



HAL
open science

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

Louise Brousseau, Sylvain Santoni, Audrey Weber, Guillaume Odonne

► To cite this version:

Louise Brousseau, Sylvain Santoni, Audrey Weber, Guillaume Odonne. Natural and human-mediated drivers of microevolution in Neotropical palms: a historical genomics approach. 2024. hal-04633302

HAL Id: hal-04633302

<https://hal.science/hal-04633302v1>

Preprint submitted on 3 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

13 **Abstract**

- 14 • Microevolution in Amazonia has been far less documented than in temperate regions
15 and still constitutes a major knowledge gap. Moreover, the actual influence of pre-
16 Columbian societies through the Holocene remains unclear, although it could provide
17 interesting insights on the relationship between natural and human-mediated
18 evolutionary processes in driving microevolution.
- 19 • Because they are widespread and traditionally managed since pre-Columbian times,
20 Neotropical palms are choice models to investigate the drivers of microevolution in
21 Amazonia. In this study, we carried out a preliminary exploration of the genomic
22 diversity in two pairs of congeneric palm species in French Guiana (*Astrocaryum* spp.
23 and *Oenocarpus* spp.). We built upon an original sampling design taking into account
24 both regional climate variations and local-scale pre-Columbian occupation, and carried
25 out an extensive target capture experiment and next-generation sequencing of hundreds
26 of palms.
- 27 • We explored patterns of genomic diversity and differentiation within and across
28 populations, bringing state-of-the-art knowledge about the roles of climate adaptation
29 and of pre-Columbian domestication. By documenting original cases of “incipient”
30 domestication, our results further open new avenues toward a better understanding of
31 microevolution in Amazonia and of the impacts of pre-Columbian societies in having
32 shaped present-day patterns of biodiversity.

33

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

36

37 **Introduction**

38

39 Evolution in Amazonia has been far less documented than in temperate regions and still
40 constitutes a major knowledge gap. Evolutionary studies in Amazonia also largely concentrated
41 on macroevolution through long-term geological periods, providing interesting insights onto
42 speciation and extinction dynamics in relation to landscape changes associated with the Andean
43 uplift and the formation of the Amazon river (Hoorn et al. 2010), to species dispersal abilities
44 (B. T. Smith et al. 2014) and to biotic interactions (Onstein et al. 2018). On the contrary, less
45 attention has been paid to the microevolution of plants and animals during the Quaternary
46 (Cavers and Dick 2013), and even less during the Holocene. This period, which started about
47 11,000 years ago, experienced both a global warming subsequent to the last glacial period
48 resulting in the present climate, and the establishment of dense networks small-scale human
49 societies subsequent to the human colonization of South America in late Pleistocene (Llamas
50 et al. 2016; Rothhammer and Dillehay 2009). A better understanding of the drivers of plant and
51 animal evolution during the Holocene could thus bring interesting insights on the interactions
52 between natural and human-mediated evolutionary processes and their role shaping present-day
53 patterns of biodiversity.

54

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

62 subsistence food and building materials – have been extensively documented at least since the
63 1850’s (Wallace 1853; N. Smith 2015). Surprisingly, they are also particularly abundant at
64 archaeological sites with evidence of long-term human occupation, and some species (in
65 particular *Astrocaryum* spp. and *Oenocarpus* spp.) are commonly credited as ‘biological
66 indicators’ of ancient settlements (Odonne et al. 2019).

67

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

79

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

87

88 We carried out a preliminary exploration of the genomic diversity in two pairs of congeneric
89 Neotropical palms in coastal French Guiana: *Astrocaryum paramaca* and *A. sciophilum* on one
90 side, *Oenocarpus bacaba* and *O. bataua* on the other side. We sampled populations following
91 an original sampling design taking into account both regional and local-scale evolutionary
92 processes related to climate variations and pre-Columbian occupation, respectively. At the
93 regional scale, we sampled populations in study sites with contrasted rainfall to catch climate
94 adaptation. At the local scale, we sampled palms both on the top of ring ditches and in the
95 surrounding forests to catch the impact of domestication associated with local anthropogenic
96 disturbances. Then, we carried out a comprehensive targeted-capture experiment by targeting
97 exonic regions within thousands of nuclear genes followed by next-generation sequencing of
98 hundreds of libraries.

99

100 Here, we provide new data on the state and distribution of genomic diversity in Neotropical
101 palms in coastal French Guiana. We investigate intra-specific patterns of diversity and
102 differentiation driven by both neutral evolutionary processes (gene flow and admixture) and
103 selective ones, disentangling the roles of (natural) climate adaptation and of (human-mediated)
104 pre-Columbian domestication that operated at the different spatial scales.

105

106

107 **Materials and Methods**

108

109 **1. Palm species**

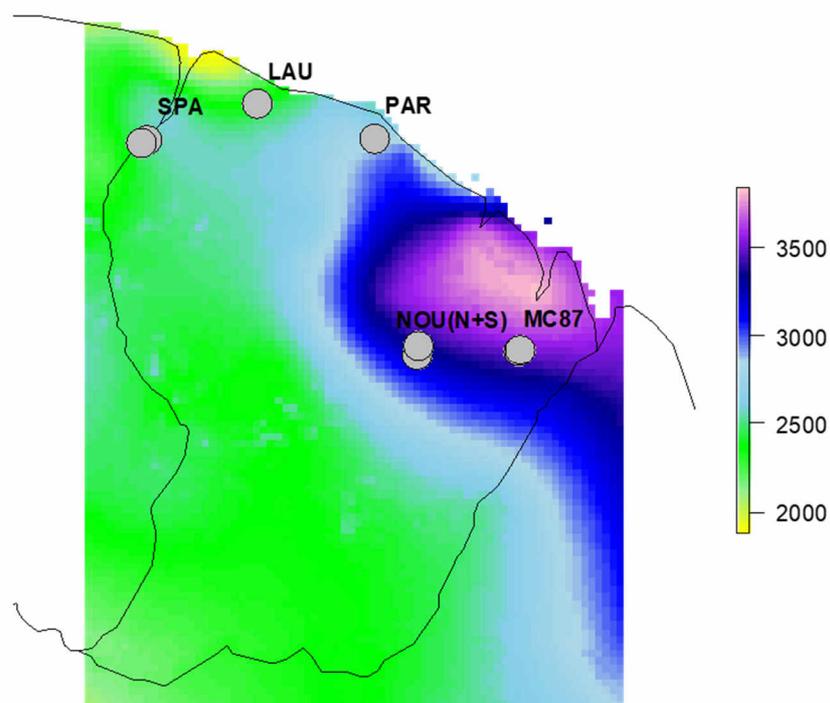
110 We focused on four palm species that belong to the Arecoideae sub-family and that are
111 widespread in forest landscapes of French Guiana: two species of the genus *Astrocaryum* (*A.*
112 *paramaca* and *A. sciophilum*) and two species of the genus *Oenocarpus* (*O. bacaba* and *O.*
113 *bataua*). *Oenocarpus bacaba* and *O. paramaca* are pan-Amazonian species that are broadly
114 distributed from Colombia to north Brazil. *Oenocarpus bataua* stands at the 7th position
115 amongst the 20 most abundant species in the Amazonian tree flora, while *O. bacaba* stands at
116 the 33th position (ter Steege et al. 2013). The distribution of *A. paramaca* and *A. sciophilum* is
117 restricted to northeast Amazonia (Suriname, French Guiana and Brazil, GBIF Secretariat: GBIF
118 Backbone Taxonomy. <https://doi.org/10.15468/39omei>). Despite this, *A. sciophilum* stands at
119 the 51th position amongst the most abundant species in forest tree inventories. Moreover, *O.*
120 *bacaba* and *A. sciophilum* are two biological indicators of pre-Columbian settlements in French
121 Guiana, as their abundance is significantly higher in forest plots with evidence of strong pre-
122 Columbian anthropogenic disturbances than in natural plots (Odonne et al. 2019). *Astrocaryum*
123 spp. and *Oenocarpus* spp. are major sources of food and of building and handicraft materials in
124 Amazonia (Sosnowska et al. 2015; Cummings and Read 2016). Their leaves are traditionally
125 used in basketry and to build houses (see for example maloca's walls and roof, García et al.
126 2015; Davy 2007; Fadiman 2008; Coomes 2004; Gutierrez 2020; Kahn 1993), while their fruits
127 are used to make oil (*e.g.* 'seje' oil) and beverages (*e.g.* so called 'comou' and 'wassai' in
128 French Guiana and 'bacaba wine' in Northwest Amazonia, Cunha et al. 2019; Castro et al.
129 2014; Puerari et al. 2015).

130

131 **2. Sampling strategy**

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

140



141

142 **Figure 1.** Localization of the study sites and annual precipitation (in mm.year⁻¹) in French
143 Guiana. Current precipitations were extracted from WorldClim version 1 using R (libraires
144 “raster” and “sp”) and are representative of the period 1960-1990
145 (<http://www.worldclim.com/version1>).

146 **Table 1.** Localization of populations in the different sampling sites and local conditions of pre-
147 Columbian occupation. “+” refers the top of the ring ditch with evidence of pre-Columbian
148 occupation, and “-” refers the surrounding sampling forest area without evidence of pre-
149 Columbian occupation. “Nouragues North” and “Nouragues South” indicate the localization of
150 two ring ditches in the north and in the south of the Nouragues Ecological research station,
151 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

152

153 3. Probes design, targeted capture and sequencing

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

162 preparation, targeted-capture and sequencing, resulting in the sampling design described in
 163 **Table 2** and totaling 405 libraries (one per palm). Genomic library preparation for multiplexed
 164 individuals and targeted capture by bait hybridization followed published protocols (Rohland
 165 and Reich 2012; Mascher et al. 2013) with some modifications. The molecular protocol is
 166 provided in **Supplementary material 2**. Briefly, for each genotype a genomic barcoded library
 167 was built; 48 barcoded libraries were mixed together for the capture by hybridization using the
 168 myBaits Custom DNA-Seq© kit “ARECO5000+”. Libraries of *Oenocarpus bataua* were
 169 sequenced (paired-end, 2 × 150bp) on an Illumina MiSeq at the AGAP genotyping platform
 170 (Montpellier, France), while libraries of *Astrocaryum* spp. and *Oenocarpus bacaba* were
 171 sequenced by the Get-PlaGe core facility (GenoToul platform, INRA Toulouse, France
 172 <http://www.genotoul.fr>) on an Illumina HiSeq3000. Sequence data are available on EMBL-
 173 ENA under accession PRJEB51800.

174

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

Site ID	Local condition	<i>Astrocaryum paramaca</i>	<i>Astrocaryum sciophilum</i>	<i>Oenocarpus bacaba</i>	<i>Oenocarpus bataua</i>
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

178

179 4. Bioinformatics

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

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

188 Reads were further cleaned using `FASTX-Toolkit` (RRID:SCR_005534) v.0.0.14:
189 they were trimmed with `fastx_trimmer` with option `-Q 33 -f 10 -l 140`, resulting
190 in reads of 140-nucleotide long. Individual bases of low quality (phred-score < 25) were then
191 masked with `fastq_masker` with options `-q 25 -Q 33`. At last, read 3'-end and 5'-end
192 were cut and orphan reads were discarded using `sickle-1.33` (Joshi and Fass 2011) with
193 options `pe -l 100` (*i.e.* default quality threshold of 20 and length threshold of 100bp). Once
194 cleaned, reads were quality controlled with `FastQC-0.11.5` (Andrews 2010).

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

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

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

205 Alignments were then sorted and indexed using `samtools sort` and `samtools index`,
206 respectively. Mapping statistics were computed using `samtools stats`, `samtools`
207 `flagstat` and `samtools bedcov`. Sequencing and mapping statistics are provided in
208 **Supplementary materials 3 and 4.**

209 Following a first mapping against *Elaeis guineensis* reference genome, the reference was
210 updated by editing genus-specific consensus (one for *Astrocaryum* spp. and one for *Oenocarpus*
211 spp.) to take into account the genomic divergence between *Elaeis guineensis* and the genera
212 under study. This step aimed to avoid detecting polymorphisms caused by inter-genus variation
213 in the following. This was achieved by calling variants using `samtools mpileup` with
214 options `-uf` for each *Astrocaryum* spp. and *Oenocarpus* spp. alignments, respectively. The two
215 consensus references were edited with `bcftools index` (v.1.3) and `bcftools`
216 `consensus` (Danecek et al. 2021). Then, the libraries of each species (*Astrocaryum*
217 *paramaca*, *A. sciophilum*, *Oenocarpus bacaba* and *O. bataua*) were mapped separately against
218 the corresponding genus-specific consensus reference following the mapping procedure
219 described above.

220 **SNP and genotype calling.** Variants were called using `samtools-1.9` and `bcftools-`
221 `1.3`: genus-specific consensus were indexed with `samtools faidx` and variant were called
222 with `samtools mpileup` with options `-g -t DP` followed by `bcftools call` with
223 options `-vc`. The resulting `.vcf` file was filtered with `grep` to discard monomorphic sites,
224 variants detected on unassembled scaffolds or chloroplast as well as indels to retain nuclear
225 single-nucleotide polymorphisms (SNPs) only. SNPs were post-filtered with R `version`
226 `3.5.1` (R Core Team 2021) to retain SNPs and genotypes with a quality threshold of 20, and

227 a per-library coverage threshold of 6X. SNPs and genotypes that did not meet the criteria were
228 edited as missing (“NA”), and SNPs that became monomorphic or contained more than 20% of
229 missing values were discarded. This resulted in 178,358 and 23,252 high-quality SNPs in
230 *Astrocaryum* spp. and *Oenocarpus* spp., respectively. Last, palm samples with more than 20%
231 of missing values were discarded. The list of samples retained for genetic analyses is provided
232 in **Supplementary material 5**.

233

234 **5. Population genetics**

235

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

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

243 **Individual ancestry.** Genome-wide genetic structure was assessed for each species by
244 estimating individual ancestry (Frichot et al. 2014) with the function `snmf` (package `LEA`) with
245 arguments `ploidy = 2`, `entropy = TRUE` and `CPU = 8`. The genetic structure was
246 explored from $K=1$ to $K=10$ ancestral populations (*A. paramaca*), to $K=8$ (*A. sciophilum*), to
247 $K=6$ (*O. bacaba*) or to $K=9$ (*O. bataua*). The most probable number of clusters was retained
248 based on cross-entropy criterion (lower values indicating better fit to the predictions) and each
249 palm sample was assigned to a given group with an ancestry threshold of 0.5.

250 **Isolation-by-distance.** Isolation-by-distance was tested across populations within each species
251 through Mantel tests between genetic and geographic distances. Genetic distance were

252 estimated with the function `dist.genpop` (package `adegenet`) with argument `method =`
253 `3` (*i.e.* coancestry coefficient or Reynolds' distance, Jombart 2008; Jombart and Ahmed 2011).
254 Geographic distances (euclidean) were computed from geographic coordinates using the
255 function `dist` with default arguments. Mantel tests between geographic and genetic distance
256 matrices were computed with the function `mantel.randtest` (package `ade4`, Mantel
257 1967; Thioulouse et al. 1997).

258 **Admixture tests.** Admixture between populations was evaluated through f_3 admixture tests for
259 all combinations of three populations (Patterson et al. 2012; Peter 2016; Reich et al. 2009):
260 $f_3(P_X; P_A, P_B)$, where P_X is the target population and P_A and P_B are two source populations.
261 Admixture tests were encoded with R `version 3.5.1`. f_3 statistic was estimated following
262 the equation by Reich et al. (2009):

$$263 \quad \mathbb{E}(p_X - p_a)(p_X - p_b),$$

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

269 **Intra-population diversity.** Observed (H_o) and expected heterozygosity (H_e) within each
270 species and populations was estimated with `adegenet` from `genind` objects. `Genind`
271 objects containing multiple populations were split into separate populations using the function
272 `seppop` and summary statistics were exported with the function `summary`. Whether H_o was
273 higher than H_e was assessed for each population through non-parametric Wilcoxon tests
274 (because of non-homogeneity of variances between H_o and H_e) with the function
275 `wilcox.test`. Inbreeding coefficients (F_{is}) were estimated following the formula: $F_{is} =$
276 $1 - \frac{H_o}{H_e}$. Hardy Weinberg equilibrium was tested for each SNP through Chi-squared tests with

277 the function `hw.test` (package `pegas`), and the proportion of SNPs showing significant
278 departure from HW equilibrium at a 1% threshold was assessed.

279 **Selection test.** We carried out a genome-wide scan for adaptive differentiation within each
280 species using `BayPass v.2.1`. Genotype data were converted into allele counts and
281 `BayPass` core model was run a first time using observed SNP counts as input. The number of
282 populations in each dataset was specified with option `-npop` and MCMC parameters were
283 specified with options `-nthreads 8 -thin 15 -burnin 2500 -npilot 15 -`
284 `pilotlength 5000`. The estimated covariance matrix (Ω), which provides an accurate
285 estimate of genome-wide allele frequency covariance across populations, was converted into a
286 correlation matrix in `R version 3.5.1` using the function `cov2cor`. The singular values
287 of the covariance matrix was decomposed using `svd` and left singular vectors (U) were
288 extracted to plot the genome-wide structure across populations. Pseudo-observed data (PODs)
289 of 10,000 SNPs (for *Oenocarpus* spp.) or 100,000 SNPs (for *Astrocaryum* spp.) were simulated
290 with the `simulate.baypass` function (`BayPass` miscellaneous R function), specifying
291 the total number of allele counts for each SNP (*i.e.* sample size) and the estimated
292 hyperparameters a_π and b_π (*i.e.* posterior mean and standard deviation of the Beta distribution
293 expressing the mean reference allele frequency, ~ancestral reference allele frequency) as input.
294 These PODs correspond to a simulated set of SNP counts that take into account the genome-
295 wide (neutral) genetic structure empirically inferred from observed SNPs counts. The core
296 model was run again using PODs as input under the same parameters to calibrate SNP-specific
297 statistics (standardized allele frequencies α^{std} , XtX and derived statistics) under neutral
298 hypothesis. `BayPass` outputs were post-processed with `R version 3.5.1`. The
299 robustness of the core model was assessed by evaluating the FMD distance (Förstner and
300 Moonen 2003) between covariance matrices estimated on observed data (Ω^{OBS}) and on
301 simulated data (Ω^{POD}) using `fmd.dist` function (`BayPass` miscellaneous R function). SNP-

302 specific departures from genome-wide (*i.e.* neutral) differentiation were detected based on the
303 joint distribution of:

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

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

$$314 \quad \alpha_{i,j}^{STD(OBS)} = \frac{\alpha_{i,j}^{STD(OBS)} - \min(\alpha^{STD})}{\max(\alpha^{STD}) - \min(\alpha^{STD})} \quad \text{and} \quad \alpha_{i,j}^{STD(POD)} = \frac{\alpha_{i,j}^{STD(POD)} - \min(\alpha^{STD})}{\max(\alpha^{STD}) - \min(\alpha^{STD})},$$

315 where:

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

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

$$321 \quad \delta_i = |\beta_i| \cdot r_i^2,$$

322 where $|\beta_i|$ and r_i^2 are the absolute regression coefficient and the r-squared of the linear model
323 between allele frequency at SNP i and precipitation: $\alpha_{i,j}^{STD} = \beta_i \cdot prec_j$, with $prec_j$ the

324 precipitation in population j (scaled between 0 and 1). Linear models were fitted for each SNP
 325 with the function `lm`.

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

$$328 \quad \Delta_{i,s} = |\alpha_{i,(j+)}^{STD} - \alpha_{i,(j-)}^{STD}|,$$

329 and

$$330 \quad \Delta_i = \sum_s (\Delta_{i,s})$$

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

334 δ and Δ statistics were separately computed on observed data (δ_i^{OBS} , $\Delta_{i,s}^{OBS}$, Δ_i^{OBS}) and on PODs
 335 (δ_i^{POD} , $\Delta_{i,s}^{POD}$, Δ_i^{POD}). A SNP was retained as being shaped by climate adaptation if it was both
 336 over-differentiated across all populations whatever their local environment based on XtX at a
 337 95% PODs threshold (*i.e.* condition 1), and if it displayed clinal allele frequency variation with
 338 regional precipitation based on δ at a 95% PODs threshold (*i.e.* condition 2). Similarly, a SNP
 339 was retained as being shaped by domestication if both condition 1 was met, and if it displayed
 340 local differences in allele frequencies in relation to pre-Columbian occupation based on Δ at a
 341 95% PODs 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: $\Delta_i^{OBS} > q_{(95)}(\Delta_i^{POD})$

342

343 where:

- 344 - XtX_i^{OBS} is the SNP-specific differentiation estimated on observed data, and
345 $q_{(95)}(XtX_i^{POD})$ the 95-quantile of the distribution (across all SNPs) of SNP-specific
346 differentiation estimated on simulated data ;
- 347 - δ_i^{OBS} is the SNP-specific cline in allele frequency along the regional precipitation
348 gradient estimated on observed data, and $q_{(95)}(\delta_i^{POD})$ the 95-quantile of the distribution
349 (across all SNPs) of SNP-specific clines in allele frequency along the regional
350 precipitation gradient estimated on simulated data ;
- 351 - Δ_i^{OBS} is the SNP-specific difference in allele frequency between local conditions of pre-
352 Columbian occupation (+ versus -) estimated on observed data, and $q_{(95)}(\Delta_i^{POD})$ the
353 95-quantile of the distribution (across all SNPs) of SNP-specific differences in allele
354 frequency between local conditions of pre-Columbian occupation estimated on
355 simulated data.

356 **Outlier blastn.** Last, we retrieved the identity of genes containing precipitation and adaptation
357 SNPs. For each species, we extracted the sequences of the genomic regions surrounding each
358 outlier SNP (+/- 200 bp) from consensus genomic sequences using `bedtools getfasta`
359 (`bedtools-2.27.1`). Each SNP-surrounding region was then blasted against the NCBI
360 nucleotide database using `blastn` (NCBI BLAST+ v. 2.10.1+), limiting blast search by
361 taxonomy (*Elaeis guineensis* taxid = 51953 and *Phoenix dactylifera* taxid = 42345) and
362 returning the first good hit to optimize computation time.

363

364 **Results**

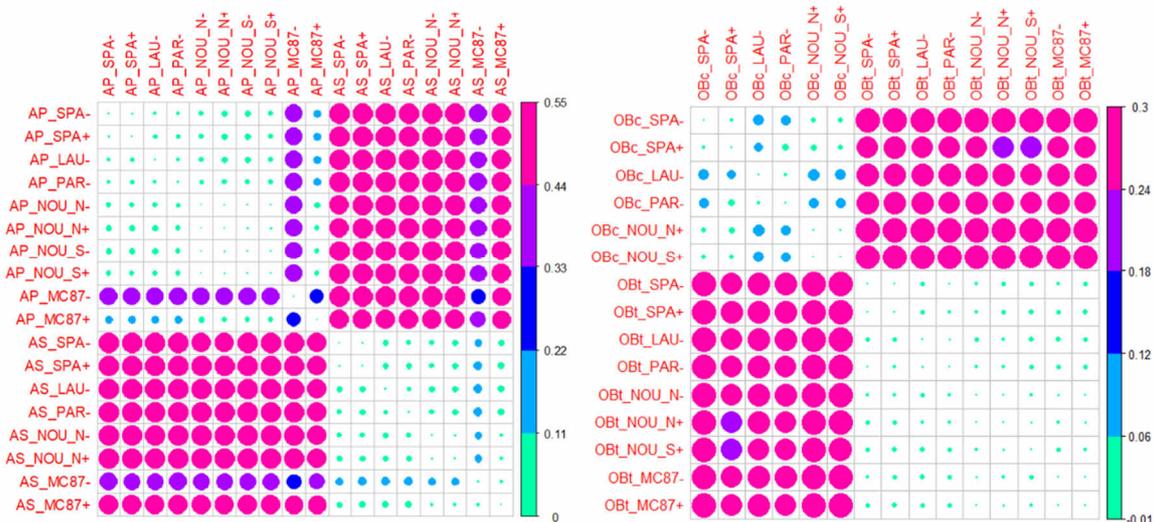
365

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

367

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

376



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

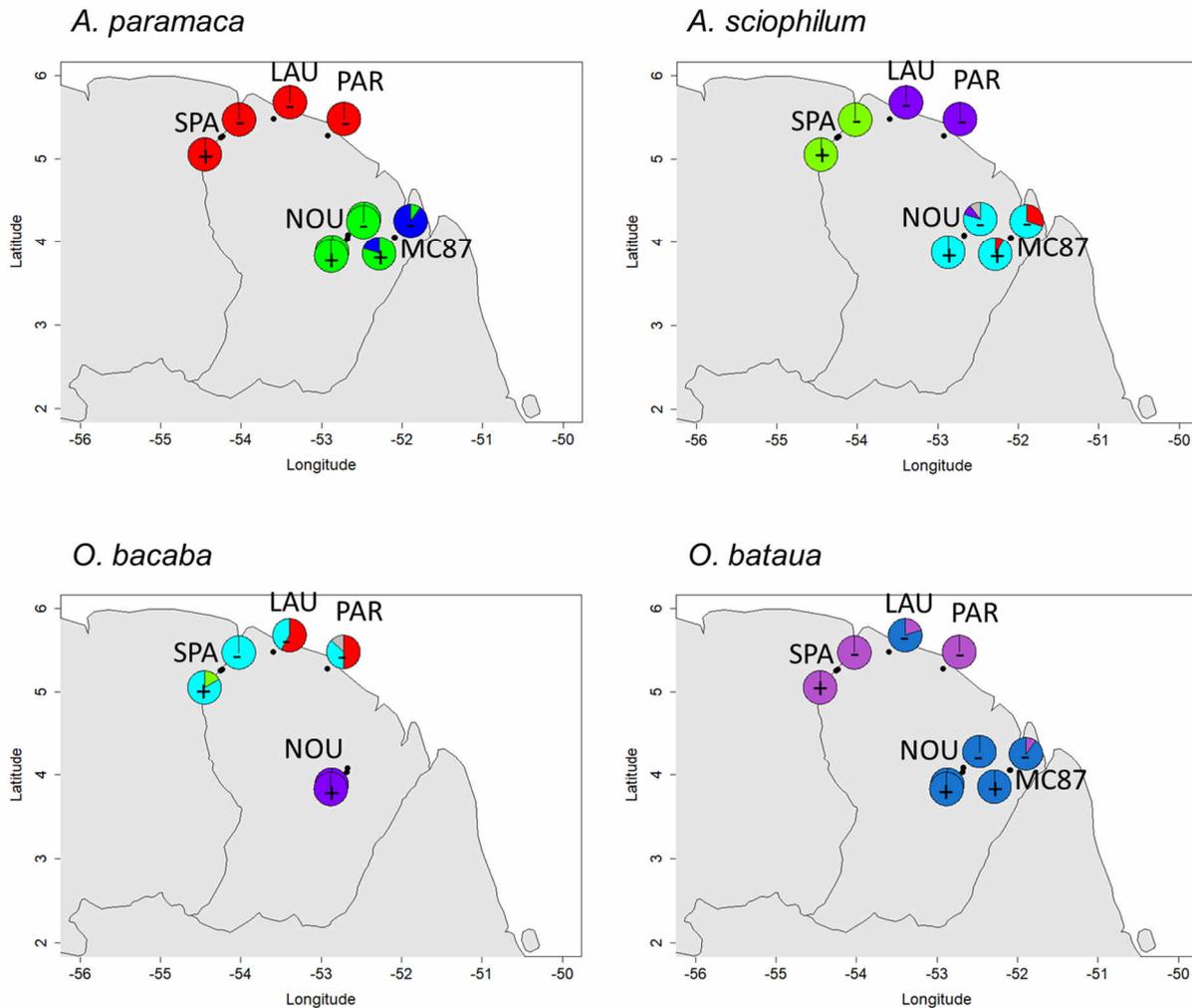
380 **Individual ancestry.** Cross-entropy criterion of individual ancestry was lower for K=3 and
381 K=4 in *A. paramaca* and *A. sciophilum*, respectively, and for K=4 and K=2 in *O. bacaba* and
382 *O. bataua*, respectively, thus indicating a genome-wide genetic structuring at a grain lower than
383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

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

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

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

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



405

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

410

411 **Isolation-by-distance.** Patterns of isolation by distance were significant in *A. sciophilum* (p-
412 value = 0.001) and *O. bataua* (p-value = 0.002), but not in *A. paramaca* (p-value = 0.20) nor in
413 *O. bacaba* (p-value = 0.14). In *A. paramaca*, however, IBD became significant (p-value =
414 0.003) after exclusion of populations from MC87 (MC87 + and MC87-), **Supplementary**
415 **material 6 (Figure S6.3).**

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

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

441 12% in *A. paramaca*, from 0 % and 14% in *A. sciophilum*, from 18% to 27% in *O. bacaba* and
 442 between 0 and 25% in *O. bataua*.

443 **Table 3.** Intra-population diversity statistics (*He*, *Ho* and *Fis*) in *Astrocaryum* spp. with
 444 Wilcoxon tests for differences between *Ho* and *He* (** : p-value <0.01 ; * : p-value<0.05).

	<i>Ho</i>	<i>He</i>	<i>Fis</i>
<i>A. paramaca</i>			
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)
<i>A. sciophilum</i>			
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)

445

446 **Table 4.** Intra-population diversity statistics (*He*, *Ho* and *Fis*) in *Oenocarpus* spp. with
 447 Wilcoxon tests for differences between *Ho* and *He* (** : p-value <0.01 ; * : p-value<0.05).

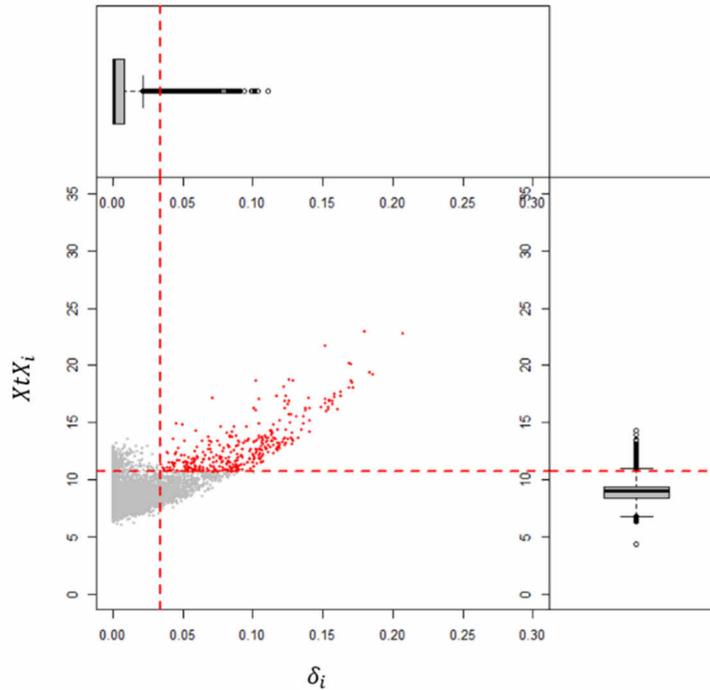
	<i>Ho</i>	<i>He</i>	<i>Fis</i>
<i>O. bacaba</i>			
SPA –	0.352 (sd=0.415)	0.205	-0.575 (sd=0.389) **
SPA +	0.373 (sd=0.393)	0.228	-0.438 (sd=0.408) **
LAU –	0.351 (sd=0.381)	0.226	-0.427 (sd=0.436) **
PAR –	0.356 (sd=0.385)	0.228	-0.453 (sd=0.442) **
NOU_N +	0.353 (sd=0.411)	0.207	-0.535 (sd=0.400) **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

448

449 2. SNP-specific differentiation

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



464

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

467

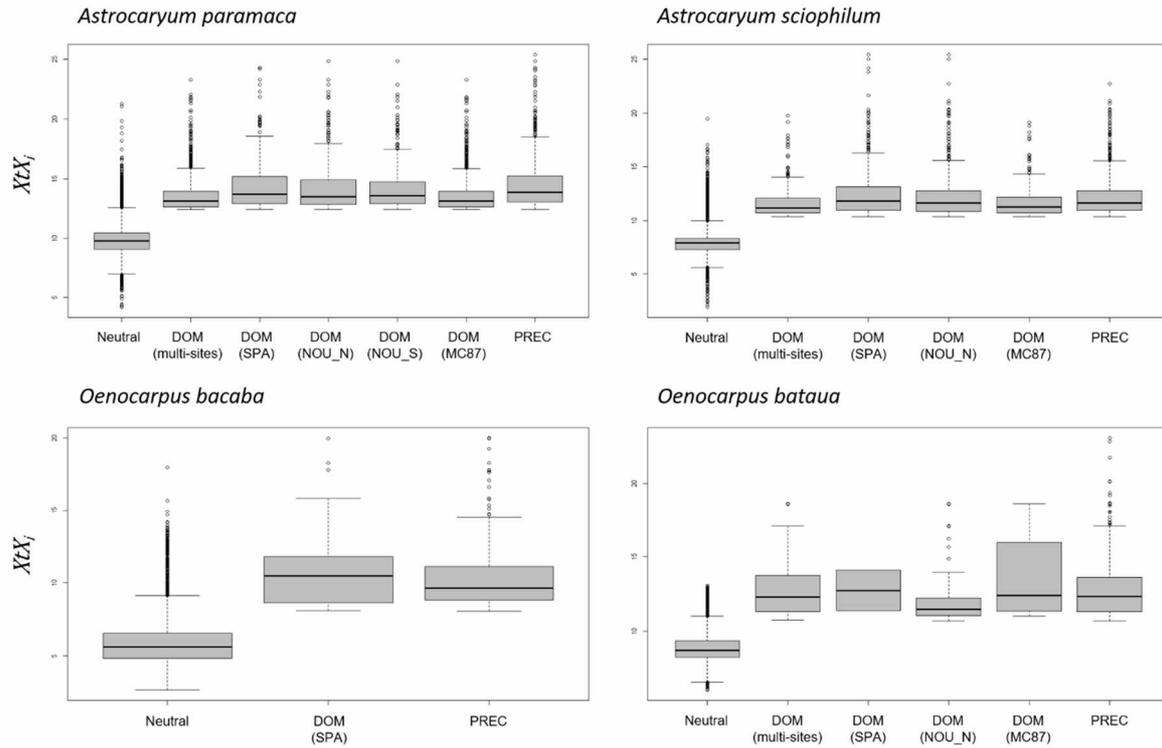
468 Between 0.74% (*A. sciophilum*) and 1.42% (*O. bataua*) climate adaptation SNPs were detected
469 in the different species, and between 0.20% (*O. bataua*) and 1.25% (*A. paramaca*)
470 domestication SNPs across all study sites. Moreover, the proportion of domestication SNPs was
471 higher when considering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

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

	<i>A. paramaca</i>	<i>A. sciophilum</i>	<i>O. bacaba</i>	<i>O. bataua</i>
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%)

483

484 The median of SNP-specific differentiation (XtX_i) for neutral SNPs was 9.8 (Q90 = [8.6 ; 12.0])
 485 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 ;
 486 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])
 487 in *A. paramaca* for climate adaptation SNP, and between 10.48 ([8.2 ; 13.8]) in *O. bacaba* and
 488 13.1 ([12.5 ; 16.0]) in *A. paramaca* without differences in the extent of genetic differentiation
 489 between climate adaptation and domestication SNPs, **Figure 6.**

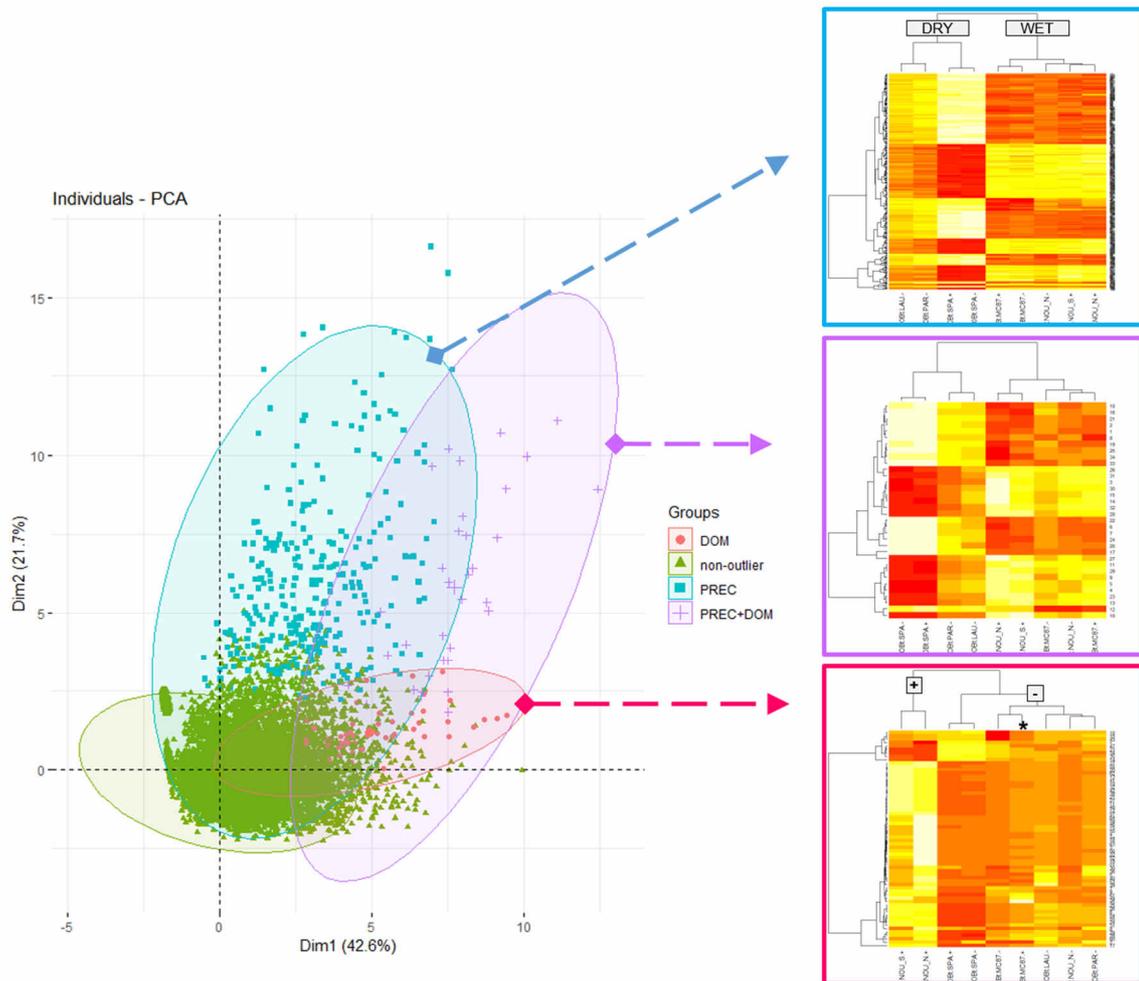


490

491 **Figure 6.** Distribution of SNP-specific differentiation (XtX_i) in neutral SNPs, domestication
492 (DOM) and climate adaptation SNPs (PREC).

493

494 Interestingly, several SNPs show variations in allele frequency in relation to both precipitation
495 and pre-Columbian occupation (**Supplementary material 6, Tables S6.3 to S6.6**), allowing
496 to distinguish four groups of SNPs within each species: (i) neutral SNPs, (ii) climate adaptation
497 SNPs showing (clinal) variations in allele frequency in relation to regional precipitation, (iii)
498 domestication SNPs showing differences in allele frequency in relation to local pre-Columbian
499 occupation, and (iv) climate and domestication SNPs showing allele frequency variations in
500 relation to both, **Figure 7**.



501

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

508 **Discussion**

509

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

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

526

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

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

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

557 Independently from site-specific hybridization patterns, estimates of intra-population genetic
558 diversity revealed high heterozygosity (H_o) and significantly negative F_{is} , indicating the
559 absence of inbreeding. These estimates are in line with previous knowledge about the breeding
560 system in Arecoideae that are known to be monoecious and predominantly outcrossed due to
561 protandry and/or genetic incompatibility (Henderson 1986; Eguiarte et al. 1992; Consiglio and
562 Bourne 2001; Picanço-Rodrigues et al. 2015; Gaiotto et al. 2003; Ottewell et al. 2012). The
563 genetic structure and diversity of populations on ring ditches was further close to populations
564 in the surrounding natural forests, indicating the absence of pre-Columbian impacts on
565 population genome-wide diversity and divergence. This result is not surprising, considering the
566 geographic proximity between “+” and “-” populations within study sites, which are probably
567 connected by extensive gene flow. The absence of genome-wide differentiation, however, does
568 not preclude the existence of SNP-specific selective processes driven by adaptation and/or
569 domestication.

570

571 **Climate adaptation and pre-Columbian domestication.**

572 Adaptation and domestication are acknowledged as important drivers of microevolution
573 worldwide, but they have been poorly documented in lowland Amazonia. While Central and
574 South America are known to be important centers of plant domestication, most studies have
575 focused on major crops such as maize (Kistler et al. 2018), papaya (Chávez-Pesqueira and
576 Núñez-Farfán 2017), manioc (Mühlen et al. 2019; Rival and McKey 2008; Alves-Pereira et al.
577 2020) or common bean (Rendón-Anaya et al. 2017), which provide the most obvious cases of
578 domestication. On the opposite, studies on minor crops are scarce although they are receiving
579 increasing attention since the last ten years (*e.g.* Brazil nut, Shepard and Ramirez 2011 ; piquiá,
580 Alves et al. 2016 ; peach palm, Clement et al. 2017 ; treegourd, Moreira et al. 2017). For its
581 part, the lack of interest in adaptation can be explained by the popularity of the neutral theory

582 under which the distribution of alleles within species (Kimura 1983) or species within
583 communities (Hubbell 2001) is assumed to be independent from fitness differences between
584 genotypes or species, being thus driven by stochastic mutation, demography and dispersal. In
585 hyperdiverse tropical rainforests, neutral models are powerful to predict global patterns of
586 diversity, as illustrated in the above section, and very useful to overcome the complexity of
587 niche differentiation modelling. Neutral and niche models are however highly complementary
588 (Chave 2004), and the neutral theory provides an elegant neutral hypothesis to test for niche
589 differentiation and adaptation (Leigh 2007). This is precisely the theoretical assumption of most
590 population genetics selection tests, which allows detecting gene- or SNP-specific departures
591 from the genome-wide, supposedly neutral, background of diversity and differentiation
592 (Beaumont and Nichols 1996).

593 By detecting many SNPs with footprints of selection, our results bring new evidence of the
594 coaction of climate adaptation and pre-Columbian domestication in shaping patterns of
595 genomic diversity across multiple spatial scales. As expected, selective processes related to
596 adaptation and domestication affected a very small fraction of the studied SNPs (always below
597 1.5%). Both the proportion of SNPs and the extent of differentiation (XtX_i) were in the same
598 order of magnitude between climate adaptation and domestication SNPs, in spite of the different
599 geographical scales involved (*i.e.* hundreds of kilometers between study sites with different
600 rainfall *versus* hundreds of meters between local populations in contrasting conditions of pre-
601 Columbian occupation). Regarding climate adaptation, the present results are in line with a
602 previous study in the hyperdominant tree *Eperua falcata* which provided another evidence of
603 climate adaptation along the sharp rainfall gradient in coastal French Guiana (Brousseau et al.
604 2021).

605 Climate and domestication SNPs were located within or in the neighborhood (< 200bp) of many
606 genes whose exhaustive list is provided in **Supplementary material 7**. Although documenting

607 the functions of these genes in *Astrocaryum* spp. and *Oenocarpus* spp. in beyond the scope of
608 this study, we can nevertheless pick a few examples to illustrate the variety of genes involved.
609 For example, we detected climate adaptation SNPs in several genes of major importance for
610 plant development and response to stresses, including ABC transporters (Yazaki 2006; Do et
611 al. 2018), leaf rust disease resistance loci (Lim et al. 2015), Apetala2-like ethylene responsive
612 transcription factor (Licausi et al. 2013), aquaporins (G. Li et al. 2014; Maurel et al. 2015), and
613 several peroxidases (Blokhina et al. 2003; Yoshida et al. 2003). Several domestication SNPs
614 were detected in genes involved in lipid and sugar metabolism that could be of importance for
615 fruit development, including phospholipase D (Sun et al. 2011; Mao et al. 2007; Pinhero et al.
616 2003), acyl-transferases, fructose-bisphosphate aldolase (Feng et al. 2020; Yin et al. 2012),
617 phospholipid-transporting ATPases, and UDP-glucose 6-dehydrogenase 4.

618 Interestingly, many SNPs showed footprints of both climate adaptation and domestication,
619 indicating that these two processes are closely intertwined. In other words, pre-Columbian
620 societies had either domesticated “*in situ*” locally pre-adapted populations or had to deal with
621 adaptation during domestication. The absence of genome-wide differentiation between palm
622 populations on ring ditches and surrounding natural populations tend to suggest that palm
623 populations were collected in the forest near to the ring ditches. Moreover, extensive gene flow
624 between wild and domesticated populations had probably continuously homogenized
625 populations, thus preventing the genome-wide differentiation and the fixation of the so-called
626 “domestication syndrome” in domesticated populations (see next section). Instead, SNP-
627 specific patterns of domestication observed here provide a new evidence of the subtle influence
628 of pre-Columbian societies and depicts what is commonly referred to as “incipient”
629 domestication (Lins Neto et al. 2014; Casas et al. 2016; Clement 1999). Gene flow since land
630 abandonment may also have partly erased patterns of differentiation caused by domestication
631 (*i.e.* de-domestication, Qiu et al. 2017). Moreover, many SNPs were detected in one study site

632 only, supporting that pre-Columbian impacts were partly site-specific. Nevertheless, many
633 SNPs were also detected in more than one study site, indicating some common trends in the
634 different study sites, probably caused by similar management practices. Whether observed
635 domestication patterns result from “conscious” domestication through direct selection for
636 desired traits or “unconscious” domestication through indirect local adaptation to
637 anthropogenic landscapes (Gepts 2010; Engelbrecht 1916; Darlington 1973) cannot however
638 be solved yet, and would require in-depth investigation at the phenotypic level to assess which
639 traits have been selected for or against in palm populations established on ring ditches.

640

641 **Moving beyond the domestication syndrome paradigm.**

642 The domestication of plants is a long-term process, with multiple possible outcomes. However,
643 most studies have focused on major crops exhibiting clear phenotypic differences with wild
644 populations, often underpinned by a sharp – genome-wide – genetic differentiation. Studies on
645 major crops led to the elaboration of an earlier model of domestication through which the
646 domestication is initiated by the isolation of a small population with reduced diversity.
647 Domesticated populations are then expected to accumulate phenotypic differences from their
648 wild ancestors under the action of strong selection (*i.e.* artificial selection for desired traits and
649 natural selection exerted by new environmental conditions) reinforced by inbreeding, both
650 resulting in the fast decrease in genetic diversity and in the fixation of phenotypic differences
651 in domesticated populations: the so-called “domestication syndrome” (Hammer 1984; Harlan
652 et al. 1973; Allaby 2014). The domestication syndrome has been extensively documented, and
653 quite over-emphasized (Kantar et al. 2017), leading to a narrow view of the domestication
654 process and to a misleading dichotomous distinction between domesticated varieties with a
655 clear domestication syndrome, and other (supposedly wild) populations. The domestication
656 process is, however, gradual through time and variable across populations (Meyer et al. 2012;
657 O. Smith et al. 2019; Kantar et al. 2017), and it is not necessarily expected to result in a sharp

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

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

672

673 **Supplementary materials**

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

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

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

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

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

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

683 **Supplementary material 7 (S7).** Genes containing climate adaptation and domestication
684 SNPs.

685

686 **Data and codes availability**

687 Sequencing data (i.e. paired-end reads) are available on the European Nucleotide Archive under
688 accession PRJEB51800. Bioinformatics codes are available on GitLab under Project ID
689 34990830 (<https://gitlab.com/LouiseBrousseau/palm-capture-bioinformatics>). Data and code
690 remains under embargo until formal acceptance of the present article.

691

692 **Competing interests**

693 The authors declared no conflict of interest.

694

695 **Authors contribution**

696 LB and GO imagined the study and carried out palm sampling. LB designed the capture kit,
697 implemented the bioinformatics pipeline, and carried out bioinformatics and population
698 genetics. SS and AW carried out the molecular experiments (gDNA extraction and purification,

699 library preparation, targeted capture and sequencing of *O. bataua* libraries). All authors wrote
700 and approved the article.

701

702 **Acknowledgements**

703 This study has been conducted as part of the research program *PalmOmix* coordinated by L.
704 Brousseau. *PalmOmix* (ID 1702-01) was publicly funded by the French National Research
705 Agency (Agence Nationale de la Recherche, ANR) under the “*Investissement d’avenir*”
706 program with reference ANR-10-LABX-001-01 LabEx Agro and coordinated by Agropolis
707 Foundation under the frame of I-SITE MUSE (ANR-16-IDEX-0006). Sampling in the
708 Nouragues research Station has been supported by the LongTime project coordinated by G.
709 Odonne and J-F. Molino and funded by an “*Investissement d’avenir*” grant from the French
710 National Research Agency (Agence Nationale de la Recherche, ANR) with reference CEBA:
711 ANR-10-LABX-25-01.

712 We are grateful to the Nouragues Ecological Research Station (CNRS) and the Réserve
713 Naturelle des Nouragues, to the Paracou Research Station (CIRAD) and to the National Forests
714 Office (ONF) in French Guiana for access to the study sites. We are also grateful to GenoToul
715 (“GeT” and bioinformatics) facilities for next-generation sequencing of *Astrocaryum* spp. and
716 *O. bacaba* libraries, and for access to computing facilities to carry out bioinformatics analyses.
717 We thank Thierry Joët, Stéphane Dussert and James Tregear for their help in pre-filtering
718 candidate proteins (“bottom-up approach”) to design the custom capture kit. We also thank T.
719 Joët for his help in palm sampling and Anaïs Prud’homme who explored *O. bataua* NGS data
720 as part of her internship at UMR DIADE under the supervision of L. Brousseau.

721 **References**

- 722 Allaby, Robin G. 2014. “Domestication Syndrome in Plants.” In *Encyclopedia of Global*
723 *Archaeology*, 2182–84. Springer New York. [https://doi.org/10.1007/978-1-4419-0465-](https://doi.org/10.1007/978-1-4419-0465-2_2416)
724 [2_2416](https://doi.org/10.1007/978-1-4419-0465-2_2416).
- 725 Alves-Pereira, Alessandro, Charles R. Clement, Doriane Picanço-Rodrigues, Elizabeth Ann
726 Veasey, Gabriel Dequigiovanni, Santiago Linorio Ferreyra Ramos, José Baldin Pinheiro,
727 Anete Pereira de Souza, and Maria Imaculada Zucchi. 2020. “A Population Genomics
728 Appraisal Suggests Independent Dispersals for Bitter and Sweet Manioc in Brazilian
729 Amazonia.” *Evolutionary Applications* 13 (2): 342–61.
730 <https://doi.org/10.1111/eva.12873>.
- 731 Alves, Rubana P., Carolina Levis, and Charles R. Clement. 2016. “Use and Management of
732 Piquiá Suggest *in Situ* Domestication along the Lower Tapajós River, Brazilian
733 Amazonia.” *Economic Botany* 70 (2): 198–202. [https://doi.org/10.1007/s12231-016-](https://doi.org/10.1007/s12231-016-9340-4)
734 [9340-4](https://doi.org/10.1007/s12231-016-9340-4).
- 735 Andrews, Simon. 2010. “FastQC: A Quality Control Tool for High Throughput Sequence
736 Data.”
- 737 Aronesty, Erik. 2013. “Comparison of Sequencing Utility Programs.” *The Open*
738 *Bioinformatics Journal* 7 (1): 1–8. <https://doi.org/10.2174/1875036201307010001>.
- 739 Baker, William J., and Thomas L.P. Couvreur. 2013. “Global Biogeography and
740 Diversification of Palms Sheds Light on the Evolution of Tropical Lineages. I. Historical
741 Biogeography.” Edited by W. Daniel Kissling. *Journal of Biogeography* 40 (2): 274–85.
742 <https://doi.org/10.1111/j.1365-2699.2012.02795.x>.
- 743 Beaumont, Mark A, and Richard A Nichols. 1996. “Evaluating Loci for Use in the Genetic
744 Analysis of Population Structure.” *Proceedings of The Royal Society B: Biological*

- 745 *Sciences* 263: 1619–26.
- 746 Blokhina, Olga, Eija Virolainen, and Kurt V. Fagerstedt. 2003. “Antioxidants, Oxidative
747 Damage and Oxygen Deprivation Stress: A Review.” *Annals of Botany* 91 (2): 179–94.
748 <https://doi.org/10.1093/aob/mcf118>.
- 749 Bodin, Stephanie C., Jean François Molino, Guillaume Odonne, and Laurent Bremond. 2020.
750 “Unraveling Pre-Columbian Occupation Patterns in the Tropical Forests of French
751 Guiana Using an Anthracological Approach.” *Vegetation History and Archaeobotany* 29
752 (5): 567–80. <https://doi.org/10.1007/s00334-019-00767-w>.
- 753 Brancier, Jeanne, Cecilia Cammas, Dominique Todisco, and Eric Fouache. 2014. “A
754 Micromorphological Assessment of Anthropogenic Features in Pre-Columbian French
755 Guiana Dark Soils (FGDS): First Results.” *Zeitschrift Fur Geomorphologie* 58
756 (October): 109–39. <https://doi.org/10.1127/0372-8854/2014/S-00159>.
- 757 Brousseau, Louise, Paul V.A. Fine, Erwin Dreyer, Giovanni G. Vendramin, and Ivan Scotti.
758 2021. “Genomic and Phenotypic Divergence Unveil Microgeographic Adaptation in the
759 Amazonian Hyperdominant Tree *Eperua Falcata* Aubl. (Fabaceae).” *Molecular Ecology*
760 30 (5): 1136–54.
- 761 Carson, John F, Jennifer Watling, Francis E Mayle, Bronwen S Whitney, José Iriarte, Heiko
762 Prümers, and J Daniel Soto. 2015. “Pre-Columbian Land Use in the Ring-Ditch Region
763 of the Bolivian Amazon.” *The Holocene* 25 (8): 1285–1300.
764 <https://doi.org/10.1177/0959683615581204>.
- 765 Carson, John Francis, Bronwen S. Whitney, Francis E. Mayle, José Iriarte, Heiko Prümers, J.
766 Daniel Soto, and Jennifer Watling. 2014. “Environmental Impact of Geometric
767 Earthwork Construction in Pre-Columbian Amazonia.” *Proceedings of the National*
768 *Academy of Sciences of the United States of America* 111 (29): 10497–502.

- 769 <https://doi.org/10.1073/pnas.1321770111>.
- 770 Casas, Alejandro, José Blancas, Adriana Otero-Arnaiz, Jeniffer Cruse-Sanders, Rafael Lira,
771 Aidé Avendaño, Fabiola Parra, et al. 2016. “Evolutionary Ethnobotanical Studies of
772 Incipient Domestication of Plants in Mesoamerica.” In , 257–85. Springer, New York,
773 NY. https://doi.org/10.1007/978-1-4614-6669-7_11.
- 774 Castro, J. W., M. S. Hernández, and R. H. Gutiérrez. 2014. “Uses for Amazonian Seje Oil
775 (*Oenocarpus Bataua*) Extracted by Mechanical and Manual Methods.” In *Acta*
776 *Horticulturae*, 1047:335–40. International Society for Horticultural Science.
777 <https://doi.org/10.17660/actahortic.2014.1047.43>.
- 778 Cavers, Stephen, and Christopher W. Dick. 2013. “Phylogeography of Neotropical Trees.”
779 *Journal of Biogeography* 40 (4): 615–17. <https://doi.org/10.1111/jbi.12097>.
- 780 Chave, Jerome. 2004. “Neutral Theory and Community Ecology.” *Ecology Letters* 7 (august):
781 241–53. <https://doi.org/10.1111/j.1461-0248.2003.00566.x>.
- 782 Chávez-Pesqueira, Mariana, and Juan Núñez-Farfán. 2017. “Domestication and Genetics of
783 Papaya: A Review.” *Frontiers in Ecology and Evolution* 5: 155.
784 <https://www.frontiersin.org/article/10.3389/fevo.2017.00155>.
- 785 Clement, Charles R. 1999. “1942 and the Loss of Amazonian Crop Genetic Resources. I. The
786 Relation between Domestication and Human Population Decline.” *Economic Botany*.
787 SpringerNew York Botanical Garden Press. <https://doi.org/10.2307/4256179>.
- 788 Clement, Charles R., Joseph M. McCann, and Nigel J. H. Smith. 2006. “Agrobiodiversity in
789 Amazônia and Its Relationship with Dark Earths.” In *Amazonian Dark Earths*, 159–78.
790 Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-2597-1_9.
- 791 Clement, Charles R, Michelly de Cristo-Araújo, Geo Coppens d’Eeckenbrugge, Vanessa
792 Maciel dos Reis, Romain Lehnebach, and Doriane Picanço-Rodrigues. 2017. “Origin and

- 793 Dispersal of Domesticated Peach Palm.” *Frontiers in Ecology and Evolution* 5: 148.
794 <https://www.frontiersin.org/article/10.3389/fevo.2017.00148>.
- 795 Clement, Charles R, Michelly de Cristo-Araújo, Geo Coppens D’Eeckenbrugge, Alessandro
796 Alves Pereira, and Doriane Picanço-Rodrigues. 2010. “Origin and Domestication of
797 Native Amazonian Crops.” *Diversity* 2 (1): 72–106. <https://doi.org/10.3390/d2010072>.
- 798 Consiglio, Trisha K., and Godfrey R. Bourne. 2001. “Pollination and Breeding System of a
799 Neotropical Palm *Astrocaryum Vulgare* in Guyana: A Test of the Predictability of
800 Syndromes.” *Journal of Tropical Ecology* 17 (4): 577–92.
801 <https://doi.org/10.1017/S0266467401001432>.
- 802 Coomes, Oliver T. 2004. “Rain Forest ‘Conservation-through-Use’? Chambira Palm Fibre
803 Extraction and Handicraft Production in a Land-Constrained Community, Peruvian
804 Amazon.” *Biodiversity and Conservation* 13 (2): 351–60.
805 <https://link.springer.com/article/10.1023/B:BIOC.0000006503.90980.e8>.
- 806 Cummings, Anthony R., and Jane M. Read. 2016. “Drawing on Traditional Knowledge to
807 Identify and Describe Ecosystem Services Associated with Northern Amazon’s Multiple-
808 Use Plants.” *International Journal of Biodiversity Science, Ecosystem Services &*
809 *Management* 12 (1–2): 39–56. <https://doi.org/10.1080/21513732.2015.1136841>.
- 810 Cunha, Vânia Maria Borges, Marcilene Paiva da Silva, Sérgio Henrique Brabo de Sousa,
811 Priscila do Nascimento Bezerra, Eduardo Gama Ortiz Menezes, Nayara Janaina Neves
812 da Silva, Débora Ariane Dornelas da Silva Banna, Marilena Emmi Araújo, and Raul
813 Nunes de Carvalho Junior. 2019. “Bacaba-de-Leque (*Oenocarpus Distichus* Mart.) Oil
814 Extraction Using Supercritical CO₂ and Bioactive Compounds Determination in the
815 Residual Pulp.” *Journal of Supercritical Fluids* 144 (February): 81–90.
816 <https://doi.org/10.1016/j.supflu.2018.10.010>.

- 817 Danecek, Petr, James K. Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O.
818 Pollard, Andrew Whitwham, et al. 2021. “Twelve Years of SAMtools and BCFtools.”
819 *GigaScience* 10 (2). <https://doi.org/10.1093/gigascience/giab008>.
- 820 Darlington, Cyril Dean. 1973. *Chromosome Botany and the Origins of Cultivated Plants*.
821 London: Allen & Unwin.
- 822 Davy, Damien. 2007. “*Vannerie et Vanniers*” *Approche Ethnologique d’une Activité*
823 *Artisanale En Guyane Française*. Université d’Orleans.
- 824 Do, Thanh Ha Thi, Enrico Martinoia, and Youngsook Lee. 2018. “Functions of ABC
825 Transporters in Plant Growth and Development.” *Current Opinion in Plant Biology*.
826 Elsevier Ltd. <https://doi.org/10.1016/j.pbi.2017.08.003>.
- 827 Eguiarte, Luis E., Nidia Perez-Nasser, and Daniel Piñero. 1992. “Genetic Structure,
828 Outcrossing Rate and Heterosis in *Astrocaryum Mexicanum* (Tropical Palm):
829 Implications for Evolution and Conservation.” *Heredity* 69 (3): 217–28.
830 <https://doi.org/10.1038/hdy.1992.119>.
- 831 Engelbrecht, Th. H. 1916. “Über Die Entstehung Einiger Feldmäßig Angebauter
832 Kulturpflanzen.” *Geographische Zeitschrift* 22 (6): 328–34.
833 <http://www.jstor.org/stable/27809914>.
- 834 Fadiman, Maria G. 2008. “Use of *Mocora*, *Astrocaryum Standleyanum* (Arecaceae), by Three
835 Ethnic Groups in Ecuador: Differences, Similarities and Market Potential.” *Journal of*
836 *Ethnobiology* 28 (1): 92–109. [https://doi.org/10.2993/0278-](https://doi.org/10.2993/0278-0771(2008)28[92:UOMASA]2.0.CO;2)
837 [0771\(2008\)28\[92:UOMASA\]2.0.CO;2](https://doi.org/10.2993/0278-0771(2008)28[92:UOMASA]2.0.CO;2).
- 838 Feng, Dan, Li Gao, Yusheng Zheng, Dongdong Li, and Peng Zhou. 2020. “Molecular Cloning
839 and Function Characterisation of a Cytoplasmic Fructose-1, 6-Bisphosphate Aldolase
840 Gene from Coconut (*Cocos Nucifera* L.)” *Journal of Horticultural Science and*

- 841 *Biotechnology* 95 (6): 703–11. <https://doi.org/10.1080/14620316.2020.1749139>.
- 842 Förstner, Wolfgang, and Boudewijn Moonen. 2003. “A Metric for Covariance Matrices.” In
843 *Geodesy-The Challenge of the 3rd Millennium*, 299–309. Berlin, Heidelberg: Springer
844 Berlin Heidelberg. https://doi.org/10.1007/978-3-662-05296-9_31.
- 845 Frichot, Eric, François Mathieu, Théo Trouillon, Guillaume Bouchard, and Olivier François.
846 2014. “Fast and Efficient Estimation of Individual Ancestry Coefficients.” *Genetics* 196:
847 973–983. <https://doi.org/10.1534/genetics.113.160572>.
- 848 Gaiotto, Fernanda A, Dario Grattapaglia, and Roland Vencovsky. 2003. “Genetic Structure,
849 Mating System, and Long-Distance Gene Flow in Heart of Palm (*Euterpe Edulis* Mart.)”
850 *Journal of Heredity* 94 (5): 399–406. <https://doi.org/10.1093/jhered/esg087>.
- 851 García, Néstor, Gloria Galeano, Laura Mesa, Nicolás Castaño, Henrik Balslev, Rodrigo
852 Bernal, Néstor García, et al. 2015. “Management of the Palm *Astrocaryum Chambira*
853 Burret (Arecaceae) in Northwest Amazon.” *Acta Botanica Brasilica* 29 (1): 45–57.
854 <https://doi.org/10.1590/0102-33062014abb3415>.
- 855 Gepts, Paul. 2010. “Crop Domestication as a Long-Term Selection Experiment.” In *Plant*
856 *Breeding Reviews*, 1–44. Oxford: John Wiley & Sons, Inc.
857 <https://doi.org/10.1002/9780470650288.ch1>.
- 858 Günther, Torsten, and Graham Coop. 2013. “Robust Identification of Local Adaptation from
859 Allele Frequencies.” *Genetics* 195 (1): 205–20.
860 <https://doi.org/10.1534/genetics.113.152462>.
- 861 Gutierrez, Maria Paz. 2020. “The Northwestern Amazon Malocas: Craft Now and Then.”
862 *Journal of Material Culture* 25 (1): 3–35. <https://doi.org/10.1177/1359183519836141>.
- 863 Hammer, Karl. 1984. “Das Domestikationssyndrom (The Domestication Syndrome).” *Die*
864 *Kulturpflanze* 32 (1): 11–34. <https://doi.org/10.1007/BF02098682>.

- 865 Harlan, Jack R, J M J de Wet, and E Glen Price. 1973. "Comparative Evolution of Cereals."
866 *Evolution* 27 (2): 311–25. <https://doi.org/10.2307/2406971>.
- 867 Henderson, Andrew. 1986. "A Review of Pollination Studies in the Palmae." *The Botanical*
868 *Review* 52 (3): 221–59. <https://doi.org/10.1007/BF02860996>.
- 869 Hoorn, Carina, Frank P. Wesselingh, Hans Ter Steege, M. A. Bermudez, A. Mora, J. Sevink,
870 I. Sanmartín, et al. 2010. "Amazonia through Time: Andean Uplift, Climate Change,
871 Landscape Evolution, and Biodiversity." *Science* 330 (6006): 927–31.
872 <https://doi.org/10.1126/science.1194585>.
- 873 Hubbell, Stephen P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*.
874 Princeton University Press.
- 875 Jombart, Thibaut. 2008. "Adegenet: A R Package for the Multivariate Analysis of Genetic
876 Markers." *Bioinformatics* 24 (11): 1403–5.
877 <https://doi.org/10.1093/bioinformatics/btn129>.
- 878 Jombart, Thibaut, and Ismaïl Ahmed. 2011. "Adegenet 1.3-1: New Tools for the Analysis of
879 Genome-Wide SNP Data." *Bioinformatics* 27 (21): 3070–71.
880 <https://doi.org/10.1093/bioinformatics/btr521>.
- 881 Joshi, N A, and Joseph Fass. 2011. "Sickle: A Sliding-Window, Adaptive, Quality-Based
882 Trimming Tool for FastQ Files (Version 1.33)."
- 883 Kahn, Francis. 1993. "Amazonian Palms: Food Resources for the Management of Forest
884 Ecosystems." *Tropical Forests, People and Food: Biocultural Interactions and*
885 *Applications to Development* 13: 153.
- 886 Kantar, Michael B, Amber R Nashoba, Justin E Anderson, Benjamin K Blackman, and Loren
887 H Rieseberg. 2017. "The Genetics and Genomics of Plant Domestication." *BioScience*
888 67 (11): 971–82. <https://doi.org/10.1093/biosci/bix114>.

- 889 Kimura, Motoo. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge
890 University Press. [https://doi.org/DOI: 10.1017/CBO9780511623486](https://doi.org/DOI:10.1017/CBO9780511623486).
- 891 Kistler, Logan, S. Yoshi Maezumi, Jonas Gregorio De Souza, Natalia A. S. Przelomska,
892 Flaviane Malaquias Costa, Oliver Smith, Hope Loiselle, et al. 2018. “Multiproxy
893 Evidence Highlights a Complex Evolutionary Legacy of Maize in South America.”
894 *Science* 362 (6420): 1309–13. <https://doi.org/10.1126/science.aav0207>.
- 895 Leigh, E. G. 2007. “Neutral Theory: A Historical Perspective.” *Journal of Evolutionary*
896 *Biology* 20 (6): 2075–91. <https://doi.org/10.1111/j.1420-9101.2007.01410.x>.
- 897 Levis, Carolina, Flávia R C Costa, Frans Bongers, Marielos Peña-Claros, Charles R Clement,
898 André B Junqueira, Eduardo G Neves, Eduardo K Tamanaha, Fernando O G Figueiredo,
899 and Rafael P Salomão. 2017. “Persistent Effects of Pre-Columbian Plant Domestication
900 on Amazonian Forest Composition.” *Science* 355 (6328): 925–31.
901 <https://doi.org/10.1126/science.aal0157>.
- 902 Li, Guowei, Véronique Santoni, and Christophe Maurel. 2014. “Plant Aquaporins: Roles in
903 Plant Physiology.” *Biochimica et Biophysica Acta - General Subjects*. Elsevier.
904 <https://doi.org/10.1016/j.bbagen.2013.11.004>.
- 905 Li, Heng. 2011. “A Statistical Framework for SNP Calling, Mutation Discovery, Association
906 Mapping and Population Genetical Parameter Estimation from Sequencing Data.”
907 *Bioinformatics* 27 (21): 2987–93. <https://doi.org/10.1093/bioinformatics/btr509>.
- 908 Li, Heng, and Richard Durbin. 2009. “Fast and Accurate Short Read Alignment with
909 Burrows–Wheeler Transform.” *Bioinformatics* 25 (14): 1754–60.
910 <https://doi.org/10.1093/bioinformatics/btp324>.
- 911 ———. 2010. “Fast and Accurate Long-Read Alignment with Burrows–Wheeler Transform.”
912 *Bioinformatics* 26 (5): 589–95. <https://doi.org/10.1093/bioinformatics/btp698>.

- 913 Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth,
914 Goncalo Abecasis, Richard Durbin, and Genome Project Data Processing Subgroup.
915 2009. “The Sequence Alignment/Map Format and SAMtools.” *Bioinformatics* 25 (16):
916 2078–79. <https://doi.org/10.1093/bioinformatics/btp352>.
- 917 Licausi, Francesco, Masaru Ohme-Takagi, and Pierdomenico Perata. 2013.
918 “APETALA2/Ethylene Responsive Factor (AP2/ERF) Transcription Factors: Mediators
919 of Stress Responses and Developmental Programs.” *New Phytologist* 199 (3): 639–49.
920 <https://doi.org/10.1111/nph.12291>.
- 921 Lim, Chae Woo, Seung Hwan Yang, Ki Hun Shin, Sung Chul Lee, and Sang Hyon Kim.
922 2015. “The AtLRK10L1.2, Arabidopsis Ortholog of Wheat LRK10, Is Involved in ABA-
923 Mediated Signaling and Drought Resistance.” *Plant Cell Reports* 34 (3): 447–55.
924 <https://doi.org/10.1007/s00299-014-1724-2>.
- 925 Lins Neto, Ernani M., Nivaldo Peroni, Alejandro Casas, Fabiola Parra, Xitlali Aguirre, Susana
926 Guillén, and Ulysses P. Albuquerque. 2014. “Brazilian and Mexican Experiences in the
927 Study of Incipient Domestication.” *Journal of Ethnobiology and Ethnomedicine*. BioMed
928 Central Ltd. <https://doi.org/10.1186/1746-4269-10-33>.
- 929 Llamas, Bastien, Lars Fehren-Schmitz, Guido Valverde, Julien Soubrier, Swapan Mallick,
930 Nadin Rohland, Susanne Nordenfelt, et al. 2016. “Ancient Mitochondrial DNA Provides
931 High-Resolution Time Scale of the Peopling of the Americas.” *Science Advances* 2 (4):
932 e1501385. <https://doi.org/10.1126/sciadv.1501385>.
- 933 Mantel, N. 1967. “The Detection of Disease Clustering and a Generalized Regression
934 Approach.” *Cancer Research*.
- 935 Mao, Linchun, Huaqing Pang, Guoze Wang, and Chenggang Zhu. 2007. “Phospholipase D
936 and Lipoyxygenase Activity of Cucumber Fruit in Response to Chilling Stress.”

- 937 *Postharvest Biology and Technology* 44 (1): 42–47.
938 <https://doi.org/10.1016/j.postharvbio.2006.11.009>.
- 939 Mascher, Martin, Todd A. Richmond, Daniel J. Gerhardt, Axel Himmelbach, Leah Clissold,
940 Dharanya Sampath, Sarah Ayling, et al. 2013. “Barley Whole Exome Capture: A Tool
941 for Genomic Research in the Genus *Hordeum* and Beyond.” *Plant Journal* 76 (3): 494–
942 505. <https://doi.org/10.1111/tpj.12294>.
- 943 Maurel, Christophe, Yann Boursiac, Doan Trung Luu, Véronique Santoni, Zaigham Shahzad,
944 and Lionel Verdoucq. 2015. “Aquaporins in Plants.” *Physiological Reviews* 95 (4):
945 1321–58. <https://doi.org/10.1152/physrev.00008.2015>.
- 946 Meyer, Rachel S., Ashley E. DuVal, and Helen R. Jensen. 2012. “Patterns and Processes in
947 Crop Domestication: An Historical Review and Quantitative Analysis of 203 Global
948 Food Crops.” *New Phytologist* 196 (1): 29–48. [https://doi.org/10.1111/j.1469-](https://doi.org/10.1111/j.1469-8137.2012.04253.x)
949 [8137.2012.04253.x](https://doi.org/10.1111/j.1469-8137.2012.04253.x).
- 950 Meyer, Rachel S., and Michael D. Purugganan. 2013. “Evolution of Crop Species: Genetics of
951 Domestication and Diversification.” *Nature Reviews Genetics* 14 (12): 840–52.
952 <https://doi.org/10.1038/nrg3605>.
- 953 Moreira, Priscila Ambrósio, Cédric Mariac, Leila Zekraoui, Marie Couderc, Doriane Picanço
954 Rodrigues, Charles R. Clement, and Yves Vigouroux. 2017. “Human Management and
955 Hybridization Shape Treegourd Fruits in the Brazilian Amazon Basin.” *Evolutionary*
956 *Applications* 10 (6): 577–89. <https://doi.org/10.1111/eva.12474>.
- 957 Mühlen, Gilda Santos, Alessandro Alves-Pereira, Cássia Regina Limonta Carvalho, André
958 Braga Junqueira, Charles R. Clement, and Teresa Losada Valle. 2019. “Genetic
959 Diversity and Population Structure Show Different Patterns of Diffusion for Bitter and
960 Sweet Manioc in Brazil.” *Genetic Resources and Crop Evolution* 66 (8): 1773–90.

- 961 <https://doi.org/10.1007/s10722-019-00842-1>.
- 962 Odonne, Guillaume, Martijn Bel, Maxime Burst, Olivier Brunaux, Miléna Bruno, Etienne
963 Dambrine, Damien Davy, et al. 2019. “Long-term Influence of Early Human
964 Occupations on Current Forests of the Guiana Shield.” *Ecology* 100 (10): e02806.
965 <https://doi.org/10.1002/ecy.2806>.
- 966 Onstein, Renske E, William J Baker, Thomas L P Couvreur, Søren Faurby, Leonel Herrera-
967 Alsina, Jens-Christian Svenning, and W Daniel Kissling. 2018. “To Adapt or Go
968 Extinct? The Fate of Megafaunal Palm Fruits under Past Global Change.” *Proceedings*.
969 *Biological Sciences* 285 (1880): 20180882. <https://doi.org/10.1098/rspb.2018.0882>.
- 970 Ottewell, K, E Grey, F Castillo, and J Karubian. 2012. “The Pollen Dispersal Kernel and
971 Mating System of an Insect-Pollinated Tropical Palm, *Oenocarpus Bataua*.” *Heredity*
972 109 (6): 332–39. <https://doi.org/10.1038/hdy.2012.40>.
- 973 Patterson, Nick, Priya Moorjani, Yontao Luo, Swapan Mallick, Nadin Rohland, Yiping Zhan,
974 Teri Genschoreck, Teresa Webster, and David Reich. 2012. “Ancient Admixture in
975 Human History.” *Genetics* 192 (3): 1065–93.
976 <https://doi.org/10.1534/genetics.112.145037>.
- 977 Peter, Benjamin M. 2016. “Admixture, Population Structure and F-Statistics.” *Genetics* 4:
978 1485–1501. <https://doi.org/10.1534/genetics.115.183913>.
- 979 Petitjean Roget, Hugues. 1991. “50 Sites de Montagnes En Guyane Française: Contribution A
980 L’inventaire Des Sites Archeologiques D’Emile Abonnenc.” In *Proceedings of the*
981 *Twelfth Congress of the International Association for Caribbean Archaeology*, 241–58.
- 982 Picanço-Rodrigues, Doriane, Spartaco Astolfi-Filho, Maristerra R Lemes, Rogerio Gribel,
983 Alexandre M Sebbenn, and Charles R Clement. 2015. “Conservation Implications of the
984 Mating System of the Pampa Hermosa Landrace of Peach Palm Analyzed with

- 985 Microsatellite Markers.” *Genetics and Molecular Biology*. scielo.
- 986 Pinhero, Reena Grittle, Kurt C. Almquist, Zuzana Novotna, and Gopinadhan Paliyath. 2003.
- 987 “Developmental Regulation of Phospholipase D in Tomato Fruits.” *Plant Physiology and*
- 988 *Biochemistry* 41 (3): 223–40. [https://doi.org/10.1016/S0981-9428\(03\)00014-7](https://doi.org/10.1016/S0981-9428(03)00014-7).
- 989 Puerari, Cláudia, Karina Teixeira Magalhães-Guedes, and Rosane Freitas Schwan. 2015.
- 990 “Bacaba Beverage Produced by Umutina Brazilian Amerindians: Microbiological and
- 991 Chemical Characterization.” *Brazilian Journal of Microbiology* 46 (4): 1207–16.
- 992 <https://doi.org/10.1590/S1517-838246420140964>.
- 993 Qiu, Jie, Yongjun Zhou, Lingfeng Mao, Chuyu Ye, Weidi Wang, Jianping Zhang, Yongyi Yu,
- 994 et al. 2017. “Genomic Variation Associated with Local Adaptation of Weedy Rice
- 995 during De-Domestication.” *Nature Communications* 8 (May): 15323.
- 996 <https://doi.org/10.1038/ncomms15323>.
- 997 R Core Team, R Core Team. 2021. “R: A Language and Environment for Statistical
- 998 Computing.” In *R Foundation for Statistical Computing, Vienna, Austria*. [https://www.r-](https://www.r-project.org/)
- 999 [project.org/](https://www.r-project.org/).
- 1000 Reich, David, Kumarasamy Thangaraj, Nick Patterson, Alkes L. Price, and Lalji Singh. 2009.
- 1001 “Reconstructing Indian Population History.” *Nature* 461 (7263): 489–94.
- 1002 <https://doi.org/10.1038/nature08365>.
- 1003 Rendón-Anaya, Martha, Josaphat M. Montero-Vargas, Soledad Saburido-Álvarez, Anna
- 1004 Vlasova, Salvador Capella-Gutierrez, José Juan Ordaz-Ortiz, O. Mario Aguilar, et al.
- 1005 2017. “Genomic History of the Origin and Domestication of Common Bean Unveils Its
- 1006 Closest Sister Species.” *Genome Biology* 18 (1): 60. <https://doi.org/10.1186/s13059-017->
- 1007 1190-6.
- 1008 Rival, Laura, and Doyle McKey. 2008. “Domestication and Diversity in Manioc (*Manihot*

- 1009 *Esculenta* Crantz Ssp. *Esculenta*, Euphorbiaceae.” *Current Anthropology* 49 (6): 1119–
1010 28. <https://doi.org/10.1086/593119>.
- 1011 Rohland, Nadin, and David Reich. 2012. “Cost-Effective, High-Throughput DNA Sequencing
1012 Libraries for Multiplexed Target Capture.” *Genome Research* 22 (5): 939–46.
1013 <https://doi.org/10.1101/gr.128124.111>.
- 1014 Rothhammer, Francisco, and Tom D. Dillehay. 2009. “The Late Pleistocene Colonization of
1015 South America: An Interdisciplinary Perspective.” *Annals of Human Genetics*. Ann Hum
1016 Genet. <https://doi.org/10.1111/j.1469-1809.2009.00537.x>.
- 1017 Shepard, Glenn H., and Henri Ramirez. 2011. “‘Made in Brazil’: Human Dispersal of the
1018 Brazil Nut (*Bertholletia Excelsa*, Lecythidaceae) in Ancient Amazonia.” *Economic*
1019 *Botany* 65 (1): 44–65. <https://doi.org/10.1007/s12231-011-9151-6>.
- 1020 Smith, Brian Tilston, John E. McCormack, Andrés M. Cuervo, Michael J. Hickerson,
1021 Alexandre Aleixo, Carlos Daniel Cadena, Jorge Pérez-Emán, et al. 2014. “The Drivers of
1022 Tropical Speciation.” *Nature* 515 (7527): 406–9. <https://doi.org/10.1038/nature13687>.
- 1023 Smith, Nigel. 2015. *Palms and People in the Amazon*. Geobotany Studies. Cham: Springer
1024 International Publishing. <https://doi.org/10.1007/978-3-319-05509-1>.
- 1025 Smith, Oliver, William V. Nicholson, Logan Kistler, Emma Mace, Alan Clapham, Pamela
1026 Rose, Chris Stevens, et al. 2019. “A Domestication History of Dynamic Adaptation and
1027 Genomic Deterioration in *Sorghum*.” *Nature Plants* 5 (4): 369–79.
1028 <https://doi.org/10.1038/s41477-019-0397-9>.
- 1029 Sosnowska, Joanna, Adam Walanus, and Henrik Balslev. 2015. “Asháninka Palm
1030 Management and Domestication in the Peruvian Amazon.” *Human Ecology* 43 (3): 451–
1031 66. <https://doi.org/10.1007/s10745-015-9745-1>.
- 1032 Steege, Hans ter, Nigel C A Pitman, Daniel Sabatier, Christopher Baraloto, Rafael P Salomão,

- 1033 Juan Ernesto Guevara, Oliver L Phillips, et al. 2013. “Hyperdominance in the
1034 Amazonian Tree Flora.” *Science* 342 (6156).
- 1035 Sun, Jian, Xiangrong You, Li Li, Hongxiang Peng, Weiqiang Su, Changbao Li, Quanguang
1036 He, and Fen Liao. 2011. “Effects of a Phospholipase D Inhibitor on Postharvest
1037 Enzymatic Browning and Oxidative Stress of Litchi Fruit.” *Postharvest Biology and
1038 Technology* 62 (3): 288–94. <https://doi.org/10.1016/j.postharvbio.2011.07.001>.
- 1039 Thioulouse, Jean, Daniel Chessel, Sylvain Dolédec, and Jean Michel Olivier. 1997. “ADE-4:
1040 A Multivariate Analysis and Graphical Display Software.” *Statistics and Computing* 7
1041 (1): 75–83. <https://doi.org/10.1023/A:1018513530268>.
- 1042 Wallace, Alfred Russel. 1853. *Palm Trees of the Amazon and Their Uses*. J. van Voorst.
- 1043 Watling, Jennifer, José Iriarte, Francis E. Mayle, Denise Schaan, Luiz C.R. Pessenda, Neil J.
1044 Loader, F. Alayne Street-Perrott, Ruth E. Dickau, Antonia Damasceno, and Alceu Ranzi.
1045 2017. “Impact of Pre-Columbian Geoglyph Builders on Amazonian Forests.”
1046 *Proceedings of the National Academy of Sciences of the United States of America* 114
1047 (8): 1868–73. <https://doi.org/10.1073/pnas.1614359114>.
- 1048 Yazaki, Kazufumi. 2006. “ABC Transporters Involved in the Transport of Plant Secondary
1049 Metabolites.” *FEBS Letters*. No longer published by Elsevier.
1050 <https://doi.org/10.1016/j.febslet.2005.12.009>.
- 1051 Yin, Yuxin, Xiaowei Zhang, Yongjun Fang, Linlin Pan, Gaoyuan Sun, Chengqi Xin,
1052 Mohammed M.Ba Abdullah, et al. 2012. “High-Throughput Sequencing-Based Gene
1053 Profiling on Multi-Staged Fruit Development of Date Palm (*Phoenix Dactylifera*, L.)”
1054 *Plant Molecular Biology* 78 (6): 617–26. <https://doi.org/10.1007/s11103-012-9890-5>.
- 1055 Yoshida, K., P. Kaothien, T. Matsui, A. Kawaoka, and A. Shinmyo. 2003. “Molecular
1056 Biology and Application of Plant Peroxidase Genes.” *Applied Microbiology and*

1057 *Biotechnology*. Springer Verlag. <https://doi.org/10.1007/s00253-002-1157-7>.

1058