 minutes with gentle agitation (mix tubes during incubation). Magnet beads and carefully transfer the supernatant fraction (eluate) to another tube and repeat elution. Combine the two eluates.

26 - Get the INPUT DNA out of the freezer. Add 500 µL of Elution buffer and include it in next steps

27 - Add 20 µL 5M NaCl to the eluate and reverse crosslink at 65 ºC for at least 6 hours or overnight.

28 - Take a 10 µL aliquot for a western analysis (SDS-PAGE) to ensure the presence of the protein. Is preferably to use a different antibody to avoid the detection of the antibody present in the sample.

29 - Add 10 µL of 0.5M EDTA, 20 µL Tris-HCl 1M pH 6.5, and 2 µL of 10mg/ml proteinase K to the eluate and incubate for one hour at 45º C

30 - Recover DNA by chloroform extraction (equal volume) and ethanol precipitation with 1/10 salt. Must split tubes for each precipitation to make room for ethanol. Add 2 µL glycogen 20mg/ml) to ethanol precipitation step. Wash pellets with 70% ethanol. Phenol extraction makes SDS insoluble defaulting the extraction. It possible to use commercial kits for a better DNA purification.

31 - For seed and silique samples a carbohydrate purification step may be applied:

- ***a.*** Resuspend the pellet in 500 µL of H_2O
- ***b.*** Add 7 µL of sodium acetate 3M pH 5.2 and 250 µL of ethanol 100% and vortex briefly
- ***c.*** Centrifuge at 4ºC for 10 min to precipitate carbohydrates
- ***d.*** Transfer the supernatant phase to a new 1.5 mL microtube
- ***e.*** Add 43 µL of sodium acetate 3M pH 5.2 and 750µL of ethanol 100% and vortex briefly
- ***f.*** Precipitate at -20ºC for 1h and centrifuge at 4 ºC for 30 min

32 - Resuspend the pellets in 50 µL of distilled water. Combine eluates form the same treatments

33 - Check IP efficiency by qPCR it there is a known DNA-binding sequence. Use from 0.5 µL to 2 µL in a 20 µL qPCR reaction. Product amplification should not exceed 200pb

34 - Use Qubit to quantify and verify the presence of DNA in the sample. Nanodrop does not work with this small amount of DNA. It is normal that the No Antibody sample present similar DNA amounts as antibodies samples, as major part of DNA is protein unspecific.

35 - Send samples to a NGS facility with dry ice.

Traducción divulgativa al Castellano

Identificación de genes relacionados con la longevidad de semilla en *Arabidopsis thaliana* mediante el uso de técnicas genómicas

Introducción general

El envejecimiento en los organismos

La vida es gloriosa, pero no es eterna. Todos los organismos vivos, desde los unicelulares a los más complejos, como las plantas y los animales, envejecen hasta la muerte. El término envejecimiento engloba todos los cambios que se producen en los organismos durante el transcurso de su vida. Parte del envejecimiento es la senescencia, término que se refiere a todos los efectos dañinos del envejecimiento que llevan a la muerte al organismo en ausencia de otros factores externos como enfermedades, estrés o daño físico.

La inmortalidad biológica no existe, aunque hay ejemplos de organismos que se rozan la vida eterna. Por ejemplo, las bacterias, comúnmente son consideradas inmortales ya que se dividen por mitosis, creando dos copias aparentemente idénticas. Sin embargo, la división no es simétrica y el daño se acumula principalmente en la célula vieja. Por otra parte, los organismos pluricelulares pueden reducir el envejecimiento celular mediante la regeneración de tejidos, produciendo individuos aparentemente inmortales, como la Hydra y ciertos tipos de medusas. Además, ciertas plantas pueden crecer indefinidamente pudiendo resultar en ejemplares milenarios.

Hay dos teorías principales sobre la razón del envejecimiento. La teoría del envejecimiento programado postula que los organismos envejecen y mueren para evitar la competencia con futuras generaciones, mientras que la teoría del daño postula que el envejecimiento se debe a acumulaciones de imperfecciones en el metabolismo, ya que se carece de selección evolutiva después de la reproducción. Experimentos en diferentes organismos modelos han mostrado evidencias de ambas teorías del envejecimiento, por lo que están aún en discusión. Un factor común a ambas teorías es que apuntan al daño oxidativo como mayor agente causante del envejecimiento, causado por el

metabolismo del organismo. Por lo tanto, el envejecimiento es el coste de la vida y las estrategias para aumentar la longevidad de los organismos, incesablemente buscadas por los humanos, pasan por reducir el metabolismo.

Diversos factores físicos externos influyen en la velocidad del envejecimiento, ya que modulan la velocidad del metabolismo o causan daño directo en los organismos. Directamente por la termodinámica, la temperatura de los organismos determina la velocidad de los procesos metabólicos, y por tanto del envejecimiento. El ciclo de vida de organismos de sangre fría directamente correlaciona con la temperatura exterior, siempre y cuando no sean temperaturas extremas y dañinas. En mamíferos y otros organismos de sangre caliente, también se observa este efecto, aunque de forma más sutil. Por ejemplo, en épocas de hibernación o si hay restricción calórica (limitación de la alimentación sin causar malnutrición), la temperatura corporal se reduce ligeramente y se reduce la velocidad del envejecimiento. De hecho, la restricción calórica es una de las estrategias más prometedoras para extender la vida humana.

Temperaturas muy altas causan desnaturalización proteica y ruptura de membranas celulares, mientras que temperaturas muy bajas inhiben a los enzimas y al desarrollo de los organismos. Bajo cero, la difusión, y por lo tanto el metabolismo, se para. Aunque el proceso es dañino por la alta concentración de solutos y ruptura de estructuras celulares por la formación de cristales, ciertos organismos y células toleran la congelación y pueden permanecer viables en zonas polares durante decenas de miles de años.

La deshidratación también ralentiza el envejecimiento de forma similar a la congelación, ya que la falta de agua hace que la difusión de moléculas ocurra lentamente. Sin embargo, son pocos los organismos que toleran la falta de agua. No obstante, los organismos que la toleran pueden permanecer viables durante décadas. En parte, la colonización de la tierra por los organismos pasó por tolerar el desecado, como pasó en las plantas. Esta tolerancia se perdió más adelante, y solo unas pocas pudieron recuperarla. No obstante, muchas especies desarrollan semillas tolerantes a la desecación. Estas semillas tienen una mayor longevidad en comparación a las que no toleran la desecación, pudiendo permanecer viables durante años, décadas y hasta miles de años, dependiendo de la especie y de las condiciones de almacenamiento.

También se ha visto que la concentración de oxígeno en la atmósfera afecta a la velocidad del envejecimiento, ya que afecta directamente al daño oxidativo. Sin embargo, el oxígeno, aunque dañino, es necesario para la vida aerobia. Otros factores que también afectan el envejecimiento son la luz ultravioleta y las radiaciones ionizantes, ya que pueden producir daños directos en macromoléculas y estructuras celulares.

La longevidad de semillas

Las semillas son un recurso interesante para investigar la longevidad, ya que los eventos de envejecimiento ocurren lentamente a baja humedad, a diferencia de otros organismos. Debido a que los mecanismos moleculares principales están conservados en eucariotas, el estudio de la longevidad en semillas puede conducir a descubrir rutas metabólicas conservadas implicadas y proveer un mayor conocimiento para la discusión de las teorías del envejecimiento.

Además, el estudio de la longevidad de las semillas es de vital importancia para la conservación de la biodiversidad vegetal *ex situ* con importantes ventajas como el reducido espacio requerido. Sin lugar a duda, la mejor forma de conservar la biodiversidad es conservando los ecosistemas en todo el globo terráqueo. Sin embargo, el ser humano viene destrozando multitud de ecosistemas reduciéndolos hasta su destrucción. Conservar las semillas en los bancos de semillas, o ejemplares de plantas en los jardines botánicos es de vital importancia para la conservación de la rica biodiversidad vegetal. La semilla, además, constituye la principal estrategia de propagación vegetal, y obtener unas semillas vigorosas y conservarlas adecuadamente es de vital importancia para la agricultura y la economía asociada.

La conservación óptima de las semillas requiere condiciones de baja humedad y temperatura para ralentizar los eventos dañinos que las envejecen. Como se puede observar en la Figura 1 en la página 17, el almacenamiento en condiciones óptimas (*Seed bank chamber*) de las semillas de la planta modelo *Arabidopsis thaliana* evita el deterioro progresivo éstas que se observa en las almacenadas a temperatura ambiente (*Ambient storage*). Esto hace que las semillas se puedan llegar a almacenar durante décadas en los bancos de semillas. El mayor de ellos es el Banco Mundial de Semillas de Svalbard, en Noruega, y se construyó con el objetivo de conservar los recursos genéticos en forma de semillas ante posibles catástrofes bélicas o globales. Su localización en zona ártica permite mantener eficientemente las bajas temperaturas incluso si la refrigeración falla. Sin embargo, todas las semillas, aunque almacenadas en condiciones óptimas, terminan perdiendo su capacidad germinativa, y deben regenerarse a través del cultivo de la planta cada cierto número de años, dependiendo de la especie.

La semilla de Arabidopsis

Las semillas son un órgano especializado de las plantas cuya función es la de mantener el embrión vivo protegiéndolo del ambiente exterior hasta que las condiciones sean favorables para la germanización. La semilla no solo es el embrión sino que esta compuesta por dos tejidos más: el endospermo, tejido triploide que resulta de la fecundación del polen y la célula gigante, y que principalmente tiene como función la de de

nutrir al embrión; y la cubierta de la semilla (llamada así en angiospermas), compuesta por distintas capas de origen materno altamente especializadas cuya función principal es la de aislar y proteger al embrión del ambiente exterior. Distintas plantas difieren en aspectos del desarrollo y de la formación de la semilla. Por ello, de ahora en adelante, se hablará en referencia a la cubierta de la semilla de la planta modelo *Arabidopsis thaliana*.

La Figura 2 en la página 19 muestra distintos aspectos de la semilla, que se han comentado, como un imagen de las semillas secas de Arabidopsis (a), un corte transversal de una semilla en desarrollo donde se distingue la cubierta de las semilla, el endospermo y el embrión en desarrollo (b), las técnicas para analizar la cubierta de las semilla como la permeabilidad de la misma medida por la reducción de las sales de tetrazolio por el embrión (c), las proantocianidinas (d), las capas aislantes de poliésteres lipídicos (e) y el mucílago (f); y las imágenes de microscopia electrónica de barrido (g) y de transmisión de una sección transversal de la cubierta de la semilla (h) dónde se distingue (de abajo a arriba) el endospermo (END), el endotelio (ETH), la capa teñida marrón (BPL) formada por los tegumentos internos colapsados, la capa de células en empalizada (PL) y la última capa de la cubierta de la semilla (COL) formada por las células con forma de volcán llamadas columelas, que le dan el aspecto a la semilla que se ve en (h).

El embrión inicialmente se desarrolla de una célula y luego progresivamente va creciendo y desarrollando los distintos tejidos de la futura plántula. Los estadios de desarrollo descritos en Arabidopsis son: globular, corazón, torpedo y cotiledón doblado. Una vez desarrollado el embrión, este pasa a la fase de maduración donde se prepara para la deshidratación. En este punto el embrión acumula proteínas y otras moléculas de reserva, antioxidantes y osmoprotectores, que protegerán al embrión de la pérdida de agua, generalmente dañina en los organismos. Finalmente, el desarrollo de la semilla concluye con su deshidratación. Este secado permite parar los procesos moleculares permitiendo así aumentar enormemente la vida media de las semillas, al evitar los procesos del metabolismo que producen daño oxidativo al embrión.

El endospermo se desarrolla de un modo curioso y distinto a otros tejidos de las plantas. Una vez fecundado conjuntamente con el embrión, el núcleo se divide múltiples veces, y es después, y no durante, cuando se celulariza produciéndose la citoquinesis. Más tarde, el endospermo va desapareciendo según el embrión se expande, nutriéndolo a su vez, hasta quedarse en una única capa celular que rodea al embrión.

La cubierta de la semilla, también llamada testa, está formada por tegumentos o capas de células que rodean al embrión. Cada capa sufre un proceso complejo de especialización y sintetiza diversos componentes moleculares con distintas funciones. De dentro de la semilla encontramos se puede diferenciar el tegumento interno, formado por

tres capas celulares, y el tegumento externo, formado por dos capas.

El endotelio es la capa más interna del tegumento interno, y es la responsable de producir las proantocianidinas, unos compuestos antioxidantes que protegen al embrión de la luz, el daño oxidativo y del ataque de patógenos y herbívoros. Son los que, al liberarse por ruptura y colapso celular, y oxidarse, dan el color marrón característico de las semillas. Semillas de plantas mutantes en genes de la ruta biosintética y de la regulación de ésta no presentan este color característico, sino tonos más amarillentos. El almacenamiento prolongado oscurece las semillas por la continuada oxidación de las proantocianidinas como se puede observar en la Figura 4 en la página 25. Además de las proantocianidinas, el endotelio produce una cutícula, formada por la cutina que se deposita entre el endospermo y la cubierta de la semilla, cuyos monómeros previamente han atravesado la pared celular. Esta capa es impermeabilizante y protege del aire y de la humedad al embrión con un efecto positivo para la longevidad de la semilla. Las dos capas siguientes, junto al endotelio colapsan y se tiñen de las proantocianidinas, formando la capa teñida marrón. Aunque estas capas parece que no producen compuestos moleculares específicos, se cree se lignifican aportando rigidez y protección a la semilla.

Envolviendo a los tegumentos internos se encuentran dos capas celulares bien distintas que forma el tegumento externo. La más interna está formada por células más pequeñas, dispuestas en empalizada. Estas células sintetizan otro poliéster lipídico, de composición similar a la cutina, pero con un patrón y localización de deposición completamente distinto. La suberina se polimeriza en la cara interna de la pared celular y su patrón de polimerización, aunque no completamente revelado, es distinguible mediante el microscopio electrónico como un bandeado adyacente a la pared celular. Las células de la capa en empalizada al deshidratarse y colapsar tras su muerte como el resto de las capas de la cubierta de la semilla, crea unos pliegos entre las intersecciones de las células creando así una estructura impermeable, que, junto a la cutina, participa en el aislamiento del embrión al ambiente exterior. Los pliegues se pueden ver en la Figura 6 en la página 32,aunque también en la Figura 2, donde se visualiza mejor el bandeado típico de la capa de suberina (SL). En la Figura 6 también se pueden ver otras estructuras como la capa de la cutícula (CL) adyacente a la pared celular del endospermo (CWE).

La capa del tegumento más externa o epitelial está compuesta por células diferenciadas de una forma especial. Tienen una forma de volcán debido a un reforzamiento de una estructura central interior llamada la columela. Estas células sintetizan el mucílago, un polisacárido peptínico altamente hidrofílico. Al embeberse la semilla en agua el mucílago se expande rompiendo la pared celular externa y formando un halo alrededor de la semilla (Figura 2f) que permite retener agua y evitar falta de esta durante las primeras etapas de la germinación de la semilla y establecimiento de la plántula. Además,

al ser un compuesto pegajoso, el mucílago puede influir positivamente en la dispersión de las semillas pegándose a los animales. Respecto a la longevidad y el mucílago, éste no parece ser determinante. Sin embargo, hasta un 30% de los recursos de la testa son usados para la producción del mucílago, indicando una función importante en las distintas especies vegetales.

La biología molecular de la longevidad de semillas

La longevidad de semillas, además de depender de las condiciones del almacenamiento de éstas, está determinada por la especie y los mecanismos moleculares desarrollados para la adaptación y reproducción de las especies vegetales en sus nichos ecológicos. Hay semillas, las llamadas recalcitrantes, que no toleran la desecación. Normalmente son semillas grandes de climas húmedos, y apenas pueden ser almacenadas. Su alta humedad, además de promover el ataque de microorganismos y hongos, no permite la parada del metabolismo, por lo que acumula en mayor medida daño oxidativo que va dañando la semilla. Además, su sensibilidad al secado hace que en climas secos no puedan ser almacenadas. Por contra, las semillas que toleran el desecado, llamadas semillas ortodoxas, son mucho más longevas. Su alta longevidad se debe a la parada de su metabolismo por el secado la planta, dejando la vida de la futura plántula en estado latente. Al parar su metabolismo, los procesos y el daño oxidativo asociado se ven altamente mermados, lo que disminuye enormemente el ratio de eventos dañinos que van envejeciendo las semillas progresivamente. Con esta estrategia las semillas pasan de durar meses hasta durar años e incluso, miles de años, como se ha datado con semillas milenarias que pudieron germinar de las especies de plantas loto sagrado (*Nelumbo nucifera*) y palmera datilera (*Phoenix dactylifera*).

Para tolerar el desecado, dañino para las estructuras y componentes celulares, el embrión se prepara durante la fase de maduración acumulando distintas proteínas y compuestos como azúcares. El citoplasma pasa de ser líquido a ser muy viscoso durante la deshidratación, reduciéndose la difusión molecular y por lo tanto el metabolismo del embrión. Este citoplasma viscoso rompe la matriz de formación de cristales evitando así daños y roturas de estructuras celulares provocadas por la pérdida de agua. Si las condiciones de secado y temperatura son los suficientemente bajas el citoplasma pasa a estar en el llamado el estado vítreo, óptimo para el almacenamiento de semillas. Por contra, si las semillas se encuentran en condiciones de humedad y temperaturas más altas, la fluidez del citoplasma aumenta y se reactiva parcialmente el metabolismo y los eventos dañinos relacionados con el mismo. Los azúcares acumulados, además de servir de reserva energética para la germanización del embrión, contribuyen a la mayor viscosidad del citoplasma y protegen al embrión frente al secado. La familia de oligosacáridos derivados de la rafinosa participa activamente a la osmoprotección

y a la formación del estado vítreo. Las proteínas LEA acumuladas al final de la embriogénesis son altamente hidrofílicas, y se postula que protegen al embrión frente a la deshidratación mediante su plegado progresivo ayudando también la formación del estado vítreo. Por otra parte, las proteínas de choque térmico participan en la protección activa de las proteínas ayudando al plegamiento de estas ante la deshidratación y durante la rehidratación.

La semilla desarrolla diversas estrategias para reducir las especies reactivas de oxígeno acumuladas progresivamente durante su almacenamiento y el daño oxidativo derivado de las mismas. Una de esas estrategias consiste en sintetizar moléculas y sistemas antioxidantes que participan en la reducción de las especies reactivas de oxígeno mediante la oxidación pasiva y activa de moléculas con poder reductor. Otra de las estrategias consiste en la producción de baterías de moléculas, incluyendo proteínas de almacenamiento, azúcares y moléculas de RNA, y que tamponan el daño oxidativo. De esta forma, aunque diversas moléculas resulten dañadas, otras no lo estarán y podrán desempeñar su función durante la germinación de la semilla, siempre que el daño oxidativo no sea excesivo.

Los cloroplastos absorben la luz del sol y a través de reacciones redox almacenan la energía de los rayos del sol en los enlaces entre carbonos, a la vez que fija los átomos de carbono del dióxido de carbono presente en el aire. Durante el desarrollo el embrión se vuelve verde debido a la acumulación de clorofila, para la realización de la fotosíntesis, necesaria para la síntesis de todos los compuestos acumulados en el embrión. Sin embargo, durante el almacenamiento la absorción de luz es deletérea para el embrión ya que conllevaría una mayor acumulación de daño oxidativo derivado de la absorción de la energía de los rayos de luz. Por ello, durante la maduración de la semilla la clorofila se degrada.

Finalmente, los sistemas de reparación pueden aliviar parte del daño producido en el embrión. Estos sistemas, específicos de cada tipo de molécula y del tipo de daño producido, actúan principalmente durante la rehidratación de la semilla. Se han descrito sistemas de reparación que actúan en proteínas y DNA. Se piensa que el retraso en la germinación observado en semillas envejecidas se debe a la actuación de estos sistemas de reparación. De hecho, el inicio de la germinación no se produce si hay daños excesivos en la integridad del DNA. Sin este control, las semillas envejecidas germinan en mayor porcentaje, pero las plántulas presentan aberraciones cromosómicas resultando en mayor número de plantas no viables que no pueden completar su ciclo vital.

La testa en la longevidad de semilla

La cubierta de la semilla aísla al embrión del ambiente exterior protegiéndolo del daño físico y oxidativo. Los diversos compuestos depositados en las distintas capas de la semilla participan en la longevidad de la semilla. Estos compuestos son las proantocianidinas, las paredes celulares reforzadas con lignina y las dos capas de poliésteres lipídicos, la cutícula y la suberina.

La especialización de los distintos tegumentos y la síntesis específica de capa celular de sus distintos componentes moleculares requiere un complejo control transcripcional estrechamente regulado. Esta regulación empieza desde la floración y la diferenciación de los tegumentos que inicialmente rodean al óvulo. Diversos factores de transcripción se han descrito en participar en esta cascada de regulación, con un importante efecto en la identidad de capa y síntesis de los compuestos asociados. Por ejemplo, hay grupos de factores de transcripción que determinan la identidad del tegumento interno mientras otros determinan la del tegumento externo. También hay factores de transcripción que regulan específicamente la producción de los distintos compuestos asociados a la cubierta, como las proantocianidinas, la suberina o el mucílago. Sin ellos, la diferenciación de la cubierta de la semilla no se logra completar. La importancia de ciertos factores de transcripción en la longevidad de semilla se ha corroborado mediante plantas mutantes que fallan en la síntesis proteica de dichos factores de transcripción, presentando alteraciones en la cubierta de la semilla con un importante efecto en su longevidad.

Dos factores de transcripción importantes en esta tesis y que participan en la longevidad de semilla son AtHB25 y COG1, cuyos mutantes de sobreexpresión se aislaron en nuestro laboratorio tras tratamientos de envejecimiento acelerado. Estos tratamientos usan alta humedad y temperaturas para incrementar el daño sufrido por las semillas y conseguir así acelerar su envejecimiento. Dentro de una colección de semillas mutantes que sobreexpresaban distintos genes, las semillas de las plantas mutantes *athb25-1D* y *cog1-2D*, con una mayor expresión de dichos genes, respectivamente, sobrevivieron y mostraron una longevidad de semilla incrementada en sucesivas generaciones comparadas a semillas silvestres. Se determinó que este efecto era de herencia materna, por lo que apuntaba a la cubierta de la semilla como responsable. Además, se vio que presentaban una mayor tinción de poliésteres lipídicos apuntado a una mayor acumulación de suberina y/o cutina en la cubierta de la semilla. Estos mutantes curiosamente presentan mayor contenido hormonal de giberelinas y ácido abscísico, hormonas que regulan el desarrollo y la maduración, respectivamente, del embrión y de la cubierta de la semilla.

Las capas de poliésteres lipídicos en la cubierta de la semilla resultan determinantes en la longevidad de las semillas, según se ha visto en esta tesis doctoral y en estudios

anteriores. Tanto la suberina como la cutina están formadas por monómeros alifáticos de larga cadena de 16 a 24 átomos de carbono (ácidos grasos (FA), alcoholes grasos (PA), ácidos grasos hidroxilados (OHPA) y ácidos dicarboxílicos DCA)); por monómeros aromáticos (ácidos hidroxicinámicos (HCA)); y glicerol 3-fosfato (G3P). Su síntesis, incluyendo las rutas de origen y los diversos enzimas involucrados; y sus estructuras moleculares se muestran de forma esquemática en la Figura 7 en la página 34, y la abundancia de estos en la cubierta de la semilla se puede ver en la Figura 8 en la página 35. Estos monómeros se polimerizan por enlaces éster, usando los grupos alcohol (OH) del glicerol 3-fosfato, de los alcoholes grasos y de los ácidos grasos hidroxilados; y el grupo carboxílico (COOH) de los ácidos hidroxicinámicos, ácidos grasos, ácidos dicarboxílicos y ácidos grasos hidroxilados. Estos monómeros pueden formar compuestos intermediarios dentro de la célula antes de su secreción a través de la membrana celular o secretarse directamente. Los compuestos intermediarios que se piensa que se forman se pueden visualizar en la Figura 9 en la página 37. Finalmente, una vez en el exterior de la célula los diversos compuestos se polimerizan unidos a la pared celular. Dicha polimerización es distinta en la suberina y la cutina, además de la abundancia relativa de los distintos compuestos y de los enzimas que participan en la biosíntesis de monómeros e intermediarios. Aunque la estructura de la suberina y de la cutina no están completamente dilucidadas, la Figura 10 en la página 39 muestra su hipotético patrón de polimerización. Tampoco están completamente dilucidados los enzimas biosintéticos que participan en dicha polimerización. En el caso de la suberina, la polimerización se produce en la cara interna de la pared celular. Se piensa que el bandeado laminar que se observa se debe a la alternancia en la deposición de monómeros aromáticos y alifáticos. Además, la mayor abundancia de monómeros aromáticos y ácidos grasos dicarboxílicos en comparación a la cutina hace pensar que el patrón de deposición crea un entramado más complejo que en la cutina, que en la semilla presenta un patrón de deposición amorfo a vista del microscopio electrónico. Dicha deposición se produce, por contra a la suberina, en la cara exterior de la pared celular. Aunque se desconoce el proceso, los compuestos hidrofóbicos tienen que atravesar la pared celular, de características hidrofílicas. Se postula que puede ocurrir por acumulación de monómeros alifáticos y separación de fases y/o mediado por enzimas de transferencia de lípidos. Una vez en el exterior los monómeros y compuestos intermediarios se polimerizan entre ellos formando la capa impermeable. Ayudando a la impermeabilidad de esta capa también se depositan ceras no polimerizadas que contribuyen a repeler el agua.

La adaptación ambiental de las semillas

Las semillas tienen una gran importancia en la adaptación de las plantas. Por ejemplo, éstas permiten la supervivencia de las plantas, aunque existan épocas de sequía o de temperaturas extremas. Cuando la planta no puede sobrevivir, las semillas preservan la vida hasta que las condiciones ambientales son favorables para la germinación y el desarrollo de la planta. Si la longevidad de semilla fuera el único aspecto de importante de las semillas para la adaptación y supervivencia de las plantas, la selección evolutiva haría que las semillas fueran cada vez más longevas. Sin embargo, no es así debido a que otras características de las semillas también participan en la adaptación de las plantas.

Una característica importante de la semilla es la dormancia. La dormancia de la semilla evita la temprana germinación. Debido a la estacionalidad es común que para el correcto desarrollo de la planta las semillas deban germinar por primavera, pasando el invierno sin germinar. Sin la dormancia, las semillas podrían germinar una vez producidas, en otoño y la nueva planta podría morir por el frío invierno. Para evitar esto la semilla percibe las señales ambientales que determinan cuando se rompe la dormancia de la semilla. Una vez rota la dormancia, las semillas germinarán en cuanto se encuentre en condiciones de temperatura, humedad y luz adecuadas. Un periodo de baja temperatura y alta humedad rompe la dormancia, ya que el paso del invierno es detectado por esta forma por las semillas. También se rompe la dormancia tras un cierto tiempo de almacenamiento. Los eventos de oxidación de las semillas, dañinas para la longevidad actúan como un reloj para la detección del paso del tiempo. La cubierta de la semilla, a través de la oxidación de las proantocianidinas regula, en parte, este tipo de dormancia. Además, las capas de suberina y la cutícula de la cubierta de la semilla evitan la oxidación, por lo que son positivas para la dormancia de semillas, al igual que para la longevidad.

La longevidad, la dormancia y otras características de las semillas no solo están determinadas genéticamente. Durante el desarrollo de la semilla, las planta madre y las propias semillas perciben las señales ambientales que determinan las diversas características de las semillas. Se ha visto que temperaturas templadas y condiciones de mucha luz durante el desarrollo de las semillas son positivas para la longevidad, y negativas para la dormancia. Este efecto puede ser debido a la adaptación geográfica: en áreas tropicales y ecuatoriales, normalmente cálidas con alta incidencia de luz, la longevidad de semilla se debe reforzar para resistir estas condiciones deletéreas. Sin embargo, las áreas más cercanas a zonas polares son normalmente lugares fríos con menor incidencia de luz, la dormancia de la semilla es más importante ya que debe evitar la germinación de la semilla en las épocas frías, además de que las condiciones no son tan deletéreas para las semillas. Otras características importantes de las semillas para la adaptación de las plantas son la cantidad de semilla producida, la dispersión y la acumulación de agua durante la rehidratación por el mucílago.

El estudio de la longevidad de semilla

El envejecimiento de las semillas es un proceso probabilístico. Es decir, no todas las semillas de un mismo lote envejecen a la vez, sino que gradualmente van perdiendo la capacidad germinativa. En la Figura 11 en la página 47 se puede ver gráficamente como las semillas almacenadas a temperatura y humedad ambiente al principio no pierden la capacidad germinativa pero luego gradualmente van perdiendola con el paso de los meses. Por ello el estado del envejecimiento de los distintos lotes de semillas a distintos tiempos se mide a través del porcentaje de germinación de las semillas. Además, para una caracterización más ajustada del envejecimiento del lote de semillas se puede estimar su longevidad a distintos tiempos y ajustar su envejecimiento a una curva sigmoidal para estimar el tiempo de almacenamiento en el cual las semillas germinan al 50%. Ese valor se le denomina P50. Entre año y medio y dos años, la capacidad germinativa de las semillas Arabidopsis del ecotipos Columbia (Col-0) se reduce sobre el 50%. Es decir, en nuestros ensayos, su P50 oscila entre 18 meses y 24 meses según condiciones y lotes.

Para estudiar la longevidad de las semillas se utilizan tratamientos de envejecimiento artificiales con condiciones de alta humedad y temperatura controladas para conseguir un rápido deterioro de las semillas y evitar la espera de largos periodos de tiempo para determinar su longevidad. Nosotros utilizamos tres tratamientos de envejecimiento artificiales. El tratamiento de envejecimiento acelerado (AAT) consiste en embeber las semillas en agua caliente sobre los 40 ºC. En un par de días las semillas de Arabidopsis (ecotipo Col-0) reducen más del 50% de su germinación. Otro tratamiento más lento que no usa semillas embebidas en agua, pero sí una alta humedad y temperatura es el tratamiento de deterioro controlado (CDT). La alta humedad, a un 75% de humedad relativa, se consigue con una solución de sal saturada encerrada herméticamente junto a las semillas y 37 ºC. Con estas condiciones, las semillas de Arabidopsis de este ecotipo pierden la capacidad de germinar sobre los 15 días de tratamiento. El último tratamiento que se ha usado es el de elevada presión parcial de oxígeno (EPPO), que no usa alta humedad o temperatura sino una elevada presión atmosférica de oxígeno. Con 5 bares de oxígeno, las semillas de Arabidopsis (Col-0) a los 5 meses pierden considerablemente su capacidad germinativa.

Para estandarizar los resultados, antes de promover la germinación, sometemos las semillas a un periodo de vernalización que consiste en dejarlas en condiciones de bajas temperaturas y alta humedad, para que la dormancia se rompa y así evitamos que ésta afecte al porcentaje de germinación. Con dos o tres días, las semillas recientes de Col-0 embebidas en agua a 4 ºC son totalmente capaces de germinar ya que su dormancia se rompe por completo

Lo que se ha hecho en esta tesis

Esta tesis se trata de una compilación de artículos y resultados por publicar en los que se han investigado diversos aspectos de la longevidad de la semilla. Los artículos completos están en inglés, y por ello la tesis también está escrita en inglés. Sin embargo, en este apartado de traducción divulgativa al castellano se explicarán de forma simplificada los resultados obtenidos en cada uno de los capítulos de la tesis, correspondientes a los distintos artículos y resultados.

Capítulo 1:

El estudio de la longevidad de semilla en los ecotipos de Arabidopsis permite descubrir nuevos genes directamente involucrados en la longevidad de semilla

Artículo completo en inglés en página 53.

La técnica genómica: el GWAS

En esta investigación se utilizaron 270 variedades naturales o ecotipos de la planta modelo *Arabidopsis thaliana* para realizar un estudio del genoma completo conocido como GWAS por sus siglas en inglés. Para ello en los 270 ecotipos se estableció la longevidad de semilla de cada uno de ellos. Se usaron los tres tratamientos artificiales de envejecimiento AAT, CDT y EPPO y un tratamiento de envejecimiento natural (NAT) que consistió en año y medio de almacenamiento en seco a temperatura ambiente del laboratorio. Se observó una gran variación en la longevidad de semillas en todos los distintos tratamientos como se puede ver en la Figura 12 página 57. Gracias a esta variación se pudo encontrar zonas del DNA del genoma de la planta que presentan una correlación estadísticamente significativa entre la distinta longevidad de semilla observada en los distintos ecotipos y las distintas variaciones del DNA, que consisten mayoritariamente en polimorfismos en única base del DNA (SNPs). Los ecotipos se crecieron simultáneamente bajo las mismas condiciones para reducir la variabilidad ambiental y obtener así una correlación mayor entre el fenotipo, en este caso, la longevidad de semilla y el genotipo, o el conjunto de DNA característico de cada ecotipo. Todos estos ecotipos pertenecen a una colección de ecotipos que han sido genotipados, por lo que hay información pública de gran parte de las variaciones presentes en su genoma. Tras computaciones estadísticas globales con correcciones de significatividad multiparamétricas, ya que se analizaban más de un 1.700.000 SNPs, se obtuvieron diversas regiones del genoma con SNPs significativos. La Figura 14 en la página 59 muestra los resultado de los distintos análisis GWAS realizados en el formato de gráfico de Manhattan. Se llama así porque se asemeja al horizonte de edificios de Manhattan. Se trata de la visualización gráfica de la significatividad de los SNPs. Aquellas zonas con “edificios” más altos se tratan de zonas genómicas donde potencialmente se encuentran determinantes de la longevidad de la semilla que explica parte de la variación observada de los ecotipos. De estas zonas, se analizaron y se seleccionaron los genes cercanos buscando que se expresaran en semilla, que pudieran tener una función en la longevidad de la semilla o que fueran cercanos a los SNPs más significativos. En la Figura 14 se han marcado aquellos SNPs cercanos a genes que luego se ha visto que tenían un papel en la longevidad de semilla tras el análisis de mutantes explicado

a continuación. No siempre corresponden con los SNPs más significativos, debido a diversos factores. Puede ser que exista redundancia genética por lo que el efecto del gen no se puede determinar mediante el análisis de mutantes realizado. También puede ser que en el ecotipo Col-0 dicho gen no tenga un papel importante pero si en otros ecotipos.

Validación de los genes mediante plantas mutantes

Para validar la implicación de dichos genes se caracterizaron líneas mutantes de inserción de T-DNA para ver si presentaban una longevidad de semilla alterada. Estas líneas T-DNA son plantas Col-0 que se han transformado con la bacteria *Agrobacterium tumefaciens*. Esta transformación consiste en la inserción aleatoria del fragmento T-DNA, contenido en un plásmido presente en la bacteria. En este caso el fragmento de T-DNA solo lleva resistencia a un herbicida o antibiótico, para la detección de las plantas con inserción de DNA o plantas transgénicas. En otros casos, el fragmento de T-DNA incluye genes y regiones promotoras que permiten la expresión del gen a estudiar o de un gen reportero, para permitir el estudio molecular del gen o del patrón de expresión del gen, respectivamente. La transformación por Agrobacteria es la más utilizada en Arabidopsis. A través de la inmersión de flores se permite la infección de óvulos o embriones que se desarrollarán en individuos transgénicos no quiméricos (no presentan células con transgén y células sin transgén). Como la inserción de T-DNA es aleatoria en el genoma de la planta, se han desarrollado colecciones de plantas mutantes que presentan inserciones de T-DNA independientes, repartidas por todo el genoma. Las diferentes líneas T-DNA se han caracterizado para determinar la posición de la inserción genética. Por ello, podemos pedir líneas T-DNA que presentan la inserción T-DNA en un gen de interés.

De las distintas líneas mutantes, se demostró que 12 genes tenían un papel directo en la longevidad de semilla. Siete de ellos mostraban un papel positivo ya que las plantas mutantes correspondientes mostraban una menor longevidad de semillas que las semillas de la planta silvestre de referencia Col-0. Cinco de ellos eran negativos para la longevidad, ya que sus semillas eran más longevas.

Los genes que se han demostrado estar involucrados en la longevidad de semilla son los siguientes:

- *SSLEA*: Se trata de una proteína LEA, que son acumuladas al final de la embriogénesis y que se piensa que ayudan a la protección de la semilla durante la deshidratación. Solo se había descrito anteriormente una proteína LEA, llamada LEA14, en influir positivamente en la longevidad de semilla. Cabe decir que no tolerar la desecación es un aspecto distinto a la longevidad. Por ello, comprobamos que las semillas secas eran capaces de germinar completamente, y que su

capacidad germinativa disminuía rápidamente con el tiempo y con lo tratamientos de envejecimiento artificiales (Figura 16 en página 61 y Tabla 3 en la página 62), demostrándose que tiene un papel en la longevidad de semilla. SSLEA fue llamada así por nosotros y las siglas en inglés de LEA de almacenamiento de semillas. El estudio de los genes de la misma subfamilia de SSLEA puede ser de gran interés para el campo de la longevidad de semilla, ya que éstos pueden ser cruciales para el almacenamiento de las semillas, pudiendo arrojar luz sobre el verdadero papel de las proteínas LEA.

- *PSAD1*: Se trata de una subunidad del fotosistema I, presente en la pared tilacoidal de los cloroplastos. Las plantas mutantes en este gen presentan un fenotipo de planta pequeña y color verde pálido. Se ha descrito que las plantas mutantes acumulan mayor daño oxidativo pudiendo ser ésta la causa de la baja longevidad de semilla observada (Figura 16 en página 61). Probablemente, al romperse la cadena de transporte de electrones de la fotosíntesis, los electrones que no pueden transportarse pasan a oxidar y dañar moléculas, siendo fatal para la longevidad de la semilla. Otra explicación de la baja longevidad de las semillas mutantes *pasd1* parte de la base de que la fotosíntesis está afectada, debido a esta disrupción de la cadena de transporte de electrones. La semilla requiere la fotosíntesis para el engrosamiento del embrión y la acumulación de macromoléculas de reserva, procesos necesarios para el establecimiento adecuado de su longevidad.

- *SSTPR*: Aunque no se conoce mucho sobre esta amplia familia de proteínas con repeticiones de tetratricopéptidos (TPR) este gen resultó importante en la longevidad de semillas ya que semillas mutantes en este gen son más sensibles al envejecimiento (Figura 16 en página 61) y presentan una permeabilidad de cubierta mayor, como se puede ver por la mayor absorción y reducción por el embrión de las incoloras sales de tetrazolio en formazán, un compuesto de color rojo fácilmente visible (Figura 17 en la página 62). Se postula que las proteínas TPR tienen un papel importante en el anclaje de proteínas y RNA con funciones esenciales en mitocondras y cloroplastos. De hecho, la proteína SSTPR, llamada así por nosotros y las siglas en inglés de TPR de almacenamiento de semillas, parece sublocalizarse en las mitocondrias. Poco más se puede concluir de este nuevo gen descrito en longevidad de semilla.

- *RBHOD, RBOHE, RBOHF*: De estos tres genes, dos de ellos salieron significativos en el distintos análisis GWAS. Se trata de una familia de 10 miembros que van de la A a la J, con alta homología a las bien descritas NADPH oxidasas de humanos. Estos enzimas catalizan la creación de superóxido de oxígeno, una especie reactiva de oxígeno muy dañina para el organismo. No obstante, el superóxido rápidamente se transforma en agua oxigenada, un poco menos dañina, que se utliza en la señalización celular de procesos de desarrollo importantes, como la

respuesta a patógenos, y la muerte celular programada, que ocurre en la cubierta de la semilla como final de su desarrollo. De hecho, estos tres *RBOH* se expresan mayoritariamente en la cubierta de la semilla y endospermo al final del desarrollo como se observa en la Figura 18 en la página 63. Semillas de plantas mutantes en estos tres genes, mostraron un longevidad aumentada en comparación a las semillas silvestres, indicando un papel negativo de estos RBOH en la longevidad de semilla. De hecho, las semillas del doble mutante *rbohd,f* presentan aun mayor longevidad indicando un efecto aditivo de las distintas mutaciones (Figura 19 en la página 63). Es de esperar que un enzima que produzca especies reactivas de oxígeno sea deletérea para la longevidad, ya que su función produce daño durante su almacenamiento. El hecho curioso es que se haya mantenido su expresión en semillas siendo ésta deletérea. Esto se podría explicar por diferentes argumentos. Podría ser que la muerte de las células de la cubierta de la semilla se retrase en los distintos mutantes, ya que la señal de muerte no estaría tan coordinada. Esto podría permitir un desarrollo de los componentes protectores de dicha cubierta por más tiempo. Sin embargo, esta explicación no parece muy plausible ya que el análisis de las cubierta de semilla de los distintos mutantes no muestras cambios en la permeabilidad o la deposición de poliésteres lipídicos aislantes como se ve en la Figura 20 en la página 64, respectivamente. La explicación que parece más plausible es que se mantengan por su función en la señalización en la repuesta a patógenos. La semilla, de esta forma, sacrificaría parte de su longevidad para poder evitar el ataque de hongos y otros patógenos, también dañino para las semillas. En la Figura 5 en la página 26 se puede observar como un hongo crece en la cubierta de la semilla. Este crecimiento de hongos finalmente romperá la cubierta de la semilla y el aislamiento que provee, y permitirá el ataque de microorganismos directamente al embrión. No obstante, las especies reactivas de oxígeno producidas por las proteínas RBOH dañan progresivamente al embrión, tanto internamente desde el embrión como por difusión desde la testa.

- *DHAR1*: Este gen sintetiza una dehidroascorbato reductasa, que regenera el ascorbato o vitamina C, desde el glutatión. Se ha descrito que hay tres isoformas activas en Arabidopsis, y *DHAR1* es la que se expresa más en la semilla en desarrollo (Figura 21 en la página 65). Aunque *DHAR1* fue el gen que salía significativo en el GWAS, se probaron también los mutantes *dhar2*, *dhar3* y el triple mutante *dhar1,2,3*. Solo resultaron sensibles al paso del tiempo y a los distintos tratamientos artificiales las semillas mutantes *dhar1* y *dhar1,2,3* (Figura 22 en la página 65), indicando que *DHAR1* es la única isoforma que tiene un efecto importante en la longevidad de semilla. El papel de DHAR1 en la longevidad de semilla debe ser el de utilizar las reservas de poder reductor almacenado en forma de glutatión reducido para regenerar el ascorbato reduciéndolo. El ascorbato es un importante antioxidante que permitirá reducir el daño de las especies reactivas

de oxígeno eliminándolas al oxidarse.

- *CYP86A8*: Este gen codifica a un citocromo P450 (CYP), enzimas con un grupo de hemo (con un átomo de hierro). Esta amplia familia génica participan en multitud de reacciones biosintéticas. La subfamila CYP86 cataliza la hidroxilación de ácidos grasos. Esta reacción es necesaria para la síntesis de ácidos grasos hidroxilados y ácidos dicarboxílicos, monómeros de cutina y suberina. Aunque en el análisis GWAS salieron remarcados los miembros de la familia *CYP86A2* y *CYP86B1*, decidimos estudiar todos los miembros que se expresaran durante el desarrollo de la semilla, según las bases de datos (Figura 23 en la página 67), por si algún otro miembro presentaba cambios en la longevidad. Y fue así, las semillas de la planta mutante *cyp86a8* envejecían rápidamente, ya que en doce meses apenas presentaban germinación cuando los otros miembros de la familia y las semillas control germinaban casi por completo como se puede ver en la Figura 24 en la página 68. En esta figura también se puede ver como las semillas del mutante *cyp86a8* (derecha) presentan menor tinción de poliésteres lipídicos (un color morado menos intenso en la cubierta de la semilla), una forma mas redondeada y una mayor permeabilidad a las sales de tetrazolio que las semillas silvestres (izquierda). La explicación del papel de CYP86A8 en la longevidad de la semilla debe ser la de sintetizar los monómeros de suberina y/o cutina de la cubierta de la semilla. Este gen se ha descrito por ser importante en la síntesis de cutina de la cutícula de la hoja, y plantas mutantes muestran un defecto pleiotrópico de crecimiento con fusión de órganos. No hay que descartar que este efecto afecte a las reservas de la planta para la producción de semillas. Sin embargo, el fenotipo visto en las semillas sugiere que se trata de la primera explicación descrita, o una combinación de ambas explicaciones

Como se ha comentado en la introducción, los factores de transcripción tienen un importante papel en el desarrollo de la semilla y, en especial, de la cubierta de la semilla. Con estos GWAS, hemos conseguido describir cuatro factores de transcripción, involucrados en el desarrollo de la cubierta de la semilla, con un papel determinante en la longevidad de semilla. Dos son positivos para la longevidad y dos son negativos para la longevidad como se puede apreciar en la Figura 26 en la página 71.

- *SPCH*: Este factor de transcripción se ha descrito crucial en el desarrollo de estomas, ya que plantas mutantes en este factor de transcripción no presentan estomas. Aunque no se expresa en gran medida durante el desarrollo de la semilla (Figura 25 en la página 70), las semillas presentan una mayor permeabilidad a las sales de tetrazolio (Figura 26), lo que podría explicar su menor longevidad. La cubierta de la semilla, sin embargo, no muestra otros fenotipos (Figura 28 en la página 72). El mecanismo por el que las semillas mutantes *spch* tienen menor longevidad queda sin resolverse.

- *KNAT7*: Este factor de transcripción se ha descrito como regulador negativo de la lignificación y se expresa en la cubierta de la semilla. También se ha relacionado con la extrusión de mucílago, ya que está reducida en las semillas mutantes *knat7*. Vimos que las semillas era más pequeñas y presentaban, además de una extrusión de mucílago menor, una menor permeabilidad de la cubierta como se puede ver en la Figura 27 en la página 72. Todos estos efectos se pueden explicar por una temprana y/o acentuada lignificación de la cubierta de la semilla. Las semillas son más pequeñas porque al lignificase la cubierta no crecen más. El mucílago no estaría reducido sino que las paredes celulares que se rompen para permitir la liberación del mucílago, al estar reforzadas, tardarían más en romperse. De hecho, en nuestras imágenes de mucílago, probablemente con mayor tiempo de incubación en agua que las previamente publicadas sí que se observa el halo de mucílago, aunque reducido (Figura 28 en la página 72). La mayor longevidad de semilla, por lo tanto, se explica por una cubierta de semilla reforzada que protegerá mejor al embrión. El papel de KNAT7 y su mantenimiento evolutivo probablemente se deba a su papel regulador en la ligninficación, que necesita ser controlada para la correcta liberación del mucílago y para la longevidad.

- *SEP3*: Este factor de transcripción está involucrado en la floración. Cuatro son los genes SEPALLATA que determinan de forma redundante la identidad del tejido de los sépalos. Además, están involucrados en la identidad de los tegumentos que darán lugar a la cubierta de la semilla. Estos factores de transcripción, y en concreto SEP3, interaccionan con multitud de factores de transcripción para regular la formación de estos tejidos. Mediante el análisis GWAS de envejecimiento natural, se remarcó una zona genómica próxima a *SEP1*, un factor de transcripción de la subfamilia *SEPALLATA*. Sin embargo, al caracterizar las semillas mutantes de los distintos genes *SEPALLATA* que teníamos disponibles (el triple mutante *sep1,2,4* y el mutante simple *sep3*) vimos que las semillas mutantes *sep3* mostraban una longevidad aumentada en comparación con las semillas silvestres, mientras que las semillas *sep1,2,4* no mostraron un cambio significativo en su longevidad. Curiosamente observamos otros fenotipos relacionados con la cubierta de la semilla en el mutante *sep3* contradictorios con su alta longevidad. Las semillas presentaban una permeabilidad a las sales de teatralizo exacerbada, como se puede ver en la Figura 26 en la página 71. Además, las semillas tienen una forma redondeada, con una muy reducida producción de mucílago (Figura 28). Estos efectos parecen estar relacionados con la función de SEP3 en el desarrollo de los tegumentos, tanto externo como interno. La alta permeabilidad puede explicarse por una presión interior del embrión frente a la cubierta de la semilla al embeberse en agua, produciendo fisuras en la cubierta de la semilla. La forma redondeada de la semilla sugiere que este efecto se deba a una cubierta menos reforzada, ya que al ejercer una presión de constricción menor el embrión tiende a crecer esféricamente. Este efecto lo hemos visto también en las semillas *cyp86a8*, que

presentaba menos poliésteres lipídicos en la testa. Las fisuras parece que sólo se producen durante la rehidratación, ya que las semillas secas no deben presentar dichas fisuras ya que son más longevas. Sin embargo, la longevidad aumentada de *sep3* no se explica de este modo. Podría ser que la cubierta presentara otras ventajas para la longevidad que no logramos detectar en nuestros ensayos. O podría ser que la producción disminuida de mucílago dejara mayor cantidad de recursos para la producción de moléculas y estructuras celulares que participan en la protección de la cubierta de la semilla frente al paso del tiempo.

- *MYB47*: El factor de transcripción de MYB47 salió remarcado en los tres GWAS correspondientes de los tratamientos de longevidad artificiales. Resultó ser un factor de transcripción positivo para la longevidad ya que las líneas mutantes mostraban menor germinación después de los tratamientos de envejecimiento. Aunque las bases de datos localizaban su expresión en la cubierta de la semilla (Figura 25 en la página 70), no vimos cambios significativos en la permeabilidad de la cubierta (Figura 26 en la página 71) ni en la deposición de poliésteres lipídicos, en las proantocianidinas ni en el mucílago (Figura 28 en la página 72). Desarrollamos distintas líneas transgénicas de sobreexpresión bajo el promotor constitutivo UBQ10 y líneas de fusión del promotor endógeno de *MYB47* (aproximadamente 2000 pares de bases anteriores al codón de inicio de la traducción del gen *MYB47*) unido a la secuencia génica codificante de la proteína verde fluorescente de medusa llamada GFP, por sus siglas en inglés. Como se puede ver en la Figura 29 en la página 73, las líneas de sobreexpresión de GFP localizaron la expresión del gen MYB47 en la cubierta de la semilla durante el desarrollo de la semilla. Primero, aproximadamente a los 3 días de la polinización la señal fluorescente de la proteína GFP se sitúa en toda la cubierta, y luego, a los siete días la señal se localiza predominantemente en la zona de la chalaza de la semilla, donde acaban y convergen los tegumentos y el endospermo de la semilla. Esta zona, durante el desarrollo de la semilla inicialmente se encuentra abierta para el paso del polen durante la fecundación y luego se cierra durante el desarrollo para el encoframiento del embrión para el aislamiento exterior. La líneas de sobreexpresión del gen MYB47, que contenían más cantidad de transcritos de dicho gen y, por lo tanto, presumiblemente más cantidad del factor de transcripción, mostraron mayor resistencia a los tratamientos de envejecimiento artificiales AAT y CDT, confirmando el papel positivo de MYB47 en la longevidad de la semilla. Sin embargo tampoco mostraron diferencias significativas en los componentes de la cubierta de la semilla que pudieran explicar su longevidad (Figura 30 en la página 74). El papel de MYB47 en la longevidad de la semilla y su mecanismo molecular quedó sin resolver, aunque apunta a una nueva característica de la cubierta de la semilla aún no descrita en la longevidad de la semilla. Estudios de secuenciación masiva de los RNA totales de las semillas de estos mutantes pueden ayudar a dilucidar estos mecanismos. Podría ser que MYB47 este invo-

lucrado en el cierre de la semilla en la zona de la chalaza, por ejemplo mediante el reforzamiento de los poliésteres lipídicos. Esta hipótesis viene en parte porque factores de transcripción de la misma familia, y muy cercanos a MYB47 según su secuencia de DNA codificante, regulan dicha deposición de poliésteres, y algunos como MYB107 son claves en la deposición de suberina en la semillas en desarrollo. La expresión de MYB47 se localiza en la chalaza en el periodo que se produce el cierre de la chalaza, y además, en la chalaza se puede ver auto fluorescencia de suberina como indicador de la deposición de este poliéster lipídico. Por ello, podría ser que MYB47 regulara genes similares a los que regula MYB107 para regular la deposición de suberina en el cierre de la chalaza, explicando así su papel positivo en la longevidad de la semilla.

Además de estos 12 genes, el fenotipado de longevidad de semilla de los ecotipos se comparó con otros fenotipos públicos de esta misma colección de ecotipos o partes de la misma (Tabla 1 en la página 58). Los fenotipos que más correlacionan con la longevidad de la semilla fueron el tiempo de floración, el número de hojas en la roseta (hojas que no son del escape floral) y el tiempo de vida de la planta. Estos tres fenotipos están altamente relacionados porque la floración temprana conlleva la parada la producción de hojas de la roseta, y la planta de Arabidopsis, que muere una vez completado la producción de semilla, morirá antes si florece antes, ya que antes producirá las semillas. Estos fenotipos correlacionan inversamente con la longevidad de la semilla, indicando que a floración más rápida, mejor longevidad de semillas. Esta correlación no había sido previamente sugerida, y podría explicarse por un mejor desarrollo o adaptación del ecotipo a las condiciones fijadas del ensayo. Cada planta proviene de un nicho ecológico distinto y, por tanto, está adaptada a condiciones ambientales distintas. Al fijar las condiciones, las que antes florecen probablemente se encuentren en condiciones de crecimiento más adecuadas y cercanas a las de su nicho que las otras. Esto produce que las semillas de las plantas que antes florecen sean de plantas que han crecido mejor, y por lo tanto, las semillas también se han desarrollado mejor. También puede ser que otros factores afecten a esta correlación inversa entre el tiempo de floración y la longevidad de la semilla.

Otro fenotipo que resultó correlacionar inversamente con la longevidad de semilla fue el contenido en hierro de la planta. Este contenido en hierro se midió en hojas, sin embargo, es presumible que las semillas de las plantas muestren cantidades de hierro proporcionalmente similares. Esta correlación, previamente sugerida, podría explicarse por el papel oxidativo del hierro ya que es un átomo aceptor de electrones, involucrado en las cadenas de transporte de electrones de mitocondria y cloroplastos. Mayor contenido en hierro presumiblemente causará mayores reacciones redox, conllevando un mayor daño oxidativo. Por lo contrario, menor contenido en hierro haría que las semillas sufrieran menos de estos eventos oxidativos causantes del envejecimiento, por lo que sería ventajoso para su longevidad.

Capítulo 2:

El factor de transcripción AtHB25, conservado en múltiples especies de plantas, regula la deposición de cutina y suberina en la testa

Artículo completo en inglés en página 81

Antecedentes

El factor de transcripción AtHB25 se caracterizó en nuestro laboratorio gracias a un rastreo en una colección de mutantes realizado por Eduardo Bueso. El mutante *athb25-1D* se aisló por presentar una fuerte resistencia al tratamiento de envejecimiento acelerado (AAT). La caracterización de este mutante, previa a mi llegada al laboratorio, desveló que dicha longevidad presentaba heredabilidad materna, sugiriendo que la sobreexpresión de *AtHB25* producía un efecto en la cubierta de la semilla, ya que es único órgano de la semilla de origen completamente materno. De los componentes de la cubierta analizados el mucílago resultó estar aumentado, aunque en este estudio hemos demostrado que el mucílago influye en la longevidad de semilla ya que semillas mutantes sin mucílago no presentan cambios en la longevidad, e incluso la sobreproducción de éste afecta negativamente a la longevidad (Figura 34 en la página 87). También, a través de tinciones de los poliésteres lipídicos, confirmó que la cubierta de la semilla presentaba una mayor acumulación de éstos, lo que explicaría la mayor longevidad de las semillas del mutante *athb25-1D*. Además las plantas mostraban mayores niveles de las hormonas giberelinas y ácido abscísico, tanto en planta como en semillas en desarrollo. Estos niveles hormonales, en especial el de las giberelinas, se utilizó como explicación inicial de la alta longevidad de las semillas, ya que estas hormonas son importantes en el desarrollo de la semilla, incluyendo la cubierta. Mutantes que fallan en la percepción de giberelinas resultó que presentaban menor longevidad, sugiriendo esta hormona como esencial para la adecuada adquisición de longevidad. La mayor cantidad de giberelinas de *athb25-1D* se explicó al observarse la sobreexpresión de un gen biosintético de giberelinas en este mutante.

Por otra parte, aunque el mutante simple de pérdida de función *athb25* no presenta menor longevidad, el doble mutante (realizado la línea mutante del gen de Arabidopsis más cercano u homólogo, *AtHB22*) sí que la presenta. Este efecto, se debe a la redundancia genética. Los genes aparecen gracias a la duplicación génica, y pueden adquirir nuevas funciones por mutación. En un modelo simple, un gen esencial permanece sin mutaciones importantes que impidan su función. Sin embargo, si se ha producido una duplicación génica, una de las dos copias podrá mutar y modificar su función, hasta adquirir otra que hará que se mantenga evolutivamente, o incluso podría llegar a mutar tanto que puede perder toda función. Por duplicación genética se forman las familias

génicas, con funciones similares pero los distintos miembros pueden ser regulados de forma distinta, por ejemplo en distintos tejidos o ante distintas señales ambientales. Sin embargo, un efecto que se ha visto es la de la redundancia génica: incluso cuando los genes se han diferenciado y especializado, al mutarse uno es posible que el otro supla su función. Incluso puede pasar que el gen redundante se sobreexprese para compensar la falta del otro gen.

En este estudio se ha procedido a caracterizar en profundizad estos mutantes, *athb25-1D* y el doble mutante *athb22 athb25*, y su papel regulador de los genes que sintetizan poliésteres lipídicos de la cubierta de la semilla. Además, se ha visto que este gen de Arabidopsis, altamente conservado en trigo y tomate, tiene un efecto similar en las semillas al expresarse de forma ectópica.

AtHB25 y la longevidad de semilla

Con los resultados mostrados en la Figura 31 en la página 85 demostramos que el factor de transcripción AtHB25 permite que las semillas sobrevivan por mas tiempo al almacenarse a temperatura ambiente (a), y que dicho efecto seguramente se deba a su menor permeabilidad de la cubierta medida con sales de tetrazolio (c), lo que influye en la oxidación del embrión, medido a través del ratio del glutatión oxidado / glutatión reducido (b). Gracias a líneas transgénicas transformadas con el promotor de *AtHB25* unido a las proteínas GUS y GFP (unidas covalentemente) se demostró que el gen *AtHB25* se expresa durante la etapa de maduración del desarrollo de las semillas en la cubierta de la semilla. La señal GUS, en azul, sitúa primero la expresión de *AtHB25* en el tegumento interno y posteriormente, también en el tegumento externo (d). La señal GFP, que puede ser analizada *in vivo*, estaba enmascarada parcialmente por la autofluorescencia de las biomoléculas que sintetiza la cubierta de la semilla, como la lignina. Esta autofluorescencia (en rojo) se consiguió separar de la señal GFP (en verde) gracias al barrido de longitudes de onda y el análisis del espectro de emisión de la proteína GFP. La expresión de *AtHB25* se confirmó en los tegumentos internos, y en los tegumentos externos incluyendo la capa en empalizada y, en menor medida, la capa exterior de células columela (e).

Para demostrar la hipótesis de que AtHB25 regula la deposición de poliésteres lipídicos en la cubierta de la semilla, se realizaron diversos experimentos. Primero, se confirmó que los mutantes *athb25-1D* y *athb22 athb25* mostraban una mayor y menor tinción de lípidos en la cubierta al delipidarse las semillas, respectivamente (Figura 34a en la página 88). Al delipidar, se quitan los lípidos no polimerizados, quedando únicamente los poliésteres lipídicos. Para la visualización de las capas de suberina y cutina se realizaron cortes ultrafinos que se visualizaron al microscopio electrónico de transmisión (TEM). La imagen se centró en la cubierta de la semilla (Figura 34b) para

poder distinguir todos los componentes de la cubierta de la semilla (SC), como las células de la columela (CO) que crean el mucílago (dMU), la capa de células en empalizada (PL) que sintetizan la capa de suberina (SL), la capa pigmentada marrón (BPL) formada por las distintas capas del tegumento interno comprimidas y teñidas por las proantocianidinas, y finalmente, junto al endospermo (END) y su pared celular (CWE) se puede distinguir la cutícula de la cubierta de la semilla (CL), junto a la pared celular del endotelio (TCW) teñida de negro. Estos componentes están esquematizados en la Figura 34c, y se pueden distinguir en las imágenes tomadas con el microscopio electrónico de los distintos mutantes del gen *AtHB25*, en la visión general de la cubierta de la semilla (Figura 34d), y las imágenes centradas en la capa de suberina (Figura 34e) pudiéndose distinguir su patrón laminado característico, y en la capa de la cutina (Figura 34f) donde se observa la deposición amorfa de la cutina formando la cutícula. Midiendo diferentes zonas pudimos estimar que, tanto la suberina (Figura 34g) como la cutícula (Figura 34h), estaban engrosadas en el mutante *athb25-1D* y eran más finas en el doble mutante *athb22 athb25*.

Bioquímicamente analizamos la composición de los poliésteres lipídicos de la cubierta de la semilla mediante espectrometría de masas, realizada en colaboración con el laboratorio de Isabel Molina en Canadá. Los resultados de este análisis se ilustran en la Figura 36 en la página 90. Como se puede observar el mutante de sobreexpresión contenía mayor cantidad de lípidos polimerizados que las semillas control, y el doble mutante contenía menos (c). El principal tipo de monómeros que mostró una clara diferencia entre los distintos mutantes fue los ácidos dicarboxílicos (DCA) (b), y específicamente el monómero con más cambios fue el 18:2 DCA (a). Estos resultados corroboran el papel regulador de AtHB25 en la acumulación de poliésteres lipídicos de la cubierta de la semilla, regulando tanto la deposición de suberina como la de cutina. Ahora, queríamos saber qué genes regulaba AtHB25 para producir este efecto.

La técnica genómica: el ChIP-seq

Para conocer los genes diana de AtHB25, realizamos un experimento complejo llamado ChIP-seq. Este experimento consta de dos partes, la inmunoprecipitación de la cromatina (ChIP) y la secuenciación masiva de los fragmentos de DNA enriquecidos tras la inmunoprecipitación. Con este experimento se obtienen las regiones de genoma a las que se une la proteína de interés. Si la proteína a analizar se trata de un factor de transcripción, como es el caso, las regiones enriquecidas obtenidas pueden indicar la zonas genómicas que puede regular dicho factor de transcripción. Si una zona en la que se une el factor de transcripción está cercana a un gen, puede ser que el gen se regule (es decir que cambie su expresión o transcripción a RNA) directamente por este factor de transcripción. Análisis de expresión génica pueden utilizarse para conseguir

confirmar si los mutantes en el factor de transcripción presentaban cambios en los niveles de RNA de dicho gen. Si es así, este gen se confirmaría como gen diana del factor de transcripción analizado. Un factor de transcripción puede ser que actúe como activador, induciendo la expresión del gen, o de represor, reprimiendo la expresión del gen. Aunque la regulación génica es compleja e intervienen muchos otros factores de transcripción y otros factores moleculares como la compactación de la cromatina, la metilación de la misma, interacción con otras proteínas, la degradación de factores de transcripción por señales específicas, o por la competencia con otros elementos por la unión al DNA, conocer los genes diana de los factores de transcripción es de gran interés para entender el efecto de estos reguladores génicos, que constituyen el principal proceso de regulación de la expresión génica.

La técnica ChIP-seq requiere del uso de un anticuerpo específico que solo se una a nuestra proteína de interés. Como normalmente no disponemos de anticuerpos para la proteína de estudio es común realizar una proteína quimérica recombinante que incluya la proteína de estudio y un fragmento de proteína que es reconocido eficientemente por un anticuerpo. Este fragmento reconocido, llamado epítopo, es preferible que provenga de otro organismo para mejor purificación de nuestra proteína quimérica. En este estudio se utilizó el epítopo HA, un fragmento de proteína de corazón humano, claramente no presente en la planta Arabidopsis. Una vez obtenida la planta transgénica, comprobamos que la proteína recombinante se expresaba y localizaba en núcleos de semillas en desarrollo y procedimos con el experimento. Dicha planta fue obtenida a través de la transformación de la planta doble mutante *athb22 athb25* para evitar competencia entre el factor de transcripción endógeno y el recombinante, y conseguir así mayor cantidad de sitios de unión del factor de transcripción libres para que se una nuestra proteína recombinante.

El experimento en sí consiste en unir la proteína covalentemente al DNA gracias a la acción del formaldehído, romper el DNA en pequeños fragmentos de 200 pares de bases mediante sonicación, inmunoprecipitar el complejo DNA-proteína recombinante gracias al uso del anticuerpo específico y bolas, magnéticas en nuestro caso, que se unen al anticuerpo, y finalmente romper la unión DNA-proteína para poder purificar el DNA y secuenciarlo. Se suele utilizar una muestra a la que no se le añade anticuerpo como control negativo. Tras la secuenciación, se analizan bioinformáticamente los resultados. Primero se eliminan las secuencias de mala calidad, luego se mapean respecto al genoma de referencia, y finalmente se comparan las muestras con la muestra control para determinar estadísticamente las zonas genómicas enriquecidas por la inmunoprecipitación. Estas zonas, o picos, llamados así debido al enriquecimiento con distribución normal de la zona, constituyen las regiones de unión de la proteína en el genoma. En la parte de arriba Figura 37e en la página 92 se pueden ver dos regiones que muestran enriquecimiento en las muestras con anticuerpo (rep1 y rep2, en azul)

respecto a la muestra sin anticuerpo (no Ab, en rojo) correspondientes a los picos 19 y 20 obtenidos tras el análisis. Arriba en verde se puede ver que los picos se encuentran cercanos a un gen. Un pico cercano a un gen sugiere la regulación de dicho gen por AtHB25.

Tras nuestro análisis ChIP-seq, obtuvimos 146 picos. Fueron pocos picos, ya que obtuvimos un porcentaje bajo de secuencias que mapeaban en Arabidopsis. Esto podría deberse a un bajo rendimiento de la inmunoprecipitación y posterior purificación de DNA. Este efecto igual se debe a la dificultad del tejido escogido ya que las semillas presentan gran cantidad de componentes que dificultan las extracciones y experimentos. Sin embargo, nuestro experimento ChIP-seq quedó ampliamente validado, ya que un gran porcentaje de regiones obtenidas se localizaban en regiones reguladoras de genes, principalmente aguas arribas del gen o región promotora (81.5%, Figura 37a y b). Además, un 83% de las regiones presentaban uno o más motivos de unión del factor de transcripción AtHB25 (Figura 37d). El motivo de unión, ilustrado en la Figura 37c, se calculó a través de las regiones genómicas enriquecidas en el ChIP-seq. Representa la probabilidad de unión de una secuencia de DNA por el factor de transcripción, representado la mayor probabilidad con letras de mayor tamaño. Sabemos que se trata del motivo del factor de transcripción AtHB25 debido a que experimentos publicados determinaron que su motivo de unión es T(G/A)ATTA, coincidiendo con el motivo más enriquecido obtenido. Otros motivos también se obtuvieron, además de otros miembros de la familia de factores de transcripción homeobox de la misma subfamilia que AtHB25 que presenta motivos de unión muy similares (Tabla 10 en la página 95, y Datos suplementarios 6 en la página 261). Los motivos de unión que no corresponden a los dominios de genes homeobox pueden estar enriquecidos debido a que comúnmente se encuentre próximos a los motivos homeobox, o debido a que el factor de transcripción directamente interaccione con otro factor de transcripción, ya que indirectamente se arrastrarían dichas secuencias, es decir, coimmunoprecipitarían. Esto último explica las regiones cromosómicas enriquecidas que no muestran motivos de unión de factor de transcripción de estudio, que corresponde a un 17% en nuestro estudio.

De las 146 regiones genómicas en las que se une AtHB25, dos de ellas, las correspondientes a los picos 19 y 20, se encuentran cercanas a genes, *LACS2* y *NHO1* respectivamente, que podrían estar implicados directamente en la biosíntesis de monómeros de los poliésteres lipídicos y explicar la acumulación de poliésteres lipídicos incrementada en la testa de las semillas del mutante *athb25-1D*. Se ratificó el enriquecimiento de estas regiones por PCR cuantitativa, de estas regiones y de zonas génicas (Figura 37e) y de otras regiones enriquecidas y no enriquecidas (Figura 38 en la página 96). Para confirmar si realmente AtHB25 regula directamente estos genes, se analizó la abundancia relativa de los transcritos en los diferentes semillas en desarrollo

de las plantas mutantes de *AtHB25* y de plantas silvestres. Los resultados apuntan a que AtHB25 directamente regula positivamente a *LACS2* y *NHO1*, es decir, *LACS2* y *NHO1* son genes diana del factor de transcripción AtHB25 (Figura 37f).

LACS2 y NHO1 son esenciales para la longevidad incrementada de *athb25-1D*

A través del uso de plantas mutantes de pérdida de función de los genes diana de AtHB25, *lacs2* y *nho1*, y los respectivos cruces con la planta mutante *athb25-1D* se pudo determinar que ambos genes, *LACS2* y *NHO1* son esenciales para la aumentada deposición de poliésteres lipídicos en mutante *athb25-1D*, responsable, al menos en parte, de la mayor longevidad de semilla. En la Figura 40 en la página 99 se puede ver como las semillas mutantes *lacs2* tienen una alta sensibilidad al tratamiento de deterioro controlado de la semilla (CDT), pero las de *nho1* no son significativamente menos longevas, aunque llega a la mitad de la germinación (P50) un par de días antes que las semillas silvestres. Los cruces de ambos mutantes con el mutante *athb25-1D* mostraron una menor longevidad que incluso las semillas silvestres (Figura 40a). La permeabilidad reducida de *athb25-1D* a las sales de tetrazolio se vio afectada (Figura 41c) y la tinción de poliésteres lipídicos se vio reducida (Figura 40d) en ambos cruces.

Para confirmar esto se hicieron análisis bioquímicos de la composición de poliésteres lipídicos. Aunque NHO1 no mostró una clara reducción, el doble mutante *athb25-1D nho1* mostró que el efecto de la sobreexpresión de *athb25-1D* quedaba anulado en ausencia de NHO1 (Figura 46 en la página 102). *NHO1* codifica una glicerol kinasa que cataliza la unión de un grupo fosfato a la molécula del glicerol formando glicerol 3-fosfato, componente de los poliésteres lipídicos. NHO1 es la única glicerol kinasa presente en el genoma de Arabidopsis, y representa un paso limitante en la síntesis de glicerol 3-fosfato. Sin embargo, éste también puede ser obtenido de la glicólisis (Figura 7 en la página 34). Esto puede explicar que el mutante *nho1* no presente una muy reducida deposición de poliésteres lipídicos. Por lo tanto, podría ser que la regulación de *NHO1* por AtHB25 ayude a controlar la mayor o menor deposición de poliésteres lipídicos. La transferencia del glicerol 3-fosfato a los monómeros de suberina mediante el enzima GPAT5, es esencial para la permeabilidad y la longevidad de las semillas silvestres y del mutante *athb25-1D* (Figura 42 en la página 100).

Por otra parte, el análisis bioquímico de las semillas *lacs2* mostró que éstas mostraban una alta reducción en los monómeros de poliésteres lipídicos (Figura 44 en la página 101). Concretamente, los monómeros que presentaban mayor reducción eran los ácidos dicarboxílicos, especialmente el 18:2, seguidamente del ferulato, y del ácido graso hidroxilado 24:0 (Figura 45 en la página 102). LACS2 es un enzima que activa los

ácidos grasos, mediante la unión del CoenzimaA. Esta activación es necesaria para el transporte de los monómeros iniciales de ácidos grasos, para la transformación de los ácidos grasos en ácidos grados hidroxilados, paso previo para la biosíntesis de los ácidos grasos dicarboxílicos y para la unión covalente del glicerol al ácido graso (Figura 7 en la página 34 y Figura 9 en la página 37). La activación directa de LACS2 por AtHB25 explica la diferencia en los niveles de los ácidos grasos dicarboxílicos en el análisis bioquímico de los mutantes del gen *AtHB25* (Figura 36 en la página 90). Los ácidos grasos dicarboxílicos al presentar doble grupo carboxílico, uno en cada extremo de la molécula deben permitir la formación de un entramado molecular de cutina y suberina de mayor complejidad, que debe crear unas barreras impermeabilizantes de mayor consistencia determinantes en la mayor longevidad de este mutante. LACS2 se ha descrito principalmente en la síntesis de la cutina, y es el único de los nueve genes LACS con un papel confirmado en la cubierta de la semilla que afecta a la longevidad de la semilla, aunque también podría estar involucrado en la síntesis de monómeros de suberina. Además, otros miembros de la familia también podrían ser determinantes en la producción de poliésteres lipídicos de la cubierta de la semilla.

También vimos que otros genes de biosíntesis de los monómeros de cutina y suberina presentaban cambios de expresión en las semillas en desarrollo de plantas mutantes de gen *AtHB25* (Figura 39 en la página 97) que podrían explicar otros cambios observados. Los cambios de expresión de estos genes, regulados indirectamente por AtHB25, podrían ser resultado de la alteración hormonal presente en el mutante de sobreexpresión *athb25-1D*.

AtHB25 funciona de forma similar en otras plantas

Para determinar si el papel de AtHB25 estaba conservado en otras plantas de interés agronómico y así conseguir mejorar la longevidad de semilla de diversos cultivos se escogieron dos especies vegetales de cultivo, el tomate y el trigo. Se determinó que ambas especies tenían proteínas con una alta homología a AtHB25, mostrando los dominios de la proteína muy conservados (Figura 50 en la página 110). Esto sugiere que la función de AtHB25 permanece conservada en las distintas especies vegetales por su importancia, que podría ser la regulación de los poliésteres lipídicos en las semillas para controlar su longevidad. La expresión ectópica del gen de Arabidopsis AtHB25 el en tomate (Figura 48 en la página 105) y trigo (Figura 47 en la página 104) también confirió a las semillas una mayor longevidad de semilla y una mayor deposición de poliésteres lipídicos (en estos casos, cutina) produciendo cubiertas de semillas más impermeables. Este resultado abre la posibilidad de utilizar el factor de transcripción de Arabidopsis AtHB25 para mejorar la longevidad de multitud de plantas de cultivo, ya que el tomate, y Arabidopsis son dicotiledóneas y el trigo, más alejado evolutiva-

mente, es una monocotiledónea, indicando que probablemente también funciones en muchas otras verduras y cereales.

Capítulo 3:

Las peroxidasas PRX2 y PRX25, reguladas por COG1 son importantes para la longevidad de la semilla

Artículo completo en inglés en página 111

Antecedentes

El mutante de sobreexpresión del gen COG1, *cog1-2D*, se obtuvo tras un rastreo de mutantes, llevado a cabo de forma similar al rastreo realizado en el que se aisló el mutante *athb25-1D* (véase página 299). COG1 se trata de un factor de transcripción al igual que AtHB25, pero pertenece a una familia distinta, la familia de factores de transcripción DOF. Curiosamente se encontraron fenotipos comunes en ambos mutantes, como el hecho de que su mayor longevidad venía dada por una mayor deposición de poliésteres lipídicos en la cubierta de semilla y que ambos presentaban mayores niveles hormonales de las hormonas giberelinas y ácido abscísico en las semillas en desarrollo. Sin embargo, otros fenotipos como la velocidad de floración, eran contrarios (*athb25-1D* presenta floración temprana, mientras que *cog1-2D* presenta floración tardía). La tardía floración de los mutantes de sobreexpresión de *COG1* se pueden ver en la Figura 69 en la página 140 (Capítulo 4). El mutante *cog1-2D* parece presentar menos cantidad de la proteína COG1, mientras que las plantas transgénicas *proUBQ10::COG1:3xHA* parece presentar mayor cantidad, ya que son más pequeñas y crecen con mayor dificultad, seguramente por el efecto incrementado de COG1. De hecho este fenotipo de planta pequeña (muy distinto al observado en la planta *cog1-2D*) se ha descrito en el mutante de sobreexpresión *cog1-1D*. Con este mutante se determinó que COG1 está involucrado en la percepción de luz, regulando negativamente la señalización del fitocromo B, involucrado en la percepción de luz roja y roja-lejana. Por ello, los mutantes de sobreexpresión de COG1 muestran elongamiento del hipocotilo, hojas de la roseta grandes (en el caso de *cog1-2D*) y mayor número de las mismas y un retraso de la floración, ya que todos estos fenotipos son fenotipos de falta de luz o de fallo de percepción de esta. En nuestro laboratorio se determinó que otros mutantes en la percepción de luz, como en mutante *phyB* mostraban también una mayor longevidad de semilla e incrementada deposición de poliésteres lipídicos, conectado la percepción de luz con la longevidad de semilla.

La técnica genómica: el RNA-seq

Aunque intentamos hacer un ChIP-seq como el realizado en el estudio de AtHB25, no pudimos obtener regiones enriquecidas debido, probablemente a la baja cantidad de proteína COG1 (en comparación a la cantidad de proteína recombinante de AtHB25,

Figura 70 en la página 141), ya que altas dosis conlleva la muerte de las plantas. Por ello, se realizó un estudio completo de RNA o RNA-seq de la planta completa *cog1-2D*. De esta forma se puede conocer todos los genes regulados directa e indirectamente por el efecto de este factor de transcripción.

La técnica consiste en la extracción de todo el conjunto de RNA la planta y su secuenciación masiva. Sin embargo, para poder obtener los genes influenciados por dicho factor de transcripción también se debe secuenciar plantas silvestres o plantas control para así comparar los transcritos de las diferentes muestras. No existe un estándar de RNA de planta control, ya que el conjunto de RNA o transcriptoma, depende de muchos factores, como el tejido, la edad, las condiciones de crecimiento, el estrés sufrido de la planta y de otros factores incontrolables. Por ello, se debe analizar mínimo de tres muestras independientes (réplicas biológicas) de cada línea, para así poder realizar estudios estadísticos adecuados para determinar los genes diferencialmente expresados entre las distintas muestras.

Como no todos los archivos obtenidos tras la secuenciación masiva del RNA presentan el mismo número de millones de lecturas, se deben estandarizar las lecturas por transcrito obtenido, tras el mapeo de las secuencias al transcriptoma de referencia, para poder comparar la abundancia relativa de los transcritos. Esta estandarización se basa en determinar la abundancia de secuencias por kilobase de transcrito y por millón de lecturas, ya que transcritos más largos, de por sí obtendrán mayor ratio de lecturas. Una vez obtenida la abundancia, se puede calcular cuánto sube o cuánto baja la expresión de todos los genes en la muestra respecto a la muestra control, la variación de la misma y la estadística (con corrección multiparamétrica por la gran cantidad de datos) asociada. Comúnmente los genes que significativamente se encuentran sobreexpresados o reprimidos en la muestra más de el doble o la mitad, respectivamente, se consideran como los genes sobreexpresados y reprimidos entre las distintas muestras.

Los resultados finales son largas listas de genes. Por ejemplo, en el RNA-seq de *cog1-2D* obtuvimos 1.360 genes sobreexpresados y 1.764 genes reprimidos. Trabajar con grandes listas de genes es difícil. Sin embargo existen estrategias para comprender el papel regulador de este factor de transcripción. Una de estas estrategias es el análisis de los términos GO. Los términos GO son descripciones técnicas, clasificadas en niveles, de las funciones moleculares de los genes, la localización de la expresión de los mismos, y sus funciones biológicas. Con estas cortas descripciones se puede analizar si hay algún término que aparece más veces de las que podría aparecer por azar, es decir si está enriquecido. Los términos GO enriquecidos nos sirven para poder intuir el papel regulador de este factor de transcripción o, en términos más generales, a qué procesos afecta las distintas condiciones analizadas mediante el RNA-seq. El análisis de los términos GO de función molecular enriquecidos del RNA-seq de *cog1-2D* se muestran en la Figura 52 en la página 115

COG1 regula las peroxidasas

Uno de los términos GO que salió sobrerepresentado en la lista de genes sobreexpresados en el mutante *cog1-2D* fue GO:001649, que corresponde a la función molecular de actividad oxidoreductasa, asociado a 273 genes de la lista (Figura 52). De estos genes vimos que trece de ellos era nperoxidasas de tipo III. Esto sugirió un papel activo de COG1 en la regulación de estos enzimas, que catalizan la creación de radicales de fenilpropainodes, permitiendo su espontánea polimerización para formar lignina y, se piensa que también suberina.

Para determinar la expresión de COG1 se usó una línea reportera transformada con el promotor de COG1 (dos kilobases previas al inicio del gen) que dirige la expresión la proteína quimérica GUS:GFP. La fluorescencia emitida por la proteína GFP se localizó en la cubierta se la semilla en desarrollo, principalmente en el tegumento externo, con gran intensidad durante el desarrollo de la columela de la capa de células más externa (Figura 54a en la página 117), que corresponde aproximadamente a los siete días de la polinización de la semilla DAP). Por ello, analizamos la expresión de las diversos genes de peroxidasas en silicuas a los siete días de la polinización. Resultó que *PRX2*, *PRX12*, *PRX22*, *PRX25* y *PRX43* se encontraron inducidos tanto en las silicuas del mutante *cog1-2D* (Figura 54b) como de plantas tratadas con giberelinas (Figura 54c). Esto sugiere que la regulación de las peroxidasas por COG1 se debe al aumentado nivel de hormonal de giberelinas presente en las semillas en desarrollo.

Previamente no se han descrito peroxidasas con un papel importante en la longevidad de semilla. En un intento de caracterizar alguna con un papel esencial en la longevidad, se obtuvieron las líneas mutantes de pérdida de función de aquellas peroxidasas que se expresaban durante el desarrollo de la semilla. Sin embargo, ninguna de estas líneas mostró sensibilidad al tratamiento de envejecimiento acelerado AAT (Figura 53 en la página 116). Como no se encontraron mutantes simples con fenotipo de longevidad, se optó por probar con los mutantes dobles y triple en los genes *PRX2, PRX25* y *PRX71*, descritos en la lignificación. Con estos mutantes se consiguió demostrar que el doble mutante *prx2 prx25* y el triple *prx2 prx25 prx71*, tenían su germinación afectada tras los tratamiento de envejecimiento artificiales AAT y CDT.(Figura 55 en la página 118). No sólo eso, las semillas mostraron también una mayor permeabilidad de la cubierta a las sales de tetrazolio (Figura 59 en la página 122). A través del microscopio electrónico y cortes ultrafinos de las semillas en desarrollo, se pudo observar que el grosor de las capa en empalizada de la cubierta de la semilla, donde se deposita la suberina, estaba disminuida en estos mutantes múltiples de estas peroxidasas (Figura 56 en la página 120). Estas imágenes fueron obtenidas previamente a las mostradas en el Capítulo 2 de los mutantes de *AtHB25*, y no supimos enfocar bien el patrón laminado de la suberina. No obstante, realizamos una análisis de los compuestos polifenólicos unidos a la pared celular y vimos que los mutantes mostraron una importante

reducción respecto a los polifenólicos cuantificados en semillas silvestres (Figura 58 en la página 121). Las cubiertas de semilla de estos mutantes no mostraban otros fenotipos de cubierta de la semilla como diferencias en la deposición de proantocianidinas o en el mucílago (Figura 57 en la página 120), indicando que la baja longevidad y alta permeabilidad se debía a la regulación de deposición por las peroxidasas PRX2 y PRX25 de los polifenoles en las paredes de la cubierta de la semilla. Queda en duda si estas peroxidasas regulan la polimerización de los polifenoles de la lignina únicamente, o también de los de la suberina. Se ha postulado teóricamente que para la polimerización de los ácidos hirdroxianímicos de la suberina es probable que se requieran peroxidasas, y por ello, pensamos que éstas pueden tener un papel clave en la polimerización de la suberina. El papel regulador de COG1 sobre las peroxidasas y en la deposición de poliésteres lipídicos, la expresión de *COG1* en las células de los tegumentos externos y las diferencias observadas en el grosor de la capa en empalizada en los mutantes doble *prx2 prx25* y triple *prx2 prx25 prx71* apoyan esta hipótesis. Sin embargo aún queda por descubrir como se polimeriza la suberina y la estructura real de la misma.

Capítulo 4:

AtHB25 se regula por frío y COG1 regula la percepción de la luz, para determinar la adaptación ambiental de la longevidad de la semilla

Artículo completo en inglés en página 127

Interacciones proteína-proteína

En el capítulo 2 caracterizamos qué genes regulaba AtHB25, y demostramos que la longevidad incrementada del mutante *athb25-1D* se debía a la incrementada deposición de poliésteres lipídicos resultante, en mayor medida, por sobreexpresión de *ACS2* y *NHO1*. Sin embargo, con este estudio no se puede saber la regulación endógena de *AtHB25*. Es decir, sabemos que regula AtHB25 pero no qué regula a *AtHB25*. Con COG1 tenemos exactamente el mismo problema.

Primero de todo realizamos un ensayo de interacción de proteína-proteína, que se realiza en levadura. Esta técnica, llamada *Split-Tpr*, es similar a la famosa técnica del doble híbrido. Ambas consisten en unir a dos proteínas fragmentos complementarios de una proteína reportera. Si las proteínas del estudio interaccionan, los fragmentos de la proteína reportera estarán cercanos y se reconstituirá dicha proteína. Por lo tanto, la interacción se determina a través de la actividad de la proteína reportera. Mientras que en el doble híbrido esta proteína se trata de un factor de transcripción, que permite la transcripción de un gen de síntesis de aminoácidos necesario para el crecimiento de la levadura (si este aminoácido no se encuentra en el medio), en *Spilt-Tpr* la proteína reconstituida si hay interacción es directamente un enzima de síntesis de el aminoácido necesario, que en este caso se trata del triptófano. Como ventaja frente al doble híbrido, el *Split-Tpr* permite que las interacción se produzca en el citosol, respecto al doble hibrido que requiere que ocurra en el núcleo. Además, a la hora de trabajar con factores de transcripción, el ensayo *Split-Tpr* permite evitar efectos de autoactivación, pues podría ser que uno de los factores de transcripción sea capaz de unirse al promotor y activar la transcripción del gen reportero directamente. Sin embargo, una de las desventajas principales es que el *Split-Tpr* tiene una menor sensibilidad que el doble hibrido presentado dificultades para encontrar interacciones más débiles. Ensayamos todos los factores de transcripción que tenían un papel descrito en el desarrollo de la cubierta de la semilla y la identidad de la misma, pero no encontramos ningún factor de transcripción que interaccionara con AtHB25 (ni tampoco con COG1 o MYB47). El único resultado positivo que obtuvimos fue que AtHB25 interacciona consigo mismo, formando heterodímeros, como ya se había descrito previamente.

La técnica genómica: el rastreo mediante el sistema Y1H

Para poder determinar qué genes regulaban a AtHB25, se decidió hacer un rastreo en levaduras de interacción DNA-proteína. Este rastreo es similar a la técnica del doble híbrido o *yeast-two-hybrid* (Y2H), por lo que al participar sólo una proteína se le ha llamado *yeast-one-hybrid* (Y1H). Esta técnica consiste en la activación transcripcional de un gen que permite la supervivencia de la planta, al igual que en el doble híbrido. No obstante, la interacción debe producirse entre la proteína quimérica y un fragmento de DNA seleccionado situado en la región promotora del gen. Para este ensayo se seleccionó un fragmento de DNA situado en el promotor de AtHB25 conservado en diversas especies de plantas no muy lejanas a Arabidopsis, resaltado en la Figura 60 página 131. Al utilizar el fragmento del promotor más conservado entre plantas, conseguimos que los resultados obtenidos (proteínas que se unen a este fragmento) probablemente estén relacionadas con la regulación más importante por la que se regula AtHB25, ya que esta secuencia reguladora ha sido conservada sin admitir mucha variación. De esta forma, podremos conocer el papel conservado de AtHB25. Comprobamos que dicho fragmento presentaba actividad transcripcional en planta mediante la actividad de un gen reportero en hojas de *Nicotiana benthamiana* (Figura 61 en la página 132).

Como proteínas a ensayar utilizamos una colección de factores de transcripción creada por Luis Oñate, en Madrid. En colaboración con ellos realizamos el rastreo probando la unión de 1.358 factores de transcripción a nuestro fragmento de DNA del promotor de *AtHB25* seleccionado. Tuvimos que usar un inhibidor competitivo de la histidina, ya que la levadura presentaba un poco de autoactivación. Obtuvimos cuatro colonias en el rastreo Y1H, que correspondían a cuatro factores de transcripción (HB52, AGL24, bZIP52 y DREB2H) (Figura 62 en la página 133). Sin embargo sólo uno de ellos se expresaba durante el desarrollo de la semilla, según las bases de datos (Figura 63a en la página 134). Este gen, *DREB2H*, podría contextualizar el papel de *AtHB25* en el desarrollo de cubierta de la semilla y su control de la longevidad. Los genes *DREB* son genes de respuesta al estrés .y justamente éste se induce por frío (Figura 63b).

AtHB25 en la respuesta al frío

Comprobamos que *AtHB25* se inducía por un choque de 24 horas a 8 ºC mediante el uso de las líneas reporteras GUS:GFP con el promotor de AtHB25 utilizadas previamente (Figura 64a en la página 135). Falta comprobar si esta inducción es dependiente de DREB2H. Además, vimos que el mutante de sobreexpresión *athb25-1D* crecía mejor, y con menos estrés, que las plantas silvestres tras un frío continuado, mientras que las plantas del doble mutante *athb22 athb25* crecían peor e incluso llegaban a mo-

rir (Figura 64b). Estos dos resultados demuestran un papel importante de AtHB25 en la regulación de la respuesta a bajas temperaturas.

Curiosamente, uno de los picos, el más enriquecido, de los obtenidos a través de ChIP-seq de AtHB25 (Capítulo 2) se encontraba en la región promotora del gen ICE1, que codifica para un factor de transcripción relacionado con la respuesta al frío. Vimos que las semillas en desarrollo del mutante *athb25-1D* presentaban más cantidad de transcritos de ICE1. En cambio, el doble mutante no mostró una importante reducción en la semilla, pero sí en la respuesta de la planta al inducir ICE1 tras un golpe de frío de 24 horas (Figura 65 en la página 136). Este factor de transcripción, además tiene un importante papel (junto a SPCH) en la diferenciación de las células de los estomas, y en la ruptura del endospermo para la asimilación de los nutrientes por el embrión y para el crecimiento de éste en la cavidad de la semilla. Además, participa en la regulación de las hormonas giberelinas y ácido abcísico. Esta regulación podría explicar la mayor acumulación de ambas hormonas en el mutante *athb25-1D*. De hecho, la planta doble mutante *athb22 athb25* presenta un fenotipo similar en tamaño y velocidad de floración al de plantas mutantes *ice1* (Figura 66 en la página 137), lo que podría deberse a esta regulación de hormonas mediante la regulación de AtHB25 a *ICE1*.

Analizamos las semillas mutantes de pérdida de función *ice1* para determinar si presentaban una menor longevidad de semilla. Y así fue, las semillas envejecían rápidamente frente a los tratamientos de envejecimiento artificiales AAT y CDT (Figura 67a n la página 138). Un fenotipo que notamos fácilmente fue que las semillas frescas presentaban un color marrón mucho más oscuro (Figura 67d), similar a semillas antiguas, lo que indicaba una alta permeabilidad de la semilla. El ensayo de permeabilidad a las sales de tetrazolio demostró que, en efecto, las semillas de *ice1* eran muy permeables (Figura 67e). Sin embargo, la deposición de poliésteres lipídicos no parecía reducida con la tinción de la Figura 67f, contrariamente a lo que podríamos esperar de la regulación de AtHB25. En esta Figura se puede ver papel que tiene ICE1 en la ruptura del endospermo, ya que al no estar este factor de transcripción el endospermo no se rompe y el embrión no puede crecer y se queda constreñido. Este efecto podría ser la causa de la mayor permeabilidad de semilla, ya que una gran presión interna podría producir fisuras en la cubierta de la semilla. Al mirar las semillas al microscopio electrónico de barrido (SEM) vimos que así era, las semillas de la planta mutante *ice1* presentaban fisuras en la cubierta de la semilla, pequeñas y grandes (Figura 68 en la página 139), por las que entra el aire que rápidamente oxida las proantocianidinas. Este efecto explica directamente la baja longevidad de las semillas *ice1*. Sin embargo, no parece que este efecto esté relacionado con la regulación de ICE1 por AtHB25, y podría ser que dicha regulación influya en algún otro aspecto relacionado con la longevidad de semilla. Sin embargo, estas fisuras enmascaran cualquier otro efecto que afecte a la longevidad de semillas.

COG1 está regulado por AP2

El estudio de la regulación de *COG1* se realizó de otro modo debido a que sólo podíamos realizar, por tiempo y por dinero, un rastreo de una zona promotora y se eligió la de *AtHB25*. Sin embargo, análisis en bases de datos de la expresión de COG1 en la cubierta de la semilla indicaban que la inducción de COG1 durante el desarrollo de la cubierta de la semilla estaba reducida en el mutante *ap2-7* (Figura 71a en la página 141). En este apartado se hablará de distintas líneas mutantes sobre el mismo gen, AP2, que muestran un efecto distinto, ya que unas presentan fenotipos débiles, porque cambia solo un aminoácido de la proteína y otros muestran fenotipos muy fuertes porque tienen una inserción de T-DNA o una mutación que trunca el gen y éste pierde completamente su función. AP2 se trata de un factor de transcripción crucial en la identidad del tegumento externo de la cubierta de la semilla, además de su papel en la formación de los pétalos. Mutantes fuertes en este gen presentan un tegumento externo completamente indiferenciado, y por lo tanto, su longevidad está altamente afectada. Estas semillas no pueden ni esterilizarse con alcohol y lejía, práctica habitual para el cultivo *in vitro* de las semillas, ya que mueren rápidamente. Por otra parte, mutantes flojos sí que presentan dicha diferenciación del tegumento externo, aunque no es completa.

Para confirmar los datos de bases de datos, la misma línea transgénica que se usó para la visualización de la expresión de COG1, que se situó en la cubierta de la semilla (Figura 54a en la página 117) se cruzó con una línea mutante del gen AP2 floja, la línea *ap2-5*. Como se puede observar la expresión de la GFP por parte del promotor de COG1 en la cubierta de semilla en desarrollo queda reducida en el fondo mutante *ap2-5* (Figura 54c). Podría ser que, AP2 no estuviera regulando a COG1, ya que al perder la identidad del tegumento externo, un efecto colateral fuera que COG1 no se expresa porque no existe el tejido donde se expresa. Para salir de dudas, se realizó un análisis de la expresión de COG1 en hojas en los mutantes flojo (*ap2-5*) y fuerte (*ap2-12*), y se comprobó que COG1 se expresaba menos (Figura 54b) como en la cubierta de la semilla, descartado que la baja expresión de COG1 en la cubierta de la semilla en los mutantes *ap2* esté relacionada con la falta de identidad de capa, y confirmando que AP2 regula la expresión de COG1 positivamente.

El tegumento externo es de vital importancia para la longevidad de semilla, ya que en la capa en empalizada se localiza la suberina. Análisis publicados de los monómeros de los poliésteres lipídicos en las semillas mutantes *ap2-7* ayudaron a la caracterización en la diferencias de los monómeros de cutina y suberina, ya que este mutante no contiene suberina, al tratarse de un mutante fuerte. Este efecto de la falta de suberina de en los mutantes *ap2* se puede ver en la Figura 74 en la página 144. El alelo fuerte, *ap2-12*, no muestra tinción en las células que deberían haberse diferenciado como tegumento externo. Es más, se puede ver como las semillas de *ap2-12* son mas grandes

y redondeadas que las demás, probablemente por la falta de consistencia de la cubierta de la semilla, permitiendo el crecimiento esférico del embrión (como se ha visto en las semillas *cyp86a8* y *sep3*, pero más exagerado). Además, en esta figura se puede ver como al cruzarse con *cog1-2D*, el cual presenta una mayor deposición de poliésteres lipídicos (Figura 72 en la página 142), se recupera parcialmente esta la deposición de poliésteres lipídicos, llegando a recuperar niveles similares a los de las semillas silvestres en el cruce *ap2-5 cog1-2D*. Este mismo efecto se puede ver en la longevidad de semilla ensayada con los tratamientos de envejecimiento artificiales AAT y CDT. Los mutantes *ap2* muestran una menor longevidad que las plantas silvestres, y *cog1-2D* muestra mayor longevidad. Al cruzarse, la sobreexpresión de COG1 permite recuperar la baja longevidad del mutante *ap2-5*, pero no la del mutante *ap2-12* que apenas germina.

Para analizar el efecto de COG1 en la capa de las células columela, la capa exterior del tegumento externo, se visualizó el halo de mucílago, pero éste, reducido en las semillas mutantes *ap2-5* y ausente en la semillas *ap2-12*, no se recuperó en los cruces con el mutante *cog1-2D* (Figura 75 en la página 145). Sin embargo, al observar las semillas y los cruces bajo el microscopio electrónico de barrido (SEM) vimos que la sobreexpresión de *COG1* podía recuperar parcialmente la formación de las células columela, completamente indiferenciadas en el mutante *ap2-12*. El cruce *ap2-12 cog1-2D* mostraba células con pseudo-estructuras de volcán, directamente influenciadas por la sobreexpresión de *COG1* (Figura 76 en la página 146). Esto demuestra que COG1, aunque no totalmente, participa en la formación y diferenciación de las estructuras de las células de columela. También se puede ver en la Figura 76 que la sobreexpresión del gen *COG1* en el mutante *cog1-2D*, muestra unas células de la mas pequeñas con la uniones celulares reforzadas. Este efecto también se ve entre las imágenes de las *ap2-5* y el cruce respectivo con *cog1-2D*, que además muestra un engrosamiento en superficie de la columela.

Por lo tanto, con estos resultados, podemos concluir que aunque AP2 es el regulador maestro de la identidad de tejido de los tegumentos externos de la cubierta de la semilla, COG1 tiene un papel en su diferenciación, promoviendo la deposición de suberina y la longevidad, y participando en la diferenciación morfológica las células columela. El papel de COG1 en la regulación de la percepción de luz, y el efecto de la luz en la determinación de la longevidad de semilla, apuntan a que COG1 regula la longevidad de semillas acorde a señales de luz.

Conclusiones de la tesis doctoral

1 - La variación en la longevidad de semilla en los ecotipos ha permitido identificar y validar 12 nuevos genes involucrados en dicha longevidad. Estos genes incluyen: una proteína LEA, una unidad del fotosistema I, proteínas productoras de especies reactivas de oxígeno, una proteína que participa en la detoxificación del daño oxidativo, y diversas proteínas de la cubierta de la semilla, relacionadas con el control transcripcional y la síntesis de poliésteres lipídicos que aíslan la cubierta de la semilla.

2 - Se ha demostrado que la incrementada longevidad de las semillas de las plantas mutantes *athb25-1D* y *cog1-2D* se debe a una acumulación de poliésteres lipídicos en la cubierta de la semilla, que impermeabilizan al embrión del agua y el aire, protegiéndolo así del daño oxidativo. El papel de la cubierta de la semilla en la longevidad se ha visto esencial a través de diferentes mutantes. Sin embargo, se ha demostrado que el mucílago no tiene un papel determinante en la longevidad de semilla.

3 - AtHB25 es un factor de transcripción, expresado en ambos tegumentos de la cubierta de la semilla, que regula tanto la deposición de suberina y cutina, a través de la regulación directa de los genes LACS2 y NHO1, involucrados en la síntesis de los monómeros de estos poliésteres lipídicos. AtHB25 participa en la regulación de la respuesta al frío, probablemente a través de su regulación por DREB2H y la regulación a *ICE1*. El efecto de esta regulación puede ser clave en la regulación de la longevidad de la semilla, como en respuesta a la temperatura. La importancia de los procesos en los que participa AtHB25 se remarcan a través de la alta conservación de la secuencia en genes ortólogos de especie lejanas, su papel activo al expresarse ectópicamente. La conservación en las regiones reguladoras apuntan a un papel crucial en la adaptación, que podría estar relacionado con las bajas temperaturas.

4 - COG1 es otro factor de transcripción, expresado principalmente en el tegumento externo de la cubierta de la semilla. Regulado por AP2, participa en la diferenciación del tegumento externo. Mayores niveles de COG1 provocan un incremento en la deposición de suberina mediante, debido, en parte, a la regulación de las peroxidasas PRX2 y PRX25, a través de la regulación de los niveles de giberelinas. Estas peroxidasas son las primeras descritas en participar en la longevidad de la semilla, y aunque puedan regular sólo la polimerización de lignina, es muy probable que también participen en la polimerización de los monómeros aromáticos de la suberina. El papel COG1 en la regulación de la percepción de luz indica que puede participar en la regulación de la longevidad de semilla acorde a las señales ambientales de luz.

Agradecimientos

A todo el laboratorio 1.10

Esta tesis ha sido realizada durante cinco años en el laboratorio de Ramón Serrano, el laboratorio 1.10 del IBMCP. Me siento muy afortunado de poder haber compartido mis primeros años de investigador con gente tan buena, profesional y personalmente, que me ha arropado y guiado por esto de la ciencia, dejándome siempre esa bonita libertad para experimentar y dirigir la investigación hacia aquello que resultaba más interesante. El estilo del laboratorio de Ramón Serrano y de Consuelo Montesinos, conocida por todos como Mariche, es puro y enriquecedor. La familia que deriva de ahí, aunque últimamente se haya visto mermada, se forma por gente con un gran corazón y espíritu investigador, que ama el laboratorio y no deja de cuestionarse como funciona la vida al nivel molecular, y que como hacen los organismos para adaptarse a las diversas situaciones de estrés.

Llegué al laboratorio unos años antes de empezar la tesis, el verano de 2015, cuando todos mis compañeros de 3ero de Biotecnología buscaban como locos prácticas en laboratorios. Como uno más, me decidí a pasar un mes de verano siguiendo yendo a la universidad. Al llegar Ramón me asignó a Gaetano Bissoli para introducirme en el arte del laboratorio. Digo arte, porque cuando Gaetano pipetea alza sus brazos al cielo como si pintara la capilla sixtina. Será algún gen de esos italianos. Y no solo eso, además de enseñar cada una de las complicadas técnicas moleculares también enseña que la ciencia de laboratorio requiere de ese toque personal para que salgan las cosas, no todo esta en los libros o protocolos. Gaetano ha estado todos los días sentado a mi lado en los sitios del ordenador, de donde me ha ido dado consejos y nos hemos ido enseñando los resultados que poco a poco han ido saliendo. De él he aprendido muchas cosas buenas, pero también alguna mala como el horario tardío o la bancada desordenada, las cuales hemos ido invadiendo a lo largo de estos años. Cabe decir que no he conseguido ganarle en la bancada, ya que él lleva más tiempo perfeccionando esa técnica. En la foto del grupo se puede ver él que me gana, su bancada (la de la izquierda) está mucho más curtida que la mía (la de la derecha). Su bancada si que parece un *manhattan plot*. Y no hablemos de su congelador, algún día explotará. Aprovecho también para decir que la hora que muestra el reloj de la foto es representativa de la estrecha franja horaria en la que coincidíamos todo el laboratorio, sin contar los días de seminario. Gaetano también me ha acompañado en todas las comidas y descansillos que hacíamos desde la cafetería de la terraza del CPI cuando estaba en prácticas, hasta las innumerables comidas primero en el tranvía y luego en la cervecería del Galileo y en la cafetería del polideportivo de Beteró, donde Edu ya se nos unió de forma permanente.

Eduardo Bueso ha sido el mejor director de tesis posible, muy cercano ya desde el principio. Realmente, él es cercano con todo el mundo, es todo un relaciones pública capaz de conseguir cualquier reactivo o instrumento de otros laboratorios. Presumo de ser su primer doctorando, y he sido durante dos años y medio su compinche en las prácticas de la carrera de las asignaturas biología molecular e ingeniería genética. Ahí he podido ver como es querido por todos los estudiantes. Hemos formado un buen equipo en las clases, pero también en la investigación. Poco a poco hemos ido avanzando en la investigación de los mutantes que un día sacó en un rastreo, hace ya unos catorce años. Aunque se diga que pa la ciencia, hay que tener paciencia, Edu no la tiene. Es todo un experto en acortar protocolos, pero preciso a la hora de hacerlos. Tanto que le he pedido muchas veces que realizara el pipeteo de las qRTs. En cierto modo yo le he explotado un poquito, pero solo para que dejara el ordenador por unas horas y que se sentara en la bancada un rato, que es lo que realmente nos mola. Estos últimos meses también hemos hecho un buen equipo a la hora de subir artículos científicos: yo hago las figuras y el escribe. Además, me hizo un gran regalo al ponerme como primer autor en el artículo de COG1 (el tercer capítulo de la tesis), ya que él debutaba por primera vez como último autor, para empezar su carrera como jefe científico investigador. Serás un gran jefe Edu, y espero un día volver a formar parte de tu equipo. Con todas las publicaciones que has conseguido estos últimos años tienen que darte tu primer proyecto nacional de una vez, y así salvar al labo de años de falta de financiación. Se nota que ya se está haciendo jefe, ya que durante los primeros años aún salia de fiesta, pero ahora ya no tanto. Lo ha cambiado por la bicicleta de montaña. Aún me acuerdo cuando le decía que cogiera una bici que se moviera en vez de hacer *spinning* en el gimnasio.

Venimos todos de un laboratorio grande, al que no le había faltado de nada gracias a la excelencia científica de nuestro jefe Ramón Serrano, que aunque sea ya profesor emérito, no le faltan ganas para seguir viniendo al laboratorio e investigar. El laboratorio de Ramón y Mariche siempre se ha caracterizado por ser un laboratorio multitudinario. De él han salido muchos científicos titulares de la universidad, con laboratorios dentro y fuera del IBMCP. Todo el mundo en el campo de la biología molecular del estrés en plantas y levaduras lo conoce, sobretodo por sus trabajos en la Protón-ATPasa. Es el profesor investigador más citado de toda la universidad. Es más, él fue fundador del IBMCP y de la carrera de biotecnología de la UPV que yo cursé. Pero además de la excelencia científica, es una excelente persona y el mejor jefe de laboratorio. Junto con Mariche, cuida a cada uno de los miembros del laboratorio, preocupándose por sus investigaciones y situación laboral. Y no solo en el laboratorio sino también fuera de éste. Todos los años han invitado a todos los del laboratorio a diversas celebraciones tanto en sus casas de Valencia, Serra o Jávea como en otros lugares. Incluso cuando les hemos invitado a comer nosotros ellos siempre han contribuido trayendo traído algo, y siempre han sido muy agradecidos.

Agradezco también enormemente a las técnico de laboratorio Irene Martínez y Lola Planes, que con su labor han hecho que todo funcione y han puesto orden en el desorden del laboratorio. Irene, además de ser la siguiente más joven del labo (sin contar alumnos) y no hacerme sentir un yogurín, ha tenido una pasión reivindicativa común a la mía: las casas okupas. Gran activista de Benimaclet, Irene siempre ha apoyado, como el resto del laboratorio, al proyecto godellense de la Figatendra, y forma parte del CSOA L'Horta, la casa okupa de carácter social más importante de Valencia. Además, aunque fue tan solo por unos días en los que no estaban los jefes, participó en la declaración del laboratorio como espacio anárquico, aunque bueno, siempre lo ha sido un poco.

Agradezco también a todos los demás miembros del laboratorio que ahora ya no están pero que han coincidido conmigo, como son Marcos Caballero, el último doctorando anterior a mí; Alessandro Rienzo, otro italiano con una gran pasión por la música y la fotografía; y Jesús Muñoz, que aunque cuando llegué al laboratorio ya se había trasladado para empezar como profesor en la Universidad de Valencia, había realizado distintas construcciones utilizadas en esta tesis.

No olvido a todos los alumnos que han pasado por el labo. Los míos, o como a mi me gustaba llamarlos, los esclavos, me hicieron los días más alegres y me hicieron despertar mis dotes de profesorillo y de dibujante de esquemas, e incluso me ayudaron a madrugar un poco más. Bea estuvo más de un año aguantándome para hacer el TFM. Junto con Sira, alumna de Gaetano, el laboratorio se llenó de vida y alegría. Julia y Moreno, más conocidos como Juleno, hicieron el TFG conmigo. En el recuerdo quedará la frase que dijo Moreno mientras sembraba *in vitro* semillas de Arabidopsis:

"Que hago aquí si podría estar cantando trap". Pero sin duda me quedo con la alegría de Julia que dijo es su TFG que había aprendido y disfrutado más de la biotecnología en el laboratorio que en cualquier curso de la carrera. Otros alumnos que recuerdo con cariño son Ainhoa, la primera alumna que tuve de custodia compartida con Regina, Alejandra y Pitu, que vinieron del curso de las excursiones de los congresillos de fertinagro, Vlada, la fan nº1 del labo y Alba, la última alumna compartida con Irene. Otros alumnos del labo que dejaron huella fueron Salva, Delo y Alex el Boliviano, que dejó símbolos y frases religiosas escondidas por el labo que a día de hoy aún aparecen.

Agradezco también a todas las personas que coincidido conmigo en el IBMCP y que han hecho la estancia más amable, tanto ayudando en el laboratorio o saliendo un rato, como simplemente saludando con una sonrisa. Puedo destacar a los compañeros del 1.09, ahora rebautizados como 1.03, Regina y Pepe mayormente, y a los de otros laboratorios, Alba, Lynne, Gustavo, Jesús Praena, Jorge Lozano, Jorge Hernández, Pepe Moya entre otros. Agradezco también al personal de servicios del IBMCP, como Javier Forment y Marisol y de invernaderos.

A mis familiares y amigos

Mis padres me han ayudado durante toda la vida. Me lo han dado todo, aunque les he salido un poco arreu, un poco como son ellos, pero algo más acentuado. Pau y Adela, tal y como les he llamado toda la vida, siempre han estado ahí y apoyado en todo, aunque en algún momento no confiaran mucho en mí. Mi hermana, Júlia ha sido todo un referente en lo que es ser un buen estudioso. Y aunque no estemos continuamente en contacto siempre estaremos el uno para ayudar al otro. Gracias también a mis primos, tíos y demás familiares por estar ahí y apoyarme.

Gran parte de esta tesis se escribió en la casa familiar, el Huerto, donde África y yo pasamos el cofitamiento. Mi compañera África ha estado a mi lado durante estos últimos casi tres años, y me ha aguantado durante toda la escritura de la tesis, que ha sido larga y tediosa. Con mucho cariño hemos pasado esta última etapa. Recientemente, poco después que volvimos al poble, Senda nos ha acompañado y alegrado cada uno de los días que ha estado a nuestro lado.

A mis amigos, a los de toda la vida que nos vemos desde el instituto e incluso desde el colegio y guardería, y a los mas recientes, que han ido sumándose a lo largo del tiempo, sobretodo desde la Figatendra, ese proyecto okupa que nos unió a los ya no tan jóvenes del pueblo y que nos llenó de es bonita ilusión de hacer las cosas entre todos y para todos. Todos ellos han conseguido distraerme y alegrarme un poco cada día.

A las revistas que permiten el uso de los artículos científicos en versión de autor en esta tesis

Agradezco a las revistas *Plant, Cell & Environment* y *New Phytologist* que permitan la utilización de los artículos aceptados en formato de versión de autor, permitiendo así su utilización en esta tesis doctoral. Ahora lo escribo en inglés para que lo puedan leer ellos.

I would like to thank to the Plant, Cell & Environment journal, the New Phytologist journal, the New Phytologist Foundation and Wiley for permitting the inclusion in this thesis of the author version of the following research articles:

Renard J., Martínez-Almonacid I., Sonntag A., Molina I., Moya-Cuevas J., Bissoli G., Muñoz-Bertomeu, J., Faus, I., Niñoles, R., Shigeto, J., Tsutsumi, Y., Gadea, J., Serrano, R., & Bueso E. (2020*a*) PRX2 and PRX25, peroxidases regulated by COG1, are involved in seed longevity in Arabidopsis. *Plant, Cell & Environment* **43**, 315–326. https://doi.org/10.1111/pce.13656

Renard J., Niñoles R., Martínez-Almonacid I., Gayubas B., Mateos R., Bissoli G., Bueso, E., Serrano, R., & Gadea J. (2020*b*) Identification of novel seed longevity genes related to oxidative stress and seed coat by genome-wide association studies and reverse genetics. *Plant, Cell & Environment* **43**, 2523–2539. https://doi.org/10.1111/pce.13822

Renard J., Niñoles R., Martínez-Almonacid I., Queralta-Castillo I., Sonntag A., Hashim A., Bissoli G., Campos, L., Muñoz-Bertomeu, J., Niñoles, R., Roach, T., Sánchez-León, S., Ozuna, C. V., Gadea, J., Lisón, P., Kranner, I., Barro, F., Serrano, R., Molina, I., & Bueso E. (2021) Apoplastic lipid barriers regulated by conserved homeobox transcription factors extend seed longevity in multiple plant species. *New Phytologist* **231**, 679–694 . https://doi.org/10.1111/nph.17399

A todo el que promulga la ciencia, desarrolla herramientas para ella y permite su acceso abierto

Agradezco a todas las personas que promulgan la ciencia, participando en ella o creando herramientas para el avance y la libre difusión de la misma, y a los medios institucionales que permiten la investigación pública en universidades.

Diversos han sido los grupos que nos han cedido lineas mutantes para la investigación, o bien han colaborado con nosotros para la generación de plantas transgénicas de interés agronómico, o en la realización de diversos experimentos. Por ejemplo, los profesores Julian I. Schroeder y June M. Kwak (San Diego, EEUU) nos cedieron el doble mutante *rbohd,f*; el profesor Graham Noctor (Orsay, Francia) nos cedió las lineas mutantes *dhar*; y la doctora Cristina Ferrándiz y el doctor Vicente Balanzà (Valencia, España) nos cedieron las líneas *sep3* y *sep1,2,4*, y los mutantes de *APETALLA2*, *ap2-5* y *ap2-12.* Gracias a este intercambio de material genético vegetal hemos podido avan-

zar y obtener resultados interesantes incluidos en esta tesis doctoral. Recientemente yo también he empezado a enviar semillas a diversos grupos de investigación, a raíz del estudio GWAS de momento. Espero que las semillas que mande sean útiles a quien las reciba. También, aunque no sea gratis, cabe reconocer el papel esencial de los bancos de semillas que mantienen las miles de líneas mutantes y ecotipos de Arabidopsis, permitiendo así la rápida obtención de plantas mutantes en distintos genes, agilizando y permitiendo el trabajo de la biología molecular de plantas. Destaco al NASC ya que principalmente nos ha provisto de multitud de plantas mutantes. Finalmente cabe remarcar, que el intercambio de información entre laboratorios, protocolos y otros productos como plásmidos, también son esenciales en nuestro trabajo de investigación.

Agradezco también a todos los que han creado herramientas esenciales para la investigación, desde equipos a programas, que han permitido el gran avance de la ciencia. Cabe destacar el papel de aquellos que no buscan lucrarse y crean herramientas gratuitas y accesibles a todos, el llamado software libre. También, aunque esté mal decirlo, agradezco a aquellos que rompen las barreras en la ciencia y permiten el acceso a las herramientas informáticas y publicaciones científicas a todo aquel que tenga conexión a Internet. Cabe mencionar a la creadora de Sci-hub, Alexandra Elbakyan, que ha conseguido mejorar este aspecto con la liberación de artículos científicos. Aunque gracias a la universidad y los convenios que tiene con multitud de revistas se puede acceder a muchas de las publicaciones que necesitas, no todas son accesibles desde la UPV. Por ello, poder simplemente abrir un artículo científico para poder leerlo u ojearlo sin pagar los costes abusivos e irrealistas que supuestamente que tendrías que pagar, ha sido crucial para el desarrollo de mi tesis doctoral, como seguramente también de muchas otras tesis y trabajos científicos. Iniciativas como esta permitirán el avance de la ciencia abierta, que es como debería ser la ciencia. Es curioso el negocio de las revistas. El investigador paga su investigación con los fondos que disponga, hace la publicación, y tiene para publicar, pero también para leer. La revista siempre gana, parece la banca. Poco a poco aumentan las publicaciones en abierto (aunque suelen costar más a la hora de publicar), y es un primer paso para compartir entre todos ese conocimiento que tendría que ser de todos. Sin embargo el modelo de publicaciones debería cambiar. Mientras tanto, iniciativas como Sci-hub nos ayudan a luchar contra este gran negocio.

Por último, agradezco a las instituciones públicas la inversión en investigación. Gracias a ella, los laboratorios, el personal y los distintos equipos se han podido financiar. Gaetano fue quien me ayudó enormemente conseguir la beca FPI del por entonces Ministerio de Economía y Competitividad BES-2015-072096, asociada al proyecto de investigación nacional BIO2014-52621-R-AR que consiguió Ramón y el laboratorio con el tema de investigación que inició y consolidó mi director de tesis Eduardo Bueso, en relación a la longevidad de las semillas. Quiero agradecer a las instituciones públicas que permiten la investigación por parte de jóvenes y no tan jóvenes investiga-

dores. Me han permitido trabajar en un trabajo que me ha encantado. Sin el apoyo de las instituciones no sería posible la investigación pública, que es aquella que no solo busca el beneficio final sino que también mira más allá, indagando en las fronteras del conocimiento que algún días nos abrirán nuevas puertas llenas de posibilidades.

No obstante hay que remarcar que aún hay mucho trabajo para conseguir que la investigación sea una salida profesional estable. Aquellos políticos que tanto remarcan que la investigación es necesaria, y están orgullosos de la alta cuantía de los fondos destinados, deberían entender mejor el sistema español de los proyectos de investigación, ya que por lo menos desde que yo empezaré la tesis se repite un patrón: siempre hay un periodo de más de medio año entre proyectos en el cual el personal de investigación no puede ser contratado, ya que el proyecto anterior se acaba y el siguiente no ha empezado todavía. Me gustaría poder pensar que no se trata de una estrategia para recuperar los fondos de investigación invertidos y el trabajo del investigador. Así pasó con Gaetano y con Irene Martínez Almonacid hasta que nos concedieron el siguiente proyecto nacional para seguir estudiando los mecanismos moleculares determinantes de longevidad de semillas, el BIO2017-88898-P. Y ahora parece que el laboratorio se queda de nuevo sin proyecto nacional, haciendo mucho más difícil la continuidad de la investigación. Esperemos que nos sea así siempre y que se estabilice la financiación y la carrera laboral de los investigadores.

y a todo el que se anime a leer esta tesis

Aunque sea largo y tedioso el contenido de esta tesis, he invertido un gran esfuerzo en la redacción, el diseño de figuras y en el maquetado. Espero que se pueda disfrutar. Gracias por el interés y espero que sirva como mi pequeña primera contribución a la ciencia. Muchas gracias.