International Statistical Genetics Workshop • Boulder, Colorado •
IBG

Introduction to Genetic Relatedness

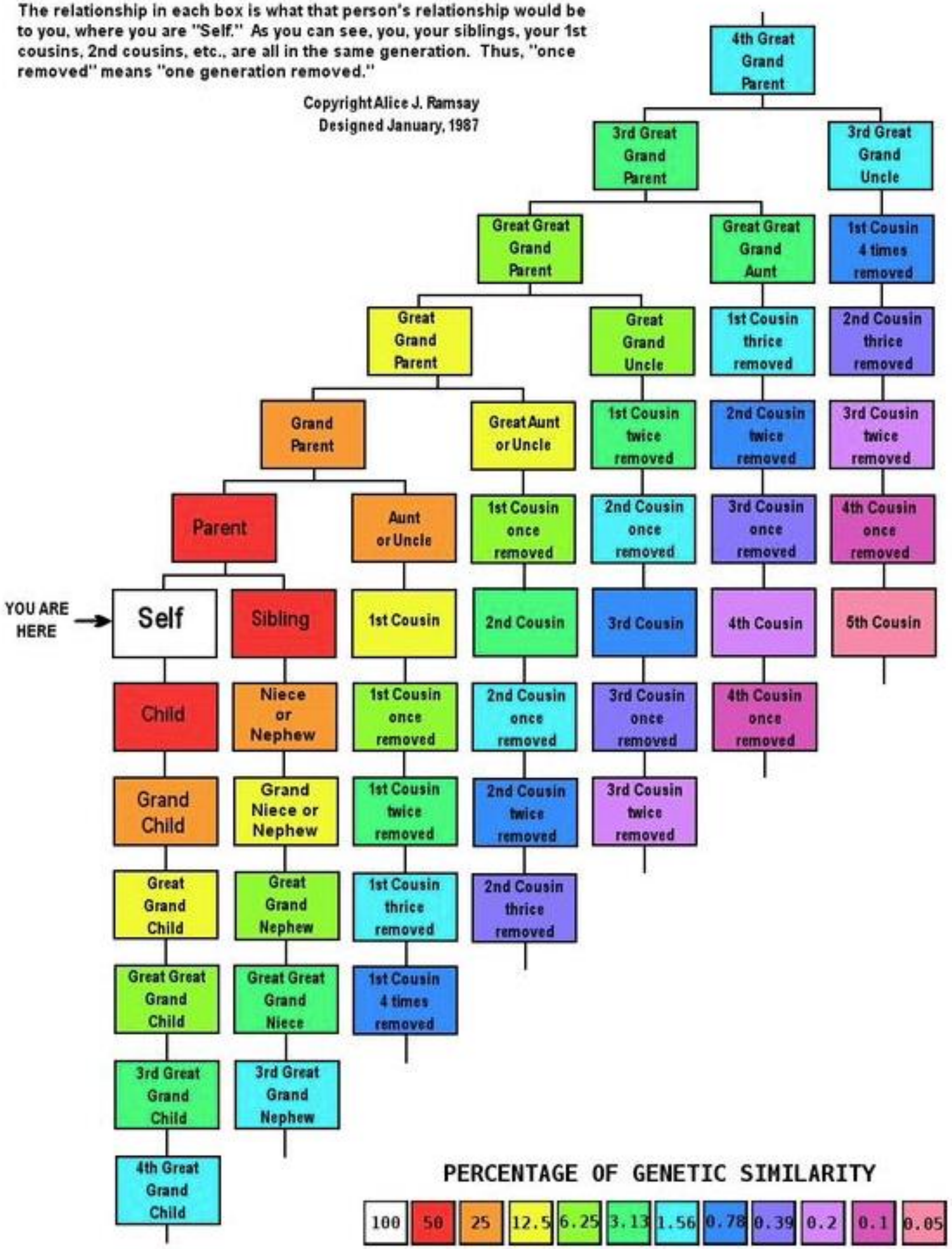
Katrina Grasby

Overview

1. Terms used in describing genetic relatedness

2. Estimating genetic relationships from genotyped data

3. Background genetic relatedness


The relationship in each box is what that person's relationship would be to you, where you are "Self." As you can see, you, your siblings, your 1st cousins, 2nd cousins, etc., are all in the same generation. Thus, "once removed" means "one generation removed."

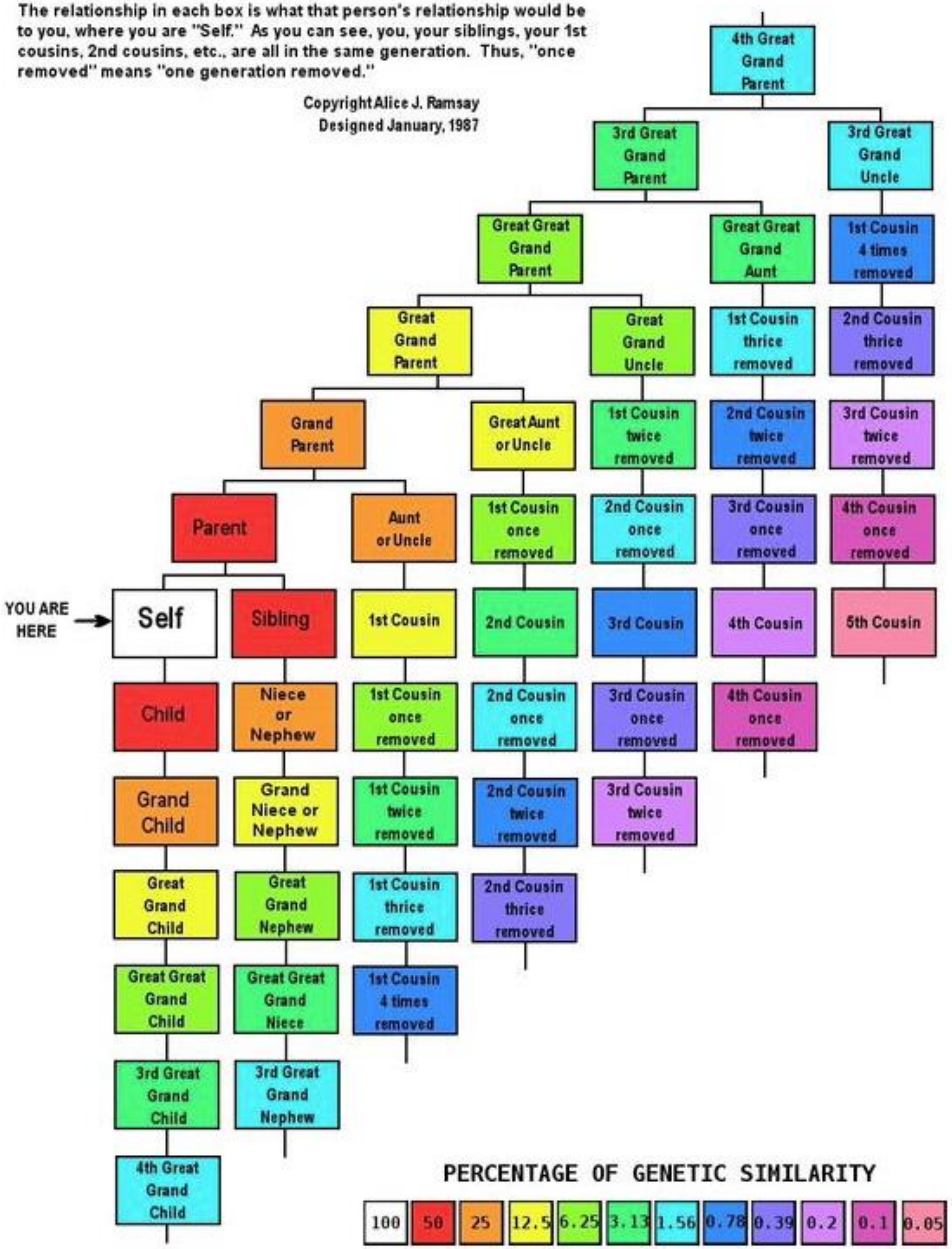
Copyright Alice J. Ramsay
Designed January, 1987

MZ twin here too →

Genetic Relatedness

Historically calculated through pedigree information.

MZ twin → here too



The relationship in each box is what that person's relationship would be to you, where you are "Self." As you can see, you, your siblings, your 1st cousins, 2nd cousins, etc., are all in the same generation. Thus, "once removed" means "one generation removed."

Copyright Alice J. Ramsay
Designed January, 1987

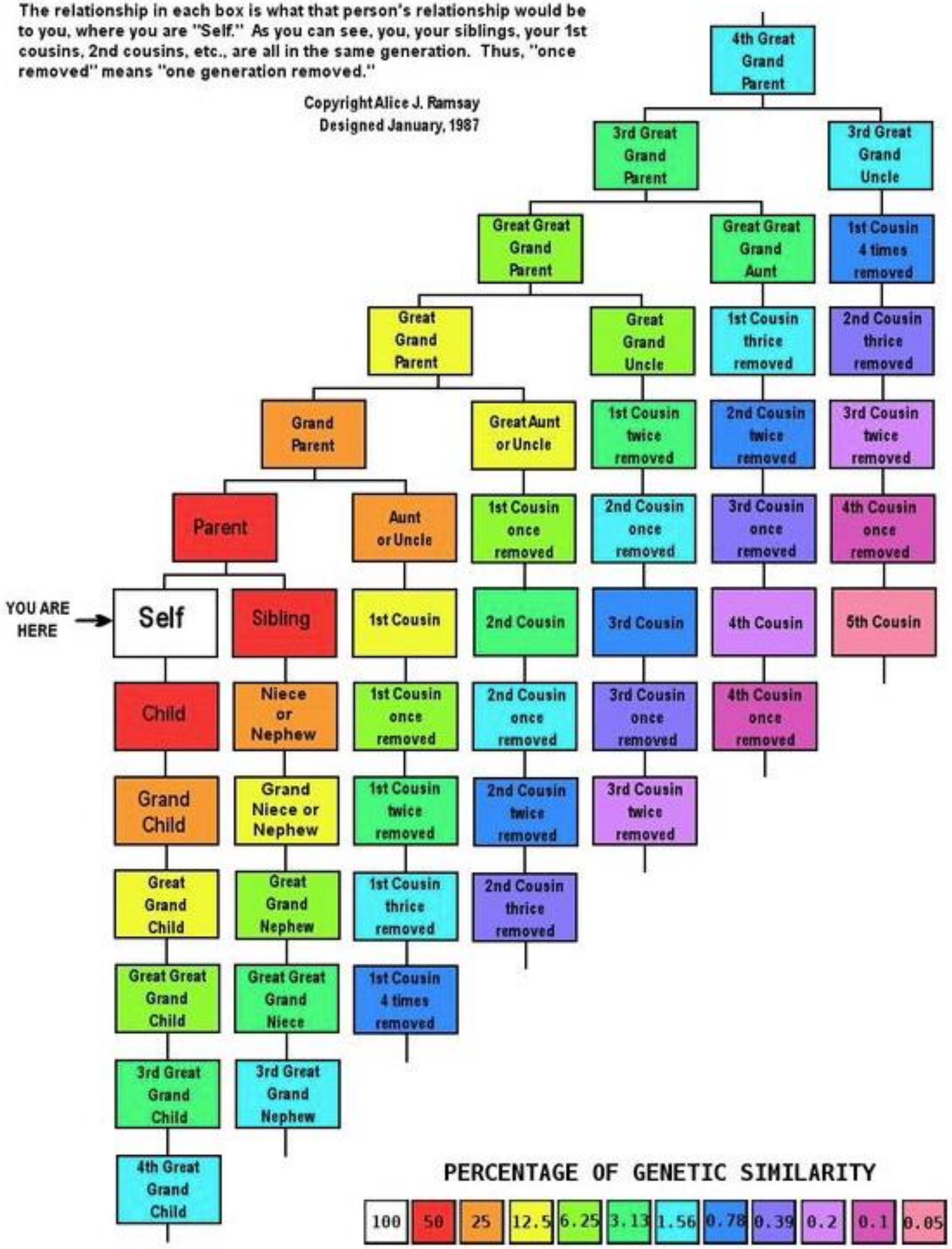
Genetic Relatedness

Historically calculated through pedigree information.

Kinship coefficient

Probability that an allele at a locus for two individuals was inherited from the same genome.

MZ twin here too



The relationship in each box is what that person's relationship would be to you, where you are "Self." As you can see, you, your siblings, your 1st cousins, 2nd cousins, etc., are all in the same generation. Thus, "once removed" means "one generation removed."

Copyright Alice J. Ramsay
Designed January, 1987

Genetic Relatedness

Historically calculated through pedigree information.

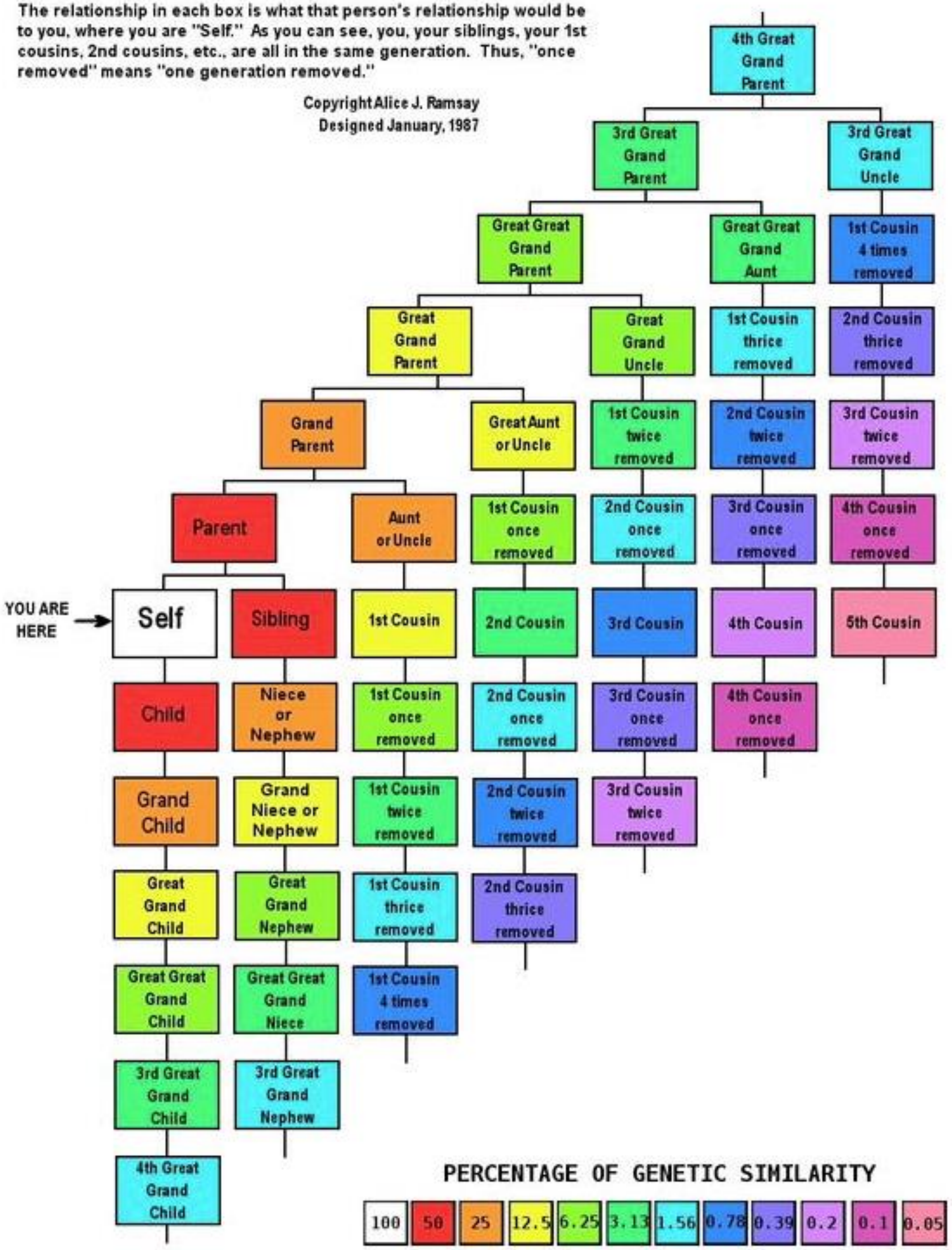
Kinship coefficient

Probability that an allele at a locus for two individuals was inherited from the same genome.

Inbreeding coefficient

Probability that the two alleles at a locus for an individual was inherited from the same genome.

MZ twin
here too


The relationship in each box is what that person's relationship would be to you, where you are "Self." As you can see, you, your siblings, your 1st cousins, 2nd cousins, etc., are all in the same generation. Thus, "once removed" means "one generation removed."

Copyright Alice J. Ramsay
Designed January, 1987

Genetic Relatedness

Historically calculated through pedigree information.

Kinship coefficient

Probability that an allele at a locus for two individuals was inherited from the same genome.

Inbreeding coefficient

Probability that the two alleles at a locus for an individual was inherited from the same genome.

Coefficient of relationship

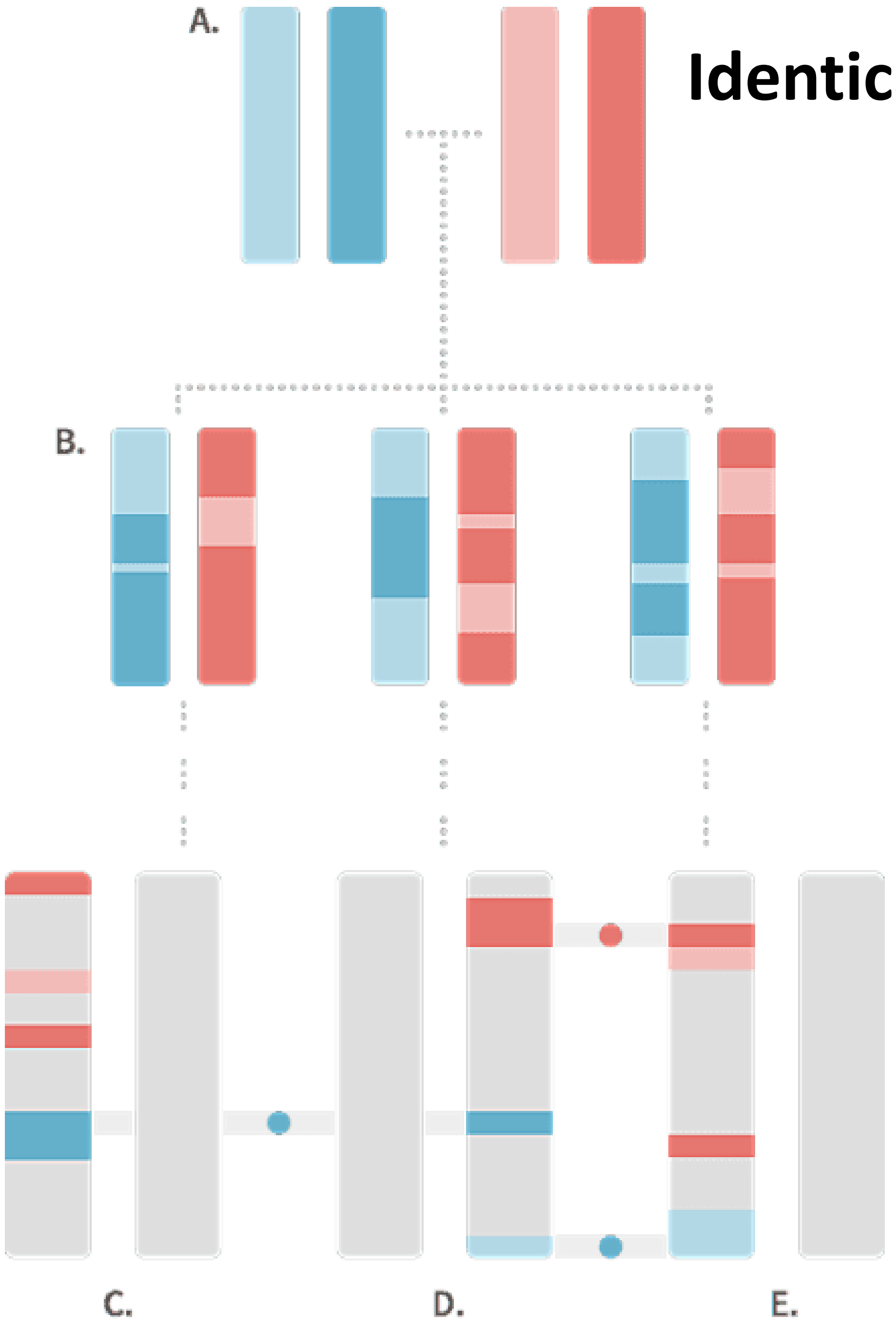
Proportion of alleles in common between two individuals.

Genotyped data can be used to estimate genetic relatedness

Identity by state
Same allele at a locus.

Identity by descent
Same allele at a locus
AND
inherited from a recent
common ancestor allele.

Identical by descent (IBD)



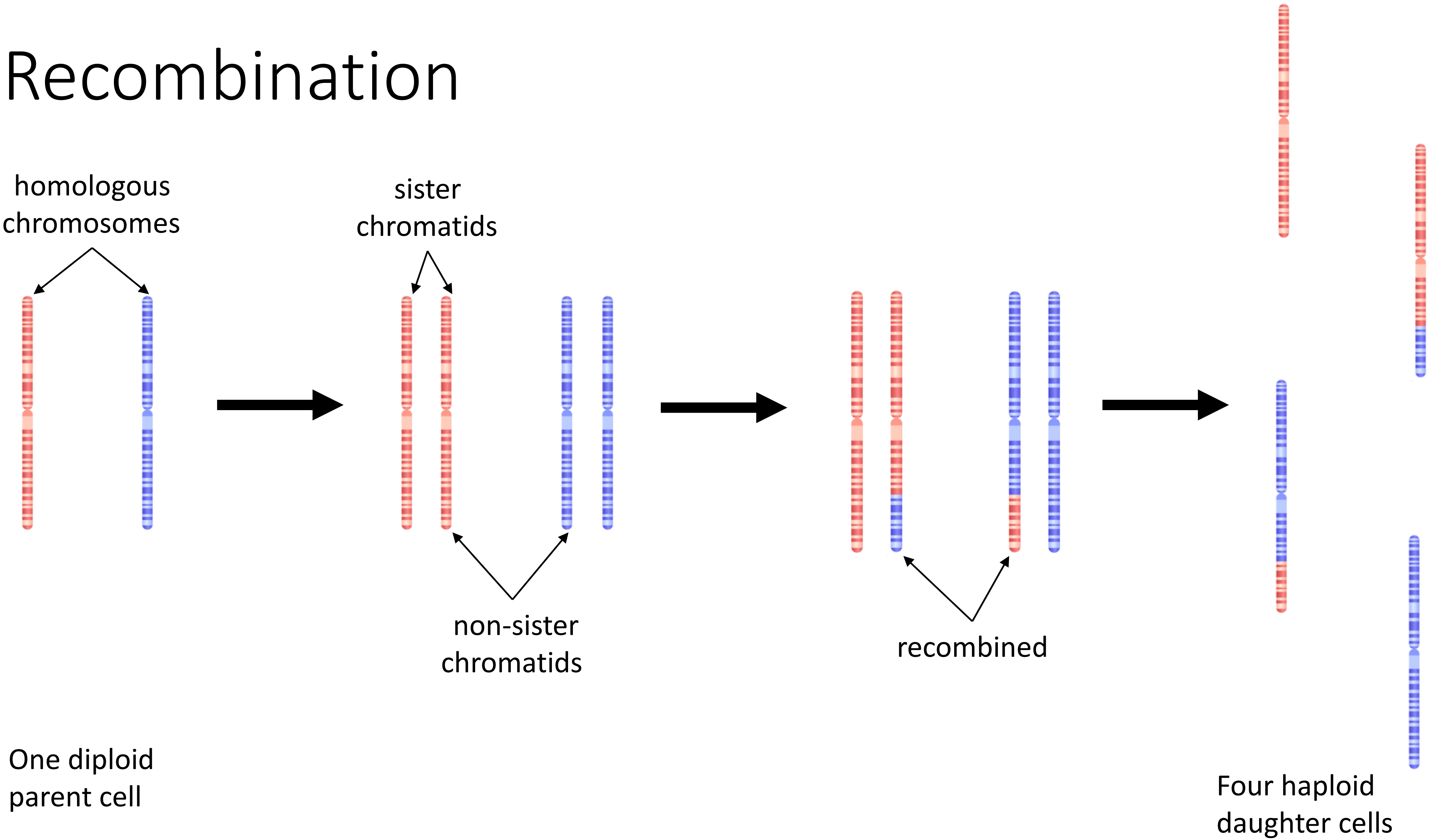
https://www.ancestry.com/cs/dna-help/matches/whitepaper

Identical by state (IBS)

Person 1	Person 2	IBS
AA	AA	2
AA	AG	1
AA	GG	0

…

Recombination



homologous chromosomes

sister chromatids

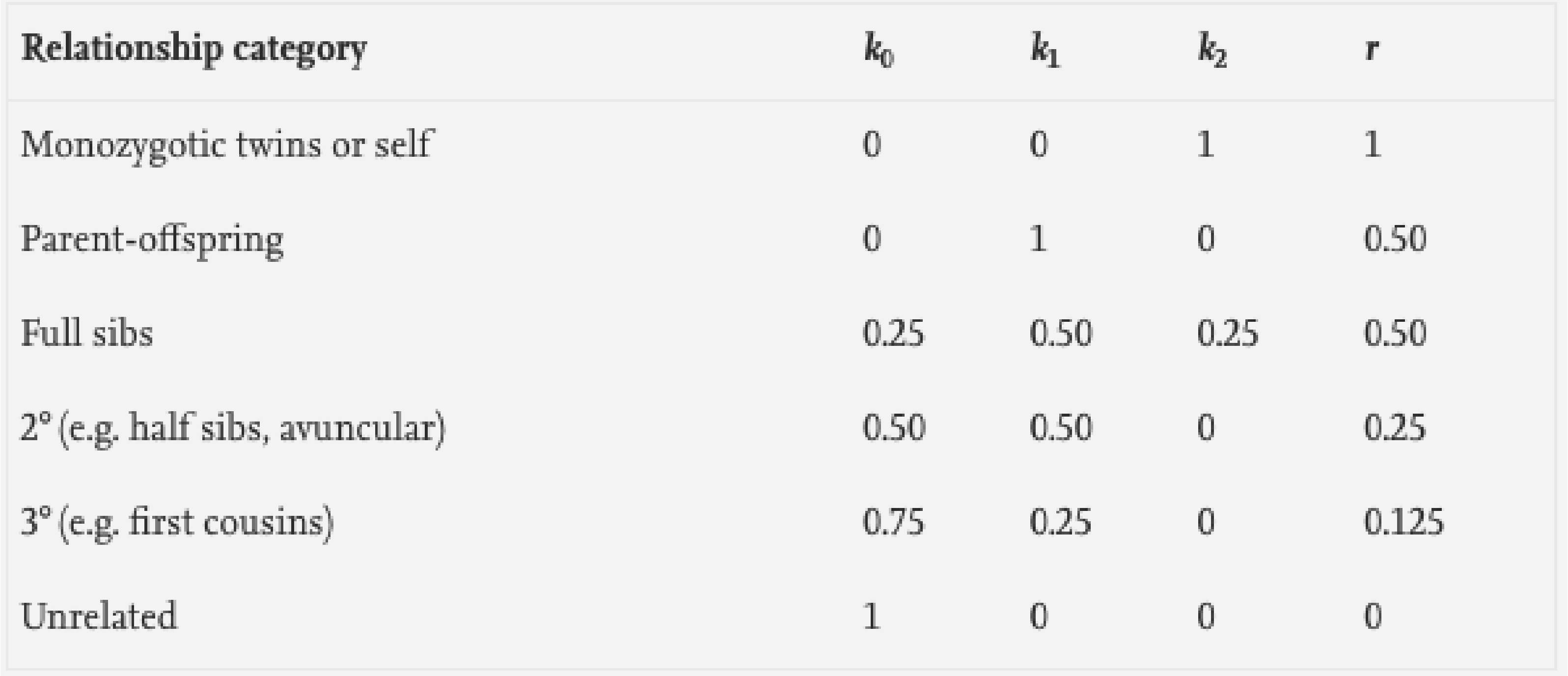
non-sister chromatids

recombined

One diploid parent cell

Four haploid daughter cells

Genetic relatedness vs relationship category



Relationship category	k_0	k_1	k_2	r
Monozygotic twins or self	0	0	1	1
Parent-offspring	0	1	0	0.50
Full sibs	0.25	0.50	0.25	0.50
2° (e.g. half sibs, avuncular)	0.50	0.50	0	0.25
3° (e.g. first cousins)	0.75	0.25	0	0.125
Unrelated	1	0	0	0

k coefficients summarize the probability that pairs of individuals share 0, 1, or 2 alleles at a locus.

The same relationship coefficient can result from different patterns of allelic sharing.

Genomic data can be used to calculate genetic relatedness

Identity by descent
Needs pedigree data.

Can be estimated.
Different methods with varying accuracy.

Typically using pattern of IBS, allele frequency, and hidden Markov models.

PLINK gives $\hat{\pi}$ as a proportion of alleles shared IBD.

Genomic data can be used to calculate genetic relatedness

Identity by state

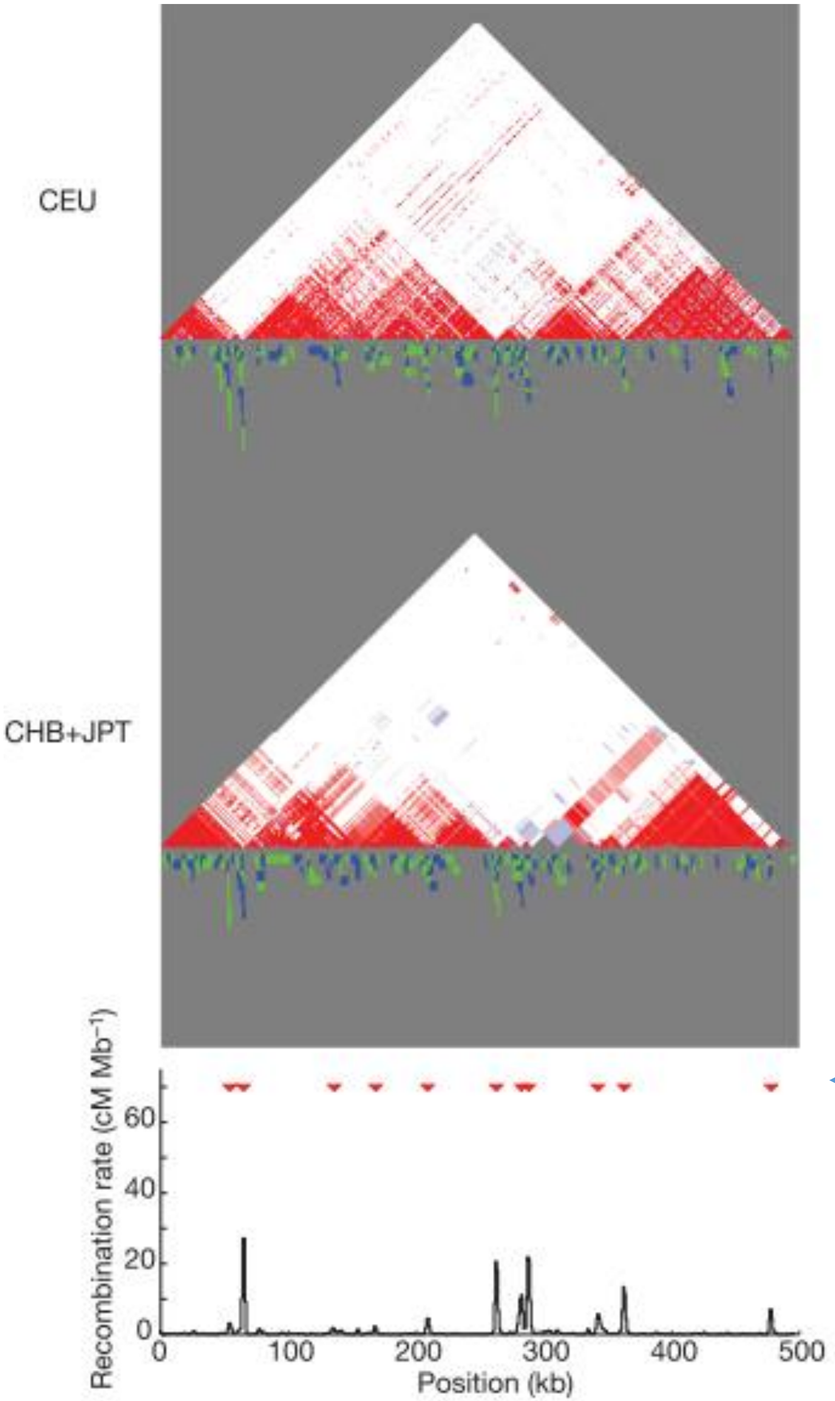
Doesn’t require pedigree information.

Observed.

Gives a coefficient of relationship *r* but doesn’t tell us about the relationship category.

Calculated in GRMs as the average correlation, weighted by allele frequency, between two individuals across genotyped variants.

Background relatedness



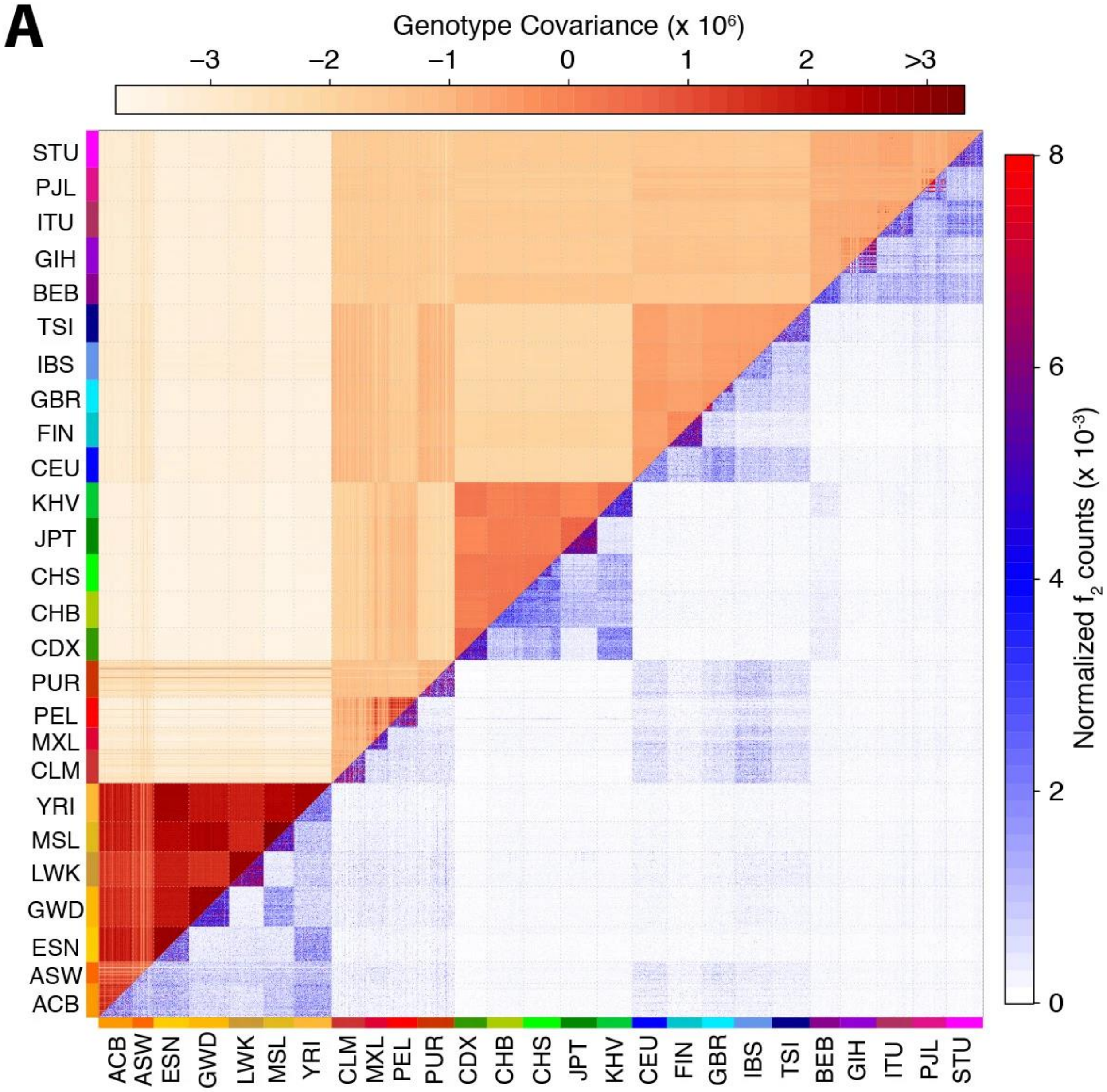
Linkage Disequilibrium

The non-random association of alleles at different loci in a population.

Recombination hot spots

The International HapMap Consortium, Nature 2005

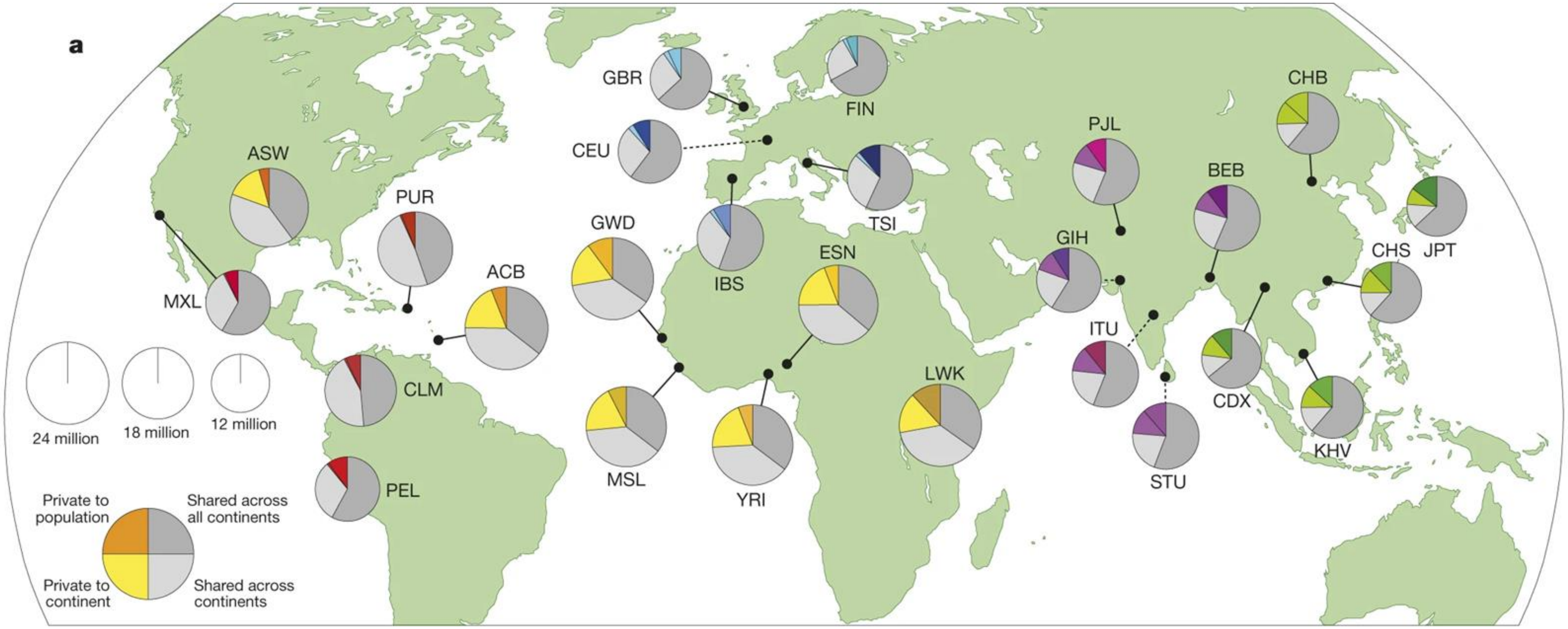
Different populations.
Different shades of the same colour indicate continental grouping of populations.



The 1000 Genomes Project Consortium, Nature 2015

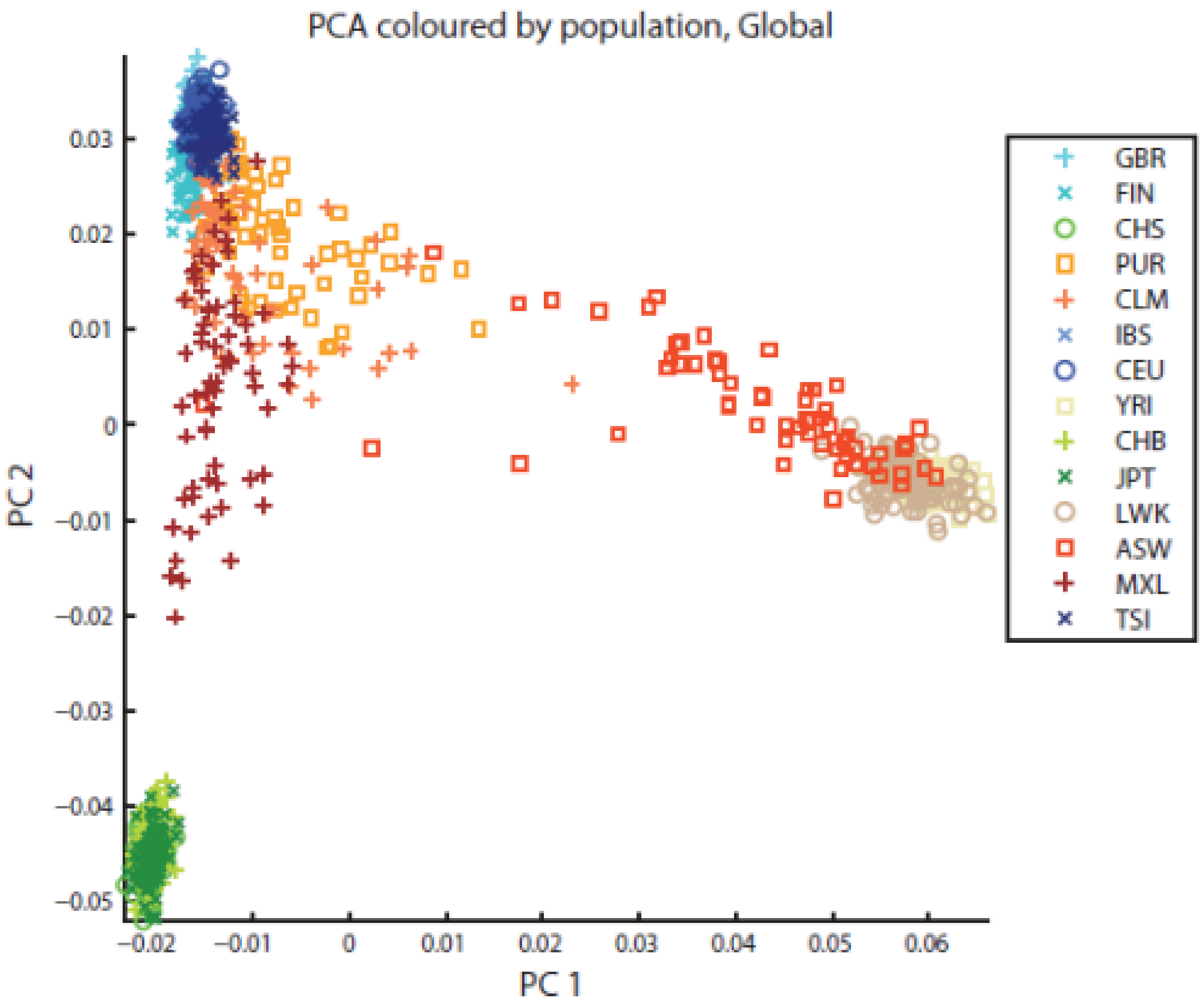
Background relatedness

Allele sharing across populations



The International HapMap Consortium, Nature 2005

Population Structure



The 1000 Genomes Project Consortium, Nature 2015