Molecular mechanisms involved in convergent crop domestication

Teresa Lenser and Günter Theißen

Department of Genetics, Friedrich Schiller University Jena, Philosophenweg 12, D-07743 Jena, Germany

Domestication has helped to understand evolution. We argue that, vice versa, novel insights into evolutionary principles could provide deeper insights into domestication. Molecular analyses have demonstrated that convergent phenotypic evolution is often based on molecular changes in orthologous genes or pathways. Recent studies have revealed that during plant domestication the causal mutations for convergent changes in key traits are likely to be located in particular genes. These insights may contribute to defining candidate genes for genetic improvement during the domestication of new plant species. Such efforts may help to increase the range of arable crops available, thus increasing crop biodiversity and food security to help meet the predicted demands of the continually growing global population under rapidly changing environmental conditions.

Crop domestication – a genetic perspective

During the process of crop domestication (see Glossary) and improvement, many morphological and physiological traits underwent dramatic modifications to meet the fastidious needs of humans (for recent reviews, see [1–5]). Some of these modifications, such as an increase in fruit number and size or a reduction of fruit abscission, were necessary to enable efficient cultivation by humans, whereas other changes were brought about merely to satisfy culinary or esthetic preferences, for example, alterations in fruit color, taste, or texture. Similar human demands led to similar adaptations of many domestication traits over a wide range of plant species (Figure 1), thereby providing numerous examples of convergent phenotypic evolution {for simplicity we use the term 'convergence' so as to include 'parallelism' following [6], even though the issue of how to define and determine convergent and parallel evolution is still actively discussed in the literature (see, e.g., [7])}. In this context, a key question is which genetic changes underlie this phenotypic convergence and, more precisely, whether it is mainly mutations at orthologous or distinct genomic loci that are involved. The term 'molecular convergence' is often used in this context if changes at orthologous loci underlie phenotypic convergence and we use it accordingly; however, care must be taken because in other studies the same term sometimes happens to be restricted to identical mutations within the same locus or expanded to include mutations within the same regulatory pathway [8,9]. For natural adaptation in animal systems it is already accepted that convergent phenotypic changes rely strikingly often on mutations at orthologous loci [6,9–11], and recently an elaborate meta-analysis gathering genetic loci of repeated evolution also found this to be the case in plant systems [12]. There has been some lively discussion regarding which factors favor the emergence of these so-called genetic hotspots [13–17]. By contrast, in crop domestication research, the importance of molecular convergence is still a controversial topic [18]. Although an often cited early study attached great importance to the role of molecular convergence [19], subsequent analyses have often dismissed or diminished its relevance because they have revealed some examples of non-orthologous loci at the basis of convergent phenotypic changes [4,20–22].

Glossary

Abscission: controlled shedding of plant organs (e.g., leaves, flowers, fruits) by means of cell–cell separation processes in specialized tissue layers (abscission zones).
Cis-regulatory mutation: mutation within a non-protein coding part of a gene influencing its level of expression.
Convergence: independent emergence of the same phenotypic trait in distinct lineages.
Determinate growth: naturally self-limited growth, for example, by the formation of a terminal flower at a predetermined developmental stage.
Domestication: process of genetic adaptation of wild species to human needs, typically including changes in appearance and lifestyle.
Fruit dehiscence: concerted process of fruit opening by cell–cell separation to enable seed dispersal.
Homology: similarity owing to common ancestry.
Indeterminate growth: plant growth continues indefinitely.
Molecular convergence: phenotypic convergence caused by changes at orthologous genetic loci.
Orthologs: genes that arose from a common ancestor sequence via speciation.
Parallelism: independent emergence of the same phenotypic trait from the same original phenotypic state in distinct lineages.
Paralogs: genes that arose from a common ancestor sequence via duplication.
Pleiotropic effects: phenomenon in which a single gene influences more than one phenotypic trait.
Polygenic: a phenotypic trait is controlled by more than one gene.
Quantitative trait locus (QTL): genomic region containing one or more genes contributing to phenotypic variation of a quantitative trait.
Standing genetic variation: a genomic locus has more than one allelic version within a population.
Stochastic process: non-deterministic, chance-dependent process whereby progression can only be predicted in terms of probability.
Vernalization: induction of flowering through exposure to low temperatures for a certain period of time.
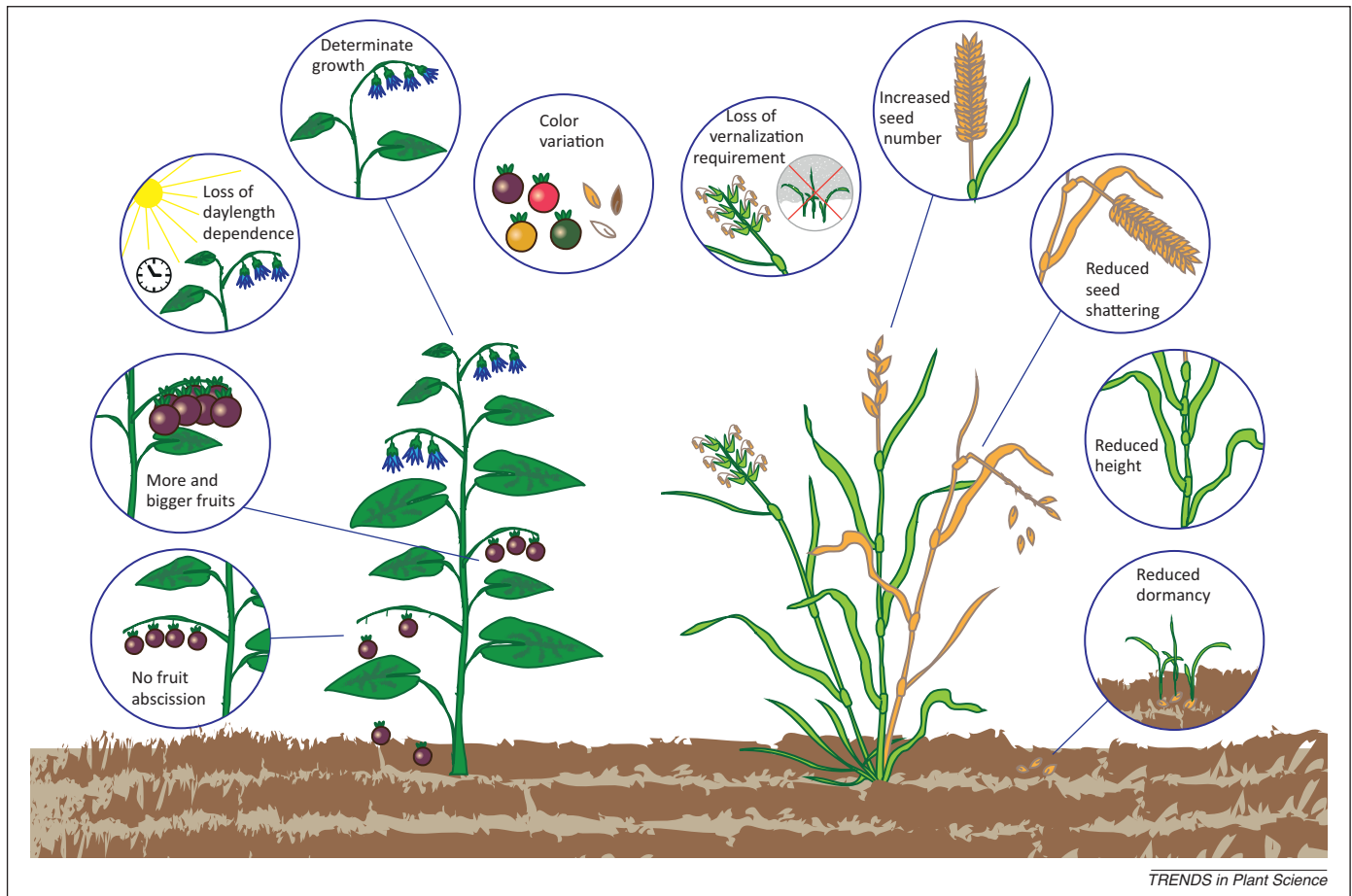
Figure 1. Convergent domestication. Convergent phenotypic changes are frequently observed in many different crops because systematic human cultivation often brings about similar demands. Attempts to maximize yield cause selective pressure for an increase in size and number of edible plant parts on the one hand and for a decrease in natural seed and fruit dispersal mechanisms to reduce yield loss on the other hand. Shifts in cultivation area often require changes in day length dependence or in the vernalization requirement and a reduction in seed dormancy is needed for synchronous germination. Small plants with a determinate growth habit are often selected because they are more robust, have a better yield to overall biomass ratio, and are better suited to mechanical harvesting methods. Finally, satisfying esthetic preferences often drives convergent adaptations, a prominent example being changes in color. Stylized examples of the major angiosperm plant lineages from which current crops originated are shown (eudicot, left; monocot, right) featuring traits of typical wild species. Characters that convergently evolved in various domesticated crops are depicted in circles.

In recent years, our knowledge about the genes involved in crop domestication has increased dramatically, enabling more in-depth questions to be asked regarding the molecular basis of domestication in a wide variety of species. In this review, we try to incorporate such recent molecular insights into the framework of genes that are already known to control domestication traits in plants. Unlike previous review articles [1,4,5], we do not distinguish between domestication and improvement genes because classification can be ambiguous. Instead, we equally consider all loci that have been artificially selected to discriminate crops from their wild ancestors as determining factors of domestication. We discuss recent findings that suggest that convergent molecular evolution played an important role in plant domestication and the suggestion that, as postulated for adaptive evolution, certain genes are particularly likely to become the target of domestication-relevant mutations. An understanding of the factors influencing this susceptibility in evolutionary biology might enable the likely course of molecular domestication to be predicted and, thus, might have great potential in the facilitation of future crop domestication and breeding procedures.

How molecular convergence contributed to crop domestication

Our knowledge of the genetic loci controlling diverse domestication phenotypes in crops is increasing. However, the picture is still far from complete, and the possibility of bias owing to the preferential investigation of candidate genes has to be taken into account when trying to evaluate the importance of molecular convergence in this context [13]. In this section, we concentrate on examining a few selected traits that have been characterized at the molecular level particularly well in several species in order to present a picture of the extent to which molecular convergence might contribute to shaping crop plants.

Plant growth

Controlling plant growth is an important aspect of domestication. Under the influence of systematic nitrogen fertilization, most wild species would grow excessively tall, making them more prone to damage by wind and rain. Moreover, the development of mechanized harvesting methods required the cultivation of plants of defined height and stature. Thus, many crop species were convergently selected for a determinate and 'dwarfing' growth

habit that increases yield at the expense of overall biomass. Although the number of effector loci as detected by quantitative trait locus (QTL) studies might differ greatly between species, plant height is a polygenic trait [23]. Some genes, such as *Ghd7* in rice (*Oryza sativa*) or *Q* in wheat (*Triticum aestivum*), have been reported to affect plant height in a certain species but to date have not been reported to influence growth habit in any other crop [4,24]. Despite this seeming plurality in domestication targets, two strategies have been repeatedly applied to cause dwarfing habit in different species: interference with hormone metabolism or signaling [25] and alteration of meristem identity causing determinate growth [26].

Loss-of-function mutations in two distinct genes of the gibberellic acid (GA) metabolic pathway cause dwarfism in rice. A change within the *ent*-kaurene oxidase gene *OsKO2* has been exclusively detected in a single Japanese cultivar [27], whereas at least nine distinct alleles of the *OsGA20 oxidase-2* (*OsGA20ox-2*) gene of varying severity were used for breeding semi-dwarfing varieties during both early domestication and the 'green revolution' in the 1960s [28,29]. The orthologous *HvGA20ox-2* gene from barley (*Hordeum vulgare*) has also been identified as an important semi-dwarfing gene in many barley varieties [30]. Contrasting this repeated occurrence of loss-of-function mutations disrupting GA biosynthesis in rice and barley, dwarfism in wheat is caused by independent gain-of-function mutations at two paralogous *Reduced height-B1/D1* loci coding GA-regulated growth repressors [1]. The mutant loci both produce proteins that lack an important GA signaling domain that normally directs GA-dependent degradation, thus becoming constitutively active growth repressors [25]. In addition to modulation of GA signaling, interference with brassinosteroid signal transduction and with polar auxin transport was shown to cause semi-dwarfing growth habit in agronomically important varieties of barley or sorghum (*Sorghum bicolor*) and pearl millet (*Pennisetum glaucum*), respectively [31–33].

To date, all the alleles that have been associated with the domestication-related switch from indeterminate to determinate growth habit are characterized by loss-of-function mutations in genes orthologous to the *Arabidopsis* (*Arabidopsis thaliana*) meristem identity gene *TERMINAL FLOWER 1* (*TFL1*) [34–36]. Furthermore, in soybean (*Glycine max*) and common bean (*Phaseolus vulgaris*), four and eight recessive alleles, respectively, were apparently selected independently [37,38], presenting an intriguing example of how orthologous genes can be repeatedly involved in convergent domestication of a given trait.
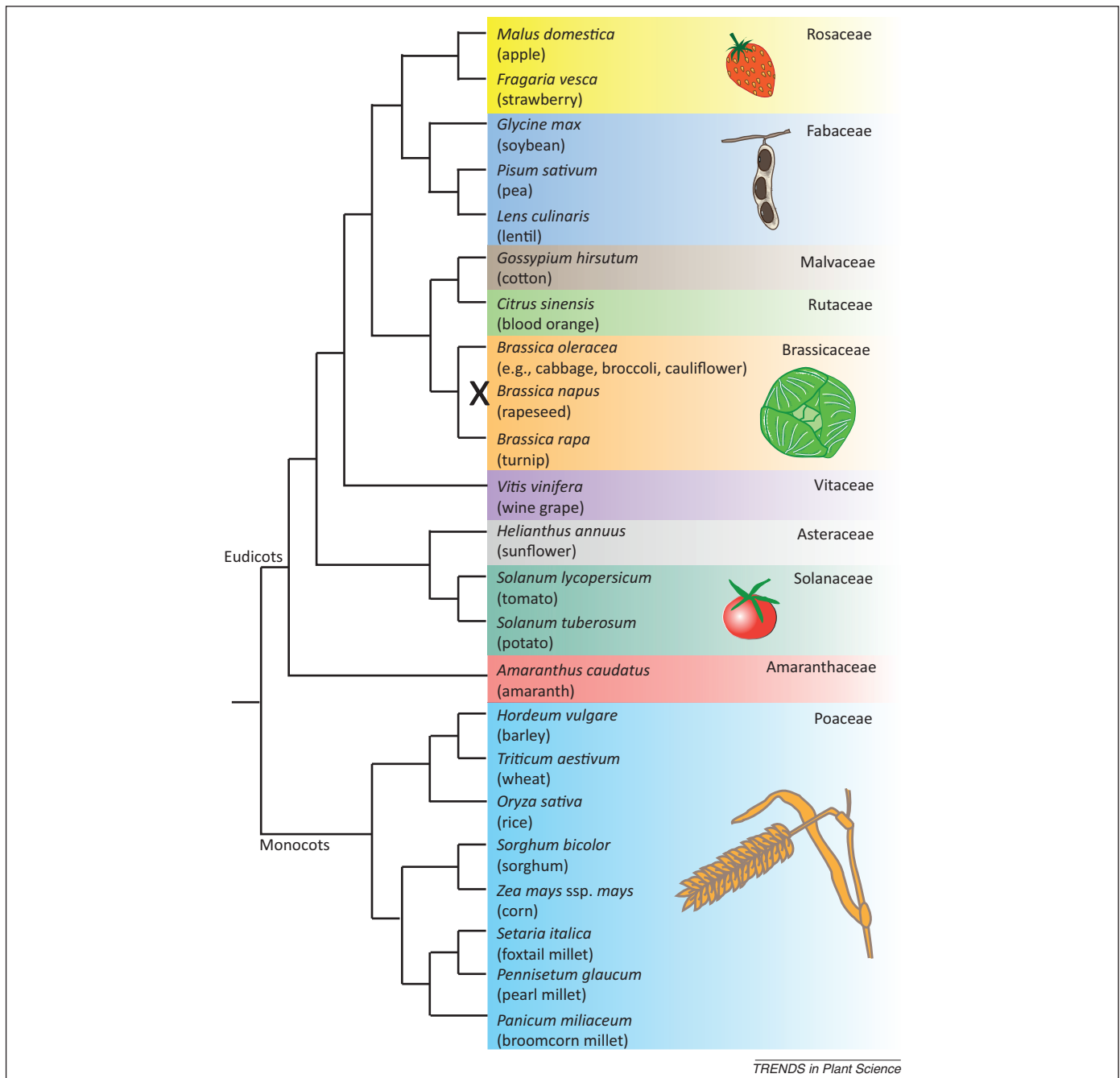
Flowering time
TFL1 orthologs are not only involved in defining the growth habit of a plant but have also been reported to induce early flowering in species as varied as barley, pea (*Pisum sativum*), and strawberry (*Fragaria vesca*) [39–41]. In the case of *TFL1*, this simultaneous influence on flowering time and plant height is probably due to its dual function in controlling the length of both the vegetative and floral phase [40]. Flowering time and plant height are also not independent of each other *per se*, because early flowering automatically reduces the time for vegetative

growth, thus leading to smaller plants, whereas late flowering increases it resulting in bigger plants. This is reflected by several domestication loci that likewise influence both traits, such as *Q* (wheat), *Ghd7* (rice), and possibly *Vgt1* in maize (*Zea mays*) [4,24,42,43]. Nevertheless, the size reduction resulting from early flowering clearly has to be distinguished from that of plants with dwarfing or determinate growth habit because it does not increase overall yield and has thus to be considered a mere side effect of the adaptation of flowering time.

Adaptation of flowering time is especially important for the global success of a particular crop species because expansion of the geographic range often also means a shift in climatic conditions and in the length of the photoperiod. A common effect observed in domesticated plants is that the onset of flowering becomes less dependent on environmental stimuli such as day length or vernalization, resulting in shorter growth cycles [39,44–46].

The molecular network controlling floral induction is rather complex and the comparative studies performed to date suggest that an ancient core pathway is conserved between distantly related plant lineages but has been modified by the recruitment of family-specific genes or pathways (see [47–49] and references therein). As a consequence of this limited conservation it has been noted that during adaptation of flowering time, molecular convergence occurs more often between closely related species (Figure 2) [49]. In the monocot lineage, for example, the *AP1*-like transcription factor (TF) and vernalization regulator gene *VRN1* was repeatedly found to induce a spring growth habit owing to dominant mutations disrupting potential repressor binding sites [49–51]. A possible candidate for the respective repressor is the zinc finger–CCT domain TF coding gene *VRN2*; loss-of-function mutations within this locus also cause a spring growth habit exclusively in monocots [49]. A member of the pseudo-response regulator (PRR) gene family in *Arabidopsis*, *PRR7*, is known to function in the circadian clock pathway [52]. Nevertheless, domestication-relevant alleles in closely related paralogs have only been reported in crops which are distantly related to *Arabidopsis*; in the closely related cereals rice [53], barley [54,55], wheat [56–58], and sorghum [45], mutations in *PRR37* orthologs influence flowering time by altering the photoperiod response, whereas in sugar beet (*Beta vulgaris* ssp. *vulgaris*) a paralogous *PRR* gene has been involved in the switch from annual to biennial growth habit [59]. By contrast, domestication-related flowering time variation by means of mutations in orthologs of *FLOWERING LOCUS C* (*FLC*) or its upstream regulator *FRIGIDA* (*FRI*) is typical for members of the Brassicaceae lineage [44,60–62].

Adaptation of flowering time in distantly related species has been reported to rely on mutations in genes that are non-homologous but that occupy analogous network positions within the floral induction pathway: for example, *FLC* and *VRN2*, which are both repressors of flowering and become downregulated by vernalization [49]. A similar non-homologous pair is *Ghd7* (rice) and *E1* (soybean), which are both photoperiod-dependent repressors of *FLOWERING LOCUS T* (*FT*) orthologs whose loss-of-function alleles induce early flowering [24,63]. However, some key genes also exist that convergently alter flowering time

Figure 2. Phylogenetic relationship of major crop plants. The topology of the tree is mainly based on data from [110] supplemented with the results of some additional recent studies. Family affiliation is indicated by colors. Branch lengths are not drawn to scale. The '×' represents a hybridization event that happened between *Brassica oleracea* and turnip (*Brassica rapa*) giving rise to rapeseed (*Brassica napus*).

in a more distantly related set of species. Besides the above-mentioned *TFL1* orthologs, loss-of-function mutations in orthologs of the *Arabidopsis* circadian clock regulator gene *EARLY FLOWERING 3* (*ELF3*) have been repeatedly involved in generating photoperiod-insensitive cultivars in grasses and legumes [46,64–66], and orthologs of the highly conserved floral inducer gene *FT* exhibit allelic variants that change flowering time across great phylogenetic distances [67–69].

Fruit and seed dispersal
Given that the evolutionary success of wild plant species relies heavily on their ability to spread their offspring,

most of them possess elaborate mechanisms to separate from their fruits and seeds. From an agronomic perspective, this is a highly undesirable trait that hampers harvest and causes considerable yield loss. As a result, there has been a convergent emergence of crop species with reduced seed dispersal capability. Fruit abscission or dehiscence is implemented by cell–cell separation at different anatomical structures, including the base of the spikelet in wheat and barley, the juncture between lemma and pedicel in rice, and the pod valve margins in soybean and *Brassica* species [22,70–72]. Thus, separation happens mostly at non-homologous structures, which appears to be associated with a high level of molecular variability.

Two tightly linked loci (*Btr1* and *Btr2*) controlling the brittle rachis phenotype have been identified in cultivated barley to date, and recessive mutations in either of these loci represent the independent emergence of shatter-resistant barley on two occasions [22,73]. In domesticated wheat species, three loci (*Br1–3*) bear recessive mutations conferring a loss of spikelet disarticulation [21,22]. None of these genes has been molecularly cloned yet and whether wheat and barley loci carry homologous genes is the subject of debate [21,22]. However, some data suggest that they are not orthologous to shattering loci identified in other grasses [22]. Equally unrelated to each other are the genes that suppress pedicel abscission in tomato (*Solanum lycopersicum*) (*jointless* [74]), pod dehiscence in soybean (*qPDH1* [70]), and the major non-shattering locus *sh4*, which is fixed in all rice cultivars [72,75]. Recently, a new shattering resistance locus was identified upstream of the rice gene *LIGULELESS* mediating the formation of closed panicles [76].

To date, only two instances of molecular convergence possibly underlying the domestication of shattering behavior have been reported. Three independent loss-of-function mutations within the YABBY-like TF gene *Sh1* were found to be responsible for a disruption in abscission layer formation and loss of seed shattering in domesticated sorghum; orthologous genes in rice and maize carry mutations that are similarly associated with shattering reduction [77]. Remarkably, single nucleotide substitutions at the same position in an upstream regulatory sequence of genes orthologous to the *Arabidopsis* gene *REPLUMLESS*, which are known to disrupt abscission layer formation in *japonica* rice were also found to alter fruit morphology in dehiscence-relevant tissues in different Brassicaceae species [78,79].

Why molecular domestication repeats itself

The question of whether or not crop plants develop similar phenotypic traits via changes at orthologous loci has been the subject of debate since investigators started to identify the genetic loci of domestication [4,18–22]. Research over the past few years has added a wealth of new genes to our list, thereby providing a solid base for further evaluation of this question. Surveying these data indicates that there is no clear-cut answer because convergent domestication at all phylogenetic levels might occur via mutations at orthologous or non-orthologous loci (Table 1). There are many examples that show that molecular convergence plays an important role during domestication (Table 1) [12]; however, why are certain genes more likely to become involved in domestication than others? To define factors that promote convergent molecular crop domestication (Figure 3), it is worth having a more in-depth look at convergent evolution because this topic has already been examined to some extent in the context of natural adaptation in animal systems [10,12,14–16].

Nodal positioning

Nodal positioning of a TF protein within a certain regulatory network has been reported to increase the probability of it becoming involved in natural adaptation [13–15]. Such evolutionary hotspot genes typically collect regulatory input from several upstream regulators and control a whole set of target genes necessary to guide the formation of an entire developmental module, such as a floral organ. Mutations in upstream regulators or downstream targets often alter only specific aspects within the developmental module, thus making the nodal gene particularly predestined for changes concerning the entire module. In addition, changes in upstream regulators might result in pleiotropic effects that might be deleterious for the organism (see section on 'minimal pleiotropic effects' below). Orthologs of the floral homeotic C-class gene *AGAMOUS* (*AG*) occupy such a nodal position in terms of controlling the formation of stamens, carpels, and determinate flowers across all angiosperms [80] (Figure 3). Changes in *AG* expression are involved in all cases of domesticated plants with double flowers that have been molecularly characterized to date, although the causative mutations have only been identified in two of these studies as loss-of-function mutations within the *AG* locus: in the ornamental 'Double White' cultivar of Rue anemone (*Thalictrum thalictroides*) and the Japanese cherry (*Prunus lannesiana*) variant 'Albo rosea' [80–83]. Double flower phenotypes have been convergently selected in various ornamental plants because the numerous extra petals increase their attractiveness for humans. As a link between upstream sensory modules and floral induction, *FT* orthologs likewise take up a nodal position in the gene regulatory network leading to flowering, which might explain their repeated involvement in flowering time variation [67–69,84]. However, given that *FT* genes are not the only mutational hotspots for flowering time manipulation, other factors have to be considered when evaluating predisposition for convergent molecular domestication, such as pathway size and complexity.

Simple pathways

Simple pathways involving only a few gene products offer limited targets for domestication-relevant mutations unlike complex pathways with a multitude of players, such as the flowering time pathway. This is likely to reflect back on the probability of orthologous genes becoming involved in the repeated occurrence of a certain phenotype [16]. Glutinous cereal varieties whose seeds develop a sticky texture when cooked were domesticated repeatedly in certain areas because of cultural culinary preferences. The glutinous character derives from seed endosperms with a reduced content of amylose, a starch molecule that is produced from a glucose precursor by a single catalytic process mediated by a granule-bound starch synthase encoded by the *Waxy* locus (Figure 3) [85]. In all studies reported to date, the glutinous character of cereal varieties has been brought about by mutations in *Waxy* orthologs [85–90]. Likewise, the loss of betaine aldehyde dehydrogenase (BADH2) enzyme activity in a metabolic pathway leading to the aroma compound 2-acetyl-1-pyrroline has repeatedly caused the occurrence of fragrant varieties in rice and soybean [91,92].

Minimal pleiotropic effects

Minimal pleiotropic effects of adaptive mutations have been reported to raise their chances of becoming evolutionarily fixed [10,12–15]. Not only might this requirement be

Table 1. Examples of molecular convergence underlying domestication-related phenotypic changes[a]

Crop species	Phylogenetic distribution[c]	Orthologous gene(s)	Class of gene product	Phenotypic effect	Causative changes[d]	Refs
Rice[b], barley	Species/family	*OsGA20ox-2* (*GA20 oxidase-2*), *HvGA20ox-2*	Metabolic enzyme	Dwarfism	Coding	[28–30]
Wheat[b]	Species	*Rht-1* (*reduced height-1*)	SH2-TF	Dwarfism	Coding	[1][e]
Sorghum, pearl millet	Family	*dw3* (*dwarfing3*), *d2*	Transporter protein	Dwarfism	Coding	[32,33]
Tomato, soybean[b], common bean[b]	Family/above family	*SP* (*SELF-PRUNING*), *Dt1* (*determinate stem locus 1*), *PvTFL1y* (*TERMINAL FLOWER 1*)	Signaling protein	Determinate growth	Coding	[1][e], [35–38]
Barley, pea, strawberry	Above family	*HvCEN* (*CENTRORADIALIS*), *PsTFL1c*, *FvTFL1*	Signaling protein	Variation in flowering time	Mixed	[39–41]
Barley, wheat[b], ryegrass (*Lolium perenne*)	Species/family	*VRN1* (*BM5*, *TmAP1*, *WAP1*, *LpVRN1*)	MADS domain TF	Variation in flowering time	Non-coding	[51][e]
Barley, wheat[b]	Species/family	*VRN2* (*ZCCT1*)	Zinc finger–CCT domain TF	Variation in flowering time	Mixed	[49][e]
Rice, barley, wheat, sorghum[b], sugar beet	Species/family/ above family	*OsPRR37* (*pseudoresponse regulator protein 37*), *Ppd-H1*, *Ppd1*, *SbPRR37*, *BvBTC1*	Regulator of the circadian clock pathway	Variation in flowering time	Mixed	[53–58]
Turnip, *Brassica oleracea*	Family	*BrFLC2* (*FLOWERING LOCUS C*), *BoFLC2*	MADS domain TF	Variation in flowering time	Mixed	[44,60,62]
Rice, barley, pea, lentil	Family/above family	*Hd17* (*Heading date 17*), *EAM8* (*EARLY MATURITY 8*)/*Mat-a* (*Praematurum-a*), *HR* (*HIGH RESPONSE TO PHOTOPERIOD*), *LcELF3* (*EARLY FLOWERING 3*)	Regulator of the circadian clock pathway	Variation in flowering time	Coding	[46,64–66]
Rice, wheat, sunflower, barley	Family/above family	*Hd3a* (*Heading date 3a*), *VRN3*/ *TaFT* (*FLOWERING LOCUS T*), *HaFT1*, *HvFT*	Signaling protein	Variation in flowering time	Mixed	[67–69]
Rice[b]	Species	*Hd1* (*Heading date 1*)	Zinc finger TF	Variation in flowering time	Coding	[12][e]
Sorghum, rice, corn	Family	*Sh1* (*Shattering 1*), *OsSh1*, *ZmSh1*	YABBY-like TF	Shatter resistance	Mixed	[77]
Rice, wheat[b], corn[b], foxtail millet[b], barley[b], amaranth, sorghum[b], broomcorn millet	Species/family/ above family	*GBSSI* (*granule-bound starch synthase I*)/*Waxy*	Metabolic enzyme	Glutinous seeds	Mixed	[85][e], [86–90]
Rice[b], soybean	Species/family	*BADH2* (*betaine aldehyde dehydrogenase gene 2*), *GmBADH2*	Metabolic enzyme	Fragrance	Coding	[91,92]
Rice[b], potato	Species/above family	*Rd/DFR* (*dihydroflavonol-4-reductase*), *DFR*	Metabolic enzyme	Coloration	Coding	[111,112]
Blood orange[b]	Species	*Ruby*	MYB-TF	Coloration	Non-coding	[97]
Rice[b]	Species	*Bh4* (*Black hull4*)	Transporter protein	Coloration	Coding	[113]
Soybean[b]	Species	*R*	MYB-TF	Coloration	Coding	[96]
Pea[b], potato	Above family	*F3'5'H* (*flavonoid 3',5'-hydroxylase*)	Metabolic enzyme	Coloration	Mixed	[12][e]
Rice[b]	Species	*Rc*	bHLH-TF	Coloration	Coding	[12][e]
Grapevine[b]	Species	*VvMYBA1-3*	MYB-TF	Coloration	Mixed	[12][e]
Corn, pearl millet, barley	Family	*tb1* (*teosinte branched 1*), *Pgtb1*, *INT-C* (*INTERMEDIUM-C*)	TCP-TF	Plant architecture	Mixed	[101,114,115]
Barley[b]	Species	*VRS1* (*six-rowed spike 1*)	Homeodomain-TF	Plant architecture	Coding	[12][e]
Rice, corn	Family	*GS3* (*QTL for grain size and length on chromosome 3*), *ZmGS3*	Putative transmembrane protein	Grain size	Mixed	[4][e], [116]
Rice[b]	Species	*GS5* (*QTL for grain size and length on chromosome 5*)	Metabolic enzyme	Grain size	Non-coding	[117]
Rice, corn, wheat	Family	*GW2* (*QTL for grain weight on chromosome 2*), *ZmGW2-CHR4/ 5*, *TaGW2*	Metabolic enzyme	Grain size	Mixed	[4][e], [118,119]
Rice[b], wheat	Species/family	*Gn1a* (*QTL for grain number on chromosome 1, a*)/*OsCKX2* (*cytokinin oxidase/ dehydrogenase*), *TaCKX6-D1*	Metabolic enzyme	Grain number	Mixed	[4][e], [120]
Corn[b]	Species	*Opaque2*	bZIP-TF	Grain quality	Mixed	[12][e]

Table 1 (*Continued*)

Crop species	Phylogenetic distribution[e]	Orthologous gene(s)	Class of gene product	Phenotypic effect	Causative changes[d]	Refs
Rice[b]	Species	*GW8* (*QTL for grain weight on chromosome 8*)/*OsSPL16* (*squamosa promoter-binding protein-like 16*)	SBP-TF	Grain size and shape	Non-coding	[12][e]
Wheat, rye (*Secale cereale*)	Family	*TaALMT1* (*Al-activated malate transporter 1*), *ScALMT1*	Transporter protein	Metal tolerance	Mixed	[12][e]
Sorghum, corn	Family	*SbMATE1* (*multidrug and toxic compound extrusion 1*), *ZmMATE1*	Transporter protein	Metal tolerance	Mixed	[12][e], [121]

[a]This list is not intended to be exhaustive.

[b]Multiple independent alleles present in this species.

[c]Species: molecular convergence within the same species, Family: molecular convergence between plants of one plant family, Above family: molecular convergence between species of different plant families.

[d]Discrimination between mutations detected exclusively in coding regions (coding), non-coding regions (non-coding) or both (mixed).

[e]Review containing original references.

fulfilled by the preferential occurrence of *cis*-regulatory mutations [13,14,93] but also by the preferential usage of mutational target genes of limited functionality [94]. The flavonoid biosynthetic pathway produces anthocyanins, thereby giving rise to blue, purple, and red coloration of plant organs among angiosperms [94,95]. Three types of TFs coordinate the activity of this pathway and it has been argued that owing to their high copy number and often tissue-specific function, evolutionary changes in members of the R2R3-MYB (MYB) TF family are likely to have fewer pleiotropic effects compared with basic helix–loop–helix (bHLH) or WD40 repeat family members, which function more broadly (Figure 3) [94,95]. Indeed, a preference for mutations in MYB-TF genes has been observed in studies dealing with the natural adaptation of floral pigment intensity [94]. During plant domestication, changes in
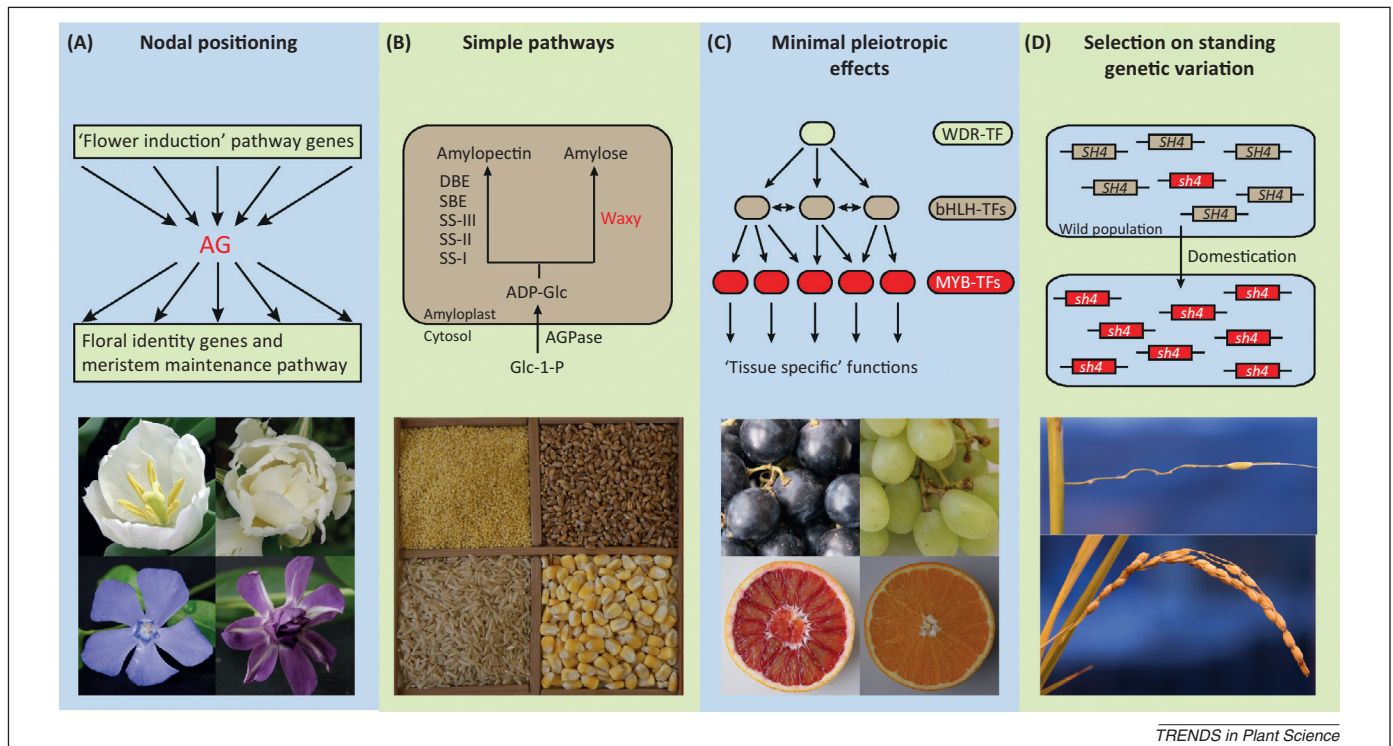
Figure 3. Factors promoting convergent molecular domestication. Convergent phenotypic changes in the course of plant domestication are often caused by mutations within orthologous genes. There are several factors explaining why mutations in such hotspot genes are more likely to become fixed in a population than mutations in other genes. (A) Genes occupying nodal positions within a given regulatory pathway incorporate inputs of several upstream regulators and in turn regulate several downstream genes, thereby often controlling self-contained developmental units. Mutations within such an input–output gene might alter a parameter value in a way that could otherwise only be achieved by concerted mutations within several upstream or downstream genes simultaneously. Mutations involving the nodal gene *AGAMOUS* (*AG*) were found at the basis of all molecularly characterized cases of domesticated plants with double flowers. (B) Simple metabolic pathways might also favor convergent molecular changes because only a minimal set of genes serves as a potential mutational target to change a given trait. Low amylose content because of mutations within *Waxy* gene orthologs is responsible for the domestication of glutinous seeds in many cereal variants. (C) Changes in fruit or seed color are often caused by mutations within MYB transcription factors because within the anthocyanin pathway they mostly have tissue-specific functions, thereby minimizing pleiotropic effects. (D) Finally, if domestication-related alleles are already present at low frequency within a wild population, as is the case for the non-shattering allele *sh4* in rice, independent selection on this standing genetic variation is likely to drive the same allele towards fixation repeatedly. Abbreviations: bHLH, basic helix–loop–helix; DBE, starch debranching enzyme; Glc, glucose; SBE, starch branching enzyme; SS, starch synthase; TF, transcription factor; WDR, WD40 repeat.

plant coloration, particularly of edible parts, also play an important role and molecular data likewise point towards mutations in MYB-TF genes as the overrepresented cause, probably to avoid unwanted pleiotropic effects [95–98]. In the same vein, the superior qualities and domestication-related success of *OsGA20ox-2* compared with *OsKO2* mutations to induce a semi-dwarfing growth habit in rice has been attributed to a difference between the pleiotropic effects of both genes [27,28]. Consequences of a defective *OsGA20ox-2* gene are restricted to stem tissue because in all other plant parts redundancy with other genes compensates for the impaired gene function, whereas in the case of *OsKO2* expression of the redundant *OsKO1* is restricted to flower organs, thus resulting in pleiotropic effects in the rest of the plant [27]. Furthermore, difficulties in breeding rape (*Brassica napus*) varieties that were resistant to pod shatter have been hypothesized to derive from fruit developmental genes acting pleiotropically in anther development [99].

Selection on standing genetic variation
Selection on the standing genetic variation might cause the repeated involvement of the same locus in independent domestication events because the selection of favorable alleles that are already present in a wild population usually proceeds faster than new mutations can arise (Figure 3) [9–12,15,100]. Naturally, this only applies in some cases because domestication-related alleles might have strong deleterious effects on wild plants; however, various domestication-related alleles with a moderate negative effect on plant fitness are known to be present in wild populations at low frequency [39,55,72,101]. Additionally, introgression of wild alleles of adaptive value into already established crop species was recently discussed as another mechanism for the evolution of adaptive changes during domestication [102]. Thus, the potential to form viable alleles with altered functionality in a wild plant population might be viewed as an additional factor designating certain genes as likely domestication targets.

Having shed light on some factors that promote molecular convergence, another important variable acting in addition with these factors seems to be phylogenetic distance between species. It was already noted elsewhere that mutations at orthologous loci cause convergent evolution more frequently within the same or between closely related species compared with distantly related ones [6,11,16,49]. Concentrating on data about plant domestication (dataset described in [12] and Table 1 from this study), we observe the same trend counting less than ten cases of molecular convergence involving species from different plant families, whereas more than 50 cases were reported for species belonging to the same family. This higher frequency of molecular convergence is probably because of higher levels of similarity on average between close relatives, both at a molecular and morphological level, thus providing more possible targets for convergent mutations affecting domestication traits. With regard to morphology, non-homologous structures at the base of a domestication trait (as for seed dispersal in soybean and rice [70,72]) should diminish (although not remove) the probability of orthologous genes controlling this trait. However, the more evolutionary

distance there is separating the respective species the more likely it is that even homologous traits are governed by a divergent set of genes (as seen for control of flowering time) [11,47–49,103]. In extreme cases, orthologs of a domestication-related gene in one species may even be absent in other species, making molecular convergence impossible. However, in most cases that we have compiled (Table 1) where different genes underlie the same domestication trait in distinct plants orthologous genes are nevertheless present in all species considered, suggesting that absence of orthologs is usually not the reason for non-convergent molecular evolution. In any case, the phenomenon of molecular convergence decreases with increasing phylogenetic distance, although this can only be considered as a probability statement given that domestication, like evolution, proceeds largely as a stochastic process.

From evolution back to domestication – on the origin of new crop species
Domestication has inspired evolutionary biologists: famous examples of domestication are outlined in Charles Darwin's seminal book *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* [104]. One of the most intriguing phenomena studied by evolutionary biology is convergent evolution, which is often based on mutations in closely related genes. In this review, we have compiled recent evidence showing that convergent evolution also applies to plant domestication and is governed by similar factors. Thus, knowledge about evolution has led to a better understanding of domestication.

However, not only can our improved understanding of domestication at the molecular level be used to improve existing crops but it can also be used to facilitate the domestication of new crops. Although several thousand plant species have been cultivated for consumption during the course of human history, at present 95% of human food energy is derived from only approximately 30 crop species (http://www.fao.org/biodiversity/components/plants/en/). Relying on this small basis of crop biodiversity seems risky and we may need more or better adapted crops to meet the challenges posed by plant pests, global warming, and the growing human population. Classical domestication and breeding is a slow process and may be too slow to meet the predicted demands of the global population. Therefore, it is good news that domestication is based on genetic hotspots and, thus, is predictable to a certain degree. Rather than starting domestication of wild plant species in a naïve way from scratch, promising hotspot genes (e.g., *TFL1* orthologs for determinate growth, *ELF3* orthologs for early flowering) might be targeted via marker-assisted breeding or transgenic technology to generate desired phenotypic changes. In principle, this rationale has already been used for crop improvement in some straightforward cases. For example, downregulating the expression of *Waxy* orthologs has been used to artificially generate potato (*Solanum tuberosum*), sweet potato (*Ipomoea batatas*), and cassava (*Manihot esculenta*) lines containing optimized starches for culinary and industrial applications [105–107]. Other desirable traits appear to be more complex or labile. For example, although salt tolerance has evolved more than

70 times in a wide range of grass species, few commercially viable salt tolerant crops have been released [108]. However, there is evidence that genes encoding membrane transporters of the HKT family (transporting sodium and potassium) can be used to generate salt tolerant cereal grasses (reviewed in [109]). Knowing why certain genes are repeatedly successfully involved in domestication might help to predict the most promising mutational targets from such newly discovered genes or pathways.

In conclusion, domestication has been a tremendous help in understanding evolution. However, now the time is ripe to use the knowledge gleaned from evolutionary biology to adapt domesticated plants to our changing environment.

References

1 Doebley, J.F. *et al.* (2006) The molecular genetics of crop domestication. *Cell* 127, 1309–1321
2 Miller, A.J. (2007) Crop plants: evolution. In *eLS* (Hetherington, A.M., ed.), pp. 1–7, John Wiley & Sons
3 Purugganan, M.D. and Fuller, D.Q. (2009) The nature of selection during plant domestication. *Nature* 457, 843–848
4 Gross, B.L. and Olsen, K.M. (2010) Genetic perspectives on crop domestication. *Trends Plant Sci.* 15, 529–537
5 Olsen, K.M. and Wendel, J.F. (2013) A bountiful harvest: genomic insights into crop domestication phenotypes. *Annu. Rev. Plant Biol.* 64, 47–70
6 Arendt, J. and Reznick, D. (2008) Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* 23, 26–32
7 Scotland, R.W. (2011) What is parallelism? *Evol. Dev.* 13, 214–227
8 Yoon, H.S. and Baum, D.A. (2004) Transgenic study of parallelism in plant morphological evolution. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6524–6529
9 Elmer, K.R. and Meyer, A. (2011) Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol. Evol.* 26, 298–306
10 Nadeau, N.J. and Jiggins, C.D. (2010) A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends Genet.* 26, 484–492
11 Conte, G.L. *et al.* (2012) The probability of genetic parallelism and convergence in natural populations. *Proc. R. Soc. B: Biol. Sci.* 279, 5039–5047
12 Martin, A. and Orgogozo, V. (2013) The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* 67, 1235–1250
13 Stern, D.L. and Orgogozo, V. (2008) The loci of evolution: How predictable is genetic evolution? *Evolution* 62, 2155–2177
14 Stern, D.L. and Orgogozo, V. (2009) Is genetic evolution predictable? *Science* 323, 746–751
15 Gompel, N. and Prud'homme, B. (2009) The causes of repeated genetic evolution. *Dev. Biol.* 332, 36–47
16 Christin, P.A. *et al.* (2010) Causes and evolutionary significance of genetic convergence. *Trends Genet.* 26, 400–405
17 Lobkovsky, A.E. and Koonin, E.V. (2012) Replaying the tape of life: quantification of the predictability of evolution. *Front. Genet.* 3, 246
18 Tang, H.B. *et al.* (2010) Domestication and plant genomes. *Curr. Opin. Plant Biol.* 13, 160–166
19 Paterson, A.H. *et al.* (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic-loci. *Science* 269, 1714–1718
20 Sood, S. *et al.* (2009) The major threshability genes soft glume (*sog*) and tenacious glume (*Tg*), of diploid and polyploid wheat, trace their origin to independent mutations at non-orthologous loci. *Theor. Appl. Genet.* 119, 341–351
21 Li, W. and Gill, B. (2006) Multiple genetic pathways for seed shattering in the grasses. *Funct. Integr. Genomics* 6, 300–309
22 Sang, T. (2009) Genes and mutations underlying domestication transitions in grasses. *Plant Physiol.* 149, 63–70
23 Salas Fernandez, M.G. *et al.* (2009) From dwarves to giants? Plant height manipulation for biomass yield. *Trends Plant Sci.* 14, 454–461
24 Xue, W.Y. *et al.* (2008) Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat. Genet.* 40, 761–767
25 Salamini, F. (2003) Hormones and the green revolution. *Science* 302, 71–72
26 McGarry, R.C. and Ayre, B.G. (2012) Manipulating plant architecture with members of the CETS gene family. *Plant Sci.* 188, 71–81
27 Itoh, H. *et al.* (2004) A rice semi-dwarf gene, *Tan-Ginbozu* (D35), encodes the gibberellin biosynthesis enzyme, *ent*-kaurene oxidase. *Plant Mol. Biol.* 54, 533–547
28 Asano, K. *et al.* (2007) Genetic and molecular analysis of utility of *sd1* alleles in rice breeding. *Breed. Sci.* 57, 53–58
29 Asano, K. *et al.* (2011) Artificial selection for a green revolution gene during *japonica* rice domestication. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11034–11039
30 Jia, Q.J. *et al.* (2009) GA-20 oxidase as a candidate for the semidwarf gene *sdw1/denso* in barley. *Funct. Integr. Genomics* 9, 255–262
31 Chono, M. *et al.* (2003) A semidwarf phenotype of barley uzu results from a nucleotide substitution in the gene encoding a putative brassinosteroid receptor. *Plant Physiol.* 133, 1209–1219
32 Multani, D.S. *et al.* (2003) Loss of an MDR transporter in compact stalks of maize *br2* and sorghum *dw3* mutants. *Science* 302, 81–84
33 Parvathaneni, R.K. *et al.* (2013) Fine-mapping and identification of a candidate gene underlying the *d2* dwarfing phenotype in pearl millet, *Cenchrus americanus* (L.) Morrone. *G3 (Bethesda)* 3, 563–572
34 Pnueli, L. *et al.* (1998) The *SELF-PRUNING* gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of *CEN* and *TFL1*. *Development* 125, 1979–1989
35 Repinski, S.L. *et al.* (2012) The common bean growth habit gene *PvTFL1y* is a functional homolog of *Arabidopsis TFL1*. *Theor. Appl. Genet.* 124, 1539–1547
36 Liu, B.H. *et al.* (2010) The soybean stem growth habit gene *Dt1* Is an ortholog of *Arabidopsis TERMINAL FLOWER1*. *Plant Physiol.* 153, 198–210
37 Kwak, M. *et al.* (2012) Multiple origins of the determinate growth habit in domesticated common bean (*Phaseolus vulgaris*). *Ann. Bot.* 110, 1573–1580
38 Tian, Z.X. *et al.* (2010) Artificial selection for determinate growth habit in soybean. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8563–8568
39 Comadran, J. *et al.* (2012) Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.* 44, 1388–1392
40 Foucher, F. *et al.* (2003) *DETERMINATE* and *LATE FLOWERING* are two *TERMINAL FLOWER1/CENTRORADIALIS* homologs that control two distinct phases of flowering initiation and development in pea. *Plant Cell* 15, 2742–2754
41 Koskela, E.A. *et al.* (2012) Mutation in *TERMINAL FLOWER1* reverses the photoperiodic requirement for flowering in the wild strawberry *Fragaria vesca*. *Plant Physiol.* 159, 1043–1054
42 Salvi, S. *et al.* (2011) Genetic dissection of maize phenology using an intraspecific introgression library. *BMC Plant Biol.* 11, 4
43 Salvi, S. *et al.* (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11376–11381
44 Wu, J. *et al.* (2012) A naturally occurring InDel variation in *BraA.FLC.b* (*BrFLC2*) associated with flowering time variation in *Brassica rapa*. *BMC Plant Biol.* 12, 151
45 Murphy, R.L. *et al.* (2011) Coincident light and clock regulation of *pseudoresponse regulator protein 37* (*PRR37*) controls photoperiodic flowering in sorghum. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16469–16474
46 Weller, J.L. *et al.* (2012) A conserved molecular basis for photoperiod adaptation in two temperate legumes. *Proc. Natl. Acad. Sci. U.S.A.* 109, 21158–21163

47 Higgins, J.A. *et al.* (2010) Comparative genomics of flowering time pathways using *Brachypodium distachyon* as a model for the temperate grasses. *PLoS ONE* 5, e10065

48 Jung, C. and Muller, A.E. (2009) Flowering time control and applications in plant breeding. *Trends Plant Sci.* 14, 563–573

49 Alonso-Blanco, C. *et al.* (2009) What has natural variation taught us about plant development, physiology, and adaptation? *Plant Cell* 21, 1877–1896

50 Trevaskis, B. *et al.* (2003) MADS box genes control vernalization-induced flowering in cereals. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13099–13104

51 Asp, T. *et al.* (2011) Comparative sequence analysis of *VRN1* alleles of *Lolium perenne* with the co-linear regions in barley, wheat, and rice. *Mol. Genet. Genomics* 286, 433–447

52 Salome, P.A. and McClung, C.R. (2005) *PSEUDO-RESPONSE REGULATOR 7* and *9* are partially redundant genes essential for the temperature responsiveness of the *Arabidopsis* circadian clock. *Plant Cell* 17, 791–803

53 Murakami, M. *et al.* (2005) Circadian-associated rice pseudo response regulators (*Os*PRRs): insight into the control of flowering time. *Biosci. Biotechnol. Biochem.* 69, 410–414

54 Turner, A. *et al.* (2005) The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* 310, 1031–1034

55 Jones, H. *et al.* (2008) Population-based resequencing reveals that the flowering time adaptation of cultivated barley originated east of the fertile crescent. *Mol. Biol. Evol.* 25, 2211–2219

56 Beales, J. *et al.* (2007) A *Pseudo-Response Regulator* is misexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 115, 721–733

57 Wilhelm, E.P. *et al.* (2009) Photoperiod insensitive *Ppd-A1a* mutations in tetraploid wheat (*Triticum durum* Desf.). *Theor. Appl. Genet.* 118, 285–294

58 Diaz, A. *et al.* (2012) Copy number variation affecting the *Photoperiod-B1* and *Vernalization-A1* genes is associated with altered flowering time in wheat (*Triticum aestivum*). *PLoS ONE* 7, e33234

59 Pin, P.A. *et al.* (2012) The role of a pseudo-response regulator gene in life cycle adaptation and domestication of beet. *Curr. Biol.* 22, 1095–1101

60 Yuan, Y.X. *et al.* (2009) A naturally occurring splicing site mutation in the *Brassica rapa FLC1* gene is associated with variation in flowering time. *J. Exp. Bot.* 60, 1299–1308

61 Wang, N.A. *et al.* (2011) Flowering time variation in oilseed rape (*Brassica napus* L.) is associated with allelic variation in the *FRIGIDA* homologue *BnaA.FRI.a*. *J. Exp. Bot.* 62, 5641–5658

62 Okazaki, K. *et al.* (2007) Mapping and characterization of *FLC* homologs and QTL analysis of flowering time in *Brassica oleracea*. *Theor. Appl. Genet.* 114, 595–608

63 Xia, Z.J. *et al.* (2012) Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2155–E2164

64 Matsubara, K. *et al.* (2012) Natural variation in *Hd17*, a homolog of *Arabidopsis ELF3* that is involved in rice photoperiodic flowering. *Plant Cell Physiol.* 53, 709–716

65 Zakhrabekova, S. *et al.* (2012) Induced mutations in circadian clock regulator *Mat-a* facilitated short-season adaptation and range extension in cultivated barley. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4326–4331

66 Faure, S. *et al.* (2012) Mutation at the circadian clock gene *EARLY MATURITY 8* adapts domesticated barley (*Hordeum vulgare*) to short growing seasons. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8328–8333

67 Yan, L. *et al.* (2006) The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19581–19586

68 Takahashi, Y. *et al.* (2009) Variations in Hd1 proteins, *Hd3a* promoters, and *Ehd1* expression levels contribute to diversity of flowering time in cultivated rice. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4555–4560

69 Blackman, B.K. *et al.* (2010) The role of recently derived *FT* paralogs in sunflower domestication. *Curr. Biol.* 20, 629–635

70 Suzuki, M. *et al.* (2010) Fine mapping and development of DNA markers for the *qPDH1* locus associated with pod dehiscence in soybean. *Mol. Breed.* 25, 407–418

71 Spence, J. *et al.* (1996) 'Pod shatter' in *Arabidopsis thaliana*, *Brassica napus* and *B. juncea*. *J. Microsc.* 181, 195–203

72 Lin, Z.W. *et al.* (2007) Origin of seed shattering in rice (*Oryza sativa* L.). *Planta* 226, 11–20

73 Azhaguvel, P. and Komatsuda, T. (2007) A phylogenetic analysis based on nucleotide sequence of a marker linked to the brittle rachis locus indicates a diphyletic origin of barley. *Ann. Bot.* 100, 1009–1015

74 Mao, L. *et al.* (2000) *JOINTLESS* is a MADS-box gene controlling tomato flower abscission zone development. *Nature* 406, 910–913

75 Li, C.B. *et al.* (2006) Rice domestication by reducing shattering. *Science* 311, 1936–1939

76 Ishii, T. *et al.* (2013) *OsLG1* regulates a closed panicle trait in domesticated rice. *Nat. Genet.* 45, 462–465

77 Lin, Z.W. *et al.* (2012) Parallel domestication of the *Shattering1* genes in cereals. *Nat. Genet.* 44, 720–724

78 Konishi, S. *et al.* (2006) A SNP caused the loss of seed shattering during rice domestication. *Plant Cell Physiol.* 47, S14

79 Arnaud, N. *et al.* (2011) The same regulatory point mutation changed seed-dispersal structures in evolution and domestication. *Curr. Biol.* 21, 1215–1219

80 Galimba, K.D. *et al.* (2012) Loss of deeply conserved C-class floral homeotic gene function and C- and E-class protein interaction in a double-flowered ranunculid mutant. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2267–E2275

81 Dubois, A. *et al.* (2010) Tinkering with the C-function: a molecular frame for the selection of double flowers in cultivated roses. *PLoS ONE* 5, e9288

82 Akita, Y. *et al.* (2011) Effect of the expression level of an *AGAMOUS*-like gene on the petaloidy of stamens in the double-flowered lily, 'Elodie'. *Sci. Hortic. (Amst.)* 128, 48–53

83 Liu, Z. *et al.* (2013) Exon skipping of *AGAMOUS* homolog *PrseAG* in developing double flowers of *Prunus lannesiana* (Rosaceae). *Plant Cell Rep.* 32, 227–237

84 Pin, P.A. and Nilsson, O. (2012) The multifaceted roles of FLOWERING LOCUS T in plant development. *Plant Cell Environ.* 35, 1742–1755

85 Jeon, J.S. *et al.* (2010) Starch biosynthesis in cereal endosperm. *Plant Physiol. Biochem.* 48, 383–392

86 Fan, L.J. *et al.* (2008) Molecular evidence for post-domestication selection in the *Waxy* gene of Chinese waxy maize. *Mol. Breed.* 22, 329–338

87 Kawahigashi, H. *et al.* (2013) A novel *waxy* allele in sorghum landraces in East Asia. *Plant Breed.* 132, 305–310

88 Kawase, M. *et al.* (2005) Diverse origins of waxy foxtail millet crops in East and Southeast Asia mediated by multiple transposable element insertions. *Mol. Genet. Genomics* 274, 131–140

89 Hunt, H.V. *et al.* (2013) Waxy phenotype evolution in the allotetraploid cereal broomcorn millet: mutations at the *GBSSI* locus in their functional and phylogenetic context. *Mol. Biol. Evol.* 30, 109–122

90 Park, Y.J. *et al.* (2012) The molecular basis of mutations at the *Waxy* locus from *Amaranthus caudatus* L.: evolution of the waxy phenotype in three species of grain amaranth. *Mol. Breed.* 30, 511–520

91 Kovach, M.J. *et al.* (2009) The origin and evolution of fragrance in rice (*Oryza sativa* L.). *Proc. Natl. Acad. Sci. U.S.A.* 106, 14444–14449

92 Juwattanasomran, R. *et al.* (2011) A SNP in *GmBADH2* gene associates with fragrance in vegetable soybean variety "Kaori" and SNAP marker development for the fragrance. *Theor. Appl. Genet.* 122, 533–541

93 Carroll, S.B. (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134, 25–36

94 Streisfeld, M.A. and Rausher, M.D. (2011) Population genetics, pleiotropy, and the preferential fixation of mutations during adaptive evolution. *Evolution* 65, 629–642

95 Petroni, K. and Tonelli, C. (2011) Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Sci.* 181, 219–229

96 Gillman, J.D. *et al.* (2011) Loss-of-function mutations affecting a specific *Glycine max* R2R3 MYB transcription factor result in brown hilum and brown seed coats. *BMC Plant Biol.* 11, 155

97 Butelli, E. *et al.* (2012) Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* 24, 1242–1255

98 Cockram, J. *et al.* (2010) Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21611–21616

99 Jarvis, M.C. *et al.* (2003) Intercellular adhesion and cell separation in plants. *Plant Cell Environ.* 26, 977–989

100 Tsiantis, M. (2011) A transposon in *tb1* drove maize domestication. *Nat. Genet.* 43, 1048–1050

101 Studer, A. *et al.* (2011) Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat. Genet.* 43, 1160–1163

102 Hufford, M.B. *et al.* (2013) The genomic signature of crop-wild introgression in maize. *PLoS Genet.* 9, e1003477

103 Wagner, M.R. and Mitchell-Olds, T. (2011) Repeated phenotypic changes highlight molecular targets of convergent evolution. *Genome Biol.* 12, 124

104 Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. J. Murray

105 Noda, T. *et al.* (2002) Physicochemical properties of amylose-free starch from transgenic sweet potato. *Carbohydr. Polym.* 49, 253–260

106 Muth, J. *et al.* (2008) Precision breeding for novel starch variants in potato. *Plant Biotechnol. J.* 6, 576–584

107 Zhao, S.S. *et al.* (2011) Development of waxy cassava with different biological and physico-chemical characteristics of starches for industrial applications. *Biotechnol. Bioeng.* 108, 1925–1935

108 Bennett, T.H. *et al.* (2013) Repeated evolution of salt-tolerance in grasses. *Biol. Lett.* 9, 20130029

109 Schroeder, J.I. *et al.* (2013) Using membrane transporters to improve crops for sustainable food production. *Nature* 497, 60–66

110 Hedges, S.B. (2002) The origin and evolution of model organisms. *Nat. Rev. Genet.* 3, 838–849

111 Furukawa, T. *et al.* (2007) The *Rc* and *Rd* genes are involved in proanthocyanidin synthesis in rice pericarp. *Plant J.* 49, 91–102

112 Zhang, Y.F. *et al.* (2009) The potato *R* locus codes for dihydroflavonol 4-reductase. *Theor. Appl. Genet.* 119, 931–937

113 Zhu, B.F. *et al.* (2011) Genetic control of a transition from black to straw-white seed hull in rice domestication. *Plant Physiol.* 155, 1301–1311

114 Remigereau, M.S. *et al.* (2011) Cereal domestication and evolution of branching: evidence for soft selection in the *Tb1* orthologue of pearl millet (*Pennisetum glaucum* [L.] R. Br.). *PLoS ONE* 6, e22404

115 Ramsay, L. *et al.* (2011) *INTERMEDIUM-C*, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene *TEOSINTE BRANCHED 1*. *Nat. Genet.* 43, 169–172

116 Li, Q. *et al.* (2010) Cloning and characterization of a putative *GS3* ortholog involved in maize kernel development. *Theor. Appl. Genet.* 120, 753–763

117 Li, Y.B. *et al.* (2011) Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. *Nat. Genet.* 43, 1266–1269

118 Yang, Z.B. *et al.* (2012) SNP identification and allelic-specific PCR markers development for *TaGW2*, a gene linked to wheat kernel weight. *Theor. Appl. Genet.* 125, 1057–1068

119 Li, Q. *et al.* (2010) Relationship, evolutionary fate and function of two maize co-orthologs of rice *GW2* associated with kernel size and weight. *BMC Plant Biol.* 10, 143

120 Zhang, L. *et al.* (2012) *TaCKX6-D1*, the ortholog of rice *OsCKX2*, is associated with grain weight in hexaploid wheat. *New Phytol.* 195, 574–584

121 Maron, L.G. *et al.* (2013) Aluminum tolerance in maize is associated with higher MATE1 gene copy number. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5241–5246

Plant Science Conferences in 2014

XVI International Congress on
Molecular Plant-Microbe Interactions
6–10 July, 2014
Rhodes, Greece
http://www.mpmi2014rhodes-hellas.gr/index.php

Plant Biology 2014
12–16 July, 2014
Portland, USA
http://www.aspb.org/meetings

GRC: Plant Molecular Biology
Decision-Making Pathways, Networks, and Models in Plant Biology
20–25 July, 2014
Holderness, USA
http://www.grc.org/programs.aspx?year=2014&program=plantmolec

EMBO Workshop: Intercellular communication in plant development and disease
24–29 August, 2014
Bischoffsheim, France
http://events.embo.org/coming-soon/index.php?EventID=w14-11