The American Journal of Human Genetics, Volume *86*

Supplemental Data

Abraham's Children in the Genome Era: Major Jewish Diaspora Populations Comprise Distinct Genetic Clusters with Shared Middle Eastern Ancestry

Gil Atzmon, Li Hao, Itsik Pe'er, Christopher Velez, Alexander Pearlman, Pier Francesco Palamara, Bernice Morrow, Eitan Friedman, Carole Oddoux, Edward Burns, and Harry Ostrer

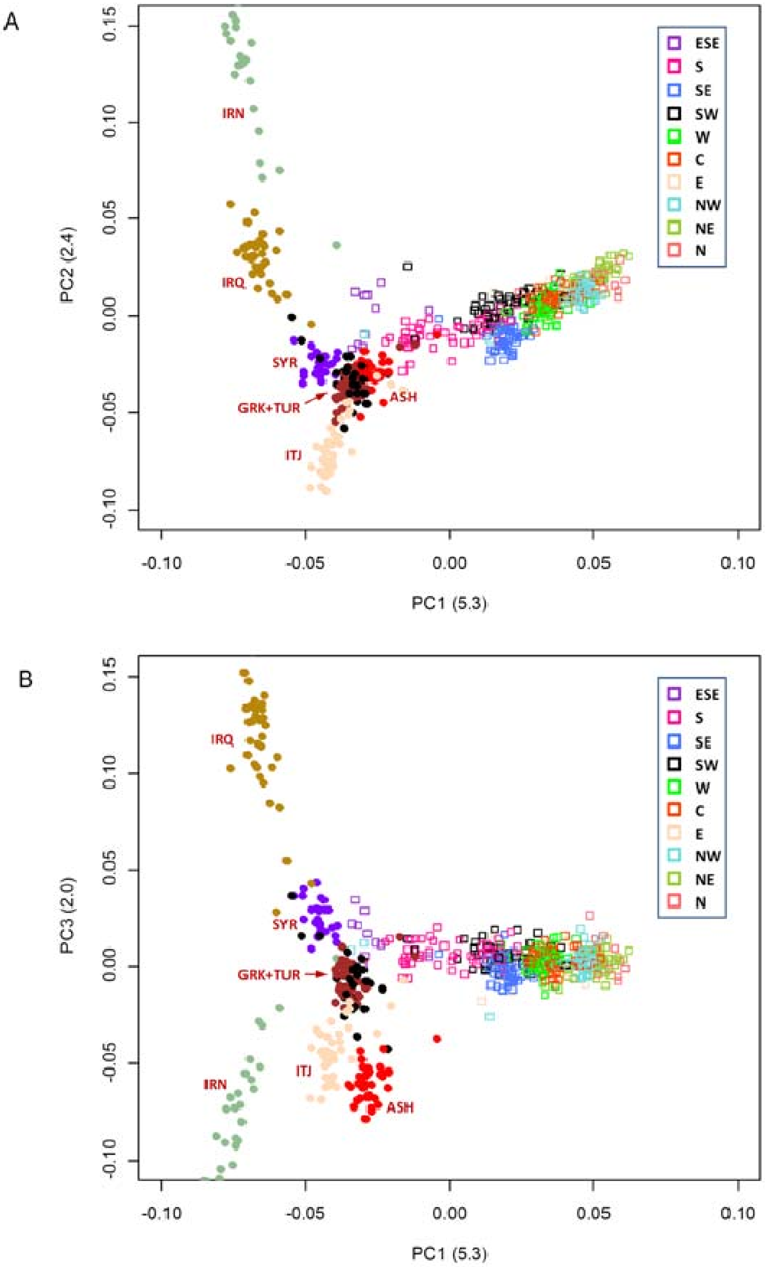
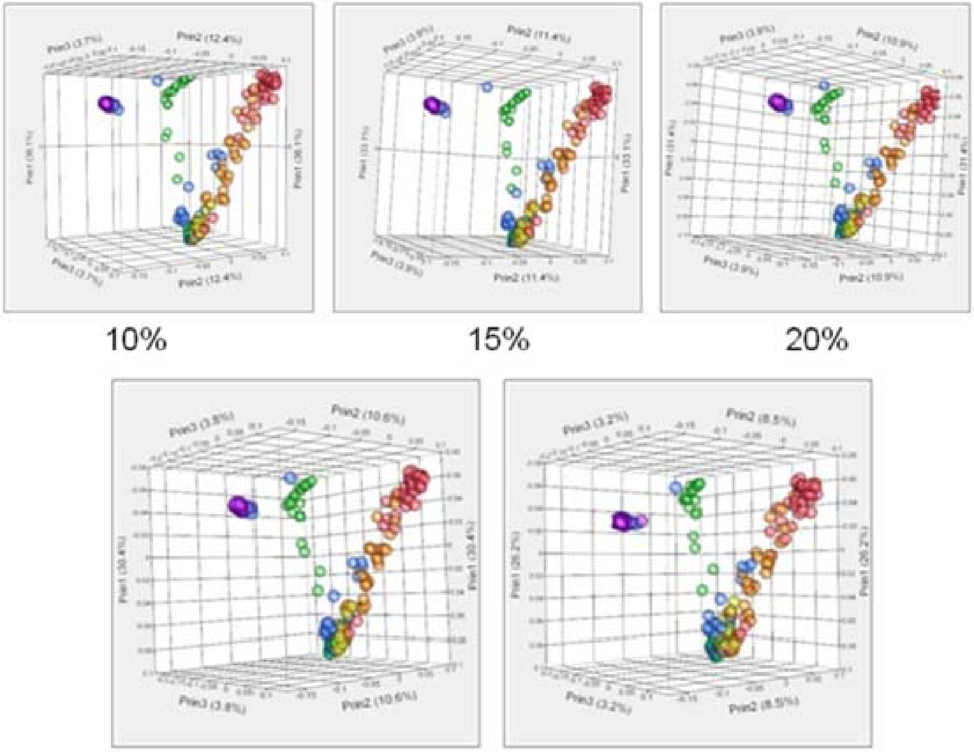
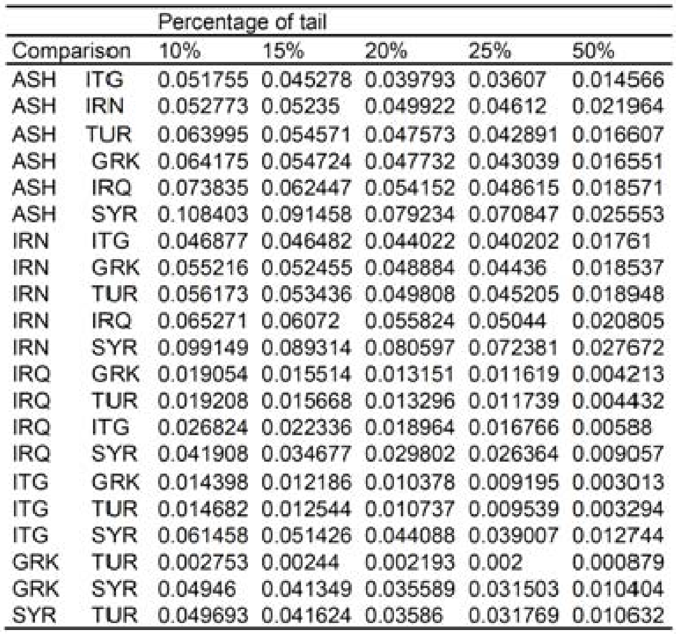


Figure S1. Principal component analysis of Jewish populations combined with selected PopRes populations. The Jewish samples are represented by large filled circles. Colored open squares represent PopRes samples grouped by different geographical origin in Europe. ESE, East Southeastern; S, Southern; SE, Southeastern; SW, Southwestern; W, Western; C, Central; E, Eastern; NW, Northwestern; NE, Northeastern; N, Northern. The list of countries from each geographical group can be found in supplemental methods.



10% 15% 20%

25% 50%



		Percentage of tail				
Comparison		10%	15%	20%	25%	50%
ASH	ITG	0.051755	0.045278	0.039793	0.03607	0.014566
ASH	IRN	0.052773	0.05235	0.049922	0.04612	0.021964
ASH	TUR	0.063995	0.054571	0.047573	0.042891	0.016607
ASH	GRK	0.064175	0.054724	0.047732	0.043039	0.016551
ASH	IRQ	0.073835	0.062447	0.054152	0.048615	0.018571
ASH	SYR	0.108403	0.091458	0.079234	0.070847	0.025553
IRN	ITG	0.046877	0.046482	0.044022	0.040202	0.01761
IRN	GRK	0.055216	0.052455	0.048884	0.04436	0.018537
IRN	TUR	0.056173	0.053436	0.049808	0.045205	0.018948
IRN	IRQ	0.065271	0.06072	0.055824	0.05044	0.020805
IRN	SYR	0.099149	0.089314	0.080597	0.072381	0.027672
IRQ	GRK	0.019054	0.015514	0.013151	0.011619	0.004213
IRQ	TUR	0.019208	0.015668	0.013296	0.011739	0.004432
IRQ	ITG	0.026824	0.022336	0.018964	0.016766	0.00588
IRQ	SYR	0.041908	0.034677	0.029802	0.026364	0.009057
ITG	GRK	0.014398	0.012186	0.010378	0.009195	0.003013
ITG	TUR	0.014682	0.012544	0.010737	0.009539	0.003294
ITG	SYR	0.061458	0.051426	0.044088	0.039007	0.012744
GRK	TUR	0.002753	0.00244	0.002193	0.002	0.000879
GRK	SYR	0.04946	0.041349	0.035589	0.031503	0.010404
SYR	TUR	0.049693	0.041624	0.03586	0.031769	0.010632

Figure S2. PCA calculation within selected tails (i.e 10,15,20,25 and 50% of genetic distance distribution) of total genomic CNVs (946k).

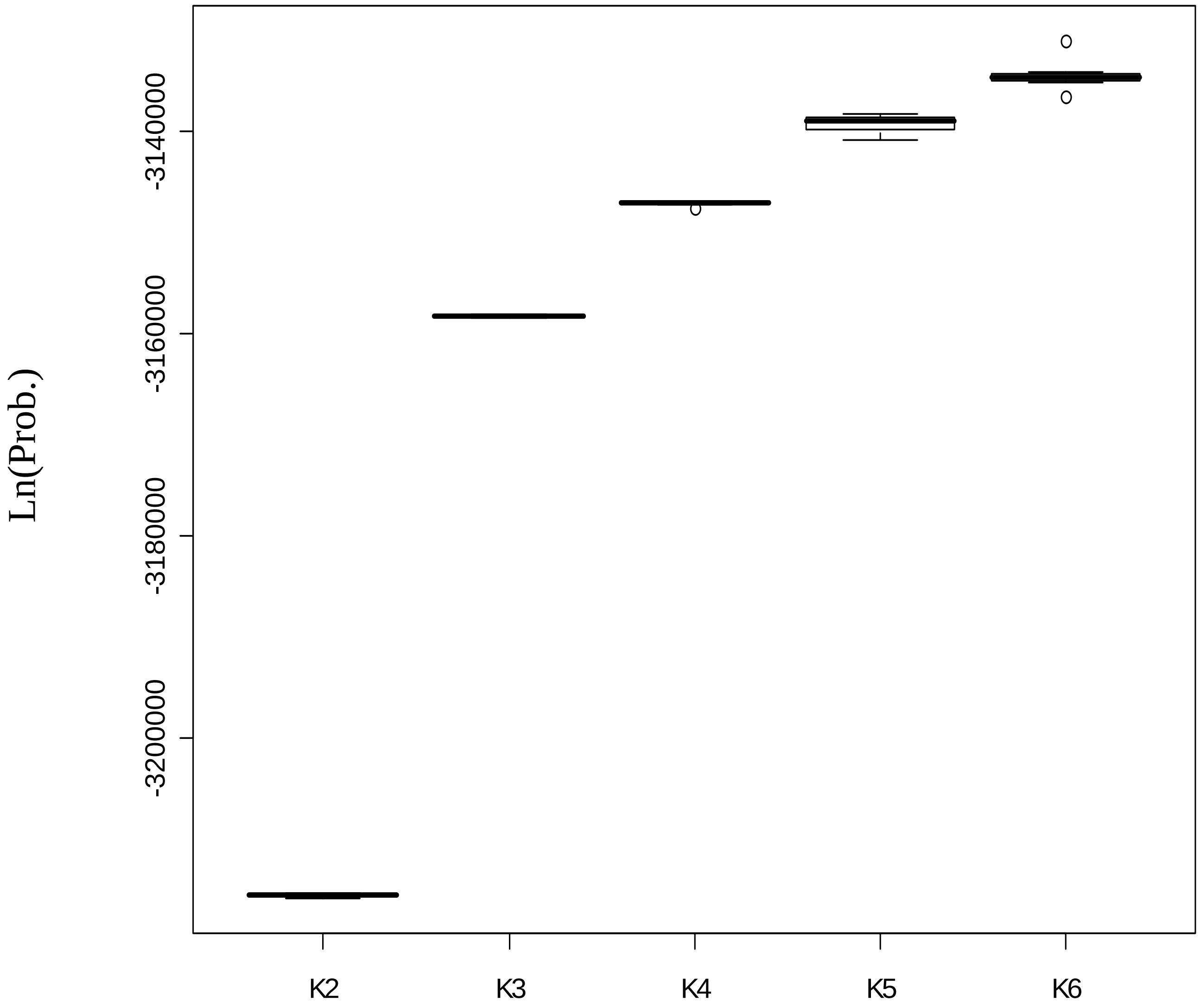


Figure S3. Likelihood scores for 10 iterations for Jewish and selected HGDP samples for K=2-6.

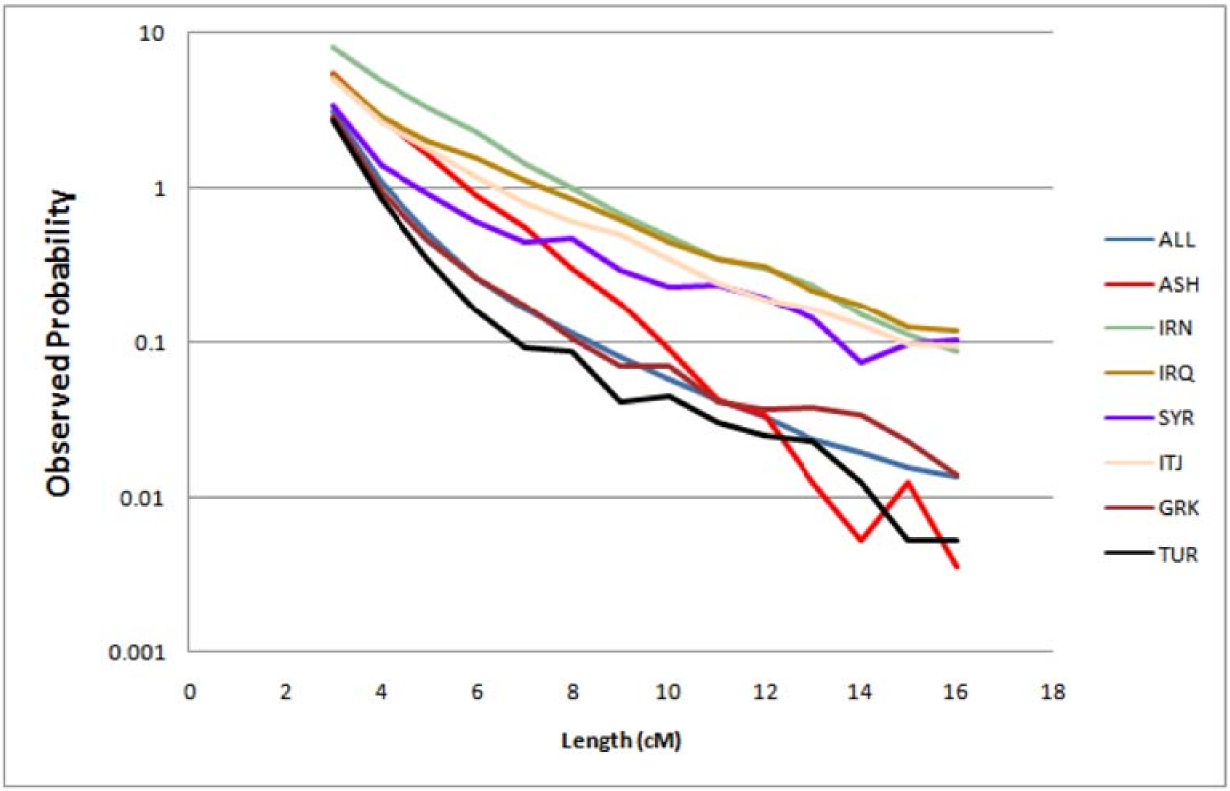


Figure S4. Distribution of segment lengths within each Jewish population. The expected number of IBD segments shared within each Jewish population (Y axis, logarithmic scale) for the discrete segment length range of 3 cM to 16 cM (X axis) were computed. An exponential decay rate (reflected by a linear behavior in the logarithmic scale) is representative of a recent population bottleneck. The behavior exhibited by the Ashkenazi population is consistent with the historical reports of a severe bottleneck followed by a rapid expansion, whereas the Greek and Turkish population decay suggests a more outbred profile.

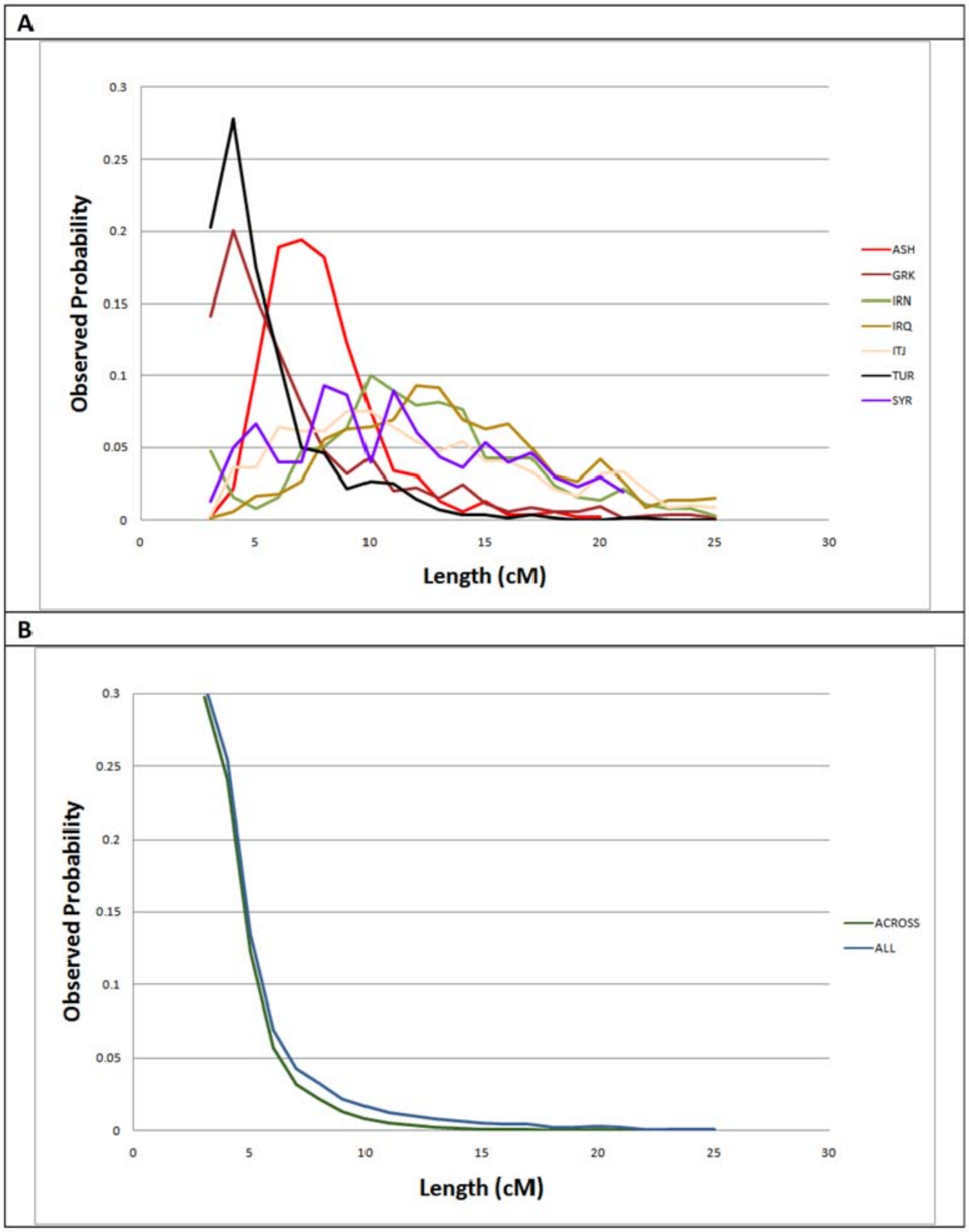


Figure S5. Distribution of longest IBD segments. The longest segments shared by the average pair of individuals as a function of the segment's length were computed. This distribution provides insights on the population's average time to most recent common ancestor. (A) Distribution of longest segments within each Jewish population. The average time to most common recent ancestor is larger for Turkish and Greek population, suggesting a more outbred profile than the other Jewish groups. (B) Distribution of longest segments for all possible pairs of Jewish individuals compared to the distribution of longest segments limited to pairs from different populations. The distribution for all possible pairs of individuals is shifted towards longer segments, reflecting the presence of closer relationships within populations. Short length for longest segments is expected across populations.

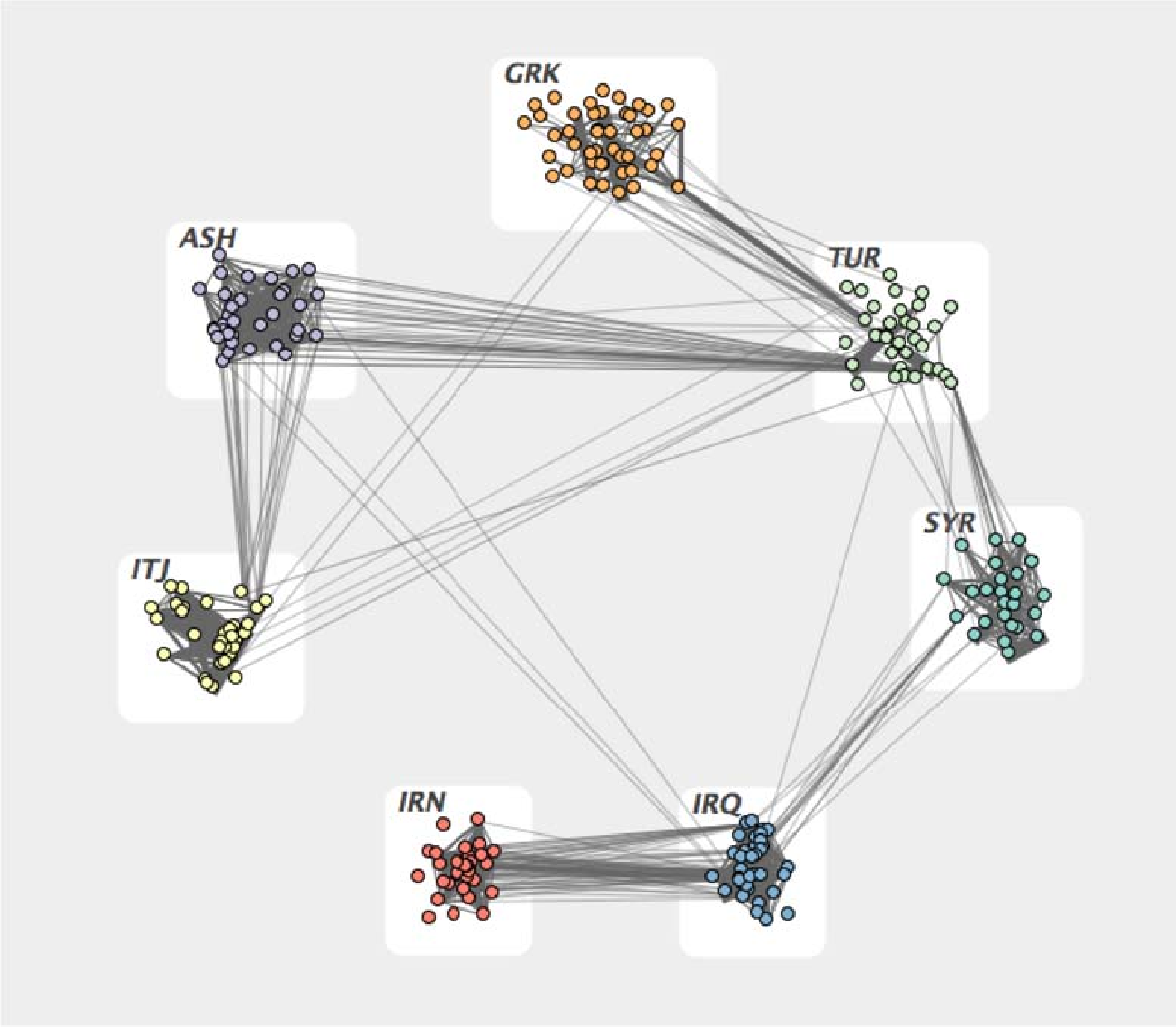


Figure S6. Network of relatedness between individuals. Each individual is represented by a node. Nodes are linked if the corresponding individuals are sharing >50cM of their genomes IBD.

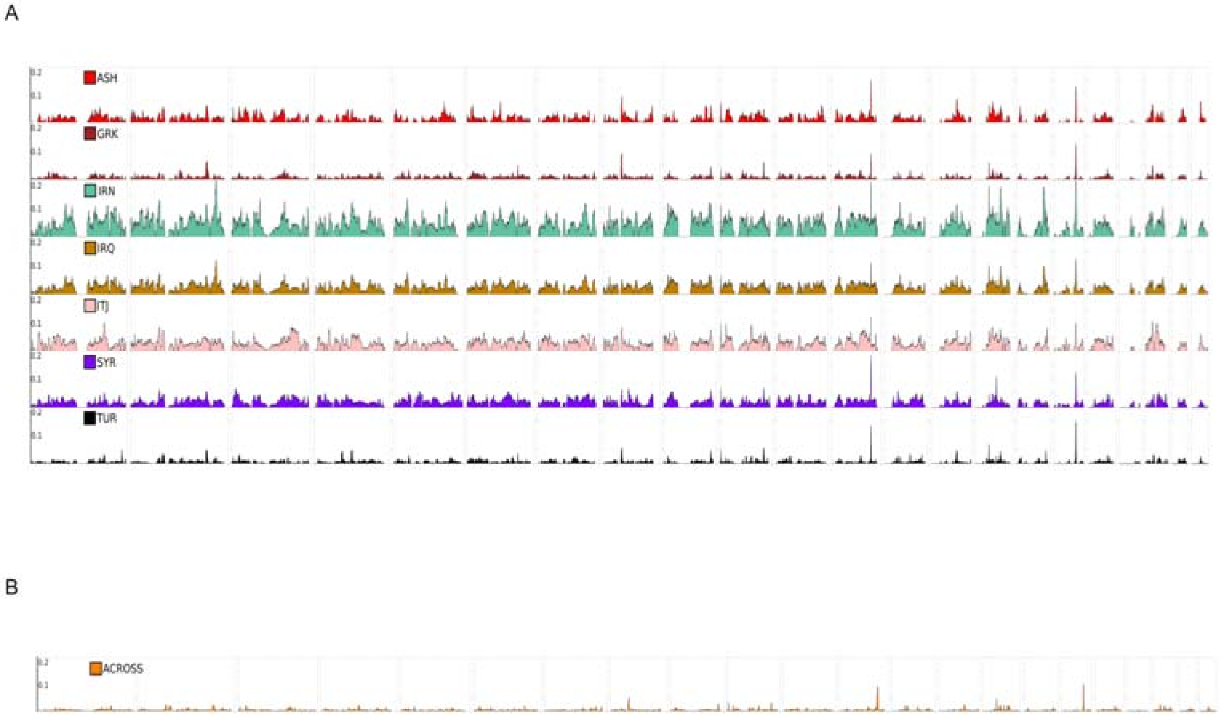


Figure S7. Manhattan plots of IBD sharing in Jewish populations. Normalized amount of sharing (Y axis) shown as a function of genomic position (X axis). (A) Sharing within populations for each Jewish group. (B) Sharing across populations.

Figure S8. LD structure of IBD sharing peaks. Regions of intense sharing across Jewish populations were processed using the Haploview software package. The structure of LD blocks in the analyzed regions does not justify the observed levels of IBD sharing.

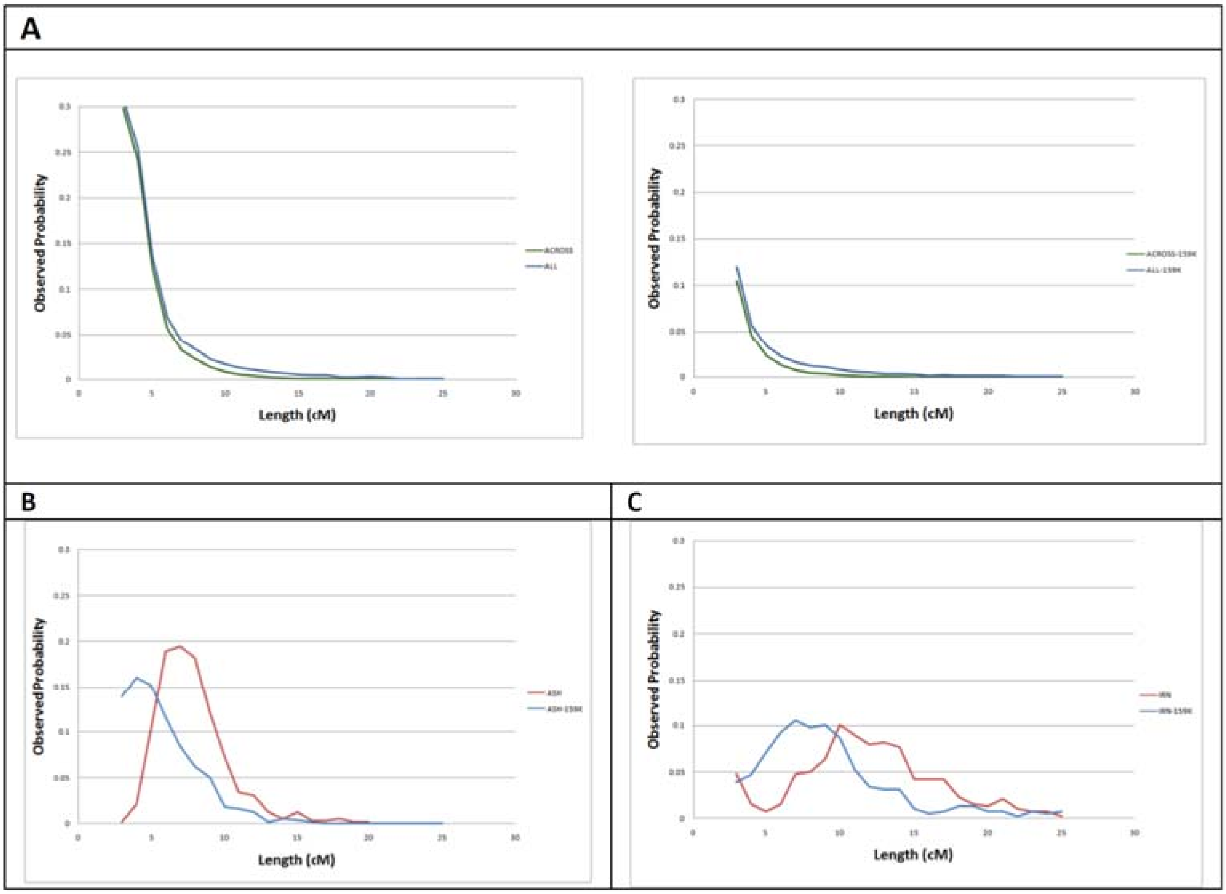


Figure S9 – Evaluation of power for IBD discovery based on a sparse set of SNPs. The results of the IBD analysis for the Jewish populations using the full set of markers and the sparser set of SNPs from the platforms' consensus were compared. (A) Distribution of longest segments for all possible Jewish pairs and limited to pairs across different populations. Comparison of the results for the full set (left) and the sparse set of markers (right). (B-C) Comparison of the results for the distribution of longest segments in the Ashkenazi and Iranian populations. Qualitatively similar results are obtained. Stricter parameters were employed in the IBD discovery phase for the sparser set of markers: long segments were excised, shifting the distributions towards shorter segments.

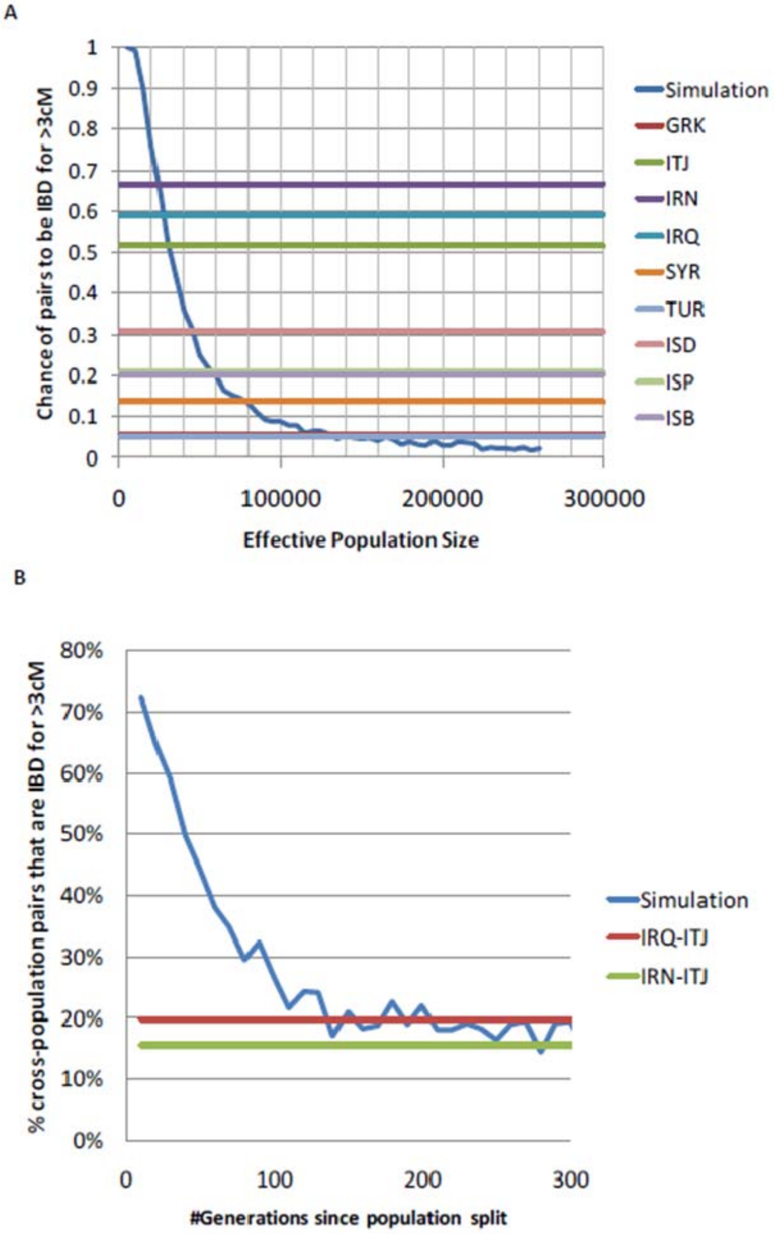


Figure S10. Inference of population history based on the observed fraction of individual pairs sharing a segment IBD (Y-axis) (A) First stage: simulation of a single population changing the ancestral population size (X-axis) to estimate the size appropriate for the Jewish and Middle Eastern non-Jewish populations. (B) Second stage : 2-population simulations changing the time of splits between Middle Eastern Jews and European (Italian) Jews is consistent with a population split 100-150 generations ago; the split between either of these groups and non-Jews is inconsistent with these simulation assumptions, as it predicts a distribution of shared segment lengths that does not fit any such scenario (data not shown), suggesting a more complex history than a simple split if a single ancestral population.[38]

Table S1. - Individuals genotyped, including country of origin and gender.

JHM ID	Pop.	Gender	JHM ID	Pop.	Gender	JHM ID	Pop.	Gender
3	IRQ	Female	307	SYR	Male	626	GRK	Female
4	IRQ	Male	308	SYR	Female	627	GRK	Male
5	IRQ	Male	311	SYR	Female	628	GRK	Male
6	IRQ	Male	312	SYR	Male	631	GRK	Female
7	IRQ	Female	313	SYR	Male	632	GRK	Female
15	IRQ	Female	315	SYR	Male	633	GRK	Male
17	IRQ	Female	316	SYR	Female	634	GRK	Female
21	IRQ	Female	317	SYR	Female	635	GRK	Male
23	IRQ	Female	318	SYR	Female	636	GRK	Female
24	IRQ	Male	323	SYR	Male	637	GRK	Female
28	IRQ	Female	325	SYR	Female	638	GRK	Male
29	IRQ	Female	330	SYR	Female	640	GRK	Female
30	IRQ	Male	331	SYR	Male	642	GRK	Male
32	IRQ	Male	332	SYR	Female	643	GRK	Female
33	IRQ	Male	333	SYR	Male	644	GRK	Female
36	IRQ	Male	336	SYR	Male	646	GRK	Male
37	IRQ	Female	341	SYR	Male	647	GRK	Male
40	IRQ	Female	351	SYR	Male	650	GRK	Male
52	TUR	Male	355	SYR	Female	651	GRK	Female
58	TUR	Male	358	SYR	Male	652	GRK	Male
59	TUR	Male	367	SYR	Male	654	GRK	Female
66	TUR	Female	368	SYR	Female	655	GRK	Male
68	TUR	Male	370	SYR	Female	657	GRK	Female
70	TUR	Female	371	IRQ	Female	662	GRK	Female
71	TUR	Male	375	IRQ	Female	663	GRK	Male
72	TUR	Male	378	SYR	Male	665	GRK	Female
76	TUR	Female	393	IRQ	Male	901	ITJ	Male
82	IRN	Female	399	IRQ	Male	902	ITJ	Female
83	IRN	Female	400	IRQ	Female	904	ITJ	Male
85	IRN	Male	402	IRQ	Male	909	ITJ	Female
87	IRN	Female	403	IRQ	Female	912	ITJ	Male
101	IRN	Male	440	IRQ	Female	913	ITJ	Female
104	IRN	Male	455	IRQ	Female	916	ITJ	Male
106	IRN	Female	457	IRQ	Male	917	ITJ	Female
129	IRN	Male	600	GRK	Male	1051	TUR	Female
136	IRN	Female	601	GRK	Female	1052	TUR	Female
148	IRN	Male	604	GRK	Male	1053	TUR	Female
151	IRN	Female	605	GRK	Male	1056	TUR	Female
157	IRN	Male	606	GRK	Male	1057	TUR	Female
158	IRN	Female	607	GRK	Female	1058	TUR	Female
159	IRN	Female	609	GRK	Male	1059	TUR	Female
164	IRN	Female	610	GRK	Male	1060	TUR	Male
167	IRN	Male	611	GRK	Male	1061	TUR	Male
180	IRN	Female	612	GRK	Female	1062	TUR	Female
188	IRN	Male	613	GRK	Male	1064	TUR	Female
196	IRN	Male	616	GRK	Male	1065	TUR	Female
268	IRN	Male	619	GRK	Male	1066	TUR	Female
273	IRN	Male	622	GRK	Male	1067	TUR	Female
276	IRN	Female	623	GRK	Female	1068	TUR	Male
305	SYR	Female	624	GRK	Female	1073	TUR	Female

Table S2. 95% confidence intervals for pairwise *Fst* between individual Jewish populations and between individual Jewish populations and populations in the HGDP panel

	IRN	**IRQ**	**SYR**	**ASH**	**ITJ**	**GRK**	**TUR**	**N_Italian**	**Sardinian**	**French**	**Basque**	**Adygei**	**Russian**	**Palestinian**	**Druze**	**Bedouin**
IRN		0.01577-0.01589	0.01618-0.01632	0.01636-0.01648	0.01886-0.01901	0.01518-0.01530	0.01521-0.01534	0.01925-0.01940	0.02839-0.02858	0.02229-0.02244	0.03074-0.03094	0.01763-0.01779	0.02835-0.02853	0.01633-0.01645	0.01700-0.01712	0.02163-0.02176
IRQ			0.00801-0.00811	0.01358-0.01370	0.01215-0.01225	0.00849-0.00858	0.00827-0.00836	0.01202-0.01215	0.02053-0.02067	0.01642-0.01654	0.02422-0.02438	0.01250-0.01263	0.02414-0.02430	0.01108-0.01118	0.01178-0.01188	0.01541-0.01553
SYR				0.00972-0.00984	0.00822-0.00831	0.00336-0.00343	0.00344-0.00352	0.00762-0.00774	0.01474-0.01488	0.01142-0.01154	0.01916-0.01932	0.00938-0.00951	0.01871-0.01886	0.00728-0.00737	0.00843-0.00854	0.01080-0.01091
ASH					0.00924-0.00934	0.00611-0.00619	0.00534-0.00542	0.00951-0.00963	0.01615-0.01629	0.01087-0.01097	0.01756-0.01772	0.01255-0.01269	0.01740-0.01754	0.01194-0.01204	0.01285-0.01296	0.01752-0.01764
ITJ						0.00508-0.00515	0.00514-0.00521	0.00877-0.00888	0.01455-0.01467	0.01175-0.01186	0.01818-0.01832	0.01287-0.01301	0.01877-0.01891	0.01054-0.01063	0.01099-0.01109	0.01483-0.01494
GRK							0.00048-0.00053	0.00417-0.00427	0.01049-0.01059	0.00730-0.00738	0.01390-0.01402	0.00853-0.00865	0.01444-0.01456	0.00591-0.00598	0.00697-0.00704	0.01052-0.01061
TUR								0.00424-0.00434	0.01054-0.01066	0.00700-0.00710	0.01323-0.01335	0.00835-0.00847	0.01436-0.01448	0.00547-0.00555	0.00612-0.00620	0.00948-0.00957
N_Italian									0.00748-0.00760	0.00190-0.00199	0.00799-0.00811	0.00702-0.00715	0.00826-0.00839	0.01112-0.01124	0.01046-0.01058	0.01744-0.01759
Sardinian										0.00971-0.00981	0.01397-0.01409	0.02039-0.02056	0.02036-0.02052	0.01759-0.01772	0.01779-0.01792	0.02331-0.02346
French											0.00726-0.00737	0.01014-0.01027	0.00554-0.00564	0.01459-0.01469	0.01424-0.01436	0.02100-0.02114
Basque												0.01810-0.01827	0.01426-0.01439	0.02050-0.02064	0.02096-0.02111	0.02733-0.02750
Adygei													0.01334-0.01349	0.01222-0.01235	0.01200-0.01213	0.01769-0.01783
Russian														0.02125-0.02139	0.02187-0.02202	0.02808-0.02824
Palestinian															0.00807-0.00815	0.00825-0.00833
Druze																0.01217-0.01227
Bedouin																

Table S3. Pairwise *Fst* between major Jewish groups (Middle-East Jews and European/Syrian Jews) and the populations in HGDP panel.

	Middle-East Jews	**European/Syrian Jews**
European/Syrian Jews	**0.00677** (0.00675-0.00680)	
N_Italian	**0.01124** (0.01119-0.01130)	**0.00462** (0.00458-0.00465)
Sardinian	**0.01995** (0.01988-0.02001)	**0.01085** (0.01081-0.01090)
French	**0.01505** (0.01500-0.01510)	**0.00738** (0.00734-0.00741)
Basque	**0.02310** (0.02302-0.02317)	**0.01395** (0.01390-0.01400)
Adygei	**0.01071** (0.01065-0.01077)	**0.00822** (0.00817-0.00827)
Russian	**0.02212** (0.02204-0.02219)	**0.01446** (0.01441-0.01451)
Palestinian	**0.00971** (0.00968-0.00975)	**0.00635** (0.00632-0.00638)
Druze	**0.01021** (0.01017-0.01025)	**0.00701** (0.00699-0.00704)
Bedouin	**0.01457** (0.01452-0.01463)	**0.01101** (0.01097-0.01105)

Table S4. ANOVA of population PCAs average.

PCA1	**IRN**	**IRQ**	**SYR**	**ASH**	**ITJ**	**GRK**	**TUR**	**N_Italian**	**Sardinian**	**French**	**Basque**	**Adygei**	**Russian**	**Palestinian**	**Druze**	**Bedouin**
IRN		0.1903	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0005	<.0001	<.0001
IRQ			<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0067	<.0001
SYR				<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
ASH					<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0349	<.0001	<.0001	<.0001	<.0001
ITJ						0.5137	0.6861	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
GRK							0.2918	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
TUR								<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
N_Italian									0.2844	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Sardinian										<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
French											<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Basque												<.0001	<.0001	<.0001	<.0001	<.0001
Adygei													0.3992	<.0001	<.0001	<.0001
Russian														<.0001	<.0001	<.0001
Palestinian															<.0001	<.0001
Druze																<.0001

PCA2	**IRN**	**IRQ**	**SYR**	**ASH**	**ITJ**	**GRK**	**TUR**	**N_Italian**	**Sardinian**	**French**	**Basque**	**Adygei**	**Russian**	**Palestinian**	**Druze**	**Bedouin**
IRN		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
IRQ			<.0001	<.0001	<.0001	0.003266	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
SYR				0.780547	0.073169	<.0001	0.011957	<.0001	<.0001	<.0001	<.0001	0.002773	<.0001	<.0001	0.757846	<.0001
ASH					0.023892	0.000415	0.002467	<.0001	<.0001	<.0001	<.0001	0.00349	<.0001	<.0001	0.977098	<.0001
ITJ						0.211799	0.400707	0.0006841	<.0001	<.0001	<.0001	<.0001	0.002627	<.0001	0.020277	<.0001
GRK							0.722138	0.015067	<.0001	0.000217	<.0001	<.0001	0.047883	<.0001	0.0003	<.0001
TUR								0.00843921	<.0001	0.000117	<.0001	<.0001	0.027345	<.0001	0.001947	<.0001
N_Italian									<.0001	0.373432	<.0001	<.0001	0.610297	0.1739838	<.0001	<.0001
Sardinian										0.000427	0.317682	<.0001	<.0001	0.00078806	<.0001	<.0001
French											<.0001	<.0001	0.138662	0.6532716	<.0001	<.0001
Basque												<.0001	<.0001	<.0001	<.0001	<.0001
Adygei													<.0001	0.04315468	0.003443	<.0001
Russian														<.0001	<.0001	<.0001
Palestinian															<.0001	<.0001
Druze																<.0001

PCA3	**IRN**	**IRQ**	**SYR**	**ASH**	**ITJ**	**GRK**	**TUR**	**N_Italian**	**Sardinian**	**French**	**Basque**	**Adygei**	**Russian**	**Palestinian**	**Druze**	**Bedouin**
IRN		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.002932	<.0001	<.0001	<.0001	<.0001
IRQ			<.0001	<.0001	<.0001	<.0001	<.0001	0.58983262	<.0001	<.0001	0.111731		<.0001	<.0001	<.0001	<.0001
SYR				0.074604	<.0001	0.062977	0.787726	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.01334626	0.024763	<.0001
ASH					<.0001	<.0001	0.02597	0.0007077	<.0001	<.0001	<.0001	<.0001	<.0001	0.48075355	0.629081	<.0001
ITJ						<.0001	<.0001	<.0001	0.00508519	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
GRK							0.08374	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
TUR								<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.00271076	0.006211	<.0001
N_Italian									<.0001	<.0001	0.059245	<.0001	<.0001	0.00410764	0.002624	<.0001
Sardinian										<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
French											<.0001	<.0001	<.0001	<.0001	<.0001	0.000368
Basque												<.0001	<.0001	<.0001	<.0001	<.0001
Adygei													<.0001	0.01904813	<.0001	<.0001
Russian														<.0001	<.0001	<.0001
Palestinian															0.828882	<.0001
Druze																<.0001

Table S5. Permutation tests for between-group identity by state (IBS) differences

Populations	***N***	**IRN**	**IRQ**	**SYR**	**ASH**	**ITJ**	**GRK**	**TUR**
IRN	28		$1.0X10^{-5}$	$1.0X10^{-5}$	$1.0X10^{-5}$	$1.0X10^{-5}$	$1.0X10^{-5}$	$1.0X10^{-5}$
IRQ	37			$1.0X10^{-5}$	$1.0X10^{-5}$	$1.0X10^{-5}$	$1.0X10^{-5}$	$1.0X10^{-5}$
SYR	25				$1.0X10^{-5}$	$1.0X10^{-5}$	$1.0X10^{-5}$	$1.0X10^{-5}$
ASH	34					$1.0X10^{-5}$	$1.0X10^{-5}$	$1.0X10^{-5}$
ITJ	37						$1.0X10^{-5}$	$1.0X10^{-5}$
GRK	42							$1.0X10^{-5}$
TUR	34							

Table S6 – Gene content for regions of high IBD sharing

Chr	**Start**	**End**	**Genes**
2	193,679,587	196,034,311	--
2	72,175,503	74,001,136	EXOC6B, EMX1, ALMS1, STAMBP, SPR, RAB11FIP5, FBXO41, EGR4, NAT8, DUSP11, ACTG2
8	50,193,145	53,996,125	SNTG1, RB1CC1
9	124,440,201	126,229,508	LHX2, STRBP, CRB2, NEK6, PDCL, GPR21, PSMB7
10	116,184,200	117,850,842	TRUB1
10	2,728,100	3,108,632	--
12	108,046,843	112,554,949	CDV1, TPCN1, ACACB, RAD9B, MYL2, SH2B3, ATXN2, PTPN11, RPH3A, FOXN4, MMAB, MVK, TRPV4, ATP2A2, TECT1, PPP1CC, ALDH2, MAPKAPK5, C12orf8, RPL6, OAS1, OAS3, OAS2, DTX1, RASAL1, DDX54, SLC24A6, LHX5
14	63,153,718	65,330,133	MTHFD1, FUT8, SYNE2, SPTB, MAX, ESR2, ZBTB25, HSPA2, GPX2, FNTB
15	37,023,662	38,192,742	THBS1, GPR176, SRP14, BMF
17	50,619,996	52,424,032	HLF, PCTP, NOG, DGKE, TRIM25, COIL
20	20,304,906	22,377,988	XRN2, NKX2-4, NKX2-2, PAX1