

## ADDITIONAL FILE 2: POSITIVE SELECTION IN ADMIXED POPULATIONS FROM ETHIOPIA

Sandra Walsh, Luca Pagani, Yali Xue, Hafid Laayouni, Chris Tyler-Smith,  
Jaume Bertranpetit

Percentile	SFselect		Mean iHS per window		-log <sub>10</sub> (p) iHS SNP based	
	99.99	99.90	99.99	99.90	99.99	99.90
Afroasiatic	2.31 (n= 20424)	1.93 (n= 20424)	2.53 (n=15809)	1.95 (n=15809)	3.62 (n=738269)	2.77 (n=738269)
Gumuz	2.24 (n= 24000)	1.56 (n= 24000)	2.54 (n=19240)	2.05 (n=19240)	3.88 (n=1055794)	2.98 (n= 1055794)

**Supplementary Table 1:** The 99.99<sup>th</sup> and 99.90<sup>th</sup> percentile thresholds of SFselect and iHS calculated after the neutral simulations. For SFselect and mean iHS we report in parenthesis the number of 30 kb simulated windows and for the SNP based iHS the number of SNPs used to calculate the thresholds.

Chr	Start	Stop	SFselect	Genes
chr2	232673695	232703695	4.89	<i>COPS7B</i>
chr11	71898574	71928574	4.77	<i>FOLR1, FOLR2</i>
chr1	147396808	147426808	4.58	<i>GPR89B</i>
chr11	71948574	71978574	4.50	<i>INPPL1, PHOX2A</i>
chr11	71923574	71953574	4.48	<i>FOLR2, INPPL1, PHOX2A</i>
* chr9	21215910	21245910	3.93	<i>IFNA16, IFNA17, IFNA14</i>
chr1	226096808	226126808	3.85	<i>LEFTY1, RP4-559A3.7, PYCR2, LEFTY2</i>
chr5	81280629	81310629	3.84	<i>ATG10</i>
chr15	32754330	32784330	3.63	<i>AC135983.2</i>
chr7	72433619	72463619	3.50	<i>TRIM74, AC005488.1</i>
chr1	149821808	149851808	3.36	<i>HIST2H2AA4, HIST2H3A, HIST2H4B</i>
chr17	30351088	30381088	3.35	<i>LRRC37B</i>
chr10	18164034	18194034	3.26	<i>MRC1</i>
chr5	140705629	140735629	3.22	<i>PCDHGA1, PCDHGA2, PCDHGA3, PCDHGB1, PCDHGA4</i>
chr8	48786376	48816376	3.09	<i>PRKDC</i>
chr2	155023695	155053695	3.06	<i>GALNT13</i>
chr3	130698294	130728294	3.04	<i>ATP2C1</i>
chr17	43676088	43706088	3.01	<i>CRHR1</i>
chr3	93598294	93628294	3.01	<i>PROS1</i>
* chr10	51214034	51244034	3.01	<i>AGAP8</i>
chr10	51189034	51219034	2.97	<i>FAM21D</i>
chr17	15526088	15556088	2.96	<i>RP11-385D13.1, TRIM16</i>
chr7	74158619	74188619	2.96	<i>GTF2I, NCF1</i>
chr17	30326088	30356088	2.88	<i>SUZ12, LRRC37B</i>
chr17	47876088	47906088	2.82	<i>KAT7</i>
* chr12	112399817	112429817	2.78	<i>TMEM116</i>
chr3	93773294	93803294	2.71	<i>ARL13B, DHFRL1, NSUN3</i>
* chr12	112374817	112404817	2.69	<i>TMEM116</i>
chr2	200773695	200803695	2.67	<i>C2orf69, TYW5</i>
chr3	93798294	93828294	2.66	<i>NSUN3</i>

**Supplementary Table 2A:** SFselect positive selection signals found in Amhara population. All the 30 kb window scores reported are above the 99.99 percentile threshold calculated after the simulations. Only protein coding genes are reported.

\* : Population-specific positive selection signals. These signals are above the 99.99 percentile threshold calculated after the simulations that are not found among the 99 percentile of the rest of the populations of study.

Chr	Start	Stop	SFselect	Genes
chr11	71923574	71953574	4.75	<i>FOLR2, INPPL1, PHOX2A</i>
chr17	15526088	15556088	4.39	<i>RP11-385D13.1, TRIM16</i>
chr10	46339034	46369034	4.39	<i>AGAP4</i>
chr11	71948574	71978574	4.26	<i>INPPL1, PHOX2A</i>
chr1	226096808	226126808	4.12	<i>LEFTY1, RP4-559A3.7, PYCR2, LEFTY2</i>
chr10	46489034	46519034	3.95	<i>PTPN20A</i>
chr2	232673695	232703695	3.83	<i>COPS7B</i>
chr1	78321808	78351808	3.72	<i>FAM73A</i>
chr1	147396808	147426808	3.67	<i>GPR89B</i>
chr13	114717755	114747755	3.62	<i>RASA3</i>
chr2	155023695	155053695	3.56	<i>GALNT13</i>
chr2	113198695	113228695	3.41	<i>RGPD8</i>
chr1	149796808	149826808	3.38	<i>HIST2H4A, HIST2H3C, HIST2H2AA3, HIST2H2AA4, HIST2H3A, HIST2H4B</i>
chr1	27096808	27126808	3.37	<i>ARID1A, PIGV</i>
chr7	74158619	74188619	3.26	<i>GTF2I, NCF1</i>
chr17	30351088	30381088	3.25	<i>LRRC37B</i>
chr5	45505629	45535629	3.21	<i>HCN1</i>
chr15	30829330	30859330	3.14	<i>GOLGA8Q</i>
chr5	81280629	81310629	3.10	<i>ATG10</i>
chr3	93698294	93728294	3.04	<i>ARL13B</i>
chr11	71898574	71928574	3.02	<i>FOLR1, FOLR2</i>
chr15	74954330	74984330	2.95	<i>EDC3</i>
chr16	22021755	22051755	2.88	<i>C16orf52</i>
chr14	36169850	36199850	2.76	<i>RALGAPA1</i>
chr15	32754330	32784330	2.76	<i>AC135983.2</i>
chr2	144973695	145003695	2.74	<i>GTDC1</i>
* chr20	32178005	32208005	2.67	<i>CBFA2T2</i>
chr11	66898574	66928574	2.63	<i>KDM2A</i>
chr15	30704330	30734330	2.62	<i>GOLGA8R</i>
chr5	36880629	36910629	2.62	<i>NIPBL</i>

**Supplementary Table 2B:** SFselect positive selection signals found in Oromo population. All the 30 kb window scores reported are above the 99.99 percentile threshold calculated after the simulations. Only protein coding genes are reported.

\* : Population-specific positive selection signals. These signals are above the 99.99 percentile threshold calculated after the simulations that are not found among the 99 percentile of the rest of the populations of study.

Chr	Start	Stop	SFselect	Genes
chr11	71923574	71953574	5.29	<i>FOLR2, INPPL1, PHOX2A</i>
chr13	53167755	53197755	4.30	<i>HNRNPA1L2</i>
chr1	147396808	147426808	4.18	<i>GPR89B</i>
chr1	78321808	78351808	4.10	<i>FAM73A</i>
chr11	71948574	71978574	3.95	<i>INPPL1, PHOX2A</i>
chr16	22021755	22051755	3.82	<i>C16orf52</i>
chr10	18164034	18194034	3.64	<i>MRC1</i>
chr2	232673695	232703695	3.48	<i>COPS7B</i>
chr3	93798294	93828294	3.38	<i>NSUN3</i>
chr1	146446808	146476808	3.34	<i>NBPF12</i>
chr15	30829330	30859330	3.26	<i>GOLGA8Q</i>
chr3	93598294	93628294	3.26	<i>PROS1</i>
chr3	93773294	93803294	3.17	<i>ARL13B, DHFRL1, NSUN3</i>
chr1	149821808	149851808	3.08	<i>HIST2H2AA4, HIST2H3A, HIST2H4B</i>
chr1	226096808	226126808	3.01	<i>LEFTY1, RP4-559A3.7, PYCR2, LEFTY2</i>
chr17	15526088	15556088	2.96	<i>RP11-385D13.1, TRIM16</i>
* chr16	22396755	22426755	2.93	<i>CDR2</i>
chr13	114717755	114747755	2.91	<i>RASA3</i>
chr7	72433619	72463619	2.91	<i>TRIM74, AC005488.1</i>
chr5	81280629	81310629	2.89	<i>ATG10</i>
chr2	113198695	113228695	2.88	<i>RGPD8</i>
chr6	50655815	50685815	2.80	<i>TFAP2D</i>
chr19	53032864	53062864	2.78	<i>ZNF808, ZNF701</i>
chr10	46339034	46369034	2.76	<i>AGAP4</i>
chr11	71898574	71928574	2.76	<i>FOLR1, FOLR2</i>
chr20	39678005	39708005	2.74	<i>TOP1</i>
chr6	83805815	83835815	2.71	<i>DOPEY1</i>
* chr7	99258619	99288619	2.68	<i>CYP3A5</i>
chr22	32476780	32506780	2.67	<i>SLC5A1</i>
chr1	161546808	161576808	2.66	<i>FCGR3A, FCGR2B</i>

**Supplementary Table 2C:** SFselect positive selection signals found in Somali population. All the 30 kb window scores reported are above the 99.99 percentile threshold calculated after the simulations. Only protein coding genes are reported.

\* : Population-specific positive selection signals. These signals are above the 99.99 percentile threshold calculated after the simulations that are not found among the 99 percentile of the rest of the populations of study.

Chr	Start	Stop	SFselect	Genes
chr11	71923574	71953574	5.12	<i>FOLR2, INPPL1, PHOX2A</i>
chr13	53167755	53197755	4.74	<i>HNRNPA1L2</i>
chr2	232673695	232703695	4.13	<i>COPS7B</i>
chr8	48786376	48816376	4.05	<i>PRKDC</i>
chr17	15526088	15556088	4.00	<i>RP11-385D13.1, TRIM16</i>
chr10	18164034	18194034	3.93	<i>MRC1</i>
chr11	71948574	71978574	3.93	<i>INPPL1, PHOX2A</i>
chr1	149796808	149826808	3.85	<i>HIST2H4A, HIST2H3C, HIST2H2AA3, HIST2H2AA4, HIST2H3A, HIST2H4B</i>
chr2	95748695	95778695	3.81	<i>MRPS5</i>
chr2	113198695	113228695	3.72	<i>RGPD8</i>
chr15	32754330	32784330	3.70	<i>AC135983.2</i>
chr15	30829330	30859330	3.63	<i>GOLGA8Q</i>
chr9	99740910	99770910	3.50	<i>HIATL2</i>
chr10	51189034	51219034	3.46	<i>FAM21D</i>
chr1	226096808	226126808	3.38	<i>LEFTY1, RP4-559A3.7, PYCR2, LEFTY2</i>
chr5	140705629	140735629	3.31	<i>PCDHGA1, PCDHGA2, PCDHGA3, PCDHGB1, PCDHGA4</i>
chr5	81280629	81310629	3.21	<i>ATG10</i>
* chr1	115121808	115151808	3.19	<i>BCAS2, DENND2C</i>
chr11	71898574	71928574	3.12	<i>FOLR1, FOLR2</i>
* chr19	57907864	57937864	3.06	<i>AC003002.4, ZNF548, AC003002.6, AC004076.7, ZNF17</i>
chr2	155023695	155053695	3.05	<i>GALNT13</i>
chr10	46489034	46519034	3.01	<i>PTPN20A</i>
chr17	47876088	47906088	2.94	<i>KAT7</i>
chr5	45355629	45385629	2.80	<i>HCN1</i>
chr1	78321808	78351808	2.77	<i>FAM73A</i>
chr16	70346755	70376755	2.73	<i>DDX19B, RP11-529K1.3</i>
chr5	36955629	36985629	2.71	<i>NIPBL</i>
chr22	32476780	32506780	2.69	<i>SLC5A1</i>
chr20	39678005	39708005	2.65	<i>TOP1</i>
chr6	118030815	118060815	2.63	<i>NUS1</i>

**Supplementary Table 2D:** SFselect positive selection signals found in Wolayta population. All the 30 kb window scores reported are above the 99.99 percentile threshold calculated after the simulations. Only protein coding genes are reported.

\* : Population-specific positive selection signals. These signals are above the 99.99 percentile threshold calculated after the simulations that are not found among the 99 percentile of the rest of the populations of study.

Chr	Start	Stop	SFselect	Genes
chr3	93798294	93828294	4.86	<i>NSUN3</i>
chr11	71948574	71978574	4.05	<i>INPPL1, PHOX2A</i>
chr10	46489034	46519034	3.97	<i>PTPN20A</i>
chr10	46539034	46569034	3.93	<i>PTPN20A</i>
chr3	93598294	93628294	3.81	<i>PROS1</i>
chr11	71898574	71928574	3.71	<i>FOLR1, FOLR2</i>
* chr9	40790910	40820910	3.41	<i>ZNF658</i>
chr2	113198695	113228695	3.38	<i>RGPD8</i>
chr13	114717755	114747755	3.33	<i>RASA3</i>
chr2	95748695	95778695	3.27	<i>MRPS5</i>
chr1	226096808	226126808	3.22	<i>LEFTY1, RP4-559A3.7, PYCR2, LEFTY2</i>
chr11	71923574	71953574	3.15	<i>FOLR2, INPPL1, PHOX2A</i>
chr17	30351088	30381088	3.03	<i>LRRC37B</i>
chr7	72433619	72463619	2.96	<i>TRIM74, AC005488.1</i>
chr19	53032864	53062864	2.93	<i>ZNF808, ZNF701</i>
chr8	48786376	48816376	2.86	<i>PRKDC</i>
* chr3	93723294	93753294	2.83	<i>ARL13B, STX19</i>
* chr2	233223695	233253695	2.79	<i>ALPP</i>
chr3	93748294	93778294	2.78	<i>ARL13B, DHFRL1</i>
chr15	32754330	32784330	2.77	<i>AC135983.2</i>
chr15	30704330	30734330	2.75	<i>GOLGA8R</i>
chr4	88110998	88140998	2.68	<i>KLHL8</i>
chr22	32451780	32481780	2.61	<i>SLC5A1</i>
chr2	200798695	200828695	2.58	<i>C2orf69, TYW5, C2orf47</i>
* chr9	40865910	40895910	2.56	<i>ZNF658</i>
chr6	25955815	25985815	2.56	<i>TRIM38</i>
chr10	77464034	77494034	2.55	<i>C10orf11</i>
* chr2	158448695	158478695	2.54	<i>ACVR1C</i>
chr9	97990910	98020910	2.53	<i>FANCC</i>
chr2	111373695	111403695	2.51	<i>BUB1</i>

**Supplementary Table 2E:** SFselect positive selection signals found in Gumuz population. All the 30 kb window scores reported are above the 99.99 percentile threshold calculated after the simulations. Only protein coding genes are reported.

\* : Population-specific positive selection signals. These signals are above the 99.99 percentile threshold calculated after the simulations that are not found among the 99 percentile of the rest of the populations of study.

Chr	Start	End	Amhara	Oromo	Somali	Wolayta	Gumuz	Genes	Type of gene
chr9	84515910	84545910	5.34	5.25	4.67	5.38	4.97	<i>RP11-383M4.2, RP11-383M4.6, SPATA31D4, SPATA31D5P</i>	lincRNA, lincRNA, PS, PS
chr11	123574	153574	5.04	4.50	2.58	5.09	4.39	<i>LINC01001, RP11-304M2.6, CICIP23, RNU6-447P</i>	lincRNA, lincRNA, PS, snRNA
chr11	71898574	71928574	4.77	3.02	2.76	3.12	3.71	<i>FOLR1, FOLR2</i>	PC
chr19	50582864	50612864	4.61	4.67	5.28	5.26	5.60	-	-
chr11	71948574	71978574	4.50	4.26	3.95	3.93	4.05	<i>INPPL1, PHOX2A, AP000593.5</i>	PC, PC, PS
chr11	71923574	71953574	4.48	4.75	5.29	5.12	3.15	<i>FOLR2, INPPL1, PHOX2A</i>	PC
chr1	226096808	226126808	3.85	4.12	3.01	3.38	3.22	<i>LEFTY1, LEFTY2, PYCR2</i>	PC
chr5	81280629	81310629	3.84	3.10	2.89	3.21	2.40	<i>ATG10, PPIAP11</i>	PC, PS
chr15	32754330	32784330	3.63	2.76	2.66	3.70	2.77	<i>AC135983.2, RP11-632K20.6</i>	PC, PS
chr3	93598294	93628294	3.01	2.47	3.26	2.54	3.81	<i>PROS1</i>	PC
chr2	96448695	96478695	2.94	3.14	2.77	2.95	2.93	<i>LINC00342, AC008268.2</i>	lincRNA, PS
chr1	571808	601808	2.78	3.00	2.62	2.99	2.85	<i>RP5-857K21.4</i>	lincRNA
chr20	39678005	39708005	2.59	2.49	2.74	2.65	2.51	<i>TOP1</i>	PC
chr15	30704330	30734330	2.52	2.62	2.38	2.45	2.75	<i>GOLGA8R, RP11-382B18.5</i>	PC, PS
chr4	89360998	89390998	2.42	2.40	2.38	2.43	2.44	<i>HERC5, HERC6</i>	PC
chr1	184021808	184051808	2.41	2.32	2.34	2.33	2.32	<i>TSEN15</i>	PC

**Supplementary Table 3:** SFselect positive selection signals shared among the five populations of study. All the reported SFselect scores of the 30 kb windows are above the 99.99 percentile threshold calculated after the simulations. We also report the genes within the window and the type of gene (PC: protein coding gene, PS: pseudogene, lincRNA: long interspersed non-coding RNA, small nuclear RNA: snRNA).

Chr	Start	End	Amhara		Oromo		Somali		Wolayta		Gumuz		Genes	Type of gene
			mean	maxP	mean	maxP	mean	maxP	mean	maxP	mean	maxP		
chr16	21746755	21776755	3.12	6.59	2.85	7.32	2.82	5.50	2.32	6.02	3.20	7.05	<i>OTOA</i>	PC
chr8	36136376	36166376	3.09	8.28	2.88	5.62	2.42	5.16	3.17	6.46	2.29	3.98	<i>MTND6P19, RP11-139F9.1, RP11-593P24.2, RP11-593P24.3</i>	PS
chr16	32296755	32326755	2.97	6.71	2.98	5.29	3.52	6.31	3.05	5.61	3.56	8.45	<i>RP11-17M15.2</i>	PS
chr16	33496755	33526755	2.54	4.87	2.53	5.34	2.88	5.11	2.46	4.70	2.06	5.21	<i>BMS1P8</i>	PS
chr7	65258619	65288619	2.45	5.09	2.66	7.49	2.13	5.39	2.66	6.28	2.74	5.27	<i>RNU6-912P, RP11-479O9.2</i>	snRNA, PS
chr16	32121755	32151755	2.35	2.88	2.53	5.37	2.08	3.06	2.62	3.38	2.89	3.67	<i>HERC2P4, RP11-1166P10.7</i>	PS
chr16	32321755	32351755	2.29	6.71	2.06	5.29	2.29	6.12	2.19	5.61	2.05	5.61	<i>RP11-17M15.2</i>	PS
chr4	510998	540998	2.24	5.67	2.30	3.71	2.68	6.25	2.12	3.54	2.35	4.53	<i>PIGG</i>	PC
chr2	398695	428695	2.14	11.89	2.95	12.85	2.55	12.36	2.36	11.78	2.45	15.05	<i>AC105393.1, AC105393.2</i>	lincRNA
chr16	31996755	32026755	1.95	2.23	2.61	2.59	2.25	3.08	2.37	3.01	3.37	4.47	<i>RP11-1166P10.1, RP11-1166P10.6</i>	PS

**Supplementary Table 4:** iHS positive selection signals shared among the five populations of study. For each population the mean iHS score for the 30 kb windows is reported and the highest  $-\log(p\text{-value})$  of a variant within the window. The windows were selected for comparison if their mean iHS score was above the 99.99 percentile threshold in at least one population and in the 99.9 threshold in the rest. In gray, the only shared window that is in all populations with mean iHS scores above the 99.99 threshold. We also report the genes within the window and the type of gene (PC: protein coding gene, PS: pseudogene, lincRNA: long interspersed non-coding RNA, small nuclear RNA: snRNA).



Chr	Start	Stop	mean <i>iHS</i>	maxP	Genes	
chr18	6247301	6277301	4.14	4.67	<i>L3MBTL4</i>	
chr1	202396808	202426808	3.57	6.71	<i>PPP1R12B</i>	
chr18	66647301	66677301	3.34	6.67	<i>CCDC102B</i>	
chr16	21746755	21776755	3.12	6.58	<i>OTOA</i>	
chr9	74490910	74520910	3.05	4.08	<i>ABHD17B</i>	
chr1	244721808	244751808	2.99	5.66	<i>C1orf101</i>	
chr18	66722301	66752301	2.98	8.46	<i>CCDC102B</i>	
chr1	244696808	244726808	2.96	5.73	<i>C1orf101</i>	
chr6	128505815	128535815	2.92	4.82	<i>PTPRK</i>	
chr11	85173574	85203574	2.92	5.16	<i>DLG2</i>	
chr12	22524817	22554817	2.89	9.50	<i>ST8SIA1</i>	
chr12	113674817	113704817	2.86	5.59	<i>TPCN1</i>	
chr1	244671808	244701808	2.77	5.18	<i>C1orf101</i>	
chr6	157480815	157510815	2.73	4.75	<i>ARID1B</i>	
*	chr6	136905815	136935815	2.71	6.33	<i>MAP3K5</i>
	chr2	158323695	158353695	2.71	4.83	<i>CYTIP</i>
	chr12	113699817	113729817	2.68	4.06	<i>TPCN1</i>
	chr11	75573574	75603574	2.66	4.31	<i>UVRAG</i>
*	chr4	47210998	47240998	2.63	4.79	<i>GABRB1</i>
	chr22	46701780	46731780	2.58	4.62	<i>GTSE1, TRMU</i>
	chr9	40815910	40845910	2.54	5.23	<i>ZNF658</i>
	chr11	85198574	85228574	2.53	4.76	<i>DLG2</i>

**Supplementary Table 5A:** *iHS* positive selection signals found in Amhara population. The mean *iHS* score for the 30 kb windows is reported and the highest  $-\log(p\text{-value})$  of a variant within the window. All the reported *iHS* scores are above the reported 99.99 percentile thresholds calculated after the simulations. Only protein coding genes are reported.

\* : Population-specific positive selection signals. These signals are above the 99.99 percentile threshold calculated after the simulations that are not found among the 99 percentile of the rest of the populations of study.

Chr	Start	Stop	mean <i>iHS</i>	max <i>P</i>	Genes
chr3	19248294	19278294	3.33	7.46	<i>KCNH8</i>
chr1	202396808	202426808	3.14	6.77	<i>PPP1R12B</i>
chr17	79776088	79806088	3.08	4.25	<i>FAM195B, AC174470.1, PPP1R27, P4HB</i>
chr2	237448695	237478695	3.07	5.37	<i>ACKR3</i>
chr12	8299817	8329817	2.96	2.77	<i>ZNF705A</i>
chr11	72023574	72053574	2.94	6.60	<i>CLPB</i>
chr3	19298294	19328294	2.91	4.79	<i>KCNH8</i>
chr2	158323695	158353695	2.87	6.16	<i>CYTIP</i>
chr11	72048574	72078574	2.86	4.86	<i>CLPB</i>
chr16	21746755	21776755	2.85	7.32	<i>OTOA</i>
chr12	82849817	82879817	2.82	7.11	<i>METTL25</i>
chr17	44351088	44381088	2.72	8.42	<i>ARL17B, LRRC37A</i>
chr3	19273294	19303294	2.69	4.79	<i>KCNH8</i>
chr11	75473574	75503574	2.60	3.09	<i>DGAT2</i>
chr11	75523574	75553574	2.60	7.06	<i>UVRAG</i>
chr3	19173294	19203294	2.59	5.31	<i>KCNH8</i>
chr12	45674817	45704817	2.58	3.82	<i>ANO6</i>
chr12	113674817	113704817	2.56	3.62	<i>TPCN1</i>
chr1	149821808	149851808	2.55	3.95	<i>HIST2H2AA4, HIST2H3A, HIST2H4B</i>
chr12	64224817	64254817	2.55	5.39	<i>SRGAP1</i>
chr10	70414034	70444034	2.54	4.44	<i>TET1</i>
* chr3	164673294	164703294	2.54	4.79	<i>SI</i>
chr15	44879330	44909330	2.54	4.07	<i>SPG11</i>
chr12	82799817	82829817	2.54	5.21	<i>METTL25</i>

**Supplementary Table 5B:** *iHS* positive selection signals found in Oromo population. The mean *iHS* score for the 30 kb windows is reported and the highest  $-\log(p\text{-value})$  of a variant within the window. All the reported *iHS* scores are above the reported 99.99 percentile thresholds calculated after the simulations. Only protein coding genes are reported.

\* : Population-specific positive selection signals. These signals are above the 99.99 percentile threshold calculated after the simulations that are not found among the 99 percentile of the rest of the populations of study.

Chr	Start	Stop	meanIHS	maxP	Genes	
chr11	75448574	75478574	3.52	7.94	<i>DGAT2</i>	
chr11	75523574	75553574	3.40	8.77	<i>UVRAG</i>	
chr2	136248695	136278695	3.32	5.99	<i>ZRANB3</i>	
chr11	75498574	75528574	3.22	6.24	<i>DGAT2, UVRAG</i>	
chr21	30808956	30838956	3.17	5.72	<i>BACH1</i>	
chr1	51596808	51626808	3.13	3.72	<i>C1orf185</i>	
chr1	70471808	70501808	3.12	6.91	<i>LRRC7</i>	
chr2	135973695	136003695	3.10	5.40	<i>ZRANB3</i>	
chr11	75548574	75578574	3.04	7.65	<i>UVRAG</i>	
*	chr2	136273695	136303695	2.98	6.50	<i>ZRANB3, R3HDM1</i>
	chr11	75723574	75753574	2.96	3.80	<i>UVRAG</i>
	chr2	136073695	136103695	2.94	5.07	<i>ZRANB3</i>
*	chr9	28615910	28645910	2.94	5.38	<i>LINGO2</i>
*	chr9	28640910	28670910	2.94	5.08	<i>LINGO2</i>
	chr2	237448695	237478695	2.92	4.86	<i>ACKR3</i>
	chr2	136023695	136053695	2.91	4.86	<i>ZRANB3</i>
*	chr19	3232864	3262864	2.88	6.38	<i>CELF5</i>
*	chr15	42654330	42684330	2.86	5.87	<i>CAPN3</i>
	chr12	44674817	44704817	2.85	6.04	<i>TMEM117</i>
	chr9	40815910	40845910	2.85	5.65	<i>ZNF658</i>
	chr11	75473574	75503574	2.83	3.54	<i>DGAT2</i>
	chr1	51696808	51726808	2.83	4.72	<i>RNF11</i>
	chr16	21746755	21776755	2.82	5.50	<i>OTOA</i>
	chr12	56724817	56754817	2.82	3.85	<i>PAN2, IL23A, STAT2, APOF</i>
	chr2	136123695	136153695	2.82	5.74	<i>ZRANB3</i>
	chr4	71635998	71665998	2.80	5.71	<i>RUFY3</i>
	chr17	58826088	58856088	2.75	4.69	<i>BCAS3</i>
	chr4	71610998	71640998	2.75	6.67	<i>RUFY3</i>
	chr2	135998695	136028695	2.73	5.40	<i>ZRANB3</i>
*	chr10	57364034	57394034	2.72	2.85	<i>PCDH15</i>
	chr2	136098695	136128695	2.72	4.33	<i>ZRANB3</i>
	chr7	137258619	137288619	2.71	6.41	<i>DGKI</i>
	chr4	510998	540998	2.68	6.25	<i>PIGG</i>
	chr9	74490910	74520910	2.68	3.93	<i>ABHD17B</i>
*	chr2	135823695	135853695	2.67	3.93	<i>RAB3GAP1</i>
	chr9	74515910	74545910	2.66	4.85	<i>ABHD17B, C9orf85</i>
	chr12	45799817	45829817	2.66	5.86	<i>ANO6</i>
	chr1	51571808	51601808	2.65	4.53	<i>C1orf185</i>
*	chr13	52942755	52972755	2.63	2.27	<i>THSD1</i>
	chr7	123658619	123688619	2.63	7.45	<i>TMEM229A</i>
	chr8	145211376	145241376	2.62	5.91	<i>MROH1</i>
*	chr1	225646808	225676808	2.61	4.72	<i>ENAH</i>
	chr2	136148695	136178695	2.59	4.94	<i>ZRANB3</i>
*	chr2	135773695	135803695	2.57	4.89	<i>MAP3K19</i>
	chr12	11174817	11204817	2.57	5.91	<i>PRR4, TAS2R14, TAS2R19, TAS2R31, AC018630.1</i>
	chr4	159235998	159265998	2.56	5.33	<i>RXFP1</i>
	chr8	145161376	145191376	2.55	7.43	<i>SHARPIN, MAF1, KIAA1875</i>
	chr11	75573574	75603574	2.54	5.48	<i>UVRAG</i>
	chr7	137233619	137263619	2.53	6.41	<i>DGKI</i>
	chr7	74558619	74588619	2.53	3.73	<i>GTF2IRD2B</i>

**Supplementary Table 5C:** iHS positive selection signals found in Somali population. The mean iHS score for the 30 kb windows is reported and the highest  $-\log(p\text{-value})$  of a variant within the window. All the reported iHS scores are above the reported 99.99 percentile thresholds calculated after the simulations. Only protein coding genes are reported.

\* : Population-specific positive selection signals. These signals are above the 99.99 percentile threshold calculated after the simulations that are not found among the 99 percentile of the rest of the populations of study.

Chr	Start	Stop	meanIHS	maxP	Genes	
chr1	202396808	202426808	3.59	6.66	<i>PPP1R12B</i>	
chr18	66647301	66677301	3.41	6.40	<i>CCDC102B</i>	
chr17	44351088	44381088	3.20	9.36	<i>ARL17B, LRRC37A</i>	
chr19	39757864	39787864	3.14	5.36	<i>IFNL2, IFNL1</i>	
chr18	66722301	66752301	3.09	7.24	<i>CCDC102B</i>	
chr1	149821808	149851808	3.03	5.01	<i>HIST2H2AAA, HIST2H3A, HIST2H4B</i>	
chr12	82849817	82879817	2.98	6.06	<i>METTL25</i>	
chr17	30276088	30306088	2.82	4.76	<i>SUZ12</i>	
*	chr11	71823574	71853574	2.82	2.76	<i>ANAPC15, FOLR3</i>
	chr17	63001088	63031088	2.82	6.13	<i>GNA13</i>
	chr8	145211376	145241376	2.81	6.46	<i>MROH1</i>
	chr17	44326088	44356088	2.81	9.36	<i>ARL17B</i>
	chr9	16565910	16595910	2.79	6.40	<i>BNC2</i>
	chr5	144980629	145010629	2.74	5.92	<i>PRELID2</i>
	chr8	124261376	124291376	2.73	5.69	<i>ZHX1-C8ORF76, ZHX1</i>
	chr18	66672301	66702301	2.72	5.42	<i>CCDC102B</i>
	chr17	62501088	62531088	2.72	2.92	<i>DDX5, CEP95</i>
	chr12	22824817	22854817	2.69	4.37	<i>ETNK1</i>
	chr8	124211376	124241376	2.68	4.99	<i>FAM83A, C8orf76, ZHX1-C8ORF76</i>
	chr2	173473695	173503695	2.65	4.94	<i>PDK1</i>
	chr1	51571808	51601808	2.63	4.00	<i>C1orf185</i>
	chr20	20678005	20708005	2.62	4.46	<i>RALGAPA2</i>
	chr20	20653005	20683005	2.62	3.97	<i>RALGAPA2</i>
	chr17	74751088	74781088	2.60	4.32	<i>MFSD11</i>
	chr8	124236376	124266376	2.60	4.61	<i>C8orf76, ZHX1-C8ORF76, ZHX1</i>
*	chr16	85621755	85651755	2.58	7.75	<i>GSE1</i>
*	chr10	45789034	45819034	2.58	3.33	<i>OR13A1</i>
	chr1	51696808	51726808	2.57	4.10	<i>RNF11</i>
	chr3	19298294	19328294	2.56	5.36	<i>KCNH8</i>
	chr12	86999817	87029817	2.56	6.20	<i>MGAT4C</i>
	chr3	19248294	19278294	2.54	4.84	<i>KCNH8</i>

**Supplementary Table 5D:** iHS positive selection signals found in Wolayta population. The mean iHS score for the 30 kb windows is reported and the highest  $-\log(p\text{-value})$  of a variant within the window. All the reported iHS scores are above the reported 99.99 percentile thresholds calculated after the simulations. Only protein coding genes are reported.

\* : Population-specific positive selection signals. These signals are above the 99.99 percentile threshold calculated after the simulations that are not found among the 99 percentile of the rest of the populations of study.

	Chr	Start	Stop	mean <i>i</i> HS	maxP	Genes
*	chr10	102014034	102044034	3.74	5.78	<i>CWF19L1, BLOC1S2</i>
*	chr10	101939034	101969034	3.45	5.57	<i>ERLIN1, CHUK</i>
*	chr10	101964034	101994034	3.37	9.85	<i>CHUK, CWF19L1</i>
	chr1	202396808	202426808	3.33	5.98	<i>PPP1R12B</i>
	chr20	20528005	20558005	3.31	7.06	<i>RALGAPA2</i>
	chr16	21746755	21776755	3.20	7.05	<i>OTOA</i>
*	chr10	101989034	102019034	3.16	6.87	<i>CHUK, CWF19L1</i>
*	chr6	128080815	128110815	3.14	6.05	<i>THEMIS</i>
*	chr10	101914034	101944034	3.11	6.23	<i>ERLIN1</i>
*	chr7	12258619	12288619	3.05	7.44	<i>TMEM106B</i>
	chr8	145161376	145191376	3.05	8.95	<i>SHARPIN, MAF1, KIAA1875</i>
	chr12	113849817	113879817	3.05	5.31	<i>SDS, SDSL</i>
*	chr6	128005815	128035815	3.00	5.00	<i>THEMIS</i>
	chr12	113824817	113854817	2.92	7.44	<i>PLBD2, SDS</i>
*	chr4	130010998	130040998	2.89	5.75	<i>SCLT1, C4orf33</i>
	chr9	16565910	16595910	2.86	7.70	<i>BNC2</i>
*	chr6	154330815	154360815	2.79	5.43	<i>OPRM1</i>
	chr2	229773695	229803695	2.75	6.00	<i>PID1</i>
*	chr16	23096755	23126755	2.75	5.04	<i>USP31</i>
*	chr16	23071755	23101755	2.73	5.11	<i>USP31</i>
	chr7	120983619	121013619	2.72	3.66	<i>FAM3C</i>
	chr6	128105815	128135815	2.70	6.05	<i>THEMIS</i>
	chr16	32046755	32076755	2.69	3.77	<i>AC142381.1</i>
	chr2	98848695	98878695	2.69	4.35	<i>VWA3B</i>
*	chr6	128055815	128085815	2.68	4.27	<i>THEMIS</i>
	chr7	74558619	74588619	2.68	3.41	<i>GTF2IRD2B</i>
	chr8	145236376	145266376	2.67	4.31	<i>MROH1</i>
*	chr7	65883619	65913619	2.67	4.09	<i>TPST1</i>
*	chr6	154305815	154335815	2.67	5.17	<i>OPRM1</i>
*	chr16	23121755	23151755	2.64	5.03	<i>USP31</i>
	chr2	98823695	98853695	2.63	4.54	<i>VWA3B</i>
*	chr7	133983619	134013619	2.62	5.29	<i>SLC35B4</i>
	chr11	16048574	16078574	2.61	5.27	<i>SOX6</i>
*	chr10	102039034	102069034	2.58	8.06	<i>BLOC1S2, PKD2L1</i>
*	chr6	33030815	33060815	2.56	6.90	<i>HLA-DPA1, HLA-DPB1</i>
	chr7	121008619	121038619	2.56	3.66	<i>FAM3C</i>
*	chr4	41985998	42015998	2.55	4.93	<i>DCAF4L1, SLC30A9</i>

**Supplementary Table 5E:** *i*HS positive selection signals found in Gumuz population. The mean *i*HS score for the 30 kb windows is reported and the highest  $-\log(p\text{-value})$  of a variant within the window. All the reported *i*HS scores are above the reported 99.99 percentile thresholds calculated after the simulations. Only protein coding genes are reported.

\* : Population-specific positive selection signals. These signals are above the 99.99 percentile threshold calculated after the simulations that are not found among the 99 percentile of the rest of the populations of study.