

Supplementary Tables and Figures for manuscript:

Genetic variation in Tunisia in the context of human diversity worldwide

Authors:

Lotfi Cherni^{1,2}, Andrew J. Pakstis³, Sami Boussetta¹, Sarra Elkamel¹, Sabehe Frigi¹, Houssein Khodjet-El-Khil¹, Alison Barton³, Eva Haigh³, William C. Speed³, Amel Ben Ammar Elgaaied¹, Judith R. Kidd³, Kenneth K. Kidd³

¹ Laboratory of Genetics, Immunology and Human Pathology, Science Faculty of Tunis, University of Tunis El Manar, 2092 Tunis, Tunisia

² High Institute of Biotechnology, University of Monastir, 5000 Monastir, Tunisia.

³ Department of Genetics, Yale University School of Medicine, New Haven, CT 06520 USA

submitted to

American Journal of Physical Anthropology

SUPPLEMENTARY TABLES AND FIGURES

Table S1A. The 65 populations studied by geographical region. Table shows the sample size for each group, population abbreviations employed in figures, and the rank ordering of populations in figures. The same relative order for these populations occurs in joint analyses of 73 populations.

Table S1B. The 8 populations from Henn et al. (2012) included in the combined analysis with the populations listed in Table S1A. Sample sizes and population abbreviations employed are shown. The order column shows the position of these 8 populations in figures reporting analyses based on the 73 populations.

Table S2. The dbSNP rs-numbers for the 399 SNPs studied are presented by chromosome and nucleotide position. They defined a total of 299 polymorphic markers in the 65 population analysis. The third column indicates the 159 SNPs that defined 59 multi-snp haplotypes; the other 240 SNPs were single site markers. The second column indicates the 102 SNPs that overlap with the dataset of Henn et al. (2012) and also had sufficiently complete genotypes so that they could be utilized in the combined analysis of 73 populations. A total of 90 polymorphic markers were included in the 73 population dataset; 21 of the 102 SNPs defined 9 multi-SNP haplotypes and the other 81 SNPs were single marker sites.

Table S3. List of 9 functional SNPs plus 9 nearby non-functional SNPs included in multi-SNP haplotypes

Figure S1. Allele frequency barplot on 65 populations for example haplotype

Figure S2. Comparing average heterozygosities of populations in 299 marker and 90 marker datasets

Figure S3. Principal Components Analysis plot for pc #1 and pc #3. PCA is based on pairwise Tau genetic distances for 65 populations.

Figure S4. Principal Components Analysis plot—73 populations, 90 markers study—pc #1 x pc #2

Figure S5. Principal Components Analysis plot—73 populations, 90 markers study —pc #1 x pc #3

Figure S6. STRUCTURE individual bar plots—73 population study, 90 markers study—displaying results for runs with highest likelihood out of 20 runs in each cluster K=5 to 11. Each individual has a separate column in the bar plot and the individuals in a population are clustered together in the display but the STRUCTURE analysis was not informed about an individual's population membership. Black vertical lines identify the population boundaries. The height extent of each color for an individual corresponds to the estimated membership of the individual in one of the clusters; each cluster is assigned a separate color. The bars with multiple colors can be interpreted as genetic admixture or as relative probabilities of belonging to the different clusters. Since there is a separate color bar for each of the more than 3000 individuals, the interval width covered by a population varies and corresponds to the number of individuals in the population.

Figures S7 to S17. Allele frequency bar plots of functional SNPs and some nearby non-functional SNPs included in multi-allelic haplotypes.

Figure S7. SNP allele frequency plot--EDAR chr. 2 rs3827760

Figure S8. Haplotype allele frequency plot—EDAR chr. 2 rs3827760 (fn), rs2169812, rs2378217, rs940928

Figure S9. SNP allele frequency plot—ADH1B chr. 4 rs1229984

Figure S10. Haplotype allele frequency plot—TAS2R1 chr. 5 rs41461, rs2234233 (fn)

Figure S11. SNP allele frequency plot—TAS2R1 chr. 5 rs2234233 (fn)

Figure S12. SNP allele frequency plot—SLC45A2 chr. 5 rs16891982 (fn)

Figure S13. Haplotype allele frequency plot—SLC45A2 chr. 5 rs11955798, rs7737265, rs16891982 (fn), rs28777

Figure S14. Haplotype allele frequency plot—SLC45A2 chr. 5 rs26722 (fn), rs35408

Figure S15. Haplotype allele frequency plot—TAS2R38 chr. 7 rs1726866 (fn), rs713598 (fn)

Figure S16. SNP allele frequency plot—HERC2 chr. 15 rs12913832

Figure S17. Haplotype allele frequency plot—MC1R chr. 16 rs3212363, rs885479 (fn)

Figures S18 to S23. Principal Components Analysis (PCA) plots focused on N.Afr and SW Asia.

Figure S18. PCA #1x#2 for 7 Tunisian, 1 Libyan populations, 299 polymorphisms

Figure S19. PCA #1x#3 for 7 Tunisian, 1 Libyan populations, 299 polymorphisms

Figure S20. PCA #2x#3 for 7 Tunisian, 1 Libyan populations, 299 polymorphisms

Figure S21. PCA #1x#2 for 8 N Afr (Tunisian, Libyan) populations & 4 SWAsian populations, 299 polymorphisms

Figure S22. PCA #1x#2 for 15 North African populations, 90 polymorphisms

Figure S23. PCA #1x#2 for 15 North African populations and 4 SW Asian populations, 90 polymorphisms

Table S1A. 65 Populations studied; order employed in tables and figures

Geographical Region	Order	Population Sample	Population Abbreviation	Number of Individuals
Africa, SubSahara	1	Biaka	BIA	66
	2	Mbuti	MBU	38
	3	Lisongo, Cen.Afr.Rep.	LIS	7
	4	Yoruba, Nigeria	YOR	77
	5	Ibo, Nigeria	IBO	47
	6	Hausa, Nigeria	HSA	38
	7	Chagga, Tanzania	CGA	45
	8	Masai, Tanzania	MAS	20
	9	Sandawe, Tanzania	SND	40
	10	Zaramo, Tanzania	ZRM	37
	11	African Americans	AAM	89
	12	Ethiopian Jews	ETJ	31
Africa, North	13	Nebeur, NW Tunisia	NEB	10
	14	Kesra, NW Tunisia	KSR	43
	15	Kairoun, Central Tunisia	KRN	41
	16	Sousse, Central Tunisia	SOU	43
	17	Mehdia, Central Tunisia	MHD	33
	18	Kerkennah, S. Tunisia	KRK	40
	19	Smar, S.Tunisia	SMR	57
	20	Libyans-from 6 cities	LYB	64
Asia, South West	21	Yemenite Jews	YMJ	41
	22	Kuwaiti	KWT	15
	23	Druze	DRU	101
	24	Samaritans	SAM	38
	25	Ashkenazi	ASH	79
Europe	26	Sardinians	SRD	34
	27	Roman Jews	RMJ	26
	28	Adygei	ADY	54
	29	Chuvash	CHV	41
	30	Hungarians	HGR	89
	31	Russians, Archangelsk	RUA	33
	32	Russians, Vologda	RUV	47
	33	European Americans	EAM	88
	34	Irish	IRI	111
	35	Danes	DAN	51
	36	Finns	FIN	34
Siberia, North West	37	Komi Zyrian	KMZ	46
Asia, South Central	38	Keralites	KER	30

	39	Thoti	THT		13
	40	Kachari	KCH		17
Siberia, North West	41	Khanty	KTY		49
Siberia, North East	42	Yakut	YAK		51
Asia, East	43	Koreans	KOR		54
	44	Japanese	JPN		44
	45	Chinese, San Francisco	CHS		56
	46	Chinese, Taiwan	CHT		48
	47	Hakka, Taiwan	HKA		41
	48	Laotians	LAO		118
	49	Cambodians	CBD		23
	50	Ami, Taiwan	AMI		40
	51	Atayal, Taiwan	ATL		41
	52	Malaysians	MLY		10
Pacific Islands	53	Samoans	SMO		9
	54	Micronesians	MCR		34
	55	Papuans, New Guinea	PNG		22
	56	Nasioi	NAS		23
America, North	57	Plains Amerindians	NPA		55
	58	Southwest Amerindians	SWA		50
	59	Pima, Mexico	PMM		53
	60	Maya, Yucatan	MAY		48
America, South	61	Guihiba, Colombia	GHB		11
	62	Quechua, Peru	QUE		22
	63	Ticuna, Amazon	TIC		65
	64	Rondonian Surui, Amazon	SUR		43
	65	Karitiana, Amazon	KAR		50
		Total Individuals			2,914

Table S1B. The 8 populations from Henn et al. (2012) included in combined analysis with the 65 populations in Table S1A.

Geographical Region	Order	Population Sample	Population Abbreviation	Number of Individuals
North Africa	12.1	Saharawi	SRW	18
	12.2	Moroccans, South	MRS	16
	12.3	Moroccans, North	MRN	18
	12.4	Algerians	ALG	19
	12.5	Tunisians—Chenini, Douiret	TNS	18
	20.1	Libyans—in/near Tripoli	LIB	17
	20.2	Egyptians	EGY	19
Europe	25.5	Spanish Basque	BSQ	20
		Total Individuals		145

Note: Extra detail on the sampling locations for the Tunisians and Libyans in the Henn et al. paper was kindly provided by Dr. David Comas. Chenini and Douiret are Berber villages in the southernmost governate (Tataouine) in Tunisia.

Table S2.
399 SNPs studied. Ordered by chromosome position;
subset of 159 SNPs organized into 59 haplotypes

SNP count	SNPs in 73 pop. analyses	Analyzed as Haplotype or Single SNP	dbSNP rs-number	Chr	GRCh38 nucleotide position	distance to next SNP in basepairs
1		H1	rs4648344	1	3,826,568	187
2	yes	H1	rs6663840	1	3,826,755	2,663,561
3		S	rs2986742	1	6,490,316	6,057,828
4		S	rs6541030	1	12,548,144	5,296,247
5		S	rs647325	1	17,844,391	9,760,796
6	yes	S	rs4908343	1	27,605,187	5,063,986
7		S	rs359955	1	32,669,173	9,225,426
8		S	rs1325502	1	41,894,599	13,303,100
9	yes	S	rs12130799	1	55,197,699	13,186,305
10		S	rs3118378	1	68,384,004	7,191,978
11		S	rs10489587	1	75,575,982	53,767
12		S	rs814856	1	75,629,749	111,056
13		S	rs12744608	1	75,740,805	45,845
14		S	rs1146635	1	75,786,650	80,468
15		S	rs5745429	1	75,867,118	25,376,889
16		S	rs3737576	1	101,244,007	49,906,006
17		S	rs7554936	1	151,150,013	8,054,880
18		H2	rs2814778	1	159,204,893	671
19		H2	rs12075	1	159,205,564	8,985,088
20	yes	S	rs1040404	1	168,190,652	17,989,248
21		S	rs1407434	1	186,179,900	18,641,949
22		S	rs2065160	1	204,821,849	7,791,692
23		S	rs4951629	1	212,613,541	3,847,545
24		H3	rs4528199	1	216,461,086	17
25		H3	rs6604596	1	216,461,103	14,248,945
26		S	rs699	1	230,710,048	489,803
27	yes	S	rs11122250	1	231,199,851	10,979,351
28		S	rs316873	1	242,179,202	
29	yes	S	rs798443	2	7,828,144	6,788,081
30		S	rs7421394	2	14,616,225	2,565,076
31		S	rs1876482	2	17,181,301	538,917
32		S	rs1834619	2	17,720,218	11,595,327
33		S	rs4666200	2	29,315,545	8,398,708
34		S	rs4670767	2	37,714,253	23,153,857
35		S	rs842639	2	60,868,110	39,612,388
36		S	rs1519654	2	100,480,498	331,159
37		S	rs7580771	2	100,811,657	163,185
38		S	rs2305160	2	100,974,842	7,922,303
39		H4	rs3827760	2	108,897,145	455
40		H4	rs2169812	2	108,897,600	136
41		H4	rs2378217	2	108,897,736	272

42		H4	rs940928	2	108,898,008	22,320
43		H5	rs260642	2	108,920,328	7,268
44		H5	rs13397666	2	108,927,596	18,443
45		H5	rs260714	2	108,946,039	17,243
46		H5	rs260690	2	108,963,282	6,575
47	yes	H6	rs260694	2	108,969,857	58
48		H6	rs11123719	2	108,969,915	66
49		H6	rs11691107	2	108,969,981	26,227,246
50		S	rs1375131	2	135,197,227	600,687
51		S	rs1807356	2	135,797,914	152,498
52		S	rs6754311	2	135,950,412	9,061,964
53		S	rs10496971	2	145,012,376	4,085,207
54		H7	rs2170607	2	149,097,583	156
55		H7	rs10497052	2	149,097,739	2,860,211
56	yes	S	rs4664511	2	151,957,950	2,085,862
57	yes	S	rs1863086	2	154,043,812	3,766,893
58	yes	S	rs10497191	2	157,810,705	119,652
59		S	rs6737672	2	157,930,357	20,811,454
60		S	rs2627037	2	178,741,811	21,415,420
61	yes	S	rs1569175	2	200,157,231	15,207,738
62	yes	S	rs11651	2	215,364,969	16,488,363
63		S	rs3098610	2	231,853,332	6,424,616
64	yes	S	rs2304672	2	238,277,948	
65		S	rs10510228	3	2,167,148	20,168,255
66		S	rs9880567	3	22,335,403	8,038,717
67	yes	S	rs4955316	3	30,374,120	8,730,818
68		S	rs9809104	3	39,104,938	7,205,926
69		H8	rs4513489	3	46,310,864	29
70	yes	H8	rs6441961	3	46,310,893	204,452
71	yes	H9	rs6808142	3	46,515,345	113
72	yes	H9	rs17030627	3	46,515,458	24,918,843
73		S	rs11713996	3	71,434,301	7,916,124
74	yes	S	rs6548616	3	79,350,425	34,800,806
75		H10	rs3773678	3	114,151,231	20,711
76		H10	rs3732783	3	114,171,942	26
77		H10	rs6280	3	114,171,968	6,631,901
78		S	rs12629908	3	120,803,869	841,457
79		S	rs1919550	3	121,645,326	95,416
80		S	rs12498138	3	121,740,742	14,454,892
81		S	rs9845457	3	136,195,634	11,836,934
82		S	rs734873	3	148,032,568	32,214,371
83		S	rs2030763	3	180,246,939	8,610,269
84		S	rs1513181	3	188,857,208	
85	yes	S	rs9291090	4	5,388,910	32,887,223
86	yes	S	rs11725412	4	38,276,133	537,748
87		S	rs4833103	4	38,813,881	2,738,466
88		S	rs10007810	4	41,552,347	4,760,229
89		S	rs279858	4	46,312,576	9,174,142
90		H11	rs11133389	4	55,486,718	15,070

91		H11	rs11133391	4	55,501,788	45,848
92	yes	S	rs4864548	4	55,547,636	28,840,289
93	yes	S	rs385194	4	84,387,925	14,756,550
94		S	rs1800760	4	99,144,475	120,744
95		H12	rs1230025	4	99,265,219	15,363
96		H12	rs975833	4	99,280,582	11,694
97	yes	H12	rs1229966	4	99,292,276	15,033
98	yes	H12	rs1042026	4	99,307,309	10,646
99	yes	H13	rs4147536	4	99,317,955	207
100		H13	rs1229984	4	99,318,162	3,690
101		H14	rs1159918	4	99,321,852	436
102		H14	rs6810842	4	99,322,288	874
103		H14	rs3811801	4	99,323,162	1,170
104		H14	rs1693439	4	99,324,332	15,300
105		H15	rs698	4	99,339,632	3,176
106		H15	rs1693482	4	99,342,808	2,168
107		H15	rs2241894	4	99,344,976	38,033
108	yes	H16	rs2584457	4	99,383,009	82
109	yes	H16	rs12648443	4	99,383,091	17,195
110	yes	H16	rs4699748	4	99,400,286	130
111		H16	rs2584461	4	99,400,416	22
112		H16	rs1442492	4	99,400,438	8,098
113	yes	H16	rs969804	4	99,408,536	6,290
114		H17	rs2851011	4	99,414,826	5,878
115	yes	H17	rs971074	4	99,420,704	7,808
116		H17	rs1573496	4	99,428,512	188,552
117	yes	S	rs1491238	4	99,617,064	4,837,202
118	yes	S	rs7657799	4	104,454,266	44,359,550
119	yes	S	rs1403454	4	148,813,816	29,664,553
120	yes	S	rs2702414	4	178,478,369	8,138,610
121		H18	rs1280100	4	186,616,979	197
122		H18	rs1280099	4	186,617,176	
123	yes	S	rs316598	5	2,364,512	4,480,392
124		H19	rs870348	5	6,844,904	18
125		H19	rs870347	5	6,844,922	2,774,871
126		H20	rs41461	5	9,619,793	9,624
127		H20	rs2234233	5	9,629,417	24,286,291
128		H21	rs11955798	5	33,915,708	14,441
129		H21	rs7737265	5	33,930,149	21,439
130		H21	rs16891982	5	33,951,588	7,266
131		H21	rs28777	5	33,958,854	4,911
132		H22	rs26722	5	33,963,765	1,068
133		H22	rs35408	5	33,964,833	80,993
134		S	rs9292521	5	34,045,826	991,184
135	yes	S	rs37369	5	35,037,010	8,674,266
136		S	rs6451722	5	43,711,276	36,078,627
137	yes	S	rs12657828	5	79,789,903	14,703,039
138		S	rs12652655	5	94,492,942	22,803,569
139		S	rs2220858	5	117,296,511	862,434

140		S	rs10079352	5	118,158,945	48,154,032
141		S	rs1500127	5	166,312,977	4,463,003
142		S	rs7722456	5	170,775,980	
143		S	rs12203592	6	396,321	4,350,604
144		S	rs1040045	6	4,746,925	7,787,954
145	yes	S	rs2504853	6	12,534,879	3,116,022
146		S	rs760761	6	15,650,901	9,739
147		S	rs909706	6	15,660,640	6,250,745
148	yes	S	rs7745461	6	21,911,385	29,835,287
149	yes	S	rs2397060	6	51,746,672	18,782,002
150		S	rs2842063	6	70,528,674	19,279,885
151		S	rs192655	6	89,808,559	31,228,078
152		S	rs1935946	6	121,036,637	15,124,952
153		S	rs3823159	6	136,161,589	8,572,606
154		S	rs4463276	6	144,734,195	18,066,565
155		S	rs4458655	6	162,800,760	5,464,320
156		S	rs1871428	6	168,265,080	
157		S	rs3779009	7	2,164,057	24,971,619
158	yes	S	rs12533947	7	27,135,676	112,746
159		H23	rs2214412	7	27,248,422	15,591
160	yes	H23	rs6943291	7	27,264,013	868,954
161		S	rs917115	7	28,132,967	4,006,545
162		S	rs32314	7	32,139,512	10,200,960
163		S	rs2330442	7	42,340,472	31,699,397
164	yes	S	rs4717865	7	74,039,869	9,863,862
165		S	rs10954737	7	83,903,731	13,200,929
166		S	rs17168174	7	97,104,660	15,040
167		S	rs7794886	7	97,119,700	946,351
168		S	rs705308	7	98,066,051	1,602,644
169		S	rs4646450	7	99,668,695	95,148
170		S	rs2242480	7	99,763,843	17,977,909
171	yes	S	rs10249419	7	117,741,752	13,315,555
172		S	rs7803075	7	131,057,307	8,690,271
173	yes	S	rs10236187	7	139,747,578	2,225,327
174		H24	rs1726866	7	141,972,905	640
175		H24	rs713598	7	141,973,545	
176		S	rs10108270	8	4,333,271	7,405,049
177	yes	H25	rs1390950	8	11,738,320	140
178	yes	H25	rs2898295	8	11,738,460	17,345,328
179		S	rs1471939	8	29,083,788	2,864,744
180		H26	rs383632	8	31,948,532	2,046
181	yes	H26	rs385396	8	31,950,578	88,498
182		S	rs1462906	8	32,039,076	10,152,774
183		S	rs732612	8	42,191,850	147,073
184		S	rs3136717	8	42,338,923	11,453,070
185		S	rs4737761	8	53,791,993	5,221,066
186		S	rs310362	8	59,013,059	4,295,037
187		S	rs7837234	8	63,308,096	22,204,291
188	yes	S	rs12544346	8	85,512,387	24,077,701

189		S	rs6990312	8	109,590,088	11,521,974
190		S	rs2196051	8	121,112,062	784,202
191		S	rs7844723	8	121,896,264	17,332,674
192	yes	S	rs2001907	8	139,228,938	5,185,359
193		S	rs1871534	8	144,414,297	
194		S	rs4741658	9	2,241,994	10,430,103
195	yes	H27	rs1408799	9	12,672,097	223
196		H27	rs1408801	9	12,672,320	36,985
197		H28	rs683	9	12,709,305	730
198		H28	rs910	9	12,710,035	1,344,583
199		S	rs10961356	9	14,054,618	2,740,670
200		S	rs10962599	9	16,795,288	11,833,214
201		S	rs10511828	9	28,628,502	40,415,862
202	yes	S	rs3793451	9	69,044,364	16,350,525
203		S	rs1867958	9	85,394,889	31,973,038
204	yes	S	rs10513300	9	117,367,927	7,137,483
205		S	rs3814134	9	124,505,410	10,019,859
206		H29	rs3118582	9	134,525,269	193
207		H29	rs10776839	9	134,525,462	
208	yes	S	rs3793791	10	49,633,658	23,907,578
209	yes	S	rs4746136	10	73,541,236	19,620,072
210		S	rs4918664	10	93,161,308	1,600,592
211		S	rs12248560	10	94,761,900	175,595
212		H30	rs4918758	10	94,937,495	4,795
213		H30	rs1799853	10	94,942,290	9,974,093
214		S	rs4917432	10	104,916,383	126,088
215	yes	H31	rs1670008	10	105,042,471	15,225
216		H31	rs6584650	10	105,057,696	213,327
217		S	rs2930457	10	105,271,023	27,565,576
218	yes	S	rs4880436	10	132,836,599	
219		S	rs6357	11	2,167,008	2,921,830
220		S	rs2499936	11	5,088,838	167,168
221	yes	S	rs2855122	11	5,256,006	2,572,763
222		S	rs10839880	11	7,828,769	7,987,822
223		S	rs1837606	11	15,816,591	8,172,393
224		S	rs2946788	11	23,988,984	21,893,078
225		S	rs2292910	11	45,882,062	15,947,678
226	yes	H32	rs174570	11	61,829,740	760
227		H32	rs1535	11	61,830,500	11,778
228	yes	H32	rs174583	11	61,842,278	5,288,743
229	yes	S	rs11227699	11	67,131,021	22,146,857
230	yes	S	rs1393350	11	89,277,878	23,962,901
231		H33	rs2303377	11	113,240,779	164
232		H33	rs2303378	11	113,240,943	89,044
233		H34	rs2303380	11	113,329,987	10,661
234	yes	H34	rs10891537	11	113,340,648	59,458
235		H35	rs1800497	11	113,400,106	8,602
236		H35	rs2242592	11	113,408,708	2,845
237	yes	H35	rs1124492	11	113,411,553	1,184

238		H35	rs6277	11	113,412,737	18
239		H35	rs6275	11	113,412,755	5,705
240	yes	H36	rs1079727	11	113,418,460	116
241	yes	H36	rs2002453	11	113,418,576	2,450
242		H36	rs2234690	11	113,421,026	4,538
243		H36	rs1079597	11	113,425,564	7,348,174
244		S	rs948028	11	120,773,738	
245		S	rs2857234	12	6,796,015	60,015
246	yes	S	rs2226955	12	6,856,030	4,692,524
247		S	rs2416791	12	11,548,554	5,706,304
248		S	rs1513056	12	17,254,858	7,951,036
249	yes	H37	rs12587	12	25,205,894	3,724
250	yes	H37	rs712	12	25,209,618	225
251		H37	rs1137282	12	25,209,843	21,036
252	yes	H37	rs17329025	12	25,230,879	1,308,560
253		S	rs708156	12	26,539,439	1,043,614
254	yes	S	rs1975920	12	27,583,053	28,186,897
255		S	rs772262	12	55,769,950	47,082,074
256		H38	rs1722383	12	102,852,024	14,832
257	yes	H38	rs2133298	12	102,866,856	185
258		H38	rs3817446	12	102,867,041	6,016,903
259		S	rs2070586	12	108,883,944	2,671,964
260		H39	rs7969300	12	111,555,908	174
261		H39	rs593226	12	111,556,082	9,497
262		H39	rs616513	12	111,565,579	40,950
263		S	rs3809274	12	111,606,529	150,643
264		H40	rs737280	12	111,757,172	16,857
265		H40	rs2238151	12	111,774,029	29,933
266		H40	rs671	12	111,803,962	6,647,721
267	yes	H41	rs1503767	12	118,451,683	71
268		H41	rs11068953	12	118,451,754	
269		S	rs7997709	13	34,273,600	6,867,546
270	yes	S	rs1572018	13	41,141,146	864,703
271		S	rs2166624	13	42,005,849	6,490,527
272		S	rs7326934	13	48,496,376	1,817,048
273		H42	rs806301	13	50,313,424	165
274		H42	rs2066700	13	50,313,589	25,106,162
275	yes	S	rs9530435	13	75,419,751	35,755,069
276		S	rs9522149	13	111,174,820	
277		S	rs1760921	14	20,349,972	31,791,277
278	yes	S	rs2357442	14	52,141,249	5,630,720
279	yes	S	rs1950993	14	57,771,969	9,648,095
280	yes	S	rs8021730	14	67,420,064	6,363,790
281		H43	rs12717560	14	73,783,854	158
282		H43	rs12878166	14	73,784,012	9,222,512
283	yes	S	rs946918	14	83,006,524	9,300,795
284		S	rs12896399	14	92,307,319	6,601,665
285	yes	S	rs200354	14	98,908,984	6,303,734
286		S	rs3784230	14	105,212,718	

287		S	rs2703969	15	27,886,487	65,404
288		H44	rs1800414	15	27,951,891	10,980
289		H44	rs11074314	15	27,962,871	18,524
290		H44	rs12914687	15	27,981,395	2,012
291		H44	rs74653330	15	27,983,407	1,765
292		H44	rs1800407	15	27,985,172	5,455
293		H44	rs1800404	15	27,990,627	3,528
294		H45	rs2015343	15	27,994,155	26,934
295		H45	rs746861	15	28,021,089	22,513
296	yes	H45	rs7170869	15	28,043,602	9,285
297	yes	H45	rs895828	15	28,052,887	37,787
298		H46	rs4778138	15	28,090,674	2,893
299		H46	rs4778241	15	28,093,567	5,525
300		H46	rs7495174	15	28,099,092	12,621
301		H46	rs1129038	15	28,111,713	7,200
302		H47	rs7494942	15	28,118,913	1,559
303	yes	H47	rs12913832	15	28,120,472	8,394
304		H47	rs3935591	15	28,128,866	53,974
305		S	rs7170852	15	28,182,840	85,378
306	yes	H48	rs916977	15	28,268,218	16,818
307		H48	rs1667394	15	28,285,036	7,642,798
308		S	rs12439433	15	35,927,834	8,932,339
309		S	rs735480	15	44,860,173	543,324
310		S	rs1153849	15	45,403,497	2,615,716
311		S	rs9788730	15	48,019,213	80,755
312		S	rs1834640	15	48,099,968	34,319
313		H49	rs1426654	15	48,134,287	11,785
314		H49	rs2469597	15	48,146,072	3,064,575
315	yes	S	rs4646	15	51,210,647	12,595,767
316		S	rs10152453	15	63,806,414	10,539,112
317	yes	S	rs11632698	15	74,345,526	96,633
318		S	rs2899826	15	74,442,159	309,738
319		S	rs2472304	15	74,751,897	16,810,581
320		S	rs8035124	15	91,562,478	
321		S	rs4984913	16	690,466	10,190,988
322		S	rs4781011	16	10,881,454	8,380,132
323		S	rs2269793	16	19,261,586	6,502,195
324	yes	S	rs17625895	16	25,763,781	3,886,115
325	yes	S	rs11859842	16	29,649,896	18,574,391
326		S	rs17822931	16	48,224,287	17,148,518
327		S	rs818386	16	65,372,805	19,777,271
328	yes	S	rs2966849	16	85,150,076	34,108
329		S	rs3924280	16	85,184,184	4,480,235
330		S	rs459920	16	89,664,419	254,614
331		H50	rs3212363	16	89,919,033	713
332		H50	rs885479	16	89,919,746	
333		S	rs1879488	17	1,498,319	2,998,742
334		S	rs333113	17	4,497,061	25,728,694
335		S	rs8071667	17	30,225,755	9,497,580

336		S	rs1136201	17	39,723,335	755,342
337		S	rs749906	17	40,478,677	1,862,192
338		H51	rs3869550	17	42,340,869	27,038
339		H51	rs957971	17	42,367,907	14,150
340	yes	H51	rs7211777	17	42,382,057	124,458
341		S	rs4411548	17	42,506,515	397,713
342	yes	S	rs2593595	17	42,904,228	41,863
343		S	rs528854	17	42,946,091	52,841
344	yes	S	rs2271539	17	42,998,932	378,635
345		S	rs2061873	17	43,377,567	505,408
346		S	rs124719	17	43,882,975	187,862
347		S	rs228758	17	44,070,837	1,324,118
348		H52	rs1059504	17	45,394,955	186
349	yes	H52	rs8327	17	45,395,141	289,349
350		S	rs2049515	17	45,684,490	45,069
351	yes	S	rs2902662	17	45,729,559	73,122
352		S	rs110402	17	45,802,681	22,950
353		H53	rs1396862	17	45,825,631	4,748
354		H53	rs1876831	17	45,830,379	160,594
355		H54	rs2435214	17	45,990,973	23,385
356		H54	rs75534191	17	46,014,358	1,604,108
357		S	rs3760370	17	47,618,466	1,429,074
358		S	rs11871214	17	49,047,540	162,165
359		H55	rs2233362	17	49,209,705	42
360		H55	rs634370	17	49,209,747	1,439,024
361		S	rs17642714	17	50,648,771	1,384,034
362		S	rs4538057	17	52,032,805	3,458,718
363		S	rs4471745	17	55,491,523	219,396
364		S	rs2033111	17	55,710,919	9,280,114
365		S	rs11652805	17	64,991,033	6,524,925
366		S	rs10512572	17	71,515,958	4,270,152
367		S	rs2125345	17	75,786,110	2,446,380
368		H56	rs11868709	17	78,232,490	59
369		H56	rs9907137	17	78,232,549	
370	yes	S	rs4798812	18	9,420,506	12,651,515
371		S	rs4800105	18	22,072,021	15,625,638
372		S	rs2042762	18	37,697,659	5,210,655
373	yes	S	rs7226659	18	42,908,314	18,757,561
374	yes	S	rs881728	18	61,665,875	8,245,820
375		S	rs3916235	18	69,911,695	288,732
376		S	rs4891825	18	70,200,427	7,143,901
377		S	rs874299	18	77,344,328	
378		S	rs7251928	19	4,077,098	775,027
379	yes	H57	rs1055919	19	4,852,125	200
380		H57	rs2271057	19	4,852,325	28,309,016
381	yes	S	rs8113143	19	33,161,341	19,237,311
382		S	rs3745099	19	52,398,652	2,704,903
383		S	rs2532060	19	55,103,555	
384		S	rs6104567	20	10,214,785	23,936,021

385	yes	S	rs1015362	20	34,150,806	118,386
386		S	rs6058017	20	34,269,192	26,859,670
387		H58	rs10854214	20	61,128,862	96
388		H58	rs10854215	20	61,128,958	1,694,449
389		S	rs1985835	20	62,823,407	
390		S	rs1041740	21	31,667,849	4,845,478
391		S	rs2835370	21	36,513,327	
392		S	rs1296819	22	17,593,780	2,347,581
393		H59	rs2020917	22	19,941,361	22,323
394		H59	rs4818	22	19,963,684	64
395	yes	H59	rs4680	22	19,963,748	4,421
396	yes	H59	rs9332377	22	19,968,169	12,002,203
397		S	rs4821004	22	31,970,372	9,330,962
398		S	rs2024566	22	41,301,334	6,510,789
399		S	rs5768007	22	47,812,123	

Figure S1. Allele frequency barplot on 65 populations for example haplotype

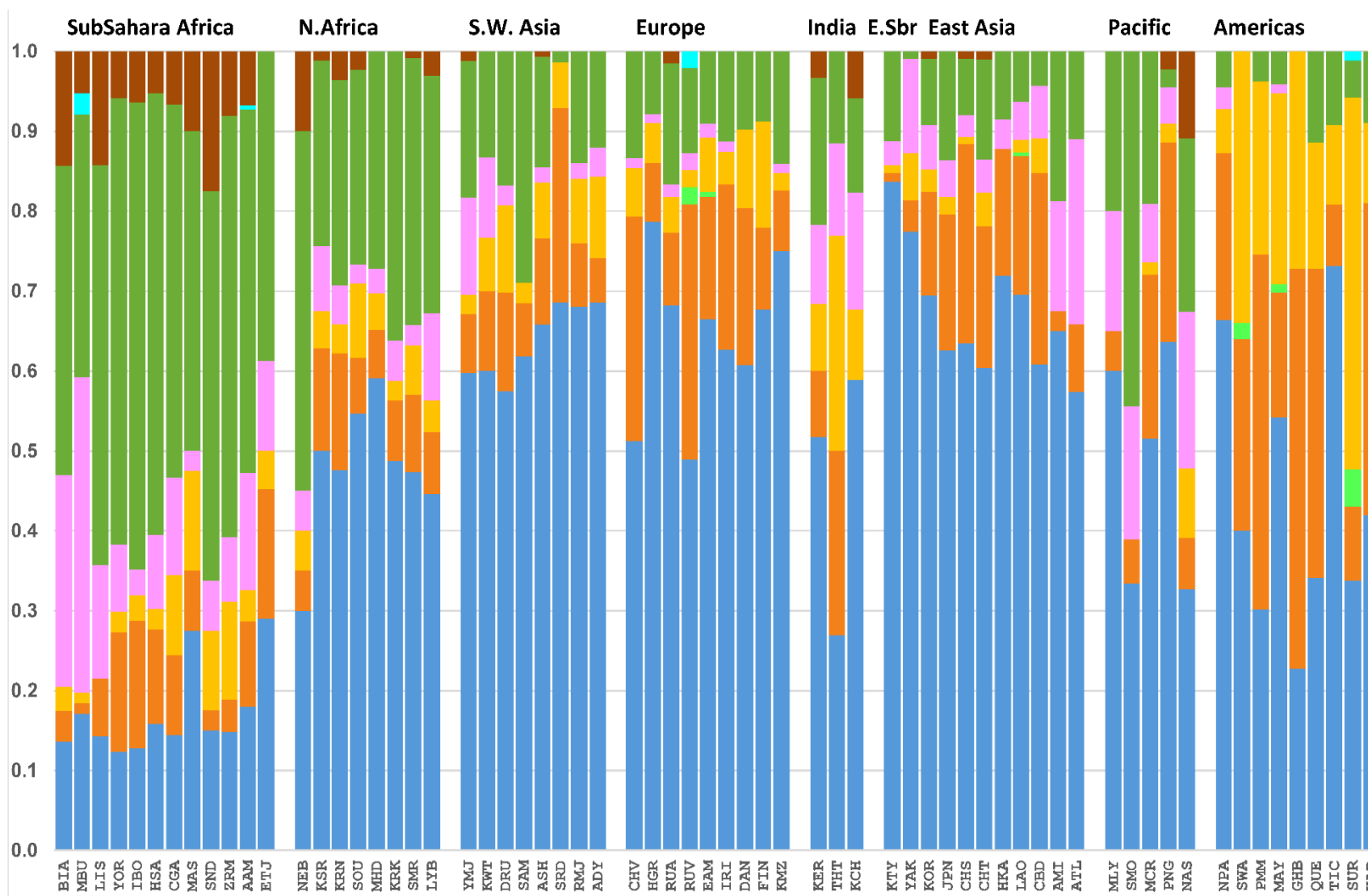


Figure S2. Comparing average heterozygosities of populations in 299 marker and 90 marker datasets

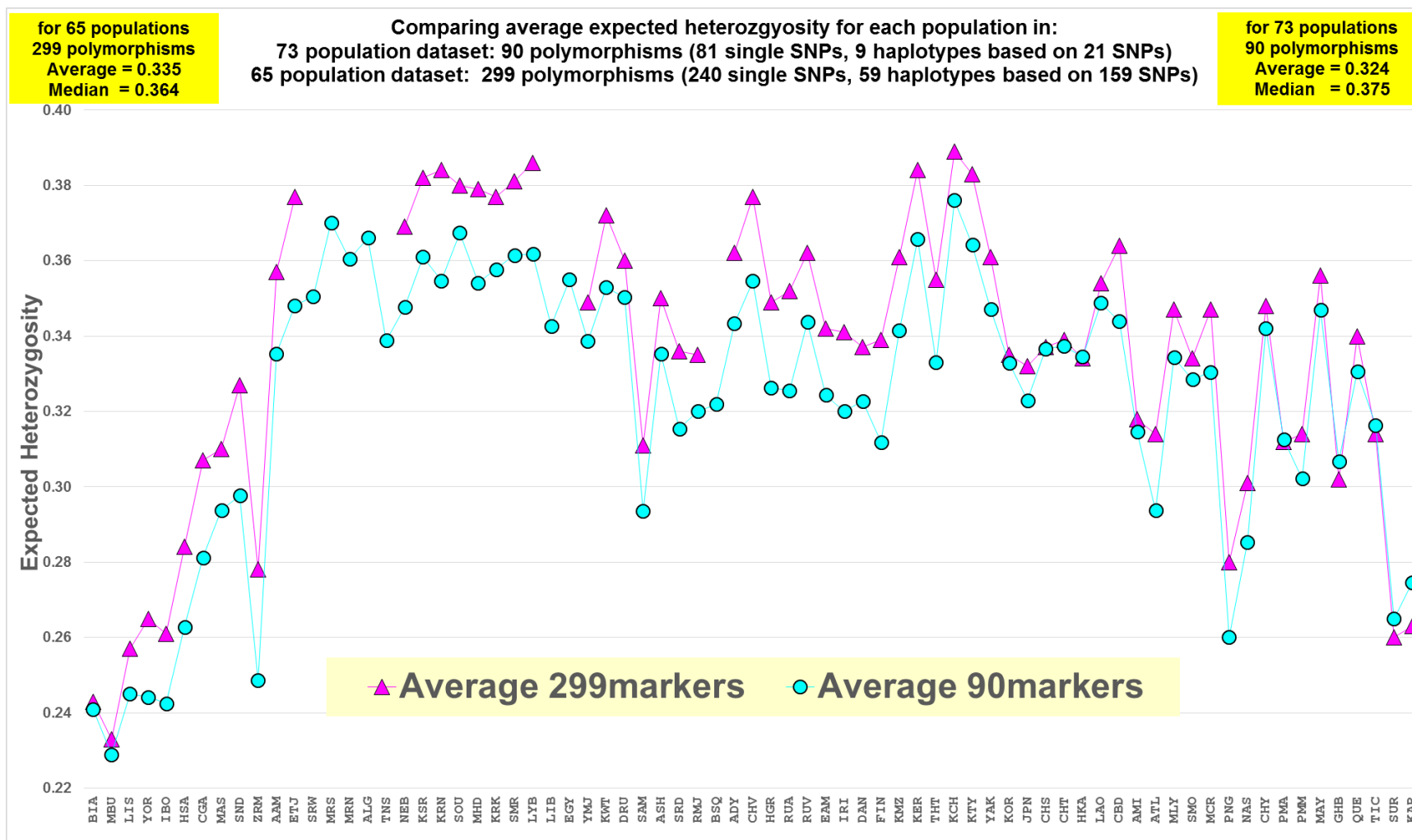


Figure S3. Principal Components Analysis plot for pc #1 and pc #3. PCA is based on pairwise Tau genetic distances for 65 populations.

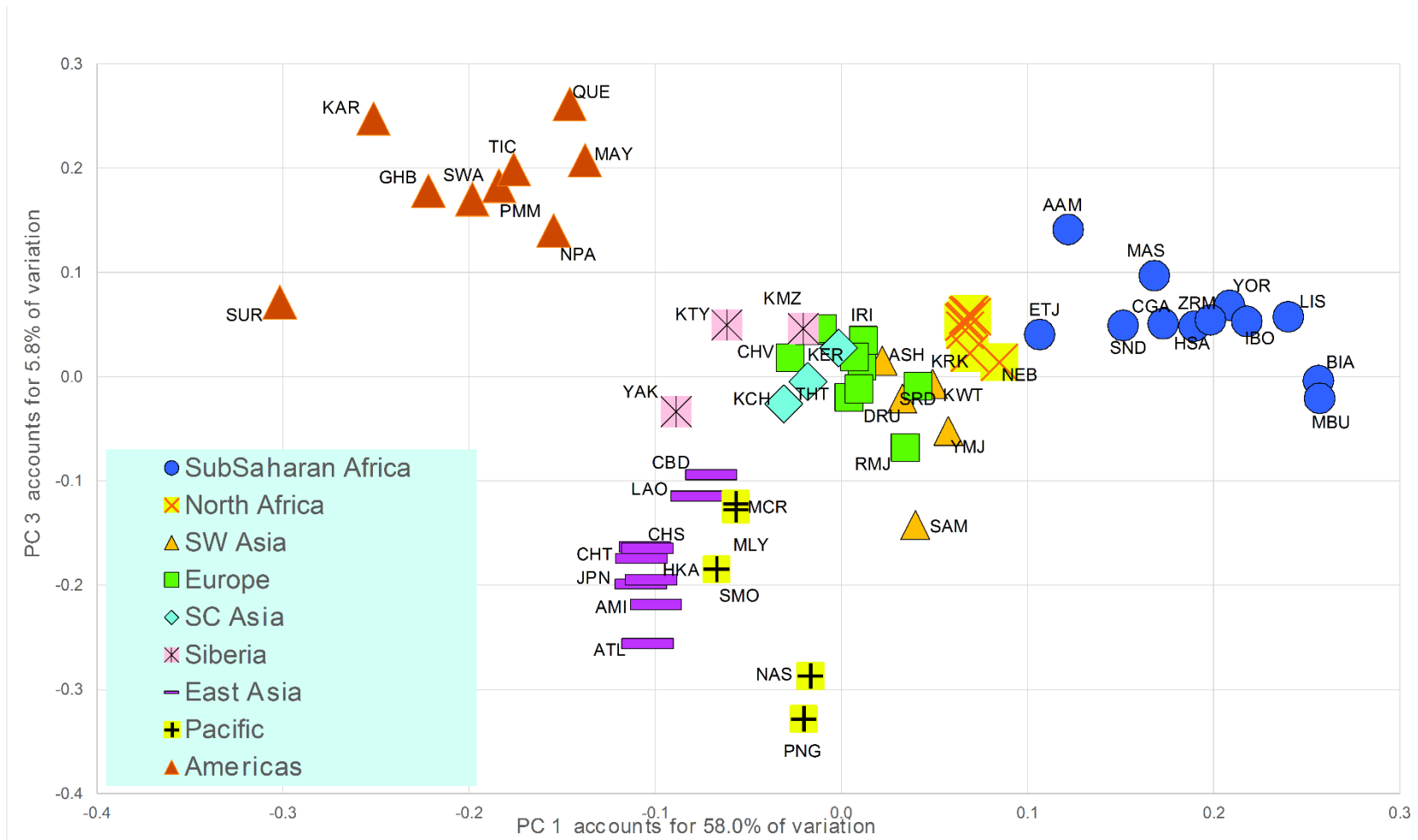


Figure S4. Principal Components Analysis plot—73 populations, 90 markers study—pc #1 x pc #2

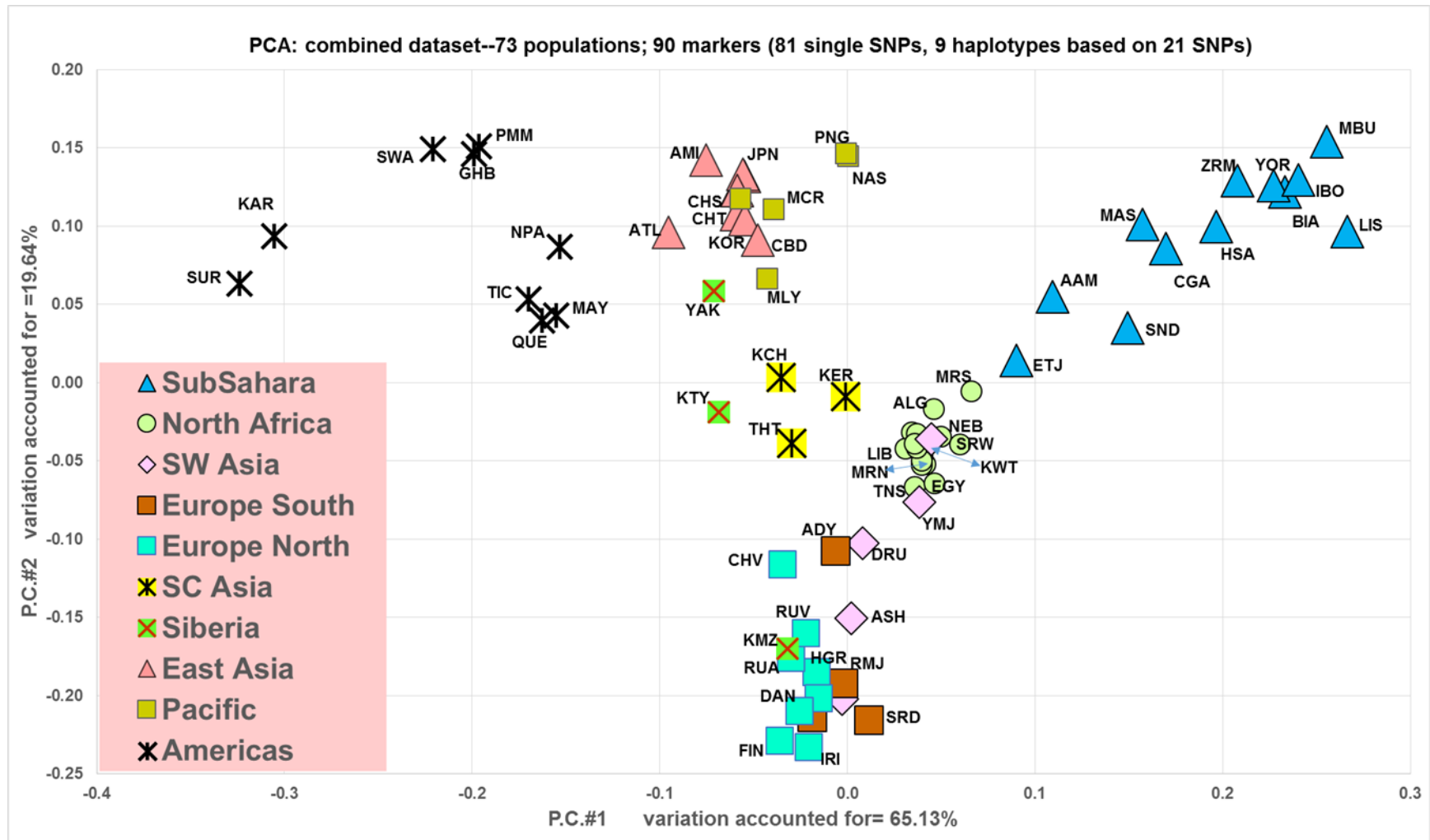


Figure S5. Principal Components Analysis plot—73 populations, 90 markers study —pc #1 x pc #3

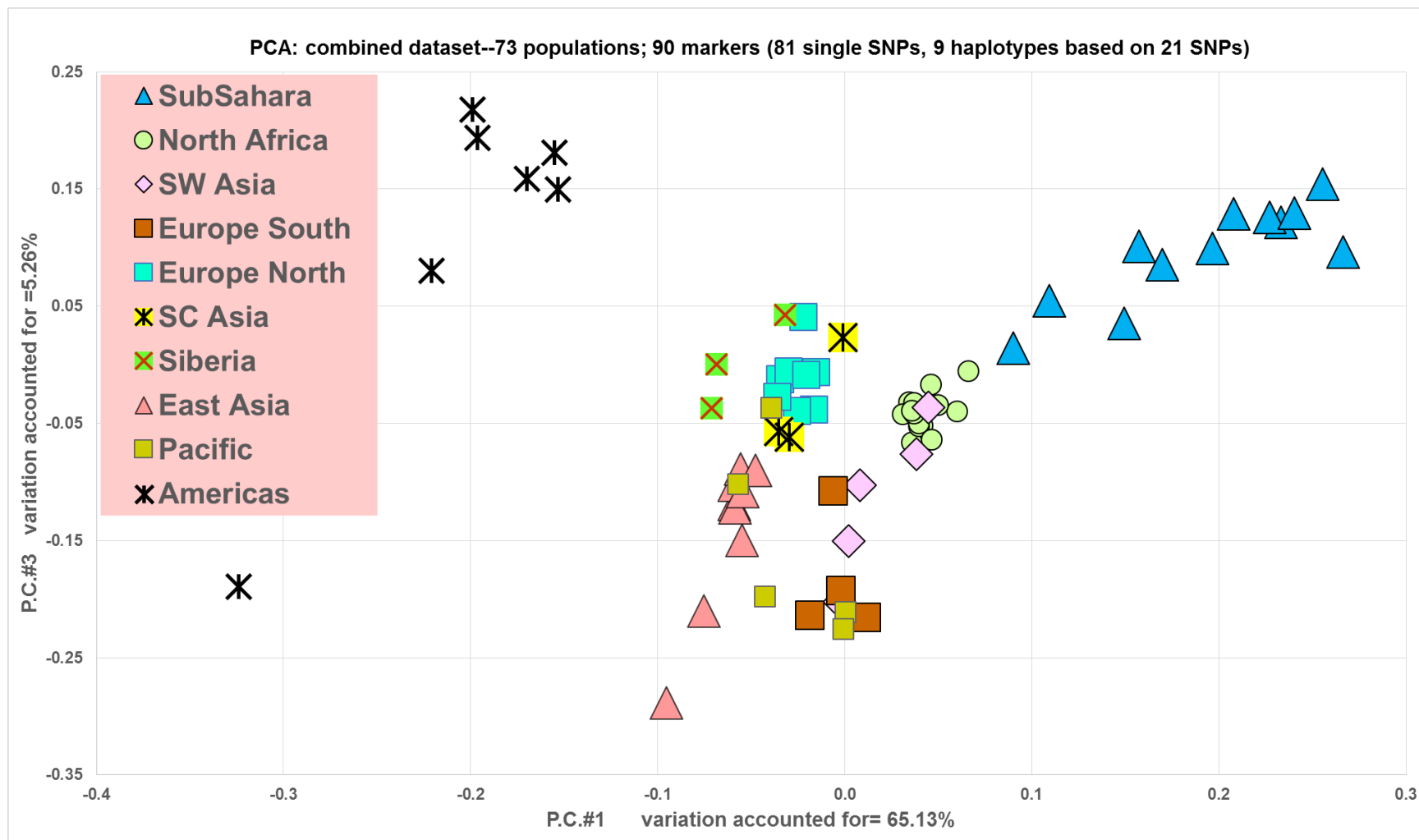
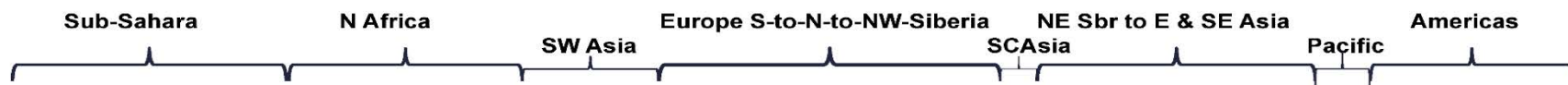
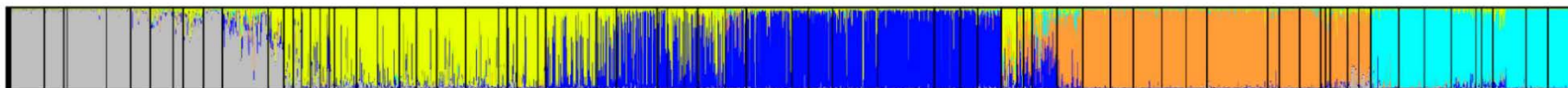


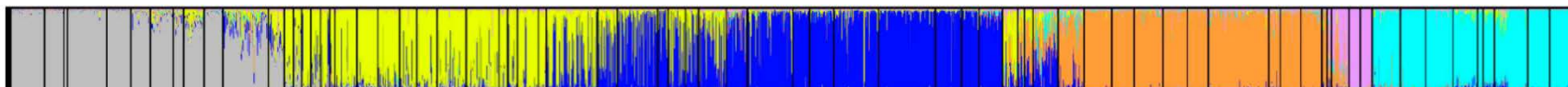
Figure S6. STRUCTURE individual bar plots—73 population study, 90 markers study—displaying results for runs with highest likelihood out of 20 runs in each cluster $K=5$ to 11. Each individual has a separate column in the bar plot and the individuals in a population are clustered together in the display but the STRUCTURE analysis was not informed about an individual's population membership. Black vertical lines identify the population boundaries. The height extent of each color for an individual corresponds to the estimated membership of the individual in one of the clusters; each cluster is assigned a separate color. The bars with multiple colors can be interpreted as genetic admixture or as relative probabilities of belonging to the different clusters. Since there is a separate color bar for each of the more than 3000 individuals, the interval width covered by a population varies and corresponds to the number of individuals in the population.



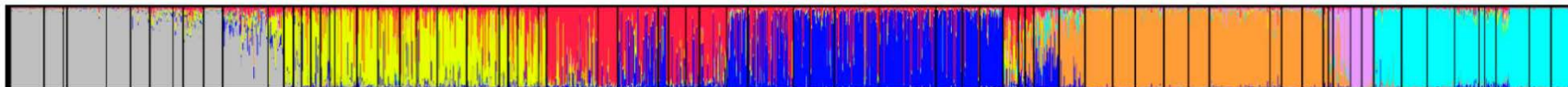
K=5



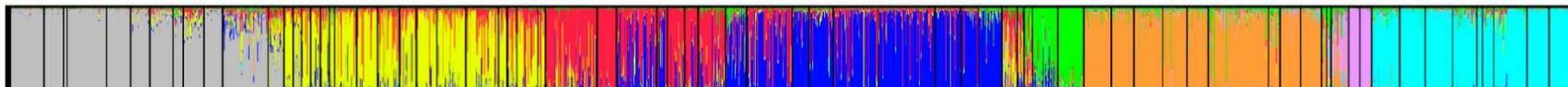
K=6



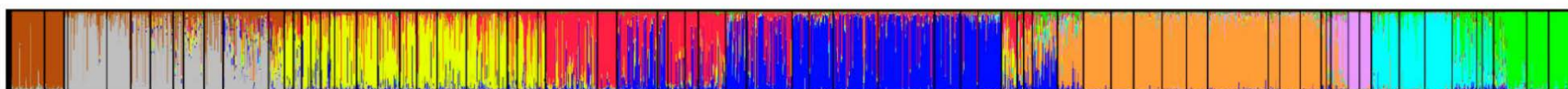
K=7



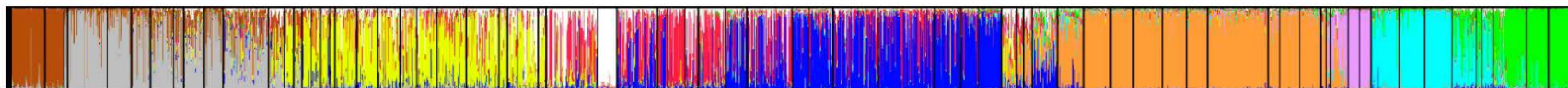
K=8



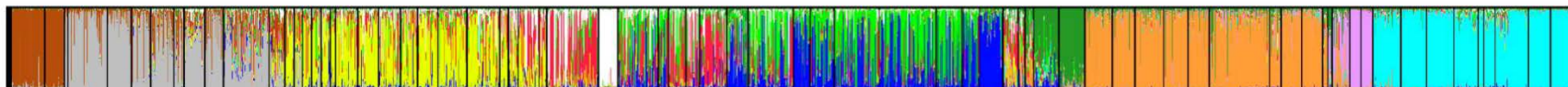
K=9



K=10



K=11



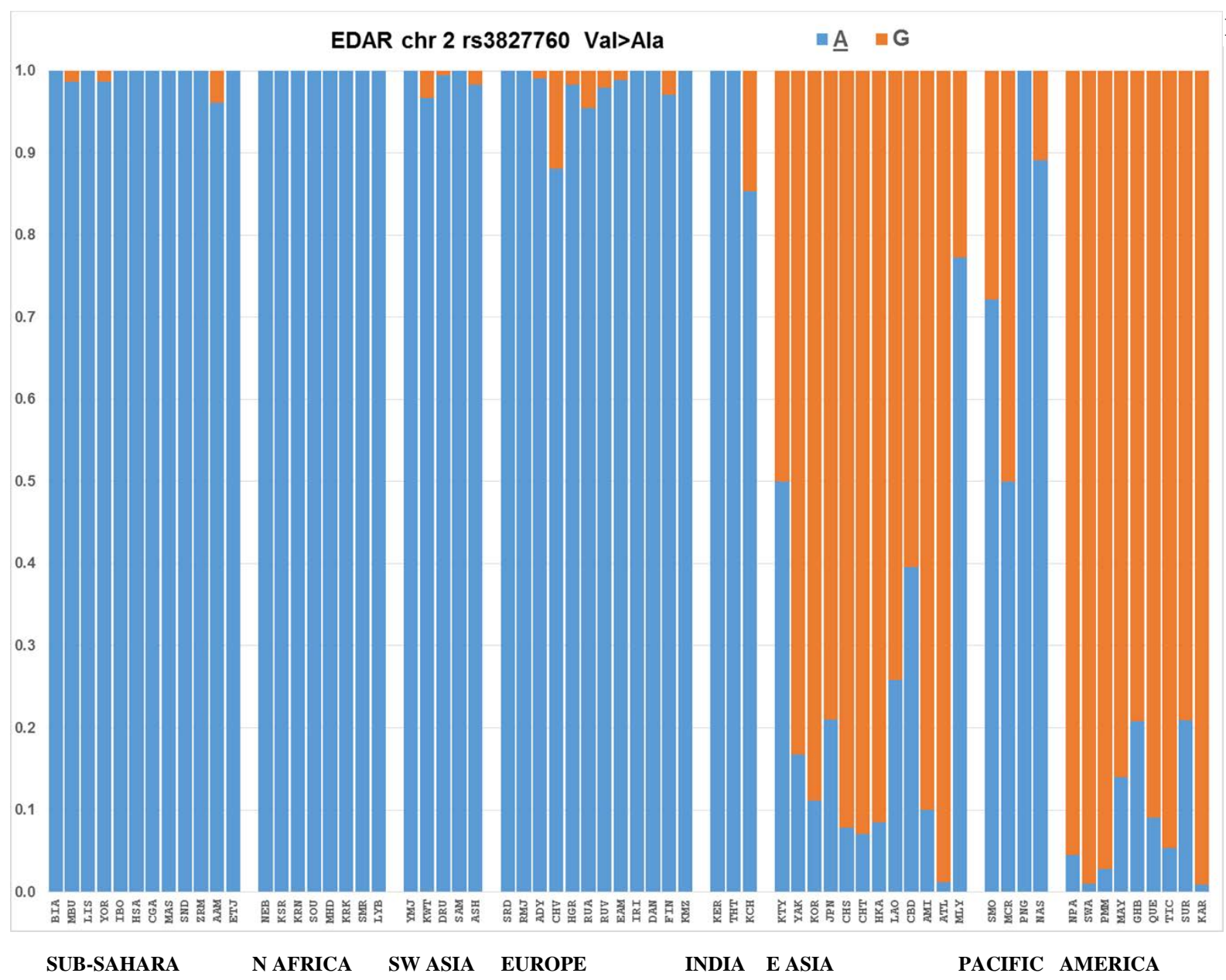
Supplementary material illustrating variation at functional genes

Table S3. List of 9 functional SNPs plus 9 nearby non-functional SNPs included in multi-SNP haplotypes

		GRCh38							
Gene	Chr	nt position	dbSNP #	Anc,Drv	Fn?	Protein change	strand	human phenotype	
EDAR	2	108,897,145	rs3827760	A,G	yes	Val > Ala	forward	hair and dental traits	
EDAR	2	108,897,600	rs2169812	G,T	no		forward		
EDAR	2	108,897,736	rs2378217	C,T	no		forward		
EDAR	2	108,898,008	rs940928	C,T	no		forward		
ADH1B	4	99,318,162	rs1229984	C,T	yes	Arg > His	forward	alcohol metabolism; flushing reaction	
TAS2R1	5	9,619,793	rs41461	C,T	no		reverse		
TAS2R1	5	9,629,417	rs2234233	C,T	yes		reverse	protein coding taste receptor	
SLC45A2	5	33,915,708	rs11955798	A,G	no		forward		
SLC45A2	5	33,930,149	rs7737265	A,G	no		forward		
SLC45A2	5	33,951,588	rs16891982	C,G	yes		forward	pigmentation	
SLC45A2	5	33,958,854	rs28777	A,C	no		forward		
SLC45A2	5	33,963,765	rs26722	C,T	yes		forward		
SLC45A2	5	33,964,833	rs35408	C,T	no		forward		
TAS2R38	7	141,972,905	rs1726866	C,T	yes	Ala > Val	reverse	"PTC"; bitter taste receptor, protein coding	
TAS2R38	7	141,973,545	rs713598	C,G	yes	Ala > Pro	reverse	"PTC"; bitter taste receptor, protein coding	
HERC2	15	28,120,472	rs12913832	A,G	yes			pigmentation	
MC1R	16	89,919,033	rs3212363	T,A	no		forward		
MC1R	16	89,919,746	rs885479	G,A	yes		forward	pigmentation	

On the next 11 pages allele frequency bar plots are presented for 9 functional SNPs that are part of the 299 marker polymorphism dataset. In each bar plot the 65 populations are in the same geographical ordering left to right. Blank bars separate some of the major geographical regions to make inspection of the plots easier. Left to right the eight regions include Sub-Saharan Africa, North Africa, SW Asia, Europe, South Central Asia (India), East Asia, Pacific Islands, Americas. See Table S1A for the detailed population list including the 3-character abbreviations. Some of the bar plots display the functional SNP frequencies in the context of multi-SNP haplotypes that include other non-functional SNPs. The tabular list at the top of this page provides details about all eighteen SNPs included in the following eleven bar plots.

Figure S7



SUB-SAHARA N AFRICA SW ASIA EUROPE INDIA E ASIA PACIFIC AMERICA

Figure S8

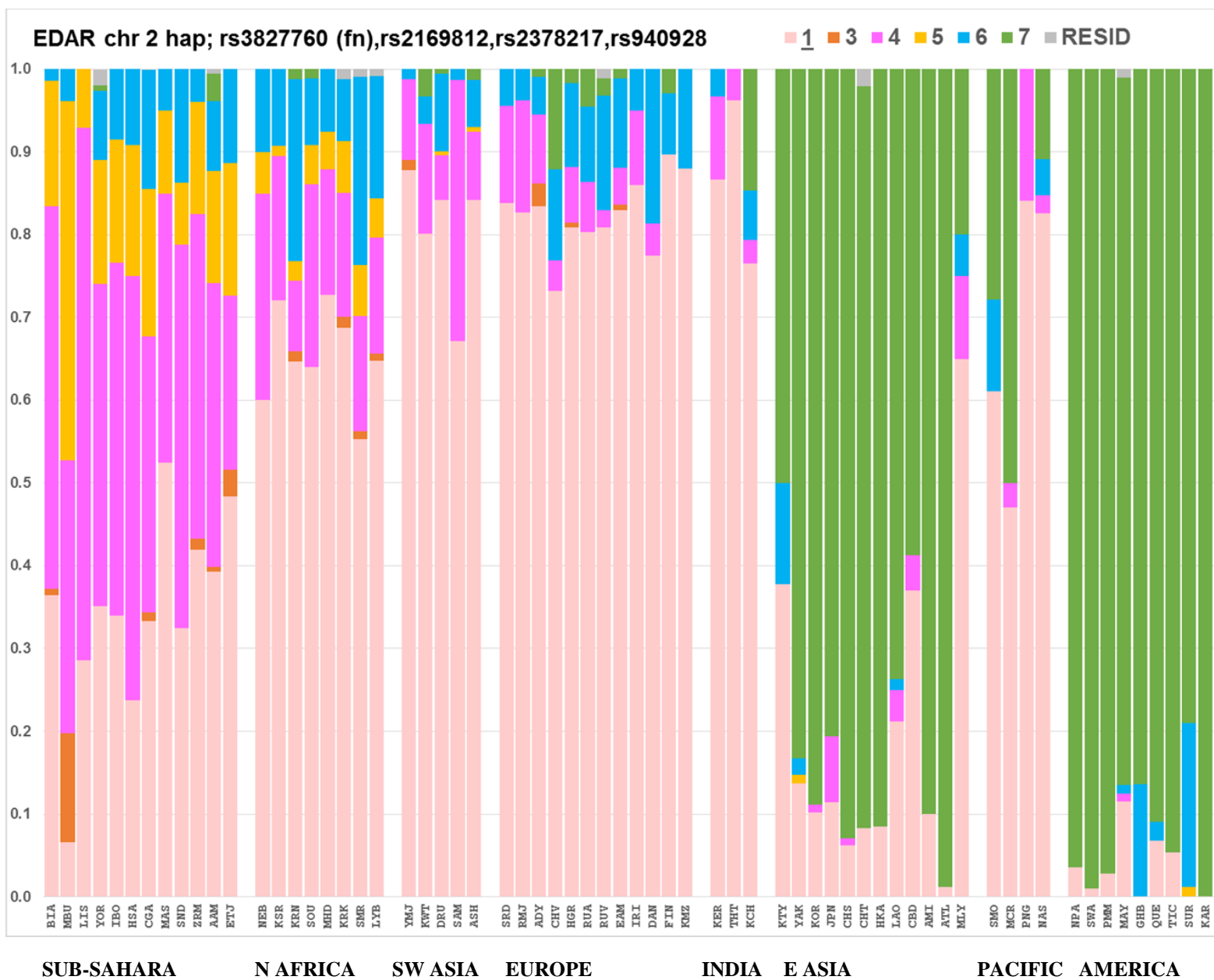


Figure S9

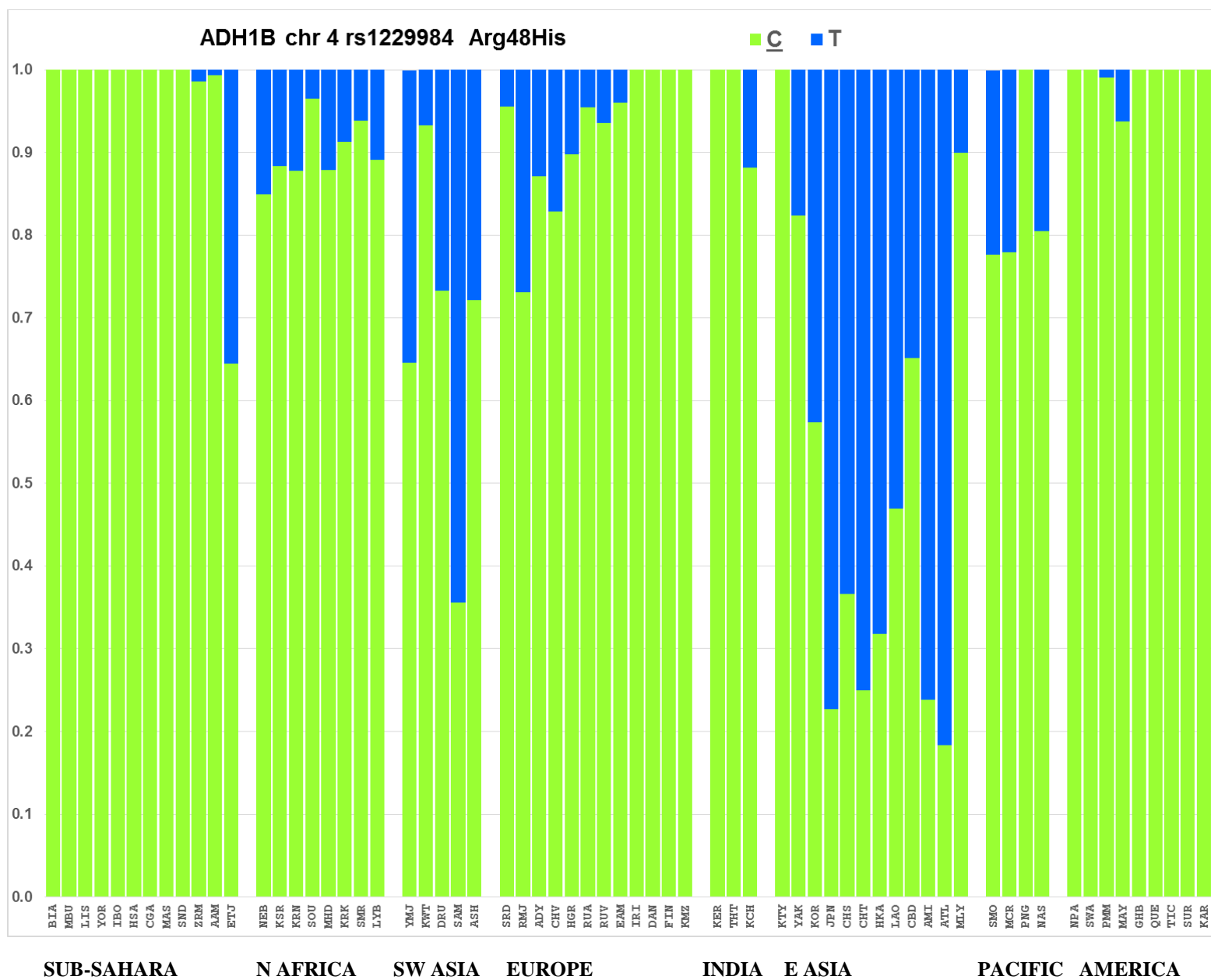


Figure S10

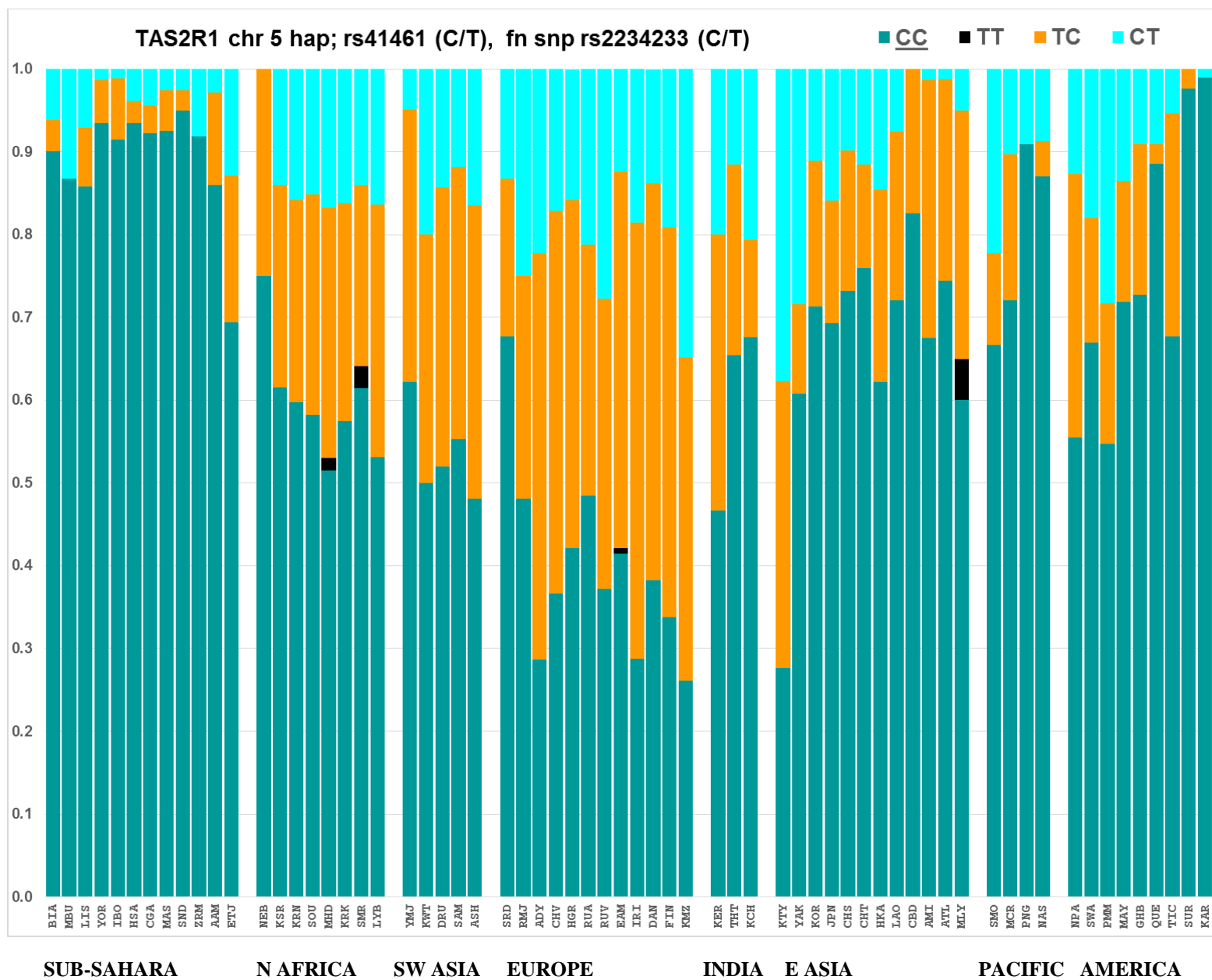


Figure S11

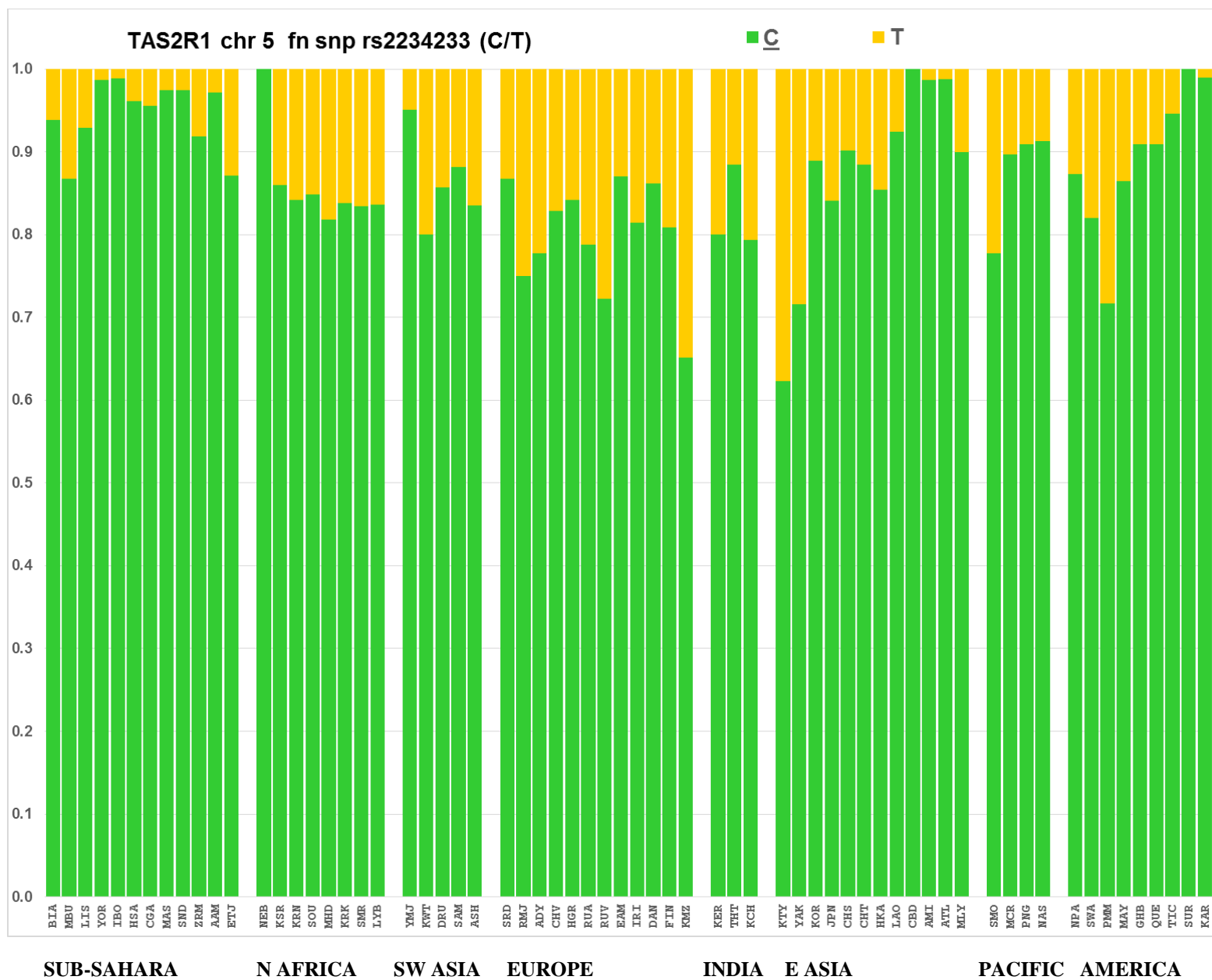


Figure S12

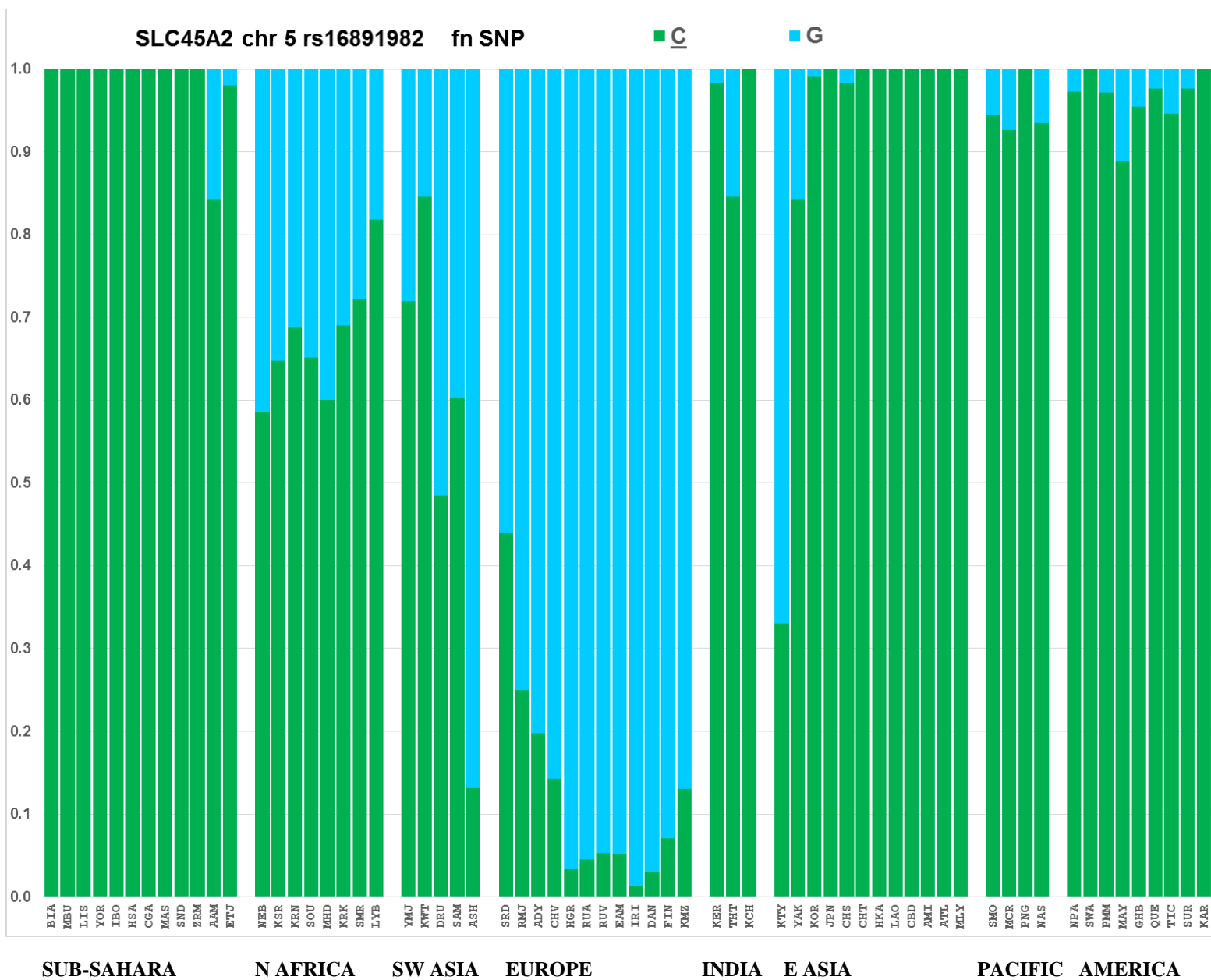


Figure S13

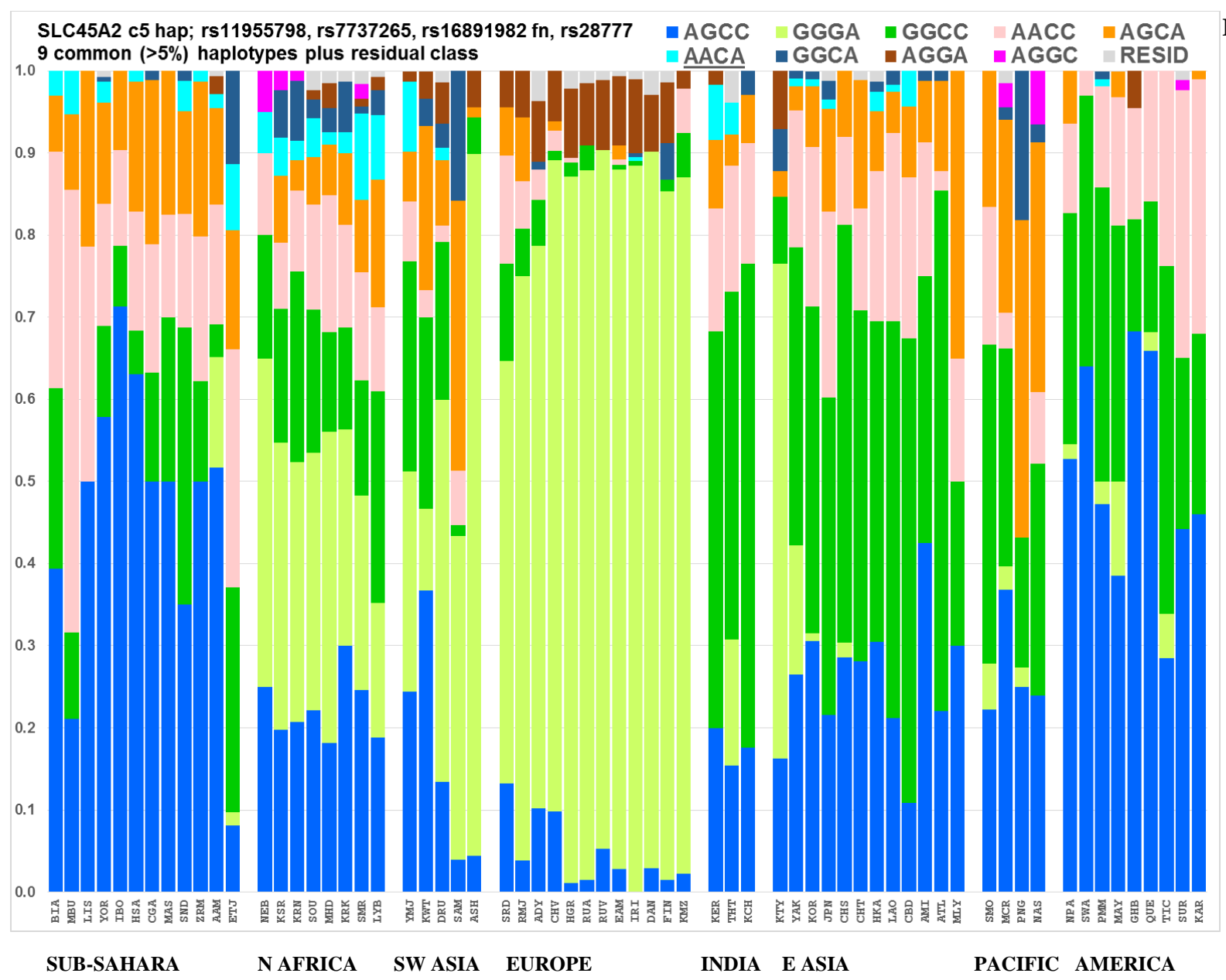


Figure S14

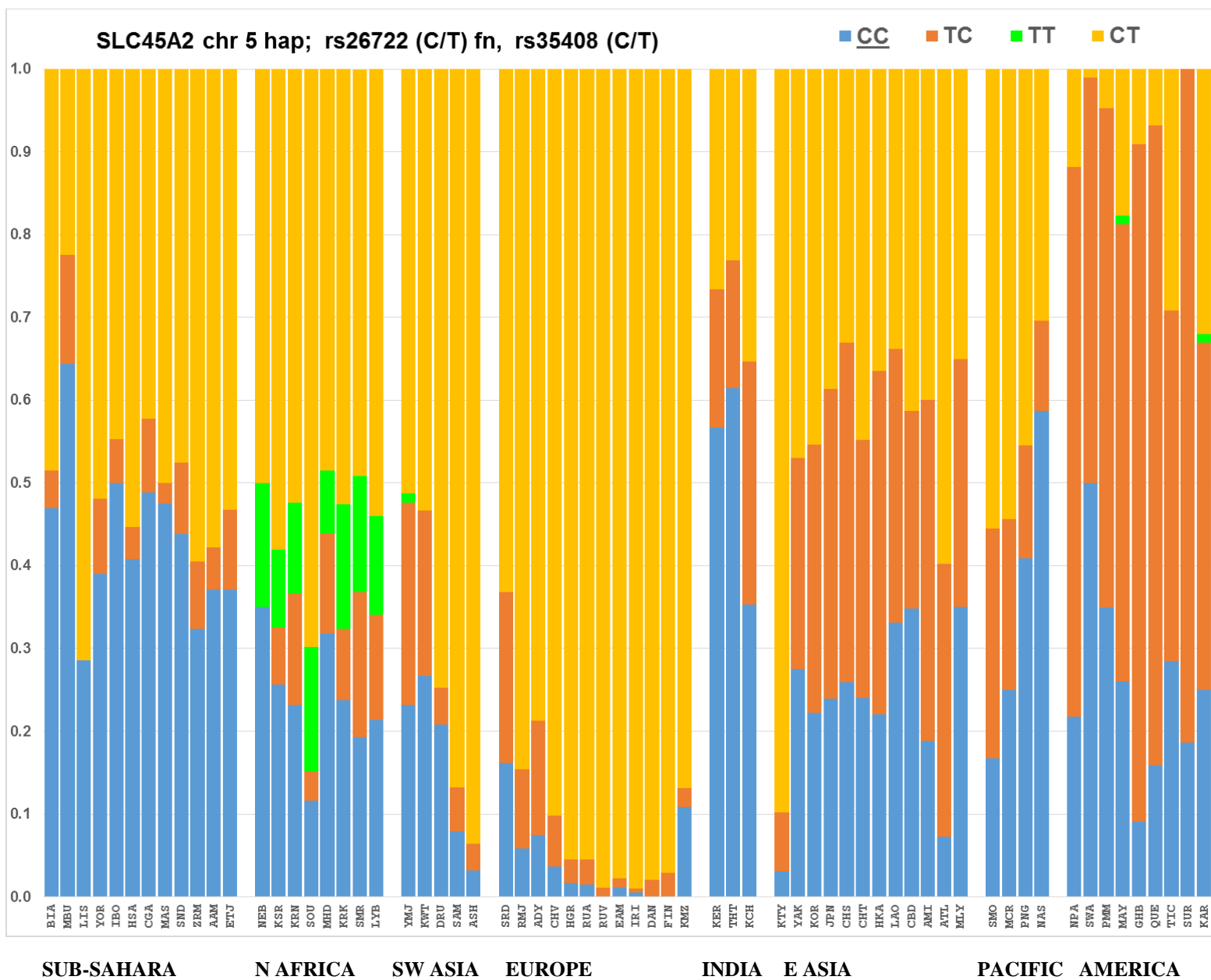
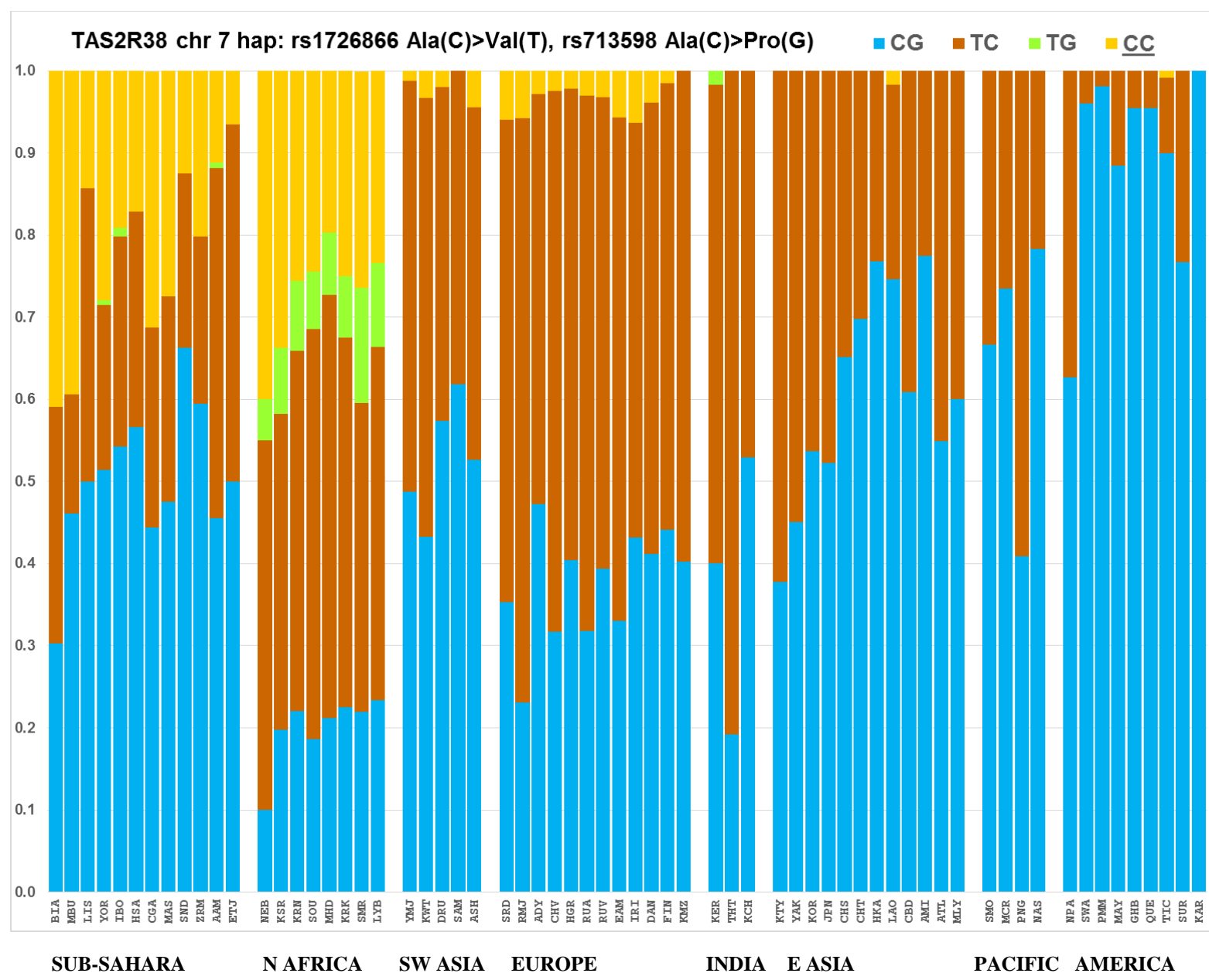
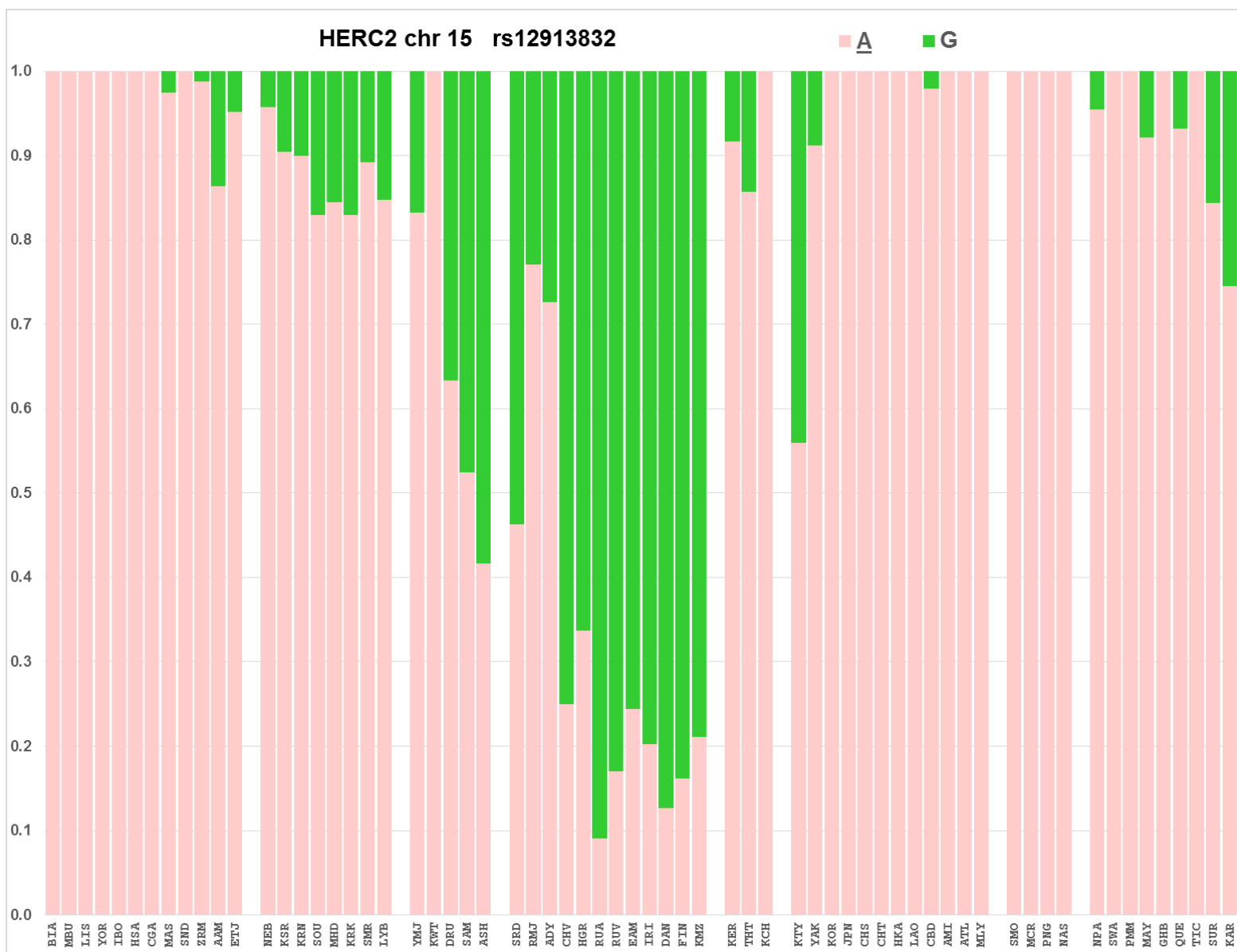


Figure S15



HERC2 chr 15 rs12913832

Figure S16



SUB-SAHARA

N AFRICA

SW ASIA

EUROPE

INDIA

E ASIA

PACIFIC

AMERICA

Figure S17

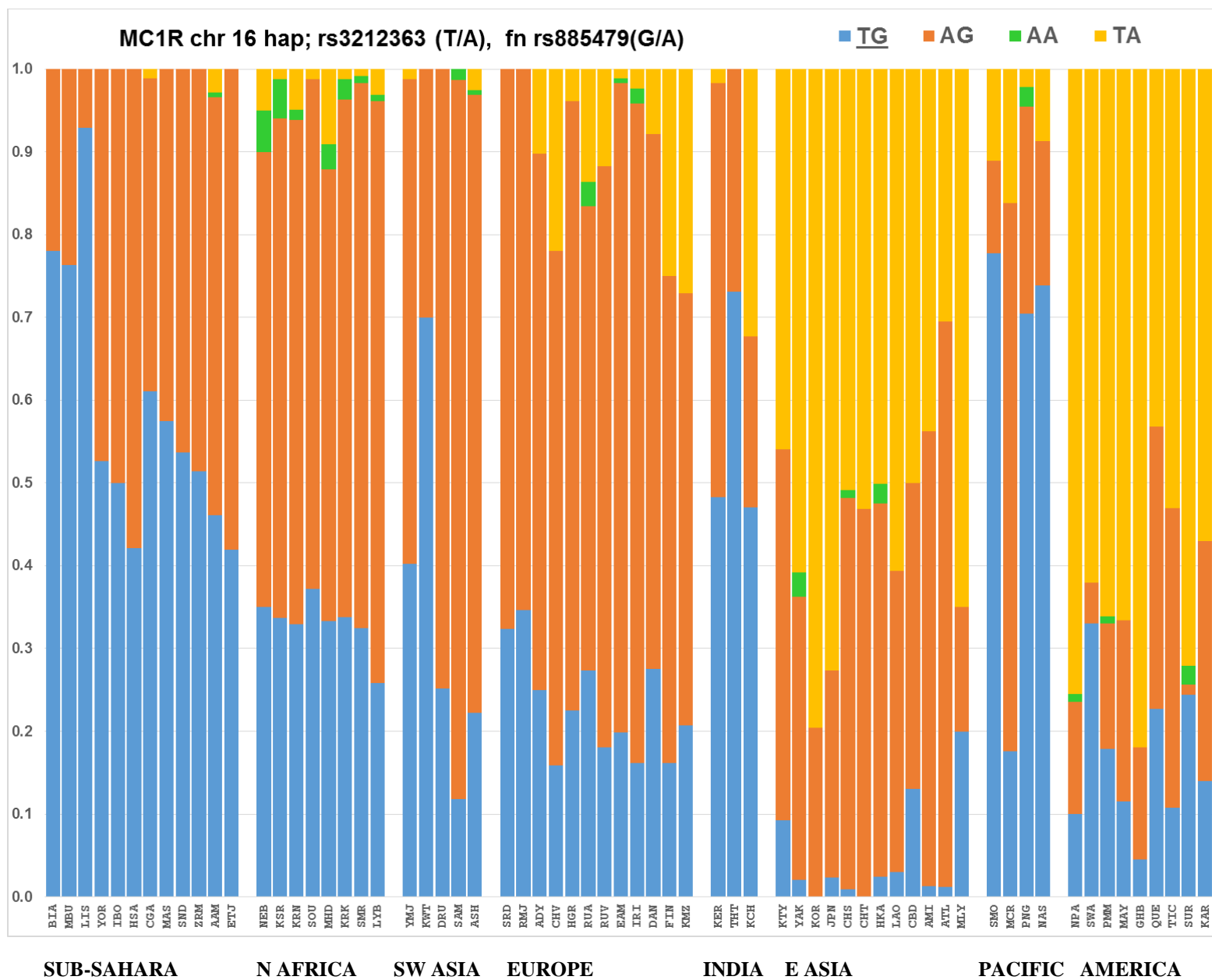


Fig. S18

PCA: 8 N. African populations; 299 genetic systems (240 single SNPs, 59 haplotypes)

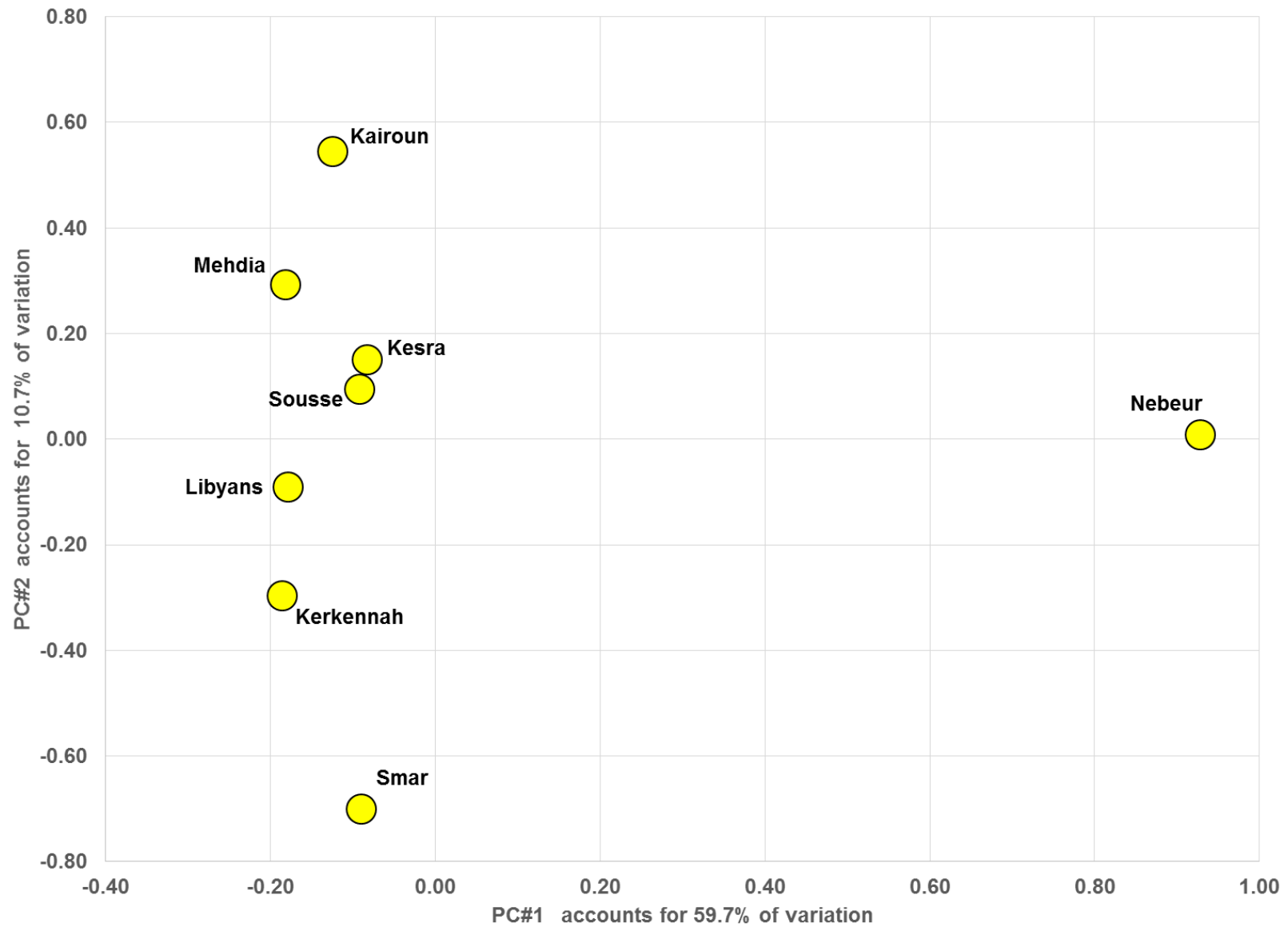


Fig. S19

PCA: 8 N. African populations; 299 genetic systems (240 single SNPs, 59 haplotypes)

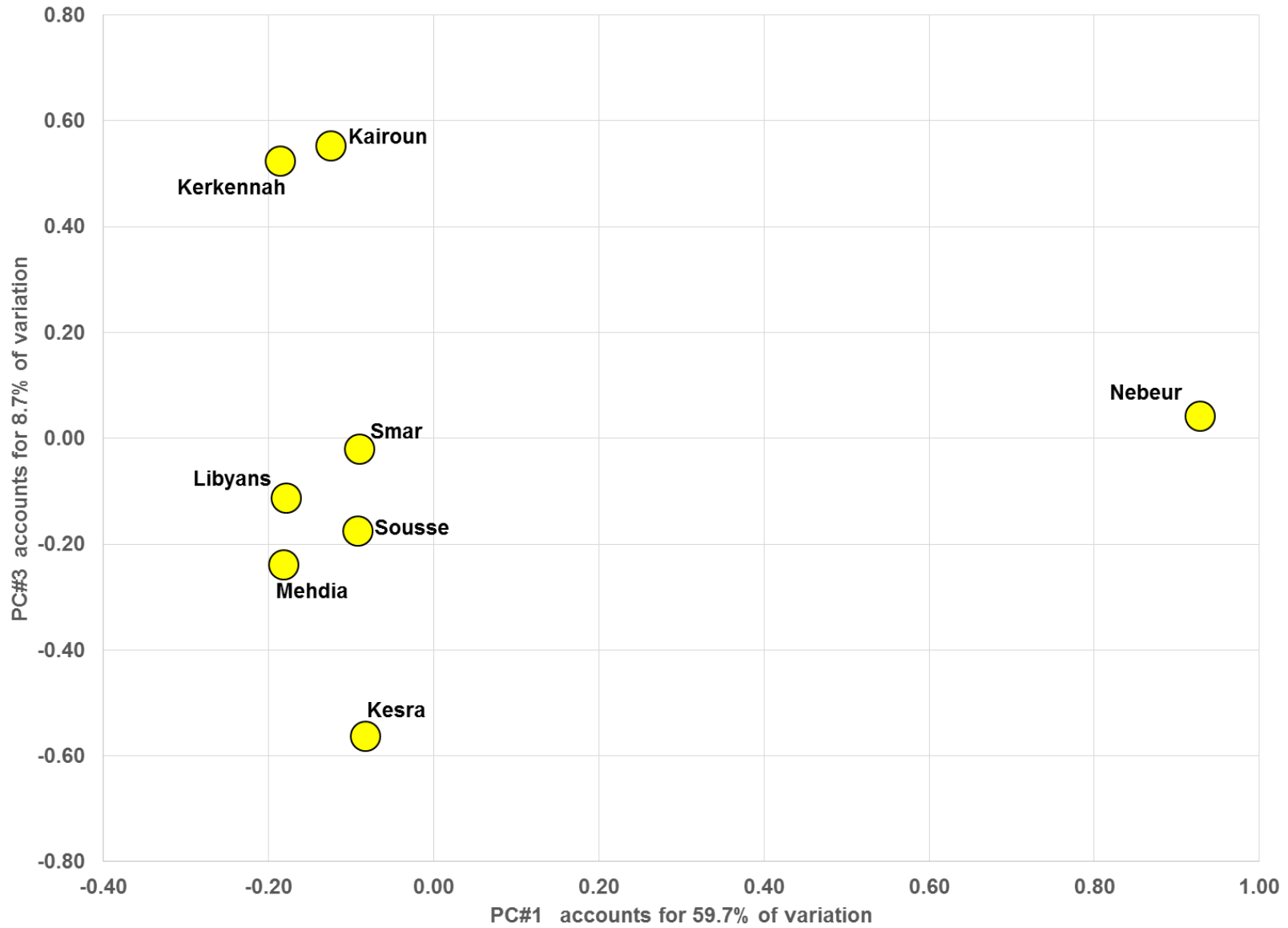


Fig. S20

PCA: 8 N. African populations; 299 genetic systems (240 single SNPs, 59 haplotypes)

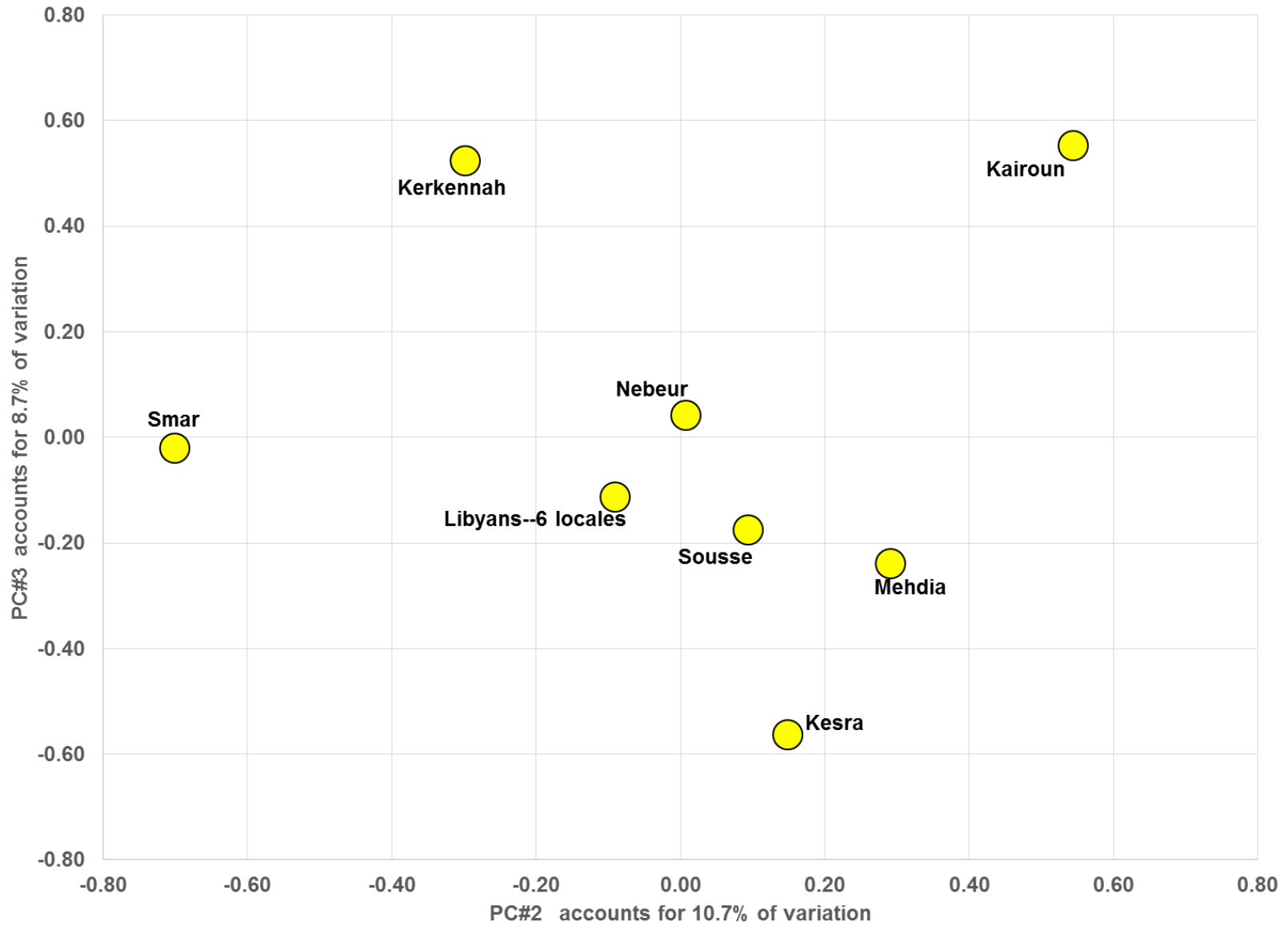


Fig. S21 PCA: N Afr & SW Asia--12 populations; 299 genetic systems (240 SNPs, 59 haplotypes)

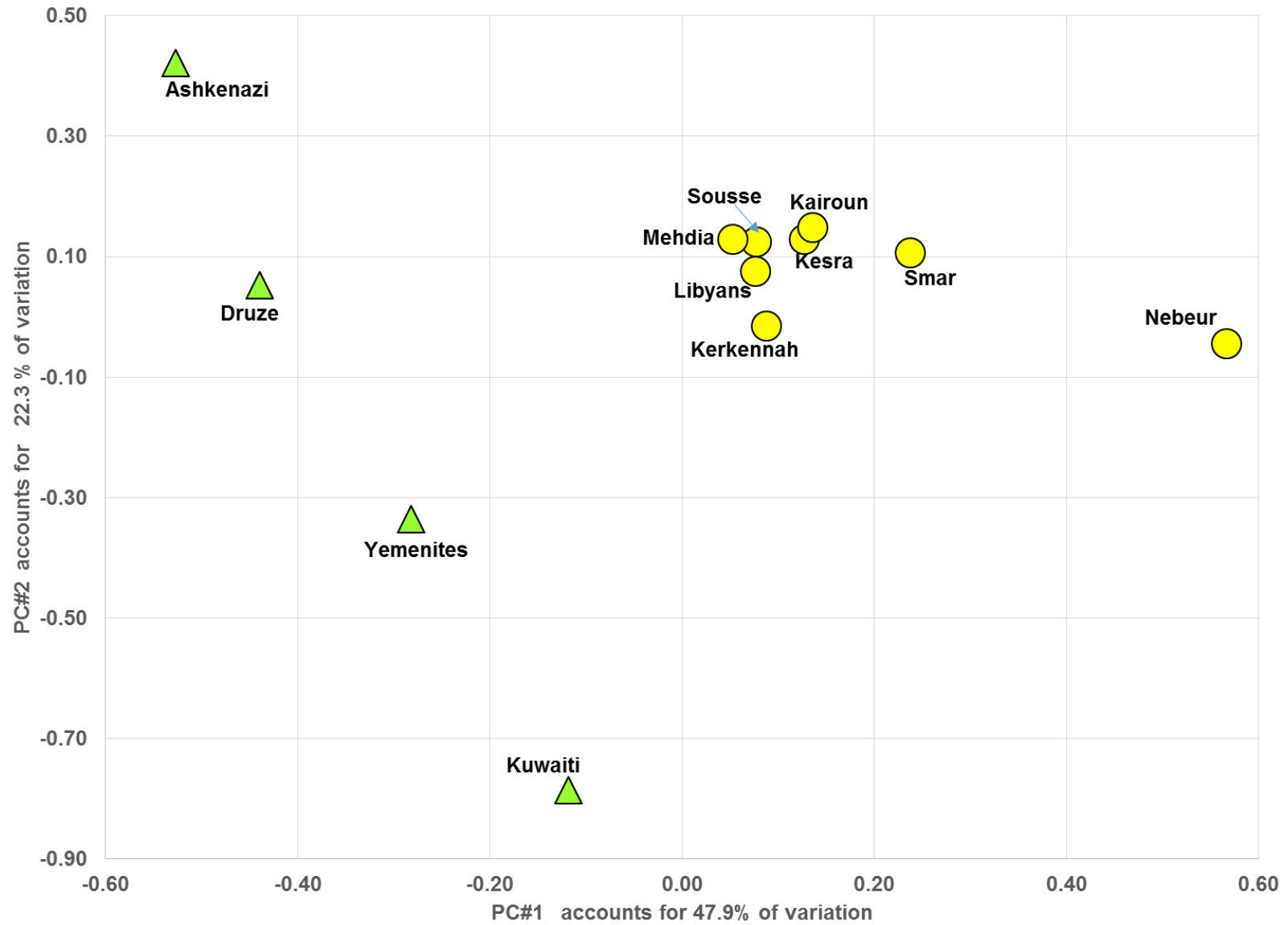


Fig. S22

PCA: 15 North African populations; 90 markers (81 single SNPs, 9 haplotypes)

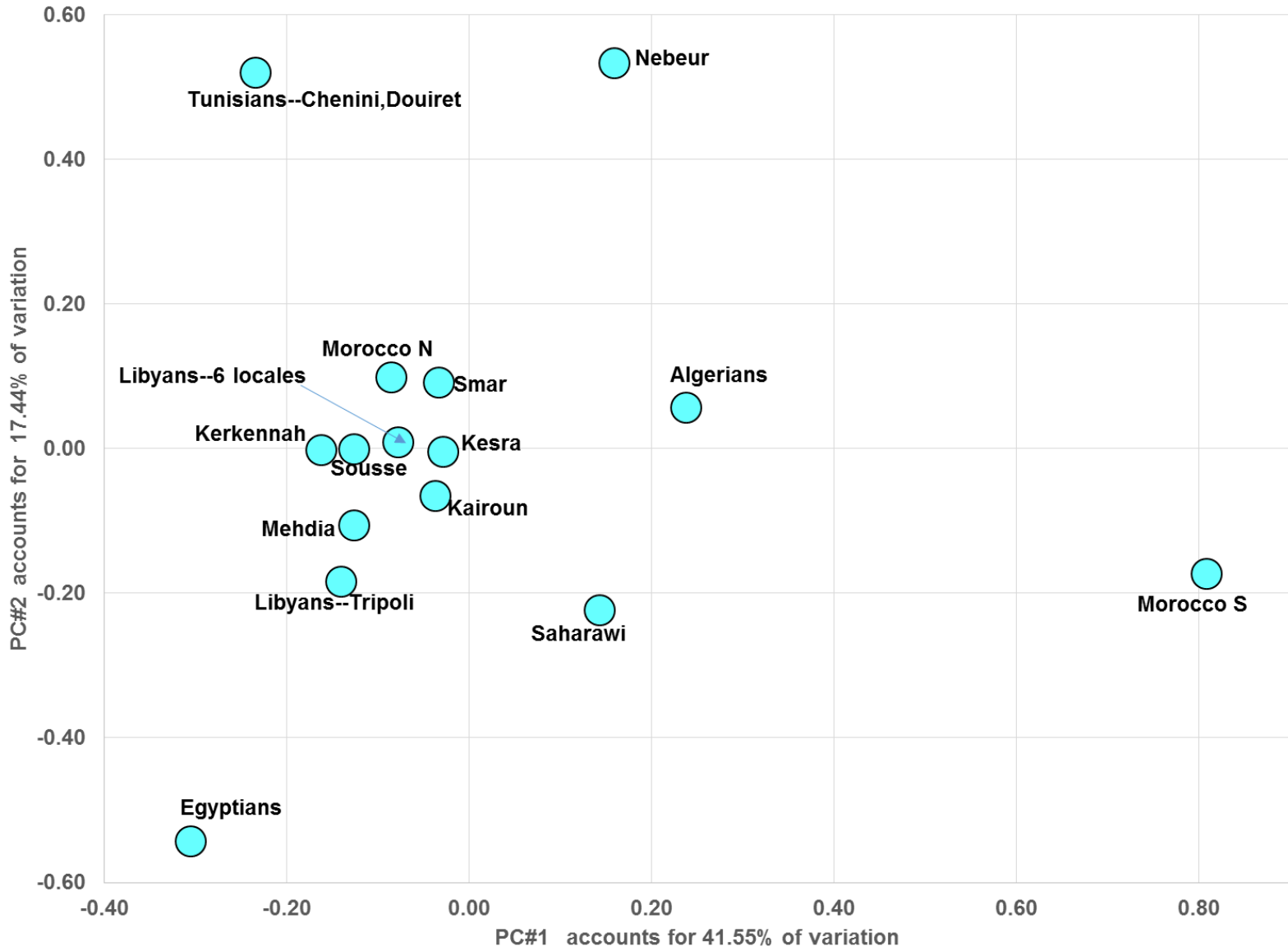


Fig. S23

PCA: N Afr, SW Asia--19 populations; 90 markers (81 single SNPs, 9 haplotypes)

