Turnitin Informe de Originalidad

Procesado el: 11-sept.-2023 12:41 a. m. CEST
Identificador: 2162330794
Número de palabras: 14262
Entregado: 1

20949655:BORRADOR.pdf Por Sara Pérez Martí

Índice de similitud	Similitud según fuente	
3%	Internet Sources:	3%
	Publicaciones:	1%
	Trabajos del estudiante:	N/A

1% match ()
Renard Meseguer, Joan. "Identification of genes related to seed longevity in Arabidopsis thaliana using genomic molecular techniques", 'Universitat Politecnica de Valencia', 2021

1% match (Internet desde 24-sept.-2022)
https://riunet.upv.es/bitstream/handle/10251/147821/Soriano%20-%20Estudio%20del%20editado%20de%20genoma%20de%20precisi%c3%b3n%20en%20Nicotiana%20benthamiana%20mediante%20el%isAllowed=y&sequence=1

1% match (Internet desde 02-jul.-2019)
http://www.rj-robbins.com/recommended/literature/crispr/inc.files/bt-flat.txt

1% match (Methods in Molecular Biology, 2014.)
Methods in Molecular Biology, 2014.

1. INTRODUCTION 1.1. Transcription Factors Gene expression and its regulation is the main mechanism that allows cells with the same genetic information to adapt to different environments or conditions: depending on their cellular cycle, to respond to stimuli, or even to be part of a multicellular organism with differentiated tissues. Basically, without proper, extremely fine-grained control of gene expression, life would not be a phenomenon that happens in the way we understand it. In this process, the proteins known as transcription factors (TFs) play a crucial role as they bind to the regulatory regions of the DNA-encoded genes and either allow or impede the RNA polymerase to start the first step of the Molecular Biology Central Dogma: Transcription. (Latchman, 1997) Thereby, TFs are characterized by presenting a DNA-binding domain (DBD) and can be categorized by the type of DBD. The number of acknowledged families has been increasing over time, as molecular high-throughput technologies such as whole genome sequencing became available. For the model plant Arabidopsis (Arabidopsis thaliana), nowadays, 58 families have been listed, including a total of 2296 TFs (Tian et al., 2019) Their pivotal function in regulation and ubiquity explain why they are a primary focus for research centering around complex mechanisms such as stress responses or developmental processes. 1.2. HOMEOBOX25: Role in seed longevity Seed longevity, or seeds' ability to remain viable after being stored for a period of time, is one of the characteristics that determine the successful reproduction of many plant species. Understanding the factors and aspects that play a role in this trait has been one of the longest-standing challenges of plant biology, being crucial to ensure the maintenance of agricultural renewal of crops and to allow the functioning of seed banks. (Zinsmeister et al., 2020). While both these activities have always been a necessity, climate change has posed a new threat that endangers the extinction of an estimated 10.4% of all flowering plants (Nic Lughadha et al., 2020), which makes preservation efforts of utmost importance and raises the necessity to avoid the loss of plant genetic resources and biodiversity. Since seeds are the plant material most thoroughly used for this endeavor, the study of the biological and molecular basis of seeds' viability could potentially help extend their long-term storage by revealing the best features and conditions that influence this polygenic, multifactorial trait. (Arif et al., 2022) In this spirit, a plethora of different genetic analyses have been performed in Arabidopsis since the 2000s, mainly through reverse genetics, but also with quantitative trait loci (QTL) mapping and the use of mutants. First, in the year 2000, four QTLs were mapped for seed storability in a Recombinant Inbred Line (RIL) when trying to find a correlation between these regions and those related to seed oligosaccharide content (Bentsink et al., 2000). Nevertheless, it was not until later that specific genes were found to be correlated with seed longevity. Many seed-specific transcription factors involved in seed development and maturation processes have been directly correlated with seed longevity, like abi3 (abscisic acid insensitive3), fus3 (fusca3), and lec2 (leafy cotyledon2). However, the complexity of their interaction has made it challenging to unravel the specificities of their role in longevity. (Sugliani et al., 2009) Furthermore, several enhancing mutations have also been revealed to impact this trait, such as grs (green seed), a modifier of abi3-1 (Clerkx et al., 2003) or abscisic acid deficient1 (aba1), which was found to cause a reduced rate of germination after storage; as so did seed-coat defective mutants, aberrant test shape (ats) (Clerkx et al., 2004), supporting previous evidence of the importance of the integrity of the testa for maintaining viability (Debeaujon et al., 2000). One of the factors deemed as most influencing in the process of seed aging is the accumulation of free radicals and Reactive Oxygen Species (ROS) since they cause damage to lipid barriers, proteins, and nucleic acids. Therefore, particular interest was put into studying mutants with impaired oxidative stress response, like the defective NADH dehydrogenase mutant frostbite1 with reduced longevity in seeds subdued to a natural aging process (Clerkx et al., 2004). Additionally, several studies corroborated oxidative stress implications in longevity through the study of antioxidant molecules, namely tocopherol (vitamin E), which prevents lipid peroxidation. The tocopherol-deficient mutant vitamin deficient 2 showed reduced germination percentage (Sattler et al., 2004), although their relation seems to be rather complex, as suggested by the low δ-tocopherol levels detected in the extended-longevity Arabidopsis thaliana T-DNA insertional line Atfahd1a-1. (Gerna et al., 2021). Low seed coat permeability has been identified as one of the causes behind reduced storability by many studies, one pointing to the altered polyphenolic content in double (prx2 prx25) and triple (prx2 prx25 prx71) Arabidopsis mutants (Renard et al., 2020a). Additionally, recent reverse genetics and genome-wide association studies showed seven novel genes positive for seed longevity (psad1, sslea, sstpr, dhar1, cyp86a8, myb47, and spch) and five negative ones (rbohd, rbohe, rbohf, knat7, and sep3) (Renard et al., 2020b). The gene Arabidopsis thaliana Homeobox25/ Zinc Finger Homeodomain2/ Zinc Finger Homeodomain1 (ATHB25/ZFHD2/ ZHD1; At5g65410) was first identified as a determinant of seed longevity through an "activation-tagging" screening of a mutant collection of Arabidopsis. (Bueso et al., 2014) This transcription factor, containing the 60 amino acid-DNA-binding motif Homeodomain (HD), hence the name 'Homeobox25' is included in the Zinc Finger-HD (ZF-HD) family, a group characterized by their distinct biochemical activities, such as floral tissue expression, redundancy, and promiscuity in interactions, being able to dimerize with practically every other ZF-HD transcription factor. It was first named ATHB25 (Tan & Irish, 2006) but also mentioned in literature as ZFHD2 (Tran et al., 2006) and ZHD1 (Hu et al., 2008); in this project; however, the name employed will be HB25. The activation-tagging mutant for HB25 showed an increased percentage of germination after subjecting the seeds to natural aging, as well as after a process of artificial aging of imbibed seeds. However, the loss-of-function mutant for the gene did not display a significantly diminished longevity, probably due to the redundancy of function exhibited by other genes in the same family, namely HB22. The action mechanism behind its involvement in the studied trait appears to lie in the Gibberellin biosynthesis pathway, as it has been shown to regulate the enzyme encoded by the gene Gibberellic Acid 3-Oxidase 2 (GA3OX2) and GA3OX3 , thus adding this phytohormone to the list of seed longevity regulators (Bueso et al., 2014). Continuing the research, a chromatin immunoprecipitation analysis identified more putative genes regulated by this TF.

Among the 146 binding sites detected, 16 genes stood out because of their role in post-embryonic seed development, especially one related to lipid polyester deposits in the seed cover: LACS2, and one encoding a glycerol kinase: NHO1. Positively, HB25 Arabidopsis mutants showed modified apoplastic lipid barriers, altered cuticle thickness, and seed permeability, corroborating its implication in the cutin or suberin biosynthesis process. Furthermore, HB25 was successfully transferred to tomato (Solanum lycopersicum) and wheat (Triticum aestivum) with similar results to those in Arabidopsis, positioning HB25 as a trans-species regulator of seed longevity. (Renard et al., 2021) 1.3. Drought-responsive element binding Abiotic stresses are thought to be the causes that represent a large percentage of average yield decline in many of the world's staple crops. Moreover, response mechanisms to these types of stresses appear to be highly convoluted, with numerous elements taking part in the signaling pathways and many overlapping functions. Those two features justify the extensive research addressed in that direction, which has led to the identification and characterization of plenty of genes, particularly in Arabidopsis; around 1500 stress-related TFs have been reported (Lata & Prasad, 2011). While most stress-related genes correspond to an Abscisic Acid (ABA)-mediated pathway, there are several genes involved in an ABA-independent response to numerous stresses. Subsequently, the extensive research conducted on them permitted the reveal of standard features, among which is found the Dehydration Response Element (DRE), a particular sequence with a 9 bp conserved core (5'-TACCGACAT-3') that lays in the promoter region of those genes related to drought, cold and salt stress responses, first identified in the gene rd29A (Yamaguchi-Shinozaki & Shinozaki, 1993). Another more generic motif, related to DRE and found in many low temperature-inducible promoters, was named C-repeat responsive element (CRT) with its conserved core comprised of the following nucleotides: 5'-CCGAC-3', also contained in the DRE motif. Those TFs capable of binding to the DRE or DRE/CRT motifs were designated as DRE binding factors or DREB. (Yamaguchi-Shinozaki & Shinozaki, 1994) These DREB proteins were classified in the significant, plant-specific APETALA2/ethylene responsive factor (AP2/ERF) family of Arabidopsis TFs, joining the other 4 subfamilies: ERF, AP2, RAV, and Soloists, which all contain at least one copy of the AP2/ERF domain in their sequence. Other conserved sequences characterize all DREBs, namely the 14th and 19th residues (valine and glutamic acid) present at the DBD, deemed determining to allow the DRE-specific binding. Also, acting as a nuclear localization signal (NLS), enabling the proteins to travel to the nuclei where they develop their function, there is a shared alkaline N-terminal amino acid region, contrasted to the acidic C-terminal region, thought to be responsible for their trans-activation activity (Lata & Prasad, 2011). Initially, two DREB genes were isolated via a yeast one-hybrid screening, both very distinct in their role: DREB1A was found to control the transduction pathway for signaling low-temperature stress, albeit DREB2A functioned in response to dehydration and high-salt concentration conditions (Liu et al., 1998). Additionally, both genes did not display significant sequence identity apart from the conserved AP2/ERF domain. Since that first identification, many other DREBs have followed, both in Arabidopsis and in a growing number of other plant species, such as rice (Oryza sativa), wheat (Triticum aestivum), barley (Hordeum vulgare), maize (Zea mays), or soy (Glycine max) (Lata & Prasad, 2011). The DREB factors were first further classified into six subgroups (A1-A6) owing to particularities in their sequences and differences in the roles they portray in plants, according to Sakuma et al. (2002). However, then a new classification system was proposed taking into account different phylogenetic information and a more in-depth sequence study, which concluded in the formation of 12 different groups, from I to X, VI-L, and Xb-L for genes in the Arabidopsis ERF family (Nakano et al., 2006). 1.3.1. DREB2 All DREB2 genes were classified into the IV group of the ERF family, further separated into subgroups IVa and IVb. While a motif is conserved among all members of the family, referred to as CMIV-1, only IVa members share high homology in the N-terminal region, which includes, apart from the CMIV-1 motif, another conserved sequence, CMIV-2. (Nakano et al., 2006) Moreover, the gene DREB2C seems to be the only one in the family presenting an intron in its sequence. (Figure 1) Figure 1. Phylogenetic relationships among group IV of Arabidopsis Ethylene Responsive Factor genes. Colored boxes represent the conserved domains, and the inverted triangle represents the presence of an intron. Bootstrap values were calculated from 100 replicates, representing only values > 50. (Nakano et al., 2006) As previously stated, DREB2 genes appear to be involved in the response to osmotic stresses. However, initial studies involving their overexpression and subsequent measure of a downstream gene showcased a weak expression of the latest, implying that a post-translational modification, or perhaps a protein-protein interaction, may be a necessary step for achieving functionality (Nakashima et al., 2000). This hypothesis was later endorsed after obtaining a drought-tolerant phenotype in transgenic Arabidopsis plants containing the modified version of the DREB2A gene (DREB2A-CA), which was constitutively active due to a 29-amino acid deletion in its negative regulatory domain. (Sakuma et al., 2006) On the one hand, genes DREB2A and DREB2B are induced by high-salt conditions in roots and dehydration stresses in roots and stems, but their expression levels do not change under low temperatures (Agarwal et al., 2006). However, they seem to be induced by high temperatures, as they tend to accumulate in the nuclei of heat-shocked cells and interact with HsfA3, an Arabidopsis well-known heat shock factor (Schramm et al., 2007; Yoshida et al., 2008). On the other hand, DREB2F and DREB2D, in contrast to their previously mentioned homologs, are only mildly induced by high-salinity treatments. While the mechanisms behind their specific induction are poorly understood, more and more advances have been made, one example being the reveal of the enzymes responsible for their repression under basal conditions: a complex formed by the coupled phosphoinositide-dependent phospholipase C (PI-PLC) and diacylglycerol kinase (DAG-kinase) (Djafi et al., 2013). Interestingly, a sunflower (Helianthus annuus L.) DREB2 homolog, HaDREB2, has been related to another heat-stress tolerance trait: seed longevity. The pathway involved requires the participation of the seed-specific HaHSFA9 heat shock factor to properly enhance longevity when using overexpression assays (Almoguera et al., 2009). 1.3.1.1. DREB2C DREB2C (At2g40340) is a seed-specific DREB2 TF involved in the late stages of heat stress response mechanisms, being able to enhance thermotolerance when overexpressed by up-regulating related downstream genes containing DRE/CRT motif (Lim et al., 2007). The preliminary approach used to identify putative targets of this TF was proteomic analyses performed in DREB2C overexpressing plants undergoing mild heat stress treatments. Even though four proteins were detected, two targets pertaining to the cyclophilin family were especially interesting, as they are known to protect plants from heat stress. (Lee et al., 2009) Although it was initially thought that abiotic stress responses that involved the participation of DREB TFs were ABA-independent, DREB2C was proven to interact with a protein capable of regulating ABA response, Abscisic Acid (ABA)-Binding Factor 2 (ABF2), thus suggesting the existence of cross-talk between pathways. Indeed, it was found that the overexpression of DREB2C under cauliflower mosaic virus' 35S promoter led to ABA sensitivity during the postgerminational growth stage. In the same study, the gene's expression pattern was deciphered, observing the most significant values in mature embryos of dried seeds and the cotyledons of germinating seedlings. Furthermore, the subcellular localization of the TF was proven to be in the nuclei. (Lee et al., 2010) Figure 2. DREB2C is a seed-specific transcription factor. Schematic view of the expression levels of DREB2C (At2g40340) detected by extensive RNA-seq map powered by BAR webservices: The color represents the relative expression levels according to the legend (left) (Fucile et al., 2011) One of the genes subject to transcriptional activation by this TF is a plant protein inhibitor of cysteine proteases, Arabidopsis thaliana phytocystatin 4 (AtCYS4), which had been linked to multiple abiotic stresses' tolerance, but principally to thermotolerance, hence revealing a new member of the cascade that culminates in this trait. (Je et al., 2013) Another link in the pathway was uncovered when Song et al. (2018) identified Vascular Plant One-Zinc-Finger 1 (VOZ1), a repressor of DREB2C under normal conditions that, under heat stress conditions, becomes ubiquitinated and subsequently degraded, freeing the AP2 DBD previously blocked by its binding. By the same token, other salt-responsive genes were proven interactors by electrophoretic mobility-shift assays, among which lay Cold-regulated 15A (COR15A) and Responsive to Dehydration29A (RD29A), both presenting up-regulated expression in salt-tolerant DREB2C overexpression mutants. Furthermore, transgenic canola plants constitutively expressing Arabidopsis' DREB2C also presented enhanced salt tolerance. (Song et al., 2014) Despite being repetitively associated with abiotic stress control, there are some pieces of evidence that suggest DREB2C has an essential role in the developmental process of flowering time, implying the need for further study of this complex, multi-functional TF. (Song et al., 2021) 1.3.1.2. DREB2H DREB2H (At2g40350), judging by their sequence identity (Ruelland et al., 2009), seems to be a paralog to DREB2C since they share most of their coding sequence (CDS), being this the part of their nucleotide sequence that will be transcribed into protein. Although being identified at the same time as other members of the family, there is significantly less research available regarding its expression pattern, putative targets, and regulators. Despite not having any clear association with any stress-inducible genes yet, owing to the lack of essays where it was included; according to

Djafi et al. (2013), DREB2H was clearly induced when inhibiting diacylglycerol kinases (DGKs) or Phosphoinositide-dependent phospholipases C (PI-PLCs), indicating that the action mechanism must be similar to those described in other members of its family. 1.4. Methods for obtaining mutants in Arabidopsis. Genetic mutants, whichever their origin, have been the cornerstone of advancements in fundamental molecular genetics research. Studying the effects of mutations is one of the tools that helps unveil the underlying mechanisms that regulate every biological process via both forward and reverse genetic approaches, as evidenced by the comments in previous sections of this introduction, where gain-of-function or loss-of-function mutants were needed to confirm suspected interaction between targets. To obtain specific mutants of different organisms, their genomic DNA sequences must be modified, and for that task, there are several fine-tuned methodologies for each organism. Arabidopsis is the most used plant model organism nowadays and, despite having a low spontaneous single-nucleotide mutation rate (6.95 × 10-9) (Weng et al., 2018), has demonstrated its facility to respond to mutagenesis procedures. Initially, the most used technique was radiation, first with ionizing rays, followed by X-ray applications. However, their effect on the Arabidopsis genome was controversial among scientists since the mutations provoked were deemed indistinguishable from spontaneous ones. (Koncz & Rbdei, 1994) 1.4.1. EMS mutagenesis In 1962, the later so-called 'supermutagen' alkalating agent Ethylmethane sulfonate (EMS) was first applied to Arabidopsis (Koncz & Rbdei, 1994). This organic compound with the chemical formula C3H8SO3 provokes only punctual random mutations, thus limiting its toxicity and making it the perfect choice for the mutagenesis needed for techniques such as 'targeting induced local lesions in genomes' (TILLING). The mutation mechanism relies upon the reaction of the ethyl group of EMS with the purine nucleobase guanine (G) to form O6-ethylguanine, an abnormal base not recognized by the DNA polymerase during replication. This can result in a transition mutation from a Guanine: Cytosine pair to an Adenine: Thymine pair. (Pegg, 2000) In Arabidopsis, these transitions lead to stop codons and missense mutations approximately 5% and 65% of the time, respectively. Therefore, EMS mutagenesis can be helpful not only for acquiring loss-of-function mutants, but also for providing information about the role and function of specific amino acids in a sequence. Nonetheless, given its randomness in the mutagenesis, the use of this method requires the screening of plenty of lines in order to achieve saturation and isolate the desired mutant. (Kim et al., 2006) 1.4.2. T-DNA insertional mutagenesis Another method commonly used to mutagenize plants is the insertional approach, which employs the natural mobile element of a plant bacterial pathogen, T-DNA from Agrobacterium tumefaciens, or rather, an unharmful, modified version of this element, to stably integrate a nucleotide sequence in a plant genome. These insertions also happen randomly, which can lead to the disruption of a CDS or the disabling of regulatory regions, among other consequences. When the insertion happens in a gene, it becomes 'tagged' with the specific sequence inserted, which could be dominant markers used for genetic mapping, reporter or enhancer genes for expression studies, or even molecular probes used for gene isolation. The number of possibilities offered by this technique makes it suitable for many research necessities, even though it entails the screening of an even larger pool of lines to find the desired mutant. That is the reason behind the proliferation of large-scale libraries of T-DNA Arabidopsis mutants. (Koncz et al., 1992) For the Arabidopsis Columbia (Col-0) accession, there are four main T-DNA insertional mutant collections: SALK, GABI-KAT, SAIL, and WISC. All the obtained mutants are characterized by DNA sequencing and stored in seed format in repositories, where they can be ordered and delivered to researchers worldwide. (O'Malley et al., 2015) Still, this method is not entirely foolproof, and, on some occasions, the seeds delivered fail to comply with the specificities of the order. 1.4.3. CRISPR/Cas9 technology The introduction of targeted genome editing tools has revolutionized the molecular biology research field, enabling the acceleration of advancements at an unprecedented rate and the emergence of new direct applications in agriculture and medicine. All precise genome engineering tools, such as Zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and the CRISPR/Cas9 system, have in common the nuclease-induced double-stranded break (DSB) of the DNA in a specific position of the genome. This DSB is known to induce different mutations when misrepaired by one of the two endogenous mechanisms the affected cell can employ: non-homologous end joining (NHEJ) or homology-directed repair (HDR). Whereas NHEJ simply joins the DNA ends without any homology requirements, leading to a disrupted sequence via nucleotide deletion or random nucleotide addition, HDR utilizes a nearby donor sequence with a certain degree of homology to the damaged site as a template to try and faithfully reconstitute the original sequence (Figure 3). This cellular machinery can be manipulated to disrupt or insert sequences in specific genome sites, provided the use of the right genetic engineering tools. (Gaj et al., 2013) Figure 3. Mechanisms of DNA double-strand break repair. Non-homologous end joining leads to the disruption of a sequence, while homology-directed repair can restore the sequence. Clustered, regularly interspaced, short palindromic repeat (CRISPR) sequences are part of many bacteria and archaea's adaptative immune system, as they, together with CRISPR-associated (Cas) proteins, form a system that defends the cell from exogenous nucleic acids, mainly from bacteriophage infections. (Mojica et al., 2005; Pourcel et al., 2005) The first step in the mechanism of action consists of the processing and acquisition of a detected foreign DNA as a new 'spacer' into an array of singular sequences, flanked by short palindromic repeats, the CRISPR locus. This protospacer is later transcribed and developed into CRISPR-RNA (crRNA), which upon hybridization with another similar small molecule known as the transactivating CRISPR RNA (tracrRNA), can function as an RNA guide that directs the Cas nuclease (i.e., Cas9) to cleave with a DSB any complementary target-DNA sequence that has immediately next to it a three-nucleotide protospacer adjacent motif (PAM) (Figure 4). (Sander & Joung, 2014) Figure 4. Close-up schematic view of the sgRNA-directed cut of the genome by the nuclease Cas9. PAM: Protospacer Adjacent Motif. The sgRNA is the whole purple molecule, the darkest part representing the target-specific sequence. Although there are several classes of CRISPR/Cas systems characterized in nature, in which many different proteins take part (Makarova et al., 2015), the type II system from Streptococcus pyogenes with the nuclease Cas9 was first adapted and programmed to use in genetic engineering applications by unifying the crRNA and tracrRNA into a single artificial chimeric molecule that could act as a guide RNA, hence the concept of single-guide RNA (sgRNA) arose (Jinek et al., 2012). After this decisive breakthrough, the use of CRISPR/Cas9 extended through many research fields, proving its functionality in different organisms. Furthermore, several modifications and extensions to the method have been proposed to fulfill a plethora of necessities: a directed Indel mutation with the simple system, a sequence insertion or replacement taking advantage of HDR, directed gene activation or repression with modified nucleases containing specific domains, epigenetic modification via defective nucleases with effector domains, visualization of specific loci by the coupling of a fluorescent protein, or even the rewriting of a gene with a prime editing guide and a reverse transcriptase (RT) - Cas fusion protein. (Ishino et al., 2018) For use in bacteria, there are also other systems proposed based on their endogenous type. Every system has its particularities, like the size of the Cas nuclease (Cas3, Cas12a, and Cas13 variants all present differences), the requirements of the PAM (5'-NGG-3' for SpCas9, but 5'-NAG-3' for others…) or the orientation of their cut (i.e., a 'nickase' is a nuclease that only cuts in one of the strands rather than producing a DSB). (Liu et al., 2020) 1.4.3.1. Dual sgRNA approach Despite the abundance of possibilities for gene editing offered by this technology, the most widespread use is directed towards the obtention of knockout mutants due to the simplicity of methodology and the valuable use of the outcome. (Sander & Joung, 2014) Although the use of one sgRNA can result in the loss of function of a gene by disrupting its sequence, the indel mutation should either produce a stop codon or be a frameshift mutation for it to succeed. Besides, detecting the mutants derived requires the use of sequencing since such small mutations cannot be detected with any other molecular techniques except when they result in an obvious phenotype. For these reasons, a dual sgRNA strategy gained momentum for the creation of loss-of-function mutants: the Cas9 nuclease produces a DSB for each sgRNA included in the editing vector, causing a large and easy-to-detect deletion in the genomic region between both sgRNAs. (Figure 5) A similar dual design has also been followed for the insertion of gene-length sequences or their total replacement, which was successful in plants and animals. (Zhao et al., 2016) However, other outcomes have been reported for this approach, namely large inversions of the sequence that was expected to be deleted when editing citrus fruits. (Salonia et al., 2022) Figure 5. Dual single-guide RNA (sgRNA) approach in CRISPR/Cas9 gene editing technology. A. The horizontal line represents the target gene, and the smaller rectangles symbolize the binding sites of the sgRNAs. 'STOP' refers to the codon with which the coding sequence finalizes. Adapted figure from (Salonia et al., 2022) 1.4.3.2. Egg cell-specific gene editing Another modification of the original strategy that has been reported to be more suitable for the editing of plants, due to the particularities of their development, is the use of egg cell-specific promoters controlling the expression of the Cas9 gene in the editing vector with which the plants are stably transformed. (Wang et al., 2015) In plants, the method for

delivering the CRISPR/Cas9 system is the Agrobacterium tumefaciens-mediated transformation of the plant material, particularly for Arabidopsis, the target for the transformation is commonly the egg cell due to the simplicity of the floral dip methodology. When fertilized, this egg cell develops into embryogenic tissue that will ultimately differentiate and grow into a new generation. This process involves multiple cell divisions, and if the editing vector is expressed in later developmental stages, where the organism is composed of more than one cell, the resulting plant will be a mosaic line with differently edited tissues and even the presence of wild-type (WT) alleles, which makes the obtention of homozygous mutant lines require the use and screening of multiple generations. That is the most common outcome when using a nuclease regulated by a strong constitutive promoter, CaMV 35S, which seems to have weak activity in egg cells and one-cell staged embryos. However, when the CRISPR/Cas9 system activates in the egg cell, complete homozygous Arabidopsis mutants can be obtained in a single generation. Wang et al. (2015) designed and optimized a new system to favor this advantageous phenomenon, thanks to employing an artificial promoter from the fusion of two germline-specific genes, EC1.1 and EC1.2, controlling the expression of the editing vector. This approach was highly successful and resulted in higher mutation rates for multiple target genes in only one generation. It is unsurprising, then, that the same strategy has been followed on plenty of occasions, mainly for Arabidopsis gene editing (Pauwels et al., 2018), but also for other plant species of agricultural significance like soy (Zheng et al., 2020). 2. ANTECEDENTS Due to the significance of the gene HB25 in seed longevity, extensive research has been directed toward finding putative interactors, in hopes of deciphering the complex signaling mechanisms that control this intricate trait. One of these studies, conducted in Bueso's lab, consisted of a Yeast One-Hybrid (Y1H) screening assay with the most conserved region of the HB25 promoter sequence against a Yeast Clone library from Oñate-Sánchez's lab (Oñate-Sánchez, 2018) containing cDNA of every transcription factor in the Arabidopsis genome. This conserved region was identified via bioinformatic studies involving the study of the gene in multiple species and the repetition of specific motives, which make the sequence the most likely to bind the trans-regulatory elements. The Y1H assay showed multiple putative interactors, such as other Homeobox gene: HB52 (At5g53980), a bZIP gene: bZIP52 (At1g06850), a MADS family-pertaining gene: AGL24 (At4g24540); and two genes from the DREB family: DREB2C (At2g40340) and DREB2H (At2g40350). Nonetheless, these putative interactions need to be further corroborated with multiple different assays, and to ensure the reliability of those experiments, multiple controls are necessary. To study gene epistasis between the candidates and HB25, we require the obtention of gain and loss-of-function mutants which will help to elucidate their individual contributions; as well as overexpression mutants. 3. OBJECTIVES The main objectives of this project are: 1. To find the specific CRISPR/Cas9 sites of edition and design the sgRNAs for the edition of the genes DREB2C and DREB2H. . 2. To obtain Arabidopsis loss-of-function mutants for both genes through the CRISPR/Cas9 dual sgRNA approach using floral-dip transformation. 3. To design plasmid constructs to obtain Arabidopsis gain-of-function mutants for the genes DREB2C and DREB2H and to obtain said mutants through using floral-dip transformation, Therefore, this project finds its relation with the United Nations' Sustainable Development Goals (SDGs) through Goal number two: Zero Hunger, since the obtention of mutants for the genes DREB2C and DREB2H could potentially aid in building knowledge about the action mechanism of HB25-related seed longevity enhancement. Specifically, this work would help reach target 2.5, which involves the maintenance of seed diversity via long-term storage. As discussed in the introduction, understanding the molecular processes behind that desirable trait could lead to improvement of the storing techniques, a decrease in diversity losses, and ultimately, food security, since genetic variability in crops is essential for preventing food scarcity in an uncertain future. The same principle applies to another relation, in particular to Goal number fifteen: Life on Land. While the plant species aimed to be protected by this goal are not crop species, their biodiversity is equally vital to Sustainable Development. Therefore, with better conservation of genetic resources, the targets 15.5 and also 15.6, could be reached more easily. 4. MATERIALS & METHODS 4.1. Single guide RNA design Following the dual single guide RNA approach on gene editing, two combinations of sgRNAs were designed for each gene: C1 and C2 for DREB2C, and H1 and H2 for DREB2H. The combinations designed for the same gene share one common upstream sgRNA but differ in the downstream one, which sums up to 6 sgRNAs in total: 3 for DREB2C, and 3 for DREB2H. (Table 1) Table 1. Combinations of the sgRNAs used for the obtention of the knockout mutants. Combination C1 C2 H1 H2 Target gene At2g40340 (DREB2C) At2g40350 (DREB2H) 340drebguia342 350drebguia166 Upstream guide Downstream guide Deletion1 340drebguia848 340drebguia919 350drebguia204 350drebguia240 505 576 38 74 1Length in base pairs of the deletion caused by the cut from the upstream and downstream sgRNAs The 20-nt sequences, adjacent to a protospacer adjacent motif, or PAM, were selected based on specificity and efficiency using the online platform CRISPR-P v 2.0 (http://crispr.hzau.edu.cn/cgi-bin/CRISPR2/CRISPR) and comparing the results of the parameters with the online tool http://crispor.tefor.net/; the criteria for selection being, in order of importance, having less than 15 potential off-targets, a calculated off-score lower than 0.50, and an average calculated efficiency higher than 0.5 (Doench et al., 2014) It is also considered the relative position of the forward guide within the gene, prioritizing those located towards the beginning of the gene. Although it is common for chosen sgRNAs to be found in the first exon, an intron-located guide was selected to avoid potential off-target effects. All the selected sgRNAs have guanosine as the first nucleotide since the promoter used is the commonly used U6, which requires it to initiate transcription. (Ranganathan et al., 2014) In Table 2, a summary of the different sgRNAs and their features can be found. Table 2. Single guide RNAs designed for the CRISPR/Cas9 editing. Name1 340drebguia342 Sequence GCTAAAATAATCGTTAAAGG %GC2 On-score3 30% 0.7146 Off-score max4 0.225 Regions5 Intergenic 340drebguia848 GTAAAGGTGGACCTGAAAAC 45% 0.6450 0.300 CDS (AT2G40350) 340drebguia919 GCTGAGATCCGTGAGCCAGA 60% 0.7417 0.462 Exon 350drebguia166 GTAAAGGTGGACCTGAAAAT 40% 0.7043 0.172 CDS 350drebguia204 GGAGTTAGACAGAGGACATG 50% 0.5394 0.283 exon 350drebguia240 GAGATCCGTGAGCCAGGCCG 70% 0.8432 0.187 CDS 1 The last three digits in the name correspond to the guide's position in reference to the ATG. 2Percentage of guanine or cytosine bases. 3Probability of accurate cleavage in the target sequence, calculated with CRISPR-P (0-1). 4Maximum probability of cleavage in a non-target sequence (0-1). 5Regions of the potential off-targets. 4.2. Vectors The name and resistance of the vectors used throughout the experimental procedure are shown in Table 3, classified by the intended use in the experimental procedure. Briefly, the plasmids pCBC-DT1T2 and pHEE-401 are employed in the construction of the editing vector, pCBC-DT1T2 as a template for a PCR, and pHEE-401 as the backbone for the final vector. The other two plasmids used, pCR8/GW/TOPO and pUBC-GFP-DEST, are used for the construction of the overexpression mutants, the first for its amplification in E. coli, and the last for integration of the overexpression cassette in the Arabidopsis plants transformed with the vector. Their molecular maps are shown in Figure 6. Table 3. Summary of the vectors used in this project. Use Name Resistance in E. coli Resistance in A. thaliana CRISPR/Cas9 pCBC-DT1T2 Chloramphenicol - pHEE401 Kanamycin Hygromycin Cloning pCR8/GW/TOPO Spectinomycin - A. thaliana overexpression pUBC-GFP-DEST Spectinomycin BASTA Figure 6. Schematic molecular maps of the plasmids used in this project. All tags finalizing in 'R' symbolize an antibiotic or herbicide resistance gene: CmR: Chloramphenicol, KanR: Kanamycin, SpecR: Spectinomycin, SmR: Streptomycin, HygR: Hygromycin, BarR: BASTA. att sites are specific for Gateway cloning, LB & RB: Left & Right border for genomic insertion. TOPO: Topoisomerase enzyme. U6-26/29-p/t: regulatory regions. Restriction enzyme sites of interest are marked, and their names are unbolded and in italics. Diagrams were made with the online tool BioRender. 4.3. Microbiological Material Escherichia coli: The chemically competent Escherichia coli DH5α strain was used for the cloning experiments. The cultures were grown at 37ºC overnight in both lysogenic Luria-Bertani (LB) broth for liquid cultures and LB-agar media in Petri dishes for solid cultures, supplemented with different antibiotics, depending on the vector's resistance for selection (Table 3). The concentrations used are kanamycin (50 µg/mL), spectinomycin (50 µg/mL), and chloramphenicol (34 µg/mL). Agrobacterium tumefaciens: The competent cells GV3101::pMP90RK (Koncz & Schell, 1986) were used for the Arabidopsis thaliana floral dip transformation. The cultures were grown in LB broth and LB-agar plates supplemented with the plasmid selection antibiotic (Table 3) and rifampicin (50 µg/mL) at 28ºC for 48-36 hours. All liquid cultures were grown under a constant agitation of 220 rpm. 4.4. Plant material Arabidopsis thaliana L. Heynh ecotype Columbia (Col-0; wild type [WT]) and hb25-hb22 Arabidopsis thaliana double mutant line (Renard et al., 2021) were used to perform the experiments described in this project. The seeds were sterilized by a 10-minute wash with 30% (v/v) bleach and then sown in phytohormone-free Murashige and Skoog media with 0,8% (w/v) agar and grown at 22 °C under long-day conditions (16 h light/8 h dark photoperiod and 100 µE m−2 s−1 light intensity). After germination, the seeds were transferred into soil and put in a growth chamber under the same conditions. While sowing

transformed seeds, the culture media consisted of MS-agar supplemented with 10 µg/µL of phosphinothricin for the selection of overexpression mutants, or 20 µg/µL of hygromycin for the selection of CRISPR/Cas9 editing vector-transformed plants, as well as 0,25 µg/µL of cefotaxime to prevent A. tumefaciens overgrowth. 4.5. Bacterial Transformation All transformations were performed by the heat shock method, combining 5-10 µL of the purified plasmid in an approximate concentration of 200 ng/µL with 50-100 µL aliquots of competent cells, which were cryopreserved in hypertonic media at -80ºC and thawed in ice before use. In order to transform E. coli, the mixture is first kept on ice for 30 minutes before being transferred to a 42ºC thermoblock for a 50-second thermal shock. Following this process, 300 µL of SOC culture broth (Super Optimal broth with Catabolite repression) is introduced to the mixture, and the cells are incubated at 37ºC with agitation. After 1 hour, 100 µL of the preculture were plated in selective LB-agar plates and grown at 37ºC for 16 hours, after which the positive colonies could be picked. For the A. tumefaciens transformation, the cells were incubated with the plasmids for 5 minutes on ice, then transferred into liquid nitrogen, where they stayed for 5 more minutes before a short 5-minute shock at 37ºC. Afterwards, 1 mL of SOC was added, and the cultures were kept for 2-4 hours at 28ºC. Finally, 200 µL of the liquid cultures were plated in LB-agar supplemented with the appropriate antibiotic. 4.6. Design of the Editing vector For the design of the CRISPR/Cas9 editing vector which will be transformed into the Arabidopsis plants, there are two main steps to follow: first, a Polymerase Chain Reaction (PCR) employing the plasmid PCR-DT1T2 as a template. Then, the product will be inserted into the final plasmid, pHEE-401, via the Golden Gate method. 4.6.1. PCR DT1T2 For this PCR, designed to assemble the sgRNAs to the scaffold and regulatory sequences, the components (Table 4) are taken with filtered pipette tips in a 0,6 mL tube and mixed thoroughly on ice before starting the reaction under the conditions: 30 seconds at 98ºC, followed by 30 cycles comprised of the following steps: 10 seconds at 98ºC, 30 seconds at 60ºC and 30 seconds at 72ºC. Then, a final cycle of 5 minutes at 72ºC. The DNA polymerase used is New England Biolabs' Phusion enzyme (2 U/µL), as it has 50 times higher copying fidelity (error rate 4.4·10-7) than Taq DNA polymerase, thanks to its 3'-5' exonuclease activity, and produces products with blunt ends, which makes it the perfect option for cloning. Table 4. List of reagents used in PCR DT1T2 with their corresponding final concentrations. Reagent Final Concentration 3 5x HF Phusion buffer 1X Phusion enzyme 0,016 U/µL dNTPs 0,05 mM each template (pCBC-DT1T2) 200 ng/µL DT1-BsF (20 µM)1 0,25 µM DT1-F0 (1 µM)1 0,0125 µM DT2-R0 (1 µM)1 0,0125 µM DT2-BsR (20 µM)1 0,25 µM Total volume 50 µL 1Primers 2Double distilled water 3Concentration of each reagent in the reaction. 4.6.2. Golden Gate To generate the desired final recombinant CRISPR/Cas9 editing vector, the Golden Gate Assembly described by Wang et al. (2015) was followed, mixing the components in a 0,6 mL tube on ice, those being: 2 microliters of purified PCR fragments, containing approximately 100 ng/µL, 2 µL of pHEE401 plasmid DNA, also at 100 ng/µL, 1,5 µL of 10× T4 DNA Ligase Buffer (New England BioLabs), 6 µL of double-distilled water and 1,5 µL of 10× BSA; along with 1 µL of the enzymes Bsa I and T4 DNA Ligase, also provided by New England BioLabs. The reaction conditions and essential components are carefully optimized to ensure efficient ligation and successful cloning. Initially, the mixture was incubated for 5 hours at 37°C; then, the temperature was raised to 50°C and maintained for 5 minutes, and finally, the reaction was terminated by a 10-minute incubation at 80ºC. 4.7. Design of the Gain-of-function vector The final plasmid that will be used for creating DREB2C and DREB2H Arabidopsis overexpression mutants is pUBC-GFP-DEST. The cloning process is comprised of the following steps: a PCR to amplify the genes, a restriction-ligation cloning into a Gateway Entry Clone, pCR8/GW/TOPO; and a Gateway LR reaction to subclone the genes into the final vector. 4.7.1. PCR for cDNA amplification With the objective of cloning the genes DREB2C and DREB2H (At2g40340 and At2g40350) using the Gateway Entry Clone plasmid pCR8/GW/TOPO, A. thaliana seed cDNA was amplified with specific primers comprised by the first and last 16 bp (Forward and Reverse primers, respectively) of the genes and the recognition sequence of the restriction enzyme EcoR I at the 5' ends. Into a 0,6 mL tube was added 10 µL of 5x GC Buffer (New England BioLabs), 1 µL of each primer (10 µM), 1 µL of dNTPs (10 mM), 1 µL of the template at 100 ng/µL , 0,25 µL of the Phusion Polymerase (2 U/µL) and water to reach a total volume of 50 µL; being the reaction conditions: 3 minutes at 98ºC, then 30 cycles comprised of the following steps: 10 seconds at 98ºC, 15 seconds at 3ºC more than the primers' lowest melting temperature and 30 seconds at 72ºC; and finalized by a cycle of 7 minutes at 72ºC. 4.7.2. Restriction-Ligation Cloning into pCR8/GW/TOPO The preparative DNA digestions were performed by mixing 25 µL of the purified plasmid or PCR product with 5 µL of the enzyme's preferred buffer at 10X (with BSA), 0,5 µL of the EcoR I endonuclease (Promega, 2 U/µL), and water up to 50 µL. The mix was then incubated overnight at 37ºC. The ligation of the PCR products/plasmids was performed with 0,5 µL of T4 ligase enzyme (New England BioTools), 1,5 µL of 10X T4 Ligase Buffer, 1,5 µL of BSA 0,1% (w/v), 0,5 µL of the destination plasmid (previously treated with alkaline phosphatase following New England BioLabs protocol) at 200 ng/µL and 11 µL of the desired insert. The mix was incubated at 16 ºC overnight since the insert, if it's a PCR product, has a single base overhang. 4.7.3. Recombination-based method: Gateway The Gateway® reaction was performed to transfer the gene sequences from the pCR8/GW/TOPO entry clone into the destination vector pUBC/GFP/DEST, following the directions of Invitrogen LR Reaction, mixing in a total volume of 2,5 µL, 1 µL of the Entry clone (50-150 ng), 0,3 µL of the Destination vector (150 ng/µl), 2 µl of LR Clonase II enzyme mix, previously thawed on ice and vortexed briefly, 4.8. PCR assays 4.8.1. Colony PCRs For the analytical PCRs, the reaction volume was 20 µL, containing: 2 µL of Buffer 10X (BioTools), 1 µL of dNTPs (10mM), 0,5 µL of each primer (10 µM), 0,5 µL of the Taq polymerase (1 U/µL) (BioTools), and 15,5 µL of double-distilled water. The quantity of the template is variable as it depends on the number of cells collected while colony picking. The reaction conditions were: 5 minutes at 95ºC, then 30 cycles comprised of 30 seconds at 95ºC, 30 seconds at 3ºC more than the primers' lowest melting temperature, and 1 minute at 72ºC. And then, a final cycle of 5 minutes at 72ºC. 4.8.2. RT-PCR The goal of this step is to obtain complementary DNA derived from the RNA of a sample. Hence, for its synthesis there needs to be an initial RNA extraction from 2-3 fresh leaves of Arabidopsis plants, using the commercial NZYTECH 'Total RNA Isolation kit'. The concentrations in the samples are subsequently normalized by dilution, using a NanoDrop spectrophotometer as an aid. After all the samples are at the same concentration, the synthesis process can follow, using the ThermoFisher Scientific's Commercial Kit 'Maxima First Strand cDNA Synthesis Kit for RT-qPCR' 4.8.3. qPCR This quantitative PCR, whose objective is to both amplify and quantify the amount of a specific type of DNA molecule in a sample, uses complementary DNA (cDNA) as a template, as it is used to measure specific mRNA in plant samples. Using ThermoFisher Scientific's SYBR® Green PCR Master Mix, the samples are loaded into a 96-well plate, and the amplification process can begin using a specialized thermocycler that measures the fluorescence emitted by the SYBRGreen component. 4.9. Restriction Enzyme Analysis To characterize the plasmids before performing an experiment and to evaluate whether the vector constructions were successful, several restriction enzyme analyses were performed by mixing 0,5 µL of the endonuclease/s (2 U/µL), 1,5 µL of 10X Digestion Buffer (composition depending on the enzyme in use), 2 µL of the purified plasmid at 200 ng/µL, 1,5 µL of BSA 0,01% (p/v) and 9,5 µL of double-distilled water. The reactions were incubated for 1 hour at 37ºC. 4.10. Plasmid extraction The plasmid extractions from a cellular culture were performed by the Alkaline Lysis method (Feliciello & Chinali, 1993) for both liquid and solid cultures of either E. coli or A. tumefaciens. The quality of each extraction and plasmid concentration were analyzed with ThermoFisher Scientific's spectrophotometer NanoDrop™. 4.11. Genomic DNA extraction The genomic DNA extraction protocol followed for the identification of mutants was performed as described by Doyle & Doyle (1987), being the material 2-3 fresh small Arabidopsis thaliana leaves and using CTAB-buffer. 4.12. Agarose gel electrophoresis In order to analyze the PCR and digestion results, agarose gel electrophoresis was carried out, as this method makes it possible to visualize the length of DNA fragments in a sample. The 0,7% (w/v) agarose gels also contained 1,5% (v/v) of TBE buffer (45mM Tris base, 45 mM boric acid, and 1 mM EDTA) and 1X GelRed® Stain (Biotium). 1/5 volumes of Loading Buffer 6x (2,5% Ficoll-400, 11mM EDTA, 3.3mM Tris- HCl, 0,017% SDS, and 0,15% Orange G) were added to each sample before loading them into the gel and run it through TBE 0,5X with a voltage between 50 – 100 V, depending on the cast size. The bands' size was inferred using the GeneRuler 1kb DNA Ladder (ThermoFisher Scientific), and the results were observed and captured through a UV light transilluminator. When needed, the bands were extracted manually and purified with the column kit NZYGelpure, following the protocol provided by the company. 4.13. Sequencing The integrity of the plasmids and the correctness of the sequences were studied through the Sequencing and Genetic Expression Analysis service of the "Instituto de Biología Molecular y Celular de Plantas" (IBMCP), which employs the Sanger sequencing methodology. The sequencing service received 10 µL of the samples, prepared at 200 ng/µL, and 5 µL of the primers for use at 5 µM, and delivered reads up to 500 nucleotides. 4.14. Plant transformation via floral dip The saturated liquid cultures of A.

tumefaciens were transferred into 50mL tubes and centrifugated for 20 minutes at 3500 rpm in a 13 cm diameter Nahita® centrifuge. The pellet was then resuspended into a 5% (w/v) sucrose and 0,02% (v/v) Silwet L-77® solution, in which young Arabidopsis plants, grown from the sterilized seeds, were submerged to impregnate the flowers and allow the A. tumefaciens infection. The treated plants were covered with plastic, and a day later, the cover was opened to prevent any humidity-related damage. Approximately 30 days after the transformation, seeds were retrieved, sterilized, and sown into selective MS-agar media plates containing 25 mg/L hygromycin after 15 days of desiccation, in order to avoid dormancy-related issues. 4.15. Primers The primers used throughout this project are presented in the table below: Table 5. List of primers used for the project Primer name Sequence (5'-3') DT1-BsF-340drebguia342 ATATATGGTCTCGATTGCTAAAATAA DT1-F0-340drebguia342 TGCTAAAATAATCGTTAAAGGGTTTTAGAGCTAGAAATAGC DT2-BsR-340drebguia848 ATTATTGGTCCTCGAAACGTTTTCAGGTCCACCTTTACAA DT2-R0-340drebguia848 AACGTTTTCAGGTCCACCTTTACAATCTCTTAGTCGACTCTAC DT2-BsR-340drebguia919 ATTATTGGTCTCGAAACTCTGGCTCACGGATCTCAGCAA DT2-R0-340drebguia919 AACTCTGGCTCACGGATCTCAGCAATCTCTTAGTCGACTCTAC DT1-BsF-350drebguia166 ATATATGGTCTCGATTGTAAAGGTGGACCTGAAAATGTT DT1-F0-350drebguia166 TGTAAAGGTGGACCTGAAAATGTTTTAGAGCTAGAAATAGC DT2-BsR-350drebguia204 ATTATTGGTCTCGAAACCATGTCCTCTGTCTAACTCCAA DT2-R0-350drebguia204 AACCATGTCCTCTGTCTAACTCCAATCTCTTAGTCGACTCTAC DT2-BsR-350drebguia240 ATTATTGGTCTCGAAACCGGCCTGGCTCACGGATCTCAA DT2-R0-350drebguia240 AACCGGCCTGGCTCACGGATCTCAATCTCTTAGTCGACTCTA U6-29p-R AGCCCTCTTCTTTCGATCCATCAAC pU6seqF AGGCATCGAACCTTCAAGAATTTG U6-29p-F TTAATCCAAACTACTGCAGCCTGAC DREB2C-F GAATTCCGGATGCCGTCGGAGATTGTTGACA DREB2C-R CGGTTATGTAGATCCATGAACATCTTTGTGAATTC DREB2H-F GAATTCCGGTCTATGCCCAGGAAACGGAA DREB2H-R AGGCTAGTCTGACCTGACCAAAACGAATTC T7 primer TAATACGACTCACTATAGGG primerGFP CGTCGCCGTCCAGCTCGACCAG DREB2Cdetect-R TTCGAGTCCTGTCCAATTCC DREB2Hdetect-R CATAACTTAGCACCTCGGCCTG qRT-DREB2C-R GGCAGCAGTCGAAGAAGAG qRT-DREB2H-F ATGGGATTTGTGACTATACAG qRT-DREB2H-R CATAACTTAGCACCTCGGCCTG 5. RESULTS AND DISCUSSION 5.1. Loss-of-function mutants 5.1.1. Construction of the editing vectors For creating a gene deletion of DREB2C and DREB2H, two combinations of sgRNAs pairs were designed for each one, as described in the Materials & Methods section. The vector selected for the gene editing is pHEE-401, which enables the assembly of two contiguous sgRNAs in the same plasmid that carries the SpCas9, the nuclease whose expression regulation falls under an egg cell-specific promoter. For the construction of the final editing vectors, the protocol consists of two main reactions: first, a PCR that uses the plasmid pCBC-DT1T2 as a template, followed by a Golden Gate assembly reaction which finalizes the process to form the whole construction. The integrity of the plasmids was checked with a restriction enzyme analysis, using the double cutters Bgl II for pHEE-401 and Sfi I for pCBC-DT1T2, obtaining a band at 690 bp for pHEE-401; and two bands at 3517 bp and 621 for pCBC-DT1T2, as seen in Figure 7, which correspond to the predictions from the in silico digest performed with the virtual tool Benchling. Figure 7. Restriction enzyme analysis of plasmids pHEE-401 and pCBC-DT1T2. A. 0.7% Agarose gel electrophoresis showing the restriction enzyme analysis of plasmids pHEE-401 (1) and pCBC-DT1T2 (2), digested with Bgl II and Sfi I, respectively. The brightness and contrast of the image were modified in order to clarify the presence of the faintest bands. B. In silico digest from Benchling, the numbers in the columns correspond to those in Figure 7A. After verifying the length of the intended template, the cloning protocol could start, first with the PCR-DT1T2. For this polymerase chain reaction assay, designed to assemble the target-specific sequences to the regulatory sequences and one of the sgRNA scaffolds, two pairs of primers are used: 'Bs' and '0'. The difference between them consists of the presence of the recognition sequence for the endonuclease type II S Bsa I in the pair 'Bs', which will be involved in further steps of the building process. Both sets of primers have in common a few nucleotides and the specific sequence of the guide; but only the set '0' has sequences complementary to those on the template, pCBC-DT1T2. Therefore, on the first cycles of the reaction, the only set of primers participating is the '0' pair: DT1-F0 and DT2-R0, and the subsequent product provides a template from which the primes 'BsF' and 'BsR' can start a parallel reaction that forms the final product. (Fig. 8) Figure 8. Diagram of the process of PCR-DT1T2. Two sets of primers (DT1F0 and DT2-R0; DT1-BsF and DT2-BsR) work together to assemble the final product. pCBC-DT1T2: plasmid used as a template. sgRNA-Sc: Scaffold portion of the single guide RNA. Target: target gene-specific portion of the sgRNA. U6-26/29-p/t are all regulatory regions. The reason for the 'Bs' primers to be 20 times more concentrated in the reaction is to ensure the competitive process concludes in the desired 626 bp product and not in the incomplete intermediary. After subjecting the PCR product to electrophoresis on an agarose gel, a distinct band with an observed size of ~650 base pairs (bp) was detected, as it laid just below the 700 bp fragment of the 1kb GeneRuler® DNA ladder. This band corresponded to the expected length of the desired product, indicating successful amplification. Subsequently, the band was purified for its employment in the Golden Gate reaction, the final step of the assembly protocol. After transforming E. coli cells with the concentrated Golden Gate product and incubating them in kanamycin-selective media, the colonies that appeared had to be screened with specific primers to find the ones with the whole 2-sgRNA construct. For the first combination of sgRNAs intended to work for the DREB2C gene, C1, numerous colonies were selected, whereas for the other combinations, C2, H1, and H2, only a few colonies grew up. However, when subjecting them to a colony-PCR, all of them seemed to have the correct construction, with very intense bands appearing at 700 bp, which corresponds to the 705 bp distance between the primers used (Fig. 9A). Figure 9. A. Colony-PCR for detecting the Escherichia coli colonies with the correct assembly of the editing vector. 0.7% Agarose gel electrophoreses with the fragments products of a Polymerase Chain Reaction. The marker used is GeneRuler 1 kb. C1, C2, H1, and H2 are the names of the constructs; subindexes represent different colonies transformed with the same constructions. The empty plasmids were employed as a negative control (pHEE-401 and pCBC-DT1T2) B. Diagram of the final construction after the Golden Gate assembly and the location of the primers. #1: pU6seqF, #2: pU6-29p-F, and #3: pU6-29p-R are the primers used for evaluation of the construction via colony-PCR (#1 & #3) and sequencing (#1 & #2). sgRNA-Sc: Scaffold portion of the single guide RNA. Target: target gene-specific portion of the sgRNA. U6-26/29-p/t are all regulatory regions. In order to further corroborate the correct assembly of the editing vector, the purified plasmids were subjected to sequencing with two different primers: pU6seqF (#1) and pU6-29p-F (#2), whose read would cover both target sequences (Fig 9B). All constructions were proven to be correctly assembled, as the lectures aligned perfectly with both the sequence of the pHEE-401 plasmid and the sequence of the designed sgRNA when subjected to a Multiple Sequence Alignment (MSA) with the online bioinformatic tool ClustalW (Fig 10) Therefore, the purified plasmids were the complete vectors ready to continue the process of editing the plant genomes. Figure 10. ClustalW alignments of the target region in the sequenced constructs. Two single guide RNAs (sgRNAs) were sequenced for each construct (combination). The first line of nucleotides is the target-specific sequence designed for each sgRNA. 'Read': sequencing result. 'pHEE401': sequence of the empty vector. 'N' represents any nucleotide, as that is the intended site for the target-specific sequence. 5.1.2. Obtention of the mutants 5.1.2.1. Transformation and selection The next step to obtain CRISPR/Cas9 edited Arabidopsis plants after completing the cloning of the editing vectors is to prepare the method for delivering this system directly to the plants' egg cells. For that task, the use of a modified strain of the plant pathogen Agrobacterium tumefaciens is the optimal choice, since this bacterium (when delivered through the floral dip method) penetrates the germinal line cells and integrates the contents of its T-DNA in the plant's genome. In this occasion, the sequence between the borders of the T-DNA corresponds to the region of pHEE-401 that codifies the sgRNAs and SpCas9 nuclease, with their own regulatory regions. Therefore, a selective plate of transformed Agrobacterium was cultured for each of the editing vectors constructed. The fastest-growing colonies of each plate were transferred into liquid media for further growth, and then the floral dip solution was prepared for the Arabidopsis transformation. For each combination of sgRNAs, 4 plants were transformed and between 0.20-0.50 g of their seeds were recollected after drying the plants. After sowing the seeds in solid media and letting them grow under long-day conditions for 3 days, the plates were covered with aluminum foil, since the privation of light would increment the differences between wild-type and transformed plants' hypocotyls. That's the most used trait for selecting mutants when using hygromycin as a selection agent. Transgenic seedlings display an elongated hypocotyl compared to the WT plants sowed in the same plate as shown in Fig. 11 Figure 11. pHEE-401 transformed Arabidopsis seedlings present an elongated hypocotyl when using hygromycin selection in dark conditions. Comparison between hygromycin-resistant

seedlings (5) and wild type (WT) seedlings grown within the same plate (1-4). For the combination H2, only one hygromycin-resistant plant was retrieved and transferred to soil for its evaluation. For the other constructions, after screening multiple plaques and repeating the sowing and selecting process 3 times, no seedlings with elongated hypocotyls were detected. This low transformation yield indicates a likely issue within this step of the experimental process since the constructions were adequate as proven by their multiple analyses. Therefore, the transformation should be repeated, possibly changing some parameters like the strain of Agrobacterium tumefaciens used to transform the plants; the developmental stage of the plants used to transform; or the exposure time during the floral dip. 5.1.2.2. Evaluation of the mutants: PCR In order to evaluate whether the transformed plants were successfully edited, a PCR able to detect the deletion allegedly present between the two sgRNAs is needed. As DREB2H is a small gene, the designed sgRNAs fell close to each other, and a change of less than 100 bp is hardly noticeable. Nonetheless, a 4% agarose gel electrophoresis to compare the products of a PCR that amplifies the whole gene would be sufficient to check the differences between the WT and an edited plant. This is not the case for the plant tested for the H2 combination, as both bands exhibit exactly the same length after amplification with the primers DREB2H-F and DREB2H-R. If a deletion had been present, the fragment obtained would have been 400 bp, as opposed to the 474 bp that the WT displays. (Fig. 12) Figure 12. The H2-transformed plant does not present a deletion. 4% Agarose gel electrophoresis comparing the amplification products of a wild type (right) genome extraction with a putatively edited plant, grown from H2 combination-transformed seeds (left). Despite not presenting a large deletion between the two sgRNAs employed, it would still be possible that the genome was edited if only one of the sgRNAs had made a cut in the sequence instead of both guides; thus the gene would contain a small InDel undetectable by this simple polymerase chain reaction. For that reason, the sample was sent to sequence, the results being still pending. Given that this process will be repeated, once multiple transformed plants are obtained, further characterization of the editing process can be done by calculating the editing efficiency of the sgRNAs. Furthermore, if in the identified mutants, only heterozygotes were identified, the screening of a second generation of Arabidopsis plants resulting from the self-fertilisation of the heterozygous edited plants should be done to obtain homozygotes, where both copies of the gene are ineffective and the mutant becomes a true loss-of-function. 5.2. Gain-of-function mutants 5.2.1. Cloning of DREB2C & DREB2H in Entry Clone The initial cloning of the genes DREB2C and DREB2H into a Gateway entry clone, pCR8/GW/TOPO was performed using the traditional restriction-ligation method with an empty plasmid and the purified product of a PCR that used seed cDNA as a template and cloning primers, both digested with the enzyme EcoR I and ligated with T4 ligase enzyme. To avoid the re-ligation of the empty plasmid, the digested backbone was previously treated with alkaline phosphatase, an enzyme that eliminates the phosphates required for ligation from both ends of the digested plasmid. Thus, the most likely ligation reaction is the one that generates the desired construct. The E. coli cells transformed with the product of said reaction were selected with Spectinomycin and screened with a colony PCR employing the forward cloning primer and the reverse primer T7 which binds to the plasmid in the T7 promoter region, next to the Multi-Cloning Site (MCS). As the CDSs of DREB2C and DREB2H are 1026 bp and 474 bp, respectively, and the T7 primer is 142 bases away from the EcoR I restriction site, the size of the bands expected for each reaction are 1168 bp and 616 bp. Following its analysis with an agarose gel electrophoresis, multiple colonies were selected (Fig 13) for the purification of their plasmids for downstream applications. Figure 13. Colony PCR of the cloning on pCR8/GW/TOPO Entry Clone. 0.7% Agarose gel electrophoresis image showcasing the DNA fragments amplified in a Polymerase Chain Reaction corresponding to the genes DREB2C (left) and DREB2H (right). The bands lay close to the marked GeneRuler1kb Ladder's 500 bp and 1000 bp bands. In order to verify the presence of the insert in the plasmid, a digestion analysis with the enzyme EcoR I was performed. As this is the enzyme with which the genes were cloned, clear bands should appear at the length of the corresponding gene: approximately 1000 bp for DREB2C and 500 bp for DREB2H. In Figures 14A and 14C, multiple positive clones can be observed. Furthermore, the purified plasmids were subjected to a restriction enzyme analysis to determine the direction of the insert since the Gateway subcloning reaction requires the insert to be in the forward orientation. For that task, two restriction enzymes were employed for each gene: one that cuts in the plasmid, close to the MCS (Bsp120), and one that cuts in near the 5' end of the insert, being Cla I for DREB2C, and Bgl II for DREB2H. Thereby, if DREB2C is correctly inserted, two bands at 3605 bp and 206 bp should appear, whereas bands at 2784 bp and 1024 bp would appear should the insert be inserted 'backward' (Fig. 14B). Similarly, the band patterns for DREB2H after analysis would be: 2240 bp and 560 bp if inserted forward; and 2687 bp and 113 bp if inserted backward. The plasmids which upon digestion, presented these bands, were selected. Figure 14. Restriction enzyme analysis of the gene DREB2C in pCR8/GW/TOPO. A. 0.7% Agarose gel electrophoresis showing the product of a restriction enzyme analysis of two different samples of DREB2C in pCR8/GW/TOPO with EcoR I ('I' is the same sample as in B.1, and II is B.2) B. 0.7% Agarose gel electrophoresis showing the product of a restriction enzyme analysis of DREB2C with Bsp120 and Cla I C. 0.7% Agarose gel electrophoresis showing the product of a restriction enzyme analysis of a sample of DREB2H in pCR8/GW/TOPO with EcoR I 5.2.2. Subcloning into Destination Vector As stated before, the cloning technique used to subclone the inserts into the destination vector is the Gateway LR reaction. pUBC-GFP-DEST is an integrative plasmid that allows the creation of a Green Fluorescent Protein (GFP) fusion protein and its overexpression in plants, it is also designed for its use in this reaction as it has in its sequence the attR1 and attR2 sites, as well as the lethal gene ccdB, which prevents the transformed cells from carrying any component of the reaction that is not the final product. (Fig. 15) In this case, the result will not be a fusion protein as it is not cloned in frame with GFP. Before the reaction, the pCR8/GW/TOPO entry clone was linearized with an EcoR V digestion and furtherly treated with alkaline phosphatase, in order to make the selection process easier. Figure 15. Diagram of the Gateway cloning LR reaction. The reaction was performed as described, and the product was transformed into competent E. coli cells which were subsequently cultured into spectinomycin selective plates. After 16h of culture, all plates presented colonies, which were screened through a directional colony PCR, that being a PCR that uses one primer hybridizing in the plasmid (primerGFP) and another located within the gene, thus only allowing clones with the insert in the correct direction to act as a template for amplification. In the case of DREB2C, several positive clones appeared (2, 3, 5, 7, and 8 in Figure 16), with a single defined band presenting below the 1500 bp mark; whereas for DREB2H, multiple bands appeared, possibly due to unspecific amplifications, but all of them presented the expected band at 700 bp. (Figure 16) Figure 16. Colony PCR of the genes DREB2C and DREB2H in pUBC/GFP/DEST. 0.7% Agarose gel electrophoresis showing the product of a colony PCR of the Escherichia coli cells transformed with the overexpression plasmid pUBC/GFP/DEST with the genes DREB2C and DREB2H. The molecular marker used is ThermoFisher Scientific's GeneRuler 1kb. The selected colonies ('2' for DREB2C and '8' for DREB2H) were collected and their plasmids were purified to transform competent Agrobacterium tumefaciens cells, which were furtherly selected and grown in liquid media. 5.2.3. Obtention of mutants For the obtention of gain-of-function mutants, Columbia-0 and double knockout mutants for the genes HB25 and HB22 Arabidopsis plants were transformed via floral dip with the Agrobacterium cultures containing the pUBC/GFP/DEST DREB2C and DREB2H clones. A total of 8 plants were transformed, two for each construction, and approximately 0,2 g of their seeds were collected and subsequently sown in BASTA-selective media. The transformed seedlings presented enlarged cotyledons when compared to the WT control seedlings, and the selected individuals were transferred to soil. 10 seedlings were selected for DREB2C and 20 for DREB2H. After a week of growth, the plants' leaves were large enough for the extraction of their RNA and subsequent evaluation through a quantitative reverse transcription polymerase chain reaction (RT-qPCR). Hence, samples were taken from 9 transformed plants for DREB2C and 14 transformed plants for DREB2H. From these extractions, cDNA was synthesized for its use as a template in the reaction. With this assay, the amount of amplification cycles needed to reach a certain intensity of fluorescence varies with the abundance of the specific template. Significant differences in the number of cycles needed between samples indicate differences in the expression of the target gene. To avoid biases, a control gene needs to be evaluated to normalize the basal expression, the housekeeping gene At5g55840. The results of the analyses are expressed in the fold-change, which is the relative difference between the gene expression in the putative mutants and the control sample, in this case, an untransformed Arabidopsis athb22-athb25 double knockout plant. The experiment, after a Student's t-test statistical analysis, revealed that there were 2 mutants that presented an overexpression profile for the gene DREB2H, specifically between 2 and 3 times higher expression than the control sample (Figure 17). For transformed WT plants, no mutants were yet detected, but more results are expected in the foreseeable future. Figure 17. Expression levels of DREB2H (At2g40350) measured

by quantitative RT PCR. Relative transcript levels of DREB2H in mutant plants (green bars) in comparison with the control, an athb22-athb25 double knockout plant (purple bar). The error bars were calculated with the standard error formula between the 3 replicates used for each sample, and the asterisk represents samples that show significant differences from the control sample. (α= 0.05). These gain-of-function mutants within the double knockout background of HB22 and HB25 are expected to be used as a control in future phenotyping assays that measure seed longevity as a way to decipher the interaction between DREB2 TFs and HB25. In this experiment, when overexpressing DREB2C or DREB2H in plants with a WT background, if the phenotype is the same as plants overexpressing HB25, it becomes evidence of interaction between both genes. Nonetheless, there needs to be a control sample that corroborates that the enhancement of seed longevity is caused by the homeobox genes instead of other factors that can be affected by the DREB2 TF abnormal expression profile. The mutants obtained in this project fill precisely that role, having a knockout version of HB25 but also of another gene with redundant function, HB22 (Renard et al., 2021) By contrast, the other mutants aimed to be obtained in this project, the knockout mutants for DREB2 TFs, would be used similarly in the phenotyping assays to determine if the loss-of-function presents any detrimental effects on the seed coat formation and ultimately in the longevity of the mutants' seeds. Furthermore, they would also be helpful for in vivo expression assays, where the values of expression of the downstream gene, HB25 would be measured in the context of loss-of-function for DREB2 TFs. A descent in the values in comparison to a control WT plant would support the hypothesis of interaction. That's the reason why the obtention of these mutants is crucial to allow further in vivo interaction assays. In this project, only the DREB2H overexpression mutants in an athb22-athb25 background were successfully obtained, indicating the need for repetition of the transformation procedures in order to detect a larger amount of mutants. 6. CONCLUSIONS The conclusions that can be extracted from this study are: - 6 different sgRNAs were designed for the edition of the genes DREB2C and DREB2H through a CRISPR/Cas9 genome editing system. - All plasmids and vectors required for the obtention of mutants were successfully cloned using diverse techniques. - Two overexpression mutants for the gene DREB2H were successfully generated - Further transformation procedures should be done with the vectors developed in this project in order to obtain additional mutants 7. BIBLIOGRAPHY AGARWAL, P. K., AGARWAL, P., REDDY, M. K., & SOPORY, S. K. (2006). Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. Plant Cell Reports, 25(12), 1263–1274. https://doi.org/10.1007/s00299-006-0204-8 ALMOGUERA, C., PRIETO-DAPENA, P., DÍAZ-MARTÍN, J., ESPINOSA, J., CARRANCO, R., & JORDANO, J. (2009). The HaDREB2 transcription factor enhances basal thermotolerance and longevity of seeds through functional interaction with HaHSFA9. BMC Plant Biology, 9(1). https://doi.org/10.1186/1471-2229-9-75 ARIF, M. A. R., AFZAL, I., & BÖRNER, A. (2022). Genetic Aspects and Molecular Causes of Seed Longevity in Plants—A Review. Plants, 11(5), 598. https://doi.org/10.3390/plants11050598 BENTSINKL., ALONSO-BLANCO, C., VREUGDENHIL, D., TESNIER, K., GROOT, S. P. C., & KOORNNEEF, M. (2000). Genetic Analysis of Seed-Soluble Oligosaccharides in Relation to Seed Storability of Arabidopsis. Plant Physiology, 124(4), 1595–1604. https://doi.org/10.1104/pp.124.4.1595 BUESO, E., MUÑOZ-BERTOMEU, J., CAMPOS, F., BRUNAUD, V., MARTÍNEZ, L., SAYAS, E., BALLESTER, P., YENUSH, P. L., & SERRANO, R. (2014). Arabidopsis thaliana HOMEOBOX25 Uncovers a Role for Gibberellins in Seed Longevity. Plant Physiology, 164(2), 999–1010. https://doi.org/10.1104/pp.113.232223 CLERKX, E. J. M., BLANKESTIJN-DE VRIES, H., RUYS, G. J., GROOT, S. P. C., & KOORNNEEF, M. (2004). Genetic differences in seed longevity of various Arabidopsis mutants. Physiologia Plantarum, 121(3), 448–461. https://doi.org/10.1111/j.0031-9317.2004.00339.x CLERKX, E. J. M., VRIES, H. B.-D., RUYS, G. J., GROOT, S. P. C., & KOORNNEEF, M. (2003). Characterization of green seed, an Enhancer of abi3-1 in Arabidopsis That Affects Seed Longevity. Plant Physiology, 132(2), 1077–1084. https://doi.org/10.1104/pp.103.022715 DEBEAUJON, I., LÉON-KLOOSTERZIEL, K. M., & KOORNNEEF, M. (2000). Influence of the Testa on Seed Dormancy, Germination, and Longevity in Arabidopsis. Plant Physiology, 122(2), 403–414. https://doi.org/10.1104/pp.122.2.403 DJAFI, N., VERGNOLLE, C., CANTREL, C., WIETRZYÑSKI, W., DELAGE, E., COCHET, F., PUYAUBERT, J., SOUBIGOU-TACONNAT, L., GEY, D., COLLIN, S., BALZERGUE, S., ZACHOWSKI, A., & RUELLAND, E. (2013). The Arabidopsis DREB2 genetic pathway is constitutively repressed by basal phosphoinositide-dependent phospholipase C coupled to diacylglycerol kinase. Frontiers in Plant Science, 4(307). https://doi.org/10.3389/fpls.2013.00307 DOENCH, J. G., HARTENIAN, E., GRAHAM, D. B., TOTHOVA, Z., HEGDE, M., SMITH, I., SULLENDER, M., EBERT, B. L., XAVIER, R. J., & ROOT, D. E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. Nature Biotechnology, 32(12), 1262–1267. https://doi.org/10.1038/nbt.3026 DOYLE , J. J., & DOYLE, J. L. (1987). A Rapid DNA Isolation Procedure for Small Quantities of Fresh Leaf Tissue . Phytochemical Bulletin, 19(11-15). FELICIELLO, I., & CHINALI, G. (1993). A Modified Alkaline Lysis Method for the Preparation of Highly Purified Plasmid DNA from Escherichia Coli. Analytical Biochemistry, 212(2), 394–401. https://doi.org/10.1006/abio.1993.1346 FUCILE, G., DI BIASE, D., NAHAL, H., LA, G., KHODABANDEH, S., CHEN, Y., … PROVART, N. J. (2011). ePlant and the 3D Data Display Initiative: Integrative Systems Biology on the World Wide Web. PLoS ONE, 6(1), e15237. https://doi.org/10.1371/journal.pone.0015237 GAJ, T., GERSBACH, C. A., & BARBAS, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends in Biotechnology, 31(7), 397–405. https://doi.org/10.1016/j.tibtech.2013.04.004 GERNA, D., ARC, E., HOLZKNECHT, M., ROACH, T., JANSEN-DÜRR, P., WEISS, A., & KRANNER, I. (2021). AtFAHD1a: A New Player Influencing Seed Longevity and Dormancy in Arabidopsis? International Journal of Molecular Science, 22(6), 2997–2997. https://doi.org/10.3390/ijms22062997 HU, W., DEPAMPHILIS, C. W., & MA, H. (2008). Phylogenetic Analysis of the Plant-specific Zinc Finger-Homeobox and Mini Zinc Finger Gene Families. Journal of Integrative Plant Biology, 50(8), 1031–1045. https://doi.org/10.1111/j.1744-7909.2008.00681.x ISHINO, Y., KRUPOVIC, M., & FORTERRE, P. (2018). History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. Journal of Bacteriology, 200(7), e00580-17. https://doi.org/10.1128/JB.00580-17 JE, J., SONG, C., HWANG, J.-W., CHUL CHUNG, W., & SEUNG LIM, C. (2013). DREB2C acts as a transcriptional activator of the thermo tolerance-related phytocystatin 4 (AtCYS4) gene. Transgenic Research, 23(1), 109–123. https://doi.org/10.1007/s11248-013-9735-2 JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J. A., & CHARPENTIER, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science, 337(6096), 816–821. https://doi.org/10.1126/science.1225829 KIM, Y., SCHUMAKER, K. , S., & ZHU, J. K. (2006). EMS mutagenesis of Arabidopsis. Methods in Molecular Biology (Clifton, N.J.), 323(2122), 101–103. https://doi.org/10.1385/1-59745-003-0:101 KONCZ, C., NÉMETH, K., RÉDEI, G. P., & SCHELL, J. (1992). T-DNA insertional mutagenesis in Arabidopsis. Plant Molecular Biology, 20(5), 963–976. https://doi.org/10.1007/bf00027166 KONCZ, C., & RBDEI, G. P. (1994). Genetic Studies with Arabidopsis: A Historical View . History of Genetics, 223–252. https://doi.org/0-87969-428-9194 KONCZ, C., & 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(3), 383–396. https://doi.org/10.1007/bf00331014 LATA, C., & PRASAD, M. (2011). Role of DREBs in regulation of abiotic stress responses in plants. Journal of Experimental Botany, 62(14), 4731–4748. https://doi.org/10.1093/jxb/err210 LATCHMAN, D. S. (1997). Transcription factors: An overview. The International Journal of Biochemistry & Cell Biology, 29(12), 1305–1312. https://doi.org/10.1016/s1357-2725(97)00085-x LEE, K. H., HAN, K. J., KWON, Y., PYO LEE, J., YEOU KIM, S., CHUL CHUNG, W., KIM, Y., CHUN, S.-S., JIN KIM, H., & BAE, D.-W. (2009). Identification of potential DREB2C targets in Arabidopsis thaliana plants overexpressing DREB2C using proteomic analysis. Molecules and Cells, 28(4), 383–388. https://doi.org/10.1007/s10059-009-0154-4 LEE, S., KANG, J., PARK, H.-J., KIM, M. D., BAE, M. S., CHOI, H., & KIM, S. Y. (2010). DREB2C Interacts with ABF2, a bZIP Protein Regulating Abscisic Acid-Responsive Gene Expression, and Its Overexpression Affects Abscisic Acid Sensitivity. Plant Physiology, 153(2), 716–727. https://doi.org/10.1104/pp.110.154617 LIM, C. J., HWANG, J. E., CHEN, H., HONG, J. K., YANG, K. A., CHOI, M. S., LEE, K. O., CHUNG, W. S., LEE, S. Y., & LIM, C. O. (2007). Over-expression of the Arabidopsis DRE/CRT-binding transcription factor DREB2C enhances thermotolerance. Biochemical and Biophysical Research Communications, 362(2), 431–436. https://doi.org/10.1016/j.bbrc.2007.08.007 LIU, Q., KASUGA, M., SAKUMA, Y., ABE, H., MIURA, S., YAMAGUCHI-SHINOZAKI, K., & SHINOZAKI, K. (1998). Two Transcription Factors, DREB1 and DREB2, with an EREBP/AP2 DNA Binding Domain Separate Two Cellular Signal Transduction Pathways in Drought- and Low-Temperature-Responsive Gene Expression, Respectively, in Arabidopsis. The Plant Cell, 10(8), 1391–1406. https://doi.org/10.1105/tpc.10.8.1391 LIU, Z., DONG, H., CUI, Y., CONG, L., & ZHANG, D. (2020). Application of different types of CRISPR/Cas-based systems in bacteria. Microbial Cell Factories, 19(1). https://doi.org/10.1186/s12934-020-01431-z MAKAROVA, K. S., WOLF, Y. I., ALKHNBASHI, O. S., COSTA, F., SHAH, S. A., SAUNDERS, S. J., BARRANGOU, R., BROUNS, S. J. J., CHARPENTIER, E.,

HAFT, D. H., HORVATH, P., MOINEAU, S., MOJICA, F. J. M., TERNS, R. M., TERNS, M. P., WHITE, M. F., YAKUNIN, A. F., GARRETT, R. A., VAN DER OOST, J., & BACKOFEN, R. (2015). An updated evolutionary classification of CRISPR–Cas systems. Nature Reviews Microbiology, 13(11), 722–736. https://doi.org/10.1038/nrmicro3569 MOJICA, F. J. M., DÍEZ-VILLASEÑOR, C., GARCÍA-MARTÍNEZ, J., & SORIA, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. Journal of Molecular Evolution, 60(2), 174–182. https://doi.org/10.1007/s00239-004-0046-3 NAKANO, T., SUZUKI, K., FUJIMURA, T., & SHINSHI, H. (2006). Genome-Wide Analysis of the ERF Gene Family in Arabidopsis and Rice. Plant Physiology, 140(2), 411–432. https://doi.org/10.1104/pp.105.073783 NAKASHIMA, K., SHINWARI, Z. K., SAKUMA, Y., SEKI, M., MIURA, S., SHINOZAKI, K., & 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(4), 657–665. https://doi.org/10.1023/a:1006321900483 NIC LUGHADHA, E., BACHMAN, S. P., LEÃO, T. C. C., FOREST, F., HALLEY, J. M., MOAT, J., ACEDO, C., BACON, K. L., BREWER, R. F. A., GÂTEBLÉ, G., GONÇALVES, S. C., GOVAERTS, R., HOLLINGSWORTH, P. M., KRISAI-GREILHUBER, I., LIRIO, E. J., MOORE, P. G. P., NEGRÃO, R., ONANA, J. M., RAJAOVELONA, L. R., & RAZANAJATOVO, H. (2020). Extinction risk and threats to plants and fungi. PLANTS, PEOPLE, PLANET, 2(5), 389–408. https://doi.org/10.1002/ppp3.10146 O'MALLEY, R. C., BARRAGAN, C. C., & ECKER, J. R. (2015). A User's Guide to the Arabidopsis T-DNA Insertional Mutant Collections. Methods in Molecular Biology (Clifton, N.J.), 1284, 323–342. https://doi.org/10.1007/978-1-4939-2444-8_16 OÑATE-SÁNCHEZ, L. (2018). Two-Hybrid Systems. In Methods in molecular biology. Springer Science. https://doi.org/10.1007/978-1-4939-7871-7 PAUWELS, L., DE CLERCQ, R., GOOSSENS, J., IÑIGO, S., WILLIAMS, C., RON, M., BRITT, A., & GOOSSENS, A. (2018). A Dual sgRNA Approach for Functional Genomics in Arabidopsis thaliana. G3 Genes|Genomes|Genetics, 8(8), 2603–2615. https://doi.org/10.1534/g3.118.200046 PEGG, A. E. (2000). Repair of O6-alkylguanine by alkyltransferases. Mutation Research/Reviews in Mutation Research, 462(2), 83–100. https://doi.org/10.1016/S1383-5742(00)00017-X POURCEL, C., SALVIGNOL, G., & VERGNAUD, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology, 151(3), 653–663. https://doi.org/10.1099/mic.0.27437-0 RANGANATHAN, V., WAHLIN, K., MARUOTTI, J., & ZACK, D. J. (2014). Expansion of the CRISPR–Cas9 genome targeting space through the use of H1 promoter-expressed guide RNAs. Nature Communications, 5(1). https://doi.org/10.1038/ncomms5516 RENARD, J., MARTÍNEZ-ALMONACID, I., QUERALTA CASTILLO, I., SONNTAG, A. G., HASHIM, A., BISSOLI, G., CAMPOS, L., MUÑOZ-BERTOMEU, J., NIÑOLES, R., ROACH, T., SÁNCHEZ-LEÓN, S., OZUNA, C. V., GADEA, J., RUBIO, C., 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(2), 679–694. https://doi.org/10.1111/nph.17399 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). PRX2 and PRX25 , peroxidases regulated by COG1, are involved in seed longevity in Arabidopsis. Plant, Cell & Environment, 43(2), 315–326. https://doi.org/10.1111/pce.13656 RENARD, J., NIÑOLES, R., MARTÍNEZ-ALMONACID, I., GAYUBAS, B., MATEOS-FERNÁNDEZ, R., BISSOLI, G., BUESO, E., SERRANO, R., & GADEA, J. (2020). Identification of novel seed longevity genes related to oxidative stress and seed coat by genome-wide association studies and reverse genetics. Plant Cell Environ, 43(10), 2523–2539. https://doi.org/10.1111/pce.13822 RUELLAND, E., VAULTIER, M.-N., ZACHOWSKI, A., & HURRY, V. (2009). Cold Signalling and Cold Acclimation in Plants. Advances in Botanical Research, Chapter 2, 35–150. https://doi.org/10.1016/s0065-2296(08)00602-2 SAKUMA, Y., LIU, Q., DUBOUZET, J. G., ABE, H., SHINOZAKI, K., & 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(3), 998–1009. https://doi.org/10.1006/bbrc.2001.6299 SAKUMA, Y., MARUYAMA, K., OSAKABE, Y., QIN, F., SEKI, M., SHINOZAKI, K., & YAMAGUCHI-SHINOZAKI, K. (2006). Functional Analysis of an Arabidopsis Transcription Factor, DREB2A, Involved in Drought-Responsive Gene Expression. The Plant Cell, 18(5), 1292–1309. https://doi.org/10.1105/tpc.105.035881 SALONIA, F., CIACCIULLI, A., DOMENICA PAPPALARDO, H., POLES, L., PINDO, M., LARGER, S., CARUSO, P., CARUSO, M., & LICCIARDELLO, C. (2022). A dual sgRNA-directed CRISPR/Cas9 construct for editing the fruit-specific β-cyclase 2 gene in pigmented citrus fruits. Frontiers in Plant Science, 13. https://doi.org/10.3389/fpls.2022.975917 SANDER, J. D., & JOUNG, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nature Biotechnology, 32(4), 347–355. https://doi.org/10.1038/nbt.2842 SATTLER, S. E., GILLILAND, L. U., MAGALLANES-LUNDBACK, M., POLLARD, M., & DELLAPENNA, D. (2004). Vitamin E Is Essential for Seed Longevity and for Preventing Lipid Peroxidation during Germination. The Plant Cell, 16(6), 1419–1432. https://doi.org/10.1105/tpc.021360 SCHRAMM, F., LARKINDALE, J., KIEHLMANN, E., GANGULI, A., ENGLICH, G., VIERLING, E., & VON KOSKULL-DÖRING, P. (2007). A cascade of transcription factor DREB2A and heat stress transcription factor HsfA3 regulates the heat stress response of Arabidopsis. The Plant Journal, 53(2), 264–274. https://doi.org/10.1111/j.1365-313x.2007.03334.x SONG, C., BAEK, J., CHOE, U., & LIM, C. S. (2021). Overexpression of DREB2C Delays Flowering in Arabidopsis thaliana via the Activation of FLC. Journal of Plant Biology, 65(2), 133–143. https://doi.org/10.1007/s12374-021-09337-y SONG, C., JE, J., HONG, J. K., & LIM, C. O. (2014). Ectopic expression of an Arabidopsis dehydration-responsive element-binding factor DREB2C improves salt stress tolerance in crucifers. Plant Cell Reports, 33(8), 1239–1254. https://doi.org/10.1007/s00299-014-1612-9 SONG, C., LEE, J., KIM, T., KUK HONG, J., & SEUNG LIM, C. (2018). VOZ1, a transcriptional repressor of DREB2C, mediates heat stress responses in Arabidopsis. Planta, 247(6), 1439–1448. https://doi.org/10.1007/s00425-018-2879-9 SUGLIANI, M., RAJJOU, L., CLERKX, E. J. M., KOORNNEEF, M., & SOPPE, W. J. J. (2009). Natural modifiers of seed longevity in the Arabidopsis mutants abscisic acid insensitive3-5 (abi3-5) and leafy cotyledon1-3 (lec1-3). New Phytologist, 184(4), 898–908. https://doi.org/10.1111/j.1469-8137.2009.03023.x TAN, Q. K.-G. ., & 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(3), 1095–1108. https://doi.org/10.1104/pp.105.070565 TIAN, F., YANG, D.-C., MENG, Y.-Q., JIN, J., & GAO, G. (2019). PlantRegMap: charting functional regulatory maps in plants. Nucleic Acids Research, 48(D1). https://doi.org/10.1093/nar/gkz1020 TRAN, L.-S. P., NAKASHIMA, K., SAKUMA, Y., OSAKABE, Y., QIN, F., SIMPSON, S. D., MARUYAMA, K., FUJITA, Y., SHINOZAKI, K., & YAMAGUCHI-SHINOZAKI, K. (2006). Co-expression of the stress-inducible zinc finger homeodomain ZFHD1 and NAC transcription factors enhances expression of the ERD1 gene in Arabidopsis. The Plant Journal, 49(1), 46–63. https://doi.org/10.1111/j.1365-313x.2006.02932.x WANG, Z.-P., XING, H.-L., DONG, L., ZHANG, H.-Y., HAN, C.-Y., WANG, X.-C., & CHEN, Q.-J. (2015). Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biology, 16(1). https://doi.org/10.1186/s13059-015-0715-0 WENG, M.-L., BECKER, C., HILDEBRANDT, J., NEUMANN, M., RUTTER, M. T., SHAW, R. G., WEIGEL, D., & FENSTER, C. B. (2018). Fine-Grained Analysis of Spontaneous Mutation Spectrum and Frequency in Arabidopsis thaliana. Genetics, 211(2), 703–714. https://doi.org/10.1534/genetics.118.301721 YAMAGUCHI-SHINOZAKI, K., & SHINOZAKI, K. (1993). The plant hormone abscisic acid mediates the drought-induced expression but not the seed-specific expression of rd22, a gene responsive to dehydration stress in Arabidopsis thaliana. Molecular and General Genetics MGG, 238-238(1-2), 17–25. https://doi.org/10.1007/bf00279525 YAMAGUCHI-SHINOZAKI, K., & SHINOZAKI, K. (1994). A Novel cis-Acting Element in an Arabidopsis Gene Is Involved in Responsiveness to Drought, Low-Temperature, or High-Salt Stress. THE PLANT CELL , 6(2), 251–264. https://doi.org/10.1105/tpc.6.2.251 YOSHIDA, T., SAKUMA, Y., TODAKA, D., MARUYAMA, K., QIN, F., MIZOI, J., KIDOKORO, S., FUJITA, Y., SHINOZAKI, K., & YAMAGUCHI-SHINOZAKI, K. (2008). Functional analysis of an Arabidopsis heat-shock transcription factor HsfA3 in the transcriptional cascade downstream of the DREB2A stress-regulatory system. Biochemical and Biophysical Research Communications, 368(3), 515–521. https://doi.org/10.1016/j.bbrc.2008.01.134 ZHAO, Y., ZHANG, C., LIU, W., GAO, W., LIU, C., SONG, G., LI, W.-X., MAO, L., CHEN, B., XU, Y., LI, X., & XIE, C. (2016). An alternative strategy for targeted gene replacement in plants using a dual-sgRNA/Cas9 design. Scientific Reports, 6(1). https://doi.org/10.1038/srep23890 ZHENG, N., LI, T., DITTMAN, J. D., SU, J., LI, R., GASSMANN, W., PENG, D., WHITHAM, S. A., LIU, S., & YANG, B. (2020). CRISPR/Cas9-Based Gene Editing Using Egg Cell-Specific Promoters in Arabidopsis and Soybean. Frontiers in Plant Science, 11(800). https://doi.org/10.3389/fpls.2020.00800 ZINSMEISTER, J., LEPRINCE, O., & BUITINK, J. (2020). Molecular and

environmental factors regulating seed longevity. Biochemical Journal, 477(2), 305–323. https://doi.org/10.1042/bcj20190165 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36