The Genetic Architecture of Maize Flowering Time

Edward S. Buckler,1,2,3* James B. Holland,1,4 Peter J. Bradbury,1,2 Charlotte B. Acharya,2 Patrick J. Brown,2,3 Chris Browne,1,3 Ethan Ersoz,1,4 Sherry Flint-Garcia,1,5 Arturo Garcia,1,5 Jeffrey C. Glabitsch,2 Major M. Goodman,4 Carlos Harjes,2 Kate Guilt,1,5 Dallas E. Kroon,2,3 Sara Larsson,2 Nicholas K. Lepak,1,3 Huili Li,8,2,9 Sharon E. Mitchell,2 Gael Pressoir,3 Jason A. Peiffer,4 Marco Oropesa Rosas,4 Torbert R. Rocheford,10,11 M. Cinta Romay,2,12 Susan Romero,2 Stella Salvo,1,4 Hector Sanchez Villeda,5,13 H. Sofia da Silva,10 Qiong Fan,14 Feng Tian,2 Narasimham Upadhyaya,10 Doreen Ware,1,15 Heather Yates,2 Jianming Yu,16 Zhiwu Zhang,2 Stephen Kresovich,2,16 Michael D. McMullen1,5*

Flowering time is a complex trait that controls adaptation of plants to their local environment in the outcrossing species Zea mays (maize). We dissected variation for flowering time with a set of 5000 recombinant inbred lines (maize Nested Association Mapping population, NAM). Nearly a million plants were assayed in eight environments but showed no evidence for any single large-effect quantitative trait loci (QTLs). Instead, we identified evidence for numerous small-effect QTLS shared among families; however, allelic effects differ across founder lines. We identified no individual QTLS at which allelic effects are determined by geographic origin or large effects for epistasis or environmental interactions. Thus, a simple additive model accurately predicts flowering time for maize, in contrast to the genetic architecture observed in the selfing plant species rice and Arabidopsis.

The nature of standing genetic variation and its relation to phenotypic variation in plants affects our understanding of evolution (1), sustainable agriculture, and preservation of inter- and intraspecific variation in times of environmental change. Maize inbred lines have an average nucleotide diversity in genic regions around 1% (π = 1 to 1.4%) (2, 3),
Flowering time reflects the adaptation of a plant to its environment by tailoring vegetative and reproductive growth phases to local climatic effects. Maize landraces vary widely, from 2 to 11 months, for the time required to mature (7). In addition, asynchrony of male and female flowering in maize may be adaptive in some cultivars, but can result in losses under drought conditions, especially in modern uniform varieties (8). Flowering time has been extensively studied in the predominantly self-fertilizing species Arabidopsis. Like maize, Arabidopsis grows across a wide range of latitudes and has flowering time controlled by the interaction of the photoperiod (light sensing and circadian rhythm), vernalization, and autonomous flowering and gibberellic acid–response pathways (9, 10). In grasses, which include maize, wheat, and rice, some of the same genes are involved in flowering, but they have different functions (11–13).

In maize, diversity-based dissection of flowering time has been hindered by tight linkage similar to the divergence between humans and chimpanzees (4). It is not uncommon to find maize haplotypes that are 5% divergent from one another (5), which indicates that the maize gene pool reaches back 2 to 4 million years (with one generation per year).

Maize is adapted to a range of environments from the lowland tropics to the Andean highlands and has been widely introduced worldwide into both temperate and tropical regions. Maize’s genetic architecture for flowering time has evolved as its wild relatives adapted to distinct ecological zones in elevations differing by more than 3000 m in Mexico and then under both natural and artificial selection over the last 7000 years, with especially intense selection over the past century. This genetic architecture has evolved under a predominantly outcrossing mating system in a species with little population differentiation (6).

Experimental design. We used the maize NAM population (16) of 200 recombinant inbred lines (RILs) from 25 crosses between diverse inbred lines and B73 (each referred to as a family), which resulted in a total of 5000 lines (17). Because maize has rapid linkage disequilibrium (LD) decay, joint linkage analysis of the maize NAM population was used to evaluate complex trait genetic architecture, as we have insufficient marker density for a genome-wide association study (GWAS). The 5000 lines plus 500 checks—nearly one million plants—were evaluated in four locations over 2 years. We scored days to silking (DS, female flowering) and days to anthesis (DA, male flowering), and we calculated the anthesis-silking interval (ASI). We estimated heritability to be 94% for DS and DA and 78% for ASI; within-cross heritability averaged 83 to 84% for DS and DA and 68% for ASI (table S1). Overall, our phenotypic data were highly heritable, and substantial replication across environments reduced environmental effects on the phenotypic mean values of the lines.

Genetic architecture of flowering. We mapped QTLs both within the 25 families separately (using stepwise regression and inclusive composite interval mapping) and in joint analyses that combined information across all families (joint stepwise regression and joint inclusive composite interval mapping, JICIM (17)). These methods produced concordant results in terms of the magnitude of effects; however, they have different power and resolution capabilities. Joint linkage QTL analysis identifies nearly twice as many significant effects compared with individual family analyses. The multiple-family joint stepwise regression method identified 36 and 39 QTLs that explained 89% of the total variance for DA or DS, respectively, whereas 29 QTLs explained 64% of the ASI variance (Fig. 1, top). JICIM generally found evidence for an additional 20 minor QTLs for each trait (Fig. 1, bottom). These findings are concordant with the evidence for 50 or more QTLs identified here.

In general, the vast majority of flowering QTLs we identified showed largely consistent results across environments. Although 59% of QTLs had significant environmental (E) interactions, the genetic variation was many times larger than QTLE interaction variation (Fig. 3A). Overall, this suggests a stable genetic architecture across environments. ASI appeared more sensitive to genotype-by-environment interactions (G×E) than the other flowering phenotypes observed here. Our testing environments had substantial differences in temperatures and rainfall, but day lengths were consistently longer than the critical photoperiod for short-day maize. Therefore, we...
expect that G×E interactions may be stronger if tested under both short and long day length environments.

This study was able to investigate pleiotropy by correlating the allelic effects on multiple traits of each QTL across a robust sample of founders. We observed that 100% of DS and DA QTLs have correlated effects on both male and female flowering (average \( r = 0.90 \) across all loci). In contrast, only about 70% of ASI QTLs had correlated effects on DS, and only 14 to 21% of the ASI QTLs had correlated effects on DA. Overall, genetic control of male and female flowers appears to involve the same set of genes (although magnitudes of effects likely vary). However, the asymmetry between DA-ASI and DS-ASI correlations resulted from the higher variation in the DS phenotype.

We used the significant NAM QTL additive effect estimates to predict the timing of flowering of the NAM founder lines and were able to accurately predict parental flowering time (\( R^2 = 87 \) to 91%) (Fig. 3B). This suggests that NAM QTL results are more reliable than individual family QTL effect estimates (25) and provides further evidence that epistasis is relatively unimportant. By including nonsignificant additive effect estimates in the model, analogous to genome selection, our predictive ability increases to \( R^2 = 95\% \). Although we cannot extend these predictions to unrelated lines because of insufficient marker density, our results suggest that with large enough samples, additive QTL models can accurately predict phenotype.

Our maize founder lines represent a wide range of latitudinal variation (tropical ↔ temperate dent ↔ Northern Flint). We tested the correlation between QTL effect estimates of each founder allele with quantitative estimates of their relation to families of origin to determine whether population structure is defined by the allelic effects at any individual QTL (26). QTL effects at 26% of the loci were correlated (\( P < 0.05 \)) with the tropical-temperatecline. The large effects observed for the QTLs on chromosome 10 were highly correlated with tropical origin, yet only 3 of the 16 lines with substantial tropical origin carried alleles at this locus that increased time to flowering by more than 0.4 days. Overall, allelic effects at many loci were weakly correlated with population structure, but tropical origin was not defined by specific QTL alleles; instead, it appears that numerous loci work in concert to produce latitudinal adaptations (Fig. 4A).

QTL and allele frequency. Thirty percent of the polymorphisms in maize were found to be unique to a single founder line (Fig. 4B), which indicates that rare sequence variants are common in diverse maize. We tested whether the phenotypic variation observed across families was mostly due to many rare variants (segregating in only one family) or to a smaller number of loci causing variation in multiple families. As each founder line gave rise to 200 offspring RILs, represented by about 40,000 plants, our design provided sufficient power to detect rare QTLs and to distinguish between these alternative
hypotheses (>90% power for 1-day effect even in single-family analysis). Most QTLs were shared among multiple families (Fig. 4B), with many QTLs showing effects among seven to eight families (30% frequency). Our data partially support the common gene hypothesis for flowering-time genetic architecture, which proposes that variation at common loci causes phenotypic variation across different families. This result is striking because the sharing of QTLs across families contrasts with the high frequency of rare single-nucleotide polymorphisms (SNPs) in maize (Fig. 4B). This discrepancy cannot be due to bias in detecting QTLs of modest frequency for several reasons. (i) Although lower than that for common QTLs, our design provides enough power to detect QTLs segregating in two to four families; nevertheless, we observed few QTLs distributed in this way. (ii) NAM can statistically detect QTLs unique to B73 (common QTLs in this reference design, but rare QTL alleles in the species). However, only 1% of the QTLs were found in 17 or more families compared with 10% of the SNPs. (iii) Additional QTLs can be identified within individual families, but when they were added to the joint family analysis models, they showed significant effects in additional families. And (iv), retesting the final joint population QTLs model by jackknifing the families (leaving one or two families out sequentially) resulted in reduced significance for some of the QTLs, but none became insignificant.

Although many QTLs appear to be shared across families, we also found evidence for allelic series at most loci. Because our founders were crossed to a common reference line, we tested for and observed allelic series, including both positive and negative effects, at the same locus for 69 to 72% of the QTLs (Fig. 4A), depending on the trait. Such allelic series have previously been observed in maize (27). Although rare alleles dispersed across multiple tightly linked QTLs may also be misclassified as an allelic series in some cases, our association analysis suggests an allelic series for flowering-time effects at vgt1 (below). Our results suggest a model of common genes with uncommon variants controlling flowering to explain our observation of a relatively small number of QTLs (e.g., <100), with many functionally distinct alleles at each locus, each occurring at low frequency. GWAS studies and fine-mapping multiple alleles per QTL will be needed to test this hypothesis.

Genes underlying this architecture. To evaluate the power and reliability of NAM, we
tested a QTL previously shown to affect flowering time in maize. The maize vgt1 locus contains an AP2-like gene, rap2.7, involved in the timing of flowering under regulatory control of an enhancer region about 70 kb from the gene (15). We confirmed the presence of a QTL in the region of vgt1 (DA: P = 4 × 10^{-44}, DS: P = 7 × 10^{-40}). Prior QTL mapping efforts (28) revealed an early flowering QTL about 5 centimorgans (cM) from vgt1 that also may be contributing to the effects detected in this region; marker saturation across this interval permitted resolution into two linked QTLs. By controlling for the rest of the genome and estimating the effects of founder alleles at vgt1, we observed a distinct allelic series at this locus (Fig. 5).

A previously identified vgt1 allele from northern germplasm associated with a miniature transposon (MITE) (15) was segregating in four of the NAM families (crosses involving II14H, P39, MS71, and Mo17). In confirmation of previous results, this allele was strongly associated with early flowering in NAM (Fig. 5). Absence of this MITE did not explain the late flowering vgt1 alleles, but sequencing of founder lines in this region identified SNPs at the rap2.7 gene itself that were associated with this late-flowering effect (Fig. 5). Associations of both the MITE at the upstream regulator region (vgt1) and the SNPs within the rap2.7 gene with flowering time were confirmed in a separate diverse maize inbred association panel (MITE: P = 6 × 10^{-4}, rap2.7 SNP: P = 0.01).

Natural variation at the zfl2 locus also significantly affects flowering time (29). Although the null mutants of this gene previously observed in natural populations (29) did not segregate among these NAM families, we nevertheless found a QTL at the zfl2 locus (DA: P = 3 × 10^{-10}, DS: P = 1 × 10^{-15}). One line, Ky21, had a phenotype associated with a 16-amino acid deletion in the proline-rich domain of this protein. In addition, other lines showing large effects have a 7-bp deletion 1 bp before the ATG start site of zfl2.

Marker saturation of RILs with recombinant chromosome blocks around the largest-effect QTL on chromosome 1 resolved the QTL to a region that includes a homolog of the recently cloned Ghd7 gene from rice (12). On chromosome 1, the bfi2 gene, which is involved in auxin transport (30), overlaps with a QTL for ASI. Association analysis in unrelated lines implicates bfi2 in the timing of flowering (31). In addition, maize flowering mutant idl and homologs of the barley photoperiod gene (Ppd-H1) (21) fall within our QTL intervals. As there are over 1000 homologs of Arabidopsis flowering-time genes in maize, our study demonstrates a means by which to reliably relate maize QTLs to candidate genes.

Implications. Our study of QTLs controlling flowering time with NAM provides insight into the genetic architecture of adaptive traits. Our results suggest that for the outcrossing species maize, the genetic architecture of flowering time is dominated by small additive QTLs with few genetic or environmental interactions (within the tested range of environments). Human height may have a similar genetic architecture (32), but in the case of flowering time, these architectures are distinct from Arabidopsis and rice, self-fertilizing plant species, where flowering-time variation is controlled by fewer genes with larger effects, epistasis, and environmental interactions (9, 11, 13, 23). This suggests that the mating system and demographics influence the genetic architecture of adaptive traits.

For maize, we now have some of the best genetic tools to conduct research of complex genetic architectures. We currently have nearly 15,000 genetic stocks for manipulation and isolation of the genetic variation throughout the entire species. We predict that soon it will be possible to more fully examine the genetic architecture of other traits of interest in maize. Such studies can be applied to improving the world’s food security and to making maize production more environmentally sustainable.

References and Notes
17. Materials and methods are available as supporting material on Science Online.
18. C. C. Laurie et al., Genetics 166, 2141 (2004).
31. We thank the Buckler, Holland, Kresovich, McMullen, and Rocheford labs for their efforts in creating the populations, phenotyping the lines, and scientific input; L. R. Liorette for technical editing of the manuscript; and J. Dobeyles for managing the NSF project. This work was supported by U.S. NSF Plant Genome Program (DBI-9872631, DBI-0321467, DBI-0820619 to E.S.B., M.M.G. [98,03], J.B.H. [03,08], M.D.M. [03,08], S.K. [DBI-9872631, DBI-0321467, DBI-0820619 to E.S.B., M.M.G. (98,03), J.B.H. (03,08), M.D.M. (03,08), S.K. (03,08), D.A.L. (98,03), S.K. (08,03) to S.K.]).
32. Materials and methods are available as supporting material on Science Online.