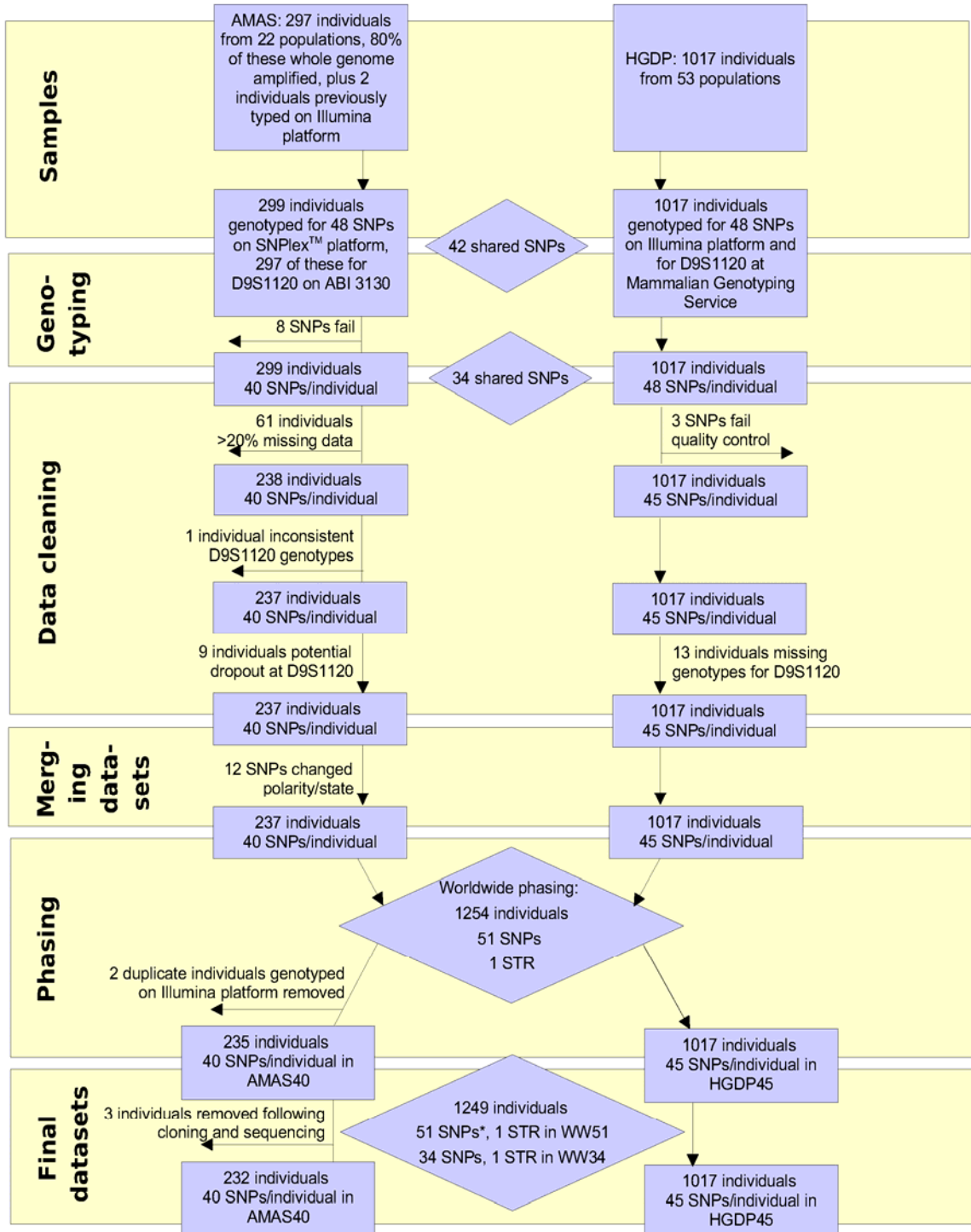


Figure S1. Flow chart describing the genotyping, data cleaning, and phasing of the two sample sets in this study.



\*Minimum of 11 imputed SNP genotypes for AMAS40 samples in WW51 and minimum of 6 imputed SNP genotypes for HGDP45 samples in WW34

Table S1. Populations sampled in this study.

<b>Population</b>	<b>Symbol</b>	<b>Sample size</b>	<b>Region</b>	<b>Dataset</b>
Aleut	34	17	Americas	AMAS40
Apache	39	19	Americas	AMAS40
Cherokee	49	10	Americas	AMAS40
Chippewa	45	11	Americas	AMAS40
Creek	47	4	Americas	AMAS40
Dogrib	37	11	Americas	AMAS40
Fox	48	2	Americas	AMAS40
Huichol	42	9	Americas	AMAS40
Inuit	53	3	Americas	AMAS40
Jemez	41	6	Americas	AMAS40
Mixtec	44	27	Americas	AMAS40
NorthernPaiute	36	8	Americas	AMAS40
Seri	38	9	Americas	AMAS40
Sioux	43	5	Americas	AMAS40
Washo	35	9	Americas	AMAS40
Karitiana	51	24	Americas	HGDP45
Maya	46	25	Americas	HGDP45
Piapoco	50	13	Americas	HGDP45
Pima	40	21	Americas	HGDP45
Surui	52	21	Americas	HGDP45
Chukchi	33	22	East-Central Asia	AMAS40
Koryaks	32	19	East-Central Asia	AMAS40
AltaiKazakh	13	14	East-Central Asia	AMAS40
Mongola-OuterMongolia	20	20	East-Central Asia	AMAS40
NorthernAltai	12	2	East-Central Asia	AMAS40
SouthernAltai	11	5	East-Central Asia	AMAS40
Cambodian	19	9	East-Central Asia	HGDP45
Dai	16	10	East-Central Asia	HGDP45
Daur	27	10	East-Central Asia	HGDP45
Han	24	34	East-Central Asia	HGDP45
Han-NorthernChina	23	10	East-Central Asia	HGDP45
Hezhen	30	9	East-Central Asia	HGDP45
Japanese	31	29	East-Central Asia	HGDP45
Lahu	15	10	East-Central Asia	HGDP45
Miao	22	10	East-Central Asia	HGDP45
Mongola-InnerMongolia	25	10	East-Central Asia	HGDP45
Naxi	14	9	East-Central Asia	HGDP45
Oroqen	28	10	East-Central Asia	HGDP45
She	26	10	East-Central Asia	HGDP45
Tu	17	10	East-Central Asia	HGDP45
Tujia	21	10	East-Central Asia	HGDP45
Uygur	9	10	East-Central Asia	HGDP45
Xibo	10	9	East-Central Asia	HGDP45
Yakut	29	25	East-Central Asia	HGDP45
Yi	18	10	East-Central Asia	HGDP45
Balochi	2	24	South Asia	HGDP45

Brahui	3	24	South Asia	HGDP45
Burusho	8	23	South Asia	HGDP45
Hazara	5	24	South Asia	HGDP45
Kalash	7	25	South Asia	HGDP45
Makrani	1	25	South Asia	HGDP45
Pathan	6	24	South Asia	HGDP45
Sindhi	4	25	South Asia	HGDP45
Bedouin	NA	47	Middle East	HGDP45
Druze	NA	45	Middle East	HGDP45
Mozabite	NA	29	Middle East	HGDP45
Palestinian	NA	48	Middle East	HGDP45
Adygei	NA	15	Europe	HGDP45
Basque	NA	24	Europe	HGDP45
French	NA	29	Europe	HGDP45
Italian	NA	12	Europe	HGDP45
Orcadian	NA	14	Europe	HGDP45
Russian	NA	24	Europe	HGDP45
Sardinian	NA	27	Europe	HGDP45
Tuscan	NA	6	Europe	HGDP45
Melanesian	NA	19	Oceania	HGDP45
Papuan	NA	16	Oceania	HGDP45
BantuKenya	NA	12	Africa	HGDP45
BantuSouthernAfrica	NA	8	Africa	HGDP45
BiakaPygmy	NA	31	Africa	HGDP45
Mandenka	NA	23	Africa	HGDP45
MbutiPygmy	NA	14	Africa	HGDP45
San	NA	7	Africa	HGDP45
Yoruba	NA	25	Africa	HGDP45

Table S2. SNPs for which the samples in this study were genotyped.

<b>Locus</b>	<b>Position (NCBI Build 35)</b>	<b>Dataset</b>
rs17088334	85053298	HGDP45
rs17337860	85084053	HGDP45
rs17088374	85096450	WW34
rs6559853	85110900	HGDP45
rs2841445	85126731	WW34
rs17088420	85141513	HGDP45
rs2841486	85155600	WW34
rs4877918	85181132	WW34
rs7032592	85193290	WW34
rs7862916	85206478	WW34
rs2814734	85222343	WW34
rs2841443	85225848	AMAS40
rs2592991	85230060	AMAS40
rs13291799	85237265	WW34
rs2593017	85240224	AMAS40
rs2841453	85244429	AMAS40
rs12685505	85247160	AMAS40
rs12555508	85250750	AMAS40
rs10512167	85253684	WW34
rs7048862	85269825	WW34
rs3849872	85284382	WW34
rs11140976	85300443	WW34
rs11140984	85314284	WW34
rs17426617	85314387	WW34
rs6559867	85316239	WW34
D9S1120	85321417	-
rs4877301	85321930	WW34
rs3849873	85322587	WW34
rs1992812	85329420	HGDP45
rs1447026	85329946	WW34
rs7042753	85331227	WW34
rs4877940	85332664	HGDP45
rs1374500	85339567	WW34
rs4082114	85339786	HGDP45
rs10512168	85345810	WW34
rs1014690	85351123	HGDP45
rs1834264	85363552	WW34
rs2278111	85378655	WW34
rs1030303	85395220	HGDP45
rs10481762	85407691	WW34
rs1863120	85420237	HGDP45
rs1469202	85435373	WW34
rs11141033	85448369	HGDP45
rs2814724	85463213	WW34
rs7025722	85480379	WW34
rs7031647	85506670	WW34
rs10117745	85525431	WW34
rs7863248	85537681	WW34
rs4877949	85551964	WW34
rs4877950	85570008	WW34
rs3916184	85582785	WW34
rs12379804	85595034	WW34

Table S3. SNPs for which we changed the minor allele or allele state in AMAS40 due to different genotyping platforms (Illumina and SNPlex<sup>Tm</sup>) used for HGDP45 and AMAS40. Two samples from HGDP45 were genotyped on the new platform (SNPlex<sup>Tm</sup>) with AMAS40 to assist in determining changes in minor alleles or allele states.

<b>Locus</b>	<b>Position</b>	<b>Change in minor allele?</b>	<b>Change in allele states?</b>	<b>HGDP45</b>	<b>AMAS40</b>	<b>Polarity apparent from homozygous HGDP controls?</b>	<b>Regional allele frequency consistent with controls?</b>
rs7042753	85331227	-	Y	A/G	T/C	Y	Y
rs11140976	85300443	-	Y	A/G	T/C	Y	Ambiguous
rs4877949	85551964	-	Y	A/G	T/C	Y	Ambiguous
rs12379804	85595034	Y	-	C/G	G/C	Y	Y
rs10481762	85407691	-	Y	G/T	C/A	Y	Y
rs2841445	85126731	-	Y	G/T	C/A	Y	Y
rs2278111	85378655	-	Y	T/C	A/G	Y	Y
rs2841486	85155600	-	Y	T/C	A/G	Y	Y
rs10117745	85525431	-	Y	T/C	A/G	N	Y
rs3849872	85284382	-	Y	T/C	A/G	Y	Ambiguous
rs2814724	85463213	-	Y	T/G	A/C	Y	Ambiguous

Table S4. Phase error rates for different geographic groupings of the populations.

<b>Grouping</b>	<b>Error rate (10% of genotypes masked)</b>
Worldwide (no grouping)	0.1022
East Asia plus Americas	0.0923
East Asia	0.1161
Americas	0.0904
East Asia minus Western Beringia	0.0847
Americas plus Western Beringia	0.0984

Table S5. PCR and sequencing primers used to determine short haplotypes for three chromosomes with the 9-repeat allele and with haplotypes very divergent from the AMH.

<b>Primer</b>	<b>Primer sequence (5' to 3')</b>	<b>Use</b>
D9-1440R <sup>2</sup>	GGGTATGCTGAGGATTAATGAG	PCR
D9-35F <sup>2</sup>	AGGATTTGAGACAAATGAAAGCA	PCR
D9S1120R <sup>1</sup>	TTAGCTGCTTCTGGGAAAGA	PCR and sequencing
D9S1120F <sup>1</sup>	TAGGATTTGAGACAAATGAAAGC	PCR and sequencing
M13-20-F	GTAAAACGACGGCCAGT	sequencing
M13-R	AACAGCTATGACCATG	sequencing
D9-630F	ATTTCCCAAATATAGTTGATCGT	sequencing
D9-1408R	ATAAAGCACTTGGTATGTCACAG	sequencing

<sup>1</sup> Short fragment <sup>2</sup> Long fragment

Table S6. Chromosomes with the 9-repeat allele and non-AMH haplotypes in WW34. The position of each SNP, in Mb, is for NCBI Build 35. From D9S1120 in either direction, cells were colored red until an allele that differed from the AMH.

Recombinant haplotypes are labeled R1 through R7.

		Position	rs SNP ID/ microsatellite	85.254	85.270	85.284	85.300	85.314	85.314	85.316	85.321	85.322	85.323	85.330
Sample	Haplotype			10512167	7048862	3849872	11140976	11140984	17426617	6559867	D9S1120	4877301	3849873	1447026
-	AMH			G	T	C	A	A	T	A	9	G	C	A
Inuit_99	R1			G	T	C	A	A	T	A	9	C	C	A
Dogrib_09	R2			G	C	C	A	A	T	A	9	G	C	A
Washo_101	R3			G	T	C	A	A	T	A	9	G	C	G
N.Paiute_046	R4			A	T	C	A	A	T	A	9	G	C	G
Koryak_114	R5			A	T	C	A	A	T	A	9	G	C	A
Koryak_66	R5			A	T	C	A	A	T	A	9	G	C	A
Apache_492	R5			A	T	C	A	A	T	A	9	G	C	A
Apache_398	R5			A	T	C	A	A	T	A	9	G	C	A
Chippewa_15	R5			A	T	C	A	A	T	A	9	G	C	A
Chippewa_03	R5			A	T	C	A	A	T	A	9	G	C	A
Washo_101	R5			A	T	C	A	A	T	A	9	G	C	A
Washo_104	R5			A	T	C	A	A	T	A	9	G	C	A
N.Paiute_078	R5			A	T	C	A	A	T	A	9	G	C	A
Cherokee_55	R5			A	T	C	A	A	T	A	9	G	C	A
Cherokee_53	R5			A	T	C	A	A	T	A	9	G	C	A
Inuit_52	R6			G	C	C	G	G	A	A	9	G	C	A
Jemez_24	R7			A	T	C	A	A	A	G	9	G	C	A