Natural and human-mediated drivers of microevolution in Neotropical palms: a historical genomics approach

Louise Brousseau[1,2]*, Sylvain Santoni[3], Audrey Weber[3], Guillaume Odonne[4]

*corresponding author: louise.brousseau@ird.fr (ORCID 0000-0003-2769-2767)

[1]UMR AMAP, IRD, Univ. Montpellier, CIRAD, CNRS, INRAE, Montpellier, France

[2]UMR DIADE, IRD, Univ. Montpellier, CIRAD, Montpellier, France

[3]UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France

[4]UAR LEEISA, CNRS, Cayenne, French Guiana

Abstract

- Microevolution in Amazonia has been far less documented than in temperate regions and still constitutes a major knowledge gap. Moreover, the actual influence of pre-Columbian societies through the Holocene remains unclear, although it could provide interesting insights on the relationship between natural and human-mediated evolutionary processes in driving microevolution.
- Because they are widespread and traditionally managed since pre-Columbian times, Neotropical palms are choice models to investigate the drivers of microevolution in Amazonia. In this study, we carried out a preliminary exploration of the genomic diversity in two pairs of congeneric palm species in French Guiana (*Astrocaryum* spp. and *Oenocarpus* spp.).
- We built upon an original sampling design, taking into account both regional climate variations and local-scale pre-Columbian occupation, and designed a new target capture kit of 20,000 molecular probes scattered across exonic regions of more than 5,000 nuclear genes in Arecoïdeae ("ARECO5000+"). Hundreds of palm libraries were sequenced through Illumina sequencing, providing a unique – high-coverage – genomic dataset in these non-model species.
- We explored patterns of genomic diversity and differentiation within and across populations, bringing state-of-the-art knowledge about the roles of climate adaptation and pre-Columbian domestication. By documenting original cases of "incipient" domestication, these pioneer results open new avenues toward a better understanding of microevolution in Amazonia and of the impact of pre-Columbian societies on present-day biodiversity patterns.

Key word: *Oenocarpus* spp., *Astrocaryum* spp., targeted capture, Illumina sequencing, climate adaptation, pre-Columbian societies, ring ditch, incipient domestication.

Introduction

Evolution in Amazonia has been far less documented than in temperate regions and still constitutes a major knowledge gap. Evolutionary studies in Amazonia also largely focused on macroevolution through long-term geological periods, providing interesting insights onto speciation and extinction dynamics in relation to landscape changes associated with the Andean uplift and the formation of the Amazon river (Hoorn et al. 2010), to species dispersal abilities (B. T. Smith et al. 2014) and to biotic interactions (Onstein et al. 2018). On the contrary, less attention has been paid to the microevolution of plants and animals during the Quaternary (Cavers and Dick 2013), and even less during the Holocene. This period, which started about 11,000 years ago, experienced both a global warming subsequent to the last glacial period resulting in the present climate, and the establishment of dense networks small-scale human societies subsequent to the human colonization of South America in late Pleistocene (Llamas et al. 2016; Rothhammer and Dillehay 2009). A better understanding of the drivers of plant and animal evolution during the Holocene could thus bring interesting insights on the interactions between natural and human-mediated evolutionary processes and their role in shaping present-day patterns of biodiversity.

Neotropical palms of the Arecoideae subfamily are choice models to investigate the drivers of microevolution in Amazonia. Dated phylogenies revealed that Arecoideae crown node diversified in mid-Cretaceous in South America, suggesting a Neotropical origin (Baker and Couvreur 2013). Today, Neotropical palms are widespread in Amazonia ; Arecaceae standing at the second position of the most abundant plant families in forest tree inventories with some “hyperdominant” species (ter Steege et al. 2013). Neotropical palms also constitute a key resource to human societies since pre-Columbian times, whose multiple uses – including

subsistence food and building materials – have been extensively documented at least since the 1850's (Wallace 1853; N. Smith 2015). Surprisingly, they are also particularly abundant at archaeological sites with evidence of long-term human occupation, and some species (in particular *Astrocaryum* spp. and *Oenocarpus* spp.) are commonly credited as 'biological indicators' of ancient settlements (Odonne et al. 2019).

This is notably the case at the so-called "ring ditches" that are abundant in interfluvial terra-firme forests in lowland Amazonia, from Bolivia (John F Carson et al. 2015) to French Guiana (Petitjean Roget 1991). These pre-Columbian earthworks provide valuable evidence of anthropogenic disturbances, with belowground enrichments in pre-Columbian ceramics, organic matter and charcoals characterizing "anthropogenic soils" that are commonly attributed to slash-and-burn agriculture (Bodin et al. 2020; Brancier et al. 2014). Even if the actual magnitude of land use at these sites remain controversial – with hypotheses varying from extensive forest clearance to the exploitation of naturally open savanna (Watling et al. 2017; John Francis Carson et al. 2014) – there is no more doubt that pre-Columbian societies have strongly impacted forest landscapes, thus contrasting with previous depictions of Amazonia as an undisturbed, primary forest (Levis et al. 2017).

These repeated and converging observations feed the long-standing hypothesis that many plant species, including Neotropical palms, may have been "incipiently-" or "semi-" domesticated through pre-Columbian times (Clement 1999; Clement et al. 2010, 2006). Whether pre-Columbian societies have actually shaped patterns of palm diversity through domestication, and how human-mediated processes interacted with natural ones, remain however largely unknown. This knowledge gap is primarily underpinned by a profound lack of observational genomic data in non-model Neotropical palms that we intend to fill in this study.

We carried out a preliminary exploration of the genomic diversity in two pairs of congeneric Neotropical palms in coastal French Guiana: *Astrocaryum paramaca* and *A. sciophilum* on one side, *Oenocarpus bacaba* and *O. bataua* on the other side. We sampled populations following an original sampling design taking into account both regional and local-scale evolutionary processes related to climate variations and pre-Columbian occupation, respectively. At the regional scale, we sampled populations in study sites with contrasted rainfall to catch climate adaptation. At the local scale, we sampled palms both on the top of ring ditches and in the surrounding forests to catch the impact of domestication associated with local anthropogenic disturbances. We designed a new capture kit composed of 20,000 molecular probes to target exonic regions within thousands of nuclear genes, and carried out an extensive targeted-capture experiment and next-generation sequencing (NGS) of hundreds of libraries.

Here, we introduce this new capture kit ("ARECO5000+") and provide new NGS data on the state and distribution of genomic diversity in Neotropical palms in coastal French Guiana. We investigate intra-specific patterns of diversity and differentiation driven by both neutral evolutionary processes (gene flow and admixture) and selective ones, disentangling the roles of (natural) climate adaptation and of (human-mediated) pre-Columbian domestication that operated at these two spatial scales.

Materials and Methods

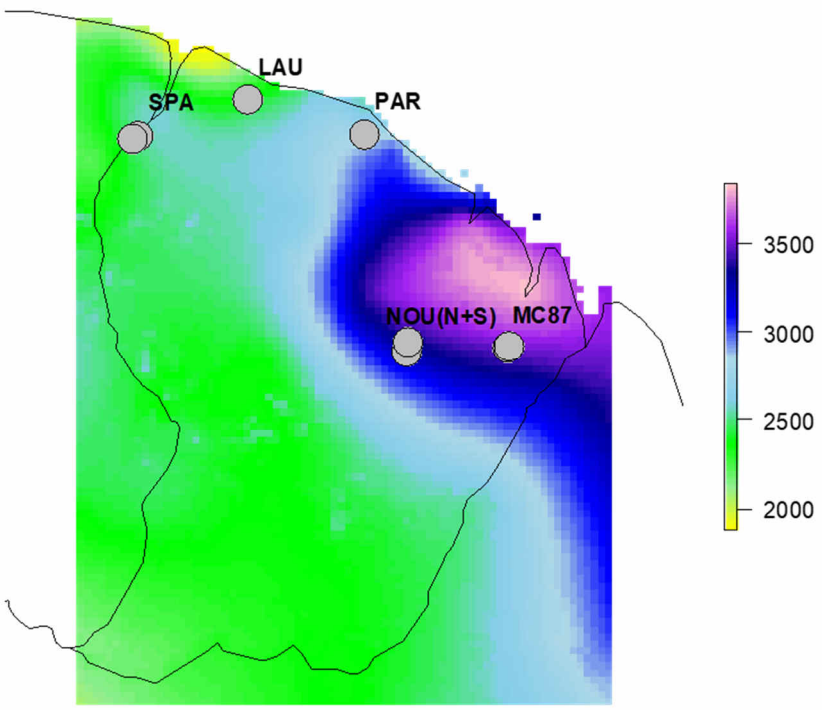
1. Palm species

We focused on four palm species that belong to the Arecoideae sub-family and that are widespread in forest landscapes of French Guiana: two species of the genus *Astrocaryum* (*A. paramaca* and *A. sciophilum*) and two species of the genus *Oenocarpus* (*O. bacaba* and *O. bataua*). *Oenocarpus bacaba* and *O. bataua* are pan-Amazonian species that are broadly distributed from Colombia to north Brazil. *Oenocarpus bataua* stands at the 7th position amongst the 20 most abundant species in the Amazonian tree flora, while *O. bacaba* stands at the 33th position (ter Steege et al. 2013). The distribution of *A. paramaca* and *A. sciophilum* is restricted to northeast Amazonia (Suriname, French Guiana and Brazil, GBIF Secretariat: GBIF Backbone Taxonomy. https://doi.org/10.15468/39omei). Despite this, *A. sciophilum* stands at the 51th position amongst the most abundant species in forest tree inventories. Moreover, *O. bacaba* and *A. sciophilum* are two biological indicators of pre-Columbian settlements in French Guiana, as their abundance is significantly higher in forest plots with evidence of pre-Columbian occupation than in natural plots (Odonne et al. 2019). Indeed, *Astrocaryum* spp. and *Oenocarpus* spp. are major sources of food and of building and handicraft materials in Amazonia (Sosnowska et al. 2015; Cummings and Read 2016). Their leaves are traditionally used in basketry and to build houses (see for example maloca's walls and roof, García et al. 2015; Davy 2007; Fadiman 2008; Coomes 2004; Gutierrez 2020; Kahn 1993), while their fruits are used to make oil (*e.g.* 'seje' oil) and beverages (*e.g.* so called 'comou' and 'wassaï' in French Guiana and 'bacaba wine' in Northwest Amazonia, Cunha et al. 2019; Castro et al. 2014; Puerari et al. 2015).

2. Sampling strategy

130 Palm populations of the four palms species were sampled in five study sites in the coastal
131 French Guiana with different rainfall ranging from about 2400 to 3500 mm.year^{-1} (**Figure 1**):
132 three archaeological sites with ring ditches (SPA, MC87 and NOU), plus two additional
133 undisturbed sites without evidence of pre-Columbian occupation (LAU and PAR). In the three
134 study sites with ring ditches, we sampled palm populations both at the top of the ring ditches
135 and in the surrounding (undisturbed) forest (< 5 km from the ring ditch), **Table 1**. When the
136 species were present, we collected a piece of leaf for 20 to 100 adult palms per species in each
137 sampling site and local condition (pre-Columbian occupation "+" and "–").

138



139

140 **Figure 1**. Localization of the study sites and annual precipitation (in mm.year^{-1}) in French
141 Guiana. Current precipitations were extracted from WorldClim version 1 using R (libraires
142 "`raster`" and "`sp`") and are representative of the period 1960-1990
143 (http://www.worldclim.com/version1).

144 **Table 1.** Localization of populations in the different sampling sites and local conditions of pre-
145 Columbian occupation. "+" refers the top of the ring ditch with evidence of pre-Columbian
146 occupation, and "–" refers the surrounding sampling forest area without evidence of pre-
147 Columbian occupation. "Nouragues North" and "Nouragues South" indicate the localization of
148 two ring ditches in the north and in the south of the Nouragues Ecological research station,
149 respectively. Coordinates are provided in decimal degrees (WGS84).

Study site	Site ID	Local condition	Latitude (N)	Longitude (E)	Rainfall (mm.year^{-1})
Sparouine	SPA	–	5.272861999	-54.22715197	2529
	SPA	+	5.253569037	-54.253331	2491
Laussat	LAU	–	5.479205009	-53.59530601	2391
Paracou	PAR	–	5.279138032	-52.92425998	2818
Nouragues South	NOU_S	–	4.038217971	-52.68241298	3246
	NOU_S	+	4.029483022	-52.681869	3246
Nouragues North	NOU_N	–	4.081238005	-52.67338903	3299
	NOU_N	+	4.079981977	-52.67142003	3299
Montagne Couronnée 87	MC87	–	4.053646959	-52.09517601	3517
	MC87	+	4.05699201	-52.086698	3517

150

151 ### 3. Probes design, targeted capture and sequencing

152 We took advantage of the availability of the chromosome-level reference genome for the oil
153 palm (*Elaeis guineensis*, Arecaceae, Arecoideae) to design a custom of capture kit of 20,000
154 molecular probes located within 5,209 orthologous nuclear genes scattered across all 16
155 chromosomes ("ARECO5000+"). Details on probes design is provided in **Supplementary**
156 **material 1**. Probe sequences are available upon request from the corresponding author. The
157 probes were then synthetized by Daicel Arbor Biosciences (myBaits Custom DNA-Seq©).
158 Molecular experiments were further performed at the AGAP genotyping platform (Montpellier,
159 France). Plant DNA was extracted from 15 mg of dried leaves. Ten to fifteen best-quality DNAs

per population were retained for library preparation, targeted-capture and sequencing, resulting in the sampling design described in **Table 2** and totaling 405 libraries (one per palm). Genomic library preparation for multiplexed individuals and targeted-capture by bait hybridization followed published protocols (Rohland and Reich 2012; Mascher et al. 2013) with some modifications. The molecular protocol is provided in **Supplementary material 2**. One genomic library was prepared for each genotype, indexed, and enriched using the myBaits Custom DNA-Seq© kit "ARECO5000+". 48 barcoded libraries were then equally mixed before sequencing. Libraries of *Oenocarpus bataua* were sequenced (paired-end, 2 × 150bp) on an Illumina MiSeq at the AGAP genotyping platform (Montpellier, France), while libraries of *Astrocaryum* spp. and *Oenocarpus bacaba* were sequenced by the Get-PlaGe core facility (GenoToul platform, INRA Toulouse, France http://www.genotoul.fr) on an Illumina HiSeq3000. Sequence data are available on EMBL-ENA under accession PRJEB51800.

Table 2. Sampling design for library preparation and sequencing. "NA" indicates that the species was absent at this sites and/or local condition.

Site ID	Local condition	*Astrocaryum paramaca*	*Astrocaryum sciophilum*	*Oenocarpus bacaba*	*Oenocarpus bataua*
SPA	–	12	12	12	12
SPA	+	12	12	12	12
LAU	–	12	12	12	12
PAR	–	12	12	12	12
NOU_S	–	12	NA	NA	NA
NOU_S	+	12	NA	12	12
NOU_N	–	15	10	NA	12
NOU_N	+	14	15	15	12
MC87	–	12	12	NA	12
MC87	+	12	12	NA	12
TOTAL		**125**	**97**	**75**	**108**

177 ## 4. Bioinformatics

178 **Reads demultiplexing and pre-processing.** Reads were demultiplexed based on their 6
179 nucleotide barcodes with `fastq-multx` of the `ea-utils-2.7` suit of tools (Aronesty
180 2013) with options `-b -x -m 0 -d 1`, and split into individual .fastq files (two files per
181 library, corresponding to "R1" and "R2" paired reads). Read sequences that failed to pass the
182 chastity filter (identified "Y" in the "@" line in CASAVA 1.8) were discarded using the
183 command:

```
184 zcat $lib_name.R1.fastq.gz | grep -A 3 '^@.* [^:]*:N:[^:]*:' |
185 grep -v "^--$" > $lib_name_pf.R1.fastq.
```

186 Reads were further cleaned using `FASTX-Toolkit (RRID:SCR_005534) v.0.0.14`
187 and `sickle-1.33`, as follow:

188 - Trimming with `fastx_trimmer` with option `-Q 33 -f 10 -l 140`, resulting
189 in reads of 140-nucleotide long.
190 - Masking of individual bases of low quality (phred-score < 25) with `fastq_masker`
191 with options `-q 25 -Q 33`.
192 - Cutting of 3'-end and 5'-end and discarding of orphan reads with `sickle-1.33`
193 (Joshi and Fass 2011) with options `pe -l 100` (*i.e.* default quality threshold of 20
194 and length threshold of 100bp).

195 Once cleaned, reads were quality controlled with `FastQC-0.11.5` (Andrews 2010).

196 **Mapping.** Cleaned reads of *Astrocaryum* spp. and *Oenocarpus* spp. libraries were separately
197 mapped against *Elaeis guineensis* reference genome (EG5) using `bwa` (H. Li and Durbin 2010,
198 2009): the reference was indexed with the command `bwa index` and paired-reads were
199 mapped using `bwa mem` with option `-t 10`.

200 Mapping were post-processed using `samtools-1.9` (H. Li 2011; H. Li et al. 2009).
201 Multiple mapped reads were discarded from .bam alignments using `samtools view` with
202 the command:

```
203 samtools view -h mapping.bam | awk '$17 !~ /XA:/ || $1 ~ /^@/'
204 | /usr/local/samtools-1.9/bin/samtools view -bS >
205 mapping_unique.bam
```

206 Alignments were then sorted and indexed using `samtools sort` and `samtools index`,
207 respectively. Mapping statistics were computed using `samtools stats`, `samtools
208 flagstat` and `samtools bedcov`. Sequencing and mapping statistics are provided in
209 **Supplementary materials 3 and 4**.
210 Following a first mapping against *Elaeis guineensis* reference genome, the reference was
211 updated by editing two genus-specific consensus (one for *Astrocaryum* spp. and one for
212 *Oenocarpus* spp.) to take into account the genomic divergence between *Elaeis guinensis* and
213 the genera under study. This step aimed to avoid detecting polymorphisms caused by inter-
214 genus variation in the following. This was achieved by calling variants using `samtools
215 mpileup` with options `-uf` for each *Astrocaryum* spp. and *Oenocarpus* spp. alignments,
216 respectively. The two consensus references were edited with `bcftools index` (v.1.3) and
217 `bcftools consensus` (Danecek et al. 2021). Then, the libraries of each species
218 (*Astrocaryum paramaca*, *A. sciophilum*, *Oenocarpus bacaba* and *O. bataua*) were mapped
219 against the corresponding genus-specific consensus reference following the mapping procedure
220 described above.
221 **SNP and genotype calling.** Variants were called using `samtools-1.9` and `bcftools-
222 1.3`, as follow:

223 - Indexing of genus-specific consensus with `samtools faidx` ;

224 • Variant calling with `samtools mpileup` with options `-g -t DP` followed by
225 `bcftools call` with options `-vc`.

226 The resulting .vcf file was filtered with `grep` to discard monomorphic sites, indels, and variants
227 detected on unassembled scaffolds or chloroplast to retain nuclear single-nucleotide
228 polymorphisms (SNPs) only. SNPs were post-filtered with `R version 3.5.1` (R Core
229 Team 2021) to retain SNPs and genotypes with a quality threshold of 20, and a per-library
230 coverage threshold of 6X. SNPs and genotypes that did not meet the criteria were considered
231 as being missing ("NA"), and SNPs that became monomorphic or contained more than 20% of
232 missing values were discarded. This resulted in 178,358 and 23,252 high-quality SNPs in
233 *Astrocaryum* spp. and *Oenocarpus* spp., respectively. Last, palm samples with more than 20%
234 of missing values were discarded. The list of samples retained for genetic analyses is provided
235 in **Supplementary material 5**.

236

237 **5. Population genetics**

238

239 Genome-wide patterns of genetic structure and diversity were computed on a random subset
240 of 5,000 SNPs in `R version 3.5.1` (`"sample"` command with argument `replace =`
241 `F)`. Genotype tables were converted from tabular (data frame) to `genind` objects using the
242 function `df2genind` (package `adegenet`).

243 **Genetic differentiation.** Pairwise genetic differentiation (F_{ST}) was estimated both between
244 congeneric species and within each species. Weir and Cockerham's F_{ST} were estimated using
245 the function `genet.dist` (package `hierfstat`) with argument: `method="WC84"`.

246 **Individual ancestry.** Genome-wide genetic structure was assessed for each species by
247 estimating individual ancestry (Frichot et al. 2014) with the function `snmf` (package `LEA`) with

arguments: `ploidy = 2`, `entropy = TRUE` and `CPU = 8`. The genetic structure was explored from `K=1` to `K=10` ancestral populations (*A. paramaca*), to `K=8` (*A. sciophilum*), to `K=6` (*O. bacaba*) or to `K=9` (*O. bataua*). The most probable number of clusters was retained based on cross-entropy criterion (lower values indicating better fit to the predictions) and each palm sample was assigned to a given group with an ancestry threshold of 0.5.

Isolation-by-distance. Isolation-by-distance (i.e. the correlation between genetic and geographic matrices across populations within each species) was tested through Mantel tests. Genetic distances were estimated with the function `dist.genpop` (package `adegenet`) with argument `method = 3` (*i.e.* coancestry coefficient or Reynolds' distance, Jombart 2008; Jombart and Ahmed 2011). Geographic distances (euclidean) were computed from geographic coordinates using the function `dist` with default arguments. Mantel tests between geographic and genetic distance matrices were computed with the function `mantel.randtest` (package `ade4`, Mantel 1967; Thioulouse et al. 1997).

Admixture tests. Admixture between populations was evaluated through *f3* admixture tests for all combinations of three populations (Patterson et al. 2012; Peter 2016; Reich et al. 2009): $f_3(P_X ;P_A,P_B)$, where P_X is the target population and P_A and P_B are two source populations. Admixture tests were encoded with `R version 3.5.1`. f_3 statistic was estimated following the equation by Reich et al. (2009):

$$\mathbb{E}(p_X - p_a)(p_X - p_b),$$

where p_x, p_a and p_b are the allele frequencies in the populations P_X, P_A and P_B, respectively, and $\mathbb{E}$ the expected value across all SNPs. Negative values indicate admixture between A and B. The significance of f_3 was assessed by bootstrap (function `sample` with argument `replace = T`) through 500 iterations. Z-scores were converted into p-values (left-tailed) with the function `pnorm`.

272 **Intra-population diversity.** Observed and expected heterozygosity within each species and
273 population (*Ho* and *He*, respectively) were estimated with `adegenet` from `genind` objects.
274 `Genind` objects containing multiple populations were split into separate populations using the
275 function `seppop` and summary statistics were exported with the function `summary`. Whether
276 *Ho* was higher than *He* was assessed for each population through non-parametric Wilcoxon
277 tests (because of non-homogeneity of variances between *Ho* and *He*) with the function
278 `wilcox.test`. Inbreeding coefficients (*Fis*) were estimated following the formula: $Fis =$
279 $1 - \frac{Ho}{He}$. Hardy Weinberg equilibrium was tested for each SNP through Chi-squared tests with
280 the function `hw.test` (package `pegas`), and the proportion of SNPs showing significant
281 departure from HW equilibrium at a 1% threshold was assessed.
282 **Selection test.** We carried out a genome-wide scan for selective differentiation within each
283 species using `BayPass v.2.1`. Genotype data were converted into allele counts and
284 `BayPass` core model was run a first time using observed SNP counts as input. The number of
285 populations in each dataset was specified with option `-npop` and MCMC parameters were
286 specified with arguments: `-nthreads 8 -thin 15 -burnin 2500 -npilot 15 -`
287 `pilotlength 5000`. The estimated covariance matrix (Ω), which provides an accurate
288 estimate of genome-wide allele frequency covariance across populations, was converted into a
289 correlation matrix in `R version 3.5.1` using the function `cov2cor`. The singular values
290 of the covariance matrix were decomposed using `svd` and left singular vectors (U) were
291 extracted to plot the genome-wide structure across populations. Pseudo-observed data (POD)
292 of 10,000 SNPs (for *Oenocarpus* spp.) or 100,000 SNPs (for *Astrocaryum* spp.) were simulated
293 with the `simulate.baypass` function (`BayPass` miscellaneous R function), specifying
294 the total number of allele counts for each SNP (*i.e.* sample size) and the estimated
295 hyperparameters a_π and b_π (*i.e.* posterior mean and standard deviation of the Beta distribution
296 corresponding to the mean reference allele frequency, ~ancestral reference allele frequency) as

297 input. These POD correspond to a simulated set of SNP counts that take into account the
298 genome-wide (neutral) genetic structure empirically inferred from observed SNP counts. The
299 core model was run again using POD as input under the same parameters to calibrate SNP-
300 specific statistics (standardized allele frequencies α^{std}, *XtX* and derived statistics) under neutral
301 hypothesis. `BayPass` outputs were post-processed with `R version 3.5.1`. The
302 robustness of the core model was assessed by evaluating the FMD distance (Förstner and
303 Moonen 2003) between covariance matrices estimated on observed data (Ω^{OBS}) and on
304 simulated data (Ω^{POD}) using `fmd.dist` function (`BayPass` miscellaneous R function). SNP-
305 specific departures from genome-wide (*i.e.* neutral) differentiation were detected based on the
306 joint distribution of:

307 - SNP-specific XtX_i (Günther and Coop 2013) estimated by `BayPass` (which describes
308 the differentiation across all populations for each SNP *i* explicitly corrected for genome-
309 wide covariance, whatever their regional and local environmental conditions), and
310 - two additional statistics (δ_i and Δ_i) derived from estimated standardized allele
311 frequencies ($\alpha_{i,j}^{STD}$, which describes allele frequencies for each SNP *i* within each
312 population *j* explicitly corrected for the genome-wide covariance matrix) and designed
313 to quantify variations in α^{STD} driven by climate adaptation and pre-Columbian
314 domestication, respectively.

315 Because α^{std} are unbounded, α^{std} were rescaled between 0 and 1 through min-max
316 normalization:

317 $$rescaled\ \alpha_{i,j}^{STD\ (OBS)} = \frac{\alpha_{i,j}^{STD\ (OBS)} - \min(\alpha^{STD})}{\max(\alpha^{STD}) - \min(\alpha^{STD})} \text{ and } rescaled\ \alpha_{i,j}^{STD\ (POD)} = \frac{\alpha_{i,j}^{STD\ (POD)} - \min(\alpha^{STD})}{\max(\alpha^{STD}) - \min(\alpha^{STD})}$$

318 ,

319 where:

- $\alpha_{i,j}^{STD\ (OBS)}$ and $\alpha_{i,j}^{STD\ (POD)}$ are the standardized allele frequency at SNP *i* in population *j* estimated from observed and simulated data respectively ;
- $\min(\alpha^{STD}) = \min\left(\alpha^{STD\ (OBS)}, \alpha^{STD\ (POD)}\right)$;
- $\max(\alpha^{STD}) = \max(\alpha^{STD\ (OBS)}, \alpha^{STD\ (POD)})$.

The statistic δ quantifies clines in α^{STD} along the regional precipitation gradient:

$$\delta_i = |\beta_i|.r_i^2 \ ,$$

where $|\beta_i|$ is the absolute regression coefficient and r_i^2 is the r-squared of the linear model between allele frequency at SNP i and precipitation: $\alpha_{i,j}^{STD} = \beta_i.prec_j$, with $prec_j$ the precipitation in population *j* (scaled between 0 and 1). Linear models were fitted for each SNP with the function `lm`.

The statistic Δ quantifies absolute differences in α^{STD} between contrasted local conditions of pre-Columbian occupation (+ *versus* –) within and across study sites (Δ_i and $\Delta_{i,s}$, respectively):

$$\Delta_{i,s} = \left|\alpha_{i,(j+)}^{STD} - \alpha_{i,(j-)}^{STD}\right|, \text{ within study sites,}$$

and

$$\Delta_i = \sum_s(\Delta_{i,s}), \text{ across study sites,}$$

where *j*+ and *j*– refer to populations located on the ring ditch and in the surrounding natural forest for each study site *s* (*i.e.* single-site contrast), respectively, and $\Delta_{i,s}$ is the sum across study sites *s* (*i.e.* multi-site contrast).

δ and Δ statistics were separately computed on observed data (δ_i^{OBS}, $\Delta_{i,s}^{OBS}$, Δ_i^{OBS}) and on POD (δ_i^{POD}, $\Delta_{i,s}^{POD}$, Δ_i^{POD}). A SNP was retained as being shaped by climate adaptation if it was both over-differentiated across all populations whatever their local environment based on XtX at a 95% POD threshold (*i.e.* condition 1), and if it displayed clinal allele frequency variation with regional precipitation based on δ at a 95% POD threshold (*i.e.* condition 2). Similarly, a SNP was retained as being shaped by domestication if both condition 1 was met, and if it displayed

344 local differences in allele frequencies in relation to pre-Columbian occupation based on Δ at a
345 95% POD threshold (*i.e.* condition 3).

Climate adaptation SNP	**Domestication SNP**
Condition 1: $XtX_i^{OBS} > q_{(95)}(XtX_i^{POD})$	
Condition 2: $\delta_i^{OBS} > q_{(95)}(\delta_i^{POD})$	Condition 3: $\Delta_i^{OBS} > q_{(95)}(\Delta_i^{POD})$

346

347 where:

348 • XtX_i^{OBS} is the observed SNP-specific differentiation, and $q_{(95)}(XtX_i^{POD})$ is the 95-
349 quantile of the distribution (across all SNPs) of expected SNP-specific differentiation
350 (estimated from POD);
351 • δ_i^{OBS} is the observed SNP-specific cline in allele frequency along the regional
352 precipitation gradient, and $q_{(95)}(\delta_i^{POD})$ is the 95-quantile of the distribution (across all
353 SNPs) of expected SNP-specific clines in allele frequency along the regional
354 precipitation gradient (estimated from POD);
355 • Δ_i^{OBS} is the observed SNP-specific difference in allele frequency between local
356 conditions of pre-Columbian occupation (+ *versus* –), and $q_{(95)}(\Delta_i^{POD})$ is the 95-
357 quantile of the distribution (across all SNPs) of expected SNP-specific differences in
358 allele frequency between local conditions of pre-Columbian occupation (estimated from
359 POD).

360 **Outlier blastn.** Last, we retrieved the identity of genes containing precipitation and
361 domestication SNPs. For each species, we extracted the sequences of the genomic regions
362 surrounding each outlier SNP (+/- 200 bp) from consensus genomic sequences using
363 `bedtools getfasta` (`bedtools-2.27.1`). Each SNP-surrounding region was then
364 blasted against the NCBI nucleotide database using `blastn` (`NCBI BLAST+ v.`

365 2.10.1+), limiting blast search by taxonomy (*Elaeis guineensis* taxid = 51953 and *Phoenix*

366 *dactylifera* taxid = 42345) and returning the first good hit to optimize computation time.

367

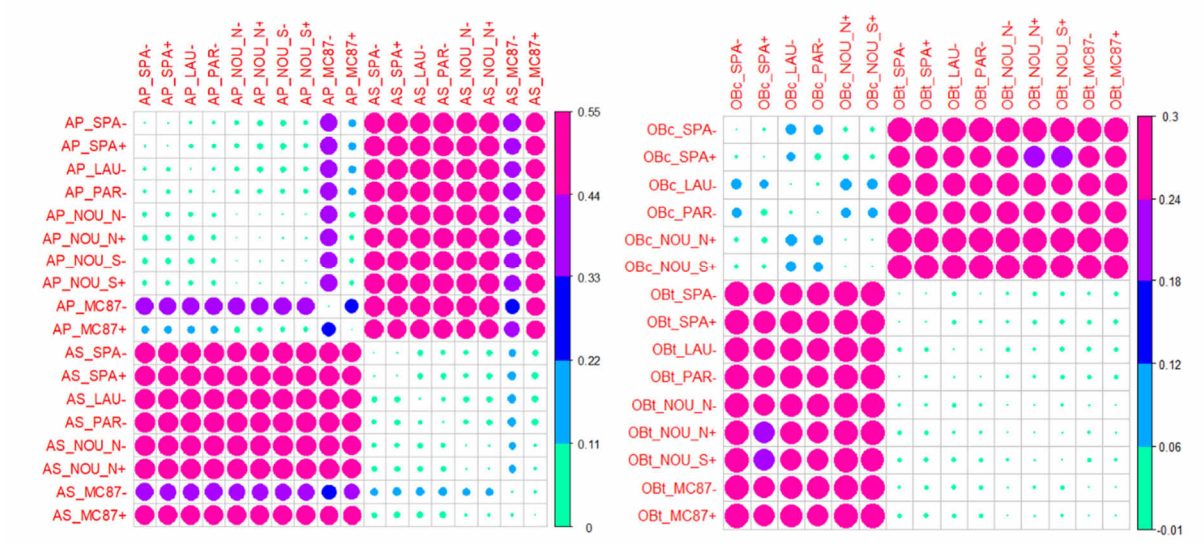
368 **Results**

369

370 **1. Genome-wide genetic structure and differentiation**

371

372 **Genetic differentiation.** At the inter-specific level, pairwise F_{ST} across populations ranged
373 between 0.30 and 0.55 in *Astrocaryum* spp. and between 0.23 and 0.30 in *Oenocarpus* spp.
374 Intra-specific F_{ST} ranged between 0.003 and 0.4 in *A. paramaca*, between 0.01 and 0.11 in *A.*
375 *sciophilum*, between 0 and 0.06 in *O. bacaba* and between 0 and 0.01 in *O. bataua*, **Figure 2**
376 and **Supplementary material 6 (Tables S5.1 & S5.2)**. In *A. paramaca*, the population
377 originating from the undisturbed area close to the ring ditch in MC87 (i.e "MC87–") showed a
378 surprisingly high differentiation – ranging between 0.39 and 0.40 – with all other population of
379 this species, while the differentiation between all other populations ranged between 0 and 0.28.



380

381 **Figure 2.** Correlation plots showing patterns of genome-wide genetic differentiation across
382 populations of *Astrocaryum* spp. (left) and *Oenocarpus* spp. (right). "AP" refers to *A.*
383 *paramaca*, "AS" to *A. sciophilum*, "OBc" to *O. bacaba* and "OBt" to *O. bataua*.

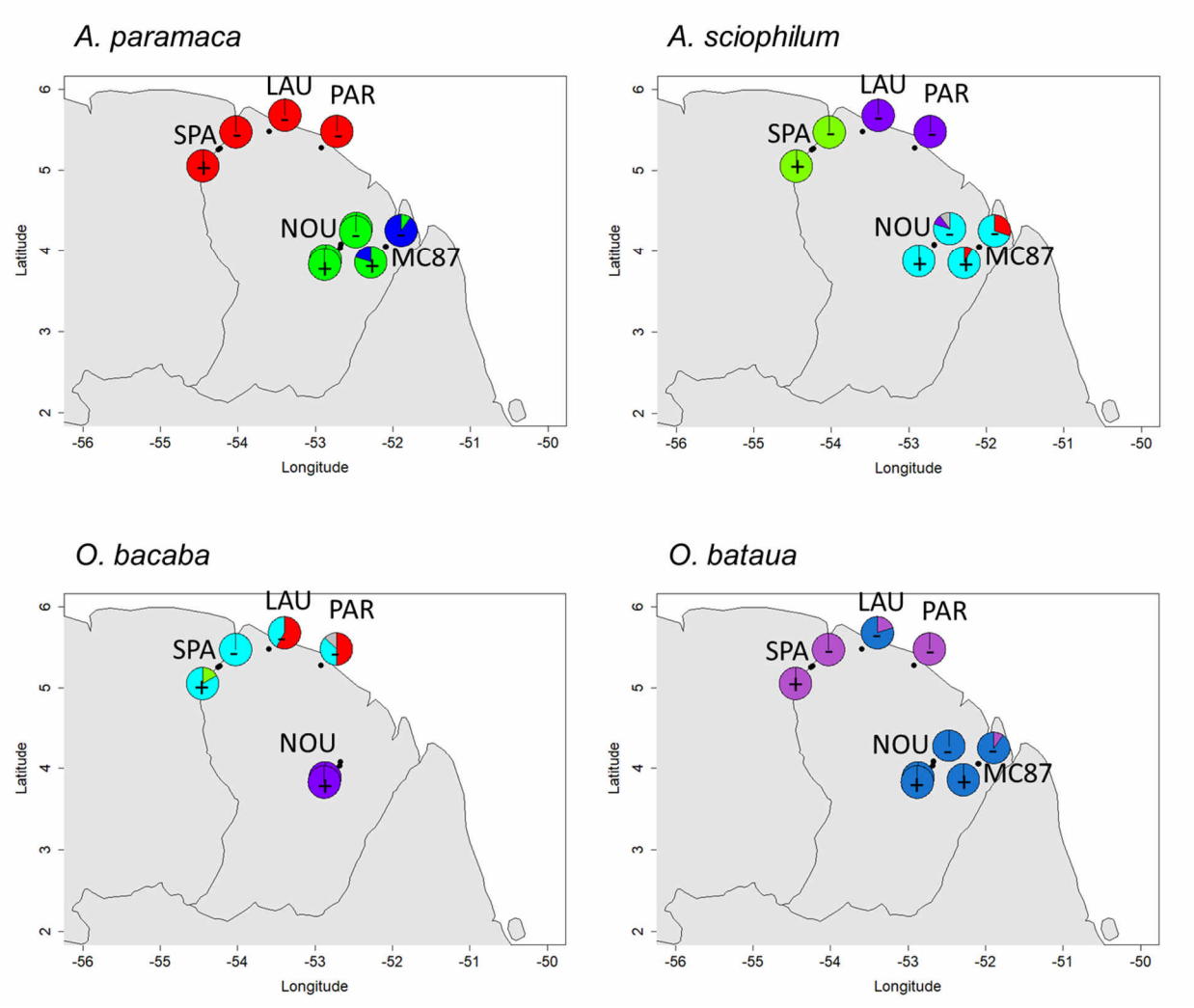
384 **Individual ancestry.** Cross-entropy criterion of individual ancestry was lower for K=3 and
385 K=4 in *A. paramaca* and *A. sciophilum*, respectively, and for K=4 and K=2 in *O. bacaba* and
386 *O.bataua*, respectively, thus indicating a genome-wide genetic structuring at a grain lower than
387 the species level (**Supplementary material 6, Figures S6.1 & S6.2**). Within each species, the
388 genetic structure was consistent with the west-east regional location of the palm populations,
389 **Figure 3**.

390 In *A. paramaca*, we detected one cluster in the west (in red) including all individuals from the
391 sites of Sparouine (SPA), Laussat (LAU) and Paracou (PAR), and two genetic clusters in the
392 east (sites of NOU and MC87), with all individuals associated to the blue group originating
393 from MC87.

394 In *A. sciophilum*, we detected two genetic clusters in the west: one corresponding to the study
395 site SPA (in green) and one (in purple) corresponding to the study sites LAU and PAR. In the
396 east (NOU and MC87), most individuals were assigned to the light blue group. Few individuals
397 originating from MC87 were assigned to a fourth group (in red). In addition, one individual was
398 assigned to the purple group and one ambiguous individual was not assigned to any group.

399 In *O. bacaba*, we detected three genetic groups in the west, and one in the east. Palms from
400 SPA were in majority assigned to the light blue group, while the populations from LAU and
401 PAR included palms assigned to both light blue and red groups. The third group (green)
402 corresponded to two individuals originating from the ring ditch (+) of SPA. The palms from
403 NOU were all assigned to the fourth group (purple).

404 In *O. bataua*, we detected two groups with one group (purple) including palms from SPA and
405 PAR, and another group (blue) corresponding to palms from NOU and MC87 except one palm
406 in the latter that clustered with the purple group with an ancestry coefficient slightly above 0.5.
407 In LAU, quite all palms clustered with the blue group except two palms clustered with the
408 purple group.



409

410 **Figure 3.** Geographic genetic structure. The pies indicate the proportion of individuals assigned
411 to genetic each cluster within each study site and local condition of pre-Columbian occupation.
412 Grey slices indicate ambiguous individuals that cannot be assigned to a given group with an
413 ancestry coefficient > 0.5.

414

415 **Isolation-by-distance.** Patterns of isolation by distance were significant in *A. sciophilum* (p-
416 value = 0.001) and *O.bataua* (p-value = 0.002), but not in *A. paramaca* (p-value = 0.20) nor in
417 *O. bacaba* (p-value = 0.14). In *A. paramaca*, however, IBD became significant (p-value =
418 0.003) after exclusion of populations from MC87 (MC87 + and MC87–), **Supplementary**
419 **material 6 (Figure S6.3)**.

420 **Admixture.** Significantly negative f_3 statistics were detected at a threshold of 5% for four target
421 populations of *Astrocaryum* spp. (*A. paramaca* from MC87 +, *A. sciophilum* from MC87 + and
422 MC87 – and *A. paramaca* from NOU_N –), and for one target population *of Oenocarpus* spp.
423 (*O.bacaba* in SPA +), **Supplementary material 6 (Figure S6.4)**. Populations of *A. paramaca*
424 from MC87 + appeared to be admixed between populations of *A. paramaca* from NOU (N and
425 S, + and –) and populations of *A. paramaca* from MC87 –, while *A. paramaca* from NOU_N –
426 appeared to be admixed between *A. paramaca* from NOU_S – and *A. sciophilum* from NOU_N
427 (+ and –), LAU – and/or PAR, thus suggesting patterns of inter-specific hybridization between
428 *A. paramaca* and *A. sciophilum* in the study site of NOU. Similar patterns of inter-specific
429 admixture also appeared in *A. sciophilum* from MC87 (+ and –), revealing admixture between
430 *A. sciophilum* from NOU_N (+ and –) and *A. paramaca* from MC87 –. *Oenocarpus bacaba*
431 from SPA + also appeared to be admixed between *O. bacaba* from SPA – and *O. bataua* from
432 different sites.

433 **Intra-population diversity.** In *Astrocaryum* spp., observed heterozygosity (*Ho*) ranged
434 between 0.136 (*A. sciophilum* in MC87 +) and 0.163 (*A. paramaca* in NOU_S –). In
435 *Oenocarpus* spp., *Ho* ranged between 0.351 (*O. bacaba* SPA – and LAU – ; *O.bataua* NOU_S
436 +, SPA – and PAR –) and 0.372 (*O. bacaba* SPA+). *Fis* was negative, indicating outcrossing.
437 *Fis* was slightly higher in *A. paramaca* than in *A. sciophilum* although non-significant
438 (Wilcoxon p-value = 0.1), and higher in *O. bacaba* than in *O. bataua* (Wilcoxon p-value =
439 0.0004), **Table 3** and **Supplementary material 6 (Figure S6.5)**. In *Astrocaryum* spp., *Fis* was
440 higher in the populations from MC 87 where it became non-significantly negative. *Fis* also tend
441 to be slightly higher in populations on ring ditches (+) than in surrounding natural populations
442 (–) in *Astrocaryum* spp. and *O. bataua* while the trend was opposite in *O. bacaba*, **Table 3** and
443 **Supplementary material 6 (Figure S6.5)**. A large proportion of SNPs shew significant
444 departure from HW equilibrium with large variations across populations ranging from 7% to

445 12% in *A. paramaca*, from 0 % to 14% in *A. sciophilum*, from 18% to 27% in *O. bacaba* and

446 from 0 to 25% in *O.bataua*.

447 **Table 3.** Intra-population diversity statistics (*He*, *Ho* and *Fis*) in *Astrocaryum* spp. with

448 Wilcoxon tests for differences between *Ho* and *He* (** : p-value <0.01 ; * : p-value<0.05).

	Ho	***He***	***Fis***
A. paramaca			
SPA –	0.156 (sd=0.292)	0.105	-0.333 (sd=0.381) **
SPA +	0.156 (sd=0.294)	0.105	-0.337 (sd=0.396) **
LAU –	0.160 (sd=0.301)	0.107	-0.365 (sd=0.401) **
PAR –	0.157 (sd=0.294)	0.106	-0.345 (sd=0.389) **
NOU_N –	0.165 (sd=0.290)	0.114	-0.273 (sd=0.362) **
NOU_N +	0.162 (sd=0.289)	0.112	-0.288 (sd=0.372) **
NOU_S –	0.163 (sd=0.296)	0.110	-0.340 (sd=0.371) **
NOU_S +	0.162 (sd=0.292)	0.111	-0.312 (sd=0.371) **
MC87 –	0.140 (sd=0.285)	0.104	-0.141 (sd=0.587)
MC87 +	0.154 (sd=0.278)	0.128	-0.101 (sd=0.556)
A. sciophilum			
SPA –	0.143 (sd=0.303)	0.089	-0.496 (sd=0.408) **
SPA +	0.139 (sd=0.289)	0.092	-0.384 (sd=0.407) *
LAU –	0.137 (sd=0.287)	0.090	-0.403 (sd=0.400) *
PAR –	0.137 (sd=0.285)	0.091	-0.363 (sd=0.402) *
NOU_N –	0.144 (sd=0.292)	0.094	-0.383 (sd=0.389) **
NOU_N +	0.137 (sd=0.285)	0.091	-0.366 (sd=0.405) *
MC87 –	0.137 (sd=0.257)	0.142	0.030 (sd=0.614) *
MC87 +	0.136 (sd=0.272)	0.110	-0.004 (sd=0.625)

449

450 **Table 4.** Intra-population diversity statistics (*He*, *Ho* and *Fis*) in *Oenocarpus* spp. with

451 Wilcoxon tests for differences between *Ho* and *He* (** : p-value <0.01 ; * : p-value<0.05).

	Ho	***He***	***Fis***
O. bacaba			
SPA –	0.352 (sd=0.415)	0.205	-0.575 (sd=0.389) **
SPA +	0.373 (sd=0.393)	0.228	-0.438 (sd=0.408) **
LAU –	0.351 (sd=0.381)	0.226	-0.427 (sd=0.436) **
PAR –	0.356 (sd=0.385)	0.228	-0.453 (sd=0.442) **
NOU_N +	0.353 (sd=0.411)	0.207	-0.535 (sd=0.400) **
NOU_S +	0.351 (sd=0.413)	0.206	-0.554 (sd=0.398) **
O. bataua			
SPA –	0.351 (sd=0.428)	0.198	-0.659 (sd=0.367) **
SPA +	0.353 (sd=0.427)	0.200	-0.639 (sd=0.376) **

LAU –	0.352 (sd=0.430)	0.198	-0.700 (sd=0.347) **
PAR –	0.352 (sd=0.427)	0.200	-0.686 (sd=0.353) **
NOU_N –	0.358 (sd=0.427)	0.203	-0.644 (sd=0.375) **
NOU_N +	0.361 (sd=0.423)	0.206	-0.614 (sd=0.384) **
NOU_S +	0.357 (sd=0.424)	0.204	-0.631 (sd=0.378) **
MC87 –	0.355 (sd=0.421)	0.204	-0.617 (sd=0.376) **
MC87 +	0.353 (sd=0.422)	0.202	-0.642 (sd=0.365) **

2. SNP-specific differentiation

`BayPass` core model provided accurate estimates of genome-wide genetic structure, SNP-specific differentiation (XtX_i) and derived statistics designed to disentangle footprints of divergent selection caused by precipitation (δ_i) at the regional scale, and pre-Columbian occupation (*i.e.* domestication) at the local scale (Δ_i). Covariance matrices estimated on observed data (Ω^{OBS}) were close to covariance matrices estimated on pseudo-observed data (Ω^{POD}), with $\overline{FMD}$ = 1.06 in *A. paramaca*, $\overline{FMD}$ = 1.25 in *A. sciophilum*, $\overline{FMD}$ = 2.07 in *O. bacaba* and $\overline{FMD}$ = 2.9 in *O. bataua*, **Supplementary material 6 (Figures S6.6 to S6.9)**. The genome-wide genetic structure estimated by `BayPass` through covariance matrices was also consistent with the genetic structure estimated by `snmf` through individual ancestry, confirming the robustness of `BayPass` to empirically infer the genomic structure across populations **Supplementary material 6 (Figures S6.6 to S6.10)**. We further detected SNPs displaying patterns of differentiation related to precipitation and/or pre-Columbian occupation based on the joint distribution of XtX_i and derived statistics (δ_i and Δ_i), **Figure 4** and **Supplementary material 6 (Figure S6.11)**.

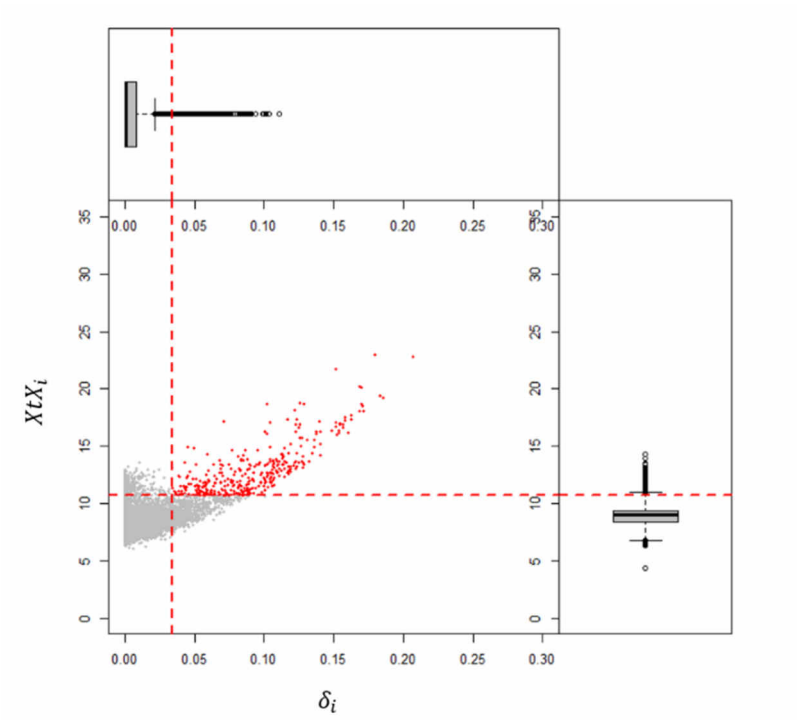


Figure 4. Example of SNP detection in relation to precipitation in *Oenocarpus bataua* based on the joint distribution of XtX_i and δ_i.

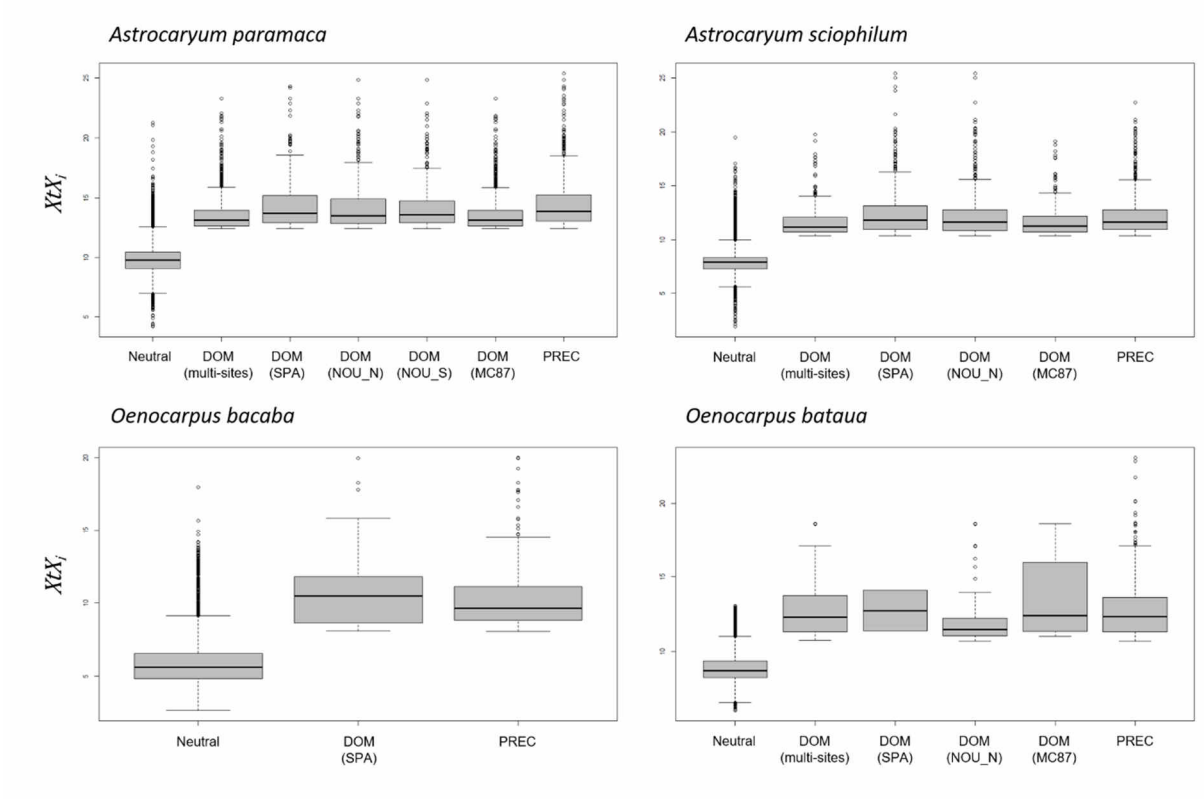
Between 0.74% (*A. sciophilum*) and 1.42% (*O. bataua*) climate adaptation SNPs were detected and between 0.20% (*O. bataua*) and 1.25% (*A. paramaca*) domestication SNPs, across all study sites. Moreover, the proportion of domestication SNPs was higher when considering all sites together (i.e. multi-site contrast) than within each site (single-site contrast), with many SNPs detected in one or two study sites only. Nevertheless, many domestication SNPs were shared between study sites: in *A. paramaca*, 1,252 domestication SNPs were detected in two sites, 285 were detected in three sites, and 28 in four sites ; in *A. sciophilum*, 564 SNPs were detected in two sites and 105 in three sites ; in *O. bataua*, only 10 SNPs were detected in two sites. The number of shared SNPs between pairs of study sites is provided in **Supplementary material 6 (Tables S6.3 to S6.6).**

482 **Table 5.** Number of over-differentiated SNPs detected in the four species in relation to climate
483 adaptation and domestication (out of 178,358 SNPs in *Astrocaryum* spp. and 23,252 SNPs in
484 *Oenocarpus* spp.). Because *O. bacaba* was only present on the ring ditches in NOU_N and
485 NOU_S and was absent from MC87, the detection of domestication SNP was based on a single
486 contrast in SPA.

	A. paramaca	***A. sciophilum***	***O. bacaba***	***O. bataua***
Climate adaptation SNPs				
Across all populations and sites	1345 (0.75%)	1320 (0.74%)	297 (1.25%)	331 (1.42%)
Domestication SNPs				
Across sites (multi-site contrast)	2238 (1.25%)	386 (0.22%)	NA	47 (0.20%)
Within sites (single-site contrast)				
• *SPA*	624 (0.35%)	640 (0.36%)	284 (1.22%)	2 (0.01%)
• *NOU N*	679 (0.38%)	810 (0.45%)	NA	70 (0.30%)
• *NOU S*	623 (0.35%)	NA	NA	NA
• *MC87*	2212 (1.24%)	374 (0.21%)	NA	16 (0.07%)

487

488 The median of SNP-specific differentiation (XtX_i) for neutral SNPs was 9.8 (Q90 = [8.6 ; 12.0])
489 in *A. paramaca*, 8.0 ([6.7 ; 9.9]) in *A. sciophilum*, 5.6 ([4.5 ; 8.1]) in *O. bacaba* and 8.7 ([7.6 ;
490 11.0]) in *O. bataua*. XtX_i ranged between 9.6 ([8.2 ; 14.6]) in *O. bacaba* and 13.9 ([12.5 ; 18.1])
491 in *A. paramaca* for climate adaptation SNP, and between 10.48 ([8.2 ; 13.8]) in *O. bacaba* and
492 13.1 ([12.5 ; 16.0]) in *A. paramaca* for domestication SNPs, without differences in the extent
493 of genetic differentiation between climate adaptation and domestication SNPs, **Figure 6**.



494

495 **Figure 6.** Distribution of SNP-specific differentiation (XtX_i) in neutral SNPs, domestication

496 (DOM) and climate adaptation SNPs (PREC).

497

498 Interestingly, several SNPs shew variations in allele frequency in relation to both precipitation

499 and pre-Columbian occupation (**Supplementary material 6**, **Tables S6.3 to S6.6**)), allowing

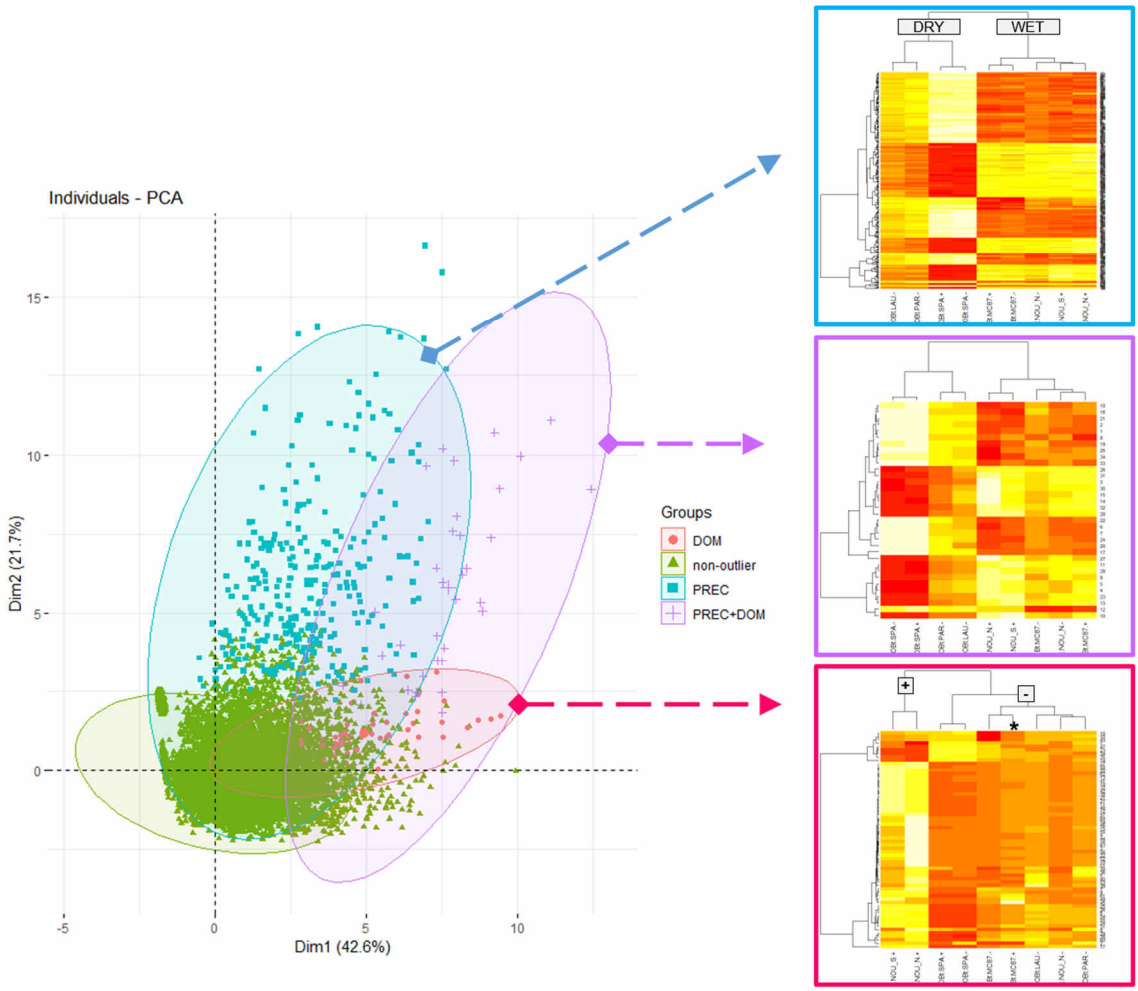
500 to distinguish four groups of SNPs within each species: (i) neutral SNPs, (ii) climate adaptation

501 SNPs showing clinal variations in allele frequency in relation to regional precipitation, (iii)

502 domestication SNPs showing contrasted allele frequency in relation to local pre-Columbian

503 occupation, and (iv) climate and domestication SNPs showing allele frequency variations in

504 relation to both, **Figure 7**.



505

506 **Figure 7.** Principal Component Analysis representing the multivariate projection of SNPs
507 based on SNP-specific estimates of XtX_i, δ_i, Δ_i and $\Delta_{i,s}$ in *Oenocarpus bataua*. Green: neutral
508 SNPs (*i.e.* non-outlier); blue: climate adaptation SNPs ("PREC") ; pink: domestication SNPs
509 ("DOM") ; violin: climate and domestication SNPs ("PREC+DOM"). Heatmaps on the right
510 show standardized allele frequencies in the different populations and for the different SNPs of
511 each group.

512 **Discussion**

513

514 This study opens a new research framework to document microevolution in non-model
515 Amazonian palms. We provide original observational genomic data describing the state and
516 distribution of genomic diversity in French Guiana, and we bring new knowledge about the
517 microevolutionary processes at play, disentangling neutral and selective processes on one side,
518 and natural and human-mediated ones on the other.

519 First of all, we must recall that all the four species were not present in all study sites and local
520 conditions of pre-Columbian occupation. This was particularly the case for *Oenocarpus* spp.
521 that were more difficult to find and to collect in natural forest areas than on ring ditches were
522 these species are frequent. As a result, we did not collect any *Oenocarpus* spp. in NOU_S –,
523 while *O. bacaba* was also absent from NOU_N – and MC87 (+ and –). This is consistent with
524 previous observations at the community level by Odonne et al. 2019 who reported an over-
525 abundance of palms in sites with strong evidence of pre-Columbian occupation compared to
526 natural forest landscapes. This constrained our study, as we had to deal with incomplete
527 sampling designs. Nevertheless, the amount of population-scale NGS data acquired was largely
528 sufficient to explore microevolutionary processes in the studied palm species, and to
529 disentangle natural and human-mediated ones.

530

531 **Neutral microevolution driven by gene flow and admixture.**

532 Patterns of genome-wide diversity revealed a west-east genetic structuring at the regional scale,
533 with similar trends between the four species. Genome-wide structure and differentiation are
534 commonly driven by neutral evolutionary processes, in particular gene flow. This is comforted
535 by significant isolation-by-distance in three species, indicating that regional patterns of
536 genomic diversity are primarily driven by gene flow, which gradually decreases with the

geographic distance between populations. We also detected intriguing patterns of genome-wide genetic structure in *Astrocaryum* spp. in the study site of MC87. In particular, the natural population (–) of *A. paramaca* revealed high levels of differentiation (about 40%) with populations from other study sites, while the population on the ring ditch was closer to populations from other study sites than to the population in the surrounding natural forest. A similar pattern was also observed in the congeneric species *A. sciophilum*, although to a lesser extent, suggesting a site-specific process probably involving inter-specific hybridization. This hypothesis was corroborated by admixture tests, which suggested inter-specific admixture between *A. paramaca* and *A. sciophilum* in the study site of MC87. Similarly, the genetic structure of *Oenocarpus bataua* revealed a site-specific structure in SPA+ corroborated by admixture tests suggesting inter-specific hybridization between *O. bacaba* and *O. bataua*. These site-specific patterns of inter-specific hybridization were neither expected nor anticipated when designing this study and carrying out the sampling. Most interestingly, *Astrocaryum* putative inter-specific hybrids were identified either as *A. paramaca* or as *A. sciophilum*, suggesting that they alternatively exhibit the morphological characteristics of one or the other species. The two putative *Oenocarpus* hybrids, on their side, were both identified as *Oenocarpus bacaba*. While there is no other species of *Oenocarpus* known in French Guiana, there is at least fifteen different species of *Astrocaryum* recorded on *GBIF* (GBIF Secretariat: GBIF Backbone Taxonomy. https://doi.org/10.15468/39omei Accessed via https://www.gbif.org/species/2738048 [30 March 2022]), raising the possibility of complex inter-specific hybridization with possible cryptic species in *Astrocaryum*. Even though we cannot conclude about the exact species involved in the pattern observed here, this preliminary observation would deserve in-depth investigation of inter-specific hybridization and introgression in the *Astrocaryum* complex.

561 Independently from site-specific hybridization patterns, estimates of intra-population genetic
562 diversity revealed high heterozygosity (H_o) and significantly negative F_{is}, indicating the
563 absence of inbreeding. These estimates are in line with previous knowledge about the breeding
564 system in Arecoideae that are known to be monoecious and predominantly outcrossed due to
565 protandry and/or genetic incompatibility (Henderson 1986; Eguiarte et al. 1992; Consiglio and
566 Bourne 2001; Picanço-Rodrigues et al. 2015; Gaiotto et al. 2003; Ottewell et al. 2012). The
567 genetic structure and diversity of populations on ring ditches was further close to populations
568 in the surrounding natural forests, indicating the absence of pre-Columbian impacts on
569 population genome-wide diversity and divergence. This result is not surprising, considering the
570 geographic proximity between "+" and "–" populations within study sites (< 5 km), which are
571 probably connected by extensive gene flow. The absence of genome-wide differentiation,
572 however, does not preclude the existence of SNP-specific selective processes driven by
573 adaptation and/or domestication.
574
575 **Climate adaptation and pre-Columbian domestication.**
576 Adaptation and domestication are acknowledged as important drivers of microevolution
577 worldwide, but they have been poorly documented in lowland Amazonia. While Central and
578 South America are known to be important centers of plant domestication, most studies have
579 focused on major crops such as maize (Kistler et al. 2018), papaya (Chávez-Pesqueira and
580 Núñez-Farfán 2017), manioc (Mühlen et al. 2019; Rival and McKey 2008; Alves-Pereira et al.
581 2020) or common bean (Rendón-Anaya et al. 2017), which provide the most obvious cases of
582 domestication. On the opposite, studies on minor crops are scarce although they are receiving
583 increasing attention since the last ten years (*e.g.* Brazil nut, Shepard and Ramirez 2011 ; piquiá,
584 Alves et al. 2016 ; peach palm, Clement et al. 2017 ; treegourd, Moreira et al. 2017). For its
585 part, the lack of interest in adaptation can be explained by the popularity of the neutral theory

586 under which the distribution of alleles within species (Kimura 1983) or species within
587 communities (Hubbell 2001) is assumed to be independent from fitness differences between
588 genotypes or species, being thus driven by stochastic mutation, demography and dispersal. In
589 hyperdiverse tropical rainforests, neutral models are powerful to predict global patterns of
590 diversity, as illustrated in the above section, and very useful to overcome the complexity of
591 niche differentiation modelling. Neutral and niche models are however highly complementary
592 (Chave 2004), and the neutral theory provides an elegant neutral hypothesis to test for niche
593 differentiation and adaptation (Leigh 2007). This is precisely the theoretical assumption of most
594 population genetics selection tests, which allows detecting gene- or SNP-specific departures
595 from the genome-wide, supposedly neutral, background of diversity and differentiation
596 (Beaumont and Nichols 1996).
597 By detecting over-differentiated SNPs in relation to precipitation and pre-Colombian
598 occupation, our results bring new evidence of the joint influence of climate adaptation and
599 ancient domestication in shaping patterns of genomic diversity across multiple spatial scales.
600 As expected, divergent selection related to adaptation and domestication affected a very small
601 fraction of the studied SNPs (always below 1.5%). Both the proportion of SNPs and the extent
602 of differentiation (XtX_i) were in the same order of magnitude between climate adaptation and
603 domestication SNPs, in spite of the different geographical scales involved (*i.e.* hundreds of
604 kilometers between study sites with different rainfall *versus* hundreds of meters between local
605 populations in contrasting conditions of pre-Columbian occupation). Regarding climate
606 adaptation, the present results are in line with a previous study in the hyperdominant tree *Eperua*
607 *falcata* which provided another evidence of climate adaptation along the sharp rainfall gradient
608 in coastal French Guiana (Brousseau et al. 2021).
609 Climate and domestication SNPs were located within or in the neighborhood (< 200bp) of many
610 genes whose exhaustive list is provided in **Supplementary material 7**. Although documenting

611 the functions of these genes in *Astrocaryum* spp. and *Oenocarpus* spp. is beyond the scope of
612 this study, we can nevertheless pick a few examples to illustrate the variety of genes involved.
613 For example, we detected climate adaptation SNPs in several genes of major importance for
614 plant development and response to stresses, including ABC transporters (Yazaki 2006; Do et
615 al. 2018), leaf rust disease resistance loci (Lim et al. 2015), Apetala2-like ethylene responsive
616 transcription factor (Licausi et al. 2013), aquaporins (G. Li et al. 2014; Maurel et al. 2015), and
617 several peroxidases (Blokhina et al. 2003; Yoshida et al. 2003). Several domestication SNPs
618 were detected in genes involved in lipid and sugar metabolism that could be of importance for
619 fruit development, including phospholipase D (Sun et al. 2011; Mao et al. 2007; Pinhero et al.
620 2003), acyl-transferases, fructose-bisphosphate aldolase (Feng et al. 2020; Yin et al. 2012),
621 phospholipid-transporting ATPases, and UDP-glucose 6-dehydrogenase 4.
622 Interestingly, many SNPs showed patterns of differentiation in relation to both climate
623 adaptation and domestication, indicating that these two processes are closely intertwined. In
624 other words, pre-Columbian societies had either domesticated "*in situ*" locally pre-adapted
625 populations or had to deal with adaptation during domestication. The absence of genome-wide
626 differentiation between palm populations on ring ditches and surrounding natural populations
627 tends to suggest that palm populations were collected in the forest near ring ditches. Moreover,
628 extensive gene flow between wild and domesticated populations had probably continuously
629 homogenized populations, thus preventing genome-wide differentiation and the fixation of the
630 so-called "domestication syndrome" in domesticated populations (see next section). Instead,
631 SNP-specific patterns of domestication observed here provide a new evidence of the subtle
632 influence of pre-Columbian societies and depicts what is commonly referred to as "incipient"
633 domestication (Lins Neto et al. 2014; Casas et al. 2016; Clement 1999). Gene flow since land
634 abandonment may also have partly erased patterns of differentiation caused by domestication
635 (*i.e.* de-domestication, Qiu et al. 2017). Moreover, many SNPs were detected in one study site

636 only, supporting that pre-Columbian impacts were partly site-specific. Nevertheless, many
637 SNPs were also detected in more than one study site, indicating some common trends in the
638 different study sites, probably caused by similar management practices. Whether observed
639 domestication patterns result from "conscious" domestication through direct selection for
640 desired traits or "unconscious" domestication through indirect local adaptation to
641 anthropogenic landscapes (Gepts 2010; Engelbrecht 1916; Darlington 1973) cannot however
642 be solved yet, and would require in-depth investigation at the phenotypic level to assess which
643 traits have been selected for or against in palm populations established on ring ditches.
644
645 **Moving beyond the domestication syndrome paradigm.**
646 The domestication of plants is a long-term process, with multiple possible outcomes. However,
647 most studies have focused on major crops exhibiting clear phenotypic differences with wild
648 populations, often underpinned by a sharp – genome-wide – genetic differentiation. Studies on
649 major crops led to the elaboration of an earlier model of domestication through which the
650 domestication is initiated by the isolation of a small population with reduced diversity.
651 Domesticated populations are then expected to accumulate phenotypic differences from their
652 wild ancestors under the action of strong selection (*i.e.* artificial selection for desired traits and
653 natural selection exerted by new environmental conditions) reinforced by inbreeding, both
654 resulting in the fast decrease in genetic diversity and in the fixation of phenotypic differences
655 in domesticated populations: the so-called "domestication syndrome" (Hammer 1984; Harlan
656 et al. 1973; Allaby 2014). The domestication syndrome has been extensively documented, and
657 quite over-emphasized (Kantar et al. 2017), leading to a narrow view of the domestication
658 process and to a misleading dichotomous distinction between domesticated varieties with a
659 clear domestication syndrome, and other (supposedly wild) populations. The domestication
660 process is, however, gradual through time and variable across populations (Meyer et al. 2012;
661 O. Smith et al. 2019; Kantar et al. 2017), and it is not necessarily expected to result in a sharp

– genome-wide – differentiation, in particular in cases of domestication "*in situ*" with extensive and continuous gene flow with wild populations (Meyer and Purugganan 2013). In a broader sense, the domestication process is better depicted as the modification of the evolutionary trajectory of a population through human-mediated evolutionary processes (including human-mediated dispersal, artificial selection for useful traits and/or indirect adaptation to anthropogenic landscapes) that can outcome – in most extreme cases – in the fixation of the domestication syndrome.

Because lowland Amazonia offers a great diversity of domestication histories and cases, from "incipiently-" or "semi-" domesticated populations to stable varieties (Lins Neto et al. 2014; Clement 1999), it offers an ideal study system to reconsider the domestication process beyond the domestication syndrome paradigm. Integrating the diversity of evolutionary histories underlying domestication further opens a new perspective toward a better understanding of microevolution in Amazonia, and the present study argues in favor of a paradigm shift to better integrate "incipiently-" and "semi-" domesticated minor crops in domestication studies.

677 **Supplementary materials**

678 **Supplementary material 1 (S1)**. ARECO5000+ Custom probe kit for targeted capture in
679 Arecoideae.

680 **Supplementary material 2 (S2)**. Molecular methods – library preparation, capture, sequencing
681 and adapters.

682 **Supplementary material 3 (S3).** Sequencing and bioinformatics (405 libraries) –
683 supplementary figures, tables and information.

684 **Supplementary material 4 (S4).** Fastq quality check.

685 **Supplementary material 5 (S5).** List of palm samples retained for genetic analyses.

686 **Supplementary material 6 (S6).** Population genetics – supplementary figures and tables.

687 **Supplementary material 7 (S7).** Genes containing climate adaptation and domestication
688 SNPs.

689

690 **Ethics statement**

691 This study was carried out in compliance with the Nagoya protocol and French regulations on
692 access to genetic resources with no direct commercial purposes. Palm samples were collected
693 for fundamental research following declaration to the French Ministry of the Ecological
694 Transition registered under the Internationally Recognized Certificate of Compliance ref.
695 ABSCH-IRCC-FR-245923-1 of 23 May 2018. The authors declare no commercial use of the
696 data generated and that no monetary benefit was derived from this study. Genomic data are
697 made available according to shared values and principles for Open Science.

698

699 **Data and codes availability**

700 Sequencing data (i.e. paired-end reads) are available on the European Nucleotide Archive under
701 accession PRJEB51800. Bioinformatics codes are available on GitLab under Project ID

702 34990830 (https://gitlab.com/LouiseBrousseau/palm-capture-bioinformatics). Data and code
703 remains under embargo until formal acceptance of the present article.
704
705 **Competing interests**
706 The authors declared no conflict of interest.
707
708 **Authors contribution**
709 LB and GO imagined the study and carried out palm sampling. LB designed the capture kit,
710 implemented the bioinformatics pipeline, and carried out bioinformatics and population
711 genetics. SS and AW carried out the molecular experiments (gDNA extraction and purification,
712 library preparation, targeted capture and sequencing of *O. bataua* librairies). All authors wrote
713 and approved the article.
714
715 **Acknowledgements**
716 This study has been conducted as part of the research program *PalmOmix* coordinated by L.
717 Brousseau. *PalmOmix* (ID 1702-01) was publicly funded by the French National Research
718 Agency (Agence Nationale de la Recherche, ANR) under the "*Investissement d'avenir*"
719 program with reference ANR-10-LABX-001-01 LabEx Agro and coordinated by Agropolis
720 Foundation under the frame of I-SITE MUSE (ANR-16-IDEX-0006). Sampling in the
721 Nouragues research Station has been supported by the LongTIme project coordinated by G.
722 Odonne and J-F. Molino and funded by an "*Investissement d'avenir*" grant from the French
723 National Research Agency (Agence Nationale de la Recherche, ANR) with reference CEBA:
724 ANR-10-LABX-25-01.
725 We are grateful to the Nouragues Ecological Research Station (CNRS) and the Réserve
726 Naturelle des Nouragues, to the Paracou Research Station (CIRAD) and to the National Forests

727 Office (ONF) in French Guiana for access to the study sites. We are also grateful to GenoToul
728 ("GeT" and bioinformatics) facilities for next-generation sequencing of *Astrocaryum* spp. and
729 *O. bacaba* libraries, and for access to computing facilities to carry out bioinformatics analyses.
730 We thank Thierry Joët, Stéphane Dussert and James Tregear for their help in pre-filtering
731 candidate proteins ("bottom-up approach") to design the custom capture kit. We also thank T.
732 Joët for his help in palm sampling and Anaïs Prud'homme who explored *O. bataua* NGS data
733 as part of her internship at UMR DIADE under the supervision of L. Brousseau.

References

Allaby, Robin G. 2014. "Domestication Syndrome in Plants." In *Encyclopedia of Global Archaeology*, 2182–84. Springer New York. https://doi.org/10.1007/978-1-4419-0465-2_2416.

Alves-Pereira, Alessandro, Charles R. Clement, Doriane Picanço-Rodrigues, Elizabeth Ann Veasey, Gabriel Dequigiovanni, Santiago Linorio Ferreyra Ramos, José Baldin Pinheiro, Anete Pereira de Souza, and Maria Imaculada Zucchi. 2020. "A Population Genomics Appraisal Suggests Independent Dispersals for Bitter and Sweet Manioc in Brazilian Amazonia." *Evolutionary Applications* 13 (2): 342–61. https://doi.org/10.1111/eva.12873.

Alves, Rubana P., Carolina Levis, and Charles R. Clement. 2016. "Use and Management of Piquiá Suggest *in Situ* Domestication along the Lower Tapajós River, Brazilian Amazonia." *Economic Botany* 70 (2): 198–202. https://doi.org/10.1007/s12231-016-9340-4.

Andrews, Simon. 2010. "FastQC: A Quality Control Tool for High Throughput Sequence Data."

Aronesty, Erik. 2013. "Comparison of Sequencing Utility Programs." *The Open Bioinformatics Journal* 7 (1): 1–8. https://doi.org/10.2174/1875036201307010001.

Baker, William J., and Thomas L.P. Couvreur. 2013. "Global Biogeography and Diversification of Palms Sheds Light on the Evolution of Tropical Lineages. I. Historical Biogeography." Edited by W. Daniel Kissling. *Journal of Biogeography* 40 (2): 274–85. https://doi.org/10.1111/j.1365-2699.2012.02795.x.

Beaumont, Mark A, and Richard A Nichols. 1996. "Evaluating Loci for Use in the Genetic Analysis of Population Structure." *Proceedings of The Royal Society B: Biological*

Sciences 263: 1619–26.

Blokhina, Olga, Eija Virolainen, and Kurt V. Fagerstedt. 2003. "Antioxidants, Oxidative Damage and Oxygen Deprivation Stress: A Review." *Annals of Botany* 91 (2): 179–94. https://doi.org/10.1093/aob/mcf118.

Bodin, Stephanie C., Jean François Molino, Guillaume Odonne, and Laurent Bremond. 2020. "Unraveling Pre-Columbian Occupation Patterns in the Tropical Forests of French Guiana Using an Anthracological Approach." *Vegetation History and Archaeobotany* 29 (5): 567–80. https://doi.org/10.1007/s00334-019-00767-w.

Brancier, Jeanne, Cecilia Cammas, Dominique Todisco, and Eric Fouache. 2014. "A Micromorphological Assessment of Anthropogenic Features in Pre-Columbian French Guiana Dark Soils (FGDS): First Results." *Zeitschrift Fur Geomorphologie* 58 (October): 109–39. https://doi.org/10.1127/0372-8854/2014/S-00159.

Brousseau, Louise, Paul V.A. Fine, Erwin Dreyer, Giovanni G. Vendramin, and Ivan Scotti. 2021. "Genomic and Phenotypic Divergence Unveil Microgeographic Adaptation in the Amazonian Hyperdominant Tree Eperua Falcata Aubl. (Fabaceae)." *Molecular Ecology* 30 (5): 1136–54.

Carson, John F, Jennifer Watling, Francis E Mayle, Bronwen S Whitney, José Iriarte, Heiko Prümers, and J Daniel Soto. 2015. "Pre-Columbian Land Use in the Ring-Ditch Region of the Bolivian Amazon." *The Holocene* 25 (8): 1285–1300. https://doi.org/10.1177/0959683615581204.

Carson, John Francis, Bronwen S. Whitney, Francis E. Mayle, José Iriarte, Heiko Prümers, J. Daniel Soto, and Jennifer Watling. 2014. "Environmental Impact of Geometric Earthwork Construction in Pre-Columbian Amazonia." *Proceedings of the National Academy of Sciences of the United States of America* 111 (29): 10497–502.

https://doi.org/10.1073/pnas.1321770111.

Casas, Alejandro, José Blancas, Adriana Otero-Arnaiz, Jeniffer Cruse-Sanders, Rafael Lira, Aidé Avendaño, Fabiola Parra, et al. 2016. "Evolutionary Ethnobotanical Studies of Incipient Domestication of Plants in Mesoamerica." In , 257–85. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-6669-7_11.

Castro, J. W., M. S. Hernández, and R. H. Gutiérrez. 2014. "Uses for Amazonian Seje Oil (*Oenocarpus Bataua*) Extracted by Mechanical and Manual Methods." In *Acta Horticulturae*, 1047:335–40. International Society for Horticultural Science. https://doi.org/10.17660/actahortic.2014.1047.43.

Cavers, Stephen, and Christopher W. Dick. 2013. "Phylogeography of Neotropical Trees." *Journal of Biogeography* 40 (4): 615–17. https://doi.org/10.1111/jbi.12097.

Chave, Jerome. 2004. "Neutral Theory and Community Ecology." *Ecology Letters* 7 (august): 241–53. https://doi.org/10.1111/j.1461-0248.2003.00566.x.

Chávez-Pesqueira, Mariana, and Juan Núñez-Farfán. 2017. "Domestication and Genetics of Papaya: A Review." *Frontiers in Ecology and Evolution* 5: 155. https://www.frontiersin.org/article/10.3389/fevo.2017.00155.

Clement, Charles R. 1999. "1942 and the Loss of Amazonian Crop Genetic Resources. I. The Relation between Domestication and Human Population Decline." *Economic Botany*. SpringerNew York Botanical Garden Press. https://doi.org/10.2307/4256179.

Clement, Charles R., Joseph M. McCann, and Nigel J. H. Smith. 2006. "Agrobiodiversity in Amazônia and Its Relationship with Dark Earths." In *Amazonian Dark Earths*, 159–78. Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-2597-1_9.

Clement, Charles R, Michelly de Cristo-Araújo, Geo Coppens d'Eeckenbrugge, Vanessa Maciel dos Reis, Romain Lehnebach, and Doriane Picanço-Rodrigues. 2017. "Origin and

Dispersal of Domesticated Peach Palm." *Frontiers in Ecology and Evolution* 5: 148. https://www.frontiersin.org/article/10.3389/fevo.2017.00148.

Clement, Charles R, Michelly de Cristo-Araújo, Geo Coppens D'Eeckenbrugge, Alessandro Alves Pereira, and Doriane Picanço-Rodrigues. 2010. "Origin and Domestication of Native Amazonian Crops." *Diversity* 2 (1): 72–106. https://doi.org/10.3390/d2010072.

Consiglio, Trisha K., and Godfrey R. Bourne. 2001. "Pollination and Breeding System of a Neotropical Palm *Astrocaryum Vulgare* in Guyana: A Test of the Predictability of Syndromes." *Journal of Tropical Ecology* 17 (4): 577–92. https://doi.org/10.1017/S0266467401001432.

Coomes, Oliver T. 2004. "Rain Forest 'Conservation-through-Use'? Chambira Palm Fibre Extraction and Handicraft Production in a Land-Constrained Community, Peruvian Amazon." *Biodiversity and Conservation* 13 (2): 351–60. https://link.springer.com/article/10.1023/B:BIOC.0000006503.90980.e8.

Cummings, Anthony R., and Jane M. Read. 2016. "Drawing on Traditional Knowledge to Identify and Describe Ecosystem Services Associated with Northern Amazon's Multiple-Use Plants." *International Journal of Biodiversity Science, Ecosystem Services & Management* 12 (1–2): 39–56. https://doi.org/10.1080/21513732.2015.1136841.

Cunha, Vânia Maria Borges, Marcilene Paiva da Silva, Sérgio Henrique Brabo de Sousa, Priscila do Nascimento Bezerra, Eduardo Gama Ortiz Menezes, Nayara Janaina Neves da Silva, Débora Ariane Dornelas da Silva Banna, Marilena Emmi Araújo, and Raul Nunes de Carvalho Junior. 2019. "Bacaba-de-Leque (*Oenocarpus Distichus* Mart.) Oil Extraction Using Supercritical CO2 and Bioactive Compounds Determination in the Residual Pulp." *Journal of Supercritical Fluids* 144 (February): 81–90. https://doi.org/10.1016/j.supflu.2018.10.010.

830 Danecek, Petr, James K. Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O.
831 Pollard, Andrew Whitwham, et al. 2021. "Twelve Years of SAMtools and BCFtools."
832 *GigaScience* 10 (2). https://doi.org/10.1093/gigascience/giab008.

833 Darlington, Cyril Dean. 1973. *Chromosome Botany and the Origins of Cultivated Plants*.
834 London: Allen & Unwin.

835 Davy, Damien. 2007. *"Vannerie et Vanniers" Approche Ethnologique d'une Activité*
836 *Artisanale En Guyane Française*. Université d'Orleans.

837 Do, Thanh Ha Thi, Enrico Martinoia, and Youngsook Lee. 2018. "Functions of ABC
838 Transporters in Plant Growth and Development." *Current Opinion in Plant Biology*.
839 Elsevier Ltd. https://doi.org/10.1016/j.pbi.2017.08.003.

840 Eguiarte, Luis E., Nidia Perez-Nasser, and Daniel Piñero. 1992. "Genetic Structure,
841 Outcrossing Rate and Heterosis in *Astrocaryum Mexicanum* (Tropical Palm):
842 Implications for Evolution and Conservation." *Heredity* 69 (3): 217–28.
843 https://doi.org/10.1038/hdy.1992.119.

844 Engelbrecht, Th. H. 1916. "Über Die Entstehung Einiger Feldmäßig Angebauter
845 Kulturpflanzen." *Geographische Zeitschrift* 22 (6): 328–34.
846 http://www.jstor.org/stable/27809914.

847 Fadiman, Maria G. 2008. "Use of Mocora, *Astrocaryum Standleyanum* (Arecaceae), by Three
848 Ethnic Groups in Ecuador: Differences, Similarities and Market Potential." *Journal of*
849 *Ethnobiology* 28 (1): 92–109. https://doi.org/10.2993/0278-
850 0771(2008)28[92:UOMASA]2.0.CO;2.

851 Feng, Dan, Li Gao, Yusheng Zheng, Dongdong Li, and Peng Zhou. 2020. "Molecular Cloning
852 and Function Characterisation of a Cytoplasmic Fructose-1, 6-Bisphosphate Aldolase
853 Gene from Coconut (*Cocos Nucifera* L.)." *Journal of Horticultural Science and*

Biotechnology 95 (6): 703–11. https://doi.org/10.1080/14620316.2020.1749139.

Förstner, Wolfgang, and Boudewijn Moonen. 2003. "A Metric for Covariance Matrices." In *Geodesy-The Challenge of the 3rd Millennium*, 299–309. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-05296-9_31.

Frichot, Eric, François Mathieu, Théo Trouillon, Guillaume Bouchard, and Olivier François. 2014. "Fast and Efficient Estimation of Individual Ancestry Coefficients." *Genetics* 196: 973–983. https://doi.org/10.1534/genetics.113.160572.

Gaiotto, Fernanda A, Dario Grattapaglia, and Roland Vencovsky. 2003. "Genetic Structure, Mating System, and Long-Distance Gene Flow in Heart of Palm (Euterpe Edulis Mart.)." *Journal of Heredity* 94 (5): 399–406. https://doi.org/10.1093/jhered/esg087.

García, Néstor, Gloria Galeano, Laura Mesa, Nicolás Castaño, Henrik Balslev, Rodrigo Bernal, Néstor García, et al. 2015. "Management of the Palm *Astrocaryum Chambira* Burret (Arecaceae) in Northwest Amazon." *Acta Botanica Brasilica* 29 (1): 45–57. https://doi.org/10.1590/0102-33062014abb3415.

Gepts, Paul. 2010. "Crop Domestication as a Long-Term Selection Experiment." In *Plant Breeding Reviews*, 1–44. Oxford: John Wiley & Sons, Inc. https://doi.org/10.1002/9780470650288.ch1.

Günther, Torsten, and Graham Coop. 2013. "Robust Identification of Local Adaptation from Allele Frequencies." *Genetics* 195 (1): 205–20. https://doi.org/10.1534/genetics.113.152462.

Gutierrez, Maria Paz. 2020. "The Northwestern Amazon Malocas: Craft Now and Then." *Journal of Material Culture* 25 (1): 3–35. https://doi.org/10.1177/1359183519836141.

Hammer, Karl. 1984. "Das Domestikationssyndrom (The Domestication Syndrome)." *Die Kulturpflanze* 32 (1): 11–34. https://doi.org/10.1007/BF02098682.

878 Harlan, Jack R, J M J de Wet, and E Glen Price. 1973. "Comparative Evolution of Cereals."
879 *Evolution* 27 (2): 311–25. https://doi.org/10.2307/2406971.

880 Henderson, Andrew. 1986. "A Review of Pollination Studies in the Palmae." *The Botanical*
881 *Review* 52 (3): 221–59. https://doi.org/10.1007/BF02860996.

882 Hoorn, Carina, Frank P. Wesselingh, Hans Ter Steege, M. A. Bermudez, A. Mora, J. Sevink,
883 I. Sanmartín, et al. 2010. "Amazonia through Time: Andean Uplift, Climate Change,
884 Landscape Evolution, and Biodiversity." *Science* 330 (6006): 927–31.
885 https://doi.org/10.1126/science.1194585.

886 Hubbell, Stephen P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*.
887 Princeton University Press.

888 Jombart, Thibaut. 2008. "Adegenet: A R Package for the Multivariate Analysis of Genetic
889 Markers." *Bioinformatics* 24 (11): 1403–5.
890 https://doi.org/10.1093/bioinformatics/btn129.

891 Jombart, Thibaut, and Ismaïl Ahmed. 2011. "Adegenet 1.3-1: New Tools for the Analysis of
892 Genome-Wide SNP Data." *Bioinformatics* 27 (21): 3070–71.
893 https://doi.org/10.1093/bioinformatics/btr521.

894 Joshi, N A, and Joseph Fass. 2011. "Sickle: A Sliding-Window, Adaptive, Quality-Based
895 Trimming Tool for FastQ Files (Version 1.33)."

896 Kahn, Francis. 1993. "Amazonian Palms: Food Resources for the Management of Forest
897 Ecosystems." *Tropical Forests, People and Food: Biocultural Interactions and*
898 *Applications to Development* 13: 153.

899 Kantar, Michael B, Amber R Nashoba, Justin E Anderson, Benjamin K Blackman, and Loren
900 H Rieseberg. 2017. "The Genetics and Genomics of Plant Domestication." *BioScience*
901 67 (11): 971–82. https://doi.org/10.1093/biosci/bix114.

Kimura, Motoo. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511623486.

Kistler, Logan, S. Yoshi Maezumi, Jonas Gregorio De Souza, Natalia A. S. Przelomska, Flaviane Malaquias Costa, Oliver Smith, Hope Loiselle, et al. 2018. "Multiproxy Evidence Highlights a Complex Evolutionary Legacy of Maize in South America." *Science* 362 (6420): 1309–13. https://doi.org/10.1126/science.aav0207.

Leigh, E. G. 2007. "Neutral Theory: A Historical Perspective." *Journal of Evolutionary Biology* 20 (6): 2075–91. https://doi.org/10.1111/j.1420-9101.2007.01410.x.

Levis, Carolina, Flávia R C Costa, Frans Bongers, Marielos Peña-Claros, Charles R Clement, André B Junqueira, Eduardo G Neves, Eduardo K Tamanaha, Fernando O G Figueiredo, and Rafael P Salomão. 2017. "Persistent Effects of Pre-Columbian Plant Domestication on Amazonian Forest Composition." *Science* 355 (6328): 925–31. https://doi.org/10.1126/science.aal0157.

Li, Guowei, Véronique Santoni, and Christophe Maurel. 2014. "Plant Aquaporins: Roles in Plant Physiology." *Biochimica et Biophysica Acta - General Subjects*. Elsevier. https://doi.org/10.1016/j.bbagen.2013.11.004.

Li, Heng. 2011. "A Statistical Framework for SNP Calling, Mutation Discovery, Association Mapping and Population Genetical Parameter Estimation from Sequencing Data." *Bioinformatics* 27 (21): 2987–93. https://doi.org/10.1093/bioinformatics/btr509.

Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows–Wheeler Transform." *Bioinformatics* 25 (14): 1754–60. https://doi.org/10.1093/bioinformatics/btp324.

———. 2010. "Fast and Accurate Long-Read Alignment with Burrows–Wheeler Transform." *Bioinformatics* 26 (5): 589–95. https://doi.org/10.1093/bioinformatics/btp698.

926 Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth,
927 Goncalo Abecasis, Richard Durbin, and Genome Project Data Processing Subgroup.
928 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25 (16):
929 2078–79. https://doi.org/10.1093/bioinformatics/btp352.

930 Licausi, Francesco, Masaru Ohme-Takagi, and Pierdomenico Perata. 2013.
931 "APETALA2/Ethylene Responsive Factor (AP2/ERF) Transcription Factors: Mediators
932 of Stress Responses and Developmental Programs." *New Phytologist* 199 (3): 639–49.
933 https://doi.org/10.1111/nph.12291.

934 Lim, Chae Woo, Seung Hwan Yang, Ki Hun Shin, Sung Chul Lee, and Sang Hyon Kim.
935 2015. "The AtLRK10L1.2, Arabidopsis Ortholog of Wheat LRK10, Is Involved in ABA-
936 Mediated Signaling and Drought Resistance." *Plant Cell Reports* 34 (3): 447–55.
937 https://doi.org/10.1007/s00299-014-1724-2.

938 Lins Neto, Ernani M., Nivaldo Peroni, Alejandro Casas, Fabiola Parra, Xitlali Aguirre, Susana
939 Guillén, and Ulysses P. Albuquerque. 2014. "Brazilian and Mexican Experiences in the
940 Study of Incipient Domestication." *Journal of Ethnobiology and Ethnomedicine*. BioMed
941 Central Ltd. https://doi.org/10.1186/1746-4269-10-33.

942 Llamas, Bastien, Lars Fehren-Schmitz, Guido Valverde, Julien Soubrier, Swapan Mallick,
943 Nadin Rohland, Susanne Nordenfelt, et al. 2016. "Ancient Mitochondrial DNA Provides
944 High-Resolution Time Scale of the Peopling of the Americas." *Science Advances* 2 (4):
945 e1501385. https://doi.org/10.1126/sciadv.1501385.

946 Mantel, N. 1967. "The Detection of Disease Clustering and a Generalized Regression
947 Approach." *Cancer Research*.

948 Mao, Linchun, Huaqing Pang, Guoze Wang, and Chenggang Zhu. 2007. "Phospholipase D
949 and Lipoxygenase Activity of Cucumber Fruit in Response to Chilling Stress."

Postharvest Biology and Technology 44 (1): 42–47. https://doi.org/10.1016/j.postharvbio.2006.11.009.

Mascher, Martin, Todd A. Richmond, Daniel J. Gerhardt, Axel Himmelbach, Leah Clissold, Dharanya Sampath, Sarah Ayling, et al. 2013. "Barley Whole Exome Capture: A Tool for Genomic Research in the Genus Hordeum and Beyond." *Plant Journal* 76 (3): 494–505. https://doi.org/10.1111/tpj.12294.

Maurel, Christophe, Yann Boursiac, Doan Trung Luu, Véronique Santoni, Zaigham Shahzad, and Lionel Verdoucq. 2015. "Aquaporins in Plants." *Physiological Reviews* 95 (4): 1321–58. https://doi.org/10.1152/physrev.00008.2015.

Meyer, Rachel S., Ashley E. DuVal, and Helen R. Jensen. 2012. "Patterns and Processes in Crop Domestication: An Historical Review and Quantitative Analysis of 203 Global Food Crops." *New Phytologist* 196 (1): 29–48. https://doi.org/10.1111/j.1469-8137.2012.04253.x.

Meyer, Rachel S., and Michael D. Purugganan. 2013. "Evolution of Crop Species: Genetics of Domestication and Diversification." *Nature Reviews Genetics* 14 (12): 840–52. https://doi.org/10.1038/nrg3605.

Moreira, Priscila Ambrósio, Cédric Mariac, Leila Zekraoui, Marie Couderc, Doriane Picanço Rodrigues, Charles R. Clement, and Yves Vigouroux. 2017. "Human Management and Hybridization Shape Treegourd Fruits in the Brazilian Amazon Basin." *Evolutionary Applications* 10 (6): 577–89. https://doi.org/10.1111/eva.12474.

Mühlen, Gilda Santos, Alessandro Alves-Pereira, Cássia Regina Limonta Carvalho, André Braga Junqueira, Charles R. Clement, and Teresa Losada Valle. 2019. "Genetic Diversity and Population Structure Show Different Patterns of Diffusion for Bitter and Sweet Manioc in Brazil." *Genetic Resources and Crop Evolution* 66 (8): 1773–90.

https://doi.org/10.1007/s10722-019-00842-1.

Odonne, Guillaume, Martijn Bel, Maxime Burst, Olivier Brunaux, Miléna Bruno, Etienne Dambrine, Damien Davy, et al. 2019. "Long-term Influence of Early Human Occupations on Current Forests of the Guiana Shield." *Ecology* 100 (10): e02806. https://doi.org/10.1002/ecy.2806.

Onstein, Renske E, William J Baker, Thomas L P Couvreur, Søren Faurby, Leonel Herrera-Alsina, Jens-Christian Svenning, and W Daniel Kissling. 2018. "To Adapt or Go Extinct? The Fate of Megafaunal Palm Fruits under Past Global Change." *Proceedings. Biological Sciences* 285 (1880): 20180882. https://doi.org/10.1098/rspb.2018.0882.

Ottewell, K, E Grey, F Castillo, and J Karubian. 2012. "The Pollen Dispersal Kernel and Mating System of an Insect-Pollinated Tropical Palm, *Oenocarpus Bataua*." *Heredity* 109 (6): 332–39. https://doi.org/10.1038/hdy.2012.40.

Patterson, Nick, Priya Moorjani, Yontao Luo, Swapan Mallick, Nadin Rohland, Yiping Zhan, Teri Genschoreck, Teresa Webster, and David Reich. 2012. "Ancient Admixture in Human History." *Genetics* 192 (3): 1065–93. https://doi.org/10.1534/genetics.112.145037.

Peter, Benjamin M. 2016. "Admixture, Population Structure and F-Statistics." *Genetics* 4: 1485–1501. https://doi.org/10.1534/genetics.115.183913.

Petitjean Roget, Hugues. 1991. "50 Sites de Montagnes En Guyane Française: Contribution A L'inventaire Des Sites Archeologiques D'Emile Abonnenc." In *Proceedings of the Twelfth Congress of the International Association for Caribbean Archaeology*, 241–58.

Picanço-Rodrigues, Doriane, Spartaco Astolfi-Filho, Maristerra R Lemes, Rogerio Gribel, Alexandre M Sebbenn, and Charles R Clement. 2015. "Conservation Implications of the Mating System of the Pampa Hermosa Landrace of Peach Palm Analyzed with

998 Microsatellite Markers." *Genetics and Molecular Biology*. scielo.

999 Pinhero, Reena Grittle, Kurt C. Almquist, Zuzana Novotna, and Gopinadhan Paliyath. 2003.
1000 "Developmental Regulation of Phospholipase D in Tomato Fruits." *Plant Physiology and*
1001 *Biochemistry* 41 (3): 223–40. https://doi.org/10.1016/S0981-9428(03)00014-7.

1002 Puerari, Cláudia, Karina Teixeira Magalhães-Guedes, and Rosane Freitas Schwan. 2015.
1003 "Bacaba Beverage Produced by Umutina Brazilian Amerindians: Microbiological and
1004 Chemical Characterization." *Brazilian Journal of Microbiology* 46 (4): 1207–16.
1005 https://doi.org/10.1590/S1517-838246420140964.

1006 Qiu, Jie, Yongjun Zhou, Lingfeng Mao, Chuyu Ye, Weidi Wang, Jianping Zhang, Yongyi Yu,
1007 et al. 2017. "Genomic Variation Associated with Local Adaptation of Weedy Rice
1008 during De-Domestication." *Nature Communications* 8 (May): 15323.
1009 https://doi.org/10.1038/ncomms15323.

1010 R Core Team, R Core Team. 2021. "R: A Language and Environment for Statistical
1011 Computing." In *R Foundation for Statistical Computing, Vienna, Austria*. https://www.r-
1012 project.org/.

1013 Reich, David, Kumarasamy Thangaraj, Nick Patterson, Alkes L. Price, and Lalji Singh. 2009.
1014 "Reconstructing Indian Population History." *Nature* 461 (7263): 489–94.
1015 https://doi.org/10.1038/nature08365.

1016 Rendón-Anaya, Martha, Josaphat M. Montero-Vargas, Soledad Saburido-Álvarez, Anna
1017 Vlasova, Salvador Capella-Gutierrez, José Juan Ordaz-Ortiz, O. Mario Aguilar, et al.
1018 2017. "Genomic History of the Origin and Domestication of Common Bean Unveils Its
1019 Closest Sister Species." *Genome Biology* 18 (1): 60. https://doi.org/10.1186/s13059-017-
1020 1190-6.

1021 Rival, Laura, and Doyle McKey. 2008. "Domestication and Diversity in Manioc (*Manihot*

Esculenta Crantz Ssp. *Esculenta*, Euphorbiaceae).” *Current Anthropology* 49 (6): 1119–28. https://doi.org/10.1086/593119.

Rohland, Nadin, and David Reich. 2012. “Cost-Effective, High-Throughput DNA Sequencing Libraries for Multiplexed Target Capture.” *Genome Research* 22 (5): 939–46. https://doi.org/10.1101/gr.128124.111.

Rothhammer, Francisco, and Tom D. Dillehay. 2009. “The Late Pleistocene Colonization of South America: An Interdisciplinary Perspective.” *Annals of Human Genetics*. Ann Hum Genet. https://doi.org/10.1111/j.1469-1809.2009.00537.x.

Shepard, Glenn H., and Henri Ramirez. 2011. “‘Made in Brazil’: Human Dispersal of the Brazil Nut (*Bertholletia Excelsa*, Lecythidaceae) in Ancient Amazonia.” *Economic Botany* 65 (1): 44–65. https://doi.org/10.1007/s12231-011-9151-6.

Smith, Brian Tilston, John E. McCormack, Andrés M. Cuervo, Michael. J. Hickerson, Alexandre Aleixo, Carlos Daniel Cadena, Jorge Pérez-Emán, et al. 2014. “The Drivers of Tropical Speciation.” *Nature* 515 (7527): 406–9. https://doi.org/10.1038/nature13687.

Smith, Nigel. 2015. *Palms and People in the Amazon*. Geobotany Studies. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-05509-1.

Smith, Oliver, William V. Nicholson, Logan Kistler, Emma Mace, Alan Clapham, Pamela Rose, Chris Stevens, et al. 2019. “A Domestication History of Dynamic Adaptation and Genomic Deterioration in *Sorghum*.” *Nature Plants* 5 (4): 369–79. https://doi.org/10.1038/s41477-019-0397-9.

Sosnowska, Joanna, Adam Walanus, and Henrik Balslev. 2015. “Asháninka Palm Management and Domestication in the Peruvian Amazon.” *Human Ecology* 43 (3): 451–66. https://doi.org/10.1007/s10745-015-9745-1.

Steege, Hans ter, Nigel C A Pitman, Daniel Sabatier, Christopher Baraloto, Rafael P Salomão,

1046 Juan Ernesto Guevara, Oliver L Phillips, et al. 2013. "Hyperdominance in the
1047 Amazonian Tree Flora." *Science* 342 (6156).

1048 Sun, Jian, Xiangrong You, Li Li, Hongxiang Peng, Weiqiang Su, Changbao Li, Quanguang
1049 He, and Fen Liao. 2011. "Effects of a Phospholipase D Inhibitor on Postharvest
1050 Enzymatic Browning and Oxidative Stress of Litchi Fruit." *Postharvest Biology and*
1051 *Technology* 62 (3): 288–94. https://doi.org/10.1016/j.postharvbio.2011.07.001.

1052 Thioulouse, Jean, Daniel Chessel, Sylvain Dolédec, and Jean Michel Olivier. 1997. "ADE-4:
1053 A Multivariate Analysis and Graphical Display Software." *Statistics and Computing* 7
1054 (1): 75–83. https://doi.org/10.1023/A:1018513530268.

1055 Wallace, Alfred Russel. 1853. *Palm Trees of the Amazon and Their Uses*. J. van Voorst.

1056 Watling, Jennifer, José Iriarte, Francis E. Mayle, Denise Schaan, Luiz C.R. Pessenda, Neil J.
1057 Loader, F. Alayne Street-Perrott, Ruth E. Dickau, Antonia Damasceno, and Alceu Ranzi.
1058 2017. "Impact of Pre-Columbian Geoglyph Builders on Amazonian Forests."
1059 *Proceedings of the National Academy of Sciences of the United States of America* 114
1060 (8): 1868–73. https://doi.org/10.1073/pnas.1614359114.

1061 Yazaki, Kazufumi. 2006. "ABC Transporters Involved in the Transport of Plant Secondary
1062 Metabolites." *FEBS Letters*. No longer published by Elsevier.
1063 https://doi.org/10.1016/j.febslet.2005.12.009.

1064 Yin, Yuxin, Xiaowei Zhang, Yongjun Fang, Linlin Pan, Gaoyuan Sun, Chengqi Xin,
1065 Mohammed M.Ba Abdullah, et al. 2012. "High-Throughput Sequencing-Based Gene
1066 Profiling on Multi-Staged Fruit Development of Date Palm (*Phoenix Dactylifera*, L.)."
1067 *Plant Molecular Biology* 78 (6): 617–26. https://doi.org/10.1007/s11103-012-9890-5.

1068 Yoshida, K., P. Kaothien, T. Matsui, A. Kawaoka, and A. Shinmyo. 2003. "Molecular
1069 Biology and Application of Plant Peroxidase Genes." *Applied Microbiology and*

1070 *Biotechnology*. Springer Verlag. https://doi.org/10.1007/s00253-002-1157-7.

1071