GENETIC DIVERSITY AND POPULATION STRUCTURE IN *Trichilia emetica* Vahl. IN WESTERN KENYA

ELIAS KIPROTICH SIRMA

A THESIS SUBMITTED TO THE SCHOOL OF SCIENCE IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE OF MASTER OF SCIENCE IN PLANT GENETICS, UNIVERSITY OF ELDORET, KENYA.

2025

DECLARATION

Declaration by the Candidate

This thesis is my original work and has not been submitted for any academic award in any institution and shall not be reproduced in part or full, or in any format without prior written permission from the author and/or University of Eldoret.

Elias Kiprotich Sirma

______________________________ **Date**________________________

SSCI/BIO/ M/ 016/19

Approval by Supervisors

This thesis has been submitted with our approval as University supervisors.

______________________________ **Date**________________________

Prof. Otto George Dangasuk

Department of Biological Sciences

University of Eldoret, Kenya

______________________________ **Date**________________________

Dr. Benson Nyongesa

Department of Biological Sciences

University of Eldoret, Kenya

DEDICATION

I dedicate this thesis to my dear parents Mr. and Mrs. Samuel Chirchir for their constant inspiration throughout the study. To my dear wife Faith and my two children Patience and Liam, I thank you for your peace and words of encouragement during the period of Study. Not forgetting my brother Eliud Mosbei and Stephen Tanui for always being there for me.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Dr. Emily Too, Head of Department of Biological Sciences for allowing me to undertake this research.

I am greatly indebted to my supervisors, Professor Otto George Dangasuk, and Dr. Benson Nyongesa for their continuous guidance, advice, effort and invertible suggestion throughout the research.

My utmost gratitude to Mr. Martin Lagat, and Cyrus Kimani, Research Scientists (KALRO), for without their continuous support this study would not have been possible. I wish to sincerely expressmy gratitude to members of KALRO for helping me carry out my research.

My appreciation also goes to Mr. Andiego Ken Patrick, for consistently encouraging me to carry on the project from data collection to thesis writing, with diligence and hope.

I take this chance also to applaud my friends Mr. Solomon Sawe, Mr. Jacob Choge, Mr. Joshua Ngetich and Mr. Nahashon Birgen for their financial Support throughout the study.

Lastly I send my sincere appreciation to my parents and family members for encouraging and supporting me throughout the study

ABSTRACT

Trichilia emetica Vahl. Commonly known as Natal Mahogany (Family: Meliaceae) is a multipurpose tree of high ecological and socio-economic value across sub-Saharan Africa. In Kenya, however, its genetic resources remain largely uncharacterized. Current knowledge focuses mainly on ethno medicinal uses, photochemistry, and propagation biology, with no molecular data available to describe its genetic diversity or population structure. Without molecular information, it is difficult to determine whether populations are genetically connected or isolated, whether inbreeding is occurring, or whether genetic erosion is already underway. Although *T. emetica* is globally classified as Least Concern, localized population declines in Kenya indicate that its genetic base may already be narrowing, reducing its adaptability to environmental change. This lack of molecular insight is a major barrier to designing evidence-based conservation, restoration, and domestication strategies. This study sought to (i) characterize phenotypic variation of *T. emetica* using eight quantitative morphological traits across six natural populations in western Kenya and (ii) determine the genetic diversity and population structure of six natural populations of *T. emetica* in western Kenya using Inter-Simple Sequence Repeat (ISSR) markers. A total of 120 leaf samples were collected from Bungoma, Kakamega, Kisumu, Siaya, Vihiga, and Nandi counties. The morphological analysis revealed substantial variation across populations, with tree height, diameter at breast height, crown diameter, flower length, seeds per fruit, seed length, and seed dry weight differing significantly ($p < 0.05$). Variation within populations (67.33%) exceeded that among populations (32.67%), and clustering consistently distinguished Nandi as a separate group, Bungoma–Kakamega as another, and Kisumu–Siaya–Vihiga as a third. At the molecular level, 171 fragments were amplified, 162 of which were polymorphic, corresponding to a polymorphism rate of 94.65%. Diversity indices revealed Nei's gene diversity (H) of 0.34 and Shannon's index (I) of 0.51, with Nandi exhibiting the highest diversity and Kisumu–Siaya the lowest. Analysis of Molecular Variance (AMOVA) indicated that 65% of the genetic variation was distributed within populations and 35% among them. Principal coordinates analysis and Nei's genetic distance dendrograms confirmed three genetic clusters, with Nandi being genetically isolated. This study is one of the first to assess the genetic diversity of *T. emetica* in Kenya, demonstrates that the species maintains considerable diversity despite fragmentation. Therefore, conservation strategies should prioritize *in situ* protection of genetically distinct populations such as Nandi, while safeguarding all populations to preserve intra-population diversity. Ex-situ measures, including representative sampling for seed banks and nurseries, are recommended to complement in situ strategies. Future research should expand to other ecological zones and integrate adaptive trait analyses to support domestication, breeding, and sustainable utilization of this valuable species.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

TERMS AND ACRONYMS

AMOVA	Analysis of Molecular Variance
ASALs	Arid and Semi-arid Lands
Bp	Base pairs
CIA	Chloroform : Isoamyl alcohol
CITES	Convention on International Trade in Endangered Species of Wild Fauna and Flora
CTAB	Cetyltrimethylammonium bromide
dH_2O	Distilled water
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic acid
IAEA	International Atomic Energy Agency
IBA	Indole-3 butyric acid
ISSR	Inter-simple sequence repeat
IUCN	International Union for Conservation of Nature
PCR	Polymerase Chain Reaction
PVC	Polyvinylpyrollidone
PVP	Polyvinylpolypyrollidone
RAPD	Randomly Amplified Polymorphic DNA
SSR	Simple Sequences Repeats
TAE	Tris- Acetate EDTA
TBE	Tris-Borate EDTA
TE	Tris- EDTA

CHAPTER ONE

INTRODUCTION

1.1 Background information

Forests and tree genetic resources provide critical ecosystem services, sustain biodiversity, and support human livelihoods. They supply timber, non-timber forest products, medicinal compounds, and oils, while also governing climate pollution through carbon isolation, maintaining hydrological cycles, and conserving soils (Hu et al., 2022). At the heart of these functions lies tree genetic diversity, which is pivotal to the resilience of populations facing threat such as habitat loss, climate variability, and overexploitation (Mugula et al., 2023). Conservation biology emphasizes the safeguarding of genetic resources as a cornerstone for sustainable use, ecological stability and long-term adaptability.

In the midst of tropical tree families, the Meliaceae stands out for its combined economic and ecological importance. Comprising over 50 genera, including *Khaya*, *Swietenia*, *Aglaia*, and *Trichilia*, this family harbours species that provide valuable timber and diverse ecosystem services. Within this group, *Trichilia emetica* (Vahl.), commonly known as Natal mahogany, is particularly notable for its multipurpose value (Figure 1). The *T. emetica* tree provides high value timber, oils such as ‘mafura oil’ and ‘mafura butter’, and bioactive compounds with both medicinal and industrial applications (Matakala et al., 2004; Perumal et al., 2020; Aldholmi et al., 2024). Nutritionally, its seed oil, rich in palmitic and oleic acids, is comparable in quality to palm and olive, highlighting its potential in pharmaceutical and cosmetic industries (Tsomele et al., 2021;

Tsomele, 2022; Aldholmi et al., 2024). The species also plays an integral role in conventional therapy, where it’s roots, leaves, barks, and seeds are used to cure diseases such as malaria, jaundice, fevers, and pigmenation disorders (Gavamukulya & Kiwuka, 2020; Mugayi & Mukanganyama, 2024). Beyond these direct uses, *T. emetica* contributes to agroforestry systems by enriching soils, providing shade, and supporting pollinators, thereby contributing to ecological resilience and sustaining rural livelihoods.

Figure 1. *Trichilia emetica* growing in its natural habitat in western Kenya.

Photo by Elias Sirma, on 2nd of December 2024.

Ecologically, *T. emetica* is well adapted to diverse environments ranging from coastal forests, riverine ecosystems, and miombo woodlands to agroforestry systems where it provides shade, soil fertility, and habitat for pollinators (Khumalo et al., 2002; Kisepa et al., 2024). In Kenya, *T. emetica* is found in ecologically diverse zones including Kisumu, Bungoma, Kakamega and Nandi forests in the west, and Arabuko-Sokoke and coastal forests in the east (PROTA et al., 2025). These ecological ecosystems are growingly terrorized by anthropogenic pressure from agricultural expansion, deforestation, and land-use change, accelerating habitat degradation and fragmentation (Maroyi, 2017). Although the species is globally classified as Least Concern, localized population declines have been observed, raising concerns about genetic erosion, reduced population connectivity, and declining adaptive potential (A., Oyama, 2016). Propagation challenges further compound this risk, as recalcitrant seeds resist conventional storage and ex-situ conservation methods (Kioko et al., 2005; Gebashe, 2015). Collectively, these pressures threaten both the ecological functions and socio-economic benefits that *T. emetica* provides.

Trichilia emetica It plays a significant ecological role in dry and moist forest ecosystems where it helps to establish the canopy structure, nutrient cycling, and habitat stability (Kenya Forestry Research Institute [KEFRI] 2021). It offers good-quality timber, which is valued to make furniture and construction, and its bark, seeds and leaves are also used in traditional ethnomedicine to treat fever, gastrointestinal disorders, and infection of the skins (Oyedeji-Amusa et al., 2021; Chebii et al., 2022). *T. emetica* also contains bioactive metabolites like limonoids, triterpenoids, and flavonoids, which have antimicrobial, antioxidant, and anti-inflammatory effects (Tsomele et al. 2021; Aldholmi et al. 2024). Furthermore, its seed oil possesses good lipid profile with excellent oxidative stability, and could be a future nutraceutical and cosmetic ingredient source (Oyedeji-Amusa et al., 2021; Mabaso et al., 2025), and is proving to have an emerging economic payoff as a pharmaceutical and cosmetic industry feedstock.

Morphological and phenological variation between the populations of *T. emetica*, such as leaf form, canopy architecture, fruit and seed size, and flowering time, has been reported in different ecological zones, which may be evidence of local adaptation and phenotypic plasticity (Akweni et al., 2021; KEFRI 2021; Tsomele et al., 2021). However, the existing literature on morphological diversity is restricted to *T. emetica* population from South Africa, with limited or no comparable data available from other regions across its distribution range. Consequently, there is insufficient understanding of how morphological traits vary among populations occurring in different ecological and climatic conditions, especially in East Africa. This gap limits the ability to relate observed morphological differences to underlying genetic variation or environmental adaptation processes and no studies has been carried out to assess population-level morphological in this species.

Although morphological and phenological variation indicates possible genetic differentiation, no population-level molecular analysis of *T. emetica* has been conducted to confirm this. Molecular studies on other genera of 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); Cui et al., 2023), and *Toona* (Nie et al., 2025), using molecular and morphological markers have revealed cryptic species boundaries, population differentiation, and adaptive divergence influenced by ecological and anthropogenic pressures. These findings demonstrate that habitat fragmentation, selective logging, and ecological gradients can lead to loss of genetic diversity and increased differentiation among populations.

Molecular markers give a precise means of revealing genetic diversity and population structure in forest trees, in particular, when the phenotypic variation is influenced by environmental factors. Co-dominant markers like simple sequence repeats (SSRs) and single-nucleotide polymorphisms (SNPs) offer high and accurate resolution of alleles but are expensive to carry out and require relatively large amount of genomic resources (Alicandri et al., 2022; Faria et al., 2024). On the contrary, dominant multilocus markers like Inter Simple Sequence Repeat (ISSR) remain affordable and applicable to non-model and under-studied species (Viswanathan et al. 2018; Borah et al. 2021; Le and Le 2024). Specifically, ISSR prove to be invaluable since they are reproducible, do not need prior sequence information, and produce high polymorphic loci that enable one to make precise estimation of the genetic variation and differentiation. They were shown to be able to identify population structure and gene flow in Meliaceae species, such as *Melia dubia* (Rawat et al., 2018). Even though ISSRs are not able to differentiate between homozygous and heterozygous loci, which is not advantageous as co-dominant markers, their high reproducibility and applicability in low-cost genotyping make them a reasonable genomic tool in genetic conservation and management (Alicandri et al., 2022). Without genomic resources in *T. emetica*, the ISSR markers will prove useful and helpful as a means to the end of creating a baseline level of knowledge about its genetic diversity and population structure.

1.2 Statement of the problem

Trichilia emetica remains poorly studied in terms of its morphological and genetic diversity. Though observable variation does exist for some of the traits such as leaf form, canopy structure, fruit and seed size, and flowering time, no systematic morphological characterization has been undertaken in Kenya or the East African region. The few diagnostic morphological features presently used for taxonomic identification are mostly influenced by environmental conditions and hence unreliable for distinguishing variation at the population level. These features limit the understanding of true morphological diversity and adaptive variation within the species. In addition, no molecular studies have been conducted to confirm if the observed phenotypic variation reflects underlying genetic differentiation or population structure. This knowledge gap is critical, considering the fact that the species occurs in fragmented habitats that are increasingly being degraded due to anthropogenic pressures, yet its genetic connectivity, isolation, or erosion at the population level is not known. In this connection, integrated morphological and molecular data have not been developed to understand its population variability, hence limiting the formulation of appropriate conservation, restoration, and domestication strategies for *T. emetica*.

1.3 Justification of the study

This study provided a scientific and practical basis for understanding and managing the morphological and genetic diversity of *T. emetica*. By investigating the morphological variation at the population level across a range of ecological zones, it identified diagnostic traits, patterns of adaptive variation, and the relationship of observable phenotypic differences to environmental conditions and potential genetic differentiation. The establishment of a morphological and molecular baseline data at the population level in Kenya brought into view the genetic structure, population connectivity, and levels of diversity of this species to allow for the detection of genetic variation and the identification of populations that are resilient. The integration of morphological and molecular data improved the reliability of conservation, restoration, and domestication strategies by ensuring that seed sourcing, propagation, and reforestation programs were based on genetically and phenotypically diverse materials. The findings also informed the development of enhanced propagation protocols, supported selection of superior genotypes for economic uses, and informed policy decisions aligned with national biodiversity priorities and global frameworks such as the Convention on Biological Diversity and the Sustainable Development Goals. The study strengthened the scientific basis for the conservation of *T. emetica*, maintenance of its adaptive potential, and its promotion toward realizing ecological and socio-economic benefits through sustainable utilization.

1.4 Objectives of the Study

1.4.1 General Objective

To determine the genetic diversity and population structure of *T. emetica* in western Kenya, in order to inform evidence-based conservation and sustainable species utilization strategies.

1.4.2 Specific Objectives

1. To characterize phenotypic variation of *T. emetica* using a set of eight quantitative morphological traits across six natural populations in western Kenya
2. To determine the genetic diversity and population structure within and among six natural populations of *T. emetica* in western Kenya using Inter-Simple Sequence Repeat (ISSR) markers, generating molecular data and diversity indices.

1.5 Research hypotheses

1. (H_1) Phenotypic variation among *T. emetica* populations in western Kenya differs significantly across eight quantitative morphological traits.
2. (H_1) Genetic diversity within and among *T. emetica* populations in western Kenya are significant and exhibits measurable population structure as revealed by ISSR markers.

CHAPTER TWO

LITERATURE REVIEW

2.1 Overview of the family Meliaceae

The Meliaceae, commonly known as the mahogany family, is a diverse and economically important group of flowering plants comprising approximately 50-60 genera and over 600 species, predominantly distributed in tropical and subtropical ecosystems (Muellner et al., 2010; Koenen et al., 2015). Members of this family are mostly trees, with some shrubs, and share characteristic features such as pinnately compound leaves and small flowers that give rise to varied fruit types (Oballa,2022). Beyond their taxonomic diversity, Meliaceae species are ecologically integral: they provide canopy cover, regulate microclimates, and contribute to nutrient cycling and habitat heterogeneity, thereby sustaining tropical forest biodiversity.

The family is also universally acknowledged for its commercial significance. High-value timbers from *Swietenia macrophylla* and *Khaya senegalensis* dominate international furniture and veneer markets, while non-timber products, such as neem (*Azadirachta indica*) extracts, have revolutionized biopesticide development (Isman et al., 2020). These examples demonstrate the family's unique position at the interface of ecology and commerce. However, this same value has driven widespread exploitation. Overharvesting of *Swietenia* and *Khaya* has resulted in severe population declines, leading to their listing under the Convention on International Trade in Endangered Species (CITES) and prompting concerns over the long-term sustainability of tropical hardwood resources (Grogan et al., 2014).

From a conservation perspective, Meliaceae exemplify the paradox of tropical resource use: while they offer irreplaceable ecological and economic services, they are simultaneously among the most threatened by logging, agricultural expansion, and climate-driven habitat loss. Thus, understanding genetic diversity and population dynamics across different genera is imperative for depicting conservation approaches that reconcile utilization with long-term survival.

2.2 Genus *Trichilia* within Meliaceae

The genus *Trichilia* belongs to the family Meliaceae, which comprises more than 70 species of trees and shrubs that are widely distributed across tropical and subtropical regions of Africa, the Americas, and parts of Asia, with distinct centers of diversity in the Neotropics and Afro-Madadascar regions (Muellner-Riehl and Rojas-Andrés,2022). Species within this genus inhabit diverse ecological zones ranging from moist lowland forests and riparian areas to dry savannah woodlands, demonstrating remarkable ecological adaptability and morphological plasticity (Orwa et al., 2009).

2.3 Taxonomy and subspecies differentiation of *Trichilia emetica*

Trichilia emetica Vahl., commonly known as Natal mahogany, is the most widespread and socio-economically significant member of the genus. The species exists in two subspecies: *T. emetica* subsp. *emetica* and *T. emetica* subsp. *suberosa* (Degen B, 2022).

The former is distributed from Eritrea and Ethiopia to South Africa, while the latter occurs from Senegal to Uganda (Szukala, 2025) The two subspecies co-occur around Lake Victoria, where hybridization is possible (Huang, 2024) Morphologically, subsp. *suberosa* is generally smaller, sometimes shrub-like, and characterized by corkybarked twigs and lax inflorescences, whereas subsp. *emetica* tends to assume larger tree forms (Hardy, 2022)

2.4 Ecological distribution and habitat adaptation

The wide distribution of *T. emetica* across Eastern, Central, and Southern Africa, demonstrates its ecological versatility. It thrives in riparian zones, savannas, and farmland margins, where it contributes multiple ecological services, including soil stabilization, microclimate regulation, and biodiversity support through its dense crown architecture (Orwa, 2014). Natural regeneration occurs both from seed and vegetative suckers, but successful establishment is strongly influenced by canopy cover and the availability of seed trees (Faria et al., 2024)

Seed dispersal is facilitated by water and frugivorous birds, particularly hornbills (Azevedo, 2024)

The species prefers well-drained, rich soil and high groundwater levels, growing well at altitudes from 0 to 1,450 meters above sea level. In dry areas, it is typically found in riparian areas (along riverbanks), while in higher altitudes, it is a component of high forests (Heuertz, 2020).

Trichilia emetica are found at major localities of community forests in Kenya, riparian areas in Bungoma, vihiga and Siaya in Western Kenya. In Kisumu the species is found along the shores of Lake Victoria. (Hines, 2020)

Research shows that *T. emetica* populations in Kenya have been drastically reduced within major community forests management regions that had greater abundance due to its over exploitation (Osewe, 2022)

2.5 Reproductive Biology and Propagation

Trichilia emetica has complex reproductive biology, with influence both from environmental and genetic factors. Generally, the species is dioecious, with separate male and female trees, although bisexual flowers have been reported in some cases, demonstrating a certain degree of sexual plasticity (Muellner-Riehl & Rojas-Andrés, 2022). Flowering phenology varies geographically, with southern African populations flowering between August and October and fruiting from December to March, while East African populations, including those in Tanzania and Kenya, flower from July to November and fruit between February and April (Handa, 2015; KEFRI, 2021).

Pollination is generally entomophilous, with bees and other small insects being attracted to the fragrant flowers, which are rich in nectar. Fruit set and seed production are often irregular and vary among individual trees and years, with seed yields ranging from 45–65 kg per tree, although yields as high as 180 kg have been reported under favorable conditions (Kamala & Rao, 2009). Upon maturity, fleshy capsules dehisce, showing seeds covered by a red sarcotesta, attractive to birds and mammals, which are key dispersal agents (Pennington, 2016; Akweni et al., 2021).

Seeds of *T. emetica* (Figure 2) are recalcitrant; losing viability rapidly when dried and thus need to be sown immediately after harvest (Medhi, 2006). Removing the sarcotesta through maceration or washing enhances germination by reducing fungal contamination and seed dormancy (Kabir & Webb, 2008). Vegetative propagation via stem cuttings, grafting, and air-layering has been successful and provides an alternative for maintaining elite genotypes and ensuring plantations of uniformity (Benjamin & Michael, 2018; KEFRI, 2021). These attributes make *T. emetica* a promising candidate for any restoration, agroforestry, and commercial propagation programs towards its sustainable utilization and genetic resources conservation.

Figure 2. Ripe seeds of *T. emetica* (Photo by Elias Sirma, on 2nd of December 2024).

2.6 Influence of dioecy and sexual plasticity on genetic diversity in *Trichilia emetica*

Dioecy, described as the presence of separate male and female individuals, is one of the major reproductive strategies that shapes plant population genetics and adaptation. It enforces obligate outcrossing and thereby maintains high intrapopulation genetic diversity with minimal inbreeding in comparison with self-compatible or monoecious systems (Muyle et al., 2021). This mating system enhances genetic variation across dioecious taxa and strengthens selection on adaptive traits- a scenario that promotes population resilience to environmental change (Muyle et al., 2021). Functional dioecy is common among Meliaceae and other tropical tree families, sustained by animal-mediated pollen and seed dispersal that maintain local gene flow. The magnitude of this effect depends, however, on dispersal distance, population size, and landscape connectivity (Montalvão et al., 2021).

Sexual plasticity, which is occasional occurrence of bisexual flowers or individuals in predominantly dioecious species, adds further complexity. Some selfing or geitonogamy may take place, giving reproductive assurance when mates or pollinators are scarce (Khanduri et al., 2021). Though such flexibility can be stabilizing in small or isolated populations, it reduces effective outcrossing and within-population genetic diversity, albeit slightly. When rare, however, bisexuality has more of an adaptive buffering against reproductive failure with little consequence for genetic structure.

At wider spatial scales, strict dioecy along with limited dispersal generally results in enhanced interpopulation genetic differentiation due to reduced gene flow. The occasional occurrence of bisexuality can modify this pattern by introducing heterogeneity in mating systems, amplifying differentiation where selfing predominates or reducing it where bisexual individuals promote pollen-mediated connectivity. Comparable reproductive mosaics in *Salix myrsinifolia* and *Cedrus deodara* create spatially variable patterns of genetic diversity driven by demography and habitat fragmentation (Mirski et al., 2017; Khanduri et al., 2021).

In *T. emetica*, predominant dioecy likely promotes high intrapopulation diversity through outcrossing, similar to what is seen in other Meliaceae taxa, such as Trichilia spp. and Swietenia macrophylla, where animal-mediated pollen dispersal maintains high gene flow and genetic diversity (Morellato, 2004; Sebbenn et al., 2012). The reported sexual plasticity (occasional bisexual flowers) may permit local selfing or geitonogamy in the case of fragmented or pollinator-limited populations. This would mean different mating outcomes across its range: some populations maintain high genetic variation while others may be more differentiated because of isolation and reduced outcrossing. These dynamics emphasize the need for integrated studies combining morphological and molecular data in order to estimate effective outcrossing rates, sex ratios, and dispersal distances, which shape the genetic diversity and structure of T. emetica (Muyle et al., 2021; Montalvão et al., 2021; Mirski et al., 2017; Morellato, 2004; Sebbenn et al., 2012).

2.7 Growth performance and ecological interactions

Growth performance varies with ecological conditions: trees attain reproductive maturity within 6–8 years under optimal environments but may take up to 20 years in shaded sites (Ling, 2014). The species coppices readily after felling, and its leaves are consumed by herbivores, including livestock and wildlife, as well as the larvae of *Charaxes* butterflies (Anil, 2012). Such ecological interactions emphasize its role in sustaining both biodiversity and livelihoods.

2.8 Socioeconomic and ethnobotanical significance

Economically, *T. emetica* is a multipurpose species with both subsistence and commercial applications. Its timber, though moderately light, is valued for furniture and carving. Seeds yield “mafurra oil,” widely used in cosmetics, soaps, and wood finishing, with potential as a biofuel feedstock (Erasmus et al., 2020). Ethnobotanical commentary points out its medicinal importance: almost every plant part is used in conventional solutions for diseases such as malaria, fever, and gastrointestinal disorders (Maroyi, 2017). Recent surveys in Western Kenya further identify it as the foremost frequently traded medicinal species, though unsustainable harvesting, especially of roots and bark, raises concern about long-term viability (Chirchir et al., 2024).

2.9 Conservation status and research imperatives

Although currently categorized as Least Concern on the (IUCN) Red List (BGCI, 2018), *T. emetica* faces significant local pressures from habitat fragmentation, market demand, and destructive harvesting practices. This paradox—global abundance masking localized decline, reflects a conservation dilemma common to many tropical multipurpose tropical trees. Genetic erosion and structural population decline are likely outcomes if unmanaged. Therefore, studies on the species’ phenotypic variation, reproductive ecology, and genetic diversity is not only educationally important but also essential for sustainable utilization. Understanding intra- and inter-population variation provide the empirical basis and scientific foundation for guiding conservation strategies, regulating harvests, and prioritizing populations for restoration and agroforestry integration.

2.10 Importance of Genetic Diversity in Forest Tree Conservation

Genetic diversity constitutes the foundation of species' adaptability and long-term survival. It represents the primal matter for natural selection, enabling populations to respond to changing environmental conditions such as climate change, habitat fragmentation, and emerging pests and diseases (Biodivers, 2023). Populace with high genetic changeability possess a broader range of traits, which increases the possibility that some characters will survive, reproduce, and sustain population continuity under novel or extreme conditions. Conversely, small or genetically uniform populations are more vulnerable to stochastic events, inbreeding, and genetic drift, which collectively erode adaptive potential and increase extinction risk (Frankham, 2015).

The adaptive significance of genetic diversity is particularly evident under rapidly shifting selective pressures. For instance, climate change alters temperature, rainfall, and soil moisture regimes. Population with broad genetic variation is more likely to harbour individuals with traits suited to such changes, thereby buffering ecosystem against instability (Oncogene, 2017). However, genetic diversity singly, is unable to guarantee persistence. Adaptive success also depends on population size, rate of environmental change, reproductive biology, and extent of gene flow among populations (Dermatol, 2017). These factors commulatively dictate whether genetic variation can be translated into functional adaptation.

Beyond species survival, genetic diversity fuels ecological resilience and continued delivery of ecosystem services. Diverse populations contribute to ecological stability by maintaining forest functions such as carbon sequestration, hydrological regulation, and soil health necessary for human livelihoods and profitable areas namely agriculture, forestry, and medicine (Integr. Agri, 2017). For multipurpose species such as *T. emetica*, conserving genetic diversity ensures biological sustainability besides the preservation of traditional knowledge, cultural values, and economic uses associated with the species.

A large gene pool acts as a buffer against environmental variability, facilitating ongoing evolution and enabling populations to withstand ecological disturbances (Zang, 2016). Assessing genetic diversity and genetic structure is therefore a central step in conservation planning, as it clarifies the species' demographic status, identifies the drivers of decline, and guides the prioritization of populations most critical for protection (Chen et al., 2012). In contexts where conservation resources are limited, such prioritization is essential to maximize the impact of interventions.

2.11 Morphological Variation in *T. emetica* and related genera in Meliaceae family

Trichilia emetica is a medium to large tropical evergreen or, depending on local conditions, a deciduous tree. The mature individuals usually reach as high as 30 m, with cylindrical boles around 80 cm in diameter (Treviño-Garza, 2012). It has imparipinnately compound leaves with 2–6 pairs of leaflets, while its inflorescence bears dehiscent capsules that contain 2–6 seeds covered with a scarlet sarcotesta (Pennington, 2016). On the whole, with a growth rate of approximately 1–2 m annually, this species represents an important part of reforestation and agroforestry systems (Ling, 2014).

Morphological diversity is salient within *T. emetica* and related species of Trichilia in the Meliaceae family, reflecting genetic differentiation and environmental adaptation. Populations show striking variation in leaf shape, hairiness of leaflets, fruit stalk length, and architecture of crown and canopy. These attributes often follow ecological gradients such as rainfall regime, soil fertility, and habitat type. For instance, *T. dregeana* has smooth leaflets with short fruit stalks, whereas *T. emetica* generally has hairy leaflets and elongated pedicels (Stefano, 2014). Such inter- and intra-specific morphological differentiation gives important taxonomic clues and suggests adaptation to different ecological conditions.

Apart from species identification, morphological variability in *T. emetica* is also a proxy of ecological plasticity: diameter at breast height (DBH), crown form, and reproductive output vary with resource availability, disturbance regimes, and competition intensity (Zuidema et al., 2009; Hernández-Díaz et al., 2008). Field observations show that anthropogenic pressures, especially selective logging and seed harvesting, have altered morphological patterns by removing larger and reproductively superior individuals, thus leaving the size-class distributions with skewed representations and reduced representation of optimal growth forms (Zuidema et al., 2009).

Morphological traits in Meliaceae result from both genetic control and phenotypic plasticity, which is the ability of the genotype to alter the phenotypic expression of a trait in reaction to environmental conditions (Tlaxcala-Méndez et al., 2016). In *T. emetica*, this morphological plasticity is manifested through variation in leaf morphology, growth rate, and tree crown structure across different ecological zones (Cunha et al., 2013). Thus, this flexibility enables the species to persist under variable environments and hence contributes to its wide geographical distribution. However, the exact degree to which such morphological traits have a genetic basis or are environmentally determined remains unclear.

Work on related genera, such as Khaya (Bouka et al. 2022), Swietenia (Alcalá et al. 2015; Limongi Andrade et al. 2022), Cedrela (Finch et al. 2022), and Melia (Rawat et al., 2018) indicates that the morphological divergences in Meliaceae reflect underlying genetic differentiation and adaptive responses to local environments. Comparable studies for *T. emetica* are few, thus limiting the understanding of the species with respect to evolutionary adaptation and taxonomy. This makes morphological diversity in *T. emetica* a very important but hardly studied aspect of its biology, which requires combined morphological and molecular analyses in order to elucidate its adaptive significance and conservation value.

2.12 Inter- Simple Sequence Repeats

Inter-Simple Sequence Repeat (ISSR) has now become one of the most useful and cost-effective molecular marker systems for assay genetic diversity, population structure, and evolutionary relationships in plants, especially non-model tropical trees such as *Trichilia emetica*. The technique amplifies genomic regions between microsatellite loci using primers anchored in simple sequence repeats, generally 16 to 5 base pairs long. This produces a multilocus banding profile that reflect genome-wide polymorphisms without requiring prior sequence information (Viswanathan et al. 2018; Borah et al. 2021; Le and Le 2024).

ISSRs are based on a single primer with di-, tri-, or tetra-nucleotide repeats targeting the inter-repeat regions across the genome (Borah et al., 2021). The longer primers and higher annealing temperatures used lead to greater reproducibility compared to Random Amplified Polymorphic DNA (RAPD) markers and maintain simplicity at low cost (Le and Le, 2024). High discriminatory power due to multiple polymorphic loci appearing from one reaction is suitable for population genetics, gene mapping, and provenance identification in both wild and cultivated species.

Although co-dominant markers such as Simple Sequence Repeats (SSRs) and Single-Nucleotide Polymorphisms (SNPs) allow for even finer allelic resolution, they are rather expensive to develop and rely on extensive genomic data (Alicandri et al., 2022; Faria, Souza & Oliveira, 2024). In contrast, ISSRs do not require any prior genomic information and as such still offer the most practical option to genetic characterization of under-studied forest trees. Their multilocus nature enables them to sample genomic variation, hence assessing both diversity and population differentiation (Viswanathan et al., 2018; Borah et al., 2021; Le and Le, 2024).

2.13 Application of ISSRs in dissection of genetic diversity and population structures in trees

ISSR markers have among the widely used DNA markers within forest tree population genetics, assessing genetic diversity, population structure, and gene flow dynamics. The multilocus and genome-wide nature of ISSR markers is able to detect polymorphisms across both coding and non-coding regions, making them particularly effective in long-lived woody taxa where recombination and mutation accumulate over time (Borah et al., 2021). Because most forest trees have large genomes and are reproductively complex, ISSRs offer a practical and cost-effective alternative to sequence-based methods like SSRs or SNP genotyping, while still capturing the evolutionary and demographic signals underlying genetic variation.

Among the Meliaceae family, ISSR markers prove to be of particular importance. For example, in *Melia dubia*, a tropical timber species which is widely spread. clear population structure and gene flow across its natural range were revealed using an ISSR-based approach (Rawat et al., 2018). High genetic differentiation among populations along with a high proportion of polymorphic loci indicated the joint effect of restricted seed dispersal and human-mediated fragmentation on genetic connectivity. These findings are in line with patterns in other Meliaceae members, such as *Swietenia macrophylla* and *Khaya anthotheca*, where ISSR and SSR studies revealed high diversity within-populations, due to outcrossing mating systems, extensive pollen-mediated gene flow, and large effective population sizes that mitigating the effects of genetic drift (Limongi Andrade et al., 2022; Bouka et al., 2022).

Beyond Meliaceae, population genetic structures were successfully dissected using ISSR across a wide range of forest taxa with different ecological and reproductive attributes. For example, in Pinus sylvestris, Sheikina and Romanov (2024) showed high levels of polymorphism through the use of ISSR markers, which clearly distinguished genetic clusters along geographic and climatic gradients. Andiego et al. (2023) reported in in Kenyan *Osyris lanceolata*, an endangered hemiparasitic tree, that ISSRs revealed both strong within-population diversity and moderate population differentiation, reflecting restricted dispersal along with habitat fragmentation. Similar results were obtained in the studies of Parashorea chinensis and Robinia pseudoacacia, where ISSRs showed high levels of polymorphism and revealed genetic structuring related to ecological variation and historical population isolation. These studies further emphasize the power of ISSR to elucidate fine-scale population structuring even in species with contrasting reproductive biology and dispersal mechanisms.

SSR-based studies also show consistent patterns between ecological context, population size, and genetic variation in forest trees. Populations from continuous or well-connected forests would usually exhibit a high Shannon's diversity index and high percentages of polymorphic loci, reflecting good gene flow and low genetic drift-for example, Memecylon subcordatum (Viswanathan et al., 2018), *Osyris lanceolata* (Mugula et al., 2023), and *Pinus sylvestris* (Sheikina & Romanov, 2024). On the other hand, populations that are more exposed to fragmentation or selective logging often show increased genetic differentiation and lower allelic richness. These trends collectively indicate that the strength of ISSR markers is not just for detecting polymorphism but also for shedding light on the ecological and demographic processes that influence genetic diversity across heterogeneous landscapes.

In summary, ISSR markers have shown high efficiency in dissection of genetic diversity and population structure in forest trees, including species of Meliaceae. Their reproducibility, inexpensiveness, and capability of revealing multilocus genetic variation make them particularly suitable for conservation genetic studies in tropical species where genomic data are limited. Consistent evidence of high within-population diversity and moderate to strong interpopulation differentiation across diverse taxa adds credence to their utility in guiding management, conservation, and sustainable use strategies of forest genetic resources.

2.14 Genetic and Morphological Studies on Trichilia and Related Species

Despite the importance of the Meliaceae family, genetic research on this family is not well-developed. The low genome contiguity of *A. indica*, which was obtained with HiSeq short Illumina reads (Krishnan et al., 2016; Kuravadi et al., 2015), limited its utility for downstream genomic research, and the chromosome-level genome of *Toona sinensis* provided limited information on phylogeny and genetic variation (Ji, 2021). Recently, a chromosome-level genome of *A. indica* was assembled to reveal terpene biosynthesis, and the results revealed that most *A. indica*-specific terpene synthase (TPS) genes and cytochrome P450 (CYP) genes were located on chromosome 13 (Du et al., 2022).

Nonetheless, many questions related to Meliaceae features have not been resolved and require attention. In general, the high-quality genomes of *A. indica* and *M. azedarach* and the genome resequencing of many other related species would help clarify the phylogenetic relationship of Meliaceae members. A wide and complete investigation would help to uncover the genetic mechanisms behind the special features of the Meliaceae family, including timber quality and limonoid biosynthesis. Additional studies on *T. emetica* focused on the allometry of and seeds biomass (Akweni et al., 2021). This assessment of plant populations based on their girth, and size-class distribution revelas voluminous information about the plant's population structure, which helps indicate survival chances of plants in one size into next size class (Cunningham, 2021). The numbers of individuals in each tree size classes can help generate baseline data critical in understanding the changes of population structure of a species occur over time (Martin, 2015). The information on plant population regeneration trend is very valuable in resource management and sets a basis for planning of sustainable forest management practices. For conservation and proper management of the existing Trichilia, the present study may need to investigate size-class distribution and regeneration status of the cultivar in disturbed forest types for it to establish the impact of human disturbance on the species' population structure.

2.15 Natural selection, local adaptation, and sustainability

Natural selection is a fundamental evolutionary process that continually shapes the genetic composition of populations by favoring individuals with traits that enhance survival and reproductive success. Over time, this process increases the average fitness of individuals within populations. Some adaptive traits -such as resistance to widespread diseases—confer advantages across multiple environments, whereas others provide benefits only under specific ecological conditions, a phenomenon known as local adaptation. Understanding the extent and mechanisms of local adaptation is critical for conservation and management because it enhances the evolutionary potential and adaptive capacity of species in the face of environmental change (Aitken & Bemmels, 2016).

In both forest restoration and reforestation, the integration of knowledge concerning local adaptation has become imperative. Genotypes from local provenances often give better survival rates, health, and growth performance compared to those from distant populations because of their evolutionary adjustment to local climatic and edaphic conditions (Aitken & Bemmels, 2016). On the other hand, failure to consider local adaptation can reduce the success in restoration, characterized by poor establishment and hence a waste of resources. Leger and Baughman (2014) pointed out the effect in the Great Basin, where poorly adapted indigenous plant materials have resulted in poor seedling establishment despite enormous investments.

Although there has been considerable research into adaptive variation among economically important temperate species, such as Quercus species (oaks) (Thomson et al., 2010; Plomion et al., 2016), the indigenous tropical tree is poorly understood beyond its taxonomy and observational descriptions. Most tropical tree species, although ecologically or economically important, are not well characterised genetically. This is a critical knowledge gap because many of the threatened tree species, which are in decline genetically, are those with minimal commercial attention and whose documented potential is limited.

Unsustainable exploitation of forest resources in Africa has resulted in widespread deforestation, habitat fragmentation, and land degradation. In this regard, countries like Nigeria and others in the continent have suffered extensive loss of forests, decline in soil fertility, and unsustainable management of natural resources. According to the United Nations (2019), most tropical trees remain undomesticated and are still mainly sourced from the wild for medicinal purposes, food, or other commercial uses. However, these trees have vast potential to contribute to the pharmaceutical, cosmetic, food, and manufacturing industries. According to Akinnifesi et al. (2017), over 1,000 tree species have been listed as Critically Endangered by the IUCN (Oldfield, 2018), hence the pressing need for the protection of forest genetic resources.

Deforestation continues to accelerate the loss of these vital species. The Millennium Ecosystem Assessment (2015) reported a 50% global reduction in forest cover over the past three centuries, with 90% of forests already lost in 29 countries. The FAO (2016) estimated annual deforestation rates of 160,000 km² between 1990–2000 and 130,000 km² between 2000–2010, with an ongoing net loss of 6.5 million hectares of natural forests annually. Between 2000 and 2012 alone, tropical forests in Africa, Asia, and South America declined by an area equivalent to that of Spain (Hansen et al., 2017). Sub-Saharan Africa and South America remain the most affected regions (United Nations, 2017).

Unsustainable exploitation of forest resources in Africa has resulted in widespread deforestation, habitat fragmentation, and land degradation. In this regard, countries like Nigeria and others in the continent have suffered extensive loss of forests, decline in soil fertility, and unsustainable management of natural resources. According to the United Nations (2019), most tropical trees remain undomesticated and are still mainly sourced from the wild for medicinal purposes, food, or other commercial uses. However, these trees have vast potential to contribute to the pharmaceutical, cosmetic, food, and manufacturing industries. According to Akinnifesi et al. (2017), over 1,000 tree species have been listed as Critically Endangered by the IUCN (Oldfield, 2018), hence the pressing need for the protection of forest genetic resources.

2.16 Research gaps

Meliaceae family encompasses woody tree species that are very useful for timber provision and ethnomedicine services. However, most of the members of this family are taxonomically under-documented, which can be been prudent in realizing their vast potentials. Past floristic studies of genus Meliaceae reveal overlaps in basic morphological traits for example number of lateral leaves, like shape, size and number of leaflets. Literature has demonstrated that little research has been conducted on taxonomic properties of not only *T. emetica* but also other members of Melieceae such as *A. indica*, *Cederella odorata*, *K. grandifoliola*, and *Khaya sensegalensis* especially their anatomical properties.

Trichilia emetica, is a green tree of Meliaceae family, that grows about 21 m in height when fully grown, and has proved to have good adaptations and responds well to the requirements of urban greening. *T. emetica* species was based on pragmatism due to its characteristics and adaptability, such as fast growth, ever greenery and wider crowns that enable them achieve green space offering ecosystem services in the cities (Akweni et al., 2021). However, like many other species in urban setting, knowledge gap on the *T. emetica* species growth quantitative parameters and structure limiting understanding of the extent of species diversity within and among populations for complete ecosystem services provision exist. For example, various tree structures such as girth, height, crown size and shape determine the services (e.g., carbon sink and shade) provided by urban trees. The information is important because, ultimately, the ecosystem services expected from urban greening can be enhanced or limited by planning, type and management of the species (Pretzsch et al., 2015).

To reveal plant diversity, the methods currently employed encompass not only research into agro-biological traits but also biochemical-molecular assays. The inter-simple sequence repeats (ISSR) technique is one of the DNA assay methods widely used in genetic diversity studies, genetic evolution and systematic specification, and gene mapping. The current study was aimed to assess individual plants of *T. emetica* varieties for morphological traits, quality parameters and genetic diversity in order to establish the correlations between the ISSR markers and characteristics of agro-biological traits.

In summary, existing research underscores substantial morphological variability within *T. emetica* and its congeners, much of which reflects adaptation to diverse ecological conditions and anthropogenic influences. However, significant gaps remain unaddressed especially regarding the genetic basis of this variation, its role in shaping population resilience, and its integration with conservation strategies. Addressing these gaps through combined morphological and molecular approaches will be essential to understand evolutionary processes within the genus and to design effective management interventions for *T. emetica* and related Meliaceae species.

CHAPTER THREE

METHODOLOGY

3.1 Study Area and Sample Collection

A field study was conducted in six populations namely Bungoma, Kakamega, Nandi, Kisumu, Siaya, and Vihiga counties representing the entire natural distribution of the species in western Kenya having different ecological variability across altitude, rainfall, temperature, and ecological zones. Twenty mature trees samples were examined per site, each spaced at least 50 meters apart to minimize the likelihood of sampling closely related individuals.

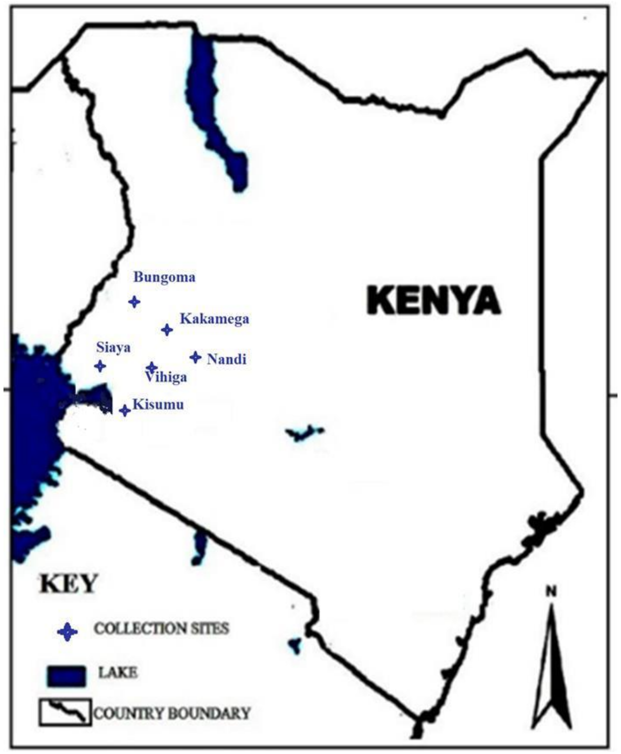


Figure 3. Location of the six sampled populations of *Trichilia emetica* in western Kenya. (Generated by Elias Sirma using ArcGIS Software)

The sites spanned altitudes ranging from 1,131 m in Kisumu to 2,047 m in Nandi, thereby covering lower midland to upper highland agro-climatic zones (Table 1; Figure 1). Annual rainfall also varied substantially, from 1,102 mm in Bungoma to 2,155 mm in Siaya, while the average temperatures for the year ranges from 20.3 °C in Bungoma to 23.9 °C in Vihiga. Bungoma (mid highland) and Kakamega (mid highland) represent cooler, high-rainfall sites, while Kisumu and Siaya (lower midland) reflect relatively warmer, moderate-rainfall environments. Vihiga (lower midland) combines higher elevation with relatively high rainfall, whereas Nandi (upper highland) represents the highest elevation population with cooler conditions and significant precipitation.

By spanning this ecological mosaic, the sampled populations provide a robust framework for examining phenotypic and genetic diversity within *T. emetica*. The variation in altitude, temperature, and rainfall across agro-climatic zones creates natural gradients that are likely to influence both morphological expression and genetic structure. Thus, this sampling criteria guarantees that the study captures not only overall diversity patterns as well as the potential role of local adaptation in shaping population differentiation.

Table 1. The characteristic of the geographic areas of *T. emetica* populations identified in western in Kenya.

Population	Trees selected	Altitude	Latitude (North)	Longitude (East)	Annual Temperature	Annual	Agroclimatic zone
Bungoma	20	1,700	0°34'10.2900	34°33’30.1536	20.3	1102	Mid highland
Kakamega	20	1,950	0°17'0.00	34°45’0	22.0	2100	Mid highland
Kisumu	20	1,131	0°5'30.12" S	34°46’4.64	23.0	1250	Lower midland
Siaya	20	1,525	0°3'45.46	34°17’16.11	23.4	2155	Lower midland
Vihiga	20	1,800	0°3'0.00	34°43’30	23.9	1900	Mid highland
Nandi	20	2,047	0°10'0.00	35°09’00	22.0	2000	Upper highland

3.2 Tree Sample Design and Sample Collection

Purposive sampling design was employed with each population to ensure unbiased representation of trees across the ecological amplitude of the sites. Sampling followed a line transect of 300-500 m, along which individual tress were selected at interval of 50 m to minimize the likelihood of sampling close relatives or clustered individuals. Each candidate tree located within the established distance threshold and meeting the inclusion criteria of having a diameter at breast height (DBH) greater than 10 cm was selected for measurement and leaf sample collection. Young healthy leaf sample were plucked from every tree for DNA analysis in laboratory using ISSR markers. Leaf samples were carefully packaged in sterile collection bags, labeled with population and tree identifiers, and refrigerated for subsequent transfer to the laboratory. These samples were later availed for DNA extraction. polymerase chain reaction (PCR) amplification using Inter-Simple Sequence Repeat (ISSR) markers was then ran. This model ensured spatially representative sampling and also provided high-quality genetic material necessary for robust assessment of intra- and inter-population genetic diversity.

3.3 Quantitative Data Generation

A total of 20 mature trees were sampled in each of the six populations, yielding 120 individuals across the study area. For each sampled tree, both vegetative and reproductive structures were collected to generate quantitative data. Tree traits recorded in the field included total height (measured in meters using hypsonometer), diameter at breast height (DBH) of trees, measured in centimeters using a diameter tape at 1.3 m above ground), and crown diameter (measured in centimeters using crown spread equipment) and height measured in meters using bole height. These traits provided baseline information on population structure and growth form.

In addition, 20 fully expanded leaves and 20 fruits were harvested per tree and transported to the laboratory for detailed morphological measurements. Leaf length and width, fruit length and diameter, and related attributes were measured in centimeters using a metric ruler. The number of seeds per fruit was recorded. To obtain seed weight data, a bulk of 1,000 seeds per population was oven dried at 27 °C for 48 hours until a constant weight was obtained and noted.

This multi-trait dataset, encompassing tree architecture, leaf morphology, fruit characteristics, and seed metrics, provides a comprehensive basis for assessing phenotypic variation within and among *T. emetica* populations. The integration of vegetative and reproductive traits ensures that both ecological adaptability and reproductive potential are captured, thereby strengthening the link between morphological diversity and potential adaptive strategies in different ecological zones.

3.4 DNA Extraction, Standardization and Quantity Determination

Extraction of genetic material was done from young, healthy leaves of *T. emetica* using a modified cetyltrimethylammonium bromide (CTAB) protocol as described by FAO/IAEA (2002). Approximately 0.5 g of fresh leaf tissue was weighed, ground with CTAB buffer using a mortar and pestle. About 100 mg of polyvinylpolypyrrolidones (pvpp) was then added to crush sample to remove polyphenolic compounds.

Six hundred microliters (600ul) of 1.5% CTAB extraction buffer (100 mM Tris-HCl, pH 7.5; 1.4 mM EDTA, pH 8.0; 1% CTAB) preheated to 60 °C, and supplemented with 0.75 µl of B – β-mercaptoethanol to prevent oxidative damage. To improve DNA quality, 100 mg of polyvinylpolypyrrolidone (PVPP) was added during grinding to remove polyphenolic compounds commonly found in woody plants. The homogenous aqueous was incubated, gentle agitated at 60 °C for 20 minutes, followed by chloroform: isoamyl alcohol (24:1, v/v) extraction to denature proteins. After centrifugation at 13,000 rpm for 25 minutes, the upper aqueous phase was carefully transferred to a new tube, treated with 10% CTAB, and re-extracted with chloroform: isoamyl alcohol.

DNA was precipitated by adding 2/3 volume of chilled isopropanol, gently mixed, and incubated until nucleic acids became visible as filaments. The DNA pellet was recovered by centrifugation at 5,000 rpm for 20 minutes, washed twice with 70% ethanol to remove salts and impurities, and air-dried at room temperature for approximately 3 hours. The dried pellet was again placed back in 100 µl of NaCl-TE buffer and incubated overnight at 55 °C to enhance solubilization.

3.5 Purification of the extracted DNA

Purification of genomic DNA followed the FAO/IAEA (2002) protocol with modifications to remove RNA and residual contaminants. Co-precipitated RNA was broken down using RNase A prepared as a 10 mg/ml stock in Tris-HCl buffer (pH 7.5, 1 mM NaCl). An aliquot of 2 µl RNase A was then added to each sample DNA, which was incubated at 65 °C for 3 hours to ensure complete RNA degradation. DNA was subsequently precipitated by adding 250 µl of cold 96% ethanol and stored overnight at –20 °C. Samples were centrifuged at 7,000 rpm for 15 minutes, the supernatant carefully discarded, and the pellet washed with 70% ethanol. A second centrifugation at 7,000 rpm for 5 minutes was performed before air-drying the pellet for 3 hours. Finally, the purified DNA was re-suspended in 100 µl sterile distilled water (SDW) and incubated at 55 °C overnight to enhance solubilization. This purification step ensured high-quality DNA suitable for downstream molecular analyses.

3.6 DNA Quantification

Quantification and quality assessment of DNA were performed using a BioPhotometer with double-stranded DNA (dsDNA) detection. DNA samples were diluted in Tris-EDTA (TE) buffer at a ratio of 2:100 µl, and absorbance was measured at 260 nm against a blank. Working samples were further diluted (1:9; DNA: PCR-grade water) and verified through electrophoresis. For gel analysis, 5 µl of diluted DNA was loaded into 2% agarose gels prepared in 0.5× Tris-Borate-EDTA (TBE) buffer. Bromophenol blue loading dye was added to track migration, and a low DNA mass ladder (Invitrogen) was included as a quantitative reference.

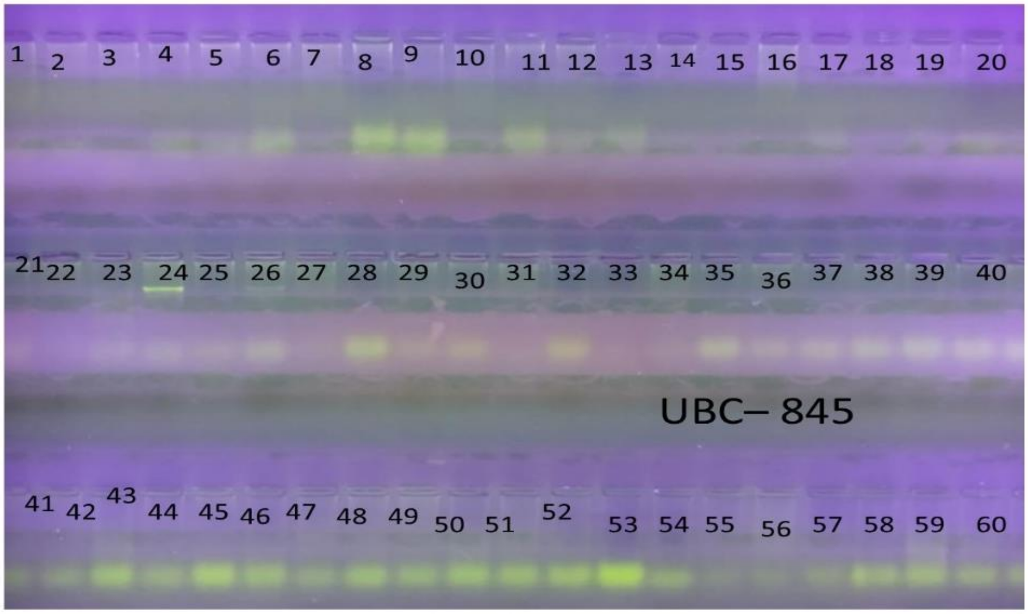


Figure 4. Gel showing DNA quantification profile

Electrophoresis was conducted at 100 V for a period of 1 hour, after which agarose gels were stained using ethidium bromide, rinsed, and visualized using a Kodak MI MT gel documentation system. DNA band intensities were compared against the reference ladder to assess concentration and integrity (Figure 2). This dual approach, spectrophotometric and gel-based quantification, ensured accurate standardization of DNA samples for polymerase chain reaction process.

3.7. PCR Amplification

The PCR was used to amplify Inter-Simple Sequence Repeat (ISSR) loci. Amplification reactions were conducted in 25 µl volumes containing 30 ng/µl of template DNA, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 2 U of Taq DNA polymerase, 10× PCR buffer, and Milli-Q water. Reactions were performed in an Eppendorf Master Cycler Gradient (Germany) conditioned as follows: initial denaturation process at 94 °C for 1 minute, followed by 39 cycles of denaturation process at 94 °C for 30 seconds, primer-specific annealing at 45–61 °C for one minute, extension at 72 °C for two minutes, and a final extension process at 72 °C for five minutes. Amplified products were mixed with bromophenol blue loading dye and separated on 2% agarose gels in 1× Tris-Acetate-EDTA (TAE) buffer at 50 V for 2.5 hours. Gels were visualized using a Herolab (Germany) gel documentation system.

To validate DNA concentration and standardize template quality, uncut, unmethylated lambda (λ) phage DNA (20 ng/µl) was used as an external reference. Serial dilutions of λ DNA (40–200 ng) were electrophoresed alongside test samples to compare band intensities and sizes. This ensured reliable standardization of DNA concentrations to 10–30 ng/µl prior to ISSR amplification. Fifteen (15) ISSR primers anchored at the 3' end were screened for their ability to generate reproducible and polymorphic banding patterns. Primers yielding clear and scorable bands were retained for genetic diversity analysis, ensuring robust assessment of variation and population differentiation among the six *T. emetica* populations.

3.8 Data analysis

3.8.1 Morphological data analysis

Quantitative morphological characters were initially tested for normality and homogeneity of variance to authenticate the assumptions of parametric tests. ANOVA was performed by employing IBM SPSS Statistics for Windows, version 28 (IBM Corp., Armonk, N.Y., USA) to quantify the extent of variation both within and among the six *T, emetica* populations. When notable variation was identified at $p < 0,05$, Duncan's Multiple Range Tests (DMRT) was applied to separate means and identify population-level differences in trait expression. Variation within and between population variations was further assessed using the Coefficient of Variations (CV %), providing an index of phenotypic diversity.

To evaluate overall patterns of morphological similarity, multivariate analysis based on Euclidean distances was performed. A dendrogram was constructed using the (UPGMA), which allowed visualization of clustering relationships inside the populations based on combined morphological traits. This integrative approach ensured robust quantification of phenotypic diversity and identification of trait-based population differentiation.

3.8.2 Molecular data analysis

Scoring of ISSR profiles was restricted to clear well-resolved and reproducible polymorphic fragments. Fragments were recorded (1) as present or (0) as absent following the binary scoring matrix (Wendel and Weeden 1989). The resulting binary data matrix was input as the basis for genetic diversity and population structure analysis.

Basic population genetic variables were analysed in POPGENE version 1.32 (Yeh et al., 2000), including (P), Shannon Information Index (I), number of effective allele (Ne), and the coefficient of differentiation (*Gst*). To partition genetic variation in populations, AMOVA) was performed in GenALEx version 1.61 (Peakall and Smouse, 2012) providing a robust measure of how diversity is distributed across hierarchical levels.

To further explore genetic relationships, pairwise genetic distances (D) were calculated based on Nei's unbiased genetic distance (Nei, 1978). These distances were used to construct a dendrogram under the (UPGMA) in MEGA version 4.0(Tamura et al., 2007), enabling visualization of population clustering patterns. Complementary multivariate analysis was conducted using Principal Coordinates Analysis (PCoA) in GenALEx version 1.61 (Peakall and Smouse, 2012) to reveal the spatial distribution of genetic variation and confirm clustering trends. This integrated analytical framework, combining population genetic indices, hierarchical variance partitioning, distance-based phylogenetic reconstruction, and multivariate ordination, ensured a comprehensive evaluation of genetic diversity and population structure in *T. emetica*.

CHAPTER FOUR

RESULTS

4.1 Morphological Variations

Quantitative assessment of eight morphological traits across six populations of *T. emetica* in western Kenya revealed substantial variation among individuals. Table 2 presents the range, mean, standard deviation (SD), and coefficient of variation (CV %) for each parameter measured. Tree height (HETG) ranged from 13.0 m to 26.3 m, with a mean of 23.97 m and a CV of 41.77%. Diameter at breast height (DBH) varied widely, ranging from 32.10 cm to 57.34 cm, with a mean of 48.58 cm and a CV of 43.98%. Crown diameter (CROD) measurements extended from 4.00 m to 8.90 m, yielding a mean of 7.71 m and a CV of 32.41%. Stem count (STEC) was relatively uniform, with values consistently recorded as 1.00, accompanied by a SD of 0.38 and a CV of 45.70%.

Reproductive traits also exhibited measurable variability. Flower length (FLLG) ranged between 1.10 cm and 1.96 cm, with a mean of 1.82 cm and the lowest coefficient of variation (62.57%). The number of seeds per fruit (SPF) varied from 3.00 to 5.00, with a mean of 3.6 and a relatively high CV of 64.94%. Seed length (SDL) ranged from 1.20 cm to 1.70 cm, with a mean of 1.45 cm and a CV of 61.62%. The dry weight of 1,000 seeds (DSWG) exhibited the widest proportional variation, ranging from 1.71 kg to 1.81 kg, with a mean of 1.76 kg and the highest CV at 65.12%.

Overall, the coefficient of variation values indicated that DSWG, SPF, and SDL exhibited the highest levels of variability among populations, while CROD showed the lowest relative variation. Across all measured parameters, the average coefficient of variation was 46.46%, highlighting the existence of notable morphological diversity within the sampled populations of *T. emetica*.

Table 2 Quantitative morphological parameters of *T. emetica* populations in western Kenya.

Parameters	Range	Mean	SD	CV %
HETG (m)	13.0 - 26.3	23.97	1.22	41.77
DBH (cm)	32.10 - 57.34	48.58	2.30	43.98
CROD (m)	4.00 - 8.90	7.71	1.31	32.41
STEC (n)	1.00 - 1.00	1.00	0.38	45.70
FLLG (cm)	1.10 - 1.96	1.82	0.19	62.57
SPF (cm)	3.00 – 5.00	3.6	2.43	64.94
SDL (cm)	1.20 – 1.70	1.45	2.56	61.62
DSWG (kg)	1.71- 1.81	1.76	1.98	65.12
				46.46

KEY: HETG= height; DBH= diameter at breast height; CROD= crown diameter; STEC= stem count; FLLG= flower length; SPF = Seeds per fruit; SDL= Seed length; DSWG= 1000 seed dry weight; SD= Standard deviation; CV= coefficient of variations

The mean values of eight morphological parameters measured in *T. emetica* across six natural populations are shown in Figure 2. Across the populations, Nandi had the highest means in most traits, including tree height (17.4%), DBH (17.2%), flower length (17.2%), seeds count per fruit (17.3%), and seed length (17.6%). Bungoma consistently followed with comparatively higher means across the same parameter. In contrast, Kisumu had the lowest parameter means with exception in crown diameter (17.4%), Stem count (16.6%). Kakamega, Siaya, and Vihiga populations showed intermediate values across most parameters, with each population displaying unique trait-specific means that positioned them between the extremes recorded in Nandi and Kisumu. Overall, the figure illustrates clear differences in mean morphological values among the six populations, with evident variation across both vegetative (height, DBH, crown diameter) and reproductive (flower length, seed number, seed length, seed weight) traits.

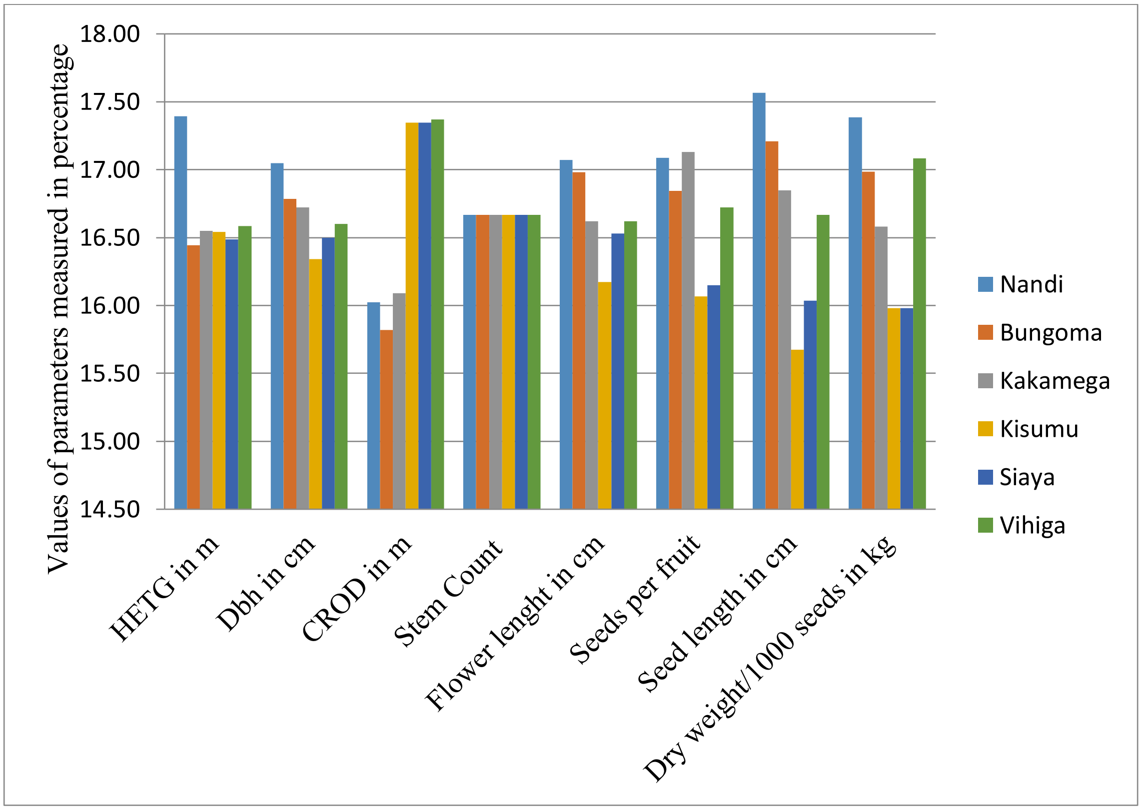


Figure 5. Mean values of eight morphological parameters of *T. emetica* populations in western Kenya

The estimated variance components revealed that 32.6% of the total morphological variation was accounted for by differences among populations, whereas a higher proportion, 67.33%, was explained by variation within populations (Table 3).

Table 3. ANOVA showing morphological variation of *T. emetica* portioned into within and among populations

Source of Variation	SS	*Df*	MS	Estimated variance	%
Among populations	1.1744	5	0.2349	9.9104	32.67
Within populations	67.2261	118	9.6037	21.7336	67.33
Total	68.4005	13		31.6440	100.00

Table 4 shows the ANOVA results for eight morphological parameters measured in *T. emetica across* six natural populations in western Kenya. The results indicate that seven of the eight traits: height, DBH, crown diameter, flower length, seeds per fruit, seeds length, and dry seed weight showed statistically significant variations among the populations at p= 0.05. In contrast, stem count parameter did not considerably vary among populations at p= 0.05.

Table 4. ANOVA for eight morphological parameters of *T. emetica* across six populations in western Kenya.

Parameter	Sum of squares	Df	F	*P*
HEGT	435.860	5	5.728 *	0.016
DBH	196.043	5	25.551*	0.010
CROD	139.877	5	21.422*	0.011
STEC	31.375	5	2.787 ^{ns}	0.127
FLLG	115.059	5	32.031*	0.050
SPF	186.179	5	36.720*	0.030
SDL	230.954	5	1.611 *	0.088
DSWG	0.352	5	1.205 *	0.091

KEY: HETG= height; DBH= diameter at breast height; CROD= crown diameter; STEC= stem count; FLLG= flower length; SPF = Seeds per fruit; SDL= Seed length; DSWG= 1000 seed dry weight. * - represents significant parameters.

Table 5 presents the portioning of mean values for the eight morphological parameters of *T. emetica* across the six populations in western Kenya where the ANOVA detected significant differences. For tree height, three distinct groupings were observed, Bungoma, Kakamega and Vihiga clustered together with similar mean values, while Nandi formed a separate cluster, and Kisumu, and Siaya grouped into a third cluster with lower mean values. For DBH, Nandi was differentiated as a single cluster with the highest mean, Bungoma and Kakamega formed the second cluster, while Kisumu, Siaya, and

Table 5. Partitioning of means for morphological parameters of *T. emetica* across six populations in western Kenya.

Population/Parameter	Nandi	Bungoma	Kakamega	Kisumu	Siaya	Vihiga	SE
HEGT	23.97^{b}	23.59^{a}	23.74^{a}	23.73^{c}	23.65^{c}	23.79^{a}	0.192
DBH	49.58^{b}	48.82^{a}	48.63^{a}	47.53^{c}	47.98^{c}	48.28^{c}	0.138
CROD	7.14^{c}	7.05^{a}	7.17^{c}	7.73^{b}	7.73^{b}	7.73^{b}	0.062
FLLG	1.90^{a}	1.89^{b}	1.87^{b}	1.80^{c}	1.84^{c}	1.86^{b}	1.079
SPF	4.19^{a}	4.18^{b}	4.12^{b}	3.94^{c}	3.96^{c}	4.10^{b}	0.971
SDL	1.95^{a}	1.91^{b}	1.87^{b}	1.74^{c}	1.78^{c}	1.85^{b}	2.453
DSWG	1.73^{a}	1.69^{b}	1.65^{b}	1.59^{c}	1.59^{c}	1.70^{a}	0.256

Mean with the same letter are similar statistically ($P<0.05$). HETG= height; DBH= diameter at breast height; CROD= crown diameter; STEC= stem count; FLLG= flower length; SPF = Seeds per fruit; SDL= Seed length; DSWG= 1000 seed dry weight; SE= standard error. Superscript letters denote grouping of DMRT.

Vihiga grouped into a third cluster with comparatively lower means. CROD partitioned the populations differently: Nandi with Kakamega grouped together, Bungoma formed an independent cluster, while Kisumu, Siaya, and Vihiga clustered into another group characterized by larger mean crown diameters. The reproductive parameters: FLLG, SPF, SDL, and DSWG showed consistent grouping patterns. In each case, Nandi formed a distinct cluster with the highest mean values. Bungoma and Kakamega were grouped together, while Kisumu, Siaya and Vihiga consistently formed a third cluster with lower mean values. Handy exception was noted in the case of DSWG, where Vihiga aligned with Nandi rather than grouping with Kisumu and Siaya. These results demonstrated that Nandi population was consistently distinct across most parameters, while Bungoma and Kakamega tended to cluster together. Kisumu, Siaya, and Vihiga generally formed a separate grouping, although Vihiga shifted clusters depending on the trait examined.

The hierarchical cluster analysis of combined morphological parameters of *T. emetica* populations revealed clear patterns of population differentiation (Figure 4). The 6 populations were divided into three major clusters. The first cluster comprised Kisumu and Siaya, which exhibited the closet morphological similarity, with Vihiga branching closely to this subgroup. The second cluster grouped Bungoma and Kakamega together, indicating their shared morphological affinities. The third cluster was represented solely by Nandi, which stood apart as a distinct group, reflecting marked divergence from the other populations.

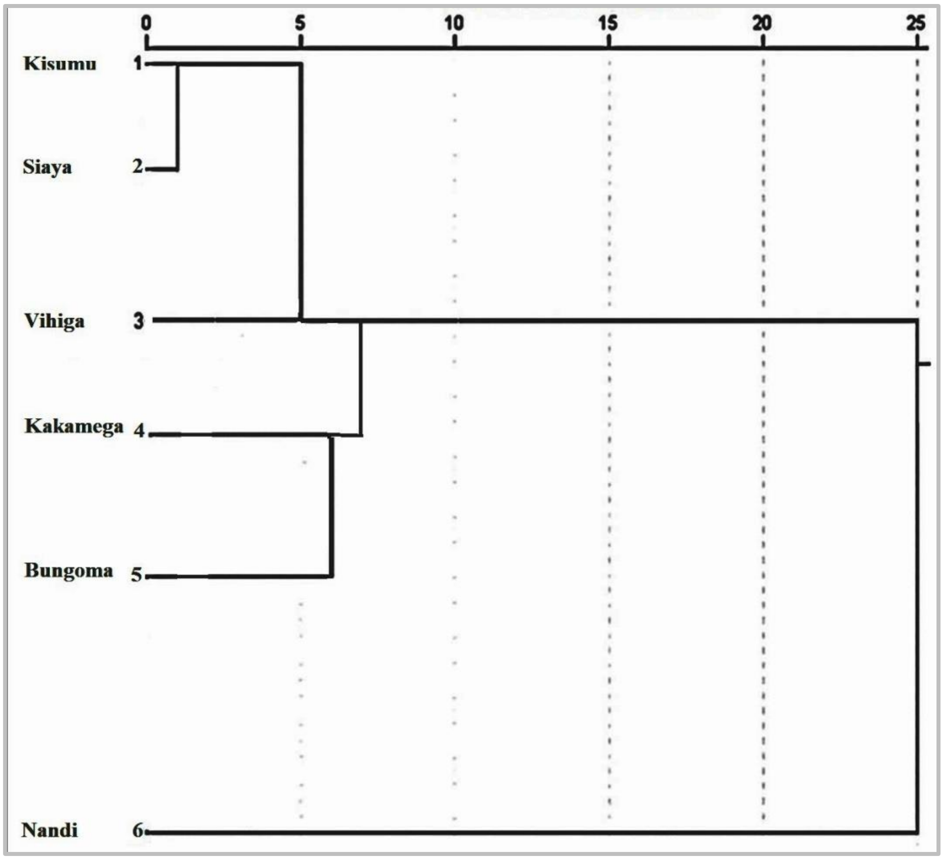


Figure 6. Hierarchical clustering based on combined morphological parameters of *T. emetica* across six populations in western Kenya.

4.2 Molecular Genetic diversity

The genetic diversity of *T. emetica* was assessed using 15 ISSR markers. The 15 primers yielded scorable bands as shown in figure 5. Table 6 provides details of the fifteen marker codes, primer sequence, melting temperature (Tm), annealing temperature (Ta), and number of amplified fragments, polymorphic fragments, percentage polymorphism (%), and the scope of amplified fragment sizes. Across all primers, a total of 171 DNA fragments were intensified ranging in size from 169 to 2,060 base pairs (bp). Of these, 162 were polymorphic, corresponding to an overall polymorphism rate of 94.65%. The polymorphic fragments numbers per primer ranged from 7 to 18.

Table 6. Fifteen ISSR marker codes, sequence, melting and annealing temperature, percentage polymorphism, and amplification range in *T. emetica*.

Marker code	Sequence (5'-3')	Tm (° C)	Optimum Ta	Total	No. of	Percentage	Range of amplifi
UBC -809	AGAGAGAGAGAGAGAGG	46.6	50.0	10	10	100	315-389
UBC -810	GAGAGAGAGAGAGAGAT	42.9	45.0	19	18	94.74	224-1420
UBC -811	GAGAGAGAGAGAGAGAC	43.3	45.0	13	11	84.62	183-1388
UBC -813	CTCTCTCTCTCTCTCTT	45.0	50.4	8	7	87.5	190-1265
UBC -823	TCTCTCTCTCTCTCTCC	47.5	50.0	8	7	87.5	213-1705
UBC -840	GAGAGAGAGAGAGAGAYT	45.8	47.0	17	16	94.12	175-1390
UBC -845	CTCTCTCTCTCTCTCTRG	43.4	47.0	8	8	100	470-1155
UBC -847	CACACACACACACACARC	54.2	53.0	9	9	100	275-1367
UBC -855	ACACACACACACACACYT	60.2	61.0	10	9	90.0	169-1368
UBC -857	ACACACACACACACACYG	57.1	58.0	7	7	100	276-1895
UBC -864	ATGATGATGATGATGATG	51.2	52.0	12	12	100	290-1487
UBC -880	GGAGAGGAGAGGAGA	49.0	44.7	15	15	100	178-2060
UBC -888	BDBCACACACACACACA	52.3	55.4	16	15	93.75	201-1654
UBC -890	VHVTGTGTGTGTGTGTG	51.8	52.0	8	7	87.5	221-679
UBC -891	VHVGTGTGTGTGTGTGT	51.8	55.0	11	11	100	216-1672
Total				**171**	**162**	**94.65**	

According to these results, all the primers revealed high polymorphism ranging from 84.62 to 100% with average of 94.65%. The primers UBC -810 generated the highest number of fragments, producing 19 total fragments of which 18 were polymorphic, with amplification sizes spanning 224-1,420 bp. UBC -840 also generated a relatively high number of polymorphic fragments (16 out of 17 total), with amplification ranging from

175-1,390 bp. Primers, including UBC-809, UBC -845, UBC -847, UBC -857, UBC 864, UBC-880, and UBC -891, yielded 100% polymorphism, though with varying number of total fragments.

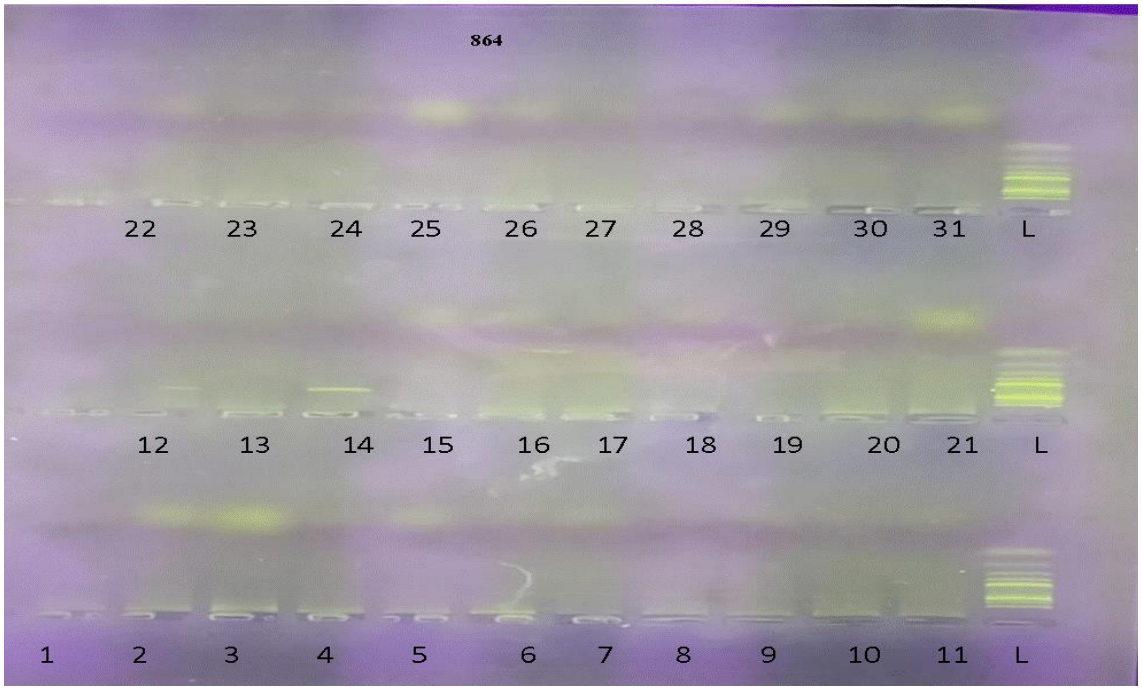


Figure 7. Gel electrophoresis pattern of Bungoma population using primers UBC (864)

By contrast, primers UBC-813, UBC-823, UBC-857, and UBC-890 amplified the fewest polymorphic bands, each producing only 7 out of 8 bands (or fewer in the case of UBC-857). Fragment size ranges also varied by primer, with UBC-880 amplifying the broadest size range (178–2,060 bp), while UBC-855 and UBC-809 amplified narrower ranges (169–1,368 bp and 315–389 bp, respectively). Overall, the combined set of primers provided a wide distribution of amplified bands and high levels of polymorphism across the six populations studied, confirming the efficiency of the ISSR markers in generating reproducible and diverse banding patterns.

Genetic diversity indices generated from six populations of *T. emetica* were presented in Table 7. Number of observed alleles (Na) changed across populations, ranging from 0.656 both Kisumu and Siaya to 1.625 in Nandi, with an overall mean of 1.021. Standard errors (SE) for Na ranged from 0.133 in Nandi to 0.178 in Vihiga, indicating differences in allele richness across populations. The number of effective alleles (Ne) showed a narrower range, with the lowest value observed in Siaya (1.166) and the highest in Nandi (1.404). The mean Ne across all populations was 1.237. Corresponding SE values ranged between 0.055 (Vihiga) and 0.067 (Nandi).

The Shannon Information Index (*I*) also varied among populations. The lowest value was observed in Siaya (0.149), followed closely by Kisumu (0.157), while the highest index was detected in Nandi (0.358). The mean Shannon Index across the six populations was 0.220, with SE values ranging from 0.042 in Vihiga to 0.052 in Bungoma. Gene diversity (He) values varied from 0.098 in Siaya to 0.237 in Nandi, with an average of 0.145. The unbiased expected heterozygosity (uHe) followed a similar trend, ranging from 0.100 in Siaya to 0.243 in Nandi, with an overall mean of 0.148. The percentage of polymorphic loci (P %) exhibited marked differences among populations, with the lowest (31.25%) reported in Kisumu and Siaya populations, while the highest of 78.31% detected in Nandi. The overall mean percentage of polymorphic loci across all populations was 48.96 %, with an associated SE of 8.94 (Table 7).

Table 7. Genetic diversity indices of six *T. emetica* populations in western Kenya generated using ISSR.

Populations		N	Na	Ne	*I*	He	uHe	%P
Bungoma	**Mean**	20.000	0.781	1.249	0.207	0.142	0.146	34.38
	SE	0.000	0.166	0.065	0.052	0.036	0.037	
Kakamega	**Mean**	20.000	1.531	1.293	0.289	0.183	0.187	75.00
	SE	0.000	0.149	0.059	0.044	0.032	0.032	
Kisumu	**Mean**	20.000	0.656	1.181	0.157	0.106	0.109	31.25
	SE	0.000	0.166	0.057	0.047	0.032	0.033	
Siaya	**Mean**	20.000	0.656	1.166	0.149	0.098	0.100	31.25
	SE	0.000	0.166	0.056	0.044	0.030	0.031	
Vihiga	**Mean**	20.000	0.875	1.169	0.161	0.102	0.104	43.75
	SE	0.000	0.178	0.055	0.042	0.030	0.030	
Nandi	**Mean**	20.000	1.625	1.404	0.358	0.237	0.243	78.31
	SE	0.000	0.133	0.067	0.048	0.035	0.036	
Mean			1.021	1.237	0.220	0.145	0.148	**48.96**
SE								**8.94**

Note: ***Na* = number of observed alleles; *Ne* = number of effective alleles; *I* = Shannon's Information Index; *He* = Nei's gene diversity; *uHe* = unbiased expected heterozygosity; *%P* = percentage of polymorphic loci. BG= Bungoma, KK= Kakamega, KSM= Kisumu, SY = Siaya, VHG= Vihiga, NA =Nandi.**

4.2.1 Molecular genetic differentiation

The analysis of molecular of variance (AMOVA) portioned genetic diversity of *T. emetica* into components attributable to variation among and within the six populations (Table 8; Figure 6). Variation within and among populations accounted for 65% and 35%, respectively.

Table 8. Analysis of molecular variance for six *T. emetica* populations based on ISSR loci.

Source	df	SS	MS	Est. Var.	%
Among Pops	5	140.717	28.143	1.288	35%
Within Pops	114	272.4	2.389	2.389	65%

The estimated coefficient of differentiation (*Gst*) across the six populations of *T. emetica* was 0.2675 (p = 0.001), indicating that 27% of the total variation was attributable to differences among populations, while the remainder was distributed within populations (Table 9). At the locus level, Gst values varied across the 15 ISSR primers. The highest coefficient of differentiation was observed at locus UBC888 (Gst = 0.457), followed closely by UBC890 (Gst = 0.4403) and UBC813 (Gst = 0.414). Other loci with relatively high Gst values included UBC809 (0.3506), UBC845 (0.3498), and UBC864 (0.3415). By contrast, loci such as UBC811 (0.042), UBC855 (0.052), and UBC880 (0.0811) exhibited comparatively low differentiation among populations.

The corresponding gene flow estimates (Nm=2.39) also displayed wide variation. Loci with high Gst values generally had lower Nm estimates, for example, UBC888 (Nm = 0.5941) and UBC890 (Nm = 0.6357). Conversely, loci with lower Gst values related with higher Nm estimates, such as UBC811 (Nm = 11.4) and UBC855 (Nm = 9.1154). The average across loci confirmed moderate population differentiation with Gst = 0.2675, supported by locus-specific values ranging between 0.042 (UBC811) and 0.457 (UBC888).

Table 9. Coefficient of genetic differentiation across different loci of *T. emetica* using ISSR primers

Locus	Sample Size	Ht	Hs	*Gst*	Nm*
UBC809	120	0.2887	0.1875	0.3506	0.9259
UBC810	120	0.4861	0.3367	0.3074	1.1264
UBC811	120	0.0165	0.0158	0.042	11.4
UBC813	120	0.2432	0.1425	0.414	0.7076
UBC823	120	0.2188	0.1558	0.2876	1.2384
UBC840	120	0.4994	0.3567	0.2859	1.249
UBC845	120	0.4499	0.2925	0.3498	0.9294
UBC847	120	0.2778	0.255	0.082	5.5976
UBC855	120	0.2778	0.2633	0.052	9.1154
UBC857	120	0.1528	0.1283	0.16	2.625
UBC864	120	0.4999	0.3292	0.3415	0.9642
UBC880	120	0.1388	0.1275	0.0811	5.6667
UBC888	120	0.4988	0.2708	0.457	0.5941
UBC890	120	0.4988	0.2792	0.4403	0.6357
UBC891	120	0.4888	0.3125	0.3606	0.8865
				0.2675	

Note: MC= Marker code, *Ht* = total genetic diversity; *Hs* = within-population genetic diversity; *Gst* = coefficient of genetic differentiation; *Nm* = estimate of gene flow.

4.3 Genetic Variations among populations

The principal coordinates (PCoA) based on ISSR marker data revealed patterns of genetic relationships among the six populations of *T. emetica* (Figure 5). The first two coordinates detailed the largest proportion of genetic variation and clearly distinguished certain population from one another. The Nandi population was positioned distinctly apart from the other groups along both coordinate 1 and Coordinate 2, indicating its separation from the remaining populations. The Vihiga population showed partial overlap with Nandi, even though it still maintains some degree of distinction. In contrast, the other populations displayed varying levels of overlap. Siaya and Kisumu clustered closely together, showing overlap within the same coordinate space. Similarly, Kakamega and Bungoma exhibited considerable overlap, forming another cluster. Despite these overlaps, the overall scatter revealed structured variation among populations, with certain groups forming compact cluster while others remained spatially distinct.

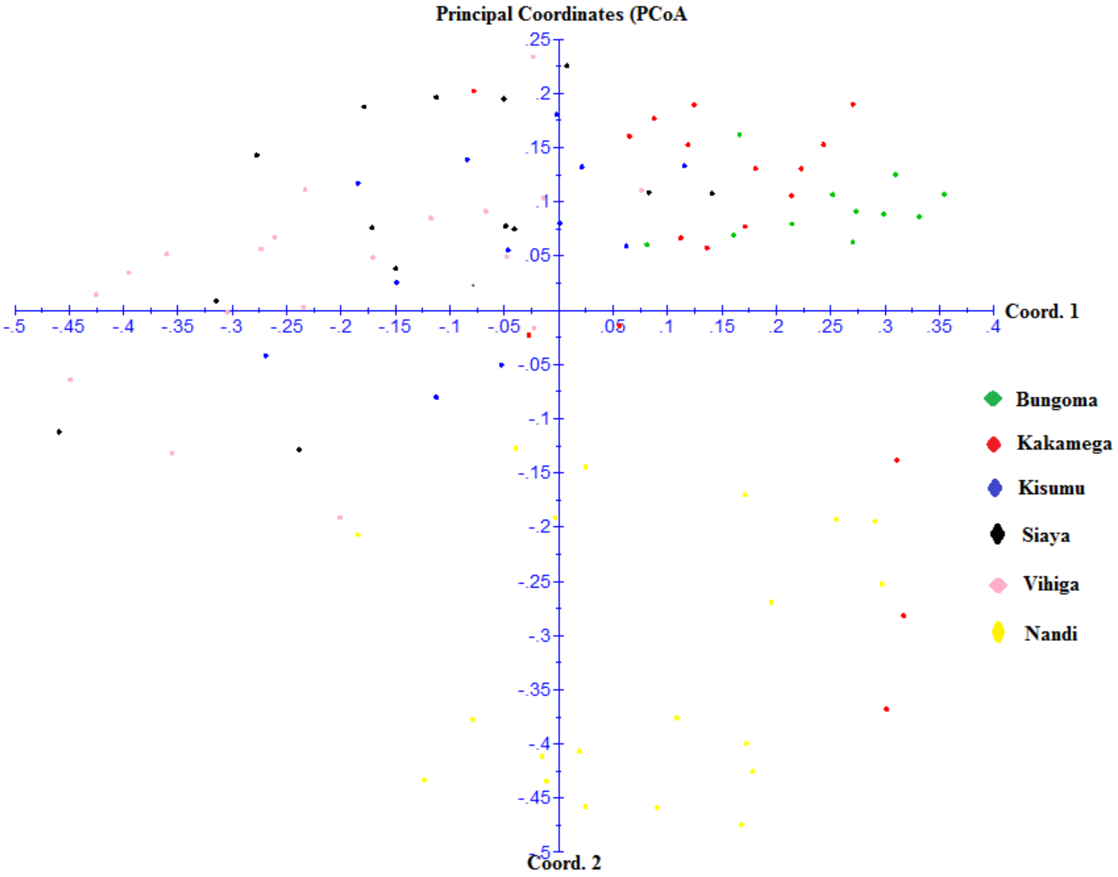


Figure 8. Principal coordinates analysis of 120 individuals in six populations of *T. emetica* in western Kenya based on ISSR markers.

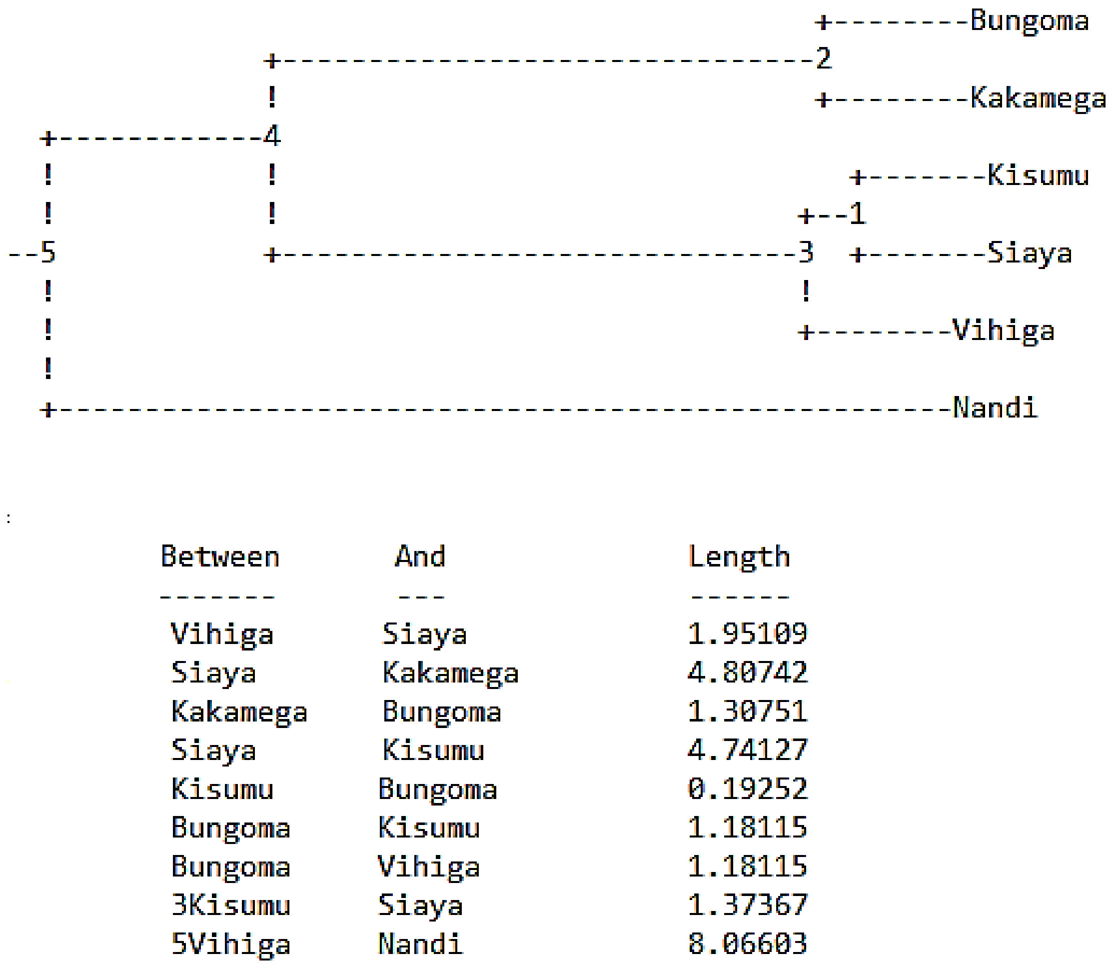
The unbiased genetic distances among the six *T. emetica* populations are presented in Table 10. Pairwise Nei's genetic distances ranged from 0.0236 to 0.2399, while corresponding values of genetic identity ranged from 0.7867 to 0.9767. The shortest genetic distance was found to be between Vihiga and Kisumu (0.0236), indicating their close genetic proximity, followed closely by Siaya and Vihiga (0.0259). Conversely, it was the highest between Bungoma and Nandi (0.2399), demonstrating substantial differentiation between these populations.

Nandi showed high distances from most other populations, including Bungoma (0.2399), Siaya (0.1494), and Vihiga (0.1494). However, it was comparatively closer to Kakamega (0.1354). Bungoma also exhibited notable divergence from Kisumu (0.1286) and Siaya (0.1640), while maintaining greater genetic affinity with Kakamega (0.0262).

Table 10. Nei's genetic identity (above) and genetic distance (below) among six populations of *T. emetica* in western Kenya.

ID	Bungoma	Kakamega	Kisumu	Siaya	Vihiga	Nandi
Bungoma	****	0.9742	0.8793	0.8487	0.833	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	****

The dendrogram derived from Nei's unbiased genetic distances further illustrates this relationship (Figure 8). Three primary clusters were evident. The first cluster grouped Siaya, Vihiga, and Kisumu, reflecting their relatively low pairwise distances. The second cluster grouped Bungoma and Kakamega, reconcilable with their observed high genetic identity (0.9742). Nandi formed the third cluster, standing unique from all other populations, hence, in agreement with observed relatively high genetic distance values.



```
                                               +--------Bungoma
               +-------------------------------2
               !                               +--------Kakamega
  +------------4
  !            !                                 +-------Kisumu
  !            !                              +--1
--5            +------------------------------3  +-------Siaya
  !                                           !
  !                                           +--------Vihiga
  !
  +----------------------------------------------------Nandi

:
         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
```

Figure 9. Dendrogram based on the Nei's unbiased genetic distance showing genetic clustering of 120 individuals in six populations of *T. emetica*.

CHAPTER FIVE

DISCUSSION

5.1 Morphological Variations

The study revealed substantial morphological variability in *T. emetica* across six natural populations in western Kenya, with higher variation observed within populations (67.3%) than among them (32.7%). This finding align with classical population genetic theory, which posits that long-lived, predominantly outcrossing tree species tend to maintain greater variability within populations than between them (Loveless & Hamrick, 1984). High within-population variation has in fact been reported in other Meliaceae species such as *Swietenia macrophylla* and *Melia azedarach* (Gillies et al., 1999; Khan et al., 2022), underscoring that morphological heterogeneity is a consistent feature of tropical tree populations.

Among the traits studied, reproductive parameters, including flower length, seeds per fruit, seeds length, and seeds dry weight, exhibited the greatest variability, while vegetative parameters such as tree height and DBH also varied significantly among populations. High variability in reproductive characters has been archived in other forest trees where reproductive organs are sensitive to both environmental and genetic influences (Athena et al., 2014; Jones et al., 2009). The significant partitioning of traits into three major clusters (Nandi alone; Bungoma and Kakamega together; and Kisumu, Siaya, and Vihiga forming a third cluster) suggests geographic structuring of morphological diversity. This clustering pattern parallels observations in other species where morphological groupings mirror geographic proximity and shared ecological conditions (Andiego et al., 2022; Tsomele et al., 2021).

The distinctiveness of Nandi, which exhibited the highest mean values across most morphological traits, is notable. Its location in the Rift Valley, often described as a biogeographic barrier influencing gene flow (Dangasuk, 2001; Chiveu et al., 2008), likely contributes to this differentiation. Similarly, the clustering of Bungoma and Kakamega may translate from phylogenetic gene flow from these populations before habitat fragmentation in Kakamega forest. The Kisumu–Siaya–Vihiga grouping reflects their geographic proximity and potential for pollinator-mediated or animal-mediated dispersal, consistent with the species' known reproductive ecology (David et al., 2014).

Overall, the morphological results highlight that *T. emetica* maintains considerable intra-population variability, suggesting phenotypic plasticity and potential for adaptation to diverse ecological conditions, while also reflecting geographic influences on morphological divergence.

5.2 Genetic Variation

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 amplifications 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 amplifications 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., 2023a; Xu et al., 2024; Al-Yasi and Al-Qthanin, 2024).

The diversity indices in this study show the existence of genetic heterogeneity among the populations of *T. emetica* in western Kenya. 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 experiences high gene flow, with reduced human activities, which allow preservation of higher genetic variability. These findings are line with the previous ISSR and SSR-studies showing relatively high within-population diversity attributed to large, continuous populations, extensive gene flow, and open pollination systems that facilitate allelic exchange and reduced genetic drift in Meliaceae trees species such as *Melia dubia* (Rawat et al., 2018), *Swietenia macrophylla* (Limongi Andrade et al., 2022) and *Khaya anthotheca* (Bouka et al., 2022). Congruent with the results in this study, populations of *Memecylon subcordatum* (Viswanathan et al. 2018), *Osyris lanceolata* (Mugula et al. 2023), *Pinus sylvestris* (Sheikina and Romanov, 2024) also had both high Shannon's diversity indices and percentage of polymorphic loci, demonstrating the influence of population size and ecological integrity in determining within-population genetic variation.

In contrast, Siaya and Kisumu *T. emetica* populations showed a lower diversity index (Na ≈ 0.66; He ≈ 0.10; %P = 31.25%), indicating reduced allelic variation and possible genetic erosion. The reduced allelic richness and heterozygosity in Siaya and Kisumu correspond with documented pressures of deforestation, agricultural expansion, and unsustainable resource extraction in western Kenya (Kogo et al., 2019; Rotich and Ojwang, 2021; Chebii et al., 2022; Osewe et al., 2022). Similar trend where habitat fragmentation and land-use change contributed to reduction in genetic diversity and gene flow in tropical tress species (Soares et al., 2019). A reduced gene diversity in isolated or heavily exploited populations has also been reported in *Melia dubia* (Rawat et al., 2018), *Khaya anthotheca* (Bouka et al., 202) and *Swietenia macrophylla* (Limongi Andrade et al., 2022). Bungoma, Kakamega, and Vihiga populations had moderate diversity indices (He = 0.14–0.19; I = 0.21–0.29; %P = 43-75%), indicating partial preservation of genetic diversity, reflecting their position along a continuum of anthropogenic distance and ecological connectivity. Therefore, the Bungoma-Kakamega-Vihiga populations may still be exchanging genes through fragmented corridors or remnant trees in agricultural landscapes, a pattern observed in *Tabebuia rosea* (Ruiz-González et al., 2023) and *Acacia senegal* (Omondi et al., 2023), where exchange of genes among isolated or remnant populations has been maintained despite fragmentation of the habitats.

Population differentiation and structure of *T. emetica*

The populations of *T. emetica* in western Kenya showed a structured yet genetically diverse pattern, influenced by limited gene flow. Sixty-five percent of the total genetic variation was observed within populations and 35% among the populations (ΦST = 0.35), showing moderate to high genetic differentiation. The *Gst* value (0.27, p = 0.001) reinforced this pattern, indicating that while diversity of alleles is preserved within the populations, a significant percentage of variation is spread among them. In comparison with other Meliaceae species, *T. emetica* displayed higher population differentiation than *Khaya anthotheca* (FST = 0.12) and *Swietenia macrophylla* (ΦST = 0.28), meaning that gene flow is restricted across its populations in western Kenya (Alcalá et al., 2015; Bouka et al. 2022). Similar reduction in gene flow due to habitat fragmentation has been reported in plant populations (Cheptou et al., 2017).

The differentiation indices in this study (ΦST ≈ 0.35; Gst ≈ 0.27) corresponded well with a mean gene flow estimate of Nm = 2.39, reflecting moderate connectivity, sufficient to reduce genetic drift but insufficient to homogenize populations completely. This balance demonstrates that the populations of *T. emetica* may be undergoing semi-independent evolutionary processes, with geographic isolation and environmental gradients acting as filters to exchange of genes. Rajarajan et al., (2024) reported comparable trends of genetic differentiation and population structuring that is influenced by environmental heterogeneity in *Azadirachta indica* and *Melia azedarach*, reinforcing the view that ecological variation drives genetic structuring within the Meliaceae. Similar levels of gene flow and spatial structuring have been reported in *Acacia senegal* (Omondi et al., 2023), *Khaya anthotheca* (Bouka et al. 2022), and *Swietenia macrophylla* (Limongi Andrade et al., 2022), where fragmentation and ecological discontinuities have produced moderate differentiation despite ongoing genetic connectivity. Variation in *Gst* values (0.04-0.46) across ISSR loci further demonstrated unevenness of genetic differentiation across the genome, differing levels of mutation, selection, or gene exchange among loci. Primers such as UBC-809 and UBC-864, which had intermediate *Gst* values, may have captured genomic regions undergoing partial isolation, while highly differentiated loci like UBC-888 likely reflect regions of genome under selection or low recombination. This trend is in line with findings in *Melia dubia* (Rawat et al., 2018) and *Juniperus excelsa*

(Al-Yasi and Al-Qthanin, 2024), where ISSR markers effectively revealed localized allelic divergence in response to ecological isolation and restricted pollen flow.

The PCoA and the UPGMA analyses supports these molecular patterns. These analyses showed three different genetic clusters: (i) Siaya–Vihiga–Kisumu, (ii) Bungoma–Kakamega, and (iii) Nandi. The similarity between UPGMA and PCoA underline the strength of these relationships and their ecological basis. The first two principal coordinates accounted for the majority of total variance and delineated populations along ecological gradients from lowland to highland zones, reflecting the regional topography and rainfall patterns of western Kenya.

The *T. emetica* population from Nandi occupied the first coordinate axis and had the highest Nei's genetic distances, ranging from 0.13 to 0.24, and had the least genetic distances of between 0.79 and 0.88 relative to other populations. This variation in classification can be linked to its ecological and geographical isolation in the highlands, where reduced pollen and seed dispersal, coupled with differences in climatic conditions, support genetic divergence. Previous studies showing distinct genetic differentiation connected to topographic and ecological discontinuities has been reported in *Swietenia macrophylla* from fragmented Mexican forests (Alcalá et al., 2015), *Osyris lanceolata* in East Africa (Mugula et al., 2023), and *Parashorea chinensis* across Southeast Asia (Li et al., 2024; Xu et al., 2024).

In contrast, the populations of *T. emetica* from Siaya, Kisumu, and Vihiga) displayed very low genetic distances, which ranged from 0.02 to 0.03, and high identity values greater than 0.97, pointing a recent common ancestry or ongoing gene flow preserved through landscape connectivity such as riverine corridors or agricultural mosaics. These findings mirror the weak spatial structuring reported in *Cedrela odorata* across Neotropical lowlands (Finch et al., 2022) and *Sonneratia caseolaris* in coastal Vietnam (Le and Le, 2024), where gene flow remains relatively unrestricted in connected habitats. On other hand, Bungoma and Kakamega clustered in a moderately distinct category, likely sustained by partial connectivity through remnant mid-altitude forest corridors. The small Nei's distance of 0.03 observed between these two populations and their close grouping on the dendrogram support the presence of ongoing but reduced gene flow across fragmented forest patches.

5.3 Comparative analysis of morphological and molecular data sets

Morphologically, *T. emetica* populations were highly variable in both vegetative and reproductive traits, while the average coefficient of variation reached 46.46%. Tree height, DBH, crown diameter, and seed traits showed significant variation ($p = 0.05$), reflecting wide phenotypic plasticity, which may be modified by ecological gradients, resource availability, and genetic constitution. Variance partitioning showed that 67.33% of morphological variation was partitioned within populations, while 32.6% was partitioned among populations, an indication that most trait variation resides within sites in agreement with widespread gene flow and environmental adaptability. Cluster analysis resulted in three major morphological groupings: i) Nandi forming a distinct cluster with high trait means, ii) Bungoma–Kakamega with intermediate values, and iii) Kisumu–

Siaya–Vihiga forming the low-trait cluster. The pattern observed reflects ecological differentiation driven by environmental gradients and anthropogenic pressures, which probably promoted genetic isolation.

In contrast, ISSR molecular data, generated from 15 primers producing 94.65% polymorphism, conferred a similar but more accurate perspective of genetic structure. The high %P of 78.31% and Nei's gene diversity He = 0.24 in the Nandi population demonstrate the maintenance of allelic richness and effective outcrossing, while dioecy, animal pollination, and high habitat integrity have likely contributed to this trend. On the other hand, the lower diversity measures in Siaya and Kisumu (%P = 31.25%; He = 0.10) indicate genetic erosion that may be linked to habitat fragmentation, deforestation, and human disturbance (Kogo et al., 2019; Rotich & Ojwang, 2021; Chebii et al., 2022). The fixation index ΦST = 0.35 revealed that the populations were substantially but not entirely differentiated genetically, confirming moderate gene flow (Nm = 2.39) across the populations. The molecular pattern of clustering was similar to the morphological pattern, grouping into three major genetic clusters: Siaya–Kisumu–Vihiga, Bungoma–Kakamega, and Nandi—suggesting consistent population structuring across the data type.

Cumulatively, these findings indicate that while morphological data yield valuable estimators of adaptive variation, molecular markers such as ISSR provide greater sensitivity in the quantification of genetic diversity and population structure. The concordance of morphological and molecular patterns of clustering confirms the reliability of population differentiation in *T. emetica*. A close correlation between morphological divergence and molecular distance matrices could therefore be anticipated to reflect deeper genetic isolation and narrow connectivity among populations in disturbed habitats. Thus, the combination of morphological and molecular assessments allows for the highlighting of a full suite of dynamics in diversity and is crucial for formulating efficient conservation strategies. Indeed, such approaches have been recommended for other tropical tree species, such as *Swietenia macrophylla* (Limongi Andrade et al., 2022), *Khaya anthotheca* (Bouka et al., 2022), and *Melia dubia* (Rawat et al., 2018), where agreement among marker types has served as a useful tool for guiding restoration and breeding programs.

In the present study, morphological and molecular analyses show complementary insights into the genetic diversity of *T. emetica*. Morphological traits underline the adaptive flexibility and ecological differentiation of the species, while the ISSR markers reveal their genetic underpinning. Integration of the two emphasizes the potential of reproductive biology, ecological integrity, and gene flow for maintaining genetic heterogeneity as the basis for resilience and sustainability in the *T. emetica* populations of western Kenya.

5.4 Implications in conservation and sustainable use

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 diversity is vital for maintaining adaptive potential and ensuring long-term evolutionary resilience, particularly under increasing environmental and anthropogenic pressures. Given projections of shifting climatic zones and increased habitat stress, maintaining such genetic diversity is crucial for forest resilience and adaptive response (Konrad et al., 2025). However, the relatively high population differentiation ($\Phi ST = 0.35$; $Gst = 0.27$) shows that genetic resources are unevenly distributed, highlighting the need for population-specific conservation strategies rather than a uniform management approach. Similar to patterns reported in fragmented alpine and tropical plant systems, where restricted gene flow and local adaptation demand site-specific conservation actions (Cheptou et al., 2017; Soares et al., 2019), these findings emphasize that the management of *T. emetica* should account for regional ecological and genetic distinctiveness.

Populations such as Nandi and Kakamega, which exhibited the highest gene diversity ($He = 0.24$ and 0.18, respectively) and the greatest percentage of polymorphic loci, represent genetic reservoirs essential for the species' survival. These populations should be prioritized as core units for in situ and ex situ conservation, serving as key sources for seed collection and the establishment of genetic resource banks. Similar conservation initiatives integrating molecular diversity and population structure have been conducted for *Vitellaria paradoxa* in Côte d'Ivoire (Attikora et al., 2024), highlighting the importance of conserving genetically diverse populations as reservoirs for breeding and restoration. In contrast, populations from Siaya and Kisumu, which recorded low gene diversity ($He = 0.10$–0.11) and only 31.25% polymorphic loci, are vulnerable to genetic erosion and stochastic loss. These populations require genetic enrichment through assisted regeneration, enrichment planting, and controlled exchange of germplasm from genetically richer populations to prevent further decline in adaptive capacity (Donkpegan et al., 2020). The moderate fixation index ($\Phi ST = 0.35$) and variable *Gst* values among ISSR loci (0.04 -0.46) reflect limited genetic exchange between populations. From a conservation perspective, this underscores the importance of preserving and restoring gene flow corridors between fragmented forest remnants. Conservation planning should prioritize connectivity conservation, linking genetically related populations such as Bungoma-Kakamega and Kisumu -Vihiga -Siaya to sustain pollen and seed dispersal. Establishing community-based forest restoration zones that utilize genetically diverse and locally adapted materials can help stabilize population structures and mitigate ongoing isolation. Furthermore, the clear delineation of populations into three genetic groups: (i) Siaya-Vihiga-Kisumu, (ii) Bungoma-Kakamega, and (iii) Nandi, provides a strong foundation for conservation and seed zoning. Such genetic zoning supports Kenya’s forest seed production and certification systems, ensuring that restoration initiatives use materials compatible with the adaptive traits of each ecological region (FAO, 2025). Integrating this molecular data into forest restoration policy will enhance seed sourcing, prevent maladaptation, and strengthen reforestation programs 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 integration is needed, given evidence that human activities are eroding genetic diversity and threatening species persistence (Soares et al., 2019; Omondi et al., 2023).

5.5 Limitations of the study

The present study provides the first population-level morphological and molecular baseline data about the genetic diversity and population structure of *T. emetica* in Kenya. In this study, morphological characterization focused essentially on quantitative traits, such as height, diameter at breast height, crown dimensions, and seed traits, reflecting environmental conditions and phenotypic plasticity rather than purely genetic differentiation. While these traits are valuable in terms of ecological adaptability, they do not provide a good representation of variation that is heritable. Therefore, employment of qualitative morphological traits, such as leaf shape, floral features, or fruit morphology, in future studies is recommended, as this will help to better capture stable, genotype-dependent variation and increase congruence with molecular data.

The analysis, using fifteen ISSR markers, showed a high level of polymorphism-a minimum of over 94%-and delineated *T. emetica* populations into distinct clusters corresponding with ecological stratification. Although the sample size of 20 individuals per population was quite sufficient to detect the general pattern of genetic diversity, finer intra- and inter-population variation may not have been captured. The dominant nature of ISSR markers has limited the distinction between homozygous and heterozygous loci, hence constraining the exact estimation of heterozygosity, inbreeding coefficients, and allelic richness.

Because ISSRs are dominant and target non-coding DNA regions, they have limited information on adaptive loci. However, recent chloroplast genomic studies in the Meliaceae family have demonstrated significant interspecific and adaptive variation that is useful in phylogenetic and evolutionary inference (Nie et al., 2025). Integration of chloroplast genomic data with co-dominant nuclear markers such as SSRs and SNPs would increase the precision of heterozygosity, inbreeding, and estimates of adaptive diversity (Bouka et al., 2022; Limongi Andrade et al., 2022; Faria et al., 2024). In general, while the ISSR markers provided useful baseline molecular data for the assessment of genetic variation and population differentiation, the integration of plastid and co-dominant markers, as well as the inclusion of qualitative morphological traits, would yield a more comprehensive understanding of the genetic structure, adaptive evolution, and conservation priorities of *T. emetica*.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusion

1. *T. emetica* populations in western Kenya exhibit significant phenotypic variation across the eight quantitative morphological traits assessed, with higher variation occurring within populations than among them. The analysis revealed three distinct clusters: Nandi as a phenotypically unique group, Bungoma and Kakamega as a closely related cluster, and Kisumu, Siaya, and Vihiga forming another cluster, reflecting both ecological and geographical structuring.
2. Molecular analysis using ISSR markers demonstrated high genetic diversity in populations, with most variation residing within populations. Gene flow was evident among geographically proximate populations, such as Siaya - Kisumu and Bungoma- Kakamega, while Nandi remained genetically distinct. These findings confirm a structured but interconnected population system with localized genetic differentiation.

6.2 Recommendations

1. The Nandi population, which harbors the highest genetic diversity and a distinct genetic structure, should be prioritized for *in situ* conservation. Targeted interventions, such as preventing habitat degradation will help preserve this critical genetic reservoir.
2. Since substantial genetic diversity is observed within populations, conservation strategies must extend to all six populations studied. Both in situ protection of natural habitats and ex situ measures, such as seed banks and controlled

propagation, is necessary to safeguard the full spectrum of species genetic diversity while ensuring the species' adaptability under changing environmental conditions

6.3 Areas to consider in future Studies

1. At the policy level, these findings should inform county governments (e.g., Kakamega, Bungoma, Nandi, Kisumu, Siaya, and Vihiga) and national agencies such as the Kenya Forest Service (KFS) in designing localized conservation action plans. Integrating *T. emetica* into existing frameworks like the Forest Conservation and Management Act (2016) and national tree-planting campaign (15 billion trees by 2032) will ensure sustainable management; align conservation efforts, and anchor climate change and biodiversity targets of Kenya as a country.
2. Community-based conservation programs should be strengthened by incorporating *T. emetica* into agroforestry and restoration initiatives, particularly in areas experiencing population decline. Promoting its multiple economic uses (timber, oil, medicine) through sustainable harvest guidelines will enhance local livelihood and reduce unsustainable exploitation pressures.

REFERENCES

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubrick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R. and Venter, J.C. 1991. “Complementary DNA sequencing: expressed sequence tags and human genome project.” Science 252: 1651–1656.

Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. 1992. “Length polymorphisms of simple sequence repeat DNA in soybean.” Genetics 132: 1131-1139.

Akweni, A. L., Zharare, G. E., & Zimudzi, C. (2021). Predicting the number of fruits and the seed biomass of *Trichilia emetica* (Vahl.) in the eastern coastal region of South Africa. *AGRIS Database*.

Albert *et al* – (2003). Importance of identifying the sub species of *Acacia nilotica* by genetic distance and the implication for biological control.

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

Aldholmi, A., Shuaibu, S., Elsaman, T., Ahmed, F., Ahmed, A., & Idris, A. (2024). Uses of *Trichilia emetica* oils: Nutritional, cosmetic, and pharmaceutical perspectives. *Journal of Ethnopharmacology*, 325, 116987.

Allaby (1989). Phylogenetic position and revised classification of acacia in south Africa.

Almeida, M. T., Silva, M. J., & Rossi, A. A. (2020). Genetic diversity in *Trichilia* species using ISSR markers: Implications for conservation. *Tree Genetics & Genomes, 16*(2), 34.

Andiego, K. P., Dangasuk, O. G., Otieno, F. D., Omondi, S. F., & Odee, W. D. (2022). Phenotypic Variation of Four Populations of Osyris lanceolata Hochst. & Steud. (African Sandalwood) in Kenya. *Africa Environmental Review Journal*, *5*(2), 168-180.

Omondi, S. O., Were, J., & Oballa, P. (2022). Genetic variation and population structure of East African Sandalwood (*Osyris lanceolata* Hochst. & Steud.). *Tree Genetics & Genomes, 18*(1), 1–14

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics, 17*(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Anil, K. K. (2012). *Observations on leaf scales and insect associations in Trichilia emetica*. National Institute of Oceanography, Dona Paulo, Goa 403 004, India.

Asudi, G. O., Mibus, H., Serek, M., & Winkelmann, T. (2010). Genetic diversity of Kenyan papaya (*Carica papaya L.*) accessions assessed with microsatellite markers. *Plant Genetic Resources: Characterization and Utilization.*

Asudi, G. O., Ombwara, F. K., Rimberia, F. K., Nyende, A. B., Ateka, E. M., Wamocho, L. S., ... & Onyango, A. (2010). Morphological diversity of Kenyan papaya germplasm. *African Journal of Biotechnology*, *9*(51), 8754-8762.

Athena, A., Gupta, R., & Singh, R. (2014). Morphological diversity and genetic variation in tree species: Implications for conservation. *Journal of Forestry Research, 25*(3), 453–462.

Awodele, O. (2015). Genetic diversity and population structure of three Meliaceae species based on RAPD markers. *African Journal of Biotechnology, 14*(21), 1783–1791.

Bardekei, F. (2001). Random Amplified Polymorphic DNA Markers. Turk J. Biol 25: 185-196.

Bassam, B.J., Caetano-Anolles, G. and Gresshoff, P.M. 1992. "DNA amplification fingerprinting of bacteria." Appl. Microbiol. Biotechnol 38:70-76.

Beckmann, J.S. and Weber, J.L. 1992. "Survey of human and rat microsatellites." Genomics 12: 627–631.

Benjamin, O., & Michael, M. (2018). *Propagation of Trichilia emetica through cuttings and coppice management*. Nature Genetics, 50(8), 1112-1121. https://doi.org/10.1038/s41588-018-0147-3

Bettinger, T., Dick, C. W., & Lemes, M. R. (2019). Seed sourcing and genetic consequences for tropical tree reforestation. *Forest Ecology and Management.*

BGCI. (2018). *Trichilia emetica*. The IUCN Red List of Threatened Species 2018.

Blake, T.K., Kadyrzhanova, D., Shepherd, K.W., Islam, A.K.M.R., Langridge, P.L., McDonald, C.L., Erpelding, J., Larson, S., Blake, N.K. and Talbert, L.E. 1996. "STS- PCR markers appropriate for wheat-barley introgression." Theor. Appl. Genet 82:715-721.

Blears, M.J., De Grandis, S.A., Lee, H. and Trevors, J.T. 1998. "Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications." J. Industrial Microbiol. Biotechnol 21: 99–114.

Bornet, B. and Branchard, M. 2001. “Nonanchored inter simple sequence repeat (ISSR) markers: Reproducible and specific tools for genome fingerprinting.” Plant Molecular Biology Reporter 19(3): 209-215.

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. 1980. “Construction of a genetic linkage map in man using restriction fragment length polymorphisms.” Am. J. Hum. Genet 32: 314-331.

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

Brumfield, R.T., Beerli, P., Nickerson, D.A., and Edwards, S.V. 2003. “The utility of single nucleotide polymorphisms in inferences of population history.” Trends Ecol. Evol 18: 249–256.

CABI. (n.d.). *Trichilia emetica* (Cape mahogany) – datasheet. *Invasive Species Compendium*.

Caetano-Anolles, G. 1994. “MAAP- A versatile and universal tool for genome analysis.” Plant Mol. Biol 25: 1011-1026.

Caetano-Anolles, G., Bassam, B.J. and Gresshoff, P.M. 1991. “DNA amplification fingerprinting using very short arbitrary oligonucleotide primers.” Biotechnol 9: 553-557.

Caetano-Anolles, G., Bassam, B.J. and Gresshoff, P.M. 1992. “Primer-template interactions during DNA amplification fingerprinting with single arbitrary oligonucleotides.” Mol. General Genet 235: 157-165.

Caetano-Anolles, G., Bassam, B.J. and Gresshoff, P.M. 1993. “Enhanced detection of polymorphic DNA by multiple arbitrary amplicons profiling of endonuclease digested DNA: identification of markers linked to the supernodulation locus in soybean.” Mol. General Genet 241: 57-64.

Chapman, J. D., & Hall, J. B. (1991). *The forest flora of Malawi*. Kew Publishing.

Chen, C.X., Zhou, P., Choi, Y.A., Huang, S. and Gmitter, F.G. 2006. “Mining and characterizing microsatellites from citrus ESTs.” Theor. Appl. Genet 112: 1248-1257.

Chirchir, R., Omondi, S., & Nyongesa, B. (2024). Market surveys of medicinal plants in Western Kenya: The case of *Trichilia emetica*. *Journal of Ethnobotany and Conservation*, 12(3), 45–58

Chivandi, E., Erlwanger, K. H., Davidson, B. C., & Madziva, M. T. (2008). Fats from non-traditional plant sources as cosmetic and pharmaceutical materials: The case of mafura butter (*Trichilia emetica* seed oil). *International Journal of Cosmetic Science*, 30(5), 339–345. https://doi.org/10.1111/j.1468-2494.2008.00462.x

Chiveu C.J, Dangasuk O G. Omunyin ME, Wachira FN (2008) Genetic diversity in Kenyan populations of *Acacia senegal* (L) wild revealed by combined RAPD and ISSR markers. Afr J Biotechnol 7 (14):2333-2340

Coenye, T., Vandamme, P., Govan, J.R.W. and LiPuma, J.L. 2001. “Taxonomy and Identification of the Burkholderia cepacia Complex.” J. Clin. Microbiol 39: 3427–3436.

Coetes, D.J. and Byrne, M. 2005. “Genetic variation in plant populations.” In Plant diversity and Evolutions-Genotypic and phenotypic variation in higher plants, edited by R.J. Henry, ch. 9: 139-164. Wallingford, Oxfordshire, UK: CABI Publishing.

Coles, N.D., Coleman, C.E., Christensen, S.A., Jellen, E.N., Stevens, M.R., Bonifacio, R.A., Rojas-Beltran, J.A., Fairbanks, D.J. and Maughan, P.J. 2005. “Development and use of an expressed sequenced tag library in quinoa (Chenopodium quinoa Willd.) for the discovery of single nucleotide polymorphisms.” Plant Soc 168: 439–447.

Collard, B.C.Y., Jahufer, M.Z.Z. and Pang, E.C.K. 2005. “An introduction to markers, quantitative trait loci (QTL) mapping and marker-assissted selection for crop improvement: The basic concepts.” Euphytica 142(1-2): 169-196.

Cooke, R.J. 1984. “The characterization and identification of crop cultivars by electrophoresis.” Elect 5: 59-72.

Corley Smith, G.E., Lim, C.J., Kalmar, G.B. and Brandhorst, B.P. 1997. “Efficient detection of DNA polymorphisms by fluorescent RAPD analysis.” BioTechniques 22: 690-692.

Cunha, T. A., Finger, C. A. G., & Schumacher, M. V. (2013). Growth and architecture of tropical tree species under different environmental conditions. *Trees, 27*(3), 877–889.

Dalmastri, C., Chiarini, L., Bevivino, A., Manno, G., Mentasti, M. and Fani, R. 2010. "Application of multiplex single nucleotide primer extension (mSNuPE) to the identification of bacteria: The Burkholderia cepacia complex case." J. Microbiol. Methods 80: 251–256.

Dalmastri, C., Chiarini, L., Bevivino, A., Manno, G., Mentasti, Winter, P. and Kahl, G. 1995. "Molecular marker technologies for plant improvement." World J. Microbiol. Biotechnol 11: 438–48.

Dangasuk OG. Panetsos kp (2004) Altitudinal and Longitudinal Variations in *Pinus brutia* (Ten) of Cret island, Greece: some needle, cone traits under natural habitats. New for 27(3): 269-284.dio:10.1023/B: NEFO.OOOO2-2227.33131.fo.

Dangasuk, O.G. 1999. Genetic diversty in *Faidherbia albida* (Del.) A. Chev. Ph.D. Thesis, Moi Universty, Eldoret, Kenya. 155 pp.

Dangasuk, O.G. 2001. Molecular Genetics of *Faidherbia albida*. European Tropical Forestry Research Network News, Biotechnology Special issue, 34: 15-17.

Dangasuk, O.G., Seurei P. and Gudu, S. 1997. Genetic Variation in seed and seedlings traits in 12 African provenances of *Faidherbia albida* (Del.) A. Chev. At Lodwar, Kenya. Agroforestry systems 37: 133-141.

David, A., Malombe, I., & Onyango, B. (2014). Pollination biology and seed dispersal of *Trichilia emetica* in coastal Kenya. *Journal of East African Natural History, 103*(1), 23–34.

David, E. T., Chhin, S., & Skole, D. (2014). Dendrochronological potential and productivity of tropical tree species in western Kenya. *Tree-ring research*, *70*(2), 119-135.

de Vicente, M.C. and Fulton T. 2003. “Using molecular marker technology in studies on plant genetic diversity.” [Online] IPGRI, Rome, Italy and Institute for Genetic Diversity, Ithaca, New York, USA.

Denslow, J. S. (1995). Disturbance and diversity in tropical rain forests: the density effect. *Ecological applications*, *5*(4), 962-968.

Du, H., et al. (2022). Chromosome-level genome assembly of *Azadirachta indica* provides insights into terpene biosynthesis. *Horticulture Research, 9*, uhac176. https://doi.org/10.1093/hr/uhac176

Dutkowski GW. Potts BM (1999) Geographic patterns of genetic variations in Eucalyptus globulus and a revised racial classification. Aust J Bot 47-237.doi: 10.1023/A: 1005865502289.

Eijk, M.J.T., Broekhof, J.L.N., van-der-Poel, H.J.A., Hogers, R.C.J., Schneiders, H., Verstege, K.J.E., van-Aart, J.W., Geerlings, H., Buntjer, J.B., van-Oeveren, A.J. and Vos, P. 2004. “SNPWave: A flexible multiplexed SNP genotyping technology.” Nucl. Acids Res 32: e47.

Eisen, J.A. 1999. “Mechanistic basis for microsatellite instability.” In Microsatellites: Evolution and applications, edited by Goldstein, D.B. & Schlötterer, C., 34-48. Oxford, UK: Oxford University Press.

Emiliani, F., & Giovannelli, G. (2023). Genetic diversity and conservation of tropical trees: Insights from molecular markers. *Plant Ecology, 224*(2), 145–160.

Emiliani, G.; Giovannelli, A. Tree Genetics: Molecular and Functional Characterization of Genes. Forests 2023, 14, 534. https://doi.org/10.3390/ f14030534

Erasmus, A., Myburgh, M., & Viljoen, A. (2020). Mafura oil (*Trichilia emetica*): A review on chemistry, bioactivity, and potential applications. *Industrial Crops and Products*, 152, 112504. https://doi.org/10.1016/j.indcrop.2020.112504

Eujayl, I., Sorrells, M., Baum, M., Wolters, P. and Powell, W. 2001. “Assessment of genotypic variation among cultivated durum wheat based on ESTSSRs and genomic SSRs.” Euphytica 119: 39–43.

Fekadu and Yohames (2004). Economic value of *Trichilia emetica,* Delanix elata, Berchemia discolor and warbugia Ugandensis in Ethiopia.

Ferri, L., Perrin, E., Campana, S., Tabacchioni, S., Taccetti, G., Cocchi, P., Ravenni, N., Dalmastri, C., Chiarini, L., Bevivino, A., Manno, G., Mentasti, M. and Fani, R. 2010. “Application of multiplex single nucleotide primer extension (mSNuPE) to the identification of bacteria: The *Burkholderia cepacia* complex case.” J. Microbiol. Methods 80: 251–256.

Figdore, S.S., Kennard, W.C., Song, K.M., Slocum, M.K. and Osborn, T.C. 1988. “Assessment of the degree of restriction fragment length polymorphism in Brassica.” Theor. Appl. Genet 75: 833–40.

Finch KN, Jones FA, Cronn RC (2022) Cryptic species diversity in a widespread Neotropical tree genus: The case of *Cedrela odorata*. *American Journal of Botany, 109*(11), 1639–1653. https://doi.org/10.1002/ajb2.16064

Foster, J.T., Allan, G.J., Chan, A.P., Rabinowicz, P.D., Ravel, J., Jackson, P.J. and Keim, P. 2010. “Single nucleotide polymorphisms for assessing genetic diversity in castor bean (Ricinus communis).” BMC Plant Biol 10: 13–23.

Gao, L.F., Jing, L.R., Huo, N.X., Li, Y., Li, X.P., Zhou, R.H., Chang, X.P., Tang, J.F., Ma, Z.Y. and Jia, J.Z. 2004. “One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat.” Theor. Appl. Genet 108: 1392–1400.

Gavamukulya, Y., & Kiwuka, C. (2020). Antimalarial plants used across Kenyan communities: A review. *Journal of Medicinal Plants Research.*

Gebashe, F. C. (2015). *Studies on the cryopreservation of shoot apices from recalcitrant-seeded Trichilia emetica Vahl. and Trichilia dregeana Sond.* (Master’s thesis, Durban University of Technology).

Germishuizen and Meyer (2003). Plants of south Africa.

Ghazoul, J. and Sheil, D. (2010) Tropical rainforest ecology, diversity and conservation.

Gillies, Navarro, Lowe, Newton, Hernández, Wilson, & Cornelius. (1999). Genetic diversity in Mesoamerican populations of mahogany (Swietenia macrophylla), assessed using RAPDs. *Heredity*, *83*(6), 722-732.

Godwin, I.D., Aitken, E.A.B. and Smith, L.W. 1997. “Application of inter simple sequence repeat (ISSR) markers to plant genetics.” Electrophoresis 18(9): 1524-1528.

Gudu S and Kojwang Ho, (1993). Biotechnology approaches to the movement of forestry and agroforestry in Kenya; Production constrains and practical solutions. In: proc of the national agricultural biotechnology workshop (eds JM Wekunda and Visser), p. 68-77.-Hamrick, J.L. 1990. Isozyme and the analysis of genetic structure in plant population. In: Isozyme in plant genetic in plant biology. Soltis, E.D and Soltis P.s (Eds), pp.87-105 Chapman and Hall, London, Uk.

Gugu Felicity Tsomele, Eudri Venter, Obiro Cuthbert Wokadala, Elize Jooste, Bhekisisa Chushuta Dlamini, Nomali Ziphorah Ngobese & Muthulisi Siwela (2021) Structural (gross and micro), physical and nutritional properties of *Trichilia emetica* and *Trichilia dregeana* seeds, CyTA - Journal of Food, 19:1, 483-492, DOI: 10.1080/19476337.2021.1915877.

Gupta, P.K. and Varshney, R.K. 2004. "Cereal genomics: an overview." In Cereal genomics, edited by Gupta P.K. and Varshney, R.K., 639. Dordrecht, The Netherlands: Kluwer Academic Press.

Gutiérrez-Vázquez, J., López-Aguillón, R., & Treviño-Garza, E. J. (2012). Morphological variation as an indicator of adaptive responses in forest tree species. *Revista Chapingo Serie Ciencias Forestales y del Ambiente, 18*(3), 389–400.

Handa, R. M. (2015). *Flowering and fruiting phenology of Trichilia emetica in East Africa*.

Hartl, D.L. 1988. "Genetic variation." In A primer of population genetics. 2nd ed., Ch. 1:1-67. Massachusetts: Sinauer Associates.

Hartl, D.L. and Clark, A.G. 1997. Principles of population genetics. 3rd ed. Sunderland, Massachsetts: Sinauer Assoc., 163.

Hartl, D.L. and Jones, E.W. 2005. "DNA Structure and DNA manipulation." In Genetics: analysis of genes and genomes. 5th ed., Ch. 2: 36-85. Sudbury: Jones and Bartlett Pub.

Hashim, A. M., Othman, R. Y., & Harun, A. R. (2015). Genetic variation in natural populations of *Neolamarckia cadamba* revealed by ISSR markers. *Journal of Tropical Forest Science, 27*(3), 391–399

Hernández-Díaz, J. C., Corral-Rivas, J. J., Quiñonez-Barraza, G., Vargas-Larreta, B., & López-Sánchez, C. A. (2008). Management strategies and their effects on forest tree morphology. *Forest Ecology and Management, 255*(3–4), 1030–1040.

Hollingsworth, W.O., Christie, C.B., Nichols, M.A. and Neilson, H.F. 1998. “Detect ion of variation among and within asparagus hybrids using random amplified DNA (RAPD) markers.” New Zealand J. Crop Hortic. Sci 26: 1-9.

Hu, S., Li, J., Wang, Z., et al. (2022). Meliaceae genomes provide insights into wood development and limonoid biosynthesis. *Plant Biotechnology Journal, 20*10), 2119–2135. https://doi.org/10.1111/pbi.13941

Huff, D.R., Peakall, R. and Smouse, P.E. 1993. “RAPD variation within and among natural populations of outcrossing buffalograss [(Buchloe dactyloides (Nutt.) Engelm].” Theor. Appl. Genet 86: 927-934.

Ibrahim, A.M. 1996. Conservation and utilization of *Faidherbia albida* (Del.) A. Chev. (Syn acacia albida) gene resources. In: Tree improvement for sustainable tropical forestry.

INTERNATIONALE CENTRE for in AGROFORESTRY (1992): A selection of useful trees and shrubs of Kenya. Kenya ICRAFT.

Jaccoud, D., Peng, K., Feinstein, D. and Kilian, A. 2001. “Diversity arrays: A solid state technology for sequence information independent genotyping.” Nucl. Acids Res 29: e25.

Jain, A., Bhatia, S., Banga, S. and Prakash, S. 1994. “Potential use of random amplified polymorphic DNA (RAPD) technique to study the genetic diversity in Indian mustard (Brassica juncea) and its relationship to heterosis.” Theor. Appl. Genet 88: 116-122.

Jain, S.M., Brar, D.S. and Ahloowalia, B.S. 2002. Molecular techniques in crop improvement. The Netherlands: Kluwer Academic Publishers.

Jarne, P. and Lagoda, P.J.L. 1996. “Microsatellites, from molecules to populations and back.” Trends in Ecol. Evol.

Ji, X., et al. (2021). Chromosome-level genome assembly of *Toona sinensis* and insights into the evolution of Meliaceae. *Horticulture Research, 8*(1),.101. T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics,*

Jones, C. G., Plummer, J.A., Barbour, E.L. & Byrne, M. (2009). Genetic Diversity of an Australian *Santalum album* Collection Implications for Tree Improvement Potential.Silvae Genetica 58: 5–6.

Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F. and van de Wiel, C. 1997. “Reproducibility testing of RAPD, AFLP, and SSR markers in plants by a network of European laboratories.” Mol. Breed 3: 381- 390.

Jones, F. A., Chen, J., Weng, G. J., & Hubbell, S. P. (2009). A genetic evaluation of seed dispersal in a neotropical tree, *Jacaranda copaia* (Bignoniaceae). *American Naturalist, 173*(5), 576–585.

Jonnala, R. S., Dunford, N. T., & Chenault, K. (2006). Tocopherol, phytosterol and phospholipid compositions of genetically modified peanut varieties. *Journal of the Science of Food and Agriculture*, *86*(3), 473-476.

Jordan, S.A. and Humphries, P. 1994. "Single nucleotide polymorphism in exon 2 of the BCP gene on 7q31-q35." Hum. Mol. Genet 3: 1915.

Kamala, K., & Rao, M. G. (2009). Seed production and variability in fruit yield of *Trichilia emetica*. *Forest Ecology and Management*, 257(4), 1036–1042.

Kantety, R.V., Rota, M. L., Matthews, D.E. and Sorrells, M.E. 2002. "Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat." Plant Mol. Biol 48: 501-510.Khanduri, V. P., Sukumaran, A., & Sharma, C. M. (2021). Gender plasticity uncovers multiple sexual morphs in natural populations of *Cedrus deodara* (Roxb.) G. Don. *Ecological Processes, 10*(35). https://doi.org/10.1186/s13717-021-00311-7

Kim, D. O., Chun, O. K., Kim, Y. J., Moon, H. Y., & Lee, C. Y. (2003). Quantification of polyphenolics and their antioxidant capacity in fresh plums. *Journal of agricultural and food chemistry*, *51*(22), 6509-6515.

Kioko, J. I., Omino, E. A., & Maundu, P. M. (2005). Seed biology and germination ecology of selected medicinal plants of Kenya. *Kenya Forestry Research Institute Report*.

Knapik, E.W., Goodman, A., Ekker, M., Chevrette, M., Delgado, J., Neuhauss, S., Shimoda, N., Driever, W., Fishman, M.C. and Jacob, H.J. 1998. "A microsatellite genetic linkage map for zebrafish (Danio rerio)." Nature Genet 18: 338-343.

Koech, G., Ofori, D., Muigai, A. W., & Makobe, M. (2014). Genetic variability and divergence of seed traits and seed germination of five provenances of *Faidherbia albida* (Delile) A. Chev. *African Journal of Plant Science*, *8*(11), 482-491.

Koech, G., Ofori, D., Muigai, A. W., & Makobe, M. (2014). Genetic variability and divergence of seed traits and seed germination of five provenances of Faidherbia albida (Delile) A. Chev. *African Journal of Plant Science*, *8*(11), 482-491.

Kojima, T., Nagaoka, T., Noda, N. and Ogihara, Y. 1998. “Genetic linkage map of ISSR and RAPD markers in Einkorn wheat in relation to that of RFLP markers.” Theor. Appl. Genet 96: 37–45.

Konieczny, A. and Ausubel, F.M. 1993. “A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers.” Plant J 4: 403–410.

Kurata, N., Umehara, Y., Tanoue, H. and Sasaki, T. 1997. “Physical mapping of the rice genome with YAC clones.” Plant Mol. Biol 35: 101–113.

Landry, B.S. and Michelmore, R.W. 1987. An Agricultural Perspective. Edited by G. Bruening, J. Harada and A. Hollaender. New York: Plenum, 134–57.

Lauterbach, D., Ristow, M., Gemeinholzer, B., & Ritz, C. M. (2019). Genetic population structure and conservation of relict species: Case study of *Sorbus domestica* in Central Europe. *Biodiversity and Conservation, 28*(1), 207–228.

Lee Jones, H. (2016). Genetic diversity in tropical plantation species: A comparison of natural and managed populations of *Neolamarckia cadamba* and *Gmelina arborea*. *Forestry Research International, 2*(1), 22–31.

Lemes, M. R., Gribel, R., Proctor, J., & Grattapaglia, D. (2003). Population genetic structure of mahogany (*Swietenia macrophylla*) across the Brazilian Amazon. *Molecular Ecology, 12*(11), 2875–2883. https://doi.org/10.1046/j.1365-294X.2003.01958.x

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. *Neotropical Biodiversity, 8*(1), 222–228. https://doi.org/10.1080/23766808.2022.2080334

Ling, P. (2014). Growth rates and management of *Trichilia emetica* in Zimbabwean agroforestry systems. *Agroforestry Systems*, 88(3), 501–509. https://doi.org/10.1007/s10457-014-9672-8

Litt, M. and Luty, J.A. 1989. “A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene.” American J. Human Genet 44: 397-401.

Liu, Z.J. and Cordes, J.F. 2004. “DNA marker technologies and their applications in aquaculture genetics.” Aquaculture 238: 1-37.

Loveless, M. D., & Hamrick, J. L. (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics, 15*(1), 65–95.

Lynch, M. and Walsh, B. 1998. Genetics and analysis of quantitative traits. Sunderland, MA: Sinauer Associates.

MACDONALD, S.E; B.R THOMAS, D.M. CHENIAWSKY and B.G PURDY (2001): Managing genetics resources of iodgepole pine in west-central Alberta: Patterns of Isozyme Variations in natural populations and effects of forests management. Forest Ecology management 152: 45-58.

Mager, P., Hamilton, M. B., & Peakall, R. (2014). Patterns of genetic diversity in tropical trees and implications for conservation. *Forest Ecology and Management,*

Maroyi, A. (2017). Diversity of use and local knowledge of wild and cultivated plants in the Eastern Cape Province, South Africa. *Journal of Ethnobiology and Ethnomedicine*, 13, 43. https://doi.org/10.1186/s13002-017-0173-8

Matakala, P., Macucule, A., & Mausse, A. (2004). Local indigenous knowledge, uses and agroforestry potentials of Mafurra (*Trichilia emetica*). Nairobi, Kenya: World Agroforestry Centre (ICRAF).

Matsumoto, A. and Tsumura, Y. 2004. "Evaluation of cleaved amplified polymorphic sequence markers." Theor. Appl. Genet 110: 80–91.

Maua, J. O., Mugatsia-Tsingalia, H., Cheboiwo, J., & Odee, D. (2020). Population structure and regeneration status of woody species in a remnant tropical forest: A case study of South Nandi Forest, Kenya. *Global Ecology and Conservation, 21*,

McCoy, M. W., Bolker, B. M., Osenberg, C. W., Miner, B. G., & Vonesh, J. R. (2006). Size correction: Comparing morphological traits among populations and environments. *Oecologia, 148*(4), 547–554.

Medhi, R. P. (2006). Seed storage behavior and germination ecology of *Trichilia emetica*. *Journal of Tropical Forestry*, 22(3), 31–37.

Michaels, S.D. and Amasino, R.M. 1998. "A robust method for detecting single-nucleotide changes as polymorphic markers by PCR." Plant J. 14: 381–385.

Micheli, M.R., Bova, R., Calissano, P. and Dambrosio, E. 1993. "Randomly amplified polymorphic DNA fingerprinting using combinations of oligonucleotide primers." Biotechniques 15: 388-390.

Michelmore, R.W., Paran, I.and Kesseli, R. V. 1991. "Identification of markers linked to disease resistance gene by bulk segregant analysis: A rapid method to detect markers in specific genome region using segregating populations." Proc. Natl. Acad. Sci 88: 9828–32.

Miller, J.C. and Tanksley, S.D. 1990. "RFLP analysis of phylogenetic relationship and variation in genus lycopersicon." Theor. Appl.

Miller, P.T. and Henry, R. 2003. "Single-nucleotide polymorphism detection in plants using a single stranded pyrosequencing protocol with a universal biotinylated primer."

Mirski, P., Brzosko, E., Jędrzejczyk, I., Hędzel, J., Kotowicz, J., Kotowicz, B., & Urbaniak, L. (2017). Genetic structure of dioecious and trioecious *Salix myrsinifolia* populations at the border of geographic range. *Tree Genetics & Genomes, 13*(15). https://doi.org/10.1007/s11295-016-1096-6

Montalvão, A. P. L., Kersten, B., Fladung, M., & Müller, N. A. (2021). The diversity and dynamics of sex determination in dioecious plants. *Frontiers in Plant Science, 11*, 580488. https://doi.org/10.3389/fpls.2020.580488

Morellato, L. P. C. (2004). Phenology, sex ratio, and spatial distribution among dioecious species of *Trichilia* (Meliaceae). *Plant Biology, 6*(4), 491-497. https://doi.org/10.1055/s-2004-817910

Msanga HP, Maghembe JA. 1993. Germination of woodland mahogany (*Trichilia emetica*) following manual scarification and potassium nitrate treatments. Journal of Tropical Forest Science.

Mueller, U.G. and Wolfenbarger, L.L. 1999. "AFLP genotyping and fingerprinting." Trends Ecol.

Muellner, A. N., Greger, H., & Pannell, C. M. (2009). Genetic diversity and geographic structure in *Aglaia elaeagnoidea* (Meliaceae, Sapindales), a morphologically complex tree species, near the two extremes of its distribution.

Muellner, A. N., Samuel, R., Chase, M. W., & Coleman, A. (2010). Molecular phylogeny of Meliaceae: Evidence from plastid and nuclear markers.

Mugayi, L. C., & Mukanganyama, S. (2024). Antimycobacterial and antifungal activities of leaf extracts from *Trichilia emetica*. *Scientifica, 2024*.

Mugula, B. B. (2024). *Population structure, parasitic associations, distribution drivers, and genetic diversity of Osyris lanceolata (Hochst. & Steud.) in Uganda and Kenya* (Master's thesis). University of Nairobi.

Mugula, B. B., Kiboi, S. K., Kanya, I. J., Egeru, A., Okullo, P., Curto, M., & Meimberg, H. (2023). Microsatellites reveal divergence in population genetic diversity and structure of *Osyris lanceolata* (Santalaceae) in Uganda and Kenya. *BMC Ecology and Evolution.*

Muhammad Idrees, Muhammad Irshad. (2014). Molecular Markers in Plants for Analysis of Genetic Diversity: A Review EUROPEAN ACADEMIC RESEARCH - Vol. II, Issue 1 / April 2014 1521

Muyle, A., Martin, H., Zemp, N., Mollion, M., Gallina, S., Tavares, R., Silva, A., Bataillon, T., Widmer, A., Glémin, S., Touzet, P., & Marais, G. A. B. (2021). Dioecy is associated with high genetic diversity and adaptation rates in the plant genus *Silene*. *Molecular Biology and Evolution, 38*(3), 805–818. https://doi.org/10.1093/molbev/msaa229

Nagaoka, T., & Ogihara, Y. (1997). Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. *Theoretical and applied genetics*, *94*, 597-602. Retrieved from https://link.springer.com/article/10.1007/s001220050456

Okeno, J. A., Mutegi, E., de Villiers, S., Wolt, J. D., & Misra, M. K. (2012). Morphological Variation in the Wild-Weedy Complex of *Sorghum Bicolor* In Situ in Western Kenya: Preliminary Evidence of Crop-to-Wild Gene Flow? *International Journal of Plant Sciences*, *173*(5), 507–515. https://doi.org/10.1086/665266

Olsen, M., Hood, L., Cantor, C. and Botstein, D. 1989. “A common language for physical mapping of the human genome.” Science 245: 1434-1435.

Orwa, C. (2014). *Trichilia emetica species profile*. World Agroforestry Centre (ICRAF).

Orwa, C., Mutua, A., Kindt, R., Jamnadass, R, & Anthony, S. (2009) Agroforestry Database: A tree reference and selection guide version 4.0

Pammi, S., Schertz, K., Xu, G., Hart, G. and Mullet, J.E. 1994. "Randomamplified-polymorphic DNA markers in sorghum." Theor. Appl. Genet 89: 80-88.

Paran, I. and Michelmore, R.W. 1993. "Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce." Theor. Appl. Genet 85: 985-993.

Pennington, T. D. (2016). *Meliaceae*. In K. Kubitzki (Ed.), *The families and genera of vascular plants* (Vol. 11, pp. 239–267). Springer.

Pennington, T. D., & Styles, B. T. (2013). *A taxonomic revision of Trichilia in continental Africa.* Kew Bulletin Monographs, Royal Botanic Gardens, Kew.

Pennington, T. D., Styles, B. T., & Taylor, D. A. H. (1981). *Meliaceae*. In R. M. Polhill (Ed.), *Flora of Tropical East Africa* (pp. 1–120). Balkema

Perumal, A., Naidu, K. S. B., Sershen, Pillay, K., & Govender, P. (2020). Phytochemical composition and biological investigation of *Trichilia emetica* seed extracts. *Letters in Applied NanoBioScience, 9*(2), 1111–1116.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLOS ONE, 7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Phillips, J. D., Gillis, D. J., & Hanner, R. H. (2019). Incomplete estimates of genetic diversity within species: Implications for DNA barcoding. *Ecology and Evolution*, *9*(5), 2996-3010.

Phillips, R.L. and Vasil, I.K. 2001. "DNA-based markers in plants." In DNA-based markers in plants. Edited by Phillips RL and Vasil IK, 497. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

PROTA. (n.d.). *Trichilia emetica Vahl.* Plant Resources of Tropical Africa (PROTA4U database). https://www.prota4u.org

Qureshi, S.N., Saha, S., Kantety, R.V. and Jenkins, J.N. 2004. "EST-SSR: a new class of genetic markers in cotton." J. Cotton Sci 8: 112–123.

Rafinski, J. and Babik, W. 2000. "Genetic differentiation among northern and southern populations of the moor frog Rana arvalis Nilsson in central Europe." Hered 84(5): 610-618.

Rapley, R. and Harbron, S. 2001. Molecular Analysis and Genome Discovery. Chichester, UK: John Wiley & Sons Ltd.

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

Reamon-Buttner, S.M. and Jung, C. 2000. "AFLP derived STS markers for the identification of sex in Asparagus officinalis L." Theor. Appl. Genet 100: 432–438.

Reddy, M.P., Sarla, N. and Siddiq, E.A. 2002. "Inter simple sequence repeat (ISSR) polymorphism and itsapplication in plant breeding." Euphytica 128: 9–17.

REED, D.H and R. FRANKHAM (2001). How closely correlated are molecular and quantitative measures of genetic variations? A mata-analysis. Evolution 55: 1095-1103.

Ridout, C. and Donini, P. 1999. "Use of AFLP in cereals research." Trends Plant Sci 4: 76-79.

Rojas-Beltran, J.A., Fairbanks, D.J. and Maughan, P.J. 2005. "Development and use of an expressed sequenced tag library in quinoa (Chenopodium quinoa Willd.) for the discovery of single nucleotide polymorphisms." Plant Soc 168: 439–447.

Rossi, A. A., Almeida, M. T., & Brandão, R. L. (2019). Genetic variation in*Trichilia subincanum* using ISSR markers: Conservation implications for Atlantic forest species. *Conservation Genetics, 20*(5), 1049–1059. https://doi.org/10.1007/s10592-019-01189-5

Russell, J.R., Fuller, J.D., Macaulay, M., Hatz, B.G., Jahoor, A., Powell, W. and Waugh, R. 1997. "Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs." Theor. Appl. Genet 95: 714-722.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. 1988. “Primer- directed enzymatic amplification of DNA with a thermostable DNA polymerase.” Nature 239: 487–97.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. 1985. “Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia.” Science 230: 1350–54.

Sankar, A.A. and Moore, G.A. 2001. “Evaluation of inter-simple sequence repeat analysis for mapping in Citrus and extension of genetic linkage map.” Theor. Appl. Genet 102: 206–214.

Sebbenn, A. M., Licona, J. C., Mostacedo, B., & Degen, B. (2012). Gene flow in an overexploited population of *Swietenia macrophylla* King (Meliaceae) in the Bolivian Amazon. *Silvae Genetica, 61*(1–6), 212-220. https://doi.org/10.1515/sg-2012-0027

Schlotterer, C. and Tautz, D. 1992. “Slippage synthesis of simple sequence DNA.” Nucleic Acids Res 20: 211–15.

Schulmann, A.H. 2007. “Molecular markers to assess genetic diversity.” Euphytica 158(3): 313-321.

Southern, E.M. 1975. “Detection of specific sequences among DNA fragments separated by gel electrophoresis.” J. Mol. Biol 98: 503–17.

Stefano, C. M., Calazans, L. S., & Sakuragui, C. M. (2014). Morphological variation in *Trichilia* species and implications for taxonomy. *Botanical Journal of the Linnean Society*, 175(2), 189–205. https://doi.org/10.1111/boj.12165

Tanaka, H., Fukuda, N. and Shoyama, Y. 2006. "Identification and differentiation of Panax species are using ELISA, RAPD and Eastern blotting." Phytochem. Anal 17: 46-55.

Tautz, D. 1989. "Hypervariability of simple sequences as a general source for polymorphic DNA markers." Nucleic Acids Res 17: 6463-6471.

Tewolde, H., & Uchimiya, M. (2013). Management and propagation of multipurpose tropical trees in agroforestry. *Agroforestry Systems*, 87(5), 1043–1052. https://doi.org/10.1007/s10457-013-9610-4

Thiel, T., Michalek, R.K., Varshney, R.K. and Graner, A. 2003. "Exploiting EST databases for the development and characterization of gene-derived SSR markers in barley (*Hordeum vulgare* L.)." Theor. Appl. Genet 106: 411– 422.

Tlaxcala-Méndez, S. I., González-Rodríguez, A., Oyama, K., & Luna-Vega, I. (2016). Phenotypic plasticity and morphological variation in tropical tree species: Implications for adaptation. *Plant Ecology & Diversity, 9*(4), 389–400. https://doi.org/10.1080/17550874.2016.1207743

TRIEST, L. (2008): Molecular ecology and biography of Mangrove trees towards conceptual Insights on gene flow and barriers: A review of Aquatic Botany 89: 138-154.

Tsomele, G. F. (2022). *Structural, nutritional and protein functional properties of Trichilia emetica and Trichilia dregeana seeds* (Doctoral dissertation, University of KwaZulu-Natal). University of KwaZulu-Natal ResearchSpace.

Tsomele, G. F., Venter, E., Wokadala, O. C., Jooste, E., Dlamini, B. C., Ngobese, N. Z., & Siwela, M. (2021). Structural (gross and micro), physical and nutritional properties of Trichilia emetica and Trichilia dregeana seeds. *CyTA-journal of Food*, *19*(1), 483-492.

Tsomele, G. F., Venter, E., Wokadala, O. C., Jooste, E., Dlamini, B. C., & Ngobese, N. Z. (2021). Structural (gross and micro), physical and nutritional properties of *Trichilia emetica* and *Trichilia dregeana* seeds. *CyTA – Journal of Food, 19*(1), 483–492. https://doi.org/10.1080/19476337.2021.1915877

Tsomele, K., Akweni, C., & Mbele, S. (2021). Morphological diversity and seed traits in *Trichilia emetica* and *Trichilia dregeana*: Implications for domestication. *South African Journal of Botany, 139*, 110–119. https://doi.org/10.1016/j.sajb.2021.01.008

Tsuchihashi, Z. and Dracopoli, N.C. 2002. “Progress in high-throughput SNP genotyping methods.” Pharmacogenomics J 2: 103–110.

Usman, A., Thoss, V., & Itodo, A. U. FLAVONOL AND ITS GLYCOSIDE FROM THE STEM OF *Trichilia emetica*.

Van Wyk, B. E., & Coe, F. G. (2021). The ethnobotany and chemistry of South African Meliaceae: A review. *Journal of Ethnopharmacology, 281*, 114530. https://doi.org/10.1016/j.jep.2021.114530

Varshney, R.K., Graner, A. and Sorrells, M.E. 2005. “Genic microsatellite markers in plants: features and applications.” Trends Biotechnol 23: 48–55.

Vejl, P. 1997. “Identification of genotypes in hop (*Humulus lupulus* L.) by RAPD analsysis using program Gel Manager for Windows.” Rostlinna Vyroba 43: 325-331.

Verstege, K.J.E., van-Aart, J.W., Geerlings, H., Buntjer, J.B., van-Oeveren, A.J. and Vos, P. 2004. “SNPWave: A flexible multiplexed SNP genotyping technology.” Nucl. Acids Res 32: e47.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van De Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. “AFLP: a new technique for DNA fingerprinting.” Nucleic Acids Res 23: 4407- 4414.

Wadasinghe, G. S., Perera, G. A. D., & Ranwala, S. M. W. (2025). Genetic variation in tropical tree species assessed by ISSR markers: A case of *Gmelina arborea* and *Pongamia pinnata*. *Journal of Forestry and Environmental Science, 41*(2), 55–68.

Wadasinghe, R. R., Kalansuriya, P., & Attanayake, A. P. (2025). Encapsulation of Gmelina arborea Roxb. and Spondias pinnata (LF) Kurz. Aqueous Extracts in Nanoliposomes: Synthesis, Characterization, and In Vitro Screening of Antidiabetic Activity. *Chemistry & Biodiversity*, *22*(5), e202402998.

Wass, P. (1995). Kenya’s Indigenous Forests: Status, Management, and Conservation: IUCN. *Gland, Switzerland & Cambridge, UK*.

Weber, J.L. and May P.E. 1989. “Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction.” American J. Human Genet 44: 388-396.

Weising, K., Nybom, H., Wolff, K. and Kahl, G. 2005. “Detecting DNA variation by molecular markers.” In DNA fingerprinting in plants: principles, methods, and applications, Ch. 2: 21-74. 2nd ed. Boca Raton, FL: Taylor and Francis.

Weller, J.W. and Reddy, A. 1997. “Fluorescent detection and analysis of RAPD amplicons using the ABI PRISM DNA sequencers.” In Fingerprinting methods based on arbitrarily primed PCR, edited by Micheli MR and Bova R., 81-92. Springer Lab Manual.

Welsh, J. and McClelland, M. 1990. “Fingerprinting genomes using PCR with arbitrary primers.” Nucleic Acids Res18: 7213-7218.

Welsh, J. and McClelland, M. 1991. “Genomic fingerprinting using arbitrarily primed PCR and a matrix of pairwise combinations of primers.” Nucleic Acids Res 19: 5275-5279.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. 1990. “DNA polymorphisms amplified by arbitrary primers are useful as genetic markers.” Nucl. Acids Res 18: 6531–6535.

Wright S, (1978) Evolution and genetics of populations. Vol 4. Variability within and among natural populations. University of Chicago press, Chicago.

Zhang JF, Kimatu JN, Guo WL, Liu B. (2009). Habitat fragmentation causes rapid genetic differentiation and homogenization in natural plant populations - A case study in Leymuschinensis. Afr J Biotech 8(150):3440–3447.

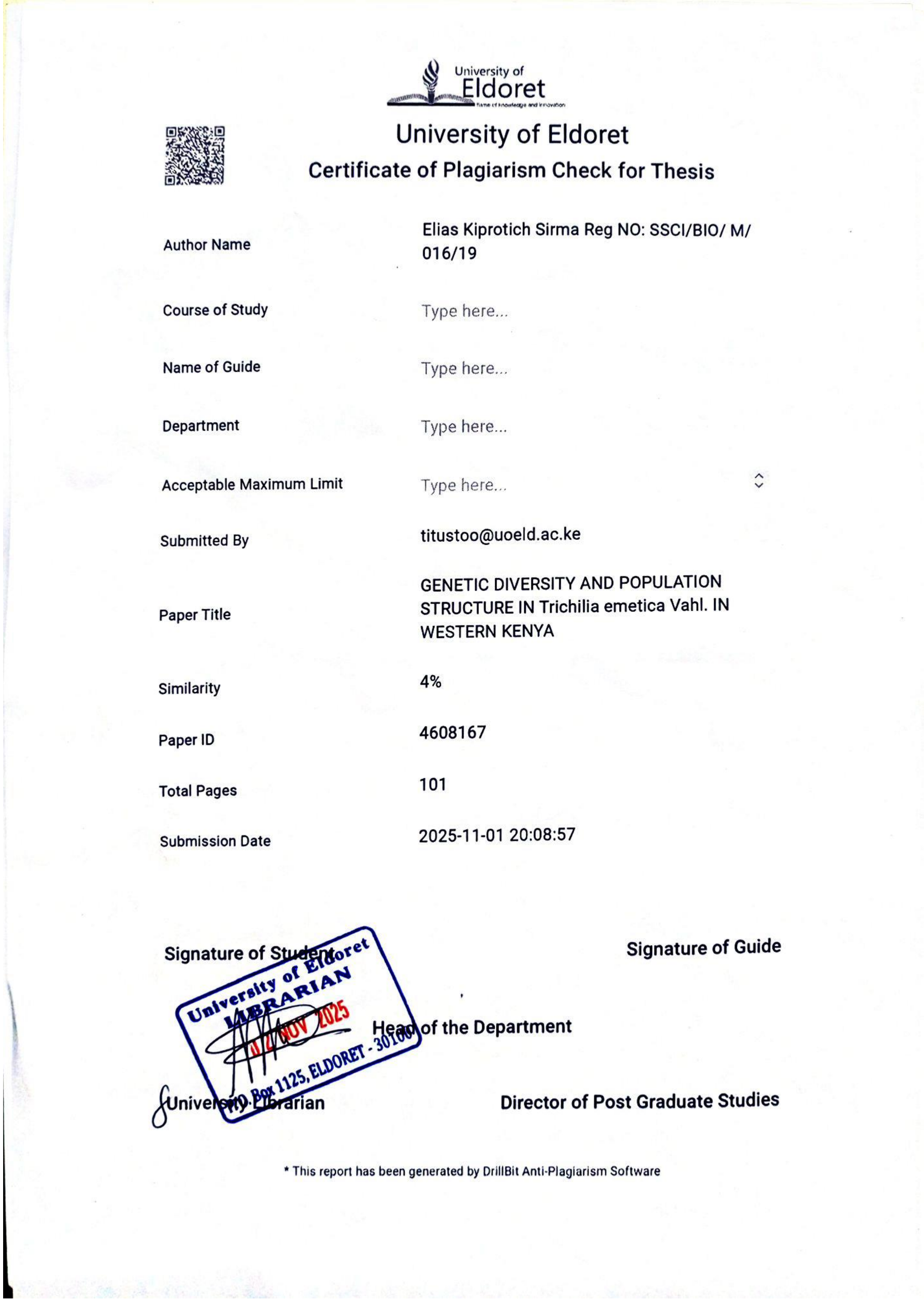
Zhang, Y., Chen, Y., & Li, D. (2009). Genetic diversity and population structure in selfing and outcrossing plant species: A comparative study using molecular markers. *Plant Biology, 11*(5), 705–712. https://doi.org/10.1111/j.1438-8677.2008.00169.x

Zietkiewicz, E., Rafalski, A. and Labuda, D. 1994. "Genome fingerprinting by simple sequence re-peats (SSR)- anchored polymerase chain reaction amplification." Genomics 20(2): 176-83.

Zuidema, P. A., Brienen, R. J. W., & During, H. J. (2009). The effects of selective logging on morphological traits and population structure of tropical trees. *Forest Ecology and Management, 257*(3), 1245–1254. https://doi.org/10.1016/j.foreco.2008.11.013

APPENDICES

Appendix II. Similarity Report



University of Eldoret

University of Eldoret

Certificate of Plagiarism Check for Thesis

Author Name	Elias Kiprotich Sirma Reg NO: SSCI/BIO/ M/ 016/19
Course of Study	Type here...
Name of Guide	Type here...
Department	Type here...
Acceptable Maximum Limit	Type here...
Submitted By	titustoo@uoeld.ac.ke
Paper Title	GENETIC DIVERSITY AND POPULATION STRUCTURE IN Trichilia emetica Vahl. IN WESTERN KENYA
Similarity	4%
Paper ID	4608167
Total Pages	101
Submission Date	2025-11-01 20:08:57

Signature of Student

Signature of Guide

Head of the Department

University Librarian

Director of Post Graduate Studies

* This report has been generated by DrillBit Anti-Plagiarism Software