

# Supporting Information

## North African Jewish and non-Jewish populations form distinctive, orthogonal clusters

C.L. Campbell, P.F. Palamara, M. Dubrovsky, L. R. Botigué, M. Fellous, G. Atzmon, C. Oddoux, A. Pearlman, L. Hao, B. M. Henn, E. Burns, C. D. Bustamante, D. Comas, E. Friedman, I. Pe'er, H. Ostrer

## SI Materials and Methods

**Recruitment of Jewish populations.** Recruitment of Jews of European and Middle Eastern origin was described previously (1, 2), and followed a New York University School of Medicine Institutional Review Board-approved protocol (07-333 “Origins and Migrations of Jewish People”). Recruitment of North African, Ethiopian, and Georgian Jews occurred at Sheba Medical Centre in Tel Hashomer, Israel following a local ethics committee and an Israeli Ministry of Health Institutional Review Board approved protocol. Recruitment of non-Jewish individuals from seven different North African locations (North Morocco, South Morocco, Western Sahara, Algeria, Tunisia, Libya and Egypt) was reported previously and followed a Universitat Pompeu Fabra, Ethics Committee-approved IRB protocol (3). In every case, subjects provided informed consent. Jewish subjects were included only if all 4 grandparents came from the same Jewish community. Subjects were excluded if they were known first or second degree relatives of other participants based on IBD analysis. Cryptic relatives were defined as pairs of individuals sharing more than a total of 800 cM and more than 10 IBD segments longer than 10 cM. These values are conservative cutoffs for the exclusion of first-degree cousins or closer relatives (4, 5). Such pairs were used for further analysis by removing one individual for each detected cryptic relationship. A list of the number of individuals for each of the populations is provided in Table 1 in the main text.

**Genotyping and Quality Control.** DNA preparation and genotyping using the Affymetrix Genome-Wide Human SNP Array 6.0 was performed as previously described (1). Our Jewish data set was then merged with selected data from the HGDP world populations (6) (run using Illumina HumanHap650K Beadchips) to enable comparisons with non-Jewish world-wide populations. Following merging of these data sets from two different platforms and filtering out low call-rate (<5%) and symmetric SNPs (alleles A-G or C-T), there were 163,199 SNPs remaining. Data management was done using PLINK (v1.07) (7). The R statistical analysis environment (v2.11.1) was used for additional analyses and to plot most of the results (8).

**Principal Component Analysis.** A further filtering step was then applied, keeping only markers with no failures, and resulted in a set of 46,324 SNPs, which were used for principal component and  $F_{ST}$  calculations. In order to confirm the self-reported ancestry of our Jewish populations we performed principal component analysis using the SMARTPCA program from the EIGENSOFT package (v3.0) (9). Outlier removal was performed using the default parameters (samples greater than 6 standard deviations from the mean in any of the top 10 eigenvectors removed over 5 iterations). Following outlier removal SMARTPCA was used for principal component analysis. Differences between subgroups were assessed using analysis of variance (ANOVA) of the top three eigenvectors.

**$F_{ST}$  calculation and Phylogenetic Analysis.**  $F_{ST}$  values were calculated for each population pair using Genepop (10) with the Weir & Cockham method (11). Confidence intervals were estimated using a bootstrap test in which markers were sampled with replacement for 500 iterations. These data were also used to generate a consensus tree using the neighbor-joining method implemented in PHYLIP (v3.69) (12). An exact test of population differentiation was also carried out using Genepop (13).

**Population Structure.** Population structure was inferred using the program STRUCTURE (v2.3) (14, 15), which uses a Bayesian clustering scheme. A subset of 5,113 markers were chosen that had high total absolute differences in allele frequencies among pairs of populations. This set was also filtered to exclude markers with high linkage disequilibrium. The main parameters for STRUCTURE included 30,000 burn-in and data collection iterations, a separate alpha estimate for each population, and assumption of an admixture model with correlated allele frequencies. Ten repeats of the program were run for each K-value (3-7), and the results were combined into a single Q matrix using CLUMPP (v1.1.2) (16) to resolve label-switching issues.

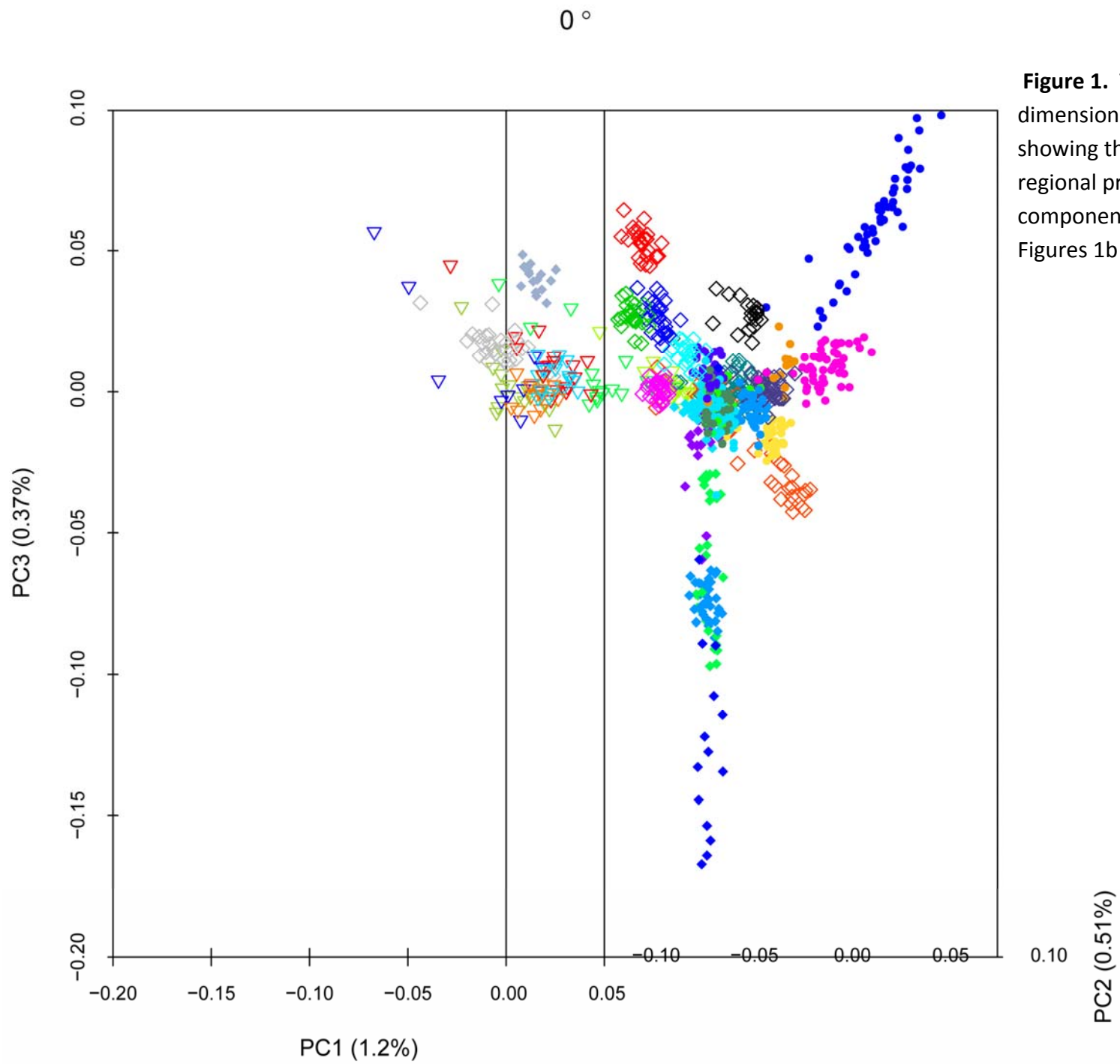
**IBD Discovery.** For IBD discovery, a total of 598,260 SNPs were used. Shared IBD segments were detected using the GERMLINE software package (17) version 1.5. All individuals were computationally phased using Beagle (18) version 3.0.1 and then processed with GERMLINE using the parameters “-min\_m 3 -err\_hom 0 -err\_het 1 -bits 64 -g\_extend”. Computational phasing is prone to errors when long-range haplotypes are considered, while phase quality is generally high for short genomic intervals. To circumvent this problem the genotype extension mode of the GERMLINE algorithm creates a dictionary of short, locally phased haplotypes for all individuals. Pairs of matching short haplotypes are then extended only considering mutually homozygous markers until a specified density of mismatching sites is encountered. Matching segments are reported if their length passes a minimum required centiMorgan length (3 cM for this analysis).

**Ancestry deconvolution.** Ancestry deconvolution was performed on a subset of North African groups using the Xplorigin software package, as described previously (5, 19). Briefly, the Xplorigin algorithm builds a database of short haplotype frequencies from three reference populations. This database is then used to probabilistically infer local ancestry for the analyzed samples, assuming each sample results from the admixture of the reference populations. Samples of North African origins were analyzed with respect to their Maghrebi, Middle Eastern and European ancestry using 36 non-Jewish Tunisian Berber, 48 Palestinian and 48 Basque reference haplotypes, respectively. Basques have experienced very low levels of gene flow from external populations, such as North Africans or Middle Easterners, compared to other neighboring European populations, and, for this reason, are good proxies of the European gene pool (3, 20). The analyzed non-Jewish sample comprised 19 Algerians, 17 Libyans, 18 Northern Moroccans, 16 Southern Moroccans and were compared to 24 Jewish Algerians, 37 Jewish Libyans and 38 Jewish Moroccan individuals. A total of 163,199 SNP markers overlapped between different platforms and were used for the ancestry analysis. Genome-wide ancestry proportions were obtained by averaging the inferred ancestry proportions for all analyzed sites within individual groups. Ancestry deconvolution of shared haplotypes was limited to regions detected to be IBD across pairs of individuals from the analyzed groups. Due to phase uncertainty, the reported values were obtained by averaging the ancestral proportions of both maternal and paternal chromosomes in genomic regions delimited by IBD segments. Permutation testing was performed by randomly dropping IBD segments on the genomes of sharing pairs then recomputing ancestry proportions, 10,000 times. We report as significant IBD ancestry values for which 97.5% of permutations result in lower/higher ancestry score.

The significance of the differences of each ancestral component across populations was assessed using a re-sampling procedure. For each individual in a population, we estimated the genome-wide ancestry proportions resulting from the Xplorigin analysis. For each population, we then created 100,000 datasets by randomly sampling with replacement individuals from the original group. For each such random dataset, we computed the average ancestry of all ancestral components, and compared all cross-population differences.

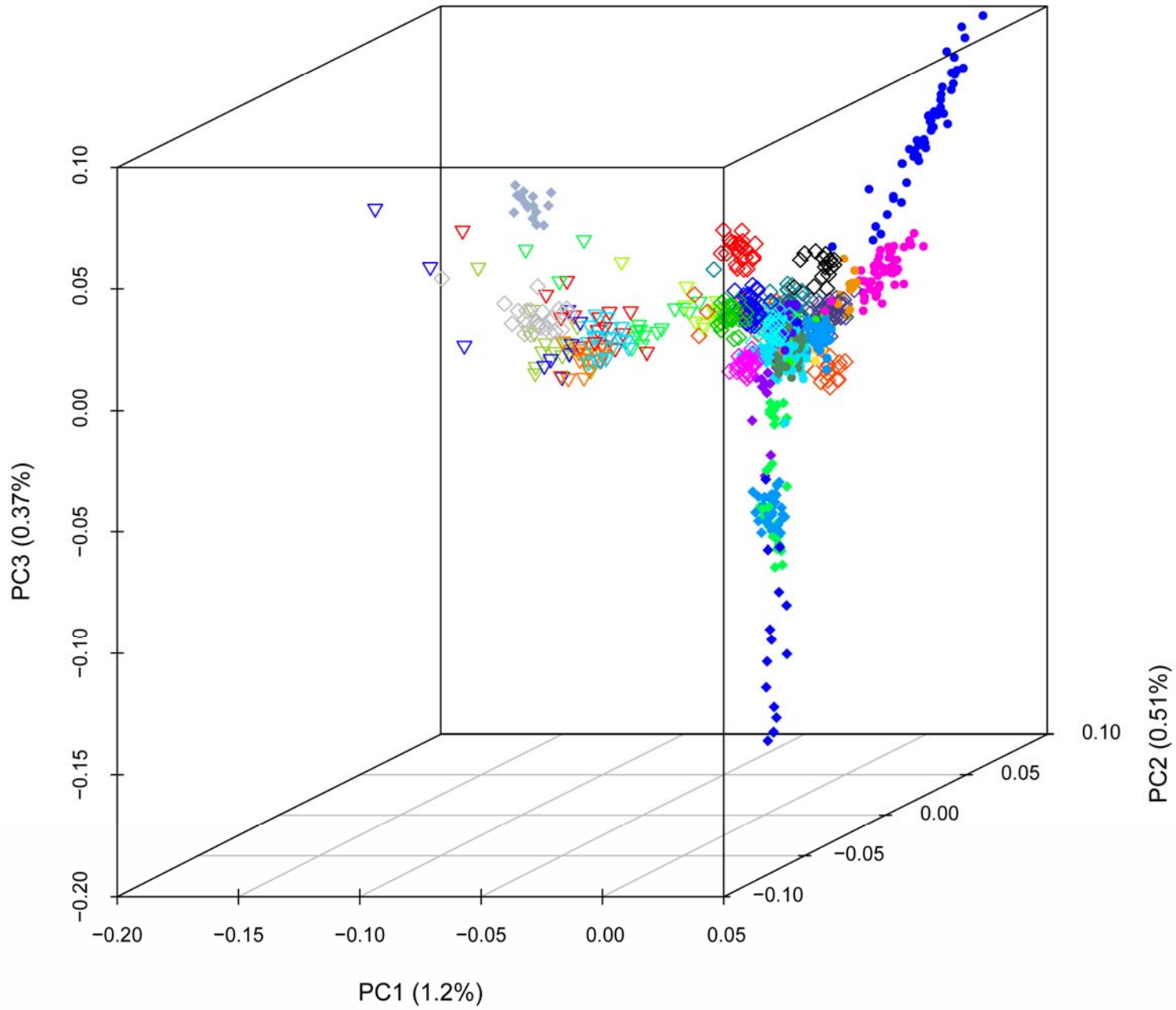
## References

1. Atzmon G, *et al.* (2010) Abraham's children in the genome era: major Jewish diaspora populations comprise distinct genetic clusters with shared Middle Eastern Ancestry. *Am J Hum Genet* 86:850-859.
2. Oddoux C, *et al.* (1999) Mendelian diseases among Roman Jews: implications for the origins of disease alleles. *J Clin Endocrinol Metab* 84:4405-4409.
3. Henn BM, *et al.* (2012) Genomic Ancestry of North Africans Supports Back-to-Africa Migrations. *PLoS genetics* 8:e1002397.
4. Kenny EE, *et al.* (2012) A genome-wide scan of Ashkenazi Jewish Crohn's disease suggests novel susceptibility loci. *PLoS genetics* 8:e1002559.
5. Velez C, *et al.* (2011) The impact of Converso Jews on the genomes of modern Latin Americans. *Hum Genet.*
6. Rosenberg NA, *et al.* (2002) Genetic structure of human populations. *Science* 298:2381-2385.
7. Purcell S, *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559-575.
8. R Development Core Team (2012) R: A language and environment for statistical computing. (R Foundation for Statistical Computing, Vienna, Austria), 2.11.1.
9. Patterson N, Price AL, & Reich D (2006) Population structure and eigenanalysis. *PLoS genetics* 2:e190.
10. Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular ecology resources* 8:103-106.
11. Weir BS & Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358-1370.
12. Felsenstein J (1989) PHYLIP - phylogeny inference package (version 3.2). *Cladistics* 5:164-166.
13. Raymond M & Rousset F (1995) An exact test for population differentiation. *Evolution* 49:1280-1283.
14. Falush D, Stephens M, & Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567-1587.
15. Pritchard JK, Stephens M, & Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.
16. Jakobsson M & Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801-1806.
17. Gusev A, *et al.* (2009) Whole population, genome-wide mapping of hidden relatedness. *Genome research* 19:318-326.
18. Browning SR & Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* 81:1084-1097.
19. Bonnen PE, *et al.* (2010) European admixture on the Micronesian island of Kosrae: lessons from complete genetic information. *Eur J Hum Genet* 18:309-316.
20. Rodriguez-Ezpeleta N, *et al.* (2010) High-density SNP genotyping detects homogeneity of Spanish and French Basques, and confirms their genomic distinctiveness from other European populations. *Hum Genet* 128:113-117.

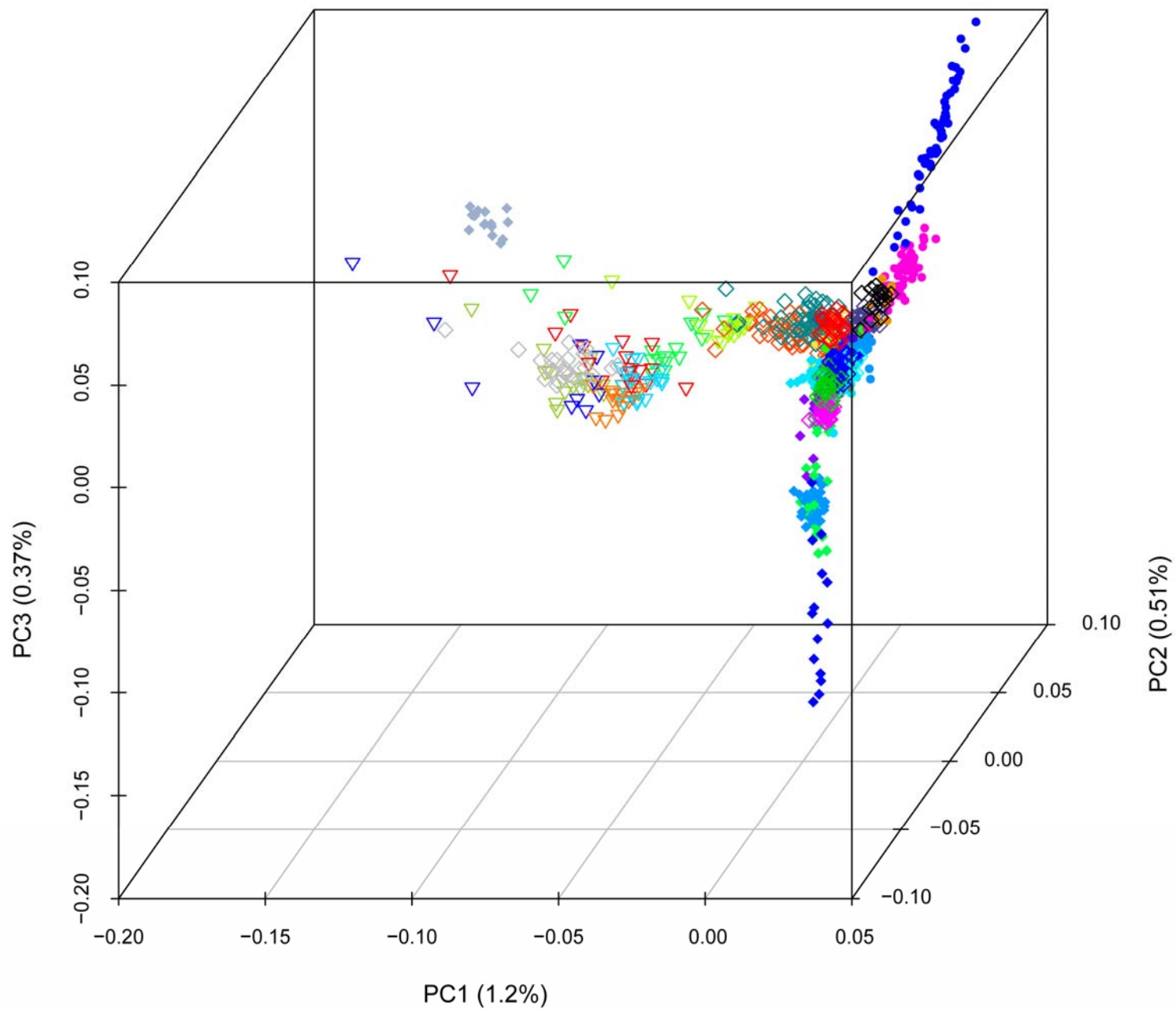


**Figure 1.** Three-dimensional PCA plots showing the first three regional principal components (as shown in Figures 1b and 1c).

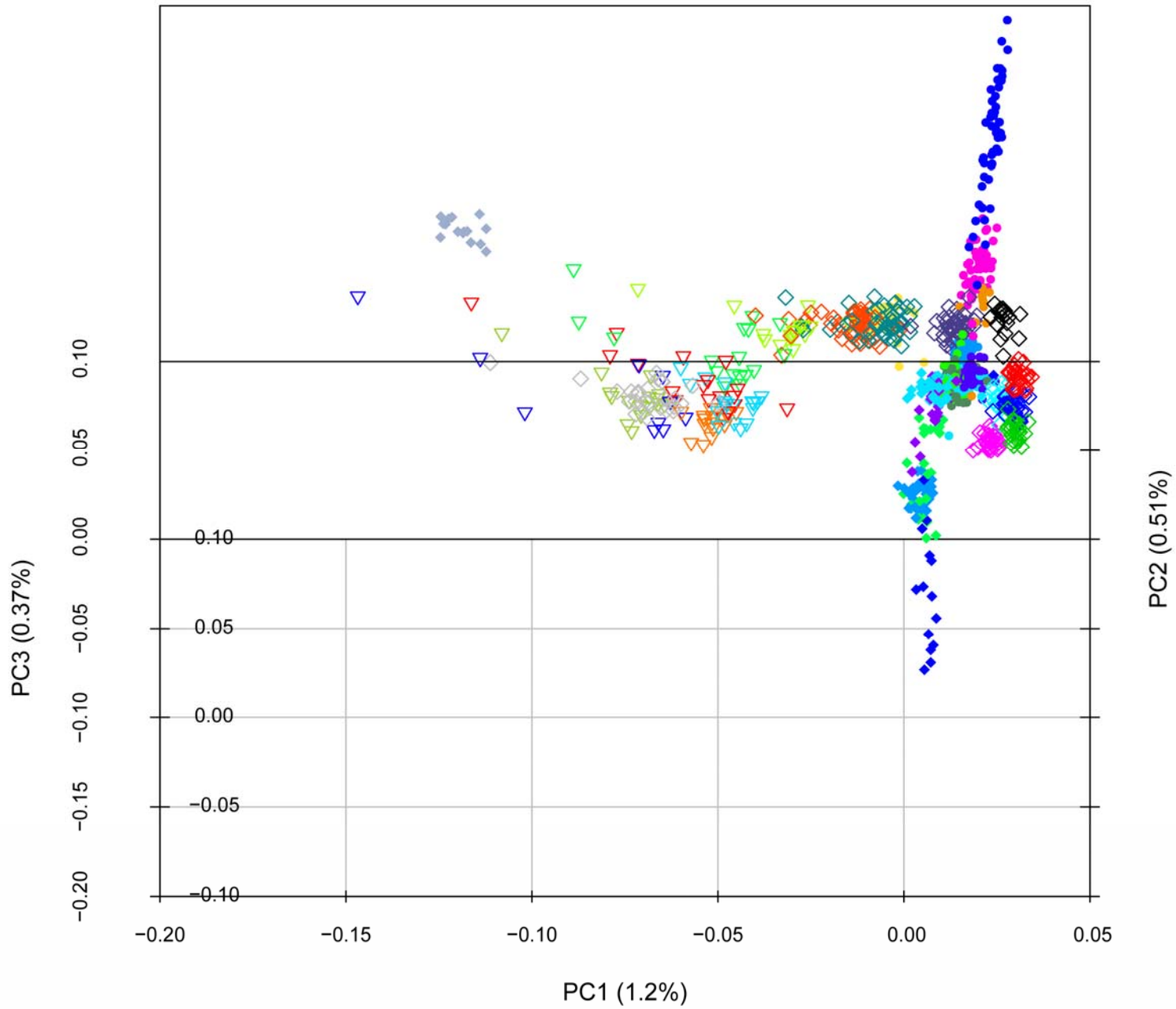
30 °



60 °

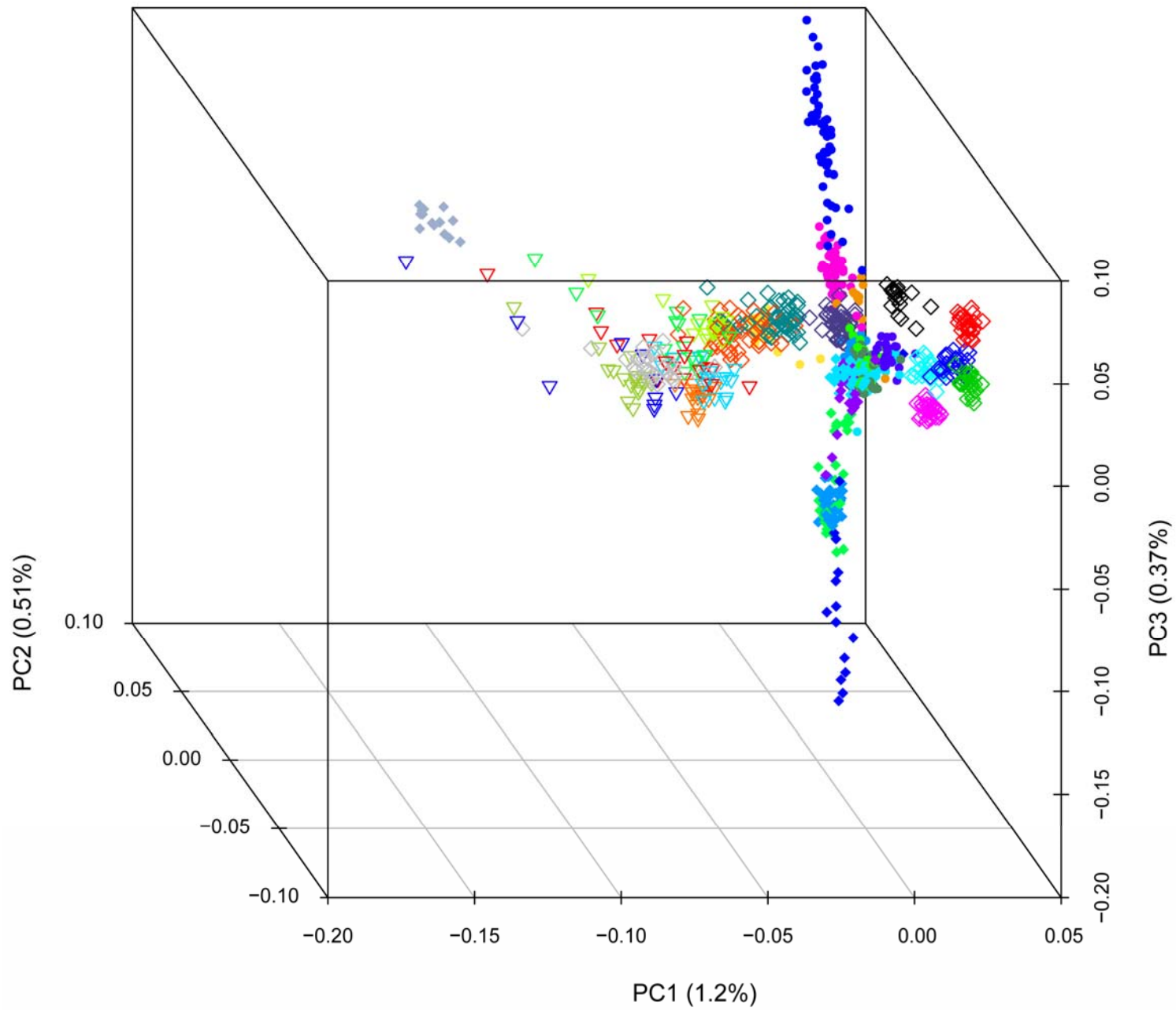


90 °

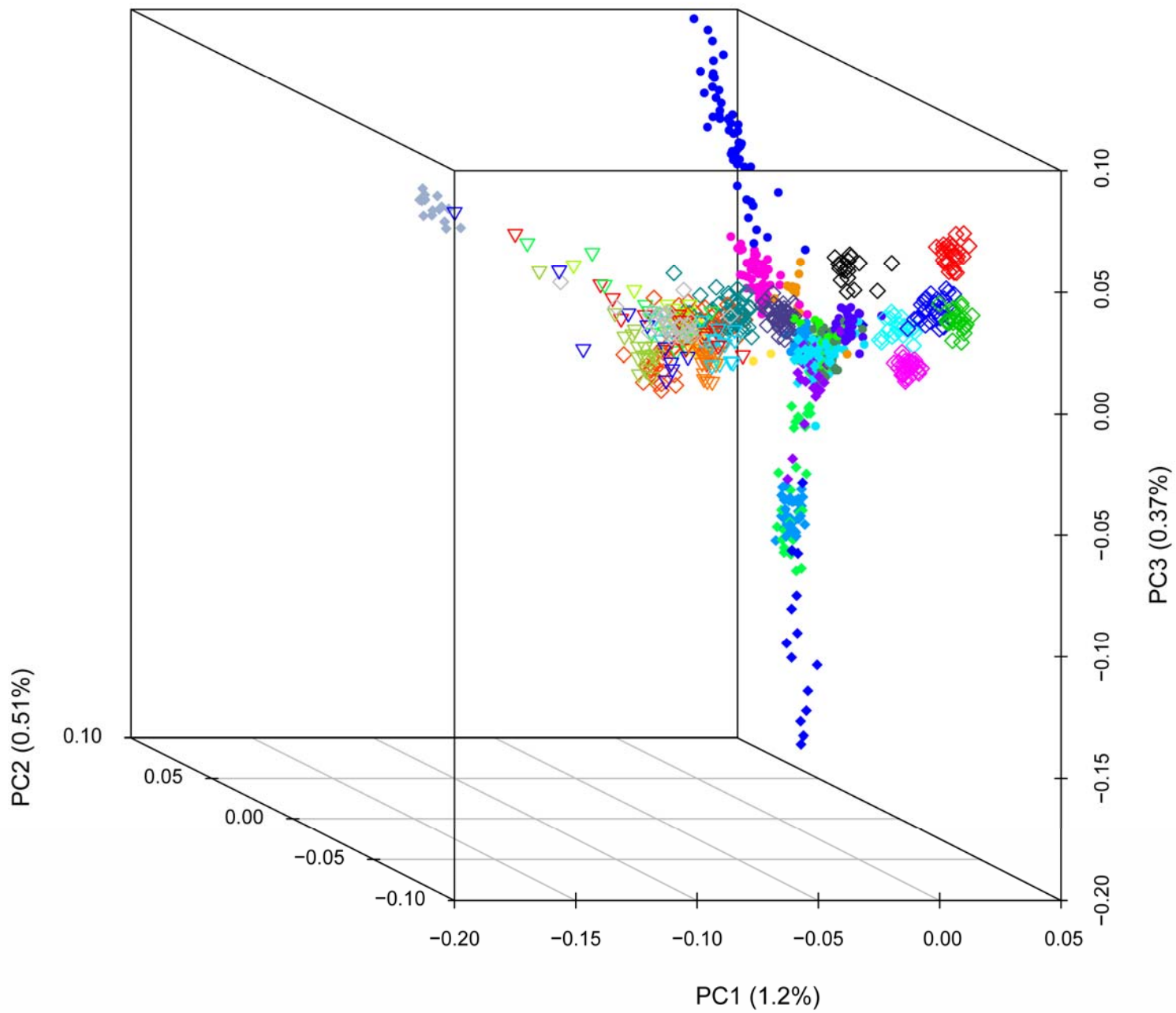


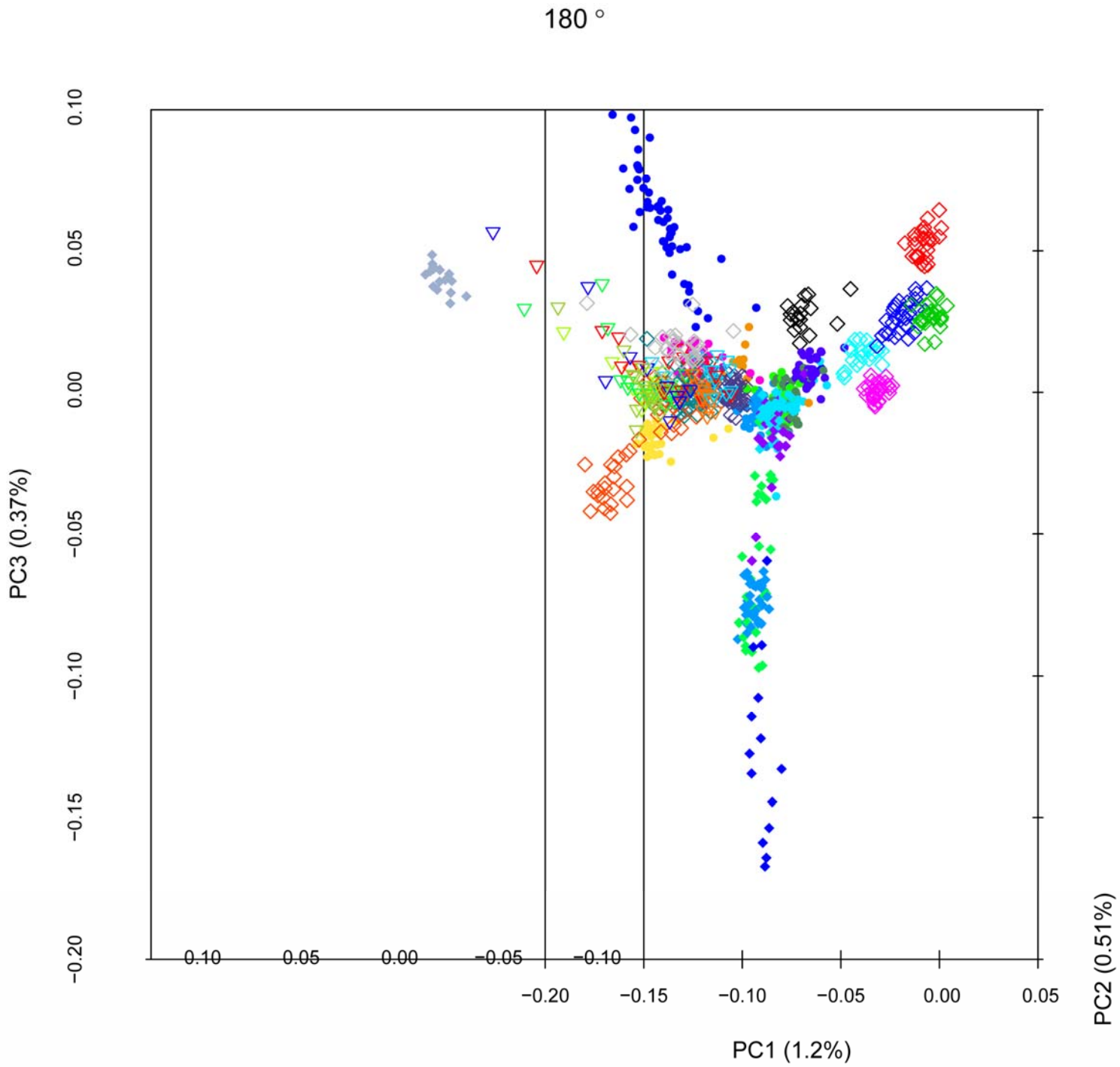


120°

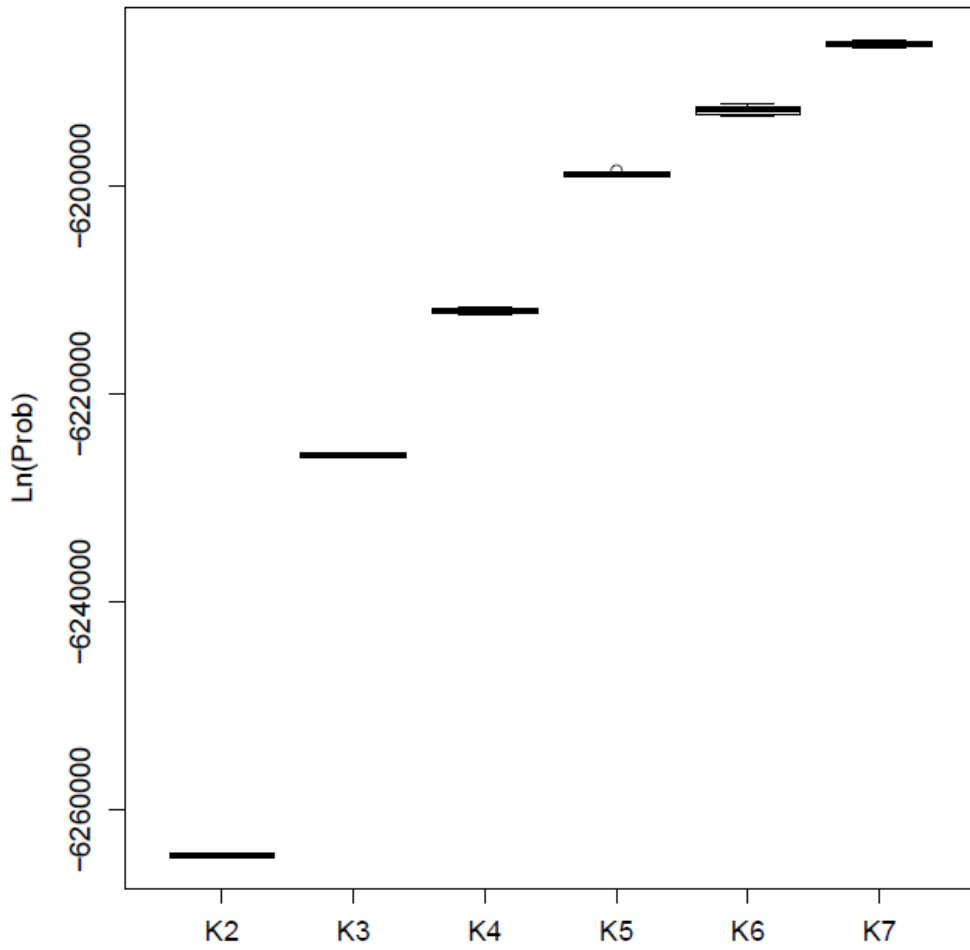


150°

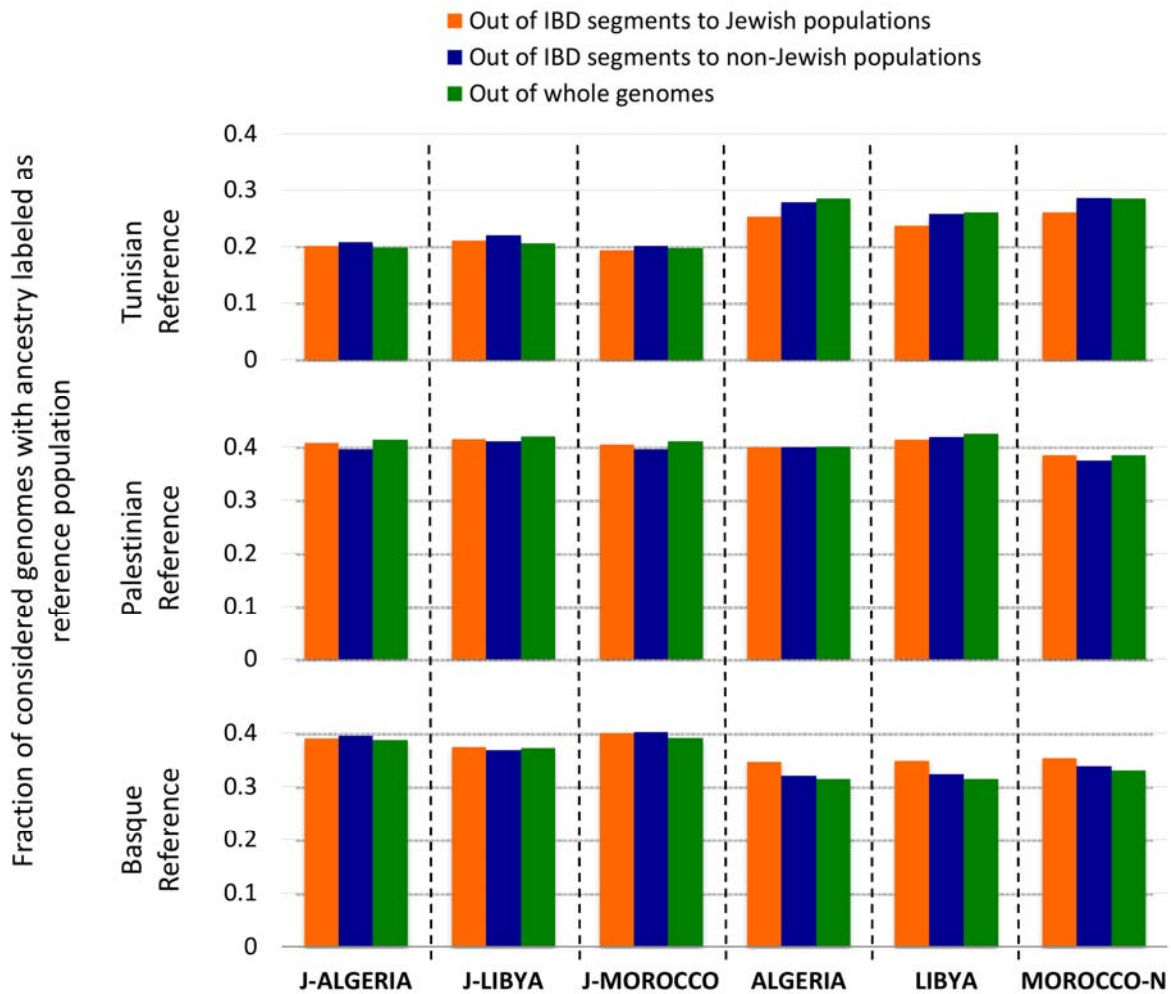




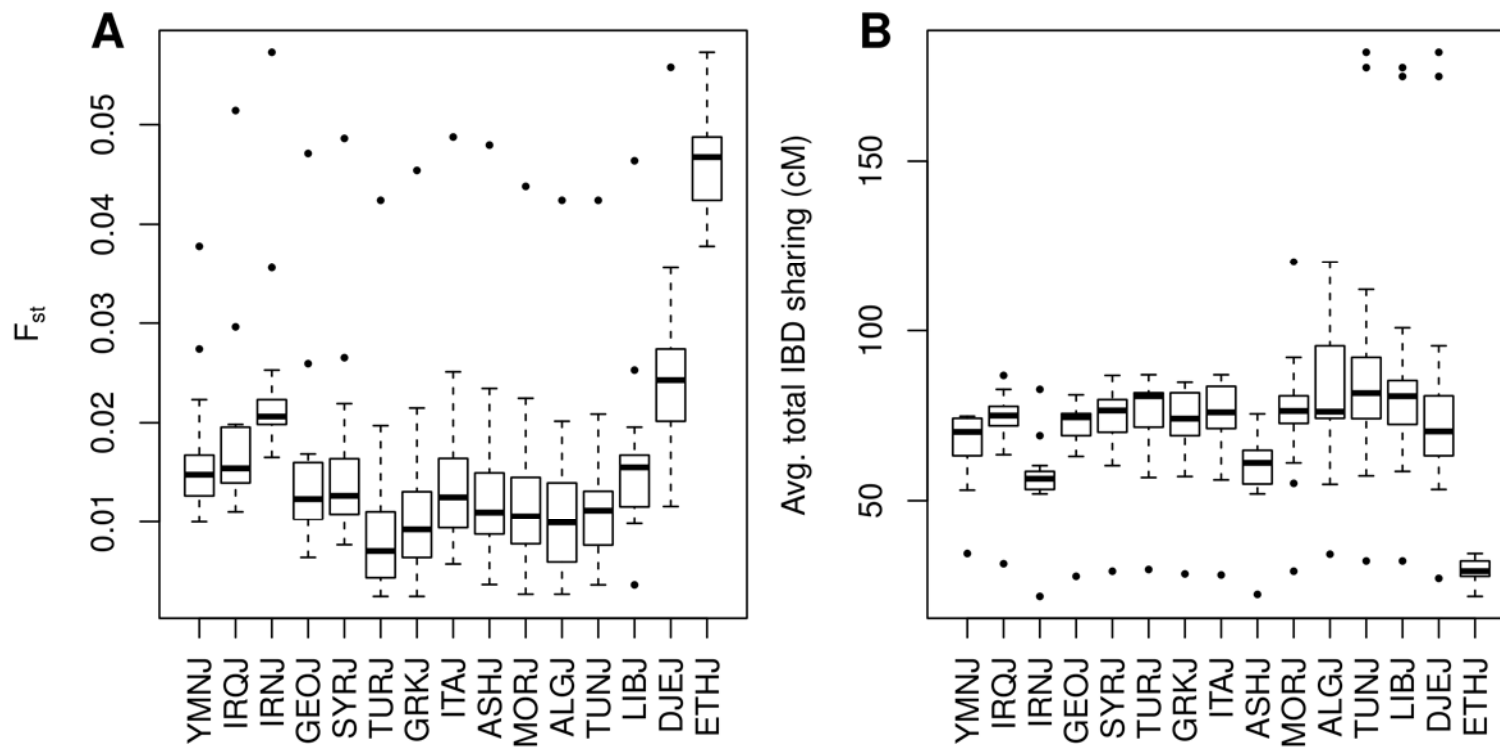
**Figure 2.** Ln probability plot showing the distribution of likelihood scores from each of the independent STRUCTURE iterations for each tested K value.



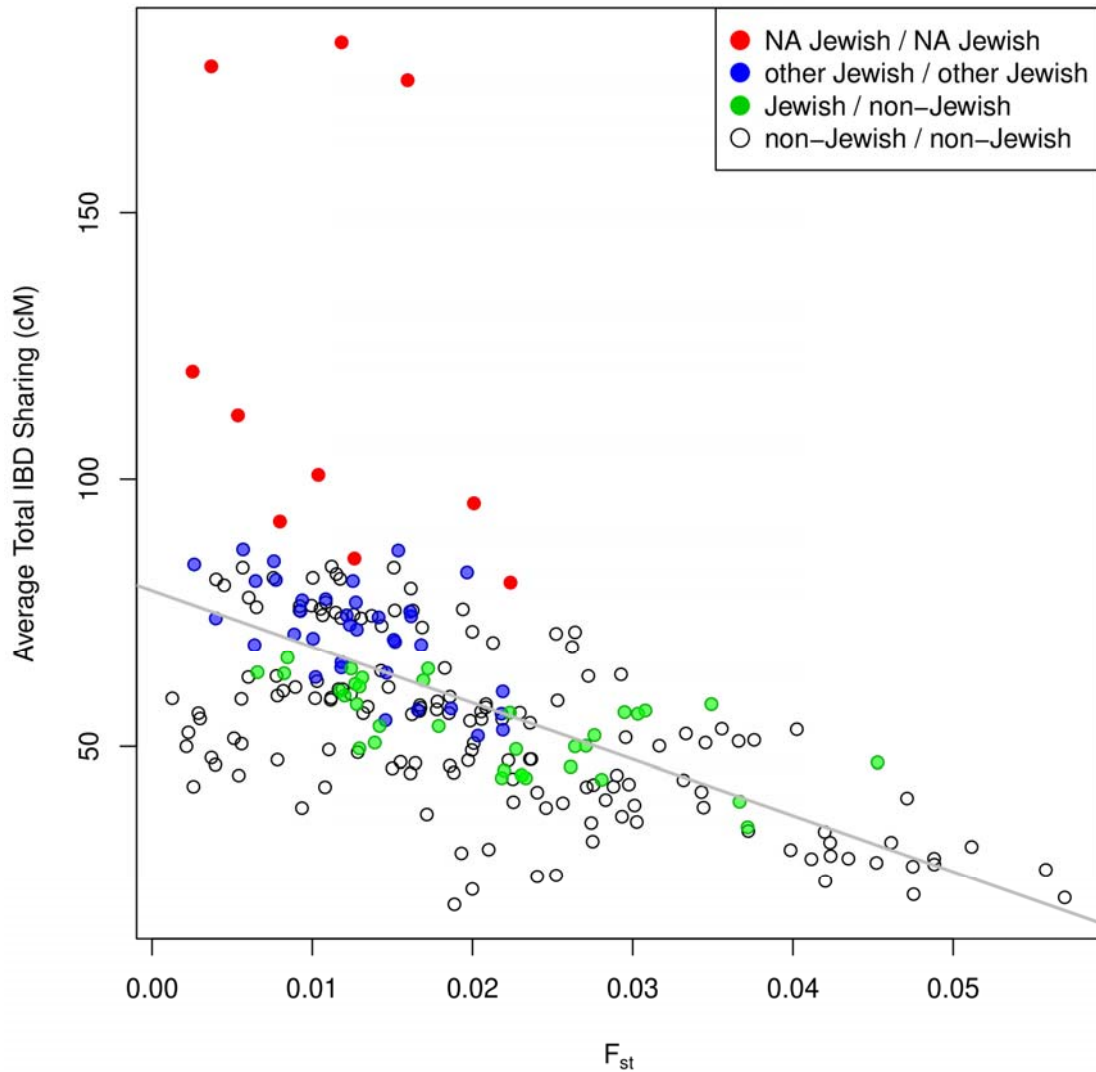
**Figure 3.** To investigate recent ancestry trends, we compared a population’s genome-wide ancestry average to the ancestry of IBD regions shared with Jewish and non-Jewish groups, for each of the analyzed ancestral components (Maghrebi - A, Middle Eastern - B, European - C). Recent admixture appears more uniform for Jewish populations, which exhibit mild fluctuations compared with genome-wide trends. Regions of non-Jewish genomes spanned by IBD segments with Jewish individuals exhibit a noticeable decrease of Maghrebi ancestry (A), compensated by an increased European ancestry (C). A similar, less evident trend is observed for regions that are shared with non-Jewish samples.



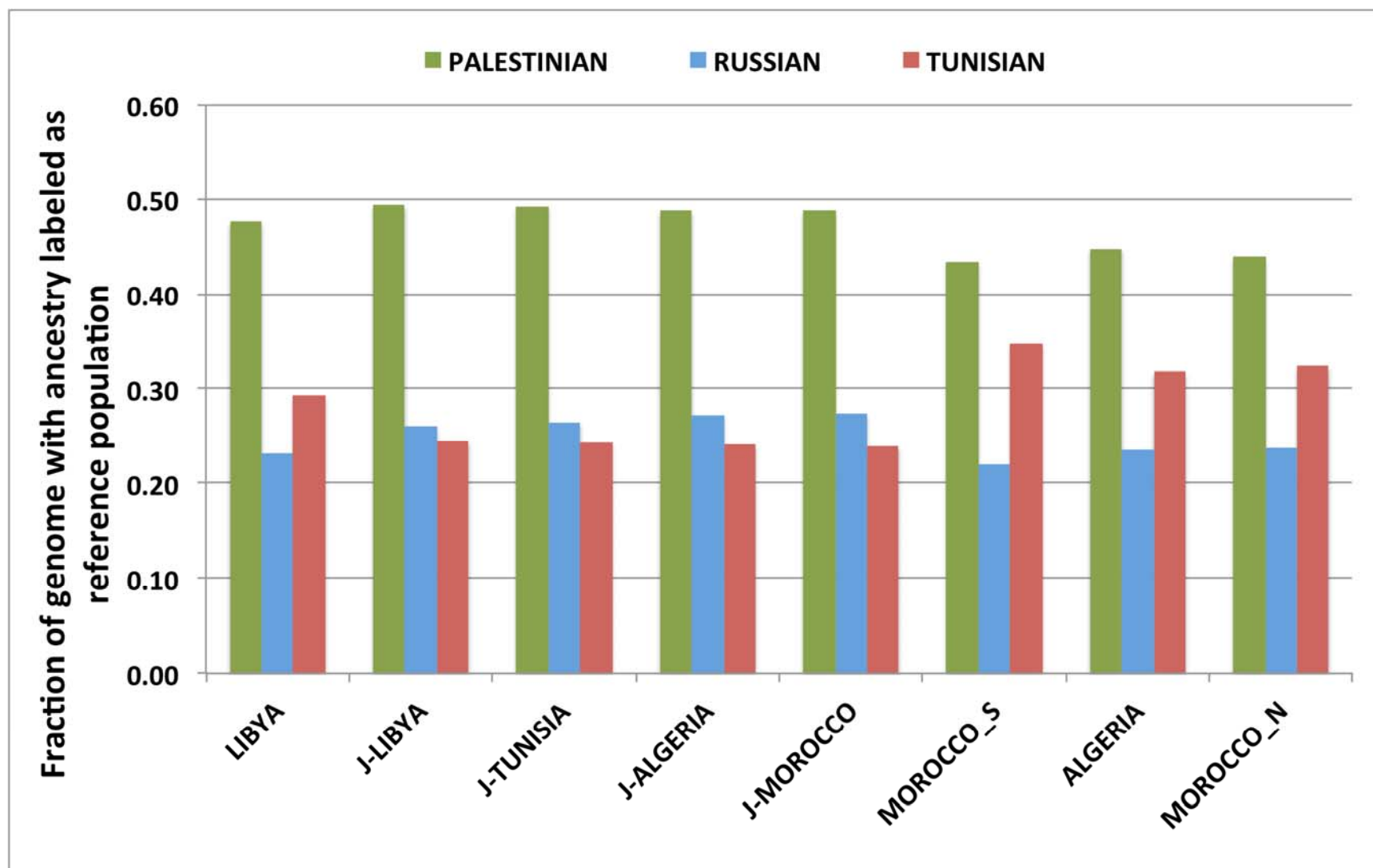
**Figure 4.**  $F_{ST}$  and IBD sharing segment distributions. Each boxplot represents a distribution of  $F_{ST}$  distance (A) or average total IBD sharing in cM (B) from the labeled population to each of the other 14 Jewish populations.



**Figure 5.** Scatterplot of IBD vs  $F_{ST}$ . Each point represents a pairwise comparison between two populations using total average IBD sharing (y-axis) and  $F_{ST}$  (x-axis). A linear regression line representing all points is shown in gray ( $r^2 = 0.30$ ). High IBD sharing and a low genetic distance ( $F_{ST}$ ) are especially evident in between-population comparisons among North African Jews (red points, mean IBD=122, mean  $F_{ST}$ =0.011), as well as in comparisons among other Jewish groups outside of North Africa (blue points, mean IBD=71.2, mean  $F_{ST}$ =0.013). Pairwise comparisons consisting of one Jewish population and another of non-Jewish origin are shown in green and exhibit lower average IBD sharing and higher  $F_{ST}$  (mean IBD=53.6 and mean  $F_{ST}$ =0.021). Comparisons among non-Jewish populations (hollow points) show a similarly low degree of relatedness. Ethiopian Jews were excluded from this analysis.



**Figure 6.** Ancestry deconvolution. The genome-wide ancestry of North African Jewish and non-Jewish populations is analyzed using Russian genomes as a reference of the European ancestry component, Tunisian non-Jewish for Maghrebi ancestry and Palestinian for Middle Eastern ancestry.





**Table 1.** Pairwise average total IBD sharing (lower triangle) and  $F_{ST}$  (upper triangle). IBD sharing across Jewish groups is generally higher than IBD sharing between other population pairs, excluding Ethiopian Jewish samples for which low IBD sharing is observed with all groups. Among Jewish populations, a strong transitive relatedness is observed for Libyan, Djerba and Tunisian samples. Among non-Jewish groups, Southern Moroccan, Saharan and Algerian samples exhibit lower than average levels of sharing with many of the analyzed populations. The shaded boxes represent comparisons between Jewish populations.

	PALN	YMNJ	IRQJ	GEOJ	SYRJ	TURJ	GRKJ	ASHJ	MORJ	ALGJ	TUNJ	LIBJ	DJEJ	ETHJ	LIBY	EGYP	ALGE	MORN	MORS	SOCC	TUNI	MOZA	NITA
PALN		0.009	0.013	0.009	0.012	0.006	0.008	0.010	0.008	0.008	0.009	0.013	0.023	0.035	0.009	0.005	0.012	0.012	0.020	0.020	0.029	0.022	0.010
YMNJ	14.3		0.016	0.014	0.015	0.010	0.012	0.015	0.013	0.012	0.013	0.017	0.027	0.037	0.012	0.008	0.016	0.016	0.026	0.024	0.033	0.026	0.016
IRQJ	13.9	15.6		0.013	0.015	0.011	0.013	0.015	0.014	0.014	0.015	0.019	0.029	0.051	0.021	0.014	0.024	0.023	0.034	0.033	0.040	0.034	0.015
GEOJ	13.8	15.2	18.2		0.012	0.006	0.009	0.010	0.011	0.009	0.011	0.016	0.026	0.047	0.017	0.010	0.019	0.019	0.030	0.030	0.038	0.031	0.010
SYRJ	15.1	16.0	24.2	18.5		0.008	0.009	0.012	0.011	0.011	0.012	0.016	0.026	0.049	0.018	0.012	0.020	0.019	0.030	0.029	0.037	0.030	0.011
TURJ	13.6	14.5	16.1	16.0	20.5		0.003	0.004	0.005	0.004	0.006	0.010	0.020	0.042	0.011	0.006	0.013	0.011	0.023	0.023	0.030	0.023	0.003
GRKJ	14.0	14.7	16.5	16.9	19.8	21.2		0.006	0.007	0.006	0.008	0.012	0.021	0.045	0.013	0.008	0.015	0.013	0.025	0.024	0.032	0.025	0.005
ASHJ	11.1	11.4	14.4	15.0	17.1	18.9	17.3		0.009	0.008	0.010	0.015	0.024	0.048	0.016	0.011	0.017	0.016	0.028	0.027	0.034	0.028	0.006
MORJ	13.8	15.6	15.9	15.8	19.8	19.3	18.4	14.8		0.003	0.008	0.013	0.022	0.043	0.013	0.008	0.014	0.013	0.023	0.023	0.030	0.024	0.007
ALGJ	14.2	15.5	16.1	15.7	20.3	19.2	20.0	14.8	37.1		0.005	0.010	0.020	0.042	0.012	0.007	0.013	0.012	0.022	0.022	0.029	0.023	0.007
TUNJ	14.1	14.7	15.7	16.3	19.7	18.8	20.6	14.1	23.7	33.9		0.004	0.012	0.042	0.013	0.008	0.014	0.013	0.023	0.023	0.031	0.024	0.009
LIBJ	13.7	14.5	15.4	15.5	19.3	18.4	20.3	12.7	20.0	28.2	68.9		0.016	0.046	0.017	0.012	0.018	0.017	0.026	0.026	0.035	0.028	0.013
DJEJ	11.9	13.8	14.9	16.4	18.3	17.0	18.4	15.1	21.5	26.8	79.4	68.5		0.056	0.027	0.022	0.028	0.028	0.037	0.037	0.045	0.038	0.023
ETHJ	6.4	8.2	7.0	5.9	6.8	5.7	6.0	4.7	6.6	7.3	6.8	5.7	5.7		0.019	0.021	0.020	0.025	0.019	0.024	0.042	0.030	0.051
LIBY	10.3	11.7	10.2	10.1	11.1	10.5	10.8	8.5	11.0	11.6	11.4	10.6	10.5	6.1		0.003	0.002	0.003	0.005	0.008	0.021	0.012	0.016
EGYP	12.8	13.4	13.6	14.6	14.9	12.6	13.7	10.9	13.3	13.6	13.4	13.1	12.7	6.5	10.4		0.004	0.005	0.009	0.011	0.022	0.015	0.011
ALGE	9.2	9.9	9.0	9.3	10.1	9.7	10.6	8.0	10.6	10.7	10.9	10.6	10.6	4.6	11.8	9.1		0.001	0.003	0.006	0.018	0.008	0.017
MORN	11.4	12.1	12.1	11.6	14.7	12.4	13.0	10.2	14.0	14.0	13.6	12.9	11.9	5.5	13.0	11.0	13.9		0.002	0.006	0.018	0.009	0.015
MORS	8.3	8.9	9.2	9.6	9.5	9.3	10.2	6.7	10.3	10.8	10.9	10.3	9.0	5.0	11.2	9.0	11.8	13.6		0.004	0.020	0.011	0.029
SOCC	9.0	10.1	10.2	10.0	11.8	10.4	10.7	7.6	11.1	11.4	11.4	10.8	10.6	5.1	12.0	9.3	13.1	16.7	16.0		0.022	0.013	0.028
TUNI	10.3	11.3	10.6	12.5	13.0	12.6	12.7	9.0	13.7	14.1	13.3	12.6	11.9	6.1	12.7	11.1	15.8	18.4	16.8	16.2		0.025	0.034
MOZA	9.7	9.8	10.1	10.5	11.2	10.3	12.0	8.6	11.5	12.4	12.3	11.6	10.1	5.6	11.6	10.3	16.5	17.6	14.3	16.7	23.4		0.028
NITA	15.2	14.7	17.8	18.5	18.4	17.3	18.3	15.8	18.1	17.7	17.7	16.5	14.9	6.6	10.8	13.9	10.9	14.4	11.0	11.0	14.0	12.5	

**Table 2.** 95% confidence intervals for pairwise  $F_{ST}$ . These values were obtained by using the bootstrapped distance matrices created for the neighbor-joining tree.

	YMNJ	IRQJ	GEOJ	SYRJ	TURJ	GRKJ	ASHJ	MORJ	ALGJ	TUNJ	LIBJ	DJEJ	ETHJ	LIBY	EGYP	ALGE	MORN	MORS	SOCC	TUNI	MOZA	NITA
PALN	.0087- .0093	.0132- .0139	.0084- .0094	.0111- .0118	.0052- .0058	.0077- .0082	.0097- .0104	.0083- .0088	.0078- .0084	.009- .0097	.0128- .0135	.0228- .024	.0336- .0352	.0085- .0094	.0045- .0052	.0112- .012	.0117- .0126	.0193- .0208	.0196- .0208	.0283- .0296	.0218- .0228	.0094- .0101
YMNJ		.0162- .017	.0135- .0148	.0146- .0154	.0096- .0104	.0115- .0122	.0142- .015	.0123- .013	.0115- .0124	.0127- .0135	.0163- .0172	.0267- .0281	.0369- .0387	.0121- .0131	.0076- .0084	.0161- .0172	.0162- .0173	.0254- .027	.0234- .0248	.0331- .0346	.0262- .0273	.0152- .0162
IRQJ			.0123- .0135	.0151- .016	.0107- .0114	.0128- .0134	.0145- .0154	.0142- .0149	.0135- .0144	.0149- .0158	.0191- .02	.029- .0303	.0504- .0524	.0203- .0215	.0145- .0154	.0235- .0246	.0231- .0243	.034- .0358	.0332- .0347	.0397- .0414	.0338- .0352	.0144- .0153
GEOJ				.012- .013	.0058- .007	.0083- .0093	.0096- .0108	.0098- .011	.0085- .0097	.0106- .0118	.0154- .0167	.025- .0268	.0459- .0483	.0152- .0168	.0094- .0108	.018- .0194	.0177- .0192	.0293- .0313	.0287- .0305	.0363- .0385	.0298- .0314	.0088- .01
SYRJ					.0074- .008	.0093- .0099	.0113- .0121	.0104- .0111	.01- .0109	.0115- .0122	.0159- .0168	.0258- .0272	.0476- .0497	.0168- .018	.0116- .0127	.0196- .0207	.0185- .0197	.0296- .0313	.0285- .0301	.036- .0378	.0293- .0306	.0107- .0117
TURJ						.0023- .0027	.0034- .004	.0041- .0046	.0034- .004	.0052- .0058	.0094- .0102	.0191- .0203	.0414- .0434	.0103- .0113	.0057- .0065	.0121- .0131	.011- .012	.022- .0237	.0216- .023	.0287- .0302	.0224- .0235	.0029- .0036
GRKJ							.0059- .0065	.0061- .0066	.0056- .0063	.0073- .008	.0112- .012	.0209- .0221	.0445- .0463	.0127- .0137	.0079- .0087	.0149- .0158	.0136- .0145	.0245- .026	.0239- .0252	.0309- .0324	.0247- .0259	.005- .0056
ASHJ								.0084- .0091	.0078- .0085	.0096- .0104	.0142- .015	.0228- .0241	.0469- .049	.0151- .0162	.0107- .0117	.0168- .0179	.0156- .0166	.0271- .0289	.0268- .0283	.0333- .035	.027- .0282	.0058- .0065
MORJ									.0024- .003	.0074- .0081	.0122- .0129	.0218- .023	.0429- .0447	.0124- .0134	.008- .0089	.0137- .0146	.0127- .0136	.0228- .0244	.0226- .0238	.03- .0314	.0237- .0249	.0065- .0073
ALGJ										.0051- .0057	.01- .0109	.0195- .0208	.0414- .0435	.0112- .0123	.0067- .0076	.0126- .0136	.0114- .0126	.0215- .0231	.0214- .0227	.0287- .0303	.0224- .0236	.0063- .0072
TUNJ											.0033- .0039	.0111- .0121	.0414- .0433	.0122- .0134	.0081- .009	.0138- .0148	.013- .0141	.0224- .024	.022- .0235	.03- .0316	.0239- .0251	.0084- .0092
LIBJ												.0153- .0164	.0454- .0474	.0162- .0174	.0122- .0132	.0173- .0185	.017- .0181	.0259- .0276	.026- .0274	.034- .0356	.0272- .0285	.013- .014
DJEJ													.0546- .0568	.0258- .0274	.022- .0234	.0273- .0288	.0269- .0284	.0368- .0388	.036- .0379	.0443- .0464	.0374- .0392	.0223- .0238
ETHJ														.0182- .0197	.0206- .0221	.0199- .0214	.0247- .0263	.0184- .0201	.0235- .025	.0414- .0435	.0286- .0302	.0499- .0521
LIBY															.0023- .0033	.0017- .0026	.0024- .0034	.0047- .0059	.0066- .0078	.0189- .0204	.011- .012	.015- .0163
EGYP																.0038- .0047	.0051- .0059	.009- .0105	.0104- .0116	.0219- .0233	.0145- .0156	.0106- .0117
ALGE																	.0013- .0021	.0023- .0035	.0047- .0057	.0175- .0188	.0079- .0089	.0164- .0177
MORN																		.0021- .0033	.0051- .0062	.0172- .0186	.0087- .0097	.0149- .0162
MORS																			.0032- .0045	.0196- .0213	.0096- .011	.0278- .0298
SOCC																				.0211- .0228	.0125- .0135	.027- .0288
TUNI																					.0243- .0258	.0325- .0343
MOZA																						.0271- .0285







**Table 4.** Total average IBD sharing (cM) among Jewish populations.

	ALGE	ALGJ	ASHJ	DJEJ	EGYP	ETHJ	GEOJ	GRKJ	IRNJ	IRQJ	ITAJ	LIBJ	LYBI	MORJ	MORN	MORS	SOCC	SYRJ	TUNI	TUNJ	TURJ	
ALGJ	49.6																					
ASHJ	37.3	60.4																				
DJEJ	43.8	95.5	54.4																			
EGYP	46.6	63.9	49.4	56.3																		
ETHJ	23.2	34.0	22.2	26.9	30.7																	
GEOJ	45.1	75.6	63.0	68.7	62.2	27.5																
GRKJ	45.9	78.0	69.1	69.5	59.5	28.2	71.1															
IRNJ	36.9	54.8	52.0	53.3	49.3	21.6	69.1	57.1														
IRQJ	47.6	74.6	63.8	63.5	64.2	31.2	81.1	72.0	82.7													
ITAJ	46.4	76.4	75.5	71.2	60.7	27.9	74.7	84.8	56.1	75.4												
LIBJ	53.8	100.8	61.1	174.9	64.6	32.0	75.6	82.4	58.6	75.8	83.6											
LIBY	50.0	59.5	45.0	50.1	55.2	30.0	57.7	56.2	42.4	57.9	56.6	62.4										
MORJ	50.7	120.3	61.1	80.8	63.7	29.0	74.7	76.2	55.1	72.7	76.5	85.3	57.9									
MORN	59.0	60.4	47.1	52.1	51.5	25.8	56.2	57.4	44.6	56.3	57.2	64.6	56.2	61.7								
MORS	42.5	44.1	32.2	34.9	38.5	20.3	35.9	38.5	28.9	38.6	40.0	46.2	44.6	44.6	52.6							
SOCC	50.5	45.5	35.7	39.7	42.4	25.6	42.9	41.4	30.6	43.7	42.8	50.0	47.5	44.1	58.9	47.9						
SYRJ	47.4	75.9	64.8	71.5	60.7	29.0	72.9	77.5	60.3	86.8	77.1	79.7	56.9	77.1	59.3	39.0	42.5					
TUNI	58.4	56.4	41.5	47.0	47.4	24.7	51.2	50.1	40.3	53.2	50.7	57.9	56.5	56.1	64.7	50.6	55.1	51.0				
TUNJ	53.8	112.1	59.0	182.0	66.7	32.0	75.2	81.7	57.3	75.6	83.8	177.5	61.2	92.1	62.9	44.7	49.5	81.5	56.7			
TURJ	48.9	81.4	74.1	71.6	63.0	29.5	81.1	84.2	56.8	77.7	87.0	81.7	58.7	80.3	59.1	39.6	43.9	81.3	51.7	83.6		
YMNJ	46.9	74.2	54.9	63.2	63.2	34.2	74.3	65.8	53.1	74.5	69.7	72.4	59.8	74.8	56.0	39.4	47.6	70.1	52.4	74.1	70.3	



**Table 6.** Ancestry deconvolution of IBD segments. Tracts shared IBD are co-inherited from common ancestors by pairs of individuals and are generally more recent than non-IBD regions of the genome. We examined the Xplorigin ancestry deconvolution of these regions and performed permutation testing to assess whether the reported ancestry is significantly increased (green) or decreased (red) compared to randomly shuffled segments of the same length. The reported increased/decreased proportions reflect recent trends in the ancestral composition of the analyzed samples.

**MAGHREBI**

	ALGERIA	LIBYA	MOROCCO_N	MOROCCO_S	J-ALGERIA	J-LIBYA	J-MOROCCO	J-TUNISIA
ALGERIA	0.291	0.291	0.292	0.294	0.259	0.268	0.253	0.264
LIBYA	0.274	0.249	0.261	0.300	0.241	0.235	0.249	0.247
MOROCCO_N	0.312	0.295	0.297	0.317	0.263	0.259	0.271	0.275
MOROCCO_S	0.323	0.349	0.338	0.352	0.345	0.329	0.348	0.312
J-ALGERIA	0.205	0.203	0.207	0.221	0.192	0.194	0.196	0.199
J-LIBYA	0.200	0.206	0.212	0.215	0.200	0.206	0.199	0.200
J-MOROCCO	0.198	0.209	0.202	0.198	0.201	0.198	0.202	0.201
J-TUNISIA	0.218	0.210	0.215	0.222	0.207	0.205	0.203	0.207

**MIDDLE-EASTERN**

	ALGERIA	LIBYA	MOROCCO_N	MOROCCO_S	J-ALGERIA	J-LIBYA	J-MOROCCO	J-TUNISIA
ALGERIA	0.388	0.392	0.382	0.372	0.386	0.387	0.397	0.383
LIBYA	0.401	0.425	0.421	0.405	0.417	0.431	0.419	0.417
MOROCCO_N	0.368	0.364	0.377	0.361	0.377	0.395	0.375	0.377
MOROCCO_S	0.371	0.362	0.362	0.367	0.350	0.355	0.344	0.350
J-ALGERIA	0.380	0.406	0.385	0.398	0.405	0.393	0.398	0.391
J-LIBYA	0.394	0.401	0.395	0.395	0.398	0.410	0.402	0.398
J-MOROCCO	0.368	0.372	0.362	0.395	0.391	0.368	0.392	0.391
J-TUNISIA	0.396	0.411	0.407	0.414	0.409	0.420	0.420	0.409

**EUROPEAN**

	ALGERIA	LIBYA	MOROCCO_N	MOROCCO_S	J-ALGERIA	J-LIBYA	J-MOROCCO	J-TUNISIA
ALGERIA	0.291	0.335	0.326	0.334	0.355	0.345	0.350	0.352
LIBYA	0.325	0.326	0.318	0.294	0.342	0.334	0.332	0.342
MOROCCO_N	0.321	0.341	0.326	0.322	0.360	0.346	0.354	0.360
MOROCCO_S	0.306	0.289	0.299	0.281	0.305	0.315	0.308	0.305
J-ALGERIA	0.415	0.391	0.409	0.381	0.403	0.414	0.406	0.410
J-LIBYA	0.406	0.393	0.393	0.390	0.402	0.384	0.399	0.402
J-MOROCCO	0.435	0.418	0.436	0.407	0.408	0.435	0.406	0.408
J-TUNISIA	0.386	0.380	0.378	0.365	0.384	0.375	0.378	0.384



**Table 7.** Ancestry deconvolution of IBD segments in different length intervals. We examined the Xplorigin deconvolution of regions shared IBD across individuals. We separately examined IBD segments of length between 3 and 4 centiMorgans, and segments longer than 4 centiMorgans, expected to be co-inherited from more recent common ancestors.

MAGHREBI	3cM ≤ Length < 4cM							
	ALGERIA	LIBYA	MOROCCO_N	MOROCCO_S	J-ALGERIA	J-LIBYA	J-MOROCCO	J-TUNISIA
ALGERIA	0.296	0.292	0.303	0.306	0.265	0.274	0.266	0.276
LIBYA	0.246	0.244	0.238	0.261	0.220	0.224	0.234	0.230
MOROCCO_N	0.304	0.295	0.295	0.314	0.270	0.272	0.275	0.283
MOROCCO_S	0.313	0.303	0.334	0.359	0.296	0.289	0.300	0.292
J-ALGERIA	0.204	0.188	0.206	0.217	0.189	0.190	0.188	0.193
J-LIBYA	0.214	0.204	0.211	0.202	0.197	0.197	0.199	0.200
J-MOROCCO	0.194	0.197	0.207	0.209	0.191	0.193	0.197	0.197
J-TUNISIA	0.220	0.200	0.223	0.213	0.201	0.201	0.203	0.198

MIDDLE-EASTERN	3cM ≤ Length < 4cM							
	ALGERIA	LIBYA	MOROCCO_N	MOROCCO_S	J-ALGERIA	J-LIBYA	J-MOROCCO	J-TUNISIA
ALGERIA	0.374	0.381	0.389	0.385	0.379	0.378	0.385	0.375
LIBYA	0.435	0.429	0.422	0.406	0.417	0.440	0.414	0.422
MOROCCO_N	0.373	0.352	0.372	0.367	0.359	0.371	0.371	0.359
MOROCCO_S	0.376	0.379	0.375	0.371	0.370	0.385	0.355	0.375
J-ALGERIA	0.397	0.406	0.384	0.381	0.406	0.400	0.404	0.401
J-LIBYA	0.409	0.418	0.414	0.411	0.404	0.418	0.406	0.419
J-MOROCCO	0.417	0.411	0.407	0.404	0.415	0.413	0.416	0.414
J-TUNISIA	0.401	0.410	0.395	0.404	0.408	0.420	0.422	0.423

EUROPEAN	3cM ≤ Length < 4cM							
	ALGERIA	LIBYA	MOROCCO_N	MOROCCO_S	J-ALGERIA	J-LIBYA	J-MOROCCO	J-TUNISIA
ALGERIA	0.330	0.327	0.308	0.310	0.356	0.349	0.349	0.350
LIBYA	0.319	0.327	0.340	0.333	0.363	0.336	0.352	0.348
MOROCCO_N	0.322	0.352	0.333	0.320	0.371	0.357	0.353	0.357
MOROCCO_S	0.310	0.318	0.292	0.270	0.334	0.326	0.345	0.333
J-ALGERIA	0.399	0.407	0.409	0.401	0.405	0.410	0.408	0.406
J-LIBYA	0.378	0.378	0.375	0.387	0.399	0.384	0.395	0.381
J-MOROCCO	0.389	0.391	0.386	0.387	0.393	0.394	0.387	0.389
J-TUNISIA	0.379	0.389	0.382	0.384	0.391	0.379	0.375	0.379

**MAGHREBI****Length  $\geq$  4cM**

	ALGERIA	LIBYA	MOROCCO_N	MOROCCO_S	J-ALGERIA	J-LIBYA	J-MOROCCO	J-TUNISIA
ALGERIA	0.273	0.248	0.246	0.297	0.248	0.265	0.248	0.251
LIBYA	0.253	0.246	0.252	0.270	0.219	0.242	0.238	0.230
MOROCCO_N	0.275	0.273	0.284	0.321	0.278	0.250	0.253	0.255
MOROCCO_S	0.304	0.303	0.358	0.337	0.261	0.245	0.269	0.273
J-ALGERIA	0.221	0.205	0.207	0.196	0.193	0.202	0.201	0.203
J-LIBYA	0.237	0.218	0.240	0.236	0.210	0.211	0.213	0.210
J-MOROCCO	0.230	0.221	0.208	0.221	0.205	0.202	0.200	0.209
J-TUNISIA	0.218	0.205	0.223	0.253	0.221	0.209	0.219	0.208

**MIDDLE-EASTERN****Length  $\geq$  4cM**

	ALGERIA	LIBYA	MOROCCO_N	MOROCCO_S	J-ALGERIA	J-LIBYA	J-MOROCCO	J-TUNISIA
ALGERIA	0.410	0.432	0.416	0.422	0.405	0.404	0.405	0.427
LIBYA	0.418	0.432	0.433	0.421	0.440	0.423	0.436	0.431
MOROCCO_N	0.378	0.375	0.386	0.343	0.367	0.407	0.391	0.402
MOROCCO_S	0.380	0.369	0.324	0.343	0.366	0.351	0.377	0.313
J-ALGERIA	0.379	0.401	0.395	0.384	0.404	0.396	0.401	0.399
J-LIBYA	0.383	0.404	0.383	0.385	0.399	0.419	0.407	0.417
J-MOROCCO	0.393	0.389	0.405	0.409	0.404	0.403	0.405	0.404
J-TUNISIA	0.417	0.434	0.420	0.417	0.405	0.422	0.413	0.421

**EUROPEAN****Length  $\geq$  4cM**

	ALGERIA	LIBYA	MOROCCO_N	MOROCCO_S	J-ALGERIA	J-LIBYA	J-MOROCCO	J-TUNISIA
ALGERIA	0.317	0.320	0.337	0.281	0.346	0.331	0.346	0.322
LIBYA	0.329	0.322	0.315	0.308	0.341	0.334	0.326	0.339
MOROCCO_N	0.347	0.352	0.330	0.335	0.355	0.342	0.356	0.343
MOROCCO_S	0.316	0.328	0.318	0.320	0.373	0.405	0.354	0.414
J-ALGERIA	0.400	0.394	0.398	0.419	0.403	0.402	0.398	0.399
J-LIBYA	0.379	0.379	0.377	0.379	0.391	0.370	0.380	0.374
J-MOROCCO	0.376	0.390	0.387	0.369	0.391	0.395	0.395	0.387
J-TUNISIA	0.365	0.360	0.358	0.330	0.373	0.370	0.368	0.371

