Comparative Genetics of Crop Plant Domestication and Evolution

Anne FRARY, Sami DOĞANLAR

Department of Biology, İzmir Institute of Technology, Gülbahçe Campus, 35437, Urla, İzmir - TURKEY

Abstract: Domesticated species differ from their wild ancestors and relatives for a set of traits that is known as the domestication syndrome. The most important syndrome traits include growth habit, flowering time, seed dispersal, gigantism and morphological diversity. This paper reviews what is known about the genetic control of domestication traits with an emphasis on comparative analyses that examine this control in two or more crop species. Such analyses indicate that although most domestication traits are quantitatively controlled, the dramatic morphological changes that accompanied domestication may be due to relatively few genes. These studies also show that domestication genes have been functionally conserved over thousands of years and have similar, although not identical, effects in various species.

Key Words: domestication syndrome, growth habit, flowering time, seed dispersal, gigantism

Kültür Bitkilerinin Evcilleştirilmesi ve Evriminin Karşılaştırılmalı Genetiği

Özet: Evcilleştirme işlemine tabi tutulmuş türler yabani atalarından ve akrabalarından evcilleştirme sindromu olarakta bilinen bir çok karakter bakımından farklılık gösterirler. En önemli evcilleştirme sindromu karakterleri arasında büyüme şekli, çiçeklenme zamanı, tohum dağıtımı, devlik ve morfolojik çeşitlilik yer almaktadır. Bu gözden geçirme çalışması, evcilleştirme karakterlerinin genetik esasları hakkında bilinenleri iki yada daha çok bitki türündeki kontrolü inceleyen kıyaslamalı analizleri göz önünde tutarak özetlemektedir. Bu analizler çoğu evcilleştirme karakterlerinin kantitatif olarak kontrol edilmelerine rağmen, evcilleştirme sonrası ortaya çıkan dramatik morfolojik değişikliklerin kısmen bir kaç gen'den dolayı olduklarını göstermiştir. Ayrıca, bu çalışmalar evcilleştirme gen'lerinin binlerce yıllık evrim içerisinde işlevsel olarak korunduklarını ve farklı bitki türlerinde aynı olmasada benzeri fenotipik etkileri gösterdiklerini ortaya koymuştur.

Anahtar Sözcükler: evcilleştirme sindromu, büyüme tarzı, çiçeklenme zamanı, tohum dökme, devlik

Introduction

Crop plant domestication began approximately 10,000 years ago at the dawn of agriculture (Harlan, 1992). During the domestication process, early agriculturalists consciously or unconsciously selected among wild germplasm for material that was better adapted to human use and cultivation. Since the transition from wild species to domesticate, crop plants have continued to change due to selection exerted by ancient and modern plant breeding and cultivation practices. These changes that occurred subsequent to the initial domestication event(s) are known as crop evolution. As a result of both crop domestication and evolution, today's domesticated species are differentiated from their wild ancestors by an assortment of morphological and physiological characteristics. Collectively these modified

traits are called the domestication syndrome (Harlan, 1992). Although the exact composition of domestication syndrome traits varies with the crop under consideration, syndrome characteristics include a more compact growth habit, increased earliness by alterations in photoperiod sensitivity and/or vernalization requirements, reduction or loss of seed dispersal and dormancy, gigantism, reduced toxicity and increased morphological diversity in the consumed portion of the plant.

The purpose of this paper is to provide a review of recent research concerning the genetic control of traits related to crop domestication and evolution (hereafter collectively called domestication traits for simplicity). Particular emphasis is placed on comparative genetic analyses that span different crop species or genera because such studies reveal the common origins of crop

Corresponding to: annefrary@iyte.edu.tr

and, indeed, all plant species. This work also shows that divergent phenotypes in different species can be controlled by genes that have been functionally conserved over thousands of years of domestication and evolution along different pathways. Examples are drawn from the plant families with the greatest numbers of domesticated species: the Leguminosae, 41 domesticates including soybean, common bean, pea and lentil; the Gramineae, 29 domesticates encompassing the world's primary crops rice, wheat and maize as well as barley, millet and sorghum; the Solanaceae, 18 domesticates including potato, tomato, pepper and eggplant; and the Cruciferae, 18 domesticates comprising canola, cabbage, cauliflower, raddish and the non-crop, preeminent model plant species *Arabidopsis thaliana* (Harlan, 1992).

Genetic Control of the Domestication Syndrome

Although the domestication syndrome is associated with dramatic alterations to plant morphology, research has shown that genetic control of the related traits is often relatively simple and involves comparatively few qualitative and quantitative genes with major phenotypic effects (Table 1). For example, in a survey of different

Table 1. Domestication syndrome traits found to be controlled by a limited number of genes.

Trait	Crop	Reference(s)
growth habit/plant architecture/height	rice	Xiong et al., 1999
	maize	Doebley et al., 1997
	pearl millet	Poncet et al., 2000
	common bean	Koinange et al., 1996
	tomato	Pnueli et al., 1998
flowering time/photoperiod sensitivity	rice	Paterson et al., 1995; Xiong et al., 1999
	maize	Paterson et al., 1995
	sorghum	Paterson et al., 1995
	common bean	Koinange et al., 1996
panicle/spike size	rice	Xiong et al., 1999; Cai and Morishima, 2002
	pearl millet	Poncet et al., 2000
spike morphology	pearl millet	Poncet et al., 2000
fruit size	tomato	Grandillo et al., 1999
	eggplant	Doganlar et al., 2002
fruit morphology	tomato	Grandillo et al., 1999
	eggplant	Doganlar et al., 2002
fruit/seed color	common bean	Koinange et al., 1996
	eggplant	Doganlar et al., 2002
seed size	rice	Paterson et al., 1995
	maize	Paterson et al., 1995
	sorghum	Paterson et al., 1995
	common bean	Koinange et al., 1996
seed morphology	rice	Cai and Morishima, 2002
seed dispersal	cereals	Ladizinsky et al., 1985
	rice	Paterson et al., 1995, Cai and Morishima, 2002
	maize	Paterson et al., 1995
	sorghum	Paterson et al., 1995
	pearl millet	Poncet et al., 2000
	common bean	Koinange et al., 1996
seed dormancy	common bean	Koinange et al., 1996
plant prickliness	eggplant	Doganlar et al., 2002

cereals Ladizinsky (1985) found that seed dispersal was controlled by only one to three major genes. The advent of molecular linkage maps and quantitative trait loci (QTL) identification and mapping techniques has allowed more detailed and comprehensive analyses. Domestication traits that appear to be controlled by a limited number of genes in various crops include flowering time, seed size and dispersal in maize, rice and sorghum (Paterson et al., 1995; Xiong et al., 1999; Cai and Morishima, 2002); growth habit, photoperiod sensitivity and seed dispersal, size, dormancy and color in common bean (Koinange et al., 1996); plant architecture in maize (Doebley et al., 1997); seed dispersal, plant architecture, spike weight and spike morphology in pearl millet (Poncet et al., 2000); growth habit, fruit weight and shape in tomato (Grandillo et al., 1999); and fruit weight, shape, color and plant prickliness in eggplant (Doganlar et al., 2002). Many of these QTL appear to be the result of major genes as their actions account for a large proportion of the phenotypic variation for the associated traits. For example, QTL on linkage group (LG) II of both mung bean and cowpea accounted for 32 to 36% of seed weight variation in the two crops, respectively (Fatokun et al., 1992). In tomato, a single locus (*fw2.2*) on chromosome 2 explained as much as 30% of the variation for fruit weight in the crop (Alpert et al., 1995). Similarly, in eggplant, approximately 75% of the variation for leaf prickliness (or spininess) was explained by a single locus on LG 6 (Doganlar et al., 2002).

When the joint action of several loci for a single trait is considered, it is evident that genetic loci alone are responsible for a large proportion of the variability seen in domestication traits. For example, Xiong et al. (1999) found that only four QTL explained 72% of the total variation for plant height in rice. In common bean, 36 and 48% of the variation for seed height and length, respectively, were each explained by the collective action of three different QTL (Park et al., 2000). Similarly, in pearl millet 77% of the phenotypic variability in spike weight was accounted for by only three loci (Poncet et al., 2000). These results indicate that, as compared to genetic factors, the environmental component of phenotypic variation for such characteristics may be relatively small. Thus, it appears that crop domestication and evolution in various species was conditioned by highly heritable mutations with major morphological effects at a limited number of loci.

Clustering of Domestication Traits

Research in several crops has shown that the genes controlling domestication-related characteristics are frequently clustered or linked within the genome. In common bean, Koinange et al. (1996) found that domestication syndrome genes were not randomly distributed across all chromosomes; instead, they were primarily concentrated in three genomic locations. One region controlled growth habit and flowering time, another was involved in seed dispersal and dormancy and a third controlled pod and seed size. Similar clustering has also been observed in the maize and rice genomes (Khavkin and Coe, 1997; Xiong et al., 1999; Cai and Morishima, 2002). In a particularly striking example, a single 16 cM interval of chromosome 1 contained loci for six different traits including shattering and plant height (Xiong et al., 1999). Pearl millet also provides examples of linked domestication syndrome traits. In this crop, nearly all of the QTL for spikelet architecture and shedding plus additional loci for plant and spike morphology and flowering time were localized to the two smallest linkage groups, LG 6 and LG 7 (Poncet et al., 2000). Two other linkage groups carried clusters of loci with major effects on growth habit, spike morphology and flowering time. Thus, it seems that domestication is often controlled by a limited number of chromosomal regions with major phenotypic effects. For some traits apparent clustering may be the result of a single gene with pleiotropic effects (Johnson et al., 1996; Koinange et al., 1996; Poncet et al., 2000; Xiong et al., 1999). However, given the increasing number of examples of linked domestication traits, it is unlikely that pleiotropy is the sole explanation for clustering of domestication syndrome loci (Park et al., 2000).

It has been hypothesized that species with linkage blocks of adaptation traits may have been favored during the domestication process because such linkages facilitated selection. In contrast to the results for common bean, rice and pearl millet, no similar clustering of domestication traits was observed in eggplant (Doganlar et al., 2002). This difference may be because eggplant, unlike rice and pearl millet, is a primarily self-pollinated crop. It has been proposed that linkage of domestication traits would only have been advantageous for the adaptation of outcrossing crops as those crops are more likely to have suffered from the introduction of unadapted genes from crosses with the wild ancestor

(Koinange et al., 1996). Such outcrossing would interfere with the continued regeneration of the cultivated type.

Overall, the findings that traits related to domestication are highly heritable, governed by relatively few loci with major phenotypic effects and are often clustered within the genome suggest that domestication may have occurred quite rapidly for some crops. According to Gepts (2002), the rapidity with which a given species can be domesticated is limited by the incidence of new, favorable phenotypes (that is, the mutation rate), the intensity of both deliberate and unconscious selection, and the rate at which the linkages between unfavorable and favorable phenotypes are broken (that is, the recombination rate). Based on the discovery that seed size, dispersal and photoperiod sensitivity in sorghum, rice and maize are controlled by a small number of shared loci with large effects, Paterson et al. (1995) hypothesized that domestication of the cereals may have occurred in as little as 100 years.

Comparative Genetic Analyses

As previously mentioned, comparative analysis of the genetic control of traits for different species and genera uncovers the functional conservation of genes over evolutionary time and the common origin of all plant species. In addition, such studies provide information that facilitates and expedites gene discovery and isolation in plant species that have large genomes and/or are less well characterized than the established model systems. Although some comparative studies with *Arabidopsis* are included, the primary subject of this section of the paper is the discussion of results from genetic comparisons among different crop species for various domestication traits. Table 2 summarizes these results.

Plant height and growth habit

As compared to their wild relatives, crop plants tend to be shorter and have a more compact growth habit. For example, wild beans are vines, whereas the most modified cultivated forms are bush types with fewer branches, a decreased number of nodes and shorter internodes (Gepts, 2002). The importance of alterations in plant height for improved crop adaptation is exemplified by the Green Revolution of the 1960s and 1970s. During this period, the worldwide acceptance and cultivation of dwarf and semidwarf wheat and rice lines

with reduced lodging was associated with dramatic, unprecedented increases in grain yield (Ikeda et al., 2001). Since then numerous comparative genetic analyses of plant height in the cereals have revealed that both gibberellin (GA) sensitive and insensitive dwarfing genes are functionally conserved. Mapping of GA sensitive dwarfing genes in wheat and rye suggests that *Rht12* on the long arm of wheat chromosome 5A is orthologous to *Ddw1* on the long arm of the homoeologous rye chromosome 5R (Börner et al., 1998). Broader functional conservation has been observed for the wheat GA insensitive semidwarfing gene *Rht-1*. This locus has orthologs that map to homoeologous regions of the maize (*D8*; Laurie and Devos, 2002) and rice (*SLR1*; Ikeda et al., 2001) genomes. These dwarfing genes also have an ortholog in Arabidopsis, the *GAI* locus, which was cloned and characterized by Peng et al. (1997). For all four genes, the mutant (dwarf) phenotype is the result of deletions in the 5' portion of the gene. Although the deletions in the different species are not identical, their effects are equivalent. These results indicate that the regulation of growth by GA has been conserved during the evolution of all flowering plants (Laurie and Devos, 2002).

Studies of plant height in other crops are at a less advanced stage; however, there is evidence that plant height QTL are also conserved among other species. Poncet et al. (2002) identified two loci on LG 7 of pearl millet that have putative orthologs on chromosomes 7 and 9 of rice. A comparison of height QTL in the eggplant and tomato genomes suggests that the two crops share a common locus on chromosome 2 (Frary et al., in press). Similar work in pepper indicates that a pepper height QTL on chromosome 3 has an ortholog in the tomato genome (Ben Chaim et al., 2001). In *Brassica oleracea*, Lan and Paterson (2001) identified five stem length (equivalent to plant height) QTL that have nine potential counterparts in *Arabidopsis*. Most of the candidate orthologs are known GA biosynthesis, auxin biosynthesis and signalling or brassinosteroid biosynthesis mutants. Although such studies are based on the localization of genes to homoeologous map positions, more precise gene cloning and/or sequence comparisons must be made to definitively identify true orthologs.

Since the domestication of tomato one of the most significant advances for cultivation of the crop was the discovery of the *self-pruning*, *SP*, mutant. Mutant plants

Table 2. Domestication traits and loci with putative conservation across crop species. The crop in which each locus was identified is in parentheses after the locus name. Multiple putative orthologs in a given crop are separated by slashes (/). Genes are named according to the cited reference(s). Loci that were not named in the associated reference are designated nn (not named) followed by the chromosome or linkage group number or letter if given.

Trait	Loci (Crop)	Reference(s)
plant height	*Rht1* (wheat), *SLR1* (rice), *D8* (maize), *GAI* (arabidopsis)	Ikeda et al., 2001; Laurie and Devos, 2002
	Rht12 (wheat), *Ddw1* (rye)	Börner et al., 1998
	ph7/ph9 (rice), *hmax7/hhead7* (pearl millet)	Poncet et al., 2002
	ht2 (tomato), *ht2.1* (eggplant)	Frary et al., in press
	ht3 (tomato), *ph3.1* (pepper)	Ben Chaim et al., 2001
	stl1a (cauliflower), *dwf3/dwf8/acl5* (arabidopsis)	Lan and Paterson, 2001
	stl1b (cauliflower), *dwf7* (arabidopsis)	Lan and Paterson, 2001
	stl3 (cauliflower), *ga1/ga4/dw1/axr1/axr3* (arabidopsis)	Lan and Paterson, 2001
growth habit	*Sp* (tomato), *Cen* (snapdragon), *tfl1* (arabidopsis)	Pnueli et al., 1998
photoperiod	*Ppd genes* (wheat), *Se1/Se3* (rice), *nn9* (maize), *Ppd genes* (barley), *nn* (sugarcane), *Ma7* (sorghum)	Paterson et al., 1995
	Ppd genes 2 (wheat), *Hd2* (rice), *PpdH1* (barley)	Laurie, 1997; Börner et al., 1998
	Hd1 (rice), *nn* (barley), *fto2.1/fto3.1/fto9.2* (broccoli), *nn* (mustard), *co* (arabidopsis)	Bohoun et al., 1998; Yano et al., 2000; Laurie and Devos, 2002
vernalization	*VrnA1* (wheat), *Sp1* (rye), *Sh2* (barley)	Laurie, 1997, Börner et al., 1998
earliness	*eps* (wheat), *eps2* (barley)	Laurie, 1997
	dtf1.1 (tomato), *dtf1.1* (eggplant)	Frary et al., in press
	dtf2.1 (tomato), *dtf2.1* (eggplant)	Frary et al., in press
	dfb1 (cauliflower), *flc/fy/tfl1/emf1/tfl2* (arabidopsis)	Lan and Paterson, 2000
	dfb3a (cauliflower), *hy2/vrn1* (arabidopsis)	Lan and Paterson, 2000
	dfb3b (cauliflower), *fha* (arabidopsis)	Lan and Paterson, 2000
	dfb4 (cauliflower), *emf2* (arabidopsis)	Lan and Paterson, 2000
	dfb5 (cauliflower), *gi* (arabidopsis)	Lan and Paterson, 2000
	dfb7a (cauliflower), *co/art1* (arabidopsis)	Lan and Paterson, 2000
	dfb7b (cauliflower), *efs* (arabidopsis)	Lan and Paterson, 2000
inflorescence size	*pl1/sd1/sp1* (rice), *Cupr* (maize), *We6/Wi6* (pearl millet)	Poncet et al., 2002
	sp7/pl7 (rice), *LoS7/Dens7* (pearl millet)	Poncet et al., 2002
	rk11/rk15 (cauliflower), *Cal/Ap1* (arabidopsis)	Lan and Paterson, 2000
fruit size	*fw2.2* (tomato), *fw2.1* (pepper), *fw2.1* (eggplant)	Ben Chaim et al., 2001; Doganlar et al., 2002
	fw9.2 (tomato), *fw9.1* (eggplant)	Doganlar et al., 2002
	fw11.1 (tomato), *fw11.1* (eggplant)	Doganlar et al., 2002
	fw3.1 (tomato), *fw3.2* (pepper)	Ben Chaim et al., 2001

Table 2. (Continued)

Trait	Loci (Crop)	Reference(s)
fruit shape	*fs10.1* (tomato), *fs10.1* (pepper), *ovs4.1* (eggplant)	Ben Chaim et al., 2001; Doganlar et al., 2002
	fs3.1 (tomato), *fs3.1* (pepper)	Ben Chaim et al., 2001
	fs8.1 (tomato), *fs8.1* (pepper)	Ben Chaim et al., 2001
	ovate (tomato), *fl2.1* (eggplant)	Doganlar et al., 2002
	fs7b (tomato), *fs7.1* (eggplant)	Doganlar et al., 2002
fruit color	*Del* (tomato), *CrtLe* (pepper)	Thorup et al., 2000
	B (tomato), *Ccs* (pepper)	Thorup et al., 2000
	l2 (tomato), *CrtLb* (pepper)	Thorup et al., 2000
	fc4.2 (tomato), *GGPPS* (pepper)	Thorup et al., 2000
	fc1.1 (tomato), *Zds* (pepper)	Thorup et al., 2000
	fc2.2 (tomato), *Ze* (pepper)	Thorup et al., 2000
	Y (potato), *crtZ* (pepper)	Thorup et al., 2000
	F/lep/lco (potato), *af/ag* (tomato), *fap10.1* (eggplant)	Doganlar et al., 2002
	Fs/u (tomato), *fst4.1* (eggplant)	Doganlar et al., 2002
seed size	*nn1* (rice), *nn3* (maize), *nnA* (sorghum)	Paterson et al., 1995
	nn2 (rice), *nn4* (maize), *nnF* (sorghum)	Paterson et al., 1995
	nn2/10 (rice), *nn5* (maize), *nnC* (sorghum)	Paterson et al., 1995
	nn1 (maize), *nnC* (sorghum)	Paterson et al., 1995
	nn3 (rice), *nn5* (maize)	Paterson et al., 1995
	nn5 (rice), *nnE* (sorghum)	Paterson et al., 1995
	nn7 (maize), *nnB* (sorghum)	Paterson et al., 1995
	nn3 (rice), *nn1* (maize)	Paterson et al., 1995
	Phs (common bean), *nn4* (mung bean)	Johnson et al., 1996
	nn2 (mung bean), *nn2* (cowpea)	Fatokun et al., 1992
	nnM (soybean), *nn7* (cowpea)	Maughan et al., 1996
seed dispersal	*nn9* (rice), *nn1/5* (maize), *nnC* (sorghum), *ct7/pb7* (pearl millet)	Paterson et al., 1995; Poncet et al., 2002
	nn1 (rice), *nn3/8* (maize), *al6* (pearl millet)	Poncet et al., 2000, 2002
	nn2 (rice), *nn4* (maize)	Paterson et al., 1995
	nn3 (rice), *nn1* (maize)	Paterson et al., 1995

have a compact, determinate growth habit which ensures synchronous fruit set and allows mechanical harvesting (Pnueli et al., 1998). The *SP* gene was isolated using a combination of map-based cloning and candidate gene techniques. Analysis revealed that the gene is the functional ortholog of the *CENTRORADIALIS (CEN)* and *TERMINAL FLOWER* 1 *(TFL1)* genes of *Antirrhinum* (snapdragon) and *Arabidopsis*, respectively. When compared to these other two species, tomato is thought to have a more primitive growth habit in which the vegetative and reproductive stages alternate. In *Arabidopsis*, however, there are distinct vegetative and reproductive periods with no alternation. Because the *CEN* and *TFL1* genes are known to be involved in the transition of the shoot meristem from vegetative to reproductive growth, it is clear that the same gene family that determines the fate of the floral meristem in *Arabidopsis* and snapdragon also controls meristem determination in tomato (Pnueli et al., 1998).

Flowering time

During domestication, variation in flowering time may have been subject to both deliberate and indirect selection. Direct selection probably resulted from the desire for a plant that produced a crop earlier in the season. Selection for yield may have indirectly resulted in the adaptation of earlier flowering types as optimization of flowering time maximizes the yield potential of a crop by providing the best environmental conditions for reproductive growth. In the cereals, flowering time is thought to be controlled by three different types of genes: photoperiod genes that respond to daylength, vernalization genes that respond to low temperature and earliness genes that are independent of light and temperature conditions (Snape et al., 2001). As flowering time in the cereals has been the subject of more than one recent review (see Laurie, 1997; Börner et al., 1998; Snape et al., 2001), only summaries of the comparative genetic analyses are provided here.

Wheat was domesticated in the region encompassed by southeastern Turkey, Iraq, Iran and Lebanon, where summers are hot and dry. Under such conditions, photoperiod insensitive wheat flowers sooner and produces larger seeds than other types (Snape et al., 2001). Paterson et al. (1995) reported that a photoperiod QTL that accounted for more than 80% of flowering time variation in sorghum has putative

orthologs/paralogs in wheat, rice, barley and sugarcane. These findings were confirmed by results that indicated that *Ppd-H1*, a major photoperiod response gene in barley, maps to a similar position as the *Ppd* genes on wheat chromosome 2 and also correlates to the rice heading date QTL, *Hd-2* (Laurie, 1997; Börner et al., 1998). Interestingly, the barley and wheat genes have been found to govern different responses to daylength, which may indicate that the genes are actually distinct and not derived by descent from a common ancestor. However, an alternative explanation is that this difference is simply a reflection of allelic variation in gene expression that has given rise to diverse phenotypes (Laurie, 1997).

The recent map-based cloning of an important QTL for photoperiod response in rice, *Hd1*, has shown that orthology for the trait is not limited to the cereals (Yano et al., 2000). *Hd1* corresponds to the rice *Se1* mutant, a homolog of the *Arabidopsis CONSTANS (CO)* gene. Putative orthologs to *CO* have also been identified in barley (Laurie and Devos, 2002) and in *B. oleracea* (broccoli and cauliflower) and *B. nigra* (mustard) (Bohuon et al., 1998), two species that are closely related to *Arabidopsis*. Because rice is a short-day plant and *Arabidopsis* is a long-day plant, these results indicate that the same genes are involved in both types of photoperiod response. Further research will be necessary to elucidate the exact molecular basis for the different daylength responses.

Vernalization is the promotion of flowering after exposure to cool temperatures. This trait is important for crop domestication and evolution because it allows adaptation to different planting times to obtain the best yield (Snape et al., 2001). Thus, winter types of cereals require a vernalization period, and are sown in the fall and harvested in the summer. In contrast, spring types do not require exposure to low temperatures, are sown in the spring and are harvested in the fall. In wheat, the difference between winter and spring types is controlled by the *Vrn* genes. In particular, *Vrn-A1* on chromosome 5A seems to be the most important locus for determining the vernalization requirements of European wheat varieties (Snape et al., 2001). This gene has putative orthologs in both rye (*Sp1*) and barley (*Sh2*) (Laurie, 1997; Börner et al., 1998) indicating that the genes controlling the vernalization response, like those for the photoperiod response, have been conserved during crop domestication and evolution.

Genes controlling earliness have not been as intensively studied as the photoperiod and vernalization loci. Despite this, there are preliminary reports that suggest that earliness genes may map to common regions in the barley and wheat genomes (Laurie, 1997), the eggplant and tomato genomes (Frary et al., in press) and the *B. oleracea* and *Arabidopsis* genomes (Lan and Paterson, 2000). Eleven earliness (days to first bud) QTL that may correpond to as few as seven ancestral genes were identified in three different *B. oleracea* populations. These loci mapped to regions that contain 13 different flowering mutants in *Arabidopsis* (Lan and Paterson, 2000). Some of these mutants correspond to known photoperiod and vernalization response genes while others have only been characterized as earliness genes. More detailed comparisons of loci that map to common genome locations are needed to determine whether these examples represent cases of true orthology.

Inflorescence size and morphology

For many crops, domestication was accompanied by modifications of inflorescence size and morphology. *Brassica* presents an outstanding example of the impact of selection on morphological diversity. Both cauliflower (*B. oleracea* subsp. *botrytis*) and broccoli (*B. oleracea* subsp. *italica*) have dramatically enlarged inflorescences as compared to their wild relatives (Purugganan et al., 2000). However, these subspecies and many of the varieties within them have distinct inflorescence phenotypes. Thus, the inflorescence of cauliflower has a curd phenotype while that of broccoli has a head phenotype. The genetic control of curding in cauliflower was analyzed by Lan and Paterson (2000). They found five QTL for the width of the largest cluster within the curd, 12 QTL for the width of the entire curd and 11 QTL for the number of first rank branches in the curd. The position of one of these curd size QTL appears to coincide with the location of the *CAULIFLOWER* and *APETALA1* mutants of *Arabidopsis*. Plants with mutations in both of these genes have a curd-like inflorescence (Kempin et al., 1995). The putative orthology between these genes and a *Brassica* QTL indicates that, although the genetic control of curding in *B. oleracea* may be much more complex than in *Arabidopsis*, it is still possible to identify genes with conserved function in the two species.

Domestication in the cereals also involved changes in inflorescence size and morphology. For example, wild species of pearl millet have short, narrow spikes while cultivated plants have much longer, wider spikes with a denser arrangement of spikelets (Poncet et al., 2000). Comparison of the locations of pearl millet spike and spikelet morphology QTL with loci for related traits in other grasses suggests that several QTL for spike weight, length, width and spikelet density on LG 6 and LG 7 of pearl millet correspond to rice loci for panicle length, spikelet density and number of spikelets per panicle and a maize QTL for ear morphology (Poncet et al., 2002).

Fruit size and morphology

During the domestication and evolution of fruit crops, selection resulted in increased fruit size and diversified fruit morphology and color. Examples of this diversity can be seen in tomato. Tomato fruit can be round, blocky, elongated, pear or heart shaped and can be found in colors ranging from green to yellow to pink, orange and red. In contrast, wild tomato species have small, round, red or green fruit. The genetic control of fruit size and shape has been compared in several solanaceous species. Grandillo et al. (1999) identified a set of fruit weight and shape loci that are shared by cultivated tomato and several of its wild relatives. A fruit weight and shape analysis was also done in pepper. Comparison of these results with those from tomato suggests that two fruit weight and three shape loci in pepper may have orthologous counterparts in tomato (Ben Chaim et al., 2001). These results are only preliminary because comparative mapping between the two species was hampered by an insufficient number of shared markers.

A similar comparison using eggplant avoided this problem by using only markers that were previously mapped and known to be single copy in tomato (Doganlar et al., 2002). All three of the fruit weight QTL identified in eggplant had putative orthologs in tomato. Among these three conserved loci, one (*fw2.1*) also appeared to have a counterpart in pepper. This QTL is of special interest because it corresponds to the *fw2.2* gene of tomato. *fw2.2* is the most important fruit weight QTL identified to date in tomato. This locus has been cloned and is thought to be a negative regulator of cell division (Frary et al., 2000). Thus, the gene is expressed at higher levels in small fruited lines than in large fruited ones. The fact that this locus appears to be conserved in all three of the major fruit domesticates of the Solanaceae provides strong evidence that the gene had a significant role in the

domestication of solanaceous crops. In addition to the fruit weight loci, two eggplant fruit shape QTL were found to have putative orthologs in tomato and an ovary shape locus has a candidate counterpart in both tomato and pepper. Included among these three loci with putative conservation is the *ovate* locus of tomato, which is one of the two most important fruit shape genes in tomato (Grandillo et al., 1999).

Organ color

Another appearance characteristic that seems to have been conserved during crop domestication and evolution is organ color. Thorup et al. (2000) mapped structural genes from the carotenoid biosynthetic pathway in pepper and compared their locations with the positions of QTL and qualitative mutations in tomato and potato. This analysis indicated that six of the pepper enzyme loci have putative orthologs in tomato. The map positions of these six genes correspond to the locations of three fruit color QTL and three tomato carotenoid mutant loci. Two of the pathway genes are apparently orthologous to tomato mutations that result in orange fruit color. Thus, the genes for lycopene e-cyclase (*CrtLe*) and capsanthin capsorubin synthase (*CcsB*) are candidates for the *Del* and *B* mutants of tomato. The third mutant locus, *lutescent-2 (l2)*, is characterized by delayed red pigment formation and reduced levels of β-carotene and xanthophylls and seems to correspond to the lycopene β-cyclase (*CrtLb*) gene. An additional carotenoid biosynthesis gene, β-carotene hydroxylase (*crtZ*), appears to have an ortholog in potato as it maps to a common region with the *yellow-flesh (Y)* mutant of potato which has yellow or orange tubers. Thus, both quantitative and qualitative variation in fruit and tuber color in pepper, tomato and potato can be attributed to genes that have been functionally conserved since the divergence of these species from their most recent common ancestor.

Similar conservation of loci for color traits was observed in a comparison between eggplant and the other Solanaceae. The most significant anthocyanin QTL in eggplant accounted for at least 75% of variation in fruit color and much of the variation in leaf, stem, prickle and flower corolla anthocyanin levels (Doganlar et al., 2002). This QTL mapped to a position on LG 10 of eggplant that is homoeologous to regions of chromosomes 5 and 10 in tomato and potato. These regions of tomato and potato

harbor five different color mutations. Three of these loci are found in potato and control flower (*F*) and tuber skin color (I_{ep} and I_{co}). The two tomato genes, *anthocyanin free (af)* and *anthocyanin gainer (ag)*, are characterized by the complete absence of and the delayed/altered distribution of anthocyanin in the plant, respectively. Although a greater phenotypic similarity suggests that *ag* is the more likely ortholog of the eggplant QTL, additional mapping and characterization of the loci are needed.

Another color trait that seems to be conserved in the Solanaceae is fruit secondary color repartition or fruit stripe. The most important fruit stripe locus in eggplant is on LG 4 and has two putative orthologs in tomato: the *Fs (fruit stripe)* and *u (uniform ripening)* mutants (Doganlar et al., 2002). As with the anthocyanin loci, additional analysis is required to verify this putative orthology. Despite this caveat, the results of the comparative genetic analysis of color traits in the Solanaceae suggest that the diversity of phenotypic expression of these traits in eggplant, tomato, potato and pepper may be explained by different mutations in an overlapping set of conserved genes.

Seed size and dispersal

In seed crops, domestication was often accompanied by dramatic increases in seed size/weight and decreases in seed dispersal. Increases in seed size are very apparent in the legumes. In early work, Fatokun et al. (1992) identified a seed weight QTL that mapped to similar locations and is apparently conserved between cowpea and mung bean. The orthologous locus was the QTL that had the greatest effect on seed weight in both species. In a comparison between common bean, cowpea and mung bean, Johnson et al. (1996) found that a seed weight QTL associated with *Phs*, the locus that codes for the most abundant seed storage protein (phaseolin) of common bean, has a putative ortholog in mung bean. Interestingly, a similar candidate for orthology was not identified in cowpea. Maughan et al. (1996) identified a different soybean seed weight locus with a potential counterpart in cowpea.

Such examples of conserved function are not limited to the legumes. In the cereals, three seed size loci have putatively orthologous counterparts in rice, maize and sorghum (Paterson et al., 1995). Other seed weight QTL have correspondence between two out of three of these

species. At the time, a total of 23 seed size QTL had been identified in rice, maize and sorghum and only four had no potential orthologs in at least one of the other species. This level of conservation for a quantitatively inherited trait is remarkable and supports the hypothesis that relatively few loci were involved in the domestication process.

The reduction of seed dispersal that is apparent in cultivated cereals was primarily accomplished by selection for non-shattering types. Analysis of the genetic control of shattering in these crops has revealed additional examples of conservation of domestication trait gene function. The single locus that controls nearly all of the shattering phenotype in sorghum has potential orthologs in both rice and maize (Paterson et al., 1995). Moreover, two other rice shattering QTL seem to have counterparts in the maize genome. Similarly, seed dispersal loci in pearl millet have putative orthologs in rice, maize and/or sorghum (Poncet et al., 2000; Poncet et al., 2002). Thus it appears that for shattering, as well as for many other traits, mutations in a limited number of genes were capable of giving rise to the domesticated phenotype.

Conclusions

Although the plant kingdom is extraordinarily diverse, humans depend on only a handful of crop species for their survival. During domestication and evolution these crop species underwent dramatic morphological and physiological changes. Genetic analysis of the control of these changes indicates that traits related to domestication are controlled by relatively few genes and that these genes often have major effects on plant phenotype. Moreover, in some species these domestication traits are clustered within the genome. These findings suggest that domestication may have occurred quite rapidly in some species as fixation of only a few genomic regions may have been sufficient for adaptation. Comparative mapping indicates that several genes for important domestication syndrome traits including plant height, flowering time, inflorescence and fruit morphology, organ color, and seed size and dispersal are conserved among different domesticated species. As exemplified by curding in *Arabidopsis* and cauliflower, such functionally conserved genes can give rise to a diversity of phenotypes. These phenotypic differences are apparently the result of varying mutations (alleles) at the common locus and/or the effects of additional loci that interact with the conserved gene to produce the final phenotype for a given trait. The limited number and apparently conserved nature of the genes controlling domestication facilitates the genetic manipulation of the related traits in both model and non-model crop species. These findings also suggest that, in the future, molecular techniques may be used to identify appropriate targets and hasten selection of new domesticates among wild germplasm. Thus, a wild species that is capable of growth under specific biotic or abiotic stresses could be rapidly domesticated for human use and cultivation.

References

Alpert, K., S. Grandillo and S.D. Tanksley. 1995. *fw2.2*: a major QTL controlling fruit weight is common to both red- and green-fruited tomato species. Theor. Appl. Genet. 91: 994-1000.

Ben Chaim, A., I. Paran, R.C. Grube, M. Jahn, R. van Wijk and J. Peleman. 2001. QTL mapping of fruit-related traits in pepper (*Capsicum annuum*). Theor. Appl. Genet. 102: 1016-1028.

Bohuon, E.J., L.D. Ramsay, J. Craft, A. Arthur, D.F. Marshall, D.J. Lydiate and M.J. Kearsey. 1998. The association of flowering time quantitative trait loci with duplicated regions and candidate loci in *Brassica oleracea*. Genetics 150: 393-401.

Börner, A., V. Korzun and A.J. Worland. 1998. Comparative genetic mapping of loci affecting plant height and development in cereals. Euphytica 100: 245-248.

Cai, H.W. and H. Morishima. 2002. QTL clusters reflect character associations in wild and cultivated rice. Theor. Appl. Genet. 104: 1217-1228.

Doebley, J., A. Stec and L. Hubbard. 1997. The evolution of apical dominance in maize. Nature 386: 485-488.

Doganlar, S., A. Frary, C.M. Daunay, R.N. Lester and S.D. Tanksley. 2002. Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. Genetics 161: 1713-1726.

Fatokun, C.A., D.I. Menancio-Hautea, D. Danesh and N.D. Young. 1992. Evidence for orthologous seed weight genes in cowpea and mung bean based on RFLP mapping. Genetics 132: 841-846.

Frary, A., S. Doganlar, C.M. Daunay and S.D. Tanksley. QTL analysis of morphological traits in eggplant and its implications for conservation of gene function during evolution of solanaceous species. Theor. Appl. Genet. (in press).

Frary, A., T.C. Nesbitt, A. Frary, S. Grandillo, E. van der Knaap, B. Cong, J. Liu, J. Meller, R. Elber, K. Alpert and S.D. Tanksley. 2000. *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. Science 289: 85-88.

Gepts, P. 2002. The evolution of crop plants. http://agronomy.ucdavis.edu/gepts/pb143 /pb143.htm.

Grandillo, S., H.M. Ku and S.D. Tanksley. 1999. Identifying the loci responsible for natural variation in fruit size and shape in tomato. Theor. Appl. Genet. 99: 978-987.

Harlan, J.R. 1992. Crops and Man. American Society of Agronomy and Crop Science Society of America, Madison, WI, USA.

Ikeda, A., M. Ueguchi-Tanaka, Y. Sonoda, H. Kitano, M. Koshioka, Y. Futsuhara, M. Matsuoka and J. Yamaguchi. 2001. *slender* rice, a constitutive gibberellin response mutant, is caused by a null mutation of the *SLR1* gene, an ortholog of the height-regulating gene *GAI/RGA/RHT/D8*. Plant Cell 13: 999-1010.

Johnson, W.C., C. Menendez, R. Nodari, E.M.K. Koinange, S. Magnusson, S.P. Singh and P. Gepts. 1996. Association of a seed weight factor with the phaseolin seed storage protein locus across genotypes, environments and genomes in *Phaseolus-Vigna* spp.: Sax (1923) revisited. J. Quan. Trait Loci 2: 5.

Kempin, S., S. Savidge and M.F. Yanofsky. 1995. Molecular basis of the cauliflower phenotype in *Arabidopsis*. Science 267: 522-525.

Khavkin, E. and E. Coe. 1997. Mapped genomic locations for developmental functions and QTLs reflect concerted groups in maize (*Zea mays* L.). Theor. Appl. Genet. 95: 343-352.

Koinange, E.M.K., S.P. Singh and P. Gepts. 1996. Genetic control of the domestication syndrome in common bean. Crop Sci. 36: 1037-1045.

Ladizinsky, G. 1985. Founder effect in crop-plant evolution. Econ. Bot. 39: 191-199.

Lan, T.H. and A.H. Paterson. 2000. Comparative mapping of quantitative trait loci sculpting the curd of *Brassica oleracea*. Genetics 155: 1927-1954.

Lan, T.H. and A.H. Paterson. 2001. Comparative mapping of QTLs determining the plant size of *Brassica oleracea*. Theor. Appl. Genet. 103: 383-397.

Laurie, D.A. 1997. Comparative genetics of flowering time. Plant Molecular Biology 35: 167-177.

Laurie, D.A. and K.M. Devos. 2002. Trends in comparative genetics and their potential impacts on wheat and barley research. Plant Molecular Biology 48: 729-740.

Maughan, P.J., M.A. Saghai-Maroof and G.R. Buss. 1996. Molecular-marker analysis of seed-weight: genomic locations, gene action and evidence for orthologous evolution among three legume species. Theor. Appl. Genet. 93: 574-579.

Park, S.O., D.P. Coyne, G. Jung, P.W. Skroch, E. Arnaud-Santana, J.R. Steadman, H.M. Ariyarathne and J. Nienhuis. 2000. Mapping of QTL for seed size and shape traits in common bean. J. Amer. Soc. Hort. Sci. 125: 466-475.

Paterson, A.H., Y.R. Lin, Z. Li, K.F. Schertz, J.F. Doebley, S.R.M. Pinson, S.C. Liu, J.W. Stansel and J.E. Irvine. 1995. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science 269: 1714-1717.

Peng, J.R., P. Carol, D.E. Richards, K.E. King, R.J. Cowling, G.P. Murphy and N.P. Harberd. 1997. The *Arabidopsis* GAI gene defines a signalling pathway that negatively regulates gibberellin responses. Genes Dev. 11: 3194-3205.

Pnueli, L., L. Carmel-Goren, D. Hareven, T. Gutfinger, J. Alvarez, M. Ganal, D. Zamir and E. Lifschitz. 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.

Poncet, V., F. Lamy, K.M. Devos, M.D. Gale, A. Sarr and T. Robert. 2000. Genetic control of domestication traits in pearl millet (*Pennisetum glaucum* L., Poaceae). Theor. Appl. Genet. 100: 147-159.

Poncet, V., E. Martel, S. Allouis, K.M. Devos, F. Lamy, A. Sarr and T. Robert. 2002. Comparative analysis of QTLs affecting domestication traits between two domesticated wild pearl millet (*Pennisetum glaucum* L., Poaceae) crosses. Theor. Appl. Genet. 104: 965-975.

Purugganan, M.D., A.L. Boyles and J.I. Suddith. 2000. Variation and selection at the *CAULIFLOWER* floral homeotic gene accompanying the evolution of domesticated *Brassica oleracea*. Genetics 155: 855-862.

Snape, J.W., K. Butterworth, E. Whitechurch and A.J. Worland. 2001. Waiting for fine times: genetics of flowering time in wheat. Euphytica 119: 185-190.

Thorup, T.A., B. Tanyolac, K.D. Livingstone, S. Popovsky, I. Paran and M. Jahn. 2000. Candidate gene analysis of organ pigmentation loci in the Solanaceae. Proc. Nat. Acad. Sci. USA 97: 11192-11197.

Xiong, L.Z., K.D. Liu, X.K. Dai, C.G. Xu and Q. Zhang. 1999. Identification of genetic factors controlling domestication-related traits of rice using an F_2 population of a cross between *Oryza sativa* and *O. rufipogon*. Theor. Appl. Genet. 98: 243-251.

Yano, M., Y. Katayose, M. Ashikari, U. Yamanouchi, L. Monna, T. Fuse, T. Baba, K. Yamamoto, Y. Umehara, Y. Nagamura and T. Sasaki. 2000. *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. Plant Cell 12: 2473-2483.