

Table S1 Measured seven quantitative and four qualitative morphological traits

Quantitative traits	Code	Unit (SI)
Trunk Circumference	TC	cm
Petiole length	PL	cm
Limb length	LL	cm
Limb width	LW	cm
Nut length	NL	cm
Nut width	NWD	cm
Nut weight	NWG	g
Qualitative traits	Code	Modalities
Tree Growth Habit	TGH	Erect, semi-erect, spreading
Leaf Apex Shape	LAS	Acute, acuminate, retuse
Adult Leaf Color	ALC	Light green, green, dark green
Seed Coat Color	SEES	Spheroid, ellipsoid, oval, ovoid

Table S2 Results of Bayesian model-based clustering implemented in Structure of 333 *Vitellaria paradoxa* accessions

Cluster No	Net nucleotide distance		<i>HE</i>	<i>Mean F_{ST}</i>	Number of	Admixed genotypes
	GP1	GP2				
GP1	0		0.2588	0.0177	167	69
GP2	0.003	0	0.2606	0.0178	95	47
GP3	0.0026	0.0025	0,2615	0.0182	71	38

F_{ST} : fixation index, *HE*: expected heterozygosity

Table S3 Studies using molecular markers to assess the genetic diversity of *V. paradoxa* in the last two decades

Subspecies	Marker	HE	F_{ST}	N	Study area	Gene. diff	Reference
<i>ssp. paradoxa</i>	SNP	0.26	0.004	333	Côte d'Ivoire	Low	Our study
<i>ssp. nilotica</i>	SNP	0.21	0.01	623	Uganda	Low	(28)
<i>ssp. paradoxa</i>	SNP	0.04	0.002	282	Ghana	Low	(50)
<i>ssp. paradoxa</i>	SSR	0.67	0.113	200	Ghana	Moderate	(26)
<i>ssp. nilotica</i>	SSR	0.73		118	Uganda	Low	(25)
<i>ssp. paradoxa</i>	SSR	0.63	0.085	673	West Africa	Low	(24)
<i>ssp. paradoxa</i>	SSR	0.51	0.21	374	Nat. range	High	(5)
<i>ssp. nilotica</i>							
<i>ssp. paradoxa</i>	SSR	0.32	0.047	169	Mali and Côte d'Ivoire	Low	(51)
<i>ssp. paradoxa</i>	RAPD	0.26	0.23	179	Nat. range	High	(19)
<i>ssp. nilotica</i>							
<i>ssp. Paradoxa</i>	RAPD	0.26	0.15	118	Nat. range	High	(18)
<i>Ssp. nilotica</i>							
<i>ssp. paradoxa</i>	SSR	0.41	0.001	441	Mali	Low	(21)

HE: expected heterozygosity, F_{ST} : fixation index, N: number of samples, Gene. diff: genetic differentiation, Nat. range: natural range

Table S4 Mean of different genetic parameters in the core set

Pop	N	Na	Ne	<i>I</i>	<i>HO</i>	<i>HE</i>	<i>uHE</i>	<i>F</i>	PPL
GP1	40.282	1.993	1.393	0.408	0.128	0.255	0.259	0.483	99.29%
GP2	25.483	1.969	1.392	0.402	0.132	0.252	0.258	0.443	96.89%
GP3	18.372	1.938	1.392	0.397	0.128	0.250	0.258	0.448	93.81
Mean	28.046	1.967	1.392	0.402	0.129	0.253	0.258	0.459	96.66%

N: number of samples, *Na*: number of different alleles, *Ne*: number of effective alleles, *I*: Shannon's index, *HO*: observed heterozygosity, *HE*: diversity index, *uHE*: unbiased diversity index, *PPL*: percentage of polymorphic loci

Table S5 Variation in quantitative morphological traits of the entire collection

Traits	Minimum	Mean $\pm$ Sd	Maximum
TC	29.7	144.6 $\pm$ 45.93	266
PL	5.32	8.29 $\pm$ 1.49	13.42
LL	9.67	14.77 $\pm$ 2.31	21.48
LW	2.94	4.61 $\pm$ 0.79	7.16
NL	1.78	3.11 $\pm$ 0.38	4.1
NWD	0.9	2.31 $\pm$ 0.3	2.9
NWG	4.33	9.55 $\pm$ 2.61	19.57

Sd: Standard deviation

Table S6 Factor loadings in the first three factor components

Factor components	PC1	PC2	PC3
Eigenvalues	2.42	1.82	0.87
Variance (%)	34.55	26.03	12.45
Cumulative variance (%)	34.55	60.58	73.03
Trunk circumference	-0.49	0.12	0.78
Petiole length	0.62	-0.38	0.2
Limb length	0.79	-0.41	0.15
Limb width	0.76	-0.38	0.07
Nut length	0.49	0.77	-0.03
Nut width	0.48	0.74	-0.12
Nut weight	0.38	0.45	0.42

PC1, PC2, PC3: principal components 1; 2; 3

Table S7 Quantitative and qualitative traits associated to the identified *Vitellaria paradoxa* groups.

Quantitative Traits		Means $\pm$ standard deviation			F-	p-
		GP1 (N=75)	GP2 (N=40)	GP3 (N=45)	value	value
Trunk Circumference (cm)		133.4 $\pm$ 49.79a	160.5 $\pm$ 38.81b	149.1 $\pm$ 41.29ab	5.04	< 0.01
Petiole length (cm)		8.23 $\pm$ 1.47	8.45 $\pm$ 1.46	8.26 $\pm$ 1.59	0.28	0.755
Limb length (cm)		15.01 $\pm$ 2.26	14.54 $\pm$ 2.41	14.58 $\pm$ 2.32	0.76	0.468
Limb width (cm)		4.8 $\pm$ 0.78a	4.46 $\pm$ 0.86ab	4.44 $\pm$ 0.67b	4.16	< 0.05
Nut length (cm)		3.09 $\pm$ 0.45	3.15 $\pm$ 0.29	3.08 $\pm$ 0.3	0.44	0.647
Nut width (cm)		2.38 $\pm$ 0.38	2.26 $\pm$ 0.2	2.34 $\pm$ 0.2	2.27	0.107
Nut weight (g)		10.62 $\pm$ 2.7a	8.89 $\pm$ 2.14b	8.36 $\pm$ 2.17b	14.11	< 0.001
Qualitative traits	Modalities	Proportions N(%)			χ^2	p-
		GP1 (N=75)	GP2 (N=40)	GP3 (N=45)		value

Tree	Growth	Erect	19 (25.33)	16 (40)	20 (44.44)	38.78	< 0.001
Habit (TGH)		Semi-erect	46 (61.33)	8 (20)	6 (13.33)		
		Spreading	10 (13.33)	16 (40)	19 (42.22)		
Leaf	Apex	Acute	20 (26.67)	9 (22.5)	18 (40)	31.93	< 0.001
Shape (LAS)		Acuminate	20 (26.67)	1 (2.5)	3 (6.67)		
		Retuse	26 (34.67)	10 (25)	13 (28.89)		
		Obtuse	9 (12)	20 (50)	11 (24.44)		
Adult	Leaf	Light green	1 (1.33)	5 (12.5)	7 (15.56)	11.9	< 0.05
Color (ALC)		Green	65 (86.67)	33 (82.5)	35 (77.78)		
		Dark green	9 (12)	2 (5)	3 (6.67)		
Seed	Coat	Creamish	13 (17.33)	0 (0)	0 (0)	25.37	< 0.01
Color (SCC)		Dull brown	3 (4)	0 (0)	0 (0)		
		Brown	31 (41.33)	18 (45)	29 (64.44)		
		Pale brown	4 (5.33)	7 (17.5)	4 (8.89)		
		Dark brown	24 (32)	15 (37.5)	12 (26.67)		
Seed	shape	Spheroid	0 (0)	1 (2.5)	3 (6.67)	33.39	< 0.001
(SEES)		Ellipsoid	24 (32)	4 (10)	9 (20)		
		Oval	40 (53.33)	14 (35)	11 (24.44)		
		Ovoid	11 (14.67)	21 (52.5)	22 (48.89)		

Values in brackets represent the percentage of modalities for each qualitative trait.

Table S8 Variation in quantitative morphological traits of the core set

Traits	Minimum	Mean $\pm$ Sd	Maximum
TC	65.5	149.7 $\pm$ 44.34	242
PL	5.32	8.12 $\pm$ 1.61	13.42
LL	10.96	14.33 $\pm$ 2.13	19.62
LW	2.94	4.44 $\pm$ 0.78	6.2
NL	1.78	3.08 $\pm$ 0.42	3.66
NWD	0.9	2.26 $\pm$ 0.41	2.76
NWG	4.33	10.09 $\pm$ 2.73	17.93

Sd: Standard deviation

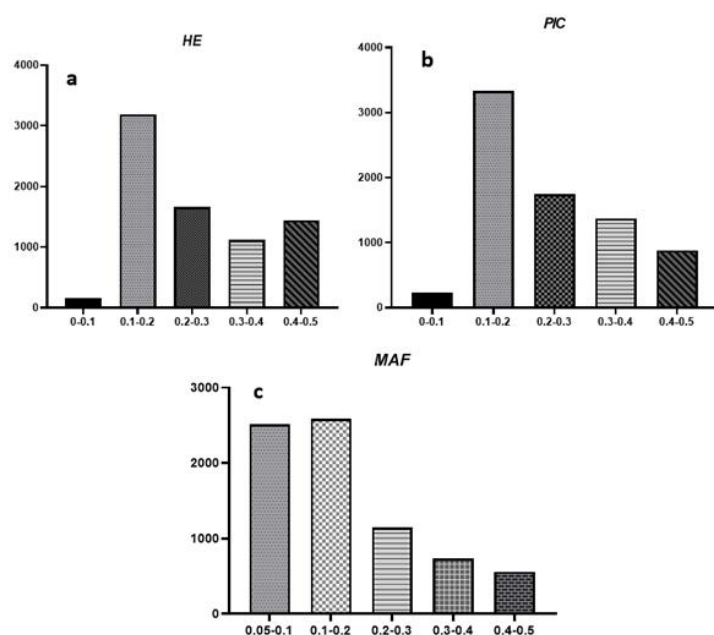


Fig. S1 Distribution of genetic diversity for the 7,559 filtered SNP markers from the 333 *Vitellaria paradoxa* genotypes. a: gene diversity (*HE*); b: Polymorphism information content (*PIC*); c: Minor allele frequency (*MAF*). X-axis indicates *HE*, *PIC* and *MAF* value classes; Y-axis indicates the number of SNP markers.

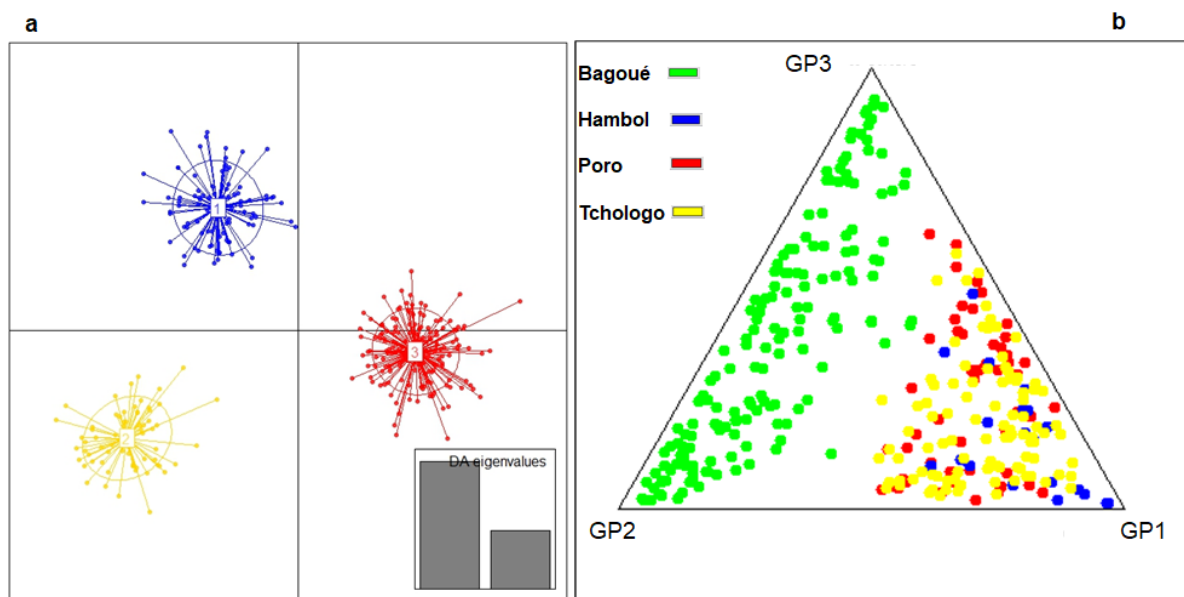


Fig. S2 (a) DAPC showing the genetic networks for the three genetic groups, with node sizes indicating genetic relationships between different accessions; (b) Structure software triangular plot showing the distribution of the four populations in the three genetic groups; dot colors represent the origin of the genotypes (green for Bagoué, blue for Hambol, red for Poro, and yellow for Tchologo).

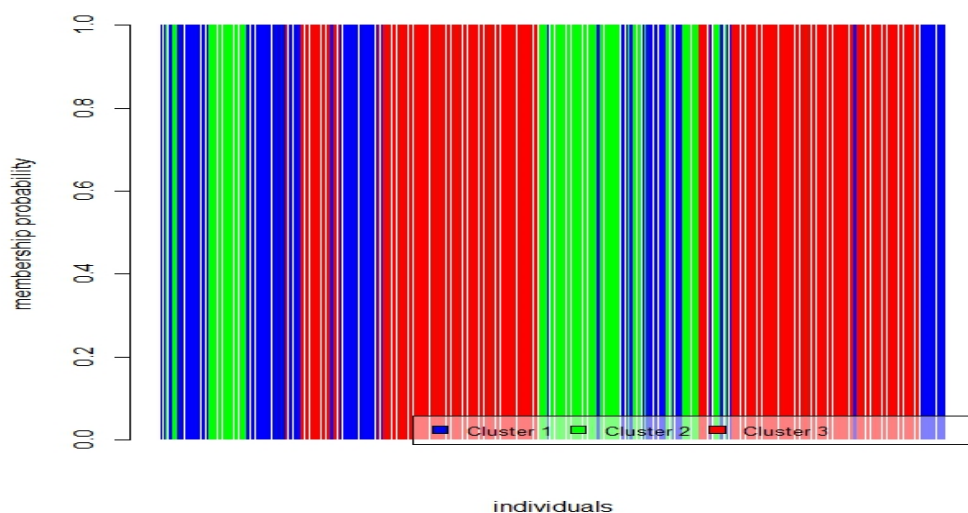


Figure S3: DAPC population structure plot showing no admixture individuals of 333 shea tree accessions, $K=3$; each color represents one cluster. X-axis represents individuals and Y-axis represents the membership probability.

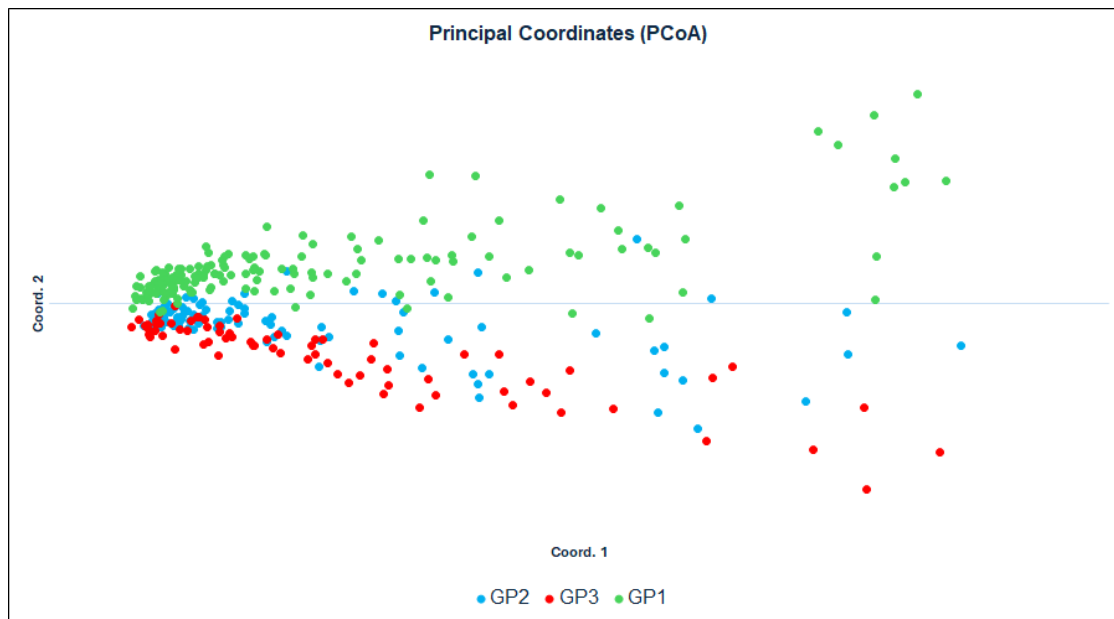


Fig. S4 Scatter plot of the 333 shea trees based on the genetic distance along the first two principal coordinates PCoA 1 and PCoA 2. The colors of the dots along the principal coordinate analysis (PCoA) coordinates 1 and 2 represent the different groups.

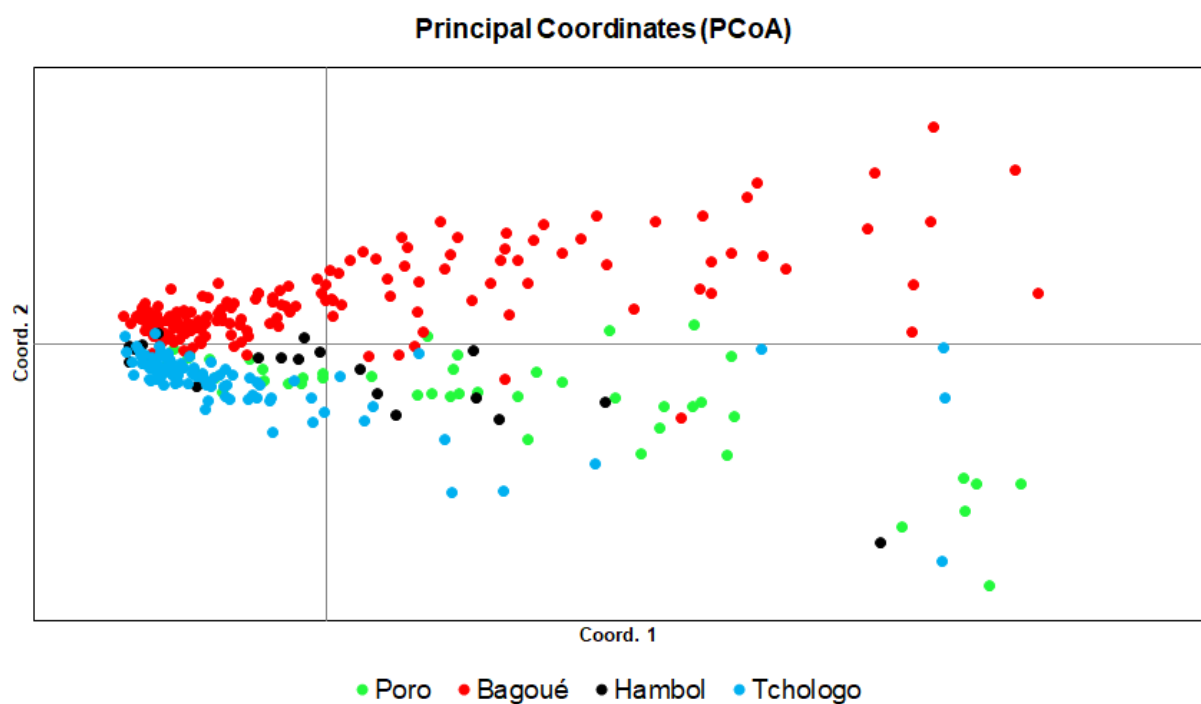


Fig. S5 Scatter plot of the 333 shea trees based on the genetic distance along the first two principal coordinates PCoA 1 and PCoA 2. The colors of the dots along the principal coordinate analysis (PCoA) coordinates 1 and 2 represent shea trees from different locations..

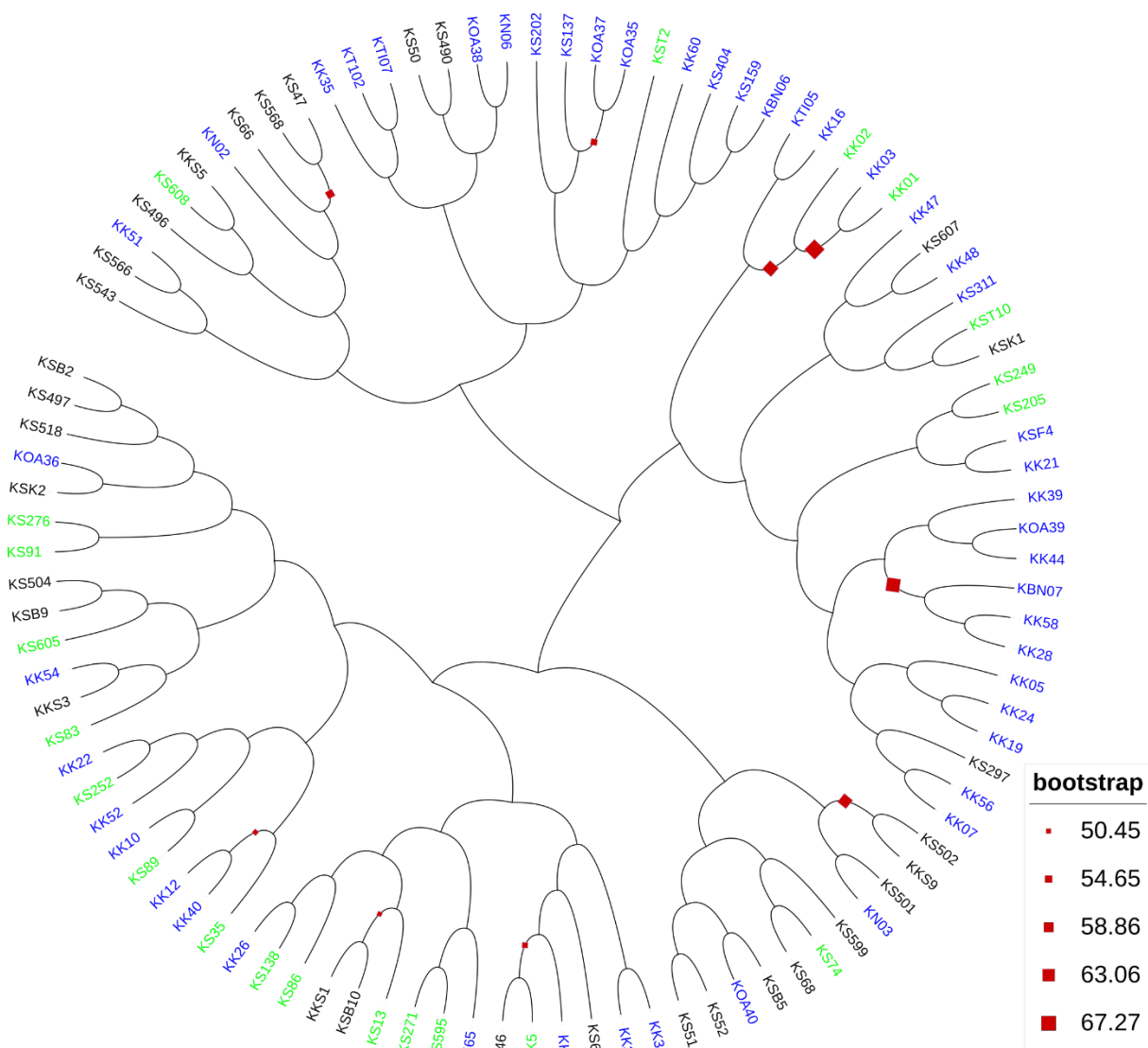


Fig. S6: Neighbor-joining phylogenetic tree with bootstrap values (red squares represent the bootstrap values; 1,000 replicates) showing relationships between the 100 core set of individual superior shea trees based on Nei's genetic distance matrix using 7,559 SNP markers; the color of the labels indicates the genetic group to which a sample belongs (blue for GP1 samples, black for GP2 samples, and green for GP3 samples); analyses were carried out in R, and the generated tree was customized using iTOL

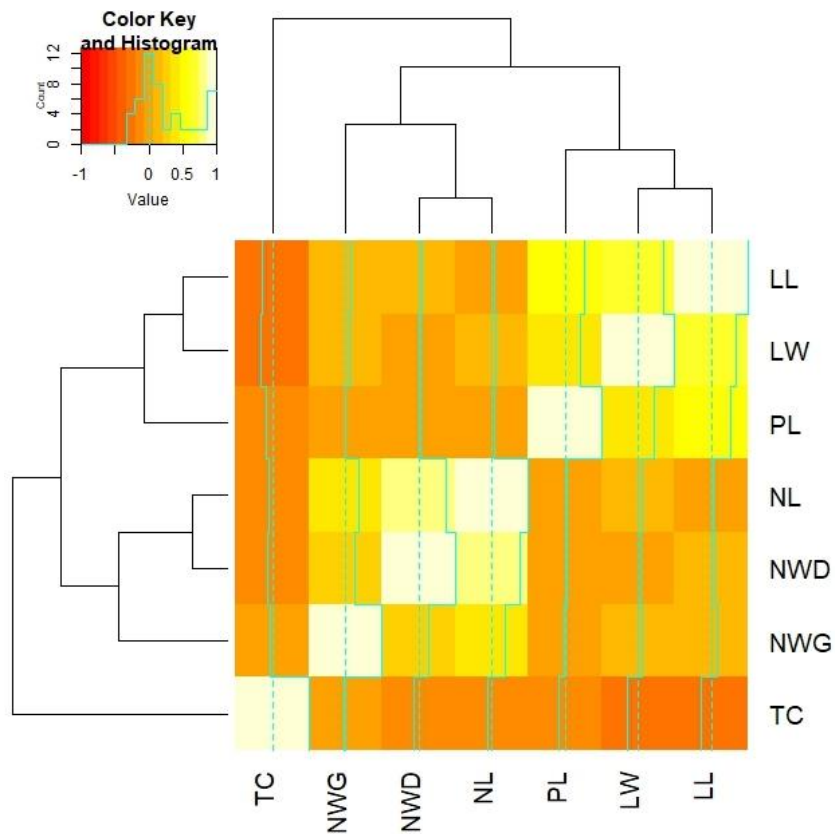


Fig. S7 Pearson correlation heat map showing relationships between morphological traits measured on elite shea trees in Bagoué, Hambol, Poro and Tchologo districts in Northern Côte d'Ivoire. TC: Circumference of trunk at 130 cm above the soil; NWG: Nut weight; NWD: Nut width; NL: Nut length; PL: Petiole length; LW: Limb width; LL: Limb length.