Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination

Caroline Roullier[a,b,1], Laure Benoit[b], Doyle B. McKey[b,c], and Vincent Lebot[a]

[a]Centre de Coopération Internationale en Recherche Agronomique pour le Développement-Systèmes Biologiques-, Unité Mixte de Recherche Amélioration Génétique et Adaptation des Plantes, 34398 Montpellier Cedex 5, France; [b]Centre National de la Recherche Scientifique, Centre d'Ecologie Fonctionnelle et Evolutive, 34293 Montpellier Cedex 5, France; and [c]Institut Universitaire de France and Université Montpellier II, 34095 Montpellier cedex 5, France

Edited by Kenneth M. Olsen, Washington University in St. Louis, St. Louis, MO, and accepted by the Editorial Board December 3, 2012 (received for review July 9, 2012)

The history of sweet potato in the Pacific has long been an enigma. Archaeological, linguistic, and ethnobotanical data suggest that prehistoric human-mediated dispersal events contributed to the distribution in Oceania of this American domesticate. According to the "tripartite hypothesis," sweet potato was introduced into Oceania from South America in pre-Columbian times and was then later newly introduced, and diffused widely across the Pacific, by Europeans via two historically documented routes from Mexico and the Caribbean. Although sweet potato is the most convincing example of putative pre-Columbian connections between human occupants of Polynesia and South America, the search for genetic evidence of pre-Columbian dispersal of sweet potato into Oceania has been inconclusive. Our study attempts to fill this gap. Using complementary sets of markers (chloroplast and nuclear microsatellites) and both modern and herbarium samples, we test the tripartite hypothesis. Our results provide strong support for prehistoric transfer(s) of sweet potato from South America (Peru-Ecuador region) into Polynesia. Our results also document a temporal shift in the pattern of distribution of genetic variation in sweet potato in Oceania. Later reintroductions, accompanied by recombination between distinct sweet potato gene pools, have reshuffled the crop's initial genetic base, obscuring primary patterns of diffusion and, at the same time, giving rise to an impressive number of local variants. Moreover, our study shows that phenotypes, names, and neutral genes do not necessarily share completely parallel evolutionary histories. Multidisciplinary approaches, thus, appear necessary for accurate reconstruction of the intertwined histories of plants and humans.

phylogeography | herbarium specimens | prehistory | early trans-Pacific travels

Efforts to understand human mobility and cultural evolution in the Pacific have encompassed archaeological, ethnographic, linguistic, and, more recently, genetic approaches (1, 2). Genetic analyses have focused not only on human populations (3, 4) but also on those of species closely associated with human movements (5–13), including commensal animals (*Rattus exulans*) (5, 6), domesticated plants [bottle gourd, *Lagenaria siceraria* (7, 8); banana, *Musa* spp. (9)] and animals [chicken, *Gallus gallus* (10, 11); pig, *Sus scrofa* (12)], and pathogens (*Helicobacter pylori*) (13). Studies of commensal species are particularly useful in contexts where mobility did not necessarily involve gene flow between human populations or where only limited archaeological data exist (14). Such studies have provided the best evidence thus far to address the controversial issue of contacts between Polynesia and the Americas in prehistoric times (8, 10, 14, 15). Sweet potato (*Ipomoea batatas* [L.] Lam.) in Oceania is undoubtedly the example most convincingly suggestive of such contacts (14).

Sweet potato originated in tropical America (16). When, from where, and how it subsequently reached Oceania have been the subjects of extensive debate. Its presence in precontact archaeological sites scattered throughout Polynesia has long been considered as direct evidence for prehistoric contact between Polynesia and America (17–21). Also, the lexical similarity between terms for sweet potato in Polynesian languages ("kuumala" and its derivatives) and the terms for this plant ("kumara," "cumar," or "cumal") found among Quechua speakers of northwestern South America supports the hypothesis that humans introduced sweet potato from South America to Polynesia (22), against the alternative hypothesis of natural long-distance dispersal of seeds (23). Finally, the tripartite hypothesis, first proposed by Barrau (24), developed by Yen (25), recently updated by Green (26) and reviewed by Clarke (27), posits both prehistoric and historic dispersal events, putting forward a tripartite origin to explain the distribution of the species (Fig. 1). In the first origin, the Polynesian sweet potato (Kumara lineage) was introduced by Polynesian voyagers who collected it somewhere from the western coast of South America, between 1000 and 1100 A.D. (26, 27). They may have rapidly diffused it throughout Polynesia, in already populated islands such as Hawaii, Easter Island, and some other islands of eastern Polynesia and then into New Zealand, around 1150–1250 A.D., with the original colonists. Other independent prehistoric introductions (from northern Colombia or even Central America) have also been hypothesized, but these lack support (23, 28). Also, the possibility of an early westward dispersal of sweet potato carried by Polynesians to Tonga, Samoa, and eastern Melanesia is suggested by early historical accounts (26, 29). The second and third origins of sweet potato in Oceania arose from European contact during the sixteenth century. Spanish "Manila-Acapulco" galleons introduced the Mesoamerican Camote to the Philippines around 1500 A.D. (Camote lineage), whereas Portuguese traders introduced to present-day eastern Indonesia the Batata from the Caribbean and Central America (Batata lineage). From these points, secondary dispersal events, mediated by local traders, European travelers, or both, may have distributed sweet potato in the western Pacific, probably early in the New Guinea highlands (around 1700), and much later (mid-19th century) into western and eastern Melanesia (Fig. 1). Two additional European introductions may also have contributed to early diffusion in the Pacific, one by Mendaña's voyage to the Marquesas and Solomon Islands in 1595 A.D. and another by Queiros in Espíritu Santo (part of present-day Vanuatu) in 1606 A.D. (30).

The tripartite hypothesis was principally based on archaeological, historical, and linguistic lines of evidence (25, 26). Regrettably,

Author contributions: V.L. designed research; C.R. and L.B. performed research; C.R. analyzed data; and C.R. and D.B.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. K.M.O. is a guest editor invited by the Editorial Board.

Data deposition: The microsatellite data have been deposited in the Dryad Digital Repository, www.datadryad.org (Data Package Identifier http://dx.doi.org/10.5061/dryad.77148).

See Commentary on page 1982.

[1]To whom correspondence should be addressed. E-mail: carolineroullier@hotmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1211049110/-/DCSupplemental.

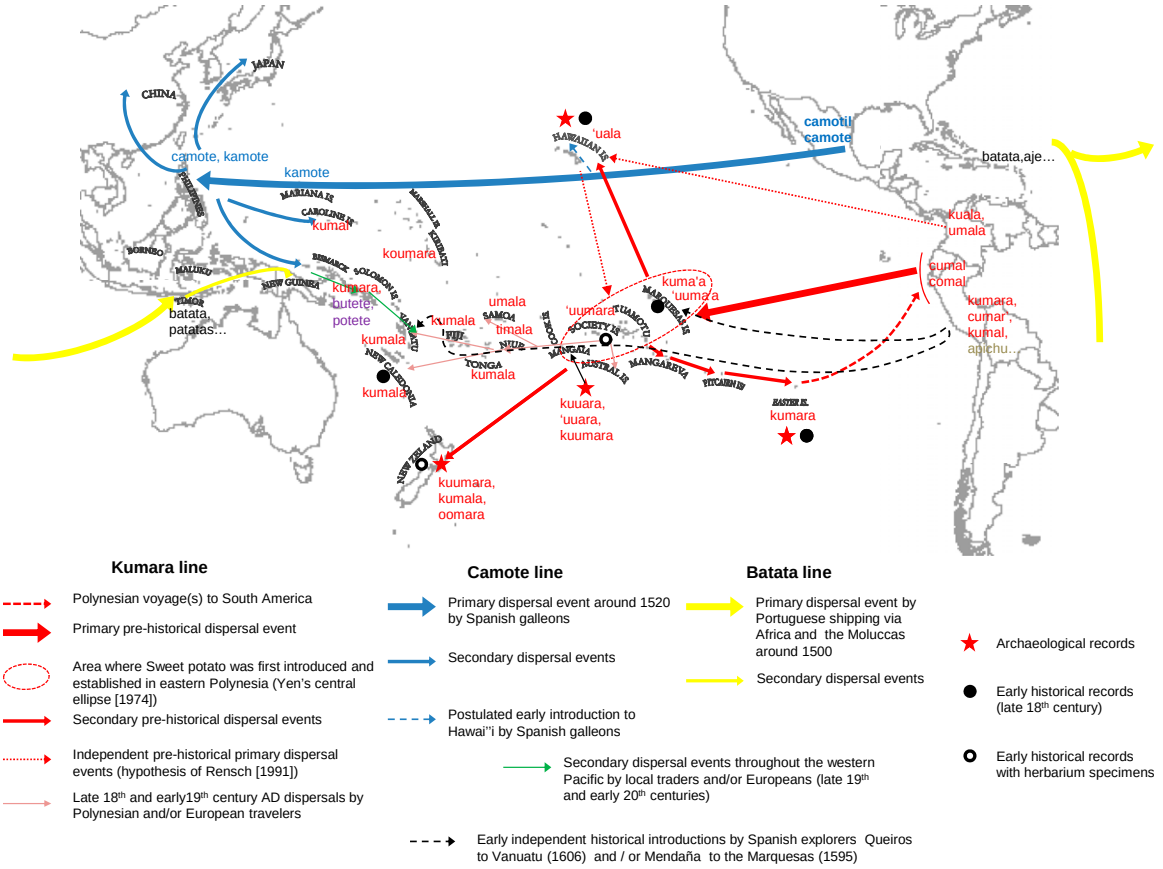

Fig. 1. Prehistoric and historical dispersal of sweet potato in Oceania, as postulated by the tripartite hypothesis. This map summarizes the tripartite hypothesis, as updated by Green (26) and reviewed by Clarke (27). Dispersal events and the terms commonly used to designate sweet potato in the different regions [compiled mostly from Yen (25)] are represented. Archaeological and early historical records providing strong evidence for the presence of sweet potato in Oceania in the prehistoric period are also indicated.

phylogeographic studies have not yet adequately tested it (27, 31–33). The genetic patterns expected under this hypothesis include a distinction between crops of the eastern and western Pacific, in which varieties from the western Pacific may be closer to those of Mesoamerica, whereas those from the eastern Pacific are closer to those of western South America. Previous attempts to elucidate origins of sweet potato in Oceania using molecular markers showed that varieties from the Pacific were more similar genetically to Mexican accessions than to those from Peru–Ecuador, supporting a Mesoamerican origin rather than early Polynesian transfers from South America (31–33). However, these studies suffered from a very restricted sample, including very few Polynesian varieties. More recently, Clarke (27), using Amplified Fragment Length Polymorphism (AFLP) markers, analyzed around 300 varieties from Oceania, including Polynesian varieties. However, that study did not use a representative sample of the tropical American source gene pools and, thus, failed to firmly trace the origin of extant Pacific lineages. Using both nuclear and chloroplast microsatellite markers, Roullier et al. (34) recently highlighted the existence of two geographically restricted gene pools in tropical America, corresponding to accessions from the Peru–Ecuador region of South America (hereafter called the Southern gene pool) and accessions from the Caribbean and Central America region (hereafter called the Northern gene pool). This set of complementary markers provides suitable tools to distinguish the Oceanian cultivars that originated from the Kumara line (derived from the Southern gene pool) from those that originated from the Batata and Camote lines (derived from the Northern gene pool).

An additional limitation to retracing the history of sweet potato in Oceania comes from the fact that contemporary patterns of its genetic variation in the region probably do not reflect initial patterns. From all putative entry points, 20th century intensification of human movements and exchanges across the Pacific probably extended the range of sweet potato in Oceania, as well as redistributing its genetic diversity (25, 35). Consequently, genetic signatures of initial introductions may be at least partially obscured by later introductions from a different genetic background. It is, thus, necessary to compare geographic patterns in genetic variation through time, by analyzing both modern and ancient samples (14, 36). Archaeological remains of sweet potato are scarce in the Pacific and are highly unlikely to contain preserved DNA. However, herbarium specimens, collected by early European explorers and naturalists and, later, by botanists and agronomists, are invaluable sources of material for tracing early movements of gene pools, complementing sampling of modern accessions, which have been more affected by the intensification of germplasm exchanges during the 20th century.

The present study attempts to fill the gap of genetic data in the reconstruction of sweet potato history in Oceania. Using a broad sampling of both modern and herbarium accessions from tropical America, Oceania, and South-East Asia, and nuclear and chloroplast markers of tropical American gene pools, we tested whether genetic evidence supports or falsifies the tripartite hypothesis.

Results and Discussion

Summary of the Situation in Tropical America. As discussed in a previous study (34) and summarized here in Fig. 2 *A–C*, combining the use of chloroplast and nuclear microsatellite markers, we established the existence of distinct gene pools in the Northern and Southern regions of the neotropics. Neotropical sweet pota-

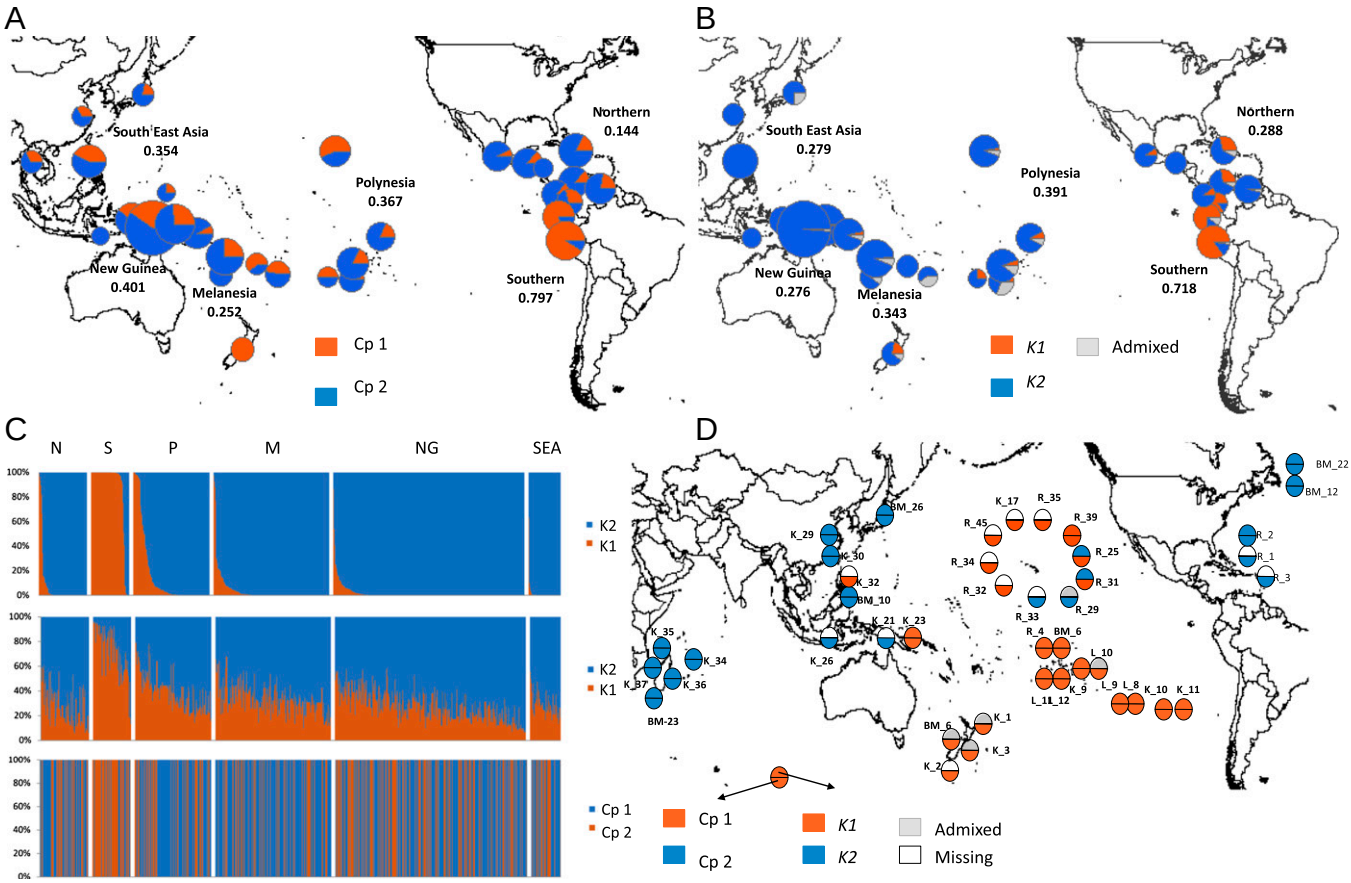

Fig. 2. Geographical distribution of nuclear and chloroplast genetic variation in sweet potato through space and over time. (*A*) Proportion of individuals belonging to each chloroplast lineage [lineage 1 in orange (Cp 1); lineage 2 in blue (Cp 2)] for each sampling site (country or archipelago). Only sites with four or more accessions were represented. Area of the circle is proportional to the square root of the sample size for the site. Values by region correspond to the frequency of chloroplast lineage 1. (*B*) Proportion of individuals belonging to each nuclear cluster (cluster *K1* in orange; *K2* in blue) for each sampling site, as determined by the DAPC analysis. Only sites with four or more accessions were represented. Area of the circle is proportional to the square root of the sample size for the site. Values by region correspond to the mean *K1* ancestry, as determined by the Bayesian clustering. (*C*) Three bar plots showing for each individual: (*i*) the probabilities of membership in nuclear clusters *K1* and *K2* as determined by DAPC (*Top*); (*ii*) the probabilities of membership in nuclear clusters *K1* and *K2* as determined by the Bayesian clustering method implemented in Structure (*Middle*); and (*iii*) the individual's chloroplast lineage (*Bottom*). Each individual is represented as a vertical bar, with colors corresponding to membership probabilities in clusters *K1* (orange), *K2* (blue), and chloroplast lineages 1 (orange) and 2 (blue). Individuals are organized by geographical origin, S for the Southern neotropical region, N for the Northern neotropical region, P for Polynesia, M for Melanesia, NG for New Guinea, and SEA for South-East Asia. (*D*) Genetic constitution of herbarium specimens collected from the 18th century to the early 20th century. Lower and upper halves of each circle represent the chloroplast lineage and nuclear cluster (as determined by DAPC), respectively.

toes are characterized by two distinct geographically restricted chloroplast lineages (in total 21 haplotypes), which correspond quite well to two nuclear genetic clusters, *K1* and *K2*, identified by *K*-means clustering grouping (Fig. 2 *A*–*C*): most accessions from the Southern region (hereafter the Southern gene pool) exhibit chloroplast haplotypes of group 1 (79.7%) and belong to the nuclear cluster *K1* (83%). Sweet potatoes from the Northern region (hereafter the Northern gene pool) carry mostly chloroplast DNA haplotypes of group 2 (85.6%) and belong to the nuclear cluster *K2* (91.46%).

Despite this clear-cut phylogeographic pattern, neotropical gene pools have secondarily come into contact, as shown by the admixture revealed with both chloroplast and nuclear markers (Fig. 2 *A*–*C*): in the Southern region, we detected some accessions with chloroplast group 2 haplotypes (16/65), some assigned to the nuclear cluster *K2* [8/65 with the discriminant analysis of principal components (DAPC) method (37), 3/65 with the Bayesian clustering method implemented in Structure 2.3.3 (38, 39)], and some that were admixed individuals (3/65 with DAPC, 34/65 with the Bayesian method). Also, in the Northern region, we identified several accessions with chloroplast group 1 haplotypes (31/82), some attributed to cluster *K1* (5/82 with DAPC), and some that were admixed individuals (2/82 with DAPC, 49/82 with the Bayesian method). This situation suggests that clones were exchanged between both regions and recombined with local material.

Assessing the Relative Contribution of Camote, Batata, and Kumara Lines of Sweet Potato Introduction into Oceania in the Modern Sample. Sweet potato varieties in Oceania represent a subset of neotropical diversity for both kinds of markers. They include representatives of both neotropical chloroplast lineages 1 and 2 (in total eight different haplotypes, including those most common in tropical America) and also some private haplotypes (eight), all rare and derived from the most common ones by only one or two mutation steps (Fig. S1 and Dataset S1). Nuclear-diversity indices [except for observed heterozygosity (*Ho*)] were slightly lower than those calculated for either of the two neotropical gene pools (Table S1).

The unrooted neighbor-joining tree revealed a relatively high degree of phylogeographic structure among Oceanian varieties (Fig. 3). As expected under the tripartite hypothesis, western Pacific varieties (varieties from South-East Asia, Melanesia, and New Guinea) are differentiated from those of Polynesia. However, incongruities between chloroplast and nuclear data prevent, in

some cases, the definitive attribution of some Oceanian varieties to an original gene pool in tropical America (Fig. 2*C*).

Using the DAPC non–model-based assignment method, most western Pacific sweet potato accessions (604 in total) were assigned to the cluster *K2*. Only four accessions were assigned to the cluster *K1*, and 14 accessions exhibited a mixed genetic constitution (probability of membership in a given cluster, <0.8). With the Bayesian clustering method, 144 accessions were assigned to the cluster *K2*, and all others had a mixed ancestry. The mean proportion of *K1* ancestry was similar (0.29 ± 0.128) to that calculated in the Northern gene pool (0.288 ± 0.145). Thus, our analysis confirms previous results (31–33) and indicates that the Northern gene pool may have been the principal contributor to the nuclear genetic diversity of extant western Pacific sweet potatoes, which, thus, appear to be derived principally from the Camote and Batata lines (31–33). Furthermore, chloroplast lineage 2 haplotypes were mainly represented throughout the western Pacific, as was the Northern nuclear gene pool. However, the frequency of chloroplast lineage 1 was generally higher (33%) than that observed in the Northern gene pool (14.4%). This suggests that South American clones may have also been directly introduced throughout the western Pacific. Alternatively, founder effects related to diffusion history and/or local selection may have increased the frequency of this haplogroup in the area of introduction, compared with that found in the already admixed original Northern gene pool.

In Polynesia, the situation is more contrasted. In New Zealand, only chloroplast lineage 1 was found, and in Hawaii, this lineage dominated, a pattern expected for varieties that originated from the Southern gene pool. In contrast, the frequency of chloroplast lineage 1 was only 19.5% in eastern Polynesia. Considering nuclear data, only 14 varieties (9 from eastern Polynesia, 1 from Hawaii, and 2 from New Zealand) were assigned to cluster *K1* and 103 to cluster *K2*, and 14 had a mixed ancestry (DAPC method). With the Bayesian clustering method, most of the accessions (127/131) had a mixed ancestry. The mean proportion of *K1* ancestry in Polynesia was 0.391 (± 0.128) [ranging from 0.332 (± 0.095) in Hawaii and 0.388 (± 0.136) in eastern Polynesia to 0.585 (± 0.186) in New Zealand], a value significantly higher than that calculated in the Northern gene pool and the western Pacific ($P < 0.01$ for both), attesting to the greater contribution of the Southern gene pool in this area. However, only seven individuals exhibited a genetic constitution clearly inherited from the Southern gene pool (chloroplast lineage 1 and assignation to cluster *K1*). Most modern accessions seem to have a mixed genetic constitution, "intermediate" between the two gene pools.

Herbarium Specimens and the Kumara Line. We recovered material from 42 herbarium accessions of sweet potato collected worldwide from the 17th century to the early 20th century and obtained data on chloroplast microsatellites for all of them and reliable data on nuclear microsatellites for 30 accessions. Almost all accessions from Polynesia exhibited a chloroplast lineage 1 haplotype (hap 14) and were assigned to nuclear cluster *K1*, a genetic background characteristic of the Southern gene pool (Fig. 2*D*). In contrast, herbarium specimens from the Caribbean, South-East Asia, and Madagascar carried a chloroplast lineage 2 haplotype (hap1_2) and were assigned to nuclear cluster *K2*, a genetic constitution likely inherited from the Northern gene pool. The DAPC and Bayesian methods gave congruent results (Fig. S2). This clear geographical pattern provides strong support for the hypothesis that the Kumara line represents a pre-Columbian diffusion of sweet potatoes from South America (Peru–Ecuador area) into Polynesia.

Of salient interest are the specimens collected by J. Banks and D. Solander during Captain James Cook's first voyage in 1769 (two plants from the Society Islands and one from New Zealand), because they may represent truly prehistoric introductions. All three are derived from the Southern gene pool (Figs. 2*D* and 3 and Fig. S3). The two plants from the Society Islands are a single clone (multilocus genotypes almost identical). The New Zealand specimen, however, appears to be a distinct variety, as attested by morphological features and confirmed by our genetic results (incompletely genotyped but still with distinct alleles). Other early 20th century specimens from eastern Polynesia also grouped with the Southern gene pool (Fig. 3 and Fig. S3). Morphologically distinguishable, they are genetically quite closely related (Fig. 3 and Fig. S3) and may be true representatives of the Kumara lineage. This result confirms the conclusion of Green (26) that distinct and identifiable lineages were already present in Polynesia at the time of early European contacts, likely having, however, a narrow genetic base (Fig. 3 and Fig. S3). Whether these lineages result from a single introduction from South America or multiple independent introductions (possibly including voyages originating from northern Colombia and Central America) (23, 28) is still unknown and the present sample of herbarium accessions does not allow us to rule out multiple prehistoric introductions.

From Ancient to Modern Patterns: Reshuffling of the Initial Genetic Background of Sweet Potato in Oceania. ***In the Eastern Pacific.*** In contrast to all herbarium specimens up to the early 20th century, most of the contemporary varieties from eastern Polynesia (80.5%) carry a chloroplast lineage 2 haplotype, supporting the long-standing contention of ethnobotanists that later introductions have replaced initial sweet potato varieties in this area (35, 40). The relative stability of genetic constitution reflected by herbarium accessions from the end of the 18th century to the beginning of the 20th century from this area may indicate that early European reintroductions (e.g., by traders, whalers, and travelers) did not immediately erase the initial genetic heritage. At first, these movements may have redistributed clones already present in the region rather than introducing new genotypes, extending the dispersal of varieties (which the Polynesians themselves had originated) throughout Polynesia and probably even further westward (24, 35). The

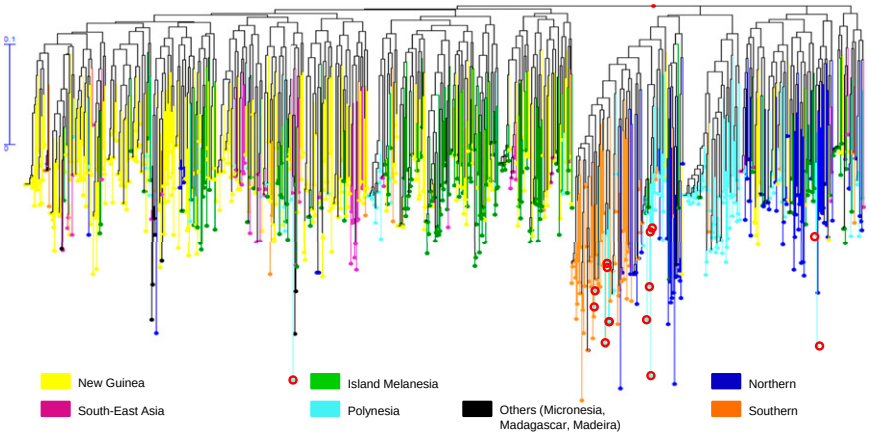

Fig. 3. Global patterns of genetic differentiation. Neighbor-joining tree based on the Lynch distance for the global dataset. Individuals are labeled according to their geographical origin, in orange, blue, cyan, green, yellow, magenta, and black for accessions from the Southern region, the Northern region, Polynesia, Island Melanesia, New Guinea, South-East Asia, and "other regions" (Micronesia, Madagascar, Madeira), respectively. Herbarium specimens from Polynesia are indicated by red open circles.

shift probably took place during the 20th century, surely before the 1960s, when Yen began his collections.

Differentiation patterns observed nowadays reveal that varieties from eastern Polynesia form a heterogeneous group, with a few accessions closely related to the Southern gene pool, some to the Northern gene pool, and some to varieties predominant in the western Pacific, whereas some others are intermediate and even form a well-differentiated cluster (Fig. 3 and Fig. S3). This group includes some clonal lineages (i.e., varieties with the same haplotype and only very few differences between genotypes) and also cultivars, which, although genetically closely related, may be derived from independent sexual recombination events. This pattern of genetic differentiation suggests that later reintroductions did not simply replace initial ones but rather reshuffled the initial genetic background. Although mainly clonally propagated, sweet potato reproduces sexually, and its volunteer seedlings are sometimes incorporated by farmers and multiplied as new clones (24, 41). In the Marquesas and Tahiti, sweet potato long remained of only secondary importance. In the early 20th century, only a few native varieties, poorly adapted to the humid tropical environments of eastern Polynesia, were recorded (42). Strong genetic bottlenecks, such as those that accompanied prehistoric diffusion of the crop by Polynesians who likely introduced a limited number of autoincompatibility groups, may have greatly limited sexual reproduction and subsequent selection of seedlings by farmers as new clones. Later European reintroductions may have allowed the initiation of local diversification. Hawaiian varieties exhibit a quite similar pattern to that shown by eastern Polynesian varieties, and most of them are part of the eastern Polynesian cluster (Fig. S3). Furthermore, some clones are shared between both regions (eastern Polynesia and Hawaii), attesting to the circulation of vegetative propagules between the two areas. Nevertheless, in Hawaii, early references by European explorers and traders described the sweet potato as a plentiful food (26), and in the early 20th century, Handy (43) characterized a great number of local varieties. Did independent prehistoric introductions, or early historical introductions by Spanish explorers (Fig. 1), broaden the initial genetic base, thus providing quite rapidly opportunities for widespread recombination and local selection of variants in this archipelago? Interestingly, early 20th Century herbarium specimens already exhibit a "mixed" constitution. Unfortunately, we identified no old Hawaiian herbarium specimens to confirm this hypothesis. Also, none of the New Zealand Kumara varieties groups with the eastern Polynesia/Hawaii cluster (Fig. S3). Two cultivars were clearly associated with the Southern gene pool on the basis of both kinds of markers and may represent true original Kumaras. Two other accessions, recognized by Maori informants to be native Kumaras (i.e., pre-European introductions) are associated with the Northern gene pool and may, instead, represent later reintroductions.

The modern genetic diversity in Polynesia, thus, results from the local reshuffling of the initial genetic background with later reintroductions. The observed pattern also reflects the likelihood that Polynesia long remained relatively isolated from the rest of the Pacific. Most notably, very few clones are shared between eastern Polynesia and the western Pacific. Movements of sweet potato germplasm between the eastern and western Pacific were quite restricted and did not homogenize sweet potato genetic diversity in Oceania, a genetic pattern that reflects well the cultural boundary between Polynesia and Island Melanesia.

In the Western Pacific. In contrast, the lack of geographic structure among western Pacific varieties (Fig. 3 and Fig. S3) suggests a common genetic background consistent with the tripartite hypothesis. Moreover, this pattern suggests a wide circulation of varieties across this region, also attested to by the presence of several shared clones (Dataset S2). Still, the mean value of *K1* ancestry found in Melanesia was slightly higher (0.343 ± 0.113) than that observed in New Guinea (0.276 ± 0.116) and in South-East Asia (0.279 ± 0.097), which both showed values very similar to that observed in the Northern gene pool (Table S1). The greater contribution of the Southern gene pool in Melanesia likely reflects early introductions of true Kumara clones, as also suggested by the common use of cognates of this term ("kumara/kumala") to designate sweet potato in eastern Melanesia (28, 35). Also, New Guinea varieties appear to be slightly differentiated from those of other areas in the western Pacific. These genetic data, combined with the presence of many names seemingly unrelated to each other (24), likely reflect multiple processes of local diversification in this island. Also, the frequency distribution of pairwise Manhattan distances for western Pacific genotypes (Fig. S4) suggests that the high diversity in this region appears to have resulted, as suspected by Yen (25), from intensive recombination among introductions and from the use by farmers of plants issued from true seed, rather than from clonal evolution and selection by farmers of somatic mutants.

Did Genes and Names Disperse Together? The mixed clonal/sexual reproductive system of sweet potato permits frequent recombination of newly introduced genotypes with local material. In these recombined genotypes, the name likely follows the phenotype. Distinctive phenotypes may reflect alleles at loci that are under strong selection and may show no correlation with the widely reshuffled neutral genetic background. Thus, old names may continue to be applied to varieties that have been strongly affected by modern plant movement and local recombination. This is what likely happened to most of the "native" Polynesian varieties originating from South America, which were progressively admixed with material introduced much later. An illustration of this evolutionary scenario may be offered by the group of ancient Polynesian sweet potatoes named as *Convolvulus chrysorrhizus dulcis* by D. Solander (and followed by J. Forster) because of the bright yellow color of the internal root flesh. In 1960, Yen identified two very similar groups of yellow varieties (yellow internal flesh and fusiform root) with a pan-Polynesian distribution (the "re'amoa" and "hererai" groups) and considered by local people to be native Kumaras (25, 26). Whereas the "native" yellow variety described by Banks and Solander carries a chloroplast lineage 1 haplotype, all of the contemporary yellow varieties bear a chloroplast lineage 2 haplotype and are very closely related, probably forming a clonal lineage or a group of genetically closely related varieties. The neutral genetic basis likely has shifted, whereas the initial phenotype, and probably the initial names, appear to have been well conserved. Thus, genotypes and names do not always migrate together, and information on each gives access to complementary parts of the plant's dispersal history.

Conclusion

Our study has shown how historical collections reveal patterns of diffusion of sweet potato in Oceania that have been obscured by modern plant movements and local recombination. Our results provide strong genetic support, previously lacking, for the tripartite hypothesis, notably concerning the Kumara line, the pre-Columbian diffusion of sweet potatoes from South America into Polynesia. We also suggest why phenotypes, names, and neutral genes do not necessarily share completely parallel evolutionary histories. Whereas old names and old phenotypes may endure, recombination may widely reshuffle their neutral genetic backgrounds and obscure the pathways of initial diffusion.

Exploiting crops as a proxy of human movements in Oceania requires multidisciplinary approaches combining linguistics, the morphological characterization of plants, and phylogeography, in which the use of herbarium collections may play a critical role.

Materials and Methods

Sampling and Genotyping. We analyzed a total of 1,245 sweet potato accessions throughout tropical America and Oceania and including some accessions from South-East Asia and Madagascar (1,188 modern accessions from several important *ex situ* collections representing the extant variability in Oceania and a total of 57 sweet potato herbarium specimens). Genetic diversity was assessed by using six chloroplast microsatellite loci and 11 nuclear microsatellite loci. For herbarium specimens, extraction and pre-PCR processing were conducted in a separate room specifically designed for degraded DNA (44) (see *SI Text* and Datasets S3 and S4 for details).

Assessing the Genetic Relationships Between Oceanian and Tropical American Gene Pools with Chloroplast and Nuclear Data. We constructed the median joining networks of haplotypes from tropical America and Oceania using the software Network 4.5.6.1 (45) (http://www.fluxus-engineering.com/sharenet.htm) and compared the proportions of haplotypes from each chloroplast lineage in the two areas.

We then resorted to two kinds of methods to assign Oceanian varieties to their neotropical source gene pools on the basis of nuclear data. First, we used a non–model-based method, the DAPC, a multivariate analysis recently developed and implemented in the Adegenet R package (38). We ran DAPC on the tropical American dataset, using Oceanian varieties as supplementary individuals, retaining five principal components (29.3% of the total variance) for prior data transformation, for $K = 2$ (prior grouping obtained following the *K*-means clustering algorithm implemented in the Adegenet package) (see *SI Appendix* for details). We compared DAPC assignment with that obtained by a Bayesian model-based method implemented in the software Structure 2.3.3, recently adapted to handle autopolyploid data (39, 40). By prespecifying source gene pools in tropical America (as defined by the *K*-means clustering for $K = 2$), the algorithm estimates ancestry for additional individuals (from Oceania), updating allele frequencies using only those from tropical America. We ran the admixture model with $K = 2$ (same grouping as for DAPC), correlated allele frequencies, 50,000 burn-in iterations and 150,000 Markov chain–Monte Carlo steps, and with data coding for handling genotype ambiguity for codominant markers in polyploids (40). An individual is considered to be well assigned to its source gene pool (*K1* or *K2*) if the associated membership probability (DAPC) or ancestry value (Bayesian clustering) is greater than 0.8; otherwise, the individual is considered "admixed."

Global genetic structure was also characterized by constructing a Lynch distance-based Neighbor-joining tree [with the APE R package (46)] and then editing it with the software Darwin 5.0.158 (http://darwin.cirad.fr/darwin). Classical diversity indices, as well as pairwise Manhattan distances, were computed using custom R scripts.

ACKNOWLEDGMENTS. We thank Marc Spencer, Jonathan Gregson, and Jacek Wajer (Natural History Museum); David Goyder and Felix Forest (Royal Botanical Gardens); Gerard Thijsse (National Herbarium Nederland); and Napua Harbottle (Bishop Museum) for providing herbarium specimens. We also thank Toru Kumagai (National Institute for Agronomical Sciences), Maurice Wong (Rural Development Services), and Wulf Schiefenhövel (Max-Planck-Institute for Ornithology) for their help in sampling "modern" sweet potato material. We also thank two reviewers for very helpful comments. Data used in this work were produced through the technical facilities of the Mediterranean Center for Environment and Biodiversity. This study was funded by grants from the "Pacific Funds," distributed by the French Ministry of European and Foreign Affairs; support was also received from the SYNTHESYS Project (Synthesis of Systematic Resources; www.synthesys.info), which is financed by European Community Research Infrastructure Action under the Seventh Framework Program. C.R. was supported by a doctoral grant from the French Ministry of Higher Education and Research.

1. Hurles ME, Matisoo-Smith E, Gray RD, Penny D (2003) Untangling oceanic settlement: The edge of the knowable. *Trends Ecol Evol* 18(10):531–540.
2. Kirch PV (2010) Peopling of the Pacific: A holistic anthropological perspective. *Annu Rev Anthropol* 39:131–148.
3. Thorsby E (2012) The Polynesian gene pool: An early contribution by Amerindians to Easter Island. *Philos Trans R Soc Lond B Biol Sci* 367(1590):812–819.
4. Kayser M (2010) The human genetic history of Oceania: Near and remote views of dispersal. *Curr Biol* 20(4):R194–R201.
5. Matisoo-Smith E, et al. (1998) Patterns of prehistoric human mobility in polynesia indicated by mtDNA from the Pacific rat. *Proc Natl Acad Sci USA* 95(25):15145–15150.
6. Matisoo-Smith E, Robins JH (2004) Origins and dispersals of Pacific peoples: Evidence from mtDNA phylogenies of the Pacific rat. *Proc Natl Acad Sci USA* 101(24):9167–9172.
7. 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.
8. Clarke AC, Burtenshaw MK, McLenachan PA, Erickson DL, Penny D; SMBE Tri-National Young Investigators (2006) Proceedings of the SMBE Tri-National Young Investigators' Workshop 2005. Reconstructing the origins and dispersal of the Polynesian bottle gourd (Lagenaria siceraria). *Mol Biol Evol* 23(5):893–900.
9. Perrier X, et al. (2011) Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proc Natl Acad Sci USA* 108(28):11311–11318.
10. Storey AA, et al. (2007) Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. *Proc Natl Acad Sci USA* 104(25):10335–10339.
11. Gongora J, et al. (2008) Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. *Proc Natl Acad Sci USA* 105(30):10308–10313.
12. Larson G, et al. (2007) Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania. *Proc Natl Acad Sci USA* 104(12):4834–4839.
13. Moodley Y, et al. (2009) The peopling of the Pacific from a bacterial perspective. *Science* 323(5913):527–530.
14. Storey A, Clarke AC, Matisoo-Smith E (2011) *Polynesians in America: Pre-Columbian Contacts with the New World*, eds Jones TL, Storey A, Matisoo-Smtih E, Ramirez-Aliaga JM (AltaMira Press, Plymouth, UK).
15. Lawler A (2010) Beyond Kon-Tiki: Did Polynesians sail to South America? *Science* 328(5984):1344–1347.
16. Austin DF (1988) The taxonomy, evolution and genetic diversity of sweet potatoes and related wild species. in *Exploration, Maintenance and Utilization of Sweet Potato Genetic Resources: Report of the First Sweet Potato Planning Conference 1987* (International Potato Center, Lima, Peru), pp 27–59. Available at http://cipotato.org/library/pdfdocs/SW19066.pdf. Accessed January 4, 2013.
17. Hather JG, Kirch PV (1991) Prehistoric sweet potato (*Ipomoea batatas*) from Mangaia Island, central Polynesia. *Antiquity* 65:887–893.
18. Ladefoged TN, Graves MW, Coil J (2005) The introduction of sweet potato in Polynesia: Early remains in Hawaii. *J Polyn Soc* 114(4):359–374.
19. Horrocks M, Rechtman RB (2009) Sweet potato (*Ipomoea batatas*) and banana (*Musa* sp.) microfossils in deposits from the Kona Field System, island of Hawaii. *J Archaeol Sci* 36(5):1115–1126.
20. Wallin P, Stevenson C, Ladefogded T (2005) *The Sweet Potato in Oceania: A Reappraisal*, eds Ballard C, Brown P, Bourke RM, Harwood T (Oceania Publications, Sydney), pp 85–88.
21. Horrocks M, Shane PA, Barber IG, D'Costa DM, Nichol SL (2004) Microbotanical remains reveal Polynesian agriculture and mixed cropping in early New Zealand. *Rev Palaeobot Palynol* 131(3-4):147–157.
22. Scaglion R (2005) *The Sweet Potato in Oceania: A Reappraisal*, eds Ballard C, Paula B, Bourke RM, Harwood T (Oceania Publications, Sydney), pp 35–42.
23. Montenegro Á, Avis C, Weaver A (2008) Modeling the prehistoric arrival of the sweet potato in Polynesia. *J Archaeol Sci* 35(2):355–367.
24. Barrau J (1957) L'énigme de la patate douce en Océanie. *Etudes d'Outre-Mer* 40:83–87.
25. Yen DE (1974) *The Sweet Potato in Oceania: An Essay in Ethnobotany* (Bishop Museum Press, Honolulu).
26. Green R (2005) *The Sweet Potato in Oceania: A Reappraisal*, eds Ballard C, Brown P, Bourke RM, Harwood T (Oceania Publications, Sydney), pp 43–62.
27. Clarke A (2009) Origins and dispersal of the Sweet potato and bottle gourd in Oceania: Implications for prehistoric human mobility. PhD thesis (Massey Univ, Palmerston North, New Zealand).
28. Rensch KH (1991) *Polynesian Plant Names: Linguistic Analysis and Ethnobotany, Expectations and Limitations*, eds Cox P, Banack S (Dioscorides Press, Portland), pp 97–111.
29. Allen MG (2005) *The Sweet Potato in Oceania: A Reappraisal*, eds Ballard C, Brown P, Bourke RM, Hardwood T (Oceania Publications, Sydney), pp 99–108.
30. Baert AMG (2001) *Histoire de la Découverte des Régions Australes: Les Salomon, Marquises, Santa Cruz, Tuamotu, Cook du Nord et Vanuatu* (l'Harmattan, Paris).
31. Zhang D, Ghislain M, Huam Z, Golmirzaie A, Hijmans R (1998) RAPD variation in sweetpotato (*Ipomoea batatas* (L.) Lam.) cultivars from South America and Papua New Guinea. *Genet Resour Crop Evol* 45(3):271–277.
32. Zhang D, Rossel G, Kriegner A, Hijmans R (2004) AFLP assessment of diversity in sweetpotato from Latin America and the Pacific region: Its implications on the dispersal of the crop. *Genet Resour Crop Evol* 51(2):115–120.
33. Gichuki ST, et al. (2003) Genetic diversity in sweetpotato [*Ipomoea batatas* (L.) Lam.] in relationship to geographic sources as assessed with RAPD markers. *Genet Resour Crop Evol* 50(4):429–437.
34. Roullier C, Rossel G, Tay D, McKey D, Lebot V (2011) Combining chloroplast and nuclear microsatellites to investigate origin and dispersal of New World sweet potato landraces. *Mol Ecol* 20(19):3963–3977.
35. Yen DE (1998) Subsistence to commerce in Pacific agriculture: Some four thousand years of plant exchange. *Plants for Food and Medicine*, eds Pendergast HDV, Etkin NL, Harris DR, Houghton PJ (Royal Botanical Garden, Kew, UK).
36. Matisoo-Smith E (2002) Something old, something new: Do genetic studies of contemporary populations reliably represent prehistoric populations of Pacific *Rattus exulans? Hum Biol* 74(3):489–496.
37. Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genet* 11:94.
38. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945–959.
39. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164(4):1567–1587.
40. Barrau J (1961) *Subsistence Agriculture in Polynesia and Micronesia* (Bernice P. Bishop Museum, Honolulu).
41. Schneider J (1995) *Indigenous Knowledge in Conservation of Crop Genetic Resources: Proceedings of an International Workshop, Held in Cisarua, Bogor, Indonesia, January 30–February 3, 1995* (Central Research Institute for Food Crops, Bogor, Indonesia), pp 63–70.
42. Brown FBH (1935) *Flora of Southeastern Polynesia* (Bernice P. Bishop Museum, Honolulu), Vol 3.
43. Handy ESC (1940) *The Hawaiian Planter. His Plants, Methods and Areas of Cultivation* (Bernice P. Bishop Museum, Honolulu), Vol 1.
44. Knapp M, Clarke AC, Horsburgh KA, Matisoo-Smith EA (2012) Setting the stage - building and working in an ancient DNA laboratory. *Ann Anat* 194(1):3–6.
45. Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16(1):37–48.
46. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20(2):289–290.