Transoceanic drift and the domestication of African bottle gourds in the Americas

Logan Kistler[a,1], Álvaro Montenegro[b], Bruce D. Smith[c], John A. Gifford[d], Richard E. Green[e], Lee A. Newsom[a,f], and Beth Shapiro[g]

[a]Department of Anthropology and [f]Institutes of Energy and the Environment, Pennsylvania State University, University Park, PA 16802; [b]Department of Geography, Ohio State University, Columbus, OH 43210; [c]Program in Human Ecology and Archaeobiology, Department of Anthropology, Smithsonian Institution, National Museum of Natural History, Washington, DC 20560; [d]Division of Marine Affairs and Policy, Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, FL 33149; and [g]Department of Ecology and Evolutionary Biology and [e]Baskin School of Engineering, University of California, Santa Cruz, CA 95064

Edited by Gayle J. Fritz, Washington University in St. Louis, St. Louis, MO, and accepted by the Editorial Board January 10, 2014 (received for review October 3, 2013)

Bottle gourd (*Lagenaria siceraria*) was one of the first domesticated plants, and the only one with a global distribution during pre-Columbian times. Although native to Africa, bottle gourd was in use by humans in east Asia, possibly as early as 11,000 y ago (BP) and in the Americas by 10,000 BP. Despite its utilitarian importance to diverse human populations, it remains unresolved how the bottle gourd came to be so widely distributed, and in particular how and when it arrived in the New World. A previous study using ancient DNA concluded that Paleoindians transported already domesticated gourds to the Americas from Asia when colonizing the New World [Erickson et al. (2005) *Proc Natl Acad Sci USA* 102 (51):18315–18320]. However, this scenario requires the propagation of tropical-adapted bottle gourds across the Arctic. Here, we isolate 86,000 base pairs of plastid DNA from a geographically broad sample of archaeological and living bottle gourds. In contrast to the earlier results, we find that all pre-Columbian bottle gourds are most closely related to African gourds, not Asian gourds. Ocean-current drift modeling shows that wild African gourds could have simply floated across the Atlantic during the Late Pleistocene. Once they arrived in the New World, naturalized gourd populations likely became established in the Neotropics via dispersal by megafaunal mammals. These wild populations were domesticated in several distinct New World locales, most likely near established centers of food crop domestication.

long-distance dispersal | New World domestication | archaeogenomics

In independent centers of plant domestication worldwide, distinct suites of food crops tend to emerge from native flora under human selection. An exception to this is the bottle gourd (*Lagenaria siceraria*, Cucurbitaceae), which is native to Africa, but was used by diverse human cultures not only in Africa, but also across Eurasia, the Pacific Islands, and the New World during pre-Columbian times (1–4). Although bottle gourd fruits are edible, they are used by humans mostly for other purposes, including as lightweight, durable containers, fishnet floats, and musical instruments (5). This variety of utilitarian applications likely explains why bottle gourds are so globally pervasive.

In the New World, bottle gourds appear in archaeological contexts as early as 10,000 BP (6) (Table 1), and become increasingly ubiquitous and widespread during the latter half of the Holocene (9). Bottle gourds were long proposed to have arrived in the Americas via long-range dispersal on ocean currents (10–13). However, an analysis of DNA from living and archaeological gourds suggested that the bottle gourd may have been transported into the New World by the first colonizing humans (6). In this scenario, the bottle gourd, like the dog (14), crossed the Bering Land Bridge with colonizing humans already in its domestic form, making the bottle gourd one of the earliest domesticated species (1, 6).

Two factors suggest that bottle gourd colonization of the Americas via the Bering Land Bridge is unlikely, however. First, bottle gourds thrive in tropical and subtropical habitats. Based on the physiological requirements of diverse modern cultivars (15), the growing season in Late Pleistocene Beringia would simply have been too cold and too short for bottle gourds to propagate and survive. Second, no archaeological or ethnographic evidence is known that supports the use of bottle gourds by humans in either Siberia or Alaska. In arctic regions, natural containers tend to be derived from animal products—hides, for example—rather than from plants (e.g., ref. 16). Given this lack of supporting evidence, the small amount of genetic data used to confirm this mode of colonization into the Americas deserves additional scrutiny. Furthermore, to explain why only pre-Columbian gourds appeared genetically Asian, authors of the previous study (6) suggested that a continent-wide replacement of New World gourd lineages by introduced varieties took place following European arrival. With no obvious explanation or mechanism for such a sweeping displacement of native varieties, however, this theory also warrants reconsideration.

We therefore returned to the previously studied archaeological gourds and used a capture-enrichment approach (17) to sequence and assemble the complete, 86,000 base pair large single-copy (LSC) region of the maternally inherited, nonrecombining plastid genome. In addition to two of the archaeological specimens previously analyzed (6), we included gourds from seven newly sampled New World archaeological assemblages and 36

Significance

Bottle gourd, one of the most cross-culturally ubiquitous crops, had a pan-tropical distribution by the beginning of the Holocene. Our findings overturn a major component of the current model for bottle gourd's early global dispersal, specifically regarding how it entered the Americas. Our findings also indicate that the domestication process itself took place in a diffuse pattern throughout the bottle gourd's New World range, explaining early and nearly contemporaneous use of bottle gourds in North, Central, and South America. Bottle gourd's weedy growth habit and the diffuse domestication pattern also suggest that early cultivation were probably not restricted to known centers of domestication. It is likely, however, that domesticated phenotypes emerged in these centers alongside food crops.

Author contributions: L.K., Á.M., B.D.S., R.E.G., L.A.N., and B.S. designed research; L.K. and Á.M. performed research; J.A.G. contributed new reagents/analytic tools; L.K., Á.M., B.D.S., R.E.G., L.A.N., and B.S. analyzed data; and L.K., Á.M., R.E.G., and B.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. G.J.F. is a guest editor invited by the Editorial Board.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. KJ399838–KJ399882) for assembled large single copy sequences, and the NCBI Sequence Read Archive (SRA Bioproject no. PRJNA236372) for short read sequence data.

[1]To whom correspondence should be addressed. E-mail: ljk187@psu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1318678111/-/DCSupplemental.

Table 1. Archaeological samples and associated AMS radiocarbon dates for gourd rind fragments used in this study

Site	Location	Accession no.	AMS Lab no.	C^{14} age (y)	Cal. age (2σ)
Loreto Cave	Baja California, Mexico	3-12793	Beta-316171	80 ± 30 BP	AD 1690–1925
Putnam Shelter	Washington County, AR	32-44-396c	Beta-316173	870 ± 30 BP	AD 1045–1244
Tularosa Cave	Catron County, NM	A246294	Beta-316172	1120 ± 30 BP	AD 824–994
Spring Branch Shelter	McCreary County, KY	aLsF2	Beta-316174	1910 ± 30 BP	AD 21–210
El Gigante	La Paz, Honduras	18-13b.3	Beta-316169	2110 ± 30 BP	203–46 BC
Alred Shelter	Benton County, AR	32-4-1176	Beta-316170	3850 ± 30 BP	2459–2206 BC
Quebrada Jaguay	Arequipa, Peru	S1-U4-PA-N1f	Beta-134112	7650 ± 50 BP	6594–6431 BC
Guila Naquitz	Oaxaca, Mexico	E10-B2	Beta-97237	7940 ± 60 BP	7043–6679 BC
Little Salt Spring	Sarasota County, FL	1408551A01	Beta-261466	8890 ± 50 BP	8241–7832 BC

Sample ages were calibrated using Oxcal 4.2 (7) assuming the IntCal09 calibration curve (8).

modern landraces and wild gourds from the Americas, Europe, Asia, Africa, and the South Pacific, representing all major geographic populations (Fig. 1 and Tables S1 and S2).

Results

Phylogenetic Analysis. Phylogenetic analysis of these sequences shows that all New World archaeological gourds fall within the diversity of their African counterparts (Fig. 2). This genetic relationship indicates that pre-Columbian bottle gourds were not derived from Eurasian gourd lineages, and were therefore not brought into North America from Asia via the Bering Land Bridge. Instead, the African (*L. siceraria* ssp. *siceraria*) and Eurasian (*L siceraria* ssp. *asiatica*) lineages (18) are distinct and strongly supported evolutionary lineages, and Africa is the clear source region of the bottle gourds that populated the Americas.

The previous conclusion that New World gourds were more closely related to Asian than to African gourds was reached by typing three variable sites within the plastid genome, two 5-bp insertion/deletions (indels) and one single nucleotide variant (SNV), that were thought to reflect fixed differences between Asian and African gourds (6). Our results highlight the risk of basing conclusions on very small genetic datasets. To resolve the discrepancy between our results and the previous findings, we used the same target-enrichment approach to capture these three variable sites from our larger dataset. Unfortunately, we were unable to capture one of the indels from any of our archaeological samples. We hypothesize that this inability was because of the large evolutionary distance separating the bottle gourd and the cucumber plastid genome, which we used to generate our probes (see *SI Methods* for discussion of additional methods). At the SNV, we captured data from four archaeological gourds, all of which we identified as carrying the African variant (Fig. S1). Although we were not able to type the SNV for either the Guila Naquitz or the Quebrada Jaguay samples, the previous study indicated that these carried the Asian variant. Because this finding indicates both variants are present in New World gourds, we conclude that this marker is not ancestry informative.

Interestingly, all of the eight archaeological gourds that we were able to type for the third marker carried the African variant (indel absent). These included the Guila Naquitz and Quebrada Jaguay specimens, both of which were reported previously to carry the Asian variant (indel present). However, further analysis of the previously published data revealed that, in the region surrounding the indel, the published archaeological gourd sequences were highly divergent both from each other and from the modern gourd sequences (6) (Fig. S1). As horizontal gene transfer from the chloroplast to the mitochondrial genome is a common phenomenon in the Cucurbitaceae (20), these data may be explained if a homologous region of the mitochondrial genome that closely matched the target in the plastid genome had been inadvertently amplified from the archaeological gourds. To test this theory, we compared this region from the bottle gourd plastid with several Cucurbitaceae plastid sequences and with the *Cucurbita pepo* mitochondrial genome. We identified a homologous mtDNA region in *C. pepo* that forms an outgroup to all plastid sequences, providing evidence for an ancient transfer event (*SI Methods* and Fig. S1). We conclude, therefore, that this marker is also not ancestry informative.

In contrast, our complete LSC dataset robustly supports an African ancestry for ancient New World gourd lineages. Assuming an evolutionary rate of 1×10^{-9} substitutions per site per year (21, 22), the African and Eurasian lineages share a most recent common ancestor (MRCA) 105–181 kya (Fig. 2, node A), and all New World gourds share a MRCA with African gourds at 60–103 kya (Fig. 2, node B), significantly predating the earliest possible domestication dates. Interestingly, archaeological gourds do not cluster together within the African clade. That is, there is no

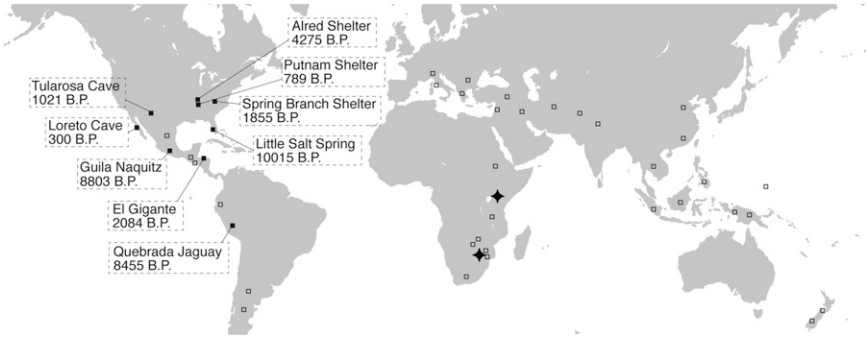

Fig. 1. Sample map showing modern domestic gourds (□), modern wild gourds (★), and archaeological gourd rind samples (■) used here. Dates reported with archaeological specimens give the weighted mean of the calibrated age invoking the IntCal.09 calibration curve (8) in Oxcal 4.2 (7). See Table 1 for complete details of archaeological samples, and Table S1 for modern sample information.

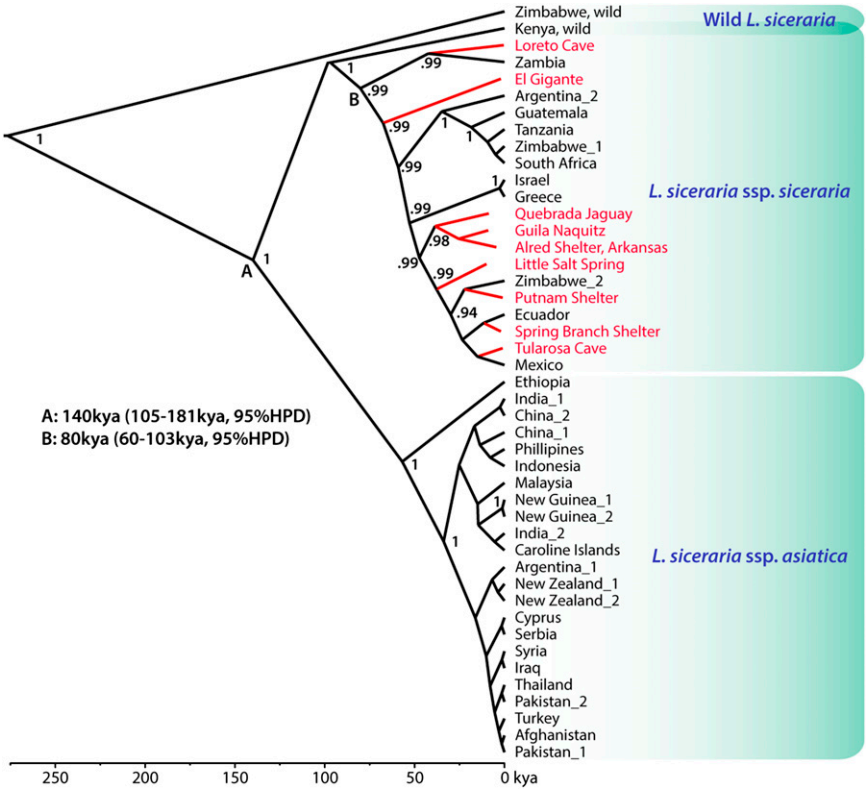

Fig. 2. Maximum clade credibility tree of modern and ancient bottle gourd LSC plastid genome data, showing Bayesian posterior probability at nodes where Bayesian posterior probability ≥0.9. Ancient branches and samples are shown in red, and the scale bar assumes an evolutionary rate of 1.0×10^{-9} substitutions site^{-1} per year^{-1}. The two subspecies, ssp. *siceraria* and ssp. *asiatica* [after Heiser (18)], are outlined, and a wild gourd from Zimbabwe [described by Decker-Walters et al. (25)] and one of only two known wild populations) forms an outgroup to all others, indicating some of the intraspecific diversity lost in *L. siceraria* during recent population declines. The Argentinian specimen within the *asiatica* lineage, from Heiser's (18) collections (Table S1), may represent a historic introduction to South America. Alternatively, it may be of particular interest regarding possible prehistoric contact, material culture exchange, and domesticated germplasm transmission between Polynesia and South America (15, 19). The *asiatica* lineage is subtended by a domestic Ethiopian landrace, whereas a wild gourd from Kenya falls at the base of the *siceraria* group, suggesting that the Horn of Africa might have been an important ancestral center of *Lagenaria* diversity, and a source region for *asiatica* gourds dispersing from Africa north and east into Eurasia, as discussed elsewhere (15).

clear founder effect within New World gourd plastid genomes. Thus, the mechanism that dispersed gourds from Africa to the Americas likely brought multiple, genetically diverse wild gourds across the Atlantic Ocean.

Oceanic Drift Modeling. To test the plausibility of oceanic dispersal on a scale sufficient to explain this level of diversity, we conducted a series of oceanic drift simulations based on wind values from the National Center for Environmental Protection/National Center for Atmospheric Research, and surface current data from the Estimating the Circulation and Climate of the Ocean consortium (23, 24). Both datasets constitute the best fit of numerical model output to empirical observations, and are used to calibrate a model that simulates drifting object displacement over time based on currents and winds. We varied model parameters—drifter sensitivity to wind speed, deflection angle between drifter and wind direction, and the intensity of wind-induced currents—to encompass the range of plausible variation in Late Pleistocene circulation patterns (*SI Methods* and Table S3). Wild bottle gourds are extremely rare, and little is known about their natural variation in either rind thickness or water-tight durability (25). However, seeds contained within the fruits of domesticated bottle gourds are known to remain completely viable after nearly a year floating in seawater (11). We therefore constrained the duration of individual drift events to 1 y, so that all transoceanic crossings would be completed within a time-frame that is compatible with seed viability.

We find strong support in our simulations for trans-Atlantic crossings in under a year, the majority of which occurred between 20° S and the equator and lasted an average of approximately 9 mo, but as little as 100 d (Fig. 3 and Table S4). For gourds leaving Africa within these latitudes, every combination of model parameters led to some successful crossings. In addition, regardless of the latitude of their departure from Africa, more than 80% of drifting gourds that arrived in the New World landed within this latitude zone. High crossing success was also observed for drifters originating between 10° N and 20° N, and this zone had the second highest proportion of successful landings (16%). Drifters originating between the equator and 10° N, however, only rarely arrived in the New World, and drifters that departed from latitudes poleward of ±20° almost never successfully crossed the Atlantic in our model. Any African gourd that reached the ocean within the tropical zone—with the help of inland river systems, for example—had a reasonable chance of making a successful trans-Atlantic crossing in sufficient time to maintain the capacity for germination upon landfall.

Discussion

It is most likely that the first bottle gourd populations in the Americas were established in the Neotropics. Like similar large, robust fruits, bottle gourds are adapted for dispersal by both water and large mammals, the latter of which may have provided important dispersal vectors within the New World (11, 21, 26–28). Wild *Cucurbita* seeds have been found in Late Pleistocene

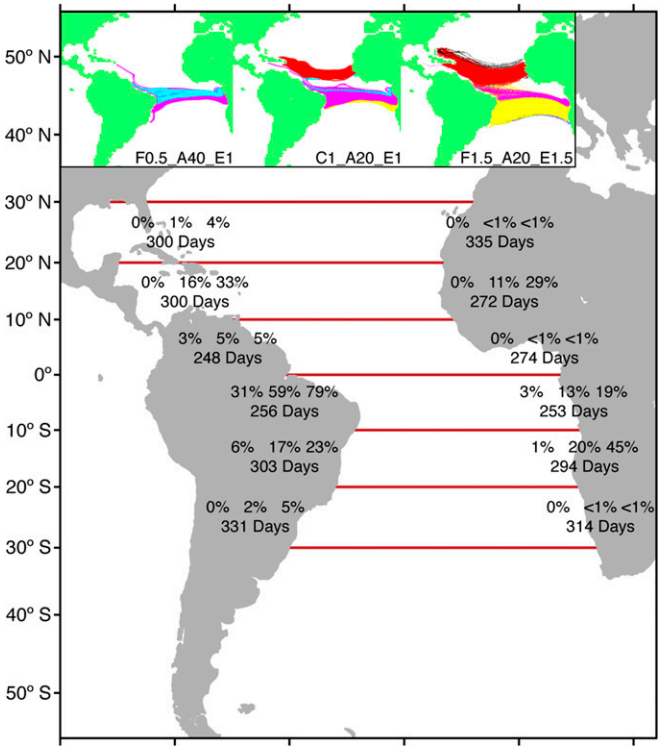

Fig. 3. Results of oceanic drift simulation experiments summarized in 10° latitude bins (Tables S3 and S4). African coastal bins give the minimum, average, and maximum drifter crossing success under all configurations of model parameters, and the average crossing times of successful drifters. New World figures indicate the minimum, average, and maximum percentage of all successful drifters arriving in a given bin, and the average crossing times of local arrivals. The *Inset* shows successful drifter paths over the course of three selected experiments (Table S3), with colors reflecting different departure latitude bands: yellow, 20°–10° S; magenta, 10°–0° S; cyan, 0°–10° N; red, 10°–20° N; black, 20°–30° N.

mastodon dung in Florida (29), revealing megafaunal dispersal of very similar fruits. Today, no wild bottle gourd populations survive in the Neotropics, and wild gourds are near extinction in Africa (5, 25). This result may reflect ecological changes since the Late Pleistocene, including the widespread disappearance of large mammal dispersers. In the absence of a natural vector for dispersal, gourds under cultivation would have had a significant reproductive advantage over their wild relatives, so that the only New World gourds likely to survive into the Holocene were the domesticates. By 10,000 y ago, humans in the Americas had begun to use bottle gourds both for food and in a variety of utilitarian capacities (Table 1) (13, 30, 31). As humans encountered free-living gourds, a level of cultivation or wild population management likely ensued, followed by selection for fruits better suited to the specific needs of local populations, and ultimately, the emergence of morphologically distinct, domesticated lineages. Although our data do not address where in the Americas these domesticated phenotypes arose, established centers of food crop domestication are strong candidates. South, Central, and eastern North America each independently gave rise to diverse suites of food crops, and early bottle gourds appear more-or-less concurrently with domestic food crops in all three of these localities.

Rather than a single domestication event followed by extensive cultural diffusion and varietal diversification, our results indicate that bottle gourd domestication was most likely a diffuse process of human intervention, taking place at multiple times across a geographically and culturally diverse landscape. Although we explore this process only in the New World, the phylogenetic intermingling of deeply divergent domestic African and New World lineages suggests a pattern of diffuse domestication in Africa similar to what we propose for the Americas. That is, the model of gourd domestication that we infer is probably not restricted to the New World, and higher-resolution analyses could help parse the details of bottle gourds' complex global movements and human interaction.

Methods

We extracted DNA from modern seed and leaf samples using Qiagen's Plant DNeasy Mini Kit (Qiagen). For ancient gourd rinds, we ground tissue using a ball mill or pellet pestle, and extracted DNA using a PTB-based extraction protocol (32) (*SI Methods*). We confirmed DNA presence using LS_INDEL1 and LS_SNP primers (6), and prepared barcoded Illumina sequencing libraries after Meyer and Kircher (33). We used the MycroArray MySelect system for in-solution RNA hybridization capture (17) of the LSC based on the cucumber (*Cucumis sativus*) plastid genome (34). We pooled and sequenced libraries in parallel on an Illumina HiSEq. 2000 at the University of California at Berkeley using 150-nt paired-end reads, and we demultiplexed reads, merged forward and reverse reads, and performed read quality control as described by Kircher (35). We used MIA (github.com/udo-stenzel/mapping-iterative-assembler) to construct a 99.8% complete LSC reference sequence from sample Ls6A based on the cucumber plastid genome, and used Burrows-Wheeler Aligner and Samtools to assemble the remaining reads for all other samples. We reconstructed a dated phylogeny (Fig. 2) using BEAST 1.7.4 (36), combining three independent Markov-chain Monte Carlo simulations, and assuming the Bayesian skyride coalescent model (37), the HKY + G nucleotide substitution model, and 1.0×10^{-9} substitutions per site per year (21, 22). We conducted oceanic-drift simulations as developed by Montenegro et al. (38, 39), integrating simulated general circulation model estimates of ocean current movement with empirical data on drifter behavior. Although observational and model-based studies indicate that large-scale wind and surface current systems in the Atlantic have remained relatively unchanged from the Late Pleistocene to the present (40–44), we varied all model inputs through parameter space to conservatively encompass reasonable behavior of drifting objects under Late Pleistocene conditions.

Short read sequence data are curated in the National Center for Biotechnology Information Sequence Read Archive (SRA Bioproject no. PRJNA236372) and assembled LSC sequences used in analysis are annotated and available in GenBank (accession nos. KJ399838–KJ399882).

ACKNOWLEDGMENTS. Modern samples were provided by B.D.S., Deena Decker-Walters, Mike Burtenshaw, the US Department of Agriculture-Agricultureal Research Service National Plant Germplasm System, and the Indiana University Herbarium. Access to ancient specimens was granted by the Phoebe A. Hearst Museum of Anthropology at University of California at Berkeley, the National Museum of Natural History at the Smithsonian Institution, the University Museum at the University of Arkansas, and the Kentucky Archaeological Survey. The El Gigante specimen was provided by Ken Hirth, and the Spring Branch Shelter specimen was provided by Jack Rossen. This work was supported in part by Pennsylvania State University College of Liberal Arts and Huck Institute of Life Sciences Grant NSF ARC-09090456, start-up funds from the University of California at Cruz (to B.S.), and the Universidade Estadual Paulista (Á.M.).

1. Richardson JBI (1972) The pre-Columbian distribution of the bottle gourd (*Lagenaria siceraria*): A re-evaluation. *Econ Bot* 26(3):265–273.
2. Decker-Walters DS, Staub J, Lopez-Sese A, Nakata E (2001) Diversity in landraces and cultivars of bottle gourd (*Lagenaria siceraria*; Cucurbitaceae) as assessed by random amplified polymorphic DNA. *Genet Resour Crop Evol* 48(4):369–380.
3. Heiser CB (1989) *Foraging and Farming: The Evolution of Plant Exploitation*, eds Harris DR, Hillman GC (Unwin Hyman, London), pp 471–480.
4. Matsui A, Kanehara M (2006) The question of prehistoric plant husbandry during the Jomon Period in Japan. *World Archaeol* 38(2):259–273.
5. Heiser CB (1979) *The Gourd Book* (Univ of Oklahoma Press, Norman, OK).
6. Erickson DL, Smith BD, Clarke AC, Sandweiss DH, Tuross N (2005) An Asian origin for a 10,000-year-old domesticated plant in the Americas. *Proc Natl Acad Sci USA* 102(51):18315–18320.
7. Bronk Ramsey C (2009) Bayesian analysis of radiocarbon dates. *Radiocarbon* 51(1): 337–360.
8. Reimer PJ, et al. (2009) IntCal09 and Marine09 radiocarbon age calibration curves, 0–50,000 years cal BP. *Radiocarbon* 51(4):1111–1150.
9. Doran GH (2002) *Windover: Multidisciplinary Investigations of an Early Archaic Florida Cemetery*, ed Doran GH (Univ Press of Florida, Gainesville), pp 1–38.

10. Camp WH (1954) A possible source for American pre-Columbian gourds. *Am J Bot* 41(9):700–701.
11. Whitaker TW, Carter GF (1954) Oceanic drift of gourds—Experimental observations. *Am J Bot* 41(9):697–700.
12. Heiser CB (1985) Some botanical considerations of the early domesticated plants north of Mexico. *Prehistoric Food Production in North America*, Anthropological Papers, ed Ford RI (Univ Of Michigan Museum of Anthropology, Ann Arbor, MI), No 75, pp 57–72.
13. Cutler HC, Whitaker TW (1961) History and distribution of the cultivated Cucurbits in the Americas. *Am Antiq* 26:469–485.
14. Leonard JA, et al. (2002) Ancient DNA evidence for Old World origin of New World dogs. *Science* 298(5598):1613–1616.
15. Clarke AC (2009) Origins and Dispersal of the Sweet Potato and Bottle Gourd in Oceania: Implications for Prehistoric Human Mobility. PhD dissertation (Massey Univ, Palmerston North, New Zealand).
16. VanStone J (1972) The first Peary collection of polar Eskimo material culture. *Fieldiana Anthropol* 63(2):31–80.
17. Gnirke A, et al. (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* 27(2):182–189.
18. Heiser CB (1973) *Tropical Forest Ecosystems in Africa and South America: A Comparative Review*, eds Meggers BJ, Ayensu ES, Duckworth WD (Smithsonian Institution Press, Washington, DC), pp 121–128.
19. Jones T, Storey A, Matisso-Smith E, Ramirez-Aliaga J, eds (2011) *Polynesians in America: Pre-Columbian Contacts with the New World* (Altamira, New York).
20. Alverson AJ, et al. (2010) Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol Biol Evol* 27(6):1436–1448.
21. Schaefer H, Heibl C, Renner SS (2009) Gourds afloat: A dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proc Biol Sci* 276(1658):843–851.
22. Gaut BS (1998) *Evolutionary Biology*, ed Hecht MK (Plenum, New York), Vol 30, pp 93–120.
23. Stammer D, et al. (2002) The Global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model. *J Geophys Res* 107(C9): 221–244.
24. Kalnay E, et al. (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77(3):437–471.
25. Decker-Walters DS, Wilkins-Ellert M, Chung SMM, Staub JEE (2004) Discovery and genetic assessment of wild bottle gourd [*Lagenaria siceraria* (Mol.) Standley; Cucurbitaceae] from Zimbabwe. *Econ Bot* 58(4):501–508.
26. Barlow C (2002) *The Ghosts of Evolution: Nonsensical Fruit, Missing Partners, and Other Ecological Anachronisms* (Basic Books, New York).
27. Janzen DH, Martin PS (1982) Neotropical anachronisms: The fruits the gomphotheres ate. *Science* 215(4528):19–27.
28. Gunn CR, Dennis JV (1976) *World Guide to Tropical Drift Seeds and Fruits* (Krieger Publishing, Malabar, FL).
29. Newsom LA, Mihlbachler MC (2006) *First Floridians and Last Mastodons: The Page-Ladson Site in the Aucilla River*, ed Webb SD (Springer, Dordrecht), pp 263–331.
30. Doran GH, Dickel DN, Newsom LA (1990) A 7,290-year-old bottle gourd from the Windover Site, Florida. *Am Antiq* 55(2):354–360.
31. Yarnell RA (1974) *Archeology of the Mammoth Cave Area*, ed Watson PJ (Academic, New York), pp 113–122.
32. Kistler L (2012) Ancient DNA extraction from plants. *Ancient DNA: Methods and Protocols*, Methods in Molecular Biology, eds Shapiro B, Hofreiter M (Humana, New York), Vol 840, pp 71–79.
33. Meyer M, Kircher M (2010) *Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing* Cold Spring Harbor Protocols (Cold Spring Harbor Lab Press, Cold Sporing Harbor, NY).
34. Plader W, Yukawa Y, Sugiura M, Malepszy S (2007) The complete structure of the cucumber (*Cucumis sativus* L.) chloroplast genome: Its composition and comparative analysis. *Cell Mol Biol Lett* 12(4):584–594.
35. Kircher M (2012) Analysis of high-throughput ancient DNA sequencing data. *Ancient DNA: Methods and Protocols*, Methods in Molecular Biology, eds Shapiro B, Hofreiter M (Humana, New York), Vol 840, pp 197–228.
36. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29(8):1969–1973.
37. Minin VN, Bloomquist EW, Suchard MA (2008) Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol Biol Evol* 25(7):1459–1471.
38. Montenegro Á, Hetherington R, Eby M, Weaver AJ (2006) Modelling pre-historic transoceanic crossings into the Americas. *Quat Sci Rev* 25(11-12):1323–1338.
39. Montenegro Á, Avis C, Weaver A (2008) Modeling the prehistoric arrival of the sweet potato in Polynesia. *J Archaeol Sci* 35(2):355–367.
40. Shi N, Schneider R, Beug H, Dupont LM (2001) Southeast trade wind variations during the last 135 kyr: Evidence from pollen spectra in eastern south atlantic sediments. *Earth Planet Sci Lett* 187(3-4):311–321.
41. Schneider RR, Müller PJ, Ruhland G (1995) Late Quaternary surface circulation in the east equatorial South Atlantic: Evidence from alkenone sea surface temperatures. *Paleoceanography* 10(2):197–219.
42. Montoya M, Von Storch H, Crowley TJ (2000) Climate simulation for 125 kyr BP with a coupled ocean-atmosphere general circulation model. *J Clim* 13(6):1057–1072.
43. Kitoh A, Murakami S, Koide H (2001) A simulation of the last glacial maximum with a coupled atmosphere-ocean GCM. *Geophys Res Lett* 28(11):2221–2224.
44. Hewitt CD, Stouffer RJ, Broccoli AJ, Mitchell JFB, Valdes PJ (2003) The effect of ocean dynamics in a coupled GCM simulation of the Last Glacial Maximum. *Clim Dyn* 20(2-3):203–218.

Supporting Information

Kistler et al. 10.1073/pnas.1318678111

SI Methods

DNA Extraction, Library Preparation, Target Capture, Sequencing, and Assembly. For modern seed samples, we extracted DNA using Qiagen's Plant DNeasy Mini Kit (Qiagen) according to the manufacturer's protocol, but with a modified tissue disruption step as follows: We used sterile razor blades to split the seeds and excise the embryos, discarding the outer tissue. We then divided the embryos, reserving half for subsequent extraction, and grinding the other half using sterile pellet pestles and a small amount of sterile sand while submerged in lysis buffer. DNA extraction success was confirmed via PCR amplification of the rpl32-trnL intergenic spacer using published primers (1). Sample 35B was a small piece of herbarium-preserved desiccated leaf tissue. We ground 50 mg of this tissue in lysis buffer and proceeded with the standard DNA extraction as above.

The archaeological samples were either desiccated ($n = 8$) or waterlogged ($n = 1$, Little Salt Spring) rind fragments that ranged in age from historic times to approximately 10,000 BP (Table 1). We carried out DNA isolation using the protocol developed previously for ancient bottle gourd rind (2) with several small modifications (3). First, we decreased the concentration of proteinase K to 0.4 mg/mL. Second, we ground rind fragments at room temperature using a Micro-Dismembrator ball mill (Sartorius), and used between 50 and 200 mg of tissue per sample. Finally, we increased the volume of extraction buffer as necessary to ensure that powdered tissue remained in a fluid suspension during agitation. For the single waterlogged specimen from Florida, we gently ground the rind tissue by hand in a 2-mL tube using a sterile pellet pestle instead of using the ball mill. We eluted DNA in 100 μL Qiagen Buffer AE with 0.05% Tween. We verified the presence of amplifiable DNA in the resulting extracts via PCR amplification and gel electrophoresis using the LS_INDEL1 and LS_SNP primers designed for plastid DNA (cpDNA) (2).

We prepared barcoded Illumina sequencing libraries using the protocol described by Meyer and Kircher (4). For modern samples, we sheared DNA with three cycles of sonication for 7 min in a Biorupter sonicator (Diagenode) using the highest energy setting available. Because of their already fragmented nature, we did not subject the archaeological samples to sonication. We used 50-μL template DNA in modern and ancient library preparations. To ensure ancient DNA short-molecule retention before adaptor ligation in the ancient libraries, we used a Qiagen MinElute PCR Purification Kits for cleanup after blunt-end repair, and SPRI bead cleanup as per the protocol for all steps after adaptor ligation.

To enrich the DNA libraries for the large single-copy (LSC) region of the plastid genome, we used in-solution RNA hybridization (5) provided by MycroArray. We designed bait libraries to target the entire plastid LSC region based on the published cucumber (*Cucumis sativus*) plastid genome (GenBank NC_007144.1) (6). We also used the published cucumber nuclear genome (7) to design bait molecules targeting roughly 900 kb of neutrally evolving nuclear regions across all seven cucumber chromosomes, and we enriched and sequenced plastid and nuclear regions simultaneously. Comparison between reads mapping to targeted versus untargeted components of the cucumber plastid genome indicates ~2,000% enrichment of targeted regions in both modern and ancient libraries. Nuclear enrichment was poor and sporadic, however, because of the evolutionary distance of the cucumber genome and the embryonic copy number discrepancy between organellar and nuclear DNA; consequently, we were unable to include analyses of nuclear DNA in this study.

We pooled and sequenced the libraries in parallel using paired-end 150-nt reads on an Illumina HiSEq. 2000 instrument at the University of California at Berkeley. We separated individual sample reads, merged forward and reverse reads, and performed read quality control as described by Kircher (8). A small proportion of reads failed to merge and were discarded. To develop a bottle gourd LSC reference sequence, we used MIA (github.com/udo-stenzel/mapping-iterative-assembler) to map reads from sample Ls6A to the cucumber LSC, iteratively correcting the reference based on the consensus sequence and remapping to convergence, and subsequently closed gaps when possible via manual read alignment. This process yielded a LSC reference sequence with only three small ambiguous regions of 10, 38, and 104 bp when aligned to the cucumber plastid genome, an estimated 99.8% complete LSC sequence. We used Burrows-Wheeler Aligner (9) to assemble the remaining sample reads to the new LSC reference sequence, and Samtools (10) to produce a consensus sequence for each LSC assembly. For ancient samples, we recovered between 65.6% and 99.58% complete LSC assemblies with nonredundant 2× coverage (Table S2).

Phylogenetic Analysis. We used Mafft (11) to align all LSC consensus sequences, and JModelTest (12, 13) to select the HKY + G nucleotide substitution model. Using BEAST 1.7.4 (14), we ran three independent Markov-chain Monte Carlo (MCMC) simulations assuming the Bayesian skyride coalescent model (15), sampling every 10,000 states until all parameters reached effective sample size of at least 200, and visually verified MCMC convergence using Tracer (http://beast.bio.ed.ac.uk/tracer), ~50,000,000 iterations per chain. We enforced a plastid substitution rate of 1.0×10^{-9} substitutions per site per year, based on empirical data within the Cucurbitaceae and other flowering plants (16, 17). We discarded the first 10% of samples as burn-in, combined the remaining posterior samples from the three chains, and summarized the maximum clade credibility tree using TreeAnnotator.

Ancient New World gourds carry considerable diversity, but they universally fall within the ssp. *siceraria* group with modern African and New World varieties with strong clade support. We artificially fixed ancient New World and modern Eurasian gourds together in assorted monophyletic configurations during analysis, and found no support for any recent affinity between the two groups, strongly suggesting a direct African origin for pre-Columbian bottle gourds in the New World. We estimate the most recent common ancestor of ancient (MRCA) New World gourds at 81 kya (95% highest posterior density 60–103 kya). Allowing this rate to vary substantially (e.g., threefold either direction) consistently places their most recent common ancestor before any accepted date estimates for the peopling of the Americas, or for gourd domestication in any locality.

Oceanic Drift Simulations. Our simulations were aimed at determining the feasibility of multiple transatlantic gourd crossings during the Late Pleistocene, the time-frame during which African and New World lineages diverged. We considered using wind and current estimates from paleoclimate simulations during that interval, but observational (18, 19) and modeling (20–22) studies indicate that large-scale wind and surface current systems in the Atlantic have not changed markedly since the Late Pleistocene. As such, we adopted data from present-day reconstructions,

providing higher spatiotemporal resolution than paleo-models by drawing on empirical observation of experimental drifters.

Surface currents and winds. Surface current data are provided by the Consortium for Estimating the Circulation and Climate of the Ocean (ECCO) funded by the National Oceanographic Partnership Program (23). We use monthly estimates from ECCO's version 3 with spatial resolution of 1° × 1° covering the period from 1993 to 2005. Surface values refer to the average over the top 5 m of the water column. The ECCO results can be understood as a numerical model constantly corrected to match observations. Before being used in the drifting model, surface current values are linearly interpolated to 5-d means. We used monthly estimates of winds at 10-m elevation generated by the National Center for Environmental Protection/National Center for Atmospheric Research (NCEP/NCAR) reanalysis (24). These are the winds adopted by the ECCO reanalysis, and the wind data were obtained at the ECCO data server with the same temporal coverage and spatial resolution as the surface current estimates. Like the currents, wind values are interpolated to 5-d means.

Drift model. Following the methodology described by Montenegro et al. (25, 26), we implemented a drift model that determines and records the displacement of individual drifters as they move under the influence of local winds and currents during a determined period. We operationally defined local winds and currents as the average value of these parameters in the 1° × 1° bins found within 0.25° of the drifter at a given time. Drifter location, along with local wind and current values, were updated every 5 d during the simulations.

Previous laboratory and in situ experiments have shown that winds tend to generate surface currents that move at 1–4% of the surface wind speed at an angle that varies from of 0° to 45° to the wind direction; the deflection occurs to the right in the northern hemisphere and to the left in the southern hemisphere (27). With the exception of regions of the western boundary and the equatorial current system, ECCO surface speeds exhibit values much smaller than 1% of the NCEP/NCAR surface winds, largely because of the ocean surface currents being averaged from the first 5 m of the water column.

The drift model compensates for this by adding to the original ECCO values a wind-induced surface speed, an approach similar to that of other ocean drift models (28). The fraction of the wind speed used to determine the wind-induced current as well and the drift angle are kept constant for each simulation. Experiments are conducted with wind-induced currents ranging from 0.5 to 1.5% of the wind speed and with drift angle varying from 0° to 40°. Because of the wind, objects drifting on the ocean can move up to twice the speed of surface currents (28). The drift model can account for this wind-induced drift speed, which is kept constant during each experiment.

Experiment description, preliminary tests, and results. Drifts began from the center of departure bins along the entire African Atlantic coast. These were comprised of data bins that had at least one side touching the coast. Drifters started once a month, every month, from January of 1993 to December of 2004, and were allowed to move under the influence of currents and winds for up to 1 y. Drifter location and velocity were updated every 5 d. If a drifter entered a coastal bin along the target areas (eastern coasts of North, South and Central America plus coasts of the Caribbean islands) the drift was considered a successful crossing. There were 196 departure bins resulting in a total of 28,224 simulated drifts for each experiment (196 drifts per 12 mo per 12 y).

Table S3 provides a list of the experiments conducted for this study and the mean simulated drifter speed for each. Experiments differed according to fraction of wind speed used to determine current velocities, angle of surface current in relationship to surface wind direction, and extra wind-induced movement on the drifter. As expected, results were sensitive to changes in surface current speed, with more successful crossings taking place under faster currents (Fig. 3 and Table S4). The number and pattern of successful crossings was also sensitive to the choice of drift angle, with crossings decreasing as angle increases.

With this in mind, we based our conclusions on the results of three experiments (Fig. 3 and Table S4): F0.5_A40_E1, a "pessimistic" scenario with low speed and extreme drift angle; F1_A20_1, a "standard" scenario, with intermediate speed and angle; and F1.5_A20_1.5, an "optimistic" scenario with higher speed and intermediate angle. Most estimates of average drifting speeds for objects with dimensions similar to those of the gourd fruit come from drift bottle experiments, and tend to range from 0.1 m/s to more than 1 m/s. When away from western boundary systems, many bottle-based speeds are centered around 0.15–0.17 m/s (29–32). After falling off ships in the North Pacific, drifting shoes and toys that made their way to the North American west coast moved with average speeds of 0.17 m/s and 0.12 m/s, respectively (28). These last two values are based on straight-line trajectories of thousands of kilometers, and probably underestimate the 5-d average speeds of these accidental drifters.

Although the average speed of all three selected simulations are within the spread of observations, the modeled speeds are closer to the lower bound of observations, and even our optimistic scenario has speeds near the center of the observed range. We believe this conservative approach is justified, given the inherent uncertainties involved in the modeling effort, and the fact that the primary goal of the simulations is to provide a stringent test of the feasibility of multiple drift crossings.

Most successful crossings originate from two segments of the African coast (Fig. 3 and Table S4). The first segment, between 20° S and the equator, show the highest probabilities of success, and is the only region with successful crossings for all experiments. The second segment is the region between 10° and 20° N. These areas are separated by a segment of the coast, between the equator and 10° N, from which almost no successful crossings originate. Although the probability of success is not as high in the northern departure band, crossings from this area tend to be slightly (7%) faster. Almost no crossings start poleward of ±20°. The vast majority of arrivals in the Americas occur between the latitudes of ±20°, with more than 80% of crossings finishing in the region between 20° S and the equator (Fig. 3 and Table S4). Modeling results show two distinct and separated possible source areas in Africa. The southern source is larger, and crossings originating from it have higher probability of crossing successfully. Successful crossings reach the Americas along a continuous band over much of the tropics, with landings most likely in the low southern latitudes.

We opted to simulate drift events originating only from the western African coast, not the eastern coast. The shortest viable route from eastern Africa into the Atlantic would be southward along the Agulhas Current, and drifters following this path would enter the Southeast Atlantic near the southwestern African coast (~35° S). This territory is already covered by our west coast simulations, and much less than 1% of successful drifters originating south of 20° S reached the Americas within a year (Table S4). Given the extra time required for displacement from the Indian Ocean into the Southeast Atlantic, we can predict that success rates of eastern African coast departures would be even smaller than that of drifters departing the southernmost latitudes of the west coast directly. Therefore, we do not regard this as a viable source area under our conservative assumptions.

Comparison with the Previous Study (2). Our results contradict the findings of a previous ancient DNA-based study, which concluded that New World gourds originated from domesticated Asian populations, most likely being introduced by colonizing Paleoindians (2). To address the discrepancy between the previous results and ours, we aligned the modern and ancient sequence data from the previous study (GenBank DQ300361–DQ300478)

to the data produced here. The previous study was based on three variable sites within intergenic spacer regions in the LSC, including one single nucleotide polymorphism, referred to as LS_SNP, which comprises a G–A transition, and two 5-bp insertion/deletions (LS_ Indel1 and LS_Indel2).

We were able to sequence LS_SNP, located in the psbM-trnY intergenic spacer region (2), for four of our nine ancient gourds. All four carry the African (G) variant (Fig. S1). Unfortunately, we were unable to retype the LS_SNP for either the Guila Naquitz or the Quebrada Jaguay samples, which the previous study indicated carried the Asian (A) haplotype. Assuming that the SNP was accurately resolved in the previous study, these results indicate that both variants are present in New World gourds, and that LS_SNP is therefore not an ancestry informative marker.

LS_Indel2 is reported by the previous study to occur in the trnS-trnG intergenic spacer. However, because the cucumber and bottle gourd sequences are highly divergent in this region, our capture probes—designed based on the cucumber plastid genome—were unable to recover LS_Indel2 reliably in any ancient sample, including Guila Naquitz and Quebrada Jaguay specimens from which this indel was reported previously (2). We are therefore unable to either confirm or reject the reliability of this region as an ancestry informative site. However, given our observations at the other two sites used by the previous analysis, we suggest that this may also have been a genomic region of high diversity within ancient New World gourds, with little power to ascertain ancestry in the absence of other markers.

Finally, we were able to type LS_Indel1, found in the psbM-trnC intergenic spacer region, in eight of nine ancient samples. In each of these samples we observed the African variant (no indel), including Guila Naquitz and Quebrada Jaguay, the two ancient gourds that were typed as having the indel in the previous study (Fig. S1). In addition to the presence of the indel, the available data from the previous study (GenBank DQ300391–DQ300400) show substantial sequence variation within this region in ancient gourds, and no variation in modern gourds. We find no variation within this region for any gourds, including the ancient gourds.

As a working explanation for the disagreement between our data and the previously published data, we hypothesize that ancient horizontal gene transfer between the plastid and mitochondrial genomes might have created a homologous region of mtDNA with a similar sequence to the target region in the plastid genome. Genomic duplication between organelles is well documented in the Cucurbitaceae (33), and has the potential to complicate organellar genome assembly and analysis. If the marker at this particular site was developed based on the amplification of longer fragments via PCR using primers placed in conserved regions of the plastid genome, moving these priming sites into the intergenic spacer region adjacent to the indel (to make the amplified fragments suitably short for use with ancient samples) may have led to the accidental amplification of a mitochondrial homolog instead of the target region of the chloroplast.

To test this hypothesis, we first confirmed that the "Asian" variant (indel present) is the ancestral character state in the plastid genome. We compared gourd data from our study and the previous study with published orthologous cpDNA sequences in four other Cucurbitaceous species: *Benincasa hispida* (GenBank DQ282074.1), *Citrullus colocynthis* (AY693733.1), *Praecitrullus fistulosus* (AY507989.1), and *Cucumis sativus* (NC_007144). All four species carry the indel, which is also shared by wild bottle gourd from Zimbabwe (our Ls6A), and the ssp. *asiatica* lineage. The wild Kenya gourd (our Ls6C), along with modern ssp. *siceraria* gourds and ancient New World specimens in our study, carry the alternate allele (indel absent), indicating that the deletion at this site took place after the split between the *asiatica* and *siceraria* subspecies ~140 kya, but before the MRCA of ssp. *siceraria* ~100 kya (95% highest posterior density 181–74 kya) (Fig. 2). We then aligned the region of Ls_Indel1 from *Benincasa*, *Citrullus*, *Praecitrullus*, *Cucumis*, and *Lagenaria* plastids with the mitochondrial genome of *Cucurbita pepo* (NC_014050), and identified a homologous *Cucurbita* mtDNA region that shares between 77.3% and 80.1% pairwise sequence identity with the corresponding cpDNA from each taxon calculated over 181 aligned base pairs. The homologous mtDNA region carries the ancestral variant at LS_Indel1 (indel present) characteristic of ssp. *asiatica* bottle gourds and other cucurbits. Furthermore, in a neighbor-joining tree (Fig. S1) the *Cucurbita* mtDNA sequence forms an outgroup to the plastid sequences, suggesting that a horizontal transfer event took place before the common ancestor of all taxa present at approximately 30 ± 4 Mya (17). Although *Lagenaria* mtDNA sequence data are not available to directly confirm the persistence of the homologous region, this provides a viable explanation for the amplification and sequencing of the Asian-type LS_Indel1 allele in ancient gourds in the previous study, surrounded by unexpected levels of variation.

1. Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *Am J Bot* 94(3):275–288.
2. Erickson DL, Smith BD, Clarke AC, Sandweiss DH, Tuross N (2005) An Asian origin for a 10,000-year-old domesticated plant in the Americas. *Proc Natl Acad Sci USA* 102(51):18315–18320.
3. Kistler L (2012) Ancient DNA extraction from plants. *Ancient DNA: Methods and Protocols*, Methods in Molecular Biology, eds Shapiro B, Hofreiter M (Humana, New York), Vol 840, pp 71–79.
4. Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc*, 10.1101/pdb.prot5448.
5. Gnirke A, et al. (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* 27(2):182–189.
6. Plader W, Yukawa Y, Sugiura M, Malepszy S (2007) The complete structure of the cucumber (*Cucumis sativus* L.) chloroplast genome: Its composition and comparative analysis. *Cell Mol Biol Lett* 12(4):584–594.
7. Huang S, et al. (2009) The genome of the cucumber, *Cucumis sativus* L. *Nat Genet* 41(12):1275–1281.
8. Kircher M (2012) Analysis of high-throughput ancient DNA sequencing data. *Ancient DNA: Methods and Protocols*, Methods in Molecular Biology, eds Shapiro B, Hofreiter M (Humana, New York), Vol 840, pp 197–228.
9. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760.
10. Li H, et al.; 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16):2078–2079.
11. Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30 (14):3059–3066.
12. Posada D (2008) jModelTest: Phylogenetic model averaging. *Mol Biol Evol* 25(7): 1253–1256.
13. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52(5):696–704.
14. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29(8):1969–1973.
15. Minin VN, Bloomquist EW, Suchard MA (2008) Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol Biol Evol* 25(7):1459–1471.
16. Gaut BS (1998) *Evolutionary Biology*, ed Hecht MK (Plenum, New York), Vol 30, pp 93–120.
17. Schaefer H, Heibl C, Renner SS (2009) Gourds afloat: A dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proc Biol Sci* 276(1658):843–851.
18. Shi N, Schneider R, Beug H, Dupont LM (2001) Southeast trade wind variations during the last 135 kyr: Evidence from pollen spectra in eastern south atlantic sediments. *Earth Planet Sci Lett* 187(3-4):311–321.
19. Schneider RR, Müller PJ, Ruhland G (1995) Late Quaternary surface circulation in the east equatorial South Atlantic: Evidence from alkenone sea surface temperatures. *Paleoceanography* 10(2):197–219.
20. Montoya M, Von Storch H, Crowley TJ (2000) Climate simulation for 125 kyr BP with a coupled ocean-atmosphere general circulation model. *J Clim* 13(6):1057–1072.
21. Kitoh A, Murakami S, Koide H (2001) A simulation of the last glacial maximum with a coupled atmosphere-ocean GCM. *Geophys Res Lett* 28(11):2221–2224.
22. Hewitt CD, Stouffer RJ, Broccoli AJ, Mitchell JFB, Valdes PJ (2003) The effect of ocean dynamics in a coupled GCM simulation of the Last Glacial Maximum. *Clim Dyn* 20(2-3): 203–218.
23. Stammer D, et al. (2002) The Global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model. *J Geophys Res* 107(C9): 221–244.

24. Kalnay E, et al. (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77(3):437–471.
25. Montenegro Á, Hetherington R, Eby M, Weaver AJ (2006) Modelling pre-historic transoceanic crossings into the Americas. *Quat Sci Rev* 25(11-12):1323–1338.
26. Montenegro Á, Avis C, Weaver A (2008) Modeling the prehistoric arrival of the sweet potato in Polynesia. *J Archaeol Sci* 35(2):355–367.
27. Chang Y, Chen G, Tseng L, Centurioni L, Chu P (2012) Observed near-surface currents under high wind speeds. *J Geophys Res Oceans* 117(C11):C11026.
28. Ingraham WJ (1997) Getting to know oscurs, refm's ocean surface current simulator. *Alaska Fisheries Science Center Quarterly Report*, April–May:1–14.
29. Gast J (1966) A drift bottle study in northern california. *Limnol Oceanogr* 11(3): 415–417.
30. Luedemann E (1967) Preliminary results of drift-bottle releases and recoveries in the western tropical atlantic. *Boletim do Instituto Oceanografico da Universidade de Sao Paulo* 16(1):13–22.
31. Duncan C, Atwood D, Duncan J, Froelich P (1977) Drift bottle returns from the caribbean. *Bull Mar Sci* 27(3):580–586.
32. Ebbesmeyer C, Belkin I, Drost H, Zimmermann S, Carmack E (2011) Wall across the atlantic: Drift bottles released by students confirm that the gulf stream prevents subarctic surface drifters from escaping south. *Oceanography* 24(1):172–174.
33. Alverson AJ, et al. (2010) Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol Biol Evol* 27(6):1436–1448.

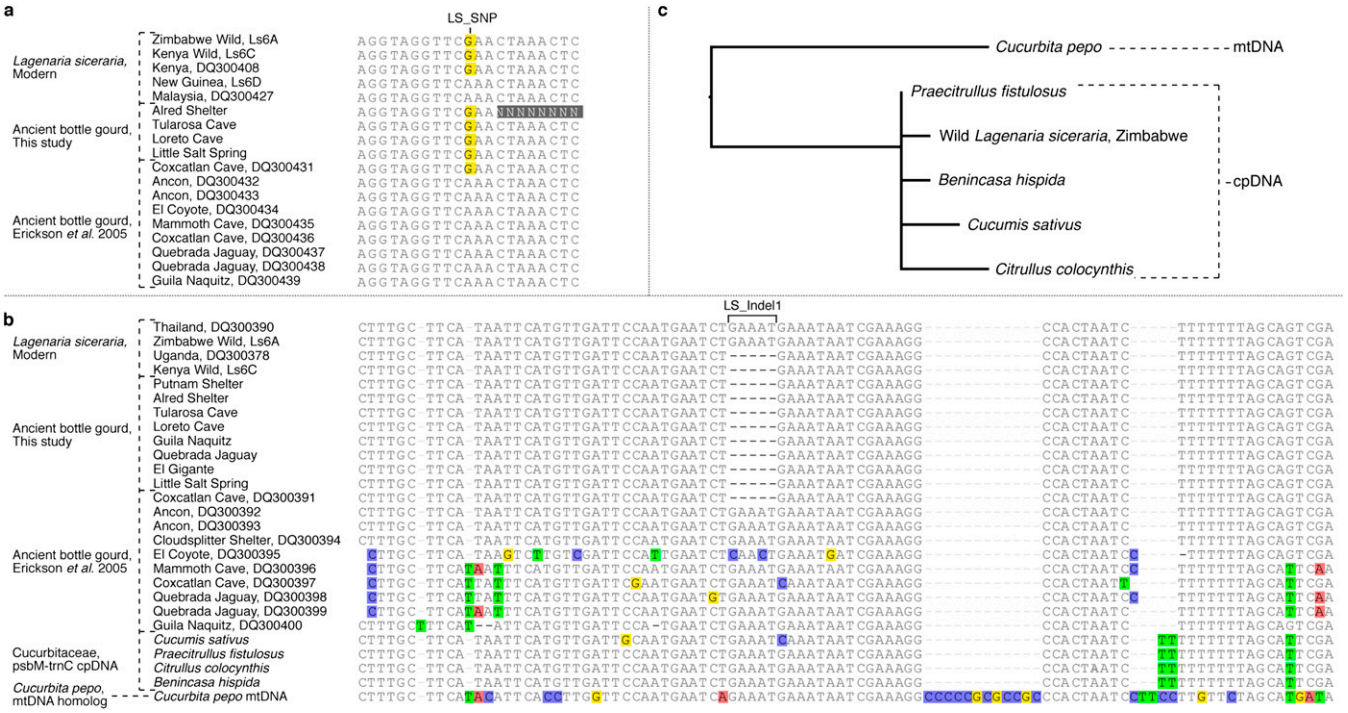

Fig. S1. Comparison between sequence data presented here and in a previous study by Erickson et al. (2). (*A*) Multiple alignment at LS_SNP in the psbM-trnY intergenic spacer region, including wild gourds from Zimbabwe and Kenya and domestic gourds from Kenya, New Guinea, and Malaysia, plus four archaeological samples typed here and nine archaeological samples typed previously. Both SNP variants (A and G) were present prehistorically in the Americas, indicating that LS_SNP(A) is not an ancestry informative marker. (*B*) Multiple alignments at LS_Indel1 in the psbM-trnC intergenic spacer region, including wild gourds from Zimbabwe and Kenya, domestic examples from Thailand and Uganda, eight ancient samples typed in this study, and 10 ancient samples typed previously. We also include four divergent species within the Cucurbitaceae, and a homologous region in the *C. pepo* mitochondrial genome. Note that the Ancon and Cloudsplitter gourd samples (from 2) appear to carry the ancestral allele (insertion present) but lack the sequence variation that surrounds the insertion in the other, previously typed archaeological specimens. If these sequence are from endogenous plastid DNA and not a mitochondrial copy (which may explain the more divergent sequences), we propose that these sequences may represent African lineages that diverged from extant members of ssp. *siceraria* before the deletion at this site. (*C*) Midpoint rooted neighbor-joining tree showing the LS_Indel1 mitochondrial homolog from *C. pepo* forming an outgroup to plastid-derived psbM-trnC spacer sequences from five diverse cucurbits, suggesting that a horizontal transfer of this region between organelles predates diversification of the plastid variant through speciation.

Table S1. Modern gourd samples used in this study, corresponding to Fig. 2

Lab No.	Location	Source	Catalog no.	GenBank accession no.
Ls3A	Argentina_2	USDA	PI 458736	KJ399849
Ls3B	Mexico	USDA	PI 438846	KJ399850
Ls3C	Syria	USDA	PI 181913	KJ399851
Ls3D	South Africa	USDA	PI 280633	KJ399852
Ls3F	Pakistan_2	USDA	PI 269508	KJ399853
Ls3G	Serbia/Montenegro	USDA	PI 368635	KJ399854
Ls4A	Turkey	USDA	PI 170463	KJ399855
Ls4C	India_2	USDA	PI 288497	KJ399856
Ls4D	Cyprus	USDA	PI 432342	KJ399857
Ls4E	Iraq	USDA	PI 435291	KJ399858
Ls4F	Indonesia	USDA	PI 470260	KJ399859
Ls4G	Zambia	USDA	PI 500836	KJ399860
Ls5A	Afghanistan	USDA	PI 256069	KJ399861
Ls5B	Pakistan_1	USDA	PI 269507	KJ399862
Ls5C	Ethiopia	USDA	PI 194994	KJ399863
Ls5D	Zimbabwe_2	USDA	PI 491266	KJ399864
Ls5E	Guatemala	USDA	PI 451857	KJ399865
Ls5F	India_1	USDA	PI 381829	KJ399866
Ls5G	Zimbabwe_1	USDA	PI 491324	KJ399867
Ls5H	China_1	USDA	PI 419089	KJ399868
Ls5J	Phillipines	USDA	PI 188809	KJ399869
Ls6A	Zimbabwe_Wild	DDW	TCN 1502	KJ399870
Ls6C	Kenya_Wild	DDW	MW 452	KJ399871
Ls6D	New Guinea_1	DDW	TCN 834–1-3	KJ399872
35B	Argentina_1	UI	35B	KJ399838
MB01	New Zealand_2	MB	MB01	KJ399873
MB02	New Zealand_1	MB	MB02	KJ399874
MB26	Ecuador	MB	MB26	KJ399877
MB29	Tanzania	MB	MB29	KJ399878
MB154	Caroline Islands	MB	MB154	KJ399875
MB191	New Guinea_2	MB	MB191	KJ399876
Ls10e	Malaysia	BDS	BDS 111	KJ399844
Ls10f	Thailand	BDS	BDS 188	KJ399845
Ls10g	China_2	USDA	PI 419215	KJ399846
Ls10h	Israel	USDA	PI 487482	KJ399847
Ls10i	Greece	USDA	PI 491252	KJ399848

BDS, B. Smith, National Museum of Natural History, Smithsonian Institution; DDW, D. Decker-Walters, the Cucurbit Network; IU, Indiana University Herbarium; MB, M. Burtenshaw, The Open Polytechnic of New Zealand; USDA, US Department of Agriculture, Agricultural Research Service, National Plant Germplasm System.

Table S2. Summary statistics of ancient gourd sequencing and assembly success

Sample	% of LSC covered (2×)	Unique mapped reads/ percentage of all reads (%)	Mean mapped read length (bp)	Mean read depth at represented sites
Putnam	72.49	19,666/0.03	63	19.9×
Alred	76.91	22,855/0.03	60	20.8×
Spring Branch	70.32	17,025/0.03	61	17.2×
Tularosa	99.58	150,335/0.12	135	236.4×
Loreto	96.01	99,739/0.10	100	120.1×
Guila Naquitz	65.55	13,759/0.03	61	14.8×
Quebrada Jaguay	73.09	20,237/0.04	62	19.8×
El Gigante	73.92	18,242/0.02	65	18.7×
Little Salt Spring	83.61	38,734/0.03	63	34.0×

Table S3. List of oceanic drift simulation experiments

Experiment ID	Wind speed (%)	Angle	Wind drift	Average speed (m/s)
F0.5_A0_E1	0.50	0	1	0.09
F0.5_A20_E1	0.50	20	1	0.09
F0.5_A40_E1	**0.50**	**40**	**1**	**0.09**
F1_A0_E1	1	0	1	0.11
C1_A20_E1	**1**	**20**	**1**	**0.11**
F1_A40_E1	1	40	1	0.11
F1.5_A0_E1.5	1.50	0	1.5	0.16
F1.5_A20_E1.5	**1.50**	**20**	**1.5**	**0.15**
F1.8_A0_E1.8	1.80	0	1.8	0.2
F2_A0_E1	2	0	1	0.15
F2_A0_E2	2	0	2	0.23
F2_A40_E2	2	40	2	0.17

"Wind speed" describes the fraction of the wind speed used to determine the wind-induced surface current. "Angle" describes the angle of deflection between surface current and wind, in degrees. In "Wind-induced drift," a value of 1 indicates that drifters move with the surface current speed, a value of 1.5 indicates that drifters move 50% faster than the surface current, a value of 1.8 indicates they move 80% faster, and a value of 2 indicates that drifters move twice the speed of the surface current. "Average speed" reports the averaged simulated drifter speed in meters per second. The conclusions discussed in the main text are based on results from the experiments in boldface, and correspond to Fig. 3.

Table S4. Compiled results of oceanic drift experiments (F0.5_A40_E1, C1_A20_E1, and F1.5_A20_E1.5) for 5° latitude bins

	Departures						Arrivals					
Latidude bins	%Suc	AveT	MinT	AveLat	MinLat	MaxLat	%Suc	AveT	MinT	AveLat	MinLat	MaxLat
35° N to 40° N	0	—	—	—	—	—	0	—	—	—	—	—
30° N to 35° N	<0.5	360	360	23	23	23	<0.5	360	360	30	30	30
25° N to 30° N	2	325	235	22	18	26	<0.5	325	235	26	25	29
20° N to 25° N	11	288	190	19	15	24	3	288	245	21	20	23
15° N to 20° N	20	299	140	16	10	22	7	299	223	16	15	18
10° N to 15° N	13	267	100	15	8	21	7	267	207	12	10	13
5° N to 10° N	<0.5	270	110	11	7	15	1	270	230	9	8	9
0° N to 5° N	<0.5	266	210	−1	−5	4	1	266	220	0	0	0
5° S to 0°	4	251	150	−3	−10	13	8	251	159	−4	−5	−2
10° S to 5° S	22	255	110	−5	−12	11	37	255	169	−9	−10	−7
15° S to 10°S	16	287	115	−7	−14	5	20	287	216	−13	−14	−11
20°S to 15° S	25	266	180	−10	−18	1	9	266	210	−18	−19	−16
25° S to 20° S	20	275	195	−14	−24	−4	6	275	212	−23	−25	−21
30° S to 25° S	2	338	250	−19	−22	−15	1	338	285	−26	−28	−26
35° S to 30° S	0	—	—	—	—	—	0	—	—	—	—	—
37° S to 35° S	0	—	—	—	—	—	0	—	—	—	—	—

AveLat, average New World arrival latitude for departing drifters (Departures), or average African source latitude of arriving drifters (Arrivals); AveT, average crossing time, in days, for all successful drifting departing/arriving in the bin; MinLat/MaxLat, minimum and maximum latitudes as described for AveLat; MinT, minimum crossing time for any individual experiment; %Suc, percentage of drifters making a successful crossing from the bin (Departures) or percentages of all successful drifters that arrived in the bin (Arrivals).