



# On the necessity of combining ethnobotany and genetics to assess agrobiodiversity and its evolution in crops: A case study on date palms (*Phoenix dactylifera* L.) in Siwa Oasis, Egypt

Muriel Gros-Balthazard, Vincent Battesti, Sarah Ivorra, Laure Paradis, Frédérique Aberlenc, Oumarou Zango, Salwa Zehdi-Azouzi, Souhila Moussouni, Summar Abbas Naqvi, Claire Newton, et al.

## ► To cite this version:

Muriel Gros-Balthazard, Vincent Battesti, Sarah Ivorra, Laure Paradis, Frédérique Aberlenc, et al.. On the necessity of combining ethnobotany and genetics to assess agrobiodiversity and its evolution in crops: A case study on date palms (*Phoenix dactylifera* L.) in Siwa Oasis, Egypt. *Evolutionary Applications*, 2020, early view, 10.1111/eva.12930 . hal-02375285v2

**HAL Id: hal-02375285**

**<https://hal.science/hal-02375285v2>**

Submitted on 29 Feb 2020 (v2), last revised 4 Sep 2020 (v3)











**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

# On the necessity of combining ethnobotany and genetics to assess agrobiodiversity and its evolution in crops: A case study on date palms (*Phoenix dactylifera* L.) in Siwa Oasis, Egypt

Muriel Gros-Balthazard<sup>1,2</sup>  | Vincent Battesti<sup>3</sup>  | Sarah Ivorra<sup>2</sup> | Laure Paradis<sup>2</sup>  |  
Frédérique Aberlenc<sup>4</sup>  | Oumarou Zango<sup>5</sup>  | Salwa Zehdi-Azouzi<sup>6</sup>  |  
Souhila Moussouni<sup>7</sup>  | Summar Abbas Naqvi<sup>8</sup>  | Claire Newton<sup>9</sup>  |  
Jean-Frédéric Terral<sup>2</sup> 

<sup>1</sup>Center for Genomics and Systems Biology, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

<sup>2</sup>Institut des Sciences de l'Évolution, UMR 5554 CNRS/Université de Montpellier/IRD/EPHE, CC065, Équipe Dynamique de la Biodiversité, Anthro-écologie, Université – Montpellier, Montpellier Cedex 5, France

<sup>3</sup>UMR 7206 Éco-anthropologie, CNRS, Muséum national d'histoire naturelle, Université de Paris: Musée de l'Homme, Paris, France

<sup>4</sup>Institut de Recherche pour le Développement (IRD), Université de Montpellier, UMR DIADE, Montpellier, France

<sup>5</sup>Université de Zinder, Zinder, Niger

<sup>6</sup>Faculty of Sciences, University of Tunis El Manar, Tunis, Tunisia

<sup>7</sup>Faculté des Sciences Biologiques, Laboratoire de Recherche sur les Zones Arides (LRZA), Université des Science et de la Technologie Houari Boumediene (USTHB), Alger, Algeria

<sup>8</sup>Institute of Horticultural Sciences, University of Agriculture, Faisalabad, Pakistan

<sup>9</sup>Rimouski, QC, Canada

## Correspondence

Muriel Gros-Balthazard, Center for Genomics and Systems Biology, New York University Abu Dhabi, Saadiyat Island, Abu Dhabi, United Arab Emirates.  
Email: muriel.grosb@gmail.com

## Abstract

Crop diversity is shaped by biological and social processes interacting at different spatiotemporal scales. Here, we combined population genetics and ethnobotany to investigate date palm (*Phoenix dactylifera* L.) diversity in Siwa Oasis, Egypt. Based on interviews with farmers and observation of practices in the field, we collected 149 date palms from Siwa Oasis and 27 uncultivated date palms from abandoned oases in the surrounding desert. Using genotyping data from 18 nuclear and plastid microsatellite loci, we confirmed that some named types each constitute a clonal line, that is, a true-to-type cultivar. We also found that others are collections of clonal lines, that is, ethnovarieties, or even unrelated samples, that is, local categories. This alters current assessments of agrobiodiversity, which are visibly underestimated, and uncovers the impact of low-intensity, but highly effective, farming practices on biodiversity. These hardly observable practices, hypothesized by ethnographic survey and confirmed by genetic analysis, are enabled by the way Isiwans conceive and classify living beings in their oasis, which do not quite match the way biologists do: a classic disparity of *etic* versus *emic* categorizations. In addition, we established that Siwa date palms represent a unique and highly diverse genetic cluster, rather than a subset of North African and Middle Eastern palm diversity. As previously shown, North African date palms display evidence of introgression by the wild relative *Phoenix theophrasti*, and we found that the uncultivated date palms from the abandoned oases share even more alleles with this species than cultivated palms in this region. The study of Siwa date palms could hence be a key to the understanding of date palm diversification in North Africa. Integration of ethnography and population genetics promoted the understanding of the interplay between diversity management in the oasis (short-time

Gros-Balthazard and Battesti contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd

Vincent Battesti, UMR 7206 Éco-anthropologie, CNRS, Muséum national d'histoire naturelle, Université de Paris: Musée de l'Homme, Paris, France.  
Email: vincent.battesti@mnhn.fr

#### Funding information

Mosaïque program, Grant/Award Number: GDR 3353 INEE CNRS; New York University Abu Dhabi Research Institute, Grant/Award Number: Biodiversity genomics; BioDivMeX (MIST RALS programm)

scale), and the origins and dynamic of diversity through domestication and diversification (long-time scale).

#### KEYWORDS

agrobiodiversity, anthropology, date palm (*Phoenix dactylifera* L.), domestication, ethnobotany, evolutionary history, farming practices, folk categorization, microsatellite markers, *Phoenix theophrasti* Greuter, population genetics, Siwa Oasis (Egypt)

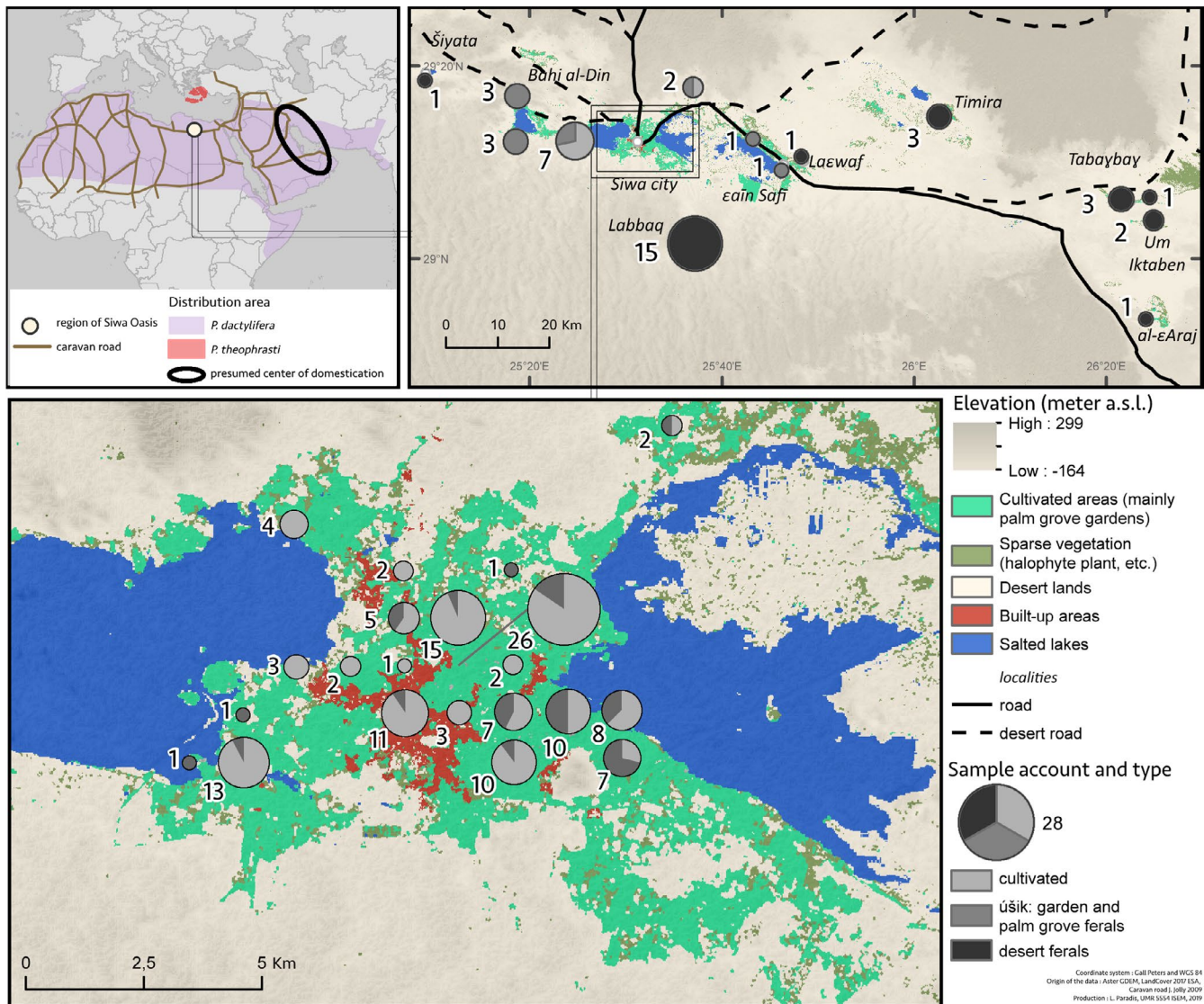
## 1 | INTRODUCTION

The date palm, *Phoenix dactylifera* L., is a major perennial crop of the hot and arid regions in the Middle East and North Africa (Barrow, 1998). Its sugar-rich fruit, the date, has been consumed for millennia (Tengberg, 2012) and has long been rooted in Berber/Amazigh and Arabic cultures. *Phoenix dactylifera* belongs to the Arecaceae family and, along with 12 or 13 other interfertile species, composes the genus *Phoenix* (Barrow, 1998). Date palms exist mostly as either cultivated or feral (i.e., uncultivated but derived from cultivated palms) domesticated forms (for review, Gros-Balthazard, Hazzouri, & Flowers, 2018). Only a few relictual populations of its wild progenitor are known today in Oman (Gros-Balthazard et al., 2017), even though Tuaregs of the Tassili n'Ajjer (Algeria) consider it to be wild in their gardens (Battesti, 2004). *Phoenix dactylifera* is dioecious (either male or female, and only females bear fruit). Today, thousands of female cultivars are reported, varying in fruit shape, color, texture, or taste, but also in their vegetative aerial architecture (Chao & Krueger, 2007). Without addressing the existence of male "cultivars" (there is little or no vegetative reproduction of identified and named males outside of research stations), male varieties are also locally identified, but are less studied. Nevertheless, local categorization of this diversity by farmers requires clarification, starting with the adequacy of the notion of "cultivar" or what it is supposed to be.

Phoeniculture (date palm cultivation) involves a mix of clonal and sexual propagation. In palm gardens, palm trees bear a name. The unquestioned and often implicit assumption is that each female date palm named type (for instance the famed Medjool or Khalas) is a cultivar, that is a clone, multiplied only through vegetative propagation. In order to obtain specifically female trees and to ascertain that fruits will be of the desired predictable quality, farmers mostly make use of the asexual reproduction abilities of this plant through offshoot multiplication. Indeed, sexual reproduction of date palms only leads in rare cases to progenies having equivalent or superior fruit qualities (4% according to Peyron, 2000), although these assessments remain subjective. For the oasis system to be efficient, despite the scarcity of water, irrigable lands, and manure, oasis communities plant and maintain 95%–99% of female palms (Battesti, 2005) instead of a natural 50:50 sex ratio (Chao & Krueger, 2007). Opting for such an artificial sex ratio requires hand pollination, a practice already used in southern Mesopotamia during the 3rd millennium BC (Landsberger, 1967), as a substitute to natural wind

pollination (Henderson, 1986). Hence, farmers almost entirely restrict sexual reproduction by seeds, with a few exceptions (e.g., in India, Newton et al., 2013). Accidental seedlings are, however, sometimes spared. A resulting male may be later used for pollination. If female, its fruits are sometimes harvested and, although rare, it can lead to a new selected and named line of clones, that is, a cultivar. Another possibility, highlighted in our previous work in the oasis of Siwa, is to integrate this new genotype into an existing named type, because, from a local perspective, it is the very same variety, the same "form," or phenotype (Battesti, 2013; Battesti et al., 2018). This farming practice challenges the presumption of a named type being a true-to-type cultivar, that is, aggregating solely vegetatively propagated individuals. This may, in turn, lead to a misinterpretation and an inaccurate estimation of local agrobiodiversity.

Siwa is a desert oasis located in the Libyan Desert 300 km south of the Mediterranean coast and the closest city Marsa Matruh, and about 30 km east of the Libyan border (Figure 1). The territory is occupied by salt lakes, the "*sabkha*," and by the cultivated area in the form of palm groves, highly concentrated around settlement areas (Figure 1; Figure S1A). With roughly 200,000 to 250,000 palms (Battesti, 2013), dates are the main commercial crop in Siwa Oasis, closely followed by olive (*Olea europaea* L.). As in most oasis systems, date palm is used to feed the oasis inhabitants, but also for drinks, fodder, building materials (beams, hedges), crafts (baskets, various utensils), and daily uses (ropes, ties, brooms, furniture) (Battesti, 2005). The date palm is of primary importance, more generally, as the keystone of the oasis ecological system (microclimate effect, Riou, 1990). It is also the cornerstone of the local economy of exportation to large urban centers. The oases have a very relative autarchy and self-sufficiency, and their inhabitants export what they have in abundance: dates (Battesti, 2013; Battesti et al., 2018). Extrapolating from early 20th century data in Siwa, an estimate of 82% of the dates in value terms was exported (and 95% of the elite cultivar *şaēidi*) (Hohler & Maspero, 1900). Several population genetics analyses focused on date palm agrobiodiversity in Siwa (Abd El-Azeem, Hashem, & Hemeida, 2011; Abou Gabal, Abedel Aziz, Hardash, & El-Wakil, 2006; El-Sharabasy & Rizk, 2019; el-Wakil & Harhash, 1998; Hemeid, Sanaa, & Abd El-Rahman, 2007; Selim, El-Mahdi, & El-Hakeem, 1970). Nevertheless, none of these studies integrate agricultural knowledge and practices, and the existing diversity in Siwa Oasis and its categorization system remain thus poorly understood.



**FIGURE 1** Localization of Siwa Oasis and sampling strategy in Siwa region

In Siwa Oasis, we previously described a complex system where named types are not necessarily cloned genotypes, but ethnovarieties or local categories (Battesti, 2013; Battesti et al., 2018). To synthesize our “objectification” of the local organization of the date palm agrobiodiversity system, we proposed these definitions:

- **Cultivar or true-to-type cultivar:** a set of clonal individuals, that is, the association of a name and a single genotype reproduced vegetatively (asexually, by offshoot) by humans. Genotypes of a same cultivar are genetically identical, except in case of somatic mutations (McKey et al., 2010).
- **Ethnovariety:** a set of similar (according to local standards) lines of clonal individuals reproduced vegetatively (asexually, by offshoot) by humans deliberately under a single local name.
- **Local category:** a set of individuals sharing some characteristics (according to local standards, i.e., fruit color, harvesting season, usage, rusticity), typically found as seedlings, sometimes reproduced vegetatively (asexually, by offshoot) by humans but

identified under a single local name. The two largest local categories are *ótem*, the male date palms, and *úšik*, which includes all date palms resulting from sexual reproduction (not to be confused with the named types with the form *úšik xxx*, such as *úšik n gubel*). Farmers can “qualify” them sometimes individually: as *úšik maεasil* or *úšik εalafī*, etc., conceiving them, as such, as qualities, not varieties.

The origins of this date palm heritage and Siwa Oasis altogether are lost in the mists of time. Siwa Oasis was well known during antiquity, in its Egyptian dynastic period, under the name of “Sexet-ām,” the land of date palms (Duemichen, 1877), that included the current oasis but also the oases now abandoned in its periphery (Kuhlmann, 2013). In the classical period, the oasis was famous throughout the Mediterranean Basin for its oracle built in the 6th century BCE (Kuhlmann, 1988, 2013; Leclant, 1950), under the name of Amon Oasis (later Hellenized in Ammon). Alexander the Great was one of the most famous to consult this desert



oracle, for him to confirm his divine ascendancy before his campaign of conquest in Persia. The dates of Siwa were mentioned as early as the 5th century BCE by Hellanicus of Mytilene (c. 480–c. 395 BCE) in his *Journey to the Oracle of Ammon* (cited by Leclant, 1950, p. 248). The date palms of Siwa were then mentioned or even celebrated by Theophrastus (c. 371–c. 288 BCE) and later Pliny the Elder (23–79 CE) and Arrian (c. 95–v. 175), before a long silence and a return under the famous Arabic authors al-Bakrī (1040–1094), al-Idrīsī (1100–1165), and al-Maqrīzī (1364–1442). A few genetic studies involved date palms from Siwa (Abd El-Azeem et al., 2011; Abou Gabal et al., 2006; Gros-Balthazard et al., 2017; Hemeid et al., 2007). Nevertheless, not all named types of the oasis were studied, levels of categorization of date palm names were confused (sometimes mistaking *úšik*, seedlings, for a cultivar), farming practices were neglected, and only date palms cultivated in the current oasis were considered, ignoring the abandoned oases scattered in the desert. A proper assessment of date palm agrobiodiversity in Siwa region, and in a broader sense an understanding of its origin, hence, is still lacking.

Beyond the peculiar history of Siwa Oasis, the scenario of the beginning of date palm cultivation in Egypt in particular, and in North Africa in general, remains incomplete. Older evidence of exploitation is found around the Persian Gulf, while it seems that in North Africa, cultivation is more recent (Flowers et al., 2019; Tengberg, 2012). In Egypt, the date palm seems exploited or cultivated sporadically since the Old Kingdom (about 2,700–2,200 BCE), but phoeniculture is only established since the New Kingdom, about 1,600–1,100 BCE (Tengberg & Newton, 2016). Genetic analyses of the current date palm germplasm identified two differentiated genetic clusters in North Africa and the Middle East, with evidence of gene flows, especially in Egypt (Flowers et al., 2019; Gros-Balthazard et al., 2017; Hazzouri et al., 2015; Mathew et al., 2015; Zehdi-Azouzi et al., 2015). A recent study showed that the cultivated North African pool has mixed ancestry from Middle Eastern date palms and the Aegean endemic wild relative *Phoenix theophrasti* Greuter, a.k.a. the Cretan date palm (Flowers et al., 2019). Nevertheless, the geographic, chronological, and historical contexts of this introgression remain enigmatic. The oasis of Siwa is located at the crossroads between Greek, Libyan, and Egyptian influences. It is on one of the rare passage points, a rare node in the network (Battesti et al., 2018), between the east and west of the distribution area of the cultivated date palm (Figure 1). A deep understanding of date palm diversity in the region could therefore enlighten the diversification history of *Phoenix dactylifera* in North Africa.

In this paper, we are taking our previous work on Siwa date palms (Battesti et al., 2018) to the next level, using a joint ethnographic study and genetic analysis. First, we greatly expanded the number of named types sampled in Siwa to further test whether named types are true-to-type cultivars or ethnovarieties or local categories. Secondly, we also increased our sampling of uncultivated individuals from the abandoned oases in the desert nearby Siwa (also known as “feral” in Battesti et al., 2018). This enabled

a full assessment of local biodiversity and potential connections between the currently cultivated pool of the oasis and the uncultivated abandoned date palms. Lastly, we used genotyping data from more than 200 cultivated date palms sampled across the entire historical range of the species in order to locate Siwa diversity located within a wider germplasm diversity. By including *Phoenix theophrasti*, we can assess a possible gene flow in this particular population of date palms. From a broader perspective, our work aims at documenting the origins of phoeniculture in Egypt and in North Africa in general.

## 2 | MATERIALS AND METHODS

### 2.1 | Plant samples and genotyping

#### 2.1.1 | Ethnobotanical study and sample collection

We sampled 176 cultivated and uncultivated date palms in Siwa Oasis (Egypt) and surrounding desert (Figure 1; Table 1; Table S1), of which 52 were included in our previous study (Battesti et al., 2018). Those samples were collected in situ with the essential cooperation of the local farmers while conducting an ethnobotanical study (Figure S1B). VB conducted social anthropological fieldwork between 2002 and 2017, including about six months dedicated to date palm categorization and naming (Battesti, 2013). While he occasionally conducted structured interviews or held focus group discussions and free listings of date palm given names, most of his data are derived from participant observation (Battesti et al., 2018). This collection is composed of 109 accessions of date palms growing in about 46 private gardens that were deliberately chosen scattered throughout the current oasis. For each named type, we collected more than one date palm, when possible (some are rare), in order to test, with genetic data, whether they represent cloned accessions or not. We also sampled accidental seedlings growing in gardens (referred to as *úšik* #1) and on the border of gardens or palm groves (referred to as *úšik* #2). Further, 27 uncultivated date palms growing in abandoned oases in the desert were sampled (Figure 1). Those oases probably already existed during the Roman/Ptolemaic period, or at least some of them, and have been presumably abandoned since the 9th or 10th century CE (Battesti, 2013).

We additionally sampled nine *Phoenix theophrasti* Greuter in their native habitat in Crete and Turkey. Two accessions of *Phoenix reclinata* Jacq., collected in the botanical garden of the Villa Thuret in Antibes, France, and originally from sub-Saharan Africa, were included as outgroup population. For each accession, a few leaflets were collected and dried in the shade, with or without silica gel.

### 2.2 | DNA extraction and microsatellite genotyping

The collection was genotyped using 17 nuclear microsatellites and one chloroplastic minisatellite (Table 2), following the protocol of

**TABLE 1** Summary of the 406 accessions of *Phoenix* spp. included in this study

Species	Origin	Status	No. of samples
<i>Phoenix dactylifera</i>	Morocco	Cultivated	15
	Mauritania	Cultivated	5
	Niger	Cultivated	20
	Algeria	Cultivated	30
	Tunisia	Cultivated	27
	Libya	Cultivated	7
	Nilotic Egypt	Cultivated	14
	Siwa, Egypt	Named types	109
		úšik #1	18
		úšik #2	22
		Uncultivated (abandoned oases)	27
	Sudan	Cultivated	3
	Iraq	Cultivated	13
	United Arab Emirates	Cultivated	20
	Oman	Cultivated	27
	Pakistan	Cultivated	38
<i>Phoenix theophrasti</i>	Crete/Turkey	Wild relative	9
<i>Phoenix reclinata</i>	Sub-Saharan Africa	Wild relative	2

Note: The species and the country/region of origin are provided along with the status. For Siwa current oasis, the status is set as follows: "named types," for cultivated accessions for which we want to check the status; "úšik #1" and "úšik #2" which refer to accidental seedlings growing respectively in gardens (#1) or on the border of gardens or palm grove (#2). Were also analyzed uncultivated date palm accessions sampled in abandoned oases scattered in the desert surrounding the current Siwa Oasis (Figure 1).

Zehdi-Azouzi et al. (2015). We crushed 40 mg of dried leaves in a fine powder using bead-mill homogenizer TissueLyser (Qiagen, Courtabœuf, France). Total genomic DNA was extracted from leaf powder using DNeasy Plant MINI Kit (Qiagen, Courtabœuf, France). The plastid dodecanucleotide minisatellite identified in the intergenic spacer psbZ-trnFM (Henderson, Billotte, & Pintaud, 2006) was genotyped. It has been previously used to define the date palm chlorotype (so-called occidental or oriental), depending on its number of repeats (three or four, respectively), and to barcode *Phoenix* species (Ballardini et al., 2013; Gros-Balthazard et al., 2017; Pintaud et al., 2010; Zehdi-Azouzi et al., 2015).

In addition to this newly generated genotyping dataset, we utilized *P. dactylifera* genotyping data from previous studies (Moussouni, Pintaud, Vigouroux, & Bouguedoura, 2017; Zango et al., 2017; Zehdi-Azouzi et al., 2015) (Table S1). Both these data and our new data relied on the same set of microsatellite markers and were generated by the same company (ADNid, Montpellier, France) with the same protocol, enabling a meta-analysis. These additional accessions are a good representation of the cultivated germplasm as their origin spans the historic date palm distribution, stretching from North Africa to the Middle East and Pakistan (Barrow, 1998).

## 2.2.1 | Genotyping data analysis

Statistical analysis was conducted with the R Statistical Programming Language (R Core Team, 2015), unless otherwise stated. To identify duplicated genotypes among the whole dataset, we performed an identity analysis using Cervus v3.0.7 (Kalinowski, Taper, & Marshall, 2007). For each pair of accessions, Cervus calculates the number of matching genotypes across the 17 nuclear loci. We used the same software for calculation of the polymorphic information content (PIC) for each locus. We estimated null allele frequencies using *null.all* function in the R package *PopGenReport* (Adamack & Gruber, 2014; Gruber & Adamack, 2015). For each locus, the number of alleles  $N_A$ , and the observed ( $H_O$ ) and expected heterozygosity ( $H_S$ ) were estimated using the R package *pegas* (Paradis, 2010). The deviation from Hardy-Weinberg equilibrium was estimated using the function *hw.test* from the same package.

## 2.2.2 | Date palm agrobiodiversity in Siwa

For each named type of Siwa, we checked whether they are actual true-to-type cultivars or rather represent a group of more or less

**TABLE 2** List of microsatellite loci and their summary statistics calculated on the 406 *Phoenix* spp. accessions included in the present study

Locus	References	% MD	N <sub>0</sub>	# All	PIC	Ho	He	HWD
cpM12	Henderson et al. (2006)	7.35	/	5	/	/	/	/
mPdCIR015	Billotte et al. (2004)	0.25	0.038	12	0.771	0.74	0.8	0.00
mPdCIR016	Billotte et al. (2004)	0.00	0.090	6	0.663	0.6	0.71	0.00
mPdCIR032	Billotte et al. (2004)	0.74	0.098	13	0.685	0.58	0.71	0.00
mPdCIR035	Billotte et al. (2004)	8.58	0.170	12	0.554	0.42	0.59	0.00
mPdCIR057	Billotte et al. (2004)	0.00	0.065	11	0.599	0.55	0.62	0.00
mPdCIR085	Billotte et al. (2004)	0.00	0.038	19	0.847	0.8	0.86	0.00
PdAG1-ssr	Billotte et al. (2004)	0.00	0.053	36	0.888	0.81	0.9	0.00
mPdCIR010	Billotte et al. (2004)	0.00	0.055	18	0.771	0.71	0.79	0.00
mPdCIR025	Billotte et al. (2004)	0.00	0.025	18	0.791	0.77	0.81	0.00
mPdCIR063	Billotte et al. (2004)	2.21	0.087	14	0.681	0.61	0.72	0.00
mPdCIR078	Billotte et al. (2004)	0.00	0.057	27	0.886	0.79	0.88	0.00
PdCUC3-ssr1	Accession number: HM622273	0.25	0.000	5	0.017	0.02	0.02	1.00
PdCUC3-ssr2	Accession number: HM622273	0.00	0.119	9	0.885	0.77	0.86	0.00
mPdIRD013	Aberlenc-Bertossi et al. (2014)	0.00	0.072	31	0.159	0.7	0.88	0.00
mPdIRD031	Aberlenc-Bertossi et al. (2014)	0.49	0.036	4	0.356	0.13	0.16	0.00
mPdIRD033	Aberlenc-Bertossi et al. (2014)	0.00	0.110	4	0.479	0.33	0.35	0.00
mPdIRD040	Aberlenc-Bertossi et al. (2014)	0.00	0.040	5	0.593	0.39	0.48	0.00

Abbreviations: % MD, Proportion of missing data; N<sub>0</sub>, null allele frequency; # All, number of alleles; PIC, polymorphic information content; Ho and He, observed and expected heterozygosity; HWD, deviation from Hardy-Weinberg p-value.

distant genotypes (ethnovariety, local category). For this purpose, we calculated a measure of identity by averaging over each named type the proportion of matching genotypes across the 18 chloroplastic and nuclear loci. For true-to-type cultivars, this number is expected to be 100%, except in case of somatic mutations or genotyping error. Second, we calculated the Euclidean distance between each of the Siwa samples (function *dist*, R package *stats*) and built a heatmap of those distances using *heatmap* function (R package *stats*).

To investigate the extent of the diversity in the whole region, and not only in the oasis, we included uncultivated date palms from the abandoned oases in the following analyses. We generated a neighbor-joining tree based on Nei's genetic distance (Nei, 1972) using *aboot* function in *poppr* package (Kamvar, Tabima, & Grünwald, 2014) with 100 bootstrap replicates.

For each line of clones identified with these analyses, we kept a single accession for downstream analyses and performed a principal component analysis using function *dudi.pca* in the *ade4* package (Dray & Dufour, 2007). Missing data were replaced by the mean allele frequencies using the function *scaleGen* from the R package *ade4* (Jombart & Ahmed, 2011).

### 2.2.3 | Extent and partitioning of worldwide date palm diversity

To draw up a picture of the overall *Phoenix* population structure, we applied two different approaches to our dataset comprising 128

unique genotypes from Siwa, 219 date palms originating from all over its historical distribution area, and nine wild *Phoenix theophrasti*. First, we performed a principal component analysis, as described above for Siwa accessions only. We further used the Bayesian clustering method based on a Markov chain Monte Carlo (MCMC) algorithm implemented in *Structure* v2.3.3 (Pritchard, Stephens, & Donnelly, 2000). Individuals are partitioned into a predefined number of clusters (K) so as to minimize linkage disequilibrium and deviation from Hardy-Weinberg within cluster. Allelic frequencies at each locus are calculated at the same time for each cluster. Genotyped individuals were allocated to one to eight clusters K. All runs were performed using a model allowing admixture and correlated allele frequencies among populations (Falush, Stephens, & Pritchard, 2003). A 100,000-iteration burn-in period was followed by 1,000,000 MCMC steps. Ten independent runs were performed for each specified K, and the convergence of likelihood values was checked for each K. The optimal value of K was estimated using both the approach of Pritchard et al. (2000) based on the maximization of the log likelihood, and the approach of Evanno, Regnaut, and Goudet (2005) based on the rate of change in the log likelihood between successive K values (delta K).

For each population pair, we estimated a measure of differentiation ( $F_{ST}$ ) using the *Genepop* software v. 4.7 (Rousset, 2008). The proportion of shared alleles between populations and subpopulations was calculated with the *pairwise.propShared* function of the R package *PopGenReport* (Adamack & Gruber, 2014; Gruber & Adamack, 2015). Further, we calculated various diversity estimates

for each population. The allelic richness and private allelic richness were calculated with a custom R script. Because the numbers of distinct alleles and private alleles depend heavily on sample size in each population, we used the rarefaction method (Petit, El Mousadik, & Pons, 1998), allowing a direct comparison among populations of different sample size. For each set of comparison, we thus used the smallest haploid sample size ( $n = 18$  in *Phoenix theophrasti*) as the number of alleles to sample, and ran 1,000 replicates of allelic and private allelic richness calculation. To test for significant differences among populations, these diversity estimates were assigned to Tukey groups (function *HSD.test*, *Agricolae* package, de Mendiburu, 2015) at a significant threshold of 0.05. Expected and observed heterozygosity ( $H_S$  and  $H_O$ ) were calculated using the *basic.stats* function of the R package *hierfstat*. The inbreeding coefficient  $F_{IS}$  was calculated with the same package, and confidence intervals were estimated by performing 1,000 bootstrap replicates over loci with the function *boot.ppfis*.

### 3 | RESULTS

A total of 406 *Phoenix* spp. samples (Table 1; Table S1) genotyped across 17 nuclear microsatellites and one chloroplastic minisatellite (Table 2) were analyzed in the present study. We report new genotyping data for 176 cultivated and uncultivated date palms from the Egyptian oasis of Siwa and the surrounding abandoned oases (Figure 1; Figure S1C), nine *Phoenix theophrasti*, and two *Phoenix recinata* (Table S2). Additional genotyping data for 98 Middle Eastern/Asian and 121 North African date palms were retrieved from previous studies (Moussouni et al., 2017; Zango et al., 2017; Zehdi-Azouzi et al., 2015).

Missing data across the full dataset were very limited with an average of 1.1%, and the mean null allele frequency was 6.8%, on average, across all 18 chloroplastic and nuclear loci (Table 2). All loci were polymorphic, with four to 31 alleles and an average polymorphic information content of 0.625. All loci deviated significantly ( $p < .05$ ) from Hardy-Weinberg equilibrium except *Cuc3-ssr1*. Except for some Siwa samples (see below), all accessions were unique, as no pair of accessions displayed 18 matches across the 18 chloroplastic and nuclear loci (Table S3).

#### 3.1 | Ethnographic and genetic analysis of date palms in Siwa region

##### 3.1.1 | On the local named types of Siwa Oasis

The ethnobotanical field survey identified the existence of 18 named types in Siwa Oasis (Table 3). Based on ethnographic work, some were hypothesized as true-to-type cultivars, as they are supposed to only arise by planting offshoots, according to farmers' accounts. Meanwhile, the fieldwork already allowed us to suppose that some of those named types are not true-to-type cultivars, as

**TABLE 3** Named type of Siwa, sampling effort, and genetic identity

Named type	No. of sample	Nuc Identity	Cp Identity
şæidi [tasutet]	13	97.59%	100%
alkak	11	100%	100%
aḡzāl	4	97.06%	100%
ḡalu en ḡanem	3	100%	100%
taṡṡagt	8	100%*	100%
úšik niqbel/ úšik en gubel	12	69.96%	92.86%
ḡrom aḡzāl	6	28.02%	88.03%
alkak wen žemb	6	28.24%	88.44%
ḡrom şæid	2	29.41%	0%
lekrawmet	4	40.20%	100%
úšik amayzuz/ amayzuz	2	23.53%	100%
úšik azzugaḡ	4	31.92%	100%
amenzu	7	24.56%	90.68%
úšik nekwayes	11	29.41%	84.04%
úšik ezzuway/ zuway [tazuwayt]	8	27.95%	46.43%
kæibī	6	20.39%	40.00%
úšik #1	18	23.97%	47.06%
úšik #2	22	27.02%	58.44%
tazuwayt	1	/	/
tažubart	1	/	/

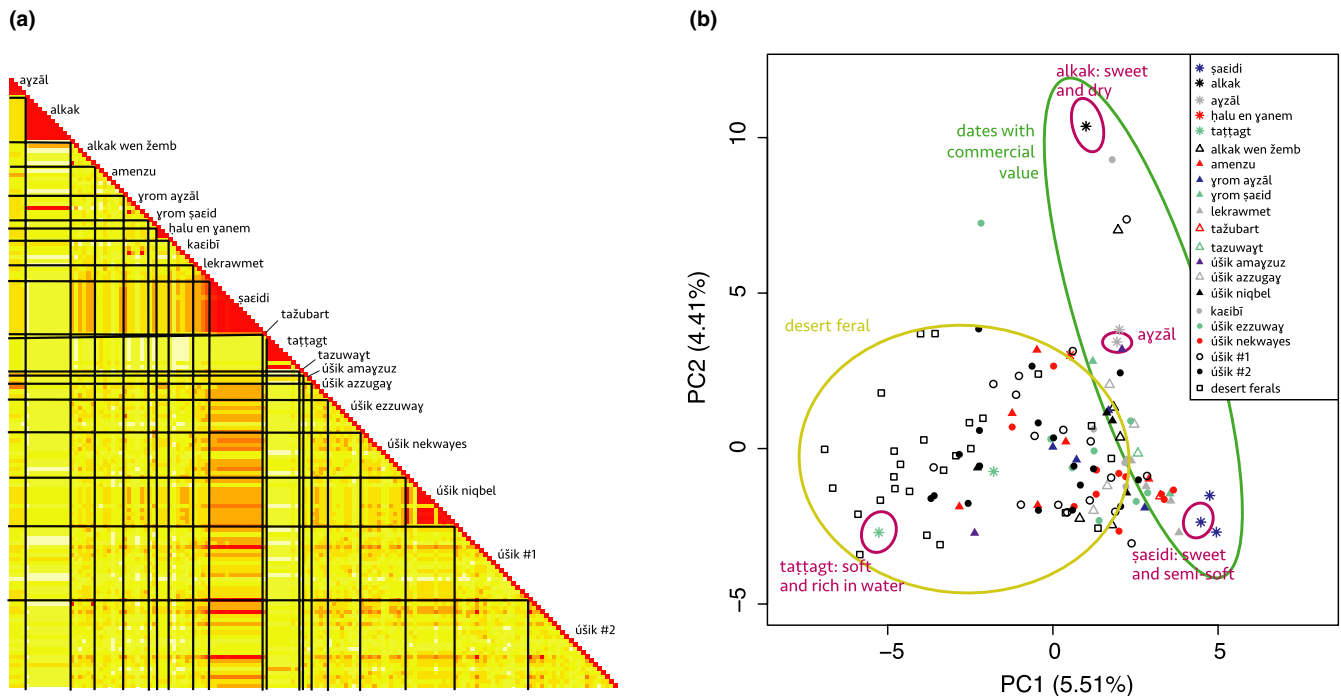
Note: For each named type, the local name of the dates is given, followed by the possible name of the palm [in brackets]. The identity is the average proportion of matching genotypes among accessions of a given named type and across the 17 nuclear microsatellites (Nuc Identity) and the chloroplastic minisatellite (Cp Identity).

\*one accession (3,377) was identified as an identification error of the informer. When included, average identity across this set of accession is 79.41%.

farmers recognized the possibility that they partially or totally arise from seeds.

To investigate this question, we calculated genetic identity among accessions of each given named type, based on 17 nuclear microsatellites and one chloroplastic minisatellite (Table S3). The named types şæidi and alkak, the main cultivated and exported dates of Siwa, show proportion of identities of 97.6% and 100% at the nuclear level, respectively, and of 100% at the chloroplastic locus (Figures 2 and 3; Table 3). The named type alkak is locally known to come only from an offshoot, but different "qualities" can be distinguished, depending on growth conditions and age. The şæidi is the emblematic "palm reproduced by offshoot," which Isiwans always oppose to úšik, the seedlings. Although slightly different (on a single locus; Table S2), the 13 şæidi accessions cluster together (Figure 3), and we hypothesize that this subtle genetic variation is due to somatic mutations or genotyping error. Hence, the two Siwa elite named types are indeed true-to-type cultivars.





**FIGURE 2** Intra-named type variation and structure in Siwa date palms. (a) Relatedness among the 149 cultivated date palms sampled in the current Siwa Oasis and genotyped across 17 nuclear microsatellite loci. Intra-named type variation is expected to be zero or near zero (red), in case of somatic mutations, while for ethnovarieties and local categories, we expect a higher intra-named type variation (yellow). (b) Principal component analysis of 128 unique date palm genotypes from both the current oasis of Siwa and the abandoned oasis of the region. Variance explained by each principal component (PC) is provided within parentheses

Have been also analyzed as true-to-type cultivars three other named types: ayzāl, tattagt, and halu en yanem. The famous ayzāl—rare but valued by the Isiwan as a tonic for lack of energy (Fakhry, 1990, p. 27) and for its aphrodisiac properties (for men)—is here represented by four accessions, three of which being identical and one showing a single allelic difference that we thus hypothesized to be a somatic mutation or genotyping error (Figure 3; Table 3; Tables S2–S3). The named type tattagt refers to a very soft date rich in water, whose name is taken from a maturation stage in the Siwa language: One half of the fruit is brown and ripe, and the other half is yellow and immature. Because it keeps very poorly, it is always eaten on the spot, but is highly appreciated locally. It was studied using eight accessions, of which a single one (3377) appears to be very different from the other genetically uniform accessions (Figures 2,3; Table 3; Tables S2–S3). The informant for this accession was a young farmer, and we presume that he was wrong about his identification. Hence, we believe that tattagt is a true-to-type cultivar, even if different qualities of tattagt are reported by our ethnography. The fifth confirmed cultivar is halu en yanem (100% identity over the three studied accessions in both nuclear and chloroplastic loci, Figure 3; Table 3; Tables S2–S3): relatively few farmers in Siwa know of its existence, but its long dates are highly appreciated by connoisseurs,

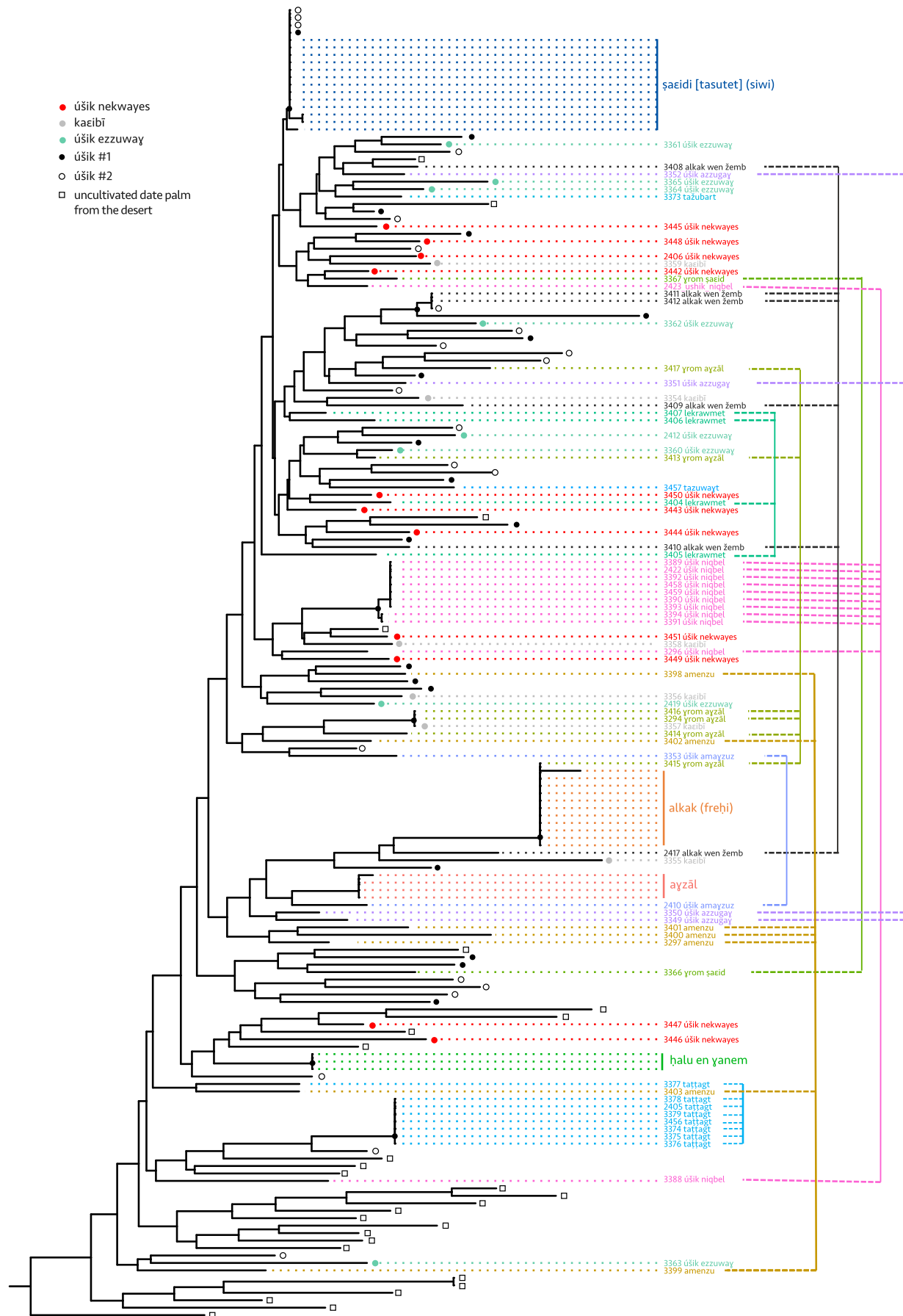
so soft and sweet that the seed remains attached to the bunch when the fruit is plucked.

Further, we can confirm some named types as ethnovarieties: yrom ayzāl and úšik niqbel indeed group some individuals having the same genotypes (clones) but also some that are isolated in different clades (Figure 3; Table 3). The first is said to resemble ayzāl and the second şaeidi, but in both cases of lower quality.

An intermediate case is alkak wen žemb: The six samples are unrelated except two which are from the same clone, but coming from the same garden a few meters away. Etymologically, alkak wen žemb is a “relegated,” “put aside,” that is, a second class alkak, but it has always been, without much ambiguity, referred to as a cultivar and cannot come from a seed, according to farmers. Under the names lekrawmet, amenzu (etymologically “early” dates, Laoust, 1932), or yrom şaeidi, we only identified unrelated individuals (Figures 2 and 3; Table 3). These are clearly not true-to-type cultivars despite them being explicitly thought by local farmers as never coming from a seedling; apparently, they often have been, although the occurrences may be too distant for local memory. They are thus interpreted as ethnovarieties.

Decisions are more difficult to reach in the following cases: úšik azzugay, úšik amayzuz, tazubart, and tazuwayt. The last two are only

**FIGURE 3** Genetic similarity of Siwa date palms based on neighbor-joining tree reconstructed from Nei's genetic distances. On the right of the tree, colored vertical lines connect samples of the same named types. Nodes supported by bootstrap values > 80% are indicated by black dots. The tree has been rooted with two *Phoenix reclinata* (not shown). The different samples of a true-to-type cultivar are expected to aggregate together, while samples of ethnovarieties would form different lines of clones, and samples of local categories would be scattered across the tree, denoting their lack of relatedness



known to a handful of farmers, and we have only one sample of each, preventing an intra-type identity analysis. Tažubart is said to look like šaēidi and tazuwayt like úšik ezzuway (zuway). The first two are more widely known, but two amayzuz (etymologically “late” dates, Laoust, 1932) samples do not allow us to reach a conclusion. Regarding úšik azzugay, we have two closely related palms from unrelated gardens and farmers, but also two isolated samples. Local farmers say úšik azzugay are not (necessarily) reproduced by offshoot and can be from a seedling (úšik). They are sometimes described as “all úšik (seedlings) that yield red dates,” “azzugay” meaning clearly “red” in *jlan en lsiwan*, the local Amazigh language. By combining genetic information and ethnographic data, these four types thus all seem to be possible “local categories.”

We are quite certain that the following named types are neither cultivars nor ethnovarieties: obviously, and as expected, the inclusive local category of úšik in general (all female seedlings), but also úšik ezzuway (úšik selected/cultivated with good reddish/dark dates), kaēibī (úšik selected/cultivated with good dry dates), and úšik nekwayes (all úšik selected/cultivated with dates suitable for human consumption). Indeed, the field survey showed they refer to individuals sharing fruit characteristics and/or originate from seedlings, and the genetic analysis confirmed that samples are unrelated (Figures 2 and 3; Table 3). Four úšik nevertheless present the same microsatellite profiles as the Siwa widespread cultivar šaēidi (Figure 3) meaning that these were actually šaēidi but thought wrongly, by the sampler, to be accidental seedlings (because they were found in abandoned areas of the palm grove).

Following this identity analysis, we removed 58 genotypes from the initial Siwa dataset in order to only keep 128 unique genotypes in downstream analyses.

### 3.1.2 | Structure of date palm diversity within Siwa region

To study the structure of the diversity of date palms in the region of Siwa, we studied named types and seedlings (úšik) from the oasis, but also uncultivated date palms from abandoned oases scattered across the desert surrounding Siwa (Figure 1). We found that the genetic diversity is mainly distributed between oasis samples versus samples from outside the oasis (i.e., the ancient abandoned oases in the desert), as the principal component (PC) 1 mostly draws apart those two types of date palms (Figure 2b). Identically, the NJ tree shows that the uncultivated date palms from the desert roughly form a distinct clade (Figure 3). Some accessions from the current oasis can nevertheless be found in the cluster formed by the desert uncultivated palms, mostly seedlings (úšik #1 and #2), but also named types, such as tažtagt. The second PC opposes the two main cultivars, namely alkak and šaēidi. Alkak, and a few other accessions, appears isolated from the other accessions of the region, including its relegated counterpart, alkak wen žemb (Figure 2b). The seedlings (úšik) do not form a distinct population. Here, we differentiated the úšik from the gardens that may be tended, depending on their use,

irrigated, as other date palms of the garden, and pollinated (úšik #1) from those that are not (úšik #2). For farmers, there is no difference and we could not separate them based on their microsatellite profiles (Figure 2b).

## 3.2 | Comparing the diversity of date palm in Siwa and worldwide

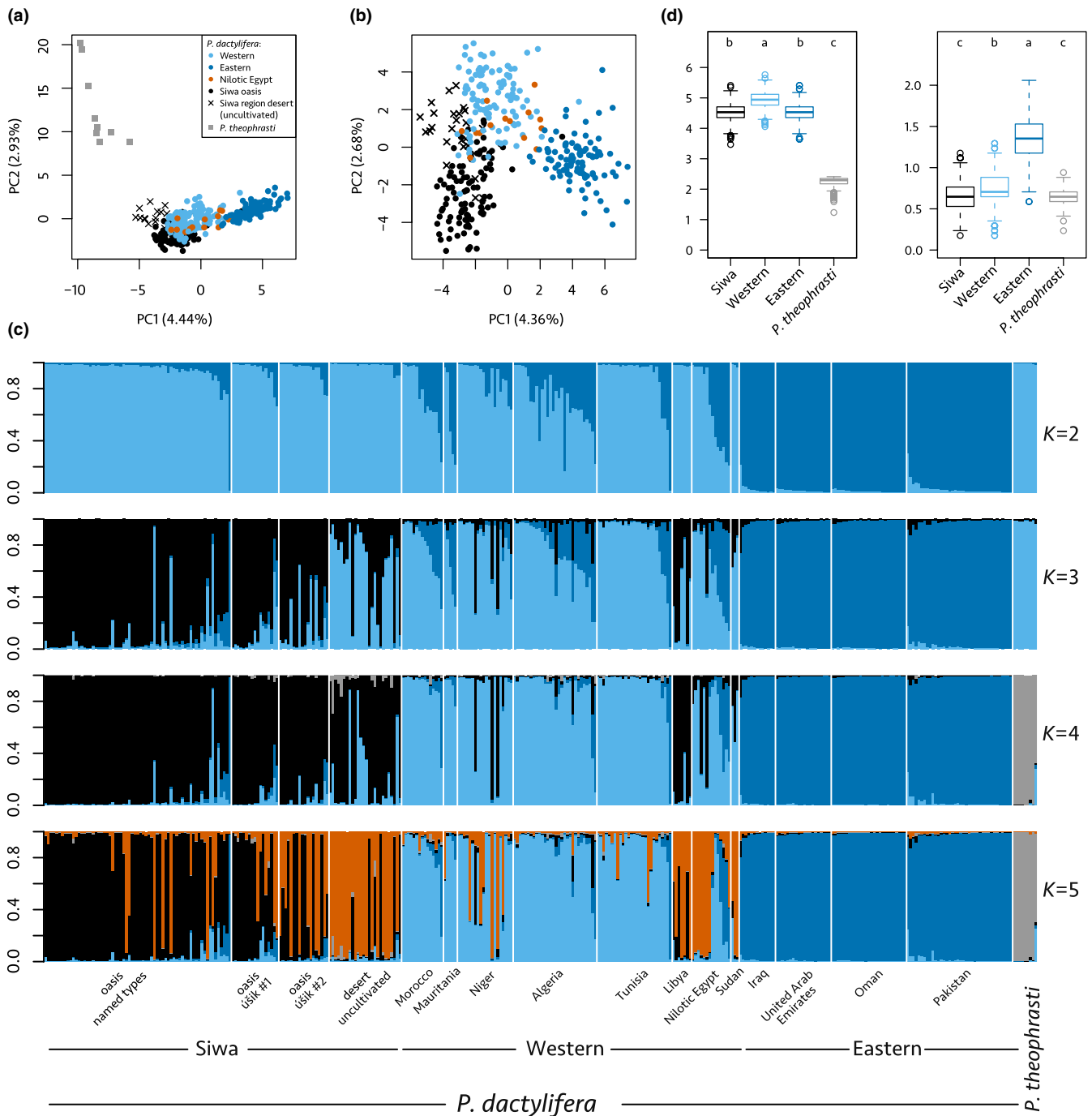
We compared 128 unique genotypes collected in Siwa Oasis and in the abandoned oases from the surrounding desert (Figure 1) with 219 date palms originating from North Africa and the Middle East (data from Moussouni et al., 2017; Zango et al., 2017; Zehdi-Azouzi et al., 2015), and nine newly genotyped *Phoenix theophrasti*. We investigated the population structure of these 358 unique genotypes with both principal component analyses and Bayesian clustering, calculated  $F_{ST}$  between various populations and subpopulations, and estimated diversity in various populations and subpopulations.

### 3.2.1 | Partitioning and extent of the diversity in *P. dactylifera* and *P. theophrasti*

The principal components (PCs) 1 and 2 both separate *Phoenix theophrasti* from *Phoenix dactylifera* accessions (Figure 4a), accounting for the large differentiation observed between these two species ( $F_{ST} = 0.32$ ; Table S4). Additional PCs do not provide notable results (Figure S2). On a side note, *Phoenix reclinata* being highly differentiated from other accessions (Figure S2), we excluded it from downstream analyses.

The PC 1 describes the geography among date palm accessions, with Middle Eastern and North African accessions (including Siwa) being stretched from left to right (Figure 4a,b). These two clusters are the first ones to emerge in the Bayesian clustering analysis (Figure 4c;  $K = 2$ ), and this partitioning is optimal according to the Evanno method (Figure S3). Eastern and western date palms are moderately differentiated, with  $F_{ST} = 0.088$  or  $0.084$ , whether Siwa accessions are included in the western cluster or not. The diversity among North African date palms is higher than among Middle Eastern date palms. Indeed, their levels of both allelic richness and private allelic richness, calculated with equal sample size, are significantly higher than that found in eastern date palms (Figure S4; Tukey's tests,  $p < .05$ ), and their expected heterozygosity reaches  $0.62$  while eastern populations display  $0.58$  (Table 4).

Interestingly, Siwa date palms overlap only slightly with the western and eastern clusters in the PCA (Figure 4a,b). Similarly, in the Bayesian clustering analysis, Siwa accessions form a distinct cluster from  $K = 3$  (Figure 4c). With an additional  $K$ , four clusters, corresponding mostly to *P. theophrasti*, eastern cultivars, western cultivars (excluding Siwa samples), and Siwa date palms, can be identified (Figure 4c; Table S5). Nevertheless, over-representation of related accessions may lead to spurious clustering in both PCA and structure analyses. In this study, we included a very large number of accessions from Siwa compared



**FIGURE 4** Worldwide structure of the diversity in date palms (*Phoenix dactylifera* L.) and *Phoenix theophrasti* Greuter. (a) Principal component analysis of 356 date palm and *P. theophrasti* accessions genotyped across 17 nuclear microsatellites. (b) Principal component analysis of 347 date palm accessions genotyped across 17 nuclear microsatellites. (c) Admixture proportion in 356 date palm and *Phoenix theophrasti* accessions with K equal 2 to 5. Each individual is represented by a vertical bar partitioned into colored segments representing the assignment coefficient or ancestry, that is, the estimated proportion of its genome derived from each cluster. (d) Allelic richness (left) and private allelic richness (right) calculated in four populations using the rarefaction method to equalize sample size across populations (haploid sample size = 18, bootstrap replicates = 1,000)

to the number of date palms from the Middle East and North Africa. Hence, we ran two other PCAs (with and without *P. theophrasti*) in which we randomly sampled ten accessions from Siwa Oasis and five accessions from the surrounding desert (Figure S5). We found that the accessions from the current oasis are found within the diversity of North African date palms. Nevertheless, when *P. theophrasti* is

included, Siwa accessions sampled in the oasis are found at the edge of the North African point cloud; when this species is not included, Siwa accessions can be differentiated from North African accessions with PC4. On the other side, most uncultivated date palms from the surrounding desert are not overlapping with North African accessions or any other palms (Figure S5).

**TABLE 4** Diversity estimates calculated in *Phoenix theophrasti* and various populations and subpopulations of *Phoenix dactylifera*

Populations	$H_o$	$H_s$	$F_{IS}$ (95% CI)
<i>P. dactylifera</i>	0.56	0.63	0.10 (0.084 – 0.14)
Siwa	0.56	0.60	0.060 (0.016 – 0.11)
Uncultivated desert	0.55	0.64	0.14 (0.072 – 0.23)
Oasis	0.56	0.57	0.015 (–0.028 – 0.066)
named type	0.55	0.56	0.0077 (–0.036 – 0.054)
úšik #1	0.59	0.60	0.025 (–0.048 – 0.078)
úšik #2	0.57	0.59	0.02 (–0.060 – 0.12)
Eastern	0.57	0.58	0.013 (–0.012 – 0.038)
Western (with Siwa)	0.56	0.62	0.095 (0.064 – 0.14)
Western (without Siwa)	0.56	0.61	0.081 (0.052 – 0.12)
<i>P. theophrasti</i>	0.17	0.42	0.59 (0.26 – 0.83)

Note: Abbreviations:  $H_o$ , observed heterozygosity;  $H_s$ , expected heterozygosity;  $F_{IS}$ , inbreeding coefficient.

Diversity estimates were calculated on western (or North African date palms) including or not Siwa accessions. Siwa population was split into two populations to oppose accessions sampled in the oasis (both named types and seedlings) with those sampled in abandoned oases in the surrounding desert. Further, accessions sampled in the oasis were further split into three populations: the named types and the seedlings collected in or nearby the gardens (úšik #1 and #2, respectively).

For the following sections, we thus consider those four populations, where Siwa and western populations are distinct. Siwa date palms appear more distinct from eastern ( $F_{ST} = 0.12$ , proportion of shared allele = 56.68%) than from western accessions ( $F_{ST} = 0.057$ , proportion of shared allele = 72.78%). In the Structure analysis, they share only, on average, 2.73% of their ancestry with the eastern cluster, while 7.66% can be traced to the western cluster (Figure 4c, Table S5). The only Siwa accession with mostly eastern ancestry is an úšik ezzuway (3363), collected in the heart of the old palm grove (in Jubba annēzi). Noteworthy, Siwa accessions are as diverse as eastern accessions in terms of allelic richness (calculated using equal sample size; Figure 4d) and even more diverse in terms of expected heterozygosity (Table 4).

On the one hand, the inferred ancestry constituting the Siwa cluster (Figure 4c, in black) is not only found in Siwa. Although it is almost absent in eastern cultivars (0.93% on average), it is substantial in western date palms (11.85% on average). It is prominent in the two Libyan accessions (Figure 4c): They share on average 89.86% of their ancestry with the Siwa cluster, the remainder being ancestry shared with the western cluster. On the other hand, we found that samples from Nilotic Egypt share ancestry with three clusters: mostly western (64.36%), but also eastern (18.36%) and Siwa clusters (16.81%). The evidence of shared ancestry with Siwa in the Nilotic Egypt cultivars is mostly driven by Hamra and Wardi cultivars (both from Aswan Governorate), having both > 70% of their ancestry shared with this cluster (a caravan road links Aswan to Siwa), while most other Nilotic

Egyptian date palms share < 1% ancestry with this cluster. The four Sudanese accessions share ancestry with both Siwa (28.30%) and western date palms (70.97%) and do not display ancestry traceable to the eastern cluster. We note that other North African countries share only a little of their ancestry with the Siwa cluster: Morocco (1.5%), Mauritania (1.26%), Algeria (3.56%), and Tunisia (1.42%), the exception being Niger, with an average of 15.55%, driven by three cultivars (Kila72, Ja5, Jahaske2, all three from Goure, Zinder region, and southwest Niger, a new region for phoeniculture) with 94.80%, 92.30%, and 73.40% of ancestry attributed to the Siwa cluster, respectively.

### 3.2.2 | Shared ancestry between *P. dactylifera* and *P. theophrasti*

Structure results show shared ancestry between *Phoenix dactylifera* and *Phoenix theophrasti* (Figure 4c). Indeed, eastern and western date palms share respectively 25.26% and 32.49% of alleles with the Cretan date palm. Structure results show shared ancestry between *P. theophrasti* and western date palms, but not with eastern date palms, except in Pakistan, probably due to recent introduction of African germplasm as reported before (Chaluvadi et al., 2019; Mathew et al., 2015). Further, western date palms appear less distinct from *P. theophrasti* ( $F_{ST} = 0.32$ ) than eastern do ( $F_{ST} = 0.40$ ). While we reported above that western accessions display more private alleles than eastern ones (Figure S4), when we include *P. theophrasti* in the analysis, they display on the opposite fewer private alleles than eastern ones (Figure 4d). Indeed, over the 227 alleles identified across the 17 loci, 17 (7.5%) are shared by *P. theophrasti* and North African date palms, but not by Middle Eastern date palms. This shared ancestry may be evidence of either common descent or gene flows or both.

Curiously, Siwa accessions share even more ancestry with *P. theophrasti* than the other date palms (cultivars) of North Africa, including Nilotic Egyptian ones. Structure results indicate they have up to 29.3% of their genetic makeup shared with *P. theophrasti* cluster (on average: 1.24%, Table S5). They are also closer, and especially the uncultivated ones, to this wild Cretan relative on PC1 than western date palms are (Figure 4c, from  $K = 4$ ) and share more alleles (35.58%) with this species than western do (32.49%).

Finally, Structure results show that there is also potential gene flow in the other direction, from *Phoenix dactylifera* to *P. theophrasti*. Indeed, two samples of the Cretan date palm display ancestry attributed to the date palm (Figure 4c).

### 3.2.3 | Partitioning and extent of the diversity within Siwa region

Siwa accessions can be further split into two distinct clusters, as seen previously (Figure 2b), and highlighted in both the Bayesian clustering (Figure 4c, from  $K = 5$ ) and the PCAs (Figure 4a,b).



These two clusters mostly fit the predefined Siwa populations, with a cluster comprising mostly accessions from the current oasis (Figure 4c, in black at  $K = 5$ ), while the second comprises mostly uncultivated desert date palms from the abandoned oases (in orange at  $K = 5$ ). Oasis and desert date palms are slightly differentiated ( $F_{ST} = 0.055$ ). Uncultivated desert date palms are less differentiated from western date palms ( $F_{ST} = 0.051$ ) than the cultivated Siwa date palms are ( $F_{ST} = 0.069$ ). The accessions from the oasis, in contrast, are less differentiated from eastern date palms (Table S4).

Remarkably, the desert uncultivated accessions appear even less differentiated from *P. theophrasti* ( $F_{ST} = 0.28$ ) than the accessions sampled in the oasis ( $F_{ST} = 0.35$ , Figure 4a,c, Table S4). Structure results indicate an average shared ancestry of 4.10% between desert uncultivated accessions and *P. theophrasti*, while solely 0.33% of current oasis accession ancestry can be traced to this species (Table S5, Structure at  $K = 4$ ). Further, they share more alleles with *P. theophrasti* than Siwa Oasis cultivated do (38.89% and 34.45%, respectively). Those abandoned desert palms have almost no trace of the eastern ancestry (0.63%), in contrast to the named types from the oasis (3.07%). They are highly diverse in terms of gene diversity (Table 4) and have, on average, more private alleles than the date palms from Siwa Oasis (Figure S4, Tukey's test,  $p < .05$ ).

The spontaneous seedlings found within and at the periphery of Siwa gardens (úšik #1 and #2, respectively) are found within the diversity of uncultivated and cultivated date palms of Siwa (Figures 3b, 3c; Figure S6). They appear intermediate between named types and uncultivated desert palms in terms of genetic makeup (Figure 4c; Table S4, Figures S7–S8 at  $K = 5$ ). Indeed, úšik #1 and #2 have 19.63% and 48.24% of ancestry shared with the Siwa desert uncultivated cluster, respectively (in orange at  $K = 5$ , Figure 4c and Figure S8), and 71.21% and 47.57% shared with the named types cluster, respectively (in black at  $K = 5$ , Figure 4c and Figure S8).

### 3.2.4 | Diversity at the chloroplast minisatellite in Siwa and worldwide

We identified five alleles at the chloroplastic locus *psbZ-trnFM* (Table S2). Two of them were restricted to *Phoenix reclinata* and one of them to *Phoenix theophrasti* (Figure S9). In *Phoenix dactylifera*, we identified the two previously reported alleles corresponding to the so-called occidental and oriental chlorotype (Gros-Balthazard et al., 2017; Zehdi-Azouzi et al., 2015). Eastern accessions mostly display the oriental chlorotype (94.90%, Figure S9). The occidental chlorotype is predominant in the western accessions (71.74%), although there is a significant presence of the oriental one, reflecting seed-mediated gene flows from east to west more prominent than in the other directions (as shown by Gros-Balthazard et al., 2017; Zehdi-Azouzi et al., 2015). Most North African countries display a higher proportion of occidental chlorotype (Figure S9), except Mauritania as reported before (Gros-Balthazard et al., 2017).

Two accessions of *Phoenix theophrasti* display the date palm western chlorotype (Figure S9). These accessions also have ancestry that is mostly attributed to date palm in the Structure analysis (Figure 4c). This could indicate potential gene flow from the date palm to this wild relative, as previously reported (Flowers et al., 2019).

In Siwa, occidental and oriental chlorotypes are found in almost equal proportion (47.65% and 52.34%, respectively). Among the named types of the oasis though, the oriental chlorotype is slightly predominant (59.70%, Figure S9), while we found the opposite pattern in the Nilotic Egyptian cultivars (42.86%). As for the uncultivated date palms from the abandoned oases of Siwa desert region, we found that they display mostly the occidental chlorotype (74.1%).

## 4 | DISCUSSION

In this paper, we studied date palms from Siwa using a combined molecular population genetic and ethnographic approach in order to (a) better understand folk categorization in conjunction with local agrobiodiversity and (b) infer the origins and the dynamic of the diversity found in this oasis and around, by comparing it to the worldwide date palm germplasm.

### 4.1 | On the folk categorization in Siwa and its implication for surveying date palm agrobiodiversity

#### 4.1.1 | Differentiating cultivar, ethnovariety, and local category

Our genetic analysis of intra-named type variability confirmed the existence of true-to-type cultivars as we found, for some named types such as the elite alkak, a 100% genetic identity across the 18 nuclear and chloroplastic loci. Nevertheless, we also pointed out the lack of genetic uniformity within other named types. Previous studies in date palms had already noted those intra-varietal variations, for instance, for the bint aisha type sampled in different localities in Egypt (El-Assar, Krueger, Devanand, & Chao, 2005, p. 606), for other named types in Libya (Racchi et al., 2013), and for samples from different countries or different regions (Chaluvadi et al., 2019; Khanamm, Sham, Bennetzen, & Aly, 2012). It has often been interpreted as resulting from somatic or somaclonal mutations (Abou Gabal et al., 2006; Devanand & Chao, 2003; Elhoumaizi, Devanand, Fang, & Chao, 2006; Gurevich, Lavi, & Cohen, 2005). Indeed, although offshoot propagation is alleged to be a true-to-type technique (Jain, 2012), somatic mutations can accumulate through generations of offshoot propagation. This has also been demonstrated in other crops, such as grape (Moncada, Pelsy, Merdinoglu, & Hinrichsen, 2007), cherry (Jarni, Jakše, & Brus, 2014), or *bractea-tus* pineapple (Chen et al., 2019). Further, when propagated through tissue culture, cultivars may accumulate somaclonal variation (El Hadrami, Daayf, Elshibli, Lain, & El Hadrami, 2011; e.g., Medjool cultivar, Elhoumaizi et al., 2006).

Here, we acknowledge that somatic mutations can indeed explain some variations. As a matter of fact, we interpreted slight genetic dissimilarity among *šaēidi* accessions as the result of somatic mutations, if not genotyping errors. Nevertheless, we also substantiate that the high genetic dissimilarity within some named types cannot solely reflect the existence of somatic mutations. Instead, it is the result of a cultivation practice that we proposed before, that is, the incorporation of new clonal lines of seedlings under an existing name (Battesti et al., 2018). In this study, we further substantiate this statement, using a larger number of named types and, for each, more samples. We also highlight that a significant number of named types, beyond the very inclusive *úšik* and *males*, are neither true-to-type cultivars nor even ethnovarieties, but of the “local category” order. This peculiar cultivation practice we observe in Siwa could also exist in other palm groves and explain previously described cases of intra-varietal genetic variation, for instance, for *Medjool*/*majhūl*, the famous Moroccan variety, which is not a “genetically uniform” clone (Elhoumaizi et al., 2006, p. 403).

The prominence and the consequences of sexual reproduction in clonal crops have often been underestimated (McKey, Elias, Pujol, & Duputié, 2012). A recent study demonstrated the occurrence of numerous recombination events in the history of *bracteatus* pineapple cultivars (Chen et al., 2019). Our study on date palms in Siwa constitutes an additional example where cultivation techniques do not silence sexual reproduction, but rather integrate it.

#### 4.1.2 | Determining the number of named types

Surveying the precise number of named types that refer to distinct local date palms in an oasis is an already complex operation, as in Siwa (Battesti, 2013) where given lists of named types can refer to varying degrees of inclusiveness (in no hierarchical order of exclusive taxa, unlike what Brent Berlin's ethnobiological theory of taxonomic categories implies: Berlin, Breedlove, & Raven, 1973; Berlin, Breedlove, & Raven, 1974). However, in a system with cultivation methods based on massive vegetative reproduction, the identification of the accurate number of named types should have been sufficient to correctly estimate the number of genotypes. Until now, therefore, and at best, the overestimation of the agrobiodiversity of the date palm at the regional level had been considered, taking into account the phenomenon of synonymy (a cultivar takes another name by changing oasis), somewhat offset by homonymy (the same name is used in different oases to designate a different cultivar) (Battesti, 2013; Battesti et al., 2018). In a single oasis, Siwa, one of the difficulties of the field survey resides in massive local synonymy, including for plant names and in particular for date palm varieties, as already mentioned (Battesti, 2013; Battesti et al., 2018). The first hypothesis addressing such a synonymy is that a landlocked Berber-speaking community should promote its export products by adding Arabic trade names. Another explanation is the co-presence on the same territory of Arabic-speaking minorities (sedentary Bedouins, especially *Awlad 'Alī*, Battesti et al., 2018) who use these Arabic

names; for example, rather than the names *tasutet*/*alkak*/*úšik*, they will systematically use (or even only know) the names *šaēidi*/*frehī*/*azzawī* to denote the same palms. Besides, we observed that local practices of categorization and integration of seedlings can lead, as explained above, to a massive phenomenon of underestimation of agrobiodiversity for an uninitiated external observer and all assessments have stumbled over this obstacle (even the recent *Atlas of date palm in Egypt*, El-Sharabasy & Rizk, 2019).

In Siwa, our ethnobotanical survey revealed the existence of 15 to 20 named types (Battesti, 2013; Battesti et al., 2018). The exact number is difficult to pinpoint as the field survey reveals that farmers offer local, more or less shared, quality distinctions, even for dates that we depicted as true-to-type cultivars (*alkak* and *taṭṭagt*, for example). Therein, a higher *alkak* is called “*alkak n amles*,” meaning smooth or wrinkle-free *alkak* dates, and a lower *alkak* with smaller dates, and three times cheaper than the upper one, is called “*alkak nifuyen*.”

Fifteen to 20 named types in Siwa is not a lot compared to the number of named types described in other oases. Assessing the agrobiodiversity of date palm is a difficult exercise and carried out using noncomparable competing methodologies (Jaradat, 2016). As a result, it is difficult to establish the terms of comparison of agrobiodiversity between oases. We can provide some comparative data. For instance, in the Jerid region of Tunisia (about twice the area of old palm groves compared to Siwa), there is a collective collection of more than 220 varieties (Battesti, 2015; Rhouma, 2005, 1994); some of them may very well be ethnovarieties or local categories, but this hypothesis still has to be checked in situ. Smaller oases have about the equivalent number of named types: 18 in Sokna (al-Jufra, Libya) for a cultivated area half as small as Siwa (Racchi et al., 2013), 18 in Kidal (Northern Mali) but for only 4,000 date palms (Babahani, Togo, & Hannachi, 2012), and 22 in el-Guettar (Tunisia) for only 3,000 date palms (Ben Salah, 2012). Hence, we could believe that the agrobiodiversity is relatively low in Siwa, if considering a conventional system where named types are thought to be genotypes. Nevertheless, with this combined ethnobotanic/genetic approach, we showed that this assumption is wrong, and that, despite a low number of named types compared to other oases, the number of genotypes in Siwa is high and so is the overall diversity.

#### 4.1.3 | Etic versus. emic categorizations

Although we confirm the validity of the notions “cultivar,” “ethnovariety,” and “local category,” we also underline their limitations: They are hardly relevant for the farmers who perform these practices. Two main difficulties arise when assessing local date palm agrobiodiversity. On the one hand, as we demonstrate here, named types are not necessarily genetically uniform contrarily to previous expectations. On the other hand, this explains the first point, the local ways of conceiving this living material and its qualities, of categorizing it (by form) does not quite match with biologists' ways of

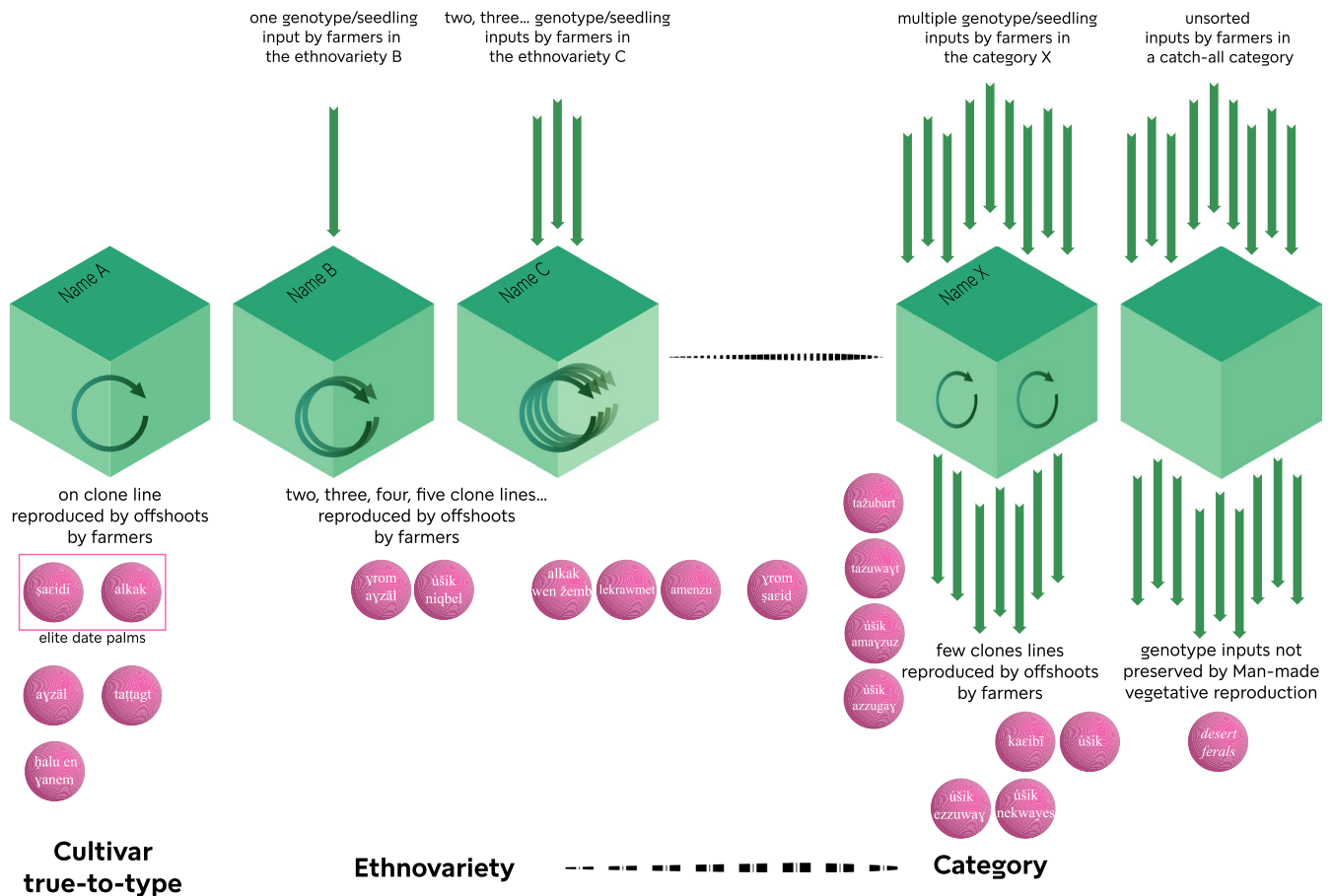
conceiving it and categorizing it (by genotype) (Battesti, 2013). This difficulty is a classic conflict for anthropologists, also known as a disparity of *etic* versus *emic* categorizations (Olivier de Sardan, 1998). Deeply characteristic of human societies, categorization processes are also at the heart of both mundane and scientific thoughts and practices. Naturally, our aim is not to use our paradigm (genetics) to try to evaluate the paradigm of indigenous local knowledge (Roué & Nakashima, 2018), but to translate the latter: For social sciences, there is no such thing as one science but several incommensurable sciences, and multiple modes of existence coexist responding to various forms of veridiction (Latour, 2013). In the *emic* version (the local point of view), the differentiation between cultivars/ethnovarieties (which farmers assimilate) and local categories is clearly thought out. However, it is not so clearly expressed: Farmers can list all named types at the same level even if they do not refer to the same object classes. It is worth bearing in mind that for local farmers, reproducing by offshoot is the rule. Naming/identifying a seedling date palm after a known cultivar (ethnovariety process) is a possible (and appropriate) practice but of an exceptional occurrence on a human life scale. Hence, the difficulty for local farmers to know whether, for instance, all the *ušik niqbel* are cultivars (which they tend to present as such) or ethnovarieties. Although they do not remember seeing an *ušik n gubel* coming from a seed, this has apparently been the case, perhaps several decades or generations ago (a date palm easily outlives a human being). The question is of little relevance to them, since the palm trees in question behave and have/produce the same form. But, the other process, categorizing a date palm from a seedling, with reddish/dark dates for instance, as having a valuable production and naming/qualifying it, as *ušik ezzuway* for instance, is a more common experience (local category process). Hence, for local categories, we have to reverse the point of view. For example, some date palms do not bear reddish/dark dates because they belong to the variety *ušik ezzuway* or *zuway*, but the fact that their dates are reddish/dark qualifies them as *ušik ezzuway* or *zuway*. As scientists, embedded in our *etic* point of view, the difference between ethnovarieties and local categories may seem thin as both result in genetically unrelated individuals (with a few clones for the former if sampling is sufficient), hence the necessity to consider cultivation practices and local knowledge. Indeed, to distinguish between what can be objectified as an ethnovariety or a local category in our samples is not easy. A clear choice for cultivar true-to-type can be made. Another clear choice is possible for an ethnovariety in the case of *x* clones plus one or two outsiders (possibly also clones, we did not face the case). When all individuals are different from each other, there is a greater chance of having a local category (to be confirmed by the way farmers talk about it), especially when there is a minimum local consensus. Assessing this consensus is obviously difficult, since no one in the community will use our *etic* way of thinking about the diversity by cultivar/ethnovariety/local category. Informants just state it as “this date palm, it is an xxxx” [a named type]. Even more complex is the case of the rare ethnovariety, which is little known, but real; so few individuals are sampled that there may not be any clones in the sample.

It is therefore preferable to consider the distance between the ethnovariety and the local category as a continuum (Figure 5). Indeed, among the 18 named types we analyzed here, we found a gradient of intra-cultivar relatedness: Some named types corresponded actually to unique genotypes while some did not, as evidenced in our previous study (Battesti et al., 2018). It is worth noting that an expanded sampling could show that what we believe today to be a true-to-type cultivar is in fact an ethnovariety, by discovering a new line of clone under the same name. Further, it could allow to identify lines of clones for the inconclusive *alkak wen žemb*, *lekrawmet*, *amenzu*, or *γrom šaid* to confirm them as ethnovarieties. To formally confirm that a given name is either a true-to-type cultivar, an ethnovariety, or a local category, it would require that the 200,000 to 250,000 date palms of Siwa to be all genotyped. This is obviously not conceivable.

## 4.2 | Usefulness of the folk categorization

What are the explanations for such a complex local system for categorizing date palms in Siwa? They might be both cognitive and agricultural. First, all “classifications are functionally linked to the effective storage, retrieval and communication of large quantities of information relating to the animal and plant worlds” (Meilleur, 1987, p. 9–10) and must be easily mobilized to guide action and communicate (knowledge, experience, etc.). This system must therefore be shared. The utilitarian nature of folk biological classifications had already been discussed by Eugene Hunn (1982). Date palm cultivation in Siwa is largely dominated by a few “elite” types (probably for centuries, an integration into the Saharan trading network, Battesti, 2018, 2013). Here, we found that they are apparently true-to-type cultivars, despite their prevalence and therefore the mechanical possibility of becoming an ethnovariety. The ethnovariety and local category system make it possible to “put in order” the profusion of all the other date palms, less commercially valued, while not multiplying the denominations for the same characteristics.

Secondly, a peculiar system allows for a peculiar action on the world, in this case, a fairly flexible management of agrobiodiversity. McKey et al. (2010) consider that cultivation of clonally propagated plants represents a singular system that has to date been largely neglected and that a key component of strategies for preserving the adaptive potential of clonal crops is the maintenance of mixed clonal/sexual systems. Their model is the cassava (*Manihot esculenta* Crantz), a eudicot grown as an annual plant in tropical and subtropical regions. To date, cassava is the only clonal crop for which in-depth information exists on how mixed clonal/sexual systems work (ibid.). We propose here another crop model, quite different as the date palm is a perennial plant that lives decades or even over a century (Chao & Krueger, 2007). This critically impacts farming practices, especially those inherent to propagation (by seed or offshoot, sexual or clonal). However, it is very likely that maintenance of a mixed clonal/sexual system is a local key strategy for managing and preserving an agrobiodiversity of



**FIGURE 5** Converging local categorization of date palms in Siwa (*emic*) and genetic categorizations (*etic*) with the distribution of the named types identified by Isiwan. For genetic analyses, there is a continuum between ethnovariety and local category, while for local farmers, the rationale is quite different (genotype vs. phenotype approaches). For farmers, ethnovarieties imply the idea of identity, when local categories refer to the idea of a qualification (to qualify a date palm of ...). Nevertheless, cultivars, ethnovarieties, and local categories are listed by Isiwan as if they were equivalent

date palms in Siwa. And this is made possible for Siwa farmers because they developed a classification system that enables them to do so.

How does it usually work—or how do we think it works? Usually, farmers have important stock of cultivars (true-to-type) reproduced vegetatively, and a possible gene pool of seedlings, with their own possible use (fodder, handicraft, etc.)—named depending on places *khalt*, *sheken*, *degla* (Battesti, 2005), *rtob* (Ben Salah, 2012), *sa'ir* (Peyron, 2000), *sayer* (Naseef, 1995), *qush* (Popenoe, 1913), *dabino* (Zango et al., 2016), etc., or *ūšik* in Siwa. From this stock is sometimes drawn on occasion a new seedling then socially recognized by the granting of a new name and then vegetatively reproduced: It enhances the initial collection.

How mixed clonal/sexual systems work in Siwa? We have demonstrated that, there, the local categorization system including ethnovarieties and local categories—part of the palm domestication at ethnographic scale—allows an even more flexible management of agrobiodiversity and gene flow to the cultivated pool, as any new seedling integrated into the procession of cultivated date palms is likely to be classified within a pre-existing named type, although not sharing the same genetic heritage in common (genotype), but sharing, from a local point of view, the same form or the same characteristics (phenotype).

### 4.3 | A unique and high genetic diversity in Siwa date palms

Siwa date palms form a partially distinct cluster from western accessions, rather than a subset of the diversity found in North Africa as it might have been expected. Some alleles are found uniquely in the region. The level of allelic richness in Siwa (a tiny region) is strikingly as high as that found in all eastern accessions, originating from a large geographic area from the Near East through the Arabian Peninsula and as far as Pakistan. We note that Libyan cultivars also belong to what we can call a third genetic cluster of modern date palm germplasm. This uniqueness of date palms from the Libyan Desert has never been reported before, although a previous study mentioned that samples from Libya and Tunisia may hold a higher diversity (Mathew et al., 2015). In fact, very few studies have compared accessions from Siwa with other date palms. Gros-Balthazard et al. (2017) mentioned that the cultivar “siwi” from Siwa (here called *sa'idi*) did not cluster with the Middle Eastern accessions nor with the North African accessions, but it was interpreted as a hybrid origin. Identically, it does not cluster with other Egyptian accessions in a study based on AFLP (El-Assar et al., 2005). A seed morphometrics

analysis shows that it clusters with date palms of mixed origins (Terral et al., 2012). Nevertheless, both a limited sampling in Siwa and an incomprehension/ignorance of the cultivation practice had hampered those previous researches to pinpoint the singularity and high diversity of Siwa date palms. We note that although there are existing studies based on microsatellite data that provide diversity estimates for date palm populations from various regions (i.e., Chaluvadi et al., 2019; Zehdi-Azouzi et al., 2015), the use of distinct loci and/or different sample size prevent comparisons between our results and that of these studies.

Three nonmutually excluding hypotheses could explain the high genetic diversity observed in Siwa date palms. First, a highly diverse but yet unidentified source of diversity may have contributed to date palms in the Libyan Desert, thus making them both highly diverse and unique, compared to other regions (see Section 4.4). Second, the cultivation practices that we here describe may have maintained/promoted this high diversity. Indeed, a sole clonal propagation leads to a loss of diversity (McKey et al., 2010). In Siwa, we found a complex system where what we thought were cultivars (clones) are in fact ethnovarieties or local categories, which in turn account for a higher diversity than expected at first sight. It is thus possible that in Siwa, more than in other oases, sexual reproduction being more common, there is a more limited loss of diversity through time. In addition, the position of Siwa, at the crossroads between western Africa and the Middle East, promotes the creation of novel hybrids, and a particularly high outcrossing rate, where many individuals contributing to the next generations may have allowed the retention of this peculiar diversity. These hypothesis remains to be tested by ethnographic survey in other oases in order to test whether sexual reproduction is more prevalent in Siwa than elsewhere. This leads us to the third hypothesis: The high diversity found in Siwa could reflect, rather than a reality, a sampling bias. Indeed, our sampling, considering practices and knowledge of farmers, may have led us to sample an appropriate representation of the existing diversity of Siwa. On the opposite, the non-Siwa date palms included here have been sampled without such an ethnobotanical survey and may in turn only be a poor representation of the actual diversity. Hence, if such a sampling methodology were applied everywhere, we may very well discover more diversity elsewhere too.

We identified that date palms in the current oasis of Siwa and in the ancient abandoned palm groves in the surrounding desert constitute two subpopulations. Changes in allele frequencies in the uncultivated palms may have been driven by a relaxation of human-induced selective constraints, natural selection, or genetic drift, following the abandonment of these groves. Additionally, while we identified potential gene flow from eastern accessions at both chloroplastic and nuclear level in the date palms from the current oasis, accessions from the abandoned palm groves may have received less diversity from this population. Uncultivated palms from the desert share more ancestry with *P. theophrasti* than both western and Siwa Oasis date palms. Locally, these desert date palms are considered as *ušik* but are also specifically called *igizzā* (sing. *agzzu*).

Our ethnographic survey revealed that they do not seem to be used by the farmers as a reservoir of diversity. Some of these abandoned oases could nevertheless constitute casual date harvest sites, and hence, seeds from the harvested fruits could potentially end up as *ušik* in the cultivated palm grove. This could explain the intermediate genetic profiles of *ušik* #1 and #2: Their diversity seems in between that found in the current oasis and the abandoned oases.

Siwa date palm origins are unknown, but their presence in Siwa dates back at least to the 5th century (mentioned by Hellanicus of Mytilene, see above); the oasis being then an independent state related to the Libyan world, well established as an essential trade and religious hub with Libyan, Egyptian, and Hellenistic influences (Kuhlmann, 1999). Declining or abandoned at the end of the first millennium, the oasis was probably recolonized in the 11th or 12th century (a priori by Amazighs from Libya, then Arabs, see Battesti, 2013), but we do not know if this was done by conserving the original stock of date palm trees or by introducing new plants/cultivars, or both. In the present study, we found that the date palm population of Siwa is unique. As deep as we can go back in Siwa's history (the realm of "two deserts"—eastern and western deserts, from a Siwa-centered point of view—Kuhlmann, 2013), opportunities for a melting pot of genetic richness of date palm were there. Nevertheless, connections are not enough. Assessment would have to be done on the dispersion flows of the domestic date palm between oases, and on the entry of genes into Siwa stock, intentionally—by offshoots?—or unintentionally—by seed? The "stepping stones" (connectivity from oasis to oasis) of the landscape ecology (Burel & Baudry, 1999) and the "functional connectivity" concept (Battesti, 2018) seem appropriate. The genetic makeup in Siwa, like that of other North African populations, may originate from a mix of Middle Eastern date palms and *P. theophrasti*; an unknown ancestral gene pool may also be involved (see below).

#### 4.4 | Revising the history of date palms origins and diffusion

Using 17 nuclear microsatellites and one chloroplastic minisatellite, we detected the previously described differentiation between date palms from North Africa (the so-called western population) and those from the Middle East and Pakistan (the so-called eastern population). Our results also corroborate the existence of gene flows between these two clusters, mostly from east to west (Gros-Balthazard et al., 2017; Hazzouri et al., 2015; Zehdi-Azouzi et al., 2015).

##### 4.4.1 | The contribution of *P. theophrasti* to North African date palms: insights from Siwa date palms

Date palms were domesticated at least in the Persian Gulf region, followed by diffusion to North Africa (for review, Gros-Balthazard et al., 2018). A recent study, based on whole-genome resequencing data, indicated that North African date palms have



been introgressed by the Aegean *Phoenix theophrasti*, and today, about 5%–18% of their genome originates from this wild relative or a *P. theophrasti*-like population (Flowers et al., 2019). Here, we also identified potential evidence of this introgression. Indeed, the Bayesian clustering identified shared alleles between these two populations, and the differentiation between western date palms and *P. theophrasti* is reduced compared to that between eastern date palm and this wild relative. We, however, note that none of the 121 North African and 156 Siwa date palms display the *P. theophrasti* unique chlorotype (five repeats of the dodecanucleotide minisatellite, Pintaud et al., 2013). This corroborates previous results showing that plastid and mitochondrial haplotypes characterizing *P. theophrasti* are absent in the 25 sequenced North African date palms (Flowers et al., 2019). This has been interpreted as an asymmetry in the direction of the interspecific cross, where gene flows from *P. theophrasti* were pollen-mediated (ibid.), but the possibility of a plastome–genome incompatibility (Greiner, Sobanski, & Bock, 2015) cannot be ruled out. Going beyond these two occidental and oriental chlorotypes, Mohamoud et al. (2019) identified four chlorotypes, using genome-wide genotyping data. Our microsatellite data do not allow us to know which of these four chlorotypes are displayed by our samples, but this represents an exciting direction for future genomic studies.

Although a major event in the diffusion and diversification of cultivated date palms to North Africa, the introgressive event by *P. theophrasti* remains puzzling, especially in terms of localization (Flowers et al., 2019). Indeed, today, this species is not distributed in North Africa, but is found in the Aegean region, mostly in Crete (Barrow, 1998), but also in coastal Turkey (for review, Boydak, 2019). One hypothesis explaining the introgression by *P. theophrasti* is that it had a much wider distribution in the past, encompassing North Africa and/or the Levantine region (Flowers et al., 2019). In the literature, authors consider it is a tertiary relict species (Vardareli, Doğaroğlu, Doğan, Taşkın, & Göçmen Taşkın, 2019). Here, we indeed found a high inbreeding coefficient that could reflect a severe population bottleneck due to habitat loss. Nevertheless, we lack positive evidence for the presence of the Cretan date palm outside of its current distribution area and this hypothesis remains unconfirmed.

Another hypothesis explaining this introgression is historical. In this study, we found that date palms from the region of Siwa share on average more alleles with *P. theophrasti* than other western varieties. This could indicate that North African date palms were introgressed by *P. theophrasti* or a *P. theophrasti*-like population in this peculiar region. As a matter of fact, tight links between Siwa and Crete did exist. Archeological and textual evidences prove that the Minoan culture that had developed in Crete interacted with the mainland Eastern Mediterranean, including New Kingdom Egypt, most intensively in the second half of the second millennium BCE (Bietak, 2005; Bietak, Maritanos, Palivou, & Brysbaert, 2007). At least from the 5th century BCE onwards, Siwa was a large hub, caught between influences exerted by Pharaonic and then Graeco-Roman Egypt on the one hand, and the Greek colonies of Cyrenaica on the other hand (Rieger, 2017, p. 52; Struffolino, 2012). Many Greek delegations and

merchants visited Siwa to present their offerings to the patron god (Colin, 1997; Kuhlmann, 2013), and Greek or even Cretan workers employed in the construction of the temples have left engraved inscriptions of their passage (Aldumairy, 2005; Kuhlmann, 2013). Those persistent connections could have set the opportunity of plant material exchanges, including of *Phoenix*.

We note that date palms from the abandoned oases of Siwa show an even closer affinity to *P. theophrasti* than do the palms from the current oasis. One speculative hypothesis for this pattern is that those palm groves were abandoned shortly after introgression by the Cretan date palm, and their genetic makeup may have evolved independently from other date palm populations, while date palms in the current oasis could have had their *P. theophrasti* ancestry diluted by gene flows from other populations, especially the eastern ones.

#### 4.4.2 | The Siwa region at the diffusion crossroads of the domestic date palm?

Siwa was in early contact with Egypt and the Libyan Sahara, a position that was essential for the diffusion of the date palm. In Egypt, unequivocal evidence for the economic and cultural importance of dates and for the local cultivation of the date palm for its fruits begins with the New Kingdom in the mid-2nd millennium BCE (Tengberg & Newton, 2016). However, there is now earlier evidence for the presence of date palm products in Egypt, but whether the dates were produced locally, in the Nile delta, or imported remains unknown. For instance, administrative texts document that two regions (nomes) located on the Mediterranean coast provided dates, of unknown origin, to workers under the reign of Khufu (around 2,600 BCE, Tallet, 2017), while a very small amount of date palm remains were found at Giza (2,700–2,100 BCE) (Malleon, 2016; Malleon & Miracle, 2018) and at the site where the texts were found (Wadi al-Jarf, a Red Sea harbor site, Newton, unpublished data).

The extension of the route leading west from Siwa, via Jaghbūb, Awjila, to Fezzan, and plausibly onwards, appears to date to the late second or early first millennium BCE (Mattingly, 2017, p. 8). In the Fezzan (Libya), the Garamantes developed into a Central Saharan state (ibid. p. 15) and were involved in separate bilateral trading arrangements with neighboring states and peoples to the north and south (ibid. p. 19). Late pastoral period sites in the Fezzan do suggest some contact with agriculturalists: Date stones were recovered dating to the end of the second millennium BCE, but could have been “exotic” imports from Egypt. From the beginning of the first millennium BCE, the occupants of Zinkekra were cultivating a range of crops characteristic of Near Eastern, Mediterranean, and Egyptian farmers, based on emmer wheat, barley, grape, fig, and date (Pelling, 2005, p. 401). These are some of the earliest evidence for oasis agriculture in North Africa (Van der Veen, & Westley, 2010). Siwa, the realm of the “two deserts” (Kuhlmann, 2013), could have been in the first millennium an important, even unavoidable node in the date palm exchange network between Egypt and Libya.

#### 4.4.3 | A missing contributor to the modern date palm genetic makeup?

North African date palms display a higher diversity than Middle Eastern ones (Gros-Balthazard et al., 2017; Hazzouri et al., 2015). Flowers et al. (2019) showed that the more the genomic regions are introgressed by *P. theophrasti* or a *P. theophrasti*-like population, the higher their diversity is, concluding that the excess of diversity found in North African date palms can be explained by the introgressive event(s) by this wild relative. Our results corroborate these findings, as we found that the fraction of alleles uniquely found in North Africa is reduced when *P. theophrasti* is included in the analysis (Figure S4).

Nevertheless, both Flowers et al. (2019) and our data indicate that there could still be an unknown population that contributed to the modern date palm germplasm. Indeed, in genomic regions with no or limited *P. theophrasti* introgression, the diversity in North African date palms remains higher than that in Middle Eastern date palms (Flowers et al., 2019). Further, of the two main chlorotypes described in date palms (Pintaud et al., 2013), we ignore the origins of the so-called occidental one, prevalent in western date palms. Indeed, the few relictual wild date palm populations (in Oman) bear the so-called oriental chlorotype (Gros-Balthazard et al., 2017), and *P. theophrasti* an even more divergent type (Pintaud et al., 2013). Here, we could also identify diversity that is unique to the North African date palms, especially in Siwa. The origins of this unique genetic makeup are unknown. It could, like other North African date palms, derive from a mix between Middle Eastern date palms and *P. theophrasti*. Indeed, demographic events, such as population bottlenecks, or selection can lead to drastic changes in allelic frequencies, while mutations may introduce new alleles. Another hypothesis, reinforced by results obtained at the chloroplastic loci, is a contribution from not only Middle Eastern date palms and *P. theophrasti*, but also from a yet unidentified North African *Phoenix* population.

So far, there is no evidence of date palms or other *Phoenix* populations in North Africa before the establishment of oasis agriculture. Predomestication *Phoenix* remains attributed to *P. dactylifera* have only been found in the Levant and in Iraq (Henry, Brooks, & Piperno, 2011; Lipshitz & Nadel, 1997; Solecki & Leroi-Gourhan, 1961). In contemporary times, wild date palm populations are exclusively known in the mountainous regions of Oman (Gros-Balthazard et al., 2017). The natural distribution of *P. dactylifera* (before its domestication and diffusion) is unknown (Barrow, 1998). It likely covered at least the Middle East and may have shrunk, in connection to climate change (Collins et al., 2017). In Egypt, it is possible that the date palm was growing wild or that it was grown for ornamental purposes (more precisely, not mainly for its fruit), perhaps from the 4th until the mid-2nd millennium BCE.

Siwa Oasis and the Libyan region in general could be a region of prime importance for the understanding of date palm origins in North Africa and the history of gene flows with the Cretan date palm *Phoenix theophrasti*. More precisely, not only date palms from

the current oasis could shed light on the history of this species, but also the uncultivated palms from the abandoned oases. So far, research has focused mainly on cultivated germplasm, neglecting uncultivated date palms, even if local oasis communities have a use for them. Here, we demonstrate that they can be of interest as, not only they can inform on past history, but they may represent untapped reservoir of diversity for future breeding programs.

#### 4.5 | Articulating the scales of ethnography and domestication over the long term

This study demonstrates that the combination of observation angles and methodologies on the cultivation of a crop, in particular perennial with a clonal/sexual reproduction, is essential. Ethnobotany hypothesized agricultural and classification practices with considerable effects on agrobiodiversity, made possible by local ways of categorizing living organisms (that result in the de facto presence of cultivars, ethnovarieties, and local categories of date palms). Nevertheless, their occurrences are rare events (on the scale of one generation of farmers) and only molecular genetic analyses could verify those hypotheses. The other way around, geneticists assessing agrobiodiversity de facto consider each named type to be a cultivar, which only a long/careful ethnography could hypothesize to be, at least partially, untrue. Without the ethnobotanists, geneticists, by collecting a single sample per named type as they consider that they are cultivars, miss a huge amount of diversity and ways to understand it. The crucial step of sampling strategy cannot also be effectively implemented without a well-established understanding of local ways of categorizing (naming) the material being studied through long-term ethnobotanical fieldwork.

In this study of the oasis of Siwa, we thus demonstrated that agrobiodiversity can be studied neither by a single genetic approach, nor by a single ethnographic approach. The combination of the two observation scales is necessary to uncover the phenomena and data necessary to test hypotheses. Indeed, these dialogues and close collaborations have resulted here in the discovery of hidden date palm diversity. Applying such an integrated approach to other oases may bring to light concealed date palm diversity elsewhere. Date palm diversity, whether wild, cultivated, or abandoned, thus still remains largely uncovered. Yet, having the right picture of the existing date palm diversity is crucial to the understanding of its origins and critical when it comes to germplasm conservation.

Finally, this interdisciplinary framework could be useful not only to date palm scientists, but also to scholars studying agrobiodiversity and its evolution in many different plant contexts.

#### ACKNOWLEDGEMENTS

The present study was funded by BioDivMeX program (Working-Group Insularities)/MISTRALS, Mosaïque program (GDR 3355 INEE CNRS), and a grant from the New York University Abu Dhabi Research Institute on Biodiversity Genomics to Michael Purugganan.

The authors are very grateful to the countless farmers of Siwa Oasis for having taken on such countless questions and for their unfailing hospitality.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

M.G.-B. and V.B. designed the research and wrote the paper. V.B. performed the ethnographic survey and studied the folk botanical categorization. M.G.-B. performed genetic data analysis and, with V.B., interpreted the results. S.I., L.P., F.A., O.Z., S.Z., S.M., S.A.N., C.N., and J.-F.T. contributed to data analysis. V.B., M.G.-B., S.I., and C.N. collected the material.

## DATA AVAILABILITY STATEMENT

Genotyping data generated for this study are available in Table S2.

## ORCID

Muriel Gros-Balthazard  <https://orcid.org/0000-0002-2587-3946>  
 Vincent Battesti  <https://orcid.org/0000-0002-5793-1098>  
 Laure Paradis  <https://orcid.org/0000-0002-2159-291X>  
 Frédérique Aberlenc  <https://orcid.org/0000-0003-1866-6807>  
 Oumarou Zango  <https://orcid.org/0000-0003-1900-6827>  
 Salwa Zehdi-Azouzi  <https://orcid.org/0000-0003-1710-0617>  
 Souhila Moussouni  <https://orcid.org/0000-0002-9989-8468>  
 Summar Abbas Naqvi  <https://orcid.org/0000-0001-8186-9055>  
 Claire Newton  <https://orcid.org/0000-0003-3791-3906>  
 Jean-Frédéric Terral  <https://orcid.org/0000-0003-1921-2161>

## REFERENCES

- Abd El-Azeem, R. M., Hashem, M. H., & Hemeida, A. A. (2011). Identification and Genetic Similarity Analysis of Date Palm (*Phoenix dactylifera* L.) Collected from Different Regions in Siwa Oasis using Morphologically Traits and Molecular Markers *Egyptian Journal of Genetics and Cytology*, 40, 281–300. <https://doi.org/10.21608/ejgc.2011.10793>
- Aberlenc-Bertossi, F., Castillo, K., Tranchant-Dubreuil, C., Chérif, E., Ballardini, M., Abdoukader, S., ... Pintaud, J.-C. (2014). In Silico Mining of Microsatellites in Coding Sequences of the Date Palm (Arecaceae) Genome, Characterization, and Transferability. *Applications in Plant Sciences*, 2, 1,300,058. <https://doi.org/10.3732/apps.1300058>
- Abou Gabal, A. A., Abedel Aziz, A. A., Hardash, M. M., & El-Wakil, H. F. (2006). Genetic diversity among seven date palm landraces in Siwa oasis. *Egyptian Journal of Genetics and Cytology*, 35, 117–128.
- Adamack, A. T., & Gruber, B. (2014). PopGenReport: Simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution*, 5, 384–387. <https://doi.org/10.1111/2041-210X.12158>
- Aldumairy, A. A. (2005). *Siwa Past and Present*. Alexandria, Egypt: Yasso.
- Babahani, S., Togo, A., & Hannachi, S. (2012). Étude sur le patrimoine phoenicole de Kidal au nord du Mali. *Fruits*, 67, 77–86. <https://doi.org/10.1051/fruits/2011071>
- Ballardini, M., Mercuri, A., Littardi, C., Abbas, S., Couderc, M., Ludeña, B., & Pintaud, J.-C. (2013). The chloroplast DNA locus psbZ-trnfM as a potential barcode marker in *Phoenix* L. (Arecaceae). *ZooKeys*, 365, 71–82. <https://doi.org/10.3897/zookeys.365.5725>
- Barrow, S. (1998). A monograph of *Phoenix* L. (Palmae: Coryphoideae). *Kew Bulletin*, 53, 513–575. <https://doi.org/10.2307/4110478>
- Battesti, V. (2004). Odeur sui generis, Le subterfuge dans la domestication du palmier dattier (Tassili n'Ajjer, Algérie). *Anthropozoologica*, 39, 301–309. <http://hal.archives-ouvertes.fr/halshs-00004025>
- Battesti, V. (2005). *Jardins au désert, Évolution des pratiques et savoirs oasiens, Jerid Tunisien*. Paris: Éditions IRD. <http://hal.archives-ouvertes.fr/halshs-00004025>
- Battesti, V. (2013). L'agrobiodiversité du dattier (*Phoenix dactylifera* L.) dans l'oasis de Siwa (Égypte): Entre ce qui se dit, s'écrit et s'oublie. *Revue D'ethnoécologie*, 4, 1–64. <https://doi.org/10.4000/ethnoecologie.1538>
- Battesti, V. (2015). Resources and appropriations: Back to the Jerid Oases (Tunisia) after the Revolution. *Études Rurales*, 2013(2), 153–175. <https://doi.org/10.4000/etudesrurales.9954>
- Battesti, V. (2018). Les possibilités d'une île, Insularités oasiennes au Sahara et genèse des oasis. In G. Tallet, & T. Sauzeau (Eds.), *Mer et désert de l'Antiquité à nos jours, Approches croisées, Histoire* (pp. 105–144). Rennes: Presses universitaires de Rennes.
- Battesti, V., Gros-Balthazard, M., Ogéron, C., Ivorra, S., Terral, J.-F., & Newton, C. (2018). Date palm agrobiodiversity (*Phoenix dactylifera* L.) in Siwa oasis, Egypt: Combining ethnography, morphometry, and genetics. *Human Ecology*, 46, 529–546. <https://doi.org/10.1007/s10745-018-0006-y>
- Ben Salah, M. (2012). *Rapport d'expertise technique sur la biodiversité oasienne en Tunisie*. RADD (Réseau associatif pour le développement durable des oasis); ASOC (Association de sauvegarde des oasis de Chenini).
- Berlin, B., Breedlove, D. E., & Raven, P. H. (1973). General Principles of Classification and Nomenclature in Folk Biology. *American Anthropologist New Series*, 75, 214–242. <https://doi.org/10.2307/672350>
- Berlin, B., Breedlove, D. E., & Raven, P. H. (1974). *Principles of Tzeltal plant classification; an introduction to the botanical ethnography of a Mayan-speaking people of highland Chiapas, Language, thought, and culture*. New York, New York: Academic Press.
- Bietak, M. (2005). The Thutmose Stronghold of Perunefer. *Egyptian Archaeology*, 26, 13–17.
- Bietak, M., Maritanos, N., Palivou, C., & Brysbaert, A. (2007). *Taureador Scenes in Tell El-Dab'a (Avaris) and Knossos, Untersuchungen der Zweigstelle Kairo des Österreichischen Archäologischen Institutes 27, Denkschriften der Gesamtakademie. Österreichische Akademie der Wissenschaften*, Wien.
- Billotte, N., Marseillac, N., Brottier, P., Noyer, J.-L.-L., Jacquemoud-Collet, J.-P.-P., Moreau, C., ... Risterucci, A.-M. (2004). Nuclear microsatellite markers for the date palm (*Phoenix dactylifera* L.): Characterization and utility across the genus *Phoenix* and in other palm genera. *Molecular Ecology Notes*, 4, 256–258. <https://doi.org/10.1111/j.1471.8286.2004.00634.x>
- Boydak, M. (2019). A new subspecies of *Phoenix theophrasti* Greuter (*Phoenix theophrasti* Greuter subsp. *golkoyana* Boydak) from Turkey. *Forestist*, 69, 133–144. <https://doi.org/10.26650/forestist.2019.19016>
- Burel, F., & Baudry, J. (1999). *Écologie du paysage. Concepts, méthodes et applications*. Paris: Tec & Doc.
- Chaluvadi, S. R., Young, P., Thompson, K., Bahri, B. A., Gajera, B., Narayanan, S., ... Bennetzen, J. L. (2019). *Phoenix* phylogeny, and analysis of genetic variation in a diverse collection of date palm (*Phoenix dactylifera*) and related species. *Plant Divers*, 41, 330–339. <https://doi.org/10.1016/j.pld.2018.11.005>
- Chao, C. T., & Krueger, R. R. (2007). The Date Palm (*Phoenix dactylifera* L.): Overview of Biology, Uses, and Cultivation. *HortScience*, 42, 1077–1082. <https://doi.org/10.21273/HORTSCI.42.5.1077>
- Chen, L.-Y., VanBuren, R., Paris, M., Zhou, H., Zhang, X., Wai, C. M., ... Ming, R. (2019). The *bracteatus* pineapple genome and domestication of clonally propagated crops. *Nature Genetics*, 51, 1549–1558. <https://doi.org/10.1038/s41588-019-0506-8>

- Colin, F. (1997). Ammon, Parammon, Poséidon, Héra et Libye à Siwa. *Bulletin De L'institut Français D'archéologie Orientale*, 97–108. <https://www.ifao.egnet.net/bifao/097/07/>
- Collins, J. A., Prange, M., Caley, T., Gimeno, L., Beckmann, B., Mülitz, S., ... Scheff, E. (2017). Rapid termination of the African Humid Period triggered by northern high-latitude cooling. *Nature Communications*, 8, <https://doi.org/10.1038/s41467-017-01454-y>
- de Mendiburu, F. (2015). *agricolae: Statistical Procedures for Agricultural Research*. Available at <https://cran.r-project.org/package=agricolae>
- Devanand, P. S., & Chao, C. T. (2003). Genetic variation in "Medjool" and "Deglet Noor" date (*Phoenix dactylifera* L.) cultivars in California detected by fluorescent-AFLP markers. *The Journal of Horticultural Science & Biotechnology*, 78, 405–409.
- Dray, S., & Dufour, A. B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, 22, 1–20.
- Duemichen, J. (1877). *Die Oasen der libyschen Wüste: Ihre alten Namen und ihre Lage, ihre vorzüglichsten Erzeugnisse und die in ihren Tempeln verehrten Gottheiten; nach den Berichten der altägyptischen Denkmäler*, Strassburg, Verlag von Karl J. Trübner. <http://digi.ub.uni-heidelberg.de/diglit/duemichen1877>
- El Hadrami, A., Daayf, F., Elshibli, S., Lain, S. M., & El Hadrami, I. (2011). Somaclonal Variation in Date Palm. In S. M. Jain et al (Eds.), *Date Palm Biotechnology* (pp. 183–203). Dordrecht: Springer B.V. [http://doi.org/10.1007/978-94-007-1318-5\\_9](http://doi.org/10.1007/978-94-007-1318-5_9)
- El-Assar, A. M., Krueger, R. R., Devanand, P. S., & Chao, C. C. T. (2005). Genetic analysis of Egyptian date (*Phoenix dactylifera* L.) accessions using AFLP markers. *Genetic Resources and Crop Evolution*, 52, 601–607. <https://doi.org/10.1007/s10722-004-0583-z>
- Elhoumaizi, M. A., Devanand, P. S., Fang, J., & Chao, C. C. T. (2006). Confirmation of "Medjool" date as a landrace variety through genetic analysis of "Medjool" accessions in Morocco. *Journal of the American Society for Horticultural Science*, 131, 403–407. <https://doi.org/10.21273/JASHS.131.3.403>
- El-Sharabasy, S. F., & Rizk, R. M. (2019). *Atlas of date palm in Egypt*. Cairo: Food & Agriculture Org. <http://www.fao.org/3/ca5205b/CA5205B.pdf>
- el-Wakil, H. E. M., & Harhash, M. M. (1998). *Evaluation of some date palm cultivars grown in Siwa oasis. First International Conference on Date Palms* (pp. 583–601). Al-Ain: UAE University.
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Fakhry, A. (1990). *Siwa Oasis*. Cairo: American University in Cairo Press.
- Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, 164, 1567–1587.
- Flowers, J. M., Hazzouri, K. M., Gros-Balthazard, M., Mo, Z., Koutrumpa, K., Perrakis, A., ... Purugganan, M. D. (2019). Cross-species hybridization and the origin of North African date palms. *Proceedings of the National Academy of Sciences*, 116(5), 1651–1658. <https://doi.org/10.1073/pnas.1817453116>
- Greiner, S., Sobanski, J., & Bock, R. (2015). Why are most organelle genomes transmitted maternally? *BioEssays*, 37, 80–94. <https://doi.org/10.1002/bies.201400110>
- Gros-Balthazard, M., Galimberti, M., Kousathanas, A., Newton, C., Ivorra, S., Paradis, L., ... Wegmann, D. (2017). The discovery of wild date palms in Oman reveals a complex domestication history involving Centers in the Middle East and Africa. *Current Biology*, 27, 2211–2218. <https://doi.org/10.1016/j.cub.2017.06.045>
- Gros-Balthazard, M., Hazzouri, K. M., & Flowers, J. M. (2018). Genomic insights into date palm origins. *Genes*, 9, 1–14. <https://doi.org/10.3390/genes9100502>
- Gruber, B., & Adamack, A. T. (2015). landgenreport: A new r function to simplify landscape genetic analysis using resistance surface layers. *Molecular Ecology Resources*, 15, 1172–1178. <https://doi.org/10.1111/1755-0998.12381>
- Gurevich, V., Lavi, U., & Cohen, Y. (2005). Genetic variation in date palms propagated from offshoots and tissue culture. *Journal of the American Society for Horticultural Science*, 130, 46–53. <https://doi.org/10.21273/JASHS.130.1.46>
- Hazzouri, K. M., Flowers, J. M., Visser, H. J., Khierallah, H. S. M. M., Rosas, U., Pham, G. M., ... Purugganan, M. D. (2015). Whole genome re-sequencing of date palms yields insights into diversification of a fruit tree crop. *Nature Communications*, 6, 8,824. <https://doi.org/10.1038/ncomms9824>
- Hemeid, A. A., Sanaa, A. R., & Abd El-Rahman, T. M. (2007). Molecular characterization of different date palm (*Phoenix dactylifera* L.) cultivars grown in Siwa Oasis. *Egyptian Journal of Genetics and Cytology*, 36, 145–162.
- Henderson, A. (1986). A review of pollination studies in the Palmae. *Botanical Review*, 52, 221–259. <https://doi.org/10.1007/BF02860996>
- Henderson, S. A., Billotte, N., & Pintaud, J. C. (2006). Genetic isolation of Cape Verde Island *Phoenix atlantica* (Arecaceae) revealed by microsatellite markers. *Conservation Genetics*, 7, 213–223. <https://doi.org/10.1007/s10592-006-9128-7>
- Henry, A. G., Brooks, A. S., & Piperno, D. R. (2011). Microfossils in calculus demonstrate consumption of plants and cooked foods in Neanderthal diets (Shanidar III, Iraq; Spy I and II, Belgium). *Proceedings of the National Academy of Sciences*, 108, 486–491. <https://doi.org/10.1073/pnas.1016868108>
- Hohler, T. B., & Maspero, G. (1900). *Report on the Oasis of Siwa*. Cairo: Waterlow and sons limited.
- Hunn, E. (1982). The utilitarian factor in folk biological classification. *American Anthropologist*, 84, 830–847. <https://doi.org/10.1525/aa.1982.84.4.02a00070>
- Jain, S. M. (2012). Date palm biotechnology: Current status and prospective - an overview. *Emirates Journal of Food and Agriculture*, 24, 400–407.
- Jaradat, A. A. (2016). *Genetic diversity and erosion in plants, Sustainable Development and Biodiversity*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-25954-3>
- Jarni, K., Jakše, J., & Brus, R. (2014). Vegetative propagation: Linear barriers and somatic mutation affect the genetic structure of a *Prunus avium* L. *Stand Forestry*, 88, 612–621. <https://doi.org/10.1093/forestry/cpv029>
- Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27, 3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>
- Kalinowski, S., Taper, M., & Marshall, T. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 16, 1099–1106. <https://doi.org/10.1111/j.1365-294X.2007.03089.x>
- Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2, e281. <https://doi.org/10.7717/peerj.281>
- Khanamm, S., Sham, A., Bennetzen, J. L., & Aly, M. A. M. (2012). Analysis of molecular marker-based characterization and genetic variation in date palm (*Phoenix dactylifera* L.). *Australian Journal of Crop Science*, 6, 1236–1244.
- Kuhlmann, K. P. (1988). Das Ammoneion: Archäologie Geschichte und Kultpraxis des Orakels von Siwa. Mainz am Rhein: P. von Zabern.
- Kuhlmann, K.-P. (1999). Siwa Oasis, Late period and Graeco-Roman sites. In K. A. Bard, & S. B. Shubert (Eds.), *Encyclopedia of the archaeology of ancient Egypt* (pp. 900–907). London and New York, NY: Routledge.
- Kuhlmann, K.-P. (2013). The realm of "two deserts": Siwah Oasis between east and west. In F. Förster, & H. Riemer (Eds.), *Desert road archaeology in ancient Egypt and beyond, Africa Praehistorica* (pp. 133–166). Köln: Heinrich-Barth-Institut.



- Landsberger, B. (1967). The date palm and its by-products according to the cuneiform sources. *Graz, Weidner, Archiv Für Orientforschung, Beiheft*.
- Laoust, É. (1932). *Siwa: I. Son parler*. Paris: Publications de l'Institut des hautes-études marocaines. Librairie Ernest Leroux.
- Latour, B. (2013). *An inquiry into modes of existence: An anthropology of the Moderns* (p. 486). Cambridge, MA: Harvard University Press.
- Leclant, J. (1950). "Per Africae Sitientia". Témoignages des sources classiques sur les pistes menant à l'oasis d'Ammon. *Bulletin De L'institut Français D'archéologie Orientale*, 49, 193–253. <https://www.ifao.egnet.net/bifao/049/10/>
- Liphshitz, N., & Nadel, D. (1997). Epipalaeolithic (19,000 B.P.) charred wood remains from Ohalo II, Sea of Galilee, Israel. *Mitekufat Haeven. Journal of the Israel Prehistoric Society*, 27, 5–18.
- Malleson, C. (2016). Informal intercropping of legumes with cereals? A re-assessment of clover abundance in ancient Egyptian cereal processing by-product assemblages: Archaeobotanical investigations at Khentkawes town, Giza (2300–2100 bc). *Vegetation History and Archaeobotany*, 25(5), 431–442. <https://doi.org/10.1007/s00334-016-0559-x>
- Malleson, C., & Miracle, R. (2018). *Giza Botanical Database [WWW Document]*. <https://doi.org/10.6078/M7JH3J99>
- Mathew, L. S., Seidel, M. A., George, B., Mathew, S., Spannagl, M., Haberer, G., ... Malek, J. A. (2015). A genome-wide survey of date palm cultivars supports two major subpopulations in *Phoenix Dactylifera*. *G3: Genes, Genomes, Genetics*, 5(7), 1429–1438. <https://doi.org/10.1534/g3.115.018341>
- Mattingly, D. J. (2017). The Garamantes and the Origins of Saharan Trade: State of the Field and Future Agendas. In D. J. Mattingly, V. Leitch, C. N. Duckworth, A. Cuénod, M. Sterry, & F. Cole (Eds.), *Trade in the ancient sahara and beyond, trans-saharan archaeology* (pp. 1–52). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108161091.002>
- McKey, D., Elias, M., Pujol, B., & Duputié, A. (2012). Ecological Approaches to Crop Domestication. In P. Gepts, T. R. Famula, R. L. Bettinger, S. B. Brush, A. B. Damania, P. E. McGuire, & C. O. Qualset (Eds.), *Biodiversity in agriculture: Domestication, evolution, and sustainability* (pp. 377–406). Cambridge: Cambridge University Press. <https://hal.archives-ouvertes.fr/hal-02349581>
- McKey, D., Elias, M., Pujol, B., Duputié, A., & Duputié, A. (2010). The evolutionary ecology of clonally propagated domesticated plants. *New Phytologist*, 186, 318–332. <https://doi.org/10.1111/j.1469-8137.2010.03210.x>
- Meilleur, B. A. (1987). Des ethnoscience à l'ethnoécologie ou du rôle des représentations écologiques populaires dans des sociétés traditionnelles. *Écologie Humaine*, 5 (2), 3–23. <http://hdl.handle.net/2042/41279>
- Mohamoud, Y. A., Mathew, L. S., Torres, M. F., Younuskunju, S., Krueger, R., Suhre, K., & Malek, J. A. (2019). Novel subpopulations in date palm (*Phoenix dactylifera*) identified by population-wide organellar genome sequencing. *BMC Genomics*, 20, 1–7. <https://doi.org/10.1186/s12864-019-5834-7>
- Moncada, X., Pelsy, F., Merdinoglu, D., & Hinrichsen, P. (2007). Genetic diversity and geographical dispersal in grapevine clones revealed by microsatellite markers. *Genome*, 49, 1459–1472. <https://doi.org/10.1139/g06-102>
- Moussouni, S., Pintaud, J., Vigouroux, Y., & Bouguedoura, N. (2017). Diversity of Algerian oases date palm (*Phoenix dactylifera* L., Arecaceae): Heterozygote excess and cryptic structure suggest farmer management had a major impact on diversity. *PLoS ONE*, 12, 1–14. <https://doi.org/10.1371/journal.pone.0175232>
- Naseef, A. S. (1995). *al-'Ulā, a study of cultural and social heritage*. unpublished translation from Arabic to English of « Naṣīf, 'Abdallāh bin Adām Ṣāliḥ, 1995 — al-'Ulā: Dirāsa fī l-thurāth ḥadārī wa-l-ijtimā'ī, Riyādh, 'Abdallāh Naṣīf, Riyādh.
- Nei, M. (1972). Genetic distance between populations. *American Naturalist*, 106, 283–292. <https://doi.org/10.1086/282771>
- Newton, C., Gros-Balthazard, M., Ivorra, S., Paradis, L., Pintaud, J.-C., & Terral, J.-F. (2013). *Phoenix dactylifera* and *P. sylvestris* in Northwestern India: A glimpse into their complex relationships. *Palms*, 57, 37–50.
- Olivier de Sardan, J.-P. (1998). Émique. *L'Homme*, 38, 151–166. <https://doi.org/10.3406/hom.1998.370510>
- Paradis, E. (2010). pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, 26, 419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Pelling, R. (2005). Garamantian agriculture and its significance in a wider North African context: The evidence of the plant remains from the Fazzan project. *The Journal of North African Studies*, 10, 397–412. <https://doi.org/10.1080/13629380500336763>
- Petit, R. J., El Mousadik, A., & Pons, O. (1998). Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, 12, 844–855. <https://doi.org/10.1046/j.1523-1739.1998.96489.x>
- Peyron, G. (2000). *Cultiver le palmier-dattier*. [Montpellier]: CIRAD.
- Pintaud, J.-C., Ludeña, B., Aberlenc-Bertossi, F., Zehdi, S., Gros-Balthazard, M., Ivorra, S., ... Bouguedoura, N. (2013). Biogeography of the date palm (*Phoenix dactylifera* L., Arecaceae): Insights on the origin and on the structure of modern diversity. *ISHS Acta Horticulture*, 994, 19–36. <https://doi.org/10.17660/ActaHortic.2013.994.1>
- Pintaud, J.-C.-C., Zehdi, S., Couvreur, T. L. P., Barrow, S., Henderson, S., Aberlenc-Bertossi, F., ... Billotte, N. (2010). Species delimitation in the genus *Phoenix* (Arecaceae) based on SSR markers, with emphasis on the identity of the Date Palm (*Phoenix dactylifera* L.). In O. Seberg, G. Petersen, A. Barfod, & J. Davis (Eds.), *Diversity, phylogeny, and evolution in the monocotyledons* (pp. 267–286). Aarhus: Aarhus Univ Press.
- Popenoe, W. (1913). *Date growing in the Old and New Worlds*. Altadena, CA: West India Gardens.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Racchi, M. L., Bove, A., Turchi, A., Bashir, G., Battaglia, M., & Camussi, A. (2013). Genetic characterization of Libyan date palm resources by microsatellite markers. *Biotech*, 3, 4–21. <https://doi.org/10.1007/s13205-013-0116-6>
- Rhouma, A. (1994). *Le palmier dattier en Tunisie, I. Le patrimoine génétique*. Tunis, Arabesques, INRA Tunisie, GRIDAO France, PNUD/FAO, 1.
- Rhouma, A. (2005). *Le palmier dattier en Tunisie, I. Le patrimoine génétique*. Rome, IPGRI, UNDP, GEF/FEM, INRAT vol. 2. <https://www.bioversityinternational.org/e-library/publications/detail/le-palmier-dattier-en-tunisie/>
- Rieger, A.-K. (2017). The various ways of being mobile: Habitual knowledge, life-strategies and the ancient route networks on the Eastern Marmarica- Plateau (Northern Libyan Desert). *Open Archaeology*, 3, 49–68. <https://doi.org/10.1515/opar-2017-0003>
- Riou, C. (1990). Bioclimatologie des oasis. In V. Dollé, & G. Toutain (Eds.), *Les systèmes agricoles oasiens* (pp. 207–220). Montpellier: CIEHAM.
- Roué, M., & Nakashima, D. (2018). Indigenous and local knowledge and science: From validation to knowledge coproduction. In H. Callan (Ed.), *The international encyclopedia of anthropology*. Oxford: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118924396.wbiea.2215>
- Rousset, F. (2008). genepop'007: A complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, 8, 103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Selim, H. H. A., El-Mahdi, M. A. M., & El-Hakeem, M. S. (1970). Studies on the evaluation of fifteen local date varieties grown under desert condition in Siwa Oasis, U. A. R. *Bulletin de L'institut du desert d'egypte XVIII*, 137–155.



- Solecki, R. S., & Leroi-Gourhan, A. (1961). Palaeoclimatology and archaeology in the near east. *Annals of the New York Academy of Sciences*, 95, 729–739. <https://doi.org/10.1111/j.1749-6632.1961.tb50073.x>
- Struffolino, S. (2012). *L'oasi di Ammone. Ruolo politico, economico e culturale di Siwa nell'antichità. Una ricostruzione critica, Scienze dell'antichità, filologico-letterarie e storico-artistiche*. Aracne, Roma
- Tallet, P. (2017). Du pain et des céréales pour les équipes royales: Le grand papyrus comptable du ouadi el-Jarf (papyrus H). *Nehet*, 5, 99–117.
- Tengberg, M. (2012). Beginnings and early history of date palm garden cultivation in the Middle East. *Journal of Arid Environments*, 86, 139–147. <https://doi.org/10.1016/j.jaridenv.2011.11.022>
- Tengberg, M., & Newton, C. (2016). Origine et évolution de la phénici-culture au Moyen-Orient et en Égypte. In M. P. Ruas (dir.), *Des fruits d'ici et d'ailleurs: Regards sur l'histoire de quelques fruits consommés en Europe*, Mouans-Sartoux, Éditions Omnisciences, *Histoire des savoirs* (pp. 83–105). Paris: Collection Histoire des Savoirs, Éditions Omniscience.
- Terral, J. F., Newton, C., Ivorra, S., Gros-Balthazard, M., Tito de Morais, C., Picq, S., ... Pintaud, J. C. (2012). Insights into the historical biogeography of the date palm (*Phoenix dactylifera* L.) using geometric morphometry of modern and ancient seeds. *Journal of Biogeography*, 39, 929–941. <https://doi.org/10.1111/j.1365-2699.2011.02649.x>
- Van der Veen, M., & Westley, B. (2010). Paleoeconomic studies. In: D. J. Mattingly (Ed.), *The Archaeology of Fazzan. Vol. 3: Excavations of C. M. Daniels* (pp. 489–519). London: Society for Libyan Studies.
- Vardareli, N., Doğaroğlu, T., Doğan, E., Taşkın, V., & Göçmen Taşkın, B. (2019). Genetic characterization of tertiary relict endemic *Phoenix theophrasti* populations in Turkey and phylogenetic relations of the species with other palm species revealed by SSR markers. *Plant Systematics and Evolution*, 305(6), 415–429. <https://doi.org/10.1007/s00606-019-01580-8>
- Zango, O., Cherif, E., Chabrilange, N., Zehdi-Azouzi, S., Gros-Balthazard, M., Naqvi, S. A., ... Aberlenc, F. (2017). Genetic diversity of Southeastern Nigerien date palms reveals a secondary structure within Western populations. *Tree Genetics and Genomes*, 13, 75. <https://doi.org/10.1007/s11295-017-1150-z>
- Zango, O., Rey, H., Bakasso, Y., Lecoustre, R., Aberlenc, F., & Pintaud, J.-C. (2016). Local practices and knowledge associated with date palm cultivation in Southeastern Niger. *Agricultural Science*, 07, 586–603. <https://doi.org/10.4236/as.2016.79056>
- Zehdi-Azouzi, S., Cherif, E., Moussouni, S., Gros-Balthazard, M., Abbas Naqvi, S., Ludeña, B., ... Aberlenc-Bertossi, F. (2015). Genetic structure of the date palm (*Phoenix dactylifera*) in the Old World reveals a strong differentiation between eastern and western populations. *Annals of Botany*, 116, 101–112. <https://doi.org/10.1093/aob/mcv068>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Gros-Balthazard M, Battesti V, Ivorra S, et al. On the necessity of combining ethnobotany and genetics to assess agrobiodiversity and its evolution in crops: A case study on date palms (*Phoenix dactylifera* L.) in Siwa Oasis, Egypt. *Evol Appl*. 2020;00:1–23. <https://doi.org/10.1111/eva.12930>

## Supplementary material for

### On the necessity of combining ethnobotany and genetics to assess agrobiodiversity and its evolution in crops: a case study on date palms (*Phoenix dactylifera* L.) in Siwa Oasis, Egypt

#### **Table S1. *Phoenix* spp. accessions included in this paper.** (Excel file):

<https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111/eva.12930&file=eva12930-sup-0002-TableS1-S3.xlsx>

The field “Country” indicates the country of origin of the sample. For some cultivars, it may be different from the place of sampling, as they may be cultivated in a region, but known to originate from another one.

The field “Reference” indicates whether it is new data (from this paper, hence found in Table S2) or if it was sourced from a published study.

For Siwa named types, the local name of the dates is given (sometimes followed, after a /, by an alternative name or pronunciation), followed by the possible name of the palm if different [in brackets], and the possible name in local Arabic (in brackets) (see Battesti, 2013 for explanation about the synonymy issue).

#### **Table S2. Microsatellite genotypes of 176 *Phoenix dactylifera* from Siwa region, nine *Phoenix theophrasti* and two *Phoenix reclinata*.** (Excel file)

<https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111/eva.12930&file=eva12930-sup-0002-TableS1-S3.xlsx>

CpM12 refers to the genotyping data of the chloroplastic minisatellite while the following columns reports genotyping data for the 17 nuclear microsatellites.

For nuclear microsatellites, the two alleles are separated by "/".

Missing data are indicated with NAs.

More information on the samples and the loci can be found in Table S1 and Table 2, respectively.

#### **Table S3. Identity analysis on each pair of samples.** (Excel file)

<https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111/eva.12930&file=eva12930-sup-0002-TableS1-S3.xlsx>

The number of nuclear microsatellite loci typed for each sample is reported (less than 17 indicates missing data), along with the number of matching and mismatching loci.

**Table S4. Pairwise  $F_{ST}$  among various populations and subpopulations.**

(A)

	Eastern <sup>+</sup>	<i>P. theophrasti</i>
<i>P. theophrasti</i>	0.4029	
Western (Siwa included) <sup>+</sup>	0.0875	0.3172

(B)

	Eastern	Siwa*	<i>P. theophrasti</i>
Siwa	0.1193		
<i>P. theophrasti</i>	0.4029	0.3257	
Western (Siwa excluded) <sup>+</sup>	0.0843	0.0566	0.3346

(C)

	Eastern <sup>+</sup>	Siwa Oasis <sup>†</sup>	Siwa uncultivated <sup>†</sup>	<i>P. theophrasti</i>
Siwa Oasis <sup>†</sup>	0.1252			
Siwa uncultivated <sup>†</sup>	0.1420	0.0547		
<i>P. theophrasti</i>	0.4029	0.3517	0.2796	
Western (Siwa excluded) <sup>+</sup>	0.0843	0.0690	0.0505	0.3346

(D)

	Eastern <sup>+</sup>	Siwa cultivated*	Siwa úšik*	Siwa uncultivated <sup>†</sup>	<i>P. theophrasti</i>
Siwa cultivated*	0.1332				
Siwa úšik *	0.1138	0.0070			
Siwa uncultivated <sup>†</sup>	0.1420	0.0673	0.0316		
<i>P. theophrasti</i>	0.4029	0.3694	0.3344	0.2796	
Western (Siwa excluded) <sup>+</sup>	0.0843	0.0781	0.0534	0.0505	0.3346

(E)

	Eastern <sup>+</sup>	Siwa cultivated*	Siwa úšik #1 <sup>∇</sup>	Siwa úšik #2 <sup>∇</sup>	Siwa uncultivated <sup>†</sup>	<i>P. theophrasti</i>
Siwa cultivated*	0.1332					
Siwa úšik #1 <sup>∇</sup>	0.1192	0.0026				
Siwa úšik #2 <sup>∇</sup>	0.111	0.0138	0.0028			
Siwa uncultivated <sup>†</sup>	0.142	0.0673	0.0392	0.0216		
<i>P. theophrasti</i>	0.4029	0.3694	0.3301	0.3566	0.2796	
Western (Siwa excluded) <sup>+</sup>	0.0843	0.0781	0.0599	0.047	0.0505	0.3346

<sup>+</sup> *Eastern* and *Western* accessions refer to date palm accessions from North Africa and the Middle East up to Pakistan, respectively.

\* *Siwa* refers to the 128 unique genotypes sampled in the oasis (named types, úšik #1 and #2) and in the surrounding desert (uncultivated date palms)

† *Siwa Oasis* and *Siwa uncultivated* refer to date palm accessions sampled in the oasis, and uncultivated accessions from the surrounding desert, respectively.

\* *Siwa cultivated* refers to named types from the oasis while *Siwa úšik* refers to both úšik #1 and #2 accessions in Siwa Oasis, respectively.

∇ Siwa úšik #1 and #2 refer to accidental seedlings growing respectively #1 in gardens, and #2 at the edge of gardens or palm groves.

**Table S5. Ancestry proportions in the Structure analysis (column) of the four original populations defined based on geographic origins of sample (row).**

Population origin	Average ancestry at $K = 4$			
	Eastern	Western	Siwa	<i>P. theophrasti</i>
Eastern	<b>97.50</b>	1.34	0.93	0.19
Western	6.48	<b>80.95</b>	11.85	0.71
Siwa	2.73	7.66	<b>88.38</b>	1.24
<i>P. theophrasti</i>	0.47	3.46	0.62	<b>95.43</b>

With  $K = 4$ , we can identify four clusters that mostly correspond to the four populations as defined based on their geographic origin.



**Figure S1. Field photographs.**



(A) View of the eastern part of Siwa Oasis (Egypt) from its urban center at nightfall. In the foreground stands out the ancient city (Shali) built in salt clay, and in the background stand out the palm grove, and, just before one of the salt lakes, the inselberg of Aghurmi where the oracle temple of Amon is erected. September 27, 2017 (6:30 pm), Vincent Battesti

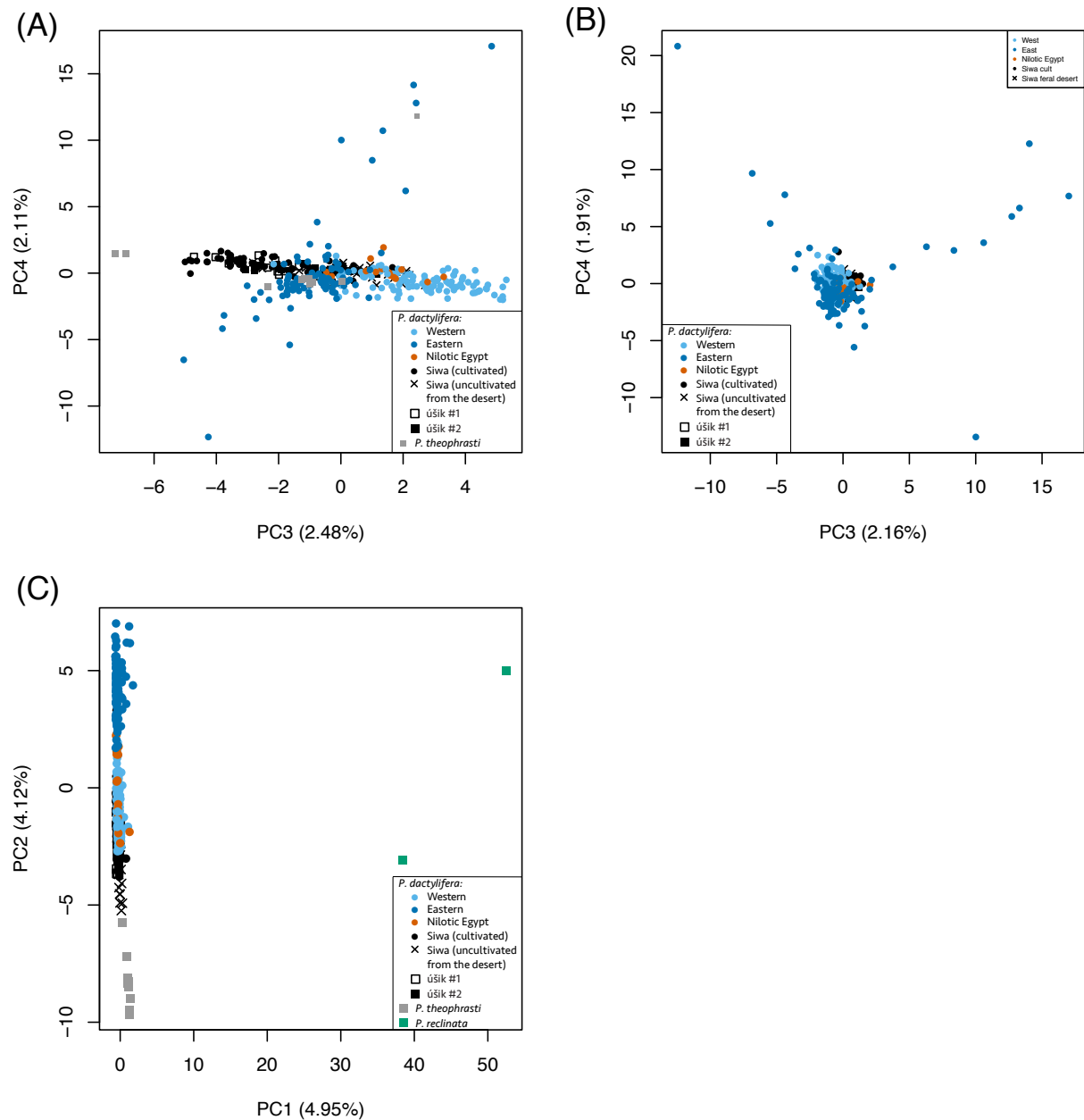


(B) In a garden in the Siwa palm grove (in the Tanyāzi area), one of the many Isiwan farmers who work on harvesting dates, here with his emblematic local sickle helped (this is not always the case) by a belt made of date tree fibers. These ūšik dates are rather intended for fodder consumption. November 11, 2014 (3:30 pm), Vincent Battesti





(C) Two examples of uncultivated date palms in an area called Labbaq south of Siwa Oasis (in the early Great Sand Sea). These two date palms are part of our sample (3467 and 3468) which have been found to be the same genotype (off-shoots of off-shoots that have separated over centuries from each other by 7 to 10 m). September 25, 2017 (4:00 pm), Vincent Battesti.

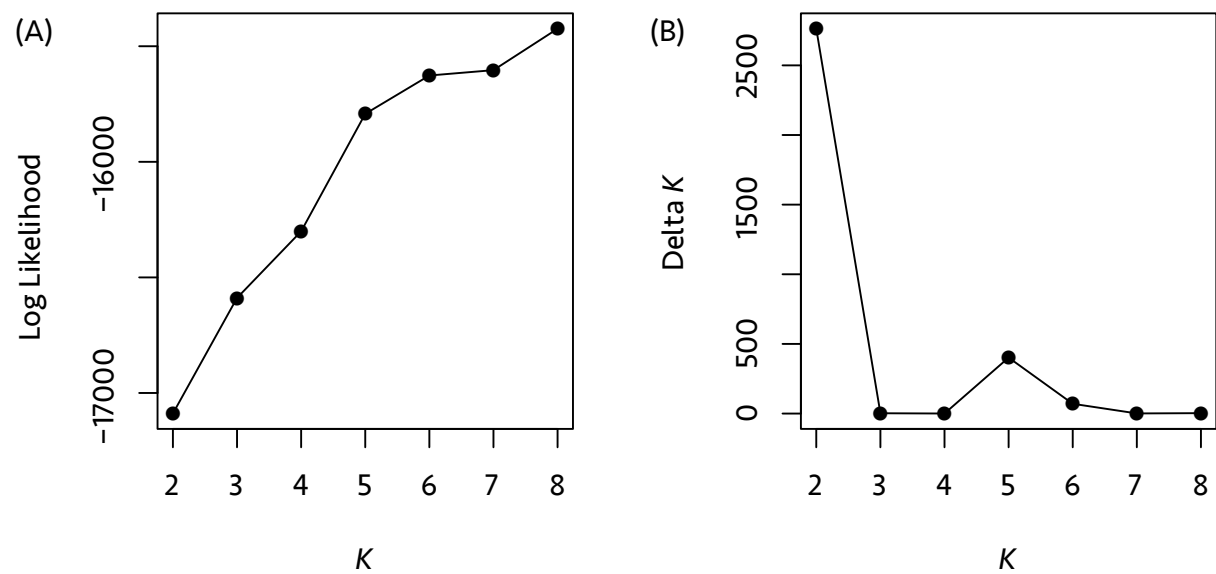


**Figure S2. Principal Component Analyses of *Phoenix* spp. accessions genotyped across 17 nuclear microsatellites.**

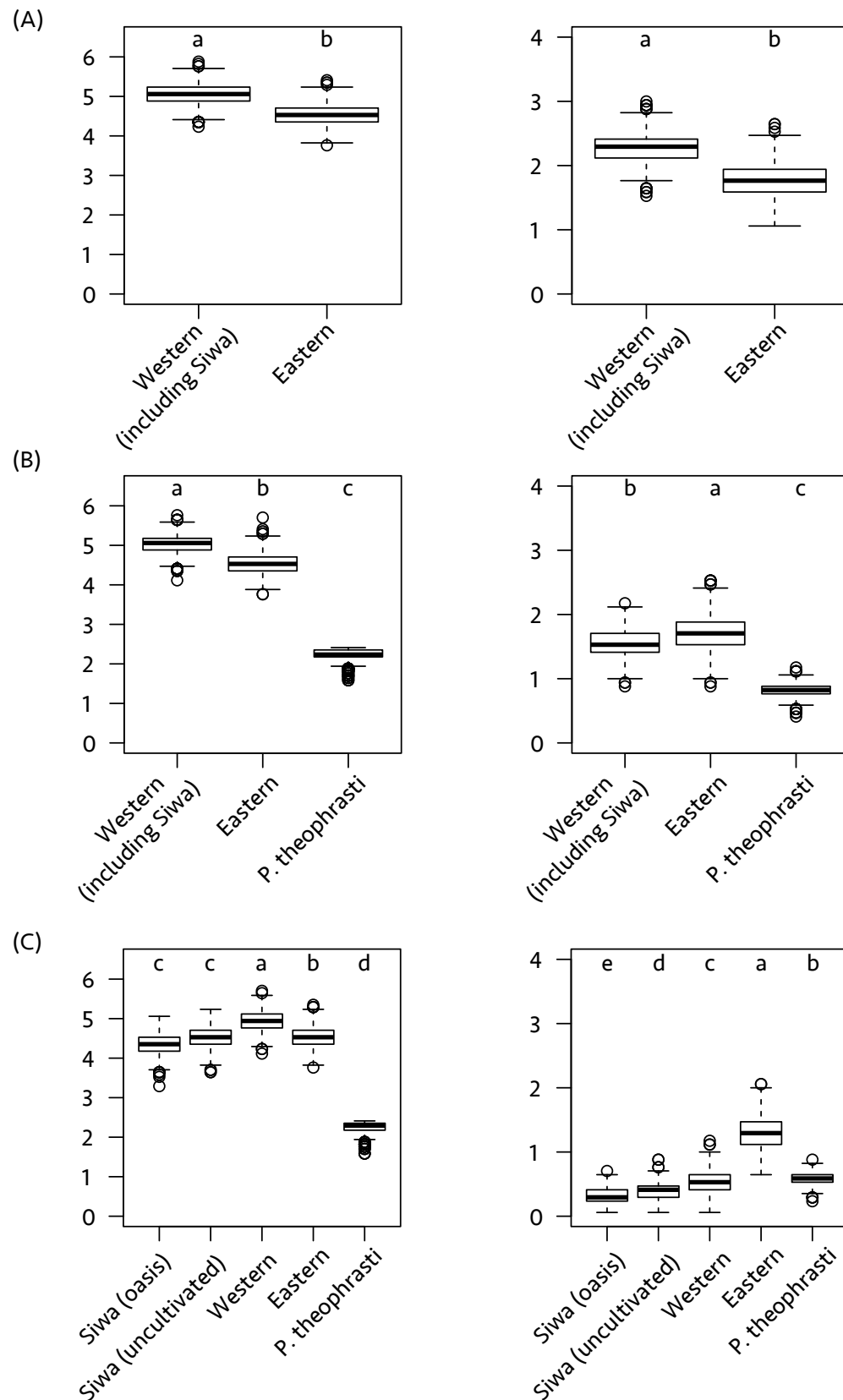
(A) PCA on nine *Phoenix theophrasti* and 347 *Phoenix dactylifera* (variance explained in brackets). Plot of PC1 and 2 can be found in Figure 4A.

(B) PCA on 347 *Phoenix dactylifera* (variance explained in brackets). Plot of PC1 and 2 can be found in Figure 4B

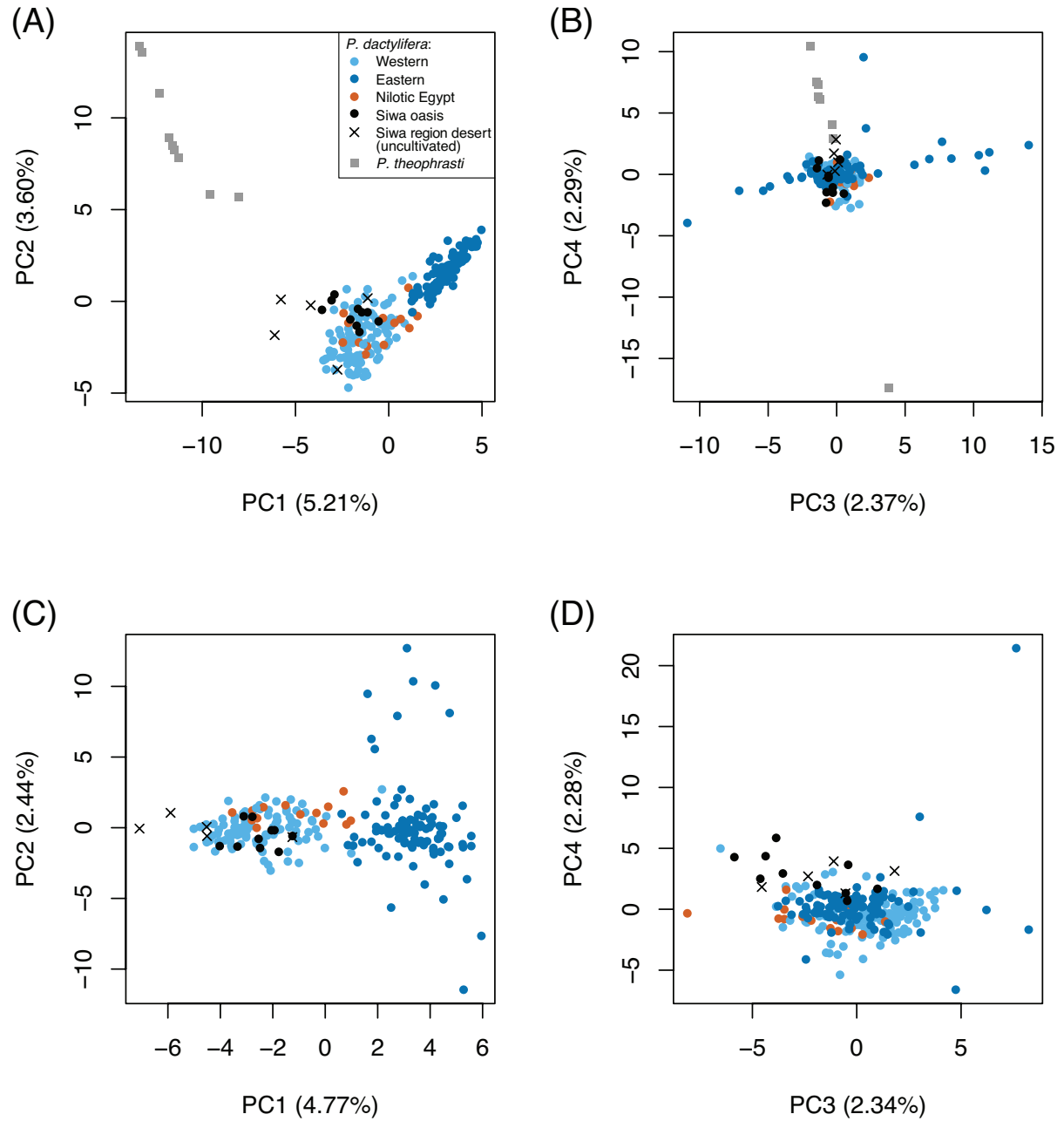
(C) PCA on two *Phoenix reclinata*, nine *Phoenix theophrasti* and 347 *Phoenix dactylifera*.



**Figure S3. Plots of (A) average maximum log likelihood over the ten runs for the Structure analysis performed on 356 date palm and *Phoenix theophrasti* accession (ancestry plot in Figure 4C and Figure S7) and (B) delta  $K$  from the same Structure analysis.**



**Figure S4. Allelic (left) and private allelic (right) richness. Calculated in two (A), three (B) and five (C) populations, using the rarefaction method (haploid sample size = 18, bootstrap number = 1,000).**

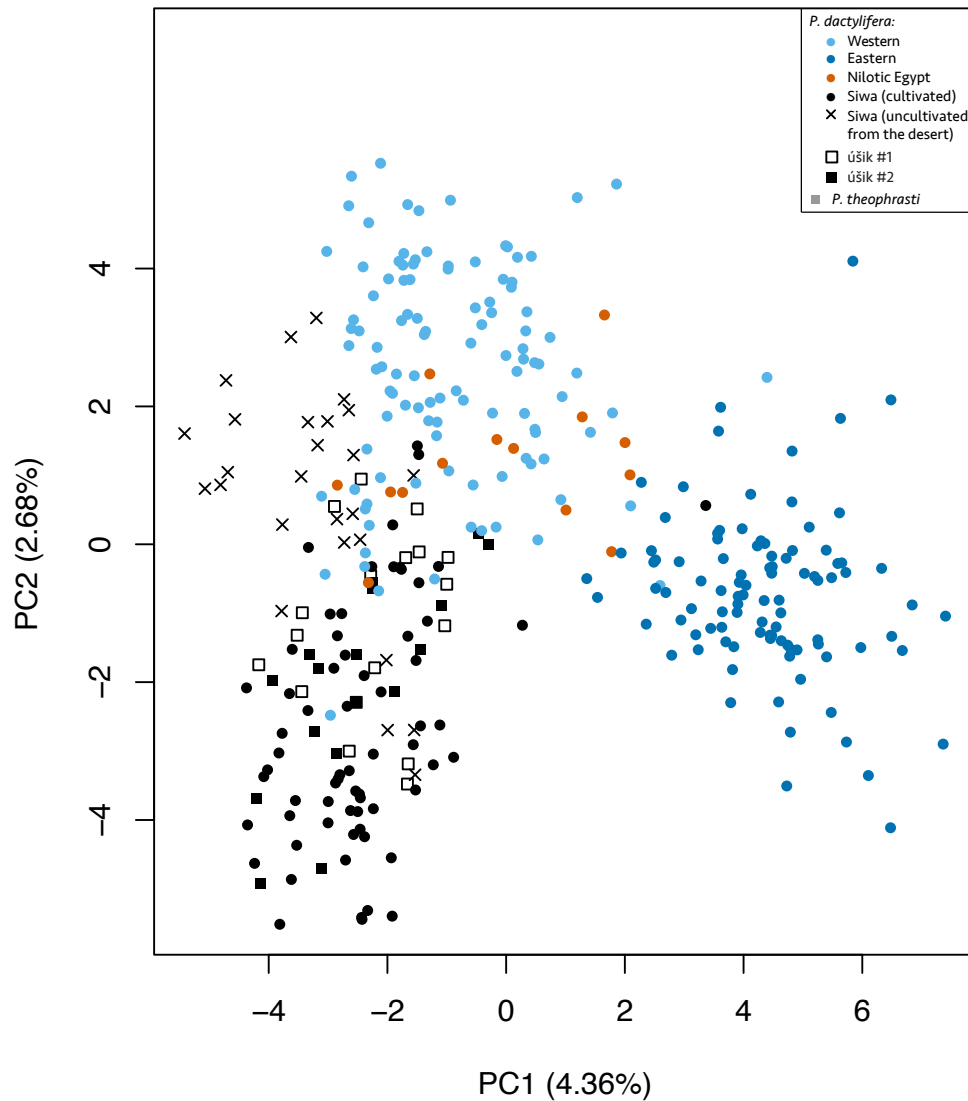


**Figure S5. Principal Component Analyses of *Phoenix* spp. accessions genotyped across 17 nuclear microsatellites. Date palms from Siwa have been randomly sub-sampled so that ten accessions from the current oasis and five from the surrounded desert (uncultivated date palms) are included.**

(A) and (B) are plots of respectively axes 1-2, and 3-4 of a PCA on nine *Phoenix theophrasti* and 234 *Phoenix dactylifera* (variance explained in brackets).

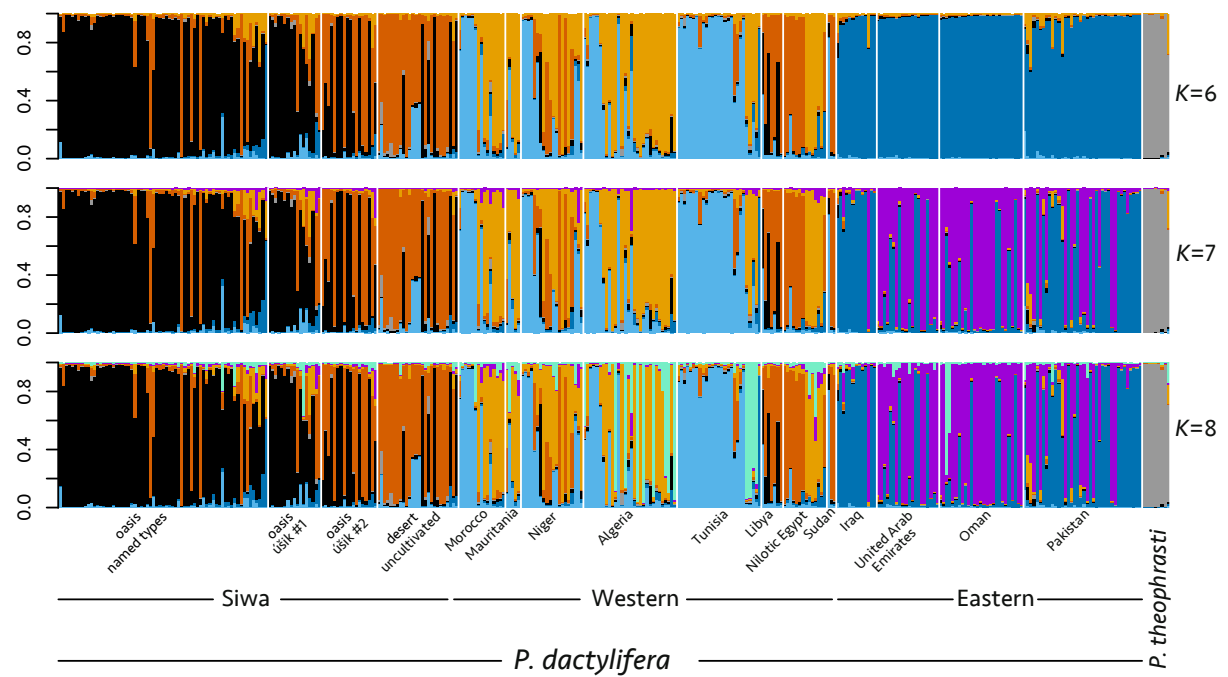
(C) and (D) are plots of respectively axes 1-2, and 3-4 of a PCA on 234 *Phoenix dactylifera* (variance explained in brackets).



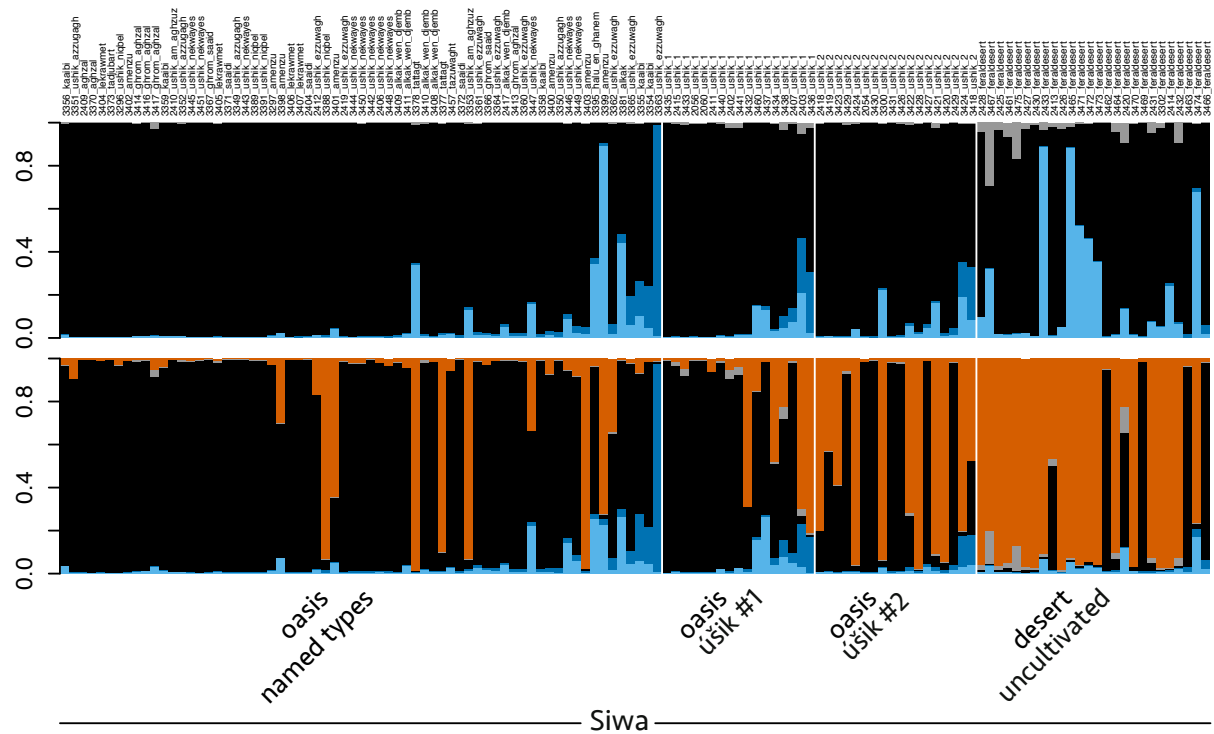


**Figure S6. Principal Component Analysis of 347 *Phoenix dactylifera* accessions genotyped across 17 nuclear microsatellites.**

This is the same Figure as in Figure 4B except that seedlings sampled in the garden (úšik #1) and on the border of the gardens or palm groves (úšik #2) are shown with distinct symbols.

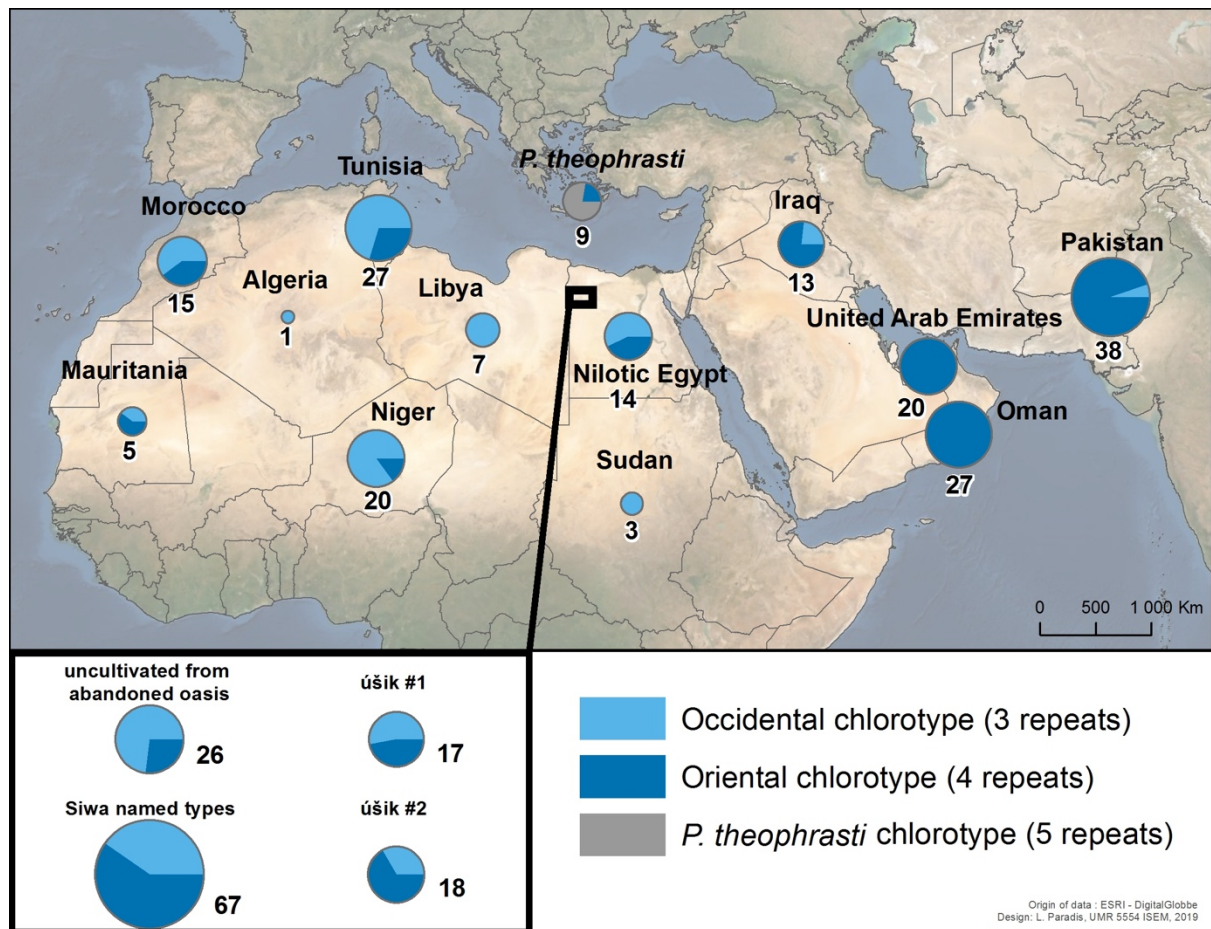


**Figure S7.** Ancestry plot derived from Structure analysis of 356 accessions of *Phoenix dactylifera* and *Phoenix theophrasti* genotyped across 17 nuclear microsatellite loci.



**Figure S8. Zoom on Siwa accessions only, from ancestry plot derived from 356 accessions of *Phoenix dactylifera* and *Phoenix theophrasti* genotyped across 17 nuclear microsatellite loci.**

The plot with all samples can be seen in Figure 5C and Figure S3. Top:  $K = 4$ ; bottom:  $K = 5$ .



**Figure S9.** Distribution of the chlorotypes in *Phoenix dactylifera* and *Phoenix theophrasti* accessions.

## **Supplementary references**

- Battesti, V. (2013). L'agrobiodiversité du dattier (*Phoenix dactylifera* L.) dans l'oasis de Siwa (Égypte): entre ce qui se dit, s'écrit et s'oublie. *Revue d'ethnoécologie*, 4. doi:10.4000/ethnoecologie.1538
- Moussouni, S., Pintaud, J., Vigouroux, Y., Bouguedoura, N., 2017. Diversity of Algerian oases date palm (*Phoenix dactylifera* L., Arecaceae): Heterozygote excess and cryptic structure suggest farmer management had a major impact on diversity. *PLoS ONE* 12(4): e0175232. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0175232>
- Zango, O., Cherif, E., Chabrillange, N., Zehdi-Azouzi, S., Gros-Balthazard, M., Naqvi, S.A., Lemansour, A., Rey, H., Bakasso, ., Aberlenc, F., 2017. Genetic diversity of Southeastern Nigerien date palms reveals a secondary structure within Western populations. *Tree Genet. Genomes* 13, 75. <https://doi.org/10.1007/s11295-017-1150-z>
- Zehdi-Azouzi, S., Cherif, E., Moussouni, S., Gros-Balthazard, M., Abbas Naqvi, S., Ludeña, B., Castillo, K., Chabrillange, N., Bouguedoura, N., Bennaceur, M., Si-Dehbi, F., Abdoukader, S., Daher, A., Terral, J.-F., Santoni, S., Ballardini, M., Mercuri, A., Ben Salah, M., Kadri, K., Othmani, A., Littardi, C., Salhi-Hannachi, A., Pintaud, J.-C., Aberlenc-Bertossi, F., 2015. Genetic structure of the date palm (*Phoenix dactylifera*) in the Old World reveals a strong differentiation between eastern and western populations. *Ann. Bot.* 116, 101–112. <https://doi.org/10.1093/aob/mcv068>