1 Kernel Evolution: From Teosinte to Maize

Sherry A. Flint-Garcia*
U.S. Department of Agriculture, Agricultural Research Service, Columbia, Missouri, USA

1.1 Introduction

Maize is the most productive and highest value commodity crop in the U.S. and around the world: over 1 billion tons were produced each year in 2013 and 2014 (FAO, 2016). Together, maize, rice, and wheat comprise over 60% of the world's caloric intake (http://www.fao.org). The importance of maize in terms of production and caloric intake is not a recent development. In fact, Native Americans have relied on maize and its ancestor for more than 9000 years. The "Columbian exchange" allowed maize to spread around the world, to adapt to new environments and become a major crop that feeds large portions of the human population. Maize, and the kernel in particular, has undergone dramatic changes over the past 9000 years. The biology of maize seed size and its starch, protein, oil content, and food characteristics, are described in other chapters of this book. Here I review the evolution of maize from teosinte (the wild ancestor) to landraces (locally adapted, open-pollinated farmer varieties) to modern maize (inbreds and hybrids), and discuss changes in kernel composition and size during this process.

1.2 Domestication

Maize, like all the world's major agricultural crop plant and animal species, underwent domestication from a wild relative. The suite of phenotypic traits that were modified during domestication is referred to as the "domestication syndrome" (Hammer, 1984) and usually includes traits related to productivity (e.g. increased seed number and size), harvestability (e.g. non-shattering and fewer seed-bearing structures), and consumption (reduced toxicity and improved palatability) among other species-specific traits (Olsen and Wendel, 2013). Evolution of the seed was central to domestication, as were traits facilitating harvest.

Genetic and archeological evidence suggest maize was domesticated from teosinte (*Zea mays* ssp. *parviglumis*) approximately 9000 years ago in the Central Balsas River Valley in southwestern Mexico in the states of Guerrero and Michoacán (Matsuoka *et al.*, 2002; Piperno *et al.*, 2009). *Zea mays* ssp. *parviglumis* (hereafter parviglumis) is an annual diploid species endemic to southwestern Mexico (Doebley and Iltis, 1980). There are several other species of teosinte with different ploidy levels, perenniality,

*Corresponding author e-mail: Sherry.Flint-Garcia@ARS.USDA.GOV

and/or special regional adaptation to higher elevations or lower latitudes (Fukunaga *et al.*, 2005), but these will not be discussed in any detail. Hereafter, whenever teosinte is mentioned, the reader may assume parviglumis unless otherwise noted.

There are dramatic differences in plant, ear, and kernel morphology between maize and teosinte (reviewed in Doebley, 2004). Parviglumis plants, when grown under the short-day conditions typical of central Mexico, are bushy and composed of many stalks (tillers) with long lateral branches ending in male inflorescences (Fig. 1.1A). In contrast, most modern maize plants are unbranched, with a single stalk and short lateral branches (ear shanks) ending in female inflorescences (Fig. 1.1B). Teosinte plants are capable of producing over 100 ear structures, each of which contains 5 to 12 seeds stacked and without a cob (Fig. 1.1C). Modern maize plants usually produce one or two ears with cobs that bear several hundred kernels in eight or more rows around the ear (Fig. 1.1D). Teosinte kernels are very small (approximately one-tenth the weight of maize kernels) and are enclosed in a hardened fruitcase (Fig. 1.1E) absent in modern maize (Fig. 1.1F). Teosinte ears shatter and disperse their seeds upon maturation, a characteristic absent in maize.

Fig. 1.1. Teosinte (A) and maize (B) differ greatly in terms of number of stalks and male and female inflorescences. Teosinte ears (C) contain 5–12 kernels without the familiar cob structure characteristic of maize (D). The small teosinte seeds (E) are enclosed in a hard fruitcase, while maize kernels (F) are naked and weigh approximately ten times more than those of teosinte.

It is something of a mystery how native peoples of Mexico used teosinte prior to domestication. There were no large domesticated animals in North America at the time, so it is unlikely teosinte was a forage crop. Modern maize is used primarily for grain, and a natural assumption is that teosinte was used similarly. However, its hard fruitcase would be a formidable deterrent, along with the limited amount of food obtained from the small seeds. George Beadle devised a method to create "teo-tortillas" using a primitive metate (grinding stone) and a water-based method to float off the broken fruitcases. Beadle also proposed that natives could have popped teosinte, similar to modern popcorn (Beadle, 1939). Others have proposed Native Americans chewed or sucked out sugars stored in the pithy teosinte stalks (Iltis, 2000) or created fermented beverages (Smalley and Blake, 2003).

1.2.1 Archeological evidence

The oldest archeological ear/cob samples are from 6200 years ago, originating in Guilá Naquitz Cave in Oaxaca (Benz, 2001), and 5500-year-old samples from the San Marcos Cave in the Tehuacán Valley in Puebla (Long *et al.*, 1989). Unfortunately, these samples are too old to bear kernels, but they do show non-shattering cobs with two to four rows of naked (no fruitcase) kernels. The oldest kernel samples, though not intact, include microfossils dated to 8700 years old and found on grinding stones from the Xihuatoxtla Shelter in Guerrero (Piperno *et al.*, 2009). Analysis of starch grains found on these stones revealed maize was the primary species processed and included popcorn and other hard/flinty kernel types. Sequence analysis of ancient DNA obtained from 660–4405-year-old ear samples from New Mexico and Mexico indicated that alleles representative of modern maize were present 4400 years ago (Jaenicke-Després *et al.*, 2003). So, it is clear primitive maize with morphologically distinct ears and kernels, though perhaps not quite resembling modern maize, was grown within a few thousand years of domestication and was an important part of the Native American diet.

1.2.2 The master regulators of domestication

Beginning in the 1800s, there were various hypotheses concerning the origin of corn that involved an extinct progenitor species, teosinte, tripsacum, pod corn, corngrass, and combinations thereof. During the 1930s, debates revolved around the extreme phenotypic differences between maize and teosinte. In an effort to understand inheritance of these differences, Beadle examined the phenotypes of over 50,000 F_2 plants derived from a cross between maize and teosinte (Beadle, 1972). He determined that approximately 1 in 500 plants looked like very teosinte-like, or very maize-like, with a ratio that suggested four or five genes control the main morphological differences between maize and teosinte.

Indeed, Beadle's calculation of a handful of genes has been largely supported by quantitative trait locus (QTL) mapping studies of morphological differences between maize and teosinte. In an F_2 population derived from a cross of a maize landrace with a more distantly related teosinte subspecies (*Zea mays* ssp. *mexicana*, hereafter mexicana), six major QTLs (chromosomes 1–5) were found to underlie key traits that differentiate maize and teosinte: lateral branch length and inflorescence architecture, and secondary sex traits such as the hard fruitcase and paired floral spikelets (Doebley *et al.*, 1990). The QTL analysis of a second F_2 population derived from a primitive landrace crossed with parviglumis revealed the same genomic regions, suggesting domestication from teosinte to a primitive maize landrace could be accomplished by modifying a few key genes or gene regions (Doebley and Stec, 1993).

Since then, several QTL have been fine mapped and cloned, revealing the importance of transcription factors controlling key steps in domestication. The important regulator of apical dominance, *teosinte branched 1*

(*tb1*), is located on the long arm of chromosome 1 (Doebley *et al.*, 1995). The domesticated allele of this transcription factor contains a *Hopscotch* transposable element 63 kb upstream of the start codon (Studer *et al.*, 2011) that results in higher expression of a lateral branch repressor (Doebley *et al.*, 1997). Thus, maize represses growth of lateral branches, resulting in fewer tillers. Also on chromosome 1 (short arm) is a QTL controlling prolificacy: in teosinte, the long lateral branches bear many ears, while the maize lateral branch bears a single terminal ear. The QTL controlling prolificacy was fine mapped to *grassy tillers 1*, a homeodomain leucine zipper transcription factor (Wills *et al.*, 2013) that was previously demonstrated to control tillering (Whipple *et al.*, 2011). The QTL on chromosome 5 originally thought to be a master controller of a number of ear-related traits (kernel row number, ear diameter, pedicellate spikelet length, and shattering) fractionated into multiple independent factors (Lemmon and Doebley, 2014). More recently, fine mapping and cloning of a shattering QTL in sorghum identified a YABBY-like transcription factor as a candidate gene for the QTL on chromosome 5 (Lin *et al.*, 2012). The genes responsible for the QTLs on chromosomes 2 and 3 have yet to be cloned.

The QTL on chromosome 4 is of particular interest to kernel evolution, since it controls development of the hardened fruitcase enclosing the teosinte seed and is absent or severely reduced in maize. The QTL underlying this trait, *teosinte glume architecture 1*, was mapped to chromosome 4 (Dorweiler *et al.*, 1993) and encodes a transcription factor in the squamosa promoter binding-protein family (Wang *et al.*, 2005); the causative lesion was later determined to be a single amino acid change affecting dimerization (Wang *et al.*, 2015). In teosinte, the fruitcase is composed of (i) a cup-shaped segment of the stem, the "cupule," in which the seed is seated, and (ii) a hardened bract or glume that is hinged onto the cupule that completely encloses the seed. The maize allele represses formation of these structures, such that the cupule and glume no longer surround the seed; these structures were evolutionarily repurposed to form the hard sections of the maize cob.

1.2.3 A thousand small effect genes underlie domestication

While QTL studies are useful as a forward genetics approach to determine genomic regions underlying a phenotype, reverse genetics approaches can be used to scan the genome for signatures of selection that could result in a phenotype related to the domestication syndrome. Selection during domestication results in a reduction of nucleotide diversity relative to the progenitor and an excess of rare variants as populations recover from selection, and can be measured using a variety of population genetic statistics. For example, an analysis of sequence diversity of 21 genes on chromosome 1 revealed only *tb1* as a target of selection (Tenaillon *et al.*, 2001).

A large-scale selection scan suggested approximately 2–4% of maize genes could have been targets of selection during domestication and/or modern breeding (Wright *et al.*, 2005). Assuming 35,000 genes in maize, this translates to 700–1400 genes that could be responsible for the transformation of teosinte into modern maize. Using the HapMap2 dataset of 55 million single nucleotide polymorphisms (SNPs) (Chia *et al.*, 2012), Hufford *et al.* (2012) found approximately 1000 genes experienced selection, with the strongest selection occurring during domestication rather than during modern breeding. The finding that so many genes were involved in domestication obviously conflicts with the five-gene hypothesis of Beadle (1939) and the early QTL mapping studies by the Doebley lab. But this paradox can be resolved by invoking the theory that a handful of master regulators can orchestrate a cascade involving intermediate and small effect genes that control a wide range of traits targeted by domestication.

1.3 Modern Breeding

As primitive corn was carried from central Mexico, north and south across the Americas, the outbreeding nature of maize and large population sizes allowed maize to adapt to new environments, e.g. day-length,

climate, soil types, and human uses (dietary preferences and religious purposes). For example, gene flow from mexicana, a highland teosinte, allowed maize to adapt to higher elevations within Mexico (van Heerwaarden *et al.*, 2011). Maize moved into the Southwestern USA by 4000 years ago, initially via a highland route through Mexico, followed approximately 2000 years later by gene flow from lowland races from the Pacific coast (Fonseca *et al.*, 2015). From the Southwestern USA, maize spread north to Canada (Vigouroux *et al.*, 2008) and became the dominant crop species of North America by 800 AD (Smith, 1989). For the southward expansion, highland maize spread to the lowland tropics of southern Mexico and Guatemala, through the Isthmus of Panama, and into Colombia. From Colombia, maize spread to the Caribbean via the Lesser Antilles and also into the rest of South America, including an independent adaptation to highlands of the Andes (Takuno *et al.*, 2015). Maize was carried to Europe, Asia, and Africa by Columbus and the early explorers, and continued to adapt (Mir *et al.*, 2013). Each landrace has distinct plant, ear, and kernel characteristics that have been used to identify and classify them (Goodman and Brown, 1988) and define their uses around the world.

Maize inbreeding began at the end the 1800s and subsequent hybridization of the early cycle inbreds (Shull, 1909) led to the hybrid seed industry and evolution of heterotic groups. Today, in the U.S. Corn Belt, there are three main heterotic groups: stiff stalks, non-stiff stalks, and iodents (Troyer, 1999). Breeding programs usually focus on specific traits relevant to the target environment: cold tolerance for northern climates, drought tolerance for the high plains, disease and insect resistance in the south, etc.

1.3.1 Dent corn

The vast majority of corn grown in the U.S. is a commodity referred to as "Number 2 Yellow Dent." In general, yield is the primary driver of dent corn, and seed quality is of secondary importance. There are regions of the USA that cater to specialty food-grade dent corn markets, such as white food corn, where producers contract their crop directly to processors and for which white food corn varieties were tested until 2002 (Darrah *et al.*, 2002). While all teosintes have white endosperm, there is wide variability in landraces and inbred lines for endosperm color, including orange and yellow (from carotenoids) and red and purple (from anthocyanins). Yellow predominates in commodity corn due to the higher nutritional value of carotenoids for animal feed, while white is preferred for human consumption in many regions around the world (Poneleit, 2001). A survey of the *y1* (*phytoene synthase*) locus revealed classic signatures of selection, in particular much lower diversity in yellow relative to white lines (Palaisa *et al.*, 2003). Anthocyanin kernel pigments appear to have been targeted by post-domestication selection for the ability to produce red and purple pigments via the *colored aleurone 1* locus (Hanson *et al.*, 1996). Together, these results suggest kernel color traits were targets of selection.

The most recognizable types of food corn are sweet corn and popcorn, where flavor and kernel quality are of highest importance. Another example, baby corn, is simply an immature ear harvested as silks begin developing; it is primarily produced in Thailand (Aekatasanawan, 2001). Each of these specialty corns has a different set of ear-kernel phenotypes and underlying genetics, some of which is discussed in detail in other chapters of this book. There has been continued evolution, breeding, and refinement of the genetics underlying these kernel phenotypes, and breeding efforts have kept the associated germplasm separate. Phylogenetic analysis of the NC7 (Ames, IA) Plant Introduction Station collection of 2800 maize inbred lines showed clear germplasm separation (Romay *et al.*, 2013): the popcorn and sweet corn accessions form very distinct germplasm groups; the stiff stalk and non-stiff stalk inbreds within the temperate germplasm have intermediate separation from each other; the tropical germplasm also forms a very distinct group. Analysis of marker data for inbred lines divided by era showed continued separation of the major heterotic groups of corn belt maize

and decreased diversity in the ancestry of the heterotic pools (van Heerwaarden *et al.*, 2012).

1.3.2 Sweet corn

Cultures across the Americas have eaten "green corn" for millennia, enjoying standard starchy corn that is picked at the "milk stage" of kernel development. Green corn is not a result of sweet corn mutations, but rather owes its low-level sweetness to sugars not yet converted to starch. Modern sweet corn is the result of precise breeding, utilizing mutations in the starch biosynthetic pathway (Chapter 12) to produce specific market classes of sweet corn ranging from the original sugary varieties to the newer synergistic, augmented, and supersweet varieties. There are only eight genes used in commercial sweet corn production, with three predominating the market at present (reviewed in Tracy, 1994): *sugary 1* (*su1*) mutations affect a starch debranching enzyme, resulting in phytoglycogen accumulation; *sugary enhancer 1* (*se1*) has an unknown function, but causes the sweet phenotype when used in conjunction with *su1* (Schultz and Juvik, 2004); *shrunken 2* (*sh2*) mutations block all complex carbohydrates (starch and phytoglycogen), causing an accumulation of sugars. While not widely grown as compared to non-sugary varieties, sweet corn (primarily *su1* types) has been grown and consumed in confections and alcoholic beverages since before the arrival of Columbus (Wellhausen *et al.*, 1952).

Among commercially important sweet corn mutations, *su1* has an interesting evolutionary history related to the diffusion of landraces across the Americas. Sequence analysis of 57 accessions of *su1* germplasm from six geographic regions of the Americas revealed five independent origins of *su1* sweet corn (Tracy *et al.*, 2006). Of these, three different alleles are caused by single amino acid changes in conserved residues of what is considered the active site of the isoamylase enzyme, and are spatially clustered in Northwestern Mexico and throughout the U.S. A fourth allele was caused by a transposon insertion in the first exon, and was found in two Mexican Maiz Dulce accessions. The causative lesions could not be determined for the fifth allele, which was identified in two Peruvian highland accessions of Chullpi. Selection for and maintenance of the first *sugary 1* mutations by Native Americans led to the success of modern breeding for additional mutations and secondary flavor and texture traits. The starch mutants were found in limited genetic resources, originating from the ancestral group of "Northern Flints" and resulting in the tight population structure of the U.S. maize germplasm collection, as discussed earlier (Romay *et al.*, 2013).

1.3.3 Popcorn

Popcorn is another favorite food corn around the world. The primary traits that make popcorn unique are the explosion of the kernel upon exposure to heat and the subsequent expansion of starch to form large "flakes" (reviewed in Ziegler, 1994). During popping, the moisture contained in the kernel expands until the pericarp can no longer withstand the pressure and bursts. Starch of the hard endosperm gelatinizes with the released steam, expands due to heat, and dries and hardens into flakes. Flake production is related to a higher ratio of hard to soft starch and a thicker pericarp that can withstand building pressure from steam, traits absent from dent corn. While popcorn kernel colors range from yellow and white (the most commercially important) to red, blue, purple and nearly black, there are only two kernel shapes: rice types with long, slender kernels and a pointed tip; and pearl types with round kernels and a smooth top. Once popped, there are two main flake shapes (with intermediate variation) that appear to be under genetic control: butterfly flakes are irregularly shaped but with many wings; mushroom flakes are round with only a few wings.

As discussed earlier, Native Americans probably enjoyed pop-teosinte prior to domestication. It is likely many primitive

landraces were popcorns selected from earlier flint types for larger popping expansion. By the time of Columbus, popcorn was prevalent in both North and South America. As popcorn became a distinct industry in the 1880s (Erwin, 1949), modern breeding methods were employed to improve agronomic traits and popcorn-specific traits: pericarp strength, popping volume, and flavor. Interestingly, a single gene has played a key role in maintaining distinct popcorn germplasm—the gametophyte factor known as *ga1*. The dominant strong allele, *Ga1-s*, which confers nearly perfect cross-incompatibility with non *Ga1-s* pollen, is present in nearly all modern popcorn germplasm (Nelson, 1952). While this gene does not affect kernel phenotypes per se, it does maintain the already distinct popcorn kernel phenotypes by preventing pollen contamination by dent maize, which typically carries the *ga1* allele.

1.4 Seed Size and Kernel Composition

It is clear that the kernel was a central focus during domestication and breeding—humans selected large seeds that are easy to harvest and consume. In the course of evolution, there have been drastic changes in seed composition. The typical chemical composition of teosinte, landraces, and inbred lines is shown in Table 1.1. Of note is the large increase in starch (34%) and large decrease in protein (–58%) during domestication (Flint-Garcia *et al.*, 2009a). Since these values are expressed as a percentage of total kernel weight, it is no surprise that various traits are correlated, regardless of the underlying biochemistry. The biology, genetics, and biochemistry of kernel composition traits and seed size are described in other chapters of this book. The objective here is to discuss evolution of these traits, which are intertwined with other traits.

1.4.1 Seed size

Increasing seed size/weight was undoubtedly valuable to the survival and prosperity of early Native Americans. Indeed, maize kernels (either landraces or modern inbred lines, excepting popcorns) weigh almost ten times more than teosinte seeds (Flint-Garcia *et al.*, 2009a), and this increase occurred during domestication. After selection to reduce and open up the fruitcase, primarily acting through *tga1*, seed volume was no longer limited by space inside the fruitcase. Enlarged seed size was probably the most important domestication trait to Native Americans, but very little is known about the genetics underlying the evolution of the process. In a QTL analysis of the same landrace × teosinte F_2 populations described earlier (Doebley *et al.*, 1990; Doebley and Stec, 1993), six and four QTL were found to control seed weight during the transition from teosinte to landraces, where all the teosinte alleles decreased seed weight (Doebley *et al.*, 1994). In a backcross 1-derived mapping population of parviglumis in the W22 background, six QTLs were identified for kernel weight (Briggs *et al.*, 2007). A similar result of a handful of QTLs controlling seed weight was also seen in a population of near isogenic lines (NILs) derived from ten parviglumis donors in the B73 background (Liu *et al.*, 2016); there was a total of eight QTLs across the entire population, with a range of

Table 1.1. Kernel composition and seed traits for a panel of teosinte (parviglumis) accessions, landraces, and inbred lines. Data summarized from Flint-Garcia *et al.* (2009a).

Germplasm	*N*	Protein %	Fat %	Fiber %	Ash %	Carbohydrate %	Seed Wt. (g)	Percent endosperm
Teosinte	11	28.71	5.61	0.91	2.24	52.92	0.03	90.18
Landraces	17	12.13	4.40	1.75	1.55	71.16	0.28	90.13
Inbred lines	27	11.11	4.12	1.80	1.40	72.37	0.26	91.85

two to six QTLs per donor. Many of the QTLs identified in these studies overlapped, and, as expected, the majority of the teosinte alleles caused a decrease in seed weight; however, one of the teosinte alleles for the QTL on chromosome 2 appears to increase seed weight (Liu *et al.* 2016). While this allelic effect remains to be validated, its potential use in breeding is attractive.

There has been limited progress identifying genes underlying teosinte kernel weight QTLs and establishing that they are related to domestication. Interestingly, *prolamin-box binding factor 1* (*pbf1*) is a strong candidate for a QTL on chromosome 2, and it will be discussed below in Section 1.4.3 on kernel proteins. For a QTL on chromosome 1, a gene with homology to *GS3* from rice was proposed as a selection candidate in maize, as *OsGS3* was found to be a domestication gene controlling grain size in rice (Takano-Kai *et al.*, 2009). Although the maize ortholog of *GS3* has lower sequence diversity in maize than teosinte, selection tests revealed it is a neutrally-evolving gene (Li *et al.*, 2010) and did not play a role in kernel evolution from teosinte, despite being a potential candidate gene underlying kernel weight.

1.4.2 Starch

Starch synthesis and accumulation in the seed involves a complex biochemical system with an array of sugars and starches, a number of plant organs and structures, and temporal regulation (Chapter 12). To explain the system briefly, and in a highly oversimplified way, a series of enzymes including sucrose synthases (e.g. *shrunken 1*) and invertases (e.g. *mn1*) break down the sucrose entering the endosperm via the basal endosperm transfer layer (BETL) into glucose and fructose; a series of enzymes including ADP-glucose pyrophosphorylase (e.g. *brittle 2=bt2* and *shrunken 2*) convert the glucose to ADP-glucose; and finally starch synthases (e.g. *waxy 1*), starch branching enzymes (e.g. *amylose extender 1=ae1*) and debranching enzymes (e.g. *su1*) act on the ADP-glucose to convert it into the two primary forms of starch (Chapters 5 and 12).

Population genetic analysis of six genes in the starch pathway revealed that three genes — *bt2*, *su1*, and *ae1* — show a signature of selection. This suggests that the starch pathway was targeted by selection (Whitt *et al.*, 2002). However, because DNA sequence data were collected from inbred lines and teosinte accessions, but no landraces, it was difficult to determine whether selection occurred during domestication or during breeding. Recently, an analysis of 348 genes in archeological landrace samples from the Southwestern USA dating back to 750–4000 years ago and Mexican samples dating back to 1400–5900 years ago showed selection for several composition genes, including *ae1* and particularly *su1* (Fonseca *et al.*, 2015). The results of this study suggest selection on *su1* was more recent, approximately 1000–1200 years ago, which coincided with the appearance of larger cobs and floury endosperm texture. Both of these genes (*ae1* and *su1*) affect the structure of amylopectin and are involved in pasting properties important for making porridge and tortillas (Whitt *et al.*, 2002; Wilson *et al.*, 2004). Again, it is not a surprise that starch synthesis was affected by domestication, because as seed size increased, starch content also increased.

1.4.3 Protein

The nature of proteins in the maize kernel is described in Chapter 14. Briefly, approximately 10–20% of the proteins are globulins found in the embryo; the remaining 80–90% occur in the endosperm. Prolamins, or zeins (α, β, γ, and δ), are the principal endosperm storage proteins and are found in protein bodies (Boston and Larkins, 2009). Native Americans developed a process called "nixtamalization," in which corn kernels were soaked in an alkaline solution (lime; calcium hydroxide) prior to cooking. This process allows easy removal of the pericarp and improves texture by gelatinizing the starch; most importantly, it improves

the nutritional value of the resulting masa by degrading the protein bodies and releasing niacin (vitamin B3) (Gomez *et al.*, 1989). Without this treatment, diets based largely on maize lead to a skin disease known as Pellagra.

Swarup *et al.* (1995) found that exotic maize and wild members of the genus *Zea* exhibit higher levels of methionine-rich δ-zeins than maize inbreds, leading the authors to hypothesize that the high methionine trait was lost in the course of domestication. Indeed, an HPLC-based survey of the zein profiles in a panel of teosinte, landrace, and inbred accessions showed higher levels of δ-zeins as well as β-zeins in landraces and teosinte (Flint-Garcia *et al.*, 2009a). A number of classical kernel mutants affect zein synthesis and/or formation of protein bodies. For example, *opaque 2* encodes a bZIP transcription factor that, when mutated, results in a severe reduction of the lysine-poor zeins and a concomitant increase in other storage proteins and free amino acids, including lysine (Schmidt *et al.*, 1990). *Opaque 1*, *floury 1*, and *floury 2* are all involved in aspects of zein trafficking in the endoplasmic reticulum. There is no evidence these genes or any of the zein genes were selected during domestication or breeding (Hufford *et al.*, 2012).

Several of the zeins (27 kDa γ-zein and 22 kDa α-zein) are regulated by *pbf1*, an endosperm-specific transcription factor (Vicente-Carbajosa *et al.*, 1997). DNA sequence analysis of *pbf1* in 660–4405-year-old ear samples from New Mexico and Mexico showed the modern maize haplotype was nearly fixed in these landrace samples (Jaenicke-Després *et al.*, 2003). This evidence of a selective sweep strongly suggests protein quality could have been under selection. The absence of a knockout mutant in *pbf1* suggests this gene is critical. Lang *et al.* (2014) used heterozygosity in a NIL carrying a teosinte *pbf1* allele to determine the target trait. They found twofold higher expression of the teosinte *pbf1* allele and a slight increase in seed weight, but no change in zein composition. This positive allelic effect on seed weight was not seen in the original maize × teosinte QTL study (Doebley *et al.*, 1994), but is consistent with the effect we observed for one of our ten donors (Liu *et al.*, 2016). The authors of the former study hypothesized that the reduction in seed weight from the maize allele was a negative pleiotropic effect of selection at *pbf1* for some unknown aspect of kernel composition.

Because zeins are so abundant, they impact the amino acid composition of the kernel, limiting the content of the essential amino acids lysine, tryptophan, and methionine (Prasanna *et al.*, 2001). However, there is variability in free amino acids (Moro *et al.*, 1996). In two large-scale selection scans, three genes involved in amino acid metabolism were identified as being selected (Wright *et al.*, 2005; Yamasaki *et al.*, 2005): *chorismate mutase*, *cysteine synthase*, and *dihydrodipicolinate synthase*. These results prompted an in-depth analysis of amino acid pathways (Flint-Garcia *et al.*, 2009b). Of the 15 additional amino acid metabolism genes tested, only four showed weak evidence of selection: *aspartate kinase – homoserine dehydrogenase 1 – AK domain*, *glutamate dehydrogenase*, *proline dehydrogenase*, and *sam synthetase II*. However, none of the selected genes cluster in pathways that make a convincing argument for evolutionary selection.

1.4.4 Oil

The typical maize kernel contains 4.3–4.5% oil, a high energy component of the grain. Generally, the mature embryo is 10% of the total kernel mass and contains about 85% of the kernel lipids, primarily as triacylglycerols (Chapter 13). In a survey of kernel traits across *Zea mays* germplasm, there was a significant decrease (–26%) in kernel oil content between teosinte and maize landraces/inbred lines (Flint-Garcia *et al.*, 2009a). Although the reduction in oil content during domestication (–21%) was small compared to the starch increase and protein decrease, it represents a major change in kernel composition. Interestingly, no change was found in the endosperm-to-embryo ratio between teosinte and landraces, suggesting it may be

possible to increase oil content by using teosinte alleles without a negative pleiotropic effect of increased embryo size.

One of the best characterized QTLs for kernel oil content is on chromosome 6 (Laurie *et al.*, 2004). It was mapped to a BAC with five genes, one of which is *DGAT1-2* (Zheng *et al.*, 2008). In the 2008 study, an association analysis identified a 3-bp insertion at position 469, resulting in an extra phenylalanine (F469) as the causative factor conferring high oil. The F469 allele was found in all teosinte accessions analyzed, and thus is considered ancestral (Zheng *et al.*, 2008). A follow-up study showed the high-oil allele is present in most of the Southwestern USA, Northern Flint, and Southern Dent landraces, at a moderate frequency in Corn Belt Dent, and nearly absent in the early inbred lines. Two hypotheses were offered to explain diversity at *DGAT1-2*: (i) the high oil F469 allele was lost due to genetic drift when a small number of Corn Belt Dent populations were chosen to develop inbred lines; or (ii) the F469 allele was selected against because of pleiotropy with other favorable agronomic traits, such as high starch content (Chai *et al.*, 2012). Indeed, *DGAT1-2* was associated with both oil and starch content in the Nested Association Mapping population (Cook *et al.*, 2012).

One unappealing aspect of using genome-wide selection scans as a reverse-genetic approach is that there may not be an immediate connection with the target trait. Among the 48 genes identified as selection candidates by Wright *et al.* (2005) and Yamasaki *et al.* (2005), most did not have obvious target traits associated with the gene. In an effort to identify the phenotypic effects associated with these selected genes, 32 genes were tested in an association analysis of two teosinte populations scored for a panel of phenotypic traits (Weber *et al.*, 2009). Interestingly, a gene with homology to an ankyrin-repeat-like protein, *AY106616*, associated most strongly with kernel oil content, but also with starch content. The ankyrin-repeat-like protein is involved in carbohydrate metabolism and allocation in tobacco and Arabidopsis (Weber *et al.*, 2009); thus, a plausible target trait for carbon cycling within the kernel has been established.

1.5 Lingering Questions and Prospects for Maize Improvement

The evolutionary history of the maize kernel presents geneticists and breeders with a series of questions from how domestication occurred to prospects for maize improvement.

1.5.1 Relationships between composition and seed size traits

As noted, there are correlations among many of the size and kernel composition traits, especially between germplasm groups: teosinte, landraces, and inbred lines (Flint-Garcia *et al.*, 2009a). For example, there is positive correlation of seed weight with kernel starch content, which begs the question from an evolutionary perspective: which came first, the chicken or the egg? Did liberation of the seed from the fruitcase allow the kernel to expand in size due to a subsequent increase in starch accumulation? Or, did selection for high starch alleles occur first and help drive expansion of the seed out of the fruitcase? Would reintroduction of all the fruitcase alleles (*tga1* and other minor QTLs, if any) limit the size of the kernel and change kernel composition, e.g. decreased starch and increased protein and oil?

The question of pleiotropy versus linkage of QTLs is not an evolution-specific one, but it is still very relevant. Because composition and seed size traits are so highly correlated, are there specific genes that mechanistically contribute to variation for multiple traits? Or are there multiple genes linked (tightly or not) in a single QTL that control different traits independently? Can these traits be manipulated independently?

1.5.2 How many of the 1000 selected genes are involved in kernel traits?

Seed size was obviously an important trait during domestication, and one would expect a large number of the 1000 selected genes could influence seed size genes (Chapter 16). Alternatively, because of the strong correlations between seed size and

composition traits, one could also expect a large number of the selected genes to be kernel composition genes. The genome-wide selection scan of Hufford *et al.* (2012) provided an excellent starting point to answer this question; however, in my opinion, poor genome annotation has been the primary impediment of progress. Of the 1000 selected genes, the vast majority are not annotated. Nevertheless, a simple query of the selection candidates in Hufford *et al.* (2012) using the 464 genes from the classical gene list (Schnable and Freeling, 2011) identified eight interesting new selection candidates that could be involved in kernel traits (Table 1.2). These genes can be tested rigorously for signatures of selection (e.g. HKA tests, coalescent simulations, etc.) and their phenotypic effects determined in both maize and teosinte germplasm.

1.5.3 Do teosinte alleles have value for improving corn?

Long ago—9000 years—humans began modifying teosinte to improve harvestability. Selection resulted in reduced genetic variation in genes underlying these traits; consequently, modern maize shows little variation. Additionally, every gene across the genome has lost some diversity because of demographic events (bottlenecks, random sampling, etc.), even if these are neutrally-evolving genes.

Today, we are growing corn in very different environments using different agronomic practices than those practiced 9000 years ago during domestication, or 1000 years ago as corn became the predominant crop in the USA or even 100 years ago when modern breeding began. Traits that were relevant 9000, 1000, or 100 years ago may not be useful today; therefore, alleles selected 9000, 1000, or 100 years ago that persist in modern germplasm may not be optimal today. This reduction in genetic variation is irreversible—especially if the current practice of recycling germplasm in breeding programs is continued—unless of course variation is reintroduced from teosinte and/or landraces.

A straightforward goal would be to try to modify our current corn for specific traits. Novel sources of genetic resistance to the foliar diseases grey leaf spot (Lennon *et al.*, 2016) and southern leaf blight (Lennon *et al.*, 2017) were identified in parviglumis. Introgression of mexicana into maize resulted in lines with significantly higher protein content, as well as higher lysine, methionine, and/or phenylalanine content (Wang *et al.*, 2008). Thus, teosinte has potential to improve many traits in maize.

If we strive for the more extreme goal of introducing large portions of the teosinte genome into modern maize germplasm, what genes/alleles should we target? Genes showing signatures of selection would provide the greatest return on investment, as they harbor allelic diversity in teosinte not present in maize. Clearly, we do not want the hard fruitcase trait back, so we will avoid *tga1*! However, perhaps a plant with a single ear is not the best ideotype in today's agronomic system where we no longer harvest

Table 1.2. Potential new selection candidates with effects on kernel traits. Results were obtained by merging the candidate gene lists from Hufford *et al.* (2012) with the Classical Gene List (Schnable and Freeling, 2011).

Gene ID	Gene name	Possible target trait
GRMZM2G348551	*su2; sugary 2*	Starch
GRMZM2G394450	*ivr1; invertase 1*	Starch
GRMZM2G089836	*ivr2; invertase 2*	Starch
GRMZM2G110175	*bm1; brown midrib 1*	Starch
AC196475.3_FG004	*bm3; brown midrib 3*	Starch
GRMZM2G098298	*ccp1; cysteine protease 1*	Protein
GRMZM2G138727	*zp27; 27-kDa zein protein*	Protein & Amino acids
GRMZM2G087612	*SDP1; sugar dependent1*	Oil

corn manually and where combines are capable of harvesting many ears per plant. Reintroducing the branching and prolificacy alleles at *tb1* and *gt1* from teosinte would be first steps to increase prolificacy. However, reintroduction of the teosinte alleles will likely disrupt the source–sink balance (see Chapter 16) that has been established in modern germplasm. Incorporating teosinte alleles of the various starch biosynthetic genes could also be useful in reprogramming corn.

One interesting question to ask: if we had a thousand years to rerun a domestication experiment, using our knowledge of plant biology, genetics, and breeding/statistics and specifically the genes that have been selected to create the crop we currently call corn, would we be able to re-domesticate a "new corn" from teosinte with the optimal alleles for our environmental conditions and agronomic practices?

References

Aekatasanawan, C. (2001) Baby corn. In: Hallauer, A.R. (ed.) *Specialty Corns* (2nd edn.). CRC Press, Boca Raton, Florida, pp. 275–292.

Beadle, G.W. (1939) Teosinte and the origin of maize. *Journal of Heredity* 30, 245–247.

Beadle, G.W. (1972) The mystery of maize. *Field Museum of Natural History Bulletin* 43, 2–11.

Benz, B.F. (2001) Archaeological evidence of teosinte domestication from Guilá Naquitz, Oaxaca. *Proceedings of the National Academy of Sciences of the United States of America* 98, 2104–2106.

Boston, R.S. and Larkins, B.A. (2009) The genetics and biochemistry of maize zein storage proteins. In: Bennetzen, J.L. and Hake, S. (eds.) *Handbook of Maize: Genetics and Genomics*. Springer, New York, pp. 715–730.

Briggs, W.H., McMullen, M.D., Gaut, B.S. and Doebley, J. (2007) Linkage mapping of domestication loci in a large maize teosinte backcross resource. *Genetics* 177, 1915–1928.

Chai, Y., Hao, X., Yang, X., Allen, W.B., Li, J., *et al.* (2012) Validation of DGAT1-2 polymorphisms associated with oil content and development of functional markers for molecular breeding of high-oil maize. *Molecular Breeding* 29, 939–949.

Chia, J.-M., Song, C., Bradbury, P.J., Costich, D., de Leon, N., *et al.* (2012) Maize HapMap2 identifies extant variation from a genome in flux. *Nature Genetics* 40, 803–807.

Cook, J.P., McMullen, M.D., Holland, J.B., Tian, F., Bradbury, P., *et al.* (2012) Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiology* 158, 824–834.

Darrah, L.L., Maddux, L.D., Hibbard, B.E., Wilmont, D.B., Lee, E.A., *et al.* (2002) White food corn 2002 performance tests. *Special Report 547*, USDA-ARS and Agricultural Experiment Station, University of Missouri-Columbia.

Doebley, J. (2004) The genetics of maize evolution. *Annual Review of Genetics* 38, 37–59.

Doebley, J. and Iltis, H.H. (1980) Taxonomy of *Zea* (Gramineae). I. A subgeneric classification with key to taxa. *American Journal of Botany* 67, 982–993.

Doebley, J. and Stec, A. (1993) Inheritance of the morphological differences between maize and teosinte: comparison of results for two F_2 populations. *Genetics* 134, 559–570.

Doebley, J., Stec, A., Wendel, J. and Edwards, M. (1990) Genetic and morphological analysis of a maize-teosinte F_2 population: implications for the origin of maize. *Proceedings of the National Academy of Sciences of the United States of America* 87, 9888–9892.

Doebley, J., Bacigalupo, A. and Stec, A. (1994) Inheritance of kernel weight in two maize-teosinte hybrid populations: implications for crop evolution. *Journal of Heredity* 85, 191–195.

Doebley, J., Stec, A. and Gustus, C. (1995) *Teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* 141, 333–346.

Doebley, J., Stec, A. and Hubbard, L. (1997) The evolution of apical dominance in maize. *Nature* 386, 485–488.

Dorweiler, J., Stec, A., Kermicle, J. and Doebley, J. (1993) Teosinte glume architecture 1: a genetic locus controlling a key step in maize evolution. *Science* 262, 233–235.

Erwin, A.T. (1949) The origin and history of popcorn, *Zea mays* L. var. *indurata* (Sturt.) Bailey mut. *everta* (Sturt.) Erwin. *Agronomy Journal* 41, 53–56.

FAO (2016) Statistics at FAO. Available at: http://faostat.fao.org (accessed December 14, 2016).
Flint-Garcia, S.A., Bodnar, A.L. and Scott, M.P. (2009a) Wide variability in kernel composition, seed characteristics, and zein profiles among diverse maize inbreds, landraces, and teosinte. *Theoretical and Applied Genetics* 119, 1129–1142.
Flint-Garcia, S.A., Guill, K.E., Sanchez-Villeda, H., Schroeder, S.G. and McMullen, M.D. (2009b) Maize amino acid pathways maintain high levels of genetic diversity. *Maydica* 54, 375–386.
Fonseca, R.R.D., Smith, B.D., Wales, N., Cappellini, E., Skoglund, P., *et al.* (2015) The origin and evolution of maize in the southwestern United States. *Nature Plants* 1, 14003.
Fukunaga, K., Hill, J., Vigouroux, Y., Matsuoka, Y., Sanchez, G., *et al.* (2005) Genetic diversity and population structure of teosinte. *Genetics* 169, 2241–2254.
Gomez, M.H., McDonough, C.M., Rooney, L.W. and Waniska, R.D. (1989) Changes in corn and sorghum during nixtamalization and tortilla baking. *Journal of Food Science* 54, 330–336.
Goodman, M.M. and Brown, W.L. (1988) Races of corn. In: Sprague, G.F. and Dudley, J.W. (eds.) *Corn and Corn Improvement Agronomy No. 18, Third Edition*. ASA-CSSA-SSSA, Madison, Wisconsin, pp. 33–79.
Hammer, K. (1984) The domestication syndrome. *Die Kulturpflanze* 32, 11–34.
Hanson, M.A., Gaut, B.S., Stec, A.O., Fuerstenberg, S.I., Goodman, M.M., *et al.* (1996) Evolution of anthocyanin biosynthesis in maize kernels: the role of regulatory and enzymatic loci. *Genetics* 143, 1395–1407.
Hufford, M.B., Xu, X., van Heerwaarden, J., Pyhajarvi, T., Chia, J.-M., *et al.* (2012) Comparative population genomics of maize domestication and improvement. *Nature Genetics* 44, 808–811.
Iltis, H.H. (2000) Homeotic sexual translocations and the origin of maize (*Zea mays*, Poaceae): a new look at an old problem. *Economic Botany* 54, 7–42.
Jaenicke-Després, V., Buckler, E.S., Smith, B.D., Gilbert, M.T.P., Cooper, A., *et al.* (2003) Early allelic selection in maize as revealed by ancient DNA. *Science* 302, 1206–1208.
Lang, Z., Wills, D.M., Lemmon, Z.H., Shannon, L.M., Bukowski, R., *et al.* (2014) Defining the role of prolamin-box binding factor1 gene during maize domestication. *Journal of Heredity* 105, 576–582.
Laurie, C.C., Chasalow, S.D., LeDeaux, J.R., McCarroll, R., Bush, D., *et al.* (2004) The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. *Genetics* 168, 2141–2155.
Lemmon, Z.H. and Doebley, J.F. (2014) Genetic dissection of a genomic region with pleiotropic effects on domestication traits in maize reveals multiple linked QTL. *Genetics* 198, 345–353.
Lennon, J.R., Krakowsky, M., Goodman, M., Flint-Garcia, S. and Balint-Kurti, P.J. (2016) Identification of alleles conferring resistance to gray leaf spot in maize derived from its wild progenitor species teosinte. *Crop Science* 56, 209–218.
Lennon, J.R., Krakowsky, M.D., Goodman, M., Flint-Garcia, S. and Balint-Kurti, P.J. (2017) Identification of teosinte (*Zea mays* ssp. parviglumis) alleles for resistance to southern leaf blight in near isogenic maize lines. *Crop Science* 57, 1973–1983. DOI:10.2135/cropsci2016.12.0979
Li, Q., Yang, X., Bai, G., Warburton, M.L., Mahuku, G., *et al.* (2010) Cloning and characterization of a putative *GS3* ortholog involved in maize kernel development. *Theoretical and Applied Genetics* 120, 753–763.
Lin, Z., Li, X., Shannon, L.M., Yeh, C.-T., Wang, M.L., *et al.* (2012) Parallel domestication of the Shattering1 genes in cereals. *Nature Genetics* 44, 720–724.
Liu, Z., Cook, J., Melia-Hancock, S., Guill, K., Bottoms, C., *et al.* (2016) Expanding maize genetic resources with predomestication alleles: maize–teosinte introgression populations. *The Plant Genome* 9. DOI: 10.3855/plantgenome2015.07.0053
Long, A.B., Benz, B.F., Donahue, D.J., Jull, A.J.T. and Toolin, L.J. (1989) First direct AMS dates on early maize from Tehuacán, Mexico. *Radiocarbon* 31, 1035–1040.
Matsuoka, Y., Vigouroux, Y., Goodman, M.M., Sanchez, G.J., Buckler, E., *et al.* (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences of the United States of America* 99, 6080–6084.
Mir, C., Zerjal, T., Combes, V., Dumas, F., Madur, D., *et al.* (2013) Out of America: tracing the genetic footprints of the global diffusion of maize. *Theoretical and Applied Genetics* 126, 2671–2682.
Moro, G.L., Habben, J.E., Hamaker, B.R. and Larkins, B.A. (1996) Characterization of the variability in lysine content for normal and opaque2 maize endosperm. *Crop Science* 36, 1651–1659.
Nelson, O.E. (1952) Non-reciprocal cross-sterility in maize. *Genetics* 37, 101–124.
Olsen, K.M. and Wendel, J.F. (2013) A bountiful harvest: genomic insights into crop domestication phenotypes. *Annual Review of Plant Biology* 64, 47–70.

Palaisa, K.A., Morgante, M., Williams, M. and Rafalski, A. (2003) Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *Plant Cell* 15, 1795–1806.

Piperno, D.R., Ranere, A.J., Holst, I., Iriarte, J. and Dickau, R. (2009) Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Sciences of the United States of America* 106, 5019–5024.

Poneleit C.G. (2001) Breeding white endosperm corn. In: Hallauer, A.R. (ed.) *Specialty Corns* (2nd edn.). CRC Press, Boca Raton, Florida, pp. 235–273.

Prasanna, B.M., Vasal, S.K., Kassahun, B. and Singh, N.N. (2001) Quality protein maize. *Current Science* 81, 1308–1319.

Romay, M.C., Millard, M.J., Glaubitz, J.C., Peiffer, J.A., Swarts, K.L., *et al.* (2013) Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biology* 14, R55.

Schmidt, R.J., Burr, F.A., Aukerman, M.J. and Burr, B. (1990) Maize regulatory gene *opaque-2* encodes a protein with a "leucine-zipper" motif that binds to zein DNA. *Proceedings of the National Academy of Sciences of the United States of America* 87, 46–50.

Schnable, J.C. and Freeling, M. (2011) Genes identified by visible mutant phenotypes show increased bias toward one of two subgenomes of maize. *PLOS ONE* 6, e17855.

Schultz, J.A. and Juvik, J.A. (2004) Current models for starch synthesis and the *sugary enhancer1* (*se1*) mutation in *Zea mays*. *Plant Physiology and Biochemistry* 42, 457–464.

Shull, G.H. (1909) A pure line method of corn breeding. *American Breeders Association Report* 5, 51–59.

Smalley, J. and Blake, M. (2003) Sweet beginnings: stalk sugar and the domestication of maize. *Current Anthropology* 44, 675–703.

Smith, B.D. (1989) Origins of agriculture in eastern North America. *Science* 246, 1566–1571.

Studer, A., Zhao, Q., Ross-Ibarra, J. and Doebley, J. (2011) Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nature Genetics* 43, 1160–1163.

Swarup, S., Timmermans, M.C., Chaudhuri, S. and Messing, J. (1995) Determinants of the high-methionine trait in wild and exotic germplasm may have escaped selection during early cultivation of maize. *Plant Journal* 8, 359–368.

Takano-Kai, N., Jiang, H., Kubo, T., Sweeney, M., Matsumoto, T., *et al.* (2009) Evolutionary history of *GS3*, a gene conferring grain length in rice. *Genetics* 182, 1323–1334.

Takuno, S., Ralph, P., Swarts, K., Elshire, R.J., Glaubitz, J.C., *et al.* (2015) Independent molecular basis of convergent highland adaptation in maize. *Genetics* 200, 1297–1312.

Tenaillon, M.I., Sawkins, M.C., Long, A.D., Gaut, R.L., Doebley, J.F., *et al.* (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proceedings of the National Academy of Sciences of the United States of America* 98, 9161–9166.

Tracy, W.F. (1994) Sweet corn. In: Hallauer, A.R. (ed.) *Specialty Corns*. CRC Press, Boca Raton, Florida, pp. 147–187.

Tracy, W.F., Whitt, S.R. and Buckler, E.S. (2006) Recurrent mutation and genome evolution: example of *Sugary1* and the origin of sweet maize. *Crop Science* 46, S49–S54.

Troyer, A.F. (1999) Background of U.S. hybrid corn. *Crop Science* 39, 601–626.

van Heerwaarden, J., Doebley, J., Briggs, W.H., Glaubitz, J.C., Goodman, M.M., *et al.* (2011) Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences of the United States of America* 108, 1088–1092.

van Heerwaarden, J., Hufford, M.B. and Ross-Ibarra, J. (2012) Historical genomics of North American maize. *Proceedings of the National Academy of Sciences of the United States of America* 109, 12420–12425.

Vicente-Carbajosa, J., Moose, S.P., Parsons, R.L. and Schmidt, R.J. (1997) A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. *Proceedings of the National Academy of Sciences of the United States of America* 94, 7685–7690.

Vigouroux, V., Glaubitz, J.C., Matsuoka, Y., Goodman, M.M., Sánchez, G., *et al.* (2008) Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. *American Journal of Botany* 95, 1240–1253.

Wang, H., Nussbaum-Wagler, T., Li, B., Zhao, Q., Vigouroux, Y., *et al.* (2005) The origin of the naked grains of maize. *Nature* 436, 714–719.

Wang, H., Studer, A.J., Zhao, Q., Meeley, R. and Doebley, J.F. (2015) Evidence that the origin of naked kernels during maize domestication was caused by a single amino acid substitution in *tga1*. *Genetics* 200, 965–974.

Wang, L., Xu, C., Qu, M. and Zhang, J. (2008) Kernel amino acid composition and protein content of introgression lines from *Zea mays* ssp. *mexicana* into cultivated maize. *Journal of Cereal Science* 48, 387–393.

Weber, A.L., Zhao, Q., McMullen, M.D. and Doebley, J.F. (2009) Using association mapping in teosinte to investigate the function of maize selection-candidate genes. *PLOS ONE* 4, e8227.

Wellhausen, E.J., Roberts, L.M. and Hernandez, X.E. (1952) *Races of Maize in Mexico*. Bussey Institute, Harvard University, Cambridge, Massachusetts.

Whipple, C.J., Kebrom, T.H., Weber, A.L., Yang, F., Hall, D., *et al.* (2011) *grassy tillers1* promotes apical dominance in maize and responds to shade signals in the grasses. *Proceedings of the National Academy of Sciences of the United States of America* 108, E506–E512.

Whitt, S.R., Wilson, L.M., Tenaillon, M.I., Gaut, B.S. and Buckler, E.S. (2002) Genetic diversity and selection in the maize starch pathway. *Proceedings of the National Academy of Sciences of the United States of America* 99, 12959–12962.

Wills, D.M., Whipple, C.J., Takuno, S., Kursel, L.E., Shannon, L.M., *et al.* (2013) From many, one: genetic control of prolificacy during maize domestication. *PLOS Genetics* 9, e1003604.

Wilson, L.M., Whitt, S.R., Ibáñez, A.M., Rocheford, T.R., Goodman, M.M., *et al.* (2004) Dissection of maize kernel composition and starch production by candidate gene association. *Plant Cell* 16, 2719–2733.

Wright, S.I., Vroh Bi, I., Schroeder, S.G., Yamasaki, M., Doebley, J.F., *et al.* (2005) The effects of artificial selection on the maize genome. *Science* 308, 1310–1314.

Yamasaki, M., Tenaillon, M.I., Vroh Bi, I., Schroeder, S.G., Sanchez-Villeda, H., *et al.* (2005) A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *Plant Cell* 17, 2859–2872.

Zheng, P., Allen, W.B., Roesler, K., Williams, M.E., Zhang, S., *et al.* (2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nature Genetics* 40, 367–372.

Ziegler, K.E. (1994) Popcorn. In: Hallauer, A.R. (ed.) *Specialty Corns*. CRC Press, Boca Raton, Florida, pp. 189–223.