New Forests (2026) 57:9
https://doi.org/10.1007/s11056-025-10154-1

Genetic diversity and population structure of *Trichilia emetica* Vahl in western Kenya using ISSR markers

Elias Kiprotich Sirma[1] · Benson Ouma Nyongesa[1] · Otto George Dangasuk[1] · Ken Patrick Andiego[2]

Received: 17 October 2025 / Accepted: 8 December 2025
© The Author(s), under exclusive licence to Springer Nature B.V. 2025

Abstract
Loss of habitat and land-use changes can alter genetic variation in tropical trees, underscoring the need for molecular information to support species conservation. *Trichilia emetica* is an indigenous multipurpose species in Kenya that is widely used in restoration programs, yet its genetic variability remains insufficiently documentedl. This study determined the genetic diversity and structure of *T. emetica* populations in western Kenya using 15 Inter-Simple Sequence Repeat (ISSR) markers. A total of 171 DNA fragments were produced, of which 94.65% were polymorphic. The overall, gene diversity was moderate (He = 0.15; I = 0.22), and diversity levels varied across populations. Relatively higher diversity was detected in Nandi (He = 0.24; %*P* = 78.31)) and Kakamega (He = 0.18; %*P* = 75.00), whereas Kisumu and Siaya showed comparatively lower estimates (He = 0.10; %*P* = 31.25%). Analysis of molecular variance (AMOVA) showed that most variation occurred within populations (65%) with moderate among population differentiation (ΦST = 0.35, $p < 0.001$). Six populations were grouped into three genetic clusters. These results provide a baseline for integrating genetic information into management decisions, including prioritizing relatively diverse populations for seed sourcing and applying targeted enrichment strategies in areas showing reduced diversity.

Keywords Genetic diversity · ISSR markers · Population structure · *Trichilia emetica*

Introduction

Most terrestrial biodiversity of the world is found in the forests that help provide ecological stability by regulating climate conditions, nutrients cycle, and supporting livelihoods (FAO 2025; Konrad et al. 2025). Nevertheless, uncontrolled, habitat fragmentation, logging, and

✉ Benson Ouma Nyongesa
bnyongesa@uoeld.ac.ke

[1] Department of Biological Science, University of Eldoret, P. O. Box 1125 - 30100, Eldoret, Kenya

[2] Department of Forestry and Wood Science, University of Eldoret, P. O. Box 1125-30100, Eldoret, Kenya

Published online: 24 December 2025

unsustainable use of resources have caused extensive loss of genetic variation of trees, particularly in the tropics (Soares et al. 2019). Reduction in intra-species diversity undermines the adaptive ability of the population to environmental change and the resilience of forest ecosystems (Konrad et al. 2025).

Despite the ecological and socio-economic importance of African forest ecosystems, genetic research on native trees is underrepresented compared to Asia and Latin America. For example, widely distributed species like *Afzelia africana* and *A. quanzensis* (Donkpegan et al. 2020), *Adansonia digitata* (Chládová et al. 2019), *Osyris lanceolata* (Mugula et al. 2023), *Acacia senegal* (Omondi et al. 2023), *Vitellaria paradoxa* (Attikora et al. 2024), demonstrating the growing application of molecular markers to assess genetic diversity and structure in tropical tree populations. Lowland humidity forests and arid woodlands are fragmenting at a high pace due to various land uses, including agriculture, logging, and settlement expansion. Such pressures interfere with gene flow and can result in iinbreeding, loss of genetic diversity, and reduced evolutionary capacity, as shown in studies of *Acacia Senegal* where human activities predominantly define spatial genetic structure (Omondi et al. 2023). Thus, understanding genetic diversity patterns and population structure is critical for evidence-based conservation, restoration, and sustainable management of African forest trees.

Trichilia emetica Vahl (Meliaceae), also known as Natal mahogany, is a multi-purpose tree species widely distributed in sub-Saharan Africa, from Senegal to the Red Sea, and extending eastwards to Congo and South Africa (PROTA 2025). It plays a significant ecological role in dry and moist forest ecosystems by contributing to canopy structure, nutrient cycling, and habitat stability (KEFRI, 2021). It provides good-quality timber for make furniture and construction, and its bark, seeds, and leaves are used in traditional ethnomedicine (Oyedeji-Amusa et al. 2021; Chebii et al. 2022). *T. emetica* also contains bioactive metabolites such as limonoids, triterpenoids, and flavonoids, which have antimicrobial, antioxidant, and anti-inflammatory effects (Tsomele et al. 2021; Aldholmi et al. 2024). Furthermore, its seed oil has favourable lipid profile with good oxidative stability and could serve as a future nutraceutical and cosmetic ingredient, demonstrating emerging economic potential source (Oyedeji-Amusa et al. 2021; Mabaso et al. 2025).

Morphological and molecular analyses of several genera in Meliaceae, such as *Khaya* (Bouka et al. 2022), *Swietenia* (Alcalá et al. 2015; Limongi Andrade et al. 2022), *Cedrela* (Finch et al. 2022), *Azadirachta* and *Melia* (Rawat et al. 2018), and *Toona* (Nie et al. 2025), have revealed cryptic species boundaries. The populations of *T. emetica* show variation in leaf form, canopy architecture, fruit and seed size, and flowering time across ecological zones, reflecting local adaptation and phenotypic plasticity (Akweni et al. 2021; KEFRI 2021; Tsomele et al. 2021). Although this variation indicates possible genetic differentiation, no population-level molecular analyses of *T. emetica* has yet been carried out. This absence of baseline molecular data for *T. emetica* represents a research gap that limits evidence-based conservation and restoration planning. Given its ecological functions, commercial value, and wide distribution, assessing genetic diversity and population structure in *T. emetica* is therefore essential for conservation and sustainable utilization.

Molecular markers give a precise means of revealing genetic diversity and population structure in forest trees, particularly when phenotypic variation is influenced by environmental factors. Co-dominant markers such as simple sequence repeats (SSRs) and single-nucleotide polymorphisms (SNPs) offer high resolution allele detection but are expensive and

require extensive genomic resources (Alicandri et al. 2022; Faria et al. 2024). Conversely, dominant multilocus markers like Inter Simple Sequence Repeats (ISSRs) are affordable and applicable to non-model and under-studied species (Viswanathan et al. 2018; Borah et al. 2021; Le and Le 2024). ISSR are reproducible, do not need prior sequence information, and produce highly polymorphic loci. enabling precise estimation of the genetic variation and differentiation. They have been shown to identify population structure and gene flow in Meliaceae species, such as *Melia dubia* (Rawat et al. 2018). Even though ISSRs cannot differentiate between homozygous and heterozygous loci, their high reproducibility and low-cost applicability makes them a suitable tool for generating baseline data on genetic diversity and population structure in *T. Emetica*.

Western Kenya is ecologically heterogenous, spanning lowland, mid-altitude, and highland agroecological zones that differ in rainfall, land-use, and soil fertility (Jaetzold et al. 2006). Extensive farming practices have led to land fragmentation and loss of forests (Kogo et al. 2019; Rotich and Ojwang 2021; Osewe et al. 2022), potentially disrupting pollen and seed dispersa, gene flow, population connectivity, and local adaptation (Cheptou et al. 2017; Soares et al. 2019). *T. emetica* occurs both as naturally regenerating populations in western Kenya, and as planted or managed trees in homesteads. Its distribution across contrasting ecological settings, coupled with varying levels of anthropogenic activities, makes it a suitable model for assessing genetic diversity. As such, this study aimed to determine genetic diversity and population structure of *T. emetica* in western Kenya using ISSR markers.

Materials and methods

Description of the study area

This study was carried out on six natural stands of *T. emetica* collected in Bungoma, Kakamega, Kisumu, Nandi, Siaya, and Vihiga in western Kenya, representing ecological and altitudinal variation in distribution of the species (Fig. 1). This region has varied agroecological zones, ranging from the lower midlands to upper highlands, with high variation in rainfall, temperature, and land use intensity (Jaetzold et al. 2006). The forest and woodland cover in western Kenya have been fragmented through agricultural land use and settlement (Kogo et al. 2019; Rotich and Ojwang 2021; Osewe et al. 2022), but *T. emetica* still occurs in both natural and modified environments. The locations were selected to capture such environmental diversity and to cover contrasting moisture and elevation, from the semi-arid Kisumu lowlands to the humid highlands of Nandi and Kakamega. Mean annual rainfall in the locations ranges from 1,100 mm in Bungoma to over 2,100 mm in Kakamega, while mean annual temperatures ranging from 20 to 24 °C (Jaetzold et al. 2006).

Sampling design and leaf collection

A systematic random sampling approach was used in each population to minimize the chance of sampling genetically related individuals. Naturally occurring trees were sampled 50 m apart on 300–500 m transects traversing representative patches of habitat. Only healthy and mature trees with a diameter at breast height (DBH) $\geq$ 10 cm was sampled. For each tree sampled, fresh young leaves were in the middle-canopy were harvested and placed in sterile

Fig. 1 Location of six sampled *T. emetica* populations in western Kenya. Map generated using ArcGIS Software (E. Sirma)

bags with silica gel and stored at −20 °C before DNA extraction. Twenty individuals from each population were sampled in Bungoma, Kakamega, Kisumu, Siaya, Vihiga, and Nandi.

DNA extraction and purification

Genomic DNA was extracted from 0.5 g of leaf tissue by using the CTAB protocol optimized for phenolic-rich tree species (Doyle 1991). Leaf tissues were homogenized in liquid nitrogen and powder dispensed in 1.5% CTAB buffer (100 mM Tris-HCl pH 7.5; 1.4 M NaCl; 20 mM EDTA) supplemented with 0.75 µl β-mercaptoethanol and 100 mg polyvinylpolypyrrolidone (PVPP) to remove polyphenolic impurities. The homogenate was incubated at 60 °C for 20 min, mixed with chloroform: isoamyl alcohol (24:1), and centrifuged at 13,000 rpm for 25 min. The aqueous phase was harvested, re-extracted with 10% CTAB, and precipitated with cold isopropanol. Nucleic acid was precipitated in ice-cold isopropanol and pellets washed in 70% ethanol, air-dried, and resuspended in 100 µl TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). RNA was removed by treatment with RNase A (10 mg/mL) at 65 °C for 3 h. Further purification was achieved by ethanol precipitation

and re-suspension in 50 µl TE buffer. DNA quality and quantity were ascertained by electrophoresis on 1.5% agarose gels using unmethylated lambda (λ)DNA markers and spectrophotometry (A260/A280 ratio). DNA was diluted to approximately 30 ng/µl for PCR amplification.

ISSR marker selection and PCR amplification

Fifteen ISSR primers (Table 1) previously optimized for other Meliaceae species like *Melia dubia* was used for polymorphsism screening (Rawat et al. 2018). Those primers that gave reproducible clear fragments duplicate reactions were selected for genetic diversity and population structure assessment. PCR amplifications were carried out in 25 µl reactions containing: 30 ng template DNA, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$,1× PCR buffer (with $(NH_4)_2$ SO_4), 2 U Taq DNA polymerase (Thermo Fisher Scientific), and 0.5 µM of each primer (Rawat et al. 2018). Amplifications were performed in an Eppendorf Master cycler (Germany) under the following conditions: initial denaturation at 94 °C for 4 min; 39 cycles of 94 °C for 30 s, primer-specific annealing (45–61 °C) for 1 min, and extension at 72 °C for 2 min; followed by a final extension at 72 °C for 5 min. Amplicons were mixed with 1× bromophenol blue loading dye and electrophoresed on 2% agarose gels in 1× TAE buffer at 50 V for 2.5 h. DNA fragments were visualized under UV light after staining with ethidium bromide and gel image captured by a Herolab Gel Documentation System (Germany).

Molecular data analysis

Only clear, well-resolved, and reproducible ISSR fragments were scored. Each distinct fragment was treated as a single locus and was scored manually for presence (1) or absence (0) in all individuals, following the binary data scoring approach of Rawat et al. (2018) and Borah et al. (2021). Consistently amplifying and polymorphic fragments were included in the final dataset. The resultant binary matrix was utilized for genetic diversity and population structure analyses. Percentage of polymorphic loci (%P), number of observed alleles

Table 1 ISSR primer sequences and amplification conditions used for PCR analysis of *T. emetica* populations in Western Kenya

Marker code	Sequence (5'−3')	Tm (º C)	Ta (º C)
UBC −809	AGAGAGAGAGAGAGAGG	46.6	50.0
UBC −810	GAGAGAGAGAGAGAGAT	42.9	45.0
UBC −811	GAGAGAGAGAGAGAGAC	43.3	45.0
UBC −813	CTCTCTCTCTCTCTCTT	45.0	50.4
UBC −823	TCTCTCTCTCTCTCTCC	47.5	50.0
UBC −840	GAGAGAGAGAGAGAGAYT	45.8	47.0
UBC −845	CTCTCTCTCTCTCTCTRG	43.4	47.0
UBC −847	CACACACACACACACARC	54.2	53.0
UBC −855	ACACACACACACACACYT	60.2	61.0
UBC −857	ACACACACACACACACYG	57.1	58.0
UBC −864	ATGATGATGATGATGATG	51.2	52.0
UBC −880	GGAGAGGAGAGGAGA	49.0	44.7
UBC −888	BDBCACACACACACACA	52.3	55.4
UBC −890	VHVTGTGTGTGTGTGTG	51.8	52.0
UBC −891	VHVGTGTGTGTGTGTGT	51.8	55.0

Tm=melting temperature, Ta=annealing temperature, B=C, G or T; R=A or G; Y=C or T; V=A, C or G and H=A, C or T

(Na), number of effective alleles (Ne), Shannon's Information Index (I), Nei's gene diversity (He), and the coefficient of genetic differentiation (Gst) were estimated using PopGene version 1.32 (Yeh et al. 2000). These genetic indices provide complementary perspectives on within- and among-population genetic diversity and allow comparisons of allelic richness and differentiation (Nei 1973).

Analysis of Molecular Variance (AMOVA) was carried out using GenAlEx version 6.5 (Peakall and Smouse 2012) to partition total genetic variation within and among populations. AMOVA was based on Φ-statistics, analogs of Wright's F-statistics, to quantify the amount of population differentiation and gene flow (Excoffier et al. 1992). Nei's unbiased genetic distance was used to compute pairwise genetic distances among individuals and populations (Nei, 1973). The gene flow (Nm) between populations was estimated indirectly from Gst using the formula $Nm = 0.5 (1 - Gst)/Gst$ (Slatkin 1987). This index was used to infer the degree of genetic exchange and isolation among the populations of *T. emetica*. Genetic differentiation among individual collections was analyzed with Principal Coordinates Analysis (PCoA) using the genetic dissimilarity matrix in GenAlEx 6.502 (Peakall and Smouse 2012). The relationship between genetic differentiation and geographic distances among *T. emetica* populations was assessed using Isolation by Distance (IBD) analysis. Pairwise genetic distances among populations were calculated using Nei's genetic distance, and geographic distances were measured in kilometers between sampling sites. The correlation between genetic and geographic distances was evaluated using the Mantel test with 999 permutations in the GenAlEx 6.5 (Peakall and Smouse 2012).

Based on the Nei genetic distance (Tamura et al. 2013), Unweighted Pair Group Method with Arithmetic Mean (UPGMA) dendrogram of the six populations was constructed using MEGA version 6.0 software at 1,000 bootstrap values. The population structure was further inferred using a Bayesian algorithm in STRUCTURE V2.3.4. Since ISSR markers are dominant, the recessive allele model was applied. The analysis was run for $K = 1$–10 clusters with an admixture ancestry model and correlated allele frequencies. For each K, 20 independent runs were performed with a burn-in period of 50,000 iterations followed by 100,000 Markov Chain Monte Carlo (MCMC) repetitions. The most likely number of clusters (K) was estimated using the Evanno method (Evanno et al. 2005). The ΔK values were calculated and plotted in Python version 3.14.1, adapting the approach implemented in STRUCTURE Harvester.

Results

Polymorphism and performance of ISSR markers

Fifteen ISSR primers produced a total of 171 scorable fragments, 162 (94.65%) of which were polymorphic (Table 2). The number of amplicons per primer varied from 7 (UBC-857) to 19 (UBC-810). Primers UBC-809, UBC-845, UBC-847, UBC-857, UBC-864, UBC-880, and UBC-891 displayed 100% polymorphism. Polymorphism was 94.74%, 94.12%, and 93.75% for UBC-810, UBC- 840, and UBC-888, respectively. In contrast, the lowest values of polymorphism were detected in UBC-811 (84.62%), UBC-813 (87.50%), UBC-823 (87.50%), UBC-855 (90.00%), and UBC-890 (87.50%). The size of amplicons differed among primers and ranged from 169 bp in UBC-855 to 2,060 bp in UBC-880. A high frag-

Table 2 Polymorphism profile of 15 ISSR primers used for the genetic analysis of *T. emetica* populations in Western Kenya

Marker code	Total fragments	Polymorphic fragments	Polymorphism (%)	Fragment size range (bp)
UBC −809	10	10	100.00	315–389
UBC −810	19	18	94.74	224–1420
UBC −811	13	11	84.62	183–1388
UBC −813	8	7	87.50	190–1265
UBC −823	8	7	87.50	213–1705
UBC −840	17	16	94.12	175–1390
UBC −845	8	8	100.00	470–1155
UBC −847	9	9	100.00	275–1367
UBC −855	10	9	90.00	169–1368
UBC −857	7	7	100.00	276–1895
UBC −864	12	12	100.00	290–1487
UBC −880	15	15	100.00	178–2060
UBC −888	16	15	93.75	201–1654
UBC −890	8	7	87.50	221–679
UBC −891	11	11	100.00	216–1672
Total	**171**	**162**	**94.65**	

Table 3 Genetic diversity parameters based on ISSR markers across six *T. emetica* populations in Western Kenya

Population	Ecological zone	Na	Ne	I	He	%*P*
Bungoma	MH	0.78 ± 0.17	1.25 ± 0.07	0.21 ± 0.05	0.14 ± 0.04	34.38
Kakamega	MH	1.53 ± 0.15	1.29 ± 0.06	0.29 ± 0.04	0.18 ± 0.03	75.00
Kisumu	LM	0.66 ± 0.17	1.18 ± 0.06	0.16 ± 0.05	0.11 ± 0.03	31.25
Siaya	LM	0.66 ± 0.17	1.17 ± 0.07	0.15 ± 0.04	0.10 ± 0.03	31.25
Vihiga	LM	0.88 ± 0.18	1.17 ± 0.06	0.16 ± 0.04	0.10 ± 0.03	43.75
Nandi	UH	1.63 ± 0.13	1.40 ± 0.07	0.36 ± 0.05	0.24 ± 0.04	78.31
Mean ± SE		1.02 ± 0.17	1.24 ± 0.06	0.22 ± 0.05	0.15 ± 0.03	48.96 ± 8.94

Na=number of observed alleles; *Ne*=number of effective alleles; *I*=Shannon's Information Index; *He*=Nei's gene diversity; *%P*=percentage of polymorphic loci. MH=Mid highland, LM=Lower Midland, UH=Upper highland

ment size range was observed with UBC-880 (178-2,060 bp), while UBC-809 produced the narrowest range of 315–389 bp.

Genetic diversity within populations

The genetic diversity indices in Table 3 were derived from ISSR marker analysis of six *T. emetica* populations in western Kenya. The number of observed alleles (Na), expressed here as the proportion of polymorphic loci, differed among populations. Na ranged from 0.66 in Siaya and Kisumu to 1.63 in Nandi, with a mean of 1.02 ± 0.17. Kakamega and Vihiga populations had intermediate Na values of 1.53 and 0.88, respectively, while Bungoma displayed a moderate value of 0.78. The number of effective alleles (Ne) ranged from 1.17 in Siaya to 1.40 in Nandi, with a mean of 1.24 ± 0.06. Kakamega and Bungoma populations registered intermediate Ne values of 1.29 and 1.25, respectively, while Kisumu (1.18) and Vihiga

(1.17) had lowest values. Shannon Information Index (I) displayed similar pattern of variation among populations, ranging from 0.15 in Siaya to 0.36 in Nandi, with a mean value of 0.22 ± 0.05. Intermediate index values of 0.29 and 0.21, were observed in Kakamega and Bungoma, respectively, while lowest values were found in Kisumu (0.16) and Vihiga (0.16). The Nei's gene diversity (He) also differed among populations with a mean of 0.15 ± 0.03. The lowest He values were found in Siaya and Vihiga at 0.10, followed by Kisumu (0.11) and Bungoma (0.14), which registered relatively high gene diversity. Kakamega (0.18) was moderate for the He value, while Nandi (0.24) had the highest gene diversity among all the populations.

The percentage of polymorphic loci (%P) varied among the six *T. emetica* populations. It ranged from 31.25% in both Siaya and Kisumu populations to 78.31% in Nandi, a mean of 48.96% ± 8.94. Kakamega population had a relatively higher percentage of polymorphism (75.00%) like that in Nandi (78.31%), but lower percentages of polymorphism were observed in Siaya and Kisumu at 31.25%. Bungoma and Vihiga populations had 34.38 and 43.75% levels of polymorphism, respectively.

Molecular genetic differentiation

65% (65%) of total genetic variation was found within populations and 35% was observed among populations (Table 4). The fixation index (ΦST) was estimated to be 0.35, confirming that approximately 35% of the total genetic diversity was attributed to differences between populations.

The overall genetic differentiation coefficient (Gst) of all loci was 0.27 ($p = 0.001$), as analyzed through 15 ISSR loci (Table 5). Uneven genetic differentiation was observed in each locus, with Gst ranging from 0.04 (UBC-811) to 0.46 (UBC-888). Intermediate Gst values were found for some primers like UBC-809 (0.35), UBC-810 (0.31), UBC-845 (0.35), UBC-864 (0.34), and UBC-891 (0.36). Lower differentiation estimates were obtained for UBC-847 (0.08), UBC-855 (0.05), UBC-857 (0.16), and UBC-880 (0.08). The corresponding total genetic diversity (Ht) among loci ranged from 0.02 (UBC-811) to 0.50 (UBC-864), while different primers such as UBC-810 (0.49), UBC-840 (0.50), UBC-888 (0.50), UBC-890 (0.50), and UBC-891 (0.49) had relatively high total heterozygosity. The within-population genetic diversity (Hs) values ranged from 0.02 in UBC-811 to 0.36 in UBC-840, averaging 0.24 across loci. Gene flow estimates (Nm), whose relationship is inversely correlated with Gst, were variable between loci, ranging from 0.59 (UBC-888) to 11.40 (UBC-811) and averaging 2.39. Loci like UBC-847 (Nm = 5.60), UBC-855 (Nm = 9.12), and UBC-880 (Nm = 5.67) displayed relatively higher Nm values, while UBC-

Table 4 Analysis of molecular variance showing partitioning of genetic variation among and within six *T. emetica* populations in Western Kenya based on ISSR marker data

Source of variation	Df	SS	MS	Estimated variance	Percentage of total variation (%)	ΦST	*p*-value
Among populations	5	140.72	28.143	1.29	35	0.35	< 0.001
Within populations	114	272.40	2.389	2.39	65	—	—

df=degrees of freedom; SS=sum of squares; MS=mean square; ΦST=fixation index representing the proportion of total genetic variation among populations. The ΦST value was calculated as 1.288÷3.677=0.35. Significance of variance components was tested using 999 random permutations ($p < 0.001$)

Table 5 Coefficients of genetic differentiation across 15 ISSR loci in six *T. emetica* populations from Western Kenya

Marker code	Ht	Hs	Gst	Nm*
UBC-809	0.29	0.19	0.35	0.93
UBC-810	0.49	0.34	0.31	1.13
UBC-811	0.02	0.02	0.04	11.40
UBC-813	0.24	0.14	0.41	0.71
UBC-823	0.22	0.16	0.29	1.24
UBC-840	0.50	0.36	0.29	1.25
UBC-845	0.45	0.29	0.35	0.93
UBC-847	0.28	0.26	0.08	5.60
UBC-855	0.28	0.26	0.05	9.12
UBC-857	0.15	0.13	0.16	2.63
UBC-864	0.50	0.33	0.34	0.96
UBC-880	0.14	0.13	0.08	5.67
UBC-888	0.50	0.27	0.46	0.59
UBC-890	0.50	0.28	0.44	0.64
UBC-891	0.49	0.31	0.36	0.89
Mean	0.34	0.21	0.27	2.39

Ht = total genetic diversity; *Hs* = within-population genetic diversity; *Gst* = coefficient of genetic differentiation; *Nm* = estimate of gene flow, calculated as Nm=0.5(1−Gst)/Gst

Table 6 Pairwise nei's genetic identity and genetic distance among six *T. emetica* populations from Western Kenya based on ISSR marker data

Population	Bungoma	Kakamega	Kisumu	Siaya	Vihiga	Nandi
Bungoma	****	0.9742	0.8793	0.8487	0.8330	0.7867
Kakamega	0.0262	****	0.9417	0.9097	0.9015	0.8734
Kisumu	0.1286	0.0601	****	0.9714	0.9767	0.8759
Siaya	0.1640	0.0946	0.0291	****	0.9744	0.8612
Vihiga	0.1827	0.1037	0.0236	0.0259	****	0.8612
Nandi	0.2399	0.1354	0.1325	0.1494	0.1494	****

Values above the diagonal represent Nei's genetic identity; values below the diagonal represent Nei's genetic distance

809 (Nm = 0.93), UBC-845 (Nm = 0.93), UBC-864 (Nm = 0.96), and UBC-890 (Nm = 0.64) had lower estimates.

Pairwise Nei's genetic distances among the six populations of *T. emetica* varied from 0.0236 between Vihiga and Kisumu (lower midland) to 0.2399 between Bungoma and Nandi (lower highland and upper highland), while the genetic identities differed from 0.79 to 0.98 (Table 6). Low genetic distances were reported for several population pairs, such as Kisumu (lower midland)–Siaya (lower midland) (0.0291), Siaya (lower midland – Vihiga (lower midland) (0.0259), and Kisumu (lower midland)–Vihiga (lower midland) (0.0236), all of which also showed high genetic identity values exceeding 0.97. Moderate distances were observed for Kakamega (Midland highland)–Siaya (lower midland) (0.0946), Kakamega (midland highland)–Vihiga (lower midland) (0.1037), and Bungoma (midland highland)–Kakamega (midland highland) (0.0262), with identity values of 0.91, 0.90, and 0.97, respectively. Larger genetic distances were found in the pairs that involved the Nandi population, especially Bungoma (midland highland)–Nandi (upper highland) (0.2399), Siaya (lower midland)–Nandi (upper highland) (0.1494), Vihiga (lower midland)–Nandi (upper highland) (0.1494), and Kisumu (lower midland)–Nandi (upper highland) (0.1325), and these had relatively low genetic identities of 0.79 and 0.88, respectively. Comparison involving

Bungoma (midland highland)–Siaya (lower midland) (0.16) and Bungoma (midland highland)–Vihiga (lower midland) (0.1827) also showed relatively large distances.

Population genetic structure

The principal coordinates analysis (PCoA) indicate that 120 individuals were divided into three groups (Fig. 2). Group I (green group) was mainly composed of Nandi, group II included individuals from Kisumu, Siaya and Vihiga, Group III mainly included the majority of individuals from Bungoma and Kakamega. The first three coordinates of the PCA explained 26.05%,17.31%, and 7.51% of the total variation, respectively, and jointly explained 50.87% of the total genetic variation.

The cluster analysis of the six populations based on Nei's unbiased genetic distances produced three clusters that clustered based geographical proximity (Fig. 3). Siaya, Vihiga, and Kisumu were grouped in the same cluster and showed minimum separation in branch length within the cluster. The second cluster consisted of Bungoma and Kakamega, and Nandi population was grouped in the third cluster, demonstrating its discreteness from the rest of the populations.

Bayesian clustering analysis using STRUCTURE was performed across K = 1 to K = 10 with 20 independent runs per K. The analysis revealed a marked peak at K = 3 (ΔK = 102.81), which represents the most likely number of genetic clusters (Fig. 4). Although mean LnP(D) values increased gradually with higher K values, the highest ΔK peak strongly supports the presence of three genetically distinct subpopulations in the dataset.

Isolation by distance (IBD) was analyzed to determine whether the observed structuring followed geographic separation. The Mantel test showed that genetic distance increased with geographic distance among populations (Fig. 5), described by the regression equation $y = 0.0126x - 0.2077y = 0.0126x - 0.2077$ with $R^2 = 0.4907$.

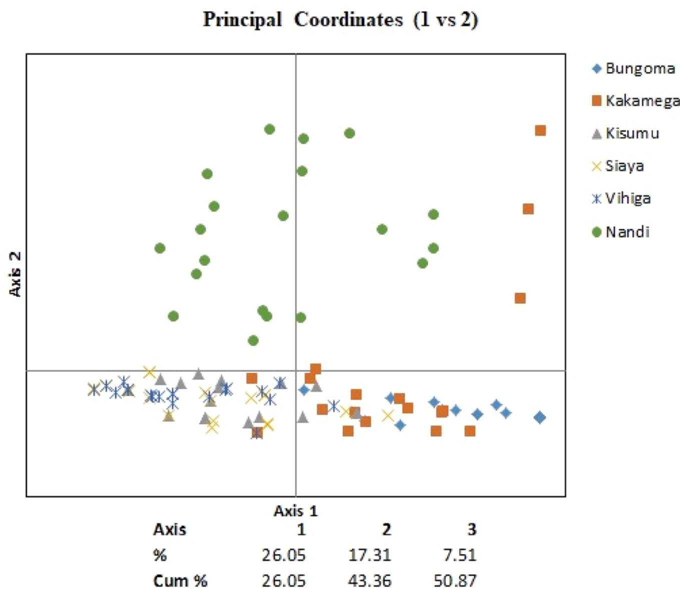
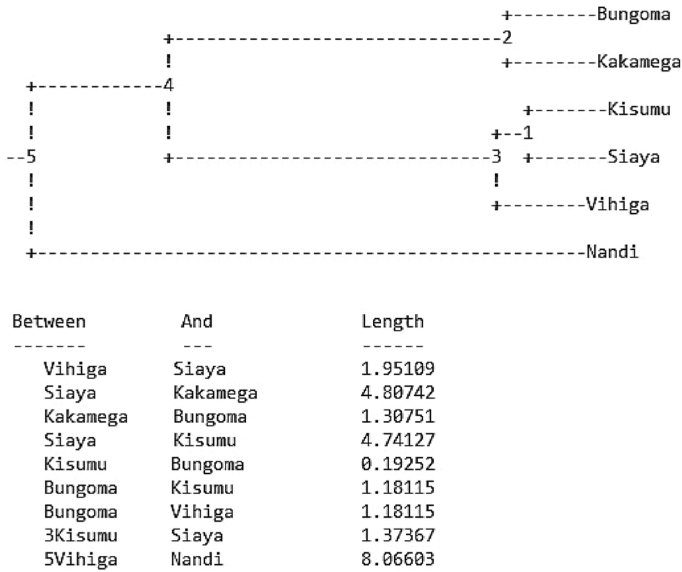


Fig. 2 Principal coordinates analysis of *T. emetica* based on the genetic distance



Between	And	Length
Vihiga	Siaya	1.95109
Siaya	Kakamega	4.80742
Kakamega	Bungoma	1.30751
Siaya	Kisumu	4.74127
Kisumu	Bungoma	0.19252
Bungoma	Kisumu	1.18115
Bungoma	Vihiga	1.18115
3Kisumu	Siaya	1.37367
5Vihiga	Nandi	8.06603

Fig. 3 UPGMA dendrogram of *T. emetica* populations based on Nei's genetic distance. The individuals were separated into three clusters

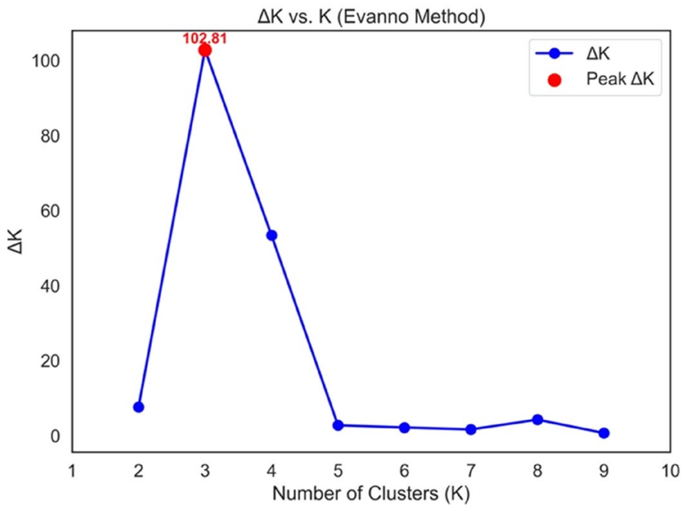


Fig. 4 Plot of ΔK values against the number of clusters (K) showing the population structure of 120 *T.emetica* individuals. The peak ΔK value indicate the most likely number of genetic clusters

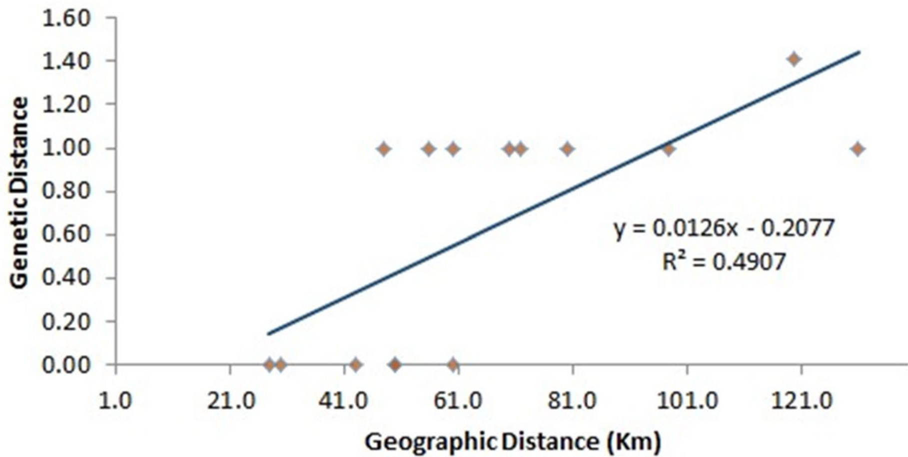


Fig. 5 Relationship between genetic distance and geographic distance mong *T. emetica* populations

Discussion

Polymorphism and performance of ISSR markers

The populations of *T. emetica* from western Kenya showed high mean percentage of polymorphic loci (94.65%), demonstrating discriminative power of ISSR markers to detect genetic variation. The high level of polymorphism observed in this study is line with 91.7% reported in *Melia dubia* (Rawat et al. 2018), 95–97% in *Pinus sylvestris* (Sheikina and Romanov 2024), 80–100% in *Osyris lanceolata* (Mugula et al. 2023) and higher than 63.8% and 46–76% in *Parashorea chinensis* and *Robinia pseudoacacia* (Li et al. 2024; Uras et al. 2024), respectively, which shows the potential of ISSR in dissecting genetic diversity in tree species. The 100% polymorphism observed in several ISSR primers such as UBC-809, UBC-845, UBC-847, UBC-857, UBC-864, UBC-880, and UBC-891 indicates that ISSR tool is highly effective in detecting genetic variation in the *T. emetica* populations. Borah et al. (2021) also reported ssimilar reliability of ISSR markers in showing high levels of genetic polymorphism in *Illicium griffithii*, further confirming their effectiveness in analyzing intra-specific variation in tree species.

The fragment sizes of amplicons ranged from 169 to 2060 bp further reinforcing the suitability of ISSR markers for assessment of genetic diversity in tree species and demonstrates wide genome variability among the *T. emetica* populations. Such high variation in fragment size shows the presence of several polymorphic loci distributed across the genome of *T. emetica* populations, indicating high allelic diversity, in line with its wide ecological distribution over lowland, mid-altitude, and highland habitats (Chebii et al. 2022; Osewe et al. 2022). Similar findings have been reported in *Khaya anthotheca* (Bouka et al. 2022) and *Cedrela odorata* (Finch et al. 2022), where wide ecological distribution was associated with high intra-specific polymorphism, indicating that tropical tree species with extensive habitats have large effective population sizes and diverse genetic backgrounds. The variation between primers in the percentage of polymorphism and the sizes of amplicons may

indicate possible underlying differences in genome structure or localized selection pressures operating throughout *T. emetica* populations. These results are in line with recent findings in other forest trees such as *Acacia Senegal*, *Parashorea chinensis*, *Juniperus* spp., and *Anadenanthera colubrina*, where variation among ISSR primers in the percentage of polymorphism and fragment sizes was attributed to differences in genome structure and localized environmental influences (Omondi et al. 2023; Xu et al. 2024; Al-Yasi and Al-Qthanin 2024).

Genetic diversity in T. emetica populations

The diversity indices in this study show the existence of genetic heterogeneity among the populations of *T. emetica* in western Kenya. The Nandi population displayed the highest diversity indices (Na = 1.63; Ne = 1.40; He = 0.24; I = 0.36; %*P* = 78.31%), indicating substantial allelic richness, evenness, and heterozygosity. These indices demonstrate that the Nandi population may be maintaining a large effective size and experiencing relatively higher gene exchange. This is possibly supported by ecological conditions such as reduced human disturbances, large canopy cover in highland forests, and species' open-pollinated reproductive system, which facilitate pollen-mediated gene exchange over wide distances. Similar trends have been reported in large continuously distributed and outcrossing Meliaceae species such as *Melia dubia* (Rawat et al. 2018), *Swietenia macrophylla* (Limongi Andrade et al. 2022) and *Khaya anthotheca* (Bouka et al. 2022), where open pollinated and ecological integrity preserved high within-population diversity.

Conversely, Siaya and Kisumu populations exhibited low genetic diversity (Na ≈ 0.66; He ≈ 0.10; %*P* = 31.25%), indicating reduced allelic variation and possible genetic erosion. These low indices correspond with documented habitat degradation, wood extraction, and land-use pressure in western Kenya (Kogo et al. 2019; Rotich and Ojwang 2021; Chebii et al. 2022; Osewe et al. 2022). Such degradation disrupts pollinator and seed dispersal routes, reduce stand density, and increase genetic drift in small, fragmented populations. Similar reductions in gene diversity caused by fragmentation has been reported in tropical tree species subjected to land use change, such as *Melia dubia* (Rawat et al. 2018), *Khaya anthotheca* (Bouka et al. 2022), and *Swietenia macrophylla* (Limongi Andrade et al. 2022). Intermediate diversity observed in Bungoma, Kakamega, and Vihiga (He = 0.14–0.19; I = 0.21–0.29; %*P* = 43–75%), indicates partial preservation of variation, likely maintained by remnant tree patches and fragmented forest corridors that support limited seed and pollen flow. Comparable patterns of diversity retention through remnant connectivity have been reported in *Tabebuia rosea* (Ruiz-González et al. 2023) and *Acacia senegal* (Omondi et al. 2023).

Genetic differentiation and population structure

T. emetica populations exhibited a structured yet genetically diverse pattern influenced by restricted gene flow. The AMOVA revealed that 65% of total genetic variation occurred within populations and 35% among the populations (ΦST = 0.35), while Gst (0.27, $p = 0.001$) further indicated moderate-to-high differentiation. Compared to other Meliaceae species, *T. emetica* displayed stronger structure than *Khaya anthotheca* (FST = 0.12) and *Swietenia macrophylla* (ΦST = 0.28), possibly indicating reduced connectivity among western Kenyan

populations (Alcalá et al. 2015; Bouka et al. 2022). Such differentiation likely results from habitat fragmentation, elevational gradients, and disrupted pollinator movement across altered landscapes, mechanisms reported in other tropical forest trees (Cheptou et al. 2017).

The differentiation indices ($\Phi ST \approx 0.35$; $Gst \approx 0.27$) were consistent with an indirect gene flow estimate of $Nm = 2.39$, indicating moderate connectivity among *T. emetica* populations. Because Nm is derived from Gst under the assumption of migration–drift equilibrium, an unlikely condition in long-lived tree species, we interpret this value qualitatively rather than as an exact migration rate. Thus, gene flow appears sufficient to counter drift but not strong enough to eliminate divergence entirely, suggesting that populations may be following semi-independent evolutionary trajectories in which gene exchange may be influenced by ecological or geographical factors. Comparable trends, where moderate differentiation coexists with incomplete homogenization, have been documented in *Azadirachta indica* and *Melia azedarach* (Rajarajan et al. 2024), and in *A. senegal*, *Khaya anthotheca*, and *Swietenia macrophylla* (Omondi et al. 2023; Bouka et al. 2022; Limongi Andrade et al. 2022). Variability in Gst values (0.04–0.46) across ISSR loci further supports uneven genome-wide differentiation, likely reflecting variation in mutation, selection or local gene exchange. Intermediate Gst values (e.g., UBC-809, UBC-864) may capture partially isolated genomic regions, while highly differentiated loci such as UBC-888 may reflect stronger selection or restricted recombination, a pattern also reported for *Melia dubia* and *Juniperus excelsa* (Rawat et al. 2018; Al-Yasi and Al-Qthanin 2024).

PCoA, UPGMA and STRUCTURE analyses consistently identified three genetic clusters: (i) Siaya–Vihiga–Kisumu, (ii) Bungoma–Kakamega, and (iii) Nandi. This portioning aligns with ecological and geographical gradients from lowland to highland regions, and likely reflects the influence of altitude, climate variation, and differential disturbance on reproductive connectivity. Nandi, positioned in upper highlands, showed the greatest divergence, likely shaped by geographical isolation, reduced pollen movement, and ecological differences in temperature and rainfall, patterns analogs to differentiated populations of *Osyris lanceolata* (Mugula et al. 2023), and *Parashorea chinensis* (Li et al. 2024; Xu et al. 2024). IBD further supported geographical separation as a key driver of this structure: genetic distance increased with geographical distance among populations (Mantel test; $y = 0.0126x - 0.2077$, $R^2 = 0.4907$), indicating that spatial isolation explains a substantial proportion of genetic divergence. This pattern reinforces that the strong differentiation of Nandi is consistent with limited pollen and seed dispersal rather than methodological bias. Conversely, the very low genetic distances and high identity values among Siaya, Kisumu and Vihiga are consistent with their short geographical separation and landscape connectivity within the lower midland zone, which likely facilitates ongoing gene flow and preserves shared ancestry. Bungoma and Kakamega, although in different highland zones, showed small genetic distances and moderate grouping, suggesting reduced but maintained gene flow that could occur through pollinator dispersal, seed movement along contiguous forest patches or riparian corridors, and other natural connectivity pathways.

Role of spatial distance in driving structure

Isolation by distance (IBD) was evaluated to determine whether differentiation followed geographic isolation. The Mantel test showed that genetic distance increased with geographic distance ($R^2 = 0.4907$), indicating that spatial separation contributes to restricted

allele sharing among populations. This spatial genetic structuring supports the hypothesis that both dispersal limitation and ecological gradients shape the population structure of *T. emetica*. The congruence between STRUCTURE clustering and IBD strengthens ecological interpretation, as has been observed in *Azadirachta indica* and *Melia azedarach* (Rajarajan et al. 2024). Thus, population subdivision in *T. emetica* likely reflects a combination of natural topographic heterogeneity and anthropogenic fragmentation that disrupts pollination and seed dispersal routes.

The present study reports the first molecular data on the genetic diversity and population structure of *T. emetica* in Kenya. Although the fifteen ISSR markers revealed over 94% polymorphism and effectively delineated populations into distinct clusters, the sample size of 20 individuals per population may have been inadequate to fully capture within-and among-population variation, particularly given the dominant nature of ISSR markers. Limited sampling can influence the precision of population genetic estimates such as He, ΦST, and Nm, and may underrepresent rare alleles. The dominant nature of ISSR markers did not allow distinction between homozygous and heterozygous loci, reducing precise estimation of heterozygosity, inbreeding coefficients, and allelic richness. This may lead to downward biases in He and Gst estimates. Since ISSRs can target non-coding regions, they do not fully capture adaptive loci. Integration of chloroplast genomic data and co-dominant markers (e.g., SSRs or SNPs) in future studies would allow more accurate estimation of heterozygosity, inbreeding, and adaptive genetic variation (Bouka et al. 2022; Limongi Andrade et al. 2022; Faria et al. 2024).

Conservation implications of genetic diversity patterns in T. emetica

The ISSR genetic patterns in this study have important conservation and sustainable management implications for *T. emetica* populations in western Kenya. The high level of polymorphism (94.65%) and the finding that 65% of the total variation occurs within populations indicate that substantial local genetic diversity still exists. This within-population variation is vital for maintaining adaptive potential and long-term resilience under increasing climatic and anthropogenic pressures. As climatic zones shift and forest stress intensifies conserving this diversity will be crucial for sustaining local adaptation, growth, and survival (Konrad et al. 2025). However, the relatively high population differentiation ($\Phi ST = 0.35$; $Gst = 0.27$), supported by IBD patterns, indicates a non-uniform distribution of genetic resources. Thus, population-specific management strategies are required, similar to fragmented alpine and tropical systems where restricted gene flow drives local adaptation (Cheptou et al. 2017; Soares et al. 2019).

Populations such as Nandi and Kakamega, which exhibited the highest gene diversity ($He \geq 0.18$) and the greatest percentage of polymorphic loci, represent core conservation units. This study proposes a quantitative threshold of $He \geq 0.18$ as a criterion for prioritizing in situ, ex situ seed banking, and seed production for restoration programs. These genetically rich populations should serve as primary sources for seed collection under Kenya's seed certification and forest restoration initiatives, including Kenya Forestry Research Institute (KEFRI) and county-level seed centers. Comparable strategies have been used for *Vitellaria paradoxa* in Côte d'Ivoire, where genetically diverse populations were designed as core seed sources for large-scale restoration (Attikora et al. 2024).

In contrast, populations from Siaya and Kisumu, which had low gene diversity (He ≤ 0.11) and only 31.25% polymorphic loci, are vulnerable to genetic erosion and stochastic loss. Conservation actions in these areas should emphasize genetic enrichment through enrichment planting, assisted natural regeneration, and controlled exchange of germplasm from genetically richer populations (e.g., Nandi and Kakamega). Enrichment planting can be carried out by selecting multiple superior genotypes from core populations and propagating then through grafting or cuttings. Introducing these propagated individuals into low-diversity populations will enhance allelic richness and adaptive potential. For areas where natural regeneration is limited, assisted regeneration can include soil preparation, protective measures against grazing, and systematic spacing of planted individuals to maximize survival and genetic mixing.

Mantel test results, coupled with STRUCTURE clustering, delineate three genetically coherent seed zones: (i) Siaya–Vihiga–Kisumu, (ii) Bungoma–Kakamega, and (iii) Nandi. These clusters provide a clear framework for genetic seed zoning and provenance selection, helping to ensure that restoration efforts use locally adapted materials. The establishment of seed zones aligns with Kenya's forest seed certification protocols and supports responsible sourcing of reproductive material for large-scale planting programs. Furthermore, the IBD pattern and moderate fixation index demonstrate limited gene flow and underscore the need to restore ecological corridors between fragmented forest remnants, particularly between genetically related populations such as Bungoma–Kakamega and Kisumu–Vihiga–Siaya. Restoration of these corridors can involve community-based restoration, riparian buffer planting, and linkage of remnant forest patches to facilitate pollen and seed dispersal, thereby sustaining long-term genetic connectivity and stability.

Integrating these molecular data into Kenya's restoration policy will strengthen seed sourcing, prevent maladaptation, and support resilience-based restoration aligned with the Bonn Challenge (IUCN 2011), the UN Decade on Ecosystem Restoration (2021–2030), and Sustainable Development Goal 15 (Life on Land) (FAO 2025). Such integrations are needed given ongoing erosion of tree genetic resource driven by human activities (Soares et al. 2019; Omondi et al. 2023). Conservation efforts should also consider farmers' and local communities' perceptions of superior trees and the agro-morphological variation among populations to identify and propagate the best phenotypes, further enhancing the adaptive potential and utility of restored populations.

Conclusion

Habitat modification and fragmentation have increasingly affected the genetic integrity of *T. emetica*, an important indigenous species in Kenya. This study provided baseline molecular information using ISSR markers to inform its conservation and seed sourcing. The 15 primers generated 171 bands, 94.65% of which were polymorphic. Overall genetic diversity was moderate (He = 0.15; I = 0.22), but its distribution was uneven. Populations from Nandi and Kakamega exhibited relatively higher variation (He ≥ 0.18), whereas populations from Siaya and Kisumu showed lower levels (He ≤ 0.11). Population structure analysis showed moderate to high differentiation ($\Phi ST = 0.35$; Gst = 0.27) and revealed three genetic clusters. These findings support the prioritisation of genetically rich populations (Nandi, Kakamega) for conservation and seed sourcing. In contrast, low-diversity populations may benefit from

enrichment planting and assisted regeneration using diverse germplasm. The identified genetic clusters offer a basis for seed zone development to minimise maladaptation. Further work incorporating adaptive traits could refine strategies for long-term management and restoration of *T. emetica*. In addition, it would be important to study the reproductive biology and pollinators, in order to better explain current results such as gene flow between populations.

Acknowledgements The authors are grateful to Kenya Agricutlural and Livestock Research Organization (KARLO) at Njoro for providing laboratory facility to undertake this study. We would like to thank Martin Lagat and Cyrus Kimani from KALRO-Njoro for technical support during the study. We would like to thank Ms. Samantha Cynthia Akinyi for editing the manuscript.

Author contributions The study was carried out as part of the first author's Master of Science thesis in Plant Genetics at the University of Eldoret. E.K.S: conceptualized the study, designed the methodology, and conducted data collection and analysis, drafted original manuscript. B.O.N and O.G.D: supervised the first author, critically reviewed the manuscript and approved version for submission. K.P.A: Supported the first author in sample collections and reviewed the manuscript.

Funding The authors did not receive funding from any organization for the submitted work.

Data availability All data supporting the findings of this study are available within the paper, and raw data can be obtained upon request from the first author.

Declarations

Competing interests The authors declare no competing interests.

Ethical approval This study was exempted from ethical review for human clinical trials or animal experiments, as it involved only the collection of plant leaves. The Ethics and Review Committee of the University of Eldoret reviewed and approved the study.

Consent for publication All authors have reviewed and approved the manuscript for submission to the *New Forests*.

References

Akweni AL, Zharare GE, Zimudzi C (2021) Predicting the number of fruits and the seed biomass of *Trichilia emetica* (Vahl.) in the eastern coastal region of South Africa. Trees For People 6:100138. https://doi.org/10.1016/j.tfp.2021.100138

Al-Yasi HM, Al-Qthanin R (2024) Comparing genetic differentiation and variation using ISSR and scot among *Juniper* plant markers in Saudi Arabia. Front Plant Sci 15:1356917. https://doi.org/10.3389/fpls.2024.1356917

Alcalá RE, Molina-Freaner F, Piñero D, Oyama K (2015) Genetic diversity and population structure of *Swietenia macrophylla* in fragmented landscapes of Mexico. Ecol Manage 344:1–9. https://doi.org/10.1016/j.foreco.2015.02.011

Aldholmi M, Althomali E, Aljishi F, Ahmad R, Alqathama A, Alaswad D (2024) A comparative study on the antidiabetic activity, cytotoxicity and lipid profile of *Trichilia emetica* oils. Plants 13(16):2234. https://doi.org/10.3390/plants13162234

Alicandri E, Vettraino AM, Agrimi M, Ciaffi M, Kuzminsky E (2022) Molecular markers dataset to assess the genetic diversity of oriental plane trees from historical sites in Lazio (central Italy). Data Brief 43:108100. https://doi.org/10.1016/j.dib.2022.108100

Attikora AJP, Yao SDM Dago ND, Silué S, De Clerck C, Kwibuka Y, Diarrassouba N, Alabi T, Achigan-Dako EG, Lassois L (2024) Genetic diversity and population structure of superior shea trees (*Vitellaria paradoxa* subsp. paradoxa) using SNP markers for the establishment of a core collection in Côte d'Ivoire. BMC Plant Biol 24:913. https://doi.org/10.1186/s12870-024-05055-8

Borah R, Bhattacharjee A, Rao SR, Kumar V, Sharma P, Upadhaya K, Choudhury H (2021) Genetic diversity and population structure assessment using molecular markers and SPAR approach in *Illicium griffithii*, a medicinally important endangered species of Northeast India. J Genet Eng Biotechnol 19:118. https://doi.org/10.1186/s43141-021-00211-5

Bouka GUD, Doumenge C, Ekué MRM, Daïnou K, Florence J, Degen B, Loumeto JJ, McKey D, Hardy OJ (2022) Khaya revisited: genetic markers and morphological analysis reveal six species in the widespread taxon K. anthotheca. Taxon 71(3):637–653. https://doi.org/10.1002/tax.12720

Chebii WK, Muthee JK, Kiemo JK (2022) Traditional medicine trade and uses in the surveyed medicine markets of Western Kenya. Afr Health Sci 22(4):684–697. https://doi.org/10.4314/ahs.v22i4.76

Cheptou P-O, Hargreaves AL, Bonte D, Jacquemyn H (2017) Adaptation to fragmentation: evolutionary dynamics driven by human influences. Philos Trans R Soc Lond B Biol Sci 372(1712):20160037. https://doi.org/10.1098/rstb.2016.0037

Chládová A, Kalousová M, Mandák B, Kehlenbeck K, Prinz K, Šmíd J, Van Damme P, Lojka B (2019) Genetic diversity and structure of baobab (*Adansonia digitata* L.) in southeastern Kenya. R Soc Open Sci 6(9):190854. https://doi.org/10.1098/rsos.190854

Donkpegan ASL, Piñeiro R, Heuertz M, Duminil J, Daïnou K, Doucet J-L, Hardy OJ (2020) Population genomics of the widespread African Savannah trees *Afzelia Africana* and *Afzelia quanzensis* reveals no significant past fragmentation of their distribution ranges. Am J Bot 107(3):498–509. https://doi.org/10.1002/ajb2.1449

Doyle JJ (1991) DNA protocols for plants. In: Hewitt GM, Johnston AWB, Young JPW (eds) Molecular techniques in taxonomy. NATO ASI Series, vol 57. Springer, Berlin, Heidelberg, pp 283–293. https://doi.org/10.1007/978-3-642-83962-7_18

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14(8):2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131(2):479–491. https://doi.org/10.1093/genetics/131.2.479

FAO (2025) State of the world's forest genetic resources—2025 update. Food and Agriculture Organization of the United Nations, Rome

Faria JCT, Konzen ER, Caldeira MVW, Godinho TO, Maluf LP, Moreira SO, Carvalho CS, Leal BSS, Azevedo CS, Momolli DR, Coelho GTCP, Oliveira CMB, Soares TCB (2024) Genetic resources of African Mahogany in Brazil: genomic diversity and structure of forest plantations. BMC Plant Biol 24:858. https://doi.org/10.1186/s12870-024-05565-9

Finch KN, Jones FA, Cronn RC (2022) Cryptic species diversity in a widespread Neotropical tree genus: the case of cedrela odorata. Am J Bot 109(11):1639–1653. https://doi.org/10.1002/ajb2.16064

International Union for Conservation of Nature (IUCN) (2011) *Bonn Challenge: A global effort on restoration.* Gland, Switzerland. https://www.bonnchallenge.org/. Accessed 11 Aug 2025

Jaetzold R, Schmidt H, Hornetz B, Shisanya CA (2006) Farm management handbook of Kenya, Vol. II. Ministry of Agriculture, Nairobi, Kenya

Kenya Forestry Research Institute (KEFRI) (2021) *Annual report 2020/2021.* KEFRI Publications, Nairobi

Kogo BK, Kumar L, Koech R (2019) Forest cover dynamics and underlying driving forces affecting ecosystem services in Western Kenya. Remote Sens Appl Soc Environ 14:100342. https://doi.org/10.1016/j.rsase.2019.02.007

Konrad H, George J-P, Szukala A (2025) The state of forest genetic diversity: Anthropogenic impacts and conservation initiatives. In: Lapin K, Oettel J, Braun M, Konrad H (eds) Ecological connectivity of forest ecosystems. Springer Nature, NY, pp 55–90. https://doi.org/10.1007/978-3-031-82206-3

Le S, Le TV (2024) Genetic diversity and population structure of natural provenances of *Sonneratia caseolaris* in Vietnam. J Genet Eng Biotechnol 22(1):100356. https://doi.org/10.1016/j.jgeb.2024.100356

Li X, Zhang Y, Liu H, Zhang Y, Huang J (2024) Geographical variation and genetic diversity of *Parashorea chinensis* germplasm resources. Front Plant Sci 15:1452521. https://doi.org/10.3389/fpls.2024.1452521

Limongi Andrade R, Pico-Mendoza J, Morillo E, Buitrón J, Meneses S, Navarrete B, Pinoargote M, Carrasco B (2022) Molecular characterization of Mahogany tree (Swietenia macrophylla King, Meliaceae) in the remnant natural forest of Ecuador. Neotrop Biodivers 8(1):222–228. https://doi.org/10.1080/23766808.2022.2080334

Mabaso M, Buthelezi LG, Zharare GE (2025) Comparative analysis of mafura butter oils from *Trichilia emetica* and *Trichilia dregeana* extracted by screw-press from seeds collected in KwaZulu-Natal Province of South Africa. Plants 14(19):3071. https://doi.org/10.3390/plants14193071

Mugula BB, Omondi S, Curto M, Kiboi SK, Kanya JI, Egeru A, Okullo P, Meimberg H (2023) Microsatellites reveal divergence in population genetic diversity and structure of *Osyris lanceolata* in Uganda and Kenya. BMC Ecol Evol 23:73. https://doi.org/10.1186/s12862-023-02182-2

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70(12):3321–3323. https://doi.org/10.1073/pnas.70.12.3321

Nie Z, Ma J, Wang C, Tang M, Jia T, Liao G, Zhang L (2025) Comparative analysis of Chloroplast genomes in meliaceae species: insights into evolution and species identification. Front Plant Sci 16:1536313. https://doi.org/10.3389/fpls.2025.1536313

Omondi SF, Githae EW, Khasa DP (2023) Long-distance gene flow in *Acacia senegal*: hope for disturbed and fragmented populations. Ecol Evol 13(9):e10292. https://doi.org/10.1002/ece3.10292

Osewe EO, Niţă MD, Abrudan IV (2022) Assessing the fragmentation, canopy loss and Spatial distribution of forest cover in Kakamega National forest Reserve, Western Kenya. Forests 13(12):2127. https://doi.org/10.3390/f13122127

Oyedeji-Amusa MO, Sadgrove NJ, Van Wyk B-E (2021) The ethnobotany and chemistry of South African meliaceae: A review. Plants 10(9):1796. https://doi.org/10.3390/plants10091796

PROTA (2025) *Plant resources of tropical Africa (Meliaceae).* Wageningen University, The Netherlands. https://prota.prota4u.org/protav8.asp?g=pe&p=Trichilia%20emetica. Accessed 11 Aug 2025

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-An update. Bioinformatics 28(19):2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Rajarajan K, Sahu S, Handa AK, Arunachalam A, Priyadarshini E, Bharti A, Yuvaraja A, Muthamilarasn M, Anuragi H (2024) Genetic diversity and population structure of *Azadirachta indica* using simple sequence repeat (SSR) markers for genetic resource conservation and utilization. Genet Resour Crop Evol 71(8):3383–3394. https://doi.org/10.1007/s10722-023-01841-z

Rawat S, Arunkumar AN, Annapurna D, Karaba NN, Joshi G (2018) Genetic diversity of *Melia dubia* using ISSR markers for natural populations and plantations. Int J Genet 10(9):490–494. https://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000226

Rotich B, Ojwang D (2021) Trends and drivers of forest cover change in the Cherangany hills forest Ecosystem, Western Kenya. Glob Ecol Conserv 30:e01755. https://doi.org/10.1016/j.gecco.2021.e01755

Ruiz-González H, Raggio M, Labrín-Sotomayor NY, Ferrer MM, Peña-Ramírez YJ (2023) Genetic variability of *Tabebuia rosea* from plantations and remnant natural populations in the Mayan Forest. Forests 14(10):2006. https://doi.org/10.3390/f14102006

Sheikina OV, Romanov EM (2024) Variability of Scots pine (*Pinus sylvestris* L.) plus trees in the middle and upper Volga region using ISSR markers. Vavilovskii Zh Genet Selekts 28(2):148–154. https://doi.org/10.18699/vjgb-24-17

Slatkin M (1987) Gene flow and the geographic structure of natural populations. Science 236(4803):787–792. https://doi.org/10.1126/science.3576198

Soares LASS, Cazetta E, Santos LR, França DS, Gaiotto FA (2019) Anthropogenic disturbances eroding the genetic diversity of a threatened palm tree: a multiscale approach. Front Genet 10:1090. https://doi.org/10.3389/fgene.2019.01090

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30(12):2725–2729. https://doi.org/10.1093/molbev/mst197

Tsomele GF, Venter E, Wokadala OC, Jooste E, Dlamini BC, Ngobese NZ, Siwela M (2021) Structural (gross and micro), physical and nutritional properties of *Trichilia emetica* and *Trichilia dregeana* seeds. CyTA J Food 19(1):483–492. https://doi.org/10.1080/19476337.2021.1915877

Uras ME, Filiz E, Sen U, Ozyigit II (2024) Genetic diversity and phylogenetic analysis of *Robinia Pseudo-acacia* L. populations using ISSR markers, ITS1 and *trnL-F* intergenic spacer sequences. J Sci 70(1):1–13. https://doi.org/10.17221/95/2023-JFS

Viswanathan MBG, Rajasekar C, Sathish Kumar P (2018) ISSR and ITS analyses to assess genetic diversity and phylogeny to conserve an endemic and critically endangered tree, *Memecylon subcordatum*, in India. Ecol Genet Genom 7:6–12. https://doi.org/10.1016/j.egg.2018.03.004

Xu Y, Liu S, Finnegan PM, Liu F, Ali I, Zhang H, Yang M (2024) Geographical variation and genetic diversity of *Parashorea chinensis* germplasm resources. Front Plant Sci 15:1452521. https://doi.org/10.3389/fpls.2024.1452521

Yeh FC, Yang RC, Boyle TJB, Ye Z, Mao JX (2000) *POPGENE version 1.32: Microsoft Windows-based freeware for population genetic analysis.* University of Alberta and Centre for International Forestry Research, Edmonton. https://www.scienceopen.com/book?vid=2d45ad78-b140-4b66-b80f-2c9f513ec997. Accessed 11 Jun 2025

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.