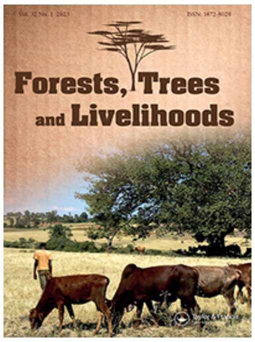

Taylor & Francis
Taylor & Francis Group

Forests, Trees and Livelihoods

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tftl20

Genetic diversity and population structure of *Irvingia species* using DArTseq generated markers

Sylvia Jepkemboi & Patroba Ojola

To cite this article: Sylvia Jepkemboi & Patroba Ojola (2023) Genetic diversity and population structure of *Irvingia species* using DArTseq generated markers, Forests, Trees and Livelihoods, 32:1, 42-54, DOI: 10.1080/14728028.2022.2160380

To link to this article: https://doi.org/10.1080/14728028.2022.2160380

FORESTS, TREES AND LIVELIHOODS
2023, VOL. 32, NO. 1, 42–54
https://doi.org/10.1080/14728028.2022.2160380

Taylor & Francis
Taylor & Francis Group

Genetic diversity and population structure of *Irvingia species* using DArTseq generated markers

Sylvia Jepkemboi and Patroba Ojola

Department of Biochemistry, Microbiology and Biotechnology, Kenyatta University, Nairobi, Kenya

ABSTRACT

Irvingia gabonensis and *Irvingia wombolu* trees indigenous to West and Central Africa are economically important owing to their valuable kernels. Massive fruit harvesting and reduction of forests land area has threatened their genetic diversity thus warranting conservation. This study reports the genetic diversity and population structure based on dominant markers of Irvingia accessions from Nigeria, Cameroon and Gabon held at World Agroforestry field genebanks. A total of 60 samples and 6532 SilicoDArT markers revealed high genetic differentiation among populations (phiPT = 0.301, P = .001). The overall genetic diversity revealed by Nei's gene diversity (He = 0.117) and percentage of polymorphic loci (39.60%) was low. *I.gabonensis* provenances from Cameroon and Nigeria were more diverse than Gabonese populations. Pairwise phiPT values calculated from AMOVA variances delineated *I. wombolu* and *I. gabonensis*. Population structure and cluster analysis based on Bayesian algorithm, DAPC and UPGMA methods placed Cameroon at the center of distribution of *I. gabonensis* accessions from Nigeria and Gabon. Gabonese accessions tend to cluster independent of Nigerian accessions while *I. wombolu* formed a distinct cluster from *I. gabonensis*. SilicoDArT markers were found suitable for population structure and differentiation analysis of Irvingia accessions and can guide future recollections to enrich genebank diversity and materials for domestication events.

KEYWORDS

Genetic diversity; *Irvingia wombolu*; *irvingia gabonensis*; population structure; germplasm conservation; SilicoDArT; bush mango

Introduction

Irvingia is a genus with seven species of trees indigenous to South East Asia (1 species), West and Central Africa (6 species). *Irvingia gabonensis* (sweet bush mango) and *Irvingia wombolu* (bitter bush mango) are the only species in the genus *Irvingia* with economic importance in Africa (Asaah et al. 2003). The distribution of these two closely related species across West and Central Africa overlaps in Nigeria and Cameroon where they co-occur unlike all the other countries in the region (Lowe et al. 2000). The trees are commonly referred by names such as bush mango, wild mango, dika nut or obono by the respective local communities. Studies of the reproductive biology of *I. gabonensis* and *I. wombolu* indicate that both species are mainly outcrossing with their pollination facilitated by insects and the seeds dispersed by large animals and humans with no evidence of hybridization in areas of sympatry (Harris 1996; Lowe et al. 2000). Reproduction influences the evolution

---

CONTACT Sylvia Jepkemboi chumbasylvia@gmail.com

Supplemental data for this article can be accessed online at https://doi.org/10.1080/14728028.2022.2160380

and ecology of plants, shaping future genomes and environmental adaptability (Schoen et al. 2019).

Fruit pericarps of *I. gabonensis* are sweet and are consumed fresh while their kernels and those of *I. wombolu* are valued and used as soup thickeners, bread and dietary supplements in weight loss regimens, medicine and as potential adjuncts in drug delivery systems (Odeku and Patani 2005; Ngondi et al. 2009; Odeku et al. 2013). The kernels are also commercialized regionally and internationally for cash exchange (e.g. Ofundem et al. 2017).

African bush mango is among the many trees with essential non-timber forest products in West and Central Africa. The high demand of their products, kernels especially, has led to massive fruit collection from forests and traditional agroforestry systems (Onyekwelu et al. 2015; Ofundem et al. 2017). This growing demand and over-harvesting of *Irvingia* fruits (seeds) from their geographical locations has waned regeneration in their habitats, resulting in a loss of genetic diversity (Vihotogbé et al. 2013). Conservation of available genetic diversity is imperative for bush mango, which is among the top prioritized trees for domestication in Africa. While tropical African forests harbor rich tree diversity, they face challenges from human activities that include unsustainable harvesting, agricultural activities, habitat loss, degradation and pollution (Davies et al. 2011; Daru et al. 2016) that endanger their diversity. These challenges can be addressed using genomics which is vital in understanding how trees cope with the threats (Plomion et al. 2016) and to uncover the potential economic benefits from forest resources. The potential benefits that can be harnessed from genomic information are tremendous (Slik et al. 2015). For instance, having reference genomes of major African tree species will help increase their utilization range by steering research in food production, environmental resilience and nutritional improvement.

Irvingia gabonensis is among the 101 species of orphan crops identified by the African Orphan Crops Consortium (AOCC). The AOCC initiative aims to address nutritional issues and hunger problems by strengthening agricultural food production, through an integrative approach involving the use of genomic tools, on under-researched African food producing woody and non-woody species. The selected species are targeted for nutritional improvement, better yields and climate resilience (Hendre et al. 2019). World Agroforestry (ICRAF) has put efforts in germplasm collection of *Irvingia* and successfully established field genebanks for them. The genebanks located at Onne and Ibadan in Nigeria and Mbalmayo, Cameroon hold germplasm conserved *ex situ* as living tree accessions. Like most woody perennials, *Irvingia* spp are not conserved in seed genebanks due to the recalcitrant nature of their seeds that only remain viable approximately four weeks after falling (Vihotogbé et al. 2014; Migicovsky et al. 2019). The genebanks are of great importance not only for conservation but also to the AOCC project since the genus *Irvingia* has one of the valuable selected African orphan crop. While genebanks are imperative to conservation, having a systematic design is important and can greatly benefit from genetic and phenotypic characterization of available germplasm for better management and future collections. Previous studies employed morphological and molecular markers on these accessions primarily for characterization, which is key for effective conservation, utilization and improvement of tree germplasm through breeding programs. These earlier studies in *Irvingia* used random amplified polymorphic DNA (RAPD) (Lowe et al. 2000), amplified fragment length polymorphism (AFLP) (Ude et al. 2006) and a combination of AFLP with chloroplast simple sequence repeats (cpSSRs) (Vihotogbé et al. 2013).These studies predicted the centers of genetic diversity and their origin. However, these markers have their limitations such as poor genome coverage, low reproducibility, low-throughput, and low marker density. In this regard, this study employed DArTseq based dominant markers that have a wide genome coverage, sequence based, generates thousands of markers in a high throughput and cost-effective method. DArTseq, developed by Diversity Arrays Technology Pty Ltd (DArT, Canberra, ACT, Australia) discovers markers through complexity reduction and next generation sequencing generating two marker types; SilicoDArT and SNPs (Killian et al. 2016). The two marker types have been applied successfully in analysis of genetic diversity and mapping of tree populations in macadamia (*Macadamia integrifolia* and *Macadamia tetraphylla)* and *Eucalyptus microcarpa* (Jordan et al. 2016; O'Connor et al. 2019).

This study presents the first application of DArTseq markers in molecular characterization of *Irvingia* accessions held at the ICRAF field genebanks. The study utilized SilicoDArT markers to describe the level of genetic diversity and structuring of genebank accessions of *Irvingia*. Half siblings were included in the study to check the suitability of DArTseq markers in characterizing *Irvingia*.

Materials and methods

Plant material and genomic DNA extraction

Fresh leaf samples were collected from field tree genebanks located at Onne, Nigeria and Mbalmayo, Cameroon. A total of 60 individuals representing two Irvingia species and three countries were sampled, including Nigeria (5 regions), Cameroon (2 regions) and Gabon (3 regions). These leaf samples represent 50 individuals from 7 populations of *Irvingia gabonensis* and 10 *Irvingia wombolu* individuals from 3 populations. The sampled trees are of fruits collected from naturally occurring stands in their respective countries, which seeds were planted in the genebanks between 1994 and 1995. Seeds from up to six fruits per tree (half siblings) were used for planting during genebank establishment. This study used up to two half siblings per tree accession from the mature trees. The leaf samples were packed in silica gel and shipped to the ICRAF, AOCC molecular and genomics laboratory from the respective field genebanks for DNA isolation.

Genomic DNA was extracted from the leaves using cetyl trimethylammoniumbromide (CTAB) based method. DNA was quantified using Qubit 2.0 fluorometer (Invitrogen Co., USA) and quality evaluated on 1% agarose gel. DNA concentration was adjusted to DArTseq's required concentration (50ng/µl). 50 µl of each DNA sample was sent to DArT Pty Ltd Canberra, Australia for sequencing and marker discovery (http://www.diversityarrays.com/).

Genotyping and marker quality

DArTseq is a proprietary genotyping by sequencing (GBS) technology that relies on complexity reduction to generate high density markers from low copy regions of the genome (Killian et al. 2012). The method was optimized for *Irvingia* samples following the protocol described by Alam et al. (2018) and Raman et al. (2014). Briefly, the initial DArTseq steps involved creation of representations from the samples for DNA library preparation. The representations were selected to capture allelic diversity of *Irvingia* samples followed by processing through a complexity reduction method that uses a combination of a rare cutting enzyme (*PstI*) and proprietary frequent cutting restriction enzymes. Sequencing was done using one of Illumina's next-generation sequencing platforms (Killian et al. 2016). The sequences were extracted and analyzed using DArTsoft v7.4.7 which aligns the sequences and provides specific SilicoDArT markers scored as present or absent (1/0) from sequence tags.

The dominant silicoDArT markers were scored in a binary fashion with each allele called as 1, 0 and '–'representing present, absent or missing alleles respectively. DArTsoft has inbuilt quality control parameters such as reproducibility, polymorphism information content (PIC) and call rate (proportion of samples that the genotype call is either 1 or 0, rather than ('-') that indicate the quality for each marker. DArTseq discovered a total of 16,322 SilicoDArT markers but only markers with call rates ≥ 95% per marker and PIC ≥ 0.05 were retained for further data analysis of genetic diversity. 6532 markers met this criteria and were therefore used for downstream analysis.

Data analysis

Analysis of molecular variance (AMOVA) was executed on GenAlEx v6.503 software to check the structuring of variation for all populations using the binary (diploid option). The variance components from AMOVA were used to calculate Phi-statistics (PhiPT) applying 999 permutations for both global and pairwise values. PhiPT is a measure of population differentiation analogous to Fst from Wright's F statistics (Peakall and Smouse 2012). GenAlEx was also used to calculate several measures of diversity; Nei's genetic distance, number of different alleles (Na) and percentage of polymorphic loci (%P) based on the 6532 markers across all populations. Nei's gene diversity (expected heterozygosity, He) was calculated on Arlequin v3.5.2.2 (Excoffier et al. 2005) with 100 steps of 10,000 Markov chain iterations assuming Hardy-Weinberg Equilibrium proportions.

To establish and describe the genetic population structure and clustering of the 60 individuals, we employed admixture ancestry model on STRUCTURE v2.3.4 (Pritchard et al. 2000) that uses Bayesian algorithm to define clusters. Ten independent runs were performed for ten K values (K = 1 to K = 10) applying an admixture ancestry model with no assumptions of prior population information. Burn in period of 50,000 and 100,000 Markov Chain Monte Carlo (MCMC) iterations were used. The results were uploaded to Structure Harvester an online program that establishes the most correct K value using Evanno method, the number of clusters (K) was plotted against delta K values (Evanno et al. 2005). Ideally, the best K value is the one with the highest delta K value (peak) which was at K = 2, that clearly distinguished the two species (T3 supplementary table). A second STRUCTURE analysis was performed only for the 50 *I. gabonensis* individuals to inspect within species population structuring. 8 populations were assumed for this second run, employing a burn in period of 10,000 and 100,000 MCMC repetitions with 3 iterations for each K value. The best K value as shown by Evanno method was at K = 2. The final run for the 50 individuals was executed with a higher length of burn in period, 150,000 and MCMC repetitions of 150,000 assuming 3 populations, with 25 iterations for each K value (K = 1 to K = 3). The results were graphically displayed as colored bar charts using membership proportions (Q values) for each individual in the inferred populations. To crosscheck genetic population structure, a scatterplot of DAPC; Discriminant Analysis of Principal Components (Jombart and Collins 2017) was plotted on R (package: Adegenet). DAPC uses Bayesian Information Criterion (BIC) to identify clusters from k-means (increasing k values) to determine the nature of clusters. The best number of clusters in a BIC plot is chosen from the lowest BIC value, which corresponded to 3 clusters for the 60 Irvingia samples (Figure 3a).

To further infer evolutionary relationships, a hierarchical approach using Jaccard dissimilarity index was employed. Here a UPGMA (unweighted pair-group method with arithmetic averaging) phylogenetic tree was constructed using 1000 bootstrap samples on DARwin v6.0.21 software.

Results

Marker quality

DArTseq generated 16,322 SilicoDArT markers, which were filtered down to 6,532 markers. All the remaining 6532 markers had less than 5% missing calls, 0.99 mean reproducibility, and 0.28 Polymorphic Information Content (PIC) (Table 1).

Genetic diversity and differentiation

Structuring of genetic variation by AMOVA indicated that much of the variation lies within populations (70%) while among populations variation was attributed 30% of the total variation (Table 2). The mean PhiPT value across all populations was PhiPT = 0.301, (*P* = .001) and it showed that populations were significantly differentiated. Pairwise PhiPT values further pointed to a higher level of genetic differentiation among *I. wombolu* and *I. gabonensis* provenances. The highest PhiPT was observed between two provenance pairs of the two species; Kogi and Mabanda, Oyo_wo and Mabanda (PhiPT = 0.589, P = .003). Lomie, Cameroon and Oyo, Nigeria provenances were the least differentiated (PhiPT = 0.043, P = .003) (Table 3).

Measures of genetic variation of provenances are presented in Table 4. Mvilla, Oyo, Rivers State and Lomie had higher values of %P (52% – 61%) and Na (1.056 –1.233). Nei's genetic diversity (He) was higher in Oyo, Mvilla and Lomie (He = 0.238, 0.231 and 0.224 respectively). The lowest %P, He and Ne were observed in an *I. wombolu* provenance, Cross River. Nei's genetic distance revealed the farthest genetic distance between Cross River and Lambarene (0.374) whereas the smallest distance was observed between Lomie and Oyo (0.033) (T1 supplementary table). Mvilla, Cameroon had the highest diversity measures across all provenances while Cross River had the least. *I. gabonensis* individuals from Cameroon and Nigeria were more variable than the Gabonese provenances and the 3 provenances of *I. wombolu* from Nigeria.

Population structure and evolutionary relationships

STRUCTURE showed the highest delta K peak at K = 2 (Figure 1) for the 60 individuals inferring two ancestry clusters. Cluster 1 contained all the *I. wombolu* accessions while the *I. gabonensis* were in cluster 2 with no admixture. Using a likelihood threshold of 0.60 in K

Table 1. Marker information for the 6,532 SilicoDArTs.

Metric	Max	Min	Mean
Call rate	1	0.95	0.977
Polymorphic information Content (PIC)	0.5	0.05	0.28
Reproducibility	1	0.95	0.99

Table 2. Summary of AMOVA Results.

Source	df	% variation	Stat	Value	P-value
Among pops	10	30%			
Within pops	49	70%			
Total	59	100%	PhiPT	0.301	0.001

PhiPT = Measure of genetic differentiation among populations

Table 3. Pairwise population PhiPT.

Cross River	Enugu	Kogi	Lambarene	Lomie	Mabanda	Mvilla	Oyo	Oyo_wo	Rivers State	Woleu Ntem	
	0.072	0.073	0.062	0.043	0.001	0.021	0.031	0.068	0.001	0.055	Cross River
0.428		0.032	0.024	0.014	0.008	0.002	0.009	0.039	0.001	0.005	Enugu
0.154	0.427		0.031	0.010	0.003	0.003	0.009	0.035	0.003	0.001	Kogi
0.539	0.309	0.530		0.006	0.003	0.005	0.002	0.028	0.001	0.036	Lambarene
0.431	0.113	0.440	0.198		0.005	0.005	0.050	0.011	0.001	0.045	Lomie
0.623	0.375	0.589	0.240	0.247		0.001	0.002	0.004	0.001	0.003	Mabanda
0.458	0.211	0.463	0.179	0.058	0.231		0.001	0.002	0.001	0.003	Mvilla
0.355	0.087	0.367	0.233	0.043	0.279	0.127		0.001	0.001	0.005	Oyo
0.249	0.419	0.137	0.527	0.437	0.589	0.461	0.359		0.001	0.001	Oyo_wo
0.450	0.116	0.440	0.347	0.168	0.373	0.249	0.156	0.441		0.001	RiversState
0.488	0.223	0.487	0.130	0.056	0.186	0.044	0.135	0.486	0.263		WoleuNtem

PhiPT values below diagonal. P- values above diagonal

= 2, the inferred clusters were generally grouped according to species, the green color represents *I. gabonensis* while red represents *I. wombolu* (Figure 1).

The second STRUCTURE run structured the 50 individuals of *I. gabonensis* species into two populations (K = 2). Cluster 1 (red) contained all the Nigerian populations while cluster 2 (green) was made up of populations from Cameroon and Gabon (Figure 2). This implied a closer relationship between *I. gabonensis* trees from Cameroon and Gabon than with those from Nigeria. K = 3 (Figure 2) shows that Nigerian and Gabonese *I. gabonensis* rarely cluster together and that Cameroon overlaps with both Nigeria and Gabon (T4 supplementary table). The inferred populations by DAPC (Figure 3) agreed with STRUCTURE results, both individuals and provenances at BIC value k = 3 (Figure 3a) were structured as in K = 2 of the 60 individuals. Memberships of each population and individual to the inferred populations for the 60 individuals by both models are available in T2 supplementary table.

The UPGMA tree constructed from Jaccard dissimilarity index clustered the 60 individuals into 2 main clusters (Figure 4). The two clusters were subdivided into 5 sub-clusters, which although not identical to DAPC and STRUCTURE, show a very similar nature of clustering and individual relationships. Here, the lowest dissimilarity (0.20) was between two Cross River individuals (both *I.wombolu*) and the highest dissimilarity (0.95) was between two individuals from Lomie (*I. gabonensis*) and Kogi (*I. wombolu*) (dissimilarity matrix available in T5 supplementary table). Distribution of individual accessions among

Table 4. Measures of variation.

Provenance	Code	Country	Species	%P	Na	He
Cross River	CrNi	Nigeria	*I.wombolu*	6.18%	0.366	0.041
Enugu	EnNi	Nigeria	*I.gabonensis*	41.82%	0.872	0.204
Kogi	KoNi	Nigeria	*I.wombolu*	23.85%	0.632	0.115
Lambarene	LaGa	Gabon	*I.gabonensis*	31.41%	0.705	0.152
Lomie	LoCa	Cameroon	*I.gabonensis*	52.37%	1.056	0.224
Mabanda	MaGa	Gabon	*I.gabonensis*	33.68%	0.724	0.141
Mvilla	MvCa	Cameroon	*I.gabonensis*	61.07%	1.233	0.231
Oyo	OyNi	Nigeria	*I.gabonensis*	60.27%	1.214	0.238
Oyo_wo	OyNi_wo	Nigeria	*I.wombolu*	24.60%	0.651	0.117
Rivers State	RiNI	Nigeria	*I.gabonensis*	53.35%	1.073	0.200
Woleu Ntem	WoGa	Gabon	*I.gabonensis*	47.01%	0.957	0.210
Mean				39.60%	0.368	0.170
SE				5.22%	0.002	0.019

%P = Percentage of polymorphic loci
Na = Number of different alleles
He = Expected heterozygosity

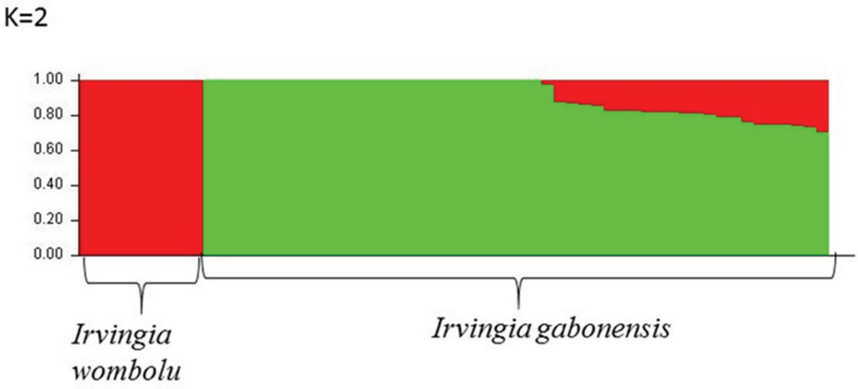

Figure 1. Structuring of the 60 individuals of I. wombolu and I. gabonensis accessions from Nigeria, Cameroon and Gabon by STRUCTURE using membership probability (Q-values) and assuming 2 populations. The colored bar charts represent the two inferred clusters which are based on the two species.

the 5 major clusters in the UPGMA tree displays a close relationship between *I. gabonensis* from Cameroon and Gabon, and from Cameroon and Nigeria. Gabon and Nigeria show some level of delineation given the rare associations in clusters and the high Pairwise PhiPT among provenances of both countries (Table 3). The *I. wombolu* were clustered independently, consistent with the first two clustering models.

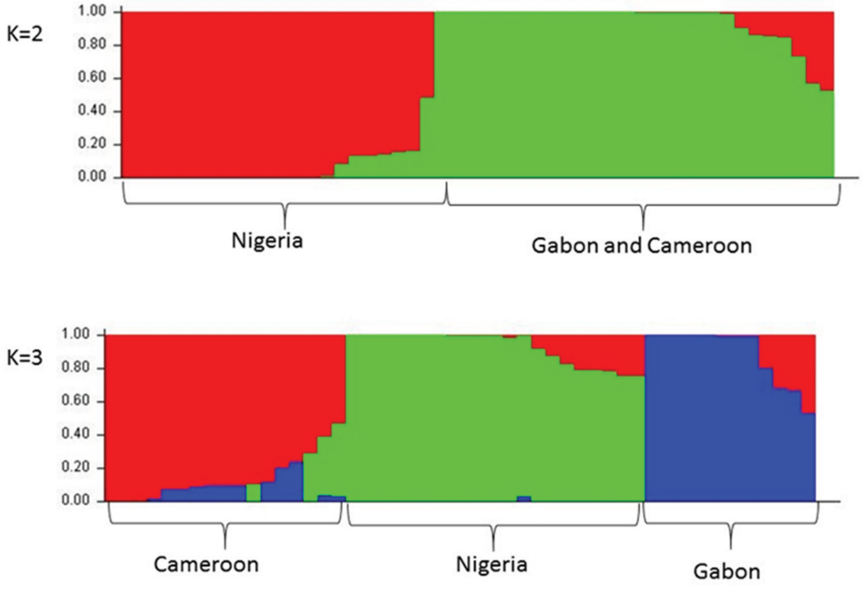

Figure 2. Population structure of 50 individuals of Irvingia gabonensis by STRUCTURE software constructed using Q-values. K = 2 shows the optimal clusters for the 50 individuals while K = 3 shows possible clustering when grouped into three populations.

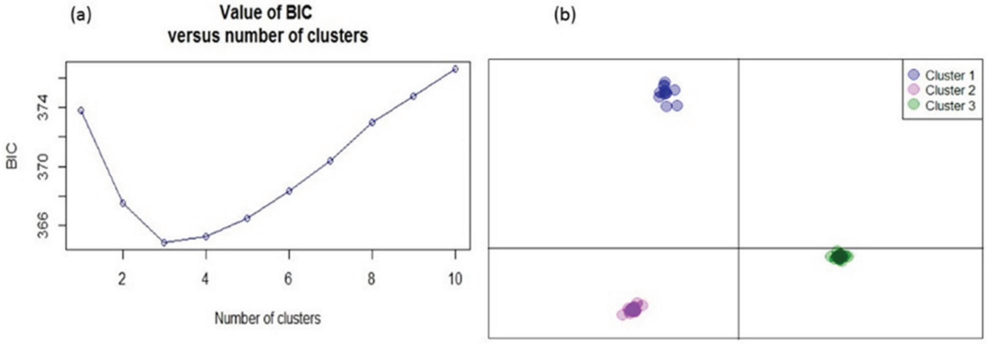

Figure 3. (a) Shows the BIC values (Bayesian Information criterion) with decreasing k values until k = 3 where the graph starts to ascend, indicating optimal clusters at k = 3 (b) Discriminant analysis of principal coordinates (DAPC) scatterplot created in R (package: adegenet) representing 86.10% variance.

Discussion

Sustainable use of available genetic resources can benefit from analysis of genetic diversity and evolutionary relationships of conserved tree germplasm. Such analyses can also provide useful insights for better management of genebanks. Previous studies have characterized both morphological and genetic diversity of *Irvingia* germplasm in ICRAF genebanks (Lowe et al. 2000; Ude et al. 2006) using RAPD and AFLP markers respectively. Analysis of our molecular variance results showed substantial genetic variation partitioned within populations (70%), the remaining 30% of the variation being attributed to among populations.

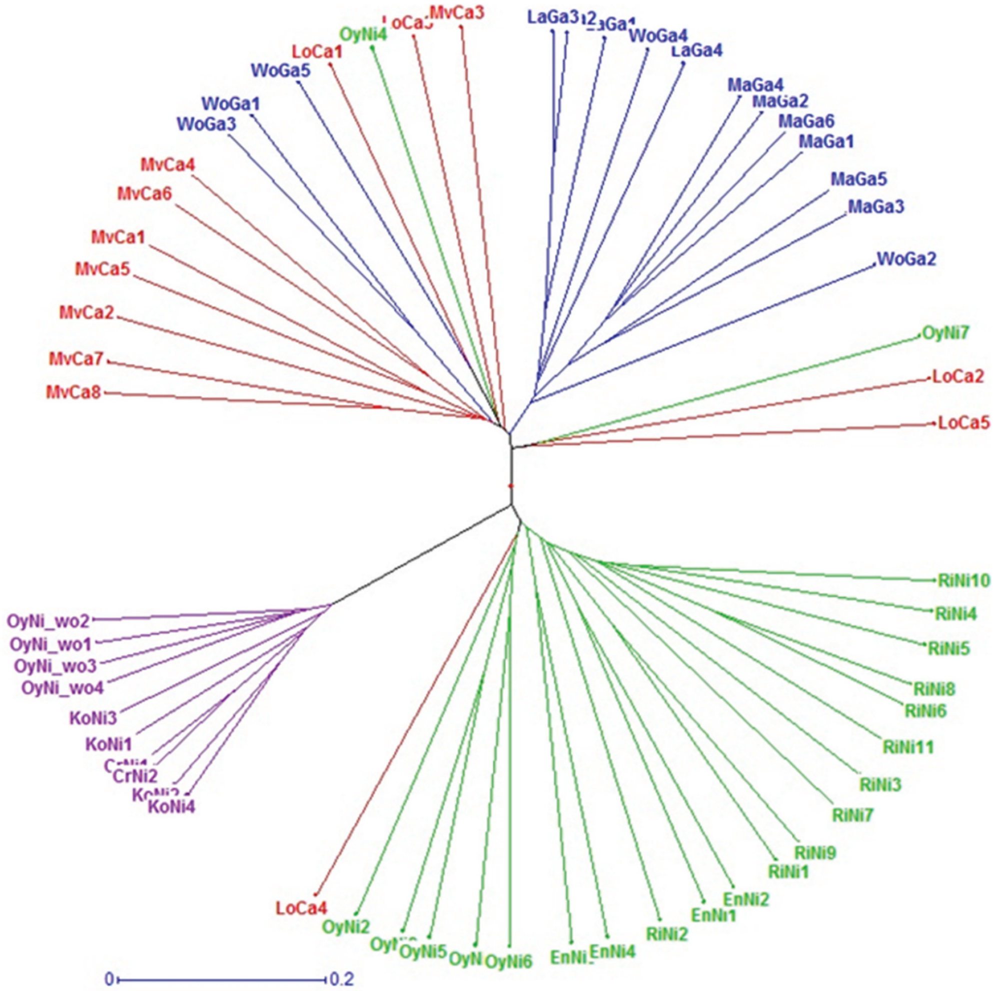

Figure 4. UPGMA tree displaying relationships among 60 individuals from 11 provenances of Irvingia gabonensis and Irvingia wombolu based on 6532 SilicoDArT markers. Sample codes correspond to population codes used in Table 4. Each color represents the species and the country of origin as follows: purple- I. wombolu from Nigeria, red- I. gabonensis from Cameroon, blue- I. gabonensis from Gabon, green- I. gabonensis from Nigeria. The I.gabonensis colors are the same with those of K = 3 bar plot in Figure 2.

Combined with phiPT = 0.301 (P = .001), this indicates moderate differentiation of the 11 provenances. DArTseq markers suggested that the provenances were to some extent evolved to form distinct units, and the pairwise phiPT values confirms this with significantly high values among some of the provenances. Clear delineation was especially evident at the species level and for Gabonese and Nigerian *I. gabonensis* groups. Within populations variation was high which is typical of outcrossing and insect pollinated species (Hamrick et al. 1992) like Irvingia. Irvingia seeds are dispersed mostly by humans and large animals who travel long distances thus facilitating a wider across population gene flow. This is also characteristic of most woody species with the same dispersal system.

The 6532 SilicoDArT markers indicated an overall low genetic diversity for the 60 *Irvingia* accessions owing to the low mean Nei's gene diversity (He = 0.117) and mean proportion of polymorphic loci (39.60%). However, Mvilla and Oyo showed higher levels of diversity compared to all other provenances while Cross River had the lowest genetic diversity (Table 4). The *I. wombolu* individuals were few (ten) and therefore had low diversity compared with *I. gabonensis*. For *I. gabonensis*, measures of diversity showed the highest variation in Cameroon followed by Nigeria and the least in Gabon. Using %P for instance, the trend for *I. gabonensis* is as follows: Cameroon (52.37%-61.07%) Nigeria (41.82–60.27%) Gabon (31.41%-47.01). This trend agrees with results reported by Ude et al. (2006). The low diversity observed in some populations could be attributed to evolutionary factors such as isolation and founder effect or forest disturbance that affect gene flow (Ravinet et al. 2017), which itself could have affected populations prior to seed collection of the accessions used in this study. Pollination also plays a fundamental role in gene flow; in particular, fragmentation hinders the movement of small insect pollinators between fragments which widely upsets the genetic equilibrium of populations (Vranckx et al. 2012). The lowland evergreen West African forests have faced deforestation for a long time. Such disturbance and fragmentation of forest areas negatively affect the genetic variation of outcrossing and insect-pollinated species trees like *Irvingia.*

In our phylogenetic analysis based on STRUCTURE, the first run grouped the 60 accessions into two main clusters (Figure 1), which showed the two species in distinct clusters. The second run (Figure 2) for the 50 individuals of *I. gabonensis* showed structuring within accessions based on the various provenances, which corresponded to the DAPC plot and the UPGMA tree. Clustering points to a shared ancestry for *I. gabonensis* accessions from Cameroon and the other two countries with more associations between Cameroon and Gabon. The same observation led to the idea by Ude et al. (2006) that Cameroon could be the center of genetic diversity and the source of *I. gabonensis* germplasm in both Nigeria and Gabon. This relationship reflects the geographic proximity of Cameroon to both countries (Cameroon is in the middle, Figure 5) implying high chances of gene flow and exchange of material. The tendency of Gabonese *I. gabonensis* accessions to cluster independently of the accessions from Nigeria could also be attributed to probable isolation by distance, which is also supported by the high Nei genetic distances and high Jaccard dissimilarity index among their provenances which could limit chances of direct natural gene flow. Coupled with high pairwise phiPT values and the high genetic dissimilarities between *I. gabonensis* accessions from Nigeria and Gabon this demonstrates the implication of selection pressure in evolution. Over time, populations adapt to immediate environmental conditions such as rainfall patterns and soil types (Kithure et al. 2015).

Conclusion

This study showed the suitability of the dominant SilicoDArT markers in distinguishing populations of Irvingia given the consistent relationships seen in the clustering methods used. Information of genetic diversity and genetic differentiation is imperative to the conservation of *Irvingia* species. This information is essential in decision making on future seed source and germplasm collections. Genebank accessions from Gabon have been consistently reported here and in previous studies to have low genetic diversity. This requires further investigation as to whether it reflects the diversity status in the Gabonese

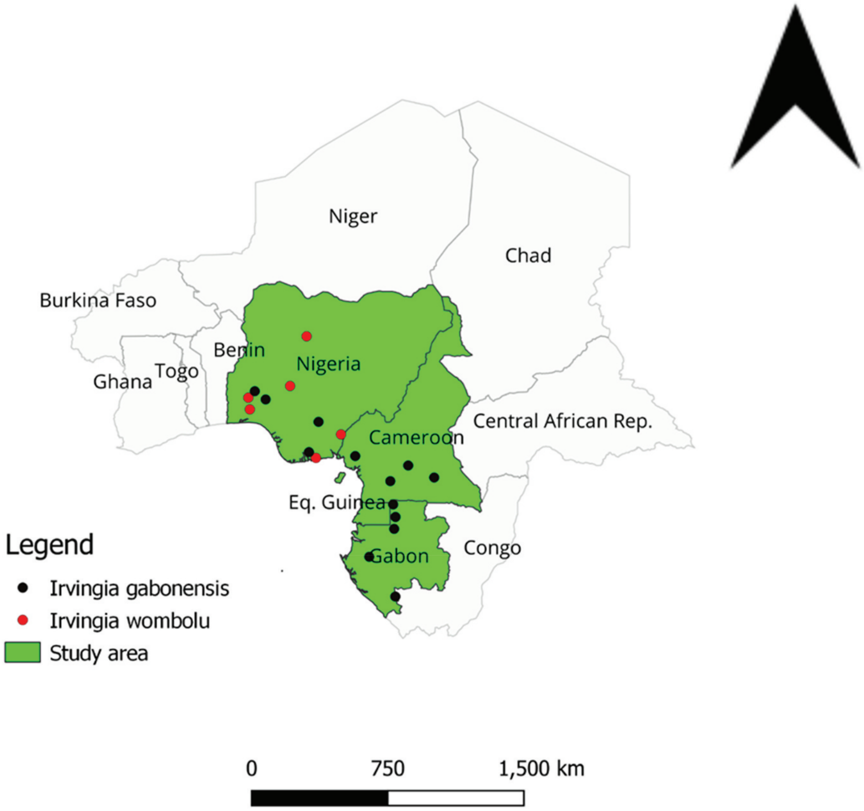

Figure 5. Study area map showing collection points of bush Mango fruits used to germinate the trees sampled for this study. Each dot represents a collection site.

wild populations or just in the genebank collections in which case a recollection to enrich the genepool in the collection should be considered. Forest resources have been shrinking considerably over time due to various anthropogenic activities and this is also the case where *Irvingia* spp are endemic. Some of the suggested ways to improve this situation involve restoration and domestication activities. Since high diversity is associated with high adaptability, populations with higher diversity such as Oyo and Mvilla are recommended to be targeted for future genebank recollections and to augment genebank conservation activities as seed sources for restoration and domestication events, being that on farm tree planting has been suggested to have positive gains in preserving genes and economically important traits.

Acknowledgements

The study was part of thesis work of the first author and she is grateful to the ICRAF molecular and genomics laboratory team for their immense dedication and support. Both authors wish to acknowledge support from the ICRAF genebank team.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was funded by CGIAR Research Program on Genebanks.

ORCID

Sylvia Jepkemboi http://orcid.org/0000-0002-9975-0308
Patroba Ojola http://orcid.org/0000-0003-0869-2550

References

Alam M, Neal J, O'Connor K, Kilian A, Topp B, Peace C. 2018. Ultra-high-throughput DArTseq-based silicoDArT and SNP markers for genomic studies in macadamia. PLoS ONE. 13(8):1–20. doi:10.1371/journal.pone.0203465.

Asaah EK, Tchoundjeu Z, Atangana AR. 2003. Cultivation and conservation status of Irvingia wombolu in humid lowland forest of Cameroon. Environment. 1:251–256.

Daru BH, Berger DK, van Wyk AE. 2016. Opportunities for unlocking the potential of genomics for African trees. New Phytol. 210(3):772–778. doi:10.1111/nph.13826.

Davies TJ, Smith GF, Bellstedt DU, Boatwright JS, Bytebier B, Cowling RM, Forest F, Harmon LJ, Muasya AM, Schrire BD, et al. 2011. Extinction risk and diversification are linked in a plant biodiversity hotspot. PLoS Biol. 9(5):e1000620. doi:10.1371/journal.pbio.1000620.

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

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolut Bioinfo. 1:117693430500100. doi:10.1177/117693430500100003

Hamrick JL, Godt MJW, Sherman-Broyles SL. 1992. Factors influencing levels of genetic diversity in woody plant species. New Forests. 6(1–4):95–124. doi:10.1007/BF00120641.

Harris DJ. 1996. A revision of the Irvingiaceae in Africa. Bulletin du Jardin Botanique National de Belgique/Bulletin van de Nationale Plantentuin van Belgie. 65:143–196. doi:10.2307/3668184.

Hendre PS, Muthemba S, Kariba R, Muchugi A, Fu Y, Chang Y, Song B, Liu H, Liu M, Liao X, et al. 2019. African Orphan Crops Consortium (AOCC): status of developing genomic resources for African orphan crops. Planta. 250(3):989–1003. doi:10.1007/s00425-019-03156-9.

Jombart T, Collins C. 2017. A tutorial for Discriminant Analysis of Principal Components (DAPC) using adegenet 2.1. 0. London (United Kingdom): Imperial College.

Jordan R, Dillon SK, Prober SM, Hoffmann AA. 2016. Landscape genomics reveals altered genome wide diversity within revegetated stands of Eucalyptus microcarpa (Grey Box). New Phytol. 212 (4):992–1006. doi:10.1111/nph.14084.

Killian A, Sanewski G, Ko L. 2016. The application of DArTseq technology to pineapple. Acta Hortic. 1111(1111):181–188. doi:10.17660/ActaHortic.2016.1111.27.

Killian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, Caig V, Heller-Uszynska K, Jaccoud D, Hopper C, et al. 2012. Diversity arrays technology: a generic genome profiling technology on open platforms Methods in molecular biologyVol. 888. Totowa (NJ): Humana Press. 67–89. 10.1007/978-1-61779-870-2.

Kithure RK, Muchugi A, Jamnadass R, Njoka FM, Mwaura L. 2015. Genetic diversity of Faidherbia albida (Del.) A. Chev accessions held at the World Agroforestry Centre. Forests Trees Livelihoods. 24(4):219–230. doi:10.1080/14728028.2015.1054439.

Lowe AJ, Gillies ACM, Wilson J, Dawson IK. 2000. Conservation genetics of bush Mango from central/West Africa: implications from random amplified polymorphic DNA analysis. Mol Ecol. 9 (7):831–841. doi:10.1046/j.1365-294X.2000.00936.x.

Migicovsky Z, Warschefsky E, Klein LL, Miller AJ. 2019. Using living germplasm collections to characterize, improve, and conserve woody perennials. Crop Sci. 59(6):2365–2380. doi:10.2135/cropsci2019.05.0353.

Ngondi JL, Etoundi BC, Nyangono CB, Mbofung CMF, Oben JE. 2009. A novel seed extract of the West African plant Irvingia gabonensis, significantly reduces body weight and improves metabolic parameters in overweight humans in a randomized double-blind placebo controlled investigation. Lipids Health Dis. 8(1):1–7. doi:10.1186/1476-511X-8-7.

O'Connor K, Kilian A, Hayes B, Hardner C, Nock C, Baten A, Alam M, Topp B. 2019. Population structure, genetic diversity and linkage disequilibrium in a macadamia breeding population using SNP and silicoDArT markers. Tree Gen Geno. 15(2). doi:10.1007/s11295-019-1331-z.

Odeku OA, Okunlola A, Lamprecht A. 2013. Microbead design for sustained drug release using four natural gums. Int J Biol Macromol. 58:113–120. doi:10.1016/j.ijbiomac.2013.03.049

Odeku OA, Patani BO. 2005. Evaluation of dika nut mucilage (Irvingia gabonensis) as binding agent in metronidazole tablet formulations. Pharm Dev Technol. 10(3):439–446. doi:10.1081/PDT-54477.

Ofundem T, Ndip NR, Awono A, Levang P. 2017. Bush Mango (Irvingia spp.): forest and on-farm resource availability and market chains in the southwest region of Cameroon. Forests Trees Livelihoods. 26(3):170–182. doi:10.1080/14728028.2017.1283250.

Onyekwelu JC, Olusola JA, Stimm B, Mosandl R, Agbelade AD. 2015. Farm-level tree growth characteristics, fruit phenotypic variation and market potential assessment of three socio-economically important forest fruit tree species. Forests Trees Livelihoods. 24(1):27–42. doi:10.1080/14728028.2014.942386.

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

Plomion C, Bastien C, Bogeat-Triboulot MB, Bouffier L, Déjardin A, Duplessis S, Fady B, Heuertz M, Le Gac AL, Le Provost G, et al. 2016. Forest tree genomics: 10 achievements from the past 10 years and future prospects. Ann For Sci. 73(1):77–103. doi:10.1007/s13595-015-0488-3.

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

Raman H, Raman R, Kilian A, Detering F, Carling J, Coombes N, Diffey S, Kadkol G, Edwards D, McCully M, et al. 2014. Genome-wide delineation of natural variation for pod shatter resistance in Brassica napus. PLoS ONE. 9(7. doi:10.1371/journal.pone.0101673

Ravinet M, Faria R, Butlin RK, Galindo J, Bierne N, Rafajlović M, Westram AM, Mehlig B, Westram AM. 2017. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. J Evol Biol. 30(8):1450–1477. doi:10.1111/jeb.13047.

Schoen DJ, Johnson MTJ, Wright SI. 2019. The ecology, evolution, and genetics of plant reproductive systems. New Phytol. 224(3):999–1004. doi:10.1111/nph.16222.

Slik JWF, Arroyo-Rodríguez V, Aiba SI, Alvarez-Loayza P, Alves LF, Ashton P, Balvanera P, Bastian ML, Bellingham PJ, Van Den Berg E, et al. 2015. An estimate of the number of tropical tree species Proc. Natl. Acad. Sci. U.S.A. 112(24), 7472–7477. 10.1073/pnas.1423147112

Ude GN, Dimkpa CO, Anegbeh PO, Shaibu AA, Tenkouano A, Pillay M. 2006. Analysis of genetic diversity in accessions of Irvingia gabonensis (Aubry-Lecomte ex O' Rorke) Baill. African J Biotech. 5(3):219–223. doi:10.5897/AJB06.041.

Vihotogbé R, Houessou LG, Assogbadjo AE, Sinsin B. 2014. Germination of seeds from earlier fruits of bitter and sweet african bush mango trees. African Crop Sci J. 22(4):291–301.

Vihotogbé R, van den Berg RG, Sosef MSM. 2013. Morphological characterization of African bush mango trees (Irvingia species) in West Africa. Gen Res Crop Evolut. 60(4):1597–1614. doi:10.1007/s10722-013-9969-0.

Vranckx G, Jacquemyn H, Muys B, Honnay O. 2012. Meta-Analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. Conserv Bio. 26(2):228–237. doi:10.1111/j.1523-1739.2011.01778.x.