Supplementary Material for

Assessment of kinship detection using RNA-seq data

Authors

Natalia Blay[1,2,3], Eduard Casas[1,2,4], Iván Galván-Femenía[1,5], Jan Graffelman[6,7], Rafael de Cid[1,5], Tanya Vavouri[1,2]

Affiliations

1. Program for Predictive and Personalized Medicine of Cancer, Germans Trias i Pujol Research Institute (PMPPC-IGTP), Badalona, Spain
2. Josep Carreras Leukaemia Research Institute (IJC), Campus ICO-Germans Trias i Pujol, Universitat Autònoma de Barcelona, Badalona, Spain
3. Masters Programme in Bioinformatics and Biostatistics, Universitat Oberta de Catalunya (UOC), Barcelona, Spain
4. Doctoral Programme in Biomedicine, Universitat de Barcelona, Barcelona, Spain
5. Genomes for Life - GCAT lab Group - Germans Trias i Pujol Research Institute, Can Ruti Campus, Ctra de Can Ruti, Camí de les Escoles s/n, Badalona, Barcelona, Spain.
6. Department of Statistics and Operations Research Universitat Politècnica de Catalunya, Barcelona, Spain.
7. Department of Biostatistics, University of Washington, Seattle, USA.

Corresponding author: Tanya Vavouri tvavouri@carrerasresearch.org

This PDF file includes:

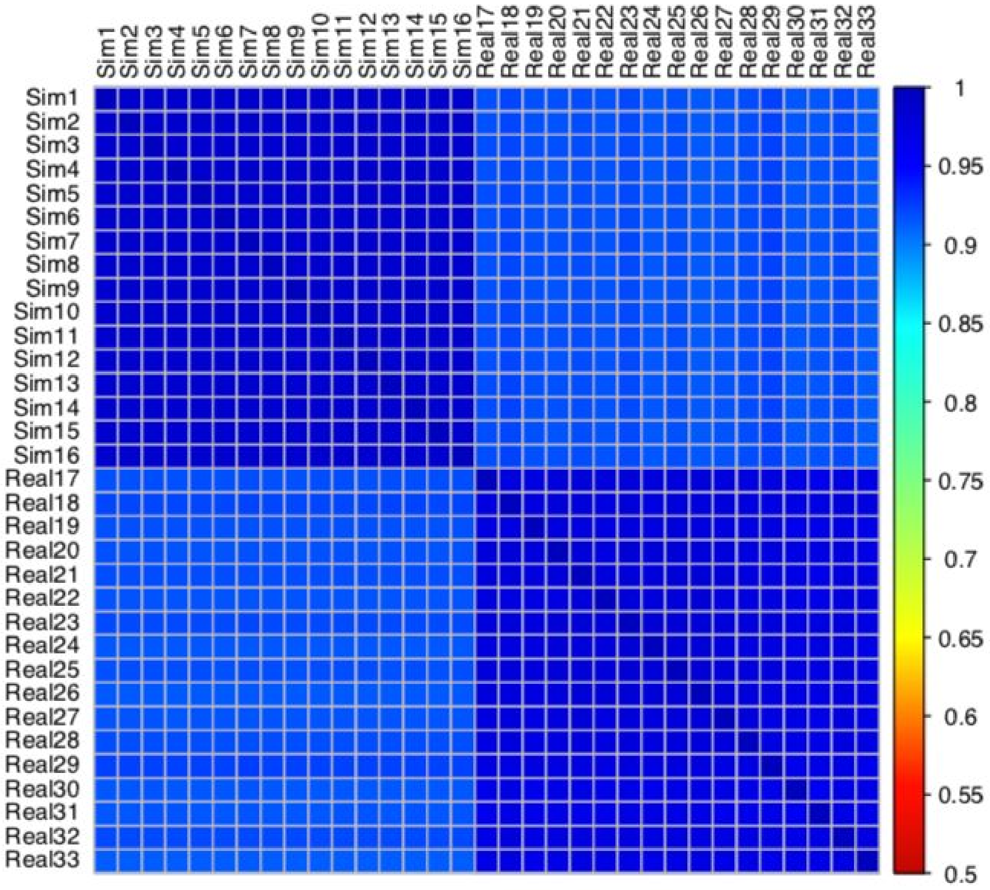
Supplementary Table 1

Supplementary Figure 1

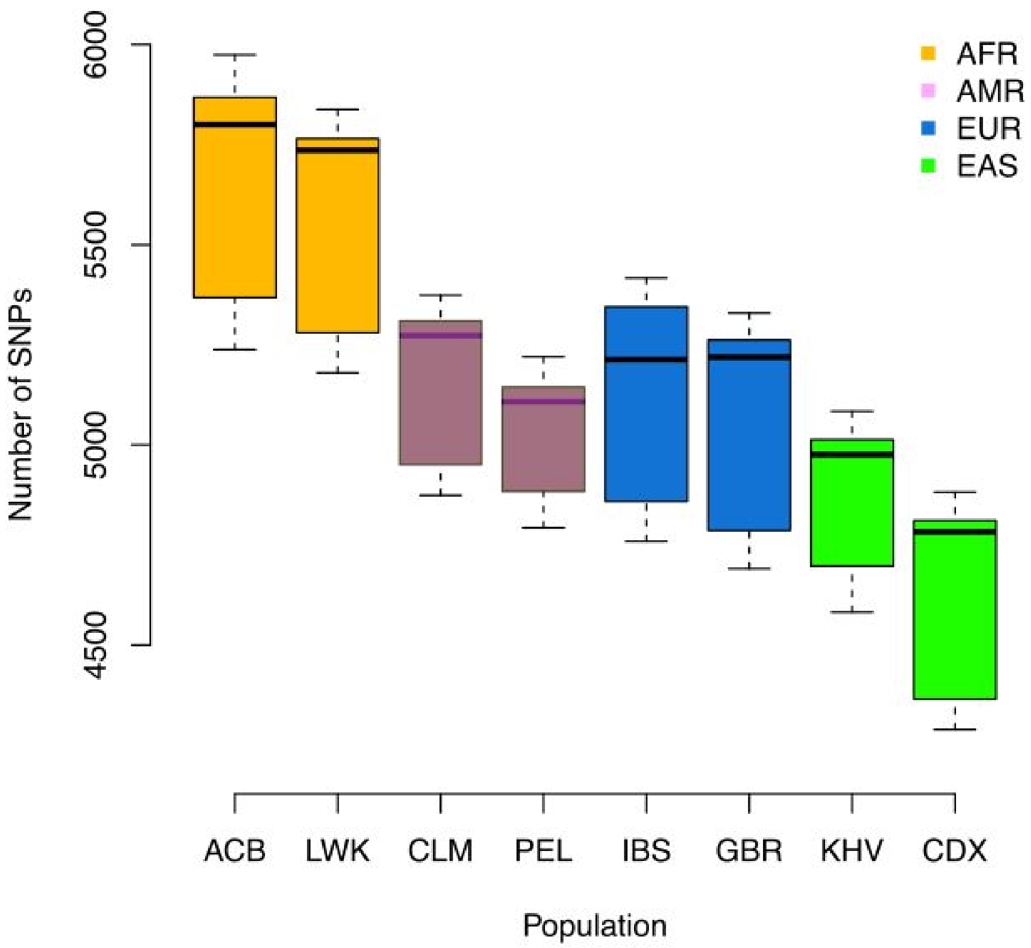
Supplementary Figure 2

Supplementary Table 1. Expected IBD probabilities of different relationships and degrees of relatedness. Z0, Z1 and Z2 represent the probabilities that a pair of individuals share 0, 1 or 2 identical by descent alleles respectively.

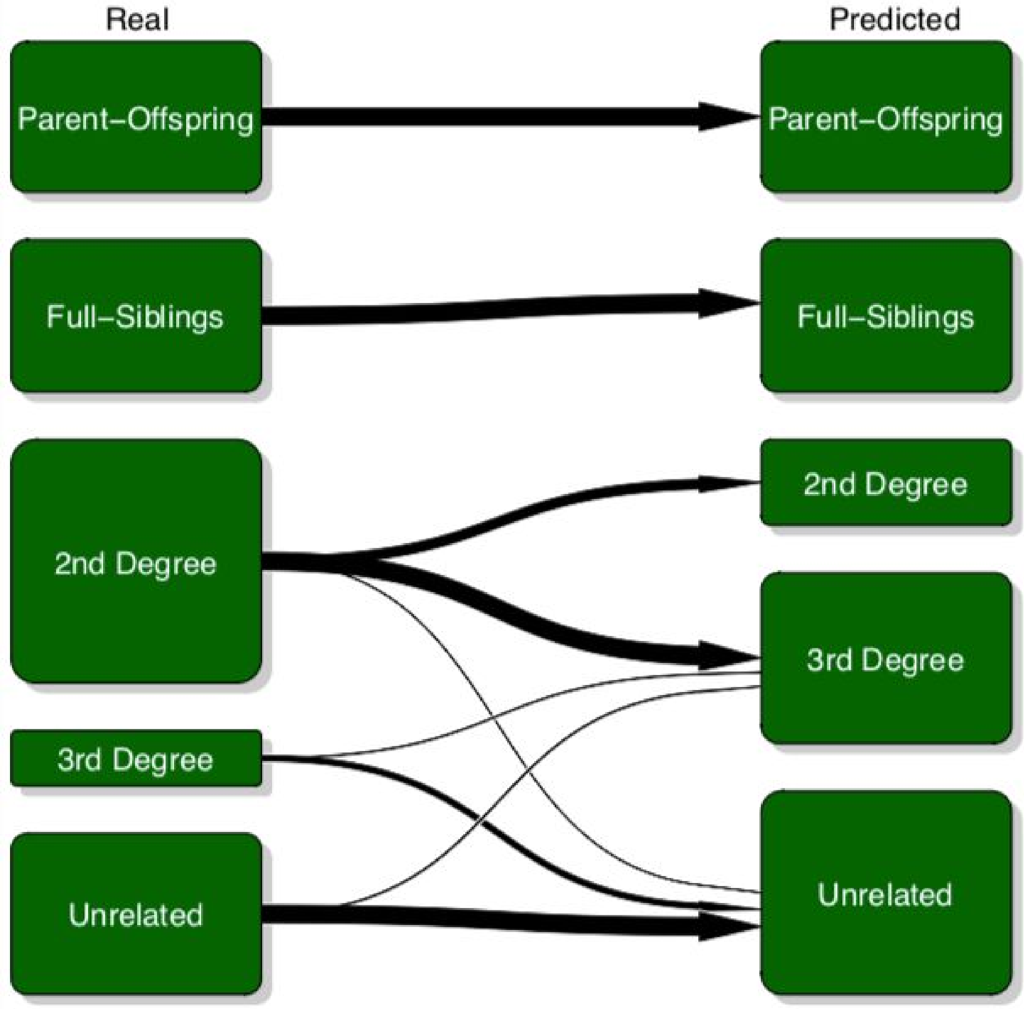
Degree	**Relationship**	**Z0**	**Z1**	**Z2**
1	parent-offspring	0	1	0
	full-siblings	0.25	0.5	0.25
2	grandparental, avuncular, half-siblings	0.5	0.5	0
3	First-cousins, great-grandparental, great-avuncular, half-avuncular	0.75	0.25	0
∞	unrelated	1	0	0



Supplementary Figure 1. Correlation matrix of the number of reads per gene in logarithmic scale between Simulated data (Sim1 - Sim16: individuals 1 - 16 of a type 1 simulated family) and Real data (Real1 - Real17: SRR1258217 - SRR1258233). Mean correlation between Real data: 0.975, between Simulated data: 0.988, between Real data and Simulated data: 0.918.



Supplementary Figure 2: Number of SNPs used for pairwise comparison of the simulated families per population. Each color represents a super population (AFR: African, AMR: Ad Mixed American, EUR: European, EAS: East Asian).



Supplementary Figure 3: Kinship classification by PRIMUS for simulated families (pairs with higher than third degree relationship are represented here as unrelated).