Genetic signals of origin, spread, and introgression in a large sample of maize landraces

Joost van Heerwaarden[a,1], John Doebley[b], William H. Briggs[c], Jeffrey C. Glaubitz[d], Major M. Goodman[e], Jose de Jesus Sanchez Gonzalez[f], and Jeffrey Ross-Ibarra[a,1]

[a]Department of Plant Sciences, University of California, Davis, CA 95616; [b]Department of Genetics, University of Wisconsin, Madison, WI 53706; [c]Syngenta Seeds, 1601 BK, Enkhuizen, The Netherlands; [d]Institute for Genomic Diversity, Cornell University, Ithaca, NY 14853; [e]Department of Crop Science, North Carolina State University, Raleigh, NC 27695; and [f]Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara, Zapopan, Jalisco CP45110, Mexico

Edited by Dolores R. Piperno, Smithsonian National Museum of Natural History and Smithsonian Tropical Research Institute, Fairfax, VA, and approved December 9, 2010 (received for review August 31, 2010)

The last two decades have seen important advances in our knowledge of maize domestication, thanks in part to the contributions of genetic data. Genetic studies have provided firm evidence that maize was domesticated from Balsas teosinte (*Zea mays* subspecies *parviglumis*), a wild relative that is endemic to the mid- to lowland regions of southwestern Mexico. An interesting paradox remains, however: Maize cultivars that are most closely related to Balsas teosinte are found mainly in the Mexican highlands where subspecies *parviglumis* does not grow. Genetic data thus point to primary diffusion of domesticated maize from the highlands rather than from the region of initial domestication. Recent archeological evidence for early lowland cultivation has been consistent with the genetics of domestication, leaving the issue of the ancestral position of highland maize unresolved. We used a new SNP dataset scored in a large number of accessions of both teosinte and maize to take a second look at the geography of the earliest cultivated maize. We found that gene flow between maize and its wild relatives meaningfully impacts our inference of geographic origins. By analyzing differentiation from inferred ancestral gene frequencies, we obtained results that are fully consistent with current ecological, archeological, and genetic data concerning the geography of early maize cultivation.

The geography of origins and diversification of agricultural species has important implications for unraveling the ecological context of Neolithic societies and for understanding current patterns of diversity in domesticated plants and animals. Traditionally the realm of archeology and botany (1, 2), the study of plant domestication has seen important contributions from genetics during the last two decades (3). Genetic data often provide evidence that is hard to obtain by other means, making it an invaluable complement to other lines of inquiry.

As a case in point, molecular markers were instrumental in establishing the single domestication of maize (*Zea mays* subspecies *mays*) from an extant wild relative (4, 5). Maize was shown to originate from annual teosinte (*Zea mays* subspecies *parviglumis*, hereafter *parviglumis*) around 9,000 y B.P., placing domestication in the mid- to lowland regions of southwest Mexico where *parviglumis* grows endemically. As predicted by this result (6), excavations in the heart of *parviglumis'* distribution have produced the earliest (8,700 y B.P.) phytolith evidence for maize cultivation (7). Other finds from Tabasco (7,300 y B.P.) (8) and Panama (7,400 y B.P.) (9) also support an early presence of maize throughout the Meso-American lowlands.

Although different types of evidence seemingly concur, questions nonetheless remain about the interpretation of the genetic data. While unequivocal with respect to maize's wild ancestor, marker evidence suggests that maize from the Mexican highlands, rather than from the lowlands, is most closely related to *parviglumis* and appears to have given rise to all cultivars currently grown throughout the Americas (5). That the oldest macrobotanical remains (6,200 y B.P.) are found in the highlands supports this observation (10). This gap between the location of maize's wild ancestor and the most ancestral maize population is paradoxical (5) and raises questions about how to reconcile the genetic ancestry of modern maize with the genetic and archeological evidence supporting domestication at lower altitudes.

Two explanations have been proposed for the ancestral position of highland maize. First, *parviglumis* may have grown in the highlands at the time of domestication (5, 11). Second, the early domesticate may have spread from the lowlands to the highlands, with a subsequent diffusion of highland maize replacing lowland populations (11). Neither resolution is particularly satisfying: *parviglumis* probably grew at lower altitudes during the cooler and dryer climatic conditions that likely existed around the time of maize domestication (7, 12, 13), and the replacement hypothesis seems unlikely given the difference in ecological adaptation between highland and lowland maize (14).

Some existing evidence suggests a third solution to the paradox. Maize in the Mexican highlands grows sympatrically with a second subspecies of annual teosinte, *Zea mays* subspecies *mexicana* (hereafter, *mexicana*). Maize and *mexicana* are interfertile (15), and there is evidence for gene flow from *mexicana* into maize (5, 16). Although not directly ancestral to maize, *mexicana* is more closely related to *parviglumis* than to maize (5), so gene flow from *mexicana* has the potential to affect the genetic similarity between highland maize populations and *parviglumis*.

Here we used a large SNP dataset of maize and teosinte (Fig. S1 and Dataset S1) to reevaluate the genetic ancestry of extant maize populations and to estimate the effects of introgression from maize's wild relatives. We overcame the limitations of crop/wild ancestor comparisons in the presence of introgression by using estimates of differentiation from ancestral gene frequencies, inferred from extant maize populations, as a measure of genetic distance from the domesticated population. We located the geographic region most closely associated with ancestral maize based on spatial estimates of gene frequency. Our results differ from previous work, identifying a region consistent with genetic and archeological evidence for maize domestication in the lowlands and suggesting that the apparent ancestral position of highland Mexican maize results from introgression from *mexicana*. These results highlight the impact of information about gene flow on historical inference from genetic data and the utility of alternative methods in reconstructing crop ancestry and geographic history.

Author contributions: J.D. designed research; W.H.B. and J.C.G. performed research; M.M.G. and J.d.J.S.G. contributed new reagents/analytic tools; J.v.H. and J.R.-I. analyzed data; and J.v.H., J.D., and J.R.-I. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. E-mail: jvheerwaarden@gmail.com or rossibarra@ucdavis.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1013011108/-/DCSupplemental.

Results

Patterns of Genetic Structure and Differentiation. Principal components analysis (PCA) (17) of the maize SNP data identifies 58 significant principal components (PCs) (explaining 37.6% of total variance), probably reflecting isolation by distance (18) and linkage effects (19). We use the first nine PCs, which present the strongest spatial autocorrelation (Fig. S2) and explain a large portion of the total variance (18.7%), to cluster the accessions into 10 geographically distinct groups (Fig. 1*A*). Meso-American maize falls into three groups: the Meso-American Lowland group, which includes predominantly lowland accessions from southeast Mexico and the Caribbean; the West Mexico group, representing both lowlands and highlands; and the Mexican Highland group, encompassing most of Matsuoka et al.'s highland Mexican accessions (5) as well as accessions from highland Guatemala. These clusters also confirm the presence of US-derived varieties in South America (20); we excluded these accessions from further analysis.

In the joint PCA analysis of the three subspecies, the first PC (10.8% of variance) separates maize from its wild relatives and confirms the similarity between maize from the Mexican Highland group and *parviglumis* (Fig. 1*B*). The second PC (4.8% of variance) mainly separates the genetic groups of maize along a north–south axis, with the Northern United States and Andean Highlands at the extremes. The third PC (2.7% of variance) predominately reflects the difference between *parviglumis* and *mexicana*. The Mexican Highland cluster extends toward *mexicana* along both PC 1 and 3, suggesting that the similarity of highland maize to *parviglumis* may reflect admixture with *mexicana*.

Admixture Analysis. Simulation of gene flow of *mexicana* into the Meso-American Lowland maize group suggests that 13% cumulative historical introgression is sufficient to explain observed differences between lowland and highland maize in terms of heterozygosity and differentiation from *parviglumis* (Fig. S3). Structure analysis (21) of all Mexican accessions lends support for this magnitude of introgression (Fig. 2). The three subspecies form clearly separated clusters, but evidence of admixture is evident in all three groups, and the two wild relatives show clear signs of bidirectional introgression at altitudes where their ranges overlap (Fig. 2). Highland maize shows strong signs of *mexicana* introgression, with 20% admixture observed in the Mexican Highland cluster, but below 1,500 m *mexicana* introgression drops to less than 1%. Introgression from *parviglumis* into maize is much lower overall, reaching its highest average value (3%) in the lowland West Mexico group.

Drift Analysis. Because introgression from *mexicana* may affect ancestry inference based on genetic distance from *parviglumis*, we took an approach that does not require reference to the wild relatives. Under models of historical range expansion, genetic differentiation increases away from the population of origin (22, 23), and estimates of drift from ancestral frequencies have been applied successfully to identify ancestral populations (24). We therefore applied the method of Nicholson et al. (25) to estimate simultaneously ancestral frequencies and F, a measure of genetic drift of away from these frequencies, for sets of predefined populations.

To illustrate the potential impact of *mexicana* introgression, we first performed a standard analysis that includes each maize population in turn in conjunction with the two wild relatives. Average drift away from the inferred common ancestor of maize, *parviglumis*, and *mexicana* is higher for maize ($F = 0.24$) than for *mexicana* ($F = 0.15$) or *parviglumis* ($F = 0.07$), probably due to changes in allele frequency following the domestication bottleneck. Because the inferred ancestral frequencies are closer to those of the wild relatives than to present-day maize, comparison with this ancestor is sensitive to introgression from these subspecies. It therefore is not surprising that estimates of F between individual maize populations and the common ancestor of all three taxa identify the Mexican Highland group as being most similar (Fig. 3*A*). This pattern is maintained in an analysis excluding *mexicana*, in which Mexican Highland maize is tied with the West Mexico group as the most ancestral population (Fig. 3*B*).

To mitigate the impact of introgression, we used a slightly modified approach that excludes both *parviglumis* and *mexicana* and calculates genetic drift with respect to ancestral frequencies inferred from domesticated maize alone. Because the genetic

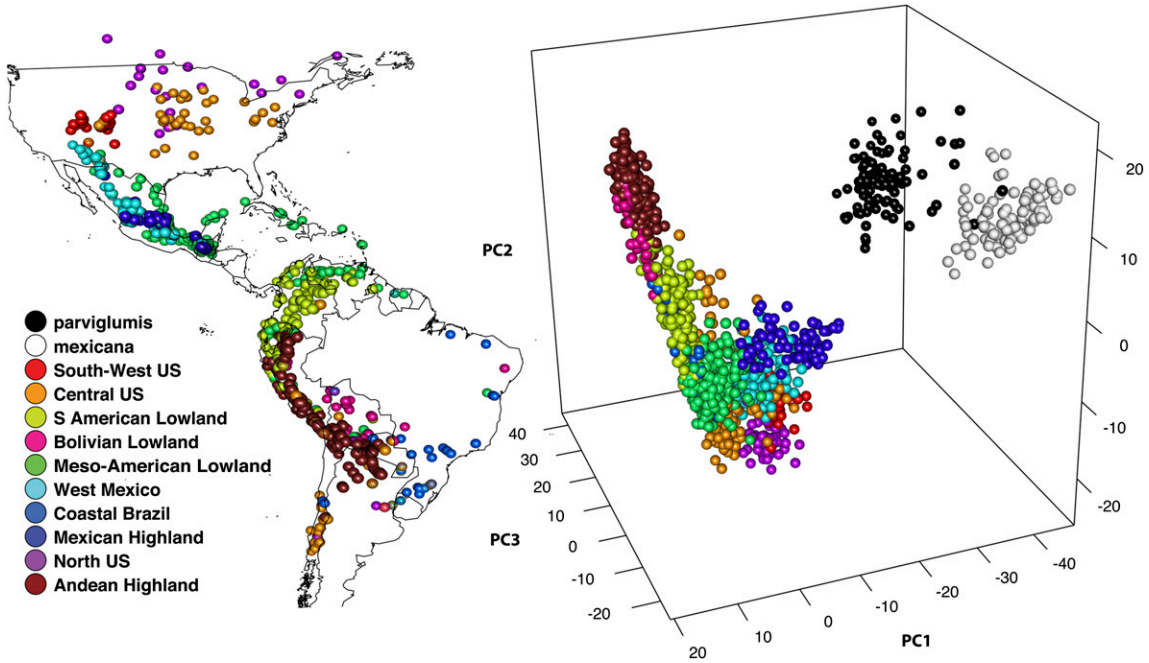

Fig. 1. (*A*) Map of sampled maize accessions colored by genetic group. (*B*) First three genetic PCs of all sampled accessions.

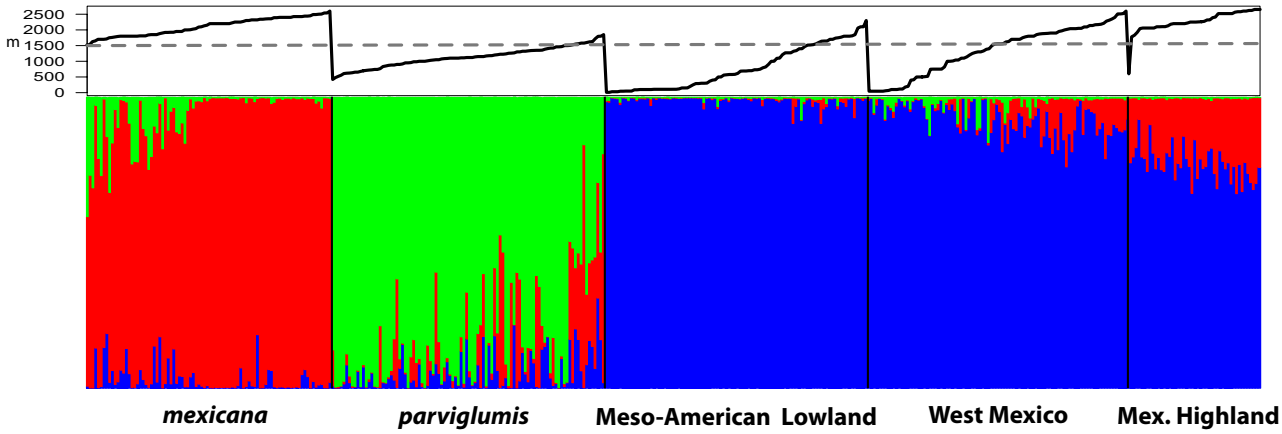

Fig. 2. (*Lower*) Bar plot of assignment values for the sample of Mexican accessions: *Mexicana* (red), *parviglumis* (green), and *mays* (blue). (*Upper*) The solid black line indicates the altitude for each sample. The dotted line marks the minimum altitude at which *mexicana* occurs.

similarity of some of our maize groups violates the assumption of independent drift, we infer ancestral frequencies by averaging over estimates obtained for pairs of diverged maize groups and calculate drift of individual populations with respect to these frequencies. In contrast to previous results, this comparison identifies the West Mexico group as being most similar to the common domesticated ancestor, followed by the Mexican Highland and Meso-American Lowland groups (Fig. 3*C*). Moreover, splitting the West Mexico group into highland (>2,000 m) and lowland (<1,500 m) components reveals that the lowland West Mexico group is most similar to the inferred ancestral maize. Direct comparison of genetic drift among the lowland West Mexico, Mexican Highland, and each of the remaining eight clusters shows further that the lowland West Mexico group is significantly closer than the Mexican Highland group to the inferred ancestor of each triplet (Fig. S4). These results strongly suggest that maize from the western lowlands of Mexico is genetically most similar to the common ancestor of maize and is more closely related to other extant populations than is maize from the highlands of central Mexico.

The ancestral position of the lowland West Mexico group is confirmed in a spatially explicit analysis of current allele frequencies in modern landraces, in which we mapped the moment estimator of F with respect to inferred ancestral allele frequencies. Mapping against allele frequencies observed in *parviglumis* (Fig. 4*A*) recapitulates earlier genetic results identifying highland maize as most similar to its wild ancestor (5). Points in the lower 0.05 quantile of F cluster in the highlands, with a mean altitude of 1,745 m. In contrast, mapping F with respect to inferred ancestral allele frequencies (Fig. 4*B*) identifies the lowest 0.05 quantile of F values in the lowlands of western Mexico, including the Balsas region and the region south of the Mexican highlands, resulting in an average altitude of 1,268 m; this analysis also clearly estimates higher values of F for maize in the Mexican highlands, particularly in areas of high inferred introgression from *mexicana* (Fig S5).

Discussion

Resolving the origins and spread of domesticated crops is a fascinating and challenging endeavor that requires the integration of botanical, archeological, and genetic evidence (26, 27, 28). Maize provides an exceptional opportunity for studying the processes of domestication and subsequent diffusion because of the wealth of existing archaeobotanical data, germplasm accessions, and molecular markers. The contradiction between evidence supporting the earliest cultivation in the lowlands and the genetically ancestral position of Mexican Highland maize is therefore of particular interest. The disagreement is important, because the adaptive differences between highland and lowland maize are profound (14, 29). In other crops, uncertainty about

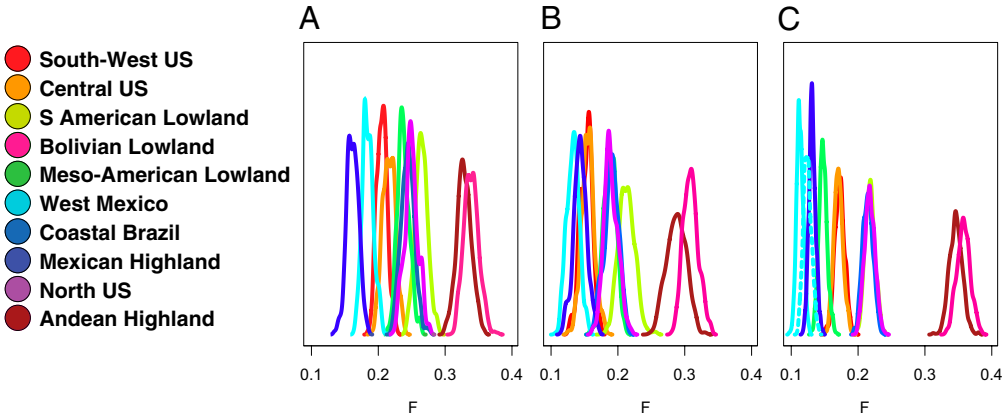

Fig. 3. Posterior densities of the genetic drift parameter F for 10 genetic groups with respect to (*A*) *mexicana* and (*B*) *parviglumis*. Only lowland accessions of the West Mexico group (light blue) were included. (*C*) Drift of all 10 genetic groups with respect to inferred ancestral frequencies. Light blue represents West Mexico; dotted line indicates the division between lowlands (<1,500 m, solid line) and highlands (>2,000 m).

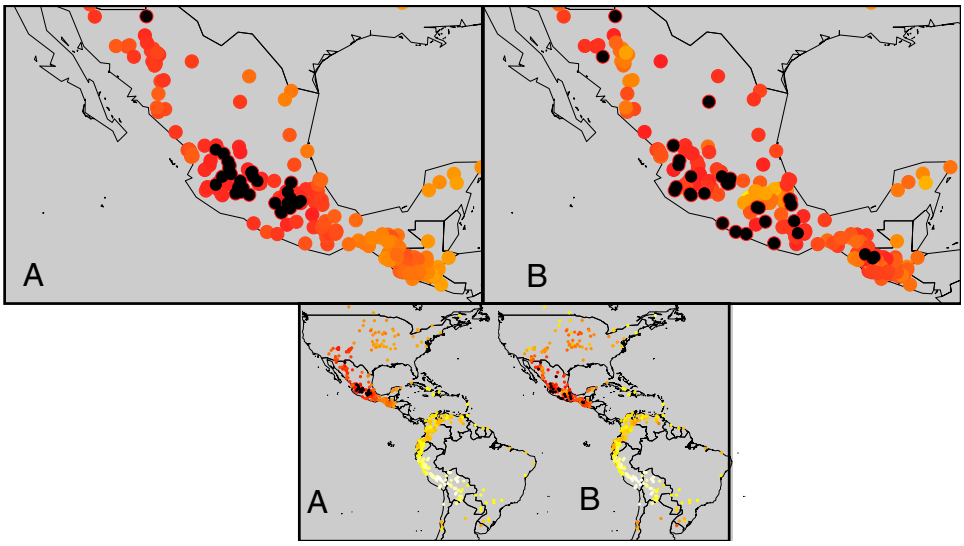

Fig. 4. Heat maps showing the amount of drift, *F*, away from (*A*) observed *parviglumis* allelic frequencies and (*B*) mean estimated ancestral frequencies. Each point is based on spatial estimation of current allele frequencies. The colors of the dots range from red for low values to white for high values of *F*. Black dots mark the lower 0.05 quantile. Upper panels *A* and *B* show enlarged sections of the lower panels *A* and *B*, respectively.

the geography of crop origins has been resolved by locating the most likely wild ancestor (24, 30–33). In the case of maize, however, the distribution of the wild ancestor does not coincide with the distribution of the cultivars most genetically similar to it.

Our results present a resolution of this paradox. We show that previous genetic evidence for an apparent highland origin of modern maize is best explained by gene flow from *mexicana* and demonstrate that admixture with a related nonancestral wild relative can interfere with analyses based on straightforward comparisons with the known ancestor. We provide a solution to the problem of admixture by using estimated ancestral frequencies without relying on the wild ancestor for inference. By taking this approach, we find support for the ancestral position of lowland maize from western Mexico, a result that is consistent with archaeobotanical data and genetic analysis of maize domestication. Our study demonstrates that reevaluating genetic evidence with improved sampling and methods can help reconcile results from the multidisciplinary efforts that are crucial to our understanding of crop origins.

Methods

Samples and Genotyping. We genotyped a single plant from each of 1,127 accessions of maize landraces (20), 100 accessions of *parviglumis*, and 96 accessions of *mexicana* (34) (Fig. S1). The maize sample contained all 351 named landraces that are known in the Americas (35), representing 689 unique geographic locations ranging from southern Chile to Canada and from the Andean mountains to the Caribbean islands. Passport data for the plants are available in Dataset S1. Genotyping methods closely follow those in refs. 36 and 19. SNPs were scored in 547 genes, the majority of which were evaluated in a diversity panel (36, 37) consisting of temperate and tropical inbred lines as well as *parviglumis* individuals. The diverse nature of this panel, combined with the robustness of the drift estimate to ascertainment (25), makes it unlikely that our results are affected by ascertainment bias. Loci with more than 15% missing data or an inbreeding coefficient F_{IS} >0.9 were removed, leaving 964 SNPs. Ten maize and two *parviglumis* individuals with >7.5% missing data were removed.

Analysis of Patterns of Differentiation. Geographic structure was evaluated by PCA analysis on the normalized matrix of SNP genotypes using a Tracy–Widom distribution to determine the number of significant PCs (17). PCs were tested for spatial autocorrelation using Moran's I. Individuals were assigned to 10 discrete groups by Ward clustering, using the Euclidean distances between individuals calculated on the nine standardized PCs showing the highest spatial autocorrelation (19).

Analysis of Admixture. We used the program Structure, version 2.3.2 (21, 38), to estimate admixture between maize and its wild relatives. Analysis was restricted to the 241 Mexican maize accessions in addition to *parviglumis* (n = 98) and *mexicana* (n = 96). We used the admixture model with k = 3, correlated allele frequencies, 100,000 burn-in iterations, and 1,000,000 Markov chain Monte Carlo steps. Use of correlated frequencies has been shown to improve estimates of admixture (38), and the accuracy of the method was evaluated by simulations of maize allele frequencies under varying levels of admixture from *mexicana*.

Drift and Ancestral Frequency Estimation. Estimates of the drift parameter *F* and ancestral allele frequencies were obtained using the Bayesian algorithm proposed by Nicholson (25) as implemented in the popdiv function in the R package popgen (39). With the assumption that drift is independent in predefined populations, the method jointly estimates the ancestral allele frequency π_i at each locus and F_j for each population such that $p_{ij} \sim$ Normal $[\pi_i, F_j\pi_i(1-\pi_i)]$, where p_{ij} is the observed allele frequency at locus i in population j. We analyzed the genetic groups identified by PCA-based Ward clustering separately for this purpose. In our modified approach, frequencies for the domesticated ancestor of all maize groups were calculated as the geometric mean over 10 ancestral frequency estimates obtained for 10 population pairs, containing each population in turn and a corresponding reference population that was chosen to maximize the value of $-\ln(1-\theta)$ (40), where θ is an estimator of F_{ST} (41). For each genetic group, we then calculated *F* with respect to this inferred ancestral frequency using a modification of the popdiv function. Direct comparison of *F* values for two candidate populations used sequential analyses of population triplets with the remaining eight maize populations, estimating ancestral frequencies separately for each triplet.

Estimation of Geographic Frequency. Maize allele frequencies across the sampled range were estimated with the method described by Wasser et al. (42) and implemented in the program SCAT. This method uses a Bayesian smoothing algorithm that infers gene frequencies underlying individual geographically mapped genotypes. To improve frequency estimation and mixing, the area containing the 689 registered sampling locations was covered with a grid of 471 interpolation points to achieve a maximum distance of 1.5° between locations. We then analyzed the geographic distribution of Nicholson's moment estimator of *F* (25) based on the allele frequency surface obtained by SCAT but including only points representative of actual sampling localities.

ACKNOWLEDGMENTS. This paper benefited from valuable discussion with G. Coop and M. B. Hufford and from the comments of two anonymous reviewers. This work was supported by National Science Foundation Grant DBI0820619 (to J.D.) and by Grant 2009-01864 from the USDA National Institute of Food and Agriculture (to J.R.-I).

1. Beadle GW (1939) Teosinte and the origin of maize. *J Hered* 30:245–247.
2. Mangelsdorf PC, Macneish RS, Galinat WC (1964) Domestication of corn. *Science* 143: 538–545.
3. Burger JC, Chapman MA, Burke JM (2008) Molecular insights into the evolution of crop plants. *Am J Bot* 95:113–122.
4. Doebley JF, Goodman MM, Stuber CW (1984) Isoenzymatic variation in *Zea* (Graminae). *Syst Bot* 9:203–218.
5. Matsuoka Y, et al. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci USA* 99:6080–6084.
6. Hastorf CA (2009) Rio Balsas most likely region for maize domestication. *Proc Natl Acad Sci USA* 106:4957–4958.
7. Piperno DR, Ranere AJ, Holst I, Iriarte J, Dickau R (2009) Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proc Natl Acad Sci USA* 106:5019–5024.
8. Pohl ME, Piperno DR, Pope KO, Jones JG (2007) Microfossil evidence for pre-Columbian maize dispersals in the neotropics from San Andres, Tabasco, Mexico. *Proc Natl Acad Sci USA* 104:6870–6875.
9. Dickau R, Ranere AJ, Cooke RG (2007) Starch grain evidence for the preceramic dispersals of maize and root crops into tropical dry and humid forests of Panama. *Proc Natl Acad Sci USA* 104:3651–3656.
10. Benz BF (2001) Archaeological evidence of teosinte domestication from Guilá Naquitz, Oaxaca. *Proc Natl Acad Sci USA* 98:2104–2106.
11. Piperno DR (2003) A few kernels short of a cob: On the Staller and Thompson late entry scenario for the introduction of maize into northern South America. *J Archaeol Sci* 30:831–836.
12. Metcalfe SE (2006) Late Quaternary environments of the northern deserts and central transvolcanic belt of Mexico. *Ann Mo Bot Gard* 93:258–273.
13. Piperno DR, et al. (2007) Late Pleistocene and Holocene environmental history of the Iguala Valley, central Balsas watershed of Mexico. *Proc Natl Acad Sci USA* 104: 11874–11881.
14. Eagles HA, Lothrop JE (1994) Highland maize from central Mexico: Its origin, characteristics, and use in breeding programs. *Crop Sci* 34:11–19.
15. Ellstrand NC, Garner LC, Hegde S, Guadagnuolo R, Blancas L (2007) Spontaneous hybridization between maize and teosinte. *J Hered* 98:183–187.
16. Wilkes HG (1972) Maize and its wild relatives. *Science* 177:1071–1077.
17. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genet* 2:e190.
18. Novembre J, Stephens M (2008) Interpreting principal component analyses of spatial population genetic variation. *Nat Genet* 40:646–649.
19. Van Heerwaarden J, et al. (2010) Fine scale genetic structure in the wild ancestor of maize (*Zea mays* ssp. *parviglumis*). *Mol Ecol* 19:1162–1173.
20. Vigouroux Y, et al. (2008) Population structure and genetic diversity of new world maize races assessed by DNA microsatellites. *Am J Bot* 95:1240–1253.
21. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
22. Austerlitz F, Jung-Muller B, Godelle B, Gouyon PH (1997) Evolution of coalescence times, genetic diversity and structure during colonization. *Theor Popul Biol* 51: 148–164.
23. Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annu Rev Ecol Evol Syst* 40:481–501.
24. Harter AV, et al. (2004) Origin of extant domesticated sunflowers in eastern North America. *Nature* 430:201–205.
25. Nicholson G, Smith AV, Donnelly P (2002) Assessing population differentiation and isolation from single-nucleotide polymorphism data. *J R Stat Soc Series B Stat Methodol* 64:695–715.
26. Smith BD (2001) Documenting plant domestication: The consilience of biological and archaeological approaches. *Proc Natl Acad Sci USA* 98:1324–1326.
27. Zeder MA, Emshwiller E, Smith BD, Bradley DG (2006) Documenting domestication: The intersection of genetics and archaeology. *Trends Genet* 22:139–155.
28. van Etten J, Hijmans RJ (2010) A geospatial modelling approach integrating archaeobotany and genetics to trace the origin and dispersal of domesticated plants. *PLoS One* 5:e12060.
29. Jiang C, et al. (1999) Genetic analysis of adaptation differences between highland and lowland tropical maize using molecular markers. *Theor Appl Genet* 99:1106–1119.
30. Gepts P, Osborn TC, Rashka K, Bliss FA (1986) Phaseolin-protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*): Evidence for multiple centers of domestication. *Econ Bot* 40:451–468.
31. Heun M, et al. (1997) Site of einkorn wheat domestication identified by DNA fingerprinting. *Science* 278:1312–1314.
32. Morrell PL, Clegg MT (2007) Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. *Proc Natl Acad Sci USA* 104: 3289–3294.
33. Olsen KM, Schaal BA (1999) Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proc Natl Acad Sci USA* 96:5586–5591.
34. Fukunaga K, et al. (2005) Genetic diversity and population structure of teosinte. *Genetics* 169:2241–2254.
35. Goodman M, Brown WL (1988) *Corn and Corn Improvement*, eds Sprague GF, Dudley JW (American Society of Agronomy, Madison, WI), 3rd Ed, pp 33–79.
36. Weber A, et al. (2007) Major regulatory genes in maize contribute to standing variation in teosinte (*Zea mays* ssp. *parviglumis*). *Genetics* 177:2349–2359.
37. Wright SI, et al. (2005) The effects of artificial selection on the maize genome. *Science* 308:1310–1314.
38. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.
39. R Development Core Team (2009) R: A Language and Environment for Statistical Computing. Available at www.R-project.org. Accessed December 19, 2010.
40. Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient: Basis for a short-term genetic distance. *Genetics* 105:767–779.
41. Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370.
42. Wasser SK, et al. (2004) Assigning African elephant DNA to geographic region of origin: Applications to the ivory trade. *Proc Natl Acad Sci USA* 101:14847–14852.