

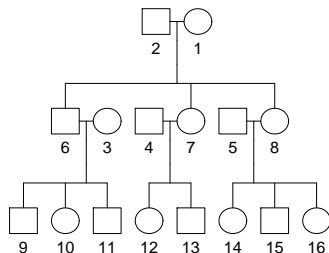
Association Testing with X Chromosome Data

An Application To HCHS/SOL

Caitlin McHugh, with Tim Thornton

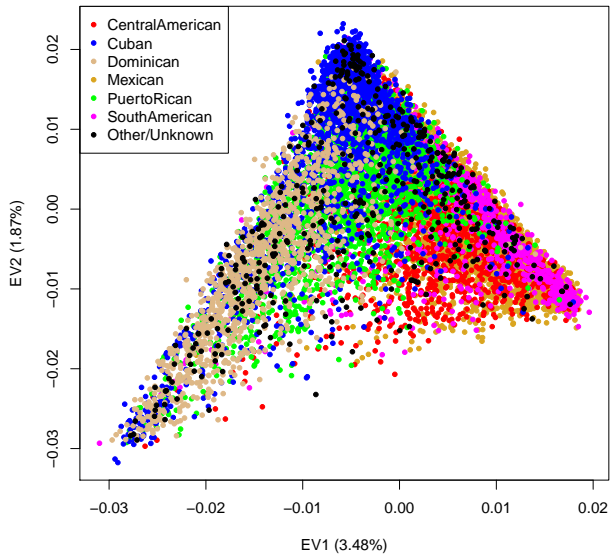
Department of Biostatistics
University of Washington

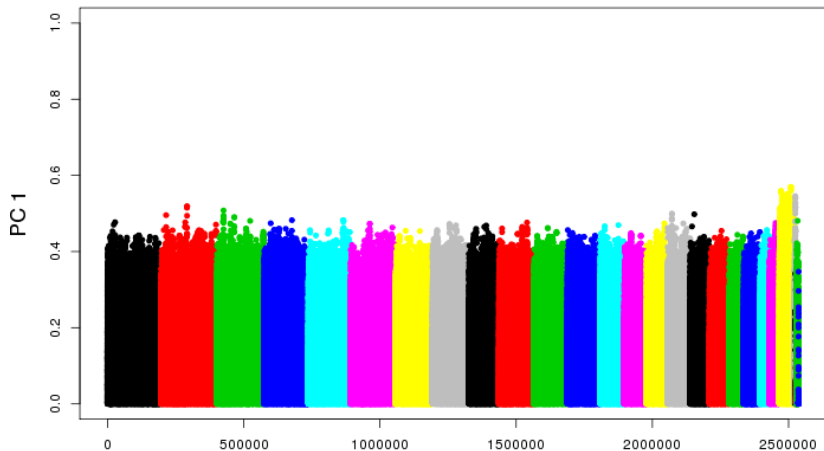
23 Feb 2015

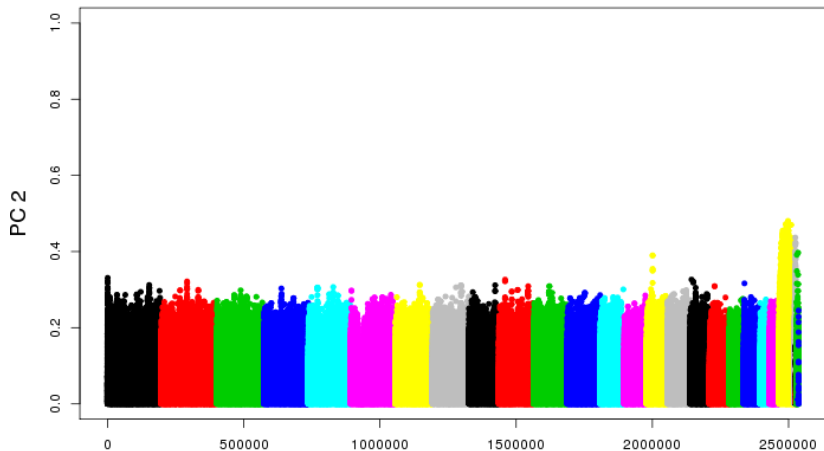


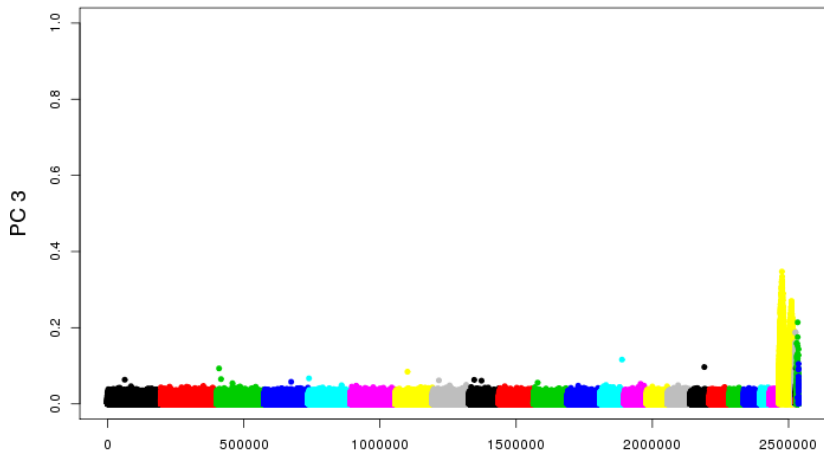
		Autosomes	X Chromosome
	Self, Female	$\frac{1}{2}$	$\frac{1}{2}$
	Self, Male	$\frac{1}{2}$	1
	Mother-Daughter	$\frac{1}{4}$	$\frac{1}{4}$
	Mother-Son, Father-Daughter	$\frac{1}{4}$	$\frac{1}{2}$
	Father-Son	$\frac{1}{4}$	0
	Full sisters	$\frac{1}{4}$	$\frac{6}{16}$
	Full brothers	$\frac{1}{4}$	$\frac{1}{2}$
	Sister-Brother	$\frac{1}{4}$	$\frac{1}{4}$
Maternal	Aunt-Niece	$\frac{1}{8}$	$\frac{3}{16}$
	Aunt-Nephew	$\frac{1}{8}$	$\frac{6}{16}$
	Uncle-Niece	$\frac{1}{8}$	$\frac{1}{8}$
	Uncle-Nephew	$\frac{1}{8}$	$\frac{1}{4}$
	Grandma-Granddaughter	$\frac{1}{8}$	$\frac{1}{8}$
	Grandma-Grandson	$\frac{1}{8}$	$\frac{1}{4}$
	Grandpa-Granddaughter	$\frac{1}{8}$	$\frac{1}{4}$
	Grandpa-Grandson	$\frac{1}{8}$	$\frac{1}{2}$
Paternal	Aunt-Niece	$\frac{1}{8}$	$\frac{1}{8}$
	Aunt-Nephew	$\frac{1}{8}$	0
	Uncle-Niece	$\frac{1}{8}$	0
	Uncle-Nephew	$\frac{1}{8}$	0
	Grandma-Granddaughter	$\frac{1}{8}$	$\frac{1}{4}$
	Grandma-Grandson	$\frac{1}{8}$	0
	Grandpa-Granddaughter	$\frac{1}{8}$	0
	Grandpa-Grandson	$\frac{1}{8}$	0

- ▶ We estimated PCs in the SOL subjects using 3,600 LD-pruned X chromosome SNPs and PC-AiR.
- ▶ The unrelated set `unrelated.pcair.deg4` of 10,272 samples as defined from the autosomes was set, and only study samples (`subj.plink & geno.cnt1==0`) excluding `gengrp6.outliers` were projected for a total of 12,747 samples.

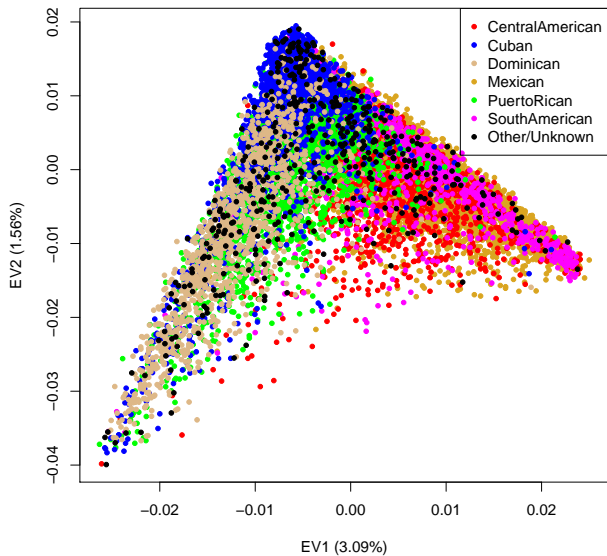


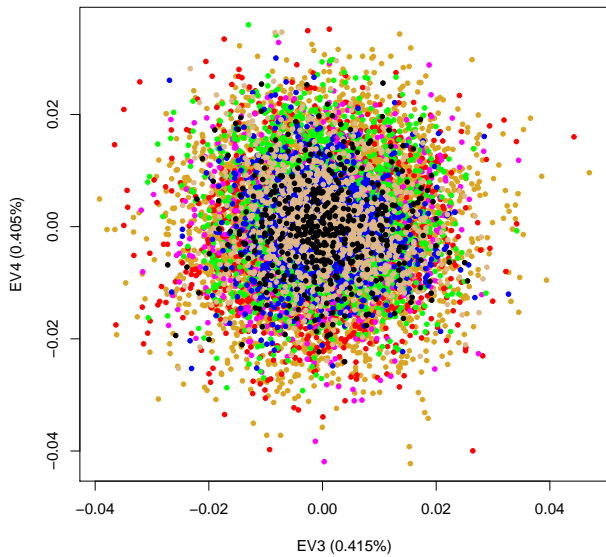


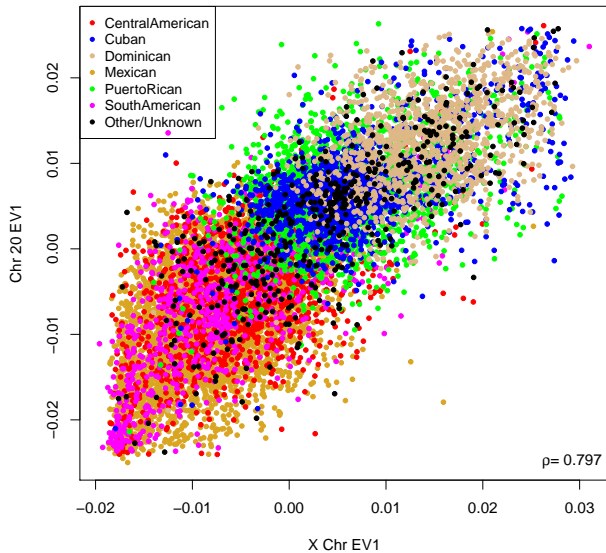


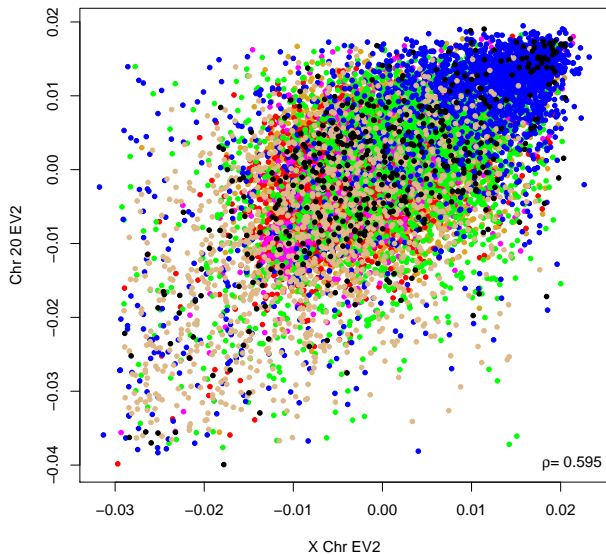


- ▶ We compare these results which use only 3,600 X chromosome SNPs to a pruned set of 4,413 chromosome 20 SNPs.

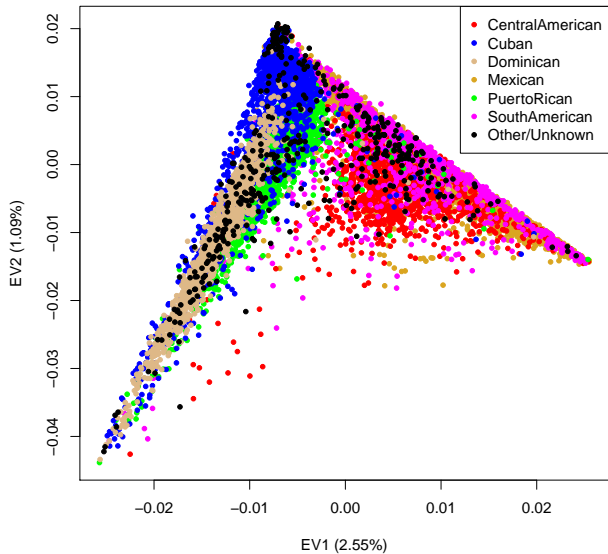


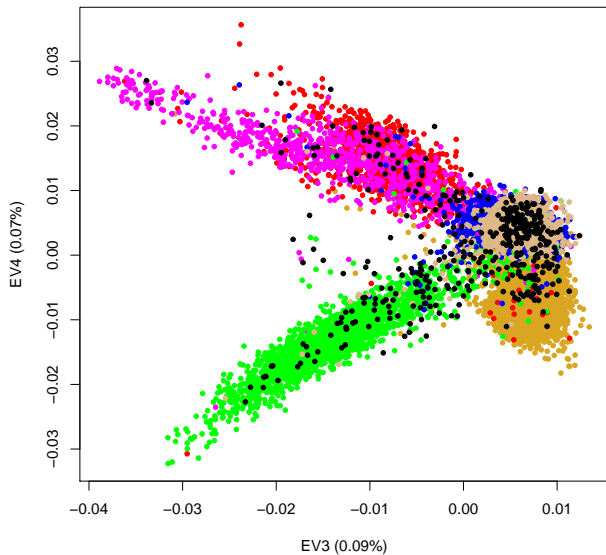


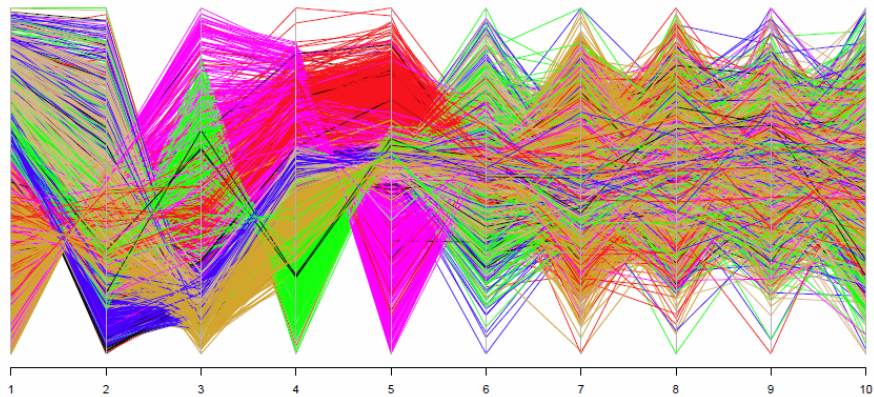


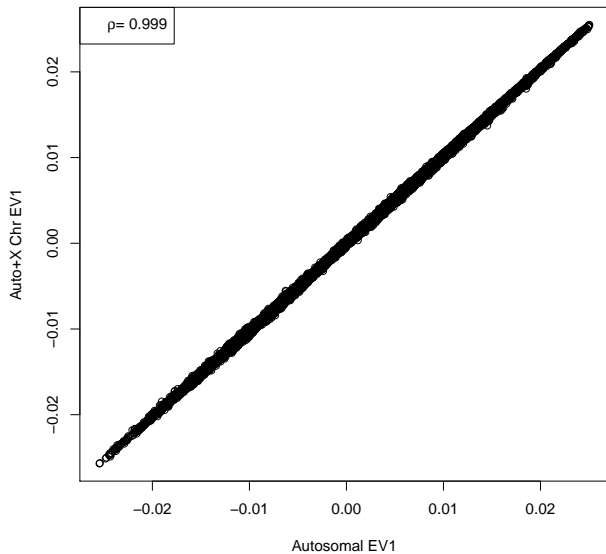


- ▶ We examine results which use both 155,196 pruned autosomal and 3,582 pruned X chromosome SNPs together, for a total SNP set of 158,778 SNPs across the genome.









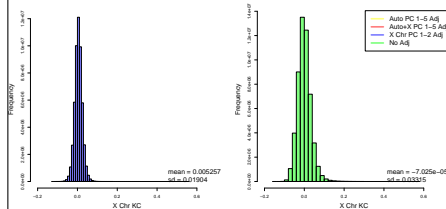
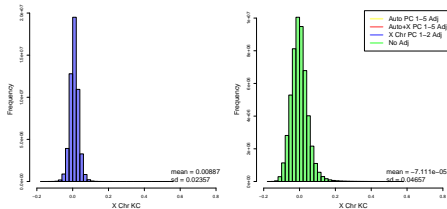
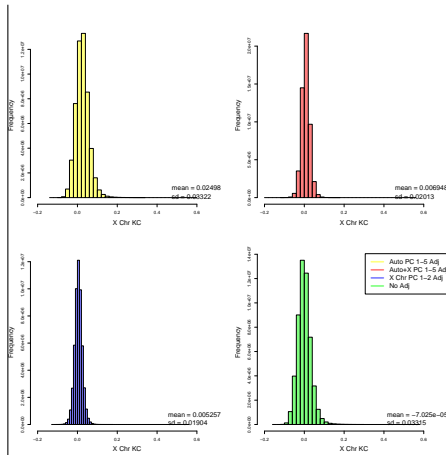
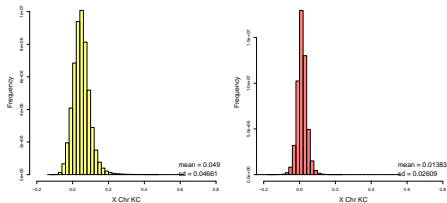
Φ_X was estimated for all OLGA samples using the following scenarios:

	SNP set	PCA run	EVs used
1	all X chr	-	-
2	all X chr	autosomes + X chr	1-5
3	all X chr	autosomes	1-5
4	all X chr	X chr	1-2
5	pruned X chr	-	-
6	pruned X chr	autosomes + X chr	1-5
7	pruned X chr	autosomes	1-5
8	pruned X chr	X chr	1-2
9	pruned autosomal	autosomes	1-5

All settings used the autosomal unrelated set of 10,272 samples and estimated Φ_X for 12,734 study samples posted to dbGaP with no X chromosome anomalies.

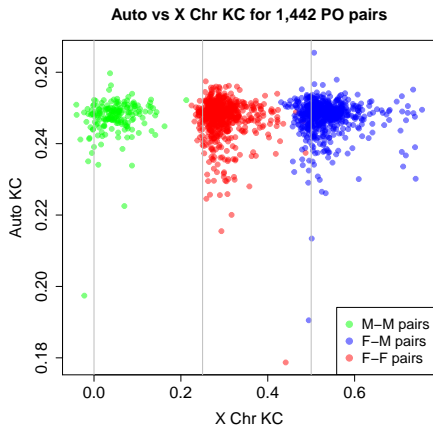
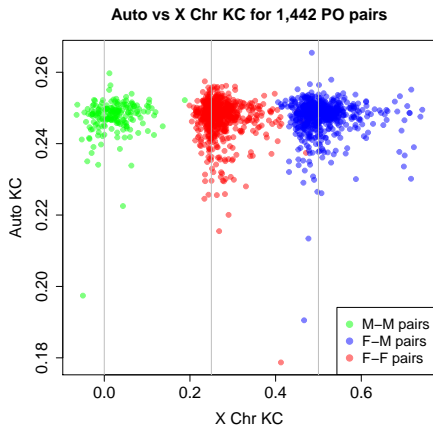
Estimate of Φ_X in 10,272 autosomal-unrelated samples unpruned X chr SNPs

pruned X chr SNPs



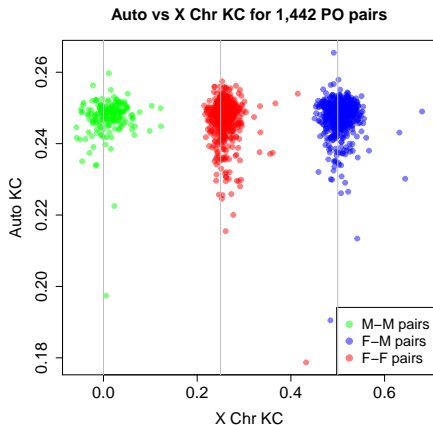
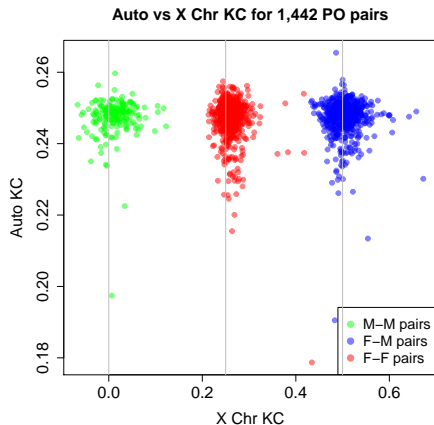
Pruned X chr SNPs, Unadjusted

Pruned X chr SNPs, Auto PC 1-5 Corr



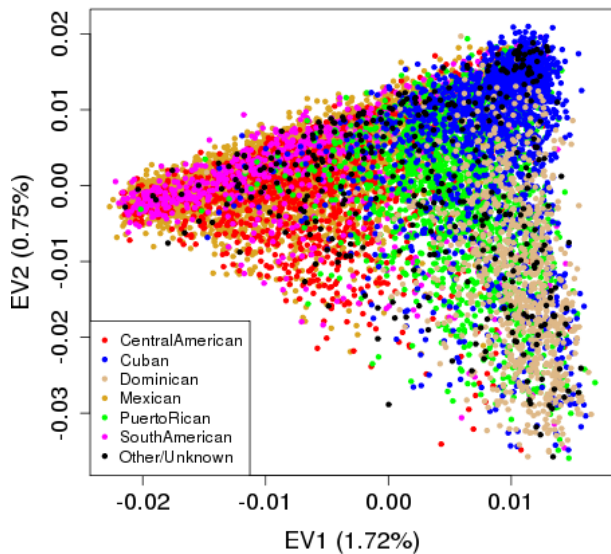
Pruned X chr SNPs, Auto+X PC 1-5 Corr

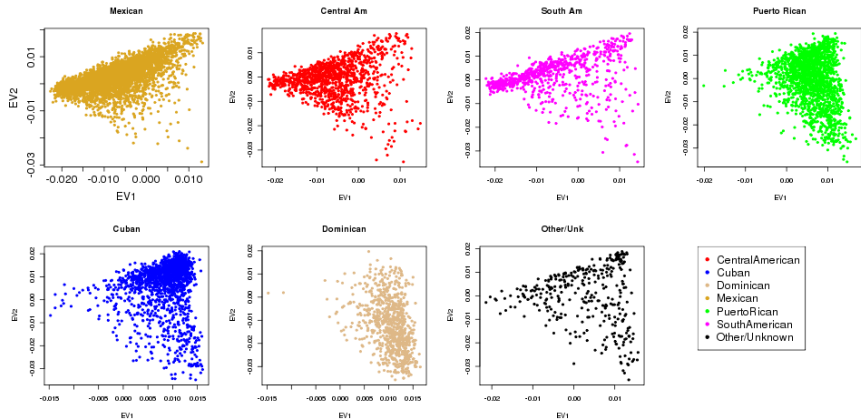
Pruned X chr SNPs, X PC 1-2 Corr

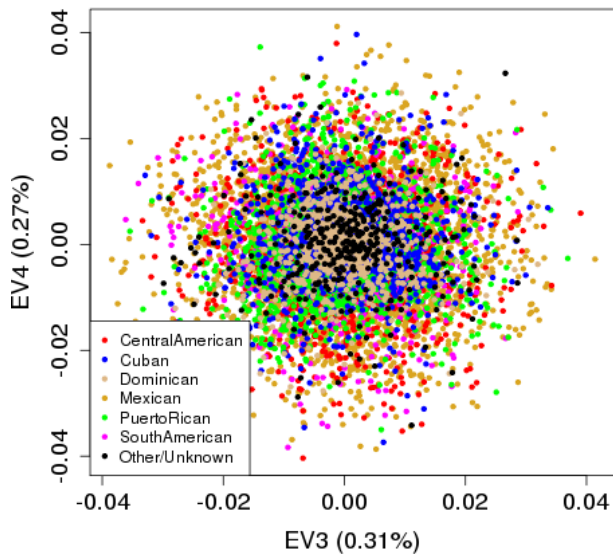


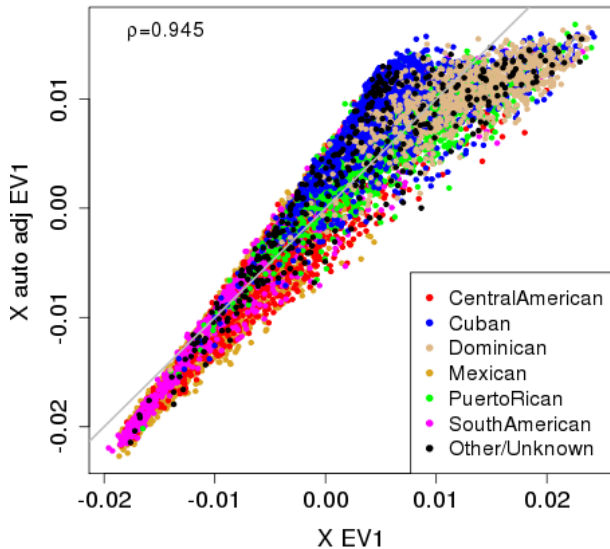
- ▶ We calculate PCs using a pruned set of autosomal SNPs.
- ▶ We then estimate Φ_X using a pruned set of X chromosome SNPs, correcting for PCs 1-5 on the autosomes {model 7}.
- ▶ Finally, we take the eigendecomposition of Φ_X , after subsetting to unrelated samples.

The eigendecomposition should yield any structure on the X chromosome that wasn't already accounted for in the correction for autosomal PCs 1-5.

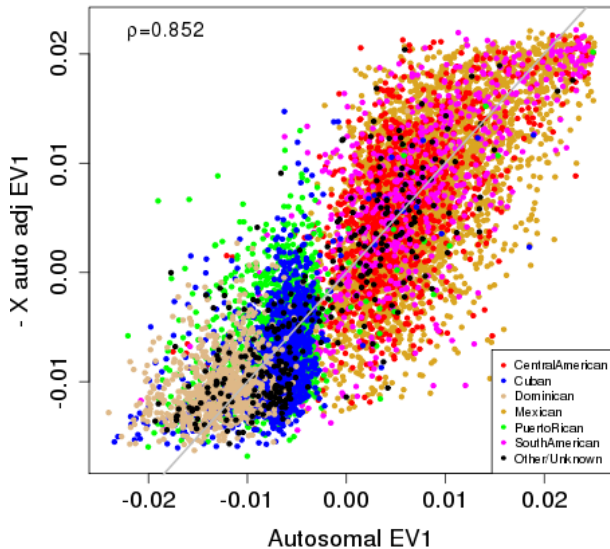








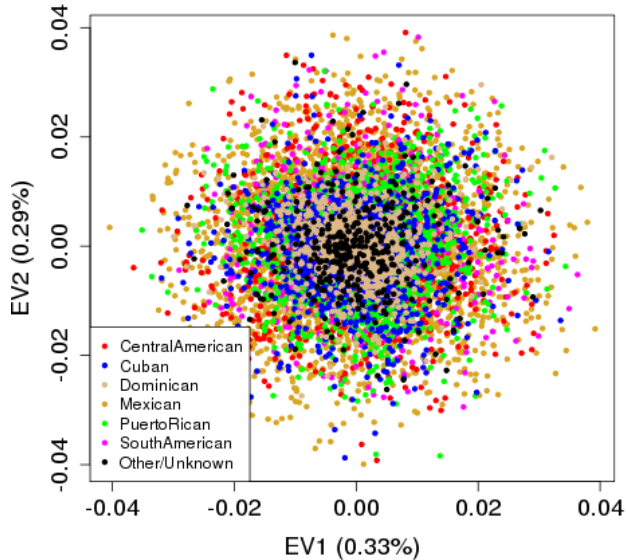
Recall $\text{cor}(X \text{ EV1, auto EV1}) = 0.870$

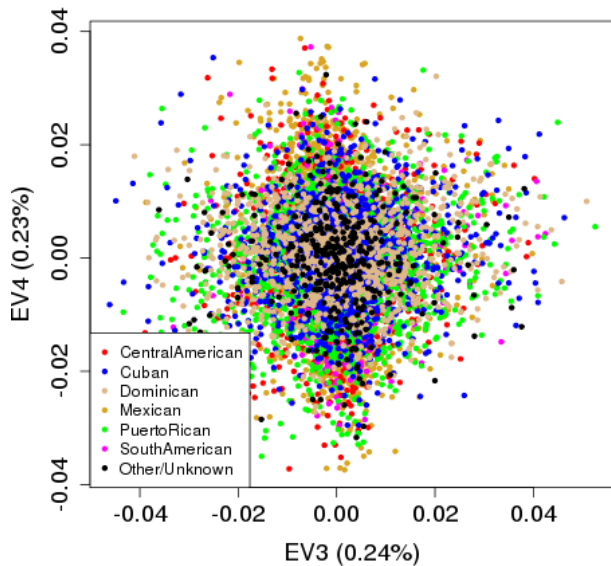


- ▶ We calculate PCs using a pruned set of X chromosome SNPs.
- ▶ We then estimate Φ_X using a pruned set of X chromosome SNPs, correcting for PCs 1-2 on the X chromosome {model 8}¹.
- ▶ Finally, we take the eigendecomposition of Φ_X , after subsetting to unrelated samples.

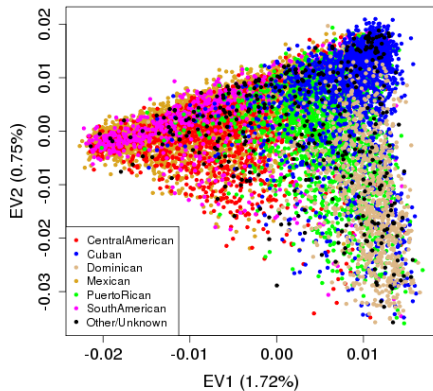
The eigendecomposition should yield any structure on the X chromosome that wasn't already accounted for in the correction for X chromosome PCs 1-2.

¹unrelated kinship threshold of 0.2





eigen(Φ_X) corr for Auto PC 1-5



eigen(Φ_X) corr for X Chr PC 1-2

