Vol. 19(7), pp.415-425, July, 2020
DOI: 10.5897/AJB2020.17150
Article Number: 4633F9864221
ISSN: 1684-5315

http://www.academicjournals.org/AJB

ACADEMIC JOURNALS
expand your knowledge

African Journal of Biotechnology

Full Length Research Paper

Genetic diversity of *Faidherbia albida* populations in the Sudano Sahelian region of Cameroon, using simple sequence repeat (SSR) markers

T. Dorothy Tchatchoua[1,2], R. Scott Poethig[3], Erin Doody[3], T. Casey Weathers[1], Krystle Swartz[1], Iain Mathieson[4], Nicole Zembower[1], Tetyana Zhebentyayeva[1] and John E. Carlson[1*]

[1]Department of Ecosystem Science and Management, Pennsylvania State University, University Park PA 16802 USA.
[2]The National Advanced School of Engineering, University of Maroua. BP 46 Maroua, Cameroon.
[3]Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA.
[4]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia. PA 19104, USA.

Received 6 May, 2020; Accepted 6 July, 2020

***Faidherbia albida* (DEL.) A. Chev is widely grown in the northern region of Cameroon because of its many benefits to local communities. However, droughts and deforestation have led to decreases in the abundance of this species, increasing the need to identify improved genotypes for conservation, management, and breeding. The genetic diversity of *F. albida* in nine populations from the Sudano-Sahelian region of northern Cameroon was characterized using microsatellite (SSRs) markers. A total of 28 alleles were recorded across 8 loci and 255 samples. The effective mean number of alleles per locus was 2.3. Observed heterozygosity ranged from 0.24 to 0.30, while expected heterozygosities ranged from 0.22 to 0.26. For most loci, F_{IS} was negative. Higher variation was observed within than among the northern Cameroon populations, and principal component and admixture analyses did not reveal any population substructure. Phenotypic diversity in 3-month-old seedlings was also characterized and significant within population variation was found for most morphological traits. Although some populations differed significantly for one or more traits, in general the populations were phenotypically similar. These results suggest little barrier to gene flow between populations of *F. albida* in northern Cameroon, and that no single provenance is likely to be superior for selection and breeding purposes.**

Key words: Faidherbia albida, phenotypic, genotypic, variation, populations, breeding.

INTRODUCTION

The North and Far North regions of Cameroon are located in the Sahel, and are characterized by Sudano-Sahelian vegetation. This vegetation contains a variety of forest tree species, such as *Adansonia digitata* (Baobab), *Acacia senegal*, *Acacia nilotica*, *Faidherbia albida, Vitellaria paradoxa* (Shea butter), *Tamarindus indica* (Tamarind) and *Azadirachta indica* (Neem), with great multipurpose potential. Collectively, these trees are

*Corresponding author. E-mail: jec16@psu.edu. Tel: (814)863-9164.

economically important sources of timber, building poles, and non-wood products such as gums and tannins, while in some cases also improve soil fertility through nitrogen fixation and organic matter production. However, the region is susceptible to droughts, degradation and desertification which has threatened, as well as caused loss of, species and genetic diversity. Additionally, reduction in land area available for cultivation due to population explosion and growing demand for forest products have resulted in the growth of trees under marginal conditions. The importance, demand and decline of these species calls for the breeding of improved and high yielding genotypes for future planting in forests, parks and farms.

Faidherbia albida (Del.) A. Chev is particularly important in this region because or its unique attribute of producing leaves and fruits in the dry season, when most other trees are leafless and dormant (World Agroforestry Center, 2009). *Faidherbia albida* also provides shelter for animals, whose dung is used as fertilizer for food crops (Boffa, 1999; Roupsard et al., 1999). Besides providing organic fertilizer and livestock fodder, *F. albida* acts as windbreak, supplies wood for fuel and construction, and reduces erosion by reducing soil compaction, which facilitates water absorption during the rainy season (Peltier, 1996). For all of these reasons, *F. albida* was selected as a priority species for domestication (Franzel, 1996) by the World Agroforestry Center (ICRAF) and by the FAO Panel of Experts on Forest Genetic Resources (FAO, 1974). To increase soil fertility in the Far-North region of Cameroon, in 1990 the Rural Development and Management Project (DPGT) lead by SODOCOTON encouraged the planting of *F. albida* on farm lands and the creation of parklands through natural and artificial regeneration (Gautier et al., 2002). Indeed, several reforestation projects in the region are currently ongoing. These projects would benefit from the availability of fast-growing genotypes resistant to biotic and abiotic stresses, but there has been little effort to identify improved genotypes of *F. albida* suitable for this region.

Forest tree genetic improvement programs typically begin with an assessment of the genetic diversity within the species of interest, to determine the amount of genetic and phenotypic variation available for selection (Muona, 1990; Wright, 1962; Franzel et al., 1996). Seed samples are collected across the entire natural range of the species, or within a geographic area of interest, with which seedling field trials are conducted, referred to as provenance tests (Zobel and Talbert, 1984). In the case of *F. albida,* genetic improvement studies started three decades ago with the morphological characterization of provenances from northwest, northcentral, east and south African populations, including field trials in Burkina Faso, Zimbabwe, Cameroon and Sudan (Sniezko and Stewart, 1989; Peltier and Eyog-Matig, 1988; Vandenbeldt, 1991, 1992; Abdelazim et al., 1997; Tchatchoua et al., 2019). These trials revealed two distinct ecotypes, one spanning the Sahel in North Africa, and a second typically found in riparian habitats in East and South Africa. Consistent with their habitats, provenances from the Sahel have small seeds, grow relatively slowly, and are drought tolerant, whereas provenances from East and South Africa have relatively large seeds, grow quickly, and are susceptible to drought (Vandenbeldt, 1991; Dangasuk et al., 2006). Analyses of isozyme variation support this distinction, and reveal that populations from northwest and northcentral Africa are genetically more diverse than populations from east and south Africa (Harris et al., 1997; Abdelazim et al., 1997; Joly et al., 1992).

Microsatellite (simple sequence repeat, SSR) markers have many advantages for analyses of genetic variation because of their high reproducibility, co-dominant inheritance, multi-allelic variation and abundance in the genome (Muchugi et al., 2008; Varshney et al., 2005). In addition, they can be evaluated relatively cheaply and at high throughput using automated platforms. In this study, the researchers took advantage of SSRs identified in *F. albida* by Russell et al. (2014) to characterize the genetic diversity within 9 provenances of *F. albida* from the Sudano-Sahelian region of Cameroon. The primary goals of this study were to determine if these markers could be used to differentiate populations and to identify populations with sufficient genetic variation to be useful for breeding programs.

MATERIALS AND METHODS

Germplasm collection

Seeds were collected indeterminately from well-spaced trees in nine populations in the distribution range of the species in the Sudano-Sahelian region of Cameroon (Table 1 and Figure 1). 50 to 100 fruit pods were collected from each of 20 trees in each population. The trees were at least 100 m apart to avoid sampling trees of the same genotype, and populations were separated by a distance of at least 100 km as proposed by Tchoundjeu et al. (1998). The seeds were extracted from the fruit by students of the National Advanced School of Engineering, University of Maroua during a practical session. Following inspection by the Plant Protection and Regulatory Services Directorate of Cameroon's Ministry of Agriculture and Rural Development (MINADER), the seeds were sent to the Schatz Center of Molecular Genetics USA (USDA import permit OMB No. 0579 - 0049) for DNA extraction, amplification and sequencing.

Phenotypic analysis

Seeds were nicked at the micropylar end using nail clippers, being careful to avoid damaging the radicle. They were germinated for two days on moist paper towels at 30°C, and viable seeds were then planted in a mixture of 1/3 bar sand and 2/3 Scott's Premium Topsoil in Stuewe TP38 pots. Four plants from each of 6 trees in each provenance were grown under fluorescent illumination in Conviron growth chambers on a diurnal cycle consisting of 12 h light (PAR 265 µm m^{-2} s^{-1}) at 30°C, and 12 h dark at 20°C. One-half teaspoon of Osmocote slow release fertilizer was added to each pot

Table 1. Populations of *Faidherbia albida* from northern Cameroon used in this study.

Population	Family code	Number of individuals	Geographical location
Adoumri	AD	24	09°16'N 13°50'E
Dinao	DIN	27	10°14'N 14°72'E
Kongola	KON	30	10°38'N 14°24'E
Mambang	MAM	30	10°65'N 14°28'E
Mokolo	MOKO	30	10°44'N 13°48'E
Moulvoudaye	MOU	30	10°23'N 14°50'E
Sirlawe	SIR	27	10°04'N 14°57'E
Tokombe	TOKO	30	10°52'N 14°09'E
Zamai	ZA	27	10°36'N 13°54'E

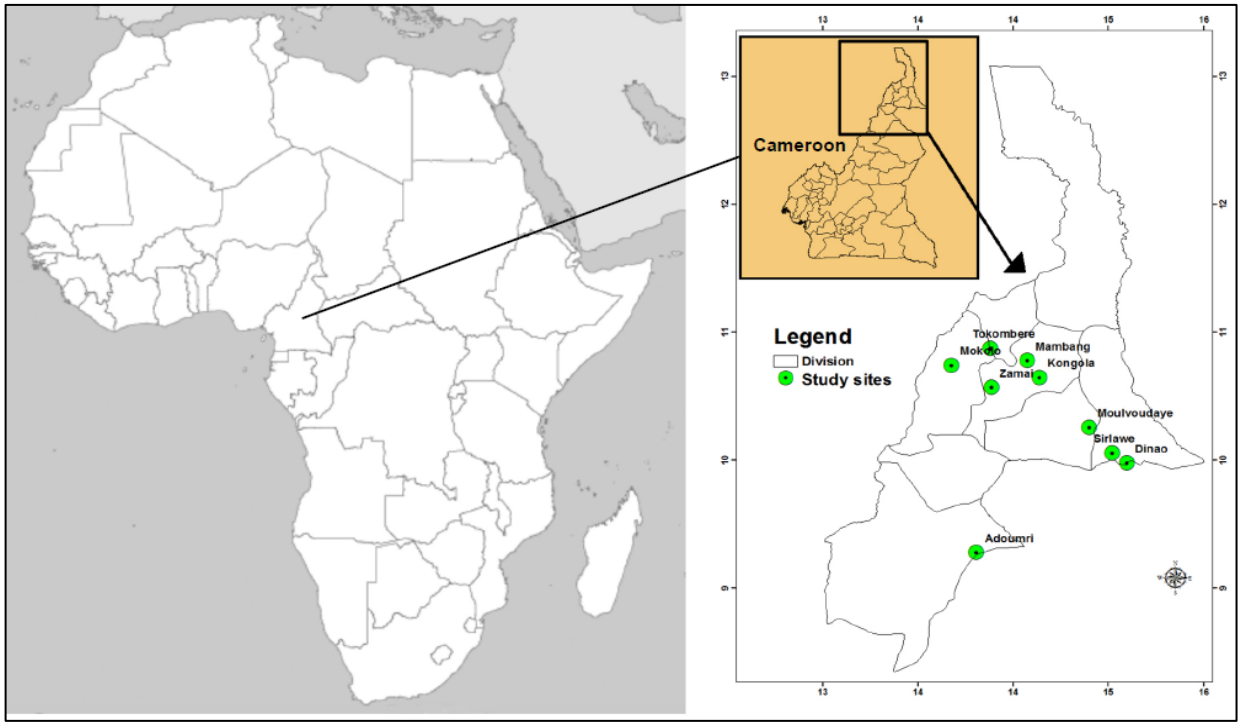

Figure 1. Maps of the geographical location of Cameroon and the study sites.

at planting and fertilized once a month with a dilute solution of 10:20:10 fertilizer. Plants were measured 3 months after planting with the help of a stereomicroscope. Plant height was measured from the soil surface to the tip of the last leaf. The number of all leaves larger than 5 mm was recorded at the same time. Stem width was measured in the third internode, using Vernier calipers. Fully expanded leaves near the tip of the shoot were attached to paper using double sided tape, scanned into a computer, and secondary leaflet area was then measured using Image J.

Growth of seeds and DNA extraction

A total of 255 seeds (that is, 3 seeds per family, 8 to 10 families per population and 9 populations) were nicked with a nail clipper and sown in polymix soil in 4 in plastic pots. Plants were grown in a growth chamber (25°C, 16 h. light: 8 h dark), and leaves were harvested for DNA extraction when plants were 4 weeks old. Approximately 100 mg of leaf tissue was macerated using a TissueLyser 11 (11.5 rmp for 1 minute), and genomic DNA was then extracted using DNeasy Plant Mini Kit (Qiagen), following the manufacturer's instructions. DNA concentrations were measured using a Qubit DNA kit (Qubit™ Assays) and a working DNA concentration of 1 ng/µl solution was prepared for all samples and stored in a -80°C freezer for further use.

Screening of SSR primers

Eight DNA samples, one from each of eight families were screened with 48 SSR primer pairs previously developed for *F. albida* (Russel et al., 2014). PCR was performed using 1 µl genomic DNA (1 ng/µl),

Table 2. Sequence of the PCR primers for the SSR markers used in this study.

Marker	Primer names	Primer sequences	Repeat motifs
5	comp19011_c0_seq1_120-F	CTTCGCAATAAGTTCTTCGT	AAG
	comp19011_c0_seq1_120-R	ATTCCTTGGTTGTGACTGAG	
8	comp15062_c0_seq1_145-F	CTACCACTGGTTCTCCTCAG	ACC
	comp15062_c0_seq1_145-R	TAGTTCGATACCAATCACCC	
11	comp20414_c0_seq1_450-F	TTGAGGTTCAAGCCAGTAGT	AGG
	comp20414_c0_seq1_450-R	TGTTTCTCAAATCTCCGTTT	
22	comp34406_c0_seq1_376-F	GATTTCATCGCAGGTAACTC	AAAG
	comp34406_c0_seq1_376-R	CTCTTATCATCGTCTCTGCC	
23	comp5175_c0_seq1_78-F	ACCCATATGAAACCCATACA	ACAT
	comp5175_c0_seq1_78-R	CTCGTGGAAATATCATTGGT	
24	comp5630_c0_seq1_260-F	CAAACTTCACACAACGTAGC	AAAT
	comp5630_c0_seq1_260-R	ATCCCAGACCCTAAAGGATA	
26	comp14514_c0_seq1_453-F	ACTCCAATTCTGGATCAGTG	AG
	comp14514_c0_seq1_453-R	GGGATTGATTCCTTCTAACA	
30	comp4810_c1_seq1_413-F	TCTTTTCCTCCTCATTCTCA	AAG
	comp4810_c1_seq1_413-R	ATGCAGATGGTTTAATCGAC	

2.5 µl each of 10 µM forward and reverse primers, 19 µl of Ultra-Pure water and 1 Edvo-tek bead. Samples were amplified in an Applied Biosystems thermocycler using the following conditions: 95°C for 15 mins, followed by 35 cycles of 94°C for 30 s, 55°C for 1.5 min and 72°C for 1 min, with a final extension at 72°C for 10 min. Eight primer pairs that produced clear, strong, polymorphic amplification products between 100 and 250 base pairs (bp) in length were chosen for further analyses (Table 2).

PCR amplification procedure

The eight selected markers were used to genotype 255 individuals from 85 families and nine populations. PCR was performed with a fluorescently labeled forward primer using the conditions described above, and the amplification products were then separated using an Applied Biosystems Illumina 3730 XL DNA sequencer at the Huck Genomic Core Facility of Pennsylvania State University, State College USA. Fragment analysis was performed with GeneMapper software at GS 500 (-200) LIZ standard size.

Data analysis

The phenotypic data were analyzed using SPSS v12 statistical package (IBM; https://www.ibm.com/products/spss-statistics) for data description and analysis of variance (ANOVA). All the measured traits assumed normal distribution and significance level for the ANOVA was at 0.05. Means comparison was conducted using the DUNCAN multiple range test (Duncan, 1955).

Population genetic analysis of the SSR marker results was performed within Excel using GenAlEx 6.5 (Peakall and Smouse, 2012). Summary statistics for allelic richness (the total number of alleles per locus and the effective number of alleles), unbiased observed and expected heterozygosities and the percentage of polymorphism in each population were calculated. The allelic richness (number of alleles), fixation index (F_{IS} = inbreeding) and geneflow information were calculated with F. Stat version 2.9 3.2 (February, 2002). GenePop (https://genepop.curtin.edu.au/) was used to conduct Hardy Weinberg Exact tests (HWE) and to estimate linkage disequilibrium at 0.05 Bonferroni corrections for pairwise comparison, genetic bottleneck, and Nei effective population size. The program ADMIXTURE v1.3.0 (Alexander et al., 2009) was used to cluster individuals into 2<K<10 clusters. The lowest cross-validation error occurred with K=2. Analysis of molecular variance (AMOVA) (Excoffier et al., 1992) was performed as a pairwise comparison of F_{ST} (genetic fixation index) values and *P*-values, using GenAlEx 6.5. Phylogenetic trees were constructed using Mega 7.027 software (Kumar et al., 1994).

RESULTS

Phenotypic variation in *F. albida* populations

A range of morphological traits was examined to determine the degree to which populations of *F. albida* in Cameroon are phenotypically distinct at the seedling stage. ANOVA revealed that there was significant variation for most traits among all 9 populations (Table 3). However, no single population was consistently different from other populations (Figure 2) and, in general; the populations were morphologically quite similar (Figure 3). At three months of age, the population means for the width of the third internode ranged from 1.5 and 1.8 mm, total leaf number ranged from 29 to 39, and plants were 19 to 29 cm in height. At this stage, populations had an average of 1 to 4.5 branches greater than 1 cm in length. The most recently formed leaves had one or two pairs of primary leaflets, each with 6 or 7 pairs of secondary leaflets, which were 4 to 5 mm^2 in area. Extrafloral nectaries (EFN) were typically absent from the first 4 to 10 leaves, but were present on later-formed leaves. One of the most obvious differences between populations was the frequency of glabrous (hairless) plants. Glabrous plants were either absent (Sirlawe) or a minor fraction of

Table 3. Analysis of variance for morphological traits.

Trait	SS	df	MS	F	P-value
Internode diameter (mm)	2.129	8	0.660	2.402	0.017
Total leaf number	1469.957	8	183.745	2.285	0.023
Plant height (cm)	2104.873	8	263.109	3.519	0.001
No. branches> 1 cm	174.667	8	21.833	2.628	0.009
No. of leaflet pairs	35.286	8	4.411	4.888	0.000
Leaflet area (mm^2)	32.184	8	4.023	1.239	0.280
Node with first extrafloral nectary	1592.439	8	199.055	2.644	0.009

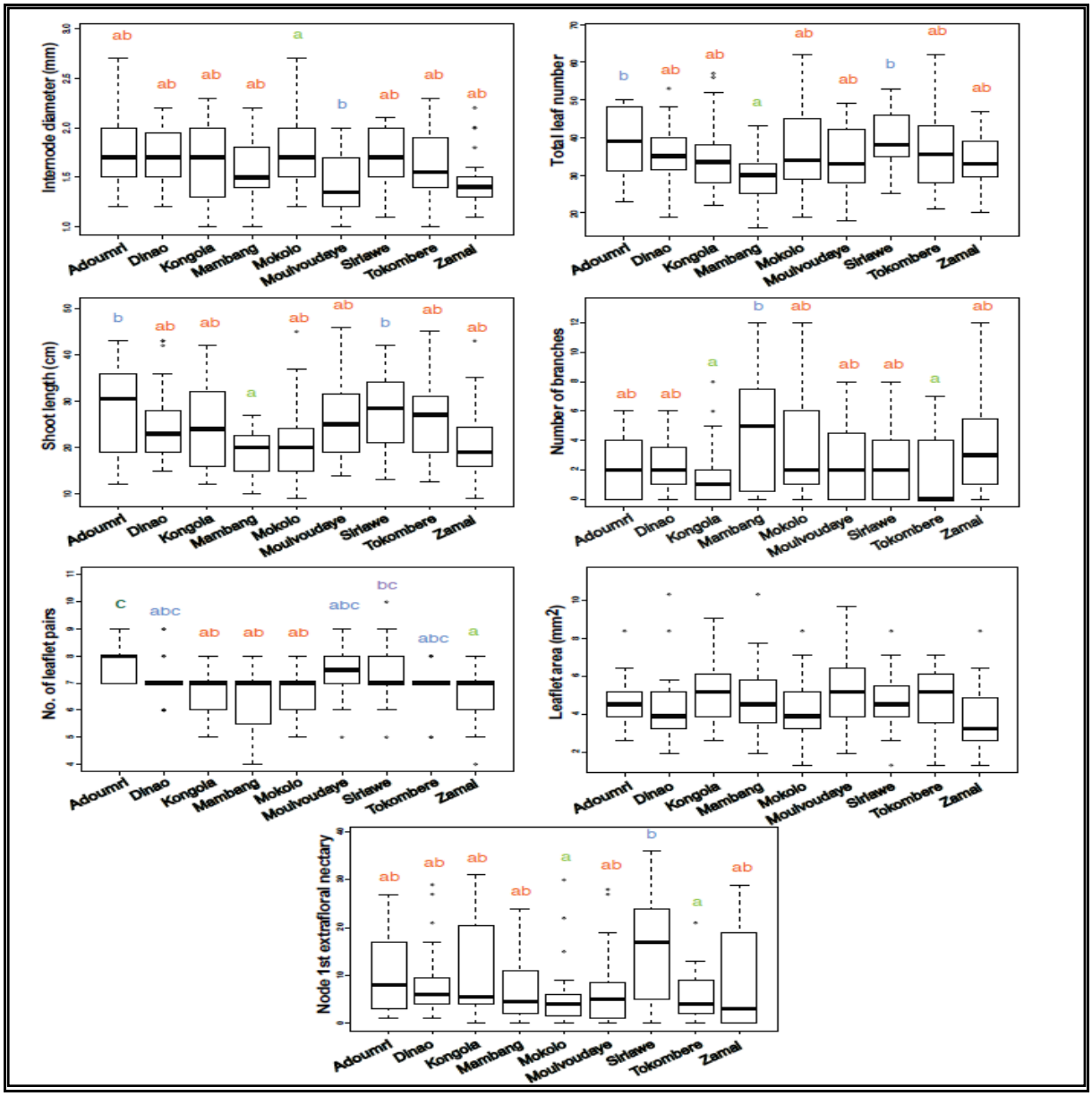

Figure 2. Comparisons of morphological traits in 3-month-old seedlings among *F. albida* populations. Center lines mark the median value, boxes outline the first and third quartiles, whiskers mark minimum and maximum values, and dots indicate outliers. Statistical significance was determined by a one-way ANOVA test followed by a family-wise Tukey HSD test. Populations with different letters are statistically different from each other at the 95% confidence level (P<0.05).

Figure 3. Three-month-old saplings of the provenances used for phenotypic analysis.

the plants in the Adoumi (11%), Dinao (22%), Kongolo (19%), Moulvoudaye (5%), and Tokombere (9%) populations, but were the major class in Mambang (87%), Mokolo (54%) and Zamai (79%) (Figures 2 and 3).

Genetic variation in *F. albida* populations

A set of 48 SSR markers were tested with DNAs from eight families to identify robust markers for assessing genetic variation in this study. Eight informative markers were selected as being most reproducible and informative based on agarose gel results, which were then used to genotype 255 individuals from 9 populations in Cameroon by high-resolution capillary electrophoresis. Seven of the SSR loci were polymorphic, while Locus_24 proved to be monomorphic in all of the populations studied based on GeneMapper analysis. The allele frequencies across all populations and loci are presented in Table 4 and the population genetic parameters calculated for these loci from the allele frequencies are presented in Tables 5 and 6. A total of 28 alleles were recorded across the 8 loci and 255 samples. The number of alleles observed at each locus ranged between 1 (marker 24) and 7 (marker 5), while the number of effective alleles ranged from 1 to 2.7 with a mean Ne of 1.5. The mean percentage of polymorphism was 73.6, while among populations polymorphism percentages ranged from 62 in Moulvodaye and Zamai to 87.5 in Sirlawe. Alleles 189 of Locus_22 and 195 of Locus_5 were only found in the Adoumri population. Alleles 180 and 198 of Locus_5 were only found in Moulvoudaye. Alleles 174 of Locus_30 and 153 of Locus_8 were unique to Kongola and allele 228 of Locus_11 was restricted to Dinao. As these alleles were only present at frequencies between 2 and 5%, further investigations should be conducted to determine if they can be used to discriminate between the Cameroon populations.

The highest number of alleles was found in the Sirlawe, Kongola and Adoumri populations, which also had the highest Shannon diversity indices (I). The Shannon diversity index ranged from 0.000 to 1.095 among the 9 populations, with a mean of 0.417. Families DIN 10, SIR 06, and TOKO 10 showed variation in 5 loci, while families AD 26, AD 36, DIN 07, DIN 11, KON 03, KON 01, KON 05, KON 08, KON 3, MOKO 08, SIR 05, SIR 28 and TOKO 05 varied at 4 loci (data not shown). These populations thus should be well-suited for selecting individuals for a tree improvement program. The average levels of observed and expected heterozygosities were 0.273 and 0.238 respectively. Observed heterozygosities ranged from 0.24 to 0.30 while expected heterozygosity ranged from 0.22 to 0.26. The highest heterozygosities were found in Sirlawe and the lowest in Moulvoudaye, but overall the heterozygosities values for all of the populations were very close to one another (Table 6).

The Hardy-Weinberg Equilibrium (HWE) values for the 9 populations ranged from 0.75 to 1.00, with an average value across populations of 0.93 and overall value of 0.67 for all 255 samples (Table 6). These HWE values indicate that the populations are largely in genetic equilibrium and that the allele frequencies can be expected to be stable in subsequent generations. The case for genetic equilibrium is supported by the lack of linkage disequilibrium (LD) observed among alleles of the 8 loci across all populations. Although LD is not necessarily expected among only 8 loci, the lack of LD does confirm that the selected loci are likely to be performing as neutral and independent loci and thus appropriate for use in genetic diversity estimates.

The sampled *F. albida* populations in Cameroon displayed higher within population variation (98%) than among population variation (2%), as shown by the Fst values for the individual SSR loci which ranged from

Table 4. Allele frequencies in populations of *F. albida* in northern Cameroon.

Primer	Alleles	AD	DIN	KON	MAMB	MOKO	MOU	SIR	TOKO	ZA	Mean
	180	0.000	0.000	0.000	0.000	0.000	0.050	0.000	0.000	0.000	0.006
	186	0.438	0.389	0.534	0.259	0.467	0.550	0.519	0.414	0.500	0.452
	189	0.292	0.167	0.207	0.431	0.217	0.117	0.037	0.207	0.148	0.202
Primer 5	192	0.250	0.389	0.259	0.276	0.317	0.250	0.370	0.379	0.352	0.315
	195	0.021	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
	198	0.000	0.000	0.000	0.000	0.000	0.033	0.074	0.000	0.000	0.012
	201	0.000	0.056	0.000	0.034	0.000	0.000	0.000	0.000	0.000	0.010
	153	0.000	0.000	0.033	0.000	0.000	0.000	0.000	0.000	0.000	0.004
Primer 8	165	0.250	0.192	0.233	0.133	0.200	0.117	0.222	0.167	0.278	0.197
	168	0.750	0.808	0.733	0.867	0.800	0.883	0.778	0.833	0.722	0.799
	209	0.348	0.333	0.267	0.167	0.283	0.333	0.327	0.328	0.333	0.300
Primer 11	217	0.152	0.130	0.183	0.167	0.133	0.133	0.038	0.190	0.296	0.159
	227	0.500	0.519	0.550	0.667	0.583	0.533	0.635	0.483	0.370	0.540
	228	0.000	0.019	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
	189	0.021	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
Primer 22	239	0.000	0.000	0.037	0.017	0.000	0.000	0.000	0.017	0.000	0.008
	243	0.958	0.979	0.963	0.983	0.931	1.000	0.960	0.983	1.000	0.973
	245	0.021	0.021	0.000	0.000	0.069	0.000	0.040	0.000	0.000	0.016
Primer 23	207	1.000	1.000	1.000	1.000	1.000	1.000	0.981	1.000	1.000	0.998
	210	0.000	0.000	0.000	0.000	0.000	0.000	0.019	0.000	0.000	0.002
Primer 24	202	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	182	0.083	0.000	0.033	0.000	0.017	0.000	0.038	0.017	0.000	0.020
Primer 26	184	0.833	0.885	0.867	0.950	0.917	0.933	0.885	0.917	0.963	0.907
	186	0.021	0.058	0.067	0.000	0.017	0.000	0.038	0.000	0.000	0.022
	188	0.063	0.058	0.033	0.050	0.050	0.067	0.038	0.067	0.037	0.051
	174	0.000	0.000	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.002
Primer 30	177	0.979	0.980	0.931	0.850	0.933	0.850	0.840	0.933	0.981	0.918
	180	0.021	0.020	0.052	0.150	0.067	0.150	0.160	0.067	0.019	0.080

Table 5. Genetic diversity statistics for each microsatellite locus used in this study.

Loci	N	Na	Ne	I	Ho	He	uHe	F	F_{IS}	F_{IT}	F_{ST}	Nm
5	28	3.667	2.734	1.095	0.694	0.632	0.644	-0.096	-0.098	-0.060	0.035	6.961
8	28	2.111	1.483	0.506	0.325	0.319	0.325	0.001	#N/A	0.000	0.017	#N/A
11	28	3.111	2.422	0.969	0.811	0.581	0.591	-0.394	-0.396	-0.363	0.023	10.468
22	27	1.889	1.056	0.117	0.054	0.052	0.053	-0.035	-0.043	-0.020	0.022	11.150
23	28	1.111	1.004	0.010	0.004	0.004	0.004	-0.019	-0.019	-0.002	0.016	14.906
24	26	1.000	1.000	0.000	0.000	0.000	0.000	#N/A	#N/A	#N/A	#N/A	#N/A
26	28	3.111	1.219	0.374	0.142	0.174	0.177	0.170	0.180	0.194	0.017	14.536
30	27	2.111	1.179	0.262	0.153	0.142	0.144	-0.061	-0.080	-0.035	0.042	5.685
Mean	27.79	2.264	1.512	0.417	0.273	0.238	0.242	-0.069	-0.068	-0.041	0.025	9.772
SD	0.26	0.122	0.078	0.048	0.036	0.029	0.029	0.027	0.060	0.058	0.004	1.859

N = Total number of alleles observed, Na = Average number of alleles per population, Ne = Number of effective alleles per population, I = Shannon Diversity index, Ho = Observed heterozygosity, He = Expected heterozygosity, uHe = Nei's unbiased gene diversity corrected for sample size, F = Overall inbreeding coefficient, F*IS* = Inbreeding coefficient of individuals relative to their subpopulation, F_{IT} = Inbreeding coefficient of individuals relative to the total population, F_{ST} = Proportion of variation in subpopulations relative to all populations sampled, N = Estimated allele migration rates among sub-populations**.**

Table 6. Genetic diversity statistics from allele frequencies of 8 SSR loci in collections from 8 northern Cameroon populations of *F. albida*.

Population	N	Ar	Ho	He	uHe	F	F_{IS}	HWE	LD
Cameroon (total)	255	1.24	0.27	0.24	0.24	-0.04	**-0.11**	**0.67**	0.00
Adoumri	24	1.26	0.29	0.26	0.26	-0.07	-0.09	1.00	0.00
Dinao	27	1.24	0.25	0.23	0.24	0.06	-0.04	1.00	0.00
Kongola	30	1.26	0.28	0.26	0.26	-0.03	-0.07	0.83	0.00
Mambang	30	1.23	0.27	0.22	0.23	-0.15	-0.19	1.00	0.00
Mokolo	30	1.24	0.28	0.24	0.24	-0.12	-0.15	0.83	0.00
Moulvoudaye	30	1.23	0.24	0.22	0.23	0.06	-0.06	1.00	0.00
Sirlawe	27	1.26	0.30	0.25	0.26	-0.15	-0.17	1.00	0.00
Tokombere	30	1.24	0.27	0.23	0.24	-0.02	-0.15	1.00	0.00
Zamai	27	1.23	0.28	0.22	0.23	-0.17	-0.25	0.75	0.00
means		1.24	0.27	0.24	0.24	-0.07	-0.13	0.93	0.00
SE		0.014	0.017	0.013	0.014	0.084	0.066	0.095	0.000

N = sample size, Ar = Average number of alleles per individual, Ho = Observed heterozygosity, He = Expected heterozygosity, uHe = Nei's unbiased gene diversity corrected for sample size, F = Overall inbreeding coefficient, FIS = Inbreeding coefficient of individuals relative to their subpopulation, HWE = Hardy-Weinberg Equilibrium, LD = Linkage disequilibrium.

Table 7. Pairwise comparison of F_{ST} (genetic fixation index) values among Cameroon collections.

Collections	Adoumri	Dinao	Kongola	Mambang	Mokolo	Moulvoudaye	Sirlawe	Tokombere	Zamai
Adoumri	–	0.230	0.400	0.010	0.399	0.056	0.036	0.398	0.201
Dinao	0.005	–	0.429	0.008	0.420	0.098	0.162	0.418	0.098
Kongola	0.000	0.000	–	0.005	0.421	0.191	0.187	0.256	0.118
Mambang	0.033	0.028	0.031	–	0.022	**0.001**	**0.000**	0.019	**0.000**
Mokolo	0.000	0.000	0.000	0.023	–	0.324	0.205	0.407	0.089
Moulvoudaye	0.017	0.011	0.006	**0.040**	0.002	–	0.182	0.318	0.031
Sirlawe	0.024	0.007	0.006	**0.057**	0.006	0.007	–	0.039	0.008
Tokombere	0.000	0.000	0.004	0.023	0.000	0.003	0.019	–	0.397
Zamai	0.008	0.013	0.010	**0.066**	0.014	0.023	0.040	0.000	–

Mean = 0.02, range: 0.00–0.07, SD ± 0.02. Pairwise F_{ST} values below the diagonal, *P*-values above. Significant differences are in bold.

0.016 (1.6%) to 0.042 (4.2%) across the populations (Table 5). This finding was reinforced by the results of a pairwise comparison of F_{ST} values between the populations which resulted in a mean Fst value overall all comparisons of 0.02 and a range of Fst values between population pairs ranging from 0.00 to 0.07, with SD of 0.02. (Table 7). An analysis of population structure using admixture detected 2 types of genotypes which were both widely distributed among all populations sampled (Figure 4). A principal component analysis (Figure 4) based on allelic variation among all 255 individuals was also unable to differentiate among the 9 populations of *Faidherbia albida* sampled in Northern Cameroon. Also, the F_{IS} values were negative for most loci and ranged from -0.396 to 0.18.

DISCUSSION

Eight informative microsatellite markers (SSRs) were selected from the markers developed by Russell et al. (2014) for *F. albida* in tropical Africa to assess genetic diversity from populations in northern Cameroon for immediate application in seed collections for reforestation and agroforestry, and to initiate a breeding program. Phenotypic data were also collected and analyzed from the populations. It was found that there was higher within population than among population genetic variation, implying a high rate of gene flow between the Cameroon populations sampled. This is consistent with the extremely phenotypic variability found within individual populations, while overall the populations were morphologically quite similar. Genetic studies of other tropical fruit and timber species from Cameroon, using SSR markers, also indicated higher within population than among population variation (Benoit et al., 2011; Garcia et al., 2004), and similar results have been reported for A*cacia senegal* in Kenya (Omondi et al., 2010). Analyses of isozyme loci showed that populations of *F. albida* from across west and north Africa are

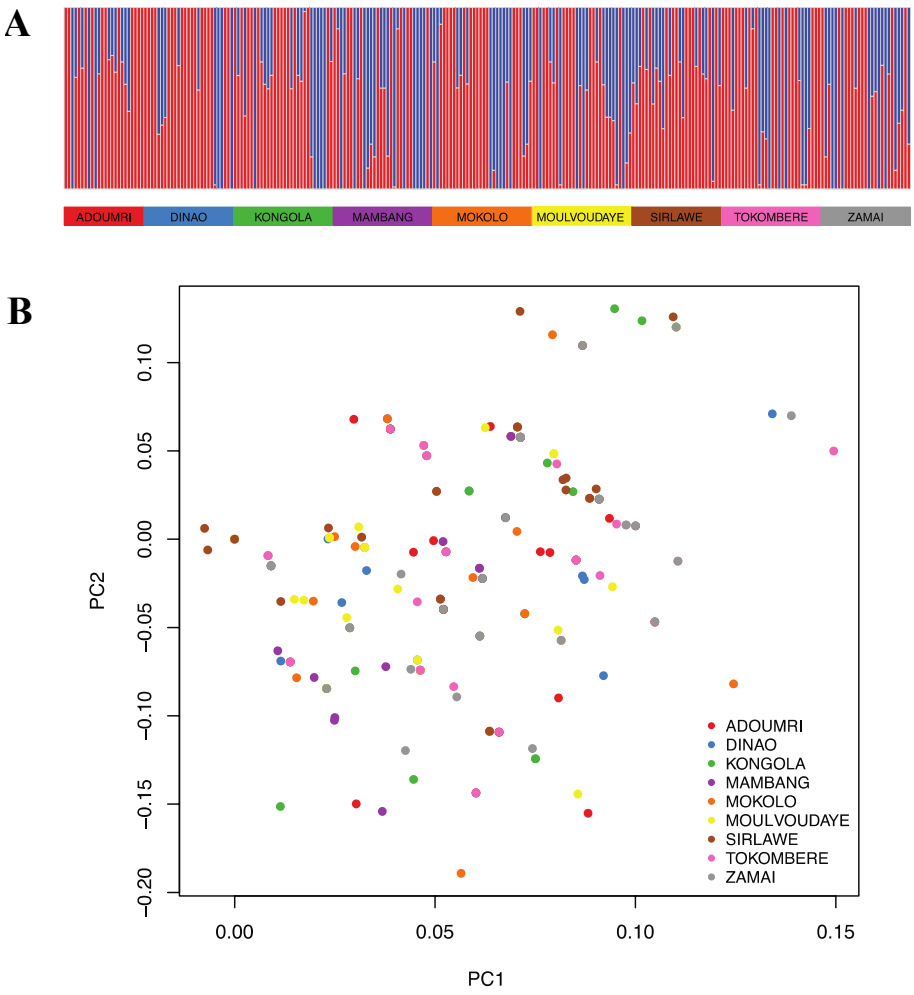

Figure 4. Results of hierarchical population structure analyses of genetic variation in populations of *Faidherbia albida* in Cameroon. (A) Admixture analysis of genotypic variation across individuals grouped by population; (B) Principal Component Analysis (PCA) based on allelic variation among all 255 individuals sampled from 9 populations of *Faidherbia albida* in Cameroon.

genetically diverse and are distinctly different from east African populations, which displayed a lower amount of genetic diversity (Joly et al., 1992; Harris et al., 1997; Dangasuk et al., 2002). In these previous studies, despite the high level of genetic diversity within the western and northern populations, within population diversity was found to be greater than among population diversity in this broad region (Harris et al., 1997). This may be attributable, in part, to exchange of seeds among farmers and transfer by animals during the nomadic life of herders in search of pasture (Hauser, 1994; Wickens, 1969).

The relatively low amount of genetic differentiation between populations of *F. albida* in the far north region of Cameroon has important implications for breeding studies. From a practical standpoint, this result means that it is impossible to predict from genetic data which populations are most likely to provide superior sources of breeding material. Given that all of the populations examined in this study are genetically similar, in principle any population is a useful starting point for the derivation of improved varieties. However, the relatively large amount of genetic variation and phenotypic variation within each of the populations reveals substantial opportunities for selection and breeding for *F. albida* improvement at the population level. Determining which populations and traits are best for breeding agronomically useful traits will require more detailed information, including which phenotypic traits segregate in families from different populations. In contrast, even with the large number of samples genotyped, the F_{IS} values were negative for most loci. Negative F_{IS} values indicate an excess of heterozygotes in populations (Guries and Ledig, 1981). This suggests that *F. albida* populations in Cameroon are either subject to high selective pressure,

or arose from genetic bottlenecks from a relatively small number of founding individuals (Hedrick and Kalinowski, 2000).

In this study, the initial steps in establishing breeding and reforestation programs for *F. albida* in Cameroon have been accomplished through the collection of seed from natural populations and assessing of genetic and morphological variation among seedlings from these populations. The high levels of variation observed within all of the source populations provides ample opportunity for making great strides in trait improvement, while retaining genetic diversity, during the next steps of selecting trees with desirable genotypes and establishing breeding orchards. The availability of the genome sequence of *F. albida* (Chang et al., 2018), will make it possible to identify many new polymorphic molecular markers useful for mapping quantitative trail loci linked to traits useful in agroforestry applications. The genetic diversity results coupled with the genome sequence should enable the use of genome-wide marker assisted selection in *Faidherbia* --an approach that is widely used in the breeding of crop plants and has the potential to dramatically accelerate the discovery and improvement of trees useful for agroforestry (Alkimim et al., 2020; Grattapaglia, 2017; Ribaut et al., 2010; van Nocker and Gardiner, 2014; Watson, 2019; Zargar et al., 2015). *F. albida* needs up to 12 years to flower (Barnes and Fagg, 2003), so breeding requires a long-term commitment on the part of both researchers and funding agencies. But, given the importance of *F. albida* in African agroforestry systems, it is an approach worth considering. Previous studies have shown that populations of *F. albida* from east and south Africa grow faster than populations from west Africa, but are also less drought tolerant than west African populations (Vandenbeldt, 1992; Dangasuk et al., 2006). This suggests that populations derived from crosses between Cameroonian and east/south Africa populations could be an important source of novel, and potentially useful, varieties with improved growth and stress-tolerance attributes.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

The visit to Penn State University by Dr. Tchatchoua to conduct this research was funded by the Fulbright Visiting Scholar Fellowship program. The financial support for this work was provided by the Department of Biology at the University of Pennsylvania, The Schatz Center of Tree Molecular Genetics at Pennsylvania State University, and through the USDA National Institute of Food and Agriculture Federal Appropriations under Project PEN04532 and Accession number 1000326. All financial sources are gratefully acknowledged.

REFERENCES

Abdelazim IM, Fagg CW, Harris SA (1997). Seed and seedling variation amongst provenances in *Faidherbia albida.* Forest Ecology and Management 97(2):197-205.

Alexander DH, Novembre J, Lange K (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Research 19(9):1655-1664.

Alkimim ER, Caixeta ET, Sousa TV, Resende MDV, da Silva FL, Sakiyama NS, Zambolim L (2020). Selective efficiency of genome-wide selection in *Coffea canephora* breeding. Tree Genetics and Genomes 16(3).

Barnes RD, Fagg CW (2003). *Faidherbia albida*. Monograph and Annotated Bibliography. Tropical Forestry Papers No. 41, Oxford Forestry Institute, Oxford, U.K.

Benoit L, Born C, Vignes H, Chevallier MH, Todou G, Debain C, Joly HI (2011). Polymorphic microsatellite loci from *Dacryodes edulis* (Burseraceae), a Central African rainforest and fruit-tree species. American Journal of Botany 98(4):74-75.

Boffa JM (1999). Agroforestry parklands in sub-Saharan Africa (No. 34). Food and Agriculture Organization Conservation Guide 230p.

Chang Y, Liu H, Liu M, Liao X, Sahu SK, Fu Y, Song B, Cheng S, Kariba R, Muthemba S, Hendre PS, Mayes S, Ho WK, Yssel AEJ, Kendabie P, Wang S, Li L, Muchugi A, Jamnadass R, Lu H, Peng S, Van Deynze A, Simons A, Yana-Shapiro HYves Van de Peer Y, Xu X, Yang H, Wang J, Liu X (2018). The draft genomes of five agriculturally important African orphan crops. *GigaScience* 8:1-16.

Dangasuk OG, Gudu S, Aman RA, Brandl T (2002). Isozyme variation and growth performance in 16 provenances of African *Faidherbia albida* (Del.) A. Chev. in Kenya." Discovery and Innovation (special edition) April: 26-36. ISSN 1015-079X

Dangasuk GO, Odago PO, Were BA, Guru S (2006). Geographical characterisation of African provenances of *Faidherbia albida*. African Crop Science Journal 14(4):263-273.

Duncan DB (1955). Multiple range and multiple F tests. Biometrics 11(1):1-42.

Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479-491.

FAO (1974). FAO Panel of Experts in Forest Gene Resources Report of the Third Session. Rome, Italy 90p.

Franzel S, Jaenicke H, Janssen W, Ayuk E, Boland D (1996). Choosing the right trees: Setting priorities for multipurpose tree improvement. ISNAR Research Report 8:87p.

Garcia F, Noyer JL, Risterucci AM, Chevallier MH (2004). Genotyping of Mature Trees of *Entandrophragma cylindricum* with Microsatellites. Journal of Heredity 95(5):454-457.

Gautier D, Mana J, Rocquencourt A, Tapsou, Njiti C (2002). Faut-il poursuivre l'opération Faidherbia du DPGT au Nord Cameroun**.** Savanes africaine: des espaces en mutation, des acteurs face à de nouveaux défis. Actes du colloque, mai 2002, Prasac, Maroua, Cameroun, N'Djamena, Tchad 9 p. ⟨hal-00133790⟩

Grattapaglia D (2017). Status and perspectives of genomic selection in forest tree breeding. In Genomic selection for crop improvement. Springer, Cham pp. 199-249.

Guries RP, Ledig TF (1981). Genetic Structure of Populations and Differentiation in Forest Trees. Symposium on Isozymes of North American Forest Trees and Forest Insects, July 27, 1979, Berkeley, California. Pacific Southwest Forest and Range Experiment Station (Berkeley, Calif.). General Technical Report PSW-48:42-47.

Harris SA, Fagg CW, Barnes RA (1997). Isozyme variation in *Faidherbia albida* (Leguminosae Mimosoideae). Plant Systematic and Evolution 207:119-132.

Hauser TP (1994). Germination, predation and dispersal of *Acacia albida* seeds. Oikos 71:421-426.

Hedrick PW, Kalinowski ST (2000). Inbreeding depression in conservation biology. Annual Review of Ecology and Systematics 31:139-162.
Joly HI, Zeh-Nlo M, Danthu P, Aygalent C (1992). Population genetics of an African Acacia: *Acacia albida*. Genetic diversity of populations from West Africa. Australian Journal of Botany 40:59-73.
Kumar S, Tamura K, Nei M (1994). MEGA: Molecular evolutionary genetics analysis software for microcomputers. Bioinformatics 10(2):189-191.
Muona O (1990). Population genetics in forest tree improvement. Plant population genetics, breeding, and genetic resources pp. 282-298.
Muchugi A, Kadu K, Kindt R, Kipruto H, Lemurt S, Olale K, Nyadoi P, Dawson I, Jamnadass R (2008). Molecular Markers for Tropical Trees. ICRAF Technical Manual 9:91p.
Omondi SF, Kireger E, Dangasuk OG, Chikamai B, Odee DW, Cavers S, Khasa DP (2010) Genetic Diversity and Population Structure of *Acacia senegal* (L) Willd. in Kenya. Tropical Plant Biology 3(1):59-70.
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.
Peltier R, Eyog-Matig O (1988). Les essais d'agroforesterie au Nord-Cameroun. Bois et Forets des Tropiques 217(3):3-31.
Peltier R (1996). Les parcs à Faidherbia. Cahier scientifique n12, CIRAD, Montpellier, France 312p.
Ribaut JM, De Vicente MC, Delannay X (2010). Molecular breeding in developing countries: Challenges and perspectives. Current Opinion in Plant Biology 13(2):213-218.
Roupsard O, Ferhi A, Granier A, Pallo F, Depommier D, Mallet B, Joly HI, Dreyer E (1999). Reverse phenology and dry-season water uptake by *Faidherbia albida (*Del.) A. Chev. in an agroforestry parkland of Sudanese West Africa. Functional Ecology 13(4):460-472.
Russell JP, Hedley PE, Dancey LCS, Morris J, Booth A, Odee D, Mwaura L, Omondi W, Angaine P, Machua J, Muchugi A, Iain Milne, Kindt R, Jamnadass R, Dawson Ian K (2014). TropiTree: An NGS-Based EST-SSR Resource for 24 Tropical Tree Species. PLoS ONE 9(7):e102502.
Sniezko RA, Stewart HTL (1989). Range-wide provenance variation in growth and nutrition of *Acacia albida* seedlings propagated in Zimbabwe. Forest Ecology and Management 27:179-197.
Tchatchoua DT, Kolyang MG, Caspa RG, Basga SD, Youri A (2019). Variation in seed and seedling traits of *Faidherbia albida* populations in the Sudano-Sahelian zone of Cameroon. International Journal for Biological and Chemical Sciences 13(4):2029-2040.
Tchoundjeu Z, Weber J, Guarino L (1998). Germplasm collections of endangered agroforestry tree species: The case of *Prosopis africana* in the semi-arid lowlands of West Africa. Agroforestry Systems 39:91-100.
van Nocker S, Gardiner SE (2014). Breeding better cultivars, faster: Applications of new technologies for the rapid deployment of superior horticultural tree crops. Horticulture Research 1(1):1-8.
Vandenbeldt RJ (1991). Rooting systems of western and southern African *Faidherbia albida* (Del.) A. Chev. (syn. *Acacia albida* Del.)--A comparative analysis with biogeographic implications. Agroforestry Systems 14:233-244.
Vandenbeldt RJ (1992). *Faidherbia albida* in the West African semi-arid tropics: Proceedings of a Workshop, 22-26 Apr 1991, Niamey, Niger. World Agroforestry Centre 212 pp. ISBN: 92-9066-220-4
Varshney RK, Graner A, Sorrells ME (2005). Genic microsatellite markers in plants: Features and applications. TRENDS in Biotechnology *23*(1):48-55.
Watson D (2019). Adaption to Climate Change: Climate Adaptive Breeding of Maize, Wheat and Rice. In Sustainable Solutions for Food Security. Springer, Cham pp. 67-89.
Wickens GE (1969). A study of *Acacia albida* DEL. (Mimosoideae). Kew Bull 23:181-202.
World Agroforestry Centre (2009). PRESS RELEASE: Unique Acacia Tree Could Nourish Soils and Life in Africa. http://www.worldagroforestrycentre.org/newsroom/press-releases/press-releaseunique-acacia-tree-could-nourish-soils-and-life-africa.
Wright JW (1962). Genetics of forest tree improvement. Genetics of forest tree improvement 399 pp.
Zargar SM, Raatz B, Sonah H, Bhat JA, Dar ZA, Agrawal GK, Rakwal R (2015). Recent advances in molecular marker techniques: Insight into QTL mapping, GWAS and genomic selection in plants. Journal of crop science and biotechnology 18(5):293-308.
Zobel BJ, Talbert JT (1984). Applied forest tree improvement. Wiley and Sons, Inc., New York 505p.