Journal of Applied Research on Medicinal and Aromatic Plants 38 (2024) 100523

Contents lists available at ScienceDirect

Journal of Applied Research on Medicinal and Aromatic Plants

journal homepage: www.elsevier.com/locate/jarmap

Research paper

Diversity analysis of southern African *Artemisia afra* using a single nucleotide polymorphism derived from diversity arrays

Moosa Mahmood Sedibe [a,*], Matumelo A. Rafiri [a], Goitsemang Mahlomola Hendry Dikane [a], Matthew Chilaka Achilonu [b], S'busiso Mfan'vele Nkosi [b], Xolile Veronica Ngubane [b], Rian Pierneef [c]

[a] Department of Agriculture, Central University of Technology, Free State, 20 President Brand Street, Bloemfontein, South Africa
[b] Technology Station in Chemicals, Mangosuthu` University of Technology, 511 Mangosuthu Highway Umlazi, Durban, South Africa
[c] University of Pretoria, Centre for Bioinformatics and Computational Biology, Lynnwood Rd, Hatfield, Pretoria 0002, South Africa

ARTICLE INFO

Keywords:
Aromatic shrub
DNA fingerprinting
Ethnomedicine
Germplasm
Next-generation sequencing
UPGMA dendrogram

ABSTRACT

Artemisia afra is an aromatic shrub used as a herbal remedy in southern Africa for various disorders, including coughs, colds, influenza, and, most notably, for the prophylaxis and treatment of *Plasmodium falciparum* infections. Modern high-throughput genotyping by sequencing, which combines next generation sequencing and diversity array technology (DArTseq) complexity reduction methods, allows for the generation of high-quality whole-genome profiles without any previous organismal genetic information that may hamper the diversity analyses. This study investigated the feasibility of incorporating DArTseq-derived single-nucleotide polymorphisms (SNPs) and bioactive compounds into a geographic information system to predict the spatial distribution of *A. afra* gene diversity in selected southern African environs. DNA extraction, sequencing, and SNP calling mineral analyses (using R Studio) were performed on a total of 80 genotypes collected around Roma and Mohale's Hoek, Lesotho, and Wepener and Hobhouse, South Africa. The unweighted pair group method with arithmetic mean (UPGMA) and neighbour-joining dendrograms revealed three clusters of variation among genotypes; however, geographic distribution did not influence *A. afra* genetic diversity. The mean heterozygosity among genotypes was 7.7%. Genotypes collected in Roma and Mohale's Hoek exhibited high levels of genetic diversity. Local *A. afra* genetic diversity was relatively low, indicating uniformity of subspecies per locality. The diversity among genotypes indicated that *A. afra* SNPs may occur despite the location. Given that genetic diversity enhances crop performance, our findings can be useful in germplasm management, industries producing by-products of *A. afra*, and plant breeding programs. For the long-term conservation of medicinal plants in South Africa, significant sampling and regional data are required. In addition to evaluating qualitative and quantitative assessments of phytochemicals, it is necessary to evaluate foliar mineral composition, as well as the agronomic and morphological characteristics of collected specimens from studied locations.

1. Introduction

Artemisia afra Jacq. ex Willd (wormwood) is part of the Mannoliphyta division, Magnoliopsida class, Asteridae sub-class, Asterales order, and Asteraceae family. Its morphology, taxonomy, distribution, ethnopharmacology, utilization, chemistry, and biological effects were reported by Raimondo et al. (2009). *A. afra* is an important medicinal shrub found in Africa and is widely distributed in South Africa, Kenya, Tanzania, Uganda, Ethiopia, Namibia, Zimbabwe, Eswatini, and Lesotho (Liu et al., 2009). *A. afra* occurs in Gauteng, Limpopo, and the eastern regions of South Africa, as well as in ESwatin and Lesotho, extending to the Western Cape. The plant also occurs widely in the KwaZulu-Natal province, ranging from the coastline to the Drakensberg Mountain range (Raimondo et al., 2009). *A. afra* is used as a fumigant against pests of stored products (such as weevils) and as a moth repellent (Nigam et al., 2019). Liu et al. (2009) highlighted its importance as a medicinal herb used by rural communities to cure infectious diseases—primarily the common cold (Van Wyk, 2008). A mixture of *A. afra* and *Eucalyptus globulus* is utilized for treating flu, while a blend of *Lippia asperifolia* and *A. afra* leaves and stems is applied as a remedy for fever, flu, measles,

* Corresponding author.
E-mail address: msedibe@cut.ac.za (M.M. Sedibe).

https://doi.org/10.1016/j.jarmap.2023.100523
Received 13 July 2023; Received in revised form 29 October 2023; Accepted 7 November 2023
Available online 10 November 2023

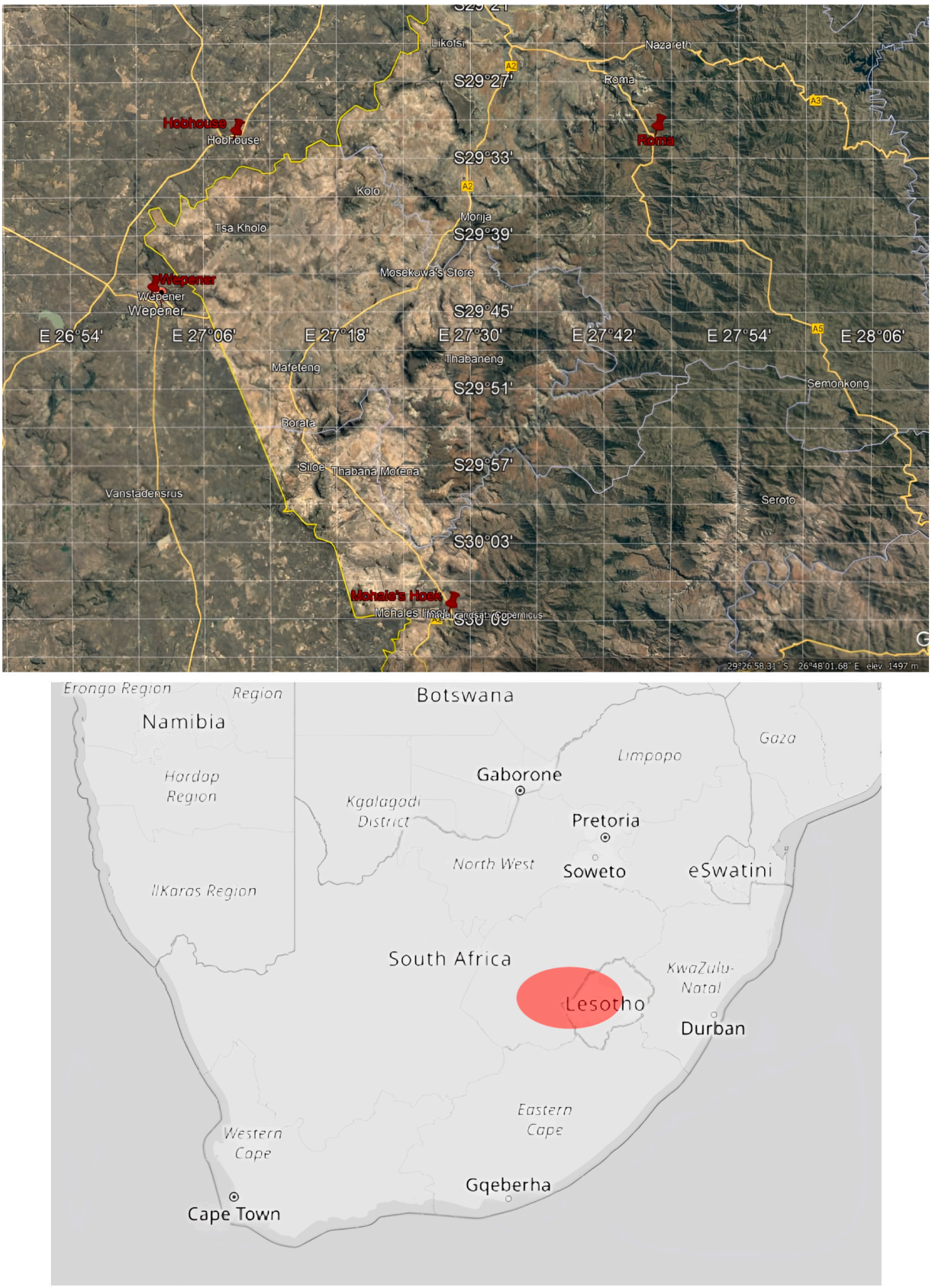

Fig. 1. Map of the study area constructed using version 7.3 of Google Earth Pro. Locations relevant to this paper are marked by red pins. Lesotho (Mohale's Hoek [MH], Roma [RO]) and South Africa (Wepener [WE], Hobhouse [HO]).

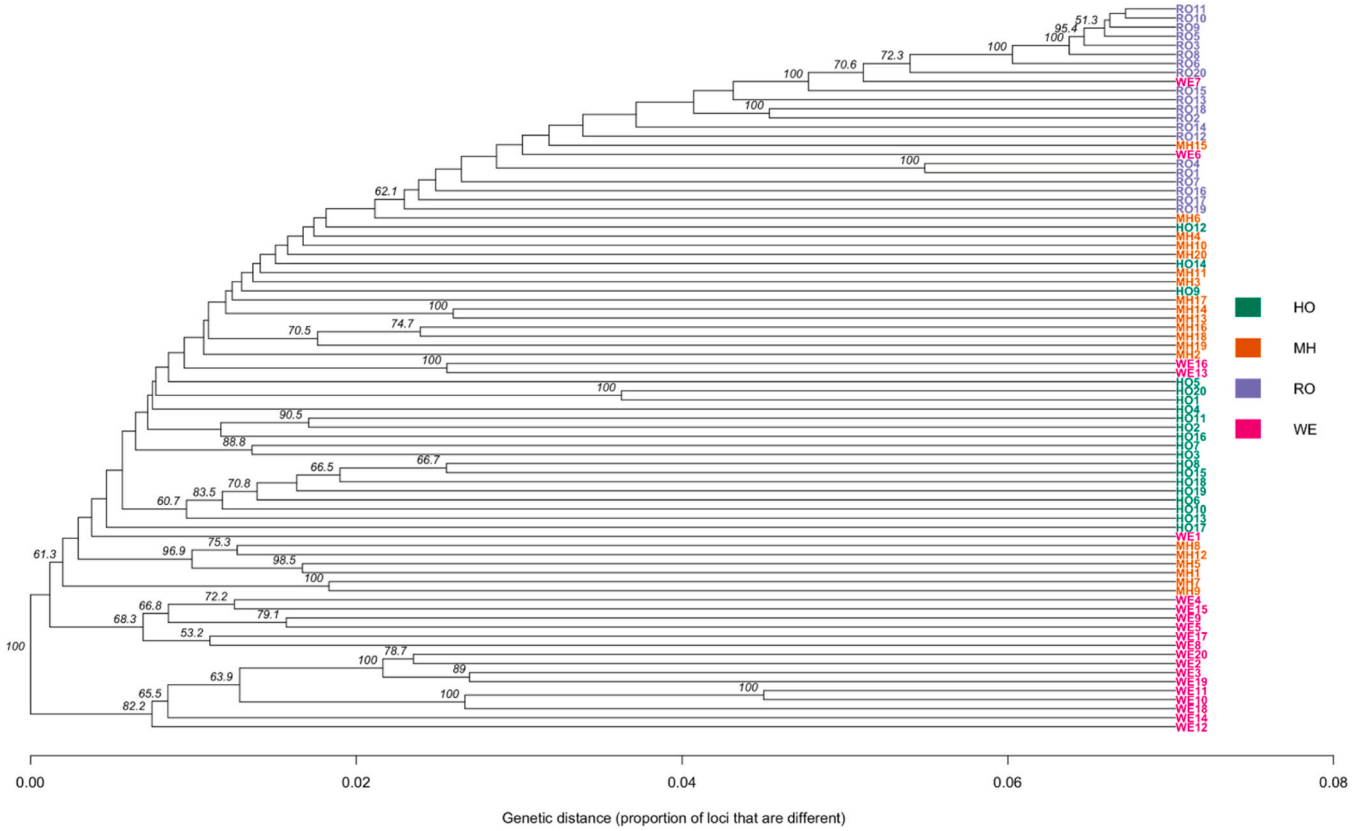

Fig. 2. Dendrogram of *Artemisia afra* collected from Lesotho (Mohale's Hoek [MH]: 30.1426° S, 27.4674° E, and Roma [RO]: 29.5212° S, 27.7755° E) and South Africa (Wepener [WE]: 29.7294° S, 27.0206° E, and Hobhouse [HO]: 29.5258° S, 27.1452° E).

and a preventative measure against lung infection (Liu et al., 2009). *A. afra* has a small, globular capitulum of tiny yellow flowers (Setshedi et al., 2022). *A. afra* is anemophilous, meaning it is pollinated by wind. However, its seeds lack a pappus, which reduces the likelihood of wind dispersal (Setshedi et al., 2022). Genetic diversity is important for the adaptability and survival of plant species (Kujane et al., 2019). While there have been several genomic studies of the *Artemisia* genus (Chen et al., 2022; Shen et al., 2018), reports of the *A. afra* genome and its sequencing remain scarce (Hu and Zhen, 2020). This study aimed to investigate the feasibility of diversity arrays technology sequencing (DArTseq)-derived single-nucleotide polymorphisms (SNPs).

2. Material and methods

2.1. Plant material collection

We collected *A. afra* specimens (20 per location) in a random fashion from natural plant communities in Lesotho (Mohale's Hoek [MH]: 30.1426° S, 27.4674° E, Roma [RO]: 29.5212° S, 27.7755° E) and the Free State Province, South Africa (Wepener [WE]: 29.7294° S, 27.0206° E, Hobhouse [HO]: 29.5258° S, 27.1452° E) (Fig. 1).

2.2. DNA extraction, sequencing, and SNP calling

Fresh leaves were frozen in liquid nitrogen and powdered with a mortar and pestle. DArTseq were used to evaluate the genetic diversity of the genotypes. The DArTseq™ protocol (Diversity Arrays Technology, Canberra, Australia) was used for the extraction and sequencing of *A. afra* DNA. Genomic DNA was extracted from 1 mg of leaf tissue from each accession. Isolation was then carried out using a modified cetyltrimethyl ammonium bromide (CTAB)/chloroform/isoamyl alcohol method (Doyle, 1987). The ground leaf tissue was mixed with 2% CTAB isolation buffer (1.4 M NaCl, 100 mM Tris [pH 8.0], and 20 mM ethylenediaminetetraacetic acid [EDTA]) that had been pre-warmed (60 °C; Sigma-Aldrich, St. Louis, MO, USA). The mixture was then transferred to a 2 mL microcentrifuge tube and incubated for 1 h at 60 °C. DNA was extracted once with chloroform-isoamyl alcohol (ChI/IAA, 24:1) and precipitated with 2 volumes of isopropanol. The pellet was washed with 70% EtOH (ethanol/ethyl), dried, and dissolved in 100 mL TE buffer (consisting of Tris and EDTA) containing 50 g/mL RNAse A (bovine pancreatic ribonuclease; Sigma-Aldrich, South Africa). For DArTseq and SNP genotyping, the extracted DNA was quantified using 0.8% agarose gel electrophoresis adjusted to 50 mg/L.

2.3. Diversity arrays technology sequencing

To process the DNA in digestion/ligation reactions, a single restriction enzyme (*PstI*)-compatible adaptor was replaced with two adaptors corresponding to two different restriction enzymes, *PstI* and *SphI*-compatible adaptors (Melville et al., 2017). An Illumina flow cell attachment sequence was combined with staggered sequences of varying length barcode regions to create the *PstI*-compatible adaptor (Melville et al., 2017). In 30 rounds of polymerase chain reaction (PCR), only mixed fragments (*PstI-SphI*) were amplified using the following reaction conditions: 94 °C for 1 min, 30 cycles of 94 °C for 20 s, 58 °C for 30 s, 72 °C for 45 s, and 72 °C for 7 min. Following PCR, equimolar amounts of amplification products from each of the 96-well microtiter plate samples were bulked and applied to a c-Bot (Illumina) bridge PCR, followed by sequencing on an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA). The single-read sequencing was repeated 77 times.

Each lane's sequences were processed using proprietary DArT analytical pipelines (PLs). In the primary PL analysis, poor-quality sequence fragments with reproducibility of $< 90\%$ and a read depth < 3.5 for SNPs, or 5 for presence-absence markers, were removed. The

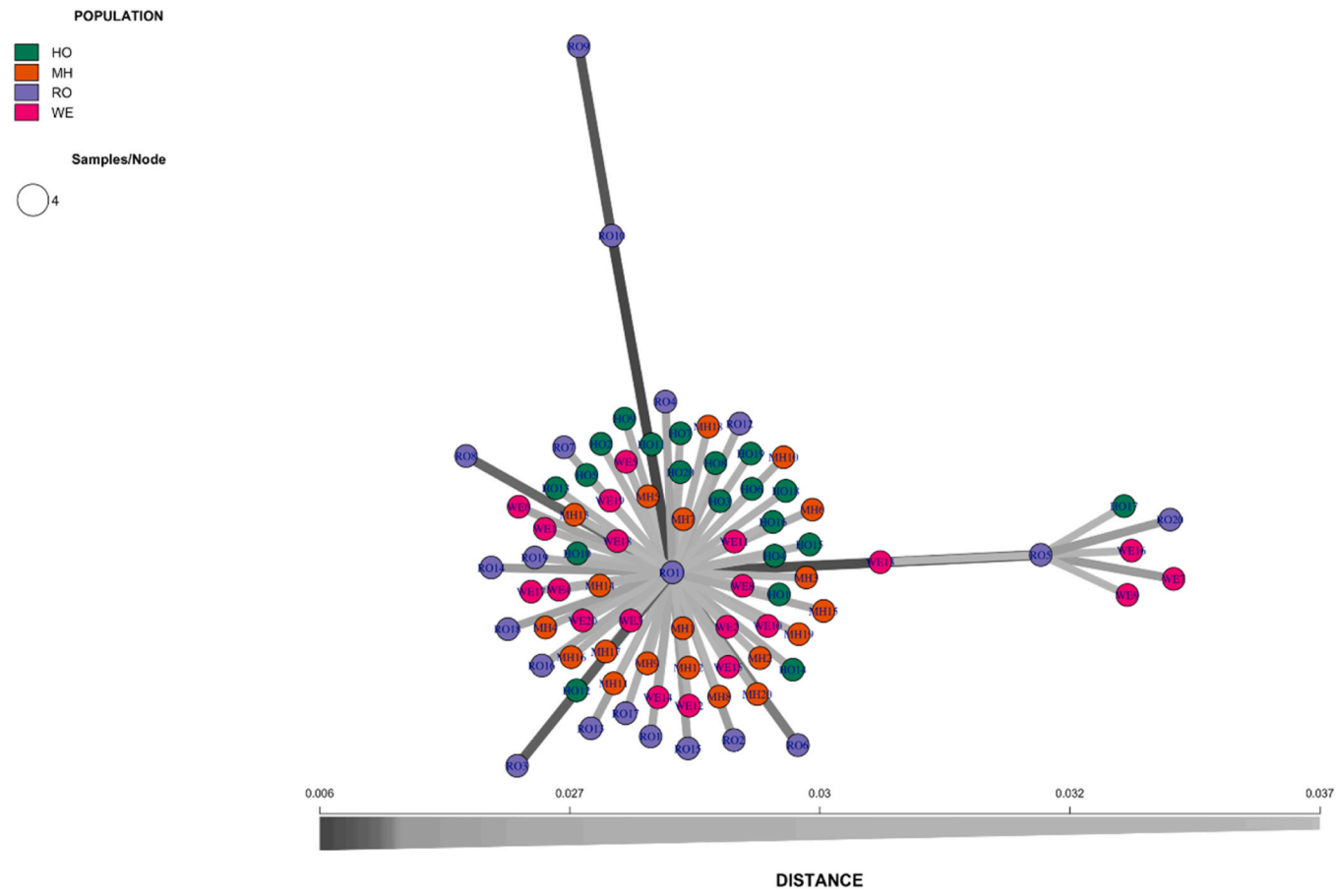

Fig. 3. Genetic distance relationships of *A. afra* populations obtained from four locations.

barcode region was subjected to more stringent selection criteria than the other sequences. Sequences were identified and used in marker callings for each barcode/sample. For reference and SNP alleles, we recorded the average browse depth across loci per individual per locus. Fastqcoll files were groomed using DArT PL's proprietary algorithm, which converts a low-quality base from a singleton tag into a correct base using collapsed tags with multiple members as a template. DArTsoft14 was used in the secondary PL to call DArT PL's proprietary SNP and SilicoDArT (presence/absence of restriction fragments in representation; PA markers) algorithms (Diversity Arrays Technology).

SNP calling was performed at a brink distance of 3 for all tags from all libraries contained within the DArTsoft14 analysis and clustered using DArT PL's C+ + algorithm. Clusters were divided into separate SNP loci using the balance of read counts for allelic pairs. To facilitate the selection of technical parameters to effectively distinguish true allelic variants from paralogous sequences, alleles were tested for deviations using the Hardy–Weinberg equilibrium. Multiple samples were processed as technical replicates from DNA to allelic calls, and scoring consistency was used as the primary selection criteria for high-quality/low-error-rate markers. The high average browse depth per locus (over 30 reads/locus across all markers) ensured for high calling quality. For genotyping by sequencing (GBS) analyses, DNA was diluted to 50 mg/L.

2.4. Data analysis

SNP analyses were conducted using DArTseq. For population genetic analyses in R Studio, large SNP sequencing sets were obtained by GBS. GBS was also used in high-throughput sequencing (HTS) experiments. The genome was reduced in GBS using restriction enzymes and sequencing with HTS (Pojskić, 2018). The analysis was performed for 80 *A. afra* genotypes. *A. annua* L., the closest relative to *A. afra*, was used as a reference in the DNA analysis.

Individual and population genetic distance matrices were created. To assess branch support, a distance tree was built using the unweighted pair group method with arithmetic mean (UPGMA) algorithm with 100 bootstrap replicates. Minimum spanning networks were used to visualize the population structure and determine the number of allelic differences that existed between two samples. Principle component analysis (PCA) was used to convert the observed SNP data into principal components, which visualizes variation between samples. A discriminant analysis of principle components, a multivariate statistical approach that maximizes variance among populations in samples, was used to investigate population assignments.

3. Results and discussion

It remains challenging to initially assemble large genomes that have a high degree of heterozygosity (Schatz et al., 2012; Kajitani et al., 2014; Shen et al., 2018). GBS DArTseq represents a useful approach to comparing the composition of genomic representations of different genotypes. DArTseq can be used for profiling whole plant genomes in broader taxonomic spectra, rather than just focusing on cultivated plants, to reveal variable hybridization signal intensities for different individuals. Despite the species being used as a multipurpose ethnomedicinal plant in many African countries, the current study serves as a benchmark for *A. afra* spp genomic profiling. We used *A. annua* as the reference species in the genomic analyses, which is one of only a few species sequenced from the family Asteraceae (Shen et al., 2018).

The UPGMA method (Fig. 2) revealed three major clusters and subclusters (Kalinowski, 2009). The dendrogram assembled 80 genotypes

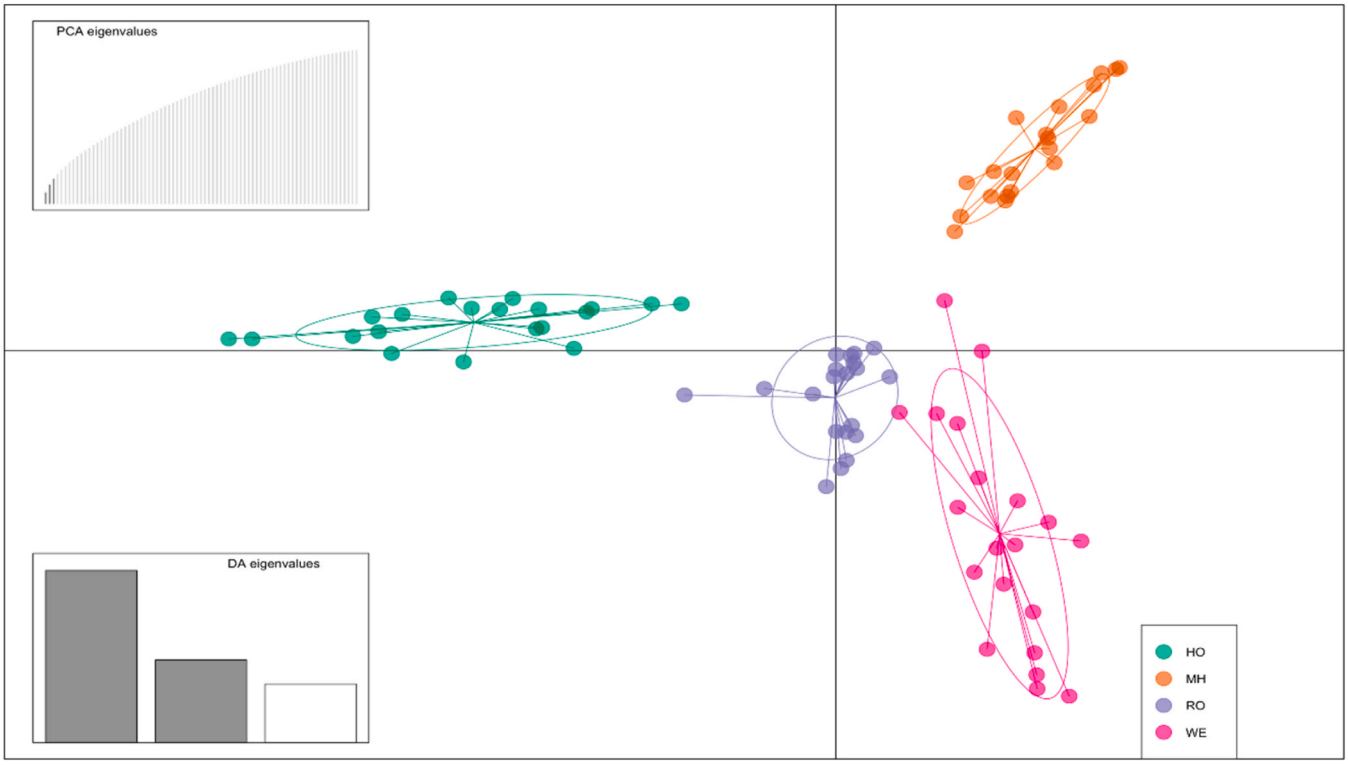

Fig. 4. Genetic variability assessed via principal component analysis (PCA) of 80 *A. afra* genotypes from four locations.

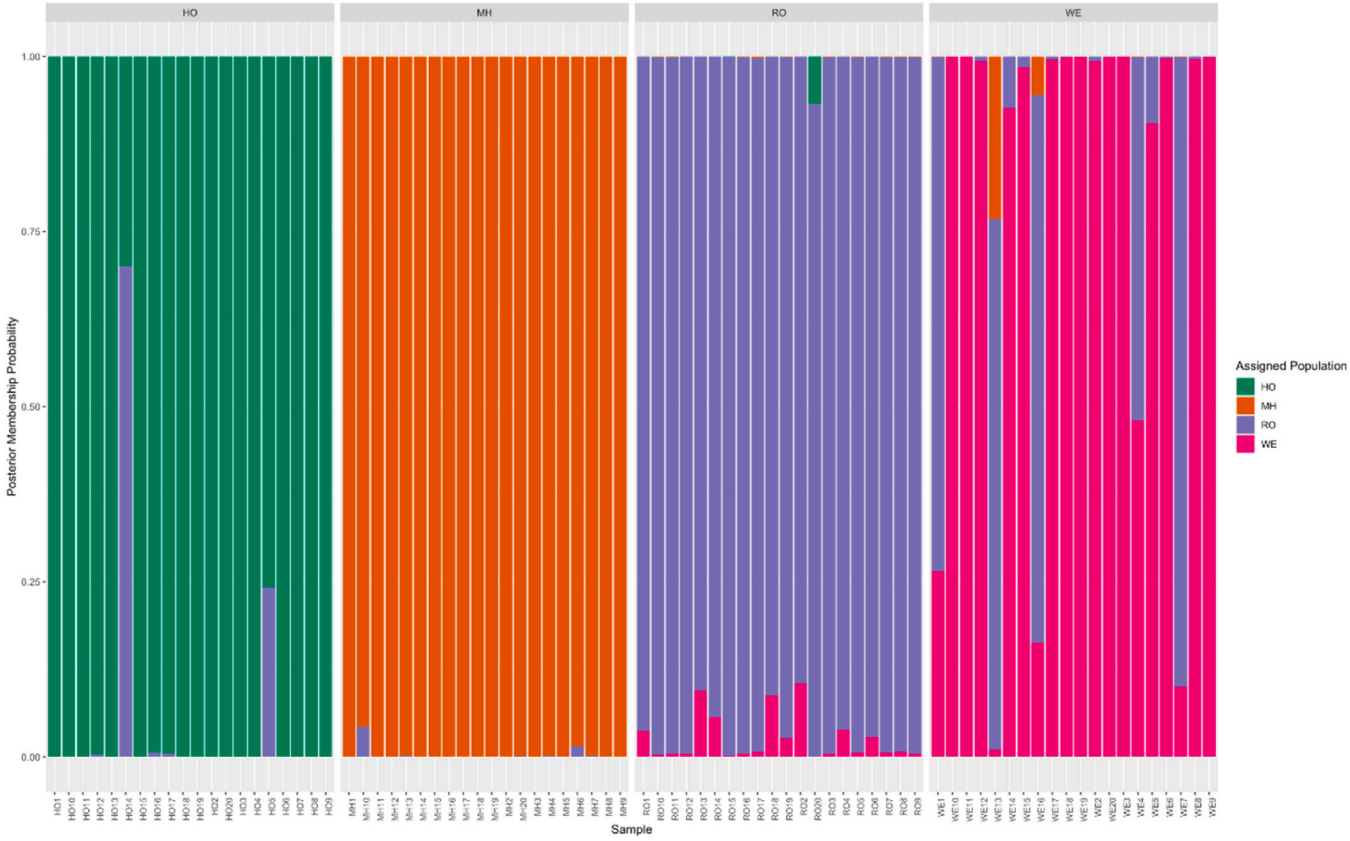

Fig. 5. Population structure and membership probability of 80 *A. afra* accessions based on diversity arrays technology. The x-axis corresponds to the genotypes and the colour coding corresponds to the sampling location. Lesotho (Mohale's Hoek [MH], Roma [RO]) and South Africa (Wepener [WE], Hobhouse [HO]).

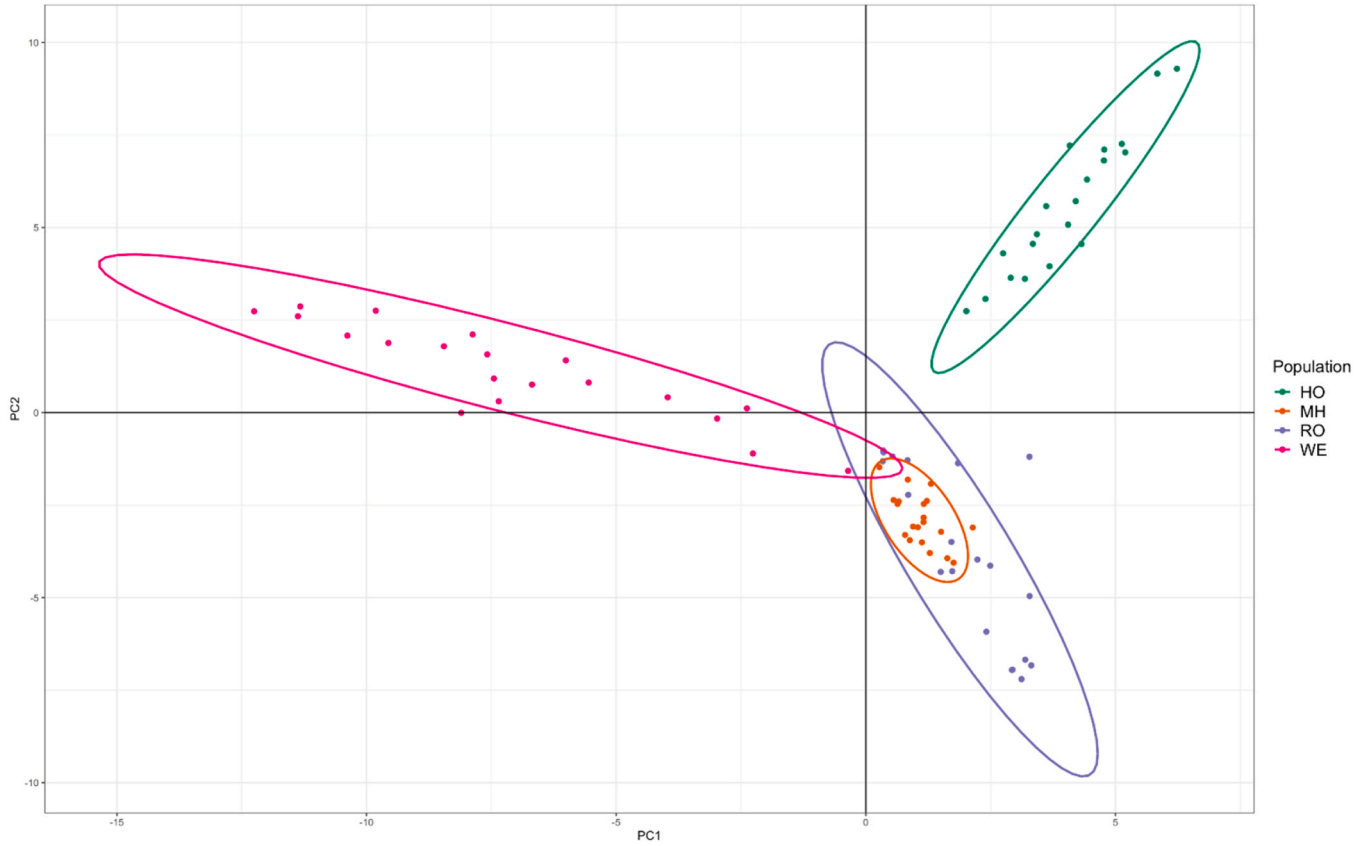

Fig. 6. PCA ordination of 80 *A. afra* genotypes from four locations.

into three clusters, grouping similar genotypes, regardless of the sampling locality. The hierarchical cluster dendrogram represents the relationships between genotypes in terms of approximate distances and is based on genetic relatedness. Cluster 1 comprised of nine genotypes, while Cluster 2 comprised of six. Clusters 1 and 2 contained only those genotypes collected from WE. However, most (n = 63) genotypes appeared in cluster 3, the largest cluster in the dendrogram. Cluster 1 included 20 genotypes from MH, RO, and HO. Five genotypes collected from WE also appeared in cluster 3. The UPGMA-constructed tree indicated variation in genetic distance (the height), ranging from 53.1 to 100. Variations in distance were noted in both clusters and clades (Fig. 2), and the observed differences between the accessions may be ascribed to genetic variation.

Genetic diversity in wild plant species is related to the life form, mating/breeding system, seed dispersal, geographic range, and population size (Li et al., 2020). Ferreira et al. (1995) reported that *A. annua* is cross-pollinated by both wind and insects, which may have influenced the genetic diversity among *Asteraceae* species. However, there is no reported information on the mode of pollination of *A. afra*.

Approximately 15 *A. afra* genotypes collected from WE were closely related to six genotypes from MH. Cluster 2 contained 17 genotypes from HO. Five genotypes from WE were not closely related to the other genotypes collected at the same location, which were grouped in cluster 3. However, they showed similarities with the genotypes from RO, MH, and HO.

In cluster 1, the genetic distance in the first clade was 81.9 while that of the second clade in clusters 2 and 3 was 59.8, further confirming the variation between these genotypes (Fig. 2). Clades within cluster 1

Table 1
Genetic diversity within and among 80 *Artemisia afra* genotypes based on 22,948 single nucleotide polymorphism markers from Lesotho (Mohale's Hoek, 30.1426° S, 27.4674° E; Roma, 29.5212° S, 27.7755° E) and South Africa (Wepener, 29.7294° S, 27.0206° E; Hobhouse, 29.5258° S, 27.1452° E).

Chromosome No.	Locus No.	Ho	Hs	Ht	Dst	Htp	Dstp	Fst	Fstp	Fis	Dest
1	5	0.011	0.039	0.041	0.002	0.042	0.003	0.053	0.069	0.716	0.003
3	7	0.041	0.067	0.068	0.001	0.068	0.002	0.019	0.025	0.381	0.002
4	2	0.098	0.107	0.116	0.009	0.119	0.013	0.081	0.106	0.081	0.014
6	3	0.093	0.083	0.083	0.000	0.083	0.000	-0.002	-0.003	-0.117	0.000
8	3	0.057	0.062	0.066	0.003	0.067	0.005	0.051	0.067	0.078	0.005
9	2	0.010	0.028	0.028	0.000	0.028	0.000	0.004	0.005	0.630	0.000
10	2	0.132	0.209	0.257	0.048	0.273	0.064	0.187	0.235	0.367	0.081
11	6	0.058	0.082	0.082	0.001	0.082	0.001	0.007	0.010	0.295	0.001
12	2	0.008	0.023	0.023	0.000	0.023	0.000	-0.002	-0.003	0.654	0.000
Mean		0.057	0.078	0.085	0.007	0.087	0.010	0.044	0.057	0.343	0.012
Standard deviation		0.044	0.056	0.071	0.016	0.076	0.021	0.061	0.077	0.290	0.026
Standard error		0.0147	0.0188	0.0236	0.0052	0.0252	0.0069	0.0204	0.0256	0.0967	0.0088

* Chr: chromosomes; Ho: heterozygosity within populations; Hs: genetic diversity within populations; Ht: overall gene diversity; Dst: gene diversity among samples; Htp: high-throughput phenotyping; Dstp: corrected Dst; Fst: fixation index; Fstp: corrected fixation index; Fis: inbreeding coefficient per overall loci; Dest: measure of population differentiation.

displayed a single-leafed branch that was distinct from the other genotypes. In cluster 2, the genetic distance between the clades was 63.4, but the genetic distance increased from 63.4 to 78.8. Cluster 3 included many genotypes, but few appeared to be different from each other. High variation in genetic distances was observed, ranging from 59.8 to 100, though some genotypes showed a shorter distance of 53.1.

Fig. 3 depicts the minimum spanning network of *A. afra* population clusters based on genetic distances between genotypes. Most *A. afra* genotypes were closely clustered, with a genetic distance ranging from 0.006 to 0.037 bp. Only a few genotypes from HO, RO, and WE were far apart from the remaining genotypes. There were three groups, one central group containing most genotypes, a second cluster in the upper left, and a third cluster in the lower left of the minimum spanning network.

The PCA analysis showed genetic variability between the *A. afra* accessions (Fig. 4). The PCA of the entire population was 11.07 with a sum of 22.15. However, the highest and lowest recorded PCA values were 21.15 and 0.79, respectively. The PCA showed that genotypes from the same location were similar. MH genotypes occurred in one quadrant, showing a positive relationship. However, genotypes from WE and HO were divergent and distant from genotypes collected at the same locations. RO genotypes occurred in the second quadrant, while most genotypes occurred in the WE quadrant. The eigenvalues revealed three clusters, consistent with the dendrogram. Similar results were obtained from a genetic diversity analysis in different populations of *A. annua* native to northern Iran (Reza et al., 2016).

Population structure and membership probability analyses grouped 80 *A. afra* genotypes into four populations (Fig. 5). Each bar in Fig. 5 signifies a distinct genotype. Admixed genotypes are indicated by multiple colours. Population 1 (HO) contained 18 pure genotypes and two admixtures. Population 2 (MH) contained 20 pure genotypes. Population 3 (WE) had 16 pure genotypes with four admixtures. Population 4 (WE) had 14 pure genotypes and six admixtures. Three groups were confirmed by the PCA ordination, where HO and WE genotypes were distinct while RO and MH had admixes with some similarities (Fig. 6).

A total of 22,948 SNPs were generated from the 80 *A. afra* genotypes (Table 1). Polymorphic loci per chromosome varied from 1 to 7. Some of chromosomes shared the same loci, with chromosomes 4, 9, 10, and 12 sharing loci 2. A similar occurrence was observed on chromosomes 6 and 8, which shared loci 3. Chromosomes 1, 3, and 11 occurred on loci 5, 6, and 7, respectively. Heterozygosity within populations ranged from 0.0078 to 0.1319 per chromosome. The highest heterozygosity within populations was found on chromosome 10 and the lowest on chromosome 12. Chromosome 10 had a higher within-population genetic diversity (Hs: 0.2085), overall gene diversity (Ht: 0.2565), sample-to-sample gene diversity (Dst: 0.0479), high-throughput phenotyping (Http: 0.2725), corrected Dst (Dstp: 0.0639), fixation index (Fst: 0.1869), corrected fixation index (Fstp: 0.2346), inbreeding coefficient per overall loci (Fis: 0.3673), and population differentiation value (Dest: 0.0808). Very low values were observed for Dst, Htp, Dstp, Fst, Fstp, Fis, and Dest on chromosomes 6 and 12. Chromosome 5 did not show any heterozygosity.

Population genetic analysis is important for detecting gene diversity within a population and identifying genetic drift. Due to their reproducibility, reliability, and independence from environmental conditions, DNA markers are used for precise genetic characterization, as well as in determining genotype origin and dispersal routes (Sütyemez et al., 2021). DArTSeq-derived SNPs are especially useful in the profiling of whole-plant genomes. In this study, a total of 22,948 SNPs were generated for *A. afra*, with a mean heterozygosity within populations of 7.7%. In studies of *A. annua*, extremely low within-population heterozygosity was observed, with variation among individual chromosomes. Gene gain or loss may be important factors influencing genetic diversity and reducing the numbers of similar genes in the host. In *Kalmia latifolia* L., the mean heterozygosity within populations was 19% (Li et al., 2020), which was higher than what we estimated here. Likewise, *A. annua*, a popular species of Asteraceae from China, showed heterozygosity of 21% (Czechowski et al., 2016).

4. Conclusions

This study outlines the importance of DArTseq-derived SNPs in characterizing the genetic diversity of *A. afra* genotypes found in different environs of southern Africa. A PCA, minimum spanning network, UPGMA-cluster dendrogram, population structure, membership probability test, and PCA ordination test confirmed the genetic relativity among genotypes from a single location. However, geographic distribution was not correlated with the genetic diversity among *A. afra* genotypes collected in the four locations. The diversity among genotypes clearly indicated that *A. afra* SNPs may occur despite the location. Given that genetic diversity enhances crop performance, our findings can be useful in germplasm management, industries producing by-products of *A. afra*, and plant breeding programs. For the long-term conservation of medicinal plants in South Africa, significant sampling and regional data are required. Furthermore, the diversity data will be utilized to conduct a comprehensive analysis of the impact of location on the medicinal properties of *A. Afra*. This will include both qualitative and quantitative assessments of phytochemicals, foliar mineral composition, as well as agronomic and morphological characteristics of the specimens collected from the studied locations.

Funding

This research was funded by the National Research Foundation (NRF) Thuthuka Post PhD Track [grant number TKK170424228587].

Declaration of Competing Interest

There is no conflict of interest to declare between authors and various institutions.

Acknowledgments

The authors thank the NRF [grant number TKK170424228587] Thuthuka Post PhD Track for their financial support as well as the late Dr. A. Mofokeng for her contributions to this project.

References

Chen, C., Miao, Y., Luo, D., Li, J., Wang, Z., Luo, M., Zhao, T., Liu, D., 2022. Sequence characteristics and phylogenetic analysis of the *Artemisia argyi* chloroplast genome. Frontiers in Plant Science 13, 906725. https://doi.org/10.3389/fpls.2022.906725.

Czechowski, T., Larson, T.R., Catania, T.M., Harvey, D., Brown, G.D., Graham, I.A., 2016. *Artemisia annua* mutant impaired in artemisinin synthesis demonstrates importance of nonenzymatic conversion in terpenoid metabolism. Proceedings of the National Academy of Sciences 113 (52), 15150–15155. https://doi.org/10.1073/pnas.1611567113.

Doyle, J.J., 1987. A rapid isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19, 11–15.

Ferreira, J.F., Simon, J.E., Janick, J., 1995. Developmental studies of *Artemisia annua*: flowering and artemisinin production under greenhouse and field conditions. Planta Medica 61 (2), 167–170. https://doi.org/10.1055/s-2006-958040.

Hu, K., Zhen, C., 2020. The complete chloroplast genome sequence of medicinal plant, *Artemisia gmelinii*. Mitochondrial DNA Part B 5 (3), 2337–2338. https://doi.org/10.1080/23802359.2020.1773955.

Kajitani, R., Toshimoto, K., Noguchi, H., Toyoda, A., Ogura, Y., Okuno, M., Yabana, M., Harada, M., Nagayasu, E., Maruyama, H., Kohara, Y., 2014. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Research 24 (8), 1384–1395. https://doi.org/10.1101/gr.170720.113.

Kalinowski, S.T., 2009. How well do evolutionary trees describe genetic relationships among populations? Heredity 102 (5), 506–513. https://doi.org/10.1038/hdy.2008.136.

Kujane, K., Sedibe, M.M., Mofokeng, A., 2019. Genetic diversity analysis of soybean ('*Glycine max* (L.)' Merr.) genotypes making use of SSR markers. Australian Journal of Crop Science 13 (7), 1113–1119. https://doi.org/10.21475/ajcs.19.13.07.p1638.

Li, H., Chappell, M., Zhang, D., 2020. Assessing genetic diversity and population structure of *Kalmia latifolia* L. in the eastern United States: an essential step towards breeding for adaptability to south eastern environmental conditions. Sustainability 12 (19), 8284. https://doi.org/10.3390/su12198284.

Liu, N.Q., Van der Kooy, F., Verpoorte, R., 2009. *Artemisia afra*: a potential flagship for African medicinal plants. South African Journal of Botany 75 (2), 185–195. https://doi.org/10.1016/j.sajb.2008.11.001.

Melville, J., Haines, M.L., Boysen, K., Hodkinson, L., Kilian, A., Smith Date, K.L., Potvin, D.A., Parris, K.M., 2017. Identifying hybridization and admixture using SNPs: application of the DArTseq platform in phylogeographic research on vertebrates. Royal Society Open Science 4 (7), 161061. https://doi.org/10.1098/rsos.161061.

Nigam, M., Atanassova, M., Mishra, A.P., Pezzani, R., Devkota, H.P., Plygun, S., Salehi, B., Setzer, W.N., Sharifi-Rad, J., 2019. Bioactive compounds and health benefits of *Artemisia* species. Natural product Communications 14 (7), 1934578×19850354. https://doi.org/10.1177/1934578×19850354.

Pojskić, N., 2018. IMAF-index of major allele frequency. Genetics & Applications 2 (2), 78–81.

Raimondo, D., von Staden, L., Foden, W., Victor, J.E., Helme, N.A., Turner, R.C., Kamundi, D.A., Manyama, P.A., 2009. Red List of South African Plants. Strelitzia. South African National Biodiversity Institute, Pretoria, p. 25.

Reza, B., Reza, N.M., Mansoor, O., Sepide, T., Shamsali, R., Farhad, H.A., Mehdi, S.A., Baharak, B.S., Saeed, P., Bagher, S.K.M., 2016. The study of genetic diversity in different plant populations of *Artemisia annua* L. native to northern Iran by molecular marker AFLP. Journal of Biodiversity and Environmental Sciences 9 (2), 51–59. https://doi.org/10.31383/ga.vol2iss2pp78-81.

Schatz, M.C., Witkowski, J., McCombie, W.R., 2012. Current challenges in de novo plant genome sequencing and assembly. Genome Biology 13 (4), 1–7. https://doi.org/10.1186/gb-2012-13-4-243.

Setshedi, K.T.A., Newete, S.W., Witkowski, E.T.F., 2022. The threat of an indigenous medicinal plant encroacher: *Artemisia afra* in klipriviersberg nature reserve, Johannesburg, South Africa. African Journal of Ecology 60, 1133–1142. https://doi.org/10.1111/aje.13057.

Shen, Q., Zhang, L., Liao, Z., Wang, S., Yan, T., Shi, P.U., Liu, M., Fu, X., Pan, Q., Wang, Y., Lv, Z., 2018. The genome of *Artemisia annua* provides insight into the evolution of Asteraceae family and artemisinin biosynthesis. Molecular Plant 11 (6), 776–788. https://doi.org/10.1016/j.molp.2018.03.015.

Sütyemez, M., Bükücü, Ş.B., Keleş, Ö., Özcan, A., Yıldırım, E., Boyacı, S., 2021. Phenological differences, genetic diversity, and population structure of genotypes obtained from seeds of kaman-1 walnut cultivar. Journal of Food Quality 1–9. https://doi.org/10.1155/2021/5579875.

Van Wyk, B.E., 2008. A broad review of commercially important southern African medicinal plants. Journal of Ethnopharmacology 119 (3), 342–355. https://doi.org/10.1016/j.jep.2008.05.029.