REVIEW PAPER

Genetic insights into the modification of the pre-fertilization mechanisms during plant domestication

Silvia Manrique1,*, James Friel2,3, Pietro Gramazio4,5, Tomas Hasing3, Ignacio Ezquer1, and Aureliano Bombarely1,3,*,

1 Department of Biosciences, Università degli Studi di Milano, Milan, Italy
2 Genetics and Biotechnology Laboratory, Plant and AgriBioscience Research Center (PABC), Ryan Institute, National University of Ireland Galway, Galway, Ireland
3 School of Plant and Environmental Sciences (SPES), Virginia Tech, Blacksburg, VA, USA
4 Instituto de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV), Universitat Politècnica de València, Valencia, Spain
5 Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan

* Correspondence: aurebg@vt.edu

Received 25 December 2018; Editorial decision 2 May 2019; Accepted 2 May 2019

Editor: Silvia Coimbra, University of Porto, Portugal

Abstract

Plant domestication is the process of adapting plants to human use by selecting specific traits. The selection process often involves the modification of some components of the plant reproductive mechanisms. Allelic variants of genes associated with flowering time, vernalization, and the circadian clock are responsible for the adaptation of crops, such as rice, maize, barley, wheat, and tomato, to non-native latitudes. Modifications in the plant architecture and branching have been selected for higher yields and easier harvests. These phenotypes are often produced by alterations in the regulation of the transition of shoot apical meristems to inflorescences, and then to floral meristems. Floral homeotic mutants are responsible for popular double-flower phenotypes in Japanese cherries, roses, camellias, and lilies. The rise of peloric flowers in ornamentals such as snapdragon and florists’ gloxinia is associated with non-functional alleles that control the relative expansion of lateral and ventral petals. Mechanisms to force outcrossing such as self-incompatibility have been removed in some tree crops cultivars such as almonds and peaches. In this review, we revisit some of these important concepts from the plant domestication perspective, focusing on four topics related to the pre-fertilization mechanisms: flowering time, inflorescence architecture, flower development, and pre-fertilization self-incompatibility mechanisms.

Keywords: Clonal propagation, domestication, flower development, flowering time, inflorescence architecture, self-incompatibility, sexual reproduction.

Introduction

Plant domestication is the process of adapting plants to human use by selecting specific traits. Domestication can also be understood as the process of selection of crucial traits by early farmers (domestication syndrome), being different from crop improvement, a later process where secondary traits are selected. For this review, we have preferred to use domestication as a synonym of human-driven active selection of useful (or desirable) traits. Under this definition, domestication can be understood as a continuous process ranging from the active growing of a wild plant with a specific goal (pre-domestication) to their
genetic modification by modern techniques such as CRISPR (clustered regularly interspaced short palindromic repeats; molecular breeding).

Many traits selected during domestication are related to plant reproduction, since most plant-derived food is the product of plant reproduction (seeds and fruits). Plant reproduction is frequently altered in the case of ornamental plants too, as often the targets of the selection are the flowers.

Charles Darwin was one of the first scientists to study the phenotypic changes related to the domestication process. The first chapter of his book ‘On the origin of species’ introduced several ideas about domestication such as the increase in phenotypic diversity and the pushing of the reproductive barriers during domestication (Darwin, 1859). Then, 9 years later in his book ‘The variation of animals and plants under domestication’, Darwin developed his ideas and observations about domestication in greater detail. Chapters X and XI summarize Darwin’s observations on the variation of flowers, buds, and reproduction modes (Darwin, 1868). He described floral homeotic mutations in which stamens and pistils are converted to petals for species such as Aquilegia vulgaris (columbine) and Primula vulgaris (hose-in-hose primroses), respectively. He also mentioned a poppy variety in which stamens have turned into pistils. DeVries (1904) also shared this observation in his book ‘Species and varieties, their origin by mutation’ where he pointed to this phenotype in the species Papaver commutatum. Another change in flower morphology described by Charles Darwin was the transition from zygomorphic to actinomorphic (peloric) flowers in Sinningia speciosa (gloxinia) and Antirrhinum majus (snapdragon). Darwin’s observations exemplify some of the changes in the reproductive mechanisms which occurred during the plant domestication process, but they are not the only ones. Changes in the transition from the vegetative to the reproductive phase, the fertilization process, fruit development, and abscission are also commonly associated with plant domestication. In this review, we present important examples of the alteration of these mechanisms. Due to space constraints, we will focus on the events occurring prior to the fertilization of the ovule: transition of the vegetative to the reproductive phase, inflorescence architecture and flower development, and self-incompatibility (SI). We summarize the main genes described in this article in Table 1.

**Changes in flowering time associated with plant domestication**

The transition from the vegetative to the reproductive stage is controlled by a complex genetic mechanism that translates changes in photoperiod, temperature, and plant hormones into the signal that induces the production of flowers. The plant domestication process involved the adaptation of human-selected populations to environments with different photoperiods and temperatures, which involves the selection of changes in the flowering time of these species. The genetic mechanisms that control flowering time have been extensively studied in several plant species, but most of the work has been done in the model species Arabidopsis thaliana, a long-day plant native to Africa and Eurasia. Hundreds of genes have been described in the flowering time pathway, but, due to space constraints, we will focus on listing genes that have been selected during domestication without giving extensive detail about the pathways, as this has already been reviewed in many excellent articles.

The central player controlling flowering time is the FLOWERING LOCUS T gene (FT), which encodes a small phosphatidyethanolamine-binding protein (PEBP) that binds to phospholipids (Kobayashi et al., 1999). FT is expressed in leaves and is induced by long-day treatment. FT is translocated to the shoot apex where it induces its own expression and activates the expression of floral determination genes to trigger flowering (Wigge et al., 2005; Corbesier et al., 2007).

Photoperiod is the most important environmental signal determining flowering time. In Arabidopsis, photoperiod information is connected to FT through the CONSTANS gene (CO), a zinc finger transcription activator expressed in leaves that activates the expression of FT (Putterill et al., 1995; Kardailsky et al., 1999). The CONSTANS (CO) gene is post-transcriptionally regulated by GIGANTEA (G), a circadian clock gene (Park et al., 1999; Huq et al., 2000). The FT protein moves to the shoot apical meristem (SAM) and activates the expression of floral determination genes to trigger flowering (Wigge et al., 2005; Corbesier et al., 2007). However, FT is tightly regulated in order to integrate other inputs, such as light quality or temperature, to ensure that flowering aligns with seasonal cues for successful reproduction (Valverde et al., 2004; Song et al., 2015). Additionally, vernalization and temperature are also important factors in regulating the expression of FT in plants that require a period of cold before flowering such as the winter-annual ecotypes of Arabidopsis (Michaels and Amasino, 1999).

The flowering time pathways in other species share some similarities with Arabidopsis, but also have distinctive features (Zhu et al., 2017). In the short-day monocot Oryza sativa, Heading date 3a (Hd3a) and FLOWERING LOCUS T 1 (RFT1) represent the orthologs of the Arabidopsis FT gene (Komiya et al., 2008). Heading date 1 (Hd1), the homolog of CO in rice, is regulated by OsGI (the GI homolog in rice) (Yano et al., 2000; Hayama et al., 2003). Hd1 promotes flowering by activating Hd3a expression under short-day conditions and delays flowering by repressing Hd3a expression under long-day conditions (Yano et al., 2000; Kojima et al., 2002; Komiya et al., 2008). The Early heading date 1 (Ehd1) pathway, which is unique to grasses and independent of Hd1 (Doi et al., 2004), also promotes the expression of FLOWERING LOCUS T 1 (RFT1) and Hd3a (Doi et al., 2004, Xue et al., 2008, Itoh et al., 2010).

A very important domestication trait related to flowering time is the ability to flower at latitudes different from the plant’s native region. Therefore, specific alleles of several important players in the flowering time pathway have been positively selected during domestication to achieve this. Cultivated rice, O. sativa L., is a good example of a plant originally from a tropical region that has been adapted to a wide range of latitudes, from 53°N to 40°S. The adaptation of rice to high latitudes was driven by the selection of natural variants of several genes from both Hd1 and Ehd1 pathways. Loss-of-function alleles of Hd1 itself (Goretti et al., 2017) or...
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Arabidopsis homolog</th>
<th>Species</th>
<th>Molecular function</th>
<th>Pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEADING DATE 3A (hd3A, Os06t0157700)</td>
<td>FLOWERING LOCUS (FT, AT1G65480)</td>
<td>Oryza sativa</td>
<td>Phosphatidylethanolamine binding (GO:0008429)</td>
<td>Flowering time</td>
</tr>
<tr>
<td>RICE FLOWERING-LOCUS T 1 (RFT1, Os06t0157500)</td>
<td>FLOWERING LOCUS (FT, AT1G65480)</td>
<td>Oryza sativa</td>
<td>Phosphatidylethanolamine binding (GO:0008429)</td>
<td>Flowering time</td>
</tr>
<tr>
<td>HEADING DATE 7 (Ghd7, Os07g0261200)</td>
<td>NA</td>
<td>Oryza sativa</td>
<td>DNA binding (GO:0003677), protein binding (GO:000551)</td>
<td>Flowering time</td>
</tr>
<tr>
<td>Heading date (QTL)-5(t) (Hd5, Os08g0174500)</td>
<td>NA</td>
<td>Oryza sativa</td>
<td>DNA binding (GO:0003677), protein binding (GO:0005515)</td>
<td>Flowering time</td>
</tr>
<tr>
<td>Os07g0261200</td>
<td>NA</td>
<td>Oryza sativa</td>
<td>Protein binding (GO:0005515)</td>
<td>Flowering Time</td>
</tr>
<tr>
<td>ZmCCT9</td>
<td>Two-component response regulator-like (APRR7, AT5G02810)</td>
<td>Oryza sativa</td>
<td>Protein binding (GO:0005515), zinc ion binding (GO:0008270)</td>
<td>Flowering Time</td>
</tr>
<tr>
<td>HAFT4</td>
<td>Pseudo-response regulator 3 (PRR3, AT5G60100)</td>
<td>Oryza sativa</td>
<td>Protein binding (GO:0005515), nuclease binding (GO:0000166)</td>
<td>Flowering Time</td>
</tr>
<tr>
<td>Heading date (QTL)-5(t) (Hd5, Os08g0174500)</td>
<td>NA</td>
<td>Oryza sativa</td>
<td>Protein binding (GO:0005515), zinc ion binding (GO:0008270)</td>
<td>Phloperiod</td>
</tr>
<tr>
<td>GmCRY1a</td>
<td>CRYPTOCHROME 2 (CRY2, AT1G04400)</td>
<td>Glycine max</td>
<td>Protein binding (GO:0005515), ubiquitin-protein ligase activity (GO:0048422)</td>
<td>Phloperiod</td>
</tr>
<tr>
<td>Glycine max CONSTANS-like 7a (GMCOL7a)</td>
<td>CONSTANS (CO, AT5G15840)</td>
<td>Glycine max</td>
<td>Protein binding (GO:0005515), ubiquitin-protein ligase activity (GO:0048422)</td>
<td>Flowering Time</td>
</tr>
<tr>
<td>Glycine max phosphatidylethanolamine-binding protein FT2a (GmFT2a, Glyma16G150700)</td>
<td>FLOWERING LOCUS T (FT, AT1G65480)</td>
<td>Helianthus annuus</td>
<td>Phosphatidylethanolamine binding (GO:0008429)</td>
<td>Flowering time</td>
</tr>
<tr>
<td>Solyc09g075080</td>
<td>Phytinome A-associated F-box protein (F-box, AT4G02440)</td>
<td>Helianthus annuus</td>
<td>Phosphatidylethanolamine binding (GO:0008429)</td>
<td>Flowering time</td>
</tr>
<tr>
<td>Solyc01g068560</td>
<td>NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED 2 (LNK2, AT3G54500)</td>
<td>Solanum lycopersicum</td>
<td>DNA binding (GO:0003677), ubiquitin-protein ligase activity (GO:0008429)</td>
<td>Flowering time</td>
</tr>
<tr>
<td>Solyc02g0733500</td>
<td>TEOSINTE BRANCHED 1 (TCP24, AT1G30210)</td>
<td>Zea mays ssp. mays</td>
<td>DNA binding (GO:0003677), ubiquitin-protein ligase activity (GO:0008429)</td>
<td>Flowering time</td>
</tr>
<tr>
<td>Solyc06g0743500</td>
<td>ABERRANT PANICLE ORGANIZATION (Os06g0665400)</td>
<td>Solanum lycopersicum</td>
<td>DNA binding (GO:0003677), ubiquitin-protein ligase activity (GO:0008429)</td>
<td>Flowering time</td>
</tr>
<tr>
<td>Self-pruning (sp, Solyc06g0743500)</td>
<td>Floral meristem identity control protein LEAFY (LFY, AT5G61850)</td>
<td>Oryza sativa</td>
<td>Ubiquitin-protein transferase activity (GO:0048422), protein binding (GO:0005515)</td>
<td>Flower development</td>
</tr>
<tr>
<td>Brassica oleracea</td>
<td>TERMINAL FLOWER 1 (TFL1, At5g03840)</td>
<td>Glycine max</td>
<td>Transcription co-regulator activity (GO:003712)</td>
<td>Flower development</td>
</tr>
<tr>
<td>Putative phosphatidylethanolamine-binding protein TFL1a (GmTFL1, Glyma03G194700)</td>
<td>CENTRORADIALIS (ATC, AT2G27550)</td>
<td>Solanum lycopersicum</td>
<td>Phosphatidylethanolamine binding (GO:0008429)</td>
<td>Flower development</td>
</tr>
<tr>
<td>Falsikora (fa, Solyc03g118160)</td>
<td>FLOWERING LOCUS T (FT, AT1G65480)</td>
<td>Helianthus annuus</td>
<td>Phosphatidylethanolamine binding (GO:0008429)</td>
<td>Flowering time</td>
</tr>
<tr>
<td>COMPOUND INFLORESCENCE (S, Solyc02g0733500)</td>
<td>CENTRORADIALIS (ATC, AT2G27550)</td>
<td>Solanum lycopersicum</td>
<td>Phosphatidylethanolamine binding (GO:0008429)</td>
<td>Flower development</td>
</tr>
<tr>
<td>LIGULELESS 1 (OsLG1, Os04g0656500)</td>
<td>Floral meristem identity control protein LEAFY (LFY, AT5G61850)</td>
<td>Solanum lycopersicum</td>
<td>Transcription factor binding (GO:0008134)</td>
<td>Flowering time, flower development</td>
</tr>
<tr>
<td>CYCLOIDEA (Cyc, Os09g02450)</td>
<td>TEOSINTE BRANCHED 1 (TCP24, AT1G30210)</td>
<td>Antirrhinum majus</td>
<td>DNA binding (GO:0003677), DNA-binding transcription factor activity (GO:0003700)</td>
<td>Flower development</td>
</tr>
</tbody>
</table>
the repressors of Ehd1, Gm
number, plant height and heading date (Ghd7) (Xue et al., 2008) and Days To Heading on chromosome 8 (DTH8/Ghd8; OsHAP3/Hd4) (Xue et al., 2008; Wei et al., 2010; Fujino et al., 2013) were selected to obtain plants with low photoperiod sensitivity. Selection of natural variation of OsPRR37, a pseudo-response regulator (PRR) gene that makes up part of the circadian clock, also contributed to the adaptation of rice to cultivation at higher latitudes (Koo et al., 2013).

In maize, the adaptation to different geographical regions follows a similar pattern to that of rice. Maize (Zea mays ssp. mays) and its ancestor teosinte are native to tropical South-western Mexico. The adaptation of maize to higher latitudes is linked with the down-regulation of the gene ZmCCT9 (homolog of the rice gene Ghd7) by the insertion of a Harbinger-like transposon in a distant regulative region, which results in photoperiod insensitivity, allowing flowering in long-day conditions (Huang et al., 2018).

Conversely, the domestication of wheat and barley followed different trajectories. While some varieties were selected as winter crops, with day-neutral behaviour and a strong verbalization requirement, other varieties were selected as short-seasoned spring varieties not requiring vernalization (Blümel et al., 2015). However, in both cases, most known variations are associated with changes in VERNALIZATION 1 (VRN1), a MADS-box transcription factor gene involved in the vernalization process that promotes inflorescence initiation (Fjellheim et al., 2014).

In the case of soybean (Glycine max), adaptation to other latitudes is associated with a strong selection of the genes GmCRY1a (homolog of the Arabidopsis blue light receptor CRYPTOCHROME 2) and GMCOL7a (homolog of CO) (Li et al., 2013). Soybean quantitative trait locus (QTL) analysis also linked four alleles (E1, E2, E3, and E4) to flowering time (Zhai et al., 2014). The major contributor, E2, is an allele of the GmGlu gene, the soybean homeolog of GI of Arabidopsis. The dominant E2 allele carries an early stop codon mutation that leads to induction of the expression of the GmFT2a gene, producing early flowering (Langewisch et al., 2014). Another case in which FT is affected is sunflower (Helianthus annuus L.), where five members of the FT gene family were selected during domestication, including a non-functionalized copy (HaFT4) (Blackman et al., 2011).

Other interesting cases are the mutations selected in the tomato genes Solyc09g075080, the homolog of the Arabidopsis F-box EID1 that functions as a negative regulator in phytochrome A-specific light signaling, and Solyc01g068560, a homolog of the Arabidopsis NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED 2 (LKN2) that functions in the integration of light signaling and the circadian clock. They are responsible for the adaptation of this crop to the long summers in temperate regions. A single amino acid deletion in the Soly09g075080/EID1 protein delays the circadian phase (Müller et al., 2016). For the Solyc01g068560/LNKN2 gene, an almost complete deletion lengthens the circadian period (Müller et al., 2018).

### Modifications in inflorescence architecture during plant domestication

The architecture of the inflorescence conditions the number of flowers and, as a consequence, fruits and seeds that are produced and their position on the plant (Wyatt, 1982; Evers et al., 2011; Iwata et al., 2012; Teo et al., 2014). Therefore, genes controlling inflorescence development are instrumental for domestication as they have a profound impact on key agronomical aspects such as yield and crop management (Benlloch et al., 2015).

Inflorescence architecture is determined by two main factors: the growth habit of the plant and the level of in/determinacy of the inflorescence meristem (IM). Regarding growth habit, in monopodial plants, such as Arabidopsis or rice, due to apical dominance, vertical growth results only from the SAM. In these plants, after floral transition, the main SAM develops into the leader inflorescence shoot with subordinate branches. Instead, in sympodial plants such as tomato, the role of the SAM is sequentially adopted by the uppermost axillary meristem, which, after a period of growth, will either terminate in reproductive structures or abort. Then, the growth will continue from a new axillary meristem that will repeat this pattern so, instead of a leader inflorescence, several inflorescences of similar size are formed along the shoot.

One of the best-known examples of genes modifying inflorescence architecture related to domestication is TEOSINTE-BRANCHED 1 (TB1) of maize/teosinte. teosinte branched1 (tb1) was identified as a major QTL contributing to the shift towards monopodial growth habit with a concomitant increase in ear size during teosinte domestication (Doebley et al., 1997). tb1 encodes a member of the TCP family of transcriptional regulators expressed in the axillary meristems (Doebley et al., 1997; Hubbard et al., 2002; Kebrom and Bruttell, 2015). The TB1 protein acts as a repressor of organ growth and contributes to apical dominance by repressing branch outgrowth. The maize allele of TB1 is more highly expressed than that of teosinte, causing greater repression of branching compared with

### Table 1. Continued

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Arabidopsis homolog</th>
<th>Species</th>
<th>Molecular function</th>
<th>Pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>DICHTOMA (DICH, Q0GJ4)</td>
<td>NA</td>
<td>Antirhinum majus</td>
<td>DNA binding (GO:0003677)</td>
<td>Flower development</td>
</tr>
<tr>
<td>DIVARICATA (DIV, Q859H7)</td>
<td>NA</td>
<td>Antirhinum majus</td>
<td>DNA binding (GO:0003677)</td>
<td>Flower development</td>
</tr>
<tr>
<td>RADIALIS (RAD, Q58FS3)</td>
<td>NA</td>
<td>Antirhinum majus</td>
<td>DNA binding (GO:0003677)</td>
<td>Flower development</td>
</tr>
</tbody>
</table>
teosinte (Doebly et al., 1997; Studer et al., 2011). Increased expression of TB1 in maize is due to an insertion of a transposable element 65 kb upstream of the TB1 coding region (Studer et al., 2011; Kebrom and Bruttell, 2015).

The balance between the maintenance of indeterminacy or commitment to flower is the main determinant of the architecture of the inflorescence. Upon perception of inducing environmental and/or internal cues, the SAM transitions from a vegetative to a reproductive identity, becoming the primary IM (Wils and Kaufmann, 2017; Cheng et al., 2018). The primary IM (I₁) can either produce flowers or remain indeterminate to produce branch meristems (I₂), which iterate the pattern of I₁ (Prusinkiewicz et al., 2007; Teo et al., 2014). Based on the activity of the I₁ and the I₂, inflorescences can be classified into four groups: when the I₁ terminates as a flower, they are determinate inflorescences; in contrast, when the I₁ grows indefinitely until senescence to produce a floral meristem (FM) or I₂, they are indeterminate inflorescences. Additionally, when the I₁ forms the FM, we are in the presence of simple inflorescences. Alternatively, if the I₁ forms further IMs, increasing the complexity of the architecture, we refer to them as compound inflorescences (Weberling, 1989; Benlloch et al., 2007, 2015; Cheng et al., 2018). A basic genetic model to explain how the balance between IM and FM identity is determined has been developed using studies in the model plant A. thaliana. Although the applicability of this model to other species is variable, it is useful to set a frame for comparison.

The development of the Arabidopsis inflorescence can be mostly explained by the function of three genes: TERMINAL FLOWER 1 (TFL1), LEAFY (LFY), and APETALA 1 (AP1) (Shannon and Meeks-Wagner, 1993; Liljegren et al., 1999; Blázquez et al., 2006; Benlloch et al., 2015). These genes coordinate to maintain the balance between IM and FM identity at the inflorescence apex. Broadly speaking, TFL1 promotes IM identity, while LFY and AP1 promote FM identity. Therefore, it has been proposed that differences in their expression patterns or function can explain much of the diversity of inflorescence architectures observed among angiosperms (Ratcliffe et al., 1999; Blázquez et al., 2006; Benlloch et al., 2007; Serrano-Mislaeta et al., 2017). Briefly, TFL1, a PEBP, is specifically expressed in the center of the I₁ and I₂, and promotes IM identity by repressing LFY and its direct target AP1 (and its paralog CAULIFLOWER, CAL) to prevent early inflorescence termination (Mandel et al., 1992; Weigel et al., 1992; Weigel and Nilsson, 1995; Parcy et al., 1998; Teo et al., 2014). Conversely, LFY, a plant-specific transcription factor gene, and AP1 and CAL, two paralog MADS-box transcription factor genes, are expressed in the lateral FM primordia produced by the IM. The joint action of LFY and AP1/CAL in the newly formed FM leads to the repression of TFL1, allowing the up-regulation of floral organ identity genes and leading to the formation of flowers (Parcy et al., 1998; Liljegren et al., 1999; Wagner et al., 1999; Kaufmann et al., 2009). However, the mechanism leading to TFL1 repression in the FM is not linear. Recent works pointed out that, actually, LFY activates TFL1 in the FM while AP1/CAL represses it (Goslin et al., 2017; Serrano-Mislaeta et al., 2017). This indicates that LFY and AP1 might be part of a feed-forward loop that could serve to ensure that flower development starts only when AP1/CAL levels are high enough to over-ride LFY inhibitory action, ensuring that the conditions for stable development of flowers are already established (Goslin et al., 2017). These results imply that LFY might also be involved in maintaining the indeterminate growth of the IM. Indeed, the activity of the TFL1 promoter is reduced in ify mutants (Serrano-Mislaeta et al., 2017). Although LFY is not expressed in the SAM, LFY protein is mobile (Sessions et al., 2000) so it can travel to the IM and bind to the TFL1 promoter. Therefore, the relationship between TFL1 and LFY is not entirely antagonistic as previously thought.

Regarding the situation in other plants, of the three genes forming the model, TFL1 is the most conserved. TFL1 orthologs exist in most land plants, and investigations in various plant species have shown that their role is mostly conserved (Chardon and Damerval, 2005; Carmona et al., 2007; Danilevskaya et al., 2010; C. Liu et al., 2013; Teo et al., 2014; Mahrez et al., 2016). Mutations of TFL1 produce a switch to more determinate inflorescences. Examples of crops where this has been selected during domestication are soybean and tomato (Wang et al., 2018). Glycine soja, the wild progenitor of soybean (G. max), is indeterminate. Instead, many cultivated soya varieties have a determinate growth habit. This trait was found to be controlled by the Drl locus encoded by GmTFL1 (Liu et al., 2010). As soybean contains several TFL1 paralogs, complementation of the Arabidopsis tfl1 mutant with GmTFL1 demonstrated that it was the functional TFL1 ortholog (Tian et al., 2010).

Instead, the presence and role of LFY and AP1 vary among species. For example, the rice ortholog of LFY, ABERRANT PANICLE ORGANIZATION 2/RICE FLORICAULA, is not expressed in FMs and, in contrast to Arabidopsis, it has a role in suppressing the transition from IM to FM (Kyozuka et al., 1998; Ikeda-Kawakatsu et al., 2012).

Another example is the tomato self-pruning (sp) mutant. sp was discovered 90 years ago and has facilitated the transformation of indeterminate tomato plants into new determinate forms (Yeager, 1927; Pnueli et al., 1998; Wang et al., 2018), leading to a more compact phenotype and synchronized growth which is adequate for mechanical harvesting (McGarry et al., 2016). For this reason, the Sp mutation was rapidly bred into all industrial tomatoes. However, it must be noted that although the product of Sp is a PEBP and the functional equivalent of that of TFL1 in tomato, the real ortholog of Sp in Arabidopsis is another PEBP-encoding gene, Arabidopsis thaliana CENTRORADIALIS (ATC) (Mimida et al., 2001; McGarry et al., 2016).

Obvious examples of inflorescence architecture re-shaping associated with domestication are the mutants of Brassica oleracea, cauliflower (B. oleracea ssp. botrytis), and broccoli (B. oleracea ssp. Italica). The cauliflower head is composed of a hypertrophied mass of IMs and FMs. In broccoli, developmental arrest happens at a later stage, so although the inflorescence also develops into a large hypertrophied structure, flower buds are eventually formed (Carr and Irish, 1997; Schilling et al., 2018). Upon the characterization of the ap1-1/dal-1 mutant of Arabidopsis, the similarity of broccoli and cauliflower with it led to speculation that the AP1 and CAL orthologs...
from *B. oleracea* might be responsible for this phenotype (Smith and King, 2000). Surprisingly, the link is not as clear as initially thought and, at present, the basis of these phenotypes is still not completely clear. Molecular and population genetic studies indicate that the function of the *B. oleracea* CAL (BoCAL) is compromised in both varieties (Kempin et al., 1995; Lowman and Purugganan, 1999; Purugganan et al., 2000; Smith and King, 2000). The situation for *API*-like genes is less clear, since several copies of *API*-like genes exist in *B. oleracea* (Lowman and Purugganan, 1999) and, although they are associated with the phenotype, they do not explain it entirely, indicating that additional genes might be involved (Labate et al., 2006; Duclos and Björkman, 2008; Schilling et al., 2018).

Besides *TFL1*, *LFY*, and *API*, many other genes involved in the network that regulates inflorescence architecture have been targeted during domestication. For example, altered inflorescence branching in *FALSIFLORA* (*FALSIFLORA*) mutants in Arabidopsis and commercial highly branched tomato varieties carry mutations on *LFY* (*LFY*). For example, in *FALSIFLORA* (*FALSIFLORA*) in tomato, cause an increase in inflorescence branching, and the phenotype, they do not explain it entirely, indicating that additional genes might be involved (Labate et al., 2006; Duclos and Björkman, 2008; Schilling et al., 2018).

Floral crop selection as an example of alterations in flower development pathways

Perfect flowers contain four types of organs arranged in concentric rings known as whorls. From the outermost to innermost whorl, the organ types are: sepals, petals, stamens, and carpels. A combinatorial model that explains how these four organ types are specified within the FM was proposed in the late 1990s, based on the observation of a series of homeotic mutants in Arabidopsis and *Antirrhinum* (*Causier et al., 2010; Moyroud and Glover, 2017*). Presently known as the ABCDE model, the model was originally proposed as the ABC model and extended later on (Coen and Meyerowitz, 1991; Weigel and Meyerowitz, 1994; Causier et al., 2010).

The model proposes that five functions named A, B, C, D, and E specify which organs form in each whorl of the flower. A+E genes specify sepals, A+B+E specify petals, B+C+E specify stamens, C+E specify carpels, and C+D+E specify ovules, and the ABCDE genes are sufficient to superimpose floral organ identity in vegetative organs of angiosperms (Percy et al., 1998; Honma and Goto, 2001; Pelaz et al., 2001). Additionally, it was observed that C-function expands into the outer whorls in A-function mutants and vice versa, so mutual repression between the A- and C-functions was integrated into the model to explain it (Causier et al., 2010). In Arabidopsis, the genes responsible for A-function are *APETALA1* (*AP1*) and *APETALA2* (*AP2*), B-function is encoded by *APETALA3* (*AP3*) and *PISTILLATA* (*PI*), and C-function is encoded by *AGAMOUS* (*AG*). Differences in flower development are involved in inflorescence architecture (Molinero-Rosales et al., 1999; Zheng et al., 2017), and the situation for *API*-like genes is less clear, since several copies of *API*-like genes exist in *B. oleracea* (Lowman and Purugganan, 1999) and, although they are associated with the phenotype, they do not explain it entirely, indicating that additional genes might be involved (Labate et al., 2006; Duclos and Björkman, 2008; Schilling et al., 2018).

Although the model is 30 years old, the conceptual framework continues to be broadly valid. Floral identity genes are also present in gymnosperms, where a ‘BC’ model has been proposed (Baum and Hileman, 2006; Theissen and Melzer, 2007; Chanderbali et al., 2016), in which C-function is expressed in male and female cones while B-function is restricted to male cones (Irish, 2017). In this context, the ABCDE model can be seen as an evolutionary extension of the ‘BC’ model. Instead, A-function has always been controversial as it seems to be much less conserved than the others and, in the last years, evidence has been accumulating pointing to the fact that outside Arabidopsis and close relatives, a classical A-function is rare (Litt, 2007; Ye et al., 2016; Morel et al., 2017; Wils and Kauffmann, 2017; Schilling et al., 2018). However, a recent study has shown that, in rice, A-function exists and it is performed by OsMADS14 and OsMADS15 genes which belong to the API/FUL clade (Wu et al., 2017), so the debate is still open.

The other functions are quite well conserved across angiosperms, particularly eudicots and monocots, although, for example, in petunia C- and D-function cannot be strictly distinguished from each other (Heijmans et al., 2012; Theissen et al., 2016; Schilling et al., 2018). Finally, some differences that do not fit in a whorl-based ABCDE model are observed in some basal angiosperms. However, those differences can be explained as modifications of the ABCDE model like the ‘fading borders model’ (Causier et al., 2010; Chanderbali et al., 2016; Wils and Kauffmann, 2017).

Regarding modification of flower morphology due to domestication, probably the most frequent and notorious examples are the alterations in *AG* that result in double flowers where stamens and carpels are replaced by petals (Bowman et al., 1989; Schilling et al., 2018). Since the determinacy of the meristem is also disturbed, this pattern is iterated multiple times, leading to flowers with very high numbers of petals (Bowman et al., 1989). In some cases, double flowers are associated with loss-of-function mutations in AG-like genes, such as, for example, in rue-anemone (*Thalictrum thalictroides* or...
Japanese cherry (Prunus lannesiana) (Galimba et al., 2012; Z. Liu et al., 2013; Schilling et al., 2018). However, the formation of double flowers in rose, Camellia, or lily is instead associated with a restriction of the expression domain of AG-like genes toward the center of the meristem (Dubois et al., 2010; Akita et al., 2011; Sun et al., 2014). The underlying molecular cause of the shrinkage of the AG expression domain was unknown, but recent data from rose and peach indicate that it might be caused by mutations in euAP2 genes, which are known to repress AG in many species (François et al., 2018; Gattolin et al., 2018). It seems that in both cases the mutations responsible for the change in the AG expression pattern were caused by the loss of the miR172-binding site in the euAP2 gene (François et al., 2018; Gattolin et al., 2018).

Besides floral organs, the symmetry of flowers is a characteristic often modified during domestication, especially in ornamental crops. Flowers can have two types of symmetry: radial or actinomorphic and bilateral or zygomorphic (Endress, 1999; Krizek and Fletcher, 2005; Smyth, 2018).Zygomorphy is thought to have evolved many times from an ancestral actinomorphic condition as a strategy apparently associated with the attraction of pollinating insects with bilateral vision (Krizek and Fletcher, 2005; Smyth, 2018). Although the classification is binary, the outcome is not, as differential development of each whorl can give rise to several intermediate situations. For example, while orchids are the classical example of zygomorphy and show a strong bilateral symmetry spanning all whorls, some species of Solanum present a restricted zygomorphy affecting only the stamen whorl, related to their interaction with pollen-collecting bees (Glover et al., 2004). Snapdragons (Antirrhinum majus) is the classic genetic model in which zygomorphy has been studied. Its symmetry is based on the expansion of the dorsal petals relative to the lateral and ventral ones and abortion of the dorsal stamen (Krizek and Fletcher, 2005). Dorsal identity is specified by two paralogous TCP-domain family transcription factor genes with overlapping functions, CYCLOIDEA (CYC) and DICHOTOMA (DICH). Ventral identity is specified by the MYB-domain transcription factor gene DIVARICATA (DIV), expressed all across the flower. Expression of CYC and DICH in the dorsal domain activates RADIALIS (RAD), which encodes a protein with a single MYB domain. RAD antagonizes DIV function in dorsal cells and limits its activity to the lateral and ventral domains by competing for DIV interaction partners, the DIV-AND-RAD-INTERACTING-FACTORs (DRIFs), which are required for DIV activity in specifying ventral symmetry (Luo et al., 1996, 1999; Galego and Almeida, 2002; Corley et al., 2005; Perez-Rodriguez et al., 2005; Gao et al., 2017). Thus, rad and cyc dich double mutants are ventralized and have a radially symmetrical appearance. Many regressions to actinomorphic symmetry, both in natural populations (Cubas et al., 1999; Reardon et al., 2009) and in domesticated crops, are caused by mutations affecting CYC expression. A recent example of zygomorphic to actinomorphic reversion in a domesticated plant caused by a mutation in a CYC ortholog are gloxinias (Sinningia speciosa), where the loss of CYC function is caused by a 10 bp deletion in the coding sequence of the gene (Dong et al., 2018).

CYC-like genes have also been associated with asymmetric pigmentation in zygomorphic flowers. For example, TfCYC2 in wishbone flower (Toireana fournieri) evolved regulatory loops to bind to the promoter region of an R2R3-MYB factor gene repressing its transcription, which under normal circumstances promotes anthocyanin-related pigmentation in the epidermal cells of petals (Su et al., 2017).

**Limitations in plant breeding driven by pre-fertilization self-incompatibility**

Plant fitness/yield has been fundamentally changed during domestication by hybridization, genetic bottlenecks, alteration of reproductive strategies, and polyploidization. Such changes have greatly modulated current plant traits in agriculture. Angiosperms exhibit a wide array of reproductive strategies, both sexual and sexual, with sexual reproduction including self-fertilization and cross-fertilization strategies. In unstable or unpredictable environments, reproductive strategies promoting cross-fertilization are fundamental to evolutionary success as they contribute to the creation of genetically diverse populations which increase the probability that at least one individual in a population will survive under changing conditions. However, in situations of low presence of sexual partners, or in stable and predictable environments, an asexual strategy and self-fertilization are effective means of reproduction that can be favored. Numerous wild plants display efficient mechanisms that ensure outcrossing promoting high levels of plant heterozygosity. SI is reported in >100 families and distributed among an estimated 39% of species (Igic and Kohn, 2006). During angiosperm evolution, different molecular mechanisms for promoting SI have evolved at least 35 times (Iwano and Takayama, 2012). Self-compatibility (SC), on the other hand, might have evolved to adapt to conditions such as the loss of pollinators (Gervasi and Schiestl, 2017).

Perennial species are generally outcrossers, while annuals are more tolerant to SC. Since annuals only have one chance to reproduce during their lifespan, for them it might be better to self-pollinate, rather than not reproducing at all. Conversely, perennial species can wait longer and avoid the detrimental effects of inbreeding over time (Pekkala et al., 2014).

Outcrossing can be achieved either through the spatiotemporal separation of the sexes via (hetero) dichogamy or dioecy, or by SI (Miller and Gross, 2011). SI is a genetically controlled mechanism that induces higher sexual selection by preventing self-fertilization in wild-type plants. In some cases, SI arises from floral morphology (heteromorphy) thanks to genetically controlled physical or temporal barriers that prevent self-pollination. This is the case for primrose (Primula) which exhibits two floral forms (morphs) that differ in morphology, primarily in the relative placement of stigmas and anthers, and pollinations succeed only between different morphs (De Nettancourt, 2001). Multiple varieties of heteromorphic SI systems are present in the plant kingdom (seen in Passifloraceae, Lythraceae, Polygonaceae, and Primulaceae) which are considered to have evolved independently (Fujii et al., 2016).
Nevertheless, most SI systems in plants belong to homomorphic systems where incompatibility is achieved by one-to-one interaction of two or a few genes. SI was originally classified into two types: (i) gametophytic SI (GSI) and (ii) sporophytic SI (SSI), based on the genetic control of the SI phenotype by pollen. A late-acting SI system (LSI) has also been described in several species recently (see review by Gibbs, 2014).

GSI has been described in Rosaceae, Plantaginaceae, Papaveraceae, and Solanaceae. According to the phylogeny and the conserved structure of the female S-RNase gene, it seems that this gene evolved only once, before the separation of the Asteridae and Rosidae, ~120 million years ago (Vieira et al., 2009). GSI has a common molecular basis across many plant families and is probably the ancestral condition for flowering plants (McClure, 2011). GSI has been most deeply characterized in Brassicaceae (Hiscock and Tabah, 2003; Kitashiba and Nasrallah, 2014; Lao et al., 2014; Iwano et al., 2015; Baldwin and Schoen, 2017). More recently, substantial advances have been made in describing SSI in Asteraceae (Gonthier et al., 2013; Koseva et al., 2017).

Finally, in LSI, both compatible and incompatible pollen grains can reach the ovary with a similar rate of pollen tube growth. However, whilst the double fertilization is completed after 24 h in compatible reactions, in incompatible reactions the male nuclei are released into the embryo sac, but they fail to achieve gamete fusion, resulting in floral abscission. LSI cases range from Malvaceae, Apocynaceae, and Bignoniaceae, to the monocot families Amaryllidaceae and Xanthorrhoeaceae (Gibbs, 2014). A well-characterized LSI case is Theobroma cacao (da Silva et al., 2016; Lanaud et al., 2017). Most of the T. cacao accessions are self-incompatible; however, some anciently domesticated varieties, such as ‘Criollo’ varieties from Central America originally cultivated by the Mayas, or ‘Comun’ from Brazil and ‘Nacional’ from Ecuador, are self-compatible (Loor Solorzano et al., 2012; Santos et al., 2015).

During domestication, SI has been introduced, or removed, according to agronomic parameters of interest (McClure, 2012). For example, in some fruit trees of the Rosaceae family, such as apple, Japanese pear, sweet cherry, or apricot, trees of different cross-compatible varieties should be planted to ensure fruit production due to SI (Sassa, 2016). Similar situations have been reported for other crops such as cabbage, chicory, or sugar beet (Ockendon, 1974; Broothaerts, 2003; Wünsch and Hörnaza, 2004; Hunt et al., 2010; Gonthier et al., 2013; Sassa, 2016; Saumitou-Laprade et al., 2017; Farinelli et al., 2018). An interesting example is the genus Prunus, a large genus in the Rosaceae family, that includes multiple domesticated crops such as almond, apricot, cherry, peach, and plum. Most of the Prunus species exhibit S-RNase-based gametophytic SI. In some species such as almond, domestication goals were exclusively focused on improving organolectic aspects such as reduced toxicity, thinner endocard, and increased seed size. In contrast, in other cases, such as peach, domestication was focused not only on improving aspects affecting fruit morphology but also on the introduction of SC (Miller and Gross, 2011). Other crops of the genus Prunus such as cherries and plums also have had SC introduced during their domestication (Spiegel-Roy, 1986). Modern breeding programs oriented towards disrupting SI in almond only began recently (Martínez-Gómez et al., 2006).

Other perennial species have evolved SC under domestication. For instance, wild grapevine is dioecious while the domesticated relative is hermaphrodite and self-compatible (De Mattia et al., 2008). In contrast to perennials where few crops derive from selfing wild populations, many annual crops have been domesticated from SC wild ancestors, such as barley, chickpea, eggplant, lentils, pea, chile, tomatoes, and wheat (Miller and Gross, 2011).

However, SI can also be a desirable trait for breeders. SI systems prevent self-fertilization, which forces outcrossing and increases genetic diversity, which is useful for the breeding of hybrid varieties of economically important plant families. Accordingly, breeding programs towards introducing functional SI have been activated in many crops (Kothen-Nakayama et al., 2010; Havlicková et al., 2014; Cheng et al., 2018; Xiao et al., 2019). In order to control mating, many advances have been made toward understanding the SI mechanisms. However, transferring these mechanisms across wide phylogenetic distances is often difficult, or even impossible, for breeders. Recently, de Graaf and co-workers introduced an SI system (from Papaver) in a species with no SI system that diverged ~140 million years ago (de Graaf et al., 2012), demonstrating that this transfer may be easier than previously thought.

Asexually (clonally) propagated crops where sexual reproduction is reduced have also been promoted by breeders. Vegetative propagation might be preferred to sexual reproduction either to avoid the segregation of traits in SI species (McKey et al., 2010) or to speed up breeding and growing cycles in perennials with long juvenile phases. For this reason, many tree crops such as avocado or olive trees are propagated clonally (Díez et al., 2015; Kuhn et al., 2019).

An interesting example of a herbaceous plant that is clonally propagated is banana. The banana domestication involved hybridizations between diverse species and subspecies that generated diploid and triploid sterile hybrids. Nonetheless, the hybrids are able to produce parthenocarpic fruits that have been thereafter dispersed by vegetative propagation (D’Hont et al., 2012).

Clonally propagated crops can potentially produce a wider range of adaptations with respect to sexual reproductive families that are easily maintained, but genetic homogeneity is an important drawback for survival if adverse conditions arise. For example, half of the 2018 banana world production relies on somaclones derived from a single triploid genotype (Cavendish) (Lescot, 2010). Since pests and diseases have gradually become adapted, at present this genetic homogeneity represents an imminent danger for global banana production (de Belleray et al., 2010; Dita et al., 2010).
Final remarks

Many of the changes selected during domestication are related to reproductive traits, either because the outcome of the reproductive process (seeds or fruits) is the desired result, or because the alteration of the reproductive process is necessary to achieve it (i.e. to be able to cross varieties/species or to avoid undesirable phenotype variation). In this review, we have summarized some of the changes associated with molecular mechanisms that human action have introduced into the reproductive structures and strategies of domesticated plants. Darwin already noticed many of the changes in reproductive structures and strategies introduced by domestication, but he lacked the tools to understand their basis. At present, many domestication-driven changes in plant reproduction are still not well understood because crop molecular biology research has developed more slowly than that based on model plants. Recently developed tools such as genome editing and next-generation sequencing are changing this scenario and quickly increasing our knowledge of the molecular basis of domestication in crop species. This will also allow for better planning of breeding strategies in the future.

Acknowledgements

We would like to acknowledge the FruitLook FP7-PEOPLE-2013-IRSES–Marie Curie Action (Project ID: 612640), the SexSeed H2020-EU.1.3.3. MSCA-RISE-Marie Skłodowska–Curie Research, Innovation Staff Exchange project (Project ID: 690946), H2020-MSCARISE Project (ExpoSeed GA-691109), and the Università degli Studi di Milano (UNIMI-RTD and Linea 2-DBS) for bringing together the team that developed this article. PG is grateful to Universitat Politècnica de València (PAID-10-18) and to the Japan Society for the Promotion of Science [FY2019 JSPS Postdoctoral Fellowship for Research in Japan (Standard)] for their respective postdoctoral grants. We would also like to thank the Society of Experimental Biology for the opportunity of summing our ideas in this article.

References


Darwin C. 1868. The variation of animals and plants under domestication. London: John Murray.


de Broothaerts W. 2003. New findings in apple S-genotype analysis resolve previous confusion and request the re-numbering of some S-alleles. Theoretical and Applied Genetics 106, 703–714.


Huq E, Tepperman JM, Quail PH. 2000. GIGANTEA is a nuclear protein involved in phytochrome signaling in Arabidopsis. Proceedings of the National Academy of Sciences, USA 97, 9789–9794.


