

*Genomic insights into population structure and conservation of wild olive (Olea europaea subsp. cuspidata) in Oman's Dhofar and Hajar Mountains*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Al Jabri, T., Halo, B. A., Culham, A. ORCID: <https://orcid.org/0000-0002-7440-0133> and Ellis, R. ORCID: <https://orcid.org/0000-0002-3695-6894> (2026) Genomic insights into population structure and conservation of wild olive (*Olea europaea* subsp. *cuspidata*) in Oman's Dhofar and Hajar Mountains. *Frontiers in Plant Science*, 17. 1813482. ISSN 1664-462X doi: 10.3389/fpls.2026.1813482 Available at <https://centaur.reading.ac.uk/129784/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.3389/fpls.2026.1813482>

Publisher: Frontiers

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in

the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

## **CentAUR**

Central Archive at the University of Reading

Reading's research outputs online



#### OPEN ACCESS

EDITED BY  
Moses C. Wambulwa,  
South Eastern Kenya University, Kenya

REVIEWED BY  
Nihel Oueslati,  
Higher Institute of Biotechnology of  
Monastir, Tunisia  
Djamila Madani,  
University of m'dila.algeria, Algeria

\*CORRESPONDENCE  
Boshra Ahmed Halo  
✉ haloboshra@yahoo.com

RECEIVED 19 February 2026  
REVISED 09 April 2026  
ACCEPTED 13 April 2026  
PUBLISHED 13 May 2026

#### CITATION

Al Jabri T, Halo BA, Culham A  
and Ellis R (2026) Genomic insights  
into population structure and  
conservation of wild olive (*Olea  
europaea* subsp. *cuspidata*) in  
Oman's Dhofar and Hajar Mountains.  
*Front. Plant Sci.* 17:1813482.  
doi: 10.3389/fpls.2026.1813482

#### COPYRIGHT

© 2026 Al Jabri, Halo, Culham and Ellis.  
This is an open-access article distributed  
under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#).  
The use, distribution or reproduction in  
other forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which does  
not comply with these terms.

# Genomic insights into population structure and conservation of wild olive (*Olea europaea* subsp. *cuspidata*) in Oman's Dhofar and Hajar Mountains

Thuraiya Al Jabri<sup>1,2,3</sup>, Boshra Ahmed Halo<sup>4\*</sup>, Alastair Culham<sup>2</sup>  
and Richard Ellis<sup>3</sup>

<sup>1</sup>Botany and Conservation Department, Oman Botanic Garden, Muscat, Oman, <sup>2</sup>School of Biological Sciences, University of Reading, Whiteknights, Reading, United Kingdom, <sup>3</sup>School of Agriculture, Policy and Development, University of Reading, Agriculture, Earley Gate, Reading, United Kingdom, <sup>4</sup>Department of Plant Sciences, College of Agricultural and Marine Sciences, Sultan Qaboos University, Al-Khoud, Muscat, Oman

**Background:** Wild olive (*Olea europaea* subsp. *cuspidata*), one of the six subspecies of the olive complex worldwide, is an ecologically and economically important tree in the mountains of Oman. It is vulnerable to various biotic and abiotic stresses. Information on its genetic diversity is needed to support the conservation of wild olive in Oman.

**Results:** A dual-genomic approach was employed, utilizing nuclear SNPs derived from Angiosperms353 target enrichment and plastid SNPs (cpSNPs) recovered from off-target reads. These markers were used to assess the genetic diversity and population structure of 48 wild olive accessions from the Dhofar, Eastern Hajar, and Western Hajar mountains of Oman. STRUCTURE, PCA, and DAPC analyses consistently revealed a clear north–south genetic split, with Dhofar populations forming a distinct and genetically isolated lineage, while populations from the Eastern and Western Hajar Mountains showed evidence of admixture and genetic connectivity. Five plastid haplotypes were identified, further supporting this phylogeographic structure. Analysis of molecular variance (AMOVA) indicated that most genetic variation occurred within populations (71.86%) rather than among populations (28.14%;  $\Phi_{ST} = 0.2814$ ), while regional divergence between northern and southern populations was higher ( $\Phi_{RT} = 0.528$ ). The overall fixation index ( $F_{ST} = 0.3093$ ) indicated moderate to high genetic differentiation, with the strongest differentiation observed between northern and southern populations.

**Conclusions:** These findings highlight the genetic uniqueness and isolation of Dhofar populations, which should be considered a conservation priority. The results provide essential guidance for *in situ* and *ex situ* conservation strategies, including germplasm preservation and habitat restoration.

#### KEYWORDS

genetic diversity, *Olea europaea* subsp. *cuspidata*, Oman, population structure, SNP markers, wild olive

## Introduction

*Olea europaea* subsp. *cuspidata* (Wall. & G. Don) Cif. (wild olive) is one of six recognized subspecies within the olive complex (Green et al., 1989). This subspecies is geographically and ecologically distinct from its Mediterranean relatives (Zohary, 1995), occurring across a broad range extending from southern and eastern Africa through the Arabian Peninsula to Iran, Pakistan, and India (Green, 2002). Its distribution may even reach the arid parts of Yunnan and Sichuan provinces in China (Green, 2002). Wild olive in Oman occurs at elevations between 900 and 2900 m (Ghazanfar, 2015) in the Western and Eastern Hajar Mountains in the north and the Dhofar Mountains in the south.

Wild olive is a monoecious, diploid species ( $2n = 46$ ) primarily pollinated by wind (McGregor, 1976) and to a lesser extent by insects (Canale and Loni, 2010). Its fruits are consumed by several bird species and small mammals, which contribute to seed dispersal (Cuneo and Leishman, 2006; Abiyu et al., 2015). Although clonal reproduction has not been recorded in this subspecies, it has been observed in related taxa such as *O. europaea* subsp. *laperrinei* (Baali-Cherif and Besnard, 2005). Wild olive provides important ecological and socio-economic benefits, offering shade, forage, and edible fruits for local communities (Ghazanfar, 2018).

In Oman, wild olive trees are often in poor condition, showing little to no new growth, and the species was previously classified as of Least Concern (LC) but is now approaching Near Threatened (NT) status (Ghazanfar, 2015). Nevertheless, it was not listed as a threatened species in Oman's Plant Red Data Book (Patzelt, 2015a). More recently, wild olive has been assessed as Vulnerable (VU) in the United Arab Emirates National Red List of Vascular Plants (Allen et al., 2021).

Wild olive typically inhabits isolated mountainous environments that are fragile and highly susceptible to various threats (Al-Charaabi et al., 2013; Patzelt, 2015b; Patzelt, 2015c). Observations in Oman indicate that anthropogenic pressures such as urban expansion, overgrazing, illegal wood harvesting, and the introduction of non-native species, along with natural stressors like pests, diseases, and climate change, have collectively contributed to habitat degradation (Al-Jabri et al., 2024). These factors are likely to accelerate genetic erosion by reducing population sizes and gene flow, as observed in *O. europaea* subsp. *laperrinei* populations of the Saharan mountains where natural regeneration has ceased (Baali-Cherif and Besnard, 2005; Anthelme et al., 2008).

**Abbreviations:** AMOVA, Analysis of Molecular Variance; cpDNA, Chloroplast DNA; cpSNPs, Plastid Single Nucleotide Polymorphisms; CTAB, Cetyltrimethylammonium Bromide; DAPC, Discriminant Analysis of Principal Components; DJQ, Jabal Qara; DJR, Jabal Qamar; DJS, Jabal Samhan; EJB, Jabal Bani Jabir; EJA, Jabal Abyad;  $F_{IS}$ , Inbreeding Coefficient;  $F_{ST}$ , Fixation Index; gVCF, Genomic Variant Call Format; Hap, Haplotype; He, Expected Heterozygosity; Hs, Genetic Diversity; NCBI, National Center for Biotechnology Information; nrDNA, Nuclear DNA; PCA, Principal Component Analysis; SNPs, Single Nucleotide Polymorphisms; TCS, Templeton-Crandall-Sing Network; VCF, Variant Call Format; WJH, Jabal Hatt; WJK, Jabal Akhdar; WJS, Jabal Shams.

Despite its ecological and cultural significance, information on the genetic diversity of wild olive in Oman remains limited. This knowledge gap restricts the understanding of its adaptive traits and hinders the development of effective conservation measures. Previous research on Omani wild olives has been geographically restricted to the Western Hajar Mountains (Habib et al., 2021), where substantial genetic variation was observed among populations. That study reported no clear evidence of population substructure, suggesting that pollen- and seed-mediated gene flow may have historically facilitated connectivity across the region.

Wild olive (*Olea europaea* subsp. *cuspidata*) plays a crucial ecological role in Omani mountain ecosystems, serving as a food source for wildlife and providing traditional medicinal and timber resources for local communities. Understanding its genetic diversity is therefore essential to evaluate its adaptive capacity and resilience to ongoing environmental change. The present study investigated the genetic diversity and population structure of wild olive populations across different regions of Oman using single nucleotide polymorphism (SNP) markers to provide a comprehensive assessment of natural genetic variation in this subspecies in support of the development of national conservation and management strategies.

## Methods

### Plant material

A total of 48 wild olive (*Olea europaea* subsp. *cuspidata*) leaf samples were collected from eight native populations distributed across three major mountain ranges in Oman - the Western Hajar Mountains, the Eastern Hajar Mountains, and the Dhofar Mountains - between June and August 2020, following standardized field sampling procedures. These eight populations were strategically chosen to cover the full latitudinal and altitudinal range of the species in Oman, ensuring representation of both major evolutionary lineages (North and South). Although the total number of individuals per population was limited due to logistical constraints, the genome-wide SNP dataset provided sufficient statistical power to accurately estimate genetic diversity and population structure, as demonstrated in previous studies where large SNP panels compensate for smaller sample sizes (Willing et al., 2012). A map showing the geographic locations of all eight sampled populations across Oman's mountain ranges is provided in Figure 1 to illustrate their distribution and spatial coverage. Within the Eastern Hajar Mountains, the Jabal Abyad (EJA) site included the Wadi Sareen Nature Reserve; accordingly, collections at this locality were conducted under a permit issued by the Office for Conservation of the Environment (Permit No. 17/SN 2020). All other sampling locations were outside designated protected areas and therefore did not require specific collection permits. Herbarium voucher specimens for all populations were collected and deposited at the University of Reading Herbarium (RNG), United Kingdom.

Leaf samples were collected from three populations in each of the Dhofar Mountains (Jabal Qara, Jabal Qamar, and Jabal Samhan) and the Western Hajar Mountains (Jabal Akhdar, Jabal Shams, and Jabal Hatt), and from two populations in the Eastern Hajar Mountains (Jabal Abyad and Jabal Bani Jabir).

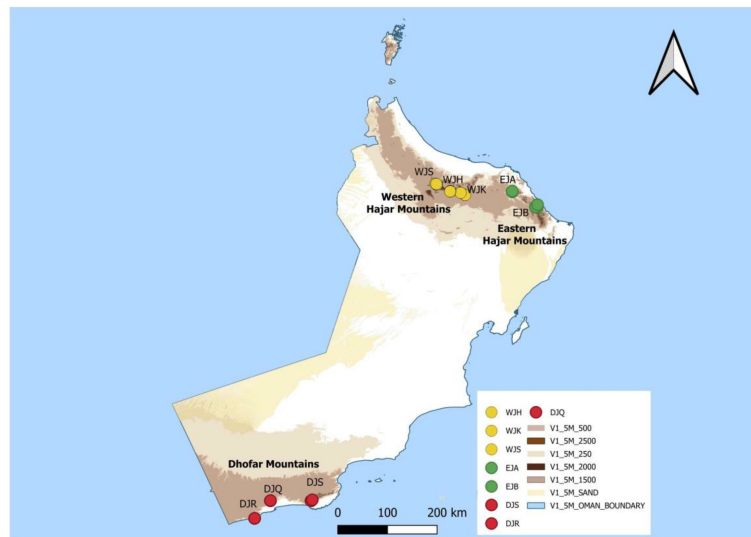


FIGURE 1

Geographic distribution of sampled wild olive (*Olea europaea* subsp. *cuspidata*) populations in Oman. Sampling was conducted across three major mountain ranges: the Western Hajar Mountains (WJH, WJK, WJS; yellow circles), the Eastern Hajar Mountains (EJA, EJB; green circles), and the Dhofar Mountains (DJQ, DJR, DJS; red circles). Elevation is indicated by shading, and the Oman national boundary is highlighted in light blue.

All collected samples consisted of fresh, healthy leaves which were immediately dried with silica gel in the field to preserve DNA integrity until extraction.

## DNA extraction

Total genomic DNA was extracted from all 48 samples. Dried leaf material (20–50 mg) was ground using a TissueLyser II (QIAGEN, Manchester, UK) for two cycles at 30 Hz for one minute each. DNA extraction followed the cetyltrimethylammonium bromide (CTAB) protocol of Doyle and Doyle (1987), with modifications based on Tel-Zur et al (Tel-Zur et al., 1999), and further optimized according to Yooprasert (2021) to improve DNA yield and purity. Specifically, three rinses in ice-cold sorbitol buffer (100 mM Tris-HCl pH 8, 5 mM EDTA pH 8, 0.35 M sorbitol) were performed prior to incubation in CTAB buffer, followed by the addition of 3 M potassium acetate (pH 5.5) before the chloroform–isoamyl alcohol extraction step. DNA quality and fragment integrity were assessed using 0.7% agarose gel electrophoresis stained with GelRed<sup>®</sup> (Biotium, Fremont, USA) and compared against HyperLadder<sup>™</sup> 1 kb (Bioline Reagents Ltd., London, UK) using either Tris-Borate-EDTA (TBE) or Sodium Borate (SB) buffers. DNA concentration and purity were quantified using a NanoDrop<sup>™</sup> Lite spectrophotometer (Fisher Scientific UK Ltd., Loughborough, UK) and a Qubit<sup>®</sup> fluorometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

## DNA library construction and sequencing

The DNA libraries were pooled and enriched using the Angiosperms353 target capture kit (myBaits<sup>®</sup> Target Sequence Capture Kit; Daicel Arbor Biosciences, Michigan, USA) following the manufacturer's protocol (version 4.0; <http://www.arborbiosci.com/mybaits-manual>). SNP markers were selected because they are bi-allelic, abundant throughout the genome, and provide high

resolution for detecting genetic variation. Additionally, SNP genotyping can be easily automated, allowing for efficient and reproducible large-scale analysis (Park et al., 2022).

## Data processing

Raw Illumina NovaSeq S4 paired-end (150 bp) reads were quality filtered and trimmed using Trimmomatic v0.39 (Bolger et al., 2014) by removing low-quality bases (Phred score < 28), discarding reads shorter than 50 bp, and eliminating sequencing adapters. The Angiosperms353 target genes were recovered from the trimmed reads using HybPiper v1.3 (Johnson and et al, 2016) and the 'mega353' target nucleotide file (McLay and et al, 2021). Briefly, HybPiper identifies potential matches between trimmed reads and target genes using BLASTX (Camacho et al., 2009), which allows for substantial phylogenetic divergence between the reference targets and sample reads. The program then sorts reads by gene and performs *de novo* assembly for each locus individually. Two sequence types were generated by HybPiper: coding sequences (CDS) corresponding to the target genes, and 'supercontigs' containing both exon sequences and flanking non-coding regions (i.e. introns and untranslated sequences). The final reference set was generated by selecting, for each gene, the longest supercontig recovered among all samples. The sample from Jabal Abyad (EA4) yielded the longest and most complete supercontigs, which were used as the reference for downstream analyses.

## Variant detection and filtering

To characterize genomic variation across different genome compartments, we generated two distinct SNP datasets: a nuclear dataset and a plastid dataset. Genetic variants among wild olive samples were identified using the Genome Analysis Toolkit (GATK v4) (McKenna et al., 2010). For the nuclear SNP dataset, the nuclear reference set

consisting of the longest Angiosperms353 supercontigs recovered from sample EA4 was used to detect genomic variation following the variantcall.sh pipeline (Slimp et al., 2021). Simultaneously, to identify plastid single nucleotide polymorphisms (cpSNPs) for haplotype and network analyses, the chloroplast genome of *Olea europaea* cultivar Bianchera (NC\_013707) was employed as a separate reference to capture “off-target” plastid reads.

The *variantcall.sh* pipeline combines Samtools and GATK to align paired-end reads to their respective reference sequences and call genotypes for each individual. The output comprised variant call format (VCF) files containing both single nucleotide polymorphisms (SNPs) and insertions/deletions (indels). Subsequently, individual genomic VCF (GVCF) files were merged into a single cohort VCF using the *GenotypesToPCA.sh* script (Johnson et al., 2021).

For the nuclear dataset, hard filtering was applied to ensure high-confidence variant calls by excluding indels and removing low-quality SNPs that did not meet the following thresholds: QualByDepth (QD) < 5.0, FisherStrand (FS) > 60.0, RMSMappingQuality (MQ) < 40.0, MappingQualityRankSum (MQRankSum) < -12.5, and ReadPosRankSum < -8.0 (Johnson et al., 2021). The resulting filtered nuclear dataset was used for downstream population structure and diversity analyses, while the cpSNP dataset was specifically utilized for the haplotype network construction (Figure 2).

The general workflow for SNP detection from target capture data is illustrated in Figure 3. Reference sequences were generated by selecting the longest supercontig from each gene for each individual. The process included read mapping, variant calling, and filtering steps as implemented through HybPiper and GATK pipelines, adapted from Slimp et al (Slimp et al., 2021).

## Population genetic and bioinformatics analysis

To assess the genetic health and connectivity of the eight *Olea europaea* populations, a suite of population genetic indices was calculated. The bioinformatics analysis was performed in R v4.3.2. The raw VCF file was processed and converted into a genomic data frame using the *vcfR* and *adegenet* packages.

Population genetic structure was evaluated using 56,410 high-quality SNPs. These were retained following a rigorous two-stage filtering process: first, GATK hard-filtering was applied to ensure high-confidence variant calls, followed by filtering for a Minor Allele Frequency (MAF) of 0.001 and a maximum missing data threshold of 30% per locus. To preserve the maximum phylogenetic signal and rare allelic variants critical for resolving the deep evolutionary divergence between the Eastern Hajar, Western Hajar, and Dhofar lineages, Linkage Disequilibrium (LD) pruning was not implemented, following the recommendations of Linck and Battey (2019). Within-population diversity was quantified by calculating the number of polymorphic sites, expected heterozygosity (Hs), and the inbreeding coefficient (F<sub>IS</sub>) for each Jabal using the *basic.stats* function in the *hierfstat* package. Nucleotide diversity (pi) was derived from the mean expected heterozygosity across all loci, serving as a robust proxy for diversity in this high-density SNP dataset.

To evaluate regional differentiation, the Global Fixation Index (F<sub>ST</sub>) was quantified using the Weir and Cockerham (1984) estimator to account for unequal sample sizes. Hierarchical Analysis of Molecular Variance (AMOVA) was performed at two scales: among mountain ranges (Table 1) and among specific geographical regions (Dhofar Mountains, Eastern Hajar Mountains, and Western Hajar Mountains; Table 2) to reconcile global F<sub>ST</sub> values with localized variance partitioning. Finally, the rate of Gene Migration (Nm) was estimated using the formula:

$$Nm = \frac{1 - F_{ST}}{4F_{ST}}$$

This calculation was used to determine whether contemporary gene flow is sufficient to counteract the effects of genetic drift (Nm > 1) or if the populations are evolving in isolation (Nm < 1).

## Population structure analysis and genetic diversity

The nuclear SNP VCF file was converted to STRUCTURE format using PGDSpider version 2.1.1.5 (Lischer and Excoffier, 2012). Population structure was inferred using STRUCTURE version 2.3.4 (Pritchard et al., 2000) under the admixture model. Analyses were conducted for K values ranging from 1 to 8, with five independent

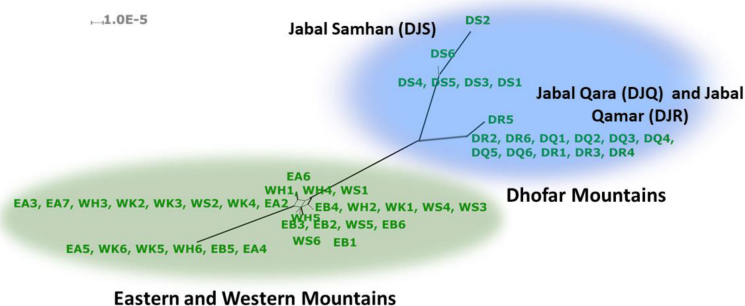
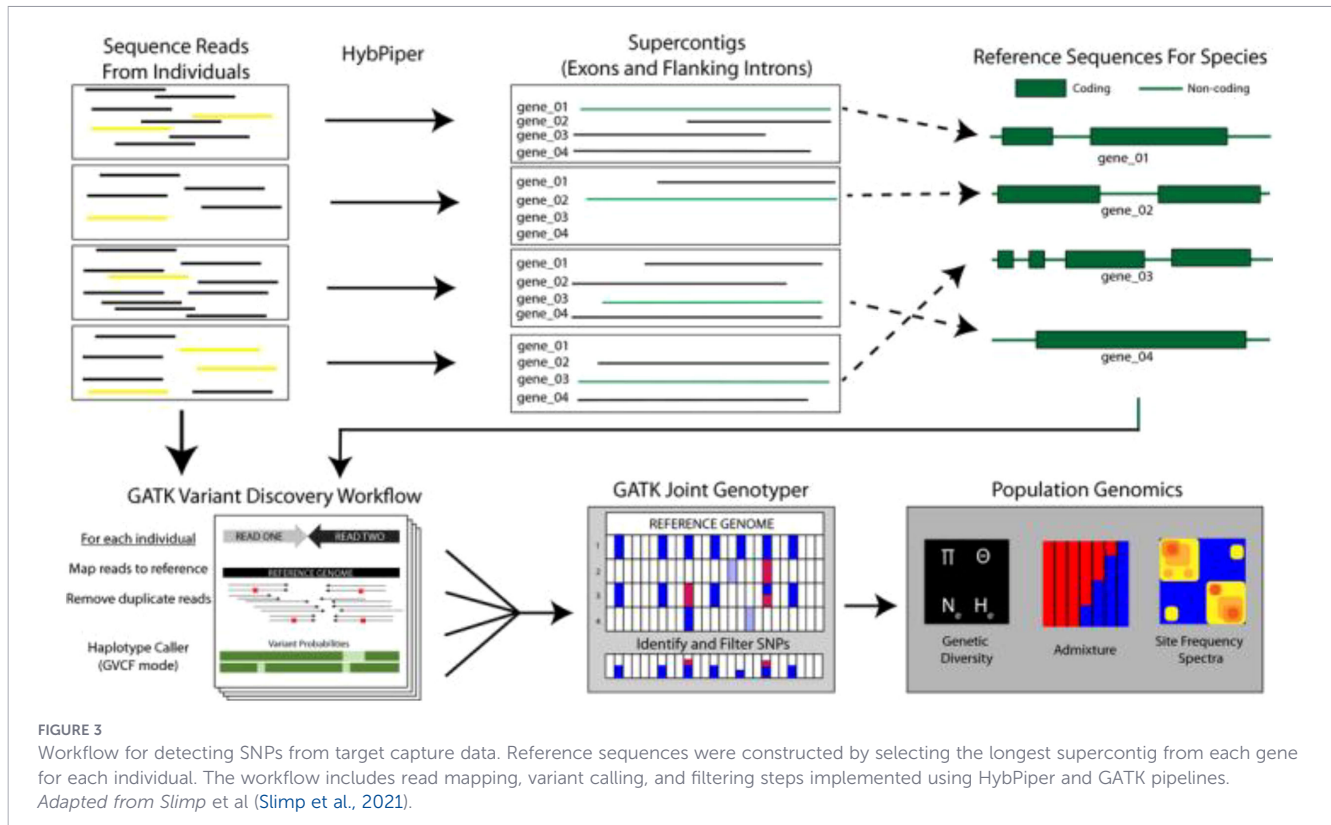


FIGURE 2

NeighborNet network generated using complete plastid sequences of 48 wild olive accessions from three mountain ranges (Eastern Hajar, Western Hajar and Dhofar) using SplitsTree version 4.17.0. Blue color indicates wild olive accessions from Dhofar mountains (Jabal Qara (DJQ), Jabal Qara (DJR) and Jabal Samhan (DJS)), green color shows the wild olive accessions from Western Hajar mountains (Jabal Hatt (WJH), Jabal Akhdar (WJK) and Jabal Shams (WJS)) and Eastern Hajar mountains (Jabal Abyad (EJA) and Jabal Bani Jabir (EJB)).



replicates for each K. Each run used a burn-in period of 100,000 generations followed by 100,000 Markov Chain Monte Carlo (MCMC) iterations.

Given the relatively large dataset and high number of SNP markers, computational efficiency was optimized by limiting the burn-in and iteration numbers to 10,000 and 50,000, respectively (Wang et al., 2013). The optimal K value was determined using the *ad hoc delta K* ( $\Delta K$ ) method (Evanno et al., 2005) implemented in the R package POPHELPER (version 2.3.1 (Francis, 2017)), which was also used to visualize STRUCTURE results.

Discriminant Analysis of Principal Components (DAPC) and Principal Component Analysis (PCA) were used to estimate the population structure of wild olives across Oman using the R package *adegenet* version 2.1.10 (Jombart, 2008). To visualize individual assignments, a composite stacked bar plot illustrating the probability of population membership on the Y-axis was generated.

The Analysis of Molecular Variance (AMOVA) and hierarchical partitioning of genetic variance were implemented using the R package *poppr* (Kamvar et al., 2014), with data integration handled via *vcfR* (Knaus and Grünwald, 2017). Genetic differentiation ( $F_{ST}$ ) between the

TABLE 1 Analysis of Molecular Variance (AMOVA) results for genetic variation among and within eight wild olive populations across Oman.

Source of variation	df	Sum of squares	Mean square	Sigma	Percentage variation	Phi	P value
Among populations	7	152,641.0	21,805.86	745.93	28.14	0.2814	0.001
Within populations	40	451,690.2	11,292.26	1,903.05	71.86		
Total	47	604,331.2	12,858.11	2,648.98	100		

Dhofar Mountains: DJQ, Jabal Qara; DJR, Jabal Qamar; DJS, Jabal Samhan; Eastern Hajar Mountains: EJA, Jabal Abyad; EJB, Jabal Bani Jabir; Western Hajar Mountains: WJH, Jabal Hatt; WJK, Jabal Akhdar; WJS, Jabal Shams. Sigma represents variance components, and *Phi* denotes population differentiation statistics.

TABLE 2 Analysis of Molecular Variance (AMOVA) for genetic variation among and within three mountain ranges: Dhofar, Eastern Hajar, and Western Hajar Mountains.

Source of variation	df	Sum of squares	Mean square	Sigma	Percentage variation	Phi	P value
Among mountains	2	275897.2	137948.588	8288.003	52.78798	0.5278798	0.001
Within mountains	45	333564.6	7412.546	7412.546	47.21202		
Total	47	609461.8	12967.271	15700.549	100.00000		

Sigma represents variance components, and *Phi* denotes population differentiation statistics.

eight Omani populations was quantified using the Weir and Cockerham unbiased estimator within the *StAMPP* package (Weir and Cockerham, 1984), which is specifically optimized for high-density SNP datasets. Estimates of within-population genetic variation, including expected heterozygosity ( $H_s$ ) and the inbreeding coefficient ( $F_{IS}$ ), were calculated using the *hierfstat* package (Purcell et al., 2007) to ensure statistical consistency across all genomic indices.

In addition, the correlation between genetic and geographic distances among the 48 Omani samples was assessed using a Mantel test (Smouse et al., 1986) with 100,000 permutations, based on chloroplast SNPs (cpDNA), as implemented in the R package *vegan* (Oksanen et al., 2018).

## Phylogenetic analysis

The filtered nuclear SNPs (nrDNA) VCF file for the Omani wild olive dataset was converted to a NEXUS format using the *vcf2phytip.py* script (Ortiz, 2019). An unrooted split tree was then constructed using the Neighbor-Net method implemented in *SplitsTree* version 4.17.0 (Huson and Bryant, 2006).

## Haplotype network analysis

The filtered plastid SNPs (cpDNA) VCF file for the Omani wild olive dataset was converted to a FASTA format using the *vcf2phytip.py* script (Ortiz, 2019). Relationships among haplotypes were inferred using the Templeton, Crandall, and Sing (TCS) network method (Templeton et al., 1992) implemented in Population Analysis with Reticulate Trees (*PopART* (Leigh et al., 2015); <https://popart.maths.otago.ac.nz/>; accessed 30 May 2022). Geographic coordinates of samples were used to visualize the spatial distribution of genetic clusters and haplotypes.

## Results

A total of 61,723 biallelic SNP markers were initially identified across the eight wild olive populations analyzed. After stringent

hard filtering, 56,410 high-quality biallelic SNPs were retained for selected downstream analyses.

## Population structure and genetic diversity analyses

The optimal  $\Delta K$  value was detected at  $K = 2$  (Figure 4), indicating the presence of two distinct genetic clusters among wild olive populations. The first cluster comprised accessions from the Dhofar Mountains (blue), whereas the second cluster included accessions from the Western and Eastern Hajar Mountains (red). The expected heterozygosity was slightly lower in Cluster 1 ( $H_e = 0.18$ ) compared to Cluster 2 ( $H_e = 0.21$ ). In contrast, genetic differentiation among populations was higher in Cluster 1 ( $F_{ST} = 0.49$ ) than in Cluster 2 ( $F_{ST} = 0.34$ ) (Table 3).

The Principal Component Analysis (PCA) revealed that the Western and Eastern Hajar Mountain populations clustered closely together, whereas the Dhofar Mountain populations were clearly distinct from both Hajar populations (Figures 5A, B). Samples from Jabal Abyad in the Eastern Hajar Mountains exhibited greater variability than those from any other location in Oman. Overall, the PCA results were consistent with the STRUCTURE analysis, supporting the presence of two major genetic clusters across Oman.

Discriminant Analysis of Principal Components (DAPC) grouped the wild olive accessions into two distinct clusters (Figure 6). Samples from the Eastern and Western Hajar Mountains were positioned on the right side of the DAPC axis, whereas the Dhofar samples were placed on the left side. Considerable overlap was observed between the Eastern and Western Hajar Mountain accessions, while the Dhofar populations were clearly separated from both Hajar groups. The population membership assignment inferred from the DAPC composite plot (Figure 7) supported the STRUCTURE and PCA results. All populations from the Western Hajar Mountains exhibited admixture with those from the Eastern Hajar Mountains (Figure 7A). Admixture was also detected within the Dhofar Mountain populations; however, no admixture was observed between Dhofar and the Eastern or Western Hajar populations (Figure 7B).

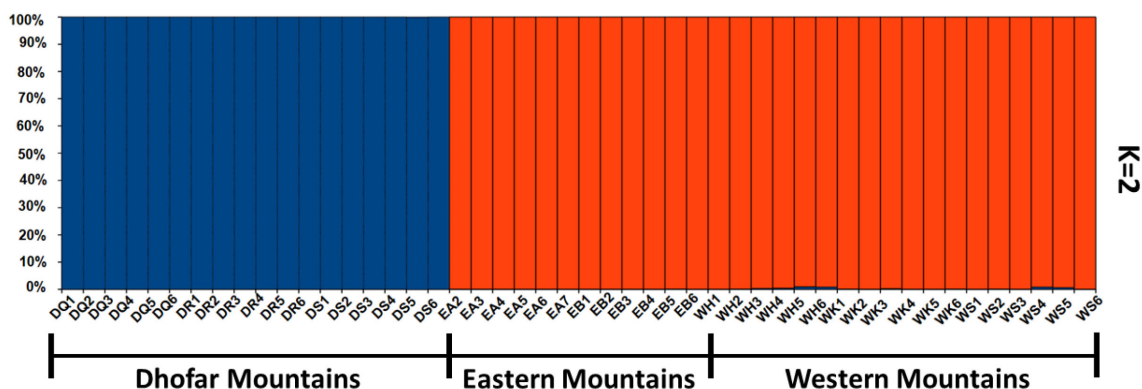


FIGURE 4

STRUCTURE bar plot of population assignment for 48 wild olive accessions from eight populations across Oman at  $K = 2$ , based on 61,723 SNPs using the longest supercontig of EA4 as the reference. Accessions represented in blue correspond to Cluster 1 (Dhofar Mountains), whereas those in red correspond to Cluster 2 (Eastern and Western Hajar Mountains). The y-axis represents the proportion of membership of each genotype to the inferred genetic clusters.

TABLE 3 Results of STRUCTURE analysis for 48 *Olea europaea* subsp. *cuspidata* accessions, showing the fixation index ( $F_{ST}$ ), expected heterozygosity ( $H_e$ ), and the number of genotypes assigned to each inferred cluster.

Population	Inferred clusters	Mean fixation index ( $F_{ST}$ )	Expected heterozygosity ( $H_e$ )	No. of genotypes
Cluster 1 (Dhofar)	0.376	0.489	0.181	18
Cluster 2 (Eastern and Western Hajar mountains)	0.624	0.339	0.210	30

Analysis of 56,410 high-quality nuclear SNPs across eight populations of *Olea europaea* subsp. *cuspidata* in Oman revealed a robust genetic structure with high genomic resolution. Within-population diversity was consistent across the mountain ranges, with Expected Heterozygosity ( $H_s$ ) and Nucleotide Diversity ( $\pi$ ) ranging from 0.0865 to 0.1082, and the number of polymorphic sites per population exceeding 9,800. Notably, all populations exhibited negative Inbreeding Coefficients ( $F_{IS}$ ), ranging from -0.1543 to -0.3104, indicating a significant excess of heterozygosity consistent with an outcrossing mating system.

Despite this internal genetic health, a high Global Fixation Index ( $F_{ST} = 0.3093$ ) and a low Gene Migration rate ( $Nm = 0.5584$ )

demonstrate restricted gene flow between the Northern Hajar and Southern Dhofar lineages. This regional isolation was further substantiated by the Hierarchical Analysis of Molecular Variance (AMOVA), which revealed that 28.14% of the total genetic variance is partitioned among mountain ranges ( $\Phi = 0.2814$ ,  $P = 0.001$ ; Table 1). Even when evaluated at a finer regional scale (Dhofar, Eastern Hajar, and Western Hajar), the genetic structure remained highly significant, confirming that the central desert acts as a formidable barrier to gene flow (Table 2).

These results confirm that while Omani wild olives are not currently suffering from inbreeding, the Southern populations (DJQ, DJR, and DJS) represent a genetically isolated and unique Evolutionarily Significant Unit (ESU) that must be managed independently to preserve the species' full evolutionary potential in the region (Table 4).

The pairwise  $F_{ST}$  values ranged from near zero (e.g., 0.00 to 0.05 among populations within the Hajar Mountains) to higher values between populations from Dhofar and the Hajar Mountains (0.38 to 0.43), indicating strong genetic differentiation (Figure 8). The highest genetic differentiation was observed between the southern Dhofar populations (DJQ, DJR, DJS) and both Eastern and Western Hajar populations, suggesting that the Dhofar populations are genetically isolated and may represent a distinct lineage or subspecies. Conversely, the lowest pairwise  $F_{ST}$  values occurred among populations within the Eastern and Western Hajar Mountains, reflecting ongoing gene flow and closer genetic connectivity in these regions.

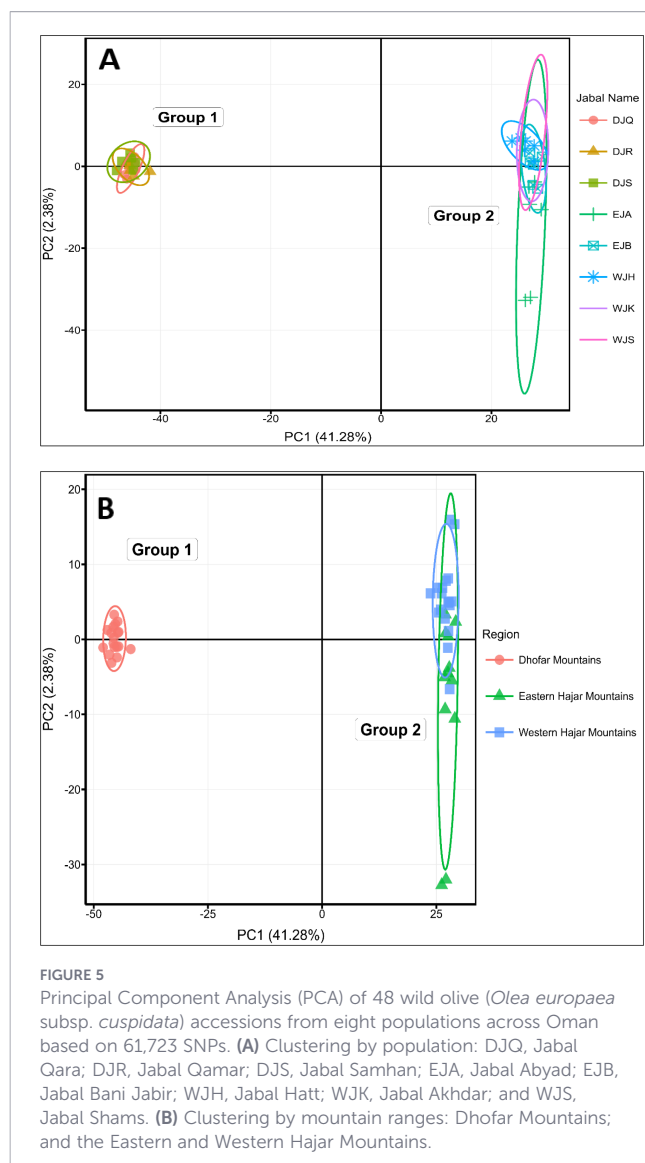
The Analysis of Molecular Variance (AMOVA; Table 1) confirmed significant genetic differentiation among the eight populations ( $\Phi_{ST} = 0.2814$ ,  $P = 0.001$ ), with the majority of genetic variation found within populations (71.86%) compared to (28.14%) among populations. This pattern supports the presence of population structure while indicating substantial genetic diversity is maintained within populations.

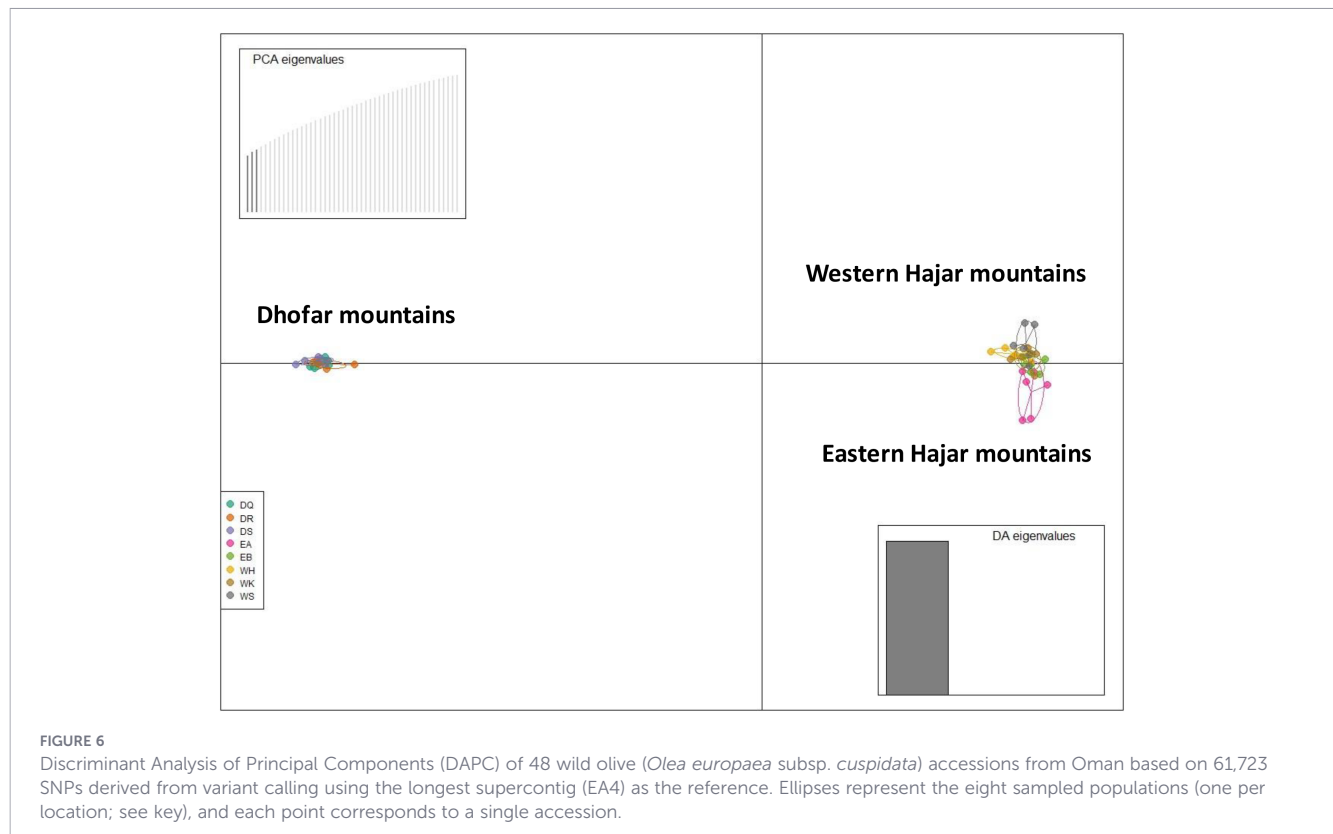
## Correlation between genetic and geographic distances (Mantel test)

The Mantel test based on complete chloroplast SNPs revealed a strong and significant positive correlation between genetic distance ( $F_{ST}/(1 - F_{ST})$ ) and geographic distance among wild olive populations across Oman ( $r = 0.6215$ ,  $P = 0.00001$ ; Figure 9). This indicates a clear isolation-by-distance pattern within the Omani populations.

## Phylogenetic analysis

The NeighborNet network analysis grouped the wild olive accessions from Oman into two distinct clusters. Cluster I comprised accessions from the Eastern and Western Hajar Mountains,





while Cluster II consisted of accessions from the Dhofar Mountains (Figure 10). This clear separation indicates a strong genetic differentiation between the Hajar and Dhofar mountain populations, supporting the results obtained from the STRUCTURE, PCA, and DAPC analyses.

## Haplotype analysis

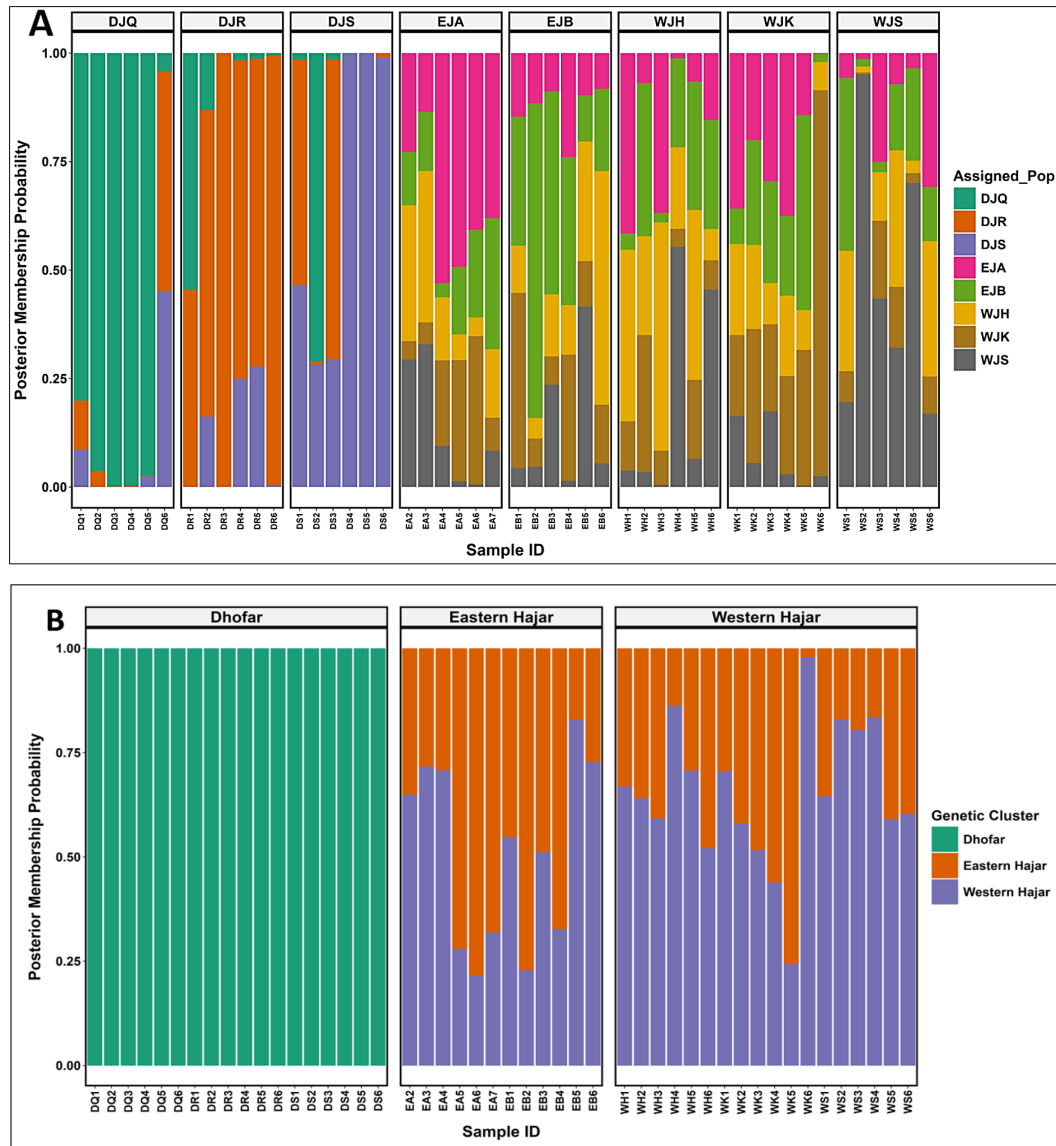
Five chloroplast haplotypes were identified among the eight wild olive populations across Oman (Figure 11). Three distinct haplotypes were detected in the Dhofar mountain populations. Jabal Qara (DJQ) and Jabal Qamar (DJR) shared the same haplotype (Hap\_1), while one accession from DJR represented a different haplotype (Hap\_2), and all accessions from Jabal Samhan (DJS) formed a unique haplotype (Hap\_3). Haplotype 4 (Hap\_4) comprised five accessions from populations in the Eastern and Western Hajar Mountains, whereas one accession from Jabal Shams (WJS) formed a separate haplotype (Hap\_5). The geographic locations of the cpDNA haplotypes and the median-joining network illustrating their relationships are presented in Figure 12.

## Discussion

This study provides a comprehensive assessment of the genetic diversity and population structure of wild olive (*Olea europaea* subsp. *cuspidata*) populations across Oman using single nucleotide polymorphism (SNP) markers generated from target capture sequencing with the Angiosperms353 probe set. SNPs are highly

stable, abundant, and widely distributed throughout the genome, making them particularly powerful for detecting fine-scale genetic variation. Their use in this study offers valuable insights into the genetic makeup of the diversity within wild olive in Oman and provides a foundation for future conservation programs there. Such genomic information can facilitate the development and characterization of core collections based on adaptive genetic traits, which in turn will support the regeneration and sustainable management of *Olea* germplasm resources (Omire et al., 2022).

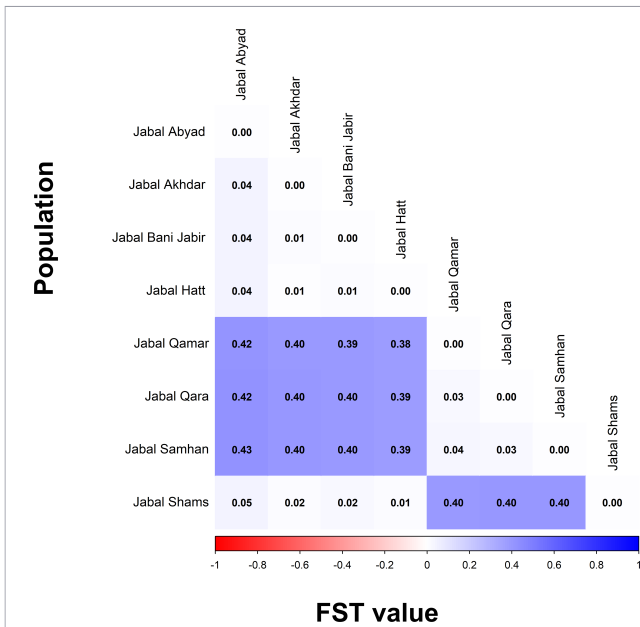
The results of the STRUCTURE, PCA, and DAPC analyses consistently identified two genetically distinct groups of wild olive samples across Oman. Accessions from the Dhofar Mountains in southern Oman formed Cluster I, whereas Cluster II comprised all accessions from the Eastern and Western Hajar Mountains. Evidence of admixture among individuals from the Eastern and Western Hajar populations indicates mixed ancestry and suggests some degree of genetic connectivity between these two regions. This pattern was also supported by the Neighbor-Net phylogenetic network, which showed overlapping relationships among accessions from both Hajar ranges. The close geographic proximity between the Eastern and Western Hajar Mountains (approximately 170 km apart) likely facilitates some level of gene flow (Besnard et al., 2007b). Olive pollen is known to disperse over long distances, even at regional scales of 200–300 km, as demonstrated by Fernández-Rodríguez et al (Fernández-Rodríguez and et al, 2014), in their study of airborne *Olea* pollen in the southwestern Iberian Peninsula. Although long-distance pollen dispersal has been documented in *Olea* species, such events are rare and unlikely to maintain effective contemporary gene flow between the Hajar and Dhofar populations.



**FIGURE 7** Composite DAPC plots of 48 wild olive (*Olea europaea* subsp. *cuspidata*) accessions from Oman based on SNPs derived from the longest supercontig reference (EA4), showing mixed ancestry between the Eastern and Western Hajar Mountains. **(A)** Composite plot of wild olive accessions from the eight sampled populations across Oman (DJQ, DJR, DJS, WJH, WJK, WJS, EJA, and EJB). **(B)** Composite plot showing population clustering among the three major mountain ranges (Western Hajar, Eastern Hajar, and Dhofar Mountains).

**TABLE 4** Mean genetic diversity indices for wild olive (*Olea europaea* subsp. *cuspidata*) accessions from eight populations across Oman based on nuclear SNPs.

Genetic index	DJQ	DJR	DJS	EJA	EJB	WJH	WJK	WJS	Overall
No. of Polymorphic Sites	9,814	9,935	10,166	11,572	14,767	14,812	14,275	13,764	56,410
Genetic Diversity (Hs)	0.0865	0.0895	0.0881	0.0925	0.1055	0.1082	0.1031	0.1020	0.0970
Inbreeding Coeff. (F <sub>IS</sub> )	-0.2922	-0.2718	-0.3104	-0.1698	-0.1543	-0.1662	-0.1564	-0.1554	-0.2100
Nucleotide Diversity (Pi)	0.0865	0.0895	0.0881	0.0925	0.1055	0.1082	0.1031	0.1020	0.0970
Fixation Index (F <sub>ST</sub> )	—	—	—	—	—	—	—	—	0.3093
Gene Migration (Nm)	—	—	—	—	—	—	—	—	0.5584



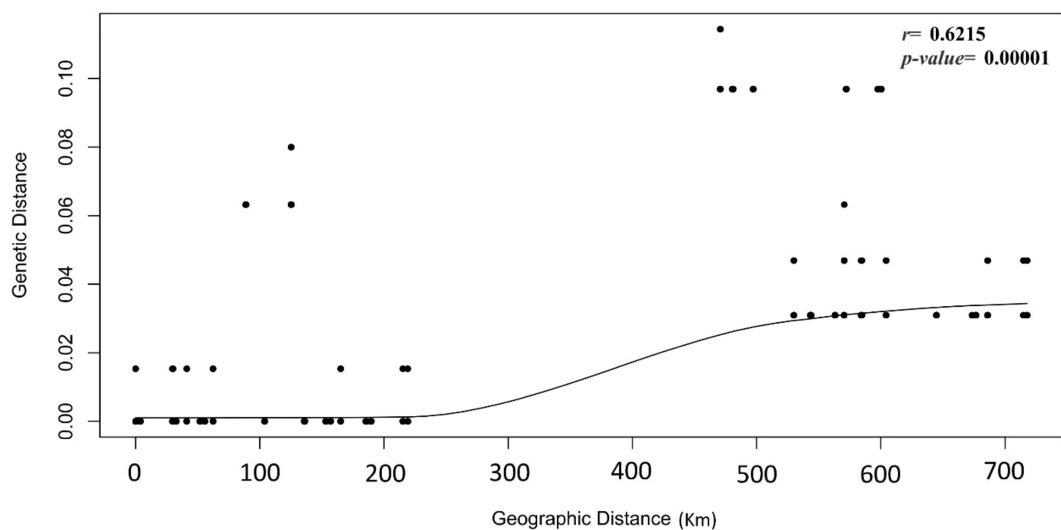
**FIGURE 8**  
Heatmap of pairwise  $F_{ST}$  values (Weir and Cockerham, 1984) among eight populations of wild olive (*Olea europaea* subsp. *cuspidata*) in Oman, calculated from 56,410 nuclear SNPs. The color scale represents the degree of genetic differentiation, ranging from low (red;  $F_{ST} < 0.05$ ) to high (blue;  $F_{ST} > 0.40$ ). Numbers within the cells indicate the specific pairwise  $F_{ST}$  estimates. A prominent genetic break is visible between the southern Dhofar Mountains (DJQ, DJR, DJS) and the northern Hajar Mountains (EJA, EJB, WJH, WJK, WJS), while populations within the Hajar range exhibit high genetic connectivity and low differentiation.

In contrast, the wild olive populations in the Dhofar Mountains were genetically isolated from those in the Hajar Mountains (Figures 4–12). This strong differentiation can be attributed to the vast arid desert -part of the Rub’ al Khali (Empty Quarter)- that separates southern Oman from the northern mountain ranges which is a major biogeographical barrier. Beyond geographic distance alone, multiple

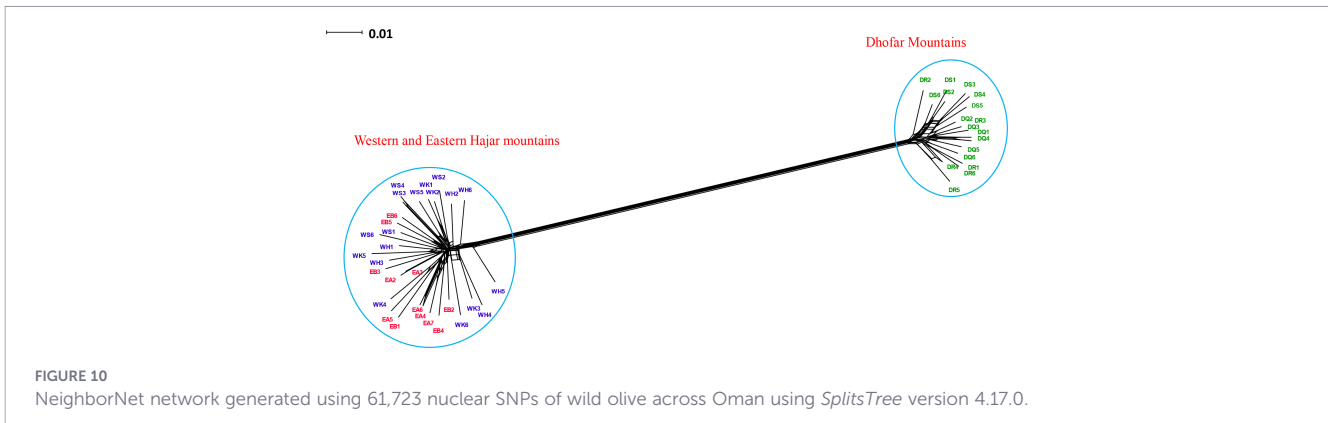
environmental factors likely contribute to the observed north–south differentiation. The Rub’ al Khali desert not only acts as a physical barrier but also represents an extreme ecological gradient limiting both pollen and seed dispersal. In addition, differences in climatic conditions between the monsoon-influenced Dhofar region and the more arid northern mountains may further reinforce genetic divergence through local adaptation. Potential anthropogenic pressures, including land use changes and habitat fragmentation, may have also contributed to reducing connectivity among populations.

A similar pattern was reported by Besnard et al (Besnard et al., 2021). in North Africa, where the Saharan Desert limited gene flow between *O. europaea* subsp. *laperrinei* populations from Niger and Algeria, leading to significant genetic divergence. Comparable patterns of strong regional differentiation have also been documented in wild olive populations across East Africa and the Arabian Peninsula, suggesting that historical isolation and environmental heterogeneity have played a key role in shaping the evolutionary history of this subspecies (Kassa et al., 2019; Habib et al., 2021). Accordingly, within Oman, the wild olive populations of the Dhofar Mountains appear to be geographically and genetically isolated, with reduced gene flow both within and beyond this mountain range. This pattern is further supported by recent genomic evidence from Oman, which revealed clear genetic differentiation between wild and cultivated olive populations, highlighting strong genetic structuring in this region (Al-Yahyai et al., 2026). Such limited connectivity is consistent with previous studies showing reduced gene flow in fragmented and isolated populations compared to those that are spatially connected (Dzialuk et al., 2014; Serrote et al., 2022).

The deep genetic divergence observed at the regional level ( $\Phi_{IRT} = 0.528$ ) between the Hajar and Dhofar lineages suggests that the current population structure reflects historical habitat fragmentation rather than simple contemporary isolation by distance (IBD). While long-distance pollen dispersal has been documented in *Olea* species, such events are likely rare and insufficient to maintain effective gene flow across the vast Omani interior.



**FIGURE 9**  
Mantel test showing the correlation between geographic and genetic distances among 48 wild olive (*Olea europaea* subsp. *cuspidata*) accessions across Oman based on chloroplast SNPs ( $r = 0.6215$ ,  $P = 0.00001$ ).



Palaeoenvironmental reconstructions for the Southern Arabian Peninsula (Preston et al., 2015) indicate a significant climatic shift toward aridity and increased regional dune emplacement between 5.9 and 5.3 ka cal. BP. This transition likely eliminated the lowland ‘corridors’ that once facilitated connectivity, effectively restricting *Olea europaea* subsp. *cuspidata* to high-altitude mountain refugia (Besnard et al., 2007b). Within these isolated ‘mountain islands,’ the cessation of migration allowed genetic drift to become the primary driver of population differentiation, shaping the distinct evolutionary lineages observed in northern and southern Oman today. These findings support a scenario in which historical connectivity was followed by fragmentation and long-term isolation, rather than ongoing gene flow between regions.

The results from PCA and DAPC analyses provided a slightly different pattern from the STRUCTURE analysis, revealing additional sub-structuring within the Omani wild olive populations. While

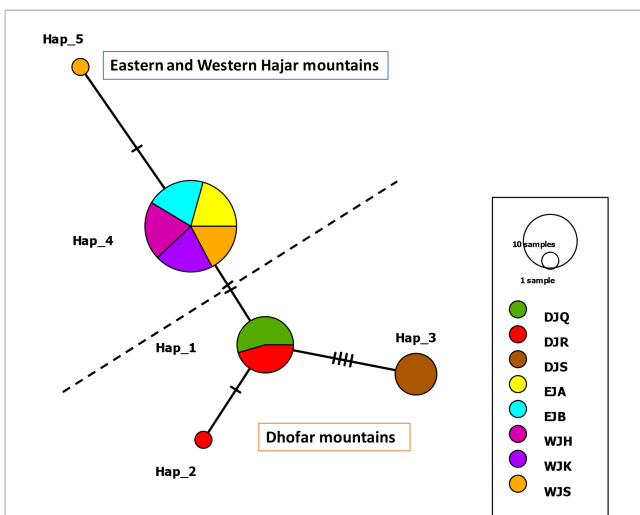
STRUCTURE identified two main clusters corresponding to the Dhofar and Hajar mountain ranges, both PCA and DAPC suggested further genetic differentiation within the Hajar populations, indicating possible subgroups between the Eastern and Western Hajar ranges. Evidence of admixture was detected between individuals from these two regions, suggesting that genetic connectivity persists across the short distance separating them, consistent with patterns observed in wind-pollinated species (Buschbom et al., 2011; Chludil et al., 2025).

The fixation index ( $F_{ST}$ ) estimates the degree of genetic differentiation among populations. In this study, the overall  $F_{ST}$  value for wild olive populations within Oman was 0.3093, indicating moderate to high genetic differentiation among populations. According to Nassiry et al (Nassiry et al., 2009),  $F_{ST}$  values of 0–0.05 indicate low differentiation, 0.05–0.15 moderate, and values above 0.15 represent very high genetic differentiation. Consistent with these thresholds, differentiation was low within the Eastern and Western Hajar Mountains ( $F_{ST} = 0.00–0.05$ ), but markedly higher between Dhofar and northern populations ( $F_{ST} = 0.38–0.43$ ), confirming strong regional genetic structuring.

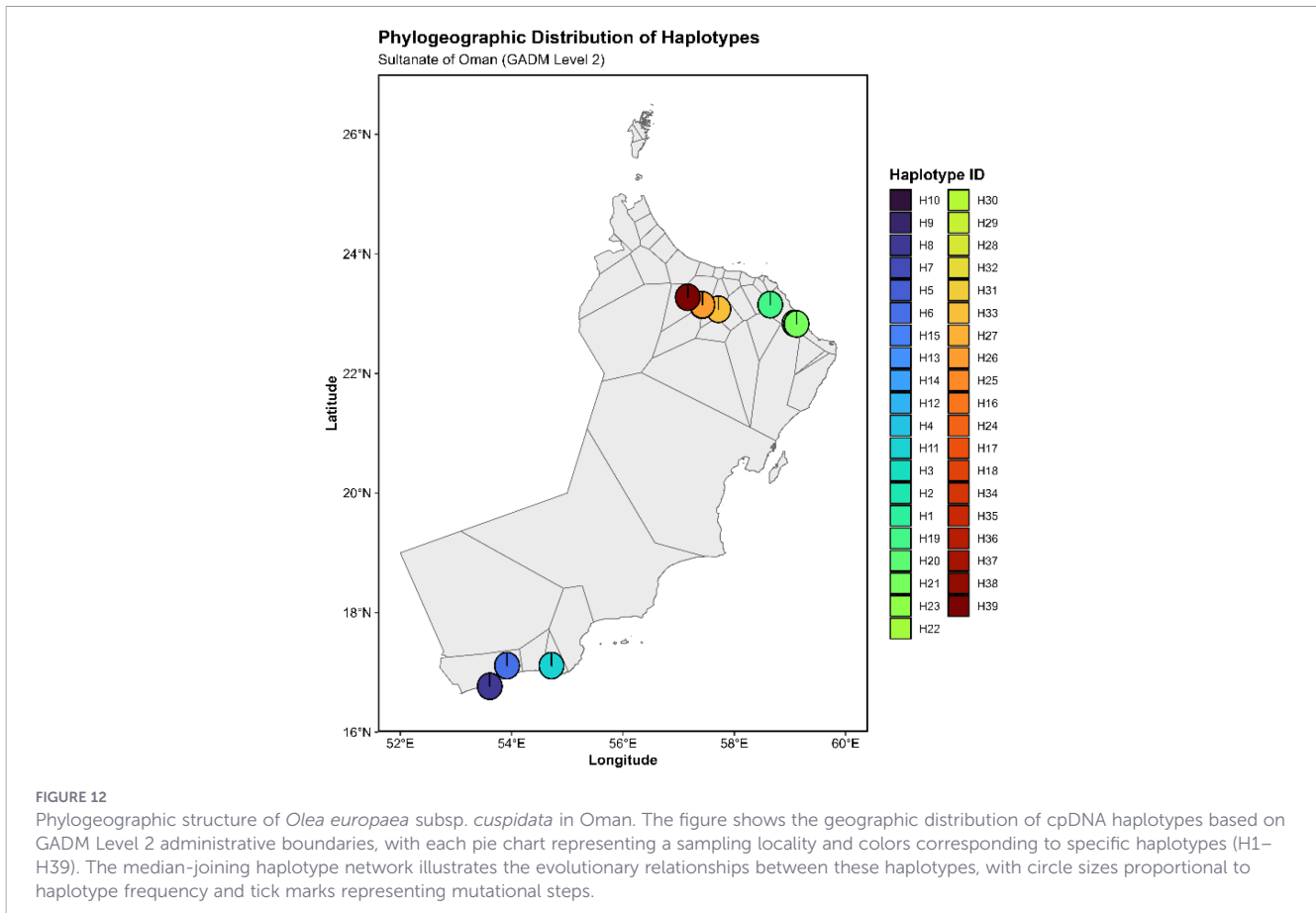
To clarify the relationship between differentiation metrics, it is important to distinguish between hierarchical levels of analysis. The global  $F_{ST}$  value (0.3093) is directly comparable to the AMOVA estimate at the population level ( $\Phi_{ST} = 0.2814$ ), and both indicate moderate to high genetic differentiation among populations. In contrast, the higher AMOVA value ( $\Phi_{RT} = 0.528$ ) reflects deeper regional divergence between the northern Hajar and southern Dhofar lineages. Therefore, the apparent difference between  $F_{ST}$  and  $\Phi$  values is expected and arises from their estimation at different hierarchical levels rather than inconsistency. Overall, both metrics consistently support strong population structuring driven by historical isolation.

Low  $F_{ST}$  values were observed among populations within the Eastern and Western Hajar Mountains, indicating limited genetic differentiation within each mountain range (Figure 8). In contrast, the Dhofar populations exhibited very high genetic differentiation ( $F_{ST} = 0.38–0.43$ ) from both Hajar Mountain populations, confirming their genetic isolation. Isolation by distance (IBD) was detected across Oman; however, this pattern may reflect historical demographic processes rather than ongoing connectivity. This pattern aligns with previous studies demonstrating that geographic barriers can strongly structure genetic diversity in olive populations (Besnard et al., 2007a; Besnard et al., 2009; Levy et al., 2016; Grigoriadou Zormpa et al., 2025).

*Olea europaea* subsp. *cuspidata* is distributed across a large geographic range, from southern Africa to the Arabian Peninsula



**FIGURE 11**  
TCS haplotype network of the complete chloroplast genome sequences (cpDNA) for wild olive across Oman based on chloroplast SNPs. The size of each circle represents the number of individuals sharing that haplotype, and crosshatches on the connecting lines indicate the number of nucleotide differences between haplotypes. Colour codes denote the eight wild olive populations across Oman: DJQ (Jabal Qara), DJR (Jabal Qamar), DJS (Jabal Samhan) from the Dhofar Mountains; EJA (Jabal Abyad) and EJB (Jabal Bani Jabir) from the Eastern Hajar Mountains; and WJH (Jabal Hatt), WJK (Jabal Akhdar), and WJS (Jabal Shams) from the Western Hajar Mountains.



and extending to north-eastern Africa and China (Green, 2002). Population genetics theory generally predicts that species with a more localized geographic range tend to exhibit lower levels of genetic variation compared to widely distributed species (Cole, 2003; Su et al., 2025). However, numerous studies have reported that localized plant species can display similar or even higher levels of genetic variation than widespread species, indicating that additional factors may influence genetic diversity (Yoichi and Tomaru, 2014; Zhu et al., 2023). Based on the isolation-by-distance model, gene flow between distant populations of a widely distributed species is expected to be lower than between more closely situated populations of geographically limited species (Godt and Hamrick, 1998). In Oman, the average distance between populations in the Eastern and Western Hajar Mountains is 313 km, whereas Dhofar populations are separated by an average of 899 km. Correspondingly, genetic divergence is lower between the Eastern and Western Hajar populations ( $F_{ST}=0.00$  to  $0.05$ ) and higher between the northern Hajar and Dhofar populations ( $F_{ST}=0.38$  to  $0.43$ ). These results indicate that gene flow among populations in the northern mountains is greater than between Dhofar populations, reflecting the effect of geographic distance on genetic connectivity.

The observed genetic patterns in Omani wild olive populations are shaped by geographic isolation, habitat structure, and other ecological and demographic factors. The Dhofar populations are both geographically distant and isolated by extensive desert areas, which restrict pollen-mediated gene flow and seed dispersal, leading to high genetic differentiation ( $F_{ST}=0.38$ – $0.43$ ) and low connectivity with the northern Hajar populations. In contrast, the Eastern and Western Hajar

Mountains are closer together and show evidence of admixture, consistent with greater gene flow and lower genetic differentiation ( $F_{ST}=0.00$  to  $0.05$ ). Additionally, the heterogeneous topography and microclimatic variability within each mountain range may maintain localized genetic variation by creating sub-habitats that promote population persistence and adaptation. These results suggest that both geographic distance and habitat structure play central roles in shaping the population genetic structure of wild olives in Oman, highlighting the importance of considering these factors in conservation planning and germplasm management.

The high genetic differentiation revealed by DAPC, PCA, and STRUCTURE analyses indicates that wild olive populations in the Western and Eastern Hajar mountains are genetically close, whereas the Dhofar populations are highly isolated due to the vast desert (part of the Empty Quarter) separating them from the northern mountains. Negative  $F_{IS}$  values obtained for wild olive populations in the three mountain ranges (Table 4) are consistent with an outcrossing mating system, but are not necessarily indicative of restricted gene flow. Moderate to high genetic differentiation among populations (global  $F_{ST}=0.3093$ ) and low gene flow ( $N_m=0.5584$ ) suggest restricted connectivity among populations. This pattern is further supported by AMOVA results, which revealed that most genetic variation occurs within populations (71.86%), while a significant proportion is distributed among populations (28.14%;  $\Phi_{ST}=0.2814$ ,  $P=0.001$ ), consistent with the global  $F_{ST}$  estimate (0.3093). The mating system of *Olea europaea* subsp. *cuspidata* likely contributes to these patterns: it is a monoecious, diploid ( $2n=46$ ) species that outcrosses, is wind-

pollinated, and its seeds are dispersed by small mammals and birds (Abiyu et al., 2015; Kassa et al., 2018). Clonal reproduction has not been observed in this subspecies, although related *O. europaea* subsp. *laperrinei* exhibits clonal growth (Baali-Cherif and Besnard, 2005).

The haplotype network based on plastid DNA SNPs identified five distinct wild olive haplotypes within Oman. The Eastern and Western Hajar mountains shared several haplotypes, reflecting gene flow and genetic connectivity between these northern populations, whereas the Dhofar populations exhibited unique haplotypes, highlighting their genetic isolation. Understanding the distribution of these haplotypes is essential for developing effective conservation strategies for wild olives in Oman. Knowledge of the genetic variation within and among populations can guide the establishment of core collections, inform habitat management, detect potential hybridization, and assess the adaptive potential of populations to environmental change (Hohenlohe et al., 2021).

From a conservation perspective, the strong genetic differentiation observed between northern and southern populations highlights the importance of region-specific management strategies. Dhofar populations should be considered a conservation priority due to their unique genetic lineages and geographic isolation. *In situ* conservation efforts should focus on the protection of natural habitats, particularly in the Dhofar Mountains, through habitat restoration and the establishment of protected areas. In addition, *ex situ* strategies, such as germplasm collection, seed banking, and breeding program development, are recommended to safeguard genetic resources and support future restoration efforts. These measures will help preserve the distinct evolutionary lineages of *O. europaea* subsp. *cuspidata* in Oman and maintain adaptive potential under environmental change.

## Conclusions

This detailed study provides essential insights into the genetic diversity and population structure of *Olea europaea* subsp. *cuspidata* across Oman. Wild olive populations exhibited relatively low genetic diversity within each mountain range but high genetic differentiation between the Dhofar mountains and both the Eastern and Western Hajar mountains ( $F_{ST} = 0.38\text{--}0.43$ ;  $\Phi_{ST} = 0.2814$ ), indicating strong isolation of southern populations and potential risk of genetic erosion. STRUCTURE, PCA, and DAPC analyses consistently revealed two major genetic clusters: one comprising the Dhofar populations, which are genetically distinct, and the other including the Eastern and Western Hajar populations, which are more genetically connected. These findings highlight the uniqueness of the Dhofar populations and the high connectivity among northern populations, emphasizing the importance of region-specific conservation strategies. Based on our analyses, we recommend prioritizing germplasm collection from the genetically distinct Dhofar populations and implementing targeted protection measures for isolated Hajar populations. Such actions will help preserve the overall genetic variation of wild olive in Oman and provide practical guidance for establishing core collections, managing fragmented populations, and supporting conservation and breeding programs.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/SRR36345062-SRR36345109>.

## Author contributions

TA: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing. BH: Data curation, Writing – original draft, Writing – review & editing. RE: Conceptualization, Writing – review & editing, Supervision. AC: Conceptualization, Writing – review & editing, Supervision.

## Funding

The author(s) declared that financial support was received for this work and/or its publication. This publication is funded by Sultan Qaboos University grant number SR/AGR/CROP/24/01.

## Acknowledgments

The authors gratefully acknowledge Salim Al Rahbi and Omar Al Amri from Oman Botanic Garden for their assistance in the collection of plant material, and Khadija Almaawali from the same institution for her support in preparing the map.

## Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Abiyu, A., Teketay, D., Glatzel, G., and Gratzer, G. (2015). Tree seed dispersal by African civets in the Afromontane Highlands: too long a latrine to be effective for tree population dynamics. *Afr. J. Ecol.* 53, 588–591. doi: 10.1111/aje.12198
- Al-Charaabi, Y., Al-Yahyai, S., Culham, A., and Ellis, R. (2013). Projection of future changes in rainfall and temperature patterns in Oman. *J. Earth Sci. Clim. Change* 4, 154. doi: 10.4172/2157-7617.1000154
- Al-Jabri, T., Culham, A., and Ellis, R. (2024). Wild olive in Oman and its conservation a review. *J. Agric. Mar. Sci.* 29, 1–14. doi: 10.53541/jams.vol29iss1pp1-14
- Allen, D., Westrip, J., Puttick, A., Harding, K., Hilton-Taylor, C., and Ali, H. (2021). *UAE national red list of vascular plants* (Dubai: Ministry of Climate Change and Environment, United Arab Emirates). Technical Report.
- Al-Yahyai, R. A., Halo, B. A., Al-Subhi, A. M., and Yaish, M. W. (2026). Genome-wide SNP analysis reveals genetic diversity and structure of wild and cultivated olives (*Olea europaea* L.) in Oman. *Sci. Rep.* 16, 11490. doi: 10.1038/s41598-026-40849-0
- Anthelme, F., Abdoukader, A., and Besnard, G. (2008). Distribution, shape and clonal growth of the rare endemic tree *Olea europaea* subsp. *laperrinei* (Oleaceae) in the Saharan mountains of Niger. *Plant Ecol.* 198, 73–87. doi: 10.1007/s11258-007-9386-6
- Baali-Cherif, D., and Besnard, G. (2005). High genetic diversity and clonal growth in relict populations of *Olea europaea* subsp. *laperrinei* (Oleaceae) from Hoggar, Algeria. *Ann. Bot.* 96, 823–830. doi: 10.1093/aob/mci232
- Besnard, G., Christin, P., Baali-Cherif, D., Bouguedoura, N., and Anthelme, F. (2007a). Spatial genetic structure in the Laperrine's olive (*Olea europaea* subsp. *laperrinei*), a long-living tree from the central Saharan mountains. *Heredity* 99, 649–657. doi: 10.1038/sj.hdy.6801051
- Besnard, G., Gorrilliot, O., Raimondeau, P., Génot, B., El Bakkali, A., Anthelme, F., et al. (2021). Contrasting genetic footprints among Saharan olive populations: potential causes and conservation implications. *Plants* 10, 1207. doi: 10.3390/plants10061207
- Besnard, G., Rubio de Casas, R., Christin, P.-A., and Vargas, P. (2009). Phylogenetics of *Olea* (Oleaceae) based on plastid and nuclear ribosomal DNA sequences: tertiary climatic shifts and lineage differentiation times. *Ann. Bot.* 104, 143–160. doi: 10.1093/aob/mcp105
- Besnard, G., Rubio de Casas, R., and Vargas, P. (2007b). Plastid and nuclear DNA polymorphism reveals historical processes of isolation and reticulation in the olive tree complex (*Olea europaea*). *J. Biogeogr.* 34, 736–752. doi: 10.1111/j.1365-2699.2006.01653.x
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Buschbom, J., Yanbaev, Y., and Degen, B. (2011). Efficient long-distance gene flow into an isolated relict oak stand. *J. Heredity* 102, 464–472. doi: 10.1093/jhered/esr023
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinf.* 10, 421. doi: 10.1186/1471-2105-10-421
- Canale, A., and Loni, A. (2010). Insects visiting olive flowers (*Olea europaea* L.) in a Tuscan olive grove. *Redia* 92, 95–98.
- Chludil, D., Čepel, J., Steffenrem, A., Stejskal, J., Sagariya, C., Pook, T., et al. (2025). A pollen-based assisted migration for rapid forest adaptation. *Global Change Biol.* 31, e70014. doi: 10.1111/gcb.70014
- Cole, C. T. (2003). Genetic variation in rare and common plants. *Annu. Rev. Ecol. Evol. Syst.* 34, 213–237. doi: 10.1146/annurev.ecolsys.34.030102.151717
- Cuneo, P., and Leishman, M. R. (2006). African Olive (*Olea europaea* subsp. *cuspidata*) as an environmental weed in eastern Australia: a review. *Cunninghamia* 9, 545–577.
- Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19 (1), 11–15.
- Dzialuk, A., Chybicki, I., Gout, R., Mączka, T., Fleischer, P., Konrad, H., et al. (2014). No reduction in genetic diversity of Swiss stone pine (*Pinus cembra* L.) in Tatra Mountains despite high fragmentation and small population size. *Conserv. Genet.* 15, 1433–1445. doi: 10.1007/s10592-014-0628-6
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- Fernández-Rodríguez, S., and et al. (2014). Identification of potential sources of airborne *Olea* pollen in the Southwest Iberian Peninsula. *Int. J. Biometeorol.* 58, 337–348. doi: 10.1007/s00484-012-0629-4
- Francis, R. (2017). pophelper: an R package and web app to analyse and visualize population structure. *Mol. Ecol. Resour.* 17, 27–32. doi: 10.1111/1755-0998.12509
- Ghazanfar, S. A. (2015). *Flora of the sultanate of Oman. Volume 3. Loganiaceae-asteraceae*. Meise, Belgium: Meise Botanic Garden.
- Ghazanfar, S. A. (2018). “Edible wild plants: a case study from Oman,” in *Global perspectives on underutilized crops* (Cham, Switzerland: Springer), 207–216. doi: 10.1007/978-3-319-77776-4\_8
- Godt, M. J. W., and Hamrick, J. (1998). Genetic divergence among infraspecific taxa of *Sarracenia purpurea*. *Syst. Bot.* 23 (4), 427–438. doi: 10.2307/2419375
- Green, P. S. (2002). A revision of *olea* L. (Oleaceae). *Kew Bull.* 50, 91–140. doi: 10.2307/4110824
- Green, P., Wickens, G., Tan, K., Mill, R., and Elias, T. (1989). *The Olea europaea complex. The Davis and Hedge Festschrift* (Scotland: Edinburgh University press).
- Grigoriadou Zormpa, O., Wilhelmi, S., Vucetic, B., Ciocirlan, M.-I.-C., Mueller, M., Ciocirlan, E., et al. (2025). Genetic diversity and fine-scale spatial genetic structure of European beech populations along an elevational gradient. *Heredity* 134, 1–10. doi: 10.1038/s41437-025-00776-8
- Habib, N. A., Müller, M., Gailing, O., Patzelt, A., Al Issai, G., Krutovsky, K. V., et al. (2021). Genetic diversity and differentiation of *Olea europaea* subsp. *cuspidata* (Wall. & G. Don) Cif. in the Hajar Mountains of Oman. *Genet. Resour. Crop Evol.* 68, 865–883. doi: 10.1007/s10722-020-01030-2
- Hohenlohe, P. A., Funk, W. C., and Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Mol. Ecol.* 30, 62–82. doi: 10.22541/au.158480040.06912807
- Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030
- Johnson, M. G., and et al. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Appl. Plant Sci.* 4, 1600016. doi: 10.3732/apps.1600016
- Johnson, M., Slimp, M., Williams, L. D., and Hale, H. (2021). “Data from: On the potential of Angiosperm353 for population genomics studies,” in *Dryad dataset*. doi: 10.5061/dryad.76hdr7sv3
- Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129
- Kamvar, Z. N., Tabima, J., and Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2, e281. doi: 10.7717/peerj.281
- Kassa, A., Konrad, H., and Geburek, T. (2018). Mating pattern and pollen dispersal in the wild olive tree (*Olea europaea* subsp. *cuspidata*). *Tree Genet. Genomes* 14, 3. doi: 10.1007/s11295-017-1215-z
- Kassa, A., Konrad, H., and Geburek, T. (2019). Molecular diversity and gene flow within and among different subspecies of the wild olive (*Olea europaea* L.): a review. *Flora* 250, 18–26. doi: 10.1016/j.flora.2018.11.014
- Knaus, B. J., and Grünwald, N. J. (2017). vcfr: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* 17, 44–53. doi: 10.1111/1755-0998.12549
- Leigh, J. W., Bryant, D., and Nakagawa, S. (2015). POPART: full-feature software for haplotype network construction. *Methods Ecol. Evol.* 6, 1110–1116. doi: 10.1111/2041-210X.12410
- Levy, E., Byrne, M., Coates, D., Macdonald, B., McArthur, S., and Van Leeuwen, S. (2016). Contrasting influences of geographic range and distribution of populations on patterns of genetic diversity in two sympatric Pilbara acacias. *PLoS One* 11, e0163995. doi: 10.1371/journal.pone.0163995
- Linck, E., and Battey, C. (2019). “Minor allele frequency thresholds strongly affect population structure inference with genomic data sets,” in *Molecular ecology resources*, vol. 19, 639–647. doi: 10.1111/1755-0998.12995
- Lischer, H. E., and Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28, 298–299. doi: 10.1093/bioinformatics/btr642
- McGregor, S. E. (1976). *Insect pollination of cultivated crop plants* (Washington, DC, USA: Agricultural Research Service, US Department of Agriculture).
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-

- generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110
- McLay, T. G., and et al. (2021). New targets acquired: Improving locus recovery from the Angiosperms353 probe set. *Appl. Plant Sci.* 9, e11420. doi: 10.1002/aps3.11420
- Nassiry, M., Javanmard, A., Tohidi, R., and Table, (2009). Application of statistical procedures for analysis of genetic diversity in domestic animal populations. 4 (4), 136–141. doi: 10.3844/ajavsp.2009.136.141
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2018). *vegan*: Community Ecology Package. R package version 2.5-3. 2018. Available online at: <https://CRAN.R-project.org/package=vegan> (Accessed January 23, 2026).
- Omire, A., Neondo, J., Budambula, N. L., Wangai, L., Ogada, S., and Mweu, C. (2022). Genetic diversity and population structure of Doum Palm (*Hyphaene compressa*) using genotyping by sequencing. *Front. Genet.* 13. doi: 10.3389/fgene.2022.762202
- Ortiz, E. M. (2019). *vcf2phyloip v2.0*: convert a VCF matrix into several matrix formats for phylogenetic analysis. doi: 10.5281/zenodo.2540861
- Park, J.-S., Kang, M.-Y., Shim, E.-J., Oh, J., Seo, K.-I., Kim, K. S., et al. (2022). Genome-wide core sets of SNP markers and Fluidigm assays for rapid and effective genotypic identification of Korean cultivars of lettuce (*Lactuca sativa* L.). *Hortic. Res.* 9, uhac119. doi: 10.1093/hr/uhac119
- Patzelt, A. (2015a). *Oman plant: red data book* (Oman, Muscat: Oman Botanic Garden).
- Patzelt, A. (2015b). *Photographic field guide to the plants of the Western Hajar mountains, sultanate of Oman* (Muscat: Sultan Qaboos University).
- Patzelt, A. (2015c). Synopsis of the flora and vegetation of Oman, with special emphasis on patterns of plant endemism. *Abhandlungen der Braunschweigischen Wissenschaftlichen Gesellschaft* 282, 317.
- Preston, G. W., Thomas, D. S., Goudie, A. S., Atkinson, O. A., Leng, M. J., Hodson, M. J., et al. (2015). A multi-proxy analysis of the Holocene humid phase from the United Arab Emirates and its implications for southeast Arabia's Neolithic populations. *Quat. Int.* 382, 277–292. doi: 10.1016/j.quaint.2015.01.054
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi: 10.1093/genetics/155.2.945
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Serrote, C. M. L., Reiniger, L. R. S., Silva, K. B., and Stefanel, C. M. (2022). Gene flow affects the genetic diversity of *Copaifera langsdorffii* Desf. in Cerrado. *Acta Botanica Brasiliica* 36, e20220007. doi: 10.1590/1677-941X-ABB-2022-0007
- Slimp, M., Williams, L. D., Hale, H., and Johnson, M. G. (2021). On the potential of Angiosperms353 for population genomic studies. *Appl. Plant Sci.* 9. doi: 10.1002/aps3.11419
- Smouse, P. E., Long, J. C., and Sokal, R. R. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst. zool.* 35, 627–632. doi: 10.2307/2413122
- Su, Z., Lin, Z., Zhuo, L., Jiang, X., and Jiang, Q. (2025). Small population size and low levels of genetic diversity in an endangered species endemic to the Western Tianshan Mountains. *Plants* 14, 3105.
- Tel-Zur, N., Abbo, S., Myslabodski, D., and Mizrahi, Y. (1999). Modified CTAB procedure for DNA isolation from epiphytic cacti of the genera *Hylocereus* and *Selenicereus* (Cactaceae). *Plant Mol. Biol. Rep.* 17, 249–254. doi: 10.1023/A:1007656315275
- Templeton, A. R., Crandall, K. A., and Sing, C. (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132, 619–633. doi: 10.1093/genetics/132.2.619
- Wang, J., Luo, M. C., Chen, Z., You, F. M., Wei, Y., Zheng, Y., et al. (2013). *Aegilops tauschii* single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New Phytol.* 198, 925–937. doi: 10.1111/nph.12164
- Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *evolution* 38 (6), 1358–1370. doi: 10.2307/2408641
- Willing, E.-M., Dreyer, C., and Van Oosterhout, C. (2012). Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. *PLoS ONE* 7 (8), e42649. doi: 10.1371/journal.pone.0042649
- Yoichi, W., and Tomaru, N. (2014). Patterns of geographic distribution have a considerable influence on population genetic structure in one common and two rare species of *Rhododendron* (Ericaceae). *Tree Genet. Genomes* 10, 827–837. doi: 10.1007/s11295-014-0723-3
- Yooprasert, S. (2021). *Taxonomy and systematics of urophyllum (Rubiaceae) in Thailand and IndoChina* (UK: University of Reading).
- Zhu, X., Tang, J., Jiang, H., Yang, Y., Chen, Z., Zou, R., et al. (2023). Genomic evidence reveals high genetic diversity in a narrowly distributed species and natural hybridization risk with a widespread species in the genus *Geodorum*. *BMC Plant Biol.* 23, 317. doi: 10.1186/s12870-023-04285-w
- Zohary, D. (1995). *Olive, Olea europaea (Oleaceae). Evolution of crop plants.* Smartt, J. & Simmonds, NW Eds (Harlow, England: Longman Scientific & Technical).