



Tansley review

A comparative view of the evolution of grasses under domestication

Author for correspondence:
Sylvain Glémin
Tel: +33 (0)4 67 14 48 18
Email: glemin@univ-montp2.fr
Received: 4 February 2009
Accepted: 1 April 2009

Sylvain Glémin¹ and Thomas Bataillon^{2,3}

¹Institut des Sciences de l'Evolution (UM2-CNRS), Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France; ²Institute of Biology, Section of Genetics and Ecology and Bioinformatics Research Center, Aarhus University, Aarhus, Denmark; ³INRA, UMR 1097 Diversité et Adaptation des Plantes Cultivées, Montpellier, France

Contents

Summary	273	VI. Mating systems and the evolutionary dynamics of domestication in grasses	284
I. Introduction	274	VII. Conclusion	286
II. A brief history of domestication in grasses	275	Acknowledgements	287
III. Domestication genes	278	References	287
IV. Models of the domestication process	280		
V. Evolutionary consequences of domestication for grass genomes	281		

Summary

New Phytologist (2009) **183**: 273–290
doi: 10.1111/j.1469-8137.2009.02884.x

Key words: domestication, genetic diversity, genome evolution, grasses, mating systems, Poaceae, selection.

Crop grasses were among the first plants to be domesticated c. 12 000 yr ago, and they still represent the main staple crops for humans. During domestication, as did many other crops, grasses went through dramatic genetic and phenotypic changes. The recent massive increase in genomic data has provided new tools to investigate the genetic basis and consequences of domestication. Beyond the genetics of domestication, many aspects of grass biology, including their phylogeny and developmental biology, are also increasingly well studied, offering a unique opportunity to analyse the domestication process in a comparative way. Taking such a comparative point of view, we review the history of domesticated grasses and how domestication affected their phenotypic and genomic diversity. Considering recent theoretical developments and the accumulation of genetic data, we revisit more specifically the role of mating systems in the domestication process. We close by suggesting future directions for the study of domestication in grasses.

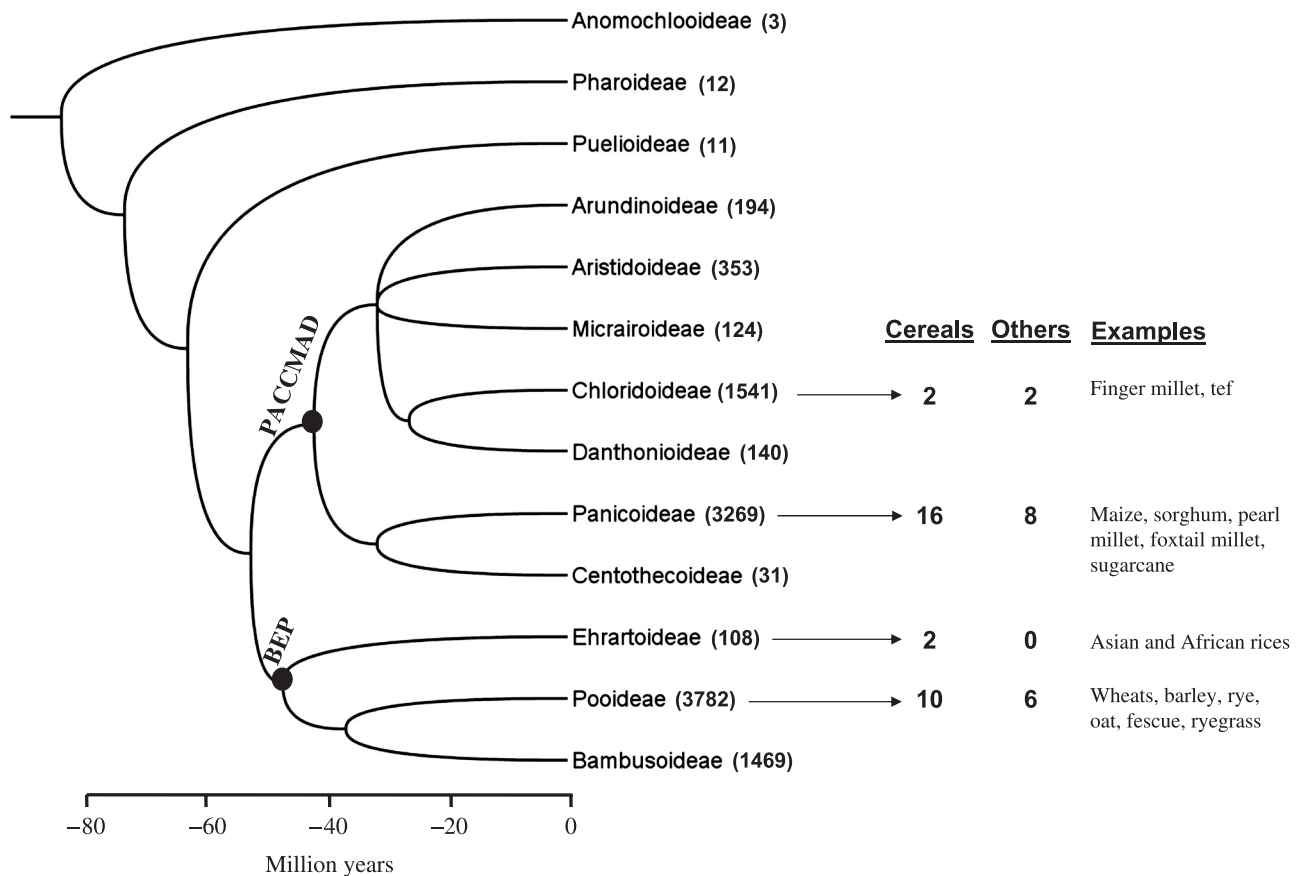


Fig. 1 Phylogeny of grasses at the subfamily level according to Bouchenak-Khelladi *et al.* (2008). An approximate evolutionary timescale is given according to fossil and molecular calibrations (Bremer, 2002; Janssen & Bremer, 2004; Leebens-Mack *et al.*, 2005; Prasad *et al.*, 2005). BEP: Bambusoideae+Ehrartoideae+Pooideae; PACCMAD: Panicoideae+Aristidoideae+Centothecoideae+Chloridoideae+Micrairoideae+Arundinoideae+Danthonioideae. The number of species in each subfamily, given in brackets, is taken from Watson & Dallwitz (1992) with slight taxonomic adjustments to match the phylogeny. The number of domesticated and cultivated species is also given (see list in Supporting Information Table S1).

I. Introduction

Plant domestication dramatically affected the fate of human societies by playing a crucial role in the shift from hunting and gathering to agriculture. Domestication is an evolutionary process whereby a population adapts, through selection, to new environments created by human cultivation, most of these adaptations being deleterious in the wild. A number of domesticated plants, at least in seed-propagated species, experienced dramatic genetic and phenotypic changes through many reproduction–selection cycles (Zohary, 2004). Domestication is a unique example of rapid evolution by selection, and was a central metaphor in Darwin's theory (Darwin, 1859, 1882). Indeed, as an evolutionary process, domestication has specific characteristics making it especially suitable for studying adaptation on a short timescale (Gepts, 2004). Over the course of a few hundreds or thousands of years, drastic environmental changes driven by human cultivation selected for striking new adaptations. These are easy to recognize at the phenotype level and to study (Purugganan & Fuller, 2009).

Among plant crops, grasses are by far the most important species. Wheats (*Triticum* sp.) and barley (*Hordeum vulgare*) were the first domesticated plants (Zohary & Hopf, 2000), and major cereals – bread wheat (*Triticum aestivum*), rice (*Oryza sativa*), maize (*Zea mays*), sorghum (*Sorghum bicolor*) and pearl millet (*Pennisetum glaucum*) – are current main staple crops for humans. Many other grasses were domesticated as cereals, although most of them are currently minor crops (Simmonds, 1976) (Figs 1, 2, Supporting Information Table S1). Many grasses are also used as fodder, such as fescue (*Festuca* sp.) and rye-grass (*Lolium* sp.), or for specific industrial production, such as sugarcane (*Saccharum officinarum*). The domestication process led to a suite of striking phenotypic changes, collectively referred to as the domestication syndrome, which many cereal grasses share, through convergent evolution (Harlan *et al.*, 1973). In this review we focus on major cereal crops, for which recent genomic data provide new tools to investigate the genetic basis of domestication changes, and more generally how crop genomes were moulded by the strong selective pressures involved in the domestication process.

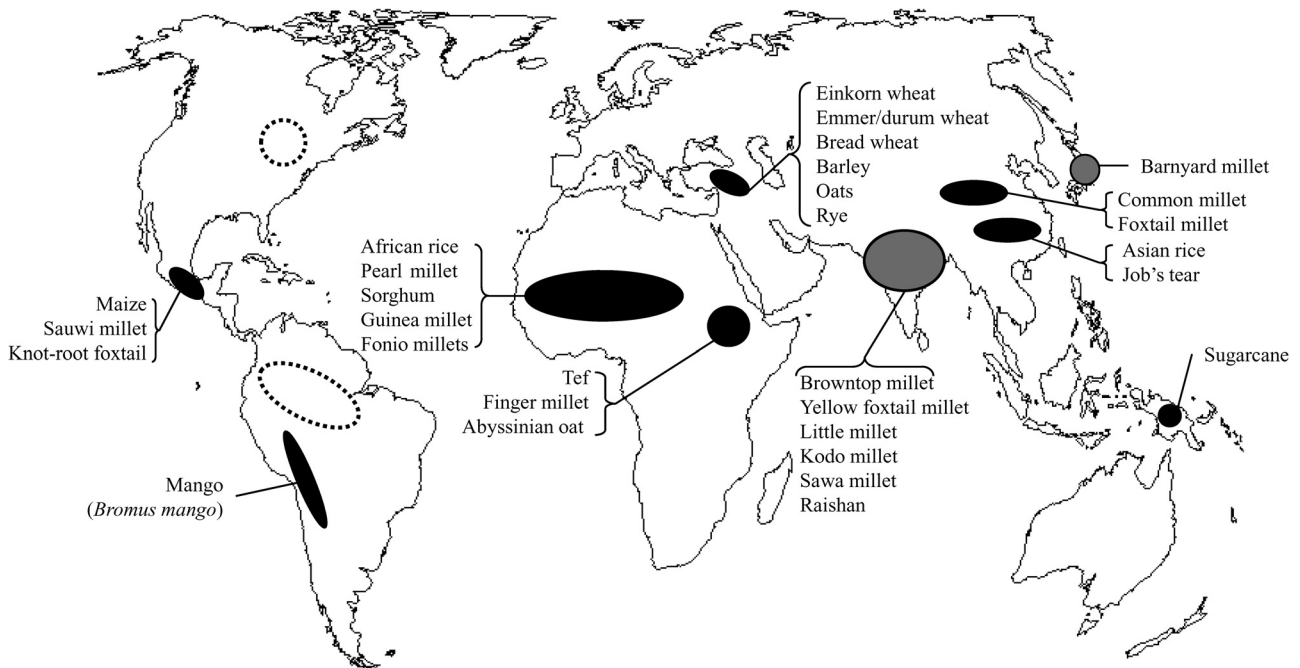


Fig. 2 World-wide distribution of domestication centres including main grass crops (black areas). Grey areas correspond to areas of origin of minor cereals, usually not considered as domestication centres. Dotted areas correspond to major domestication centres without domesticated grass species.

After sketching briefly the history of domestication in grasses, we review and discuss how domestication affected phenotypic and genomic evolution in these species. Previous reviews mainly addressed specific aspects of plant domestication (e.g. Paterson, 2002; Ross-Ibarra, 2005; Doebley *et al.*, 2006; Doust, 2007; Ross-Ibarra *et al.*, 2007; Burger *et al.*, 2008) or detailed the domestication process in one or a few species (e.g. Salamini *et al.*, 2002; Doebley, 2004; Kovach *et al.*, 2007). In this review, we use grasses to illustrate key patterns and mechanisms in domestication. We also use a comparative approach to discuss more specifically the role of genetic systems in the domestication process, comparing verbal arguments, predictions based on explicit population genetics models, and available data.

II. A brief history of domestication in grasses

1. Timescale of grass evolution and domestication

The grass family (Poaceae) includes approx. 10 000 extant species (Watson & Dallwitz, 1992). It originated in the late Cretaceous *c.* –80 Mya according to both recent fossil data (Prasad *et al.*, 2005) and molecular calibrations (Janssen & Bremer, 2004). The core Poaceae split into two major clades, the BEP and the PACCMAD clades (Fig. 1; GPWG, 2001; Bouchenak-Khelladi *et al.*, 2008), which diverged around, or even before, –55 Mya (Bremer, 2002; Leebens-Mack *et al.*, 2005; Prasad *et al.*, 2005). Domesticated and cultivated

species belong to these two clades within four subfamilies, two in each clade: Ehrartoideae (rices) and Pooideae (wheats, barley, rye and oats – *Avena* sp.) in the BEP clade, and Panicoideae (maize, sorghum, pearl millet and sugarcane) and Chloridoideae (finger millet – *Eleusine coracana* – and tef – *Eragrostis tef*) in the PACCMAD clade (Fig. 1). All these species except rices come from the three most diverse subfamilies of the 13 belonging to the Poaceae (Fig. 1). Pooideae diversified in cool temperate and boreal regions and include C3 species only, while Panicoideae and Chloridoideae diversified in the tropics and subtropics and include many C4 species. Among these subdomesticated and cultivated crops belong to a few genera or tribes, forming clusters of closely related species, such as Triticeae (wheats, barley and rye), which emerged *c.* 12 Mya (Gaut, 2002; Huang *et al.*, 2002), or Andropogoneae (maize, sorghum and sugarcane), which emerged between 9 and 16.5 Mya (Gaut & Doebley, 1997; Gaut, 2002).

Compared with the evolution of the grass family, the timescale of the domestication process is much more recent. The transition from foraging to agriculture began *c.* 12 000 yr ago in the Fertile Crescent (now spanning Israel, Jordan and parts of Turkey) and more recently in other parts of the world. It was probably associated with warmer and dryer climatic conditions following the Younger Dryas period, a brief cold climate period at the end of the Pleistocene, *c.* 11 500 to 13 000 BP (Wright, 1976), and with the hunting to extinction of large mammalian game (Diamond, 1999, 2002). While it happened rapidly

on an evolutionary timescale, domestication was still a gradual process involving a stage of cultivation of wild plants that preceded morphological domestication (Weiss *et al.*, 2006). The rate of transition from initial cultivation to full domestication is still debated. Domestication experiments in einkorn wheat (*Triticum monococcum* ssp. *monococcum*), the first domesticated wheat, and barley as well as population genetic simulations suggest that acquisition of a tough rachis, a key trait for domesticated cereals (see II.3), could potentially evolve very rapidly in 20 to 100 yr, even with unconscious selection (Hillman & Davies, 1990). However, archaeological data indicate that the establishment of this trait probably took over 1500 yr in wheat and 2000 yr or more in barley (Tanno & Willcox, 2006; Fuller, 2007). Similar timescales of *c.* 1000–1500 yr have also been proposed for Asian rice (*Oryza sativa*) (Fuller, 2007). The domestication process was thus probably slower than previously assumed, involving a rather long pre-domestication cultivation stage and a protracted transition towards full domestication (see Section V) (Wright, 1976; Fuller, 2007; Allaby *et al.*, 2008).

2. Origin of domesticated species

A limited number of domestication centres have been described (Fig. 2; Diamond, 2002). Within these centres, the distribution of domesticated species can be linked to the biogeographical distribution of grass subfamilies. In tropical and subtropical centres, species from the Panicoideae and the Chloridoideae were domesticated while Pooideae species were domesticated in the more temperate Fertile Crescent (Simmonds, 1976; Frankel *et al.*, 1995). Interestingly, in America, domestication of cereal grasses was mainly confined to maize, and agriculture originated with staple crops other than grass cereals in two or three out of four centres. Some minor cereals were also probably domesticated in other areas such as the Indian peninsula and Japan, usually not considered as major domestication centres (Fig. 2).

For most domesticated species, the wild progenitor has been identified by combining data on the morphology, biogeographical distribution and ecology of extant wild species and archeobotanical data (see Table S1). The origin of maize (*Zea mays* ssp. *mays*) was initially much debated because, unlike most crops, it has no morphologically equivalent wild form (reviewed in Doebley, 2004). Molecular evidence now clearly shows that teosinte (*Z. mays* ssp. *parviglumis*), while it differs strikingly from maize in lateral branching and the female inflorescence, is the direct ancestor (Doebley *et al.*, 1984; Matsuoka *et al.*, 2002).

Knowledge of the wild progenitors offers a unique opportunity to assess the phenotypic and genetic consequences of the domestication process by contrasting initial and final states. For most species, extant wild populations are still known. Comparison of domesticated grasses with their wild relatives is the basis for identifying the set of traits adapted to human cul-

tivation, collectively referred to as the domestication syndrome. Most of these traits are shared across species through convergent evolution in response to similar selective pressures. Below, we detail the domestication syndrome associated with cereal grasses; some of these traits are shared with other grain crops.

3. Morphological changes during domestication: the domestication syndrome

Archaeological evidence suggests that wild grasses were harvested before cultivation (Weiss *et al.*, 2006). Under harvest alone, selection favours wild-type traits as only seeds escaping the harvester contribute to the next generation. The situation was reversed as soon as humans started to sow what they had harvested: harvested seeds are now those contributing to the next generation. Strong disruptive selection then occurred between wild and cultivated forms (see Fig. 3b). This transition was associated with selection-targeted traits related to harvesting conditions, seed production, and seedling competition (Harlan *et al.*, 1973).

Traits associated with harvesting Nonshattering of seeds at maturity is often regarded as the most important domestication trait and it is often diagnostic of domesticated forms for archaeologists. Most seeds that do not shatter are harvested, while shattering seeds are dispersed and lost. This makes crop propagation fully dependent upon farmer harvesting and sowing so that conscious selection is not necessary to explain the evolution of such a trait (Harlan *et al.*, 1973; Zohary, 2004). Selection for other traits associated with seed dispersal was also relaxed. For instance, domesticated wheats are less hairy than wild forms, and awns are reduced or absent.

Under cultivation, space became the limiting factor, while time was limiting for foraging (Fuller, 2007). This selected for more determinate and compact growth habit. This is especially striking for crops belonging to the PACCMAD clade (maize, sorghum and different millets). Vegetative branching differs between the PACCMAD and the BEP clades. In the BEP clade, tiller production is favoured over axillary branching, whereas in the PACCMAD clade, in addition to tillers, plants produce axillary branches under nonlimiting space and light conditions (Doust, 2007). During domestication, the outgrowth of multiple axillary branching was strongly selected against, as exemplified by the difference between teosinte and maize (Doebley *et al.*, 1997).

Harvesting also selected for synchronization of maturation. The life cycle of domesticated forms is less plastic than that of wild plants, and vernalization and control by day length can be weaker or even lost. Selection for synchronization may also lead to a shift from perennial to annual, as in Asian rice (Cheng *et al.*, 2003) or Kodo millet (*Paspalum scrobiculatum*) (de Wet *et al.*, 1983).

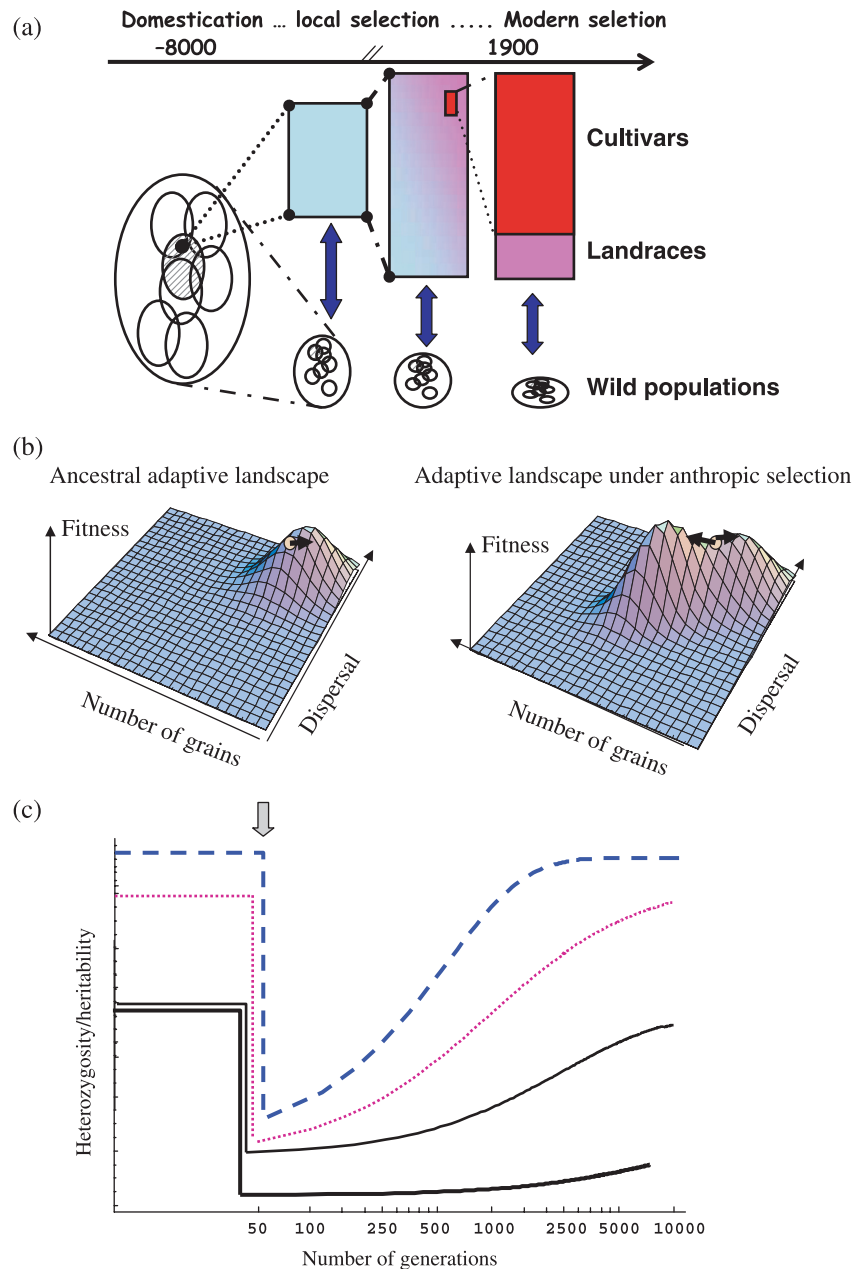


Fig. 3 An overview of the domestication process in grasses, and the demographic, genetic and phenotypic changes involved. (a) A simplified sketch of the demographic history of a crop including a series of nested bottlenecks and gene flow between the crop and wild gene pools. Both bottleneck and gene flow are integral parts of grass domestication history and have opposing effects on genetic diversity. Quantification of the relative importance of these two components in crop history is an area of active research. (b) A sketch of the evolutionary consequences of domestication for two quantitative traits. Domestication by opening a new adaptive niche generates strong disruptive selection (arrows). (c) Expected timescales for regenerating genetic diversity by mutation after a massive loss of polymorphism. Curves are numerical iterations of a theoretical model where mutation regenerates neutral variation in a single population (assuming an effective size of $N_e = 10\,000$). Dashed line, quantitative variation (assuming heritability increases by 1% per generation); dotted line, variation at microsatellite loci (assuming a mutation rate of 0.001 per locus and per generation); thin line, nucleotide variation at anonymous loci (assuming a mutation rate of 0.00001 per locus and per generation); bold line, nucleotide variation at domestication loci.

Traits associated with seedling competition Cultivation also selected for a general increase in seedling vigour through an increase in seed size and in carbohydrate relative to protein content. The increase in seed size was mostly a result of an increase in the size of the endosperm, which is richer in carbohydrate but lower in proteins than the embryo (Harlan *et al.*, 1973). Rapid germination was also selected for through the reduction or loss of dormancy and the reduction of glumes. In wild oats, einkorn and emmer (*Triticum turgidum* ssp. *dicoccum*), the erratic rainfall of the Mediterranean region selected for an interesting pre-adaptation to cultivation under competitive conditions. Each spikelet contains a dormant seed and a nondormant, much bigger one, which germinates

with the first rains in the autumn and which can compete with dense populations of other annual plants (Harlan *et al.*, 1973).

Traits associated with seed production and use In addition to unconscious selection associated with seedling competition, seed size might have been selected consciously by early farmers, together with other traits increasing seed production. For instance, changes in inflorescence structure were selected to produce higher yields. In crops from the PACCMA clade, where branching is reduced, larger inflorescences were selected for, especially in maize, sorghum and pearl millet. In crops from the BEP clade, such as wheats, barley and rice, the number of

inflorescences was preferentially selected for through denser tillering (Doust, 2007). More compact spikes with more fertile flowers were also selected for. In maize, the female inflorescence is strikingly different between teosinte, with two ranks of single spikelets, and maize, with multiple ranks of paired spikelets (Doebley, 2004). The transition from two-rowed to six-rowed barleys resulted from similar selective pressures (Komatsuda *et al.*, 2007).

Finally, a naked kernel became a desirable trait as it allowed free-threshing. In the first domesticated cereals, grains were hulled and the naked grain was sometimes selected later. No naked form was selected in einkorn wheat. In emmer wheat, which is hulled, selection for naked grain yielded durum (pasta) wheat. Similarly, hulled spelt wheat led to naked bread wheat (Zohary & Hopf, 2000). However, hulled varieties of wheat and barley were sometimes preferred, maybe for improved storability or for specific uses, such as beer brewing (Purugganan & Fuller, 2009).

Crops then spread out of their centre of origin. While minor crops remained cultivated close to their centre of origin, major crops spread world-wide, mostly after the advent of world travel in the 16th century. This came along with diversifying selection and adaptation to new habitats, leading to locally and culturally adapted landraces, which phenotypically diverged from the initial crop. Here, we will not discuss further post-domestication evolution (see, for instance, Doebley *et al.*, 2006; Burger *et al.*, 2008).

III. Domestication genes

Recent genetic and genomic studies allow identification of the genes underlying the phenotypic evolution described above. Here, we review strategies that have been used so far to identify domestication genes and how these studies shed new light on the genetic basis of the domestication process in grasses.

1. How to identify domestication genes: from phenotypes to genotypes, and back

The traditional method of identifying domestication genes is to examine co-segregation of genetic markers and phenotypes in crosses between cultivated and wild genotypes. Modest genetic divergence between a crop and its wild relatives rarely prevents such types of crossing design being used. This approach, combined with the wealth of molecular markers that have recently become available, has yielded a very large number of quantitative trait loci (QTLs) or regions harbouring genes controlling the variation of interest (reviewed in Paterson, 2002). Going from QTL mapping, typically involving mapping at best within 1–2 centiMorgans, to actual genes is often a heroic task, especially in genomes of grasses such as polyploid wheats, which may have huge genomes and many intergenic regions filled with repetitive sequences. In this respect, the availability of the genome sequences of at least two rice varieties, the *indica*

and *japonica* cultivated types, is a formidable tool for going from QTL mapping to the gene level. Combining traditional QTL studies with rice transgenic transformants has allowed rapid progress in identifying domestication genes (see Table 1). An alternative to strict linkage mapping is to perform so-called 'association mapping', which relies on the statistical associations between marker loci and QTLs in large populations. One theoretical advantage of the latter approach is the exploitation of patterns of linkage disequilibrium generated by deeper genealogical links between individuals, resulting in the accumulation of more recombination events than in traditional crossing designs (Nordborg & Tavaré, 2002). This approach has so far mostly been employed for cultivated populations (for maize see Thornsberry *et al.*, 2001) and remains to be implemented in heterogeneous populations mixing cultivated and wild ancestors.

A radically different approach is to rely solely on patterns of polymorphism and seek genomic regions exhibiting footprints of selection in samples of genotypes contrasting the extant diversity of crops and their wild ancestors (Wright *et al.*, 2005). Following the advent of cheap high-throughput sequencing and genotyping, this population genetics approach can be seen as an attractive alternative to the forward genetic strategies. In practice, several factors may limit the efficacy of such a strategy. One limiting factor is the power of tests to detect selection footprints and the need to specify a plausible null hypothesis for the distribution of these tests that incorporates a possibly complex demographic history (see below). The second limiting factor is that, whenever a selection footprint is detected, it may be hard to date unambiguously the selection event and whether domestication or subsequent artificial selection triggered the selection. To date, this approach has only yielded a few new candidates for domestication genes and has mainly identified already known domestication genes or QTL segments (Wright *et al.*, 2005). However, use of such genome scans may be a powerful way to narrow the search to candidate regions that can be further investigated for their phenotypic effects (Ross-Ibarra *et al.*, 2007).

2. The nature of domestication genes and the nature of selective changes

Inspection of the actual domestication genes and the molecular basis for phenotypic changes reveals a wide variety of molecular changes, including almost ubiquitous changes in gene expression, amino acid changes, and so far only two single 'loss of function' mutations (frameshift deletions in the *Rc* – *brown pericarp and seed coat* – and *vrs1* – *six-rowed spike 1* – genes; see Table 1). This contrasts with earlier views of domestication proceeding mostly through recessive loss-of-function mutants. Patterns of dominance of mutations have been found to be quite variable both at the gene and the QTL levels (see Burger *et al.*, 2008 for a detailed account at the QTL level). However, an interesting pattern emerges across the genes cloned so far. In Table 1, seven genes out of 10 are regulatory genes, mainly

Table 1 A selection of domestication genes in grasses

Gene(s)	Crop	Phenotypic function	Molecular function	Selection footprints	Causative ^a	Dominance	References
<i>tb1</i> (<i>teosinte branched1</i>)	Maize	Plant and inflorescence structure	Transcriptional regulator	Yes	Reg	Partly dominant	Wang <i>et al.</i> (1999)
<i>tga1</i> (<i>teosinte glume architecture</i>)	Maize	Seed casing	Transcriptional regulator	Yes	AC	Partly dominant	Wang <i>et al.</i> (2005)
<i>qSH1</i> (QTL of seed shattering in chromosome 1)	Rice	Abscission layer formation, shattering	Transcriptional regulator (homeodomain)	Not tested	Reg	?	Shomura <i>et al.</i> (2008)
<i>Rc</i> (<i>brown pericarp and seed coat</i>)	Rice	Change from the red seeds of wild rice to the white seeds of cultivated rice	Transcriptional regulator	Yes	Two mutations; frame shift deletion (98%)/AC (2%)	?	Sweeney <i>et al.</i> (2006, 2007)
<i>sh4</i> (QTL 4 responsible for the reduction of grain shattering)	Rice	Abscission layer formation, shattering	Transcriptional regulator	Not tested	Reg + AC	Recessive	Li <i>et al.</i> (2006)
<i>GIF1</i> (<i>grain incomplete filling 1</i>)	Rice	Controls grain filling	Cell-wall invertase, changing expression patterns of <i>GIF1</i>	Not tested	Reg	Codominant	Wang <i>et al.</i> (2008)
<i>qSW5</i> (QTL for seed width on chromosome 5)	Rice	Increase in sink size as a result of an increase in cell number in the outer glume of the rice flower		Yes?	Deletion	?	Shomura <i>et al.</i> (2008)
<i>PROG1</i> (<i>prostrate growth 1</i>)	Rice	Affects prostrate vs erect growth habit. Can decrease the number of primary and secondary branches, grain number and grain yield	Cys(2)-His(2) zinc-finger protein	Yes	Indel and SNPs causing 23 ACs	Codominant	Tan <i>et al.</i> (2008a)
<i>vrs1</i> (<i>six-rowed spike 1</i>)	Barley	Controls transition from two-rowed to six-rowed barley	Transcription factor that includes a homeodomain with a closely linked leucine zipper motif			Recessive	Komatsuda <i>et al.</i> (2007)
Q	Wheat	Inflorescence structure	Transcriptional regulator	Not tested	Reg/AC	?	Simons <i>et al.</i> (2006)

^aReg, regulatory changes (in the level and/or pattern of gene expression); AC, amino acid change in the protein.

transcription factors. If this tendency is confirmed, it would support the view that morphological variations between wild and crop species can be easily created through the regulation of development pathways. The pleiotropic nature of these genes could also explain how few genetic changes may have such strong pleiotropic phenotypic effects.

3. Number of domestication genes and convergent evolution

Theoretical (Le Thierry d'Ennequin *et al.*, 1999) and QTL studies (Burger *et al.*, 2008) suggest that few gene regions of large effect are likely to underlie trait changes under domestication. However, a recent genomic scan in maize and rice suggested that many more genes could have been involved. Wright *et al.* (2005) suggested that 2–4% of genes, that is, *c.* 1200 genes, experienced artificial selection in maize. In rice, selection affecting polymorphism throughout the genome was also invoked to explain the peculiar pattern of polymorphism (Caicedo *et al.*, 2007): many derived single nucleotide polymorphism (SNP) alleles segregate at high frequency, a molecular signature of positive selection.

The issue of the number of genes needed to 'build' a crop is also related to the issue of convergent evolution at the molecular level. More than a decade ago, Paterson *et al.* (1995) proposed that the convergent phenotypic evolution observed for a variety of domesticated grasses had occurred through independent selection of mutations at orthologous genetic loci. Since then, the accumulation of QTL data has not strongly supported this hypothesis. One should note, however, that the imprecision typically associated with QTL location, 5–10 cM and sometimes more, makes a direct comparison less straightforward. More precise analyses using molecular mapping did not support Patterson *et al.*'s hypothesis. For instance, the comparative mapping of seed shattering genes in wheats, barley, rice, maize and sorghum suggests multiple genetic routes for the selection of the nonshattering phenotype (Li & Gill, 2006). In barley, different nonorthologous loci control the nonbrittle rachis phenotype (Komatsuda *et al.*, 2004), which suggests that different nonorthologous loci can be recruited, even within a single species.

Recently, the comparison of genes cloned in different species has become possible. So far, the findings of such studies also partly contradict Patterson *et al.*'s hypothesis. *tb1* (*teosinte branched1*), the major gene controlling branching phenotype in maize, has only a minor and variable effect in foxtail millet (*Setaria italica*). In this species, other candidate genes seem to be involved, which suggests that orthologous loci may not be involved in phenotypic convergence (Doust *et al.*, 2004). The evolution of the glutinous or waxy phenotype is another compelling example of convergent evolution, associated with long-standing cooking practices. Evidence of convergent selection on the same *waxy* gene in rice (Olsen *et al.*, 2006) and maize (Fan *et al.*, 2008) supports Patterson *et al.*'s hypothesis. However,

the selection of the same phenotype in foxtail millet is a result of multiple insertion and deletion of a transposable element in the granule-bound starch synthase 1 (*GBSSI*) gene (Kawase *et al.*, 2005). Finally, analysis of the gene controlling the number of rows in barley, *vs1*, shows that at least three independent mutations in this same gene underlay the convergent transition from the two-rowed to the six-rowed phenotype (Komatsuda *et al.*, 2007).

IV. Models of the domestication process

In combination with archaeological studies, genomic data are powerful tools that can be used to unravel domestication scenarios. Supposedly neutral markers, such as microsatellites, or anonymous gene fragments were initially used, but tracing back the history of the very genes underlying domestication changes is pivotal to determining the tempo and the mode of the domestication process. This helps to elucidate not only the demographic history of the domesticated traits but also the dynamics of the selection acting on them.

1. Single versus multiple domestications

Whether wild progenitors were taken into cultivation only once or more than once is still debated in many species (for instance Zohary, 1999; Allaby *et al.*, 2008; Ross-Ibarra & Gaut, 2008). A single origin is consistent with a rapid emergence of domesticated crops, associated with the selection of nonshattering, commonly viewed as the very first domestication trait, followed by diffusion from their centre of origin. This scenario was often considered as the default scenario for most species. Maize is the most convincing example of a single domestication event, as was clearly proved by microsatellite analyses (Matsuoka *et al.*, 2002). Genetic studies also indicated a precise location for the early diversification of maize in Mexico, in the highlands near Oaxaca close to the oldest known archaeological remains (Piperno & Flannery, 2001). Pearl millet is another example of a species that was probably domesticated only once, in an area ranging from the interior delta of the Niger to the Aïr mountains (Oumar *et al.*, 2008). The history of domestication in a number of major grass crops, such as rice and barley, is still a subject of debate and extensive investigation.

In rice (*Oryza sativa*), two main groups, currently called *japonica* and *indica*, were recognized as early as the Chinese Han dynasty (approx. AD 100) (reviewed in Sweeney & McCouch, 2007). Phylogenetic analyses of these groups suggest a polyphyletic origin from the wild species *Oryza rufipogon* (Cheng *et al.*, 2003; Zhu & Ge, 2005; Londo *et al.*, 2006). The *indica* and *japonica* groups may have diverged 0.4 Mya, long before domestication (Zhu & Ge, 2005). However, *indica* and *japonica* genotypes share the same haplotype at two important domestication loci, controlling the nonshattering (Li *et al.*, 2006) and the white pericarp phenotypes (Sweeney *et al.*, 2006), suggesting a common origin for these

genes. Under a multiple domestication scenario, these results provide evidence for a single origin of these domestication alleles, followed by introgression driven by strong directional selection (Kovach *et al.*, 2007; Sang & Ge, 2007; Sweeney *et al.*, 2007). A genome-wide microsatellite survey also indicates that completely independent domestication is less likely than a domestication scenario involving partial sharing of an ancestral population or recent gene flow (Gao & Innan, 2008).

Domesticated barleys are morphologically diverse, spikelets having either two rows (*Hordeum vulgare* ssp. *distichum*) like the wild form (*Hordeum vulgare* ssp. *spontaneum*) or six rows (*Hordeum vulgare* ssp. *vulgare*). A single domestication from the Jordan–Lebanon region was initially inferred through AFLP (Amplified Fragment–Length Polymorphism) analyses (Badr *et al.*, 2000). However, more recent analyses suggested the possibility of multiple domestication events, one of them perhaps located east of the Fertile Crescent (Morrell & Clegg, 2007; Saisho & Purugganan, 2007), in agreement with distinct genetic control of the nonbrittle rachis phenotype between oriental and occidental lines (Komatsuda *et al.*, 2004). The western/eastern separation is also compatible with the morphological distinction between the two-rowed and six-rowed spikelet phenotypes (Saisho & Purugganan, 2007), in agreement with their distinct genetic clustering (Kilian *et al.*, 2006). Finally, the independent mutations in the *vrs1* gene that led to the six-rowed phenotype (see III.3) also suggest a recurrent domestication process (Komatsuda *et al.*, 2007).

2. Towards a protracted model of domestication

These examples suggest that multiple domestications and complex scenarios seem to be more frequent than initially thought. The slow rate of the domestication process inferred on the basis of recent archaeological data (see II.1) also challenges the rapid transition model (Tanno & Willcox, 2006; Fuller, 2007). Under a protracted model with a long timescale, multiple local domestications with gene flow between localities of cultivation are likely, making the geographical origin of the crops more diffuse and the origins of the different cultivated pools nonindependent (Allaby *et al.*, 2008), as exemplified by rice and barley. In rice, the scenario proposed to resolve the paradox of multiple domestications with the sharing of the nonshattering haplotype (Kovach *et al.*, 2007; Sang & Ge, 2007) also suggests that the domestication process began before the acquisition of the nonshattering traits, at least in one of the two rice varieties. This is consistent with new archaeological data suggesting that grain size and shape evolved before nonshattering in rice, but also in einkorn wheat and barley (Fuller, 2007).

Einkorn wheat (*Triticum monococcum* ssp. *monococcum*) also offers a good example of the evolving view of the domestication process. The first genetic analyses suggested a monophyletic origin from the wild subspecies *Triticum monococcum* ssp. *boeoticum*, localized in the Turkish Karaçadag mountains

(Heun *et al.*, 1997) close to some of the earliest sites of agricultural settlements in the Near East (Zohary & Hopf, 2000). However, a recent analysis, based on a very wide sample of wild populations, modified the scenario (Kilian *et al.*, 2007). The wild subspecies was found to be genetically structured into three groups, one of them, called β , being the sister clade of the domesticated form. Because the domesticated form is as polymorphic as this β form, the authors suggested a protracted ‘dispersed-specific’ model to explain their genetic data. The wild β form would have been initially cultivated in the Karaçadag region, then dispersed and domesticated locally several times. If a single origin scenario appeared to be a good approximation of the data on a rough scale, more detailed genetic studies recently showed that a more complex and protracted model is more appropriate.

V. Evolutionary consequences of domestication for grass genomes

1. Loss of diversity ...

The domestication scenarios discussed above imply that the amount of genetic diversity in a crop is a variable fraction of the initial diversity present in its wild ancestor’s gene pool. As expected, studies comparing gene nucleotide polymorphism in crops vs wild relatives document substantial loss of diversity during domestication (Table 2, Fig. 3). Bottleneck effects on crop genetic diversity have often been quantified using a simple demographic model featuring an instantaneous change in population size at the time of domestication. This model is, in our opinion, not to be taken literally or meant to be historically correct but aims merely to quantify the net effect of the cumulative impact of domestication (and subsequent selection) on current crop diversity. It can also serve subsequently as a null model to locate genome fragments that exhibit atypical loss of diversity suggestive of selection. In this context, the intensity of the bottleneck can be quantified by the ratio of (long-term effective) population size before (N_{wild}) and after domestication (N_{dom} ; see Table 2), given that in principle only scaled mutation rates (product $\theta = 4N_e\mu$, where N_e is the effective population size, and μ the mutation rate) are estimable from the data (Table 2). One exception is the situation where patterns of polymorphisms are surveyed using loci with known mutation rates (Thuillet *et al.*, 2005). This allows direct estimation of the long-term effective size realized in each gene pool and may also provide a more accurate picture of the intensity of successive bottlenecks during domestication and subsequent selection.

Inspection of Table 2 reveals wide ranges of bottleneck intensities. Variation in this ratio can be caused by several nonmutually exclusive factors, including the intensity of selection associated with domestication and with subsequent breeding, and mating systems. Although einkorn shows the lowest loss of diversity, which is consistent with its limited use

Table 2 Loss of genetic diversity in some grasses during domestication

Common name	Genus	Wild species	Domesticated species	Markers	$n_{\text{wild}}^{\text{a}}$	$n_{\text{domesticated}}^{\text{a}}$	$\theta_{\text{wild}}^{\text{b}}$	$\theta_{\text{domesticated}}^{\text{b}}$	Ratio	References
Maize	<i>Zea</i>	<i>Zea mays</i> ssp. <i>parviglumis</i>	<i>Zea mays</i> ssp. <i>mays</i>	774 nuclear genes	16	14	10.91	6.27	0.57	Wright <i>et al.</i> (2005)
Pearl millet	<i>Pennisetum</i>	<i>Pennisetum glaucum</i> ssp. <i>monodii</i>	<i>Pennisetum glaucum</i>	1 nuclear gene	6	4	3.6	2.4	0.67	Gaut & Clegg (1993)
Einkorn wheat	<i>Triticum</i>	<i>Triticum monococcum</i> ssp. <i>boeoticum</i>	<i>Triticum monococcum</i> ssp. <i>monococcum</i>	17 nuclear + 1 chloroplastic gene	321	84	3.45	2.45	0.71	Kilian <i>et al.</i> (2007)
		<i>Triticum monococcum</i> ssp. <i>boeoticum</i> (subsample β) ^c	<i>Triticum monococcum</i> ssp. <i>monococcum</i>	17 nuclear + 1 chloroplastic gene	43	2.05	84	2.45	1.20	Kilian <i>et al.</i> (2007)
Emmer wheat	<i>Triticum</i>	<i>Triticum turgidum</i> ssp. <i>dicoccoides</i>	<i>Triticum turgidum</i> ssp. <i>dicoccum</i>	21 nuclear genes	28	12	3.52	1.25	0.36	Haudry <i>et al.</i> (2007)
			<i>Triticum turgidum</i> ssp. <i>durum</i>	21 nuclear genes	28	20	3.52	0.52	0.15	Haudry <i>et al.</i> (2007)
Bread wheat	<i>Triticum</i>	<i>Triticum turgidum</i> ssp. <i>dicoccum</i> (A and B genomes)	<i>Triticum aestivum</i>	21 nuclear genes	12	41	1.25	0.64	0.51	Haudry <i>et al.</i> (2007)
		<i>Triticum turgidum</i> ssp. <i>tauschii</i> (D genomes)	<i>Triticum aestivum</i>	1 nuclear gene ^d	528	198	27.09	1.16	0.04	Caldwell <i>et al.</i> (2004)
Barley	<i>Hordeum</i>	<i>Hordeum vulgare</i> ssp. <i>spontaneum</i>	<i>Hordeum vulgare</i> ssp. <i>vulgare</i> + <i>distichum</i> (total)	7 nuclear genes	25	13	7.28	1.85	0.25	Kilian <i>et al.</i> (2006)
			Modern cultivars	4 nuclear genes	34	74	14.41	3.77	0.26	Caldwell <i>et al.</i> (2006)
			Landraces	4 nuclear genes	34	15	14.41	5.03	0.35	Caldwell <i>et al.</i> (2006)
Asian rice	<i>Oryza</i>	<i>Oryza rufipogon</i>	<i>Oryza sativa</i> (total)	11 nuclear genes	21	60	3.7	2.11	0.57	Caicedo <i>et al.</i> (2007)
			<i>Oryza sativa</i> ssp. <i>indica</i>	11 nuclear genes	21	21	3.7	1.58	0.43	Caicedo <i>et al.</i> (2007)
			<i>Oryza sativa</i> ssp. <i>japonica</i>	11 nuclear genes	21	39	3.7	1.23	0.33	Caicedo <i>et al.</i> (2007)

^aSample size.^bWatterson estimator ($\times 10^3$).^cKilian *et al.* (2007) considered that the loss of diversity must be computed only from the subsample that directly produced the domesticated species.^dOne gene of the study with only one single nucleotide polymorphism (SNP) in both the wild and the domesticated samples is not included here.

and nonintense breeding, the claim that einkorn underwent no reduction of diversity during domestication seems overstated (Kilian *et al.*, 2007). The domesticated form is as diverse as the β form from which it arose; however, the β form contains only a fraction of the whole wild gene pool diversity (Table 2). With the exception of einkorn, it is worth noting that selfing species suffer from the greatest loss of diversity. While the number of species surveyed is too low to allow definitive conclusions to be drawn, this finding is consistent with two predictions from population genetics theory. First, selection acting on domestication genes might affect a bigger fraction of the genome in selfing species because of stronger genetic linkage (e.g. Caicedo *et al.*, 2007). Secondly, restoration of genetic diversity after domestication through wild-to-crop pollen gene flow is more likely in outcrossing than in selfing species.

While levels of diversity have often been analysed using coalescent simulations and the simple bottleneck population model described above, a protracted model, featuring mild bottlenecks and recurrent gene flow that can mitigate the loss of diversity on a local or regional scale, seems closer to the historical truth (Brown *et al.*, 2009). Future research should explore whether parameters of a coalescent model incorporating a realistic gene flow component can be estimated independently from the bottleneck intensity and whether such gene flow contributes substantially to current observed levels of diversity. To go further, a systematic comparison of several scenarios ranging from single short bottleneck models to a fully protracted model to explain current patterns of diversity in many species would be enlightening.

2. ... and regeneration

Another open question regards the amount of diversity regained after domestication. Studies probing genetic diversity solely by using nucleotide polymorphism in genic regions are not surprisingly finding that virtually all nucleotide variation currently segregating in crops is merely a subset of the variation present in the ancestral wild gene pool (see the expected theoretical patterns of differential regeneration in Fig. 3c). It is thus difficult to distinguish between ancestral wild variation retained in the crop and allele reintroduction through gene flow from the wild populations. In barley landraces, photoperiod response follows a latitudinal cline, with the responsive forms, which flower early under long days, in the Mediterranean region, and nonresponsive forms predominating in the north of Europe (Jones *et al.*, 2008). The nonresponsive form is associated with an SNP present in some wild Iranian and Israeli populations. This suggests that the nonresponsive form evolved after domestication through introgression of a wild haplotype during the spread of agriculture in Europe, or that two distinct domesticated pools spread through the north and the south of Europe (Jones *et al.*, 2008).

Studies focusing on microsatellite variability (Thuillet *et al.*, 2002; Vigouroux *et al.*, 2002; Thuillet *et al.*, 2005) or struc-

tural variation (Morgante *et al.*, 2005) can document ongoing production of new mutations in both maize and wheat. Estimates for the rate of genome-wide mutation affecting quantitative traits are scarce in grasses but the few studies available suggest that new mutations have the potential to increase heritable phenotypic variation by as much as 0.1–1% per generation (Sprague *et al.*, 1960; Lynch, 1988; Houle *et al.*, 1996; Bataillon, 2000). Whether such variation is primarily attributable to unconditional deleterious alleles or can provide heritable variation for adaptation to a variety of conditions is still not resolved. We expect that new beneficial mutations affecting phenotypes are exceedingly rare in well-adapted populations. However, new mutations can have a sizable probability of being beneficial when adapting to a new phenotypic optimum (Martin & Lenormand, 2006). There is thus hope for a resolution of the apparent paradox of variation in crops relative to their wild ancestors: at the molecular level variation is typically low in crops relative to their wild relatives, while the reverse is found at the phenotypic level, as emphasized by Darwin (1878). A recent study suggests that this scenario may have occurred in sorghum (de Alencar Figueiredo *et al.*, 2008). Of six candidate genes analysed, two displayed novel variations derived from post-domestication mutations, suggesting that neo-diversity contributed to new adaptations for human uses.

3. The genomic cost of domestication

The evolutionary history of domestication in grasses is characterized by a genome-wide loss of diversity attributable to rather intense selection on a subset of wild genotypes exhibiting desirable phenotypic changes. This means that the genome-wide effects of random genetic drift have probably been magnified relative to the wild populations. Moreover, strong selection on domestication genes is likely to have swept away pre-existing variations in neighbouring regions in a window spanning 10–100 kb (see review by Purugganan & Fuller, 2009). Theory predicts that selection at one locus can greatly diminish the efficacy of selection at neighbouring loci (Kimura, 1962), especially in regions of the genome that are poorly recombining. Although there are wide variations in the rate of recombination we expect that centromeric regions and chloroplast and mitochondrial genomes will be particularly affected (Gordo & Charlesworth, 2001). One way to document such effects is to look for signs of relaxed selection against slightly deleterious mutations. This is typically done by examining the frequency of nonsynonymous (amino acid changing), and thus most often deleterious, SNP variation. The rationale is that, in populations where natural selection is maximally efficient, new slightly deleterious mutations at nonsynonymous sites will be either weeded out or kept at very low frequency. A genome-wide survey of patterns of non-synonymous versus synonymous divergence between two rice subspecies crops (*Oryza sativa* ssp. *indica* and *Oryza sativa* ssp. *japonica*) and a wild relative (*Oryza brachyantha*) suggests acceleration in

protein evolution between rice cultivars (Lu *et al.*, 2006). A total of 15 406 genes were compared between cultivars and 4640 genes between *Oryza sativa* ssp. *japonica* and *Oryza brachyantha*. The authors then binned genes in fragments of 2 megabases and found that the ratio of amino-acid versus synonymous divergence values, K_A/K_S , for each 2-Mb fragment between *O. sativa* ssp. *japonica* and *O. sativa* ssp. *indica* was negatively correlated with the average recombination rates of that fragment ($r = -0.192$, $P = 0.008$). A detailed analysis of the site frequency spectrum of SNPs in 111 gene fragments also found a very high frequency of derived SNPs (Caicedo *et al.*, 2007). The author explored the fit of various demographic scenarios that were either neutral or included some form of selection. Patterns of variation could not be explained by a neutral model featuring a simple bottleneck. Other studies have also documented the accumulation of slightly deleterious mutations in domesticated species, such as in the mitochondrial genome of the domestic dog (Cruz *et al.*, 2008) as well as in industrial and laboratory yeast strains (Gu *et al.*, 2005). In several grass species, including maize, sorghum and wheat, growing data on nucleotide polymorphism could be used to investigate systematically this phenomenon and provide additional insights into the evolutionary history of crops that will complement our knowledge about patterns of diversity loss during domestication.

VI. Mating systems and the evolutionary dynamics of domestication in grasses

Mating systems are thought to have played a key role in domestication. It is usually assumed that self-fertilization made domestication easier (for instance Zohary & Hopf, 2000), and it was argued that mating systems could partly explain differences in domestication rates between the Old and the New Worlds (wheats/barley vs maize; see Diamond, 1999). However, these ideas have not really been tested. Grasses present a wide range of breeding systems (Connor, 1981), and, while rather few outcrossing rate estimates are available, they are expected to be strongly bimodal according to the distribution found in anemophilous species in general (Vogler & Kaliz, 2001). Grasses are thus ideal for comparisons of the effects of strongly contrasted mating systems on the domestication process.

1. Mating systems and the tempo and mode of domestication

Theoretical predictions As discussed above, a key step in the domestication process is the selection, and eventually fixation, of domestication alleles. If alleles enabling domestication are mainly recessive (a point now much debated; see VI.2), we expect domestication rates to strongly depend on mating systems. In outcrossing populations, recessive alleles can be selected for only once they reach sufficiently high frequencies to be revealed in homozygotes. In large populations, the pro-

bability of fixation of a single recessive mutation is thus extremely low, about $\sqrt{s/2N}$, where s is the selection coefficient in favour of homozygotes and N the population size (Kimura, 1962). By contrast, in selfing populations, the probability of fixation is independent of the dominance level, and is equal to s (Charlesworth, 1992). Advantageous mutations also go to fixation more quickly in selfing than in outcrossing populations; the time to fixation may even be reduced by an order of magnitude (Damgaard, 2000). Initial domestication steps should thus be easier and faster in selfing than in outcrossing species. However, the picture is less simple if adaptation proceeds from standing variation attributable to (possibly slightly deleterious) alleles already available in the wild populations. In contrast to the case of a single mutant, adaptation from standing variation is easier for recessive than for dominant mutations, simply because recessive mutations are expected to be found initially at higher frequencies (Orr & Betancourt, 2001).

Because of disruptive selection between wild and cultivated populations (Fig. 3c), gene flow from the wild is expected to bring genotypes that are poorly adapted to cultivation (Lenormand, 2002). In particular, once a suite of domesticated alleles has been selected from the wild (through mutants or deleterious variants segregating at low frequency), further gene flow would recurrently bring wild maladapted alleles, limiting the efficiency of human selection. In that context, selfing also 'protects' cultivated populations from (pollen) gene flow from wild populations or from different cultivars (Zohary & Hopf, 2000).

Empirical evidence According to theoretical predictions, we expect to find more selfing species among domesticated grasses than among wild species. Indeed, among major cereals, only maize, rye, and pearl millet are outcrossing, while wild ancestors of Asian and African rice (*Oryza rufipogon* and *Oryza barthii*, respectively) have mixed-mating systems. However, cereals are annual and many annuals are self-fertilizing in grasses (Barrett *et al.*, 1996). So, it is not clear if domesticated grasses are truly exceptional as far as mating systems are concerned or merely reflect the association between annuality and selfing.

We compared the distribution of mating systems in domesticated and wild species using information available in the literature (see Table S3 for the full list). Quantitative estimates of selfing rates reveal a shift towards selfing from wild perennials to wild annuals, as generally documented in seed plants (Barrett *et al.*, 1996), and from wild to domesticated annuals (Fig. 4). Using a rougher characterization of mating systems (outcrossing/mixed-mating/selfing) this picture is strengthened, and the difference is statistically significant ($P = 0.0077$, Fisher's exact test; Fig. 4). This simple analysis supports a bias towards selfing among domesticated species. However, we suggest that the fact that almost exclusively annual species were domesticated also explains the prominence of selfing species among domesticated cereals. Extension of this analysis to other groups, such as legumes, should bring sufficient statistical power to confirm or refute our conclusions in grasses.

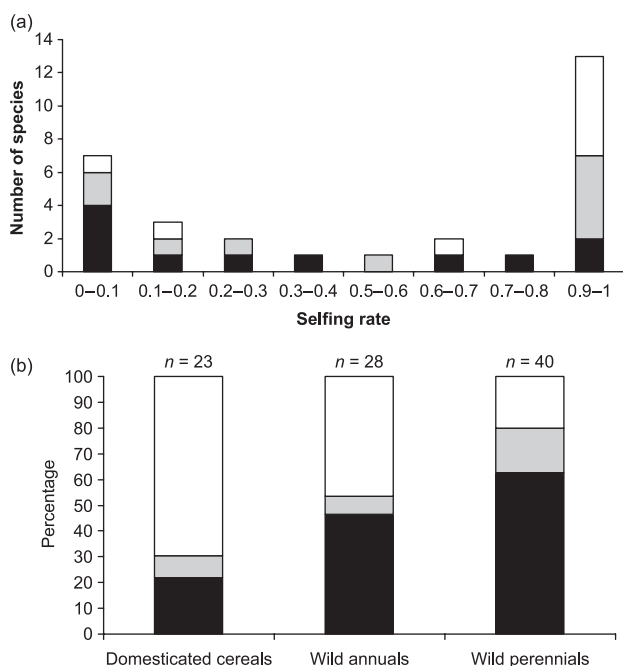


Fig. 4 Distribution of mating systems in wild and domesticated grasses. (a) Distribution of selfing rates estimated through genetic markers. Wild perennials, black bars; wild annuals, grey bars; domesticated cereals, white bars. (b) Distribution of species according to three qualitative categories (selfing, white bars; mixed mating, grey bars; outcrossing, black bars). (See detailed list in Supporting Information Table S3.) The difference between the distributions in domesticated and wild annual species was tested by Fisher's exact test ($P = 0.0077$).

Another prediction regards domestication rates. Selfing species should have been domesticated earlier and more rapidly than outcrossing ones. This hypothesis is difficult to assess because too few outcrossing grasses have been domesticated, and domestication dynamics are poorly known except for major crops. The first grasses domesticated, *c.* 10 000 BP, are self-fertilizing (wheats and barley). However, maize (outcrossers) and Asian rice (selfer) were both domesticated a bit later, *c.* 8000 BP. Within each centre of origin, the order of domestication also seems independent of mating systems. In Mesoamerica, maize was domesticated before common bean (*Phaseolus vulgaris*) (selfers; 4000 BP); in Africa, pearl millet (outcrosser) was domesticated *c.* 3000 BP, after sorghum (selfer; 4000 BP) but before African rice (selfer; 2000 BP) (dates reviewed in Doebley *et al.*, 2006). Finally, the fact that multiple domestications occurred in most selfing species but not in outcrossing ones supports the idea that domestication should be easier in selfing species.

2. Mating systems and the genetic architecture of domestication traits

Origin and dominance of domestication alleles Mating systems should also affect patterns of dominance and linkage

among traits selected during the domestication process. Most domestication alleles are thought to be strongly deleterious in the wild. In selfing populations, such alleles are maintained at very low frequencies because of efficient purging (Ohta & Cockerham, 1974). Selection should thus proceed from new mutations. In selfing species, any type of mutation could be equally selected for, which seems compatible with the variability of dominance levels detected at the gene and QTL levels (see Table 1 and Ross-Ibarra, 2005; Burger *et al.*, 2008).

In outcrossing populations, selection of domestication alleles should mainly proceed from standing variation (a so-called 'soft sweep') or from new dominant mutations, because it is very unlikely to select for new recessive mutation (see VI.1). It has been argued that it is difficult to distinguish adaptation resulting from a single new mutation (a 'hard sweep') from adaptation resulting from standing variation (Orr & Betancourt, 2001; Innan & Kim, 2004; Przeworski *et al.*, 2005). However, recent theoretical work suggests that soft sweeps may leave their own specific imprint on neighbouring neutral polymorphisms: under recurrent mutation, they are expected to leave a strong signature in patterns of linkage disequilibrium in the region around the sweep (Pennings & Hermisson, 2006).

In the outcrossing maize, Innan & Kim (2004) interpreted the weak signature of selection on domestication candidate genes involved in the starch (Whitt *et al.*, 2002) and anthocyanin (Hanson *et al.*, 1996) pathways as evidence of a 'soft sweep' from standing variation. More interestingly, in two major domestication genes involved in maize domestication – *tb1*, controlling apical dominance, and *tga1*, controlling the naked kernel trait – dominant or partially dominant alleles have been selected for (Dorweiler *et al.*, 1993; Doebley *et al.*, 1995, 1997). In pearl millet, another outcrossing species, the domesticated phenotype for plant architecture seems also to be dominant over the wild type (Poncet *et al.*, 1998).

Clustering of domestication QTLs Mating systems should also affect the patterns of physical linkage among domestication genes. In a simulation study, Le Thierry d'Ennequin *et al.* (1999) predicted that the number of QTLs involved in domestication should be higher in selfing than in outcrossing species. Moreover, QTLs should be more tightly linked in outcrossing species. In maize, QTL (Briggs *et al.*, 2007) and genome scan (Wright *et al.*, 2005) approaches showed that a small number of genomic regions control several phenotypic traits involved in the domestication process and further selection. In pearl millet, major QTLs are also grouped (Poncet *et al.*, 1998, 2000, 2002). However, this feature is not specific to outcrossers. QTLs have also been found to be grouped within a few genomic regions in rice (Cai & Morishima, 2002; Tan *et al.*, 2008a,b), in wheat (Peng *et al.*, 2003), and in barley (Gyenis *et al.*, 2007), while some cases of apparent QTL clustering may be attributable to pleiotropic effects of individual genes.

Current data in grasses yield a mixed picture. Thanks to high-throughput sequencing technologies, we hope that it

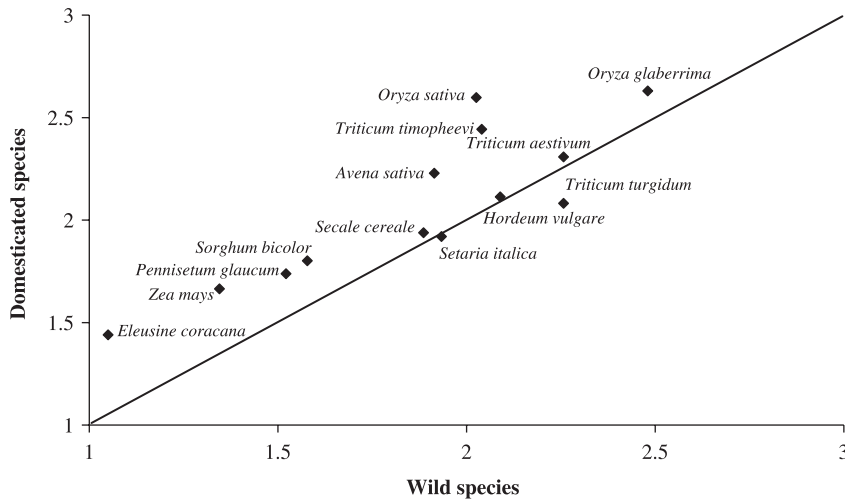


Fig. 5 Relative differences in chiasma frequency per bivalent between domesticated grasses and their wild relatives (wild progenitors or close relatives). Data are taken from Ross-Ibarra (2004); see details in Supporting Information Table S2.

will soon be possible to perform genome scans to estimate the number and the distribution of domestication genes in many grasses and other important crop species. Comparative population genomics is a promising approach to address in much greater generality the question of the genetic effect of mating systems on the domestication process.

3. Domestication and the evolution of genetic systems

Finally, the domestication process may also affect the evolution of mating systems themselves, and genetic systems in general. An increase in selfing rates during domestication has been documented in several species. In both Asian and African rices, the mating system evolved from mixed-mating ancestors (*O. rufipogon* and *O. barthii*) to highly selfing domesticated species (*O. sativa* and *Oryza glaberrima*, respectively) (Caicedo *et al.*, 2007; Sweeney & McCouch, 2007). Such a transition also probably occurred in finger millet between the wild (*Eleusine africana*) and the cultivated (*Eleusine coracana*) forms (Ganeshaiiah & Umashaanker, 1982).

Conscious selection for selfing is hardly possible, but selfing could evolve under domestication for two reasons. First, selfing could be selected under directional selection because it increases additive genetic variance and thus the response to selection. This argument is similar to that for the evolution of recombination (Barton & Charlesworth, 1998), though there is as yet no formal model. Secondly, gene flow from wild or weed species recurrently introduces wild deleterious alleles into the cultivated pool. If selection and local adaptation are strong (which is the case for domesticated vs wild or weed species), it generates strong outbreeding depression, which will select for a higher selfing rate (Epinat & Lenormand, in press).

However, strong directional selection could also prevent the evolution of selfing rates that are too high. Despite the two-fold cost of outcrossing which should promote the complete fixation of selfing (Fisher, 1941), low outcrossing rates can theoretically be selected for in predominantly selfing popula-

tions under strong directional selection (David *et al.*, 1993). This has been found in an experimental population of barley, which was constructed from intercrosses among 30 varieties and then put into cultivation according to agricultural practices (Kahler *et al.*, 1975). The outcrossing rate, initially very low (0.57%), reached 0.88% 11 generations later and 1.24% 20 generations later ($P < 0.001$). More generally, rapid evolutionary changes and strong directional selection are expected to select for increased recombination rates during the domestication process irrespective of the underlying genetic basis of cultivation alleles (Burt & Bell, 1987; Otto & Barton, 1997). This theory has been confirmed for some crops, which show an excess of chiasma frequencies compare with wild species (Ross-Ibarra, 2004). Most cereals offer a good example of this evolutionary change under domestication (Fig. 5). Eleven out of 13 species exhibit an increase in chiasma frequency from the wild progenitor, by up to 40%.

VII. Conclusion

The advent of genome-wide studies of nucleotide polymorphism in a series of grasses will yield unprecedented amounts of comparative data with which to revisit some of the questions outlined in this review (see for instance the recent publication of the sorghum genome in Paterson *et al.*, 2009). Beyond dissecting the specific story of a given species, more general understanding of the factors affecting the domestication process (e.g. mating systems) could be achieved through comparative domestication approaches. In particular, the rice genome can now be used as a source of candidate genes for domestication in a host of domestic grasses with much more complex genomes such as polyploid wheat or barley. The hypothesis of convergence at the gene level during domestication could now be revisited in a lot more detail.

Some progress has been made in using current patterns of polymorphism within and between wild/cultivated grasses to unravel the evolutionary history of grasses under domestication.

However, making sense of the wealth of empirical comparative and population genomics data will require careful model-based approaches (likelihood or Bayesian) combining more realistic demographic models and incorporating independent information from archaeological data (e.g. Allaby *et al.*, 2008; but see Ross-Ibarra & Gaut, 2008). In this respect, a hierarchical modelling approach would allow the specification of both a historical scenario with a given mating system and a background level of negative selection affecting the whole genome as well as the incorporation of the local effects of directional selection (through hard and soft sweeps). This will hopefully be the way to resolve the current paradox where even fairly different evolutionary scenarios can hardly be distinguished based on the data (e.g. Caicedo *et al.*, 2007). It will also shed more light on the relative importance of new versus standing variation. For instance, reanalysis of large population genomics data sets with the aim specifically of distinguishing among neutrality and hard and soft sweeps, in relation to dominance level analyses, may hold the promise of a better understanding of the processes of selection during domestication. Although progress has been impressive for simple traits such as qualitative differences at the phenotypic level, the challenge now will be to achieve the same degree of understanding for more quantitative traits (although some progress has been made regarding flowering phenology). The model-based population genomics approaches discussed above should be a starting point to solve this difficult issue.

Last but not least, a major challenge in the application of this evolutionary knowledge will be to use it to target efforts to introgress new genetic variation in the gene pools of elite cultivars, which are often lacking variation, and adapt these to a wide range of agrosystems and new challenging growing conditions.

Acknowledgements

This is publication ISE-M 2009-041. TB acknowledges financial support from a Steno fellowship provided by the Danish Research Council (FNU). We thank J. David and Eva H. Stukenbrock for helpful discussions and comments on the manuscript.

References

- Allaby RG, Fuller DQ, Brown TA. 2008. The genetic expectations of a protracted model for the origins of domesticated crops. *Proceedings of the National Academy of Sciences, USA* 105: 13982–13986.
- Badr A, Muller K, Schafer-Pregl R, El Rabey H, Effgen S, Ibrahim HH, Pozzi C, Rohde W, Salamini F. 2000. On the origin and domestication history of barley (*Hordeum vulgare*). *Molecular Biology and Evolution* 17: 499–510.
- Barrett SC, Harder LD, Worley AC. 1996. The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society London B* 351: 1271–1280.
- Barton NH, Charlesworth B. 1998. Why sex and recombination? *Science* 281: 1986–1990.
- Bataillon T. 2000. Estimation of spontaneous genome-wide mutation rate parameters: wither beneficial mutations? *Heredity* 84: 497–501.
- Bouchenak-Khelladi Y, Salamin N, Savolainen V, Forest F, Bank M, Chase MW, Hodkinson TR. 2008. Large multi-gene phylogenetic trees of the grasses (Poaceae): progress towards complete tribal and generic level sampling. *Molecular and Phylogenetic Evolution* 47: 488–505.
- Bremer K. 2002. Gondwanan evolution of the grass alliance of families (Poales). *Evolution* 56: 1374–1387.
- Briggs WH, McMullen MD, Gaut BS, Doebley J. 2007. Linkage mapping of domestication loci in a large maize teosinte backcross resource. *Genetics* 177: 1915–1928.
- Brown TA, Jones MK, Powell W, Allaby RG. 2009. The complex origins of domesticated crops in the Fertile Crescent. *Trends in Ecology and Evolution* 24: 103–109.
- Burger CJ, Chapman MA, Burke JM. 2008. Molecular insight into the evolution of plant crops. *American Journal of Botany* 95: 113–122.
- Burt A, Bell G. 1987. Mammalian chiasma frequencies as a test of two theories of recombination. *Nature* 326: 803–805.
- Cai W, Morishima H. 2002. QTL clusters reflect character associations in wild and cultivated rice. *Theoretical and Applied Genetics* 104: 1217–1228.
- Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR *et al.* 2007. Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genetics* 3: 1745–1756.
- Caldwell KS, Dvorak J, Lagudah ES, Akhunov E, Luo MC, Wolters P, Powell W. 2004. Sequence polymorphism in polyploid wheat and their d-genome diploid ancestor. *Genetics* 167: 941–947.
- Caldwell KS, Russell J, Langridge P, Powell W. 2006. Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. *Genetics* 172: 557–567.
- Charlesworth B. 1992. Evolutionary rates in partially self-fertilizing species. *The American Naturalist* 140: 126–148.
- Cheng C, Motohashi R, Tsuchimoto S, Fukuta Y, Ohtsubo H, Ohtsubo E. 2003. Polyphyletic origin of cultivated rice: based on the interspersed pattern of SINEs. *Molecular Biology and Evolution* 20: 67–75.
- Connor HE. 1981. Evolution of Reproductive Systems in the Gramineae. *Annals of the Missouri Botanical Garden* 68: 48–74.
- Cruz F, Vila C, Webster MT. 2008. The legacy of domestication: accumulation of deleterious mutations in the dog genome. *Molecular Biology and Evolution* 25: 2331–2336.
- Damgaard C. 2000. Fixation of advantageous alleles in partially self-fertilizing populations. The effect of different selection modes. *Genetics* 154: 813–821.
- Darwin C. 1859. *The origin of species*. London, UK: John Murray.
- Darwin C. 1878. *The effects of cross and self-fertilization in the vegetal kingdom*. London, UK: John Murray.
- Darwin C. 1882. *The variation of animals and plants under domestication*. London, UK: John Murray.
- David J, Savy YB, Brabant P. 1993. Outcrossing and selfing evolution in populations under directional selection. *Heredity* 71: 642–651.
- de Alencar Figueiredo LF, Calatayud C, Dupuits C, Billot C, Rami JF, Brunel D, Perrier X, Courtois B, Deu M, Glaszmann JC. 2008. Phylogeographic evidence of crop neodiversity in sorghum. *Genetics* 179: 997–1008.
- de Wet J, Prasada Rao K, Mengesha M, Brink D. 1983. Diversity in Kodo millet, *Paspalum scrobiculatum*. *Economic Botany* 37: 159–163.
- Diamond J. 1999. *Guns, germs, and steel*. New York, NY, USA: W. W. Norton & Company.
- Diamond J. 2002. Evolution, consequences and future of plant and animal domestication. *Nature* 418: 700–707.
- Doebley J, Stec A, Gustus C. 1995. teosinte branched1 and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* 141: 333–346.

- Doebley J. 2004. The genetics of maize evolution. *Annual Review of Genetics* 38: 37–59.
- Doebley JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. *Cell* 127: 1309–1310.
- Doebley JF, Goodman MM, Stuber CW. 1984. Isozymic variation in *Zea* (Gramineae). *Systematic Botany* 9: 203–218.
- Doebley JF, Stec A, Hubbard L. 1997. The evolution of apical dominance in maize. *Nature* 386: 485–488.
- Dorweiler J, Stec A, Kermicle J, Doebley J. 1993. Teosinte glume architecture 1: a genetic locus controlling a key step in maize evolution. *Science* 262: 233–235.
- Doust A. 2007. Architectural evolution and its implications for domestication in grasses. *Annals of Botany* 100: 941–950.
- Doust AN, Devos KM, Gadberrry MD, Gale MD, Kellogg EA. 2004. Genetic control of branching in foxtail millet. *Proceedings of the National Academy of Sciences, USA* 101: 9045–9050.
- Epinat G, Lenormand T. (in press). The evolution of assortative mating and selfing with in- and outbreeding depression. *Evolution*. doi:10.1111/j.1558-5646.2009.00700.x
- Fan L, Quan L, Leng X, Guo X, Hu W, Ruan S, Ma H, Zeng M. 2008. Molecular evidence for post-domestication selection in the *Waxy* gene of Chinese waxy maize. *Molecular Breeding* 22: 329–338.
- Fisher RA. 1941. Average excess and average effect of a gene substitution. *Annals of Eugenism* 11: 53–63.
- Frankel OH, Brown AHD, Burdon JJ. 1995. *The conservation of plant biodiversity*. Cambridge, UK: Cambridge University Press.
- Fuller DQ. 2007. Contrasting patterns in crop domestication and domestication rates: recent archaeobotanical insights from the Old World. *Annals of Botany* 100: 903–924.
- Ganeshiah K, Umashaanker R. 1982. Evolution of reproductive behaviour in the genus *Eleusine*. *Euphytica* 31: 397–404.
- Gao LZ, Innan H. 2008. Nonindependent domestication of the two rice subspecies, *Oryza sativa* ssp. *indica* and ssp. *japonica*, demonstrated by multilocus microsatellites. *Genetics* 179: 965–976.
- Gaut BS. 2002. Evolutionary dynamics of grass genomes. *New Phytologist* 154: 15–28.
- Gaut BS, Clegg MT. 1993. Nucleotide polymorphism in the *Adh1* locus of pearl millet (*Pennisetum glaucum*) (Poaceae). *Genetics* 135: 1091–1097.
- Gaut BS, Doebley JF. 1997. DNA sequence evidence for the segmental allotetraploid origin of maize. *Proceedings of the National Academy of Sciences, USA* 94: 6809–6814.
- Gepts P. 2004. Crop domestication as a long-term selection experiment. In: Janick J, ed. *Plant breeding reviews*. New York, NY, USA: John Wiley & Sons, 1–44.
- Gordo I, Charlesworth B. 2001. Genetic linkage and molecular evolution. *Current Biology* 11: R684–R686.
- (GPWG) GPWG. 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). *Annals of the Missouri Botanical Garden* 88: 373–457.
- Gu Z, David L, Petrov D, Jones T, Davis RW, Steinmetz LM. 2005. Elevated evolutionary rates in the laboratory strain of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences, USA* 102: 1092–1097.
- Gyenis L, Yun SJ, Smith KP, Steffenson BJ, Bossolini E, Sanguineti MC, Muehlbauer GJ. 2007. Genetic architecture of quantitative trait loci associated with morphological and agronomic trait differences in a wild by cultivated barley cross. *Genome* 50: 714–723.
- Hanson MA, Gaut BS, Stec AO, Fuerstenberg SI, Goodman MM, Coe EH, Doebley JF. 1996. Evolution of anthocyanin biosynthesis in maize kernels: the role of regulatory and enzymatic loci. *Genetics* 143: 1395–1407.
- Harlan J, de Wet J, Price E. 1973. Comparative evolution of cereals. *Evolution* 27: 311–325.
- Haudry A, Cenci A, Ravel C, Bataillon T, Brunel D, Poncet C, Hochu I, Poirier S, Santoni S, Glémin S *et al.* 2007. Grinding up wheat: a massive loss of nucleotide diversity since domestication. *Molecular Biology and Evolution* 24: 1506–1517.
- Heun M, Schäfer-Pregl R, Klawan D, Castagna R, Accerbi M, Borghi B, Salamini F. 1997. Site of Einkorn wheat domestication identified by DNA fingerprinting. *Science* 278: 1312–1314.
- Hillman GC, Davies MS. 1990. Domestication rates in wild-type wheats and barley under primitive cultivation. *Biological Journal of the Linnean Society* 39: 39–78.
- Houle D, Morikawa B, Lynch M. 1996. Comparing mutational variabilities. *Genetics* 143: 1467–1483.
- Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P. 2002. Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proceedings of the National Academy of Sciences, USA* 99: 8133–8138.
- Innan H, Kim Y. 2004. Pattern of polymorphism after strong artificial selection in a domestication event. *Proceedings of the National Academy of Sciences, USA* 101: 10667–10672.
- Janssen T, Bremer K. 2004. The age of major monocot groups inferred from 800+ *rbcl* sequences. *Botanical Journal of the Linnean Society* 146: 385–398.
- Jones H, Leigh FJ, Mackay I, Bower MA, Smith LM, Charles MP, Jones G, Jones MK, Brown TA, Powell W. 2008. Population-based resequencing reveals that the flowering time adaptation of cultivated barley originated east of the Fertile Crescent. *Molecular Biology and Evolution* 25: 2211–2219.
- Kahler AL, Clegg MT, Allard RW. 1975. Evolutionary changes in the mating system of an experimental population of barley (*Hordeum vulgare* L.). *Proceedings of the National Academy of Sciences, USA* 72: 943–946.
- Kawase M, Fukunaga K, Kato K. 2005. Diverse origins of waxy foxtail millet crops in East and Southeast Asia mediated by multiple transposable element insertions. *Molecular Genetics and Genomics* 274: 131–140.
- Kilian B, Ozkan H, Kohl J, von Haeseler A, Barale F, Deusch O, Brandolini A, Yucel C, Martin W, Salamini F. 2006. Haplotype structure at seven barley genes: relevance to gene pool bottlenecks, phylogeny of ear type and site of barley domestication. *Molecular Genetics and Genomics* 276: 230–241.
- Kilian B, Ozkan H, Walther A, Kohl J, Dagan T, Salamini F, Martin W. 2007. Molecular diversity at 18 loci in 321 wild and 92 domesticate lines reveal no reduction of nucleotide diversity during *Triticum monococcum* (Einkorn) domestication: implications for the origin of agriculture. *Molecular Biology and Evolution* 24: 2657–2668.
- Kimura M. 1962. On the probability of fixation of mutant genes in a population. *Genetics* 47: 713–719.
- Komatsuda T, Maxim P, Senthil N, Mano Y. 2004. High-density AFLP map of nonbrittle rachis 1 (*btr1*) and 2 (*btr2*) genes in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 109: 986–995.
- Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicker T, Tagiri A *et al.* 2007. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proceedings of the National Academy of Sciences, USA* 104: 1424–1429.
- Kovach MJ, Sweeney MT, McCouch SR. 2007. New insights into the history of rice domestication. *Trends in Genetics* 23: 578–587.
- Le Thierry d'Ennequin M, Toupance B, Robert T, Godelle B, Gouyon PH. 1999. Plant domestication: a model for studying the selection of linkage. *Journal of Evolutionary Biology* 12: 1138–1147.
- Leebens-Mack J, Raubeson LA, Cui L, Kuehl JV, Fourcade MH, Chumley TW, Boore JL, Jansen RK, depamphilis CW. 2005. Identifying the basal angiosperm node in chloroplast genome phylogenies: sampling one's way out of the Felsenstein zone. *Molecular Biology and Evolution* 22: 1948–1963.
- Lenormand T. 2002. Gene flow and the limit to natural selection. *Trends in Ecology & Evolution* 17: 183–189.
- Li C, Zhou A, Sang T. 2006. Rice domestication by reducing shattering. *Science* 311: 1936–1939.

- Li W, Gill BS. 2006. Multiple genetic pathways for seed shattering in the grasses. *Functional & Integrative Genomics* 6: 300–309.
- Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA. 2006. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences, USA* 103: 9578–9583.
- Lu J, Tang T, Tang H, Huang J, Shi S, Wu CI. 2006. The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication. *Trends in Genetics* 22: 126–131.
- Lynch M. 1988. The rate of polygenic mutation. *Genetical Research* 51: 137–148.
- Martin G, Lenormand T. 2006. A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution* 60: 893–907.
- Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez GJ, Buckler E, Doebley J. 2002. A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences, USA* 99: 6080–6084.
- Morgante M, Brunner S, Pea G, Fengler K, Zuccolo A, Rafalski A. 2005. Gene duplication and exon shuffling by helitron-like transposons generate intraspecific diversity in maize. *Nature Genetics* 37: 997–1002.
- Morrell PL, Clegg MT. 2007. Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. *Proceedings of the National Academy of Sciences, USA* 104: 3289–3294.
- Nordborg M, Tavaré S. 2002. Linkage disequilibrium: what history has to tell us. *Trends in Genetics* 18: 83–90.
- Ohta T, Cockerham CC. 1974. Detrimental genes with partial selfing and effects on a neutral locus. *Genetical Research* 23: 191–200.
- Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan MD. 2006. Selection under domestication: evidence for a sweep in the rice waxy genomic region. *Genetics* 173: 975–983.
- Orr HA, Betancourt AJ. 2001. Haldane's sieve and adaptation from the standing genetic variation. *Genetics* 157: 875–884.
- Otto SP, Barton NH. 1997. The evolution of recombination: removing the limits to natural selection. *Genetics* 147: 879–906.
- Oumar I, Mariac C, Pham JL, Vigouroux Y. 2008. Phylogeny and origin of pearl millet (*Pennisetum glaucum* [L.] R. Br) as revealed by microsatellite loci. *Theoretical and Applied Genetics* 117: 489–497.
- Paterson AH. 2002. What has QTL mapping taught us about plant domestication? *New Phytologist* 154: 591–608.
- Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitors T, Poliakov A *et al.* 2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457: 551–556.
- Paterson AH, Lin Y-R, Li Z, Schertz KF, Doebley JF, Pinson SRM, Liu S-C, Stansel JW, Irvine JE. 1995. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science* 269: 1714–1717.
- Peng J, Ronin Y, Fahima T, Roder MS, Li Y, Nevo E, Korol A. 2003. Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. *Proceedings of the National Academy of Sciences, USA* 100: 2489–2494.
- Pennings PS, Hermisson J. 2006. Soft sweeps III: the signature of positive selection from recurrent mutation. *PLoS Genetics* 2: e186.
- Piperno DR, Flannery KV. 2001. The earliest archaeological maize (*Zea mays* L.) from highland Mexico: new accelerator mass spectrometry dates and their implications. *Proceedings of the National Academy of Sciences, USA* 98: 2101–2103.
- Poncet V, Lamy F, Devos KM, Gale MD, Sarr A, Robert T. 2000. Genetic control of domestication traits in pearl millet (*Pennisetum glaucum* L., Poaceae). *Theoretical and Applied Genetics* 100: 147–159.
- Poncet V, Lamy F, Enjalbert J, Joly H, Sarr A, Robert T. 1998. Genetic analysis of the domestication syndrome in pearl millet (*Pennisetum glaucum*): inheritance of the major characters. *Heredity* 81: 648–658.
- Poncet V, Martel E, Allouis S, Devos M, Lamy F, Sarr A, Robert T. 2002. Comparative analysis of QTLs affecting domestication traits between two domesticated × wild pearl millet (*Pennisetum glaucum* L., Poaceae) crosses. *Theoretical and Applied Genetics* 104: 965–975.
- Prasad V, Stromberg CA, Alimohammadian H, Sahni A. 2005. Dinosaur coprolites and the early evolution of grasses and grazers. *Science* 310: 1177–1180.
- Przeworski M, Coop G, Wall JD. 2005. The signature of positive selection on standing genetic variation. *Evolution* 59: 2312–2323.
- Purugganan MD, Fuller DQ. 2009. The nature of selection during plant domestication. *Nature* 457: 843–848.
- Ross-Ibarra J. 2004. The evolution of recombination under domestication: a test of two hypotheses. *The American Naturalist* 163: 105–112.
- Ross-Ibarra J. 2005. Quantitative trait loci and the study of plant domestication. *Genetica* 123: 197–204.
- Ross-Ibarra J, Gaut BS. 2008. Multiple domestications do not appear monophyletic. *Proceedings of the National Academy of Sciences, USA* 105: E105.
- Ross-Ibarra J, Morrell PL, Gaut BS. 2007. Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proceedings of the National Academy of Sciences, USA* 104 (Suppl. 1): 8641–8648.
- Saisho D, Purugganan MD. 2007. Molecular phylogeography of domesticated barley traces expansion of agriculture in the Old World. *Genetics* 177: 1765–1776.
- Salamini F, Ozkan H, Brandolini A, Schafer-Pregl R, Martin W. 2002. Genetics and geography of wild cereal domestication in the near east. *Nature Review Genetics* 3: 429–441.
- Sang T, Ge S. 2007. The puzzle of rice domestication. *Journal of Integrative Plant Biology* 49: 760–768.
- Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M. 2008. Deletion in a gene associated with grain size increased yields during rice domestication. *Nature Genetics* 40: 1023–1028.
- Simmonds NW. 1976. *Evolution of crop plants*. New York, NY, USA: Longman Inc.
- Simons KJ, Fellers JP, Trick HN, Zhang Z, Tai YS, Gill BS, Faris JD. 2006. Molecular characterization of the major wheat domestication gene Q. *Genetics* 172: 547–555.
- Sprague GF, Russell WA, Penny LH. 1960. Mutations affecting quantitative traits in the selfed progeny of doubled monoploid maize stocks. *Genetics* 45: 855–866.
- Sweeney M, McCouch S. 2007. The complex history of the domestication of rice. *Annals of Botany* 100: 951–957.
- Sweeney MT, Thomson MJ, Cho YG, Park YJ, Williamson SH, Bustamante CD, McCouch SR. 2007. Global dissemination of a single mutation conferring white pericarp in rice. *PLoS Genetics* 3: e133.
- Sweeney MT, Thomson MJ, Pfeil BE, McCouch S. 2006. Caught red-handed: Rc encodes a basic helix-loop-helix protein conditioning red pericarp in rice. *The Plant Cell* 18: 283–294.
- Tan L, Li X, Liu F, Sun X, Li C, Zhu Z, Fu Y, Cai H, Wang X, Xie D *et al.* 2008a. Control of a key transition from prostrate to erect growth in rice domestication. *Nature Genetics* 40: 1360–1364.
- Tan L, Zhang P, Liu F, Wang G, Ye S, Zhu Z, Fu Y, Cai H, Sun C. 2008b. Quantitative trait loci underlying domestication- and yield-related traits in an *Oryza sativa* × *Oryza rufipogon* advanced backcross population. *Genome* 51: 692–704.
- Tanno K, Willcox G. 2006. How fast was wild wheat domesticated? *Science* 311: 1886.
- Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES. 2001. *Dwarf8* polymorphisms associate with variation in flowering time. *Nature Genetics* 28: 286–289.
- Thuillet AC, Bataillon T, Poirier S, Santoni S, David JL. 2005. Estimation of long-term effective population sizes through the history of durum wheat using microsatellite data. *Genetics* 169: 1589–1599.
- Thuillet AC, Bru D, David J, Roumet P, Santoni S, Sourdille P, Bataillon T. 2002. Direct estimation of mutation rate for 10 microsatellite

- loci in durum wheat, *Triticum turgidum* (L.) Thell. ssp *durum* Desf. *Molecular Biology and Evolution* 19: 122–125.
- Vigouroux Y, Jaqueth JS, Matsuoka Y, Smith OS, Beavis WD, Smith JS, Doebley J. 2002. Rate and pattern of mutation at microsatellite loci in maize. *Molecular Biology and Evolution* 19: 1251–1260.
- Vogler DW, Kaliz S. 2001. Sex among the flowers: the distribution of plant mating systems. *Evolution* 55: 202–204.
- Wang E, Wang J, Zhu X, Hao W, Wang L, Li Q, Zhang L, He W, Lu B, Lin H *et al.* 2008. Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nature Genetics* 40: 1370–1374.
- Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Faller M, Bomblies K, Lukens L, Doebley JF. 2005. The origin of the naked grains of maize. *Nature* 436: 714–719.
- Wang R, Stec A, Hey J, Lukens L, Doebley J. 1999. The limits of selection during maize domestication. *Nature* 398: 236–239.
- Watson L, Dallwitz MJ. 1992. Grass genera of the world: descriptions, illustrations, identification, and information retrieval, including synonyms, morphology, anatomy, physiology, phytochemistry, cytology, classification, pathogens, world and local distribution, and references. <http://biodiversity.uno.edu/delta/>.
- Weiss E, Kislev ME, Hartmann A. 2006. Autonomous cultivation before domestication. *Science* 312: 1608–1610.
- Whitt SR, Wilson LM, Tenaillon MI, Gaut BS, Buckler EST. 2002. Genetic diversity and selection in the maize starch pathway. *Proceedings of the National Academy of Sciences, USA* 99: 12959–12962.
- Wright HE Jr. 1976. The environmental setting for plant domestication in the Near East. *Science* 194: 385–389.
- Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS. 2005. The effects of artificial selection on the maize genome. *Science* 308: 1310–1314.
- Zhu Q, Ge S. 2005. Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytologist* 167: 249–265.
- Zohary D. 1999. Monophyletic vs. polyphyletic origin of the crops on which agriculture was founded in the Near East. *Genetic Resources and Crop Evolution* 46: 133–142.
- Zohary D. 2004. Unconscious selection and the evolution of domesticated plants. *Economic Botany* 58: 5–10.
- Zohary D, Hopf M. 2000. *Domestication of plants in the old world*. Oxford, UK: Oxford University Press.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 Lists of domesticated cereal grasses (+sugarcane), and of cultivated forage grasses

Table S2 Chiasma frequency data in domesticated grasses and wild progenitors or close relatives

Table S3 List of grass species with known mating systems

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.



About New Phytologist

- *New Phytologist* is owned by a non-profit-making **charitable trust** dedicated to the promotion of plant science, facilitating projects from symposia to open access for our Tansley reviews. Complete information is available at www.newphytologist.org.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as-ready' via *Early View* – our average submission to decision time is just 29 days. Online-only colour is **free**, and essential print colour costs will be met if necessary. We also provide 25 offprints as well as a PDF for each article.
- For online summaries and ToC alerts, go to the website and click on 'Journal online'. You can take out a **personal subscription** to the journal for a fraction of the institutional price. Rates start at £139 in Europe/\$259 in the USA & Canada for the online edition (click on 'Subscribe' at the website).
- If you have any questions, do get in touch with Central Office (newphytol@lancaster.ac.uk; tel +44 1524 594691) or, for a local contact in North America, the US Office (newphytol@ornl.gov; tel +1 865 576 5261).