

Megabase-Scale Inversion Polymorphism in the Wild Ancestor of Maize

Zhou Fang,* Tanja Pyhäjärvi,[†] Allison L. Weber,^{*,1} R. Kelly Dawe,[§] Jeffrey C. Glaubitz,**
 José de Jesus Sánchez González,** Claudia Ross-Ibarra,[†] John Doebley,[‡] Peter L. Morrell,^{*,2}
 and Jeffrey Ross-Ibarra^{*,†,2}

*Department of Agronomy and Plant Genetics, University of Minnesota, Saint Paul, Minnesota 55108, [†]Department of Plant Sciences and ^{**}Center for Population Biology and the Genome Center, University of California, Davis, California 95616, [‡]Department of Genetics, University of Wisconsin, Madison, Wisconsin 53706, [§]Department of Plant Biology, University of Georgia, Athens, Georgia 30602, ^{**}Institute for Genome Diversity, Cornell University, Ithaca, New York 14853, and ^{††}Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara, Zapopan, Jalisco, Mexico CP45110

ABSTRACT Chromosomal inversions are thought to play a special role in local adaptation, through dramatic suppression of recombination, which favors the maintenance of locally adapted alleles. However, relatively few inversions have been characterized in population genomic data. On the basis of single-nucleotide polymorphism (SNP) genotyping across a large panel of *Zea mays*, we have identified an ~50-Mb region on the short arm of chromosome 1 where patterns of polymorphism are highly consistent with a polymorphic paracentric inversion that captures >700 genes. Comparison to other taxa in *Zea* and *Tripsacum* suggests that the derived, inverted state is present only in the wild *Z. mays* subspecies *parviglumis* and *mexicana* and is completely absent in domesticated maize. Patterns of polymorphism suggest that the inversion is ancient and geographically widespread in *parviglumis*. Cytological screens find little evidence for inversion loops, suggesting that inversion heterozygotes may suffer few crossover-induced fitness consequences. The inversion polymorphism shows evidence of adaptive evolution, including a strong altitudinal cline, a statistical association with environmental variables and phenotypic traits, and a skewed haplotype frequency spectrum for inverted alleles.

THE evolutionary role of chromosomal inversions has been studied in a wide array of organisms, from insects (Ayala *et al.* 2011; Stevison *et al.* 2011) to birds (Huynh *et al.* 2011) and plants (Hoffmann and Rieseberg 2008; Lowry and Willis 2010). Examination of inversion polymorphism was fundamental to the early study of selection and adaptive diversity, as well as the basis for understanding the maintenance of neutral polymorphism within populations (Dobzhansky 1950; Hoffmann *et al.* 2004). Homologous pairing of an inverted and a noninverted chromosome in heterozygotes leads to the formation of an inversion loop,

and crossing over in an inversion loop can cause the formation of a dicentric chromosome and an acentric fragment at meiosis I, resulting in terminal deletions of the affected chromosome and gamete death at frequencies that correlate with the size of the inversion (Burnham 1962). Because of the difficulty of homologous pairing and the deleterious effects of homologous crossing over in inversions, inversions are typically observed to disrupt recombination in heterozygous individuals, leading to measurable effects on nucleotide sequence polymorphism, including the generation of extended linkage disequilibrium (LD). Inversion-induced LD has been reported in a variety of organisms, including humans (Bansal *et al.* 2007), *Drosophila subobscura* (Munte *et al.* 2005), and several other species (reviewed in Hoffmann and Rieseberg 2008). Strong differentiation between chromosomal arrangements (as measured by F_{ST}) has also been used as evidence of inversions in *Drosophila* (Andolfatto *et al.* 1999; Depaulis *et al.* 1999; Nóbrega *et al.* 2008).

A variety of circumstances can favor the maintenance or spread of an inversion polymorphism. The inversion may be

Copyright © 2012 by the Genetics Society of America
 doi: 10.1534/genetics.112.138578

Manuscript received January 10, 2012; accepted for publication April 20, 2012
 Supporting information is available online at <http://www.genetics.org/content/suppl/2012/04/27/genetics.112.138578.DC1>.

¹Present address: Department of Genetics, Box 7614, North Carolina State University, Raleigh, NC 27695.

²Corresponding authors: Department of Agronomy and Plant Genetics, 411 Borlaug Hall, University of Minnesota, Saint Paul, MN 55108. E-mail: pmorrell@umn.edu; and Department of Plant Sciences, 262 Robbins Hall, University of California, Davis, CA 95616. E-mail: rossibarra@ucdavis.edu

selected for if the structural rearrangement itself has fitness consequences (Castermans *et al.* 2007). Natural selection can also favor the spread of an inversion if it contains locally adapted alleles, because inversions can suppress recombination and thus protect adaptive alleles from gene flow (Kirkpatrick and Barton 2006; Machado *et al.* 2007). Some inversion polymorphisms display strong patterns of geographic structure, consistent with local adaptation to ecological factors such as temperature regimes or water availability (White *et al.* 2009; Lowry and Willis 2010; Ayala *et al.* 2011). For example, strong differentiation among ecological zones was observed for an inversion in the mosquito *Anopheles funestus* (Ayala *et al.* 2011). In the yellow monkeyflower, *Mimulus guttatus* (Lowry and Willis 2010), an inversion is involved in local adaptation to Mediterranean habitats through several morphological and phenological traits, while the standard arrangement appears in a perennial ecotype from habitats with high year-round soil moisture. In addition to selection, inversion polymorphisms without strong deleterious effects may also increase in frequency through genetic drift and migration, potentially resulting in fixation in small populations (Bengtsson and Bodmer 1976; Lande 1984).

Here we examine the population-level diversity of a newly discovered 50-Mb inversion found in the wild subspecies of *Zea mays* (known collectively as teosinte). Inversions in both wild and domesticated *Z. mays* and related taxa have been reported previously (McClintock 1931; Morgan 1950; Ting 1965, 1967; Kato 1975; Ting 1976), but these were detected cytologically and little is known about their evolution in natural populations. We examine genome-wide patterns of LD in *Z. mays*, using 941 single-nucleotide polymorphism (SNP) markers genotyped in a diverse sample of 2782 individuals, including representatives of three *Z. mays* subspecies: domesticated maize (*Z. mays* ssp. *mays*), its wild progenitor *Z. mays* ssp. *parviglumis*, and the weedy taxon *Z. mays* ssp. *mexicana* (hereafter *mays*, *parviglumis*, and *mexicana*, respectively). A region spanning ~50 Mb on the short arm of chromosome 1 in *parviglumis* and *mexicana* demonstrates the highest level of LD in the genome and coincides with a region of high differentiation between *mays* and *parviglumis* reported by Hufford *et al.* (2012). Comparison to other taxa in *Zea* and *Tripsacum* suggests that the inverted arrangement is derived. The inverted arrangement is present at population frequencies up to 90% in *parviglumis*, but completely absent in domesticated maize. We present evidence that the inversion is relatively ancient and has persisted in teosinte populations for sufficient time to permit its widespread occurrence in all 33 natural populations of subspecies *parviglumis* investigated. Our data further suggest the inversion may be adaptive, as the inverted arrangement shows a strong altitudinal cline and is associated with multiple environmental and phenotypic traits, and the haplotype frequency spectrum of inverted alleles appears inconsistent with neutral evolution.

Materials and Methods

Plant materials and genetic data

Plant materials (Supporting Information, Table S1) included accessions of all four subspecies of *Z. mays* (1573 ssp. *mays*, 975 ssp. *parviglumis*, 161 ssp. *mexicana*, and 10 ssp. *huehuetenangensis*), as well as *Z. luxurians* (17), *Z. diploperennis* (15), and *Z. perennis* (9). The panel also included 22 *Tripsacum* accessions used as outgroups. The 975 *parviglumis* accessions include 33 populations with at least 10 individuals in each population. The *mays* samples include 1283 accessions representing ~250 traditional open-pollinated landraces (including 27 inbred landraces) and 290 modern inbred lines (Table S1).

Genotyping for the genome-wide set of 959 SNPs followed previously described methods (Weber *et al.* 2007; van Heerwaarden *et al.* 2010). The SNP discovery panel consisted of 14 *mays* inbred lines and 16 teosinte partial inbreds (Wright *et al.* 2005; Weber *et al.* 2007). We excluded accessions (7 *parviglumis*, 1 *mexicana*, and 20 *mays*) and loci (3 SNPs) with >15% missing data. We also removed SNPs where BLAST searches of context sequence identified multiple locations in the *mays* reference genome (Release 5a.59) (Schnable *et al.* 2009). This resulted in a final panel of 941 SNPs (Table S2) from 542 mapped genes. SNP genotypes and their contextual sequences are available at <http://www.panzea.org> and <http://www.rilab.org>. A subset of these data has been published elsewhere, including 706 SNPs from 584 *parviglumis* accessions in Weber *et al.* (2007); 123 SNPs from 817 *parviglumis* in Weber *et al.* (2008); and 468 SNPs from 1127 *mays*, 100 *parviglumis*, and 96 *mexicana* in van Heerwaarden *et al.* (2011).

The majority of the accessions represent open-pollinated populations that are highly heterozygous, resulting in genotypic data of unknown phase. We computationally phased each of the three subspecies separately using the software fastPHASE (Scheet and Stephens 2006) with 20 random starts and 25 iterations of the expectation-maximization algorithm. For each subspecies, inbred lines (Table S1) were used as training data.

We made use of published Sanger resequencing data from Wright *et al.* (2005). Wright and coauthors sequenced PCR products from teosinte accessions that had been inbred for two generations. Within the inversion, two individuals with a sequence from one arrangement at one locus and the alternate arrangement at a different locus were deemed heterozygous and removed from the analysis. We limited our analyses to reliably aligned loci on chromosome 1 with $n \geq 10$ total sequences and $n \geq 3$ of each arrangement. In total, we analyzed 95 loci, including 7 inside the inversion (Table S3; data available at www.panzea.org). We used the top BLAST hit from the *Sorghum* genome (Paterson *et al.* 2009) as an outgroup for each locus. Loci were annotated by BLAST comparison to the *mays* reference genome (Release 5a.59) (Schnable *et al.* 2009).

Data analysis and divergence time

Summary statistics of the SNP genotyping data and the resequencing data were calculated using the analysis package of the libsequence library (Thornton 2003). Although the absolute values of some summary statistics of the SNP genotyping data may be affected by ascertainment bias (Clark *et al.* 2005a), the relative values of these statistics are expected to be more robust.

We tested for Hardy–Weinberg equilibrium by treating the inversion as a single biallelic locus. LD (as measured by r^2) was calculated for SNPs with a minor allele frequency (MAF) >5%, using the LDheatmap package (Shin *et al.* 2006) in R (R Development Core Team 2011). To assign individuals to haplotype clusters at the inversion, we used the genetic assignment software STRUCTURE (Pritchard *et al.* 2000; Falush *et al.* 2003). We estimated haplotype clusters for values of K ranging from 1 to 5. For each value of K , we used 10 replicate runs of the admixture model, with a burn-in of 100,000 iterations and a run length of 100,000 steps. To compare differentiation inside and outside of the inversion, we divided the sample of *parviglumis* into the two clusters identified by STRUCTURE and calculated F_{ST} (Weir and Cockerham 1984) between these two groups along chromosome 1.

Genetic (Manhattan) distance among inbred *parviglumis* lines was estimated with the software TASSEL (Bradbury *et al.* 2007) and calculated separately for SNPs inside and outside of the inversion. We used a Fitch–Margoliash least-squares approach (Fitch and Margoliash 1967) as implemented in the software package PHYLIP (Felsenstein 2005) to estimate a dendrogram for all taxa using the 17 SNPs inside the inversion.

We applied two common tests of neutrality (Hudson *et al.* 1987; McDonald and Kreitman 1991) to the Sanger resequencing data. For the McDonald–Kreitman (MK) test we used the seven resequenced loci inside the inversion and compared polymorphism at the inverted arrangement to divergence between the inverted arrangement and a *Sorghum* outgroup. For the Hudson–Kreitman–Aguade (HKA) test, we used 74 loci with ancestral information. HKA tests were performed both for the combined set of sequences and for sequences from each of the chromosomal arrangements separately. Because loci within the inversion are unlikely to be independent, we summed polymorphism and divergence data across loci within the inversion. We used the maximum-likelihood approach of Wright and Charlesworth (2004), running 100,000 Markov chain Monte Carlo (MCMC) iterations, and a starting *parviglumis*–*Sorghum* divergence of 60N generations.

We estimated divergence time between the arrangements, using sequences at seven loci from the two observed haplotype groups inside the inversion. We treated samples of the two chromosomal arrangements as distinct populations and estimated divergence time under an isolation-with-migration model as implemented in the software MIMAR (Becquet and Przeworski 2007). We set the inheritance and

the mutation rate variation scalars both to 1 and the recombination inheritance and rate variation scalar to $(Z_{ri} - 1)/(Z_i - 1)$, where Z_{ri} is the initial length of locus i and Z_i corresponds to the number of base pairs in locus i after filtering out indels and missing data. The mutation rate per generation per base pair was assumed constant across loci and set to 3×10^{-8} (Clark *et al.* 2005b). The population mutation rate per base pair, divergence time, and the natural logarithm of the population migration parameter were sampled from the uniform distributions $U(0, 0.08)$, $U(0, 10^6)$, and $U(-2, 1)$, respectively. The exponential growth parameter was set to 6 (other values did not change results considerably). We ran the Markov chain for 5000 burn-in steps followed by 10,000 steps for parameter estimation, repeating our analysis with two independent seeds. Three hundred genealogies were generated per locus for each step of the MCMC. We inferred that convergence was reached when the posterior distributions of both runs were very similar; results reported are the average of both runs.

Association analyses

We used a Bayesian approach (Coop *et al.* 2010) to test for associations between the inversion and 22 geographical (altitude, latitude, and longitude) and bioclimatic (worldclim.org) variables in the 33 *parviglumis* populations with ≥ 10 samples (Table S4). The analysis explicitly accounts for population structure, using a covariance matrix of allele frequencies estimated by 50,000 MCMC steps using all SNPs. We assessed association genome-wide using all SNPs and using a single-marker test, treating the inversion as a single locus. In each case, five separate runs with 50,000 iterations were performed to control for differences among MCMC runs.

To examine whether the inverted arrangement was associated with phenotypic variation, we used phenotype data from Weber *et al.* (2008). Both phenotype and genotype data were available for 811 individuals. A kinship matrix was estimated from all 941 SNPs, using the options “all” and “additive” in the EMMA R package (Kang *et al.* 2008). Both genome-wide association and single-marker association (treating the inversion as a single locus) were performed for each of 37 phenotypes. Associations were tested using a mixed linear model as implemented in the software TASSEL (Bradbury *et al.* 2007). The R package qvalue was used to estimate the false discovery rate (FDR) and identify SNPs that were significant at an FDR of 5% (Storey and Tibshirani 2003).

Cytology

To assess the potential cytological impacts of the inversion, we screened 174 meiocytes from immature tassels of six *parviglumis*/mays F_1 progeny resulting from the cross of two inbred *parviglumis* each with a single *mays* inbred line. Meiocytes were collected and staged following Li *et al.* (2010) and the chromosomes stained with either 1% aceto-orcein or DAPI. Recombination within the inversion was scored as previously described, noting chromosome

bridges and acentric fragments at anaphase I (Dawe and Cande 1996).

Results

We examined the level of LD in each of the three subspecies of *Z. mays* with a genome-wide set of 941 SNPs from 2782 samples. Using computationally phased genotypic data, we searched for pairs of markers in high LD ($r^2 > 0.6$) and separated by >1 Mb. Our scan identified two such regions, an ~ 50 -Mb region on chromosome 1 and an ~ 15 -Mb span of chromosome 8. Because the region on chromosome 8 is near a likely assembly error in the reference genome (J. Glaubitz, unpublished data), we focused our analysis on chromosome 1. The region of high LD on chromosome 1 in our data corresponds closely to the 65- to 115-Mb region on the physical map of the reference *mays* genome (B73 RefGen v2, release 5a.59, 2010–2011) recently reported by Hufford *et al.* (2012) as a putative inversion. Our data reveal high LD (mean $r^2 = 0.24$) among the 17 SNPs from Mb 65.09 to 106.16 (Figure 1), compared to a genome-wide average of 0.004. Gametic disequilibrium, as estimated from unphased SNP genotyping data, also demonstrates this excess of LD (data not shown). Finally, high levels of LD are also evident in genotypic data from a panel of 13 individuals of *parviglumis* genotyped using the 55,000 SNPs on the MaizeSNP50 Illumina Infinium Assay (Hufford *et al.* 2012), suggesting that the LD observed is not an artifact of the genotyping platform used.

The extended region of high LD on chromosome 1 is a putative inversion

Because *mays* and the teosintes are outcrossing taxa with large effective population sizes, LD in the genome generally declines rapidly with distance ($r^2 < 0.1$ within 1500 bp in domesticated *mays*) (Remington *et al.* 2001). The region of high LD is distinct from both the centromere (Wolfgruber *et al.* 2009) and known heterochromatic knobs (Buckler *et al.* 1999) and exhibits relatively low recombination (Figure 1). An ~ 50 -Mb span of high LD is unexpected, and while *parviglumis* and *mexicana* show evidence of high LD in this chromosomal region, levels of LD in our large sample of domesticated *mays* are similar to genome-wide averages (Figure 1). Other wild taxa also do not show an excess of LD on the short arm of chromosome 1, although our power to measure LD in these samples is likely hampered by smaller sample size and SNP ascertainment bias. Finally, a recent genetic map from a BC2S3 population derived from a cross between a *mays* line and a *parviglumis* line with the putatively inverted arrangement shows no crossovers inside the ~ 50 -Mb span in the 881 progeny genotyped, consistent with the putative inversion suppressing recombination in heterozygotes (L. Shannon and J. Doebley, unpublished data). Although final validation will require demonstrating differential marker order in the progeny of self-fertilized individuals

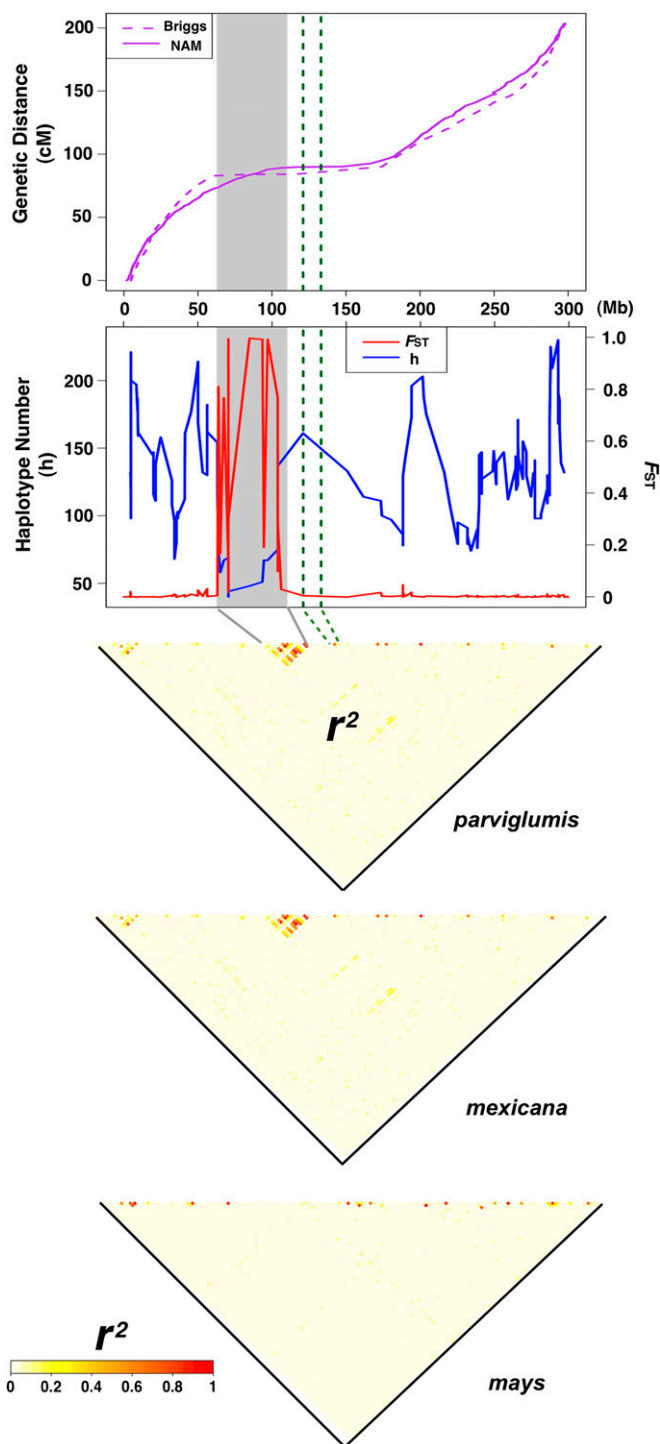


Figure 1 Population genetic evidence for the *Inv1n* inversion. Top, cumulative genetic distance by physical position along chromosome 1. The dashed curve is based on the teosinte–maize backcross map of Briggs *et al.* (2007) and the solid curve is from the maize nested association-mapping (NAM) population (Yu *et al.* 2008). Bottom, haplotype number (blue curve) and F_{ST} between the inverted and standard arrangements (red curve). The number of haplotypes present across chromosome 1 was calculated in overlapping 10-SNP windows with 1-SNP increments. The inverted region is marked in gray and the centromere in green dashed lines. Below, LD (r^2) is plotted across the chromosome for *parviglumis*, *mexicana*, and *mays*.

homozygous for alternate arrangements (Mano *et al.* 2012), we view these multiple lines of evidence as a strong case that recombination is suppressed due to an inversion in this region, henceforth identified as *Inv1n*.

To test for evidence of pairing and recombination within the large *Inv1n* region, we examined male meiocytes from six F₁ plants derived from two crosses between *mays* and an inbred *parviglumis* line containing *Inv1n*. Both hybrids revealed a low frequency of dicentric bridge formation at ~4% (7/167), but no acentric fragments were observed (Table S5). Although such bridges were rare, an anaphase I bridge in a plant heterozygous for *Inv1n* was observed (Figure S1). In addition, we observed no obvious reduction in pollen viability or seed set in a total of five F₁ plants (data not shown).

Haplotype variation and divergence time

STRUCTURE (Pritchard *et al.* 2000; Falush *et al.* 2003) analysis of SNPs on all 1936 *parviglumis* chromosomes inside *Inv1n* shows the highest likelihood for $K = 2$ clusters, a pattern not seen from the full set of genome-wide SNPs (data not shown). These groups are hereafter referred to as *Inv1n-I* and *Inv1n-S* for the inverted and standard arrangements, respectively (Figure 2). Recombination among loci within a chromosomal arrangement should be unaffected, and levels of LD within *Inv1n-I* (mean $r^2 = 0.11$) and *Inv1n-S* arrangements (mean $r^2 = 0.07$) are indeed low and similar to background levels (Figure S2). Average F_{ST} between chromosomes with alternate arrangements is notably higher inside the *Inv1n* region (0.54) than across the rest of the genome (0.01) (Figure 1). Genetic distance among accessions for SNPs along chromosome 1 outside the *Inv1n* region shows little evidence of haplotype structure (Figure 3A), while genetic distance for SNPs inside *Inv1n* divides *parviglumis* into two clear haplotypic groups representing *Inv1n-I* and *Inv1n-S* (Figure 3B). The *Inv1n-S* cluster includes all taxa of *Zea* and *Tripsacum* investigated, and it is parsimonious to assume that the *Inv1n-I* cluster, present only in *parviglumis* and *mexicana*, represents the derived inverted arrangement (Figure 3C). Despite strong differentiation, the two arrangements share polymorphic SNPs (Figure 2), even in homozygous individuals unaffected by haplotype phasing (data not shown). Among the 968 *parviglumis* samples, 345 (35.6%) are heterozygous at *Inv1n*, while 369 (38.1%) and 254 (26.3%) are homozygous for the *Inv1n-I* and *Inv1n-S* arrangements, respectively. *Inv1n-I* consists of a smaller number of distinct haplotypes and shows a paucity of rare haplotype variants compared to *Inv1n-S* (Figure 2).

Resequencing data from seven loci within *Inv1n* mirror these results (Table 1). Four loci (PZA00692, PZA00593, PZA03014, and PZA00146) show distinct haplotype clusters consistent with the SNP genotyping data (data not shown), dividing *parviglumis* into two groups representing *Inv1n-I* and *Inv1n-S*. A comparison of the two groups reveals a higher number of fixed differences, fewer shared derived SNPs, and higher average F_{ST} (0.53 vs. 0.05) inside the *Inv1n* region than outside. Average Tajima's D of the entire sample is

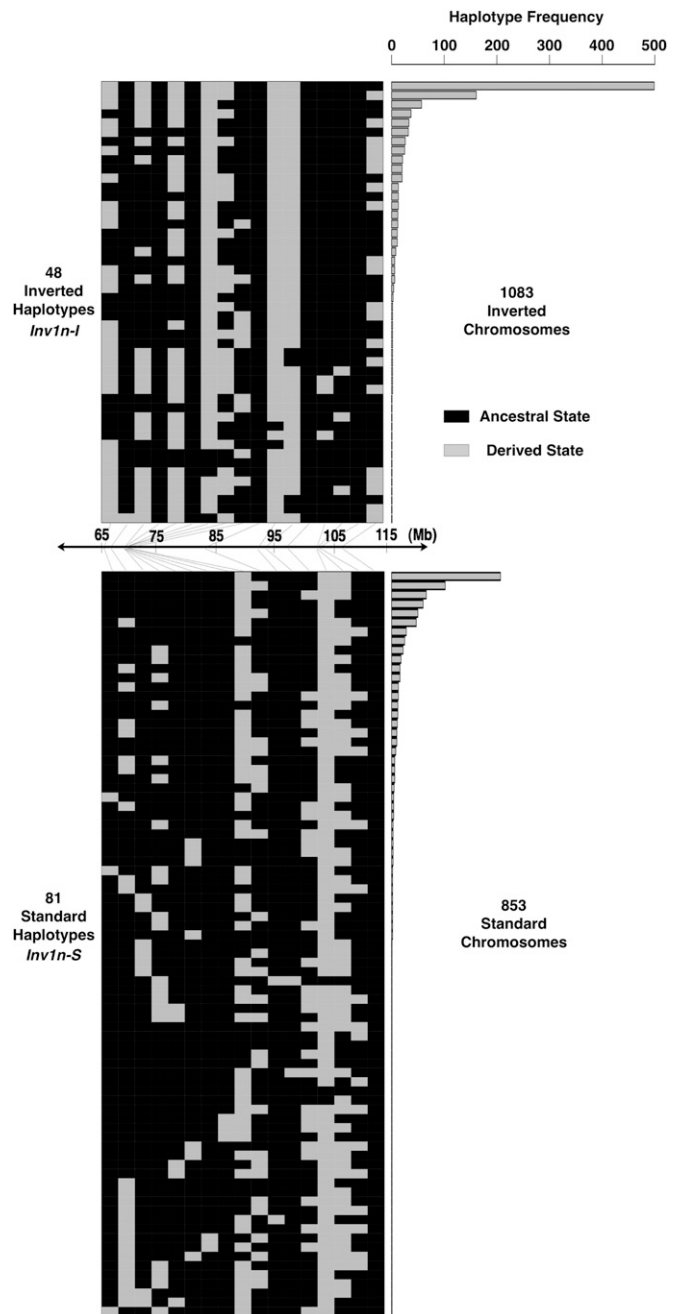


Figure 2 Diagram of haplotype diversity in *parviglumis* based on the 17 SNPs within *Inv1n*. Haplotypes are divided into the two clusters identified by STRUCTURE. Each SNP is represented by either the ancestral state (solid) or the derived state (shaded). The frequency of each of the haplotypes from the inverted (top) and standard (bottom) arrangements is shown on the right. The middle bar shows the physical position of each of the 17 SNPs inside *Inv1n*.

higher inside *Inv1n* (0.58 vs. -0.29), and the lack of rare haplotypes on the *Inv1n-I* background observed in the SNP data is reflected in the positive Tajima's D at sequences from these chromosomes (Table 1). All alleles private to *Inv1n-I* are derived on the basis of the *Sorghum* outgroup sequence, but 30% of the alleles private to *Inv1n-S* are ancestral.

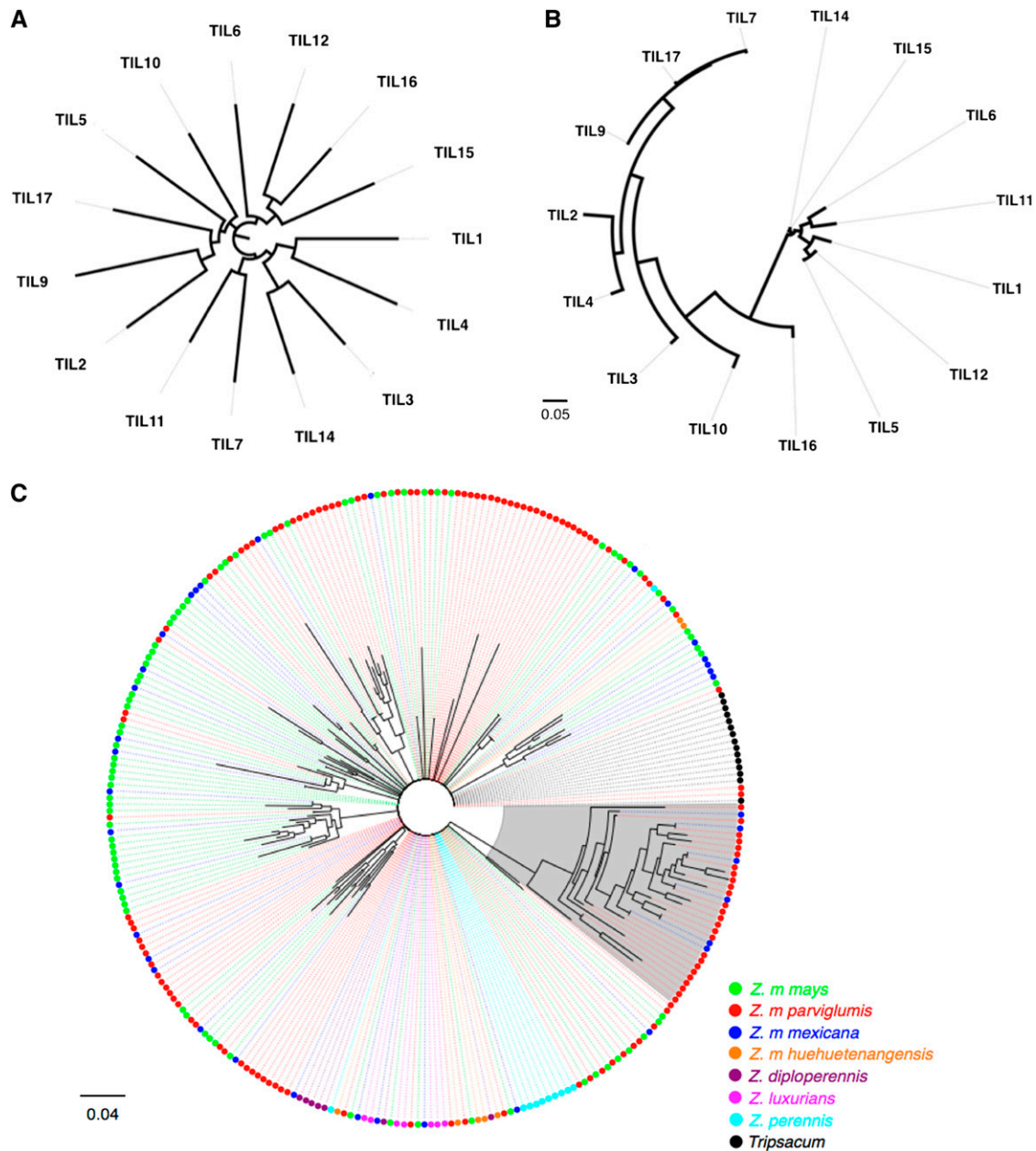


Figure 3 (A) Neighbor-joining tree for all SNPs outside *Inv1n*, using 15 *parviglumis* inbred lines. (B) Neighbor-joining tree for all SNPs inside *Inv1n*, using 15 *parviglumis* inbred lines. (C) Neighbor-joining tree for all unique haplotypes in each taxon, using all SNPs inside *Inv1n*. The haplotypes in the gray region represent the *Inv1n-I* arrangement.

We used multiple approaches to estimate the age of *Inv1n-I* from the resequencing data. Using the MCMC approach of Becquet and Przeworski (2007), which estimates divergence time from patterns of shared polymorphism under an isolation model, divergence was estimated to be ~296,000 generations, with a 95% confidence interval

Table 1 Mean (and standard deviation) of summary statistics for 7 resequenced loci inside and 88 loci outside *Inv1n*

	No. loci	<i>n</i>	<i>L</i>	<i>S_{sh}</i>	<i>S_f</i>	<i>S_p</i>	<i>h</i>	<i>H</i>	θ_π	Taj <i>D</i>	Fay and Wu's <i>H</i>
Inside (<i>Inv1n-I</i>)	7 (6)	14.6 (1.5)	307 (88)	0.3 (0.8)	4.3 (3.1)	2.9 (3.2)	2.3 (1.4)	0.49 (0.41)	0.004 (0.004)	0.37 (1.21)	-0.001 (0.003)
Inside (<i>Inv1n-S</i>)	7 (6)	14.6 (1.5)	307 (88)	0.3 (0.8)	4.3 (3.1)	2.9 (2.5)	3.3 (1.7)	0.59 (0.36)	0.004 (0.003)	-0.70 (0.56)	0 (0.003)
Outside (<i>Inv1n-I</i>)	88 (68)	13.5 (1.9)	414 (107)	4.7 (4.6)	0.1 (0.9)	4.1 (4.1)	3.9 (1.2)	0.89 (0.21)	0.011 (0.009)	-0.22 (0.65)	-0.003 (0.017)
Outside (<i>Inv1n-S</i>)	88 (68)	13.5 (1.9)	414 (107)	4.7 (4.6)	0.1 (0.9)	4.7 (3.6)	4.7 (1.8)	0.88 (0.21)	0.010 (0.007)	-0.34 (0.62)	-0.008 (0.029)

The number of loci with an outgroup is listed in parentheses in the "No. loci" column. The numbers in parentheses in other columns are standard deviations: *n*, number of samples; *L*, length of the locus; *S_{sh}*, number of shared SNPs between *Inv1n-I* and *Inv1n-S*; *S_f*, number of fixed SNPs; *S_p*, number of private SNPs; *h*, number of haplotypes; *H*, haplotype diversity; θ_π , pairwise difference per base pair.

(C.I.) between 221,000 and 398,000 generations. Assuming a constant rate of substitution of 3×10^{-8} per generation (Clark *et al.* 2005b), we can also calculate divergence time from net differences in nucleotide diversity (Nei and Li 1979) between *Inv1n-I* and *Inv1n-S*, which gives an estimate of 260,000 generations. Finally, estimates of the time to the most recent common ancestor (TMRCA) (Thomson *et al.* 2000; Hudson 2007) of the complete sample inside *Inv1n* (308,100 generations, 95% C.I. of 272,800–345,600 generations) and of *Inv1n-I* alone (133,200 generations, 95% C.I. of 96,500–175,800 generations) are consistent with other methods.

Neutrality tests

On the basis of standard tests of neutrality, there is limited evidence of selection on *Inv1n*. HKA tests on *Inv1n* did not detect evidence of balancing selection caused by environmental heterogeneity (P -value = 0.46, divergence to diversity ratio = 2.58). MK (P -value = 0.65) and HKA (P -value = 0.15) tests on resequencing data from the *Inv1n-I* arrangement failed to reject a neutral model, and Fay and Wu's H (Fay and Wu 2000) did not differ markedly between the *Inv1n-I* arrangements or compared to loci outside of *Inv1n* (Table 1).

Population frequencies and association analyses

All 33 *parviglumis* populations sampled were polymorphic for both arrangements at the *Inv1n* locus, with a mean *Inv1n-I* frequency of 55%. The frequency of *Inv1n-I* is negatively correlated with altitude ($r^2 = 0.34$) (Figure 4, Figure S3, and Table S4), ranging from 90% in the Quenchendio population at an altitude of 653 m to 18.4% in Ahuacatitlan at 1528 m. Consistent with this, *Inv1n-I* occurs at a frequency of only 9.7% in subspecies *mexicana*, which is found at higher altitudes than subspecies *parviglumis* (mean altitude of 2091 m vs. 1087 m for our *mexicana* and *parviglumis* samples). The most common *Inv1n-I* haplotype makes up 46% (499/1083) of *parviglumis* chromosomes with the *Inv1n-I* variant and does not vary significantly in frequency among populations ($\chi^2 = 2.27$, d.f. = 32, P -value = 1).

Using a model-based approach (Coop *et al.* 2010) to control for population structure, we examined the association between *Inv1n-I* frequency and 22 environmental variables (Table S6). Among environmental variables, altitude was most strongly correlated with *Inv1n-I* frequency and consistently obtained the highest Bayes factors among runs (mean 136). High Bayes factors were also observed for other bioclimatic variables, including temperature (mean temperature of driest quarter, mean Bayes factor = 49) and precipitation (precipitation of driest month, mean Bayes factor = 48). Genome-wide analysis of all SNPs produced mean Bayes factors for association with altitude ranging from 0.29 to 1048 (Table S6). SNPs in the inversion are 10-fold enriched in the top 5% tail of Bayes factors for altitude and >20-fold enriched in the top 1% tail, strongly suggesting a link between *Inv1n* and altitude (Figure 5).

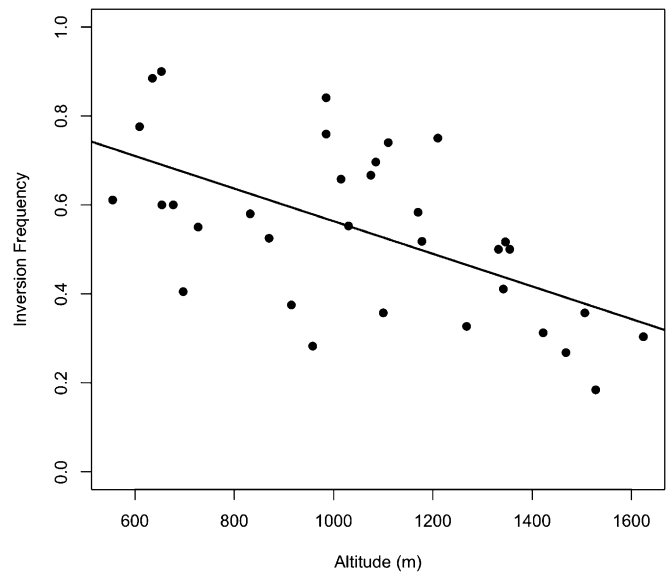


Figure 4 *Inv1n-I* frequency in *parviglumis* populations is negatively correlated with altitude. Each point corresponds to a population, with its altitude on the x -axis and *Inv1n-I* frequency on the y -axis.

We also used a mixed linear model analysis to test for associations between *Inv1n-I* and phenotypic data from the same populations (Weber *et al.* 2008). *Inv1n-I* appears to be associated with the percentage of male internodes (PSIN) (P -value = 0.0055, $r^2 = 0.024$), percentage of staminate spikelets (STAM) (P -value = 0.0069, $r^2 = 0.023$), culm diameter (CULM) (P -value = 0.0137, $r^2 = 0.011$), and leaf number (LFNM) (P -value = 0.0232, $r^2 = 0.010$) (Table S7), but none of the associations are significant after Bonferroni correction for phenotypes tested and effect sizes for all phenotypes are very small. In addition to testing the inversion as a single locus, we also investigated associations between individual SNPs and phenotypes. For both PSIN and STAM, none of the 17 SNPs in *Inv1n* were among the 1% of SNPs most strongly associated with the two phenotypes. However, SNPs in *Inv1n* were enriched in the 1% tail of P -values for both CULM (15 \times enrichment) and LFNM (3 \times). None of the SNPs were significantly associated with PSIN or LFNM at a FDR of 5%, while four SNPs outside of the inversion were significantly associated with STAM. Of the seven SNPs significantly associated with CULM at an FDR of 5%, four (PZA00263.14, PZD00077.7, PZA00692.5, and PZA03014.24) are inside *Inv1n*.

Discussion

Using a genome-wide set of SNPs in a large panel of wild and domesticated *Zea*, we provide evidence of an ~50-Mb inversion on the short arm of chromosome 1 of subspecies *parviglumis* and *mexicana*. While our cytological data are not directly diagnostic for an inversion, population genetic data preclude alternative explanations. For example, the dramatic reduction in haplotype number in the *Inv1n* region (Figure 1) could be indicative of a selective sweep (Kim and

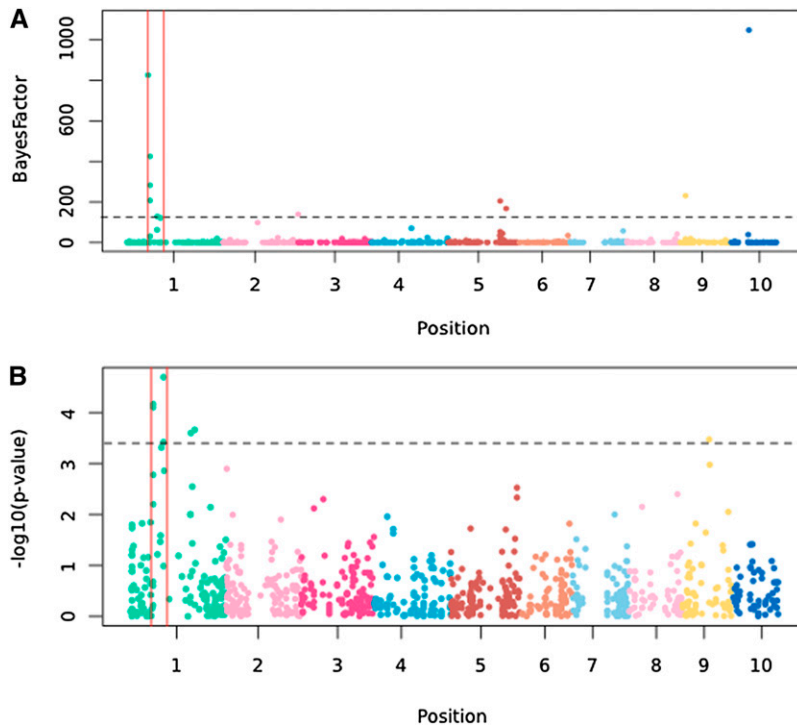


Figure 5 (A) Bayes factors for correlation between allele frequencies and altitude in 33 natural *parviglumis* populations. *Inv1n* is indicated by red vertical lines. The 99th percentile of the distribution of Bayes factors is indicated by a horizontal dashed line. Chromosomes 1–10 are plotted in order and in different colors. (B) Association between all SNPs and culm diameter. SNPs significant at 5% FDR are above the dashed line.

Nielsen 2004; Nielsen *et al.* 2005; McVean 2007). However, the largest sweep identified in maize to date is only 1.1 Mb (Tian *et al.* 2009), and both the age of the inversion and common tests for departures from neutrality do not provide evidence of strong selection. Another alternative explanation would be the presence of strong negative interactions between distantly linked loci, potentially due to synthetic lethality (Boone *et al.* 2007). Such interactions should not generate extended patterns of elevated LD among intervening SNPs, as crossing over among haplotypes not carrying alleles involved in the negative interaction should not be affected. Both selective sweeps and negative interactions are inconsistent with the presence of only two major haplotypes in the *Inv1n* region and fail to explain the clinal variation in haplotype frequencies seen at *Inv1n-I*.

To our knowledge, the only prior evidence for *Inv1n* is a report of high LD and high F_{ST} from a much smaller sample of *parviglumis* (Hufford *et al.* 2012), but a number of other large inversions have been previously reported in *mays* and its wild relatives (Ting 1965, 1967, 1976; Maguire 1966; Kato 1975). These include an ~50-Mb inversion on the long arm of chromosome 3 in *Z. luxurians* (Ting 1965) and an ~35-Mb inversion that covers most of the short arm of chromosome 8 in both *mays* (McClintock 1960) and *mexicana* (Ting 1976). While some of these inversions were experimentally induced (McClintock 1931; Morgan 1950), several have also been identified in natural populations of multiple taxa (Kato 1975; Ting 1976).

One of the factors that may limit the geographic spread of large inversions is the potential fitness cost of crossing over. The frequency of chromosome loss is dependent on the inversion size and efficiency of synapsis over the inverted

region (Burnham 1962; Maguire and Riess 1994; Lamb *et al.* 2007). When gene density is low, such as in pericentromeric regions, or there is a lack of continuous homology, chromosomes will often synapse in a nonhomologous manner without recombination (McClintock 1933). In maize, for example, an inversion on the long arm of chromosome 1 similar in size to *Inv1n* (19 cM) was seen to undergo homologous pairing in only about one-third of cases (Maguire 1966). Since *Inv1n* is located in a pericentromeric region with low gene density and covers a short genetic distance (2–13 cM), we anticipated that it would rarely pair and recombine with a noninverted chromosome. Our data are consistent with these arguments. We observed repressed recombination around *Inv1n* and no cytological evidence of crossing over in inversion heterozygotes. SNP data indicate no deviations from expected Hardy–Weinberg genotype frequencies at *Inv1n*, and we see no obvious evidence of effects on fertility. Given these observations, we suspect that inversion polymorphisms may be relatively common in natural plant populations, especially in regions of the genome with low recombination rates such as pericentromeres. Low recombination has also been offered as an explanation for the lack of underdominance in many pericentromeric inversions in *Drosophila* (Coyne *et al.* 1993). As dense genotyping becomes more cost effective, we predict that numerous common inversions will be identified in natural populations of *Zea* and other organisms.

Origin and age of *Inv1n*

Our evidence suggests that *Inv1n-I* is the derived, inverted arrangement. *Inv1n-I* is not found in *Tripsacum* or *Zea* taxa except for *parviglumis* and *mexicana* (Figure 3C), and, unlike in *Inv1n-S*, all SNPs private to *Inv1n-I* are derived in

resequencing data. Both SNP and resequencing data show strong differentiation between the two arrangements (Figure 1 and Table 1). Multiple methods of estimating the age of the *Inv1n-I* haplotype point to an origin ~300,000 generations ago. This predates both the split between *mexicana* and *parviglumis* and the split between *Z. luxurians* and *Z. mays* and is similar to the estimated age of divergence of most species in the genus *Zea* (Ross-Ibarra *et al.* 2009). Several considerations suggest that these numbers are plausible. First, the proportion of SNPs shared between *parviglumis* and *mexicana* on chromosome 1 does not differ inside or outside of *Inv1n* (15/17 vs. 134/139, Fisher's exact test P -value = 0.48), suggesting that the presence of the inversion in both subspecies is likely due to shared ancestral polymorphism rather than recent gene flow. Second, while the estimated age of *Inv1n-I* is similar to the estimated divergence of species, other species in *Zea* have narrow distributions (Fukunaga *et al.* 2005) and presumably small effective population sizes, increasing the potential for loss of variants at low frequency in the ancestral population. Third, *Inv1n-I* could consist of multiple independent inversions, similar to the inversion polymorphisms identified in the white-throated sparrow (Thomas *et al.* 2008). In this case, estimates of the age of the inversion would be biased upward, as each inversion would have arisen independently on distinct backgrounds. Such a scenario might also explain the observation of shared polymorphisms between *Inv1n-I* and *Inv1n-S*. Our data cannot distinguish the number of independent inversions in the region, however, which would instead require analysis of progeny derived from crosses of multiple individuals homozygous for different haplotypes of *Inv1n-I* and a more dense set of markers.

While small effective population size may explain the absence of *Inv1n-I* from other taxa in *Zea*, its complete absence in our sample of 1573 *mays* requires additional explanation. Sampling alone is unlikely to play a role, as the vast majority of our *mays* accessions are landraces, collected from across the Americas, including accessions collected within the range of *parviglumis* and *mexicana*. Estimates of the domestication bottleneck and observed levels of diversity in domesticated *mays* (Tenailon *et al.* 2004; Wright *et al.* 2005) also suggest that drift during domestication is not a compelling explanation, especially given that *Inv1n-I* occurs at frequencies of up to 90% in the lowland areas where domestication is thought to have occurred (Matsuoka *et al.* 2002; Piperno *et al.* 2009; van Heerwaarden *et al.* 2011). We speculate instead that *Inv1n-I* may have been selected against in domesticated *mays*. Our association analysis provides limited evidence in support of this idea, as *Inv1n-I* is negatively associated with culm width in *parviglumis*, while domesticated *mays* has more robust culms than its wild progenitor (Briggs *et al.* 2007).

Selection on *Inv1n*

While standard tests of neutrality do not provide evidence for selection, there is reason to believe that *Inv1n* is not

evolving neutrally. First, the *Inv1n-I* arrangement is widely distributed, segregating in all 33 populations investigated; only two SNPs on chromosome 1 are also polymorphic in all populations. Second, pairwise F_{ST} inside *Inv1n* appears uncorrelated to pairwise F_{ST} genome-wide ($r^2 = 0.04$; Figure S4), suggesting that the frequency of *Inv1n* is not entirely due to isolation by distance. Third, even after correcting for population structure, *Inv1n-I* frequency is associated with a number of environmental variables (Table S6), including a strong altitudinal cline (Figure 4 and Figure S3). Latitudinal (Anderson *et al.* 2005; Santos *et al.* 2005; Umina *et al.* 2005) and altitudinal clines (Levitan 2001) are commonly observed for inversion polymorphisms and are often thought to be related to temperature adaptation (Levitan 2001; Umina *et al.* 2005; Balanyá *et al.* 2006). Fourth, the *Inv1n-I* arrangement is, to our knowledge, the first inversion in *Zea* shown to be associated with phenotypic differences (Table S7). These include culm diameter, a trait that differentiates maize from teosinte (Briggs *et al.* 2007), and tassel morphology (Table S7), which is known to differ between *parviglumis* and *mexicana* (Doebley 1983) and between lowland and highland maize (Anderson 1946; Bretting and Goodman 1989). Fifth, the lack of rare variants and the high frequency of the most common *Inv1n-I* haplotype (Figure 2) suggest that this haplotype may have recently risen to high frequency due to a partial sweep. The observed lack of rare variants is especially striking given the genome-wide pattern of an excess of low-frequency variants (Table 1), an observation reported in multiple studies (Tenailon *et al.* 2004; Wright *et al.* 2005; Moeller *et al.* 2007; Ross-Ibarra *et al.* 2009). While the most common haplotype at *Inv1n-I* does not show signs of extended homozygosity beyond the borders of the inversion (Figure S5) as might be expected if it has been recently swept to higher frequency, the nearest flanking SNPs are 1.1 and 14.6 Mb distant and our power to detect an extended haplotype is low. Sixth, the absence of *Inv1n-I* from domesticated maize, in spite of recurrent gene flow from both *parviglumis* and *mexicana* (Wilkes 1967; Fukunaga *et al.* 2005; Ross-Ibarra *et al.* 2009; van Heerwaarden *et al.* 2011), suggests that the inverted arrangement was selected against at some point during *mays* domestication or breeding. Finally, we note that, aside from strong divergence between chromosomes of different arrangements, selection may be difficult to detect in diversity data from inversions of an age similar to ours (Guerrero *et al.* 2012).

Selection may act on inversions because of the fitness consequences of the structural rearrangement itself or of adaptive alleles at loci inside the inversion (Kirkpatrick and Barton 2006). While these scenarios predict somewhat different patterns of diversity (Guerrero *et al.* 2012), our SNP genotyping data are of insufficient density to distinguish between them. Regardless of the mechanism of selection, the observed altitudinal cline and absence in domesticated *mays* suggest that *Inv1n-I* is not ubiquitously adaptive. Even in low-altitude populations where it seems to be favored, a large inversion such as *Inv1n* may have captured several

recessive deleterious alleles, effectively preventing its fixation (Kirkpatrick and Barton 2006).

Acknowledgments

We thank Kate Hodges for help with cytological scoring; Kevin Volz for help with geographic information systems data and for producing Figure S3; and Graham Coop, Loren Rieseberg, and two anonymous reviewers for helpful comments on earlier versions of the manuscript. This work was carried out using computing resources at the University of Minnesota Supercomputing Institute. We acknowledge funding from the Academy of Finland (to T.P.) and start-up funds from the University of Minnesota Department of Agronomy and Plant Genetics (to P.L.M.) and from the National Science Foundation (NSF) (DBI-0820619 to J.D. and NSF IOS-0922703 to R.K.D. and J.R-I.).

Literature Cited

- Anderson, A. R., A. A. Hoffmann, S. W. Mckechnie, P. A. Umina, and A. R. Weeks, 2005 The latitudinal cline in the *In(3R) Payne* inversion polymorphism has shifted in the last 20 years in Australian *Drosophila melanogaster* populations. *Mol. Ecol.* 14: 851–858.
- Anderson, E., 1946 Maize in Mexico a preliminary survey. *Ann. Mo. Bot. Gard.* 33: 147–247.
- Andolfatto, P., J. D. Wall, and M. Kreitman, 1999 Unusual haplotype structure at the proximal breakpoint of *In(2L)t* in a natural population of *Drosophila melanogaster*. *Genetics* 153: 1297–1311.
- Ayala, D., M. C. Fontaine, A. Cohuet, D. Fontenille, R. Vitalis *et al.*, 2011 Chromosomal inversions, natural selection and adaptation in the malaria vector *Anopheles funestus*. *Mol. Biol. Evol.* 28: 745–758.
- Balanyá, J., J. M. Oller, R. B. Huey, G. W. Gilchrist, and L. Serra, 2006 Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science* 313: 1773–1775.
- Bansal, V., A. Bashir, and V. Bafna, 2007 Evidence for large inversion polymorphisms in the human genome from HapMap data. *Genome Res.* 17: 219–230.
- Becquet, C., and M. Przeworski, 2007 A new approach to estimate parameters of speciation models with application to apes. *Genome Res.* 17: 1505–1519.
- Bengtsson, B. O., and W. F. Bodmer, 1976 On the increase of chromosome mutations under random mating. *Theor. Popul. Biol.* 9: 260–281.
- Boone, C., H. Bussey, and B. J. Andrews, 2007 Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* 8: 437–449.
- Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss *et al.*, 2007 TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23: 2633–2635.
- Bretting, P. K., and M. M. Goodman, 1989 Karyotypic variation in Mesoamerican races of maize and its systematic significance. *Econ. Bot.* 43: 107–124.
- Briggs, W. H., M. D. McMullen, B. S. Gaut, and J. Doebley, 2007 Linkage mapping of domestication loci in a large maize teosinte backcross resource. *Genetics* 177: 1915–1928.
- Buckler, E. S., T. L. Phelps-Durr, C. S. K. Buckler, R. K. Dawe, J. F. Doebley *et al.*, 1999 Meiotic drive of chromosomal knobs reshaped the maize genome. *Genetics* 153: 415–426.
- Burnham, C. R., 1962 *Discussions in Cytogenetics*. Burgess Publishing, Minneapolis.
- Castermans, D., J. R. Vermeesch, J. P. Fryns, J. G. Steyaert, W. J. M. Van De Ven *et al.*, 2007 Identification and characterization of the *TRIP8* and *REEP3* genes on chromosome 10q21.3 as novel candidate genes for autism. *Eur. J. Hum. Genet.* 15: 422–431.
- Clark, A. G., M. J. Hubisz, C. D. Bustamante, S. H. Williamson, and R. Nielsen, 2005a Ascertainment bias in studies of human genome-wide polymorphism. *Genome Res.* 15: 1496–1502.
- Clark, R. M., S. Tavaré, and J. Doebley, 2005b Estimating a nucleotide substitution rate for maize from polymorphism at a major domestication locus. *Mol. Biol. Evol.* 22: 2304–2312.
- Coop, G., D. Witonsky, A. Di Rienzo, and J. K. Pritchard, 2010 Using environmental correlations to identify loci underlying local adaptation. *Genetics* 185: 1411–1423.
- Coyne, J. A., W. Meyers, A. P. Crittenden, and P. Sniegowski, 1993 The fertility effects of pericentric inversions in *Drosophila melanogaster*. *Genetics* 134: 487–496.
- Dawe, R. K., and W. Z. Cande, 1996 Induction of centromeric activity in maize by suppressor of meiotic drive 1. *Proc. Natl. Acad. Sci. USA* 93: 8512–8517.
- Depaulis, F., L. Brazier, and M. Veuille, 1999 Selective sweep at the *Drosophila melanogaster* suppressor of hairless locus and its association with the *In(2L)t* inversion polymorphism. *Genetics* 152: 1017–1024.
- Dobzhansky, T., 1950 Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. *Genetics* 35: 288–302.
- Doebley, J. F., 1983 The maize and teosinte male inflorescence: a numerical taxonomic study. *Ann. Mo. Bot. Gard.* 70: 32–70.
- Falush, D., M. Stephens, and J. K. Pritchard, 2003 Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.
- Fay, J. C., and C. I. Wu, 2000 Hitchhiking under positive Darwinian selection. *Genetics* 155: 1405–1413.
- Felsenstein, J., 2005. PHYLIP (Phylogeny Inference Package), version 3.6 (distributed by the author). Department of Genome Sciences, University of Washington, Seattle.
- Fitch, W. M., and E. Margoliash, 1967 Construction of phylogenetic trees. *Science* 155: 279–284.
- Fukunaga, K., J. Hill, Y. Vigouroux, Y. Matsuoka, G. J. Sanchez *et al.*, 2005 Genetic diversity and population structure of teosinte. *Genetics* 169: 2241–2254.
- Guerrero, R. F., F. Rousset, and M. Kirkpatrick, 2012 Coalescent patterns for chromosomal inversions in divergent populations. *Philos. Trans. R. Soc. B* 367: 430–438.
- Hoffmann, A. A., and L. H. Rieseberg, 2008 Revisiting the impact of inversions in evolution: From population genetic markers to drivers of adaptive shifts and speciation? *Annu. Rev. Ecol. Evol. Syst.* 39: 21–42.
- Hoffmann, A. A., C. M. Sgro, and A. R. Weeks, 2004 Chromosomal inversion polymorphisms and adaptation. *Trends Ecol. Evol.* 19: 482–488.
- Hudson, R. R., 2007 The variance of coalescent time estimates from DNA sequences. *J. Mol. Evol.* 64: 702–705.
- Hudson, R. R., M. Kreitman, and M. Aguadé, 1987 A test of neutral molecular evolution based on nucleotide data. *Genetics* 116: 153–159.
- Hufford, M. B., X. Xun, J. van Heerwaarden, T. Pyhäjärvi, J. M. Chia *et al.*, 2012 Comparative population genomics of maize domestication and improvement (in press).
- Huynh, L. Y., D. L. Maney, and J. W. Thomas, 2011 Chromosome-wide linkage disequilibrium caused by an inversion polymorphism in the white-throated sparrow (*Zonotrichia albicollis*). *Heredity* 106: 537–546.
- Kang, H. M., N. A. Zaitlen, C. M. Wade, A. Kirby, D. Heckerman *et al.*, 2008 Efficient control of population structure in model organism association mapping. *Genetics* 178: 1709–1723.

- Kato, T. A., 1975 Cytological studies of maize and teosinte in relation to their origin and evolution. Available at <http://www.maizegdb.org/cooperators.php>. Accessed June 2012.
- Kim, Y., and R. Nielsen, 2004 Linkage disequilibrium as a signature of selective sweeps. *Genetics* 167: 1513–1524.
- Kirkpatrick, M., and N. Barton, 2006 Chromosome inversions, local adaptation and speciation. *Genetics* 173: 419–434.
- Lamb, J. C., J. M. Meyer, and J. A. Birchler, 2007 A hemicentric inversion in the maize line knobless Tama flint created two sites of centromeric elements and moved the kinetochore-forming region. *Chromosoma* 116: 237–247.
- Lande, R., 1984 The expected fixation rate of chromosomal inversions. *Evolution* 38: 743–752.
- Levitan, M., 2001 Studies of linkage in populations. XIV. Historical changes in frequencies of gene arrangements and arrangement combinations in natural populations of *Drosophila robusta*. *Evolution* 55: 2359–2362.
- Li, X., C. N. Topp, and R. K. Dawe, 2010 Maize antibody procedures: immunolocalization and chromatin immunoprecipitation (ChIP), pp. 271–286 in *Plant Cytogenetics, Genome Structure and Chromosome Function*, edited by H. W. Bass and J. A. Birchler. Springer-Verlag, Berlin/Heidelberg, Germany/New York.
- Lowry, D. B., and J. H. Willis, 2010 A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* 8: e1000500.
- Machado, C. A., T. S. Haselkorn, and M. A. F. Noor, 2007 Evaluation of the genomic extent of effects of fixed inversion differences on intraspecific variation and interspecific gene flow in *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 175: 1289–1306.
- Maguire, M. P., 1966 The relationship of crossing over to chromosome synapsis in a short paracentric inversion. *Genetics* 53: 1071–1077.
- Maguire, M. P., and R. W. Riess, 1994 The relationship of homologous synapsis and crossing over in a maize inversion. *Genetics* 137: 281–288.
- Mano, Y., F. Omori, and K. Takeda, 2012 Construction of intraspecific linkage maps, detection of a chromosome inversion, and mapping of QTL for constitutive root aerenchyma formation in the teosinte *Zea nicaraguensis*. *Mol. Breed.* 29: 137–146.
- Matsuoka, Y., Y. Vigouroux, M. M. Goodman, G. Sanchez, E. S. Buckler *et al.*, 2002 A single domestication for maize shown by multilocus microsatellite genotyping. *Proc. Natl. Acad. Sci. USA* 99: 6080–6084.
- McClintock, B., 1931 Cytological observations of deficiencies involving known genes, translocations and an inversion in *Zea mays*. *Mo. Agric. Expt. Sta. Bull.* 163: 1–30.
- McClintock, B., 1933 The association of non-homologous parts of chromosomes in the mid-prophase of meiosis in *Zea mays*. *Z. Zellforsch. Mikrosk. Anat.* 19: 191–237.
- McClintock, B., 1960 Chromosome constitutions of Mexican and Guatemalan races of maize. *Carnegie Inst. Wash. Yearbook* 59: 461–472.
- McDonald, J. H., and M. Kreitman, 1991 Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351: 652–654.
- McVean, G., 2007 The structure of linkage disequilibrium around a selective sweep. *Genetics* 175: 1395–1406.
- Moeller, D. A., M. I. Tenaillon, and P. Tiffin, 2007 Population structure and its effects on patterns of nucleotide polymorphism in teosinte (*Zea mays* ssp. *parviglumis*). *Genetics* 176: 1799–1809.
- Morgan, D. T., 1950 A cytogenetic study of inversions in *Zea mays*. *Genetics* 35: 153–174.
- Munte, A., J. Rozas, M. Aguade, and C. Segarra, 2005 Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *Drosophila subobscura*. *Genetics* 169: 1573–1581.
- Nei, M., and W. H. Li, 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* 76: 5269–5273.
- Nielsen, R., S. Williamson, Y. Kim, M. J. Hubisz, A. G. Clark *et al.*, 2005 Genomic scans for selective sweeps using SNP data. *Genome Res.* 15: 1566–1575.
- Nóbrega, C., M. Khadem, M. Aguadé, and C. Segarra, 2008 Genetic exchange vs. genetic differentiation in a medium-sized inversion of *Drosophila*: the A2/AST arrangements of *Drosophila subobscura*. *Mol. Biol. Evol.* 25: 1534–1543.
- Paterson, A. H., J. E. Bowers, R. Bruggmann, I. Dubchak, J. Grimwood *et al.*, 2009 The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457: 551–556.
- Piperno, D. R., A. J. Ranere, I. Holst, J. Iriarte, and R. Dickau, 2009 Starch grain and phytolith evidence for early ninth millennium BP maize from the Central Balsas River Valley, Mexico. *Proc. Natl. Acad. Sci. USA* 106: 5019–5024.
- Pritchard, J. K., M. Stephens, and P. Donnelly, 2000 Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- R Development Core Team, 2011 *R: a language and environment for statistical computing*, version 2.12.2. R Foundation for Statistical Computing, Vienna. <http://www.r-project.org>.
- Remington, D. L., J. M. Thornsberry, Y. Matsuoka, L. M. Wilson, S. R. Whitt *et al.*, 2001 Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci. USA* 98: 11479–11484.
- Ross-Ibarra, J., M. Tenaillon, and B. S. Gaut, 2009 Historical divergence and gene flow in the genus *Zea*. *Genetics* 181: 1399–1413.
- Santos, M., W. Céspedes, J. Balanyà, V. Trotta, F. C. F. Calboli *et al.*, 2005 Temperature-related genetic changes in laboratory populations of *Drosophila subobscura*: evidence against simple climatic-based explanation for latitudinal clines. *Am. Nat.* 165: 258–273.
- Scheet, P., and M. Stephens, 2006 A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* 78: 629–644.
- Schnable, P. S., D. Ware, R. S. Fulton, J. C. Stein, F. Wei *et al.*, 2009 The B73 maize genome: complexity, diversity, and dynamics. *Science* 326: 1112–1115.
- Shin, J. H., S. Blay, B. Mcnoney, and J. Graham, 2006 LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J. Stat. Softw.* 16: 1–10.
- Stevison, L. S., K. B. Hoehn, and M. A. F. Noor, 2011 Effects of inversions on within- and between-species recombination and divergence. *Genome Biol. Evol.* 3: 830–841.
- Storey, J. D., and R. Tibshirani, 2003 Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* 100: 9440–9445.
- Tenaillon, M. I., J. U'Ren, O. Tenaillon, and B. S. Gaut, 2004 Selection vs. demography: a multilocus investigation of the domestication process in maize. *Mol. Biol. Evol.* 21: 1214–1225.
- Thomas, J. W., M. Cáceres, J. J. Lowman, C. B. Morehouse, M. E. Short *et al.*, 2008 The chromosomal polymorphism linked to variation in social behavior in the white-throated sparrow (*Zonotrichia albicollis*) is a complex rearrangement and suppressor of recombination. *Genetics* 179: 1455–1468.
- Thomson, R., J. K. Pritchard, P. Shen, P. J. Oefner, and M. W. Feldman, 2000 Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. *Proc. Natl. Acad. Sci. USA* 97: 7360–7365.
- Thornton, K., 2003 Libsequence: a C++ class library for evolutionary genetic analysis. *Bioinformatics* 19: 2325–2327.

- Tian, F., N. M. Stevens, and E. S. Buckler, 2009 Tracking footprints of maize domestication and evidence for a massive selective sweep on chromosome 10. *Proc. Natl. Acad. Sci. USA* 106: 9979–9986.
- Ting, Y. C., 1965 Spontaneous chromosome inversions of Guatemalan teosintes (*Zea mexicana*). *Genetica* 36: 229–242.
- Ting, Y. C., 1967 Common inversion in maize and teosinte. *Am. Nat.* 101: 87–89.
- Ting, Y. C., 1976 Chromosome polymorphism of teosinte. *Genetics* 83: 737–742.
- Umina, P. A., A. R. Weeks, M. R. Kearney, S. W. Mckechnie, and A. A. Hoffmann, 2005 A rapid shift in a classic clinal pattern in *Drosophila* reflecting climate change. *Science* 308: 691–693.
- van Heerwaarden, J., J. Ross-Ibarra, J. Doebley, J. C. Glaubitz, J. Gonzalez Jde *et al.*, 2010 Fine scale genetic structure in the wild ancestor of maize (*Zea mays* ssp. *parviglumis*). *Mol. Ecol.* 19: 1162–1173.
- van Heerwaarden, J., J. Doebley, W. H. Briggs, J. C. Glaubitz, M. M. Goodman *et al.*, 2011 Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc. Natl. Acad. Sci. USA* 108: 1088–1092.
- Weber, A., R. M. Clark, L. Vaughn, J. De Jesus Sánchez-Gonzalez, J. Yu *et al.*, 2007 Major regulatory genes in maize contribute to standing variation in teosinte (*Zea mays* ssp. *parviglumis*). *Genetics* 177: 2349–2359.
- Weber, A. L., W. H. Briggs, J. Rucker, B. M. Baltazar, J. De Jesus Sanchez-Gonzalez *et al.*, 2008 The genetic architecture of complex traits in teosinte (*Zea mays* ssp. *parviglumis*): new evidence from association mapping. *Genetics* 180: 1221–1232.
- Weir, B. S., and C. C. Cockerham, 1984 Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.
- White, B. J., C. Cheng, D. Sangare, N. F. Lobo, F. H. Collins *et al.*, 2009 The population genomics of trans-specific inversion polymorphisms in *Anopheles gambiae*. *Genetics* 183: 275–288.
- Wilkes, H. G., 1967 Teosinte: the closest relative of maize. Ph.D. Thesis, Bussey Institute, Harvard University, Cambridge, MA.
- Wolfgruber, T. K., A. Sharma, K. L. Schneider, P. S. Albert, D. H. Koo *et al.*, 2009 Maize centromere structure and evolution: sequence analysis of centromeres 2 and 5 reveals dynamic loci shaped primarily by retrotransposons. *PLoS Genet.* 5: e1000743.
- Wright, S. I., and B. Charlesworth, 2004 The HKA test revisited: a maximum likelihood ratio test of the standard neutral model. *Genetics* 168: 1071–1076.
- Wright, S. I., I. V. Bi, S. G. Schroeder, M. Yamasaki, J. F. Doebley *et al.*, 2005 The effects of artificial selection on the maize genome. *Science* 308: 1310–1314.
- Yu, J., J. B. Holland, M. D. McMullen, and E. S. Buckler, 2008 Genetic design and statistical power of nested association mapping in maize. *Genetics* 178: 539–551.

Communicating editor: O. Savolainen

GENETICS

Supporting Information

<http://www.genetics.org/content/suppl/2012/04/27/genetics.112.138578.DC1>

Megabase-Scale Inversion Polymorphism in the Wild Ancestor of Maize

Zhou Fang, Tanja Pyhäjärvi, Allison L. Weber, R. Kelly Dawe, Jeffrey C. Glaubitz,
José de Jesus Sánchez González, Claudia Ross-Ibarra, John Doebley, Peter L. Morrell,
and Jeffrey Ross-Ibarra

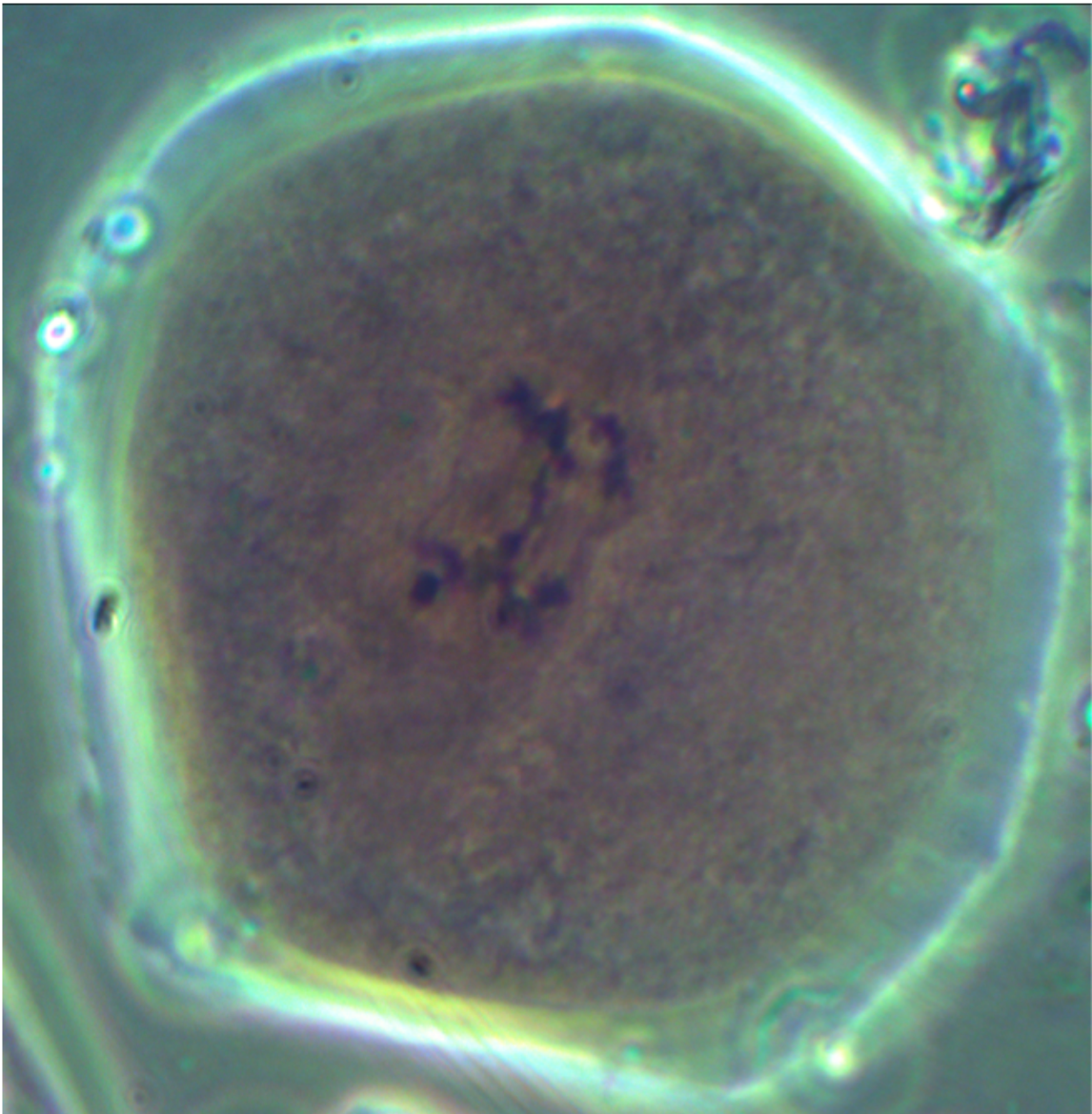


Figure S1 An anaphase I bridge in a plant heterozygous for *Inv1n*. Such bridges were rare, observed in only ~4% of the meiocytes undergoing anaphase I.

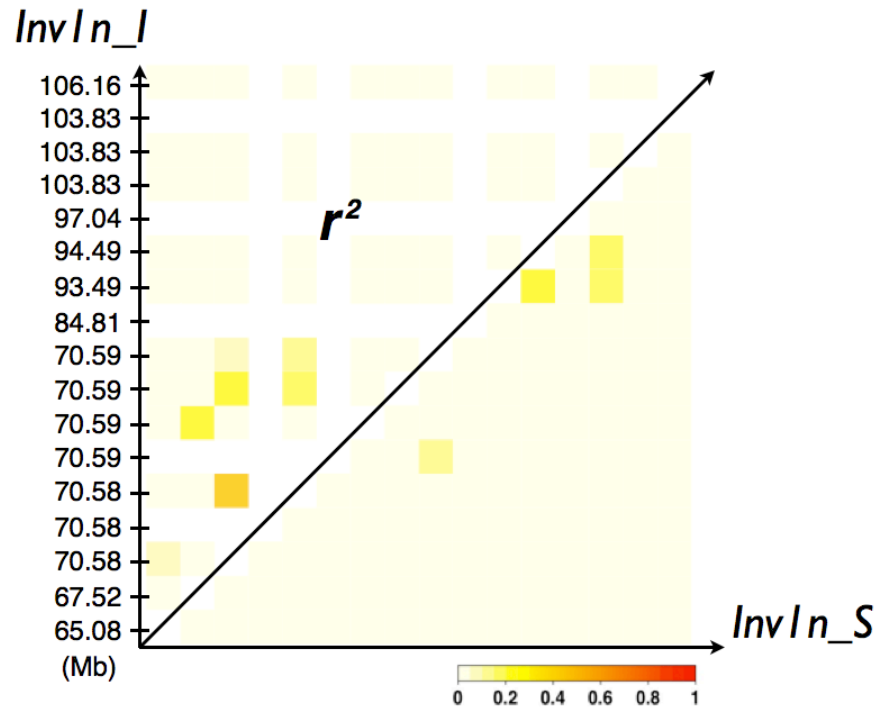


Figure S2 LD (r^2) among the 17 SNPs inside *Inv1n* in *parviglumis*. The physical positions of the 17 SNPs are shown on the left. The upper triangle represents LD in *Inv1n-I*, and *Inv1n-S* is shown in the lower triangle.

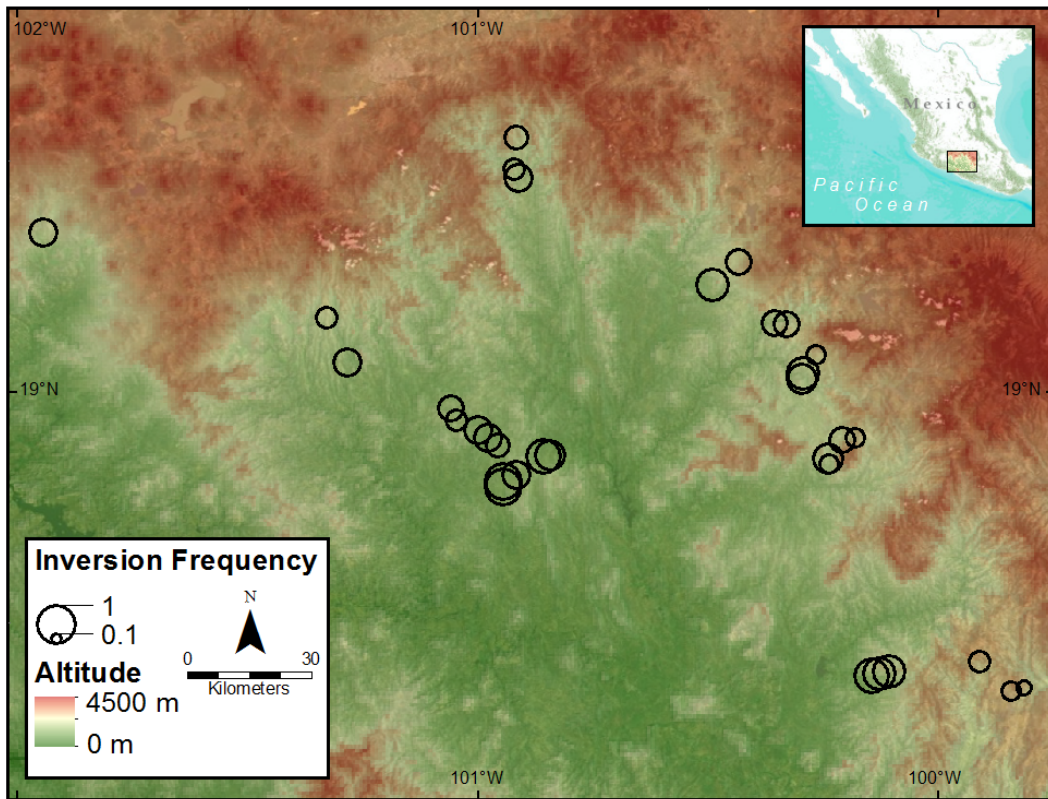


Figure S3 Geographic distribution of the 33 *parviglumis* populations. The size of the circle is proportional to the *Inv1n* frequency, and color represents elevation. The study area in Mexico is shown in the inset.

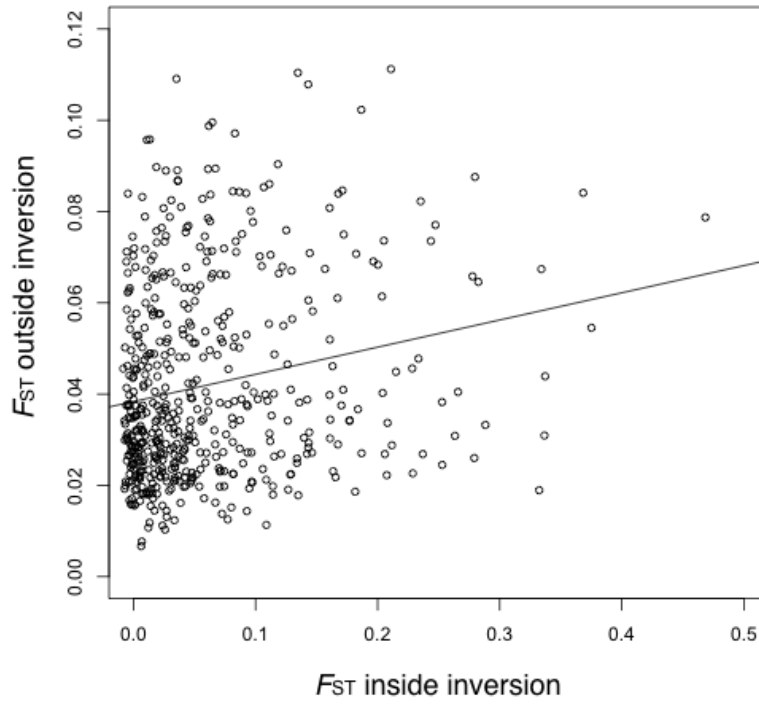


Figure S4 Pairwise F_{ST} among 33 *parviglumis* natural populations at SNPs inside *Inv1n* compared to SNPs outside *Inv1n*.

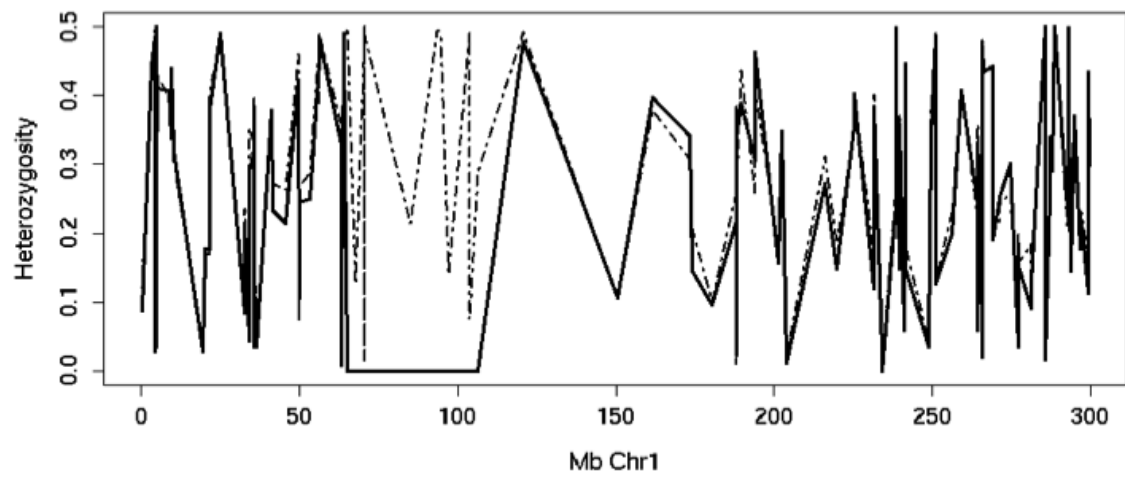


Figure S5 Expected SNP heterozygosity across chromosome 1 for all *parviglumis* (dashed line) and the most common *Inv1n-1* haplotype (solid line).

Table S1 Sample locality including germplasm type (in a separate file).

Table S1 is available for download at <http://www.genetics.org/content/suppl/2012/04/27/genetics.112.138578.DC1> as an Excel file.

Table S2 Data on 941 SNPs used in this paper

SNP	Chr	RefGen v2 Position
abph1.20	2	27,482,289
abph1.22	2	27,482,777
ae1.3	5	168,468,441
ae1.4	5	168,468,333
ae1.5	5	168,468,472
an1.4	1	241,222,410
ba1.6	3	183,084,850
ba1.9	3	183,085,136
bt2.5	4	58,957,188
bt2.7	4	58,956,243
bt2.8	4	58,956,482
Fea2.1	4	133,663,613
Fea2.5	4	133,663,191
id1.3	1	239,609,554
lg2.11	3	176,805,886
lg2.2	3	176,805,320
pbf1.1	2	153,513,462
pbf1.2	2	153,513,364
pbf1.3	2	153,512,659
pbf1.5	2	153,512,386
pbf1.6	2	153,511,544
pbf1.7	2	153,510,870
pbf1.8	2	153,510,843
PZA00003.11	6	121,697,500
PZA00004.2	5	13,621,123
PZA00005.8	4	237,391,332
PZA00005.9	4	237,391,476
PZA00006.13	6	86,257,528
PZA00006.14	6	86,257,555
PZA00008.1	2	12,024,277
PZA00010.5	1	120,848,790
PZA00013.10	1	264,163,949
PZA00013.11	1	264,163,707
PZA00013.9	1	264,164,037
PZA00015.4	9	113,420,505
PZA00017.1	1	35,578,083
PZA00018.5	8	156,627,475

PZA00029.11	2	147,311,129
PZA00029.12	2	147,311,042
PZA00031.5	5	123,964,207
PZA00041.3	6	122,710,163
PZA00042.2	6	156,981,860
PZA00042.5	6	156,981,744
PZA00043.7	7	172,926,856
PZA00045.1	8	118,191,980
PZA00049.12	8	134,066,246
PZA00050.9	1	296,575,998
PZA00051.2	4	239,429,365
PZA00058.5	8	6,023,537
PZA00058.6	8	6,023,577
PZA00060.2	9	129,011,584
PZA00061.1	2	33,346,033
PZA00065.2	1	49,621,367
PZA00069.4	5	215,340,163
PZA00070.5	9	88,242,881
PZA00078.2	4	18,682,174
PZA00079.1	10	18,859,495
PZA00081.17	1	45,513,286
PZA00084.2	7	41,732,007
PZA00084.3	7	41,731,951
PZA00088.3	3	230,181,537
PZA00090.2	8	138,794,764
PZA00092.1	4	114,693,724
PZA00092.5	4	114,693,700
PZA00093.2	8	124,657,588
PZA00096.26	7	149,603,394
PZA00097.13	8	146,033,256
PZA00098.14	1	161,615,948
PZA00100.10	3	5,673,891
PZA00100.12	3	5,673,605
PZA00100.14	3	5,673,926
PZA00100.9	3	5,673,572
PZA00106.9	1	10,096,325
PZA00107.18	3	182,088,854
PZA00108.12	2	13,851,566
PZA00108.14	2	13,851,376
PZA00108.15	2	13,851,427
PZA00109.3	3	113,066,409
PZA00109.5	3	113,066,444
PZA00111.2	7	139,366,525

PZA00111.4	7	139,366,225
PZA00111.5	7	139,366,472
PZA00111.6	7	139,366,469
PZA00111.8	7	139,366,372
PZA00114.3	3	146,252,963
PZA00116.2	9	129,409,633
PZA00119.4	1	84,814,783
PZA00120.4	8	148,515,290
PZA00123.1	9	153,738,351
PZA00125.2	4	36,210,500
PZA00131.14	1	203,947,049
PZA00132.17	7	31,773,777
PZA00132.18	7	31,773,878
PZA00132.3	7	31,773,709
PZA00135.6	10	109,712,602
PZA00137.2	1	248,738,069
PZA00139.14	4	20,200,937
PZA00140.10	2	143,045,557
PZA00140.6	2	143,045,483
PZA00140.9	2	143,045,604
PZA00142.6	8	144,271,083
PZA00163.4	2	223,310,842
PZA00164.1	2	220,832,970
PZA00164.2	2	220,833,009
PZA00164.3	2	220,832,898
PZA00166.1	3	173,587,461
PZA00166.3	3	173,587,211
PZA00170.1	2	69,704,230
PZA00170.3	2	69,704,287
PZA00170.4	2	69,704,335
PZA00174.1	10	66,638,175
PZA00174.2	10	66,638,311
PZA00175.2	1	8,558,256
PZA00176.8	2	10,624,968
PZA00177.4	8	156,772,795
PZA00178.3	7	170,990,302
PZA00182.3	6	113,558,220
PZA00182.4	6	113,558,165
PZA00184.1	9	155,683,609
PZA00184.4	9	155,683,487
PZA00188.1	7	2,998,218
PZA00188.3	7	2,998,557
PZA00191.5	5	2,120,221

PZA00192.6	1	35,579,277
PZA00192.7	1	35,579,161
PZA00193.2	4	192,205,587
PZA00198.39	6	164,372,830
PZA00200.11	2	5,819,590
PZA00200.17	2	5,820,190
PZA00200.9	2	5,819,731
PZA00201.2	3	199,274,401
PZA00204.1	10	127,348,939
PZA00210.1	3	30,057,289
PZA00210.6	3	30,057,176
PZA00211.7	7	131,186,613
PZA00212.1	6	72,869,016
PZA00213.19	9	130,918,999
PZA00214.1	6	92,813,618
PZA00216.9	4	2,999,034
PZA00218.1	4	71,640,363
PZA00218.6	4	71,640,430
PZA00219.7	3	220,883,383
PZA00220.11	3	208,344,179
PZA00220.12	3	208,344,329
PZA00221.7	1	173,317,166
PZA00225.8	9	107,625,055
PZA00226.7	2	42,792,015
PZA00227.8	7	167,241,354
PZA00230.5	1	299,528,078
PZA00232.24	3	174,582,549
PZA00234.21	3	223,307,712
PZA00235.6	1	293,795,824
PZA00235.8	1	293,795,893
PZA00237.2	2	6,049,080
PZA00237.7	2	6,048,795
PZA00237.8	2	6,048,913
PZA00238.3	2	197,595,234
PZA00240.9	1	41,104,364
PZA00241.6	6	150,304,658
PZA00243.27	1	296,891,619
PZA00245.14	1	285,522,115
PZA00245.16	1	285,522,140
PZA00245.17	1	285,521,947
PZA00245.18	1	285,521,975
PZA00245.19	1	285,522,004
PZA00249.2	5	100,929,495

PZA00250.1	6	132,889,427
PZA00251.1	10	122,801,848
PZA00254.3	3	143,488,417
PZA00255.15	5	170,416,067
PZA00255.17	5	170,416,167
PZA00256.16	7	17,520,123
PZA00256.21	7	17,520,225
PZA00256.23	7	17,520,294
PZA00257.11	10	10,882,576
PZA00257.22	10	10,882,651
PZA00261.6	5	75,935,660
PZA00263.14	1	65,086,477
PZA00266.5	6	161,454,751
PZA00273.1	5	84,064,839
PZA00274.7	1	41,377,563
PZA00277.17	2	144,729,592
PZA00277.9	2	144,729,626
PZA00280.14	2	62,991,918
PZA00287.1	9	75,780,129
PZA00289.11	1	216,050,825
PZA00294.20	1	63,891,664
PZA00297.2	3	40,322,808
PZA00297.3	3	40,322,831
PZA00297.4	3	40,322,907
PZA00298.4	8	168,273,519
PZA00298.5	8	168,273,438
PZA00299.2	8	79,007,430
PZA00300.12	5	172,395,470
PZA00300.13	5	172,395,364
PZA00300.14	5	172,395,199
PZA00300.16	5	172,395,504
PZA00301.3	1	106,159,660
PZA00303.19	3	189,347,439
PZA00303.21	3	189,347,297
PZA00307.12	1	293,041,636
PZA00307.14	1	293,041,693
PZA00309.2	3	1,298,636
PZA00310.5	10	20,409,576
PZA00314.6	8	154,651,600
PZA00314.8	8	154,651,315
PZA00315.1	4	40,157,629
PZA00315.6	4	40,157,532
PZA00318.2	1	256,084,651

PZA00323.3	9	146,915,036
PZA00323.4	9	146,915,233
PZA00326.16	7	159,131,005
PZA00326.18	7	159,131,080
PZA00326.19	7	159,131,038
PZA00332.8	4	189,042,100
PZA00332.9	4	189,042,120
PZA00334.2	8	155,483,084
PZA00337.3	10	86,424,870
PZA00337.4	10	86,424,768
PZA00337.5	10	86,424,803
PZA00342.9	10	130,512,105
PZA00344.10	4	187,355,764
PZA00345.15	5	62,595,345
PZA00346.1	7	145,729,154
PZA00346.2	7	145,729,102
PZA00346.3	7	145,729,100
PZA00349.3	2	21,283,667
PZA00349.5	2	21,283,640
PZA00350.2	7	152,832,335
PZA00352.22	5	191,590,886
PZA00355.1	6	79,915,144
PZA00355.2	6	79,915,132
PZA00356.9	1	264,533,418
PZA00364.5	1	269,025,539
PZA00364.6	1	269,025,550
PZA00367.2	2	164,723,975
PZA00369.1	1	219,807,376
PZA00370.1	7	128,121,229
PZA00370.5	7	128,121,322
PZA00380.5	3	39,822,327
PZA00380.7	3	39,822,411
PZA00381.3	1	238,343,204
PZA00381.4	1	238,343,143
PZA00381.5	1	238,343,616
PZA00382.17	6	118,604,616
PZA00385.3	1	292,796,839
PZA00386.3	7	170,656,392
PZA00390.6	2	209,927,372
PZA00391.2	3	158,354,531
PZA00392.3	7	144,981,423
PZA00392.4	7	144,981,553
PZA00393.1	1	4,193,258

PZA00393.4	1	4,193,538
PZA00394.11	2	56,966,318
PZA00395.1	5	203,679,356
PZA00395.2	5	203,679,327
PZA00396.12	2	4,692,070
PZA00401.11	5	56,019,395
PZA00401.6	5	56,019,323
PZA00406.1	5	92,183,810
PZA00407.9	5	49,522,205
PZA00408.7	5	215,357,992
PZA00409.3	10	62,249,775
PZA00411.1	1	251,088,838
PZA00411.4	1	251,088,969
PZA00411.5	1	251,088,935
PZA00413.17	3	128,907,269
PZA00413.18	3	128,907,125
PZA00413.21	3	128,907,248
PZA00417.2	8	45,262,229
PZA00417.3	8	45,262,014
PZA00419.1	8	126,550,861
PZA00420.4	3	228,448,528
PZA00422.2	8	154,682,448
PZA00422.5	8	154,682,386
PZA00422.6	8	154,682,317
PZA00423.16	7	172,769,215
PZA00423.17	7	172,769,301
PZA00424.1	7	174,785,900
PZA00425.4	1	21,468,390
PZA00425.9	1	21,468,287
PZA00429.1	8	145,842,587
PZA00433.5	5	214,074,969
PZA00436.7	4	6,408,499
PZA00439.6	6	141,610,864
PZA00440.1	6	32,969,931
PZA00442.3	2	63,668,565
PZA00442.4	2	63,668,537
PZA00442.5	2	63,668,622
PZA00442.6	2	63,668,616
PZA00444.1	10	106,952,925
PZA00444.5	10	106,952,878
PZA00445.18	4	49,917,528
PZA00449.2	5	38,201,854
PZA00458.6	3	5,674,594

PZA00459.5	2	17,094,806
PZA00460.3	8	164,859,711
PZA00460.5	8	164,859,744
PZA00460.7	8	164,859,811
PZA00462.2	6	44,008,563
PZA00463.3	10	13,493,461
PZA00466.1	9	12,104,064
PZA00468.11	1	50,096,641
PZA00468.7	1	50,096,930
PZA00470.1	5	170,023,744
PZA00471.2	2	22,504,931
PZA00471.3	2	22,505,161
PZA00471.4	2	22,505,011
PZA00472.2	5	63,229,797
PZA00477.10	1	299,197,598
PZA00477.11	1	299,197,721
PZA00477.5	1	299,197,562
PZA00477.9	1	299,197,628
PZA00478.10	2	165,667,714
PZA00478.11	2	165,667,750
PZA00478.7	2	165,668,066
PZA00478.9	2	165,667,779
PZA00480.10	3	23,306,494
PZA00481.7	8	24,530,674
PZA00484.5	1	180,308,158
PZA00485.2	2	124,057,780
PZA00487.16	2	225,160,494
PZA00487.24	2	225,160,213
PZA00487.26	2	225,160,158
PZA00489.1	6	51,226,441
PZA00495.3	2	173,250,676
PZA00495.4	2	173,250,689
PZA00495.6	2	173,250,638
PZA00496.1	7	167,915,230
PZA00497.1	2	19,918,031
PZA00497.4	2	19,917,977
PZA00498.4	8	52,299,554
PZA00499.10	5	35,276,873
PZA00499.12	5	35,276,915
PZA00499.3	5	35,276,907
PZA00501.12	6	34,440,016
PZA00501.14	6	34,440,097
PZA00502.5	9	143,803,192

PZA00503.5	6	22,667,086
PZA00504.1	7	147,716,994
PZA00504.2	7	147,716,948
PZA00505.4	7	168,254,993
PZA00505.8	7	168,255,179
PZA00510.2	6	131,323,502
PZA00510.3	6	131,323,440
PZA00514.1	9	107,313,968
PZA00514.6	9	107,313,892
PZA00514.7	9	107,314,063
PZA00515.14	2	172,144,187
PZA00516.3	8	164,866,616
PZA00517.6	5	18,348,509
PZA00522.12	5	58,603,542
PZA00523.2	4	149,284,134
PZA00525.16	2	2,539,802
PZA00525.2	2	2,539,982
PZA00527.6	2	219,860,018
PZA00527.9	2	219,860,155
PZA00529.3	4	235,165,818
PZA00531.1	8	125,316,626
PZA00533.3	2	190,430,209
PZA00533.4	2	190,430,143
PZA00533.5	2	190,430,284
PZA00533.6	2	190,430,239
PZA00534.2	7	21,636,438
PZA00536.2	6	123,777,598
PZA00538.12	3	208,472,870
PZA00538.16	3	208,473,135
PZA00538.8	3	208,472,941
PZA00543.2	6	81,692,897
PZA00543.4	6	81,693,175
PZA00543.5	6	81,693,243
PZA00545.21	5	207,708,730
PZA00545.22	5	207,708,616
PZA00545.4	5	207,708,479
PZA00547.13	5	93,579,904
PZA00547.18	5	93,579,495
PZA00552.4	10	26,987,438
PZA00562.4	10	61,954,391
PZA00565.3	3	73,964,566
PZA00568.19	9	133,422,395
PZA00573.3	5	160,830,866

PZA00578.1	6	63,950,351
PZA00579.6	9	151,724,182
PZA00582.4	6	62,187,756
PZA00586.1	3	187,713,152
PZA00587.3	10	84,610,447
PZA00587.6	10	84,610,480
PZA00588.2	9	61,356,589
PZA00588.4	9	61,356,555
PZA00589.10	9	43,679,421
PZA00589.8	9	43,679,500
PZA00589.9	9	43,679,530
PZA00593.2	1	93,492,705
PZA00595.3	5	165,658,227
PZA00600.11	8	85,123,811
PZA00603.1	5	184,494,852
PZA00608.1	3	68,056,795
PZA00608.5	3	68,056,885
PZA00610.18	1	281,206,778
PZA00610.9	1	281,206,931
PZA00613.22	2	4,193,245
PZA00614.12	7	151,523,785
PZA00615.3	3	1,885,957
PZA00615.6	3	1,885,816
PZA00615.8	3	1,885,793
PZA00617.16	1	53,270,492
PZA00618.22	3	190,889,590
PZA00620.2	2	10,516,860
PZA00621.2	3	69,856,000
PZA00622.1	6	114,439,165
PZA00622.2	6	114,438,752
PZA00623.2	1	294,729,591
PZA00626.3	5	66,831,722
PZA00626.4	5	66,831,656
PZA00630.9	4	96,312,649
PZA00637.4	2	172,403,167
PZA00639.12	2	210,078,347
PZA00639.13	2	210,078,538
PZA00639.15	2	210,078,233
PZA00641.7	8	167,465,970
PZA00641.8	8	167,465,876
PZA00644.11	4	25,326,089
PZA00647.9	10	130,080,970
PZA00650.8	1	189,588,779

PZA00654.10	1	32,543,238
PZA00654.12	1	32,543,175
PZA00655.1	7	139,358,571
PZA00656.15	10	17,602,152
PZA00656.16	10	17,602,091
PZA00656.18	10	17,601,973
PZA00656.4	10	17,601,894
PZA00658.19	1	225,130,798
PZA00658.23	1	225,131,040
PZA00662.3	4	124,359,583
PZA00665.6	10	69,031,285
PZA00667.1	3	162,743,710
PZA00672.6	2	232,004,707
PZA00672.8	2	232,004,713
PZA00673.2	1	259,012,580
PZA00674.3	5	169,236,723
PZA00676.2	2	217,281,403
PZA00680.1	2	1,094,384
PZA00680.3	2	1,094,394
PZA00682.2	4	239,370,840
PZA00684.12	5	920,922
PZA00686.8	8	148,078,863
PZA00692.5	1	94,489,227
PZA00693.3	9	28,399,313
PZA00695.1	7	171,232,753
PZA00698.4	1	274,531,768
PZA00700.3	4	3,262,444
PZA00704.11	4	129,738,534
PZA00705.5	1	97,038,983
PZA00706.16	8	161,188,592
PZA00710.1	5	62,200,260
PZA00710.16	5	62,200,400
PZA00712.4	10	117,991,835
PZA00715.3	5	193,301,336
PZA00717.14	8	70,199,281
PZA00719.1	3	144,344,237
PZA00719.2	3	144,344,055
PZA00719.3	3	144,343,938
PZA00720.2	2	153,510,503
PZA00720.3	2	153,510,547
PZA00721.4	6	147,137,475
PZA00721.5	6	147,137,547
PZA00725.4	2	204,879,868

PZA00726.6	4	91,996,108
PZA00726.7	4	91,996,159
PZA00726.9	4	91,996,053
PZA00727.11	10	143,259,325
PZA00727.12	10	143,259,404
PZA00729.18	2	188,980,715
PZA00729.19	2	188,980,910
PZA00730.2	2	37,867,237
PZA00731.6	1	9,304,156
PZA00731.7	1	9,304,303
PZA01104.1	6	126,148,975
PZA01149.1	9	34,845,337
PZA01149.3	9	34,845,357
PZA01182.1	5	162,837,167
PZA01240.1	4	19,597,341
PZA01240.2	4	19,597,206
PZA01420.1	3	88,528,316
PZA01420.2	3	88,528,312
PZA01420.3	3	88,528,283
PZA01474.2	2	47,506,839
PZA01637.2	4	181,187,885
PZA01637.3	4	181,187,743
PZA01637.4	4	181,187,733
PZA01725.1	1	20,115,876
PZA01725.2	1	20,115,867
PZA01782.2	9	19,325,456
PZA01782.3	9	19,325,169
PZA01782.4	9	19,325,560
PZA02789.31	3	223,691,724
PZA02789.36	3	223,691,657
PZA02791.6	1	67,524,140
PZA02792.16	5	21,893,193
PZA02792.9	5	21,893,329
PZA02806.4	1	34,255,672
PZA02806.9	1	34,255,898
PZA02807.5	5	168,532,545
PZA02808.12	2	44,168,169
PZA02808.16	2	44,168,094
PZA02819.35	4	173,601,472
PZA02820.6	5	201,667,378
PZA02822.2	5	149,459,255
PZA02824.1	3	218,898,837
PZA02824.3	3	218,898,805

PZA02825.8	4	174,596,692
PZA02831.5	2	133,841,755
PZA02837.5	1	288,742,212
PZA02844.1	9	17,751,142
PZA02850.18	7	174,836,174
PZA02850.4	7	174,836,372
PZA02853.10	10	25,976,137
PZA02853.7	10	25,976,183
PZA02856.1	3	156,482,524
PZA02865.11	4	215,239,863
PZA02869.2	1	4,608,173
PZA02869.8	1	4,608,015
PZA02872.1	7	13,174,365
PZA02872.3	7	13,174,261
PZA02878.12	9	37,409,502
PZA02888.3	3	138,624,226
PZA02890.3	2	190,160,002
PZA02890.4	2	190,159,942
PZA02890.5	2	190,159,924
PZA02894.1	5	945,545
PZA02897.12	9	97,246,219
PZA02906.12	3	148,289,175
PZA02906.7	3	148,289,102
PZA02921.9	1	24,941,000
PZA02923.7	3	161,729,466
PZA02927.1	8	34,033,391
PZA02938.5	4	68,402,538
PZA02939.6	2	159,987,717
PZA02940.3	7	150,550,725
PZA02941.3	10	71,361,885
PZA02941.6	10	71,361,989
PZA02941.8	10	71,362,041
PZA02947.2	6	56,218,574
PZA02949.22	4	186,040,517
PZA02949.26	4	186,040,520
PZA02952.10	3	196,048,525
PZA02954.2	1	19,460,965
PZA02955.3	8	14,778,018
PZA02958.17	6	151,822,737
PZA02959.7	7	157,280,779
PZA02961.1	10	16,212,441
PZA02962.13	1	3,274,647
PZA02963.5	1	173,991,765

PZA02966.11	2	130,956,103
PZA02968.4	4	39,092,497
PZA02969.11	10	144,061,127
PZA02970.9	6	98,315,872
PZA02972.1	4	3,257,744
PZA02982.5	4	149,276,720
PZA02982.6	4	149,276,909
PZA02983.38	7	4,238,932
PZA02988.2	4	134,622,435
PZA02993.5	10	106,792,585
PZA02997.16	9	150,813,060
PZA02997.19	9	150,813,124
PZA03001.15	1	231,701,106
PZA03001.18	1	231,701,045
PZA03001.9	1	231,701,259
PZA03009.5	3	11,649,287
PZA03009.6	3	11,649,362
PZA03009.7	3	11,649,402
PZA03009.8	3	11,649,449
PZA03011.6	3	203,579,387
PZA03012.10	8	116,625,263
PZA03013.7	4	139,753
PZA03013.8	4	139,810
PZA03014.10	1	103,826,380
PZA03014.21	1	103,826,237
PZA03014.24	1	103,826,094
PZA03017.10	4	237,610,824
PZA03017.11	4	237,610,827
PZA03024.16	5	199,380,293
PZA03024.18	5	199,380,332
PZA03024.7	5	199,380,207
PZA03028.5	1	150,682,455
PZA03032.16	3	173,016,042
PZA03034.1	5	6,823,291
PZA03035.5	8	77,590,436
PZA03037.8	1	287,063,325
PZA03037.9	1	287,063,398
PZA03041.8	6	115,129,610
PZA03042.1	5	65,114,161
PZA03042.5	5	65,114,362
PZA03046.2	5	3,206,772
PZA03046.3	5	3,206,775
PZA03047.12	6	31,412,258

PZA03047.20	6	31,412,046
PZA03047.22	6	31,412,177
PZA03049.23	5	89,317,871
PZA03051.1	10	96,330,328
PZA03051.3	10	96,330,666
PZA03052.15	8	130,730,669
PZA03054.3	3	32,026,203
PZA03054.5	3	32,026,168
PZA03058.17	9	19,393,926
PZA03062.15	9	3,927,674
PZA03062.7	9	3,927,845
PZA03063.17	6	35,896,303
PZA03063.18	6	35,896,343
PZA03064.6	1	249,417,407
PZA03067.17	7	14,460,265
PZA03067.20	7	14,460,194
PZA03068.11	6	140,409,576
PZA03068.13	6	140,409,675
PZA03069.6	6	82,967,424
PZA03073.23	3	170,110,501
PZA03073.24	3	170,110,525
PZA03074.24	1	201,486,818
PZA03078.33	10	6,594,010
PZA03081.1	4	209,799,999
PZA03081.10	4	209,800,055
PZA03081.11	4	209,800,051
PZA03081.13	4	209,800,213
PZA03081.6	4	209,800,094
PZA03083.7	2	162,471,818
PZA03089.12	6	165,631,271
PZA03090.31	6	71,014,658
PZA03092.7	5	12,038,114
PZA03094.18	5	173,805,293
PZA03094.6	5	173,805,347
PZA03095.1	5	66,432,514
PZA03095.2	5	66,432,703
PZA03095.3	5	66,432,443
PZA03097.4	9	44,336,288
PZA03097.7	9	44,336,417
PZA03097.9	9	44,336,534
PZA03102.10	6	145,649,127
PZA03102.2	6	145,649,201
PZA03102.9	6	145,649,075

PZA03137.1	10	28,603,151
PZA03172.2	5	188,967,385
PZA03223.3	7	143,204,647
PZA03258.2	7	143,203,880
PZA03283.2	3	7,774,330
PZA03284.3	5	92,369,229
PZA03290.1	6	9,144,627
PZA03290.2	6	9,144,765
PZA03295.4	7	172,723,384
PZA03296.6	6	84,226,532
PZA03296.7	6	84,226,556
PZA03298.1	5	21,933,689
PZA03298.2	5	21,933,666
PZA03301.2	1	241,298,148
PZA03301.4	1	241,298,319
PZA03302.1	3	119,175,697
PZA03305.6	1	287,728,642
PZA03305.7	1	287,728,786
PZA03311.2	4	154,166,345
PZA03311.3	4	154,166,386
PZA03311.4	4	154,166,658
PZA03311.5	4	154,166,660
PZA03312.1	5	193,819,059
PZA03312.2	5	193,819,018
PZA03316.2	8	4,792,016
PZA03319.3	2	31,493,278
PZA03319.4	2	31,493,137
PZA03320.3	5	190,746,629
PZA03320.4	5	190,746,566
PZA03328.5	2	36,843,149
PZA03329.1	7	158,299,030
PZA03329.2	7	158,298,929
PZA03333.3	6	143,961,781
PZA03335.2	5	211,666,584
PZA03335.3	5	211,666,758
PZA03337.1	4	47,242,400
PZA03340.2	5	20,310,727
PZA03342.2	10	59,878,382
PZA03344.4	7	22,014,951
PZA03344.5	7	22,015,062
PZA03344.6	7	22,015,028
PZA03345.1	1	36,509,795
PZA03345.2	1	36,509,841

PZA03345.4	1	36,509,719
PZA03347.1	4	154,166,530
PZA03348.1	6	159,694,555
PZA03349.1	3	211,536,618
PZA03349.9	3	211,536,839
PZA03767.1	1	56,289,498
PZA03767.4	1	56,289,530
PZA03767.5	1	56,289,258
PZA03773.2	3	88,529,456
PZA03773.3	3	88,529,363
PZA03774.1	5	181,588,954
PZA03774.10	5	181,591,578
PZA03774.2	5	181,589,966
PZA03774.4	5	181,590,507
PZA03774.5	5	181,590,649
PZA03774.6	5	181,590,922
PZA03774.8	5	181,591,195
PZA03774.9	5	181,591,369
PZA03775.1	2	47,509,672
PZA03775.11	2	47,507,587
PZA03775.2	2	47,509,355
PZA03775.3	2	47,509,287
PZA03775.4	2	47,508,979
PZA03775.6	2	47,508,409
PZA03775.7	2	47,508,297
PZA03775.8	2	47,508,224
PZA03775.9	2	47,508,186
PZA03781.1	5	162,834,983
PZA03781.2	5	162,835,430
PZA03781.3	5	162,835,530
PZA03781.4	5	162,835,750
PZA03781.5	5	162,835,982
PZA03781.6	5	162,836,268
PZA03781.7	5	162,836,471
PZA03781.8	5	162,836,647
PZA03782.1	5	80,719,678
PZA03782.3	5	80,720,324
PZA03786.1	9	34,845,691
PZA03786.2	9	34,845,657
PZA03789.1	8	70,201,052
PZA03789.2	8	70,201,552
PZA03789.4	8	70,201,822
PZB00011.4	1	234,058,603

PZB00011.5	1	234,058,440
PZB00041.2	8	44,554,499
PZB00041.4	8	44,554,252
PZB00049.2	8	21,102,553
PZB00049.4	8	21,102,159
PZB00049.7	8	21,102,626
PZB00055.1	10	140,854,286
PZB00060.4	3	133,090,703
PZB00062.6	7	150,661,551
PZB00062.7	7	150,661,408
PZB00062.8	7	150,661,711
PZB00067.2	1	277,152,522
PZB00067.3	1	277,152,084
PZB00067.4	1	277,152,753
PZB00067.5	1	277,152,107
PZB00078.1	2	32,588,115
PZB00081.2	5	7,001,502
PZB00081.4	5	7,001,339
PZB00081.5	5	7,002,138
PZB00081.7	5	7,001,555
PZB00092.1	4	157,210,342
PZB00092.4	4	157,210,317
PZB00093.3	4	123,980,253
PZB00093.4	4	123,980,408
PZB00093.6	4	123,980,230
PZB00096.2	8	21,100,594
PZB00096.3	8	21,100,482
PZB00136.3	5	5,886,221
PZB00140.1	7	2,062,178
PZB00145.2	8	44,553,085
PZB00149.2	1	188,182,081
PZB00149.4	1	188,182,112
PZB00153.1	4	123,979,525
PZB00153.2	4	123,979,477
PZB00153.3	4	123,979,555
PZB00153.5	4	123,979,344
PZB00160.1	8	158,759,297
PZB00160.2	8	158,759,347
PZB00160.4	8	158,759,510
PZB00165.2	2	43,558,423
PZB00165.6	2	43,558,519
PZB00169.4	7	133,170,656
PZB00169.6	7	133,170,485

PZB00175.1	7	110,332,040
PZB00175.2	7	110,331,886
PZB00175.3	7	110,332,043
PZB00175.4	7	110,332,075
PZB00175.5	7	110,331,867
PZB00180.1	5	80,718,715
PZB00180.2	5	80,719,063
PZB00183.3	2	43,558,626
PZB00207.3	9	135,803,288
PZB00221.3	9	146,686,210
PZB00221.8	9	146,686,602
PZB00229.3	4	157,210,028
PZB00232.1	5	64,118,796
PZB00232.2	5	64,119,272
PZB00232.4	5	64,119,116
PZB00232.5	5	64,119,253
PZB00379.3	9	26,715,605
PZB00379.4	9	26,715,588
PZB00379.5	9	26,715,442
PZB00393.7	1	202,305,125
PZB00409.3	10	84,096,438
PZB00416.2	6	131,158,004
PZB00416.5	6	131,157,450
PZB00454.2	1	188,182,192
PZB00454.3	1	188,182,145
PZB00454.4	1	188,182,274
PZB00454.5	1	188,182,373
PZB00498.2	5	88,793,779
PZB00498.4	5	88,793,994
PZB00598.1	7	167,238,971
PZB00598.2	7	167,238,920
PZB00603.3	7	171,703,464
PZB00603.4	7	171,703,461
PZB00603.5	7	171,703,857
PZB00607.2	2	108,634,037
PZB00761.1	9	86,863,679
PZB00761.2	9	86,864,051
PZB00849.2	10	59,516,563
PZB00849.3	10	59,516,587
PZB00849.4	10	59,516,627
PZB00859.1	1	157,104
PZB01109.2	3	196,440,909
PZB01109.3	3	196,441,027

PZB01110.1	9	24,069,690
PZB01110.2	9	24,069,762
PZB01110.3	9	24,069,506
PZB01111.6	10	134,413,629
PZB01111.7	10	134,413,440
PZB01111.8	10	134,413,161
PZB01112.3	5	69,120,089
PZB01112.4	5	69,119,579
PZB01112.5	5	69,119,926
PZB01112.6	5	69,119,810
PZB01113.4	4	147,122,242
PZB01114.1	9	63,685,643
PZB01114.3	9	63,685,291
PZB01115.1	5	61,672,319
PZB01115.5	5	61,672,126
PZB01115.6	5	61,672,027
PZB01116.2	5	120,484,698
PZB01221.1	6	83,673,652
PZB01222.1	6	164,321,896
PZB01222.3	6	164,321,834
PZB01223.3	3	194,674,364
PZB01223.7	3	194,674,207
PZB01225.1	1	63,212,578
PZB01225.2	1	63,212,688
PZB01225.4	1	63,212,638
PZB01228.1	7	15,457,461
PZB01228.3	7	15,457,543
PZB01228.4	7	15,457,589
PZB01233.2	2	3,388,933
PZB01233.3	2	3,388,514
PZB01238.5	7	166,860,469
PZB01238.6	7	166,860,367
PZB01427.1	1	271,346,409
PZB01427.3	1	271,346,495
PZB01463.2	3	158,567,940
PZB01463.3	3	158,568,190
PZB01463.4	3	158,568,045
PZD00003.1	3	162,800,088
PZD00003.3	3	162,800,459
PZD00007.1	3	165,178,026
PZD00008.3	3	165,176,818
PZD00011.1	6	131,832,400
PZD00011.3	6	131,831,855

PZD00011.4	6	131,831,919
PZD00012.1	6	131,825,470
PZD00012.2	6	131,826,708
PZD00012.3	6	131,826,731
PZD00012.4	6	131,826,486
PZD00012.5	6	131,826,355
PZD00013.3	3	137,233,322
PZD00013.4	3	137,233,307
PZD00014.3	3	137,230,750
PZD00017.1	5	196,556,802
PZD00019.1	1	4,877,306
PZD00020.2	1	4,862,839
PZD00020.3	1	4,863,420
PZD00020.4	1	4,863,098
PZD00020.6	1	4,863,339
PZD00021.2	1	4,862,689
PZD00021.4	1	4,862,621
PZD00021.5	1	4,862,745
PZD00022.1	2	235,853,547
PZD00022.3	2	235,853,500
PZD00022.4	2	235,852,911
PZD00024.2	8	22,248,100
PZD00025.1	8	22,245,644
PZD00025.2	8	22,246,020
PZD00030.1	4	178,888,718
PZD00030.4	4	178,889,674
PZD00030.5	4	178,889,949
PZD00030.6	4	178,889,456
PZD00034.3	8	22,978,660
PZD00043.1	7	164,417,453
PZD00043.2	7	164,417,394
PZD00043.3	7	164,417,373
PZD00043.4	7	164,417,293
PZD00044.2	7	164,412,086
PZD00044.3	7	164,412,214
PZD00044.4	7	164,412,165
PZD00045.1	7	164,413,574
PZD00045.2	7	164,413,483
PZD00045.3	7	164,413,785
PZD00045.4	7	164,413,134
PZD00049.3	9	17,019,874
PZD00049.4	9	17,019,800
PZD00049.5	9	17,019,677

PZD00051.1	1	194,013,409
PZD00052.3	1	194,021,359
PZD00052.4	1	194,021,224
PZD00062.2	1	265,690,409
PZD00066.1	1	266,095,785
PZD00067.1	4	133,663,998
PZD00067.2	4	133,663,533
PZD00067.3	4	133,663,860
PZD00068.1	1	239,609,530
PZD00069.2	1	292,891,821
PZD00069.3	1	292,891,218
PZD00069.4	1	292,892,381
PZD00069.5	1	292,891,797
PZD00073.1	7	110,331,963
PZD00073.2	7	110,331,259
PZD00073.6	7	110,331,230
PZD00074.1	2	3,387,857
PZD00075.1	10	147,059,868
PZD00075.2	10	147,060,198
PZD00076.1	1	70,581,280
PZD00076.2	1	70,581,470
PZD00076.4	1	70,581,587
PZD00077.10	1	70,592,301
PZD00077.5	1	70,593,006
PZD00077.7	1	70,592,373
PZD00077.8	1	70,592,832
sh2.5	3	216,418,060
sh2.6	3	216,418,000
sh2.7	3	216,417,426
sh2.9	3	216,418,124
su1.4	4	41,369,511
su1.5	4	41,378,878
su1.7	4	41,376,540
tb1.17	1	265,747,286
tb1.18	1	265,747,334
tb1.19	1	265,747,374
tb1.5	1	265,745,267
te1.3	3	165,177,210
te1.4	3	165,176,963
zagl1.1	1	4,877,511
zagl1.6	1	4,862,498
zap1.2	2	235,852,993
zfl2.6	2	12,642,670

Table S3 Data on 95 resequencing loci used

Locus	Length	Sample size	RefGen v2 position
PZA00624	519	15	1,177,470
PZA02851	347	16	3,834,244
PZA02869	360	12	4,607,915
PZA00181	271	15	8,351,398
PZA00528	577	15	8,352,809
PZA00175	329	14	8,558,092
PZA00447	398	15	9,052,384
PZA00106	369	16	10,096,199
PZA00697	433	16	17,856,957
PZA00358	546	14	18,980,277
PZA00631	554	13	20,113,871
PZA00491	372	14	20,419,065
PZA00688	290	16	21,461,304
PZA00425	296	14	21,468,171
PZA02921	450	14	24,940,726
PZA00691	323	11	25,415,740
PZA03004	359	16	26,434,743
PZA00021	373	16	28,799,398
PZA00654	328	15	32,543,070
PZA02790	420	15	33,722,666
PZA00017	547	16	35,577,839
PZA00192	542	13	35,579,032
PZA00240	487	16	41,104,182
PZA00014	363	13	41,278,262
PZA00605	544	15	42,476,893
PZA00661	267	15	47,506,412
PZA00065	459	12	49,621,287
PZA00468	597	16	50,096,486
PZA00617	397	14	53,270,379
PZA02826	355	11	56,220,687
PZA00328	633	10	59,581,784
PZA00378	685	13	63,008,454
PZA00263	249	13	65,086,382
PZA02791	233	13	67,523,911
PZA00593	357	16	93,492,381
PZA00692	233	15	94,489,012
PZA00146	249	16	100,583,349
PZA03014	477	13	103,826,023
PZA00301	323	16	106,159,360
PZA00205	331	12	119,652,952
PZA00010	426	14	120,848,690
PZA03086	722	14	125,074,531
PZA00075	307	16	138,146,925
PZA00083	391	16	151,738,698
PZA00098	396	12	161,615,693
PZA03101	359	12	161,775,074
PZA00660	420	13	170,476,533

PZA02963	476	12	173,991,638
PZA00455	377	12	179,989,018
PZA00484	272	12	180,308,017
PZA00375	312	16	181,496,388
PZA00421	453	11	182,092,438
PZA00068	559	10	183,985,724
PZA00544	383	12	188,182,045
PZA00650	305	11	189,588,623
PZA00469	494	12	195,187,825
PZA00435	317	15	199,902,446
PZA03074	497	13	201,486,648
PZA00131	324	13	203,947,004
PZA00369	383	15	219,807,209
PZA02960	310	15	223,825,184
PZA00403	276	16	224,079,008
PZA03091	376	15	224,522,055
PZA02823	346	15	226,625,187
PZA00664	247	12	228,194,890
PZA03001	639	10	231,700,907
PZA00137	327	11	248,737,932
PZA00411	447	15	251,088,838
PZA00318	306	16	256,084,575
PZA00035	380	16	260,619,385
PZA00339	348	11	261,177,655
PZA00709	562	15	262,720,110
PZA02985	368	14	263,200,608
PZA00013	566	15	264,163,639
PZA00364	328	13	269,025,379
PZA00698	327	13	274,531,678
PZA00036	350	13	275,249,952
PZA02957	383	14	282,796,262
PZA02837	303	11	288,742,129
PZA02935	409	13	289,217,209
PZA00150	240	13	289,336,993
PZA00313	517	13	289,566,801
PZA00520	379	14	291,254,752
PZA00242	341	11	292,089,628
PZA03033	355	11	292,875,379
PZA00235	415	10	293,795,684
PZA00623	535	10	294,729,356
PZA00343	563	12	295,762,441
PZA00703	582	10	295,790,731
PZA03018	431	12	296,020,928
PZA00050	403	13	296,575,866
PZA00243	528	14	296,891,356
PZA03006	501	12	298,408,692
PZA00477	411	14	299,197,483
PZA00230	323	15	299,527,976

Table S4 Location of the 33 *parviglumis* study populations with mean per-SNP values of summary statistics

Population	State/	Latitude	Longitude	Altit-	N	F	HWE_			Fay/	
	Province			ude			P-value	θ_H	π	TajD	Wu H
Crustel	Guererro	18.383	-100.145	985	22	0.84	0.48	0.32	0.26	1.31	-0.06
Amates 1	Guererro	18.388	-100.128	1110	25	0.74	0.47	0.31	0.28	1.36	-0.03
Amates 2	Guererro	18.394	-100.108	1210	24	0.75	0.39	0.30	0.28	1.31	-0.02
Iguala	Guererro	18.414	-99.909	1506	28	0.36	0.96	0.31	0.28	1.51	-0.03
Rincon	Guererro	18.350	-99.841	1624	28	0.30	0.20	0.31	0.27	1.68	-0.04
Ahuacatitlan	Guererro	18.356	-99.814	1528	19	0.18	0.59	0.29	0.27	1.43	-0.02
Huetamo 1	Michoacan	19.063	-101.283	832	25	0.58	0.63	0.31	0.25	1.28	-0.06
Puerto 1	Michoacan	18.963	-101.058	870	20	0.53	0.66	0.32	0.26	1.26	-0.06
Zapote	Michoacan	18.938	-101.048	915	24	0.38	0.59	0.31	0.26	1.33	-0.05
Puerto 2	Michoacan	18.916	-101.000	727	20	0.56	0.58	0.31	0.24	1.12	-0.07
Huetamo 2	Michoacan	18.900	-100.979	677	20	0.60	0.46	0.30	0.25	1.09	-0.05
Cuirindalillo	Michoacan	18.883	-100.957	697	21	0.40	0.69	0.31	0.25	1.03	-0.06
Crucero	Michoacan	18.794	-100.946	653	25	0.90	0.02	0.30	0.23	1.17	-0.07
Quenchendio	Michoacan	18.805	-100.946	635	26	0.88	0.51	0.31	0.25	1.09	-0.06
Potrero	Michoacan	18.820	-100.916	654	20	0.60	0.85	0.30	0.26	1.12	-0.04
Crucita	Michoacan	18.858	-100.857	609	29	0.78	0.56	0.31	0.25	1.25	-0.06
Guayabo	Michoacan	18.862	-100.844	555	27	0.61	0.12	0.31	0.25	1.23	-0.06
Toluca 1	Mexico	18.899	-100.181	1422	24	0.31	0.74	0.31	0.27	1.33	-0.04
Toluca 2	Mexico	18.895	-100.209	1355	23	0.50	0.86	0.31	0.28	1.22	-0.03
Toluca 3	Mexico	18.854	-100.239	1015	19	0.66	0.82	0.32	0.27	1.06	-0.05
Salitre-Monte	Mexico	18.842	-100.238	958	23	0.28	0.23	0.31	0.27	1.17	-0.04
Taretan	Michoacan	19.344	-101.944	1170	18	0.58	0.90	0.29	0.21	0.98	-0.08
Los Guajes	Michoacan	19.231	-100.491	985	27	0.76	0.65	0.31	0.27	1.36	-0.04
Norte	Michoacan	19.281	-100.434	1332	27	0.50	0.85	0.30	0.27	1.33	-0.03
Zuluapan 1	Mexico	19.148	-100.355	1178	28	0.53	0.46	0.31	0.28	1.33	-0.03
Zuluapan 2	Mexico	19.146	-100.329	1346	30	0.52	0.73	0.31	0.27	1.30	-0.04
Zacazonapan 1	Mexico	19.079	-100.266	1468	28	0.27	0.99	0.30	0.26	1.33	-0.04
Zacazonapan 2	Mexico	19.039	-100.295	1085	28	0.70	0.71	0.32	0.27	1.38	-0.05
El Puente	Mexico	19.029	-100.296	1075	24	0.67	0.76	0.31	0.27	1.30	-0.04
Queretanillo	Michoacan	19.551	-100.918	1342	28	0.41	0.32	0.32	0.26	1.33	-0.06
Temascal 1	Michoacan	19.483	-100.921	1100	28	0.36	0.80	0.32	0.27	1.33	-0.05
Temascal 2	Michoacan	19.464	-100.912	1030	19	0.55	0.27	0.31	0.27	1.18	-0.04
Casa Blanca	Michoacan	19.161	-101.329	1268	26	0.33	0.00	0.33	0.26	1.40	-0.07

N: sample size; F: frequency of *Inv1n-l*.

Table S5 Counts of anaphase and telophase pollen meiocytes showing dicentric bridges or normal segregation during meiosis

	Line	Normal	Bridge	Sum
1	B73 x TIL5	17	0	17
2	B73 x TIL5	45	3	48
3	OH43 x TIL11	48	2	50
4	OH43 x TIL11	36	1	37
5	OH43 x TIL11	4	0	4
6	OH43 x TIL11	17	1	18

Table S6 Mean Bayes factors for all environmental variables and inversion as single marker, all the SNPs in *Inv1n* and all SNPs. T: temperature.

	Inversion	SNPs in <i>Inv1n</i>	All SNPs
Longitude	0.36	0.69	1020.37
Latitude	1.64	0.87	34.59
Altitude	136.37	124.63	5.86
Annual Mean T	18.93	12.57	5.75
Mean Diurnal T Range	0.82	4.35	28.61
Isothermality	1.26	0.84	2.77
T Seasonality	0.92	0.87	2.15
Max T of Warmest Month	22.46	16.42	5.76
Min T of Coldest Month	11.47	5.69	7.43
T Annual Range	1.90	9.75	28.75
Mean T of Wettest Quarter	21.04	17.20	9.33
Mean T of Driest Quarter	48.98	26.87	3.35
Mean T of Warmest Quarter	21.27	13.83	5.30
Mean T of Coldest Quarter	17.19	10.34	6.66
Annual Precipitation	0.88	1.26	2.69
Precipitation of Wettest Month	0.33	0.53	19.25
Precipitation of Driest Month	47.29	17.46	2.40
Precipitation Seasonality	10.54	3.67	2.24
Precipitation of Wettest Quarter	0.46	0.70	2.89
Precipitation of Driest Quarter	34.48	17.86	4.53
Precipitation of Warmest Quarter	0.44	0.51	1.73
Precipitation of Coldest Quarter	5.21	2.45	1.43

Table S7 Results of association analysis. P-value, marker r^2 , a (genotypic value of inversion homozygote) and d (genotypic value of heterozygote) are based on analysis of *Inv1n* as single marker. Number of significant SNPs at FDR 5% is reported separately for all SNPs and for SNPs inside *Inv1n*.

Trait	Inversion as single marker				Association analysis for all SNPs	
	p	marker r^2	a	d	Significant SNPs at FDR 5%	Significant SNPs in inversion at FDR 5%
Blade Length ^a	0.7065	0.001	-0.33	0.15	0	0
Culm diameter	0.0137	0.011	-0.55	0.54	7	4
Days to Pollen	0.2595	0.004	-0.50	0.69	0	0
Days to Silk	0.4831	0.002	-0.33	0.54	0	0
Female ear length ^b	0.3304	0.005	1.18	-0.80	0	0
Fruitcase compression ^b	0.0662	0.007	-0.11	0.03	0	0
Fruitcase length ^b	0.1393	0.008	0.12	-0.03	0	0
Fruitcase weight	0.5961	0.001	0.00	0.00	0	0
Lateral branch internode number ^a	0.663	0.001	-1.05	0.56	0	0
Lateral Branch Length ^a	0.8508	0.001	-0.02	-0.06	0	0
Lateral inflorescence branch number ^a	0.5653	0.002	-0.62	0.42	0	0
Lateral inflorescence length	0.3641	0.005	-0.65	0.94	0	0
Leaf Number	0.0232	0.01	-0.12	0.68	0	0
Leaf Width	0.1243	0.005	-0.13	0.11	2	0
Maize Introgressed	0.5062	0.002	-0.01	0.00	4	0
Mean lateral branch internode length ^a	0.7455	0.001	-0.27	0.32	0	0
Number of Barren nodes	0.7907	0.001	0.00	0.01	0	0
Number of female cupules ^b	0.7164	0.001	0.03	-0.13	0	0
Number of female internodes ^b	0.8604	0.001	0.06	-0.09	0	0
Oil Content, Wet	0.7653	0.001	0.01	0.04	0	0
Paired Spikelets	0.5523	0.002	0.00	0.00	5	0
Pedicellate Spikelet	0.1717	0.006	0.00	0.01	15	1
Percent of Male Internodes in the lateral inflorescence	0.0069	0.023	-2.62	2.75	0	0
Percent staminate spikelets in the lateral inflorescence	0.0055	0.024	-1.96	2.08	4	0

Plant Height	0.1175	0.005	-1.21	6.42	0	0
Polystichous	0.1162	0.007	-0.01	0.02	2	0
Prolificacy ^a	0.3762	0.002	-0.41	0.28	0	0
Proportion of female cupules ^b	0.2483	0.007	0.02	-0.01	0	0
Proportion of female ear length ^b	0.3717	0.005	0.02	-0.01	0	0
Proportion of female internodes ^b	0.2604	0.007	0.02	-0.01	0	0
Protein Content	0.5198	0.002	-0.22	0.07	0	0
Sexual Identity of the Lateral Inflorescence	0.294	0.003	-0.02	0.05	0	0
Sheath Length	0.9239	0	-0.03	0.04	0	0
Starch Content ^c	0.4999	0.002	0.21	-0.12	0	0
Tassel branch number	0.1425	0.01	-5.80	0.64	0	0
Tiller number	0.2321	0.004	0.55	0.58	0	0
Yoked Cupules	0.7034	0.001	0.00	0.00	21	0

^a On the second lateral branch from the top of the plant

^b In the basal ear

^c Adjusted to percent dry matter using a pooled moisture content estimate