Common SNPs explain a large proportion of the heritability for human height

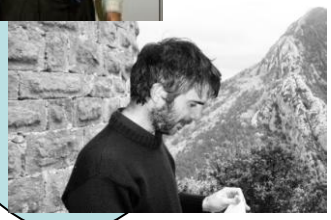
Jian Yang[1], Beben Benyamin[1], Brian P McEvoy[1], Scott Gordon[1], Anjali K Henders[1], Dale R Nyholt[1], Pamela A Madden[2], Andrew C Heath[2], Nicholas G Martin[1], Grant W Montgomery[1], Michael E Goddard[3] & Peter M Visscher[1]

but also, and first:

le parenté (génomique): cet inconnu

peu de généa

Henri IV
(de Navarre
évidemment)



Measurements of relationships

- La matrice de parenté additive (a_{xy}, numerator relationship matrix)
  - n'est pas une matrice de probabilités,
  - mais de 2 * coancestries (proba d'apparénté de Malécot, r_{xy})
- La consanguinité et les apparentés
  - sont relatives à une population de base
  - où l'on définie un apparentement arbitraire (normalement 0).

Molecular relationships

- In conservation genetics, molecular markers have often been used to estimate relationships
  - Either estimates of r_{xy}, or estimates of « the most likely relation » (son-daughter, cousins, whatever)
  - Not very accurate
  - e.g. Ritland, 1996

- Some formulae pop out in later works

The genomic relationship matrix

- But we can say **g** = **Za**

(genetic value = sum of SNP effects).

- If we assume Var(**a**)=**I**σ^2_a , it follows that
  - Var(**g**)=**ZZ'**σ^2_a
- Standardizing
  - Var(**g**)= **ZZ'** σ^2_u /k = **G**σ^2_u

  - Where σ^2_u is « the » additive variance
  - and k = σ^2_u /σ^2_a

The genomic relationship matrix

- How do we get the variance of SNP effects from an estimate of the polygenic variance?

$\sigma^2_a = \sigma^2_u / k$ $\qquad k = 2 \sum_{all\ SNPs} p_i \left(1 - p_i\right)$

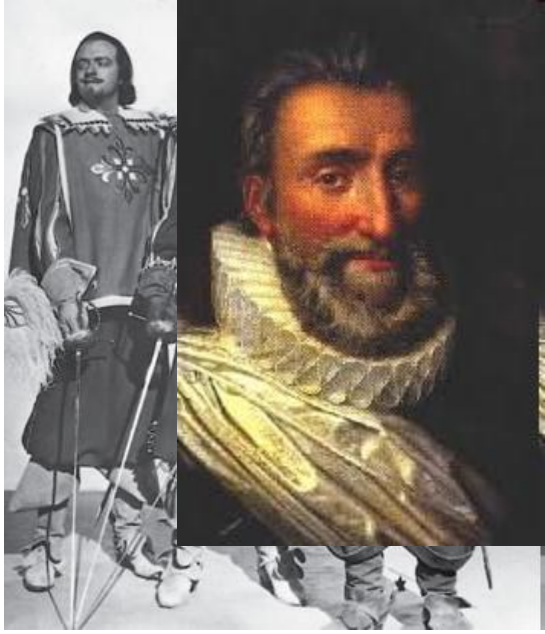
- This formula assumes HW, linkage equilibrium of SNPs (which is false) Gianola et al. (Genetics, 2009)
- *k* is (in HW) equal to trace(**ZZ**')/ number of individuals in data

- *k* is *not* the number of SNPs

The genomic relationship matrix

- The other way around
  - Les SNPs sont des génotypes qui sont transmis selon des règles mendéliennes
  - Donc on peut également appliquer ces lois aux different génotypes
  - et calculer des « vrais » apparentés
- Digression: c'est quoi un « vrai » apparenté?

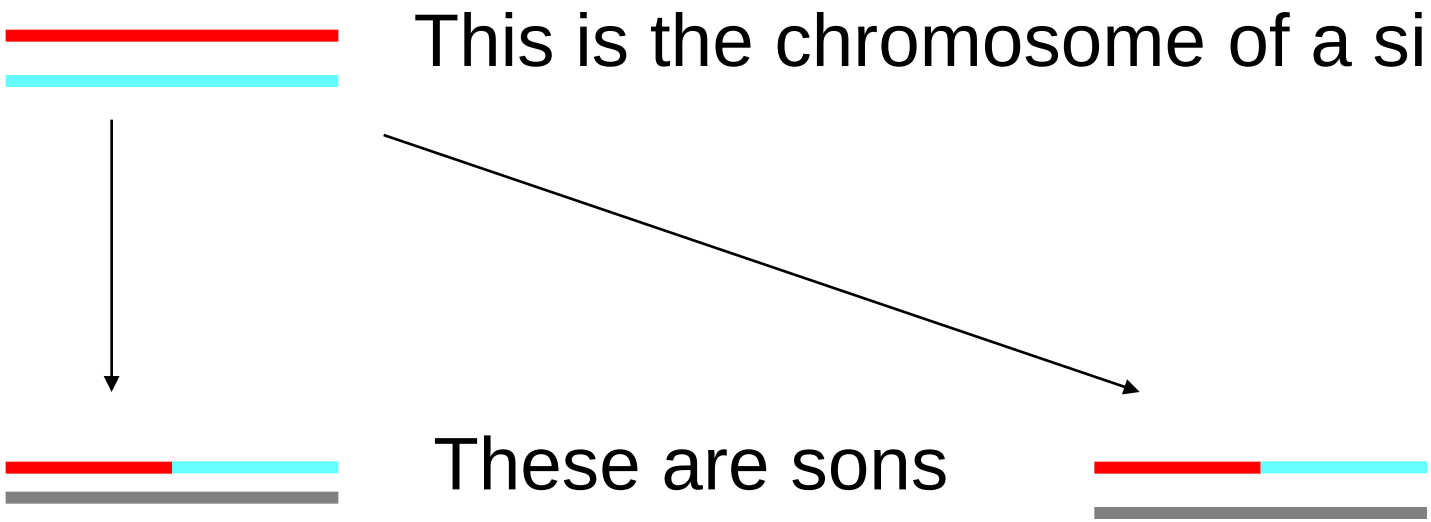
- Ou



The genomic relationship matrix

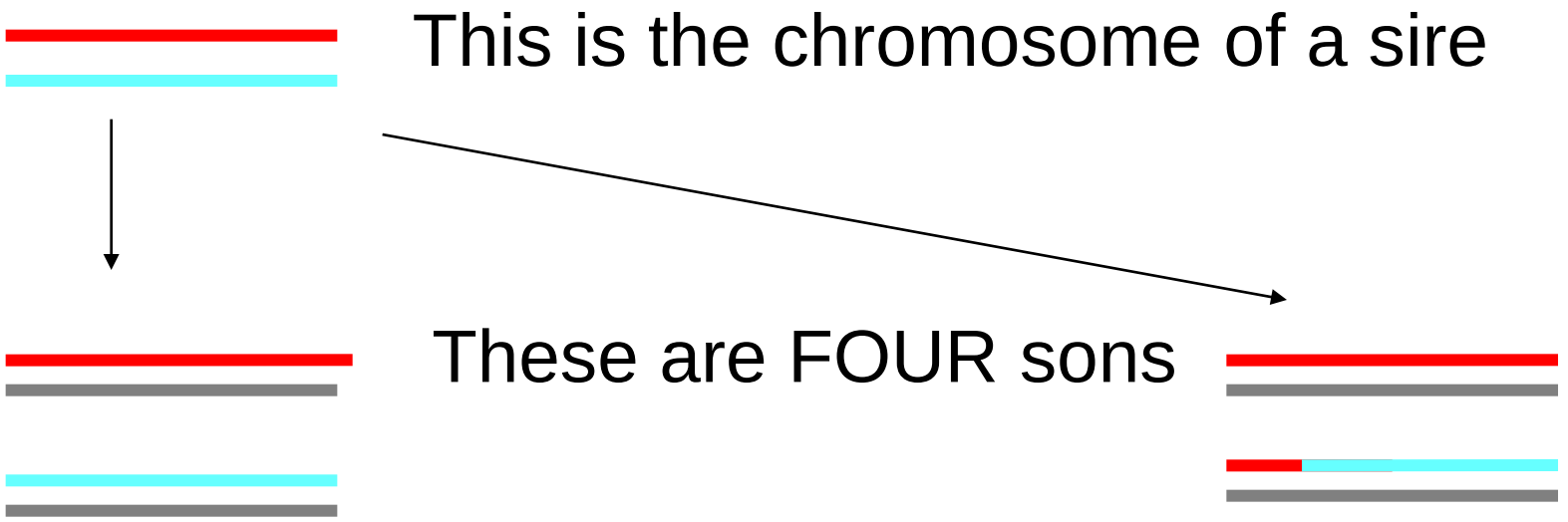
- SNPs are very informative on « true » relationships
- The relationship matrix **A** based on pedigree is an average relationship which assumes many unlinked genes, deviations of which do exist in reality
- SNPs more informative than **A**.
  - Two fullsibs might have a correlation of 0.6 or 0.4
- You need many markers to get these « fine relationships »

Example



In the infinitesimal model, each son receives exactly half the sire.

Example



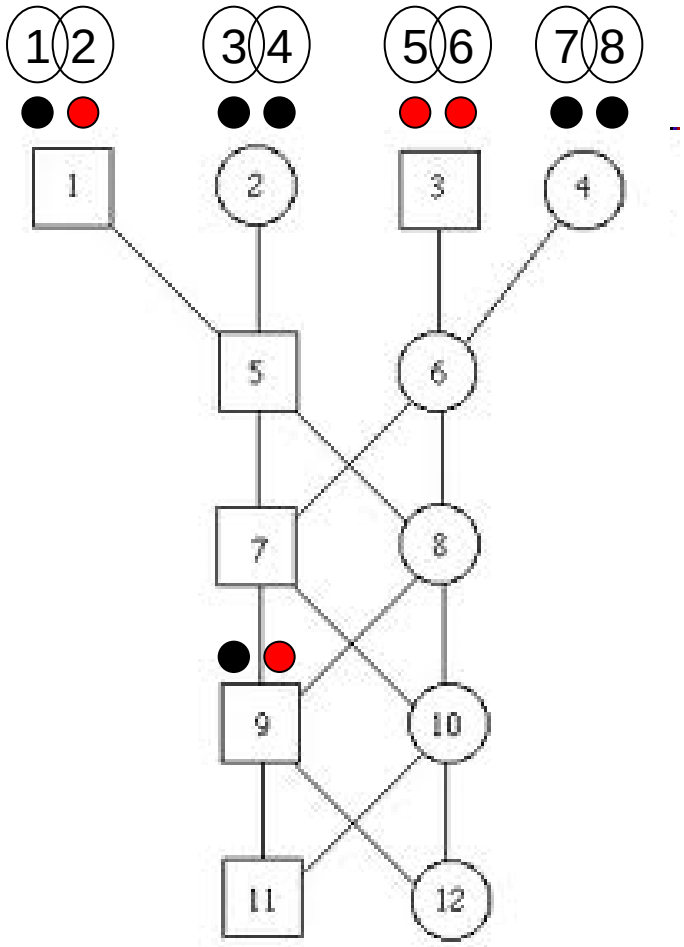
•In reality, two sons are identical and other two are very different from the first two but alike among them.

First derivation

- PVR (2008) explains (without much detail) that **G** (if derived properly) and the pedigree relationship (**A**) are somehow « compatible »
- He provides three derivations
  - I will provide first the rationale why this is true

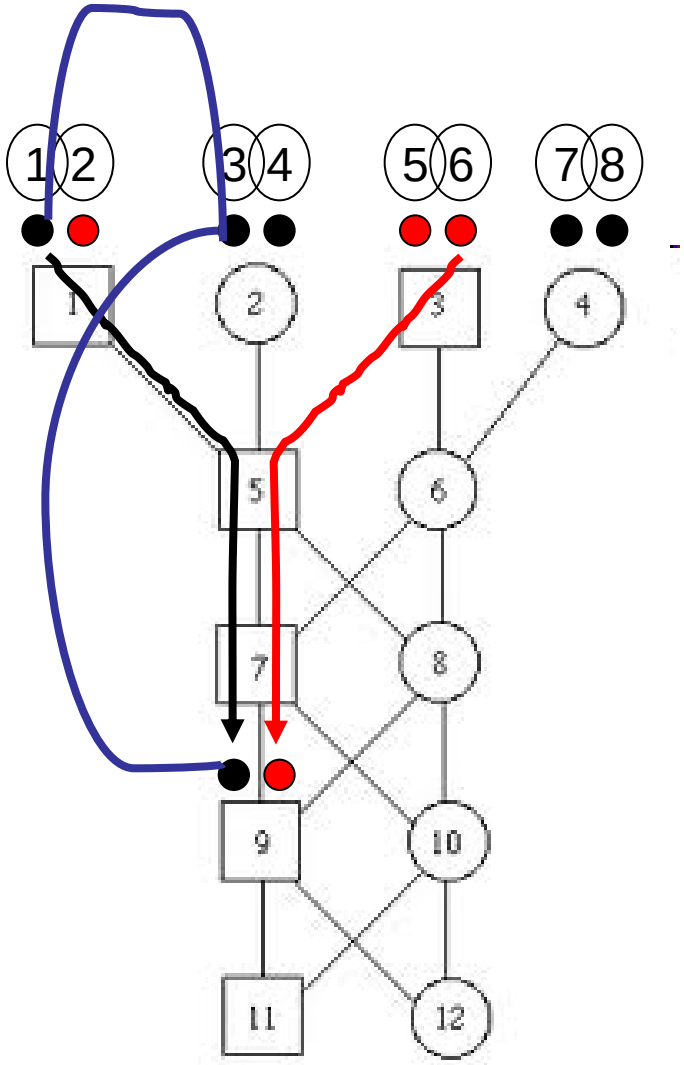
Formal derivation (MA Toro

- Let us imagine that to each one of the 2M founder alleles we assign at random a tag saying if the allele is A or a with probability p and q=1-p
- Then we genotype 9
- Can we say which ancestral allele (1 to 8) inherited 9?



Formal derivation (MA Toro)

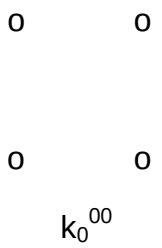
- The molecular coancestry between two individuals x and y will be
  - probability that two alleles are equal (alike in state),
    - either because they have become identical by descent or
    - either because they are not identical by descent but equal in the base population.



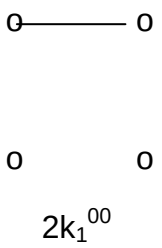
Formal derivation (MA Toro)

- There is a random variable *g* (gene content) with values 0, ½ and 1 for AA, Aa and aa
- We can derive covariances for g in two individuals *i* and *j*
- In a general population, there are nine ways in which relatives can be IBD

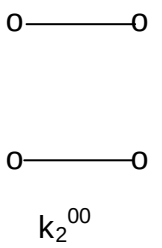
Nine ways in which pair of relatives can share genes identical by descent, with frequencies k_i



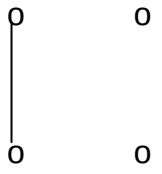
k_0^{00}



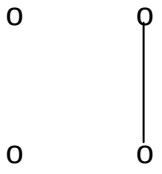
$2k_1^{00}$



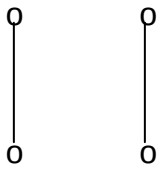
k_2^{00}



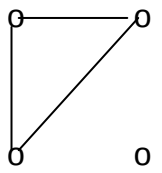
k_0^{10}



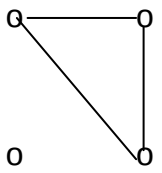
$2k_1^{01}$



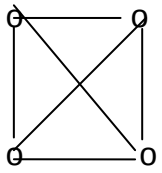
k_2^{11}



$2k_1^{10}$



$2k_1^{01}$



k_2^{11}

- With probabilities (Crow and Kimura)

x	y	f_M	p_X	p_Y	Frequency
AA	AA	1.	1.	1.	$k_0^{00}p^4+(2k_1^{00}+k_0^{10}+k_0^{01})p^3+(k_2^{00}+k_0^{11}+2k_1^{10}+2k_1^{01})p^2+k_2^{11}p$
AA	Aa	0.5	1.	0.5	$k_0^{00}2p^3q+2k_1^{00}p^2q+k_0^{10}2p^2q+2k_1^{10}pq$
Aa	AA	0.5	0.5	1.	$k_0^{00}2p^3q+2k_1^{00}p^2q+k_0^{01}2p^2q+2k_1^{01}pq$
AA	aa	0.	1.	0.	$k_0^{00}p^2q^2+k_0^{10}pq^2+k_0^{01}p^2q+k_0^{11}pq$
aa	AA	0.	0.	1.	$k_0^{00}4p^2q^2+k_0^{10}p^2q+k_0^{01}pq^2+k_0^{11}pq$
Aa	Aa	0.5	0.5	0.5	$k_0^{00}p^2q^2+2k_1^{00}pq+k_2 2pq$
Aa	aa	0.5	0.5	0.5	$k_0^{00}2pq^3+2k_1^{00}pq2+k_0^{01}2pq^2+2k_1^{01}pq$
aa	Aa	0.5	0.	0.5	$k_0^{00}2pq^3+2k_1^{00}pq^2+k_0^{10}2pq^2+2k_1^{10}pq$
aa	aa	1.	0.	0.	$k_0^{00}q^4+(2k_1^{00}+k_0^{10}+k_0^{01})q^3+(k_2^{00}+k_0^{11}+2k_1^{10}+2k_1^{01})q^2+k_2^{11}q$

- and it follows that

$$f_{xy} = \frac{1}{pq}\mathrm{cov}_{Mxy}$$

Coancestry

Covariance of gene content

- In other words
  - $Cov(g_i,g_j) = r_{ij}/pq$

$r_{ij} = A_{ij}/2$

- This holds « on expectation » for each locus
  - p's are those in the base population!!
- The question is how we « pool » information across loci

The genomic relationship matrix

- I will show three parameterizations
  - Malécot coefficient of identity by state
  - Paul Van Raden's 2008 relationships
- All three correspond to different linear models

Malécot (IBS)

- 2*Malécot coefficients of identity (by state)
- It considers that every allele of every SNP is a gene
- Corresponds to a linear model in which every allele of every SNP has an effect, and this SNP has « a priori » 0 mean *(this is a problem)*
  - (size of **a** = 2 * number of SNPs)

Most common G Van Raden (2008), Amin et al. (2008), Astle & Balding (2009), Yang et al. (2010) (second **G**)

- Estimator of relationship

$$G_{ij} = 2\frac{1}{n}\sum \frac{\left(g_{ik} - p_k\right)\left(g_{jk} - p_k\right)}{p_k\left(1 - p_k\right)}$$

- We estimate a relationship by locus, and then we estimate its average
- Less polymorhic locus have more weight

Paul Van Raden (2008) »first **G** «

- Compute a covariance by each locus
- And divide by average variance (implicitely in H-W, linkage equilibrium)

$$G_{ij} = 2\frac{1}{n}\frac{\sum\left(g_{ik}-p_k\right)\left(g_{jk}-p_k\right)}{\sum p_k\left(1-p_k\right)} \qquad \mathbf{G} = \frac{\mathbf{Z}\mathbf{Z}'}{2\sum p_i\left(1-p_i\right)}$$

- More intuitive as a linear mixed model
  - Corresponds to the work of Gianola (2009)

Some properties

- In H-W, Linkage equilibrium
  - Average of Diag(G) = 1
  - Average off-diagonal(G) =0
  - Average genetic value of genotyped individuals =0
  - This corresponds to the definition of base population
- With average inbreeding F,
  - Average of Diag(G) = 1+F

Mixing molecular & pedigree relationships

- Many animals do not have genotypes and it would be nice to include them in the genomic relationship matrices
- There are two attempts to do so (Legarra et al., 2009; Christensen & Lund, 2010)
- Both use pedigree-based “predictions” (and their variances) of genetic values or SNP genotypes and arrive to the same result

$$Var\begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{pmatrix} = \mathbf{H} = \begin{bmatrix}\mathbf{H}_{11} & \mathbf{H}_{12}\\ \mathbf{H}_{21} & \mathbf{H}_{22}\end{bmatrix} =$$

$$\begin{bmatrix}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{G}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} + \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{G}\\ \mathbf{G}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{G}\end{bmatrix}$$

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1}\end{bmatrix}$$

- $\mathbf{H}^{-1}$ has been used in one-step genetic evaluation (Aguilar et al., 2010)
- Still not well understood

Unsolved problems

- Full compatibility of « genomic » and « pedigree » relationships
  - Only important if we want to mix both informations (as in the single-step procedure)
  - We need thus the same genetic base:
    - Same constraint on the genetic values (average breeding value of the base = 0)
    - Same genetic variance
- Achieved using base allelic frequencies
  - But these are impossible to estimate (well)

Unsolved problems

- Ad-hoc corrections:
  - Scaling: divide ZZ' by its trace and not $2\sum p_i\left(1-p_i\right)$
    - Useful if there is not H-W
  - Sum to achieve same average coancestry

$$\mathbf{G}^{\dagger}=\mathbf{G}+\mathbf{1}\mathbf{1}'\alpha \qquad \alpha=\frac{1}{n^2}\left[\sum_i\sum_j \mathbf{A}_{22\ (i,j)}-\sum_i\sum_j \mathbf{G}_{i,j}\right]$$

    - Very useful if there is selection (Vitezica)
  - Regress **G** on **A** (Van Raden)

$$\mathbf{MM}' = g_0\mathbf{1}\mathbf{1}' + g_1\mathbf{A} + \mathbf{E},$$

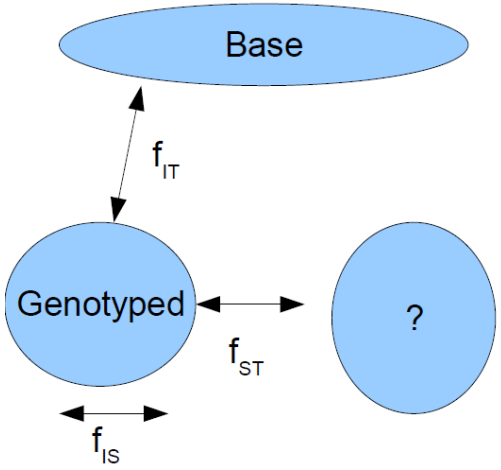
    - Multiple breed version (Harris & Johnson)

$$\mathbf{G} = \mathbf{L}_1\hat{\mathbf{F}}_1^{-1}\left(\mathbf{ZZ}' - \sum_{k\le l}\hat{b}_{1(kl)}\mathbf{J}_{(kl)}\right)_1\hat{\mathbf{F}}_1'^{-1}\mathbf{L}_1'$$

Unsolved problems

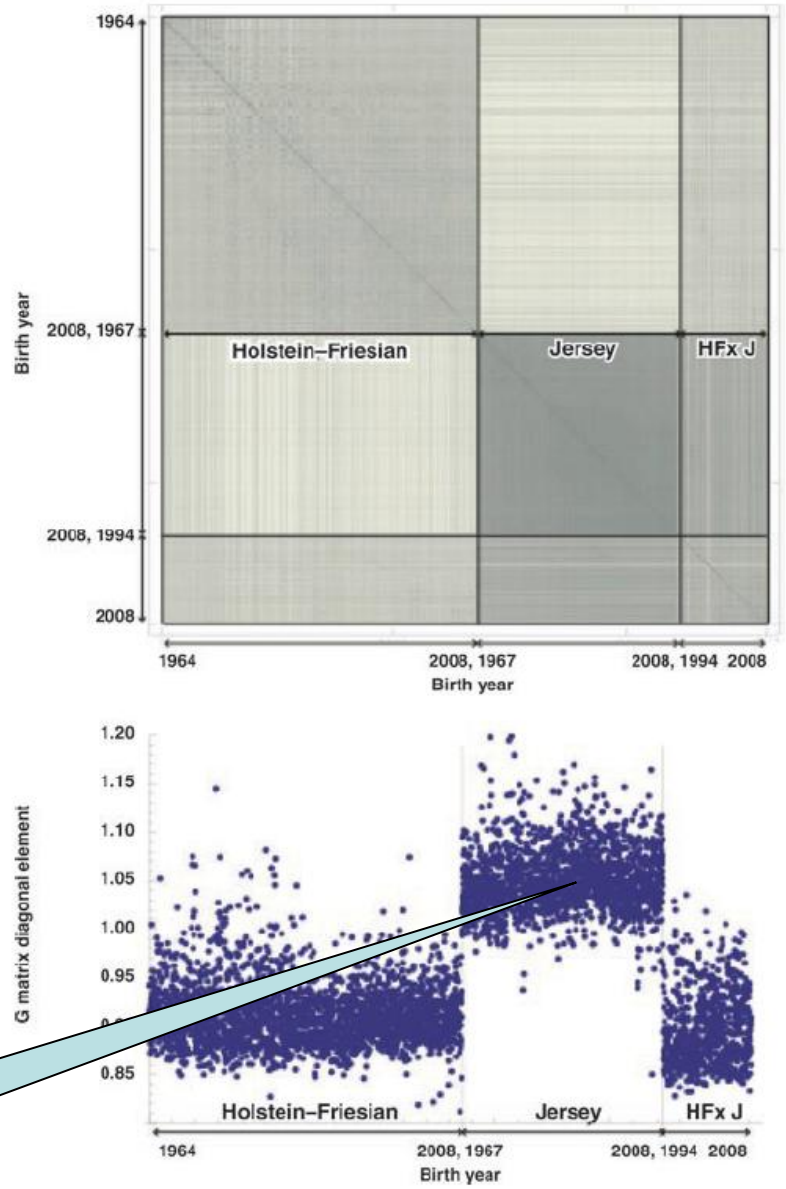
- Possibly, a correction based on Wright's Fst can be achieved (suggestion by ME Goddard)

$$(1 - f_{IT}) = (1 - f_{IS})(1 - f_{ST})$$



G for a crossbred population (Harris & Johnson)

- Before correction



Too high inbreeding

Figure 2. Heat map of genomic relationship matrix estimated ignoring breed and using whole-population SNP frequencies; darker areas correspond to a greater degree of relationship. The lower graph displays diagonal elements. HF = Holstein-Friesian; J = Jersey.

G for a crossbred population (Harris & Johnson)

- After correction

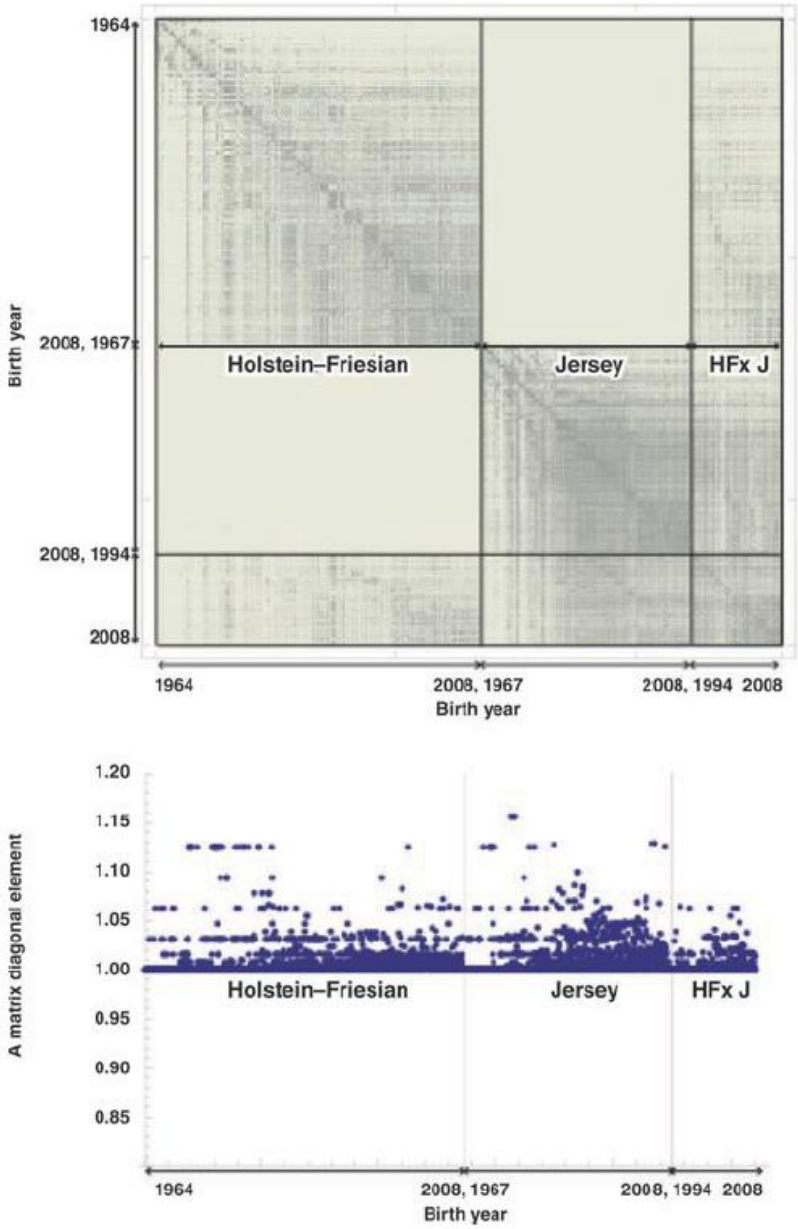
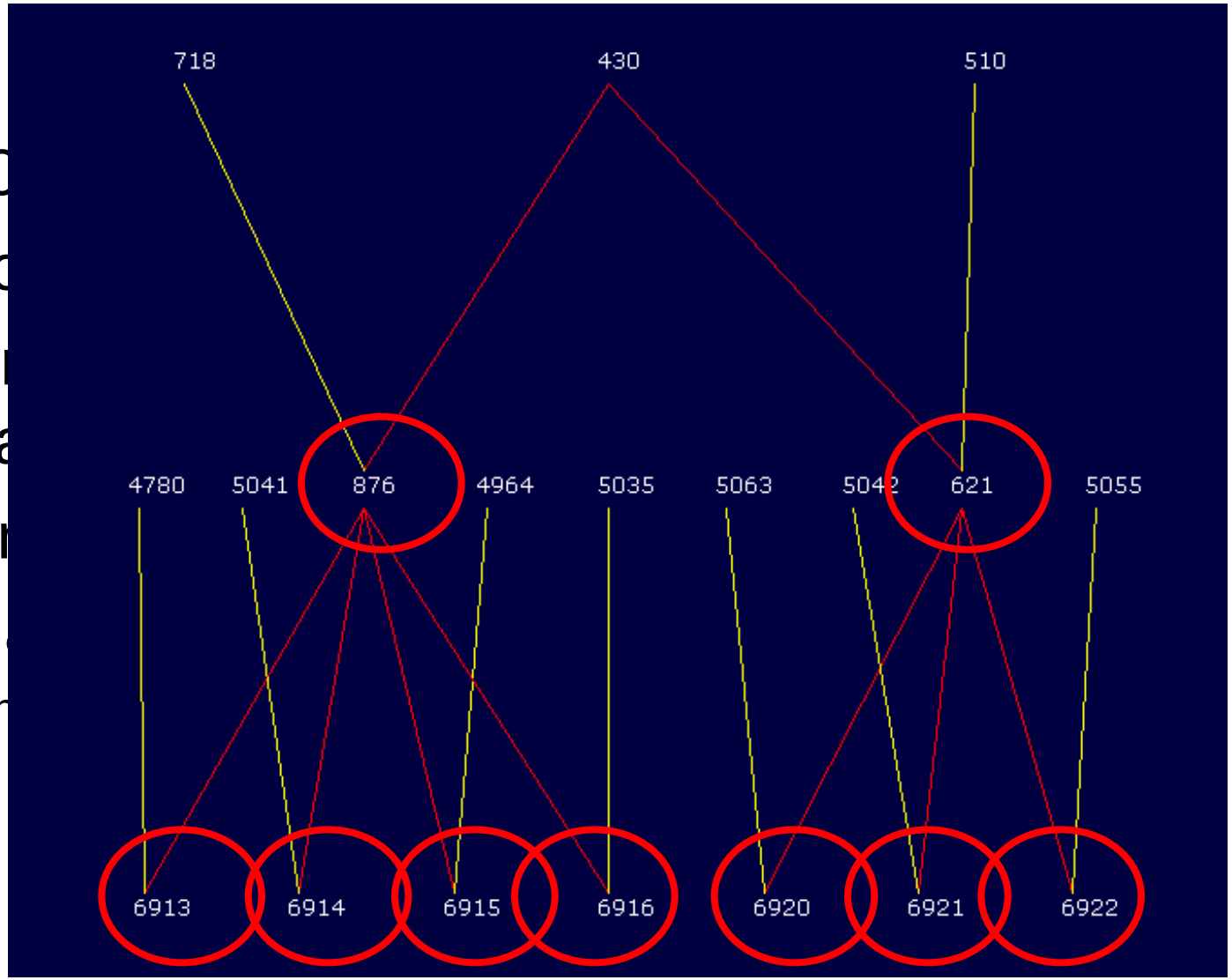


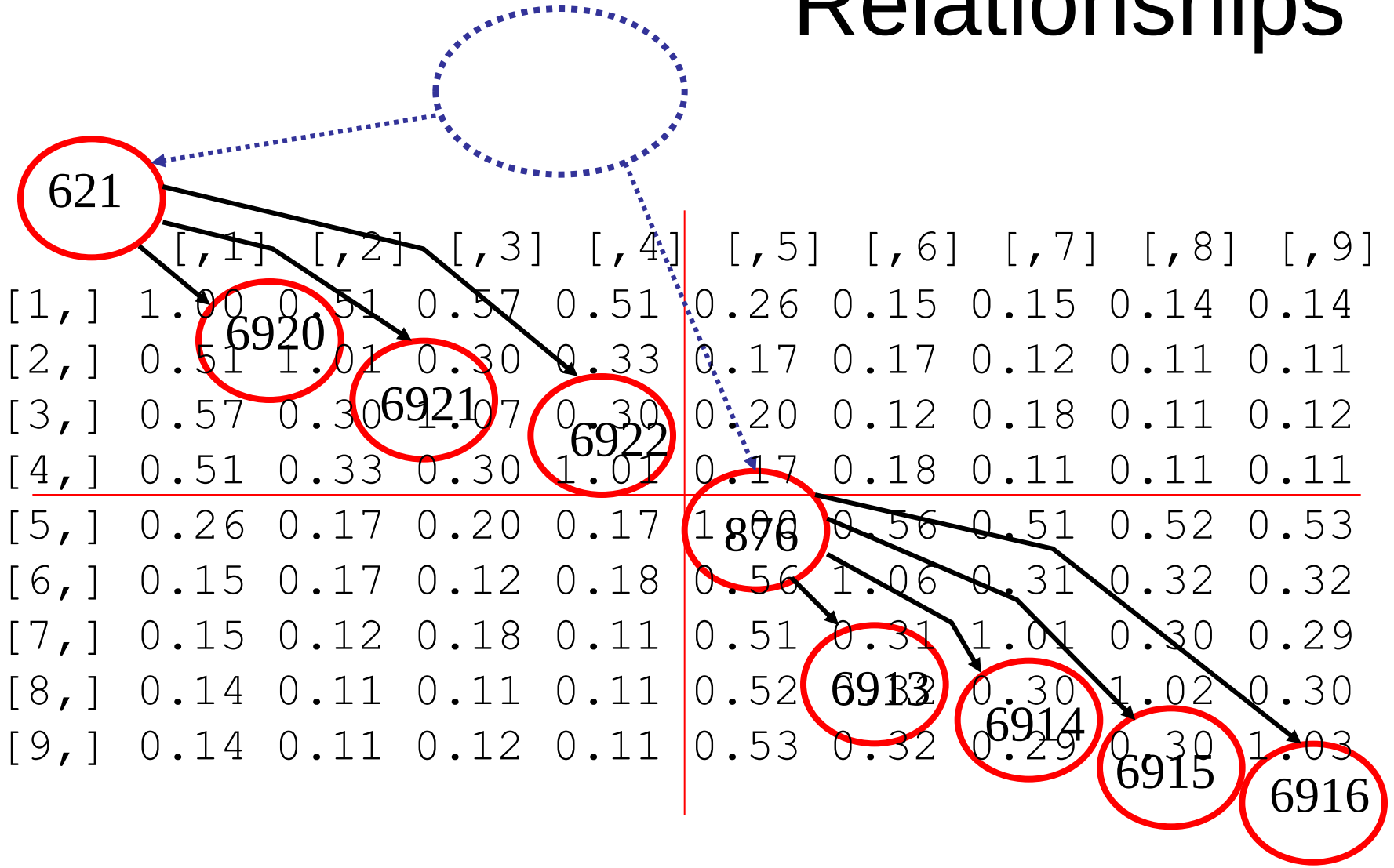
Figure 1. Heat map of genotyped block of average relationship matrix; darker areas correspond to a greater degree of relationship. The lower graph displays diagonal elements. HF = Holstein-Friesian; J = Jersey.

Real results (AMASGEN)

- 9 real
  ~5000
- Very c
- All ge
  estima
- Genor
- Popul
- Program



Relationships



	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]
[1,]	1.00	0.51	0.57	0.51	0.26	0.15	0.15	0.14	0.14
[2,]	0.51	1.01	0.30	0.33	0.17	0.17	0.12	0.11	0.11
[3,]	0.57	0.30	1.07	0.30	0.20	0.12	0.18	0.11	0.12
[4,]	0.51	0.33	0.30	1.01	0.17	0.18	0.11	0.11	0.11
[5,]	0.26	0.17	0.20	0.17	1.00	0.56	0.51	0.52	0.53
[6,]	0.15	0.17	0.12	0.18	0.56	1.06	0.31	0.32	0.32
[7,]	0.15	0.12	0.18	0.11	0.51	0.31	1.01	0.30	0.29
[8,]	0.14	0.11	0.11	0.11	0.52	0.32	0.30	1.02	0.30
[9,]	0.14	0.11	0.12	0.11	0.53	0.32	0.29	0.30	1.03

(whole) Pedigree-based relationship

Little inbreeding

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]
[1,]	1.00	0.51	0.57	0.51	0.26	0.15	0.15	0.14	0.14
[2,]	0.51	1.01	0.30	0.33	0.17	0.17	0.12	0.11	0.11
[3,]	0.57	0.30	1.07	0.30	0.20	0.12	0.18	0.11	0.12
[4,]	0.51	0.33	0.30	1.01	0.17	0.18	0.11	0.11	0.11
[5,]	0.26	0.17	0.20	0.17	1.00	0.56	0.51	0.52	0.53
[6,]	0.15	0.17	0.12	0.18	0.56	1.06	0.31	0.32	0.32
[7,]	0.15	0.12	0.18	0.11	0.51	0.31	1.01	0.30	0.29
[8,]	0.14	0.11	0.11	0.11	0.52	0.32	0.30	1.02	0.30
[9,]	0.14	0.11	0.12	0.11	0.53	0.32	0.29	0.30	1.03

Relationships among cousins are ~ 0.125

“Second G” genomic relationship

Less than 1 in the diagonal

Negative coefficients

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]
[1,]	0.82	0.40	0.43	0.38	0.12	0.04	0.04	0.01	0.10
[2,]	0.40	0.91	0.18	0.24	0.02	0.05	-0.04	-0.04	0.04
[3,]	0.43	0.18	0.88	0.19	0.07	0.00	0.07	-0.02	0.05
[4,]	0.38	0.24	0.19	0.86	0.02	-0.01	-0.02	0.01	0.03
[5,]	0.12	0.02	0.07	0.02	0.73	0.34	0.30	0.31	0.35
[6,]	0.04	0.05	0.00	-0.01	0.34	0.85	0.15	0.14	0.18
[7,]	0.04	-0.04	0.07	-0.02	0.30	0.15	0.80	0.14	0.17
[8,]	0.01	-0.04	-0.02	0.01	0.31	0.14	0.14	0.80	0.17
[9,]	0.10	0.04	0.05	0.03	0.35	0.18	0.17	0.17	0.85

Relationships among cousins are ~0

$$G_{ij} = 2\frac{1}{n}\frac{\sum\left(g_{ik}-p_k\right)\left(g_{jk}-p_k\right)}{\sum p_k\left(1-p_k\right)}$$

“First **G**” genomic relationship

Closer to 1 in the diagonal

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]
[1,]	0.91	0.44	0.47	0.42	0.14	0.05	0.05	0.02	0.11
[2,]	0.44	1.01	0.20	0.27	0.02	0.06	-0.04	-0.04	0.04
[3,]	0.47	0.20	0.98	0.21	0.07	0.00	0.08	-0.02	0.05
[4,]	0.42	0.27	0.21	0.96	0.02	-0.01	-0.02	0.01	0.04
[5,]	0.14	0.02	0.07	0.02	0.81	0.37	0.33	0.35	0.39
[6,]	0.05	0.06	0.00	-0.01	0.37	0.94	0.16	0.15	0.20
[7,]	0.05	-0.04	0.08	-0.02	0.33	0.16	0.88	0.15	0.19
[8,]	0.02	-0.04	-0.02	0.01	0.35	0.15	0.15	0.88	0.18
[9,]	0.11	0.04	0.05	0.04	0.39	0.20	0.19	0.18	0.94

Very similar but more “exaggerated”

$$G_{ij} = 2\frac{1}{n}\sum \frac{(g_{ik} - p_k)(g_{jk} - p_k)}{p_k(1 - p_k)}$$

Malécot genomic relationship

Large coefficients

This is because it assumes that the two alleles at one locus are independents

					[,6]	[,7]	[,8]	[,9]	
					1.34	1.34	1.33	1.36	
					1.34	1.30	1.30	1.33	
					1.32	1.35	1.31	1.33	
[4,]	1.45	1.39	1.38	1.63	1.34	1.32	1.31	1.32	1.33
[5,]	1.38	1.34	1.36	1.34	1.65	1.48	1.46	1.47	1.48
[6,]	1.34	1.34	1.32	1.32	1.48	1.66	1.39	1.39	1.40
[7,]	1.34	1.30	1.35	1.31	1.46	1.39	1.64	1.39	1.40
[8,]	1.33	1.30	1.31	1.32	1.47	1.39	1.39	1.64	1.40
[9,]	1.36	1.33	1.33	1.33	1.48	1.40	1.40	1.40	1.66

“Second G” genomic relationship after Yang et al. correction for the diagonal

Very close to 1 in the diagonal

Negative coefficients

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]
[1,]	0.93	0.40	0.43	0.38	0.12	0.04	0.04	0.01	0.10
[2,]	0.40	1.00	0.18	0.24	0.02	0.05	-0.04	-0.04	0.04
[3,]	0.43	0.18	0.98	0.19	0.07	0.00	0.07	-0.02	0.05
[4,]	0.38	0.24	0.19	0.96	0.02	-0.01	-0.02	0.01	0.03
[5,]	0.12	0.02	0.07	0.02	0.93	0.34	0.30	0.31	0.35
[6,]	0.04	0.05	0.00	-0.01	0.34	0.99	0.15	0.14	0.18
[7,]	0.04	-0.04	0.07	-0.02	0.30	0.15	0.95	0.14	0.17
[8,]	0.01	-0.04	-0.02	0.01	0.31	0.14	0.14	0.95	0.17
[9,]	0.10	0.04	0.05	0.03	0.35	0.18	0.17	0.17	0.98

Relationships among cousins are ~0

G for a crossbred population (Harris & Johnson)

- Before correction

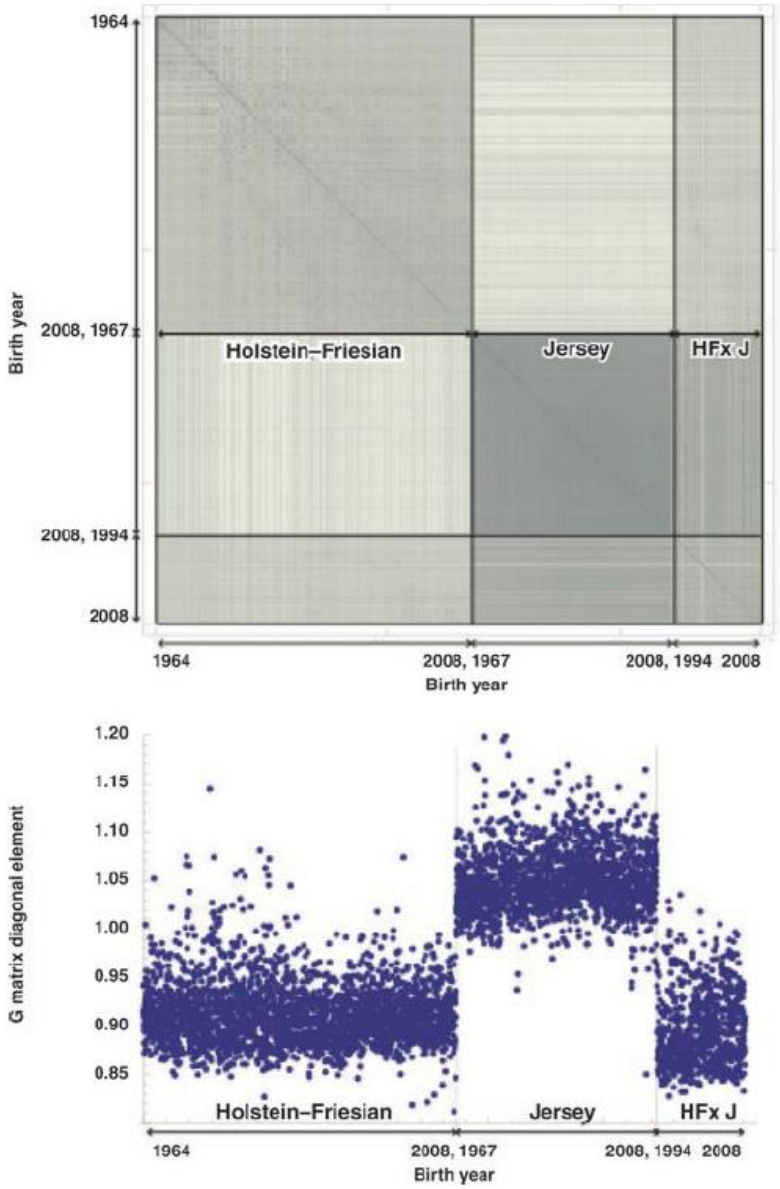


Figure 2. Heat map of genomic relationship matrix estimated ignoring breed and using whole-population SNP frequencies; darker areas correspond to a greater degree of relationship. The lower graph displays diagonal elements. HF = Holstein-Friesian; J = Jersey.

G for a crossbred population (Harris & Johnson)

- After correction

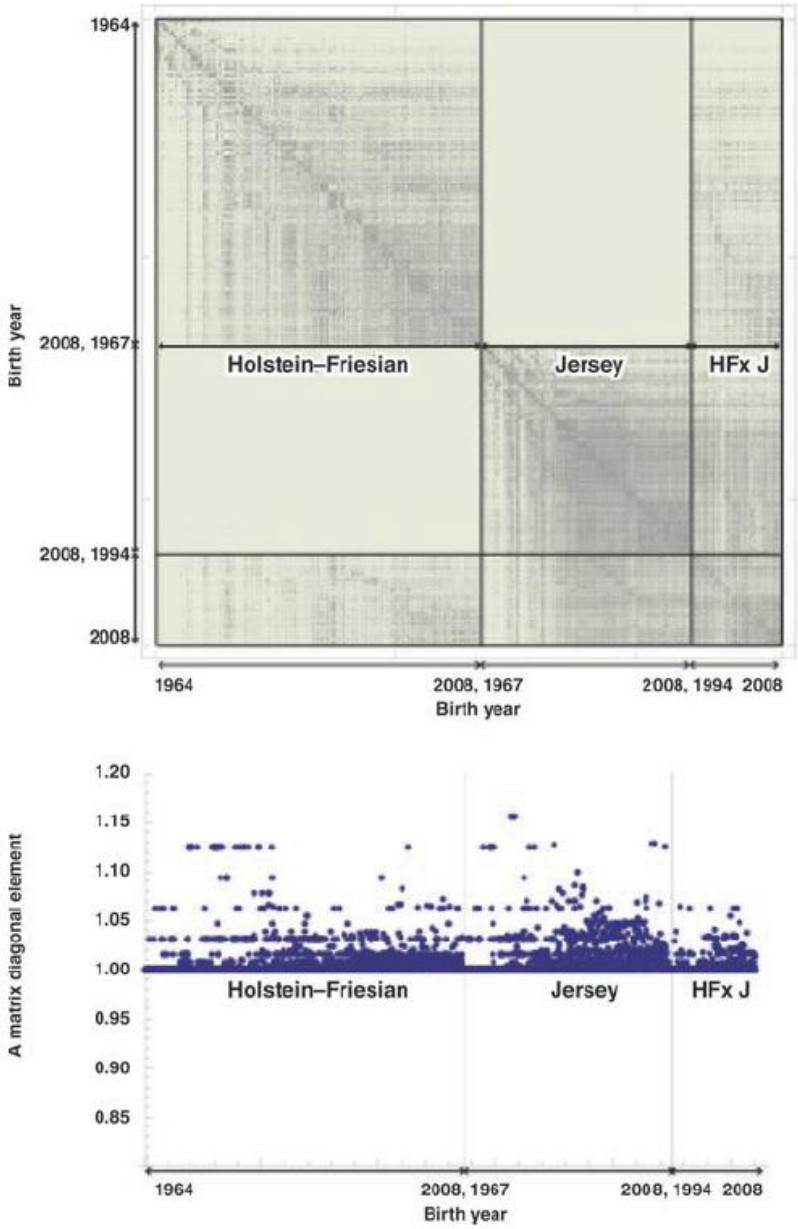


Figure 1. Heat map of genotyped block of average relationship matrix; darker areas correspond to a greater degree of relationship. The lower graph displays diagonal elements. HF = Holstein-Friesian; J = Jersey.

Use of **G**

- Genomic selection (GBLUP)
- Estimation of genomic parameters (GREML)
  - In populations with no pedigree recording
  - How much variance due to SNPs, how to pedigree
- Improved association analysis model (Yu et al…)
  - $\mathbf{y} = SNP_i + \mathbf{g} + \mathbf{e}$, $\mathbf{g} \sim N(0, \mathbf{G}\sigma_g^2)$

Conclusions

- Genomic relationships work very well and are (now) well defined
- The exact formula depends on the interpretation but results do not change much
- Unless somebody wants to combine pedigree and molecular relationships

Common SNPs explain a large proportion of the heritability for human height

Jian Yang[1], Beben Benyamin[1], Brian P McEvoy[1], Scott Gordon[1], Anjali K Henders[1], Dale R Nyholt[1], Pamela A Madden[2], Andrew C Heath[2], Nicholas G Martin[1], Grant W Montgomery[1], Michael E Godd Peter M Visscher[1]

- Or: The « missing » heritability was always there

Missing heritability

- Found SNP variants explaining height explain a *very* small fraction of heritability
- Most likely explanation lots of variations and little power

In the paper

- Use a mixed model to estimate heritability
- Explain we do they found less than expected
- They say it's because typical QTLs have <0.1 MAF

- What I think
  - I don't fully believe their explanation
  - But it is a possibility
  - And the methods are very interesting

Methods

- Estimate heritability by REML using SNPs in « unrelated » population and a genomic relationship matrix
- Kinship estimated using slightly modified formula with correction for the diagonal

$$A_{jk} = \frac{1}{N}\sum_i A_{ijk} = \begin{cases} \frac{1}{N}\sum_i \frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1 - p_i)}, & j \neq k \\ 1 + \frac{1}{N}\sum_i \frac{x_{ij}^2 - (1 + 2p_i)x_{ij} + 2p_i^2}{2p_i(1 - p_i)}, & j = k \end{cases}$$

- « Unrelated » individuals: relationships from -0.025 to 0.025
  - Is this not a problem?

Results

- Estimate of h2 = 0.45 (+- 0.08)
- Usual estimate is 0.8
- Why?

Is « relationship » a « true » relationship?

- Hypothesis: SNP do not provide realistic estimates of relationships because they are not « true » QTLs
  - What if QTLs have smaller MAF than SNPs?
  - Then relationships are « under » estimated
  - Can be checked by comparing A_{ij} estimated with SNPs at low MAF and A_{ij} estimated with all

$$A^*_{jk} = \begin{cases} \beta A_{jk}, & j \neq k \\ 1 + \beta(A_{jk} - 1), & j = k \end{cases} \qquad \beta = 1 - \frac{(c + 1/N)}{\mathrm{var}(A_{jk})}$$

- Assume MAF of QTLs is <0.1, then re-compute **A***

Results 2

- Estimate of h2 = 0.84 (+- 0.16)
- Usual estimate is 0.8
- Are we happy?

This does not prove that the causal variants have MAF < 0.1, but it shows that if this were the case, they could explain the estimated heritability of height (~0.8).

Conclusions

- Missing heritability is there, but GWAS tests are just too stringent. Random models overcome this problem.
- Possibly, not all causal variants are well tagged by SNPs
  - (problem of SNP chip but also of amount of data)

Criticism

- Why do we need to correct the genomic matrix?
  - Estimates of 0.8 can possibly be obtained with « uncorrected » pedigre relationship matrix?
- Is the second heritability « the same »?
  - Do they refer to the same genetic base?

Variance of the base population

Short example:

$$\mathbf{g} \sim \mathbf{G}\,\sigma^2_g$$

- These two formulations parent-son are equivalent
- Is the first less inbred with more variance or the second less inbred with more variance?

$$\begin{pmatrix} u_s \\ u \end{pmatrix} \sim \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix} 1.1$$

- If we manipulate **G**, we *possibly* refer to different things

$$\begin{pmatrix} u_s \\ u \end{pmatrix} \sim \begin{pmatrix} 1.1 & 0.55 \\ 0.55 & 1.1 \end{pmatrix} 1$$

Real example (mice data)

- I took one **G** computed for the mice data and estimated variance components with **G**, and with **G*** = **G***0.5
- The heritability increases artificially

	varg	varu	varc	vare	h^2
		Body length			
A		0.038	0.048	0.147	0.16
G	0.035		0.050	0.149	0.15
G* = G*0.5	0.071		0.050	0.149	0.26

Criticism

- Is this just a problem of wrong estimation?
- Large standard error in estimation of h^2
- If we have very little genetic information (individuals are unrelated), how can we estimate heritabilities?
  - Low relationships -> possible bias
  - Bias of heritability depends on the relationship (Ponzoni and James, 1978):

$$E(\hat{t}-t) \simeq \frac{-2(1-t)\left(t+\frac{1-t}{n}\right)\left(t+\frac{1-t}{sn}\right)}{s-1}$$

  - For s=100 couples of n=2 individuals related by 0.001 expected bias of h^2 is -0.26

(My) Conclusion

- Very interesting paper
- They are right that heritability is not missing and that mixed models can estimate it correctly
- I think that using « unrelated » individuals causes them problems in estimation

- I also think that SNP do not completely trace causal variants, but not only because of MAF (small effects, epistasis)