
Article:

Cid Vian, Jerónimo, Pérez-Escobar, Oscar A., Baker, William J. et al. (2025) Genomics and morphometry of herbarium specimens elucidate the origin of the Cape Verde date palm (Phoenix atlantica A.Chev.) and highlight its agronomic potential. Plants People Planet. ISSN: 2572-2611

https://doi.org/10.1002/ppp3.10632

Received: 5 July 2024 | Revised: 28 November 2024 | Accepted: 5 December 2024
DOI: 10.1002/ppp3.10632

Plants People Planet PPP
Open Access

RESEARCH ARTICLE

Genomics and morphometry of herbarium specimens elucidate the origin of the Cape Verde date palm (*Phoenix atlantica* A. Chev.) and highlight its agronomic potential

Jerónimo Cid-Vian[1,2] | **Oscar A. Pérez-Escobar**[1] | **William J. Baker**[1,3] | **Vincent Bonhomme**[4] | **Sylvie Ferrand**[5] | **Sarah Ivorra**[4] | **Thomas Kaczmarek**[4] | **Clémence Pagnoux**[6] | **Jean-Frédéric Terral**[4] | **Nathan Wales**[7] | **Michael D. Purugganan**[5,8,9] | **Simon Creer**[2] | **Alexander S. T. Papadopulos**[2] | **Sidonie Bellot**[1] | **Muriel Gros-Balthazard**[10]

[1]Royal Botanic Gardens, Kew, Richmond, UK

[2]Molecular Ecology and Evolution Group, Environment Centre Wales, Bangor University, Bangor, UK

[3]Department of Biology, Aarhus University, Aarhus C, Denmark

[4]ISEM, Institut des Sciences de l'Evolution-Montpellier, Université de Montpellier/CNRS/IRD/EPHE, Montpellier, France

[5]Center for Genomics and Systems Biology (CGSB), New York University Abu Dhabi, Abu Dhabi, UAE

[6]UMR 7209 AASPE (Archéozoologie, Archéobotanique: Sociétés, Pratiques, Environnement), Muséum national d'Histoire Naturelle, CNRS, Paris, France

[7]Department of Archaeology, University of York, York, UK

[8]Center for Genomics and Systems Biology (CGSB), New York University, New York, USA

[9]Institute for the Study of the Ancient World, New York University, New York, USA

[10]DIADE, Univ Montpellier, CIRAD, IRD, France

Correspondence
Jerónimo Cid Vian, Royal Botanic Gardens, Kew, Richmond TW9 3AE, UK.
Email: j.cid@kew.org

Muriel Gros-Batlthazard, 10DIADE, Univ Montpellier, CIRAD, IRD, France.
Email: muriel.gros-balthazard@ird.fr

Societal Impact Statement

As climate change accelerates, breeding resilient crops is urgent. The date palm (*Phoenix dactylifera* L.), a crucial 18 billion USD fruit crop, underpins North African and West Asian oasis agroecosystems. This study investigates the genetics and morphology of its closest wild relative, the endemic and endangered Cape Verde date palm (*Phoenix atlantica* A.Chev.). We reveal that the latter originated through a feralisation process from domesticated date palms, evolving back into a wild state with unique genetic diversity. This highlights the need to protect its distinctiveness and opens the possibility of enhancing date palm resilience through its inclusion in breeding programs.

Summary

- The date palm (*Phoenix dactylifera* L.). is a major crop but its sustainability is threatened by climate change. This challenge could be mitigated by breeding with wild relatives or feral populations. The endangered and endemic Cape Verde date palm (*Phoenix atlantica* A.Chev.) is thought to be among the closest relatives of the date palm, but its origin and taxonomic status remain unclear, with studies having recovered it either as a separate species or a feral date palm. Better understanding the differentiation of *P. atlantica* and *P. dactylifera* is needed to fully unlock the potential of the Cape Verde date palm as a candidate for date palm improvement.
- In this study, we employ an integrative approach to clarify the origin of *P. atlantica*. Specifically, we investigate the phylogenetic placement, genetic structure, and seed morphology of *P. atlantica* by generating extensive genomic and

---

Sidonie Bellot and Muriel Gros-Batlthazard jointly supervised the study.

Funding information
Natural Environment Research Council, Grant/Award Number: NE/S007423/1; New York University Abu Dhabi Research Institute, Grant/Award Number: 1205H

[Correction added on 18 February 2025, after first online publication: The surname of first and last authors has been corrected.]

seed morphology datasets from expert-verified historical and modern specimens of *Phoenix*, including the type specimen of *P. atlantica*.

- Our findings indicate that *P. atlantica* is monophyletic but nested within *P. dactylifera*. Its genetic diversity overlaps substantially but not completely with North African date palm varieties, and its seeds share characteristics with those of cultivated and feral date palms, notably elongation, a trait linked to domestication.
- These results suggest that *P. atlantica* may be an incipient species that originated from a date palm feralisation event. We emphasise the need for conserving its natural genetic variation and the traditional knowledge associated with its name to support breeding programs.

KEYWORDS
conservation, crop wild relatives, date palm, feral crops, herbariomics, *Phoenix*, seed morphometrics, taxonomy

1 | INTRODUCTION

Population growth, climate change, and shifts in consumption patterns are creating increasing demand for higher food production under more challenging environments (Tilman et al., 2011). Conserving and improving crops for food security is imperative, with crop wild relatives (CWR) and feral plants (escaped domesticated crops that grow and sustain themselves in the wild) playing a crucial role as a source of useful traits and genetic diversity in crops (Dempewolf et al., 2017; Mabry et al., 2023). For instance, the incorporation of CWR into breeding programs has enhanced yield in rice (e.g. Xiao et al., 1996) and tomato (e.g. Tanksley et al., 1996), and improved drought tolerance in barley (Honsdorf et al., 2014). Feral plants are gaining attention as underutilised resources (Mabry et al., 2023; Pisias et al., 2022; Zsögön et al., 2022) as their close relationship with established crops makes them an ideal focus for crop improvement and neo-domestication (the introduction of domestication traits into uncultivated species). However, although numerous cases of feralisation have been reported, few feral crops have been characterised genetically (reviewed by Wu et al., 2021).

The date palm (*Phoenix dactylifera* L.) is a pivotal crop in arid regions of North Africa and West Asia, with 9 million tons of dates produced globally in 2022 (FAO, 2024). Its dependency on significant irrigation, especially in areas projected to face intensified drought in the coming decades (Allan & Douville, 2024; Carr, 2018), combined with its vulnerability to diseases and pests threaten its long-term sustainability (Hazzouri et al., 2020). While phylogenetic relationships within the *Phoenix* genus remain unclear, the Cape Verde date palm, or 'tamareira' (*Phoenix atlantica* A.Chev.; Figure 1), has been identified as the closest wild living relative to the cultivated date palm (reviewed by Gros-Balthazard, et al., 2021) and described as resistant to the Bayoud disease (*Fusarium oxysporum* f. sp. *albedinis*) (Chevalier, 1952). It inhabits arid zones from sea level up to 300 m (Brochmann et al., 1997), making it a potentially interesting target species for improving resilience to disease and climate change in the date palm. *Phoenix atlantica* is endemic to Cape Verde, an island country located approximately 570 km off the coast of Senegal in Africa, in the Atlantic Ocean (Chevalier, 1935; Govaerts & Dransfield, 2005). Due to its fragmented distribution across the drier eastern islands of Sal, Boa Vista, Maio, Santiago, and possibly Santa Luzia (Brochmann et al., 1997), coupled with ongoing population decline, the species is classified as Endangered on the International Union for Conservation of Nature (IUCN) Red List (Duarte et al., 2017). Locally, the Cape Verde date palm is a valued source of shade and construction materials, and its yellow (sometimes red) fruits are occasionally eaten (Chevalier, 1935; Henderson et al., 2003).

The evolutionary origins of *Phoenix atlantica* and its taxonomic status with respect to the date palm remain subject of ongoing debate. Studies have regarded it variously as an imperfectly known taxon in the latest taxonomic monograph of *Phoenix* (Barrow, 1998), a wild relative (Chevalier, 1935, 1952; Henderson et al., 2006) or a feral derivative of the date palm (Flowers et al., 2019; Gros-Balthazard et al., 2017; Sarmiento Cabello et al., 2024). Distinguishing *P. atlantica* and *P. dactylifera* based on morphology is difficult because of a lack of consistent differences (see 'Discussion'; Chevalier, 1935; Barrow, 1998; Henderson et al., 2003, 2006). Furthermore, genetic investigations have not enabled further resolution of the issue, with some studies suggesting that *P. atlantica* and the date palm are different (e.g. Henderson et al., 2006; Sarmiento Cabello et al., 2024) and others suggesting that they are indistinguishable (e.g. Flowers et al., 2019; Gros-Balthazard et al., 2017). Currently, no genetic studies have used the type specimen for *P. atlantica*, raising the possibility that some of the samples employed belonged to *P. dactylifera* or were hybrid individuals. To address this issue, a study of the *P. atlantica* samples used by Chevalier in his original description of the species, particularly the type specimen, would be beneficial. This is now feasible thanks to

FIGURE 1 *Phoenix atlantica* in Ribeira Sao Martinho Grande, Santiago in one of the locations sampled by A. Chevalier in 1935. Photo taken by William J. Baker in May 2002.

herbariomic approaches, which involve the genetic analysis of herbarium specimens. This technique facilitates the examination of historical specimens and has demonstrated its potential for advancing our understanding of palm taxonomy and systematics (e.g. Brewer et al., 2019; Pérez-Calle et al., 2024). Additionally, an important characteristic that remains unexplored in investigating the origin of *P. atlantica* is seed morphology. Seed size and shape are regarded as traits associated with domestication that significantly differ between cultivated and wild date palms (Gros-Balthazard et al., 2016, 2017) making them highly relevant for testing the hypothesis that *P. atlantica* is a feral date palm.

Here, we harness the potential of herbarium collections to clarify the origins of *P. atlantica* and to evaluate its utility as a source of genetic material for date palm crop improvement. We generate new genomic data from the type specimen of *P. atlantica* and from expert-verified specimens from this and six other *Phoenix* species, including the date palm. We analyse these together with genome sequencing data from previous studies using phylogenetic and population genetic methods. In parallel, we perform a comparative morphometric analysis of the seed elongation of *P. atlantica* and other *Phoenix* species. Combining these new genetic and morphological data, we elucidate the evolutionary origins of the Cape Verde date palm, and we discuss its taxonomic status and its potential to support the breeding of more resilient date palm cultivars.

2 | MATERIALS AND METHODS

2.1 | Genomic analyses

2.1.1 | Taxon sampling

Our sampling comprised eight accessions representing *P. atlantica* (two of them from the same individual) and 46 accessions representing six other *Phoenix* species from the western distribution range of the genus (two to 22 accessions per species; Table 1; Dataset S1; Methods S1). Sixteen accessions were sequenced as part of this study, notably three herbarium specimens of *P. atlantica* collected by A. Chevalier in 1934 (Chevalier, 1935), including the type specimen. Genome sequencing data from the other accessions were obtained from GenBank (Dataset S1).

TABLE 1 Number of accessions included in the genetic analyses. Numbers are divided into those newly sequenced for this study (from herbarium or fresh/silica specimens) and those for which data was obtained from GeneBank. 'Herbarium' refers to DNA obtained from plant material that was stored as a herbarium specimen, while 'fresh/silica' refers to DNA obtained from fresh leaf material or from leaf material that was immediately dried in silica gel at the time of collection. Herbarium specimens were collected between 1934 and 2002, while silica-dried material dated from between 1994 and 2017.

	Number of accessions		
Species	Herbarium	Fresh/silica	GenBank
Phoenix atlantica	5	0	3
Phoenix caespitosa	1	1	0
Phoenix canariensis	0	0	4
Phoenix dactylifera	0	0	22
Phoenix reclinata	1	1	1
Phoenix sylvestris	1	1	1
Phoenix theophrasti	1	4	7

2.1.2 | DNA extraction, whole-genome sequencing and genomic data processing

The protocols for DNA extraction, library preparation and sequencing are described in Methods S2. DNA for nine accessions (hereafter 'herbarium' accessions) was extracted from herbarium specimens collected between 1934 and 2002, while the DNA from the other accessions (hereafter 'non-herbarium' accessions) was obtained from fresh or silica-dried material (Dataset S1). DNA libraries were sequenced on an Illumina NovaSeq system to produce 150-bp long paired-end reads. Reads from herbarium specimens were processed with PALEOMIX v.1.3.7 (Schubert et al., 2014), while reads from non-herbarium samples were processed following Gros-Balthazard, Flowers, et al. (2021) (Methods S2). Filtered reads were aligned to a *P. dactylifera* reference genome of length ~772 Mb, containing 18 primary contigs (corresponding to the 18 nuclear chromosomes) and 2371 unplaced contigs (Hazzouri et al., 2019). For reads from herbarium accessions, the alignment was performed using BWA v. 0.7.17-r1188 (Li & Durbin, 2009) through PALEOMIX, whereas for the non-herbarium accessions, it was done with bwa mem v.0.7.15-r1140 (Li, 2014) following Gros-Balthazard, Flowers, et al. (2021). Alingment statistics can be found in Dataset S2.

2.1.3 | Phylogenetic and structure analyses

To account for the low sequencing depth inherent to herbarium specimens, we employed a pseudohaploid approach (Pérez-Escobar et al., 2022) implemented with ANGSD, v.0.940 (Korneliussen et al., 2014) to generate separate nuclear and plastid alignments of all samples (Methods S3). A plastid phylogenetic tree was inferred with IQ-TREE v. 2.2.5 (Minh et al., 2020) based on the whole plastome alignments. The best fitting substitution model was selected with ModelFinder (Kalyaanamoorthy et al., 2017), and 1000 ultrafast bootstrap replicates were performed. To infer a nuclear phylogenetic tree, the nuclear genome alignments were segmented into 10-Kb windows, and maximum likelihood phylogenies were inferred from 7709 of these windows using IQ-TREE as above (details in Methods S3). A single species tree was then inferred based on the resulting 7709 trees using the summary multispecies coalescent approach implemented in wASTRAL v1.16.3.4 (Zhang & Mirarab, 2022) with branch length and support weighting. Both nuclear and plastid species trees were rooted on *P. reclinata*. To examine population structuring and the genetic relationships of the Cape Verde date palm with other *Phoenix* species, we performed population clustering analyses with NGSAdmix v.32 (Skotte et al., 2013) and principal component analysis (PCA) with PCAngsd v.1.10 (Meisner & Albrechtsen, 2018) (Methods S3). In both cases, we used genotype likelihoods to account for the low coverage of some of the accessions.

2.2 | Seed morphometric analyses

Seed morphometrics is a powerful tool which can help determine if a species is wild or feral, as domesticated date palms have elongated seeds, a trait not found in wild *Phoenix* species, which typically have round seeds (Gros-Balthazard et al., 2016). For the first time, *P. atlantica* was included in a morphometric analysis of *Phoenix* seeds. Sampling for morphometric analysis encompassed 30 *P. atlantica* seeds collected by Chevalier in 1934, along with a reference collection of seeds from all *Phoenix* species (7795 seeds from 397 individuals; Table S1; Dataset S3; Methods S4).

Seed shape and length appraisal followed protocols described by Terral et al. (2012) and Gros-Balthazard et al. (2016) (Methods S4). Briefly, each seed was photographed in dorsal and lateral views, seed length was measured on the images using ImageJ (Abramoff et al., 2004) and seed shape appraised by the elliptical Fourier transforms with the package Momocs v.1.4.0 (Bonhomme et al., 2014). We explored the diversity of seed morphology using a PCA applied to length as well as the 56 Fourier coefficients. We then used balanced sample size linear discriminant analysis (LDA) on the same data to classify *Phoenix atlantica* seeds into predefined categories: (1) domesticated date palm (including cultivated and feral date palm) versus wild date palm versus wild *Phoenix* species, (2) cultivated date palm versus feral date palm versus wild date palm versus wild *Phoenix* species, and (3) cultivated date palm versus wild date palm versus wild *Phoenix* species.

3 | RESULTS

3.1 | Genome sequencing of historical and modern *Phoenix* specimens

We sequenced herbarium and silica-dried specimens to elucidate the origin of *P. atlantica* and its relationship to the date palm and other

Phoenix species (Table 1; Dataset S1). An average of 140,535,134 reads per sample was retrieved (Dataset S2). Following alignment to the date palm reference genome, the read depth of the nine datasets obtained from herbarium specimens was on average 3.85×. For the three *P. atlantica* specimens collected by A. Chevalier in 1934, read depth ranged from 0.026 × to 0.65 × (Dataset S2). Datasets obtained from fresh or silica-dried material had an average read depth of 17.72×, ranging from 4.67 × to 44.53 × .

3.2 | Phylogenetic placement of *P. atlantica*

Nuclear and plastid phylogenetic trees were mostly concordant (Figure 2), showing that *P. caespitosa*, *P. canariensis*, *P. sylvestris* and *P. theophrasti* were each monophyletic with strong support (local posterior probability [LPP] in the nuclear tree: 0.897–1; ultrafast bootstrap [UFB] support in the plastid tree: 90%–100%). There were two exceptions to the monophyly of *P. sylvestris* and *P. theophrasti*. PTHE49 is most likely a case of chloroplast capture, and PSYL41 possibly of admixture (see Note S1). In both trees, *P. dactylifera* and *P. atlantica* formed a clade (LPP = 1.0 and UFB = 100%) that made *P. dactylifera* paraphyletic and *P. theophrasti* was recovered as sister to these two species with moderate support (LPP = 0.9; UFB= 90%). *Phoenix sylvestris* was recovered as sister to this clade in the nuclear tree (Figure 2a) or branching inside *P. dactylifera* in the plastid tree (Figure 2b), in both cases with maximal support.

In the nuclear tree (Figure 2a), *P. atlantica* was recovered as a maximally supported (LPP = 1) monophyletic group nested inside *P. dactylifera* and clustering closer to the North African than to the West Asian date palms. In the plastid tree, *P. atlantica* was also part of the *P. dactylifera* clade with maximal support, but it was not recovered as monophyletic (Figure 2b). Instead, all but one *P. atlantica* individuals fell inside the North African date palm clade with maximal support, where their intermixing with West African date palm individuals was poorly to moderately supported (UFB ≤ 84%). The last individual (PATL7) grouped in the West Asian date palm clade with maximal support, even though it was maximally supported as grouping with the other *P. atlantica* individuals in the nuclear tree (Figure 2a).

3.3 | Genetic structure in selected *Phoenix* species

Genetic clustering and PCA consistently differentiated *Phoenix* species, except *P. atlantica* and *P. dactylifera*. In the PCA, each species formed a distinct cluster, except these two species, which showed overlapping positions across the first four components, together explaining 97.5% of the genetic variance (Figure 3a; Figure S1). The genetic clustering showed *P. atlantica* as indistinguishable from Northwestern African *P. dactylifera* and revealed that both share ancestry with *P. theophrasti* (Figure 3b; Figure S2).

To further explore the genetic differentiation of *P. atlantica* and *P. dactylifera*, we performed PCA and genetic clustering solely on these species (Figure 4; Figure S3). This confirmed the greater similarity of *P. atlantica* to North-western African, rather than West Asian, date palms and revealed differences between *P. atlantica* and *P. dactylifera*. Indeed, PC2 (9.5% of the variation) broadly sorted the date palm accessions along an east to west gradient (from right to left), with *P. atlantica* accessions positioned at the same level as, or to the left of, the most westerly date palm accessions from Morocco and Algeria (Figure 4a). This was also visible in the results of the genetic clustering analysis, where *P. atlantica* accessions were most similar to westernmost date palm accessions (Figure 4b; Figure S4). Moreover, accessions of *P. atlantica* were separated from date palm accessions along PC1 (44.7%), with their distribution along this axis broadly reflecting their geographic provenance and distance from mainland Africa (Figure 4a). Specifically, accessions from Sal Island were the closest to Northwestern *P. dactylifera*, followed by accessions from Boa Vista and Maio, and then those from Santiago (except for PATL3, which clustered with Sal specimens).

3.4 | Morphological differences among *Phoenix* species

Our comparative morphological analyses show that *P. atlantica* seeds are elongated and display a pointed apex, similar to those of domesticated date palms, whereas wild *Phoenix* species have rounded seeds (Figure 5a). PCA of seed shape and seed length showed differences in seed morphology between domesticated date palms and wild *Phoenix* (Figure S5; Dataset S4). Seed length varied among *Phoenix* species (Figure S6) and between wild and domesticated types (Figure 5b). On average, domesticated date palms (cultivated and feral, $n = 260$ individuals/5189 seeds) had longer seeds than wild *Phoenix* (wild date palm and wild relative species, $n = 137$ individuals/2636 seeds; one-way ANOVA: $F[4, 7{,}850] = 1777$, p-value < 2×10^{-16}). *Phoenix atlantica* ($n = 30$ seeds) was indistinguishable from domesticated date palm (pairwise t-tests with Bonferroni correction; $t = -1.73$, df = 37.8, adjusted p-value = .912), and it had longer seeds than wild *Phoenix* ($t = 42.6$, df = 41.4, adjusted p-value $= 8.54 \times 10^{-35}$).

LDA consistently predicted seeds of *P. atlantica* as more similar to domesticated date palms than wild *Phoenix*. In the first LDA, discriminating four groups, namely cultivated, feral, wild date palms and wild relative species, the model global accuracy reached 81%, mostly due to a poor accuracy for the feral group (Figure 5c; Figure S7A). Of the 30 *P. atlantica* seeds, 11 were classified into the feral category, six into wild relatives, and three into the cultivated category (Figure S1). In a LDA differentiating two categories, namely domesticated (cultivated and feral date palms) and wild *Phoenix* (wild date palms and wild relative species), the model accuracy reached 92.8%, and 20 of the *P. atlantica* seeds were assigned to the domesticated category by at least 80% of the votes and three to the non-domesticated category (Figure 5e). Building the same LDA model after excluding ferals from the training set (i.e. from the 'domesticated' category) yielded a slightly higher accuracy (94.1%; Figure S8). With this model, the allocation (based on at least 80% of the votes) of feral date palm seeds

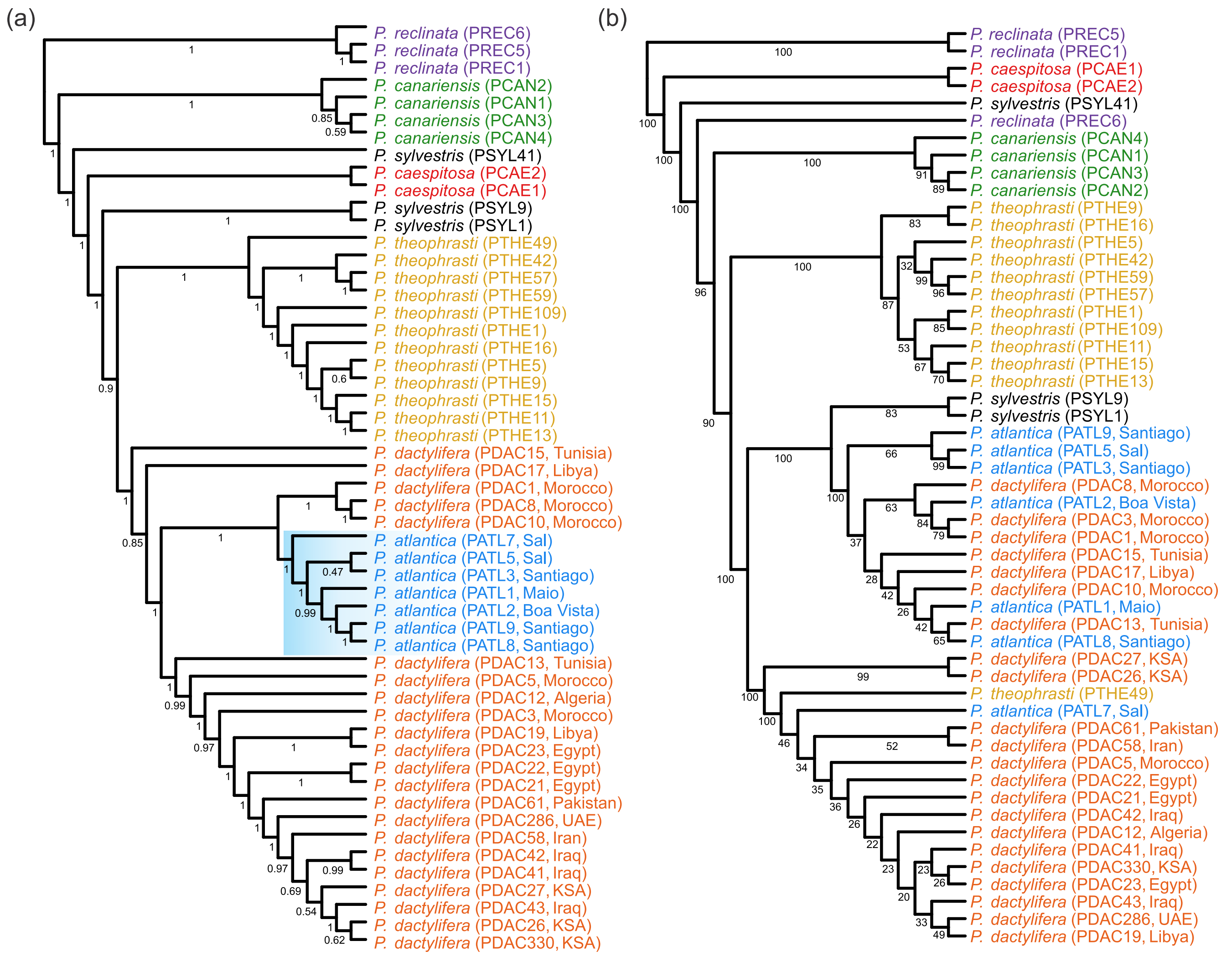

FIGURE 2 Nuclear and plastid phylogenetic trees of 54 western *Phoenix* accessions belonging to *Phoenix atlantica* (8 accessions), *P. dactylifera* (22 accessions), *P. sylvestris* (three accessions), *P. caespitosa* (three accessions), *P. reclinata* (three accessions), *P. canariensis* (four accessions) and *P. theophrasti* (12 accessions). (a) Nuclear phylogenetic tree obtained with wASTRAL based on 7709 gene trees generated from 10 kbp-long nuclear sequence alignments using IQ-TREE. Numbers below branches indicate their local posterior probabilities. (b) Plastid phylogenetic tree produced with IQ-TREE based on aligned whole plastome pseudohaplotypes. Numbers below branches indicate their ultrafast bootstrap percentages. A sample with extremely low coverage, PATL4, was excluded from the phylogenetic analyses. KSA, Kingdom of Saudi Arabia; UAE, United Arab Emirates.

and of *P. atlantica* seeds to the domesticated category was comparable in performance, with 68.2% and 63.3% of the seeds allocated to this category, respectively.

4 | DISCUSSION

This study investigates whether *P. atlantica* is the closest wild relative of the date palm or a feral derivative, analysing the type specimen at a molecular level for the first time. The inclusion of the type specimen of *P. atlantica* in our study has enabled us to certify the identity of other samples, including those used in previous studies that had alluded to a possible feral origin of this species (Flowers et al., 2019; Gros-Balthazard et al., 2017; Henderson et al., 2006). Our results, based on phylogenetic, population genetic, and seed morphology analyses, point to an origin of *P. atlantica* from domesticated date palm through a feralisation event, but also highlight a degree of differentiation since this event.

4.1 | *Phoenix atlantica* is a feral date palm

The nesting of *P. atlantica* within a clade of North African cultivated date palms in our phylogenetic trees supports a feral origin of this species. This is corroborated by our analysis of genetic diversity, which showed that the genetic variation of *P. atlantica* overlaps substantially with that of *P. dactylifera* from Northwestern Africa. In addition, both *P. atlantica* and African date palms have shared ancestry

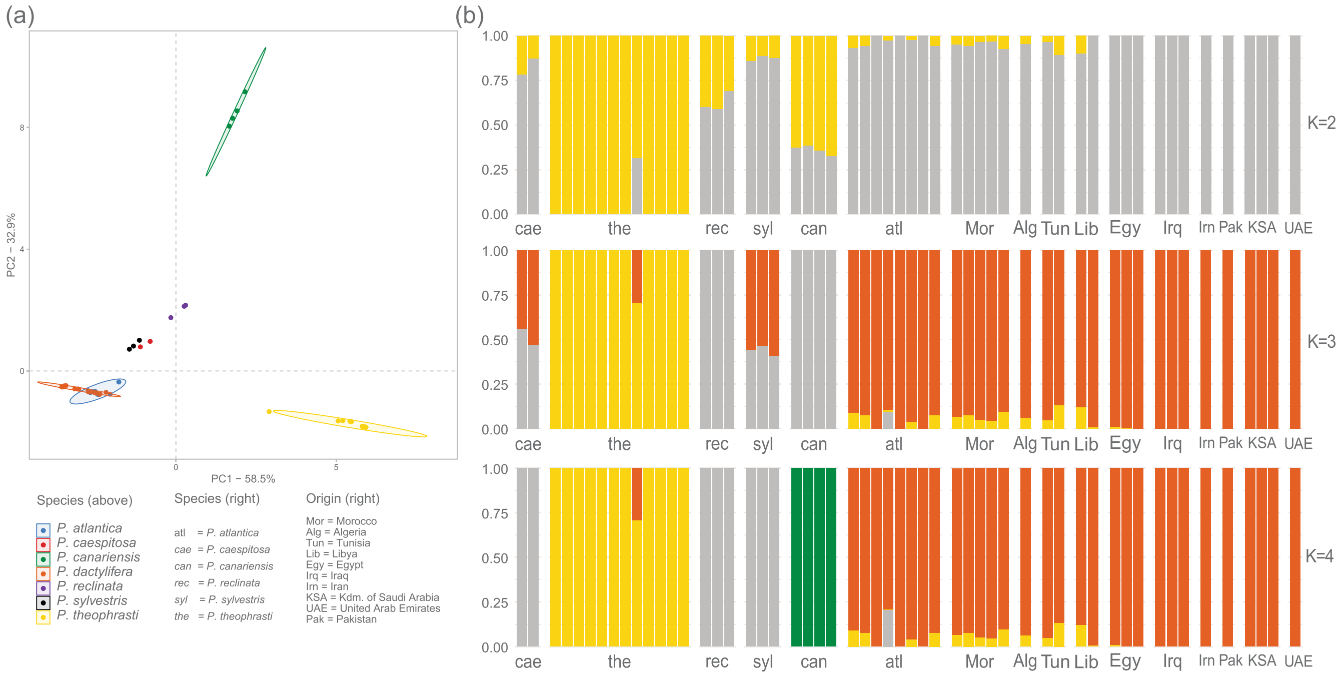

FIGURE 3 Principal component analysis (PCA) and admixture plot of 54 western *Phoenix* accessions belonging to *Phoenix atlantica* (8 accessions), *P. dactylifera* (22 accessions), *P. sylvestris* (three accessions), *P. caespitosa* (three accessions), *P. reclinata* (three accessions), *P. canariensis* (four accessions) and *P. theophrasti* (12 accessions). (a) PCA (variance explained in axis label) based on 15,308,583 nuclear genotype likelihood sites. (b) Admixture proportions based on genotype likelihoods for the same accessions with $K = 2$ (top), $K = 3$ (middle) and $K = 4$ (bottom).

with *P. theophrasti* (Figure 3b). This pattern is consistent with an origin of African *P. dactylifera* by hybridisation of West Asian *P. dactylifera* with *P. theophrasti* (Flowers et al., 2019). The presence of this shared ancestry in *P. atlantica*, despite *P. theophrasti* growing thousands of kilometres away (in Greece and Turkey), further supports the hypothesis that *P. atlantica* originated from African introgressed date palms. Our analysis of the plastid genome shows that *P. atlantica* shares the two chlorotypes (so-called 'oriental' and 'occidental') previously evidenced in date palms (Pintaud et al., 2013), with the occidental chlorotype being more frequent (Figure 2b), similar to what is found among West African date palms (Pintaud et al., 2013; Gros-Balthazard et al., 2017: Figure S4). Finally, the feralisation hypothesis is also supported by our morphometric results, which show that *P. atlantica* seeds are long and elongated, similar to those of domesticated date palms, rather than small and round as found in wild *Phoenix* species (Gros-Balthazard et al., 2016, 2017). Because *P. atlantica* grows in the wild, is not cultivated and is rarely exploited for its fruits (Chevalier, 1935; Henderson et al., 2003), the identification of seed traits associated with domestication can best be explained by a feral origin from domesticated date palms.

The date palm crop arrived in North Africa as the result of human-mediated diffusion from Asia, where *P. dactylifera* was first domesticated, and there is no evidence of wild *P. dactylifera* in Africa (Kaczmarek et al., 2024; Tengberg, 2012). Our analyses substantiate the clear diffusion axis of domesticated date palms, from West Asia to Northwestern Africa, observable in both the nuclear and plastid genomes, and consistent with an isolation-by-distance pattern (Flowers et al., 2019; Gros-Balthazard et al., 2017). *Phoenix atlantica* sits on the western most extent of West African date palms on this axis. The chronology of the date palms arrival in Africa is unclear: its presence is demonstrated by the first millennium BCE in Central Sahara (Kaczmarek et al., 2024; Van der Veen, 1992, 2010), but the first evidence in Morocco may solely be Medieval (Ros et al., 2024). The Cape Verde islands were first populated by the Portuguese in the 15th century (Castilla-Beltrán et al., 2021), so it is possible that the date palm arrived with them (Chevalier, 1935) and has since evolved in the wild into *P. atlantica*, although natural dispersal of *P. dactylifera* or *P. atlantica* from Africa cannot be ruled out.

4.2 | *Phoenix atlantica* evolves independently from other species

Our analyses of genetic diversity and structure confirmed the high similarity between *P. atlantica* and the date palm, but also revealed unique diversity, with some *P. atlantica* samples extending beyond the typical west–east diffusion axis seen in date palms (Figure 4a). This is corroborated by previous studies which, despite reporting a near absence of private diversity, still acknowledge the existence of alleles found only in *P. atlantica* (Gros-Balthazard et al., 2017; Sarmiento Cabello et al., 2024). This diversity is also visible in our nuclear phylogenetic tree, where *P. atlantica* formed a well-supported monophyletic group, indicating shared genetic variation among all Cape Verde palm accessions, encompassing specimens (including the species type)

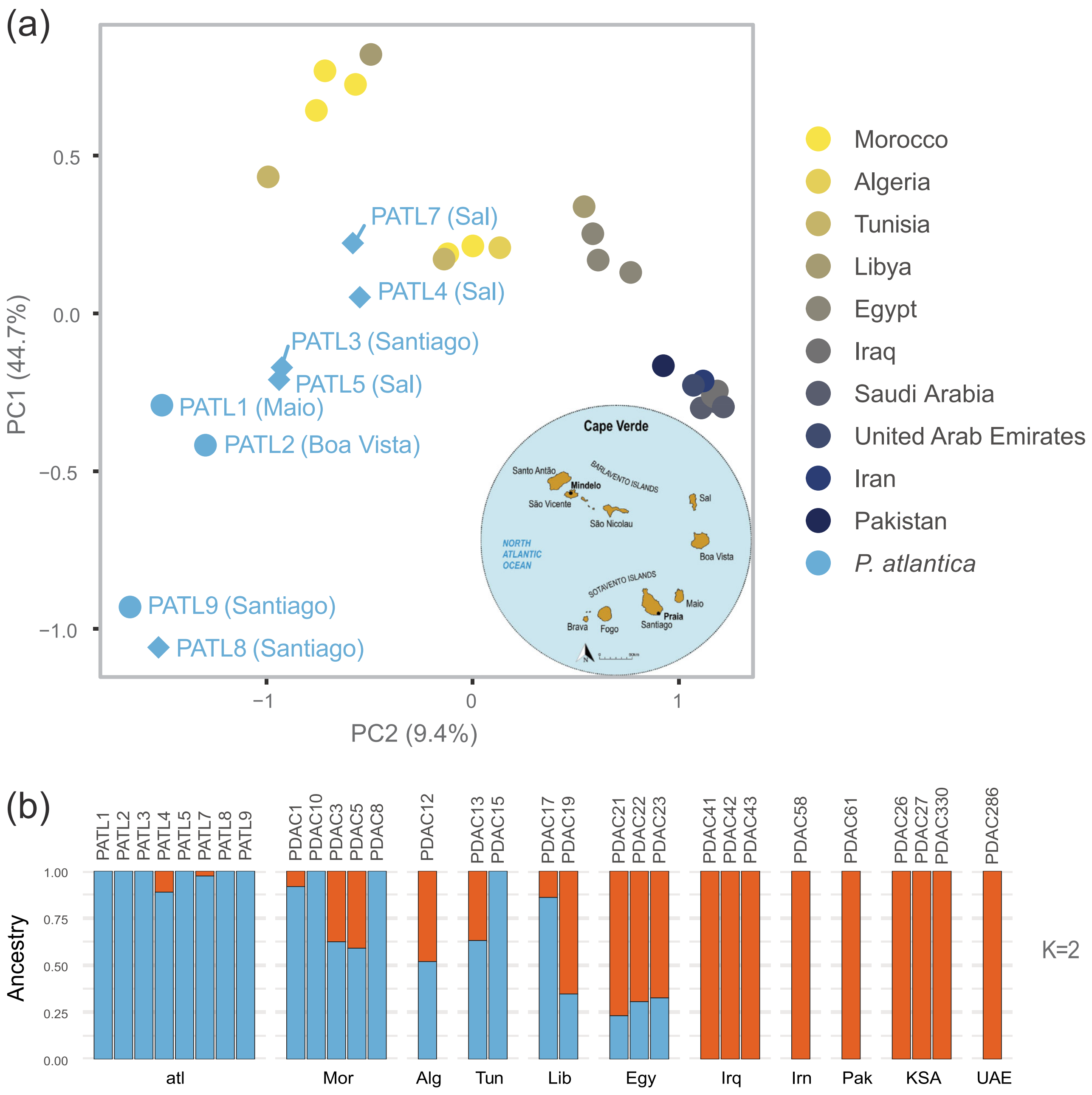

FIGURE 4 Principal component analysis (PCA) and admixture plot for *Phoenix atlantica* (eight accessions) and *P. dactylifera* (22 accessions). (A) PCA (variance explained in parentheses) inferred based on 6,884,062 nuclear genotype likelihood sites. Axes have been inverted to facilitate geographic interpretation. Diamonds mark herbarium specimen accessions. Inset map adapted from Pina-Araujo et al. (2018), available via license CC BY 4.0. (B) Admixture proportions based on genotype likelihoods of the same samples with $K = 2$.

sampled by Chevalier in 1934 as well as more recent specimens (Flowers et al., 2019; Gros-Balthazard et al., 2017; Henderson et al., 2003).

Our study unveiled an isolation-by-distance pattern across the Cape Verde islands, with increasing genetic and geographical distance from Moroccan date palms. This pattern supports an arrival of *P. atlantica* on the easternmost islands (possibly Sal) followed by a northeast-to-southwest dispersal through Boa Vista, Maio and then Santiago. According to Chevalier (1935) and Varela et al. (2022), *P. atlantica* appears spontaneously only on these eastern, driest islands of the archipelago (Santiago, Maio, Boa Vista and Sal), but other studies imply that it could be present in the islet of Santa Luzia (Nogueira, 1978) and possibly in Santo Antão and San Nicolau (Barbosa, 1961). In light of its feral origin, genome-wide studies with more samples from more islands, including *P. dactylifera* from Cape Verde, are needed to fully understand the spread of *P. atlantica* across Cape Verde and to fully characterise its genetic diversity.

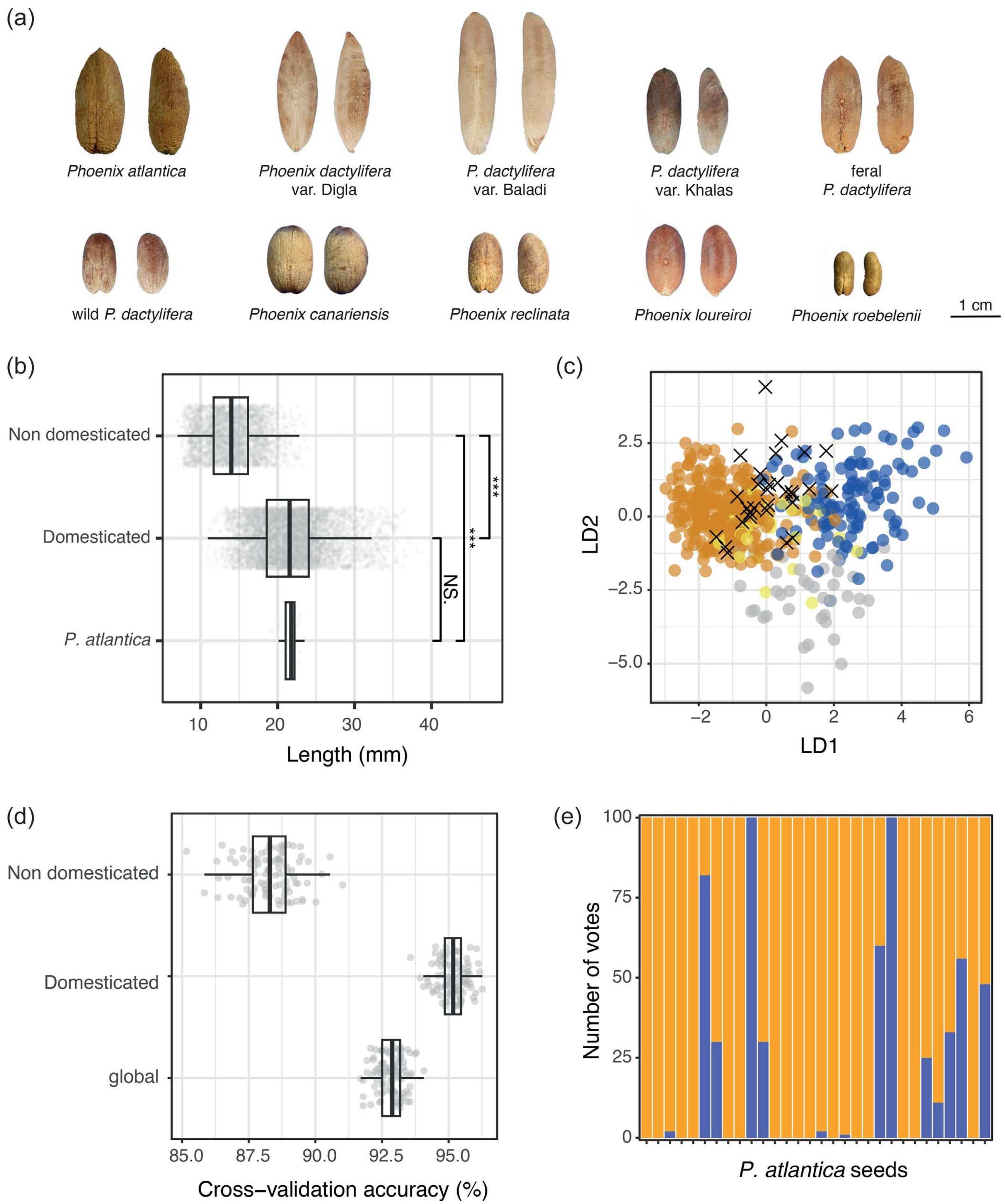

FIGURE 5 Legend on next page.

FIGURE 5 Morphometric analyses of *Phoenix atlantica* seeds in regard with other *Phoenix* species. (a) Example of seeds in the various groups/species. (b) Seed length variation in domesticated versus non-domesticated *Phoenix* and in *P. atlantica*. The lower and upper hinges of the boxplot correspond to the first and third quartiles (the 25th and 75th percentiles). Comparisons were performed with a t-test where asterisks denote statistical significance: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$ and 'NS' means no significance. (c) Linear discriminant analysis on four *Phoenix* groups: cultivated date palms (orange), feral date palms (yellow), wild date palms (grey), and wild relative species (blue). One dot corresponds to one accession (average of the seeds). *Phoenix atlantica* seeds are added to the model (black cross). (d) Model accuracy of a linear discriminant analysis differentiating domesticated (cultivated and feral date palms) from undomesticated *Phoenix* (wild date palms and wild relative species). (e) LDA allocation proportions of individual seeds to domesticated (orange) versus wild Phoenix (blue), keeping feral date palms as part of the training set.

4.3 | Taxonomic status and conservation of *P. atlantica*

This study examined three lines of evidence that can be used to evaluate the taxonomic status of *P. atlantica*, following (De Queiroz, 2007; Gros-Balthazard, et al., 2021): (i) fixed morphological differences, (ii) genetic composition and (iii) monophyly. We showed that there are no differences in seed morphology between *P. atlantica* and domesticated *P. dactylifera*, that both species overlap, yet differ in their genetic composition, that *P. atlantica* shows geographic structure within the Cape Verde archipelago and that it is monophyletic but nested inside *P. dactylifera*. Previous authors have also stressed that morphological differences used by A. Chevalier to distinguish *P. dactylifera* and *P. atlantica* are only subtle (leaf colour, petal shape) or resulting from the anthropic selection of *P. dactylifera* (fruit size, habit; summarised in Henderson et al. (2003)). Other lines of evidence that can be used to make a taxonomic decision include the existence of intrinsic, ecological or geographic reproductive isolation (De Queiroz, 2007). *Phoenix dactylifera* occurs sporadically on Cape Verde where it can occasionally hybridise with *P. atlantica* (Chevalier, 1935; Henderson et al., 2006) suggesting that the reproductive isolation between these species is mainly due to their distinct distribution ranges. Taken together, these lines of evidence show that *P. atlantica* is very similar yet evolving independently from the date palm, suggesting an incipient budding speciation process. Such mode of speciation, also known as progenitor-derivative, occurs when the progenitor species does not go extinct and instead becomes paraphyletic, with the derivative (monophyletic) species nested inside it (Anacker & Strauss, 2014; Gottlieb, 2004). Interestingly, in the case of *P. atlantica*, the progenitor is a domesticated species, highlighting the fact that domestication may not necessarily be an evolutionary dead-end.

Because *P. atlantica* appears to represent an early phase of the speciation process, at least three decisions could be taken regarding its taxonomic status. First, *P. atlantica* can be retained as a separate species according to De Queiroz's (2007) species concept, because it evolves independently from the date palm in two lines of evidence, monophyly and genetic composition. This would make the date palm paraphyletic, which may not be an issue given that many species are paraphyletic (Bateman, 2021; Chandler & Crisp, 1996; Rieseberg & Brouillet, 1994). However, it would remain problematic from a practical perspective, as *P. atlantica* is challenging to distinguish from *P. dactylifera* based on morphology (Henderson et al., 2003), even though people in Cape Verde refer to each species by a different name (Chevalier, 1935) likely based on morphology but also on where *P. atlantica* populations have been growing historically. A second option would be to synonymise *P. atlantica* under *P. dactylifera* so that the latter remains monophyletic. However, the fact that *P. atlantica* is a separately evolving lineage would not be reflected in the taxonomy, and the conservation of its distinct germplasm could be at a greater risk owing to the IUCN's greater focus on species. A third option would be to classify *P. atlantica* as a subspecies of *P. dactylifera*. This represents a potential compromise between the two other options, as it reflects both the integrity of the *P. atlantica* metapopulation and its relationship with *P. dactylifera*. It would allow *P. atlantica* to still be recognised as a distinct and accepted taxon that could be the focus of conservation efforts, which is crucial given that this species is currently assessed as Endangered due to habitat fragmentation, overexploitation and notably potential hybridisation with *P. dactylifera* (Duarte et al., 2017; Varela et al., 2022). However, the impact that ranking *P. atlantica* as a subspecies may have on its conservation would be uncertain until the subspecies is assessed for the Red List and local and international stakeholders take conservation action. As one of only two trees (in a broad sense) native to the archipelago, *P. atlantica* has strong cultural and ecological significance in the Republic of Cape Verde, where it is economically important as a source of animal forage, construction materials, handicrafts and shade (Henderson et al., 2003; Nogueira, 1978). Before making a taxonomic decision, careful assessment of the cultural and conservation implications will therefore be needed, in concertation with relevant stakeholders.

4.4 | Prospects for date palm improvement

Our study highlights *P. atlantica* as a potential source of genetic variation for crop improvement of the date palm. First, as a feral date palm, *P. atlantica* may have conserved some agriculturally desirable traits (e.g. large seeds, thus fruits and potentially appealing taste qualities of its fruits) and should have a higher genomic compatibility with the date palm than other wild relatives, owing to its more recent diversification from it (Zhang et al., 2016). This genomic compatibility could make breeding more straightforward and less prone to introducing deleterious alleles while preserving desired agricultural characteristics in the offspring (Pisias et al., 2022). Second, the independent evolution of *P. atlantica* may have generated genetic novelties that could be relevant for improving the date palm crop if they are underpinning traits of interest. Wu et al. (2021) reviewed how feralisation is not a mere reversal of domestication: different genomic loci are often involved. Indeed, previous studies on weedy rice (*Oryza sativa* f. *spontanea*) have demonstrated that feral lineages can evolve novel traits relevant for the improvement of the crop they derived from (e.g. Zhou et al., 2021). Although this possibility remains to be validated for *P. atlantica*, it is supported by the fact that other, possibly more recent, feral date palms around the oasis of Siwa (Egypt) show a remarkable and unexpected genetic diversity compared with surrounding crops (Gros-Balthazard et al., 2020). *Phoenix atlantica* can grow and reproduce in coastal areas without human intervention (pesticides, irrigation) and is expected to fare better than other endemic taxa of Cape Verde under changing climates (Varela et al., 2022). Consequently, traits of interest may include a greater resistance than cultivated date palms to drought, salinity and pests (e.g. the Bayoud disease), that is, three of the key challenges to the sustainability of date palm cultivation in the face of a warming climate and increasing globalisation (Dhawi & Aleidan, 2024; Hazzouri et al., 2020). Such traits could be introduced using CRISPR editing once phenotypic effects have been mapped to genomic regions, or with more traditional processes of crossing date palm cultivars with *P. atlantica* (and their successive hybrids) to create new varieties with the desired traits, in addition to elevated genetic diversity. Acknowledging the increasing pressures faced by oasis agriculture, that include water scarcity and competition for resources associated with growing urbanisation (Dhawi & Aleidan, 2024), breeding or genomic improvements in crops, along with sustainable resource use practices, is particularly timely for sustainable date palm production.

5 | CONCLUSION

By using an integrative approach to characterise the identity of *P. atlantica* with respect to that of the date palm and other *Phoenix*, we have illuminated the potential of the Cape Verde date palm as a reservoir of variation for date palm improvement. Genotype–phenotype association studies and breeding experiments are now required to make use of this potential. Our study also revealed *P. atlantica* as a model system to investigate natural evolution following the release of directional selection of anthropogenic origin. This adds to a series of date palm feralisation events that can be used to study the evolutionary consequences of feralisation through time and under different degrees of geographical isolation. Taken together, these highlight the need of conserving the unique germplasm of *P. atlantica* in the wild.

AUTHOR CONTRIBUTION

Jerónimo Cid Vian: Conceptualisation; methodology; software; data curation; formal analysis; writing—original draft; review and editing; visualisation; project administration. **Oscar A. Pérez-Escobar:** Conceptualisation; methodology; software; formal analysis; supervision; writing—original draft; review and editing. **William J. Baker:** Conceptualisation; writing—review and editing. **Vincent Bonhomme:** Methodology; software; formal analysis; data curation; writing—review and editing. **Sylvie Ferrand:** Investigation. **Sarah Ivorra:** Resources; software; data curation. **Thomas Kaczmarek:** Resources; writing—review and editing. **Clémence Pagnoux:** Resources; writing—review and editing. **Jean-Frédéric Terral:** Resources; methodology. **Nathan Wales:** Lab work; writing—review and editing. **Michael D. Purugganan:** Resources; funding acquisition. **Simon Creer:** Conceptualisation; supervision; writing—review and editing. **Alexander S. T. Papadopulos:** Conceptualisation; supervision; methodology; writing—original draft; review and editing. **Sidonie Bellot:** Conceptualisation; investigation; funding acquisition; project administration; visualisation; supervision; writing—original draft; review and editing. **Muriel Gros-Batlthazard:** Conceptualisation; methodology; software; formal analysis; resources; data curation; writing—original draft; review and editing; visualisation; supervision.

ACKNOWLEDGMENTS

We thank Marc Arnoux, Nizar Drou, and the New York University Abu Dhabi Bioinformatics Core for assistance with DNA sequencing and bioinformatic analyses. This work was supported in part through the NYU IT High Performance Computing resources, services, and staff expertise. The authors acknowledge Research Computing at the James Hutton Institute for providing computational resources and technical support for the 'UK's Crop Diversity Bioinformatics HPC' (BBSRC grants BB/S019669/1 and BB/X019683/1), use of which has contributed to the results reported within this paper. We acknowledge the support of the Supercomputing Wales project, which is part-funded by the European Regional Development Fund (ERDF) via Welsh Government. The assistance of the Bangor University eResearch team is also gratefully acknowledged. We also thank Laszlo Csiba and Tom Wells for help with DNA extractions and library preparation, Peter Moore from Infohackit for graphical design assistance and Dr. Cassia Bitencourt for the Portuguese social impact statement translation.

We thank the Institute of Evolutionary Science of Montpellier, the Millennium Seed Bank and the Royal Botanic Gardens of Kew, the Herbarium Palmarum of Florence, the seed reference collection of Florence, the Centro Studi Erbario Tropicale in Florence and the herbarium of Montpellier for access to their seed collections. We thank the Herbarium of the Muséum national d'Histoire naturelle (MNHN P), Paris, France, and in particular Marc Jeanson. The MNHN provides access to the collections in the framework of the RECOLNAT national Research Infrastructure. We acknowledge the Atelier Iconographie Scientifique, UAR 2700 2AD, BAOBAB facilities, supported by DIM-MAP Île-de-France, CNRS and MNHN, for their assistance with seed image acquisition and analysis. This work was supported by the Natural Environment Research Council (grant number NE/S007423/1), with the studentship of Jerónimo Cid Vian, the first author of this study through the Envision Doctoral Training Partnership. Envision DTP is a consortium consisting of Bangor University, British Geological Survey, Centre for Ecology and Hydrology, Lancaster University, Rothamsted Research and the University of Nottingham. The project was funded by the New York University Abu Dhabi Research Institute (1205H).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the supporting information of this article. Sources for all downloaded genomic data are stated in Dataset S1. Sequencing reads for the newly sequenced accessions can be found in the public sequence database GenBank under the BioProject ID PRJNA1177830, and BioSample IDs can be found in Dataset S1. Sources for all the morphometric data are stated in Dataset S3. The morphometric data can be found in Dataset S4.

ORCID

Jerónimo Cid-Vian https://orcid.org/0000-0003-0390-7698
William J. Baker https://orcid.org/0000-0001-6727-1831
Thomas Kaczmarek https://orcid.org/0000-0001-8526-4651
Michael D. Purugganan https://orcid.org/0000-0002-9197-4112

REFERENCES

Abramoff, M. D., Magelhaes, P. J., & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics International*, *11*(7), 36–42.

Allan, R. P., & Douville, H. (2024). An even drier future for the arid lands. *Proceedings of the National Academy of Sciences*, *121*(2), e2320840121. https://doi.org/10.1073/pnas.2320840121

Anacker, B. L., & Strauss, S. Y. (2014). The geography and ecology of plant speciation: Range overlap and niche divergence in sister species. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20132980. https://doi.org/10.1098/rspb.2013.2980

Barbosa, L. A. G. (1961). Subsídios para um dicionário utilitário e glossário dos nomes vernáculos das plantas de arquipélago de Cabo Verde. Garcia de Orta, Sér. Bot.

Barrow, S. C. (1998). A monograph of Phoenix L. (Palmae: Coryphoideae). *Kew Bulletin*, *53*(3), 513–575. https://doi.org/10.2307/4110478

Bateman, R. M. (2021). Challenges of applying monophyly in the phylogenetic shallows: Taxonomic reappraisal of the Dactylorhiza maculata group. *Kew Bulletin*, *76*(4), 675–704. https://doi.org/10.1007/s12225-021-09971-2

Bonhomme, V., Picq, S., Gaucherel, C., & Claude, J. (2014). Momocs: Outline analysis using R. *Journal of Statistical Software*, *56*(13), 1–24. https://doi.org/10.18637/jss.v056.i13

Brewer, G. E., Clarkson, J., Maurin, O., Zuntini, A., Barber, V., Bellot, S., Biggs, N., Cowan, R., Davies, N., Dodsworth, S., Edwards, S., Eiserhardt, W., Epitawalage, N., Frisby, S., Grall, A., Kersey, P., Pokorny, L., Leitch, I., Forest, F., & Baker, W. (2019). Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. *Frontiers in Plant Science*, *10*, 1102. https://doi.org/10.3389/fpls.2019.01102

Brochmann, C., Rustan, Ø. H., Lobin, W., & Kilian, N. (1997). The endemic vascular plants of the cape Verde islands, W Africa. *Sommerfeltia*, *24*(1), 1–363. https://doi.org/10.2478/som-1997-0001

Carr, P. H. (2018). What is climate change doing to us and for us? *Zygon*®, *53*(2), 443–461. https://doi.org/10.1111/zygo.12410

Castilla-Beltrán, A., de Nascimento, L., Fernández-Palacios, J.-M., Whittaker, R. J., Willis, K. J., Edwards, M., & Nogué, S. (2021). Anthropogenic transitions from forested to human-dominated landscapes in southern Macaronesia. *Proceedings of the National Academy of Sciences*, *118*(40), e2022215118. https://doi.org/10.1073/pnas.2022215118

Chandler, G., & Crisp, M. (1996). Paraphyletic species. *Telopea*, *6*, 813–844. https://doi.org/10.7751/telopea19963037

Chevalier, A. (1935). Les Iles du Cap Vert. Géographie, biogéographie, agriculture. Flore de l'Archipel. *Journal d'agriculture Traditionnelle et de Botanique appliquée*, *15*(170), 733–1090. https://doi.org/10.3406/jatba.1935.5553

Chevalier, A. (1952). Recherches sur les Phœnix africains. *Journal d'agriculture traditionnelle et de botanique appliquée*, *32*(355), 205–236. https://doi.org/10.3406/jatba.1952.6498

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*(6), 879–886. https://doi.org/10.1080/10635150701701083

Dempewolf, H., Baute, G., Anderson, J., Kilian, B., Smith, C., & Guarino, L. (2017). Past and future use of wild relatives in crop breeding. *Crop Science*, *57*(3), 1070–1082. https://doi.org/10.2135/cropsci2016.10.0885

Dhawi, F., & Aleidan, M. M. (2024). Oasis agriculture revitalization and carbon sequestration for climate-resilient communities. *Frontiers in Agronomy*, *6*, 1386671. https://doi.org/10.3389/fagro.2024.1386671

Duarte, M. C., Catarino, S., & Romeiras, M. (2017). Phoenix atlantica. The IUCN Red List of Threatened Species. IUCN Red List of Threatened Species. Accessed: 17 March 2024. https://www.iucnredlist.org/en

FAO. (2024). Accessed: 15 March 2024. https://www.fao.org/faostat/en/#data/QCL

Flowers, J. M., Hazzouri, K. M., Gros-Balthazard, M., Mo, Z., Koutroumpa, K., Perrakis, A., Ferrand, S., Khierallah, H. S. M., Fuller, D. Q., Aberlenc, F., Fournaraki, C., & Purugganan, M. D. (2019). Cross-species hybridization and the origin of nNorth African date palms. *Proceedings of the National Academy of Sciences*, *116*(5), 1651–1658. https://doi.org/10.1073/pnas.1817453116

Gottlieb, L. D. (2004). Rethinking classic examples of recent speciation in plants. *New Phytologist*, *161*(1), 71–82. https://doi.org/10.1046/j.1469-8137.2003.00922.x

Govaerts, R., & Dransfield, J. (2005). World checklist of palms. The Board of Trustees of the Royal Botanic Gardens, Kew. Accessed: 14 June 2024. https://www.abebooks.co.uk/9781842460849/World-Checklist-Palms-Govaerts-Dransfield-1842460846/plp

Gros-Balthazard, M., Baker, W. J., Leitch, I. J., Pellicer, J., Powell, R. F., & Bellot, S. (2021). Systematics and evolution of the genus Phoenix: Towards understanding date palm origins. In J. M. Al-Khayri, S. M. Jain, & D. V. Johnson (Eds.), *The date palm genome, Vol. 1: Phylogeny, biodiversity and mapping*. Compendium of plant genomes (pp. 29–54). Springer International Publishing. https://doi.org/10.1007/978-3-030-73746-7_2 Accessed: 15 March 2022

Gros-Balthazard, M., Battesti, V., Ivorra, S., Paradis, L., Aberlenc, F., Zango, O., Zehdi-Azouzi, S., Moussouni, S., Naqvi, S. A., Newton, C., & Terral, J. F. (2020). On the necessity of combining ethnobotany and genetics to assess agrobiodiversity and its evolution in crops: A case study on date palms (phoenix dactyliferaL.) in siwa oasis, Egypt. *Evolutionary Applications*, *13*(8), 1818–1840. https://doi.org/10.1111/eva.12930

Gros-Balthazard, M., Flowers, J. M., Hazzouri, K. M., Ferrand, S., Aberlenc, F., Sallon, S., & Purugganan, M. D. (2021). The genomes of ancient date palms germinated from 2,000 y old seeds. *Proceedings of the National Academy of Sciences*, *118*(19), e2025337118. https://doi.org/10.1073/pnas.2025337118

Gros-Balthazard, M., Galimberti, M., Kousathanas, A., Newton, C., Ivorra, S., Paradis, L., Vigouroux, Y., Carter, R., Tengberg, M., Battesti, V., Santoni, S., Falquet, L., Pintaud, J. C., Terral, J. F., & 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*(14), 2211–2218.e8. https://doi.org/10.1016/j.cub.2017.06.045

Gros-Balthazard, M., Newton, C., Ivorra, S., Pierre, M.-H., Pintaud, J.-C., & Terral, J.-F. (2016). The domestication syndrome in Phoenix dactylifera seeds: Toward the identification of wild date palm populations. *PLoS ONE*, *11*(3), e0152394. https://doi.org/10.1371/journal.pone.0152394

Hazzouri, K. M., Flowers, J. M., Nelson, D., Lemansour, A., Masmoudi, K., & Amiri, K. M. A. (2020). Prospects for the study and improvement of abiotic stress tolerance in date palms in the post-genomics era. *Frontiers in Plant Science*, *11*, 293. https://doi.org/10.3389/fpls.2020.00293

Hazzouri, K. M., Gros-Balthazard, M., Flowers, J. M., Copetti, D., Lemansour, A., Lebrun, M., Masmoudi, K., Ferrand, S., Dhar, M. I., Fresquez, Z. A., Rosas, U., Zhang, J., Talag, J., Lee, S., Kudrna, D., Powell, R. F., Leitch, I. J., Krueger, R. R., Wing, R. A., … Purugganan, M. D. (2019). Genome-wide association mapping of date palm fruit traits. *Nature Communications*, *10*(1), 4680. https://doi.org/10.1038/s41467-019-12604-9

Henderson, S., Gomes, I., Gomes, S., & Baker, W. (2003). Phoenix in the Cape Verde Islands. 47 10.

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*(2), 213–223. https://doi.org/10.1007/s10592-006-9128-7

Honsdorf, N., March, T. J., Berger, B., Tester, M., & Pillen, K. (2014). High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. *PLoS ONE*, *9*(5), e97047. https://doi.org/10.1371/journal.pone.0097047

Kaczmarek, T., Van der Veen, M., Ivorra, S., Mattingly, D., Terral, J.-F., & Gros-Balthazard, M. (2024). Origins and evolution of oasis agriculture in the Sahara: Evidence from morphometric analyses of archaeological date palm seeds. *The Holocene*, *34*, 353–365. https://doi.org/10.1177/09596836231211879

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Li, H. (2014). Aligning sequence reads, clone sequences and assembly con*gs with BWA-MEM. p. 0 Bytes. https://doi.org/10.6084/M9.FIGSHARE.963153.V1

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Mabry, M. E., Bagavathiannan, M. V., Bullock, J. M., Wang, H., Caicedo, A. L., Dabney, C. J., Drummond, E. B. M., Frawley, E., Gressel, J., Husband, B. C., Lawton-Rauh, A., Maggioni, L., Olsen, K. M., Pandolfo, C., Pires, J. C., Pisias, M. T., Razifard, H., Soltis, D. E., Soltis, P. S., … McAlvay, A. C. (2023). Building a feral future: Open questions in crop ferality. *Plants, People, Planet*, *5*(5), 635–649. https://doi.org/10.1002/ppp3.10367

Meisner, J., & Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics*, *210*(2), 719–731. https://doi.org/10.1534/genetics.118.301336

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Nogueira, I. (1978). *Plantas colhidas pelo Eng. 0 L. A. Grandvaux Barbosa no arquipelago de Cabo Verde. -VII. Spermatophyta (Cannaceae -Cyperaceae).* Garcia de Orta. Ser. Bot.

Pérez-Calle, V., Bellot, S., Kuhnhäuser, B. G., Pillon, Y., Forest, F., Leitch, I. J., & Baker, W. J. (2024). Phylogeny, biogeography and ecological diversification of New Caledonian palms (Arecaceae). *Annals of Botany*, *134*(1), 85–100. https://doi.org/10.1093/aob/mcae043

Pérez-Escobar, O. A., Tusso, S., Przelomska, N. A. S., Wu, S., Ryan, P., Nesbitt, M., Silber, M. V., Preick, M., Fei, Z., Hofreiter, M., Chomicki, G., & Renner, S. S. (2022). Genome sequencing of up to 6,000-year-old Citrullus seeds reveals use of a bitter-fleshed species prior to watermelon domestication. *Molecular Biology and Evolution*, *39*(8), msac168. https://doi.org/10.1093/molbev/msac168

Pina-Araujo, I., Spitz, N., Soares, C., Niel, C., Lago, B., & Gomes, S. (2018). Hepatitis B virus genotypes A1, A2 and E in Cape Verde: Unequal distribution through the islands and association with human flows. *PLoS ONE*, *13*, e0192595. https://doi.org/10.1371/journal.pone.0192595

Pintaud, J. C., Ludeña, B., Aberlenc-Bertossi, F., Zehdi, S., Gros-Balthazard, M., Ivorra, S., Terral, J. F., Newton, C., Tengberg, M., Abdoulkader, S., Daher, A., Nabil, M., Saro Hernández, I., González-Pérez, M. A., Sosa, P., Santoni, S., Moussouni, S., Si-Dehbi, F., & Bouguedoura, N. (2013). Biogeography of the date palm (Phoenix dactylifera L., Arecaceae): Insights on the origin and on the structure of modern diversity. *Acta Horticulturae*, *994*, 19–38. https://doi.org/10.17660/ActaHortic.2013.994.1

Pisias, M. T., Bakala, H. S., McAlvay, A. C., Mabry, M. E., Birchler, J. A., Yang, B., & Pires, J. C. (2022). Prospects of feral crop de novo redomestication. *Plant and Cell Physiology*, *63*(11), 1641–1653. https://doi.org/10.1093/pcp/pcac072

Rieseberg, L. H., & Brouillet, L. (1994). Are many plant species paraphyletic? *Taxon*, *43*(1), 21–32. https://doi.org/10.2307/1223457

Ros, J., Badri, F.-E., & Pelling, R. (2024). Agrobiodiversity and crop diffusion in Morocco from antiquity to the early Modern period (8th century bce–17th century ce): An archaeobotanical review. *Vegetation History and Archaeobotany*. https://doi.org/10.1007/s00334-024-01013-8

Sarmiento Cabello, S., Rodríguez-Rodríguez, P., Arbelo Ramírez, G., Naranjo-Cigala, A., Curbelo, L., da Graca Gomes, M. M., Brito, J., Aberlenc, F., Zehdi-Azouzi, S., & Sosa, P. A. (2024). A comparative genetic analysis of Phoenix atlantica in Cape Verde. *Plants*, *13*(16), 2209. https://doi.org/10.3390/plants13162209

Schubert, M., Ermini, L., Sarkissian, C. D., Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M. D., Fernández, R., Kircher, M., McCue, M., Willerslev, E., & Orlando, L. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nature Protocols*, *9*(5), 1056–1082. https://doi.org/10.1038/nprot.2014.063

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, *195*(3), 693–702. https://doi.org/10.1534/genetics.113.154138

Tanksley, S. D., Grandillo, S., Fulton, T. M., Zamir, D., Eshed, Y., Petiard, V., Lopez, J., & Beck-Bunn, T. (1996). Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative L. pimpinellifolium. *Theoretical and Applied Genetics*, *92*(2), 213–224. https://doi.org/10.1007/BF00223378

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

Terral, J.-F., Newton, C., Ivorra, S., Gros-Balthazard, M., de Morais, C. T., Picq, S., Tengberg, M., & 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*(5), 929–941. https://doi.org/10.1111/j.1365-2699.2011.02649.x

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, *108*(50), 20260–20264. https://doi.org/10.1073/pnas.1116437108

Van der Veen, M. (1992). Garamantian agriculture: The plant remains from Zinchecra, Fezzan. *Libyan Studies*, *23*, 7–39. https://doi.org/10.1017/S0263718900001722

Van der Veen, M. (2010). Plant Remains from Zinkekra - Early Evidence for Oasis agriculture. 489–519.

Varela, D., Romeiras, M. M., & Silva, L. (2022). Implications of climate change on the distribution and conservation of Cabo Verde endemic trees. *Global Ecology and Conservation*, *34*, e02025. https://doi.org/10.1016/j.gecco.2022.e02025

Wu, D., Lao, S., & Fan, L. (2021). De-domestication: An extension of crop evolution. *Trends in Plant Science*, *26*(6), 560–574. https://doi.org/10.1016/j.tplants.2021.02.003

Xiao, J., Grandillo, S., Ahn, S. N., McCouch, S. R., Tanksley, S. D., Li, J., & Yuan, L. (1996). Genes from wild rice improve yield. *Nature*, *384*(6606), 223–224. https://doi.org/10.1038/384223a0

Zhang, C., & Mirarab, S. (2022). Weighting by gene tree uncertainty improves accuracy of quartet-based species trees. *Molecular Biology and Evolution*, *39*(12), msac215. https://doi.org/10.1093/molbev/msac215

Zhang, H., Mittal, N., Leamy, L. J., Barazani, O., & Song, B. (2016). Back into the wild—Apply untapped genetic diversity of wild relatives for crop improvement. *Evolutionary Applications*, *10*(1), 5–24. https://doi.org/10.1111/eva.12434

Zhou, C., Feng, Y., Li, G., Wang, M., Jian, J., Wang, Y., Zhang, W., Song, Z., Li, L., Lu, B., & Yang, J. (2021). The new is old: Novel germination strategy evolved from standing genetic variation in weedy rice. *Frontiers in Plant Science*, *12*, 699464. https://doi.org/10.3389/fpls.2021.699464

Zsögön, A., Peres, L. E. P., Xiao, Y., Yan, J., & Fernie, A. R. (2022). Enhancing crop diversity for food security in the face of climate uncertainty. *The Plant Journal*, *109*(2), 402–414. https://doi.org/10.1111/tpj.15626

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Cid-Vian, J., Pérez-Escobar, O. A., Baker, W. J., Bonhomme, V., Ferrand, S., Ivorra, S., Kaczmarek, T., Pagnoux, C., Terral, J.-F., Wales, N., Purugganan, M. D., Creer, S., Papadopulos, A. S. T., Bellot, S., & Gros-Balthazard, M. (2025). Genomics and morphometry of herbarium specimens elucidate the origin of the Cape Verde date palm (*Phoenix atlantica* A.Chev.) and highlight its agronomic potential. *Plants, People, Planet*, 1–14. https://doi.org/10.1002/ppp3.10632