Inferring Relationships from Genetic Data

Selma Chihab

Background Information: Key Terminology

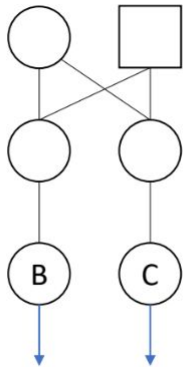
- Identity-by-descent (IBD)
  - Alleles inherited from a common ancestor
- Kinship Coefficient (ψ)
  - Probability of IBD

Pairwise relationship	κ_0	κ_1	κ_2	ψ
Unrelated	1.00	0	0	0
Parent-offspring	0	1.00	0	0.25
Monozygous twin	0	0	1.00	0.50
Full Sib	0.25	0.50	0.25	0.25
Half sib, grandparent, aunt	0.50	0.50	0.00	0.125
First cousin	0.75	0.25	0	0.0625
Double first cousin	0.5625	0.375	0.0625	0.125
Quadruple half first cousin	0.5312	0.4375	0.0312	0.125

Background Information: Path Counting

- Formula: $\psi(B,C) = \sum_A \sum_{\mathcal{P}(A)} (1 + f_A) \left(\frac{1}{2}\right)^{n(\mathcal{P}(A))+1}$
  - A: the common ancestor
  - B, C: the individuals in question
  - f_A : the kinship coefficient between the mother of A and father of A (inbreeding coefficient)
  - *n(P(A))* : number of meiosis
- Example: Cousins

$$\psi(B,C) = \frac{1}{32} + \frac{1}{32} = \frac{1}{16}$$



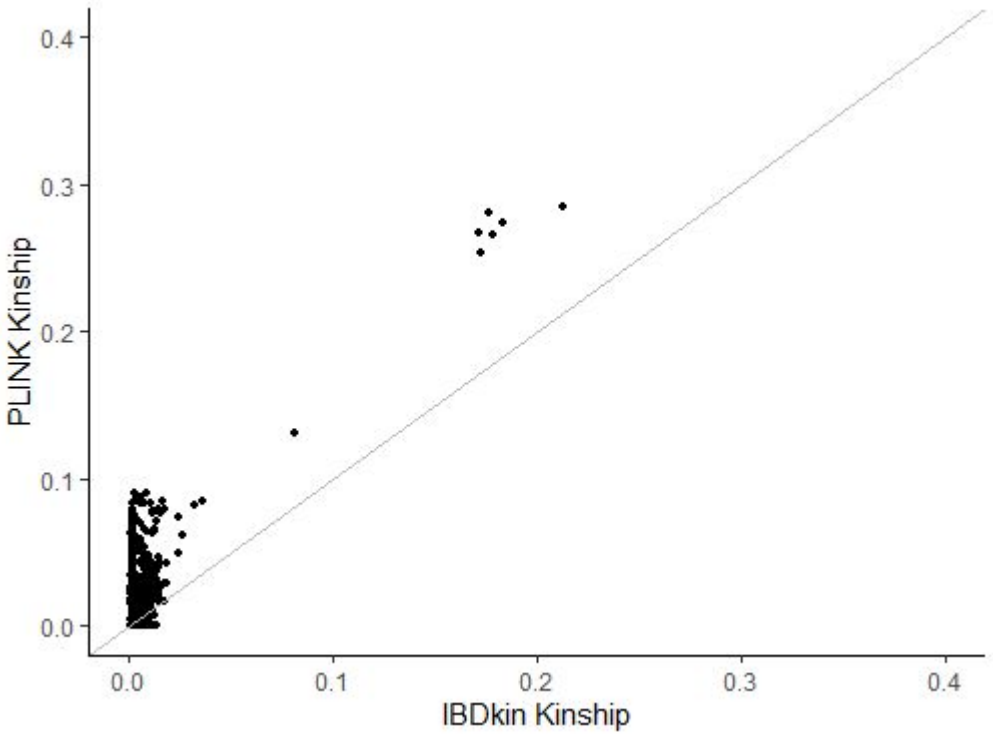
Project Overview: Software

- Compare the kinship estimations of two softwares:
  - PLINK
    - Uses SNPs (single nucleotide polymorphisms)
  - IBDkin
    - Uses IBD segments
- IBDkin more accurate but requires more steps
  - Ramstetter, et al. (2017)

Project Overview: Data

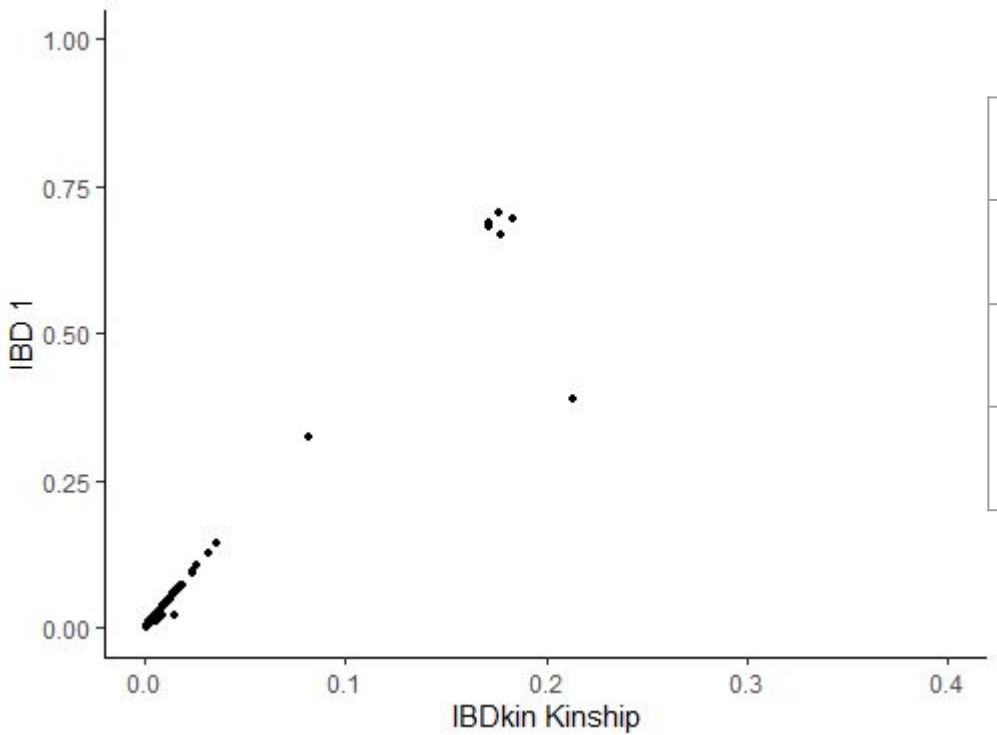
- Sample of 300 individuals
- Made up of 4 different populations
  - Utah Residents with Northern and Western European Ancestry
  - African Ancestry in Southwest USA
  - Puerto Rican in Puerto Rico
  - African Caribbean in Barbados
- Assumption : unrelated

Results: IBDkin Kinship vs. PLINK Kinship Plot



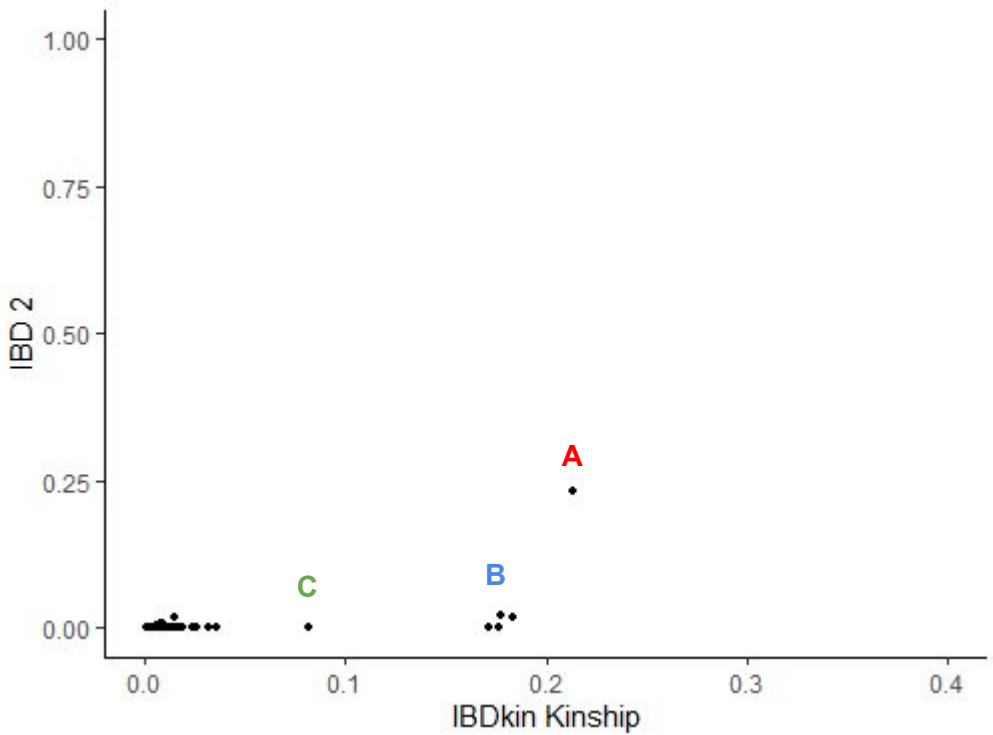
- PLINK : overestimation
- IBDkin : underestimation

Results: Kinship vs. IBD 1 Plot



Relationship	IBD 0	IBD 1	IBD 2	Kinship
Parent/child	0	1	0	0.25
Siblings	0.25	0.5	0.25	0.25
Cousins	0.75	0.25	0	0.0625

Results: Kinship vs. IBD 2 Plot



- **A** : siblings
- **B** : parent/child relationships
- **C** : cousins

Acknowledgements

- Seth Temple, PhD Statistics Student
- Sharon Browning, Biostatistics Researcher and Professor
- Elizabeth Thompson, Emeritus Professor and Author of *Statistical Inference from Genetic Data on Pedigrees*