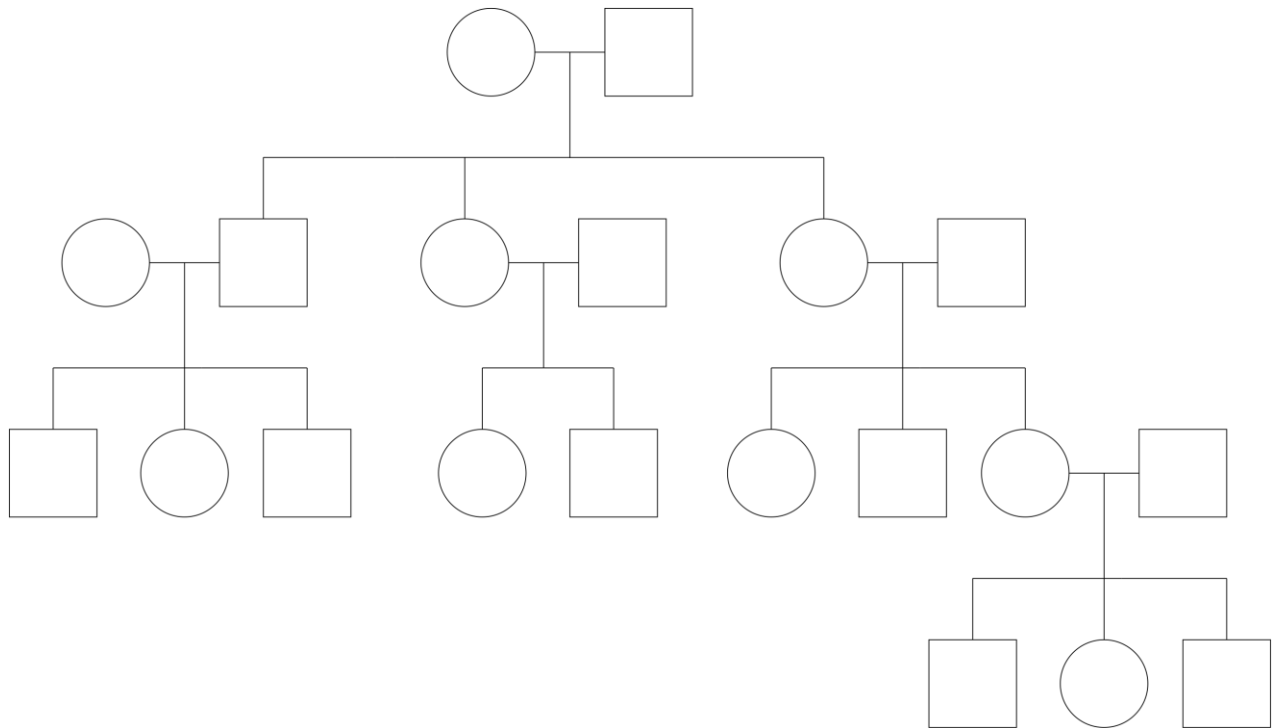


## Supplemental Data

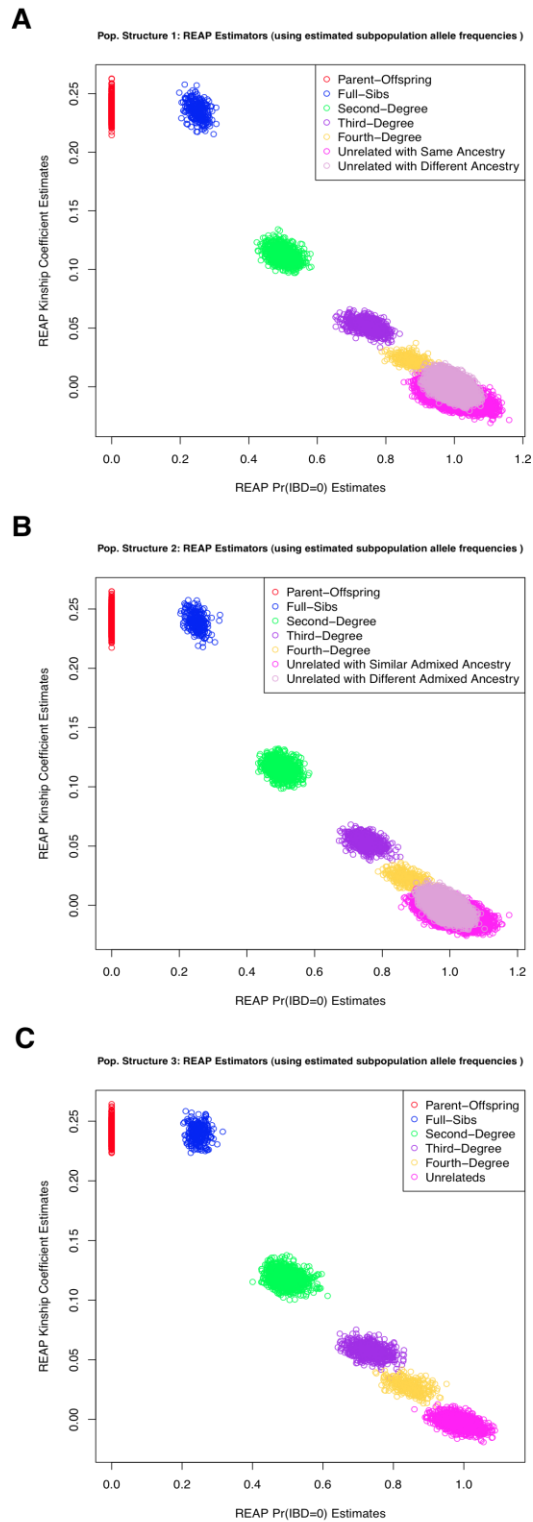
### Estimating Kinship in Admixed Populations

Timothy Thornton, Hua Tang, Thomas J. Hoffmann, Heather M. Ochs-Balcom, Bette J. Caan, and Neil Risch



**Figure S1. Extended Pedigree Configuration for the Simulation Studies**

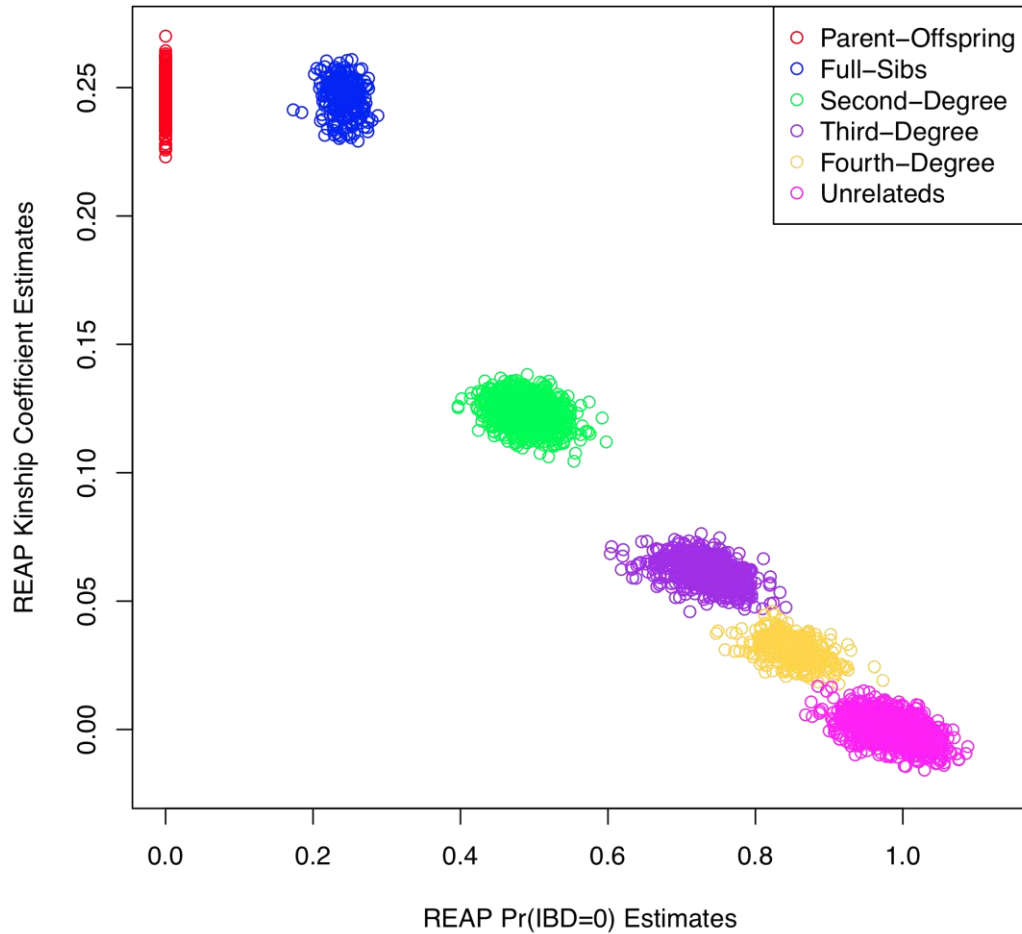
The pedigree configuration for each of the 40 outbred, four-generation pedigrees included in the simulation studies, where the overall structure of each pedigree is as depicted, but the pattern of ancestry admixture varies according to the specified population structure setting.



**Figure S2. REAP Kinship Coefficients versus Zero IBD Sharing Probabilities with Estimated Allele Frequencies**

(A), (B), and (C) are REAP estimated kinship coefficients plotted against zero IBD sharing probability estimates for the population structure settings 1, 2, and 3, respectively, using subpopulation allele frequencies that were estimated using the *frappe* software program. Kinship coefficients and zero IBD sharing probabilities were calculated using the REAP estimators of equations (3) and (4), respectively, with 10,000 simulated random SNPs.

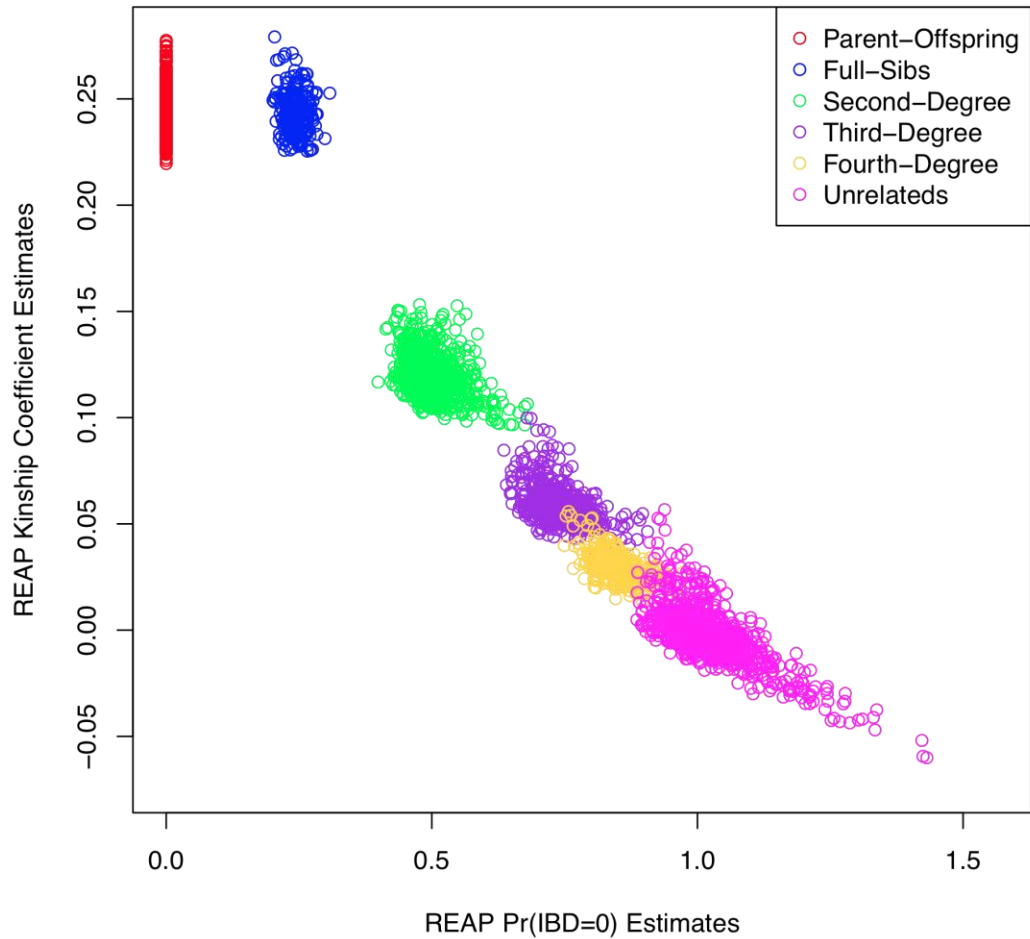
### Ancestry Misspecification Setting 1: REAP Estimators



**Figure S3. REAP Kinship Coefficients versus Zero IBD Sharing Probabilities with Misspecification of the Ancestral Populations**

REAP estimated kinship coefficients plotted against zero IBD sharing probability estimates for misspecification setting 1, which corresponds to population structure setting 3, where the sample individuals in this population structure setting have admixture from three subpopulations, but with the number of ancestral populations incorrectly set to  $K=4$  in the *frappe* software analysis for individual ancestry estimation. Kinship coefficients and zero IBD sharing probabilities were calculated using the REAP estimators of equations (3) and (4), respectively, with the estimated ancestry proportions and subpopulation allele frequencies from the *frappe* software program, for 10,000 random SNPs.

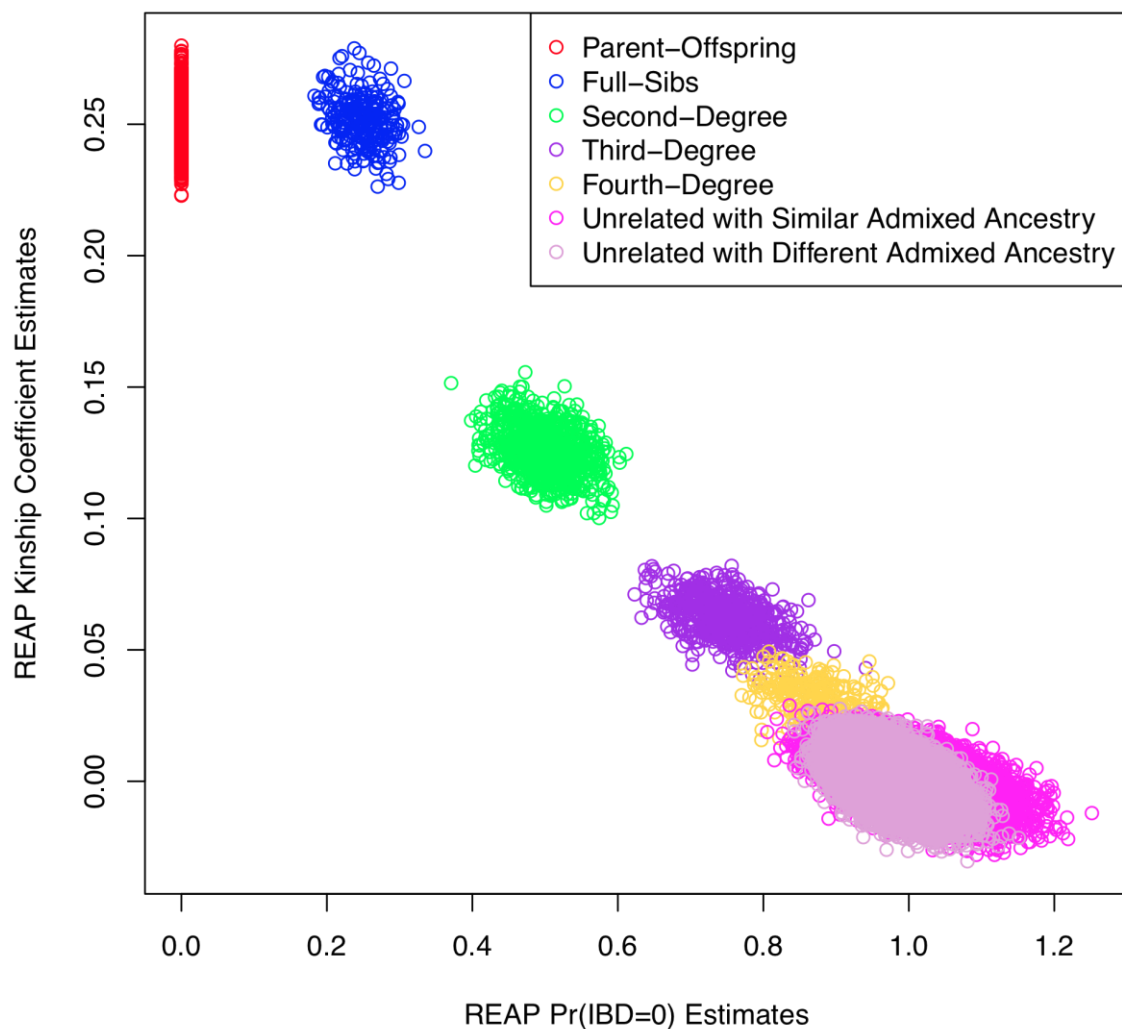
### Ancestry Misspecification Setting 2: REAP Estimators



**Figure S4. REAP Kinship Coefficients versus Zero IBD Sharing Probabilities with Misspecification of the Ancestral Populations**

REAP estimated kinship coefficients plotted against zero IBD sharing probability estimates for misspecification setting 1, which corresponds to population structure setting 3, where the sample individuals in this population structure setting have admixture from three subpopulations, but with the number of ancestral populations incorrectly set to  $K=2$  in the *frappe* software analysis for individual ancestry estimation. Kinship coefficients and zero IBD sharing probabilities were calculated using the REAP estimators of equations (3) and (4), respectively, with the estimated ancestry proportions and subpopulation allele frequencies from the *frappe* software program, for 10,000 random SNPs.

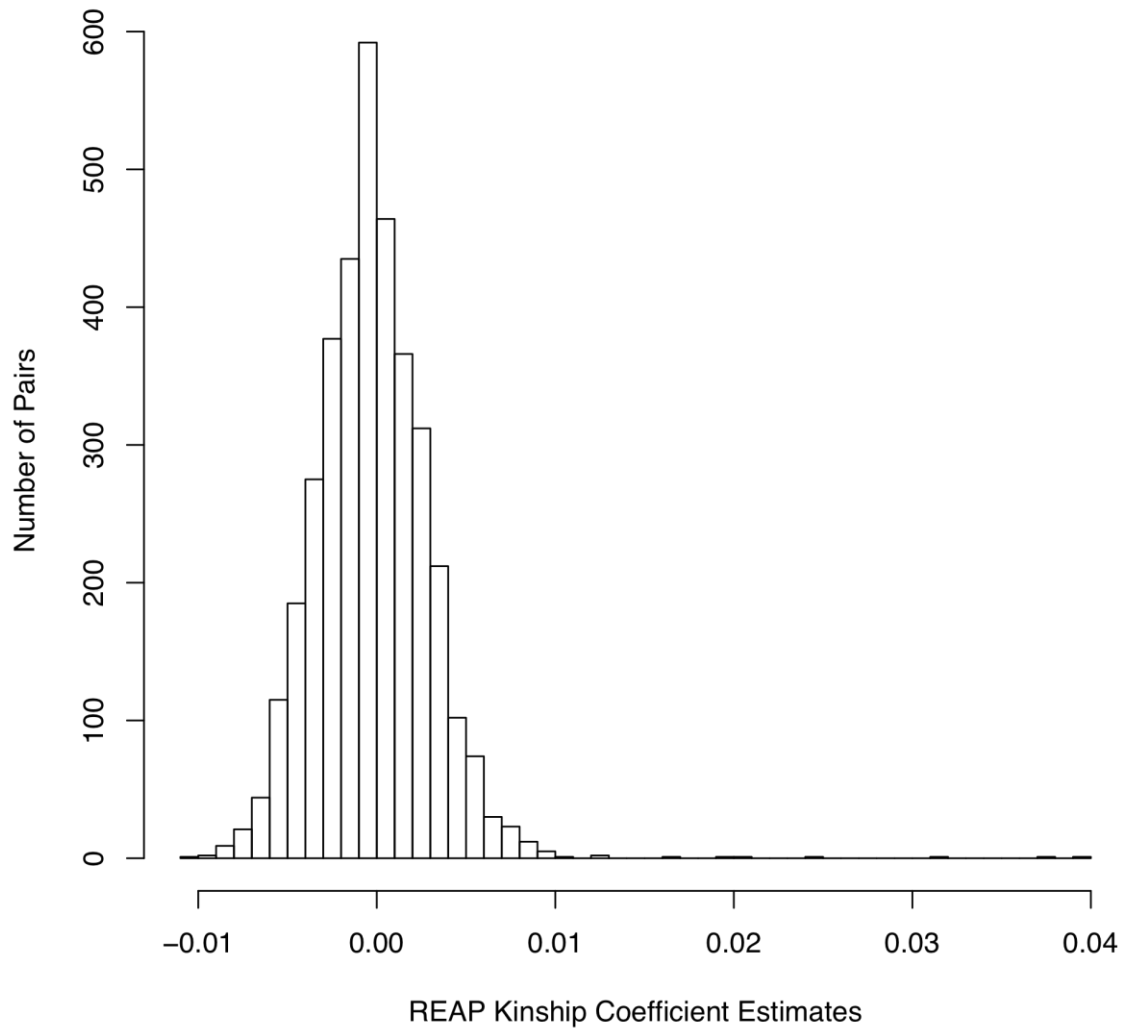
### Pop. Structure 2: REAP Estimators (5000 SNPs)



**Figure S5. REAP Kinship Coefficients versus Zero IBD Sharing Probabilities with 5000 SNPs**

REAP kinship coefficient estimates plotted against REAP zero IBD sharing probability estimates for population structure setting 2. REAP estimates were calculated using the kinship coefficient and zero IBD sharing probability estimators of equations (3) and (4), respectively, with 5,000 simulated random SNPs.

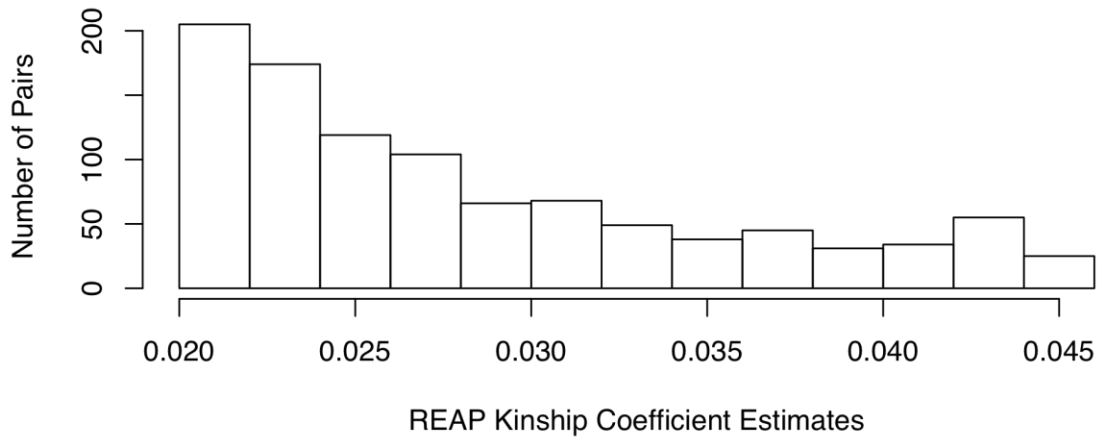
### HapMap MXL: Kinship Coefficients less than .045



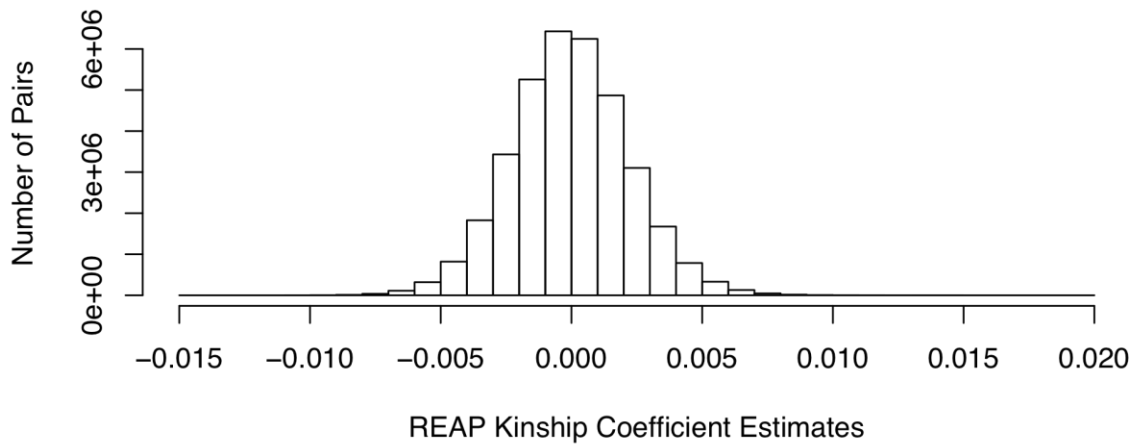
**Figure S6. Histogram of REAP Estimated Kinship Coefficients that Are Less than .045 for HapMap MXL**

The REAP kinship coefficient estimator of equation (3) was used to estimate kinship coefficients for all pairs of individuals in HapMap MXL. The histogram given is for REAP kinship coefficient values that are less than 0.045. All REAP kinship coefficients were calculated using 150,872 SNPs from the autosomal chromosomes.

### WHI-SHARe African Americans: Kinship Coefficients from .02 to .045



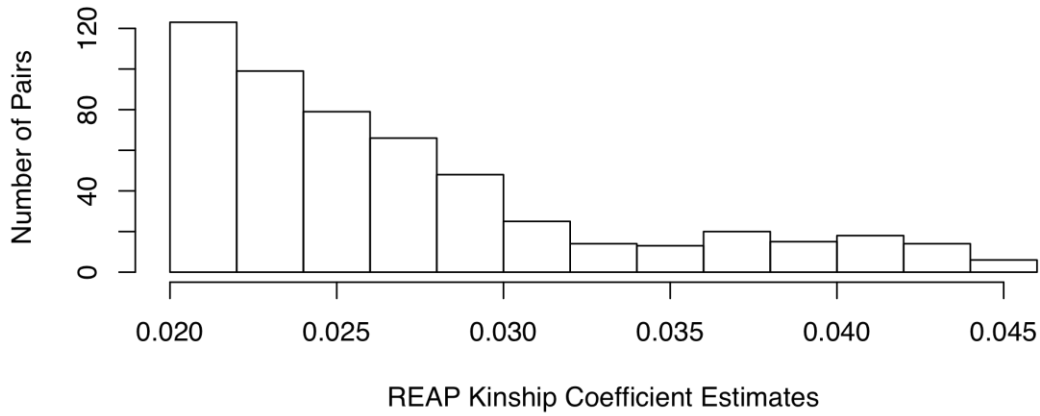
### WHI-SHARe African Americans: Kinship Coefficients less than .02



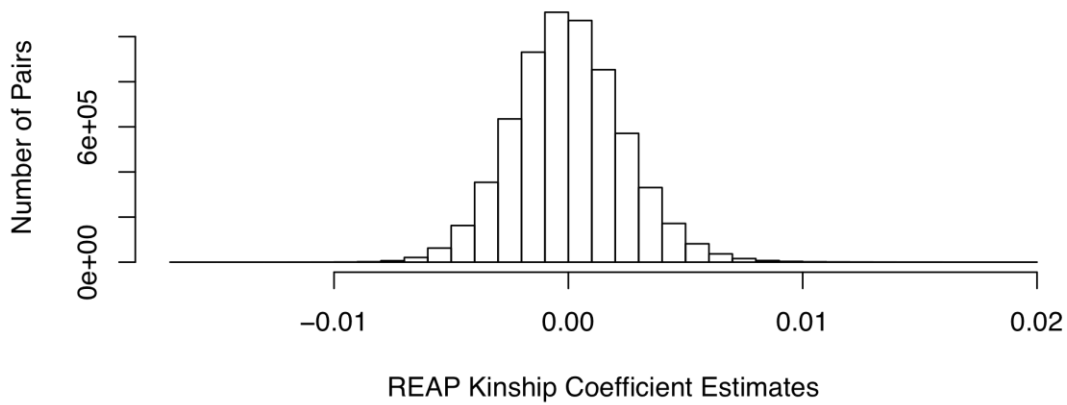
**Figure S7. Histograms of REAP Estimated Kinship Coefficients that Are Less than .045 for WHI-SHARe African Americans**

The REAP kinship coefficient estimator of equation (3) was used to estimate kinship coefficients for all pairs of self-reported African Americans in the WHI-SHARe study. Histograms are given for REAP kinship coefficient values between 0.045 and 0.025 and for REAP kinship coefficient values that are less than 0.025. All REAP kinship coefficients were calculated using 656,852 SNPs from the autosomal chromosomes.

### WHI-SHARe Hispanics: Kinship Coefficients from .02 to .045



### WHI-SHARe Hispanics: Kinship Coefficients less than .02



**Figure S8. Histograms of REAP Estimated Kinship Coefficients that Are Less than .045 for WHI-SHARe Hispanics**

The REAP kinship coefficient estimator of equation (3) was used to estimate kinship coefficients for all pairs of self-reported Hispanics in the WHI-SHARe study. Histograms are given for REAP kinship coefficient values between 0.045 and 0.025 and for REAP kinship coefficient values that are less than 0.025. All REAP kinship coefficients were calculated using 656,852 SNPs from the autosomal chromosomes.



**Table S1. REAP Kinship Coefficient Estimates for Misspecification Setting 2 and with Ancestry Correctly Specified**

<b>Outbred Relationship</b>	<b><math>\phi</math>(Kinship Coefficient)</b>		
	<b>Theoretical Kinship Coefficient</b>	<b>REAP Estimate (SD): Ancestral Populations Correctly Specified</b>	<b>Reap Estimate (SD): Misspecification Setting 2</b>
Parent-offspring	0.25	0.247 (0.007)	0.245 (0.010)
Full siblings	0.25	0.245 (0.007)	0.244 (0.010)
Second-degree	0.125	0.124 (0.006)	0.120 (0.010)
Third-degree	0.0625	0.062 (0.005)	0.059 (0.009)
Fourth-degree	0.03125	0.032 (0.005)	0.030 (0.007)
Unrelated	0	0.00002 (0.005)	-0.003 (0.009)

REAP kinship coefficients were estimated for population structure 3, where all sample individuals have admixture from three subpopulations. In the table, REAP with the ancestral populations correctly specified and REAP under misspecification setting 2 correspond to the REAP estimator of equation (3) calculated using individual ancestry estimates from the *frappe* software program with the number of ancestral populations correctly set to three and with the number of ancestral populations incorrectly set to two, respectively. The REAP kinship coefficients were estimated using 10,000 simulated random SNPs.

**Table S2. REAP-Inferred Relative Pairs in Release 3 of Phase III of the HapMap MXL Population Sample**

First Individual			Second Individual			REAP-Inferred Relationship	HapMap Reported Relatives
Family Number	Identification Number	Sex	Family Number	Identification Number	Sex		
M001	NA19649	M	M001	NA19650	M	parent-offspring	yes
M002	NA19669	F	M002	NA19671	F	parent-offspring	yes
M002	NA19671	F	M002	NA19670	M	parent-offspring	yes
M004	NA19675	F	M004	NA19677	F	parent-offspring	yes
M004	NA19676	M	M004	NA19677	F	parent-offspring	yes
M004	NA19675	F	M009	NA19678	F	parent-offspring	no
M004	NA19675	F	M009	NA19680	F	full-siblings	no
M004	NA19675	F	M009	NA19679	M	parent-offspring	no
M004	NA19677	F	M009	NA19678	F	second-degree	no
M004	NA19677	F	M009	NA19680	F	second-degree	no
M004	NA19677	F	M009	NA19679	M	second-degree	no
M005	NA19651	F	M005	NA19653	F	parent-offspring	yes
M005	NA19653	F	M005	NA19652	M	parent-offspring	yes
M006	NA19656	F	M006	NA19654	F	parent-offspring	yes
M007	NA19657	F	M007	NA19659	F	parent-offspring	yes
M007	NA19658	M	M007	NA19659	F	parent-offspring	yes
M007	NA19657	F	M032	NA19785	F	fourth-degree	no
M007	NA19657	F	M032	NA19787	M	complex <sup>a</sup>	no
M007	NA19657	F	M032	NA19786	M	third-degree	no
M007	NA19659	F	M032	NA19787	M	complex <sup>b</sup>	no
M007	NA19659	F	M032	NA19786	M	third-degree	no
M008	NA19660	F	2382	NA19672	F	full-siblings	no
M008	NA19662	F	2382	NA19672	F	second-degree	no
M008	NA19661	M	M008	NA19662	F	parent-offspring	yes
M008	NA19660	F	M008	NA19662	F	parent-offspring	yes
M008	NA19661	M	M011	NA19685	M	parent-offspring	no
M008	NA19660	F	M011	NA19685	M	parent-offspring	no
M008	NA19662	F	M011	NA19685	M	full-siblings	no
M009	NA19678	F	M009	NA19680	F	parent-offspring	yes
M009	NA19680	F	M009	NA19679	M	parent-offspring	yes
M010	NA19683	F	M010	NA19682	M	parent-offspring	yes
M010	NA19683	F	M010	NA19681	F	parent-offspring	yes
M011	NA19686	F	2382	NA19672	F	third-degree	no
M011	NA19685	M	2382	NA19672	F	second-degree	no
M011	NA19686	F	M008	NA19661	M	second-degree	no
M011	NA19686	F	M008	NA19660	F	second-degree	no
M011	NA19686	F	M008	NA19662	F	second-degree	no
M011	NA19686	F	M011	NA19684	F	parent-offspring	yes
M011	NA19686	F	M011	NA19685	M	parent-offspring	yes
M012	NA19664	M	2382	NA19672	F	second-degree	no
M012	NA19665	F	2382	NA19672	F	third-degree	no
M012	NA19664	M	M008	NA19660	F	second-degree	no
M012	NA19664	M	M008	NA19662	F	third-degree	no
M012	NA19665	F	M008	NA19660	F	third-degree <sup>c</sup>	no
M012	NA19665	F	M008	NA19662	F	fourth-degree	no
M012	NA19664	M	M011	NA19686	F	fourth-degree	no
M012	NA19664	M	M011	NA19685	M	third-degree	no
M012	NA19665	F	M011	NA19685	M	fourth-degree	no
M012	NA19663	F	M012	NA19665	F	parent-offspring	yes
M012	NA19664	M	M012	NA19665	F	parent-offspring	yes

**Table S2. Continued**

First Individual			Second Individual			REAP-Inferred Relationship	HapMap Reported Relatives
Family Number	Identification Number	Sex	Family Number	Identification Number	Sex		
M014	NA19718	F	M014	NA19716	F	parent-offspring	yes
M015	NA19719	F	M015	NA19721	F	parent-offspring	yes
M015	NA19720	M	M015	NA19721	F	parent-offspring	yes
M016	NA19722	F	M016	NA19724	M	parent-offspring	yes
M016	NA19723	M	M016	NA19724	M	parent-offspring	yes
M017	NA19726	M	M017	NA19727	M	parent-offspring	yes
M017	NA19725	F	M017	NA19727	M	parent-offspring	yes
M023	NA19747	M	M023	NA19748	F	parent-offspring	yes
M023	NA19746	F	M023	NA19748	F	parent-offspring	yes
M024	NA19749	F	M024	NA19751	M	parent-offspring	yes
M024	NA19751	M	M024	NA19750	M	parent-offspring	yes
M026	NA19755	F	M026	NA19757	M	parent-offspring	yes
M026	NA19756	M	M026	NA19757	M	parent-offspring	yes
M027	NA19759	M	M027	NA19760	F	parent-offspring	yes
M028	NA19761	F	M028	NA19763	F	parent-offspring	yes
M028	NA19762	M	M028	NA19763	F	parent-offspring	yes
M031	NA19772	M	M031	NA19771	M	parent-offspring	yes
M031	NA19772	M	M031	NA19770	F	parent-offspring	yes
M032	NA19785	F	M032	NA19787	M	parent-offspring	yes
M032	NA19787	M	M032	NA19786	M	parent-offspring	yes
M033	NA19773	F	M033	NA19775	F	parent-offspring	yes
M033	NA19774	M	M033	NA19775	F	parent-offspring	yes
M034	NA19776	F	M034	NA19778	M	parent-offspring	yes
M034	NA19777	M	M034	NA19778	M	parent-offspring	yes
M035	NA19780	M	M035	NA19781	F	parent-offspring	yes
M035	NA19779	F	M035	NA19781	F	parent-offspring	yes
M036	NA19783	M	M036	NA19784	M	parent-offspring	yes
M036	NA19784	M	M036	NA19782	F	parent-offspring	yes
M037	NA19788	F	M037	NA19790	F	parent-offspring	yes
M037	NA19790	F	M037	NA19789	M	parent-offspring	yes
M039	NA19796	M	M039	NA19794	F	parent-offspring	yes
M039	NA19796	M	M039	NA19795	M	parent-offspring	yes

<sup>a</sup> NA19657 is related to NA19787 through NA19787's maternal and paternal lineage based on constructed pedigree(s), so the relationships is labeled as complex.

<sup>b</sup> NA19659 is related to NA19787 through NA19787's maternal and paternal lineage based on constructed pedigree(s), so the relationship is labeled as complex in the table.

<sup>c</sup> The relationship for NA19660 and NA19665 given in the table was changed from the REAP inferred fourth-degree to third-degree based on constructed pedigree(s).