



# Origin and Dispersal of Domesticated Peach Palm

Charles R. Clement<sup>1,2\*</sup>, Michelly de Cristo-Araújo<sup>1,2</sup>, Geo Coppens d'Eeckenbrugge<sup>3</sup>, Vanessa Maciel dos Reis<sup>2</sup>, Romain Lehnebach<sup>4,5</sup> and Doriane Picanço-Rodrigues<sup>2</sup>

<sup>1</sup> Department of Technology and Innovation, Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil, <sup>2</sup> Laboratório de Evolução Aplicada, Instituto de Ciências Biológicas, Universidade Federal do Amazonas, Manaus, Brazil, <sup>3</sup> Centre de Coopération Internationale en Recherche Agronomique Pour le Développement, UMR AGAP, Montpellier, France, <sup>4</sup> Université Montpellier II, UMR AMAP (Botanique et Bio-Informatique de L'architecture des Plantes), Montpellier, France, <sup>5</sup> Centre de Coopération Internationale en Recherche Agronomique Pour le Développement, UMR AMAP, Kourou, French Guiana

## OPEN ACCESS

### Edited by:

Eike Luedeling,  
Consultative Group on International  
Agricultural Research, United States

### Reviewed by:

Paul Gepts,  
University of California, Davis,  
United States  
Rachel Meyer,  
University of California, Los Angeles,  
United States

### \*Correspondence:

Charles R. Clement  
cclement@inpa.gov.br

### Specialty section:

This article was submitted to  
Agroecology and Land Use Systems,  
a section of the journal  
Frontiers in Ecology and Evolution

**Received:** 18 May 2017

**Accepted:** 13 November 2017

**Published:** 29 November 2017

### Citation:

Clement CR, Cristo-Araújo M,  
Coppens d'Eeckenbrugge G,  
Reis VM, Lehnebach R and  
Picanço-Rodrigues D (2017) Origin  
and Dispersal of Domesticated Peach  
Palm. *Front. Ecol. Evol.* 5:148.  
doi: 10.3389/fevo.2017.00148

Peach palm (*Bactris gasipaes* Kunth) is a Neotropical palm domesticated by Native Americans. Its domestication resulted in a set of landraces (var. *gasipaes*), some with very starchy fruit used for fermentation, others with an equilibrium of starch and oil used as snacks. Which of the three wild types (var. *chichagui*) was involved and where the domestication process began are unclear, with three hypotheses under discussion: an origin in southwestern Amazonia; or in northwestern South America; or multiple origins. We reevaluate one of the wild types, defining it as the incipient domesticate, and then evaluate these hypotheses using the Brazilian peach palm Core Collection and selected herbaria samples to: (1) model the potential distributions of wild and domesticated populations; (2) identify the probable origin of domestication with a phylogeographic analysis of chloroplast DNA sequences; and (3) determine the dispersal routes after domestication using spatial analysis of genetic diversity based on 17 nuclear microsatellite loci. The two very small-fruited wild types have distinct distributions in the northern Andes region and across southern Amazonia, both under moderately humid climates, while the incipient domesticate, partly sympatric with the southern wild type, is also found along the Equatorial Andes, in a more humid climatic envelope, more similar to that of the domesticated landraces. Two distribution models for Last Glacial Maximum conditions (CCSM4, MIROC) also suggest distinct distributions for the two wild populations. The chloroplast DNA phylogeographic network confirms the area of sympatry of the incipient domesticate and the southern wild type in southwestern Amazonia as the origin of domestication. The spatial patterns of genetic diversity confirm the proposal of two dispersals, one along the Ucayali River, into western Amazonia, northwestern South America and finally Central America; the other along the Madeira River into central and then eastern Amazonia. The first dispersal resulted in very starchy fruit for fermentation, while the second may have been later and resulted in snack fruits. Further explorations of southwestern Amazonia are essential for more precise identification of the earliest events, both with new archeological methods and genetic analyses with larger samples.

**Keywords:** *Bactris gasipaes*, chloroplast phylogeography, ecological niche models, landrace biogeography, microsatellite markers

## INTRODUCTION

The peach palm (*Bactris gasipaes* Kunth, Palmae) is a Neotropical palm with populations domesticated by Native Americans (Clement, 1988), and presents impressive morphological diversity in its wild and cultivated populations, since these occur in different environments and exhibit different degrees of domestication (Mora-Urpí et al., 1997). At the time of European conquest, peach palm was an important food crop and the basis of a fermented drink, both of which featured in community festivals from western Amazonia to southern Central America (Mora-Urpí et al., 1997; Patiño, 2002). It was less important in the rest of humid-lowland northern South America (Patiño, 1963, 2002).

The origin of domesticated peach palm from wild populations remained a matter of speculation for more than a century, until the systematic analysis of *Bactris* presented by Henderson (2000). Since then several hypotheses have proposed a single origin (Morcote-Rios and Bernal, 2001; Rodrigues et al., 2005; Cristo-Araújo et al., 2013; Galluzzi et al., 2015) or multiple origins (Mora-Urpí, 1999; Hernández-Ugalde et al., 2011). We will identify a problem in the systematic analysis that has influenced many of these hypotheses and is essential to understanding the origin of domesticated peach palm, then model the ecological niches of the wild populations, and expand our genetic analyses of the origin of the domesticated populations (Cristo-Araújo et al., 2013) to understand their dispersal.

Henderson (2000) reduced the previously recognized nine species and three varieties in Martius' genus *Guilielma* into synonymy with *B. gasipaes*, and proposed two varieties: *chichagui* (H. Karsten) A.J. Henderson, including wild populations with small fruits (1.2–2.3 × 1.1–1.8 cm); and *gasipaes*, including domesticated populations of peach palm with large fruits (3.5–6.5 × 3–4.5(–6) cm) (p. 71). This revision allowed phylogenetic hypotheses about the origin of var. *gasipaes* and the subsequent dispersal of its cultivated populations and landraces (Cristo-Araújo et al., 2013). However, there is a disjunction in fruit sizes between var. *chichagui* and var. *gasipaes* that should not exist if var. *gasipaes* was domesticated from var. *chichagui*. This may be due to lack of herbarium samples that fill the gap, or to attributions of synonymy during Henderson's revision, or both.

Within var. *chichagui*, Henderson (2000) proposed the existence of three wild morphotypes, without attributing synonymy of previously accepted species to morphotype. Here we expand on Ferreira (1999) and Clement et al. (2009b), and propose that type 1 is synonymous with *Guilielma mattogrossensis* Barbosa Rodrigues and *G. microcarpa* Huber, type 2 with *G. macana* Martius and *Bactris caribaea* H. Karsten, and type 3 with *B. speciosa* Martius var. *chichagui* H. Karsten, hence the varietal name in Henderson's combination. Two other previously accepted species also have small to very small fruit: *G. insignis* Martius and *Martinezia ciliata* Ruiz & Pavon. Both were attributed to var. *gasipaes* by Henderson (p. 71); however, both are more likely to be synonymous with var. *chichagui* type 3 (Clement et al., 2009b).

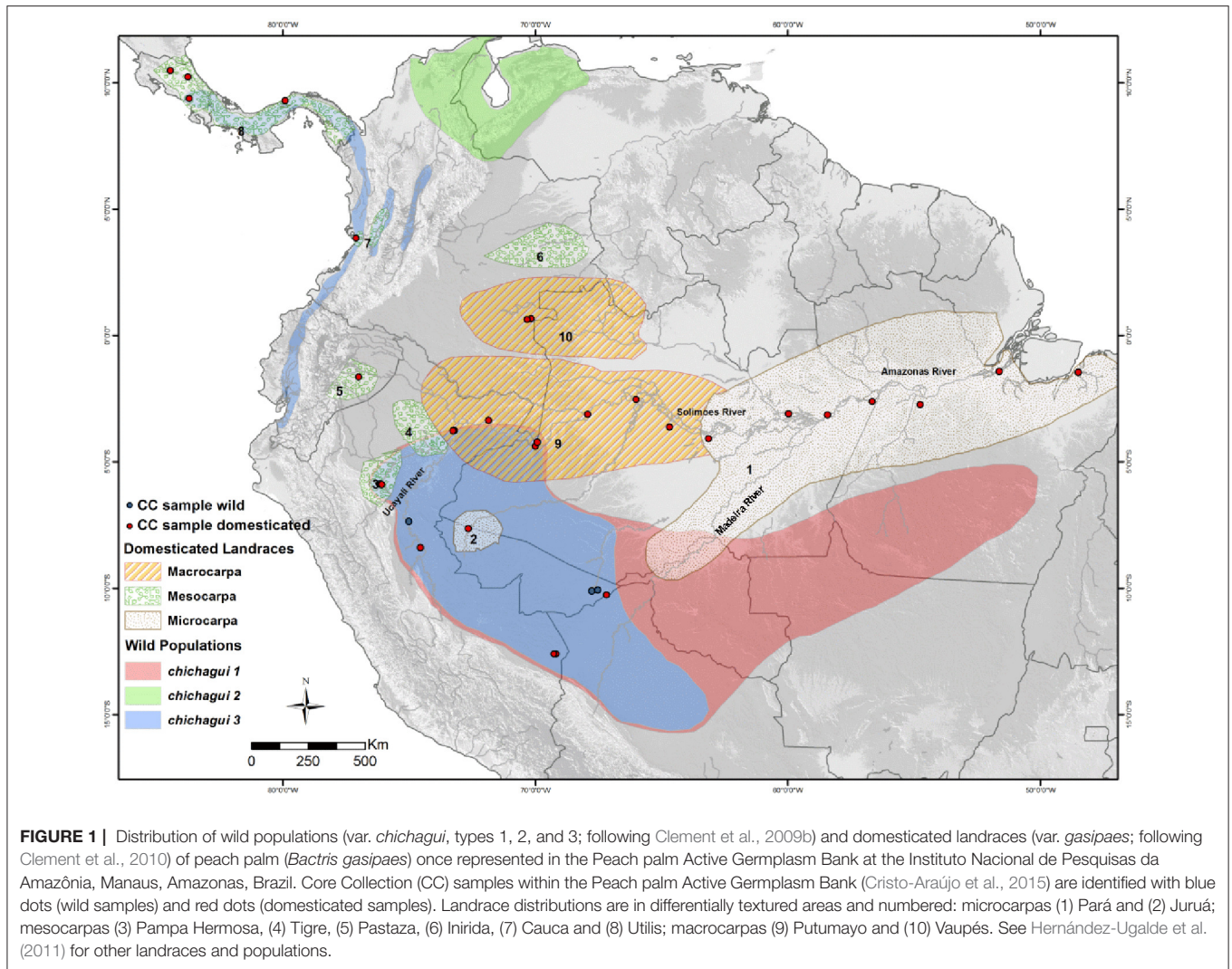
When mapping the distribution of wild types 1 and 3 in southern Amazonia, Clement et al. (2009b) identified sympatry

in southwestern Amazonia, while type 2 is isolated in northern South America (**Figure 1**). Hernández-Ugalde et al. (2011) describe the geological history of northern South America, and how this contributed to isolate the wild types 1 and 2 into their current distributions. Type 3 is the most variable of Henderson's wild types, with fruits that range from 2 to 10 g, rarely 15 g, whereas types 1 and 2 both have fruits that range from 0.5 to 2 g. Sympatry of types 1 and 3 was also noted by Huber (1904), who suggested that hybridization between small-fruited *G. insignis* and his very small-fruited *G. microcarpa* could explain the origin of cultivated peach palm in southwestern Amazonia. Observe that Huber apparently considered *G. insignis* fruits to be smaller than typical cultivated var. *gasipaes*, hence our suggestion that it should be synonymous with var. *chichagui* type 3, contrary to Henderson. Sympatry has additional significance: gene flow inhibits population divergence (Futuyma, 2005; p. 216) and this suggests that one type is not valid as a wild type.

These observations allow us to define var. *chichagui* type 3 as the incipient domesticate, following Clement et al. (2009a), i.e., it represents the beginning of domestication of peach palm from type 1 (Cristo-Araújo et al., 2013; Galluzzi et al., 2015). Saldías-Paz (1993) observed exactly the expected type of variation in fruit size and human propagation in lowland Bolivia, where very small fruits similar to type 1, which he equated to *G. microcarpa*, were observed in open forests, small fruits similar to type 3, which he equated to *G. insignis*, were observed in anthropogenic forests and were tolerated when they appeared in swiddens, and only var. *gasipaes* microcarpa-type fruits were intentionally propagated.

This redefinition of type 3 as the incipient domesticate explains the variability in fruit size, from the 2 g of type 1 to the minimum 15 g of a microcarpa population of var. *gasipaes*, as well as the disjunct distribution of type 3 (**Figure 1**), since dispersal by humans is effective for crossing barriers such as the Andes (see Supplementary Material 1.1. for more about type 3). This proposal is also consistent with a hypothesis of Graefe et al. (2012) that some "natural populations are in reality feral populations, i.e., material from cultivated populations that have gone wild," even though they doubted its validity because of the advanced degree of domestication of most var. *gasipaes* populations. Type 3 has generally been confused for a wild type precisely because it survives in little-disturbed ecosystems, i.e., it can become feral quite successfully.

Domestication is a co-evolutionary process in which human selection, both conscious and unconscious, interact with natural selection and result in changes in the population's genotypes and phenotypes that make them more useful to humans and better adapted to human intervention in the landscape (Rindos, 1984; Clement, 1999). Consequently, different populations may present different modifications due to selection, which Clement (1999) organized along a continuum from incipiently domesticated to semi-domesticated to domesticated. At the origin of domestication, incipient domesticates exhibit small differences from the local wild populations and hybridize freely with them, inhibiting changes due to human selection (Miller and Gross, 2011), just as observed by Saldías-Paz (1993). As the incipiently domesticated populations become more useful, their numbers increase in anthropogenic landscapes,



permitting greater responses to selection. When humans disperse domesticates beyond the distribution of wild populations, response to selection will be freed from frequent introgression with wild types (Clement et al., 2009a; Miller and Gross, 2011). The distribution of peach palm's landrace complex (**Figure 1**) demonstrates the expected trends, with the macrocarpa landraces exhibiting dramatic changes in fruit size and little sympatry (Putumayo) or no sympatry (Vaupés) with wild populations.

Before Henderson's revision, some of the domesticated populations of peach palm had been grouped into landraces (Mora-Urpí et al., 1997) and their distribution was mapped (**Figure 1**). This landrace classification was based on fruit size, as it reflects the degree of change due to human selection during domestication (Mora-Urpí, 1984; Clement, 1988; Meyer et al., 2012). Microcarpa landraces have small fruits (<25 g), mesocarpa landraces have intermediate sized fruits (25–70 g), and macrocarpa landraces have large fruits (>70 g) (Mora-Urpí et al., 1997).

One result of the domestication process that is very important for our discussion is ecological adaptation. Fully domesticated populations have reduced ecological adaptation in their original ecosystems, having lost defensive mechanisms, reproductive success, competitive ability etc., generally due to natural selection for adaptation to human created agroecosystems (Harlan, 1992; Clement, 1999; Purugganan and Fuller, 2009). By definition the incipient domesticate has not lost ecological adaptation (Clement, 1999), partly because the domestication process is only starting and partly as a consequence of hybridization with local wild plants (Miller and Gross, 2011). *Var. chichagui* type 3's adaptation to advanced secondary succession in anthropogenic landscapes, as observed by Saldías-Paz (1993) in Bolivia, and its survival in naturally open forests, explain why so many botanists have considered it to be wild. This also explains why hypotheses about the origin of domestication of *var. gasipaes* from *var. chichagui* type 3 outside of the distribution of *var. chichagui* type 1 are problematic, e.g., Mora-Urpí (1999), Morcote-Rios and Bernal (2001), and Hernández-Ugalde et al. (2011), and why



secondary domestication events (Galluzzi et al., 2015) outside of the distribution of var. *chichagui* type 1 are also.

Identifying the origin of domestication and tracing subsequent dispersal routes of cultivated plants is a multidisciplinary task, involving botanical, biogeographical, historical, archeological, linguistic and genetic evidence. In the case of peach palm, there are some historical references and numerous indigenous names (Patiño, 1963, 2002), but little archeological information (Morcote-Rios and Bernal, 2001). However, there is abundant botanical and biogeographic information synthesized by Henderson (2000), and an increasing abundance of genetic information. Early molecular genetic studies found deep divergence between populations in southwestern to eastern Amazonia (Pará landrace, **Figure 1**), and those in western Amazonia, northern South America and Central America (the other landraces, **Figure 1**) (Rojas-Vargas et al., 1999; Rodrigues et al., 2005; Hernández-Ugalde et al., 2011). Genetic introgression between adjacent cultivated and supposedly wild populations (in fact, populations of var. *chichagui* type 3) was reported (Couvreur et al., 2006). Hernández-Ugalde et al. (2011) interpreted these relationships among cultivated populations and adjacent wild populations in at least three regions as independent domestications. Clement et al. (2010) reviewed the molecular evidence and kept the most parsimonious hypothesis: a single domestication event in southwestern Amazonia with two dispersals. They reasoned that because nuclear DNA markers are inherited from both parents and undergo recombination, these markers are not ideal for identifying origins. Analysis with chloroplast DNA avoids the problems of meiotic recombination and biparental inheritance, and is more suitable for phylogeographic analysis (Avise, 2004). The first analysis with a chloroplast sequence (Cristo-Araújo et al., 2013) strongly suggests a single domestication event in southwestern Amazonia. We will expand on this analysis here.

Ecological Niche Modeling (ENM), which allows approximating both current and past distributions of species, has been used in a number of cases to test biogeographical hypotheses with cultivated species. Galluzzi et al. (2015) modeled wild and domesticated peach palm distributions, without a clear understanding of the implications of var. *chichagui* type 3, the incipient domesticate. As has been shown in cotton (*Gossypium hirsutum* L.), the distribution range of cultivated plants is considerably wider than that of their wild ancestor, because their climatic envelope is essentially delimited by the fundamental niche of the species (defined by abiotic constraints), as farmers, through common agricultural practices, control most biotic components of the ecological niche, particularly competition and parasitism (Coppens d'Eeckenbrugge and Lacape, 2014). In contrast, the distribution of wild populations is delimited by the species' realized niche, as competition and predation fully constrain the species' distribution. Finally, the distribution of feral populations is intermediate between that of wild and cultivated populations, as the impact of biotic factors depends on the degree of landscape anthropization. Thus, feral cotton's distribution is very similar to that of cultivated landraces; however, they tend to disappear in a few generations after fields are replaced by secondary vegetation.

Referring explicitly to the niche concept underlying the distribution models of wild, feral and cultivated plants allows evaluating the status of particular populations. In the particular case of peach palm, we can examine the ecological niche similarities for the three var. *chichagui* morphotypes. Under our working hypothesis that type 3 populations are incipiently domesticated forms, we expect that many of them are found outside of the niches of the two other morphotypes. Furthermore, once the wild status of the two morphotypes is assessed, we can evaluate hypotheses about the species' distribution at the time when humans began to interact with native Amazonian plants at the end of the last glacial period.

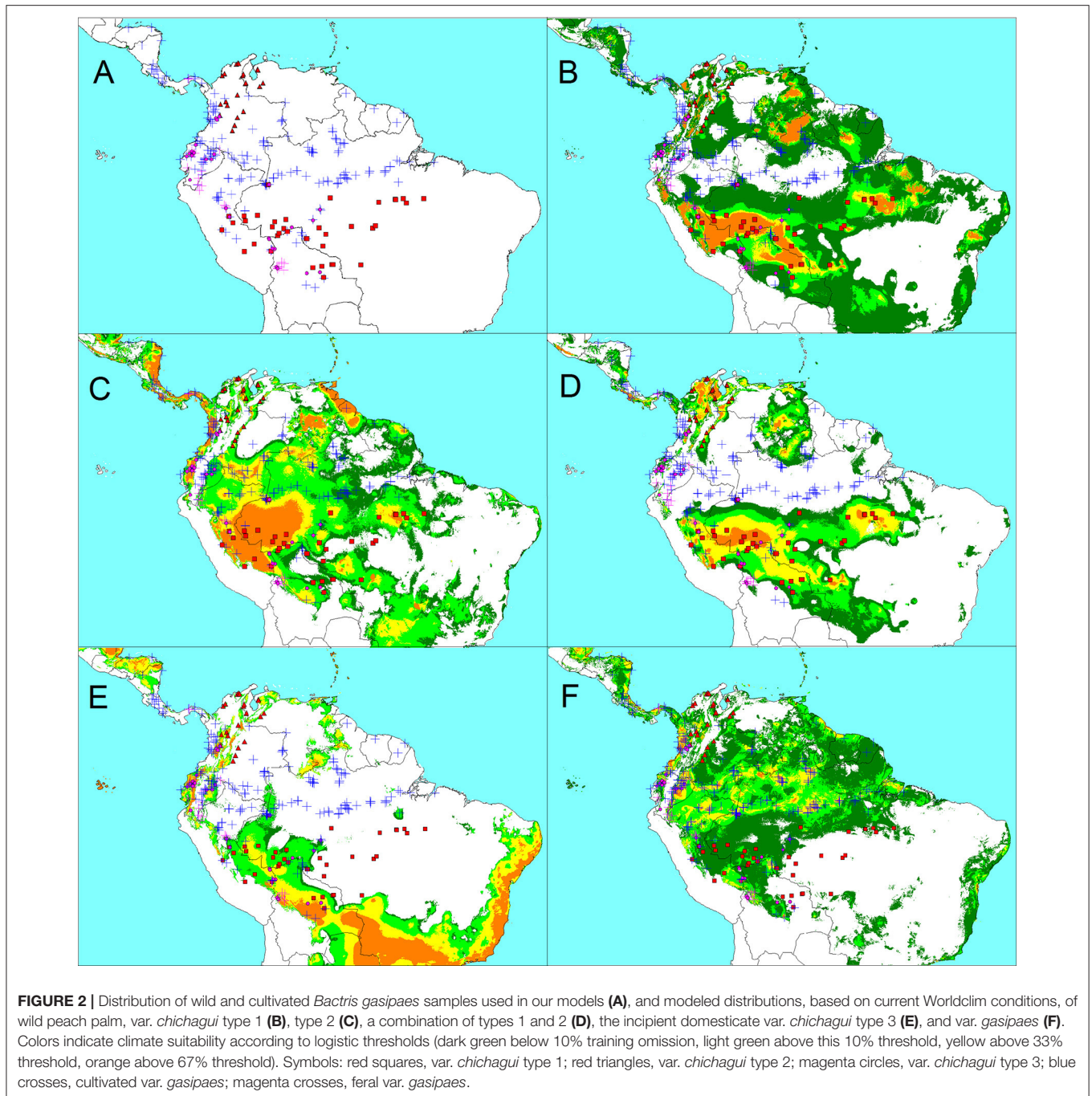
This study aimed to evaluate hypotheses about the origin of domesticated peach palm using: (1) ecological niche models to identify the potential distributions of wild and domesticated populations, to compare these with known distributions, to assess the status of var. *chichagui* type 3 in light of types 1 and 2, and then project models of these distributions on climatic models for the last glacial maximum; (2) phylogeographic analysis of chloroplast DNA sequences to determine the relationship among wild and domesticated populations, as well as the probable origin of domestication; and (3) phylogenetic analysis and spatial distribution of genetic diversity of peach palm, based on nuclear microsatellite loci, to determine the location of areas with greater genetic diversity and likely dispersal routes after domestication.

## RESULTS

### Distributions of Wild and Domesticated Peach Palm

The geographic distribution of our sample for modeling (**Figure 2A**) is similar to that shown in **Figure 1**. The two wild morphotypes of var. *chichagui* are clearly separated by an Equatorial band, with type 1 in southern Amazonia, and type 2 along the Caribbean coastal region of Colombia and western Venezuela, the Andean foothills of the Colombian and Venezuelan Orinoquia, and the Andean valleys in Colombia, reaching elevations above 1,000 m, south to the Quindío region in the Cauca Valley. Type 3 is sympatric with type 1 in southwestern Amazonia and it is also found on both sides of the Andes of Ecuador, as well as in the Cauca Valley in Colombia, where it appears marginally sympatric with type 2. The main differences between **Figures 1, 2A** concern the presence of type 2 further south in Colombia, confirmed by Rodrigo Bernal, Universidad Nacional de Colombia. Our sample shows limited sympatry between var. *chichagui* and var. *gasipaes*. The few cases where var. *chichagui* is found in close proximity to var. *gasipaes* (Cauca department in Colombia, parts of Ecuador, Peru, and the upper Solimões River in Brazil) involve type 3. Interestingly, samples of var. *gasipaes* in secondary vegetation (magenta crosses) are more common in regions where type 3 is found.

The Principal Component Analysis (PCA) characterizing the climatic envelopes of the different taxa (**Table 1, Figure 3**) did not use bioclimatic variables 1 (mean annual temperature), 5 (maximal temperature of warmest month), 12 (annual



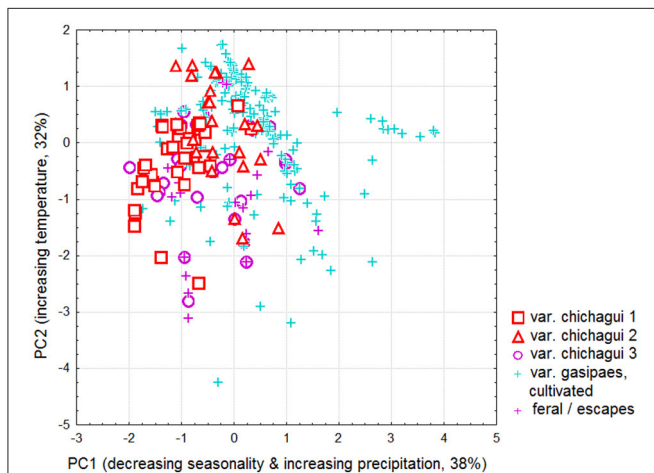
precipitation), and 16 (precipitation of wettest quarter), as they contributed little to the different Ecological Niche Models and/or can be deduced directly from other variables. The first axis is related to decreasing seasonality and increasing precipitation. The second axis is positively correlated with temperatures. Representatives of var. *chichagui* types 1 and 2 found in open tropical forests are concentrated in the upper left of the principal plane (warm tropical climate with more pronounced seasonality). A similar trend is observed for var. *gasipaes*; however, its climatic space is wider as it can be grown under

more humid equatorial conditions, which is fully consistent with its common presence along the Equator. The climatic envelope of var. *chichagui* type 3 is intermediate between that of type 1 (its putative wild ancestor), and that of domesticated peach palm (var. *gasipaes*) (Supplementary Material 1.2; Figure S2), which is consistent with our working hypothesis that this type is the incipient domesticate. This is further supported by the fact that observations of feral var. *gasipaes* in secondary or disturbed vegetation occupy the same climatic space.

**TABLE 1** | Principal component analysis of bioclimatic variables for *Bactris gasipaes* and contributions of the bioclimatic variables to the first two components.

| Bioclimatic variable          | PC1          | PC2         |
|-------------------------------|--------------|-------------|
| Mean diurnal range            | -0.48        | -0.39       |
| Isothermality                 | <b>0.79</b>  | 0.04        |
| Temperature seasonality       | <b>-0.61</b> | -0.23       |
| Min. temp. coldest month      | 0.37         | <b>0.91</b> |
| Temp. annual range            | <b>-0.73</b> | -0.33       |
| Mean temp. wettest quarter    | -0.16        | <b>0.89</b> |
| Mean temp. driest quarter     | 0.13         | <b>0.97</b> |
| Mean temp. warmest quarter    | -0.14        | <b>0.95</b> |
| Mean temp. coldest quarter    | 0.10         | <b>0.97</b> |
| Precipitation wettest month   | <b>0.67</b>  | 0.04        |
| Precipitation driest month    | <b>0.92</b>  | -0.07       |
| Precipitation seasonality     | <b>-0.69</b> | 0.03        |
| Precipitation driest quarter  | <b>0.93</b>  | -0.07       |
| Precipitation warmest quarter | <b>0.63</b>  | -0.29       |
| Precipitation coldest quarter | <b>0.83</b>  | 0.18        |
| Explained variance            | 5.67         | 4.84        |
| % Total variance              | 38           | 32          |

Values in bold face contribute significantly to the principal component.



**FIGURE 3** | Principal component analysis of bioclimatic parameters for *Bactris gasipaes*. Projection of observations of var. *chichagui* types 1, 2, and 3, cultivated var. *gasipaes*, and feral or escaped specimens found in disturbed or secondary vegetation in the principal plane. Axis legends identify the main relations of principal components with original variables and percentage of explained variance.

### Distributions of var. *Chichagui* Types 1 and 2

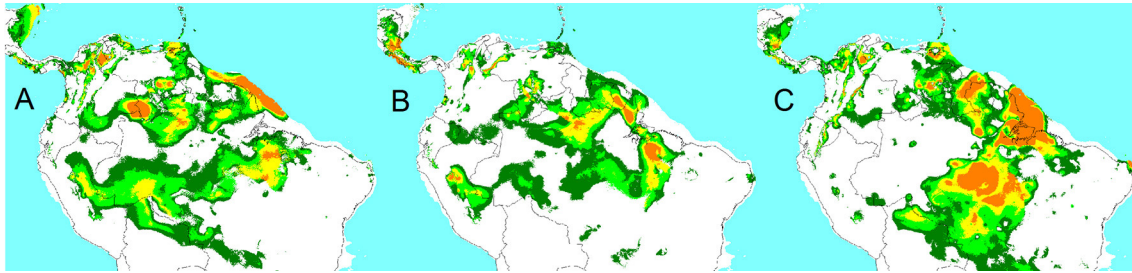
The limited overlap between the climatic envelopes of var. *chichagui* types 1 and 2 raises the question of their ecological differentiation. If any, it could be related to the topographical contrasts between their home regions, as suggested by their relative distribution in the principal plane: in southwestern Amazonia, conditions for type 1 get both warmer and wetter close to the Equator, while type 2, in the northern Andes, finds wetter

conditions at higher elevations, i.e., under cooler conditions. To assess whether such sources of climate variation may have resulted in significant ecoclimatic adaptation, we modeled the distribution of both wild types separately. Extrapolating the type 1 distribution to tropical South America (**Figure 2B**) allowed us to predict the main features of type 2's distribution, despite the topographic differences and low number of type 1 observations. The reciprocal extrapolation from the even smaller sample of type 2 data (**Figure 2C**), which does not discard the equatorial region east of Ecuador and along the Amazon river, is less convincing; however, it is still consistent with most observations of type 1. Interestingly, both models predict presence in the inter-Andean valleys of Colombia, the Caribbean coast, and the Andean foothills to the Orinoco basin. On the other hand, both **Figures 2B,C** lack specificity and indicate suitable regions where wild peach palm is absent, for example on the Guiana shield and in southeastern Brazil. Finally, the best distribution map results from the combination of types 1 and 2 (**Figure 2D**). This model is much more specific, with an excellent correspondence between observations and potential distribution. The current distribution of type 1 is well represented and explained with a considerable extension of suitable climates in the southern Amazon basin and to southeastern Brazil. The current distribution of type 2 appears less massive, but is equally well explained by the network of large Andean valleys and foothills in Colombia and Venezuela, as well as parts of their Caribbean coastal regions and around Lake Maracaibo in Venezuela. A third important and well separated suitable area exists in eastern Venezuela and Roraima; however, no wild peach palm has been reported there during the last 100 years of exploration.

This wild peach palm distribution model is highly consistent with our views on type 3 as the incipient domesticated, introduced by man and feralized in Ecuador, on both sides of the Andes. In some cases, as in southern Colombia and western Ecuador, feralization was favored by suitable climates, associated with more open vegetation (and less competition). In other cases, as on the eastern side of the Ecuadorian Andes, feralization was possible mostly in anthropogenic landscapes, as indicated by the observations of fully domesticated peach palms in the neighborhood, and their remnants in secondary vegetation. This is likely to be the case for Panamanian representatives of type 3 also, including the Azuero population mentioned by Hernández-Ugalde et al. (2011), as they appear in areas that are climatically unsuitable for wild peach palm.

The distribution model for the var. *chichagui* types 1 and 2 combination was projected for the Last Glacial Maximum climate, as predicted by the CCSM4, MIROC-ESM, and MPI-ESM-P climate models (**Figure 4**). The first two models project similar distributions of var. *chichagui* types 1 and 2, which remain separated during glacial periods, especially in western Amazonia. The third model is not consistent with the modern distribution of var. *chichagui* types 1 and 2, which casts doubts on its validity. Thus, these LGM simulations clearly suggest that populations of types 1 and 2 have long been separated, which helps explain the strong genetic differentiation (Hernández-Ugalde et al., 2011). They also





**FIGURE 4** | Maxent-generated potential distribution models based on Last Glacial Maximum (A) CCSM4, (B) MIROC-ESM, and (C) MPI-ESM-P estimated climatic conditions for *B. gasipaes* var. *chichagui* types 1 and 2. Color code as in **Figures 2B–D**.

suggest that there was suitable habitat for var. *chichagui* type 1 in southwestern Amazonia when humans arrived in the late Pleistocene.

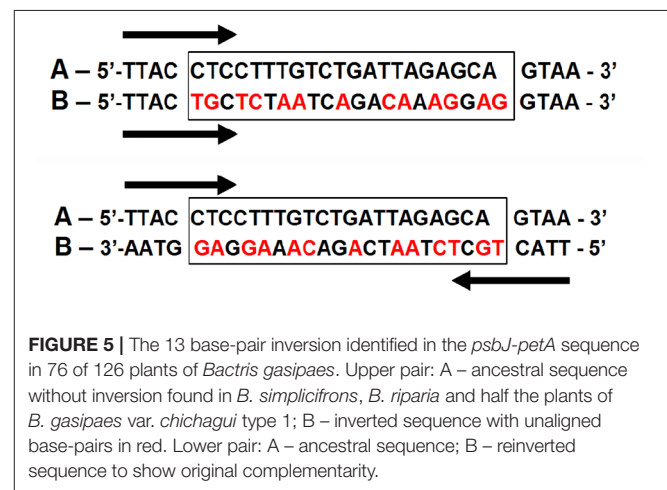
### Distributions of var. *Chichagui* Type 3 and var. *Gasipaes*

The number of field collections and herbarium samples that could be attributed to var. *chichagui* type 3 is quite small ( $n = 29$ ). The distribution model (**Figure 2E**) is consequently less reliable. Nonetheless, both the PCA (**Figure 3**) and the ecological niche model indicate that type 3's climatic envelope is slightly more humid than that of type 1, based on the expansion from Southwestern Amazonia northwards.

The modeled distribution of var. *gasipaes* (**Figure 2F**) represents reasonably well what is expected for cultivated peach palm from the literature and anecdotes. The highest probabilities are observed across central and western Amazonia, which is where a significant amount of collecting occurred in the late twentieth century (see also **Figure S1**) and where peach palm was most important at the time of European conquest (Patiño, 1963, 2002) (**Figure S3**). This area has much higher precipitation than the area where var. *chichagui* type 1 is distributed, even in the western part of its distribution. Also, var. *gasipaes*' niche encompasses var. *chichagui* type 3's niche, whereas the niche of wild types 1 and 2 (**Figure 2**) only encompasses type 3's niche in the southwestern Amazon, where types 1 and 3 are sympatric. Hence, peach palm's fundamental niche is much ampler than the realized niches of var. *chichagui* types 1 and 2, and type 3 shows the beginnings of this change.

### The Origin of Domestication Identified with a Chloroplast DNA Sequence

Only two of the 12 hypervariable sequences identified by Shaw et al. (2007) were variable in our study: *psbJ-petA* and *psaI-accD*. The *psbJ-petA* sequence had 1,040 base pairs, of which 26 were variable. The *psaI-accD* sequence had 622 base pairs, but was only useful for discriminating *B. simplicifrons* from *B. gasipaes*/*B. riparia* and was not used for further analysis. Seventy six of the 126 plants analyzed presented a 13 base-pair inversion in *psbJ-petA* at the same position (**Figure 5**), which distinguishes eastern and western Amazonian landraces and populations of peach palm. The population of var. *chichagui* type 1 near Rio Branco, Acre, Brazil, was polymorphic for this inversion, which



**FIGURE 5** | The 13 base-pair inversion identified in the *psbJ-petA* sequence in 76 of 126 plants of *Bactris gasipaes*. Upper pair: A – ancestral sequence without inversion found in *B. simplicifrons*, *B. riparia* and half the plants of *B. gasipaes* var. *chichagui* type 1; B – inverted sequence with unaligned base-pairs in red. Lower pair: A – ancestral sequence; B – reinverted sequence to show original complementarity.

has implications for the origin of domestication and subsequent dispersal of var. *gasipaes*. Neither var. *chichagui* nor *Bactris riparia* were discriminated from *B. gasipaes* var. *gasipaes* with the information in this sequence.

In this set of 126 plants, 12 haplotypes were identified and organized into a network with maximum parsimony analysis and the Median Joining algorithm (**Figure 6A**). Three haplotypes were very common, with one very common in southwestern to eastern Amazonia, and two very common in western Amazonia to Central America, with one mutational difference between the two common western haplotypes. Nine less common haplotypes were specific to a landrace (Juruá) or shared by a landrace (Putumayo), a population (upper Madeira River) or a species (*B. simplicifrons*). The eastern and western groups of haplotypes were differentiated by the inversion (**Figure 5**), with interesting exceptions. As mentioned, var. *chichagui* type 1 was polymorphic for the inversion. So was the Putumayo landrace, which has the largest distribution of the western landraces and extends eastwards along the Solimões River to contact with the Pará landrace in Central Amazonia (**Figure 1**). Hence, this polymorphism in the Putumayo landrace is probably due to introgression, unlike the polymorphism in var. *chichagui* type 1 in Acre. Because of small sample sizes and the conserved nature of the chloroplast sequences, estimates of chloroplast diversity were not very informative (**Table 2**), although var. *chichagui* type 1

**TABLE 2 |** Genetic diversity parameters estimated for the two chloroplast sequences for *Bactris gasipaes* var. *chichagui* types 1 and 3, and seven landraces and two non-designated populations of var. *gasipaes* in the Core Collection of the Peach palm Active Germplasm Bank, Manaus, Brazil, and two *Bactris* species used as outgroups.

| Category     | Landrace or population       | n   | h | #Subs | #InDels | S  | Hd     | $\pi$<br>( $\times 10^{-2}$ ) |
|--------------|------------------------------|-----|---|-------|---------|----|--------|-------------------------------|
| Wild         | var. <i>chichagui</i> 1      | 9   | 3 | 13    | 3       | 16 | 0.7778 | 0.00706                       |
| Incipient    | var. <i>chichagui</i> 3      | 6   | 1 | 0     | 0       | 0  | 0      | 0                             |
| Domesticated | var. <i>gasipaes</i>         | 103 | 9 | 14    | 5       | 19 | 0.7757 | 0.00731                       |
| Microcarpa   | Pará                         | 18  | 1 | 0     | 0       | 0  | 0      | 0                             |
|              | Upper Madeira River          | 9   | 2 | 0     | 3       | 3  | 0.2222 | 0                             |
|              | Juruá                        | 6   | 1 | 0     | 0       | 0  | 0      | 0                             |
| Mesocarpa    | Ucayali River                | 6   | 2 | 0     | 1       | 1  | 0.3333 | 0                             |
|              | Pampa Hermosa                | 12  | 1 | 0     | 0       | 0  | 0      | 0                             |
|              | Pastaza                      | 1   | 1 | 0     | 0       | 0  | 0      | 0                             |
|              | Cauca                        | 6   | 1 | 0     | 0       | 0  | 0      | 0                             |
|              | Utilis                       | 15  | 1 | 0     | 0       | 0  | 0      | 0                             |
| Macrocarpa   | Putumayo                     | 27  | 5 | 14    | 1       | 15 | 0.6667 | 0.00529                       |
|              | Vaupés                       | 3   | 1 | 0     | 0       | 0  | 0      | 0                             |
| Outgroups    | <i>Bactris riparia</i>       | 4   | 1 | 0     | 0       | 0  | 0      | 0                             |
|              | <i>Bactris simplicifrons</i> | 4   | 2 | 1     | 3       | 4  | 0.5000 | 0.00054                       |

*n*, number of plants analyzed; *h*, number of haplotypes; #Subs, number of nucleotide substitutions; #InDels, number of insertions and deletions; S, number of polymorphic sites; Hd, haplotypic diversity;  $\pi$ , nucleotypic diversity.

had approximately the same haplotypic diversity as var. *gasipaes*, principally because var. *chichagui* type 1 and the Putumayo landrace are polymorphic for the *psbJ-petA* inversion.

The phylogenetic tree estimated with Bayesian methods is similar to the haplotype network and the bootstrapped confidence values are high for all relationships (Figure 6B). As in the network, there is a clear separation between eastern and western groups. There is one important difference: the plants of var. *chichagui* 1 that grouped with the eastern Amazonian populations in the network are grouped with var. *chichagui* 3—the incipient domesticate—among the western Amazonian populations, even though half of them contain the inversion.

The phylogenetic network and tree support a single domestication event in southwestern Amazonia, probably in the upper Madeira River basin of modern Bolivia. The argument is most readily observed in the haplotype network (Figure 6A): the out-group (*B. simplicifrons*) is most closely related to the upper Madeira River populations, most of which have only one mutational difference from all of the other eastern Amazonian

populations. Although var. *chichagui* type 1 does not have exactly the same haplotype as the upper Madeira River populations, both populations occur in the same general area and share the ancestral *psbJ-petA* chloroplast sequence.

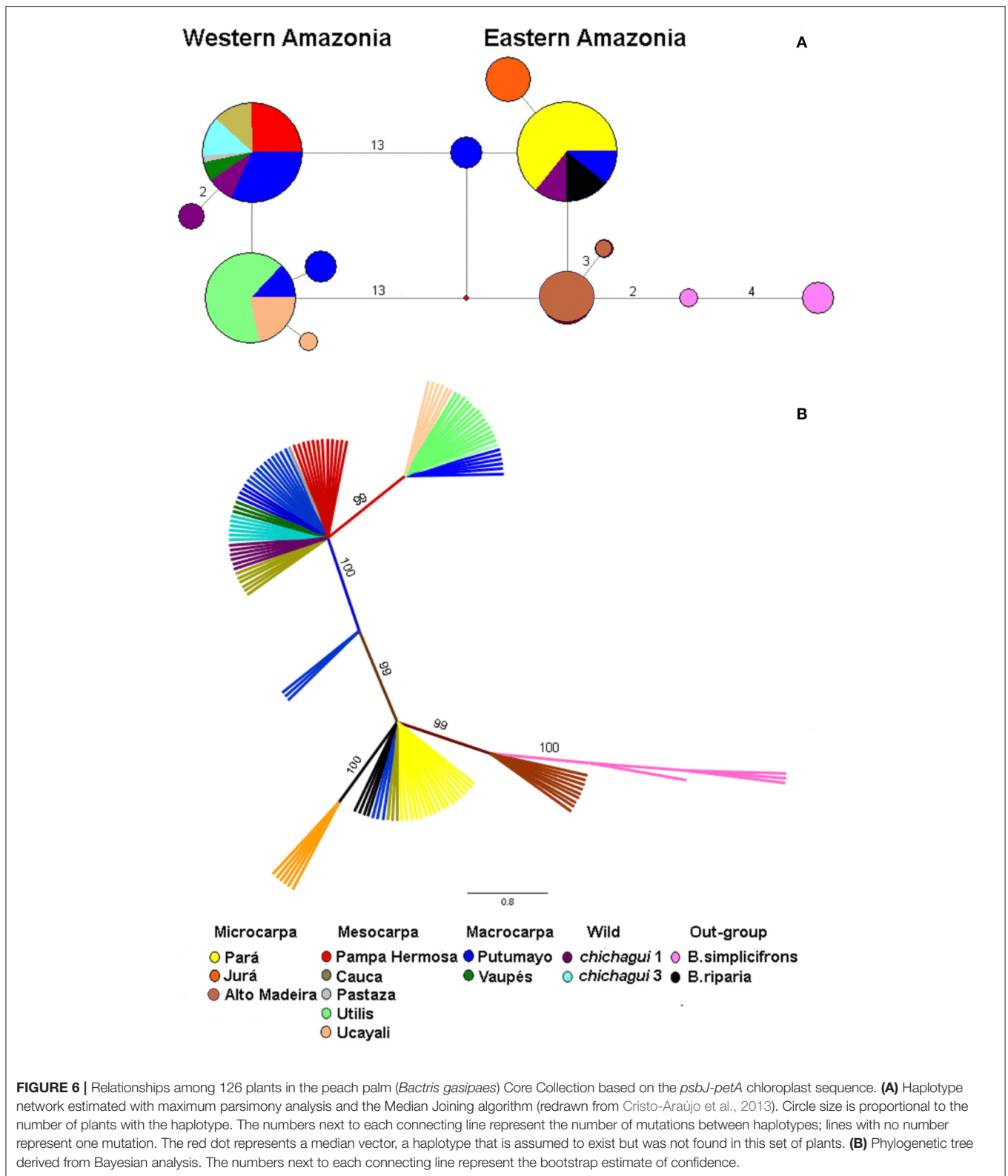
## Dispersal of the Landrace Complex Interpreted with Nuclear Microsatellites

The 17 most informative SSR loci among the 39 tested detected 302 alleles in the 173 plants analyzed, with a mean of 17.8 alleles per locus. The two accessions of var. *chichagui* type 1 had slightly lower heterozygosities than the two type 3 accessions, probably because the type 3 accessions are from different populations, both in sympatry with type 1, whereas the type 1 accessions are from the same population (Table 3). This explains the difference in inbreeding coefficients also. The highest values of observed heterozygosities are in landraces or undesignated populations within the distributions of var. *chichagui* types 1 and 3, which suggest introgression. The lowest heterozygosities occur in the two landraces furthest from the center of domestication in southwestern Amazonia: Utilis in Central America and Pará in eastern Amazonia.

Based on 173 plants of the Core Collection, the best grouping of accessions with the Structure program was found for  $K = 2$ , with interesting groupings at  $K = 3$  and 4 (Figure 7). At  $K = 2$ , the southwestern to eastern populations were distinguished from all other populations (Figure 8A). At  $K = 3$ , the Utilis landrace of Central America was discriminated from the other western populations (Figure 8B). At  $K = 4$ , the western Amazonian populations were divided into two groups (Figure 8C): a southern group containing the Pampa Hermosa and Juruá landraces and the Ucayali River populations; a northern group containing the two macrocarpa landraces (Putumayo and Vaupés) and the mesocarpa Cauca landrace of western Colombia. Note that the southern group is in sympatry with var. *chichagui* types 1 and 3, while the macrocarpa Putumayo has only minor areas of sympatry and the macrocarpa Vaupés has none (Figure 1). The origin of domestication encompasses the eastern half of the southern western group (populations 7, 8, 9) and the western part of the eastern group (populations 10, 11). At  $K = 10$  (data not shown) some landraces in Table 3 are relatively well distinguished, but others present considerable admixture with adjacent landraces and populations, as is already evident at  $K = 4$  (Figure 8C).

Although Structure offers robust simulations, it is based on the presuppositions of the Hardy-Weinberg equilibrium (Pritchard et al., 2000), many of which do not hold for small populations, especially for domesticated populations, nor for groups of gene bank accessions. Hence, we used spatial Analysis of Principal Components (sPCA), which does not rely on HWE presuppositions, to examine the relationships of the 156 var. *gasipaes* plants in the Core Collection with 17 SSR. The high variance and Moran's I recorded for the first three global principal components highlight the existence of global structure (data not shown) and is corroborated by the significance of the Monte-Carlo simulations ( $p < 0.001$ ). Due to the low variances and Moran's I index (data not shown), as well as the





weak significance of the Monte-Carlo simulations ( $p = 0.015$ ) within the set of local components, the local structure is not discussed.

The global structure using this SSR data presented good interpretation of the spatial distribution of genetic diversity and is similar to the Structure analysis (**Figure 8**). The first

**TABLE 3** | Genetic parameters estimated with 17 SSR loci for *Bactris gasipaes* var. *chichagui* types 1 and 3, and seven landraces and two non-designated populations of var. *gasipaes* in the Core Collection of the Peach palm Active Germplasm Bank, Manaus, Brazil.

| Category   | Landrace or population  | n   | A   | P  | Ho   | He   | Fis   |
|------------|-------------------------|-----|-----|----|------|------|-------|
| Wild       | var. <i>chichagui</i> 1 | 10  | 110 | 6  | 0.66 | 0.78 | 0.13  |
| Incipient  | var. <i>chichagui</i> 3 | 7   | 98  | 5  | 0.76 | 0.82 | 0.05  |
| Microcarpa | Pará                    | 34  | 161 | 9  | 0.57 | 0.77 | 0.28  |
|            | Upper Madeira River     | 12  | 124 | 4  | 0.65 | 0.83 | 0.20  |
|            | Juruá                   | 10  | 82  | 4  | 0.68 | 0.70 | -0.02 |
| Mesocarpa  | Ucayali River           | 10  | 111 | 4  | 0.66 | 0.80 | 0.16  |
|            | Pampa Hermosa           | 18  | 141 | 11 | 0.74 | 0.83 | 0.10  |
|            | Cauca                   | 8   | 84  | 2  | 0.64 | 0.73 | 0.09  |
|            | Utilis                  | 24  | 121 | 5  | 0.55 | 0.70 | 0.19  |
| Macrocarpa | Putumayo                | 35  | 184 | 10 | 0.70 | 0.83 | 0.15  |
|            | Vaupés                  | 5   | 67  | 2  | 0.73 | 0.76 | -0.02 |
| Overall    |                         | 173 | 302 |    | 0.66 | 0.76 | 0.21  |

*n*, number of plants analyzed; *A*, number of alleles; *P*, private alleles; *Ho*, observed heterozygosity; *He*, expected heterozygosity; *Fis*, intra-population inbreeding coefficient.

spatial principal component differentiated eastern Amazonia, as in the Structure analysis at  $K = 2$  (Figure 9A), the second differentiated Central America from the other western populations (Figure 9B), as in the Structure analysis at  $K = 3$ , and the third was less efficient at differentiating northern western Amazonia from southern western Amazonia (Figure 9C), probably because of the abundant gene flow. The spatial synthetic projection of the 3 global components (Figure 9D) suggests that eastern Amazonia is not as clearly related to southwestern Amazonia as in the Structure analysis, although this may be due to the lack of sampling along the middle and lower Madeira River. The relationship among populations in southwestern Amazonia is also much clearer than in the Structure analyses in that the upper Madeira and Ucayali Rivers are more clearly related.

Although the Core Collection is quite small and there appears to be abundant gene flow among these populations at different scales (Figures 8, 9), the Nei genetic distances among these groups are informative (Figure 10). The deepest divergence is between the southwestern to eastern populations, including var. *chichagui* 1, and all of the western populations, as in Figure 8A. However, the var. *chichagui* 1 population from Rio Branco, Acre, is not at the root of this group, suggesting that it is not the original source population for domestication. The western cluster contains the three other groups defined by Structure (Figure 8), with a very interesting organization. The cluster is rooted in var. *chichagui* type 3, the incipient domesticate, and has the microcarpa Juruá landrace and Ucayali River populations in sequence, followed by the mesocarpa Pampa Hermosa landrace.

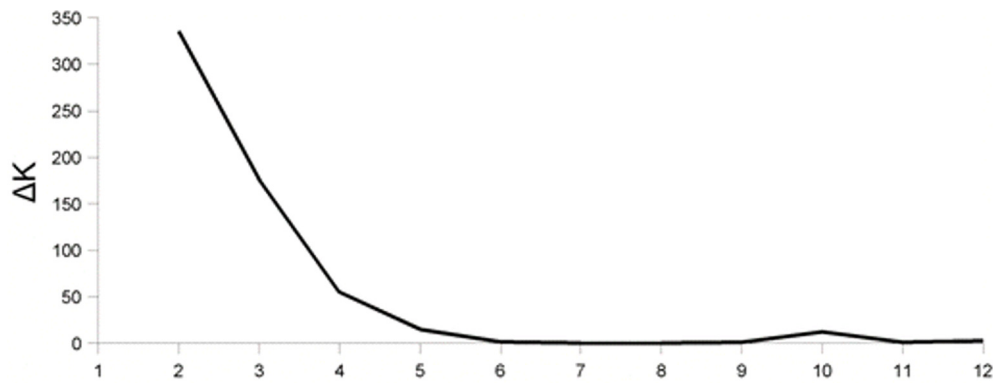
All of these populations are sympatric with var. *chichagui* type 1 also. The next cluster is derived from the previous, as expected by dispersal of domesticated types northward. The macrocarpa Putumayo landrace in western Amazonia is associated with the mesocarpa Cauca landrace in western Colombia, suggesting that there is gene flow over the Andes, perhaps in southern Colombia. Both Cauca and the western part of Putumayo are sympatric with the incipient domesticate (var. *chichagui* type 3). Also in the northwestern Amazonia Structure group (Figure 8C), the Vaupés landrace is the only one that is not sympatric with any wild populations, which may explain why it is the larger-fruited of the two macrocarpa landraces, since there is no introgression to slow response to selection. What is curious in this cluster is that the mesocarpa Utilis landrace of Central America is derived from the same lineage that gave rise to Vaupés, but this may only be an artifact of small sample sizes, although there is gene flow (or gene bank error?) visible in the Structure analyses (Figures 8B,C). Another possibility is that Utilis is derived from a different dispersal than Cauca, with the latter a dispersal over the Andes from the upper Putumayo River and the former a dispersal along the northeastern flank of the Andes and then into Central America.

AMOVA estimated that 87% of the total genetic variation accessed with these 17 SSR in the Core Collection is found within landraces and populations, while 13% is found among them. Other genetic divergence indices [Fst (0.13), Rst (0.19) and Gst (0.13)] agree with the AMOVA estimate of variation among populations. When comparing only the var. *gasipaes* accessions at  $K = 2$  (Figure 8A) and the deep dichotomy in the dendrogram (Figure 10), AMOVA estimated 92% within and 8% among landraces and populations.

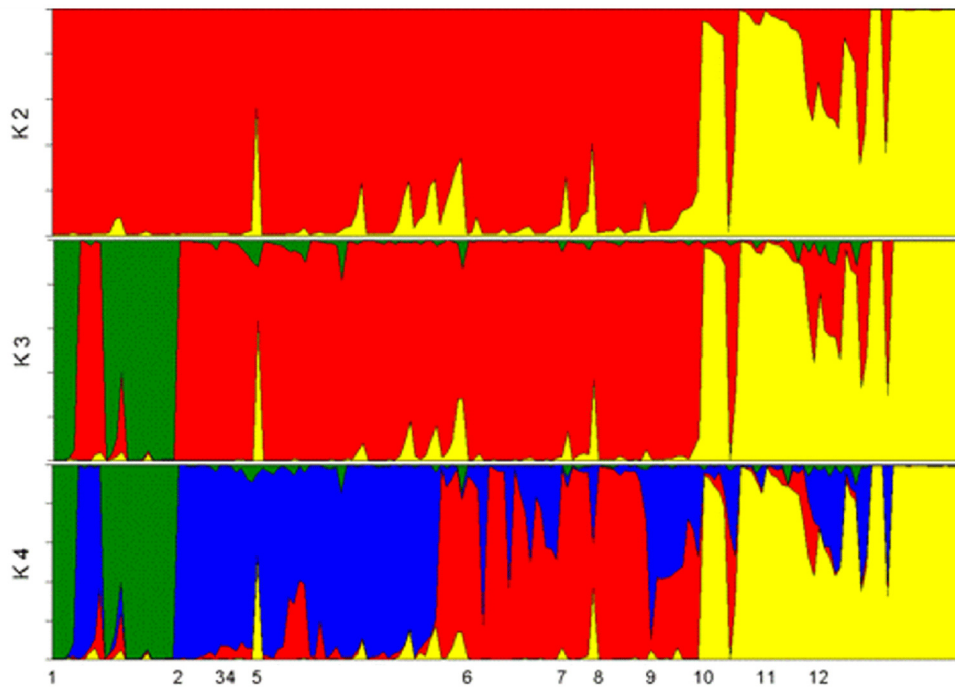
## DISCUSSION

### Distributions of Wild and Domesticated Peach Palm

Galluzzi et al. (2015) were the first to use ecological niche modeling with *Bactris gasipaes*, but our results cannot be compared with theirs for several reasons. Although they accepted the hypothesis that var. *chichagui* type 3 represents the incipient domesticate, they pooled all three types of var. *chichagui* into a 55-record “wild” sample, rather than maintain type 3 separate. They then added all samples of var. *gasipaes* that fall in the same climatic envelope (polygon in their Figure 2) to allow increased precision for LGM modeling, without considering that the niches of wild, feral, and cultivated plants cannot be interpreted in the same way. Furthermore, the resulting sample is biased, because their choice of a climatic envelope was determined partly by the two most extreme var. *chichagui* outliers, as well as the great majority of observations concentrated on the opposite convex side of the polygon in their PCA. This is important because distribution models are determined not only by the overall climatic space, but also by the distribution of observations within it.



**FIGURE 7 |** The most probable number of groups ( $K$ ) in the 173 plant data set from the Peach palm Core Collection analyzed with 17 SSR based on the *ad hoc*  $\Delta K$  of Evanno et al. (2005).



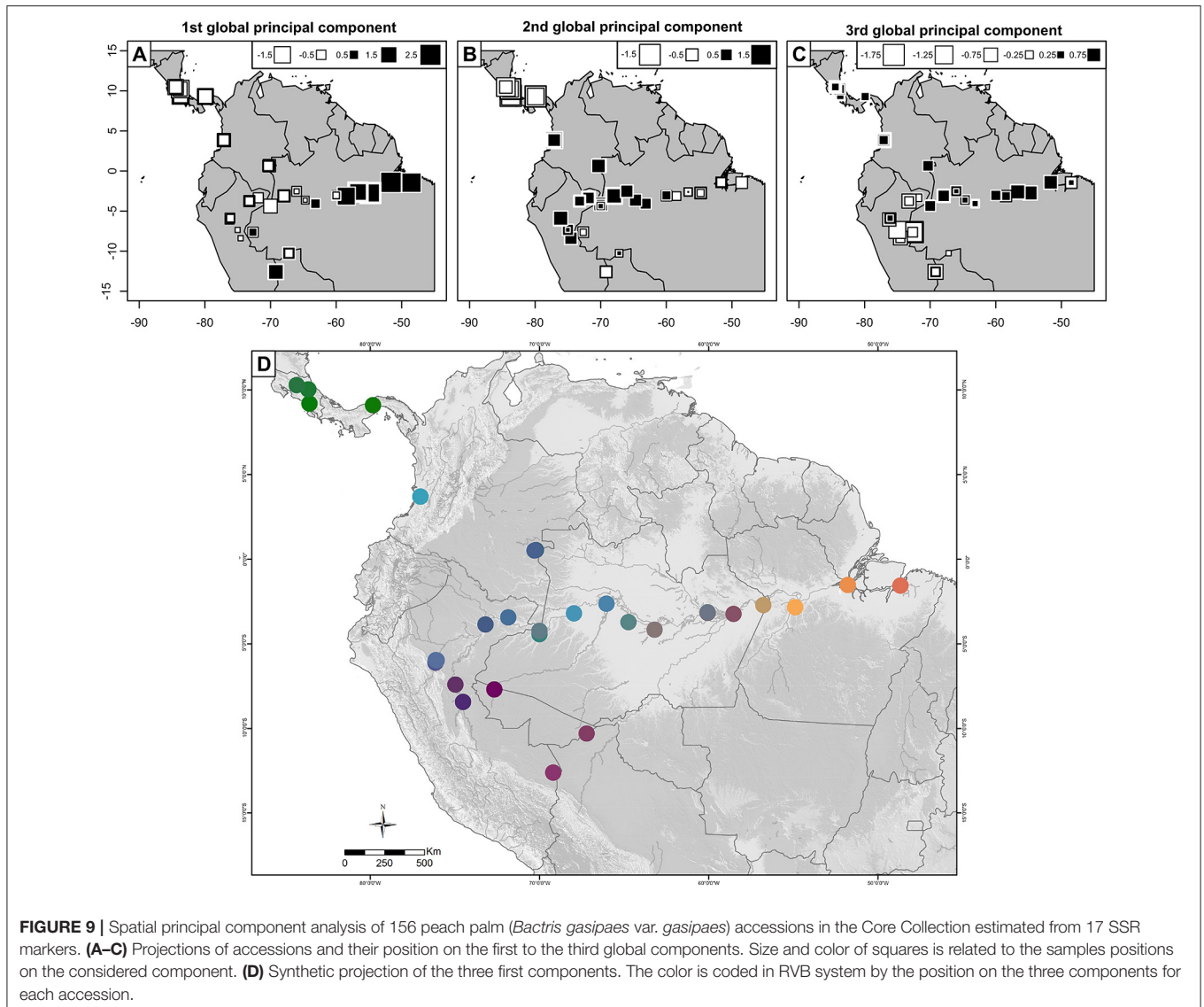
**FIGURE 8 |** Structure analysis of 173 peach palms (*Bactris gasipaes*) in 36 accessions of var. *gasipaes* and 4 accessions of var. *chichagui* in the Core Collection without prior information. Plants are organized by landrace and population from west (left) to east: 1, Utilis; 2, Cauca; 3, Pastaza; 4, Vaupés; 5, Putumayo; 6, Pampa Hermosa; 7, var. *chichagui* type 3; 8, Ucayali River; 9, Juruá; 10, upper Madeira River; 11, var. *chichagui* type 1; 12, Pará. **(A)** Group assignments for each palm for  $K = 2$  groups. **(B)** Group assignments for each palm for  $K = 3$  groups. **(C)** Group assignments for each palm for  $K = 4$  groups, with yellow representing the southwestern to eastern Amazonian populations, red the southern western Amazonian populations, blue the northern western Amazonian populations and green the Central American populations.

We followed a different approach, where wild peach palm's distribution was modeled from the observations of truly wild peach palm, i.e., a subsample including only observations of *B. gasipaes* var. *chichagui* that could be assigned to types 1 and 2, even though their number is modest. Observations of feral and escapes (respectively from var. *chichagui* and var. *gasipaes*), as well as cultivated populations, were only kept for purposes of comparison, using

them in the PCA comparative climatic characterization, and contrasting their realized distribution with that of wild peach palm.

Although the geographic distributions of var. *chichagui* types 1 and 2 are widely separated, the characterization and projections of their respective climatic spaces appear consistent with their infra-specific taxonomic status as two morphotypes of the same botanical variety (Henderson, 2000), with little apparent





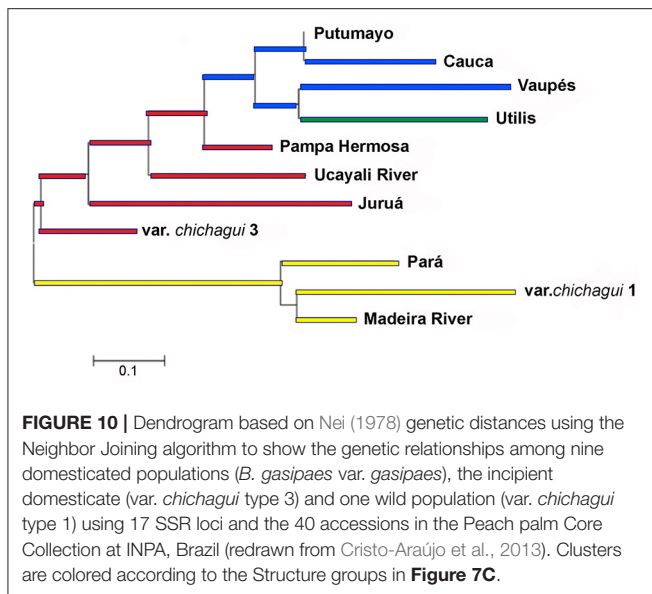
divergence in their ecologies. The fact that both models predict presence in the inter-Andean valleys of Colombia supports this hypothesis of limited ecological differentiation. Sample size is less of a problem when the truly wild types of var. *chichagui* are pooled in the analysis, for a total of 60 observations. The resulting model shows an excellent correspondence between these observations and their disjunct potential distribution (Henderson, 2000; Figure 29B, p. 71), and allows visualizing the more humid equatorial geographic barrier hampering gene flow between them.

Our extrapolation of the wild peach palm distribution during the LGM gave variable results according to the climatic model. In terms of suitable habitat across southern Amazonia, our CCSM4 and MIROC-ESM modeled distributions fit reasonably well into the ecotone between the evergreen broad-leaf and the deciduous broad-leaf forests modeled by Mayle et al. (2004) for the LGM, which reflects the ecological adaptation of type 1 (Clement et al., 2009b). The CCSM4-based modeled LGM distribution was the

most consistent with the modern wild peach palm distribution, showing the same potential separation of favorable habitats of types 1 and 2. Varela et al. (2015) caution that different general circulation models offer different predictions for the tropics, which explains why MPI-ESM-P produced such a divergent modeled distribution.

## The Origin of Domestication

The origin of domestication of cultivated populations of any species should be sought in the distribution of its wild populations. In the case of peach palm, var. *chichagui* types 1 and 2 have the smallest fruits and are considered truly wild. The molecular analyses that included type 2 concluded that it was not involved in the domestication of peach palm (Hernández-Ugalde et al., 2011), as did cladistic analysis of morphological traits (Ferreira, 1999). For reasons presented in the Introduction and re-enforced by the ecological niche models, type 3 is not wild. Hence, the origin of domestication is expected in the geographic



area of sympatry between type 1 and the incipient domesticate (Figure 1). Chloroplast sequences were used to examine this expectation.

Intergenic spacers in the chloroplast genome are important sources of information in plant systematics, but are often insufficiently variable at low taxonomic levels to differentiate populations (Shaw et al., 2005, 2007). This is clear in *Bactris*, where interspecific chloroplast variation is scarce in the *Bactris* species closest to *B. gasipaes* and even scarcer within the species, as Couvreur et al. (2007) failed to discriminate between var. *chichagui* and var. *gasipaes* with commonly used chloroplast sequences *trnD-trnT* and *trnQ-rps16*, and a sequence that they designed (*psbC-trnfM*), all located in non-coding regions. We also failed to discriminate var. *chichagui* from var. *gasipaes* or *B. gasipaes* from *B. riparia* with the *psbJ-petA* and *psaI-accD* sequences, although we did find variation in *psbJ-petA*, but this sequence is less variable than *trnD-trnT* (1,066 base pairs with 36 variable) and *trnQ-rps16* (1,046 base pairs with 47 variable) (Couvreur et al., 2007). The 13 bp inversion that we found (Figure 5) is intermediate in size between the minute (4 bp) and a middle-sized (20 bp) inversions in *Bactris trnD-trnT* (Couvreur et al., 2007), although neither allowed discrimination within the *Bactris gasipaes-riparia* complex.

Both the haplotype network (Figure 6A) and the tree (Figure 6B) show a clear separation between eastern and western populations of var. *gasipaes*, reflecting the deep divergence found in all previous molecular analyses (Rojas-Vargas et al., 1999; Rodrigues et al., 2005; Cristo-Araújo et al., 2010; Hernández-Ugalde et al., 2011). Since this analysis is with a single chloroplast sequence, it is evident that this inversion does not explain the deep divergence observed with nuclear markers, but it does provide a parallel marker.

The network and the tree also support a single domestication event in southwestern Amazonia, probably in the upper Madeira River basin of modern Bolivia, which had already

been proposed as the center of domestication (Huber, 1904; Clement, 1995; Rodrigues et al., 2005; Cristo-Araújo et al., 2010, 2013; Galluzzi et al., 2015). This is also one of the areas identified by Mora-Urpí (1993, 1999) and Hernández-Ugalde et al. (2011). This chloroplast analysis does not provide support for additional domestication events outside of southwestern Amazonia, contrary to the hypotheses of Mora-Urpí (1993, 1999), Morcote-Rios and Bernal (2001), and Hernández-Ugalde et al. (2011), nor the idea of secondary domestications suggested by Galluzzi et al. (2015), although this may be because of the small amount of variation found to date. These unsupported hypotheses all depend upon the distribution of var. *chichagui* type 3, the incipient domesticate.

## Dispersal of the Landrace Complex

In domesticated peach palm, two dispersals out of the center of domestication in southwestern Amazonia were hypothesized (Rodrigues et al., 2005), one down the Ucayali River into western Amazonia and beyond, and one down the Madeira River into eastern Amazonia, and should exhibit these trends when examined with neutral molecular markers. The highest values of observed heterozygosity are in landraces or undesignated populations within the distributions of var. *chichagui* types 1 and 3 (Table 3), which suggests introgression (Couvreur et al., 2006; Hernández-Ugalde et al., 2011). The lowest heterozygosity values occur in the two landraces at the extremes of the two hypothesized dispersals: Utilis in Central America at the end of the western dispersal and Pará in eastern Amazonia at the end of the eastern dispersal. The low heterozygosity in the Utilis landrace is surprising, given the existence of type 3 populations in the region and observed introgression (Hernández-Ugalde et al., 2011), suggesting that Utilis represents an independent dispersal and not an *in situ* development from the local var. *chichagui* type 3 populations.

While validation of morphometrically defined landraces has been a preoccupation of the molecular genetic analyses in the Brazilian germplasm bank (Sousa et al., 2001; Clement et al., 2002; Rodrigues et al., 2005; Cristo-Araújo et al., 2010; Santos et al., 2011), these analyses often had insufficient or unbalanced numbers of each population to work with. When sufficient numbers were available, they validated some landraces and did not validate others, specifically the Guatuso and Tuira landraces in Central America, and the Solimões landrace in central-western Amazonia (Rodrigues et al., 2005). Hence, Structure was used to study the relationships among accessions in the Core Collection.

$K = 2$  (Figure 8A) identified the deep divergence detected in previous molecular analyses (Rojas-Vargas et al., 1999; Rodrigues et al., 2005; Cristo-Araújo et al., 2010; Hernández-Ugalde et al., 2011). In  $K = 3$ , the Utilis landrace of Central America was distinguished from the other western populations (Figure 8B), probably because allelic richness is lower (Table 1) and some alleles are locally common (Hernández-Ugalde et al., 2011), although this was not detected by Galluzzi et al. (2015) in their reanalysis of Hernández-Ugalde et al.'s dataset. At  $K = 4$  a considerable amount of admixture was detected (Figure 8C), certainly the reason for the poor discrimination between the groups.

The  $K = 4$  grouping also identifies either long distance dispersal events or germplasm bank errors; the latter have been detected before in the INPA germplasm bank with molecular analyses (Sousa et al., 2001; Rodrigues et al., 2005; Cristo-Araújo et al., 2010). Two accessions from the relatively spineless Guatuso populations of the Utilis landrace in Central America are assigned to the northern western Amazonia group, suggesting that spinelessness may not have been selected *in situ* but may be due to long distance dispersal, and the accessions from Coari, previously classified with the Putumayo landrace (Rodrigues et al., 2005), are assigned to the southern western Amazonia group, some 500 km to the west. Seed exchange networks had previously been identified within the Pampa Hermosa landrace in southern western Amazonia (Adin et al., 2004), as well as long distance gene flow between the southern and the northern western Amazonian groups (Cole et al., 2007), so this admixture is not surprising and may not be due to germplasm bank error.

Overall, the Structure analyses confirm part of the landrace hierarchy proposed originally by Mora-Urpí and Clement (1988) and validated by previous molecular analyses (Rodrigues et al., 2005; Cristo-Araújo et al., 2010). They also confirm the commonness of both long and middle distance gene flow and introgression reported by various authors (Adin et al., 2004; Couvreur et al., 2006; Cole et al., 2007; Hernández-Ugalde et al., 2011). The fact that they do not fully validate the landrace hierarchy can be attributed to the design of the Core Collection (Cristo-Araújo et al., 2015), since this was not designed primarily to study the origin and dispersal of peach palm, but to support the management of peach palm germplasm at INPA.

The spatial analysis of principal components (Figure 9) generally agreed with the Structure analyses (Figure 8), but also suggested a very interesting relationship among the upper Madeira River populations and the Ucayali River populations, which was not evident in the Structure analyses. This relationship is the primary region of sympatry between var. *chichagui* types 1 and 3 (Figure 1), and where var. *gasipaes* microcarpa populations have the smallest fruit. In fact, the headwater tributaries of the Ucayali River and those of the Madre de Dios River, the major northern tributary of the upper Madeira River, are quite close in southern Peru, allowing relatively easy human passage. It follows that the upper Ucayali River basin cannot be ruled out as part of the center of origin of domestication. Further prospection of peach palm in the upper parts of both basins will allow better resolution of these analyses.

The Neighbor-joining dendrogram of Nei's genetic distances (Nei, 1978) among the landraces and populations of the Core Collection (Table 3, Figure 10) is quite similar to all previous dendrograms based on molecular analyses. The eastern cluster contains the upper Madeira River populations and the Pará landrace, as observed by Rojas-Vargas et al. (1999) and Hernández-Ugalde et al. (2011), and in agreement with morphological similarities (Mora-Urpí, 1999), since both have microcarpa fruit types. This cluster is associated with var. *chichagui* type 1, as observed by Rodrigues et al. (2005) and Hernández-Ugalde et al. (2011), but is not rooted in the Rio Branco population of type 1. Hence, the exact origin of

domestication remains to be identified, although the general region of origin is clear.

As expected, the western cluster in Figure 10 is similar to that reported previously (Rodrigues et al., 2005; Cristo-Araújo et al., 2010), since it uses some of the same accessions, and is quite different from Hernández-Ugalde et al. (2011), whose analyses identified the divergence between the Cauca and Utilis landraces and the western Amazonian landraces at a lower level in the dendrogram. However, given the numerous additional analyses included here (Table 1, Figures 6–8), plus the reinterpretation of var. *chichagui* type 3 as the incipient domesticate, this interpretation of Figure 10 appears to represent a more robust hypothesis of the origin and dispersal of domesticated peach palm.

The two dispersals proposed here also resulted in two quite different fruit types with different cultural importance. The western dispersal down the Ucayali soon generated starchy fruit, possibly quite early, since Mora-Urpí (1984) observed that very starchy microcarpa fruit were common in Pucallpa and Contamana, along the Ucayali River in Peru. Starchy fruit are easily fermented, much like sweet manioc (*Manihot esculenta*) or maize (*Zea mays*) (Patiño, 1963, 1992, 2002). Because starch is much less energy intensive than oil, even unconscious selection for starchy fruit quickly results in increases in fruit size (Clement et al., 2009a), resulting in the mesocarpa Pampa Hermosa and Tigre landraces in central Peru. As the dispersal of this type of fruit continued down the Ucayali and Amazonas into northwestern Amazonia, the cultivated peach palms were taken out of sympatry with type 1 and the previously distributed incipient domesticate (type 3), and the very starchy macrocarpa fruits of the Putumayo and Vaupés landraces could appear. However, the oldest archeological record in Colombian Amazonia is the Abejas site, along the Caquetá River, with pollen dated to 1,535 BP (Morcote-Rios and Bernal, 2001), which suggests that this dispersal may have been rather late. Throughout central-western and northwestern Amazonia peach palm was cultivated both in homegardens and in swiddens, yielding large amounts of starchy fruit for fermentation that became the centerpiece of yearly harvest festivals (Patiño, 1992). As the dispersal continued northwestward into Central America, the cultivated populations were again sympatric with the incipient domesticate (type 3) and the starchy mesocarpa Cauca and Utilis landraces appeared. During the conquest of Panamá and Costa Rica, European adventurers felled tens of thousands of peach palms in the Sixaola River valley in order to subdue the native peoples, which resulted in the first court case of the Spanish crown against a group of conquistadors and is the reason that such dramatic numbers of palms are known to have existed (Patiño, 1963). Given the enormous numbers of palms involved, we can assume that peach palm was as important in southern Central America as in western and northwestern Amazonia.

The eastern dispersal appears to have been quite different, since the fruit retained considerable quantities of oil and was never selected even to mesocarpa size, even though the majority of the dispersal was outside of the distribution of the wild type. Oily fruit do not ferment well and there are no early



historical records of harvest festivals with abundant fermented peach palm, as there are in western Amazonia. Bates (1962) observed fruits typical of the Pará landrace during his trip along the Amazon River, and commented that they increased in size once he started up the Solimões River, confirming the confluence of the two dispersals in Central Amazonia mentioned above. It is even possible that Bates' observations represent the final expansion of the eastern dispersal, since Patiño's (1963) analysis of the earliest European reports from eastern Amazonia seldom mention peach palm. Patiño's (1963) map (p. 131) supports this supposition of a late expansion into eastern Amazonia. At the Hatahara archeological site, lower Solimões River, 20 km from its confluence with the Negro River to form the Amazon River, *Bactris-Astrocaryum* phytoliths increase in number continually from the lowest levels (~1,000 BP) to the time of conquest (500 BP) in terra preta middens; since it is the only cultivated palm in Central Amazonia, these phytoliths may represent peach palm (Bozarth et al., 2009), although managed *Astrocaryum aculeatum* cannot be ruled out. It follows that the eastern dispersal may have started later than the western dispersal.

## CONCLUSIONS

The argument developed here starts from the reanalysis of Henderson's (2000) taxonomic revision of *Bactris gasipaes* and is based on expectations that arise from the domestication process. Precisely because domestication is a process, gradual changes from the wild type to the domestication continuum of incipient to semi-domesticated to domesticated are expected, and a species with abundant domesticated populations, such as peach palm, is expected to contain populations along the whole continuum. The reinterpretation of var. *chichagui* type 3 from wild to the incipient domesticate fills the gap in the continuum that had been lacking. The ecological niche modeling and climatic PCA suggest that var. *chichagui* types 1 and 2 do not differ significantly in their climatic space, which contrasts with the wider adaptation of type 3 and the even wider adaptation of var. *gasipaes*. The incipient domesticate (type 3) was able to maintain populations in anthropized forests, under more humid conditions as it was dispersed from southwestern Amazonia into western Amazonia and beyond. The ecological niche models of wild peach palm's potential distribution during the Last Glacial Maximum suggest that it was present in southwestern Amazonia when people arrived. This identification of the incipient domesticate also narrowed the search for the origin of its domestication to southwestern Amazonia, where it is sympatric with var. *chichagui* type 1, as expected if it originated there.

Although only one of the 12 chloroplast sequences tested was informative within peach palm, the inversion in *psbJ-petA* paralleled the deep divergence in nuclear molecular genetic variability observed in all previous analyses. The patterns observed in the geographic distribution of nuclear genetic diversity are those expected during dispersal from the origin of domestication to peach palm's present distribution throughout the lowland Neotropics from Bolivia to Nicaragua, even though the INPA Core Collection does not have samples from numerous

areas in this ample distribution. It does, however, have enough samples in strategic locations to confirm two major dispersals: the first out of southwestern Amazonia down the Ucayali River into western Amazonia and beyond, which resulted in the complex landrace hierarchy of that region and the very starchy fruit that could be fermented and become important to pre-conquest indigenous cultures; the second out of the same region down the Madeira River into eastern Amazonia, which did not result in a complex landrace hierarchy, perhaps because the starchy-oily microcarpa fruit were used more for snacks than as a starchy staple.

The current analysis obviously has limitations, principally due to the modest number of samples for such an ample distribution, even though these were carefully chosen to be representative of the samples available in the Brazilian peach palm collection via the creation of the Core Collection. Further studies of the two wild types and the incipient domesticate are needed. A revision of the infra-specific relationships and nomenclature of *Bactris gasipaes* will be required to assist botanists and plant breeders with this new proposal.

## MATERIALS AND METHODS

### Core Collection for Genetic Analysis and Niche Modeling

We used 174 plants from 36 accessions (3–5 plants per accession) of domesticated peach palm (var. *gasipaes*) and four accessions (2–5 plants per accession) of wild peach palm (var. *chichagui* types 1 and 3). These accessions belong to the Core Collection designed by Cristo-Araújo et al. (2015) within the Peach palm Active Germplasm Bank, maintained by the Instituto Nacional de Pesquisas da Amazônia (INPA), located at km 38 of the BR-174 highway, Manaus, Amazonas, Brazil (latitude 2° 38' 34.28" S and longitude 60° 2' 33.63" W). An accession is the progeny obtained from seed of a single open-pollinated bunch from a palm sampled in a traditional farmer's property. All sampling was done with prior informed consent before the Convention of Biological Diversity. Five samples each of *Bactris riparia*, a very close relative of peach palm, and *B. simplicifrons*, a distant relative (Henderson, 2000; Couvreur et al., 2007), were also genetically characterized to serve as out-groups.

### Geographic Coordinates for Niche Modeling

In addition to the geo-referenced samples of the Core Collection, we used some *B. gasipaes* var. *gasipaes* from the Peach palm Active Germplasm Bank and downloaded geo-referenced occurrence records of var. *gasipaes* and var. *chichagui* from the Global Biodiversity Information Facility (GBIF) data-portal (<http://data.gbif.org>) on 21 August 2013 (see Supplementary Material 2.1). The Instituto de Ciencias Naturales, Bogotá, the Herbario Nacional de Bolivia and the Herbarium of the University of Aarhus kindly supplied additional coordinates and/or information to confirm the type of var. *chichagui* contained in the GBIF database. Only the samples that could reasonably be classified to a specific type of var. *chichagui* were

used for wild peach palm ENM. We also used geo-referenced samples reported in Clement et al. (2009b) for which we have personal information, i.e., we did not use possible var. *chichagui* from the RADAM database, because these could not be identified as to type. The data set selected for wild peach palm niche modeling includes 38 type 1 and 22 type 2 (Supplementary Material 2.1). The information gathered on other observations of *B. gasipaes* was used for comparison with feral and cultivated materials, including 29 type 3, 202 var. *gasipaes* dataset, as well as 25 observations involving feral peach palms that could not be assigned to a particular morphotype, and are probable escapes from cultivation. All geographic coordinates were assigned or verified, using the Geonames gazetteer (<http://www.geonames.org/>) and Google Earth.

## Ecological Niche Characterization and Modeling

For the environmental layer input, we used 19 bioclimatic variables at a 2'30" grid resolution (corresponding roughly to 4.4 × 4.6 km at the Equator), for current conditions (~1950–2000) (<http://www.worldclim.org/current>) (Hijmans et al., 2005) and for past conditions—Last glacial maximum (LGM; ~21,000 years BP) (<http://www.worldclim.org/past> from <http://pmip2.lscce.ipsl.fr/>). For the LGM models, we used CCSM4.0 (Community Climate System Model), MIROC-ESM (Model for Interdisciplinary Research on Climate-Earth Model System), and MPI-ESM-P (Max Planck Institute-Earth Model System).

The geo-referenced samples were used to model the geographic area that would be most likely to meet the climatic requirements of wild and cultivated peach palm (Phillips et al., 2006). The Maxent program identifies potential distribution areas on the basis of their similarity in climatic conditions compared to those at the sites where the species has already been observed, hence modeling where conditions are suitable for their survival. It infers the probability distribution of maximum entropy (i.e., closest to uniform) subject to the constraint that the expected value of each environmental variable (or its transform and/or interactions) under this estimated distribution matches its empirical average (Phillips et al., 2006; Thomas et al., 2012). To model the distribution of the realized niche of wild peach palm, Maxent was run on the following subsamples: (1) var. *chichagui* type 1 (Bioclim coverage 5–17°S, 49–76°W), (2) var. *chichagui* type 2 (2–12°N, 70–76°W), and (3) var. *chichagui* types 1 and 2 (18°S–16°N, 48–86°W). It was also run on the whole sample (18°S–16°N, 48–86°W), dominated by cultivated peach palm data, to approach the distribution of the fundamental niche (Supplementary Material 1.2). A logistic threshold value equivalent to the 10th percentile training presence was retained to separate climatically favorable areas from marginally fit areas. Thresholds of 33 and 67% training presence were used to discriminate “very good” and “excellent” climates for the production of comparable climate suitability maps. For the LGM distribution models, we used the combined var. *chichagui* types 1 and 2 sample, and excluded Bioclim variables 14 and 15 that show a high level of discrepancy between LGM climate models (Varela et al., 2015).

We performed a principal component analysis (PCA) on the whole dataset, to characterize and compare the climatic envelopes of wild, feral, and cultivated peach palms, retaining those variables that contributed to the Maxent model and applying a varimax normalized rotation. The maximal temperature from the warmest month (Bio5) was discarded as it can be deduced from the minimal temperature of the coldest month (Bio6) and the annual range (Bio7). The different categories of peach palm populations were then plotted onto the principal plane.

## Analysis Using cpDNA

Fourteen chloroplast sequences (Shaw et al., 2007) were tested and two were informative (*psbJ-petA* and *psaI-accD*), but only *psbJ-petA* was used, because *psaI-accD* was only useful for discriminating *B. simplicifrons* from the *B. gasipaes/riparia* complex. The alignment of sequences obtained in both directions (forward and reverse) and the creation of the consensus sequence of each pair were performed using BioEdit 7.0.5.3 (Hall, 1999). Bayesian phylogenetic reconstructions were conducted in MrBayes v2.01 (Huelsenbeck and Ronquist, 2001), which uses the Metropolis-coupled Markov Chain Monte Carlo (MCMC) method to estimate the posterior probability distribution (Schmidt, 2009). Two runs with 10 million generations applied substitution models determined for each partition in MrModeltest v.2.2 (Nylander <https://www.abc.se/~nylander/>). In order to estimate posterior probabilities, 25% of the trees were discarded as a burn-in stage, observing when average standard deviation of split frequency (ASDSF) values dropped below 0.01. Phylogenetic Network v.4.5.1.6, developed to estimate phylogenetic networks with maximum parsimony, was used to build a network of haplotypes with the Median Joining algorithm (Bandelt et al., 1999). This method combines features of Kruskal's algorithm that finds the best tree while favoring short connections, the heuristic algorithm of maximum parsimony of Farris, and adds vertices called median vectors that represent extinct or un-sampled haplotypes in populations (Bandelt et al., 1999). Chloroplast genetic diversity across taxa was estimated with DNAsp 5 (Librado and Rozas, 2009).

## Analysis Using SSR Markers

We tested 39 SSR loci developed for peach palm (Martinez et al., 2002; Billotte et al., 2004; Rodrigues et al., 2004) for selection of loci with clear and informative amplification profiles. PCR reactions were performed according to Rodrigues et al. (2004). The SSR data is in Supplementary Material 2.2. Allele frequencies and private alleles of all loci were calculated using the Convert program (Glaubitz, 2004). We estimated the genetic distances of Nei (1978) between landraces defined by Rodrigues et al. (2005).

## Bayesian Analysis of Population Structure

Pritchard et al. (2000) developed a Bayesian method (implemented in the program Structure) to model the number of groups (K) of individuals based on their multi-locus genotypes. One advantage of this method is that populations need not be defined *a priori*, but will be identified by the data generated with SSR markers. This is very important when studying samples from germplasm banks, which often do not contain samples

that can be considered representative of populations. Hence, it is even more important when analyzing core collections. The parameters used were: burn-in of 10,000 permutations, Markov Chain Monte Carlo (MCMC) with 100,000 permutations, the admixture ancestry model, where each individual can have more than one ancestral population, and independent allele frequencies ( $\lambda = 1$ ). The best K was identified by LNP values (D) and  $\Delta K$ , following Evanno et al. (2005), from 15 simulations for each possible K from 1 to 12 (the number of hypothesized landraces and non-designated populations).

## Spatial Principal Components Analysis

We used spatial principal components analysis (sPCA) (Jombart et al., 2008), implemented in adegenet 1.3–2 (Jombart, 2008) with R (R Development Core Team, 2011) to visualize continental-scale gradients in the genetic diversity of peach palm. This method uses a matrix with allele frequencies of genotypes and a spatial weighting matrix containing measurements of spatial proximity among entities based on a connection network to produce scores that summarize the spatial structure and the genetic variability among groups of individuals across geographic space (Jombart et al., 2008). Various types of connection networks are available in adegenet. Given the continental scale of our study, we used the Gabriel graph network, because (1) it avoids unlikely connections (e.g., between eastern Amazonia and Central America), unlike Delaunay triangulation, and (2) allows possible connections at regional scale (e.g., southern and northern Western Amazonia) unlike relative neighbors network.

Moran's I is used to measure spatial autocorrelation in allele frequency values of samples. More specifically, sPCA optimizes the product of the variance of a few synthetic variables and of Moran's I, and generates two sets of axes: one with positive eigenvalues and the other with negative eigenvalues. Positive eigenvalues correspond to global structures, while negative eigenvalues are indicative of local patterns (Jombart et al., 2008). Abrupt decreases in both sets of eigenvalues indicate that global or local structure should be interpreted. The significance of "global" and/or "local" spatial structure was assessed using monte-carlo simulations implemented in global.rtest() and local.rtest() functions respectively, from adegenet 1.3–2 (Jombart et al., 2008). We proceeded with 9999 permutations per test. After having identified local and/or global structure and selected the

number of components to consider, samples position on each component was plotted onto the geographic space. As several principal components were retained we also used the colorplot() function from adegenet 1.3–2 (Jombart et al., 2008) to summarize spatial gradient of peach palm genetic diversity. This function uses the score of 1 to 3 components to compose a color per sample based on the RGB (Red Green Blue) system.

## AUTHOR CONTRIBUTIONS

CC designed research, obtained funding, contributed to analysis, wrote the article. MC-A, VMR, and DP-R executed research, contributed to analysis, wrote the article. GC and RL contributed to analysis, wrote the article.

## ACKNOWLEDGMENTS

We thank the Brazilian National Research Council (CNPq), Universal projects 47.6189/2003-9 and 47.5219/2006-6, and CT-Amazonia project 57.5588/2008-0, for financial support, the Foundation for the Support of Science in the State of Amazonas (FAPEAM) (project 062.00831/2015) and the Agence Inter-établissements de Recherche pour le Développement (AIRD) bilateral cooperation agreement, project 2074/2011, for financial support and training, the National Research Institute for Amazonia (INPA) for a PCI scholarship for MC-A (2008–2013), Rodrigo Bernal, Universidad Nacional de Colombia, Birgitte Bergmann and Henrik Balslev, Herbarium AAU, Aarhus University, Viviana Vargas and Monica Moraes, Herbario Nacional de Bolivia de la Universidad Mayor de San Andrés, for information on var. *chichagui*, Yves Vigouroux, Institut de Recherche pour le Développement (IRD), and Evert Thomas, Bioversity International, for useful criticism and suggestions on earlier versions of this paper, and CNPq for a research fellowship for CC.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fevo.2017.00148/full#supplementary-material>

## REFERENCES

- Adin, A., Weber, J. C., Sotelo Montes, C., Vidaurre, H., Vosman, B., and Smulders, M. J. M. (2004). Genetic differentiation and trade among populations of peach palm (*Bactris gasipaes* Kunth) in the Peruvian Amazon - implications for genetic resource management. *Theor. Appl. Genet.* 108, 1564–1573. doi: 10.1007/s00122-003-1581-9
- Avise, J. C. (2004). *Molecular Markers, Natural History, and Evolution*. Sunderland, MA: Sinauer Associates.
- Bandelt, H.-J., Forster, P., and Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48. doi: 10.1093/oxfordjournals.molbev.a206036
- Bates, H. (1962). *The Naturalist on the River Amazons [Reprinted from the 2nd Edn., 1864, John Murray, London]*. Berkeley, CA: University of California Press.
- Billotte, N., Couvreur, T., Marseillac, N., Brottier, P., Perthuis, B., Vallejo, M., et al. (2004). A new set of microsatellite markers for the peach palm (*Bactris gasipaes* Kunth); characterization and across-taxa utility within the tribe Cocoeae. *Mol. Ecol. Notes* 4, 580–582. doi: 10.1111/j.1471-8286.2004.00741.x
- Bozarth, S. R., Price, K. P., Woods, W., Neves, E. G., and Rebellato, L. (2009). "Phytoliths and terra preta: the Hatahara site example," in *Amazonian Dark Earths: Wim Sombroek's Vision*, eds W. Woods, W. G. Teixeira, J. Lehmann, C. Steiner, A. M. WinklerPrins, and L. Rebellato (Heidelberg: Springer), 85–98. doi: 10.1007/978-1-4020-9031-8\_4



- Clement, C. R. (1988). "Domestication of the pejobaye palm (*Bactris gasipaes*): past and present," in *The Palm - Tree of Life. Biology, Utilization and Conservation*, ed M. J. Balick (New York, NY: The New York Botanical Garden), 155–174.
- Clement, C. R. (1995). "Pejobaye (*Bactris gasipaes*)," in *Evolution of Crop Plants, 2nd Edn.*, eds J. Smartt and N. W. Simmonds (London: Longman Scientific & Technical), 383–388.
- Clement, C. R. (1999). 1492 and the loss of Amazonian crop genetic resources. I. The relation between domestication and human population decline. *Econ. Bot.* 53, 188–202. doi: 10.1007/BF02866498
- Clement, C. R., Cristo-Araújo, M., Coppens d'Eeckenbrugge, G., Alves Pereira, A., and Picanço-Rodrigues, D. (2010). Origin and domestication of native Amazonian crops. *Diversity* 2, 72–106. doi: 10.3390/d2010072
- Clement, C. R., Rival, L., and Cole, D. M. (2009a). "Domestication of peach palm (*Bactris gasipaes* Kunth): the roles of human mobility and migration," in *Shifting Spaces, Changing Times: Mobility, Migration and Displacement In Indigenous Lowland South America*, ed M. N. Alexiades (Oxford: Berghahn), 117–140.
- Clement, C. R., Santos, R. P., Desmouliere, S. J. M., Ferreira, E. J. L., and Farias Neto, J. T. (2009b). Ecological adaptation of wild peach palm, its *in situ* conservation and deforestation-mediated extinction in southern Brazilian Amazonia. *PLoS ONE* 4:e4564. doi: 10.1371/journal.pone.0004564
- Clement, C. R., Sousa, N. R., Rodrigues, D. P., Astolfi-Filho, S., Moreno, Y. N., Pascual, V. T., et al. (2002). Use of AFLPS to distinguish landraces of pejobaye (*Bactris gasipaes*) in Brazilian Amazonia. *Sci. Agric.* 59, 743–753. doi: 10.1590/S0103-90162002000400019
- Cole, D., White, T., and Nair, P. (2007). Maintaining genetic resources of peach palm (*Bactris gasipaes* Kunth): the role of seed migration and swidden-fallow management in northeastern Peru. *Genet. Resour. Crop Evol.* 54, 189–204. doi: 10.1007/s10722-005-3134-3
- Coppens d'Eeckenbrugge, G., and Lacape, J.-M. (2014). Distribution and differentiation of wild, feral, and cultivated populations of perennial upland cotton (*Gossypium hirsutum* L.) in Mesoamerica and the Caribbean. *PLoS ONE* 9:e107458. doi: 10.1371/journal.pone.0107458
- Couvreur, T. L. P., Billotte, N., Risterucci, A.-M., Lara, C., Vigouroux, Y., Ludeña, B., et al. (2006). Close genetic proximity between cultivated and wild *Bactris gasipaes* Kunth revealed by microsatellite markers in Western Ecuador. *Genet. Resour. Crop Evol.* 53, 1361–1373. doi: 10.1007/s10722-005-5033-z
- Couvreur, T. L. P., Hahn, W. J., de Granville, J. J., Pham, J. L., Ludena, B., and Pintaud, J. C. (2007). Phylogenetic relationships of the cultivated neotropical palm *Bactris gasipaes* (Arecaceae) with its wild relatives inferred from chloroplast and nuclear DNA polymorphisms. *Syst. Bot.* 32, 519–530. doi: 10.1600/036364407782250526
- Cristo-Araújo, M., Picanço-Rodrigues, D., Astolfi-Filho, S., and Clement, C. R. (2010). Genetic variability in the peach palm genebank with RAPD markers. *Crop Breed. Appl. Biotechnol.* 10, 211–217. doi: 10.1590/S1984-70332010000300005
- Cristo-Araújo, M., Reis, V. M., Rodrigues, D. P., and Clement, C. R. (2013). Domestication of peach palm in southwestern Amazonia. *Tipiti* 11, 74–80. Available online at: <https://digitalcommons.trinity.edu/tipiti/vol11/iss2/9/>
- Cristo-Araújo, M., Rodrigues, D. P., Astolfi-Filho, S., and Clement, C. R. (2015). Peach palm core collection in Brazilian Amazonia. *Crop Breed. Appl. Biotechnol.* 15, 18–26. doi: 10.1590/1984-70332015v15n1a3
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- Ferreira, E. J. L. (1999). "The phylogeny of pupunha (*Bactris gasipaes* Kunth, Palmae) and allied species," in *Evolution, Variation, and Classification of Palms; Memoirs of the New York Botanical Garden*, vol. 83, eds A. Henderson and F. Borchsenius (New York, NY: The New York Botanical Garden), 225–236.
- Futuyama, D. J. (2005). *Evolution*. Sunderland, MA: Sinauer Associates.
- Galluzzi, G., Dufour, D., Thomas, E., van Zonneveld, M., Escobar Salamanca, A. F., Giraldo Toro, A., et al. (2015). An integrated hypothesis on the domestication of *Bactris gasipaes*. *PLoS ONE* 10:e0144644. doi: 10.1371/journal.pone.0144644
- Glaubitz, J. (2004). Convert: a user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Mol. Ecol. Notes* 4, 309–310. doi: 10.1111/j.1471-8286.2004.00597.x
- Graefe, S., Dufour, D., van Zonneveld, M., Rodriguez, F., and Gonzalez, A. (2012). Peach palm (*Bactris gasipaes*) in tropical Latin America: implications for biodiversity conservation, natural resource management and human nutrition. *Biodivers. Conserv.* 22, 269–300. doi: 10.1007/s10531-012-0402-3
- Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.
- Harlan, J. R. (1992). *Crops & Man*. Madison, WI: American Society of Agronomy & Crop Science Society of America.
- Henderson, A. (2000). *Bactris* (Palmae). *Flora Neotropica* 79, 1–181.
- Hernández-Ugalde, J. A., Mora-Urpí, J., and Rocha, O. J. (2011). Genetic relationships among wild and cultivated populations of peach palm (*Bactris gasipaes* Kunth, Palmae): evidence for multiple independent domestication events. *Genet. Resour. Crop Evol.* 58, 571–583. doi: 10.1007/s10722-010-9600-6
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965–1978. doi: 10.1002/joc.1276
- Huber, J. (1904). A origem da pupunha. *Boletim do Museu Paraense Emilio Goeldi* 4, 474–476.
- Huelsenbeck, J., and Ronquist, F. (2001). MrBAYES: Bayesian inference of phylogenetic trees. *Bioinforma* 17, 754–755. doi: 10.1093/bioinformatics/17.8.754
- Jombart, T. (2008). ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinforma* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129
- Jombart, T., Devillard, S., Dufour, A. B., and Pontier, D. (2008). Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* 101, 92–103. doi: 10.1038/hdy.2008.34
- Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187
- Martinez, A. K., Gaitán-Solis, E., Duque, M. C., Bernal, R., and Tohme, J. (2002). Microsatellite loci in *Bactris gasipaes* (Arecaceae): their isolation and characterization. *Mol. Ecol. Notes* 2, 408–410. doi: 10.1046/j.1471-8286.2002.00256.x
- Mayle, F. E., Beerling, D. J., Gosling, W. D., and Bush, M. B. (2004). Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 359, 499–514. doi: 10.1098/rstb.2003.1434
- Meyer, R. S., DuVal, A. E., and Jensen, H. R. (2012). Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. *New Phytol.* 196, 29–48. doi: 10.1111/j.1469-8137.2012.04253.x
- Miller, A. J., and Gross, B. L. (2011). From forest to field: perennial fruit crop domestication. *Am. J. Bot.* 98, 1389–1414. doi: 10.3732/ajb.1000522
- Mora-Urpí, J. (1984). "El pejobaye (*Bactris gasipaes* H. B. K.): origen, biología foral y manejo agronómico," in *Palmeras Poco Utilizadas de América Tropical* (Turrialba: Food and Agriculture Organization/Centro Agronómico Tropical de Investigación y Enseñanza), 118–160.
- Mora-Urpí, J. (1993). "Diversidad genética en pejobaye. II-Origen y domesticación," in *Anales del IV Congreso Internacional sobre Biología, Agronomía e Industrialización del Pijuayo*, eds J. Mora-Urpí, L. T. Szott, M. Murillo, and V. M. Patiño (San José: Editorial de la Universidad de Costa Rica), 21–30.
- Mora-Urpí, J. (1999). "Origen y domesticación," in *Palmito de Pejobaye (*Bactris gasipaes* Kunth): Cultivo e Industrialización*, eds J. Mora-Urpí and J. Gainza Echeverría (San José: Universidad de Costa Rica), 17–24.
- Mora-Urpí, J., and Clement, C. R. (1988). "Races and populations of peach palm found in the Amazon basin," in *Final Report (Revised): Peach palm (*Bactris gasipaes* H.B.K.) Germplasm Bank. US AID, Proj. no. DAN-5542-G-SS-2093-00*, eds C. R. Clement and L. Coradin (Manaus: INPA, Embrapa Cenargen, US AID), 78–94.
- Mora-Urpí, J., Weber, J. C., and Clement, C. R. (1997). *Peach Palm (*Bactris gasipaes* Kunth). Promoting the Conservation and Use of Underutilized and Neglected Crops*, vol. 20. Gatersleben; Rome: IPK, IPGRI.
- Morcote-Rios, G., and Bernal, R. (2001). Remains of palms (Palmae) at archaeological sites in the New World: a review. *Bot. Rev.* 67, 309–350. doi: 10.1007/BF02858098
- Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89, 583–590.

- Patiño, V. M. (1963). *Plantas Cultivadas y Animales Domésticos en América Equinoccial. Frutales*. Cali: Imprenta Departamental.
- Patiño, V. M. (1992). An ethnobotanical sketch of the palm *Bactris (Guilielma) gasipaes*. *Principes* 36, 143–147.
- Patiño, V. M. (2002). *Historia y Dispersión de los Frutales Nativos del Neotrópico*. Cali: Centro Internacional de Agricultura Tropical.
- Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Modell.* 190, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. Available online at: <http://www.genetics.org/content/155/2/945>
- Purugganan, M. D., and Fuller, D. Q. (2009). The nature of selection during plant domestication. *Nature* 457, 843–848. doi: 10.1038/nature07895
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing, Version 2.14*. Vienna: R Foundation for Statistical Computing.
- Rindos, D. (1984). *The Origins of Agriculture: An Evolutionary Perspective*. San Diego, CA: Academic Press.
- Rodrigues, D. P., Astolfi Filho, S., and Clement, C. R. (2005). Molecular marker-mediated validation of morphologically defined landraces of pejibaye (*Bactris gasipaes*) and their phylogenetic relationships. *Genet. Resour. Crop Evol.* 51, 871–882. doi: 10.1007/s10722-005-0774-2
- Rodrigues, D., Vinson, C., Ciampi, A., Farias, I., Lemes, M., Astolfi-Filho, S., et al. (2004). Novel microsatellite markers for *Bactris gasipaes* (Palmae). *Mol. Ecol. Notes* 4, 575–576. doi: 10.1111/j.1471-8286.2004.00739.x
- Rojas-Vargas, S., Ramírez, P., and Mora-Urpí, J. (1999). Polimorfismo isoenzimático en cuatro razas y un híbrido de *Bactris gasipaes* (Palmae). *Rev. Biol. Trop.* 47, 755–761.
- Saldias-Paz, M. (1993). “La chonta de castilla (*Bactris gasipaes* H.B.K.): taxonomía y algunos datos económicos em Santa Cruz y su distribución en Bolivia,” in *Anales del IV Congreso Internacional sobre Biología, Agronomía e Industrialización del Pijuayo*, eds J. Mora Urpí, L. T. Szott, M. Murillo, and V. M. Patiño (San José: Editorial Univ Costa Rica), 115–126.
- Santos, R. P., Cristo-Araújo, M., Picanço-Rodrigues, D., Astolfi-Filho, S., and Clement, C. R. (2011). Variabilidade genética e fluxo gênico em populações híbridas e silvestres de pupunha acessada com marcadores RAPD. *Rev. Bras. Frutic.* 33, 1200–1208. doi: 10.1590/S0100-29452011000400019
- Schmidt, H. A. (2009). “Testing tree topologies,” in *The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing*, eds P. Lemey, M. Salemi, and A.-M. Vandamme (Cambridge: Cambridge University Press), 381–403. doi: 10.1017/CBO9780511819049.014
- Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., et al. (2005). The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.* 92, 142–166. doi: 10.3732/ajb.92.1.142
- Shaw, J., Lickey, E. B., Schilling, E. E., and Small, R. L. (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, 275–288. doi: 10.3732/ajb.94.3.275
- Sousa, N. R., Rodrigues, D. P., Clement, C. R., Nagao, E. O., and Astolfi-Filho, S. (2001). Discriminação de raças primitivas de pupunha (*Bactris gasipaes*) na Amazônia brasileira por meio de marcadores moleculares (RAPDS). *Acta Amazonica* 31, 539–545. doi: 10.1590/1809-43922001314545
- Thomas, E., van Zonneveld, M., Loo, J., Hodgkin, T., Galluzzi, G., and van Etten, J. (2012). Present spatial diversity patterns of *Theobroma cacao* L. in the Neotropics reflect genetic differentiation in Pleistocene refugia followed by human-influenced dispersal. *PLoS ONE* 7:e47676. doi: 10.1371/journal.pone.0047676
- Varela, S., Lima-Ribeiro, M. S., and Terribile, L. C. (2015). A short guide to the climatic variables of the Last Glacial Maximum for biogeographers. *PLoS ONE* 10:e0129037. doi: 10.1371/journal.pone.0129037

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Clement, Cristo-Araújo, Coppens d’Eeckenbrugge, Reis, Lehnebach and Picanço-Rodrigues. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.