Supplementary Material for: Cultural Evolution of Central African hunter-gatherers reflects a deep history of interconnectivity

Cecilia Padilla-Iglesias, Javier Blanco-Portillo, Alexander G. Ioannidis, Andrea Manica, Lucio Vinicius and Andrea Bamberg Migliano

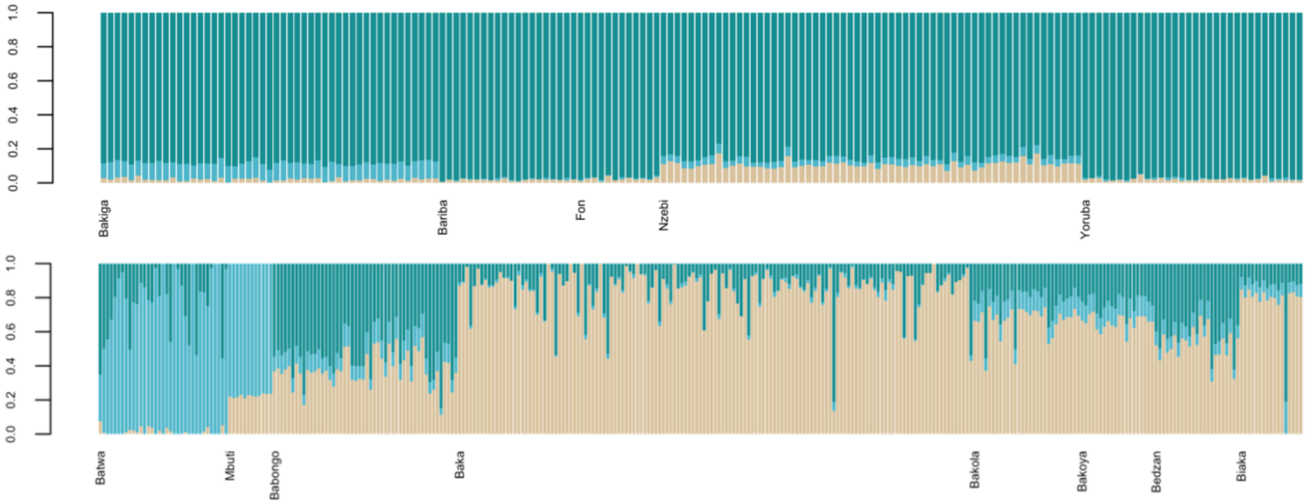


Figure 1. Results from the programme ADMIXTURE on our reduced pruned SNP dataset. Each bar corresponds to an individual and colours represent the proportion of inferred ancestry components from K=3 ancestral populations.

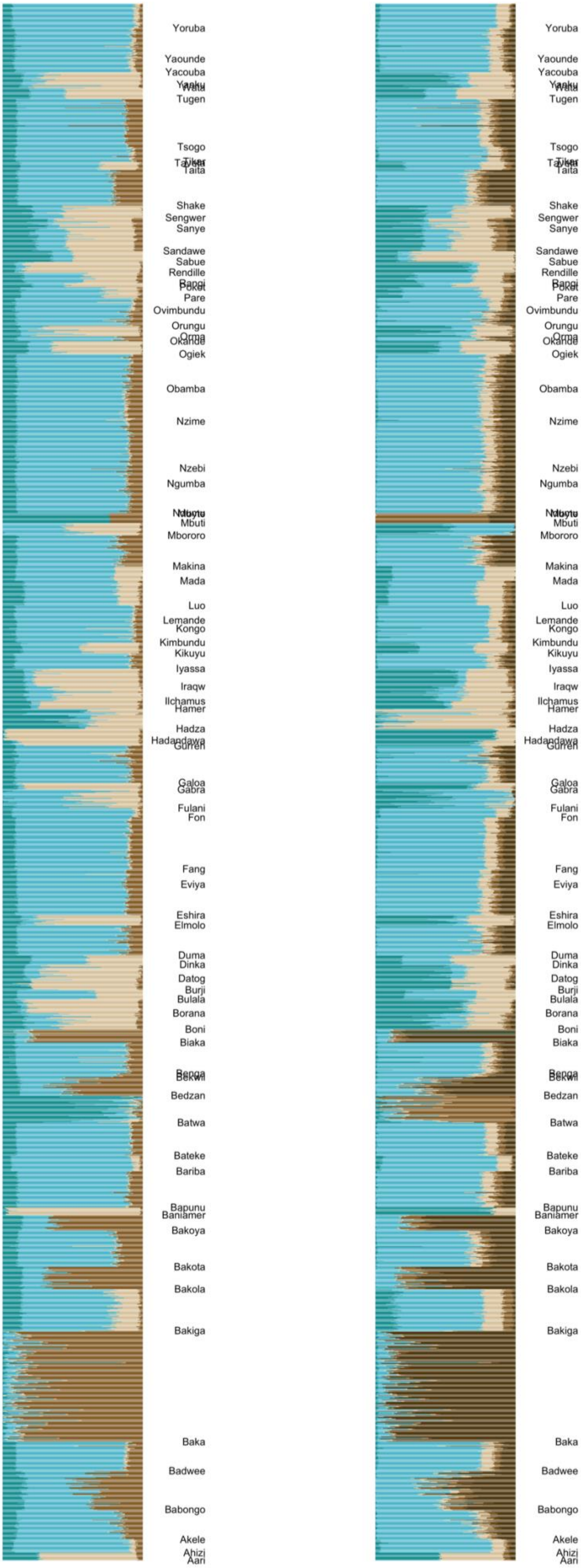


Figure 2. Admixture results on full pruned dataset at K=4 (left) and K=5 (right)

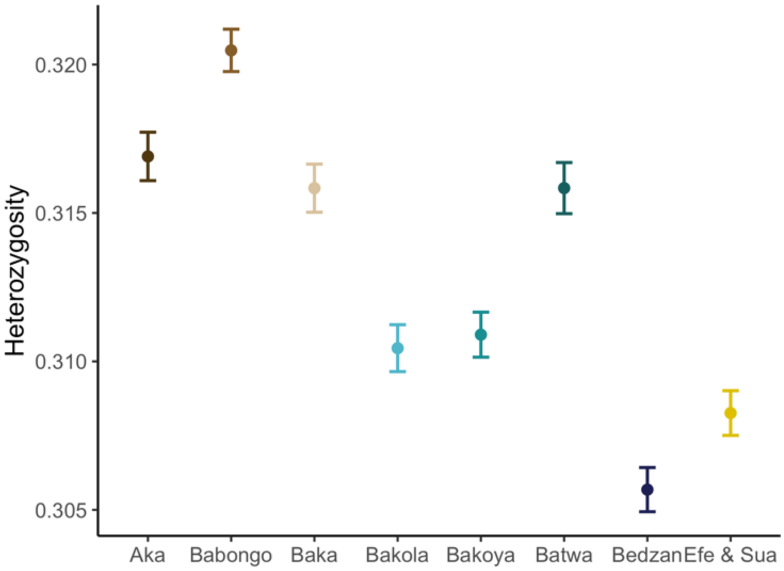


Figure 3. Average heterozygosity proportion of CAHG ancestry components in the populations in our sample.

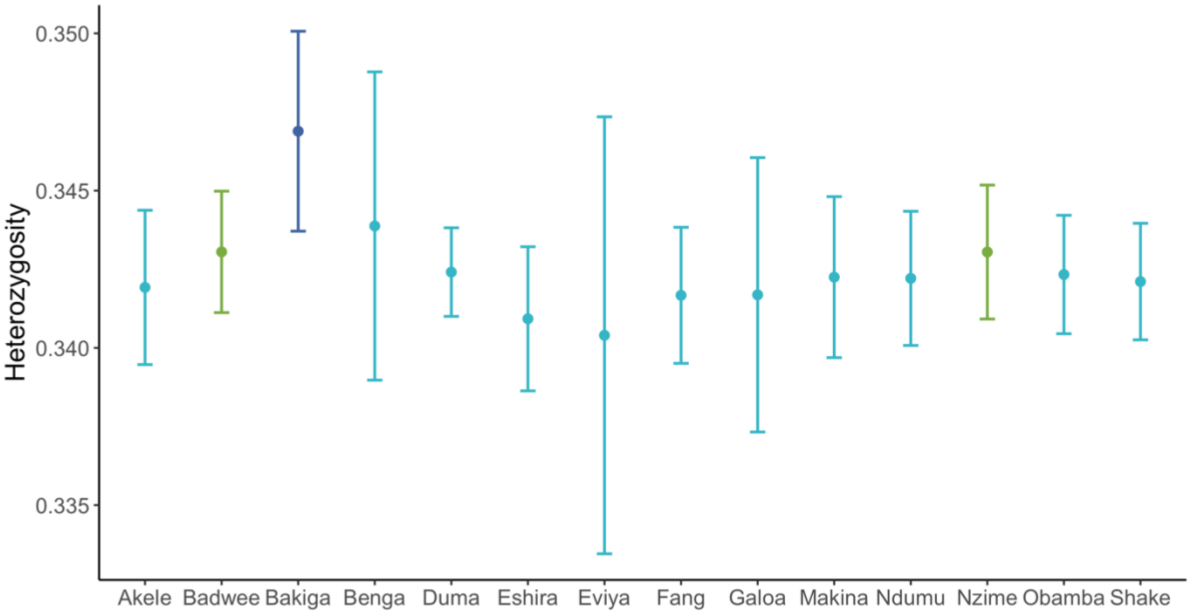


Figure 4. Average heterozygosity of farmer populations neighbouring CAHG in our sample. Dark blue designates Ugandan population, light blue Gabonese populations and green Cameroonian populations.

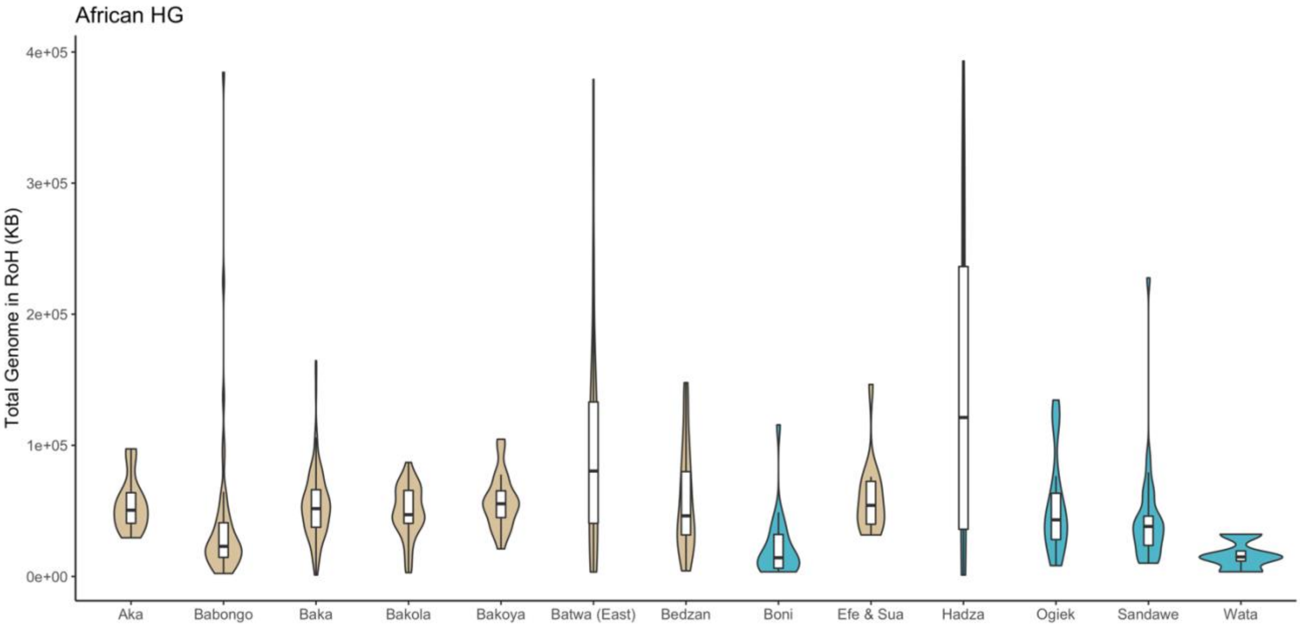


Figure 5. Average length of the genome in runs of homozygosity (RoH) in all African hunter-gatherer populations included in our sample

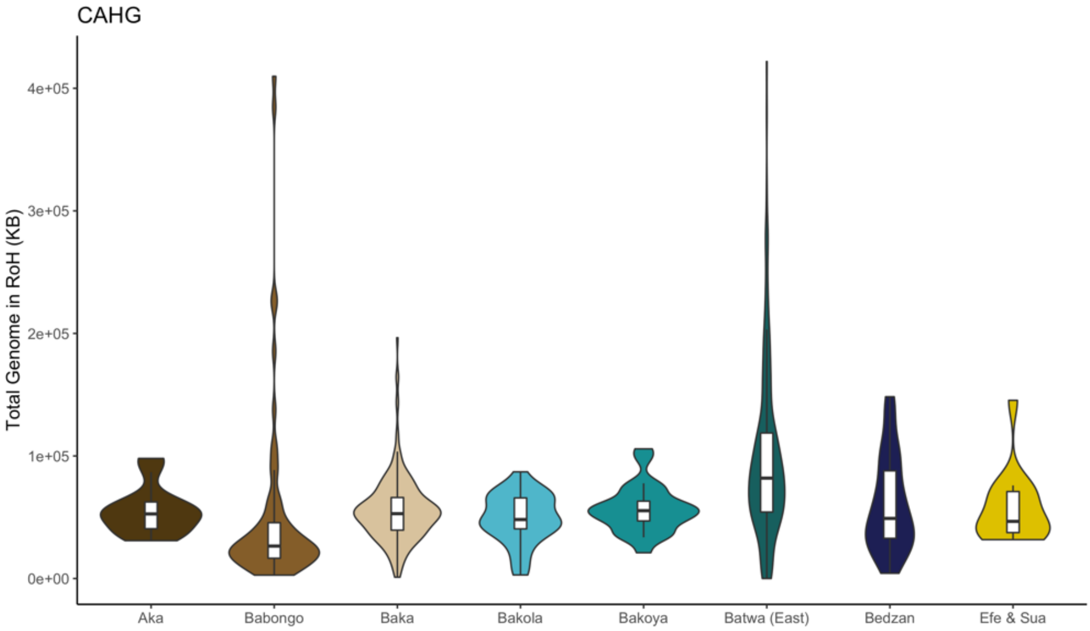


Figure 6. Average length of genome in runs of homozygosity for CAHG populations in our sample

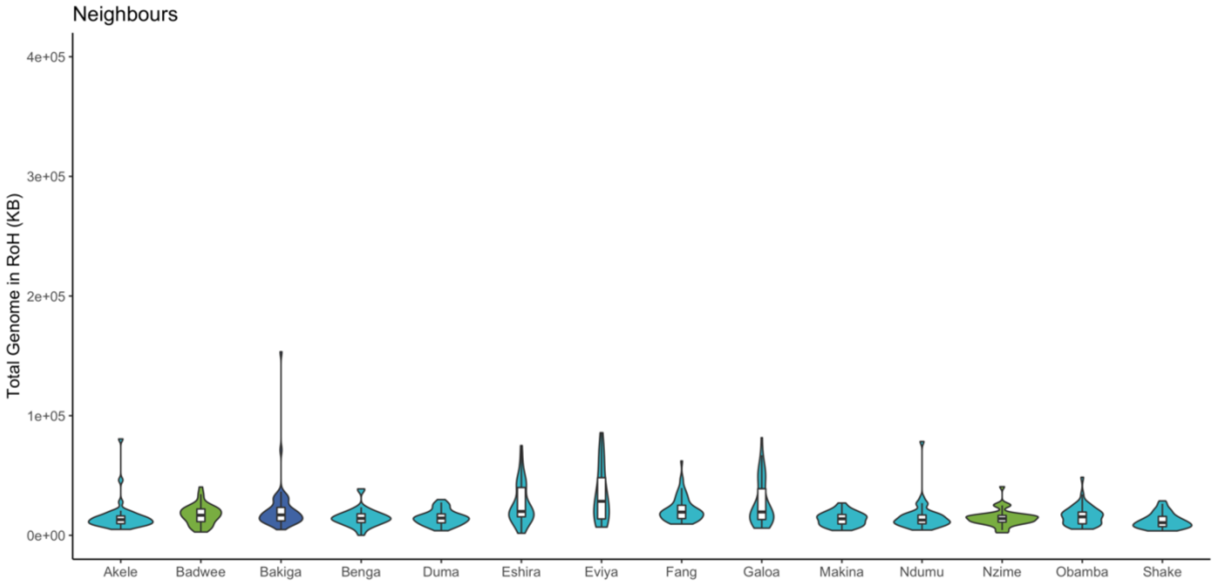


Figure 7. Average length of the genome in runs of homozygosity (RoH) in neighbours of CAHG populations included in our sample.

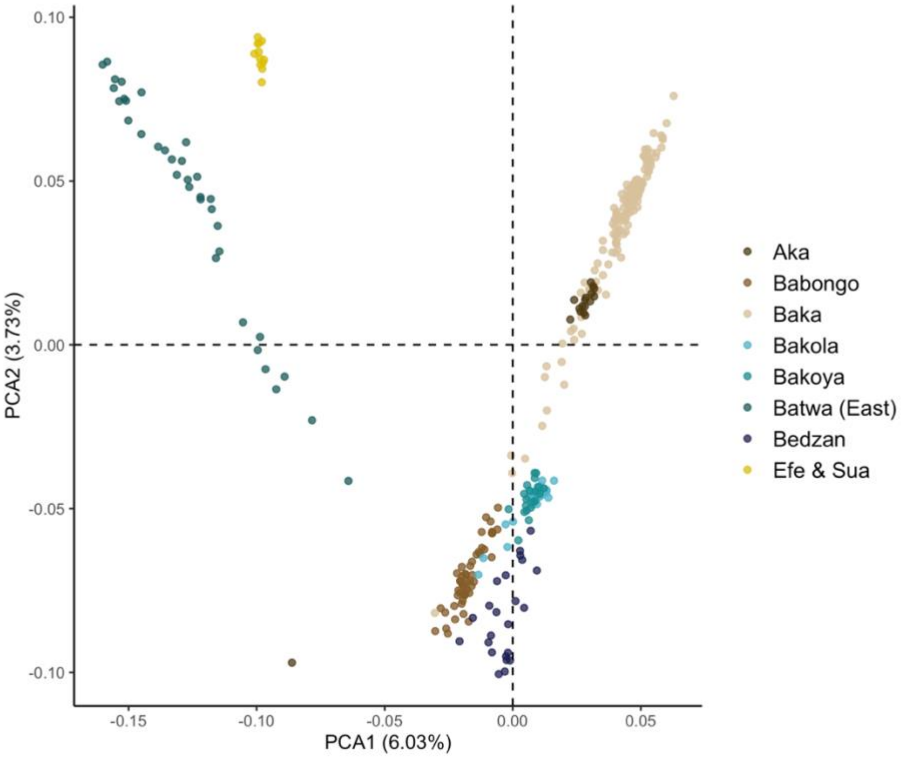


Figure 8. PCA of CAHG with available genome-wide SNP data using the full (unmasked) dataset

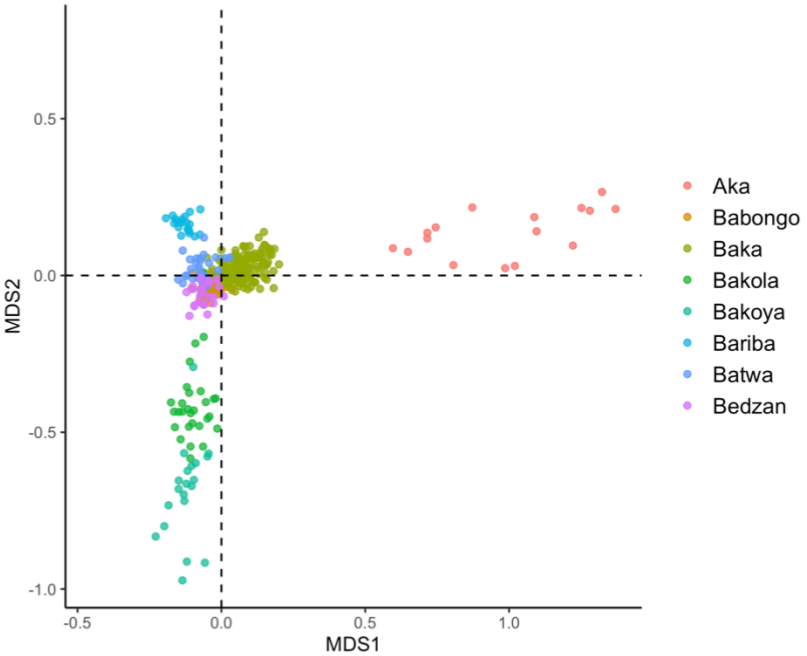


Figure 9. PCA of CAHG with available genome-wide SNP data using exclusively Bantu-associated genomic segments. Note that the Mbuti are not included as they do not carry a sufficient amount of Bantu-associated DNA.

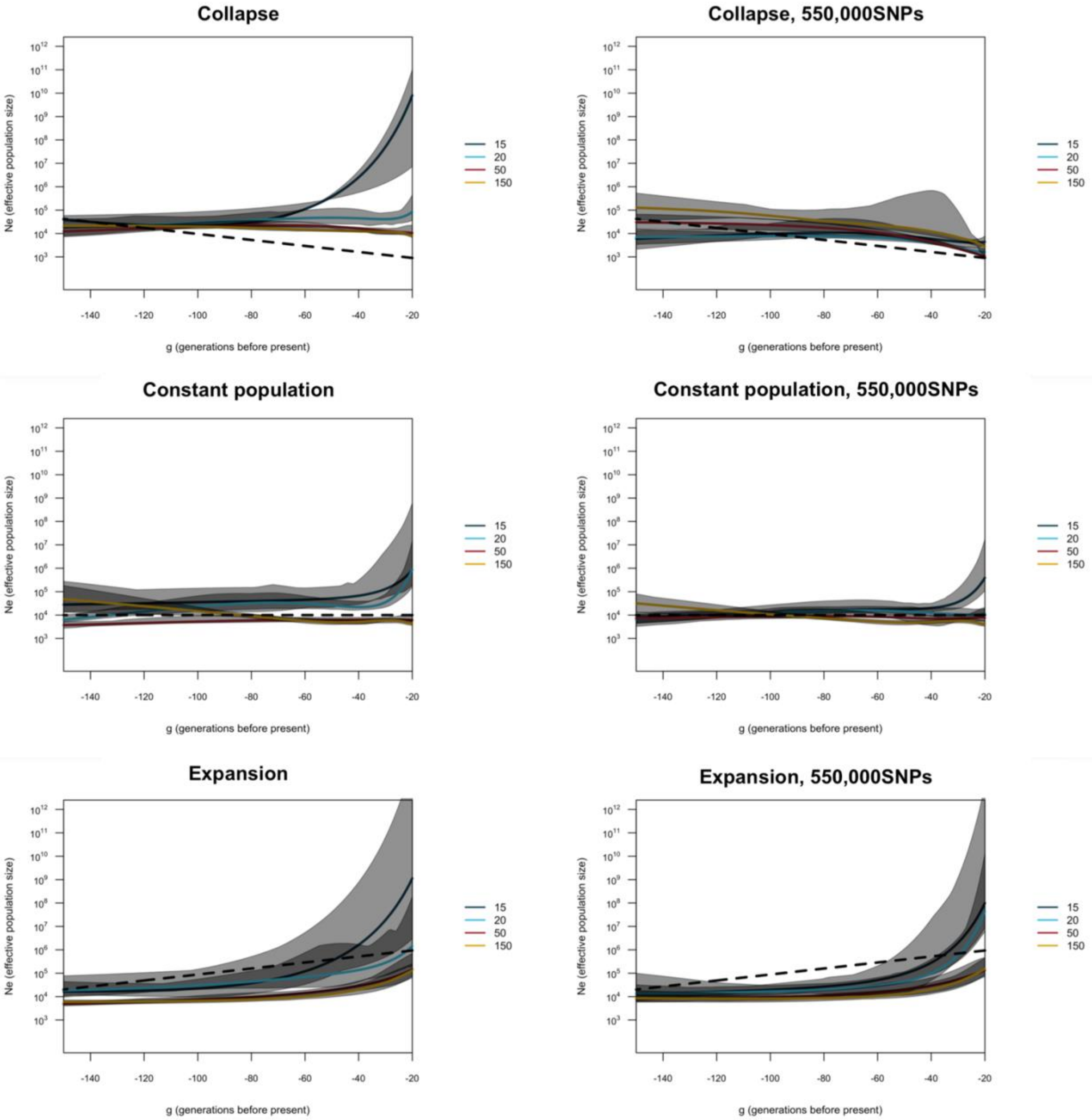


Figure 10. Population size (Ne) estimates from 4,500 years B.P until 600 years B.P inferred from simulated data. Left column shows results from simulated sequence data filtered for alleles with a frequency <10% and right column from simulated SNP array data. Top row shows results from a declining effective population of initially 45,000 individuals declining at a rate of -0.03, middle row from a constant effective population of 10,000 individuals and bottom row from a growing effective population of initially 11,000 individuals at a rate of 0.03. Coloured lines show Ne for simulations containing sample sizes of 15, 20, 50 and 150 individuals. Shaded regions show bootstrap 95% confidence intervals around such estimates and dashed lines show real simulated Ne.

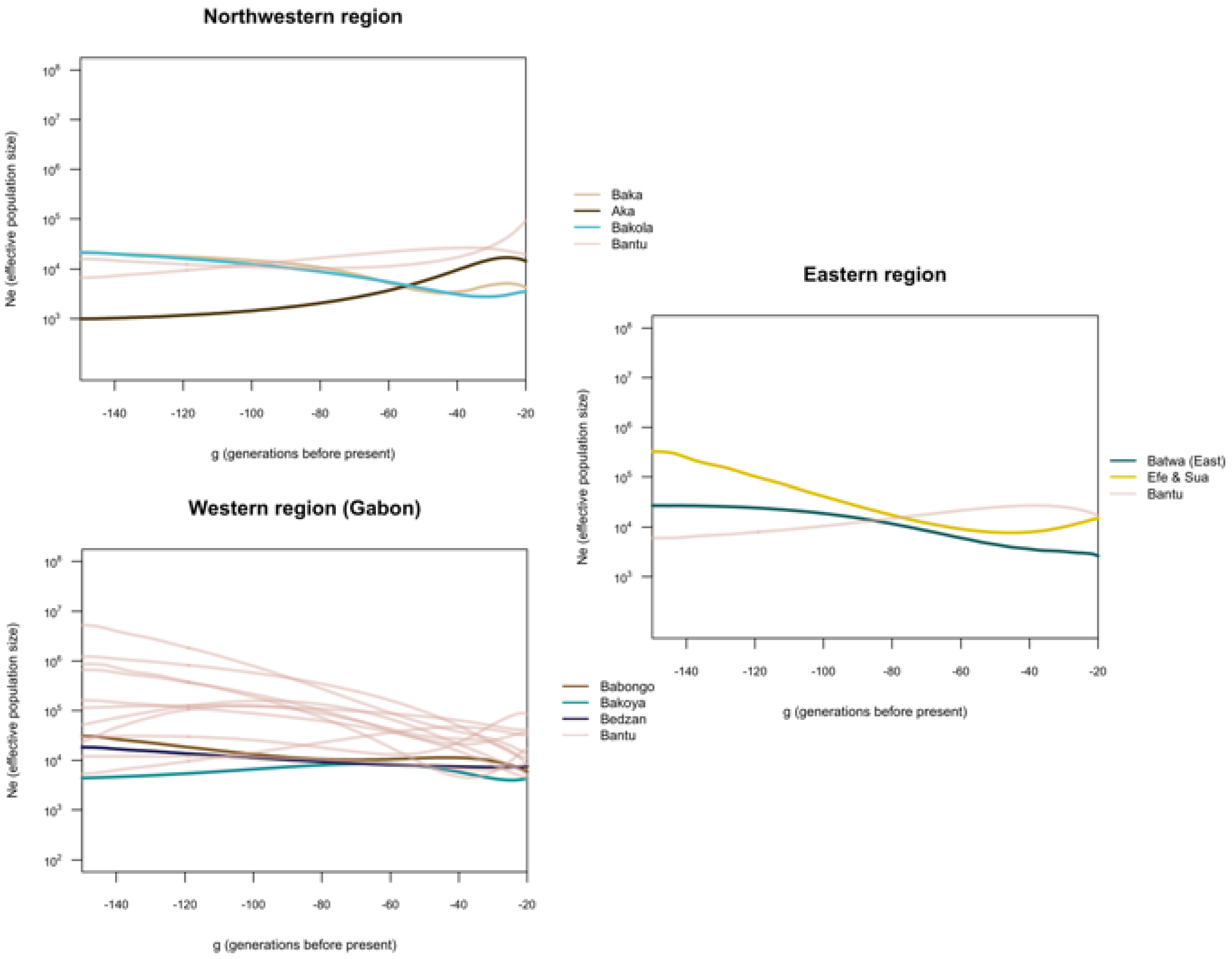


*Figure 11. Effective population size (N_e) estimates for CAHG and Bantu-speaking populations inhabiting the same regions from 4,500 years B.P. until 600 years B.P. (**A**) N_e estimates for northwestern groups include Baka, Aka and Bakola (CAHG) and Bantu-speaking Badwee and Nzime from Cameroon and the Central African Republic. (**B**) N_e estimates for Eastern groups Mbuti and Eastern Batwa (CAHG) and Bantu-speaking population Bakiga from the Eastern part of the Democratic Republic of Congo and Uganda. (**C**) N_e estimates Babongo, Bakoya and Bedzan (CAHG) and Bantu-speaking Akele, Benga, Duma, Eshira, Eviya, Fang, Galoa, Makina, Ndumu, Obamba and Shake from Gabon. Each generation represents 30 years.*

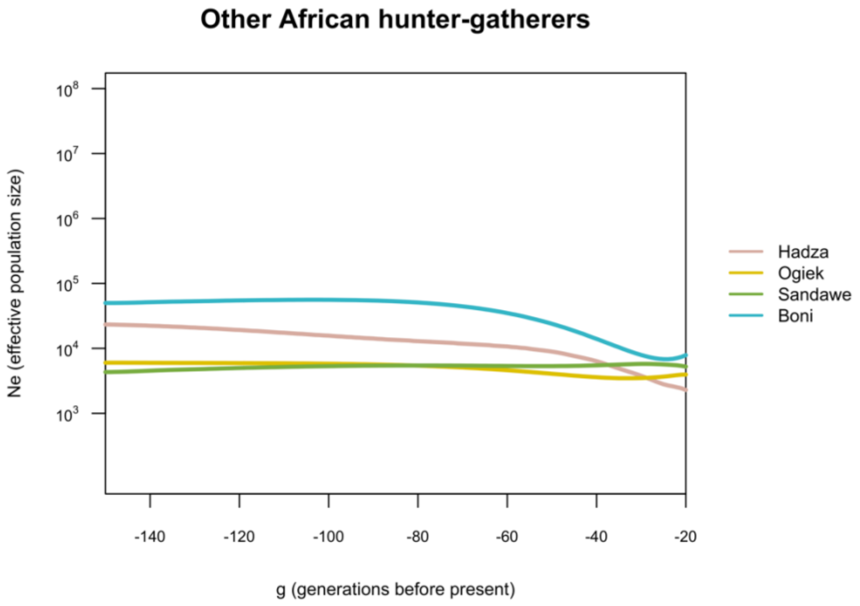
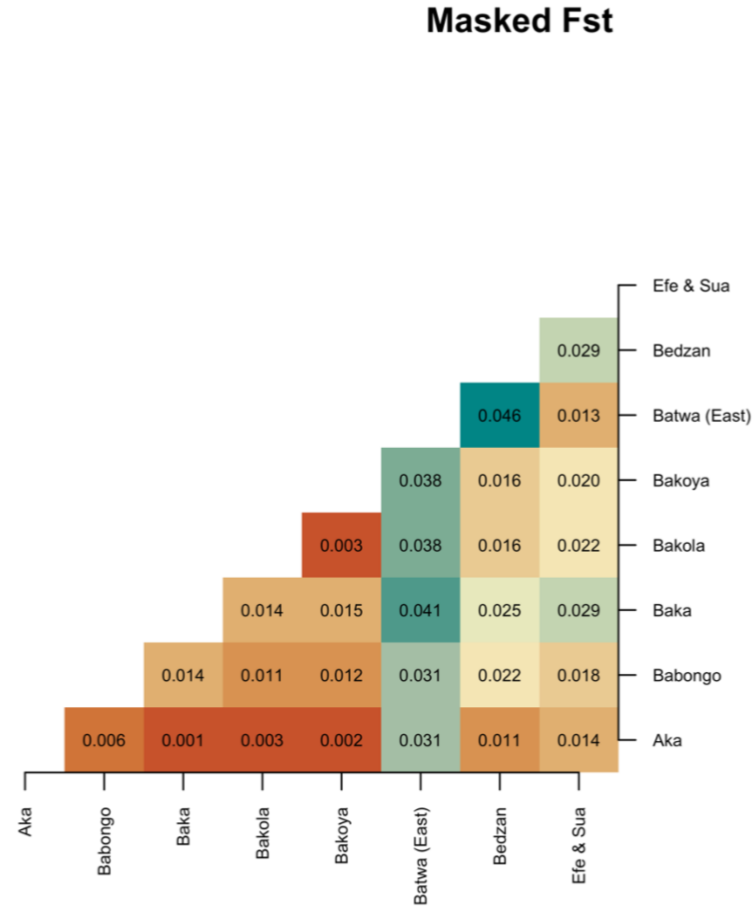
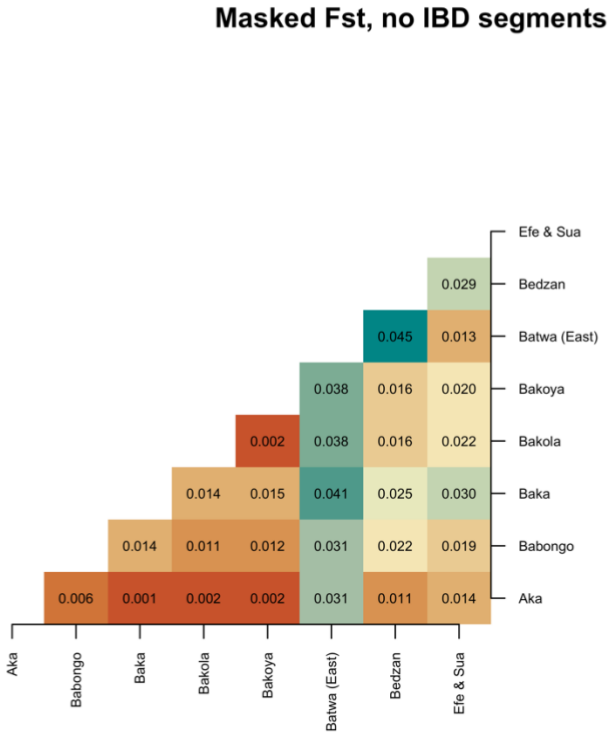


Figure 12. IBDNe results from hunter-gatherer populations from Eastern Africa included in our sample



	Aka	Babongo	Baka	Bakola	Bakoya	Batwa (East)	Bedzan	Efe & Sua
Efe & Sua								
Bedzan								0.029
Batwa (East)							0.046	0.013
Bakoya						0.038	0.016	0.020
Bakola					0.003	0.038	0.016	0.022
Baka				0.014	0.015	0.041	0.025	0.029
Babongo			0.014	0.011	0.012	0.031	0.022	0.018
Aka		0.006	0.001	0.003	0.002	0.031	0.011	0.014

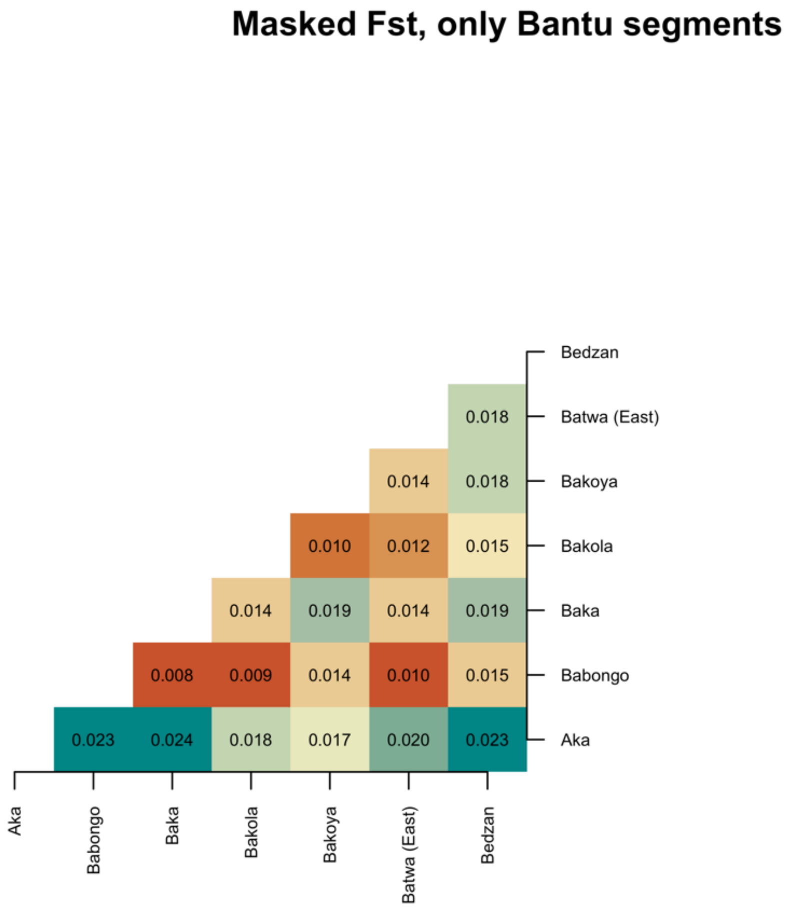
Figure 13. Inter-population F_{ST} using only CAHG ancestry components



Masked Fst, no IBD segments

	Aka	Babongo	Baka	Bakola	Bakoya	Batwa (East)	Bedzan	Efe & Sua
Efe & Sua								
Bedzan								0.029
Batwa (East)							0.045	0.013
Bakoya						0.038	0.016	0.020
Bakola					0.002	0.038	0.016	0.022
Baka				0.014	0.015	0.041	0.025	0.030
Babongo			0.014	0.011	0.012	0.031	0.022	0.019
Aka		0.006	0.001	0.002	0.002	0.031	0.011	0.014

Figure 14. Pairwise Fst matrix using only CAHG ancestry components and excluding shared IBD segments between populations



	Aka	Babongo	Baka	Bakola	Bakoya	Batwa (East)	Bedzan
Bedzan							
Batwa (East)							0.018
Bakoya						0.014	0.018
Bakola					0.010	0.012	0.015
Baka				0.014	0.019	0.014	0.019
Babongo			0.008	0.009	0.014	0.010	0.015
Aka		0.023	0.024	0.018	0.017	0.020	0.023

Figure 15. Pairwise Fst matrix using only Bantu ancestry components

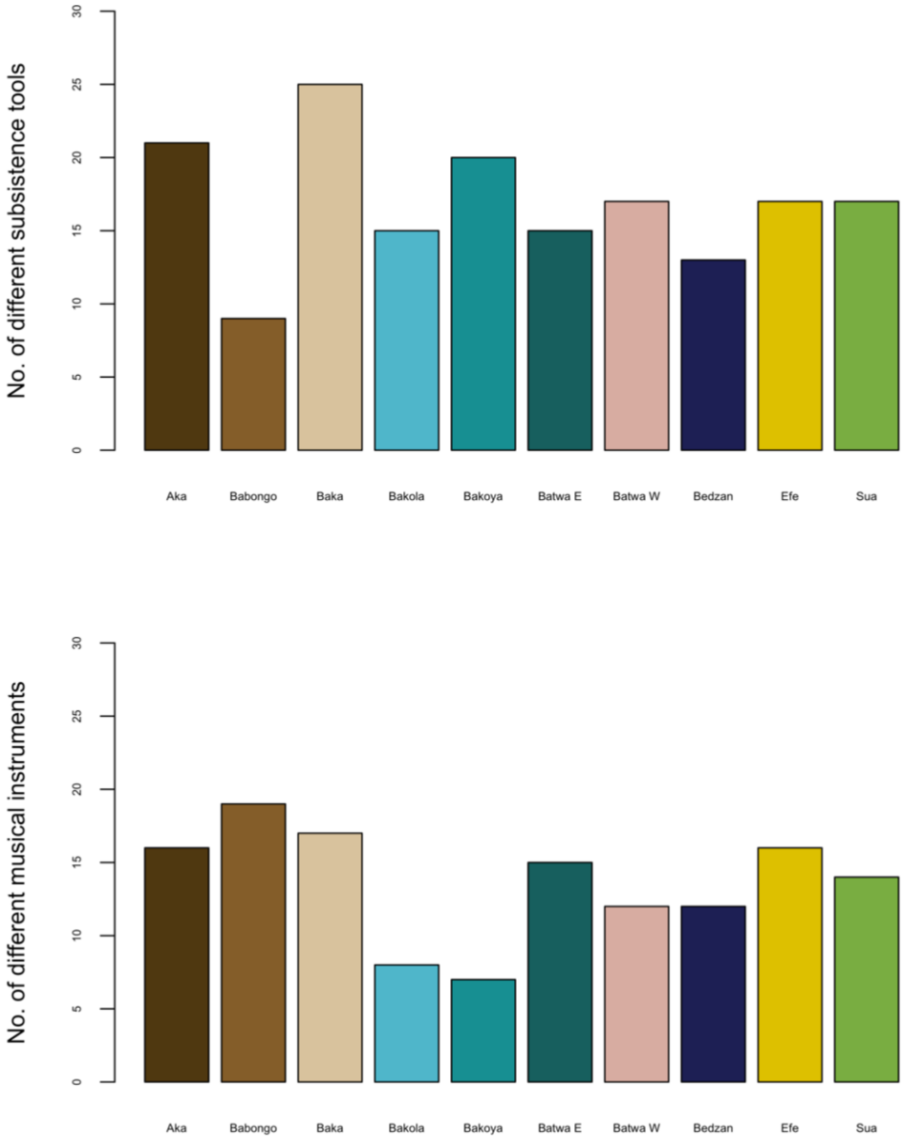
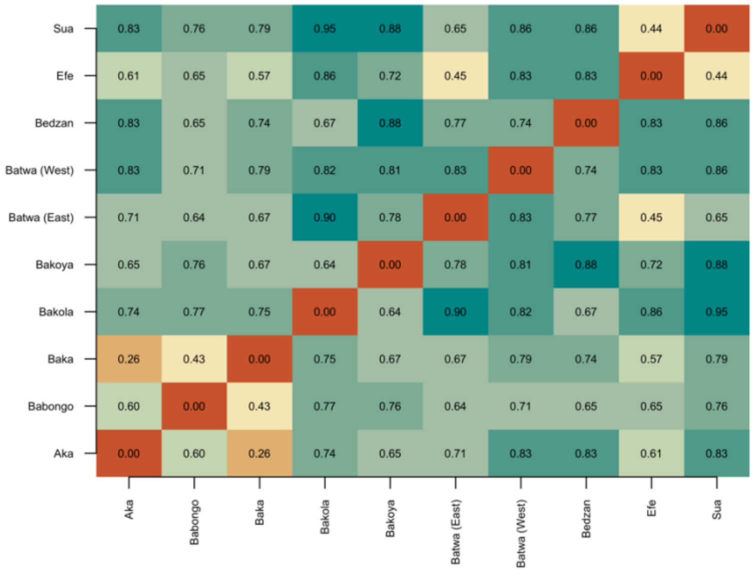


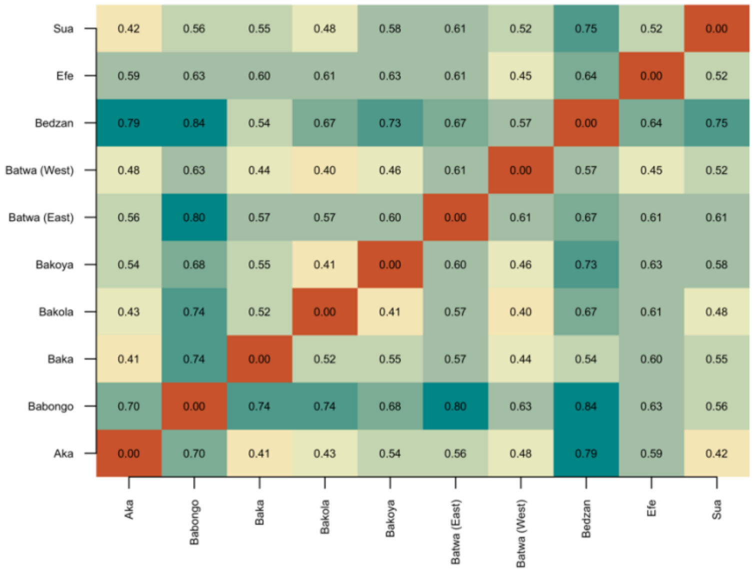
Figure 16. Count of different subsistence tools (top) and musical instruments (bottom) for the 10 CAHG populations in our sample.

Musical Instruments



	Aka	Babongo	Baka	Bakola	Bakoya	Batwa (East)	Batwa (West)	Bedzan	Efe	Sua
Sua	0.83	0.76	0.79	0.95	0.88	0.65	0.86	0.86	0.44	0.00
Efe	0.61	0.65	0.57	0.86	0.72	0.45	0.83	0.83	0.00	0.44
Bedzan	0.83	0.65	0.74	0.67	0.88	0.77	0.74	0.00	0.83	0.86
Batwa (West)	0.83	0.71	0.79	0.82	0.81	0.83	0.00	0.74	0.83	0.86
Batwa (East)	0.71	0.64	0.67	0.90	0.78	0.00	0.83	0.77	0.45	0.65
Bakoya	0.65	0.76	0.67	0.64	0.00	0.78	0.81	0.88	0.72	0.88
Bakola	0.74	0.77	0.75	0.00	0.64	0.90	0.82	0.67	0.86	0.95
Baka	0.26	0.43	0.00	0.75	0.67	0.67	0.79	0.74	0.57	0.79
Babongo	0.60	0.00	0.43	0.77	0.76	0.64	0.71	0.65	0.65	0.76
Aka	0.00	0.60	0.26	0.74	0.65	0.71	0.83	0.83	0.61	0.83

Subsistence tools



	Aka	Babongo	Baka	Bakola	Bakoya	Batwa (East)	Batwa (West)	Bedzan	Efe	Sua
Sua	0.42	0.56	0.55	0.48	0.58	0.61	0.52	0.75	0.52	0.00
Efe	0.59	0.63	0.60	0.61	0.63	0.61	0.45	0.64	0.00	0.52
Bedzan	0.79	0.84	0.54	0.67	0.73	0.67	0.57	0.00	0.64	0.75
Batwa (West)	0.48	0.63	0.44	0.40	0.46	0.61	0.00	0.57	0.45	0.52
Batwa (East)	0.56	0.80	0.57	0.57	0.60	0.00	0.61	0.67	0.61	0.61
Bakoya	0.54	0.68	0.55	0.41	0.00	0.60	0.46	0.73	0.63	0.58
Bakola	0.43	0.74	0.52	0.00	0.41	0.57	0.40	0.67	0.61	0.48
Baka	0.41	0.74	0.00	0.52	0.55	0.57	0.44	0.54	0.60	0.55
Babongo	0.70	0.00	0.74	0.74	0.68	0.80	0.63	0.84	0.63	0.56
Aka	0.00	0.70	0.41	0.43	0.54	0.56	0.48	0.79	0.59	0.42

Figure 17. Pairwise Jaccard distances in repertoires of musical instruments and foraging-related toolkits

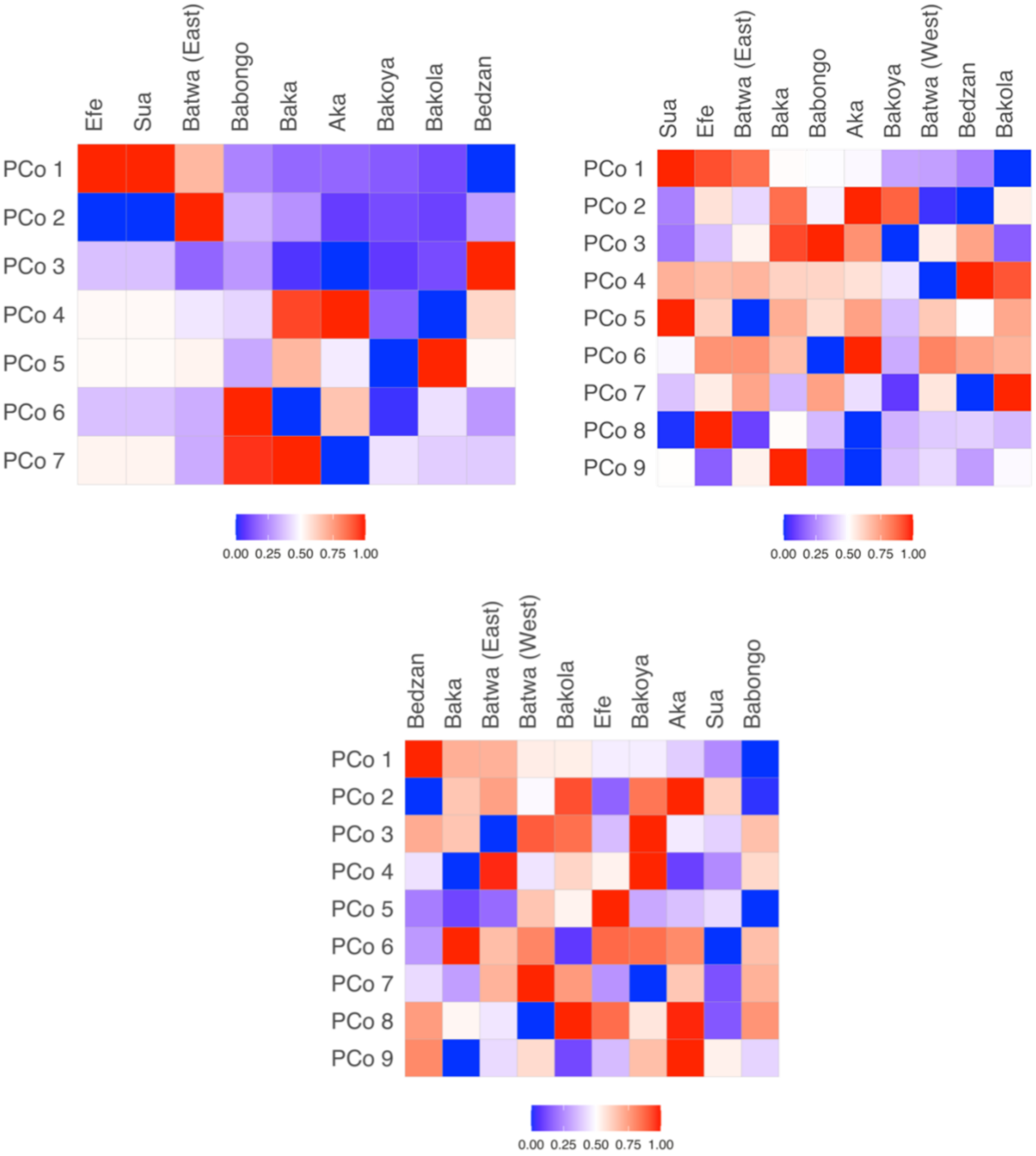


Figure 18. Contribution to PCoA heatmaps in genes (top left), music (top right) and subsistence toolkits (bottom)

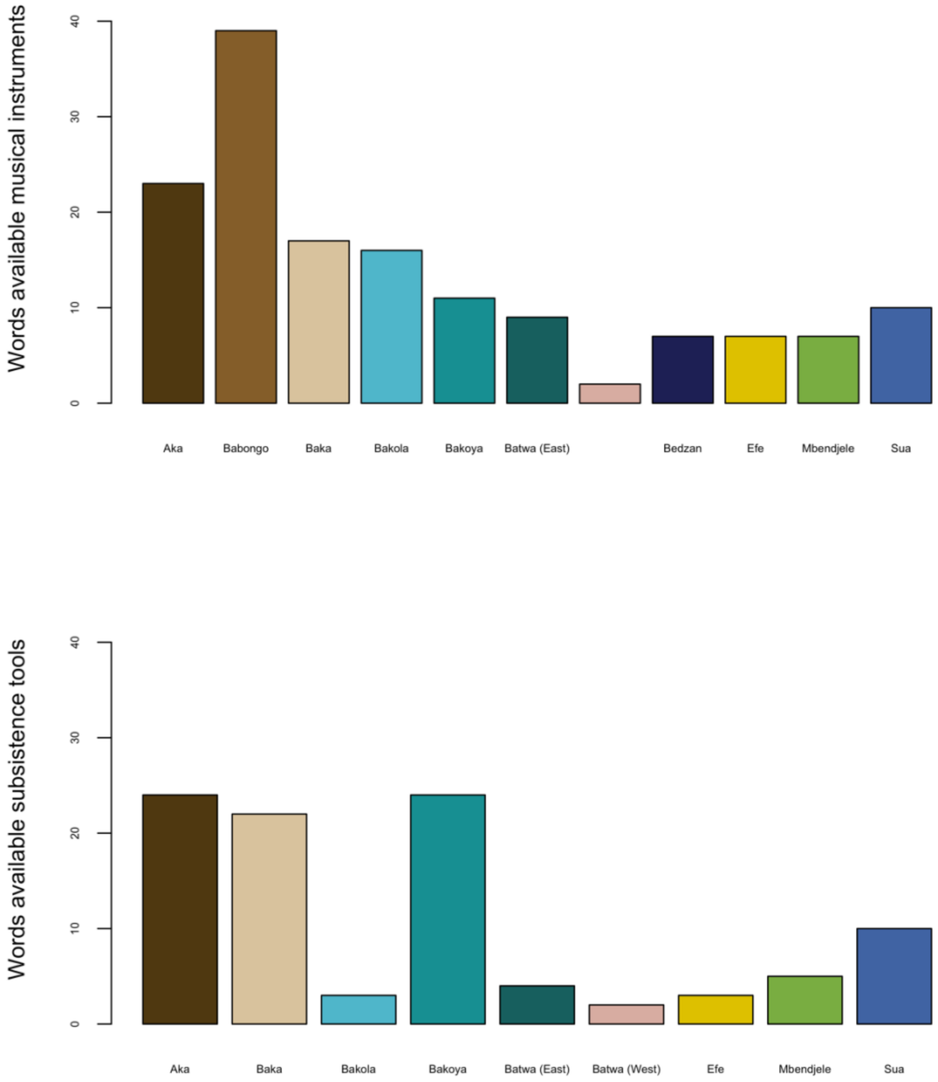


Figure 19. Number of different words compiled per population to designate musical instruments (top) and subsistence tools (bottom)

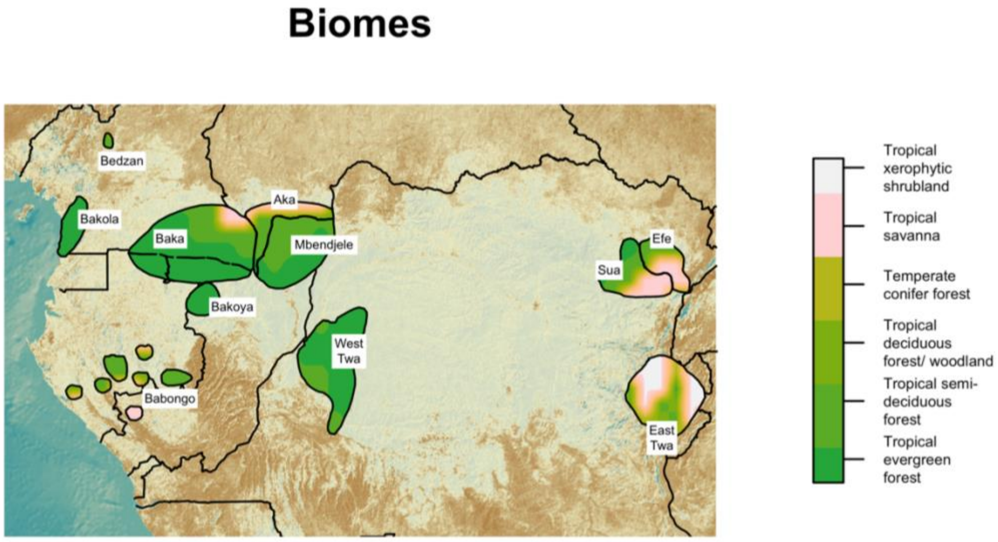


Figure 20. Biomes in the territories of the cultural groups included in our analyses obtained from Beyer et al. (2020)

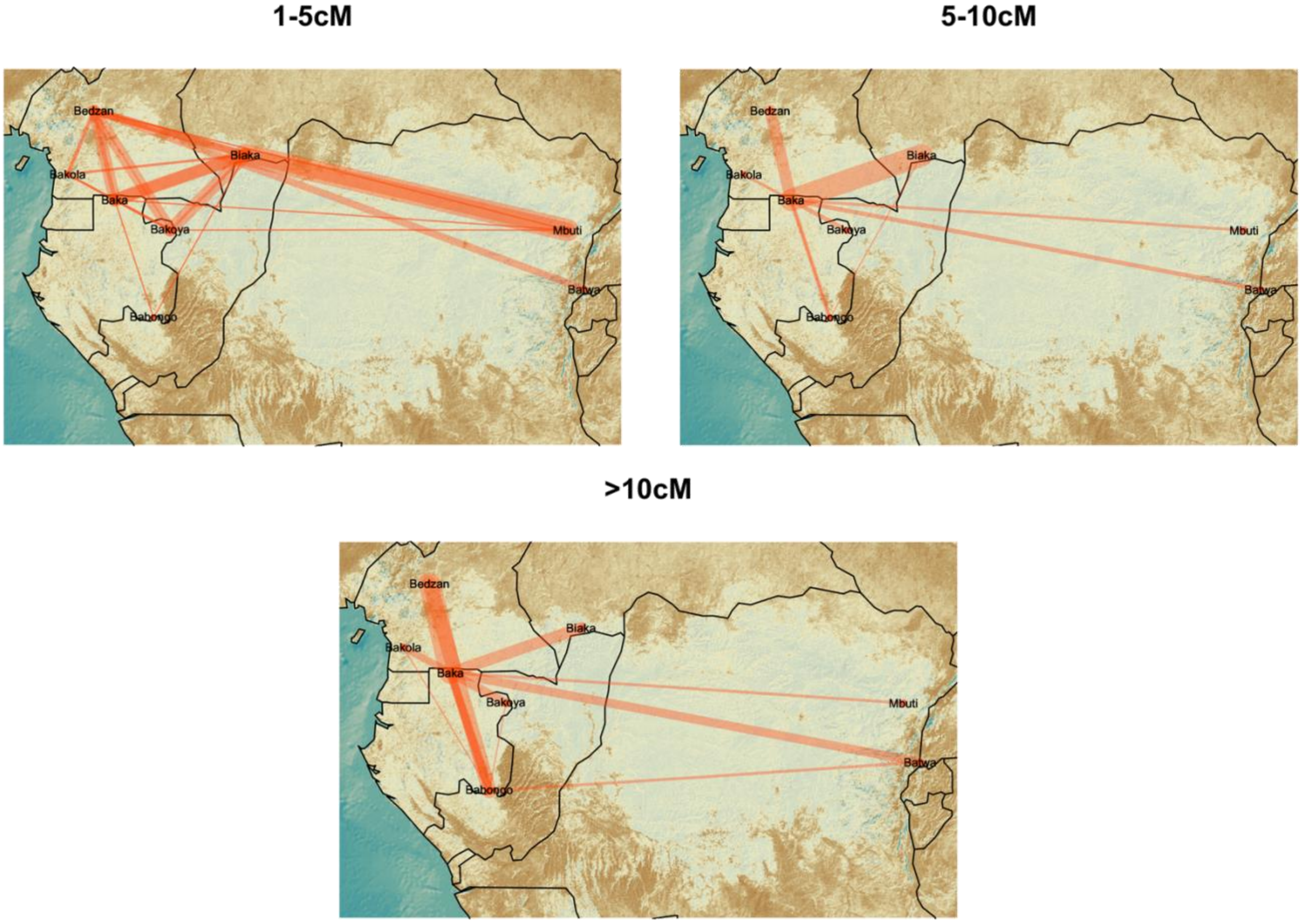


Figure 21. Recent genetic connectivity between CAHG populations. Network visualizations of the average number of IBD segments shared per crosspopulation individual pairs using only CAHG genetic segments with identified IBD blocks in the range of: (A) 1 to 5 cM (2,500 to 1,500 y ago), (B) 5 to 10 cM (1,500 to 500 y ago), (C) and over 10cM (500 to 0 y ago). Thicker lines indicate greater gene flow as identified by a higher probability of sharing IBD blocks of the specified length.

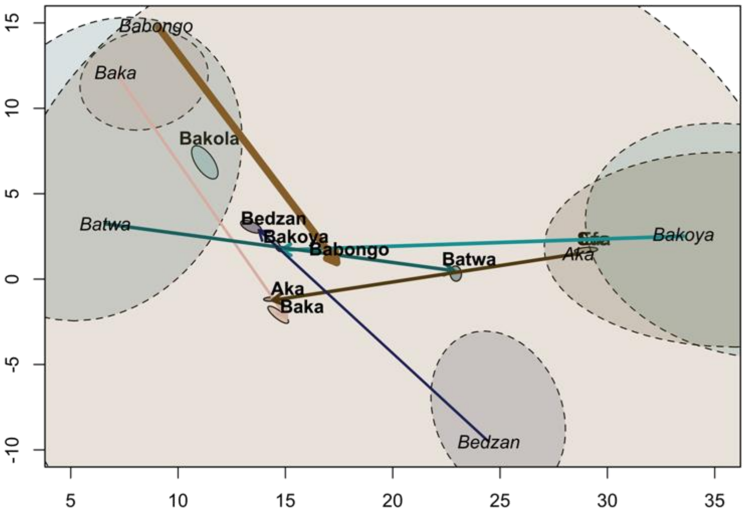


Figure 22. Estimated recent gene flow between CAHG by SpaceMix. Bold letters indicate geo-genetic coordinates of populations in our sample, Italic letters indicate sources of admixture, and arrows indicate direction of admixture. Ellipses bound by solid lines indicate 95% confidence interval around the estimated geo-genetic location of each population and ellipses bound by dashed lines 95% confidence intervals around the hypothesised admixture source. Width of arrows is proportional to estimated proportion of ancestry of a population deriving from admixture source location (origin of the arrows).

Table 1. Area, and location of each group after bahichet 2006;2012 and estimated population size of each of the groups by Bahuchet 2006; Hewlett and Fancher, 2013 (Batwa west), and Lewis, 2000 (for the Batwa of the great lakes). For Babongo, we merged his estimate for central and southern Babongo.

Culture	Longitude	Latitude	Area	Estimated census pop size	Country
Aka	17.22	2.43	6.09	30,000-50,000	CAR, Congo
Babongo	11.98	-2.00	1.84	9,500	Gabon
Baka	14.14	2.35	8.01	30,000-40,000	Cameroon, Gabon, Congo
Bakola	10.16	3.01	1.12	4,000	Cameroon
Bakoya	14.49	0.70	0.88	2,600	Gabon, Congo
Batwa (East)	29.35	-2.38	4.45	70,000-87,000	DRC, Uganda, Rwanda
Batwa (West)	18.53	-1.33	4.89	14,000	DRC
Bedzan	11.36	5.87	0.11	400	Cameroon
Efe	29.33	1.81	1.73	10,000	DRC
Sua	28.38	1.62	2.44	26,000	DRC

Table 2. Populations included in our genetic analyses. Asterisks indicate populations were included in the reduced dataset we created for running ADMIXTURE as well as in our IBDNe analyses

Population	N	Meta group	ReducedData	IBDNe
Boni	21	HG		*
Hadza	26	HG		*
Ogiek	17	HG		*
Sandawe	29	HG		*
Wata	4	HG		
Babongo	50	West CAHG	*	*
Baka	144	West CAHG	*	*
Bakola	29	West CAHG	*	*
Bakoya	20	West CAHG	*	*
Bedzan	24	West CAHG	*	*
Aka	17	West CAHG	*	*
Batwa	35	East CAHG	*	*
Mbuti	13	East CAHG	*	*
Aari	9	Non_HG		
Ahizi	17	Non_HG		
Akele	39	Non_HG		*
Badwee	38	Non_HG		*
Bakiga	54	Non_HG	*	*
Bakota	47	Non_HG		

Baniamer	10	Non_HG		
Bapunu	47	Non_HG		
Bariba	20	Non_HG	*	
Bateke	43	Non_HG		
Bekwil	5	Non_HG		
Benga	40	Non_HG		*
Borana	18	Non_HG		
Bulala	12	Non_HG		
Burji	15	Non_HG		
Datog	18	Non_HG		
Dinka	12	Non_HG		
Duma	39	Non_HG		*
Elmolo	13	Non_HG		
Eshira	40	Non_HG		*
Eviya	20	Non_HG		*
Fang	67	Non_HG		*
Fon	12	Non_HG	*	
Fulani	24	Non_HG		
Gabra	9	Non_HG		
Galoa	48	Non_HG		*
Gurreh	7	Non_HG		
Hadandawa	15	Non_HG		
Hamer	10	Non_HG		
Ilchamus	19	Non_HG		
Iraqw	23	Non_HG		
Iyassa	20	Non_HG		
Kikuyu	14	Non_HG		
Kimbundu	18	Non_HG		
Kongo	11	Non_HG		
Lemande	19	Non_HG		
Luo	32	Non_HG		
Mada	19	Non_HG		
Makina	40	Non_HG		*
Mbororo	16	Non_HG		
Ndumu	38	Non_HG		*
Ngumba	20	Non_HG		
Nzebi	61	Non_HG	*	
Nzime	42	Non_HG		*
Obamba	45	Non_HG		*
Okande	6	Non_HG		
Orma	14	Non_HG		
Orungu	20	Non_HG		

Ovimbundu	16	Non_HG		
Pare	14	Non_HG		
Pokot	5	Non_HG		
Rangi	14	Non_HG		
Rendille	14	Non_HG		
Sabue	14	Non_HG		
Sanye	14	Non_HG		
Sengwer	16	Non_HG		
Shake	46	Non_HG		*
Taita	9	Non_HG		
Taveta	2	Non_HG		
Tikar	19	Non_HG		
Tsogo	62	Non_HG		
Tugen	15	Non_HG		
Yaaku	16	Non_HG		
Yacouba	17	Non_HG		
Yaounde	40	Non_HG		
Yoruba	32	Non_HG	*	

Table 3. Mean proportion of ancestry components from out ADMIXTURE run at K=3. CAHG_1 and CAHG_2 refer to ancestry components associated with Western and Eastern CAHG respectively. AGR refer to ancestry components associated with African agriculturalist populations

Population	**CAHG_1**	**CAHG_2**	**AGR**
Babongo	0.378	0.100	0.523
Baka	0.851	0.012	0.137
Bakola	0.663	0.109	0.228
Bakoya	0.656	0.110	0.234
Batwa	0.014	0.768	0.219
Bedzan	0.522	0.078	0.400
Aka	0.763	0.080	0.157
Efe & Sua	0.225	0.775	0.000001

Table 4. Average heterozygosity of populations in our sample using unmasked dataset

Population	**Heterozygosity**
Babongo	0.326
Baka	0.319
Bakola	0.318
Bakoya	0.318
Batwa (East)	0.315
Bedzan	0.323
Aka	0.317
Efe & Sua	0.291

Table 5. Mantel tests predicting structure of genetic diversity (Fst) using CAHG ancestry components. P values are adjusted with Benjamini Hochberg procedure.

	Genetic diversity (F_{ST})	
	Statistic	*Adj p*
Geography	**0.818**	**0.001**
Ecology	**0.434**	**0.022**
Geography (cont. ecology)	**0.746**	**0.001**
Ecology (cont. geography)	0.201	0.182

Table 6. Number of simulated SNPs for sequence data generated to validate IBDNe results before and after minor allele frequency filtering.

Model	**Sample size**	**SNPs**	**SNPs MAF=0.05**	**SNPs MAF=0.1**
collapse	15	27765599	20664486	16895493
collapse	20	26680735	22137250	16543917
collapse	50	35811285	21580546	15931823
collapse	150	42370877	21056808	15674009
constant	15	6279121	4635268	3788050
constant	20	6187033	5112965	3693104
constant	50	8195115	4845000	3573145
constant	150	9943843	4728448	3514107
expansion	15	7102193	5149201	4207589
expansion	20	7666131	5681552	4122339
expansion	50	9553205	5379826	3969557
expansion	150	12376707	5243012	3896819

Table 7. Estimated admixture proportion from hypothesised admixture source location in Spacemix (taken from MCMC run with highest posterior probability)

Population	**Proportion of genome from inferred admixture source**
Baka	0.110
Bakoya	0.054
Bakola	0.002
Efe & Sua	0.002
Batwa (East)	0.063
Aka	0.061
Bedzan	0.128
Babongo	0.010

Table 8. Multiple matrix regressions assessing the relationship between geographical, genetic, ecological and cultural distance with masked dataset (CAHG genomic segments).

	Musical instruments		**Subsistence tools**	
	Coefficient	*p value*	Coefficient	*p value*
Genes (cont. Geography)	0.185	0.191	0.244	0.164
Genes (cont. ecology)	**0.414**	**0.018**	0.142	0.274
Geography (cont. Genes)	0.269	0.088	-0.031	0.578
Geography (cont. Ecology)	**0.395**	**0.008**	0.0457	0.373
Ecology (cont. Geography)	-0.029	0.559	**0.496**	**0.015**
Ecology (cont. Genes)	-0.033	0.541	0.309	0.119

Table 9. Total number of shared IBD segments >1cM between individuals belonging to CAHG populations (only within CAHG genomic components)

	Babongo	Baka	Bakola	Bakoya	Batwa	Bedzan	Aka	Efe & Sua
Babongo	1385	2790	394	330	378	639	527	68
Baka	2790	71819	4549	4118	1714	12155	28922	339
Bakola	394	4549	981	835	390	1718	1325	32
Bakoya	330	4118	835	882	284	1623	1386	31
Batwa	378	1714	390	284	4245	393	213	100
Bedzan	639	12155	1718	1623	393	9904	4735	127
Aka	527	28922	1325	1386	213	4735	7358	77
Efe & Sua	68	339	32	31	100	127	77	3015

Table 10. Total length of IBD segments >1cM among shared all individuals from each CAHG population pair (only within CAHG genomic components)

	Babongo	Baka	Bakola	Bakoya	Batwa	Bedzan	Aka	Efe & Sua
Babongo	16893.098	12601.295	1968.582	1650.36	2423.201	1664.461	1114.199	128.321
Baka	12601.295	254052.515	15724.828	13071.769	8204.875	29727.654	78632.294	683.507
Bakola	1968.582	15724.828	4172.076	3090.363	2955.117	4285.595	2980.787	66.201
Bakoya	1650.36	13071.769	3090.363	3590.363	2328.025	3976.952	3209.948	57.697
Batwa	2423.201	8204.875	2955.117	2328.025	20928.371	1379.269	443.288	214.79
Bedzan	1664.461	29727.654	4285.595	3976.952	1379.269	64284.796	11427.79	247.443
Aka	1114.199	78632.294	2980.787	3209.948	443.288	11427.79	31292.914	149.019
Efe & Sua	128.321	683.507	66.201	57.697	214.79	247.443	149.019	9804.437

Table 11. Linguistic classification of the languages spoken by the CAHG populations included in this study from Bahuchet 2006;2012 (Following classification by Greenberg, 1966; Ruhen, 1991 and Guthrie, 1967-71).

CAHG group	Phylum	Stock	Family	Group	Sub-group
Baka	Nigel-Kordofanian	Niger-Congo	North Central	Adamawa-Ubangian	Gbanzili-Sere
Bedzan	Nigel-Kordofanian	Niger-Congo	South Central	Bantoid, Non-Bantu	Tikar
Bakola	Nigel-Kordofanian	Niger-Congo	South Central	Northwest Bantu	A80
Bakoya	Nigel-Kordofanian	Niger-Congo	South Central	Northwest Bantu	B20
Babongo	Nigel-Kordofanian	Niger-Congo	South Central	Northwest Bantu	B30
Babongo	Nigel-Kordofanian	Niger-Congo	South Central	Northwest Bantu	B60
Babongo	Nigel-Kordofanian	Niger-Congo	South Central	Northwest Bantu	B70
Mbendjele	Nigel-Kordofanian	Niger-Congo	South Central	Northwest Bantu	C10
Aka	Nigel-Kordofanian	Niger-Congo	South Central	Northwest Bantu	C10
Batwa (West)	Nigel-Kordofanian	Niger-Congo	South Central	Northwest Bantu	C60
Batwa (East)	Nigel-Kordofanian	Niger-Congo	South Central	Central Bantu	J11
Batwa (East)	Nigel-Kordofanian	Niger-Congo	South Central	Central Bantu	J60
Sua	Nigel-Kordofanian	Niger-Congo	South Central	Central Bantu	D30
Efe	Nilo-Saharan	Central-Sudanic	East Central	Mangbetu-Efe	-

Table 12. Shared lexical items between CAHG groups

Word	**Meaning**	**Populations sharing word and meaning**
Musical instruments		
ndumu	drum (conical)	Aka, Baka, Bakoya, Mbendjele
ngom	drum (cylyndrical)	Bakoya, Bakola, Batwa (West), Babongo
hindehu	flute	Aka, Mbendjele
ngombi	harp	Babongo, Baka, Aka
bogongo	zither harp	Aka, Baka
ligbebe	rattle	Aka, Baka
soko	rattle	Aka, Babongo
baka	struck beam	Babongo, Baka
mobio	flute	Aka, Baka
ruma	flute	Efe, Sua
sanze	lamellaphone	Aka, Babongo
mokinda	single-skinned drum	Aka, Baka
kembe	lamellaphone	Baka, Batwa (East), Efe, Sua
Subsistence tools		
apekulo	cable snare	Baka, Bakoya
api	arrow	Sua, Efe
sàwàla	leather sac	Baka, Aka
mokobe	honey container	Baka, Aka