Table S1

Summary of Latino populations and assembled reference panels

ADMIXED LATINO ADMIXED LATINO	Population Cuban in South Florida Colombian in South Florida Dominican in South Florida Puerto Rican in South Florida Honduran in South Florida Haitian in South Florida Haitian in South Florida TOTAL SOUTH FLORIDA Puerto Rican in New York Dominican in New York Colombian in New York Colombian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	Pop ID CUB COL (FL) DOM (FL) PUR (FL) HON HAI PUR (NY) DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	Sample size ¹ 80 85 34 27 19 6 251 27 26 20 70 70 80 571 30 25 25	Subset in low density ADMIXTURE ² 52 50 18 18 13 4 155 26 26 20 70 70 48 441	Subset in high density ADMIXTURE ³ 52 50 18 18 13 4 155 48 203 29 24	Region/country	Array platform affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 illumina 650K illumina 650K illumina 650K illumina 650K illumina 0mni 2.5 illumina 0mni 2.5 affymetrix 6.0	Study/source Present study Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b Holo Genomes 1000 Genomes HapMap3 Mao et al. 2007
ADMIXED LATINO ADMIXED LATINO	Cuban in South Florida Colombian in South Florida Dominican in South Florida Puerto Rican in South Florida Honduran in South Florida Haitian in South Florida Haitian in South Florida Haitian in South Florida TOTAL SOUTH FLORIDA Puerto Rican in New York Dominican in New York Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	CUB COL (FL) DOM (FL) PUR (FL) HON HAI PUR (NY) DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	80 85 34 27 19 6 251 27 26 20 70 80 571 30 25 25	ADMIXTURE ² 52 50 18 18 18 13 4 155 26 26 20 70 70 48 441 -	ADMIXTURE ³ 52 50 18 18 18 13 4 155 48 203 29		affymetrix 6.0 illumina 650K illumina 650K illumina 650K illumina 0mni 2.5 illumina 0mni 2.5 affymetrix 6.0	Present study Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
ADMIXED LATINO ADMIXED LATINO	Cuban in South Florida Colombian in South Florida Dominican in South Florida Puerto Rican in South Florida Honduran in South Florida Haitian in South Florida Haitian in South Florida Haitian in South Florida TOTAL SOUTH FLORIDA Puerto Rican in New York Dominican in New York Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	CUB COL (FL) DOM (FL) PUR (FL) HON HAI PUR (NY) DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	80 85 34 27 19 6 251 27 26 20 70 70 80 571 30 25 25	52 50 18 18 13 4 155 26 26 20 70 70 48 441	52 50 18 18 13 4 155 - - - - - 48 203 29		affymetrix 6.0 illumina 650K illumina 650K illumina 650K illumina 0mni 2.5 illumina 0mni 2.5 affymetrix 6.0	Present study Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
ADMIXED LATINO ADMIXED LATINO	Colombian in South Florida Dominican in South Florida Puerto Rican in South Florida Honduran in South Florida Haitian in South Florida Haitian in South Florida TOTAL SOUTH FLORIDA Puerto Rican in New York Dominican in New York Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Duerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	COL (FL) DOM (FL) PUR (FL) HON HAI PUR (NY) DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	85 34 27 19 6 251 27 27 26 20 70 70 80 571 30 25 25	50 18 18 13 4 155 26 26 26 20 70 70 48 441	50 18 18 13 4 155 - - - - - 48 203 29		affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 illumina 650K illumina 650K illumina 650K illumina 650K illumina 0mni 2.5 illumina Omni 2.5 affymetrix 6.0	Present study Present study Present study Present study Present study Present study Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
ADMIXED LATINO	Dominican in South Florida Puerto Rican in South Florida Honduran in South Florida Haitian in South Florida Haitian in South Florida TOTAL SOUTH FLORIDA Puerto Rican in New York Dominican in New York Colombian in New York Colombian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	DOM (FL) PUR (FL) HON HAI PUR (NY) DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	34 27 19 6 251 27 27 26 20 70 70 80 571 30 25 25	18 18 13 4 155 26 26 26 20 70 70 48 441	18 18 13 4 155 - - - - - - - - 48 203 29		affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 illumina 650K illumina 650K illumina 650K illumina 650K illumina 0mni 2.5 illumina Omni 2.5 affymetrix 6.0	Present study Present study Present study Present study Present study Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
ADMIXED LATINO	Puerto Rican in South Florida Honduran in South Florida Haitian in South Florida Haitian in South Florida TOTAL SOUTH FLORIDA Puerto Rican in New York Dominican in New York Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	DOM (FL) PUR (FL) HON HAI PUR (NY) DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	27 19 6 251 27 27 26 20 70 70 80 571 30 25 25	18 18 13 4 155 26 26 26 20 70 70 48 441	18 13 4 155 - - - - - - - - 48 203 29		affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 illumina 650K illumina 650K illumina 650K illumina 650K illumina 0mni 2.5 illumina 0mni 2.5 affymetrix 6.0	Present study Present study Present study Present study Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
ADMIXED LATINO	Honduran in South Florida Haitian in South Florida Haitian in South Florida TOTAL SOUTH FLORIDA Puerto Rican in New York Dominican in New York Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	PUR (FL) HON HAI PUR (NY) DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	27 19 6 251 27 27 26 20 70 70 80 571 30 25 25	18 13 4 155 26 26 26 20 70 70 70 48 441	18 13 4 155 - - - - - - - - 48 203 29		affymetrix 6.0 affymetrix 6.0 affymetrix 6.0 illumina 650K illumina 650K illumina 650K illumina 650K illumina 0mni 2.5 illumina 0mni 2.5 affymetrix 6.0	Present study Present study Present study Present study Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
ADMIXED LATING	Haitian in South Florida TOTAL SOUTH FLORIDA Puerto Rican in New York Dominican in New York Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	PUR (NY) DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	6 251 27 27 26 20 70 70 80 571 30 25	4 155 26 26 26 20 70 70 48 441	13 4 155 - - - - - - - 48 203 29		affymetrix 6.0 affymetrix 6.0 illumina 650K illumina 650K illumina 650K illumina 650K illumina Omni 2.5 illumina Omni 2.5 affymetrix 6.0	Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
ADMIXED LATING	Haitian in South Florida TOTAL SOUTH FLORIDA Puerto Rican in New York Dominican in New York Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	PUR (NY) DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	6 251 27 27 26 20 70 70 80 571 30 25	4 155 26 26 26 20 70 70 48 441	4 155 - - - - - - - - 48 203 29		affymetrix 6.0 illumina 650K illumina 650K illumina 650K illumina 650K illumina 0mni 2.5 illumina Omni 2.5 affymetrix 6.0	Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
F N T	TOTAL SOUTH FLORIDA Puerto Rican in New York Dominican in New York Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	PUR (NY) DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	251 27 27 26 20 70 70 80 571 30 25	155 26 26 26 20 70 70 48 441	155 - - - - - - 48 203 29		illumina 650K illumina 650K illumina 650K illumina 650K illumina Omni 2.5 illumina Omni 2.5 affymetrix 6.0	Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
F N T	Puerto Rican in New York Dominican in New York Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	27 27 26 20 70 70 80 571 30 25	26 26 26 20 70 70 48 441	- - - - - - 48 203 29		illumina 650K illumina 650K illumina 650K illumina 0mni 2.5 illumina 0mni 2.5 affymetrix 6.0	Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
F N T	Dominican in New York Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	DOM (NY) COL (NY) ECU CLM (1KG) PUR (1KG) MXL	27 26 20 70 70 80 571 30 25 25	26 26 20 70 70 48 441	- - - - 48 203 29	-	illumina 650K illumina 650K illumina 650K illumina 0mni 2.5 illumina 0mni 2.5 affymetrix 6.0	Bryc et al. 2010b Bryc et al. 2010b Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
F N T	Colombian in New York Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	COL (NY) ECU CLM (1KG) PUR (1KG) MXL	26 20 70 70 80 571 30 25 25	26 20 70 70 48 441	- - - 48 203 29	-	illumina 650K illumina 650K illumina Omni 2.5 illumina Omni 2.5 affymetrix 6.0	Bryc et al. 2010b Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
F N T	Ecuadorian in New York Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	ECU CLM (1KG) PUR (1KG) MXL	20 70 70 80 571 30 25 25	20 70 70 48 441	- - 48 203 29		illumina 650K illumina Omni 2.5 illumina Omni 2.5 affymetrix 6.0	Bryc et al. 2010b 1000 Genomes 1000 Genomes HapMap3
F N T N	Colombian in Medellin Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa	CLM (1KG) PUR (1KG) MXL	70 70 80 571 30 25 25	70 70 48 441 -	- - 48 203 29		illumina Omni 2.5 illumina Omni 2.5 affymetrix 6.0	1000 Genomes 1000 Genomes HapMap3
F N T N C	Puerto Rican in Puerto Rico Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa Warao	PUR (1KG) MXL	70 80 571 30 25 25	70 48 441 -	- 48 203 29	-	illumina Omni 2.5 affymetrix 6.0	1000 Genomes HapMap3
N T N C	Mexican in Los Angeles TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa Warao	MXL	80 571 30 25 25	48 441 - -	48 203 29		affymetrix 6.0	НарМар3
N N C	TOTAL ADMIXED LATINOS Nahua Maya Quechua Aymara Bari Yukpa Warao		571 30 25 25	441 - -	203 29		,	
N C A	Nahua Maya Quechua Aymara Bari Yukpa Warao		30 25 25	-	29	Mesoamerica	affymetriy 500K	Mao et al. 2007
N C A	Maya Quechua Aymara Bari Yukpa Warao		25 25	-				
A B	Quechua Aymara Bari Yukpa Warao	-	25	-		Mesoamerica	affymetrix 500K	Mao et al. 2007
A	Aymara Bari Yukpa Warao	-			24	Andes	affymetrix 500K	Mao et al. 2007
В	Bari Yukpa Warao	-		_	25	Andes	affymetrix 500K	Mao et al. 2007
	Yukpa Warao		29	27	27	Venezuela	affymetrix 6.0	Present study
E AMERICA	Warao	-	25	25	25	Venezuela	affymetrix 6.0	Present study
E AMER		_	25	23	23	Venezuela	affymetrix 6.0	Present study
E AM	Eskimo-Aleut	-	23	23	-	North America	illumina 650K	Reich et al. 2012
I ш ''	Na-Dene	-	15	15	_	North America	illumina 650K	Reich et al. 2012
· > \	Northern Amerind		93	93	-	North America	illumina 650K	Reich et al. 2012
	Central Amerind	-	108	108	-	North America	illumina 650K	Reich et al. 2012
2 2	Chibchan-Paezan	-	65	65	-	Central America	illumina 650K	Reich et al. 2012
	Andean	-	97	97	-	South America	illumina 650K	Reich et al. 2012
	Ge-Pano-Carib	-	12	12	-	South America	illumina 650K	Reich et al. 2012
		-	80	80	-	South America	illumina 650K	Reich et al. 2012
	Equatorial-Tucanoan	-				South America	IIIuIIIIIa 050K	Reich et al. 2012
	TOTAL NATIVE AMERICAN ⁴		677	568	177			_
	European (North West)	EUR NW	266	40	40	-	affymetrix 500K	POPRES ⁶
E	European (North/Norh East)	EUR NNE	76	15	15	-	affymetrix 500K	POPRES
E	European (South East)	EUR SE	96	8	8	-	affymetrix 500K	POPRES
E	European (East/South East)	EUR ESE	8	-	-	-	affymetrix 500K	POPRES
_ E	European (Central)	EUR C	186	20	20	-	affymetrix 500K	POPRES
	European (West)	EUR W	259	60	60	-	affymetrix 500K	POPRES
B E	European (South)	EUR S	232	20	20	-	affymetrix 500K	POPRES
EUROPEAN	European (South West)	EUR SW	264	40	40	-	affymetrix 500K	POPRES
╷╙│т	TOTAL POPRES ⁵		1387	203	203			
	lberian in Andalusia	IBE AND	17	-	-	Spain	affymetrix 6.0	Botigue et al. 2013
_	Iberian in Galicia	IBE GAL	17	_	_	Spain	affymetrix 6.0	Botigue et al. 2013
l li	Iberian in the Basque Country	IBE BAS	20	_	_	Spain	affymetrix 6.0	Henn et al. 2012
	TOTAL EUROPEAN		1441	203	203			
	Yoruba	YRI	167	50	58	West Africa	affymetrix 6.0	НарМар3
	Luhya	LWK	90	-	-	East Africa	affymetrix 6.0	НарМар3
	Bamoun	-	20	-	-	West Africa	affymetrix 500K	Bryc et al. 2010a
	Fang	-	18	-	-	West Africa	affymetrix 500K	Bryc et al. 2010a
.:	lgbo	-	17	-	-	West Africa	affymetrix 500K	Bryc et al. 2010a
l 뉴 Lk	Kongo	-	9	-	-	West Africa	affymetrix 500K	Bryc et al. 2010a
_	Brong	-	8	-	-	West Africa	affymetrix 500K	Bryc et al. 2010a
N	Mandenka	-	24	-	-	West Africa	illumina 650K	HGDP
T	TOTAL AFRICAN		353	50	58			
T	TOTAL		3042	1262	641			

¹Total of samples included in the study. For Native American, European, and African populations, this is the maximum number of samples used to construct the reference panels (e.g., for ASPCA analyses). ²The low-density dataset consists of 30,860 SNPs from a representative subset of most populations with Affymetrix SNP array data merged with populations with Illumina SNP array data. Numbers vary from the initial sample size either due to QC filtering (e.g., trios offspring, cryptic related individuals, and PCA outliers) or to broadly equalize sample sizes across populations (e.g., within POPRES collection).

³The high-density dataset consists of 389,225 SNPs from a representative subset of most populations with Affymetrix SNP array data. Only a subset of YRI samples were included in order to focus on European and Native American sub-continental structure (other analyses, such as ASPCA, were performed focusing on sub-continental African ancestry including all West African populations). ⁴The subset of 493 Native American samples from Reich et al (2012) represent a total of 52 populations. However, for summary purposes, these are shown grouped by linguistic families as in the original publication.

⁵We restricted to 1,387 POPRES European samples with four grandparents from the same country as in Novembre et al. (2008) to ensure replication of the PCA map of Europe. Geographic groups are as in Auton et al (2009). Full details of studied populations by country are available in Novembre et al. (2008).

⁶The collections and methods for the Population Reference Sample (POPRES) are described by Nelson et al. (2008). The datasets used for the analyses described in this manuscript were obtained from dbGaP at

http://www.ncbi.nlm.nih.gov/projects/gap/cgibin/study.cgi?study_id=phs000145.v1.p1 through dbGaP accession number phs000145.v1.p1.