The Nelson Mandela AFrican Institution of Science and Technology

NM-AIST Repository **https://dspace.mm-aist.ac.tz**

Life sciences and Bio-engineering Research Articles [LISBE]

2023

Phylogenetic Diversity of Allspice (Pimenta Dioica) Collections from Tanzania Using Chloroplast (Cp) Rbcl Gene

Lutege, Raymond

Heliyon

https://dspace.nm-aist.ac.tz/handle/20.500.12479/2328

Provided with love from The Nelson Mandela African Institution of Science and Technology

Tittle: Phylogenetic diversity of Allspice (Pimenta dioica) collections from Tanzania using chloroplast (cp) rbcL gene

Name: Mr. Raymond Lutege

Email: luteger@nm-aist.ac.tz

School: School of life science and Bio-engineering,

Department: Sustainable Agriculture, Biodiversity and Ecosystem

Institution: The Nelson Mandela African Institution of science and Technology Arusha (NM-AIST).

Name: Dr Pavithravani Venkataramana

Email: pavithravani.venkataramana@nm-aist.ac.tz

School: School of life science and Bio-engineering,

Department: Sustainable Agriculture, Biodiversity and Ecosystem

Institution: The Nelson Mandela Institute of science and Technology Arusha (NM-AIST).

Name: Dr Joseph Ndunguru

Email: jndunguru2003@yahoo.co.uk

Institution: Tanzania Plant Health and Pesticide Authority (TPHPA)

Name: Dr Laura Boykin

Email: lboykin@mac.com

Institution: Bio team USA

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

1. Introduction

Myrtaceae is a highly diverse family of plants that includes over 130 genera and 4000 species that are distributed around the world [1]. One such species, Allspice (*Pimenta dioica)* grow up to 20 meters tall and has sweet-smelling leaves that measures between 9 and 20 centimeters.[2]. Allspice thrives in semitropical lowland regions with an average temperature of 15°C to 32°C, 600 meters above sea level, and an average rainfall of 1,500 to 2,500mm annually [3]. This species is indigenous to Southern Mexico and Central America, and was introduced to Tanzania by Arab traders in the 19th century, and it quickly became popular in the country's coastal areas [4]. Allspice has spread to many countries due to the trade of spices, as well as seeds dispersal by birds [5]. Allspice is found in various countries around the world, and in Africa, the tree has been reported in 27 countries [1]

The tree produces a fragrant fruit that measures between 4-8 mm in size, which is used as an ingredient in the food industry [6]. Allspice is highly valued economically, as it is used in the manufacture of spices in the food industry (65-70%), pimento oil production (20-25%), and for domestic purpose (5 - 10 %) [7]. Allspice is known for its pesticidal properties, and it has been used as an insecticide for plant protection in various pests such as *Reticulitermes speratus, Acanthoscelides obtectus, and Sitophilus zeamais*[8], [9]. Oils derived from this has antifungal [10], antibacterial [11], and antinematicidal properties[12]. Additionally, the use of Allspice in traditional medicine has been prevalent for its therapeutic properties. The powder extracted from its fruit is known to be effective in treating several health problems, including menstrual discomfort, inflammation, stomach aches, and muscle pain [13]. Therefore, it is imperative to accurately identify it to guarantee its safe and effective use. Also, there is a need to create awareness among local communities and promote proper identification and utilization of Allspice.

The complexity of plant diversity is shaped by various biotic and abiotic factors [14]. Environmental factors, including climate, topography, resource availability, and disturbance like wildfires, flooding, deforestation, and agriculture., play a critical role in determining species composition and shaping evolutionary processes [15]. Two of the most important climatic factors, temperature and rainfall, are strong predictors of plant diversity. Ecosystems with average temperatures above the thresholds of 25°C-30°C (77-86°F) for temperate and boreal areas and 28°C-32°C (82-90°F) for tropical regions exhibit lower plant diversity. Conversely, low temperature stress on plant diversity is usually observed at temperatures around -10°C to -15°C

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

(14°F to 5°F) in temperate and boreal ecosystems and between 5°C and 10°C (41°F to 50°F) in tropical ecosystems. [16]. Plant diversity is highest at intermediate rainfall levels (1000-1500 mm per year) and declines when rainfall falls below 500 mm or rises above 1500 mm in tropical forests, and below 400 mm per year in grasslands [17]. . Also, edaphic factors, such as soil type and nutrient availability, also influence plant diversity[18]. On the other hand, genetic variation, mutation, and genetic drift are important mechanisms that drive evolution and shape genetic diversity within populations [19].

However, Allspice has been misidentified as *Pimenta racemosa* and other similar trees in Myrtaceae family. This confusion is primarily due to the striking similarities in their morphological characteristics, which can make it difficult to distinguish between the two [20]. DNA barcoding is a useful method for identifying and classifying plant species, especially when morphological features are insufficient or ambiguous[21]. It utilizes a specific genetic marker, such as the chloroplast (cp) *rbcL* gene, which encodes the large subunit of ribulose bisphosphate carboxylase [22]. This gene has average length of 1400 bp [23], and is commonly used as a DNA barcode in plants due to its high variability between species, highly conserved, experiencing low levels of mutations and its presence in nearly all plants [24]–[29]. Therefore, cp*rbcL* gene is a potentially tool that can be used in the study of evolutionary, intraspecies diversity, and phylogenetic variations among species.

This study aims to investigates the phylogenetic relationship between various collection of Allspice from different locations in Tanzania and other related members of the Myrtaceae family using cp*rbcL* gene, and the outcomes of this research are anticipated to offer valuable understandings into the evolutionary relationships among species and the intraspecies diversity within the Myrtaceae family.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

2. Materials and Methods

2.1 Plant Material

The study sampled a total of fifty nine (59) collections of Allspice from various locations in Tanzania. These samples were collected from different areas: Kizimbani – Zanzibar (23), Kizugu botanical garden (25), Zigi Amani forest (9), and World Vegetable Center (AVRDC) –Arusha (2). Fresh leaves were collected from the field and then subjected to a drying process using silica gels. All samples were kept in storage at a temperature of -80 ◦C until they were processed. Table 1 and figure 1 provide more detailed information about collections and locations where the samples were collected.

Table 1 Location and sample identity of Allspice collections from various areas in Tanzania

Location	Sample Identity	Longitude	Latitude	Altitude (masl)	No. of collections
Kizimbani – Zanzibar	TZnK	39°12'90"E	6°5'34" S	119	23
World Vegetable Center (AVRDC) – Arusha	TAR	36°41'15"E	3°23'19"S	1400	2
Kizugu Botanical garden, Tanga	TTKB	38°39'52"E	5°6’49"S	2289	25
Zigi Amani Forset, Tanga	TTZF	38°38'59"E	5°3’49"S	2289	9
Total					**59**

2.2 DNA extraction

The CTAB procedure was used to extract genomic DNA [30]. Leaves were dried overnight at a temperature of 70°C. Then, CTAB buffer was warmed at 65°C for 15 minutes. After that, pieces of dried leaves were ground into a fine powder using Geno grinding machine (2010 Geno/Grinder®) with two steel grinding media at 1200rpm for 40 seconds, two times. 1ml of warm CTAB buffer was added to the eppendorf tube containing the fine powder and then warmed at 65°C for 30 minutes, while shaking the tube in the interval of 10 minutes. Next, centrifugation was done at 13,000rpm for 15 minutes at room temperature, and the supernatant was transferred to a new eppendorf tube. An equal volume of chloroform-isoamyl alcohol (24:1, v/v) was added and centrifuged again for 15 minutes. The upper layer was collected, and 0.7 of the total volume of cold isopropanol was added and stored at -20°C for 1 hour, then centrifuged at 13,000 rpm for 30 minutes. The pellet was collected, washed with 500µl of ethanol, and centrifuged again at 13,000 rpm for 10 minutes. The pellet was air-dried for 45 minutes and suspended in nuclease-free water, which was then treated with RNase (10 mg/mL) and incubated at 37°C for 1 hour. Finally, the

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

DNA concentration was measured and purity was determined by agarose electrophoresis and estimating the ratio of absorbance at 260 nm to that at 280 nm (A260/A280), respectively. The spectrophotometer (Metter Toledo UV/VIS) was used to measure the purity of DNA that was isolated. The quality assessment was performed by utilizing agarose gel electrophoresis with a 1.5% concentration (Bio-Rad's Mini-sub cell GT electrophoresis system). Subsequently, the gels were observed via a UV transilluminator manufactured by Invitrogen under which the image was taken using a Gel documentation system.

2.3 PCR amplification and sequencing of cp*rbcL* gene

The Allspice cp*rbcL* gene marker from cpDNA was used in this study. Primer pairs used in PCR amplification were P610 as forward primer 5'-ATGTCACCACAAACAGAGACTAAAGC-3' and P609 as reverse primer 5'-GTAAAATCAAGTCCACCRCG-3'[31]. The PCR program was as follows: denaturation at 94°C for five minutes was followed by 35 cycles at 94°C for 30 seconds, annealing at 52°C for 30 seconds, an extension of 72°C for 1 minute and final extension of 72°C for 10 minutes. Each reaction contained 10.5µl of water, 12.5µl Taq DNA Polymerase 2x Master Mix RED (Ampliqon A/S, Stenhuggervej 22, Denmark), 0.5µl of each primer and 1.0µl of DNA template were added to make a final volume of 25µl [32]. The PCR products of the cp*rbcL* gene were visualized by running them on a 1.5% agarose gel with ethidium bromide staining. The single band of PCR amplicon were purified using a column-based DNA purification kit. Thenafter, the purified PCR product was sent to Inqaba Biotech Pretoria, South Africa for Sanger sequencing using Sanger's di-deoxy sequencing method on an ABI prism 3700 DNA analyzer. The resulted cp*rbcL* sequence obtained from Allspice was submitted to the NCBI database and the accession numbers obtained range from OP985342 to OP9885400.

2.4 DNA sequence alignment and phylogenetic analyses

Geneious Prime version 2023.0.1 (www.geneious.com) was used to assess the quality of the nucleotide sequences, which allowed for obtaining consensus arrangements of nucleotides. The study of the evolution of allspice involved constructing a phylogenetic tree with major species from the Myrtaceae family. To perform this analysis, a BLASTn search (http://www.ncbi.nlm.nih.gov) was conducted to compare the query sequence with the subject sequence available in the NCBI database. For each event, six to ten closely related sequences were selected and multiple sequence alignment was performed using MUSCLE 5.1, which is integrated

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

into Geneious Prime software [33]. In order to examine the differences among the sequences, a distance matrix was created, and using the dissimilarities expressed in the matrix, a phylogenetic tree was constructed using Geneious Prime software. To assess the structure of the phylogenetic tree, the bootstrap method was employed with 1000 replicates for all nodes[34]. *Eucalyptus behriana* (MW446388) was used as an out group to position the root of tree. Bootstrap analysis was used to re-evaluate the resulting rooted tree topologies using 1,000 resampling of the data.

3. Results

3.1 PCR amplification and sequence analysis

The electrophoresis of the allspice cp*rbcL* gene amplified by PCR was successful on a 1.5% agarose gel. In the agarose gel (Fig.3 a – c), the first well shows ladder DNA and the remaining wells indicate amplified cp*rbcL* gene product. In this case, the gel was able to separate the amplified cp*rbcL* gene product from other DNA fragments. The thick and single band present in the gel confirms that the amplification was successful and that the size of the amplified cp*rbcL* gene was 560 bp. The nucleotide content of the allspice cp*rbcL* gene was analyzed using Geneious Prime software version 2023.0.1. The nucleotide statistics for all 59 sequences were calculated, and it was determined that the amplified cp*rbcL* gene contained a total of 33,064 nucleotides with a mean molecular weight of 173kDa for ssDNA and 346.234kDa for dsDNA. The nucleotide composition of the amplified allspice cprbcL gene was analyzed, and it was found to consist of 9,381 bases of adenine (A), 7,149 bases of cytosine (C), 7,420 bases of guanine (G), and 9,114 bases of thymine (T). These bases correspond to 28.4%, 21.6%, 22.4%, and 27.6% of the gene, respectively with a GC content of 44%.

3.2 Multiple sequence alignment and Phylogenetic tree analyses

The BLASTn searches conducted in this study yielded significant results, with a high degree of similarity ranging from 96% – 100% and E-value of 0.0 between the query sequence and subject sequences obtained from GenBank. Specifically, the 59 sequences of allspice cp*rbcL* genes were found to be similar to 17 species within the Myrtaceae family, as shown in Tables 2. Results of the study indicate that the Allspice isolate with accession number TZnK2_OP985343 showed 100% similarity with six members of the Myrtaceae family, including *Eucalyptus torquata* (NC_022401), *Eucalyptus spathulata* (NC_022400), *Eucalyptus torquata* (KC180794), *Eucalyptus spathulata* (KC180793), *Syzygium polyanthum* (OQ355361), and *Syzygium*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

aromaticum (ON920513). Additionally, three isolates from Kizimbani – Zanzibar and Kizugu botanical garden exhibited genetic similarity ranging from 96.77% to 96.59% with 11 members of the Myrtaceae family, including *Luma apiculata* (KX162972), *Eugenia aggregata* (OP650216), *Eugenia Selloi* (MN095411), *Myrcianthes pungens* (MN095409), *Campomanesia xanthocarpa* (KY392760), *Acca sellowiana* (KX289887), *Syzygium samarangense* (NC_060657), *Lophomyrtus bullata* (MW214669), *Lenwebbia prominens* (MW214668), *Lenwebbia lasioclade* (MW214667), and *Syzygium nervosum* (NC_053907). Furthermore, twenty-three Allspice isolates exhibited 100% to 99.82% similarity with the 11 members of the Myrtaceae family, while 22 isolates showed 98.94% to 98.76% similarity to the same group of 11 members. Additionally, five other isolates exhibited 98.92% to 98.75% and 98.91% to 98.73% similarity to the same members of the Myrtaceae family. The isolates was collected from various locations, including Kizimbani – Zanzibar, Kizugu botanical garden, Amani Zigi Forest, and World vegetable center – Arusha.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

Table 2: BLASTn results for allspice cp*rbcL* gene with their closest match in GenBank

No	Isolate/AN	Size (bp)	Closest Match/AN	% Sequence Similarity	Source	Reference
1	TZnK2_OP985343	463	Eucalyptus torquata (NC_022401)	100	Plastid chloroplast, Australia	Baly *et al* (2013)
			Eucalyptus spathulata (NC_022400)	100	Plastid chloroplast, Australia	Baly *et al* (2013)
			Eucalyptus torquata (KC180794)	100	Plastid chloroplast, Australia	Baly et al (2013)
			Eucalyptus spathulata (KC180793)	100	Plastid chloroplast, Australia	Baly et al (2013)
			Syzygium polyanthum (OQ355361)	100	Plastid chloroplast	Unpublished
			Syzygium aromaticum (ON920513)	100	Plastid chloroplast, China	Unpublished
2	TZnK22_OP985362	558	Eugenia aggregata (OP650216)	96.77	Plastid organelle, China	Unpublished
3	TTKB2_OP985368	562	Luma apiculata (KX162972)	96.77	Plastid organelle, UK	Unpublished
4	TZnK24_OP985364	552	Eugenia Selloi (MN095411)	96.59	Plastid organelle, Brazil	Rodrigues et al (2020)
			Mycianthes pungens (MN095409)	96.59	Plastid organelle, Brazil	Rodrigues et al (2020)
			Campomanesia xanthocarpa (KY392760)	96.59	Plastid Chloroplast, Brazil	Unpublished
			Acca sellowiana (KX289887)	96.59	Plastid Chloroplast, Brazil	Machado et al (2017)
			Syzygium samarangense (NC_060657)	96.59	Plastid organelle, China	Wei, (2021)
			Lophomyrtus bullata (MW214669)	96.59	Plastid organelle, New Zealand	Maurin (2020)
			Lenwebbia prominens (MW214668)	96.59	Plastid organelle, New Zealand	Maurin (2020)
			Lenwebbia lasioclade (MW214667)	96.59	Plastid organelle, New Zealand	Maurin (2020)
			Syzygium nervosum (NC_053907)	96.59	Plastid Chloroplast, China	Unpublished

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

Table 2: (cont.)

No	Isolate/AN	Size (bp)	Closest Match/AN	% Sequence Similarity	Source	Reference
5	TZnK1_OP985342	568	Luma apiculata (KX162972)	100	Plastid organelle, UK	Unpublished
6	TZnK3_OP985344	565	Eugenia aggregata (OP650216)	100	Plastid organelle, China	Unpublished
7	TZnK4_OP985345	565	Eugenia Selloi (MN095411)	99.82	Plastid organelle, Brazil	Rodrigues et al (2020)
8	TZnK5_OP985346	564	Mycianthes pungens (MN095409)	99.82	Plastid organelle, Brazil	Rodrigues et al (2020)
9	TZnK6_OP985347	569	Campomanesia xanthocarpa (KY392760)	99.82	Plastid chloroplast, Brazil	Unpublished
10	TZnK7_OP985348	565	Acca sellowiana (KX289887)	99.82	Plastid organelle, Brazil	Machado et al (2017)
11	TZnK8_OP985349	561	Syzygium samarangense (NC_060657)	99.82	Plastid organelle, China	Wei, (2021)
12	TZnK9_OP985350	563	Lophomyrtus bullata (MW214669)	99.82	Plastid organelle, New Zealand	Maurin (2020)
13	TZnK10_OP985351	567	Lenwebbia prominens (MW214668)	99.82	Plastid organelle, New Zealand	Maurin (2020)
14	TZnK11_OP985352	559	Lenwebbia lasioclade (MW214667)	99.82	Plastid organelle, New Zealand	Maurin (2020)
15	TZnK12_OP985353	563	Syzygium nervosum (NC_053907)	99.82	Plastid chloroplast, China	Unpublished
16	TZnK13_OP985354	566				
17	TZnK14_OP985355	563				
18	TZnK15_OP985356	565				
19	TZnK20_OP985360	567				
20	TTKB3_OP985369	563				
21	TTKB4_OP985370	563				
22	TTKB5_OP985371	553				
23	TTKB7_OP985373	563				
24	TTKB8_OP985374	567				
25	TTZF3_OP985394	562				
26	TTZF4_OP985395	556				
27	TTZF6_OP985397	562				

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

Table 2: (cont.)

No	Isolate/A N	Size (bp)	Closest Match/AN	% Sequence Similarity	Source	Reference
28	TZnK18_OP985358	565	Eugenia aggregata (OP650216)	98.94	Plastid organelle, China	Unpublished
29	TAR2_OP985366	564	Luma apiculata (KX162972)	98.93	Plastid organelle, UK	Unpublished
30	TTKB1_OP985367	564	Eugenia Selloi (MN095411)	98.76	Plastid organelle, Brazil	Rodrigues et al (2020)
31	TTKB6_OP985372	564	Mycianthes pungens (MN095409)	98.76	Plastid organelle, Brazil	Rodrigues et al (2020)
32	TTKB9_OP985375	568	Campomanesia xanthocarpa (KY392760)	98.76	Plastid Chloroplast, Brazil	Unpublished
33	TTKB12_OP985378	567	Acca sellowiana (KX289887)	98.76	Plastid Chloroplast, Brazil	Machado et al (2017)
34	TTKB16_OP985382	566	Syzygium samarangense (NC_060657)	98.76	Plastid organelle, China	Wei, (2021)
35	TTKB17_OP985383	569	Lophomyrtus bullata (MW214669)	98.76	Plastid organelle, New Zealand	Maurin (2020)
36	TTKB18_OP985384	564	Lenwebbia prominens (MW214668)	98.76	Plastid organelle, New Zealand	Maurin (2020)
37	TTKB21_OP985387	566	Lenwebbia lasioclade (MW214667)	98.76	Plastid organelle, New Zealand	Maurin (2020)
38	TTKB23_OP985389	564	Syzygium nervosum (NC_053907)	98.76	Plastid Chloroplast, China	Unpublished
39	TTKB24_OP985390	564				
40	TTZF1_OP985392	564				
41	TTKB11_OP985377	561				
42	TTKB13_OP985379	560				
43	TTKB14_OP985380	560				
44	TTKB15_OP985381	560				
45	TTKB20_OP985386	560				
46	TTKB25_OP985391	561				
47	TTZF8_OP985399	559				
48	TTZF9_OP985400	560				
49	TZnK19_OP985359	563				

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

Table 2: (cont.)

No	Isolate/AN	Size (bp)	Closest Match/AN	% Sequence Similarity	Source	Reference
50	TZnK21_OP985361	558	Eugenia aggregata (OP650216)	98.92	Plastid organelle, China	Unpublished
51	TAR1_OP985365	554	Luma apiculata (KX162972)	98.92	Plastid organelle, UK	Unpublished
52	TZnK23_OP985363	560	Eugenia Selloi (MN095411)	98.75	Plastid organelle, Brazil	Rodrigues et al (2020)
53	TTKB10_OP983776	563	Mycianthes pungens (MN095409)	98.75	Plastid organelle, Brazil	Rodrigues et al (2020)
54	TZnK17_OP985357	553	Campomanesia xanthocarpa (KY392760)	98.75	Plastid Chloroplast, Brazil	Unpublished
			Acca sellowiana (KX289887)	98.75	Plastid organelle, Brazil	Machado et al (2017)
			Syzygium samarangense (NC_060657)	98.75	Plastid organelle, China	Wei, (2021)
			Lophomyrtus bullata (MW214669)	98.75	Plastid organelle, New Zealand	Maurin (2020)
			Lenwebbia prominens (MW214668)	98.75	Plastid organelle, New Zealand	Maurin (2020)
			Lenwebbia lasioclade (MW214667)	98.75	Plastid organelle, New Zealand	Maurin (2020)
55	TTKB19_OP985385	556	Luma apiculata (KX162972)	98.91	Plastid organelle, UK	Unpublished
56	TTKB22_OP985388	553	Eugenia Selloi (MN095411)	98.73	Plastid organelle, Brazil	Rodrigues et al (2020)
57	TTZF7_OP985398	568	Mycianthes pungens (MN095409)	98.73	Plastid organelle, Brazil	Rodrigues *et al* (2020)
58	TTZF5_OP985396	557	Campomanesia xanthocarpa (KY392760)	98.73	Plastid Chloroplast, Brazil	Unpublished
59	TTZF2_OP985393	556	Acca sellowiana (KX289887)	98.73	Plastid organelle, Brazil	Machado et al (2017)
			Syzygium samarangense (NC_060657)	98.73	Plastid organelle, China	Wei, (2021)
			Lophomyrtus bullata (MW214669)	98.73	Plastid organelle, New Zealand	Maurin (2020)
			Lenwebbia prominens (MW214668)	98.73	Plastid organelle, New Zealand	Maurin (2020)
			Lenwebbia lasioclade(MW214667)	98.73	Plastid organelle, New Zealand	Maurin (2020)
			Syzygium nervosum (NC_053907)	98.73	Plastid Chloroplast, China	Unpublished

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

Key: TZnK = Kizimbani – Zanzibar, TTB = Kizugu Botanical Garden Tanga, TTZF = Amani Zigi Forest Tanga, TAR = World Vegetable Center (AVRDC) – Arusha, bp = base pair, AN = Accession number.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

A total of 17 sequences from different species of Myrtaceae family were selected for the analysis, and multiple sequence alignment was performed using the Geneious prime software. Table 3 show the phylogenetic parameters obtained from sequences alignment.

Table 3: Multiple sequence alignment phylogenetic parameters.

Data	**Value**
Identical number of bases	539
Percentage of identical number of bases (%)	94.90
Pairwise identity between sequence (%)	99.20
Mean of alignment	553
Standard deviation of alignment	26.2
Patristic distance range	0.01 - 0.04

The degree of genetic divergence between sequences ranges from 0.01 – 0.04 at gene level. The results presented in table 4 of this study indicate that a total of 24 isolates exhibited a patristic distance of 0.01 to 17 members of the Myrtaceae family, which represents the minimum level compared to isolate 32, which showed an intermediate level of 0.02. However, three allspice isolates showed a significantly higher level of 0.04, which is the maximum level observed. However, the results indicate that the Myrtaceae family members have a relatively low level of genetic diversity, as evidenced by the small range of patristic distances observed. In contrast, the allspice isolates exhibited a higher level of genetic diversity, with some isolates showing a significantly higher patristic distance.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

Table 4. Patristic distance between allspice (isolate) cp*rbcL* gene and GenBank sequences

Isolate/AN	GenBank sequences accession numbers																
	KX162972	OP650216	MN095411	MN095409	KY392760	KX289887	NC_060657	MW214669	MW214668	MW214667	NC_053907	NC_022401	NC_022400	KC180794	KC180793	OQ355361	ON920513
TZnK1_OP985342	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK2_OP985343	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK3_OP985344	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK4_OP985345	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK5_OP985346	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK6_OP985347	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK7_OP985348	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK8_OP985349	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK9_OP985350	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK10_OP985351	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK11_OP985352	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK12_OP985353	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK13_OP985354	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK14_OP985355	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK15_OP985356	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TZnK20_OP985360	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TTKB3_OP985369	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TTKB4_OP985370	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TTKB5_OP985371	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TTKB7_OP985373	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TTKB8_OP985374	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TTZF3_OP985394	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TTZF4_OP985395	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
TTZF6_OP985397	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

Table 4: (cont.)

Isolate/AN	GenBank sequences accession numbers																
	KX162972	OP650216	MN095411	MN095409	KY392760	KX289887	NC_060657	MW214669	MW214668	MW214667	NC_053907	NC_022401	NC_022400	KC180794	KC180793	OQ355361	ON920513
TZnK17_OP985357	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TZnK18_OP985358	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TZnK19_OP985359	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TZnK21_OP985361	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TZnK23_OP985363	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TAR1_OP985365	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TAR2_OP985366	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB1_OP985367	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB6_OP985372	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB9_OP985375	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB10_OP985376	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB11_OP985377	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB12_OP985378	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB13_OP985379	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB14_OP985380	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB15_OP985381	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB16_OP985382	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB17_OP985383	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB18_OP985384	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB19_OP985385	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB20_OP985386	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB21_OP985387	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB22_OP985388	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB23_OP985389	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

Table 4: (cont.)

Isolate/AN	GenBank sequences accession number																
	KX162972	**OP650216**	**MN095411**	**MN095409**	**KY392760**	**KX289887**	**NC_060657**	**MW214669**	**MW214668**	**MW214667**	**NC_053907**	**NC_022401**	**NC_022400**	**KC180794**	**KC180793**	**OQ355361**	**ON920513**
TTKB24_OP985390	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB25_OP985391	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTZF1_OP985392	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTZF2_OP985393	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTZF5_OP985396	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTZF7_OP985398	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTZF8_OP985399	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTZF9_OP9853400	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
TTKB2_OP985368	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
TZnK22_OP985362	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
TZnK24_OP985364	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04

Key:

Accession numbers		Species name	Accession numbers		Species name
KX162972	=	*Luma apiculata*	MW214669	=	*Lophomyrtus bullata*
OP650216	=	*Eugenia aggregata*	MW214668	=	*Lenwebbia prominens*
MN095411	=	*Eugenia Selloi*	MW214667	=	*Lenwebbia lasioclade*
MN095409	=	*Myrcianthes pungens*	NC_053907	=	*Syzygium nervosum*
KY392760	=	*Campomanesia xanthocarpa*	NC_022401	=	*Eucalyptus torquata*
KX289887	=	*Acca sellowiana*	NC_022400	=	*Eucalyptus spathulata*
NC_060657	=	*Syzygium samarangense*	KC180794	=	*Eucalyptus torquata*
ON920513	=	*Syzygium aromaticum*	KC180793	=	*Eucalyptus spathulata*
AN	=	Accession number	OQ355361	=	*Syzygium polyanthum*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

The phylogenetic tree (Fig. 4) shows one main clade that contains four subclades. The first subclade is composed of *Eucalyptus torquata* (KC180794), *Eucalyptus spathulata* (NC_022400), *Syzygium aromaticum* (ON920513), and *Eucalyptus spathulata* (KC180793), which are very closely to allspice isolate with accession numbers TZnK2_OP985343 and TTKB5_OP985371 from the Kizimbani – Zanzibar and Kizugu botanical gardens respectively. Additionally, *Syzygium polyanthum* (OQ355361) was closely related to allspice isolates with accession numbers TTZF4_OP985395, TTZF3_OP985394, and TTZF6_OP985397 from Amani Zigi Forest, TZnK8_OP985349, TZnK9_OP985350, and TZnK11_OP985352 from Kizimbani –Zanzibar. *Luma apiculata* (KX162972) and *Eugenia aggregata* (OP650216) clustered with allspice isolates with accession numbers TZnK12_OP985353 and TZnK14_OP985355 from Kizimbani – Zanzibar, TTKB3_OP985369, TTKB4_OP985370, and TTKB7_OP985373 from Kizugu botanical garden. *Campomanesia xanthocarpa* (KY392760) was closely related to isolate with accession numbers TZnK1_OP985342, TZnK6_OP985347, TZnK3_OP985344 to TZnK15_OP985356, and TZnK20_OP985360 from Kizimbani –Zanzibar and TTKB8_OP985374 from Kizugu botanical garden.

The allspice isolate gene sequences with accession numbers TTKB1_OP985367, TTKB6_OP985372, TTKB9_OP985375 to TTKB25_OP985391 from the Kizugu botanical garden, TAR1_OP985365 and TAR2_OP985366 from World vegetable center – Arusha, TZnK17_OP985357, TZnK18_OP985358, TZnK19_OP985359, TZnK21_OP985361, and TZnK23_OP985363 from Kizimbani –Zanzibar, as well as TTZF1_OP985392, TTZF2_OP985393, TTZF5_OP985396, TTZF7_OP985398, TTZF8_OP985399, and TTZF9_OP9853400 from Amani Zigi Forest, were found to be closely related to one another within the second subclade.

In the third subclade *Eucalyptus torquata* (NC_022401) showed a closely relationship with allspice isolate with accession numbers TZnK22_OP985362, TZnK24_OP985364 from Kizimbani – Zanzibar and TTKB2_OP985368 from Kizugu botanical garden.

The fourth subclade was consisting only of the following Myrtaceae species *Acca sellowiana* (KX289887), *Syzygium samarangense* (NC_060657), *Eugenia Selloi* (MN095411), *Myrcianthes pungens* (MN095409), *Lophomyrtus bullata* (MW214669), *Lenwebbia prominens* (MW214668), *Lenwebbia lasioclade* (MW214667), and *Syzygium nervosum* (NC_053907).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

4. Discussion

Studying the phylogenetic diversity of plants is crucial to understand its evolutionary history, identifying genetic variations and traits, and developing effective conservation [35]. In this study, the amplification of the allspice cp*rbcL* gene was observed at a length of 560bp. The yield and appearance of the DNA band were attributed to the gel's capacity to sort DNA molecules by size and the primer's efficacy in amplifying the conserved region of the allspice cp*rbcL* gene, as reported previously[36], [37]. Notably, the investigation on amplification and sequence analysis of cp*rbcL* in three distinct plant species yielded results that underscore the importance of appropriate primer design in achieving precise and dependable outcomes in genetic studies[38]. The nucleotides statistics of cp*rbcL* gene was found to have GC contents of 44%, however, there is a difference in GC content between genes within a genome and between genomes of one species and another [39]. It has been reported that GC levels in plants range from 28.81% to 42.14% [40]. Therefore, in this study GC levels were found to be 44% which is a relatively high level and indicate the thermal stability of the cp*rbcL* gene in allspice.

BLASTn search as presented in table 2 shows a high level of similarity between the query and subject sequence, with a range of 96% to 100%. Additionally, the number of identical bases was 539, representing 94.90% of the total bases, with a pairwise identity between sequences of 99.20%. The mean alignment was 553, and the standard deviation was 26.2. The patristic distance ranged from 0.01 to 0.04. These findings suggest that the query and subject sequences are highly similar, indicating a possible evolutionary relationship between them. The high level of similarity also implies that the sequences may be functionally related. These results are consistent with previous studies that have used BLASTn searches to identify sequence similarities between organisms and genes[41], [42]. For example, a study that compared the genetic information of loblolly pine and *Arabidopsis thaliana*, it was found that there was a higher level of apparent homology between their expressed genes from wood-forming tissues [43].

The patristic distance in table 4 shows a range of 0.01 to 0.04 which indicates that the query and subject sequences are relatively closely related, with a low degree of divergence between them. This may suggest a recent common ancestor or a relatively short evolutionary distance between

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

the sequences [44]. This finding is in line with previous studies that have shown that patristic distance is a useful metric for measuring the evolutionary distance between DNA sequences[45].

The phylogenetic tree presented in Figure 4 depicts a common ancestor and cluster of plant species that are closely related to allspice, as evidenced by their cp*rbcL* genes. The tree represents a shared genetic origin for the plant species within a subclade, indicating their evolutionary relatedness and the diversification of their genetic material over time. The identified plant species are derived from various geographic regions, including Australia, China, the United Kingdom, Brazil, and New Zealand, as shown in Tables 2.

The first subclade in the phylogenetic tree (Figure 4) identifies a group of plant species that show less genetic heterogeneity between members of the Myrtaceae family and the allspice isolate from Tanzania. The observation of a close evolutionary relationship between allspice and members of the Myrtaceae family is a significant finding, as it suggests that these plant species share a common ancestry and have likely undergone similar evolutionary processes. This is consistent with previous studies that have investigated the evolutionary relationships between different plant taxa based on molecular information [46]. A study discovered that the plant families Boraginaceae and Convolvulaceae exhibit a close relationship based on genetic data, in accordance with their comparable floral characteristics and habitat preferences [47]. The closeness among allspice and other Myrtaceae family members also implies potential ecological relationships among them. This is supported by studies that have shown that closely related plant species tend to share ecological niches and adaptive traits [48]. A research conducted on oak species in California revealed that leaf characteristics linked with resistance to drought and efficient use of resources are analogous among closely related species [49]. The identification of plant species from different geographical regions and their placement on the phylogenetic tree has significant implications for plant taxonomy, ecology, and conservation [50].

However, the second subclade of the phylogenetic tree in figure 4 indicate a lower level of genetic diversity within the allspice population from Tanzania. This finding is consistent with previous studies that have shown that small, isolated populations are more susceptible to genetic drift and inbreeding, which can reduce genetic diversity [51], [52].

The third subclade in the phylogenetic tree of allspice and related species reveals valuable insights into the evolution and genetic variation of these plant species. Specifically, the genetic diversity

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

observed between allspice isolates from Kizimbani – Zanzibar with accession numbers TZnK22_OP985362 and TZnK24_OP985364, and from Kizugu botanical garden with accession number TTKB2_OP985368. and *Eucalyptus torquata* (NC_022401) suggests that these species have diverged significantly over time. The level of genetic diversity observed in this subclade aligns with findings from prior research that have shown that genetic diversity can be influenced by various factors such as geographic distance, environmental conditions, and reproductive isolation[53]. A study on Eucalyptus species in southeastern Australia revealed high genetic diversity, likely due to geographic isolation and adaptation to diverse environmental conditions[54]. Another study on the genetic diversity of allspice in Jamaica found that the population structure was influenced by the geographic distribution of the species and the type of soil in which it grows [1]. Conservation strategies such as habitat protection and restoration can help preserve the genetic diversity of these plant species and maintain their ecological relationships for future generations [55].

The phylogenetic tree (Fig 4) reveals that there is substantial genetic diversity among the population of Allspice from Tanzania. This finding is supported by the presence of distinct subclades that share a common ancestor but display significant genetic variations. The existence of these subclades provides evidence that Allspice population in Tanzania has undergone genetic differentiation, which may have occurred due to various factors such as geographic isolation, founder effects, or natural selection. In a recent study on a rare plant species known as Silene tatarica, it was discovered that the species exhibited considerable genetic differentiation as a result of both geographic isolation and founder effects [56]. Another study on the genetic diversity of *Pinus koraiensis* populations in China also found evidence of genetic differentiation within populations, which was attributed to natural selection and genetic drift [57].

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

5. Conclusion

The study aims to investigate the evolutionary divergence and relatedness of Allspice using the cp*rbcL* gene, a widely used marker for plant phylogenetic analysis. The obtained sequence was compared to other related species in the Myrtaceae family using multiple sequence alignment. The study found that Allspice had 96% - 100% similarity in its cp*rbcL* gene to members of Myrtaceae family. The evolutionary divergence of Allspice was determined to range from 0.01 as minimum to 0.04 as highest among other species in Myrtaceae. Phylogenetic analysis of the cp*rbcL* gene data revealed genetic diversity within allspice population, and had a strong evolutionary relationship with other species in the Myrtaceae family. The study shows the usefulness of the cp*rbcL* gene for inferring evolutionary relationships among plant species and the importance of the evolutionary approach in estimating the evolutionary divergence of species. The obtained results provide new insights into the evolutionary history of Allspice and its relationship with other species in the Myrtaceae family, which can be used to better understand the biology and ecology of Allspice and aid in conservation efforts.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

6. Acknowledgment

I would like to thank the Nelson Mandela Institution of Science and Technology, Arusha (NM-AIST) Tanzania, for providing institutional support during the study including CREATES for financial support to this study.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

7. References

[1] D. Grattapaglia *et al.*, "Progress in Myrtaceae genetics and genomics: Eucalyptus as the pivotal genus," *Tree Genet. Genomes*, vol. 8, no. 3, pp. 463–508, 2012, doi: 10.1007/s11295-012-0491-x.

[2] M. S. Mérida-Reyes *et al.*, "Composition and Antibacterial Activity of the Essential Oil from Pimenta dioica (L.) Merr. from Guatemala," *Medicines*, vol. 7, no. 10, p. 59, 2020, doi: 10.3390/medicines7100059.

[3] P. L. Merr and C. L. Mill, "Pimenta dioica Scientifi c Name," vol. 3, pp. 655–664, 2012, doi: 10.1007/978-94-007-2534-8.

[4] T. Sunseri, "The political ecology of the copal trade in the Tanzanian coastal Hinterland," *J. Afr. Hist.*, vol. 48, no. 2, pp. 201–220, 2007, doi: 10.1017/S0021853707002733.

[5] L. Zhang and B. L. Lokeshwar, "Medicinal Properties of the Jamaican Pepper Plant Pimenta dioica and Allspice," *Curr. Drug Targets*, vol. 13, no. 14, pp. 1900–1906, 2012, doi: 10.2174/138945012804545641.

[6] A. C. Lorenzo-Leal, E. Palou, A. López-Malo, and H. Bach, "Antimicrobial, Cytotoxic, and Anti-Inflammatory Activities of Pimenta dioica and Rosmarinus officinalis Essential Oils," *Biomed Res. Int.*, vol. 2019, 2019, doi: 10.1155/2019/1639726.

[7] M. George and L. Joseph, "Pharmacognostical and Phytochemical Characterization of Pimento Leaves," vol. 7, no. 1, pp. 75–80, 2013, doi: 10.5829/idosi.gjp.2013.7.1.1102.

[8] M. B. Isman, "Plant essential oils for pest and disease management," *Crop Prot.*, vol. 19, no. 8–10, pp. 603–608, 2000, doi: 10.1016/S0261-2194(00)00079-X.

[9] S. M. Seo, J. Kim, S. G. Lee, C. H. Shin, S. C. Shin, and I. K. Park, "Fumigant antitermitic activity of plant essential oils and components from ajowan (Trachyspermum ammi), allspice (Pimenta dioica), caraway (Carům carvi), dill (Anethum graveoiens), geranium (Pelargonium graveoiens), and litsea (Litsea cubeba) oils against," *J. Agric. Food Chem.*, vol. 57, no. 15, pp. 6596–6602, 2009, doi: 10.1021/jf9015416.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

[10] H. Hitokoto, S. Morozumi, T. Wauke, S. Sakai, and H. Kurata, "Inhibitory effects of spices on growth and toxin production of toxigenic fungi," *Appl. Environ. Microbiol.*, vol. 39, no. 4, pp. 818–822, 1980, doi: 10.1128/aem.39.4.818-822.1980.

[11] M. Friedman, P. R. Henika, and R. E. Mandrell, "Bactericidal activities of plant essential oils and some of their isolated constituents against Campylobacter jejuni, Escherichia coli, Listeria monocytogenes, and Salmonella enterica," *J. Food Prot.*, vol. 65, no. 10, pp. 1545–1560, 2002, doi: 10.4315/0362-028X-65.10.1545.

[12] J. Kim, S. M. Seo, S. G. Lee, S. C. Shin, and I. K. Park, "Nematicidal activity of plant essential oils and components from coriander (Coriandrum sativum), oriental sweetgum (Liquidambar orientalis), and valerian (Valeriana wallichii) essential oils against pine wood nematode (Bursaphelenchus xylophilus)," *J. Agric. Food Chem.*, vol. 56, no. 16, pp. 7316–7320, 2008, doi: 10.1021/jf800780f.

[13] Y. Yareni Andrade Avila, J. Cruz-Olivares, and C. Pérez-Alonso, "Antioxidant Effect and Medicinal Properties of Allspice Essential Oil," 2022, doi: 10.5772/intechopen.103001.

[14] Nurhasanah, Sundari, and N. Papuangan, "Amplification and Analysis of Rbcl Gene (Ribulose-1,5-Bisphosphate Carboxylase) of Clove in Ternate Island," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 276, no. 1, 2019, doi: 10.1088/1755-1315/276/1/012061.

[15] W. Zhao *et al.*, "Evaluation of environmental factors affecting the genetic diversity, genetic structure, and the potential distribution of Rhododendron aureum Georgi under changing climate," *Ecol. Evol.*, vol. 11, no. 18, pp. 12294–12306, 2021, doi: 10.1002/ece3.7803.

[16] W. Thuiller *et al.*, "Predicting global change impacts on plant species' distributions: Future challenges," *Perspect. Plant Ecol. Evol. Syst.*, vol. 9, no. 3–4, pp. 137–152, 2008, doi: 10.1016/j.ppees.2007.09.004.

[17] K. Guan, S. P. Good, K. K. Caylor, H. Sato, E. F. Wood, and H. Li, "Continental-scale impacts of intra-seasonal rainfall variability on simulated ecosystem responses in Africa," *Biogeosciences*, vol. 11, no. 23, pp. 6939–6954, 2014, doi: 10.5194/bg-11-6939-2014.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

[18] S. A. Shameem and I. N. Kangroo, “Comparative assessment of edaphic features and phytodiversity in lower Dachigam National Park , Kashmir Himalaya , India,” *Sci. Technol.*, vol. 5, no. November, pp. 972–984, 2011, doi: 10.5897/AJEST11.099.

[19] N. H. Barton, “Mutation and the evolution of recombination,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 365, no. 1544, pp. 1281–1294, 2010, doi: 10.1098/rstb.2009.0320.

[20] R. Mcvaugh, “THE GENERA OF AMERICA:;\" MYRTACEAE - A.\" IVfERIM REPORT Rogers McVaugh ",” vol. 17, 1828, doi: 10.2307/1217393.

[21] C. Y. Zhang, F. Y. Wang, H. F. Yan, G. Hao, C. M. Hu, and X. J. Ge, “Testing DNA barcoding in closely related groups of Lysimachia L. (Myrsinaceae),” *Mol. Ecol. Resour.*, vol. 12, no. 1, pp. 98–108, 2012, doi: 10.1111/j.1755-0998.2011.03076.x.

[22] P. D. N. Hebert, A. Cywinska, S. L. Ball, and J. R. DeWaard, “Biological identifications through DNA barcodes,” *Proc. R. Soc. B Biol. Sci.*, vol. 270, no. 1512, pp. 313–321, 2003, doi: 10.1098/rspb.2002.2218.

[23] Y. Chen *et al.*, “Identification of rubisco rbcL and rbcS in Camellia oleifera and their potential as molecular markers for selection of high tea oil cultivars,” *Front. Plant Sci.*, vol. 6, no. MAR, pp. 1–11, 2015, doi: 10.3389/fpls.2015.00189.

[24] Y. Wang, J. Yu, Y.-K. Chen, and Z.-C. Wang, “Complete Chloroplast Genome Sequence of the Endemic and Endangered Plant Dendropanax oligodontus: Genome Structure, Comparative and Phylogenetic Analysis,” *Genes (Basel).*, vol. 13, no. 11, p. 2028, 2022, doi: 10.3390/genes13112028.

[25] W. Dong, C. Xu, J. Wen, and S. Zhou, “Evolutionary directions of single nucleotide substitutions and structural mutations in the chloroplast genomes of the family Calycanthaceae,” *BMC Evol. Biol.*, vol. 20, no. 1, pp. 1–12, 2020, doi: 10.1186/s12862-020-01661-0.

[26] R. G. Olmstead and J. D. Palmer, “Chloroplast DNA systematics: a review of methods and data analysis,” *Am. J. Bot.*, vol. 81, no. 9, pp. 1205–1224, 1994, doi: 10.1002/j.1537-2197.1994.tb15615.x.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

[27] M. T. Clegg, “Chloroplast gene sequences and the study of plant evolution,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 90, no. 2, pp. 363–367, 1993, doi: 10.1073/pnas.90.2.363.

[28] J. Cros *et al.*, “Phylogenetic Analysis of Chloroplast DNA Variation inCoffeaL,” *Mol. Phylogenet. Evol.*, vol. 9, no. 1, pp. 109–117, 1998, doi: 10.1006/mpev.1997.0453.

[29] M. Hasebe, T. Omori, M. Nakazawa, T. Sano, M. Kato, and K. Iwatsuki, “rbcL gene sequences provide evidence for the evolutionary lineages of leptosporangiate ferns,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 91, no. 12, pp. 5730–5734, 1994, doi: 10.1073/pnas.91.12.5730.

[30] M. G. Murray and W. F. Thompson, “Rapid isolation of high molecular weight plant DNA,” *Nucleic Acids Res.*, vol. 8, no. 19, pp. 4321–4326, 1980, doi: 10.1093/nar/8.19.4321.

[31] S. Y. Vanson Liu, T. P. Kumara, and C. H. Hsu, “Genetic identification and hybridization in the seagrass genus Halophila (Hydrocharitaceae) in Sri Lankan waters,” *PeerJ*, vol. 8, no. 2011, pp. 1–16, 2020, doi: 10.7717/peerj.10027.

[32] T. C. Lorenz, “Polymerase chain reaction: Basic protocol plus troubleshooting and optimization strategies,” *J. Vis. Exp.*, no. 63, pp. 1–15, 2012, doi: 10.3791/3998.

[33] R. C. Edgar, “MUSCLE: Multiple sequence alignment with high accuracy and high throughput,” *Nucleic Acids Res.*, vol. 32, no. 5, pp. 1792–1797, 2004, doi: 10.1093/nar/gkh340.

[34] J. Felsenstein, “Confidence Limits on Phylogenies: An Approach Using the Bootstrap,” *Evolution (N. Y).*, vol. 39, no. 4, p. 783, 1985, doi: 10.2307/2408678.

[35] M. R. Gostel and W. J. Kress, “The Expanding Role of DNA Barcodes: Indispensable Tools for Ecology, Evolution, and Conservation,” *Diversity*, vol. 14, no. 3, pp. 1–23, 2022, doi: 10.3390/d14030213.

[36] M. Kiran Kumar and B. V. Sandeep, “Phylogenetic study of mangrove associate grass Myriostachya wightiana (Nees ex Steud.) Hook. F. using rbcL gene sequence,” *Plant Sci. Today*, vol. 8, no. 3, pp. 590–595, 2021, doi: 10.14719/PST.2021.8.3.1133.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

[37] J. A. Raven and J. F. Allen, “Genomics and chloroplast evolution: What did cyanobacteria do for plants?,” *Genome Biol.*, vol. 4, no. 3, pp. 1–5, 2003, doi: 10.1186/gb-2003-4-3-209.

[38] C. H. W. M. R. Bhagya Chandrasekara, D. N. U. Naranpanawa, B. S. Bandusekara, D. K. N. G. Pushpakumara, D. S. A. Wijesundera, and P. C. G. Bandaranayake, “Universal barcoding regions, rbcL, matK and trnH-psbA do not discriminate Cinnamomum species in Sri Lanka,” *PLoS One*, vol. 16, no. 2 February, pp. 1–16, 2021, doi: 10.1371/journal.pone.0245592.

[39] J. F. Wendel, J. Greilhuber, J. Doležel, and I. J. Leitch, “Plant genome diversity volume 1: Plant genomes, their residents, and their evolutionary dynamics,” *Plant Genome Divers. Vol. 1 Plant Genomes, their Resid. their Evol. Dyn.*, pp. 1–279, 2012, doi: 10.1007/978-3-7091-1130-7.

[40] J. Kusumi and H. Tachida, “Compositional properties of green-plant plastid genomes,” *J. Mol. Evol.*, vol. 60, no. 4, pp. 417–425, 2005, doi: 10.1007/s00239-004-0086-8.

[41] Z. Zhang, S. Schwartz, L. Wagner, and W. Miller, “A greedy algorithm for aligning DNA sequences,” *J. Comput. Biol.*, vol. 7, no. 1–2, pp. 203–214, 2000, doi: 10.1089/10665270050081478.

[42] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, “Basic local alignment search tool,” *J. Mol. Biol.*, vol. 215, no. 3, pp. 403–410, 1990, doi: 10.1016/S0022-2836(05)80360-2.

[43] M. Kirst *et al.*, “Apparent homology of expressed genes from wood-forming tissues of loblolly pine (Pinus taeda L.) with Arabidopsis thaliana,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 12, pp. 7383–7388, 2003, doi: 10.1073/pnas.1132171100.

[44] J. H. Burns and S. Y. Strauss, “More closely related species are more ecologically similar in an experimental test,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 13, pp. 5302–5307, 2011, doi: 10.1073/pnas.1013003108.

[45] M. Fourment and M. J. Gibbs, “PATRISTIC: A program for calculating patristic distances and graphically comparing the components of genetic change,” *BMC Evol. Biol.*, vol. 6,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

pp. 1–5, 2006, doi: 10.1186/1471-2148-6-1.

[46] B. Bremer *et al.*, “An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III,” *Bot. J. Linn. Soc.*, vol. 161, no. 2, pp. 105–121, 2009, doi: 10.1111/j.1095-8339.2009.00996.x.

[47] E. Långström and M. W. Chase, “Tribes of Boraginoideae (Boraginaceae) and placement of Antiphytum, Echiochilon, Ogastemma and Sericostoma: A phylogenetic analysis based on atpB plastid DNA sequence data,” *Plant Syst. Evol.*, vol. 234, no. 1–4, pp. 137–153, 2002, doi: 10.1007/s00606-002-0195-z.

[48] J. J. Wiens *et al.*, “Niche conservatism as an emerging principle in ecology and conservation biology,” *Ecol. Lett.*, vol. 13, no. 10, pp. 1310–1324, 2010, doi: 10.1111/j.1461-0248.2010.01515.x.

[49] J. Cavender-Bares, S. Kothari, J. E. Meireles, M. A. Kaproth, P. S. Manos, and A. L. Hipp, “The role of diversification in community assembly of the oaks (Quercus L.) across the continental U.S,” *Am. J. Bot.*, vol. 105, no. 3, pp. 565–586, 2018, doi: 10.1002/ajb2.1049.

[50] D. A. N. Rosauer, S. W. Laffan, M. D. Crisp, and S. C. Donnellan, “Phylogenetic endemism : a new approach for identifying geographical concentrations of evolutionary history,” pp. 4061–4072, 2009, doi: 10.1111/j.1365-294X.2009.04311.x.

[51] C. T. Cole, “Genetic Variation in Rare and Common Plants,” *Annu. Rev. Ecol. Evol. Syst.*, vol. 34, pp. 213–237, 2003, doi: 10.1146/annurev.ecolsys.34.030102.151717.

[52] E. M. Leffler *et al.*, “Revisiting an Old Riddle: What Determines Genetic Diversity Levels within Species?,” *PLoS Biol.*, vol. 10, no. 9, 2012, doi: 10.1371/journal.pbio.1001388.

[53] A. Hampe, “Some Evolutionary Consequences of Being a Tree,” 2006, doi: 10.1146/annurev.ecolsys.37.091305.110215.

[54] C. G. M. Ekomono, C. B. S. V. Loubassou, M. P. Mbama, G. J. L. Panzou, and P. Vigneron, “Adaptability and interspecific variability in growth and leaf traits of eucalypt,” *IForest*, vol. 14, pp. 560–568, 2021, doi: 10.3832/ifor3660-014.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

[55] H. Mouhib, D. Jelisavac, W. Stahl, R. Wang, I. Kalf, and U. Englert, “The conformation of odorants in different states of aggregation: A joint venture in microwave spectroscopy and X-ray diffraction,” *ChemPhysChem*, vol. 12, no. 4, pp. 761–764, 2011, doi: 10.1002/cphc.201000986.

[56] N. Tero, J. Aspi, P. Siikamäki, A. Jäkäläniemi, and J. Tuomi, “Genetic structure and gene flow in a metapopulation of an endangered plant species, Silene tatarica,” *Mol. Ecol.*, vol. 12, no. 8, pp. 2073–2085, 2003, doi: 10.1046/j.1365-294X.2003.01898.x.

[57] X. Li, M. Zhao, Y. Xu, Y. Li, M. Tigabu, and X. Zhao, “Genetic diversity and population differentiation of pinus koraiensis in China,” *Horticulturae*, vol. 7, no. 5, pp. 1–18, 2021, doi: 10.3390/horticulturae7050104.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

Figures

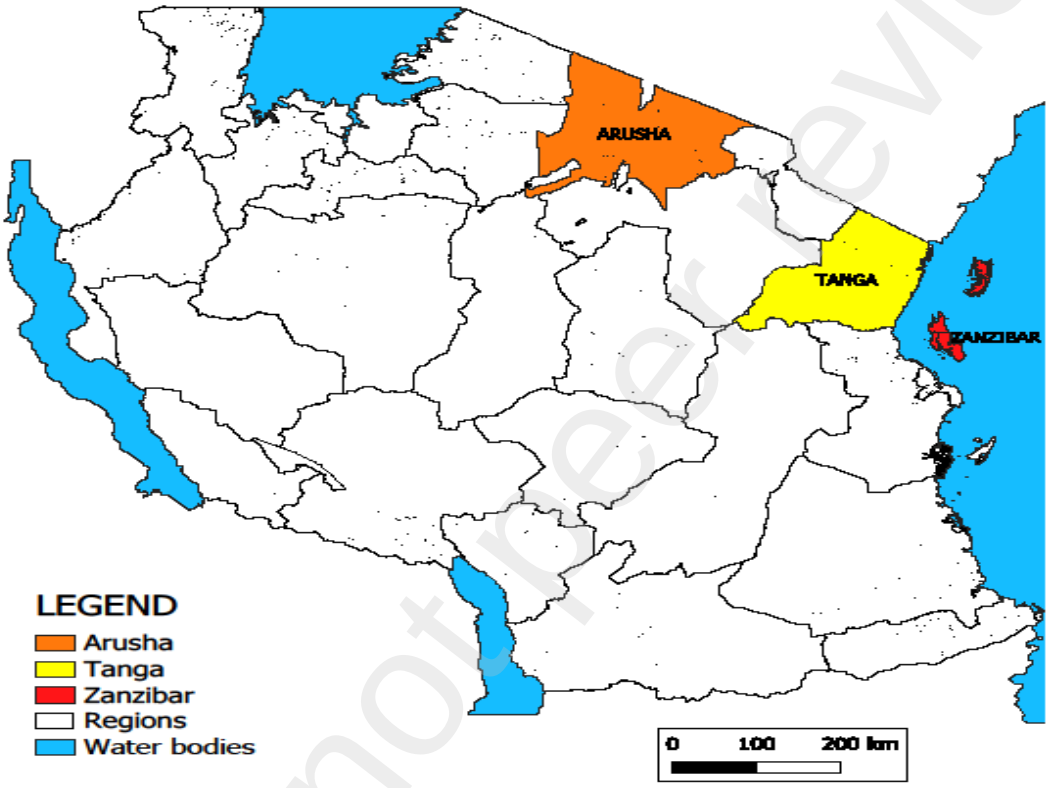

Figure 1. Map of Tanzania showing the locations where allspice samples were collected, particularly in Arusha, Tanga, and Zanzibar

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

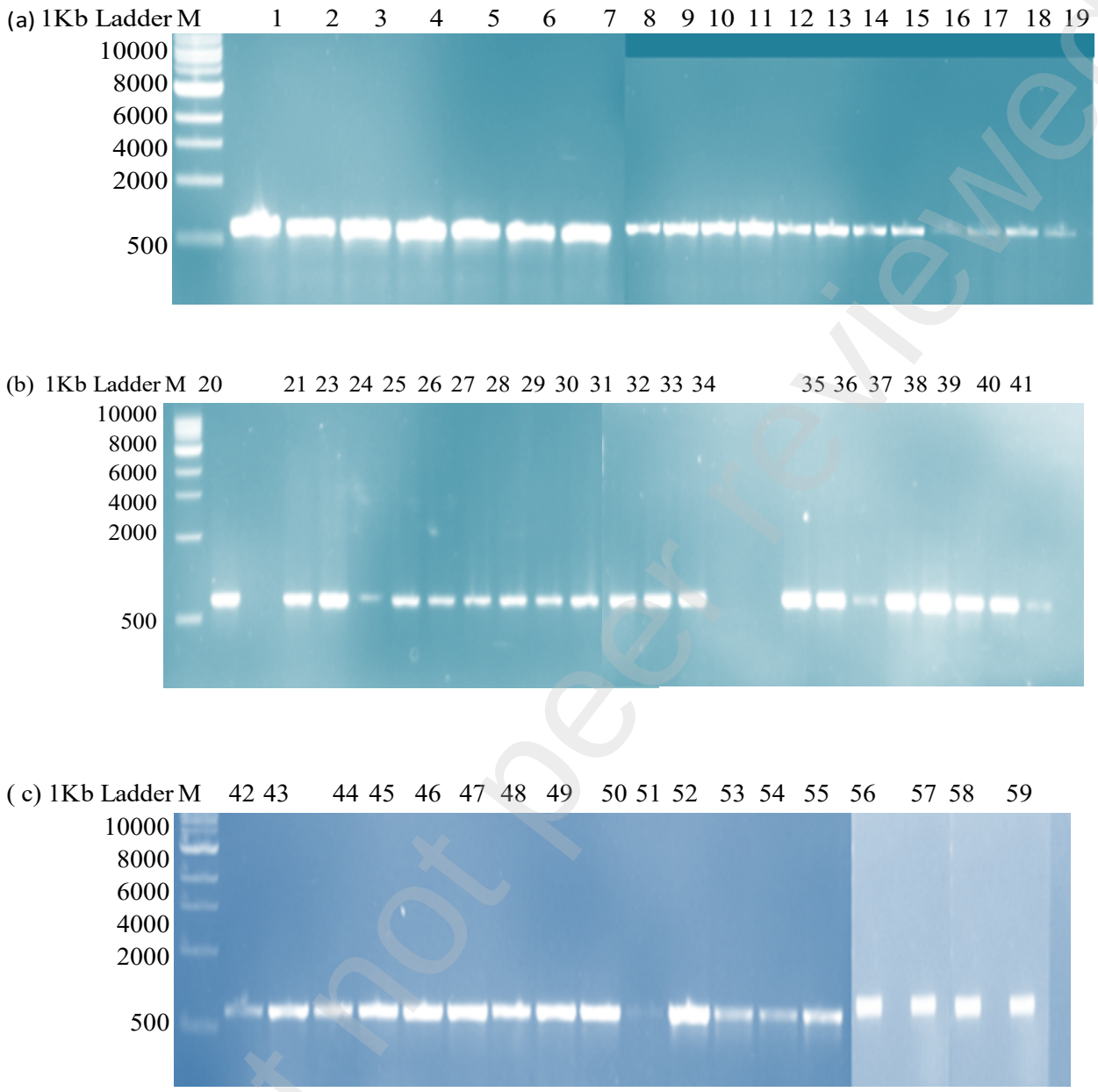

Figure 2. PCR amplified products for the cprbcL gene by Allspice collections a (1-19), b (20-41) c (42-59).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

Figure 3. Allspice tree in Kizugu botanical garden Tanga

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340

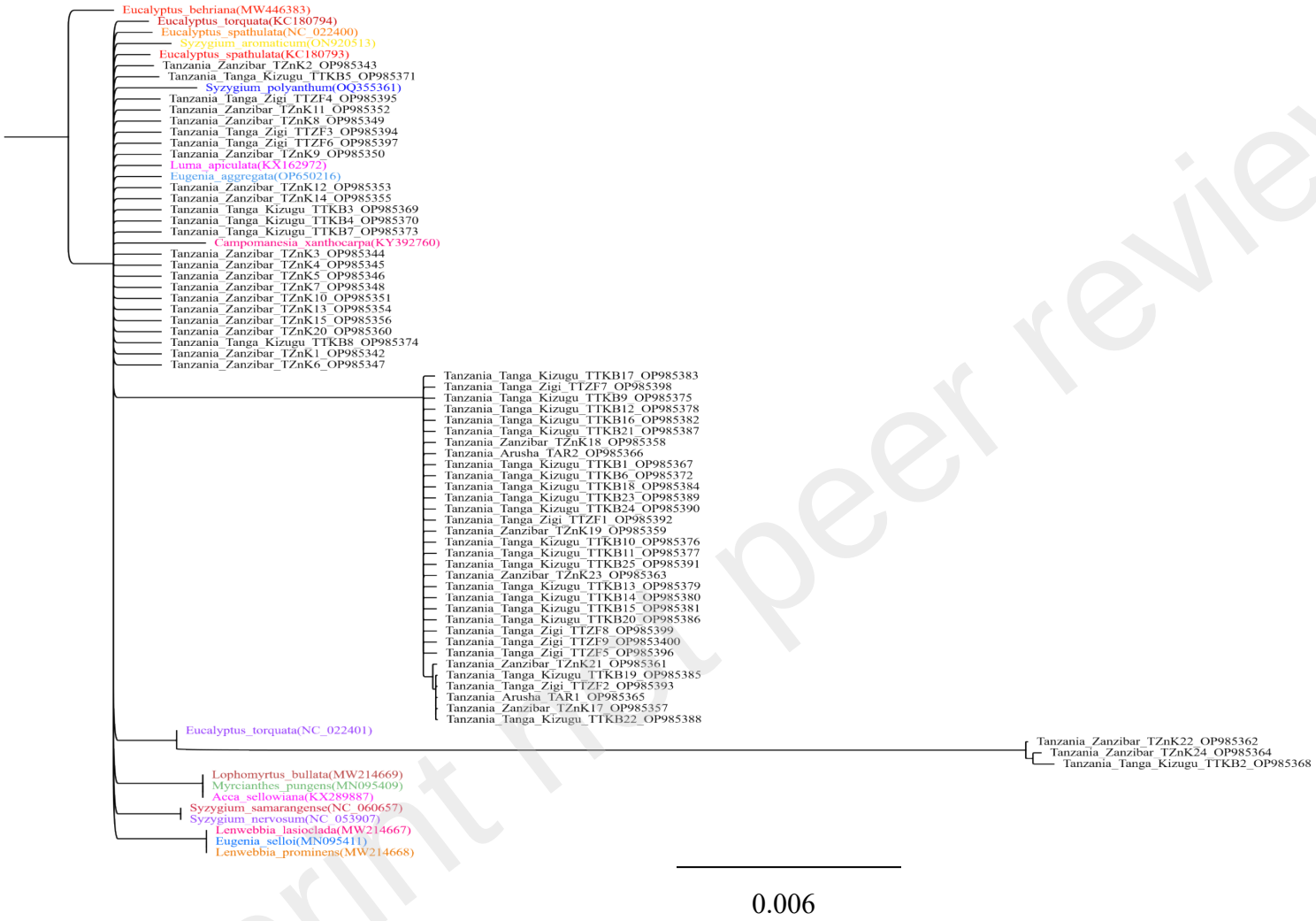

Figure 4. Phylogenetic relationship of cprbcL gene of Allspice from Tanzania. The inference tree was constructed through Geneious Prime 2023.0.1 using Tamura-Nei, maximum likelihood as a statistical method and neighbor-joining method. Eucalyptus behriana (MW446383) was used as an outgroup to position the root of the tree.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4480340