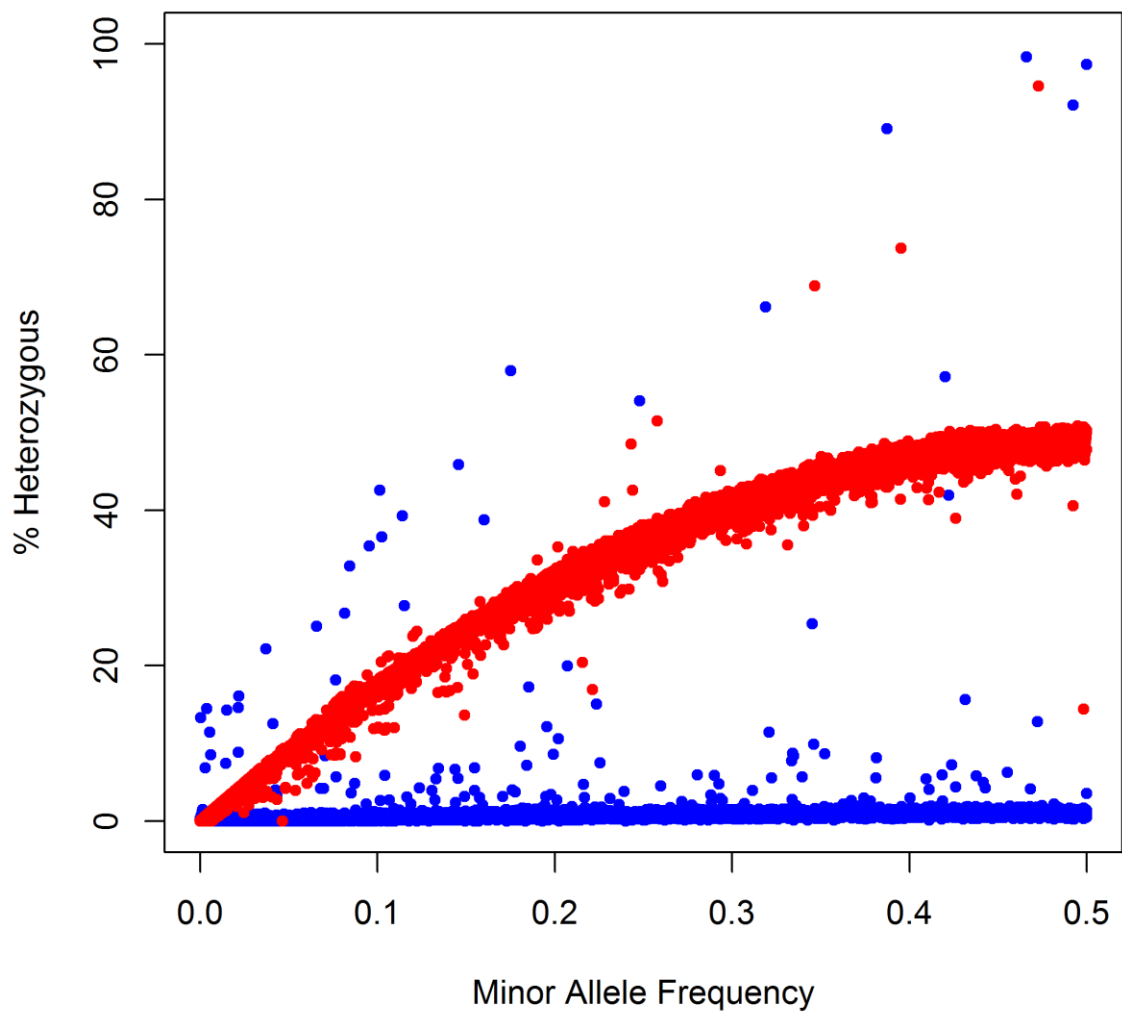


Supplementary Figure S1. Chromosome Y Missingness versus Chromosome X Heterozygosity.

The fraction of Y chromosome markers with missing calls and fraction of X chromosome markers with heterozygous calls were plotted for 889 males (blue) and 5,361 females (red). Each point represents an individual.



Supplementary Figure S2. Heterozygous Prevalence versus Minor Allele Frequency.

Across 889 males (blue) and 5,361 females (red), the percentage of heterozygous genotypes and minor allele frequency of 15,258 X chromosome markers were plotted. Each marker is represented by 2 points, one for males and one for females.

Table 1. Comparison of sex inference accuracy using PLINK versus seXY.

	PLINK		seXY	
	Male	Female	Male	Female
50/50 CV round 1	99.8	96.8	100.0	99.8
50/50 CV round 2	99.8	96.4	100.0	99.9
80/20 CV round 1	99.4	96.6	100.0	99.8
80/20 CV round 2	100.0	96.4	100.0	99.7
80/20 CV round 3	100.0	96.7	100.0	99.9
80/20 CV round 4	99.4	96.8	100.0	99.9
80/20 CV round 5	100.0	96.7	100.0	99.6

Accuracies are displayed as percent of test set individuals whose sexes were correctly predicted. CV, cross-validation.

Supplementary Table S1. Comparison of sex inference accuracy using PLINK versus seXY.

	PLINK		seXY	
	Male	Female	Male	Female
<i>A. Using 100% of genotyped markers</i>				
50/50 CV round 1	99.8	96.8	100.0	99.8
50/50 CV round 2	99.8	96.4	100.0	99.9
80/20 CV round 1	99.4	96.6	100.0	99.8
80/20 CV round 2	100.0	96.4	100.0	99.7
80/20 CV round 3	100.0	96.7	100.0	99.9
80/20 CV round 4	99.4	96.8	100.0	99.9
80/20 CV round 5	100.0	96.7	100.0	99.6
<i>B. Using randomly selected 90% of genotyped markers</i>				
50/50 CV round 1	99.8	96.9	100.0	99.8
50/50 CV round 2	99.8	96.4	100.0	99.9
<i>C. Using randomly selected 75% of genotyped markers</i>				
50/50 CV round 1	100.0	96.2	100.0	99.9
50/50 CV round 2	99.8	96.7	100.0	99.8
<i>D. Using randomly selected 50% of genotyped markers</i>				
50/50 CV round 1	99.8	97.0	100.0	99.8
50/50 CV round 2	100.0	96.6	100.0	99.9

Accuracies are displayed as percent of test set individuals whose sexes were correctly predicted. CV, cross-validation.