Advanced course 2005 -Faculty

- Boulder (IBG)
  - John Hewitt (host)
  - Jeff Lessem (IT)
- Richmond (VIPBG)
  - Mike Neale
  - Hermine Maes
- Amsterdam (VU)
  - Dorret Boomsma
  - Danielle Posthuma
  - Meike Bartels
- Brisbane (QIMR)
  - Nick Martin
  - Sarah Medland
  - Manuel Ferreira
- Oxford (WTCHG)
  - Lon Cardon (director)
  - Stacey Cherny
  - Andrew Morris
  - David Evans
  - Jonathan Flint
- Ann Arbor (U Mich)
  - Gonzalo Abecasis
- London (IoP)
  - Pak Sham
  - Ben Neale
- Boston (Harvard)
  - Shaun Purcell

	Year	Location	Type	#Faculty	# Students
TC1	1987	Leuven	Introductory	10	24
TC2	1989	Leuven	Introductory	11	41
TC3	1990	Boulder	Introductory	11	28
TC4	1991	Leuven	Introductory	14	49
			Advanced	12	55
TC5	1993	Boulder	Introductory	13	49
TC6	1994	Boulder	Introductory	16	43
TC7	1995	Helsinki	Introductory	10	29
TC8	1996	Boulder	Introductory	10	49
TC9	1997	Boulder	Introductory	10	55
TC10	1998	Boulder	Introductory	12	57
TC11	1998	Leuven	Introductory	10	55
			Advanced	13	62
TC12	1999	Boulder	Advanced	12	37
TC13	2000	Boulder	Introductory	12	63
TC14	2001	Boulder	Advanced	18	65
TC15	2002	Boulder	Introductory	18	95
TC16	2003	Boulder	Advanced	15	81
TC17	2004	Boulder	Introductory	16	93

Attendance at International Workshops on Methodology of Twin and Family Studies

Frequency of Attendance	1	2	3	4	5	6	7	8	9	10	16	18	19	
Faculty	9	4	2	4	3	2	2	2	3	1	2	2	4	**40**
Student	502	147	34	12	6	2	1							**704**

Introductory	**730**
Advanced	**300**
Total	**1030**

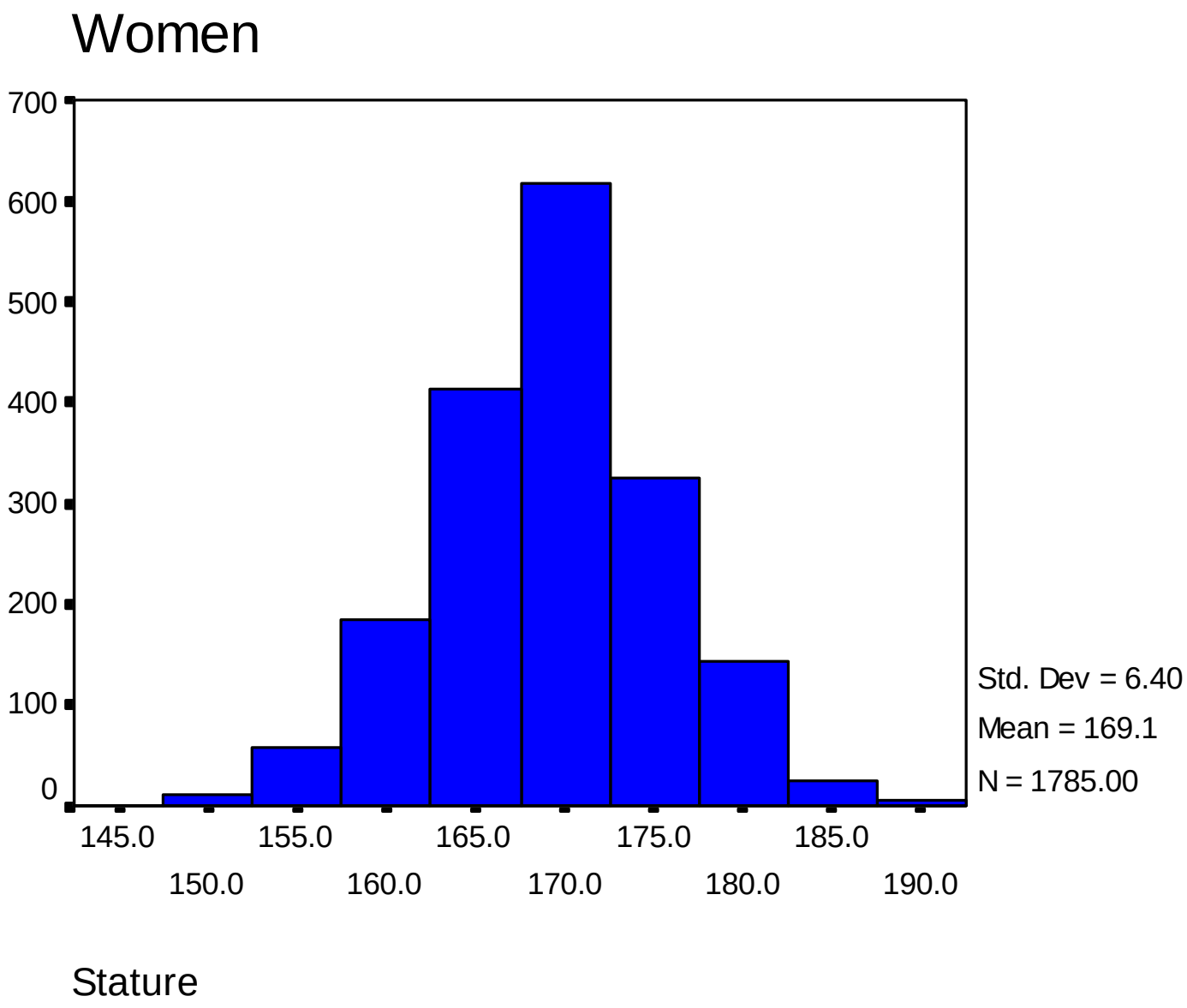
Hunting QTLs

Nick Martin

Queensland Institute of Medical Research

Boulder workshop: March 7, 2005

Stature in Dutch adolescent twins



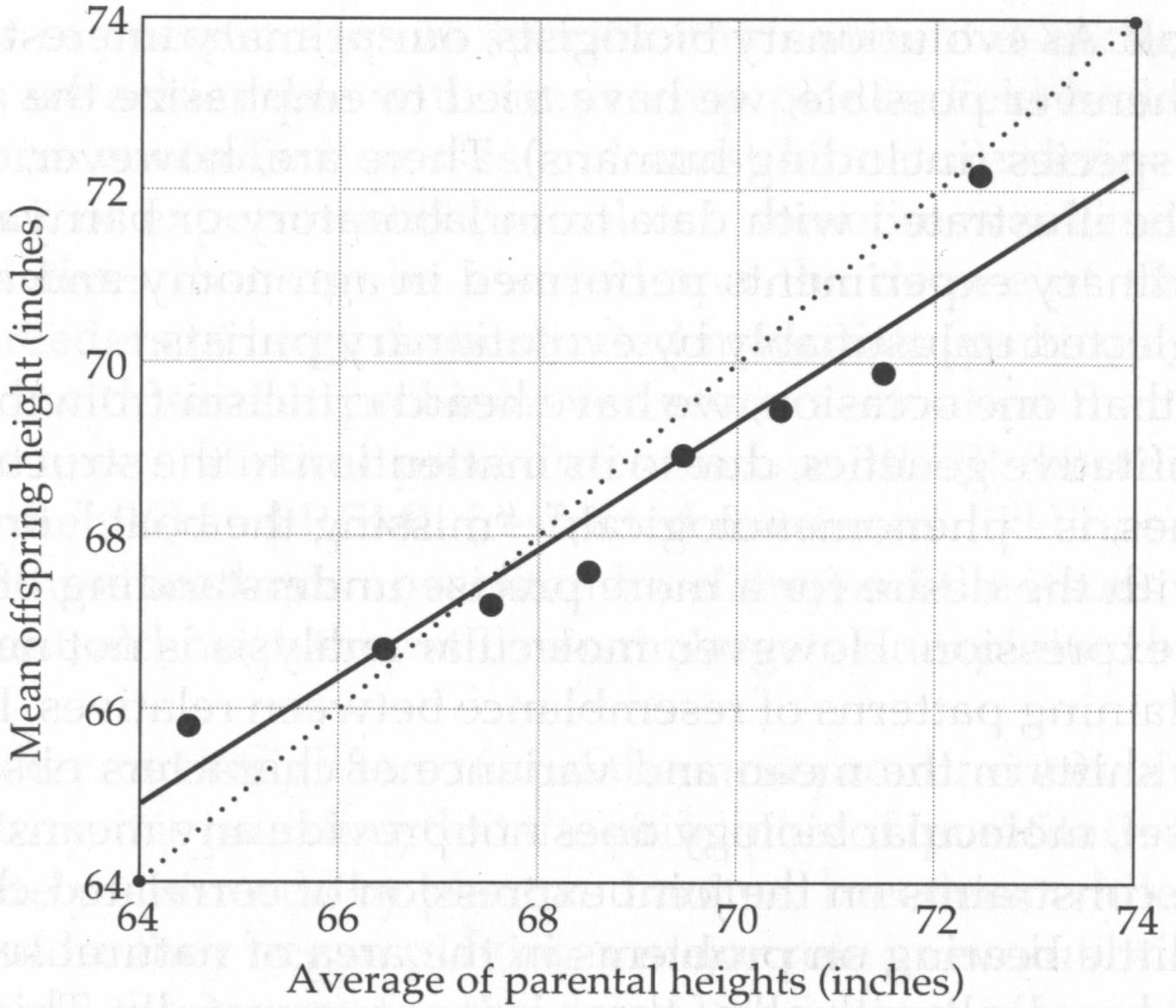
Francis Galton

Sir Francis Galton F.R.S. 1822-1911

Victorian polymath: geographer, meteorologist, tropical explorer, founder of differential psychology, inventor of fingerprint identification, pioneer of statistical correlation and regression, convinced hereditarian, eugenicist, proto-geneticist, half-cousin of Charles Darwin and best-selling author.

I have no patience with the hypothesis occasionally expressed, and often implied, especially in tales written to teach children to be good, that babies are born pretty much alike, and that the sole agencies in creating differences between boy and boy, and man and man, are steady application and moral effort. It is in the most unqualified manner that I object to pretensions of natural equality. The experiences of the nursery, the school, the University, and of professional careers, are a chain of proofs to the contrary.

-- Francis Galton, *Hereditary Genius*


74
72
70
68
66
64
Mean offspring height (inches)
64
66
68
70
72
74
Average of parental heights (inches)

[Galton, 1889]

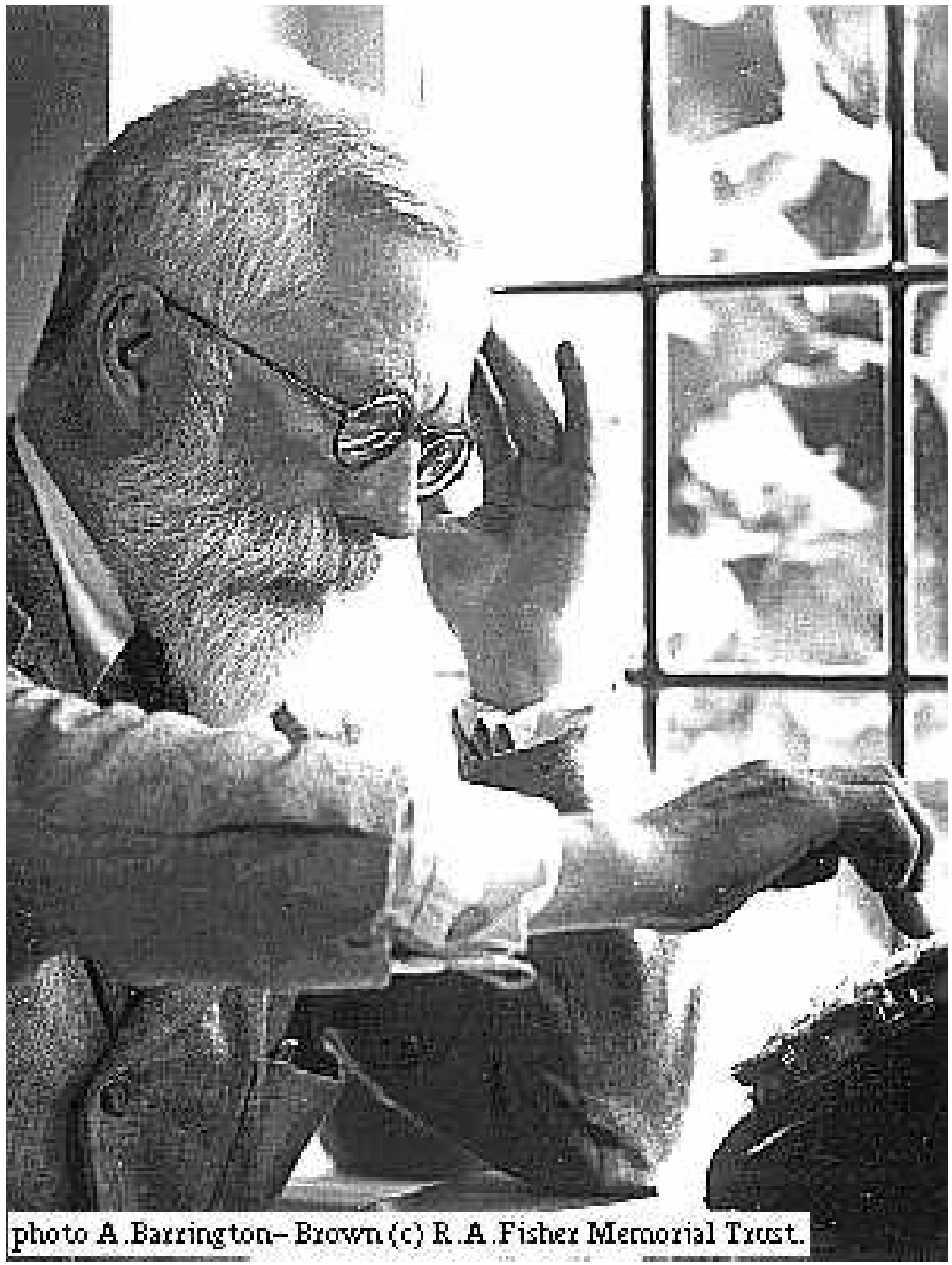
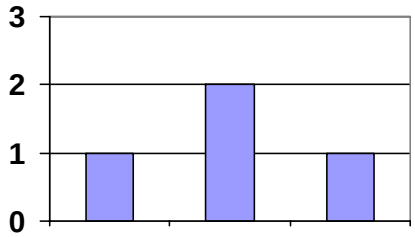
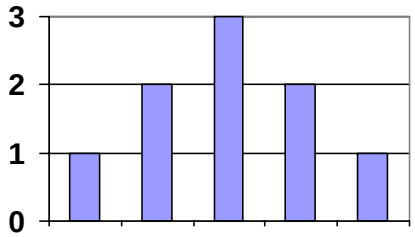


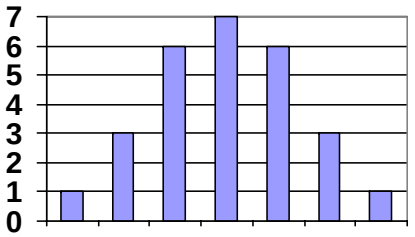
photo A.Barrington-Brown (c) R.A.Fisher Memorial Trust.

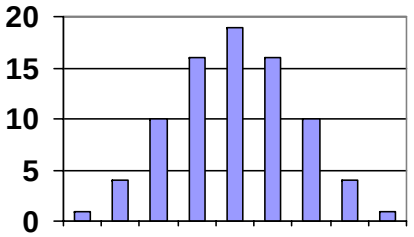
Polygenic Traits

1 Gene	2 Genes	3 Genes	4 Genes
→ 3 Genotypes → 3 Phenotypes	→ 9 Genotypes → 5 Phenotypes	→ 27 Genotypes → 7 Phenotypes	→ 81 Genotypes → 9 Phenotypes



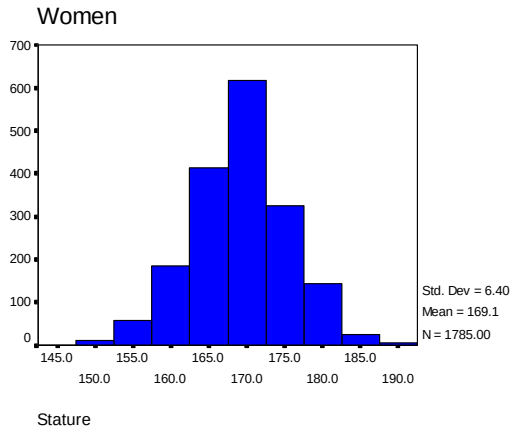


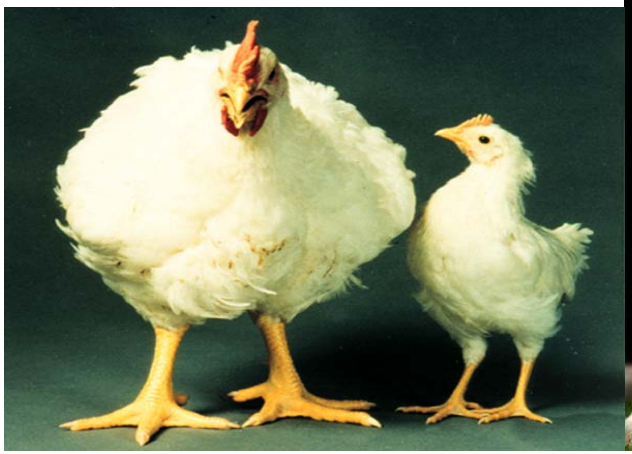




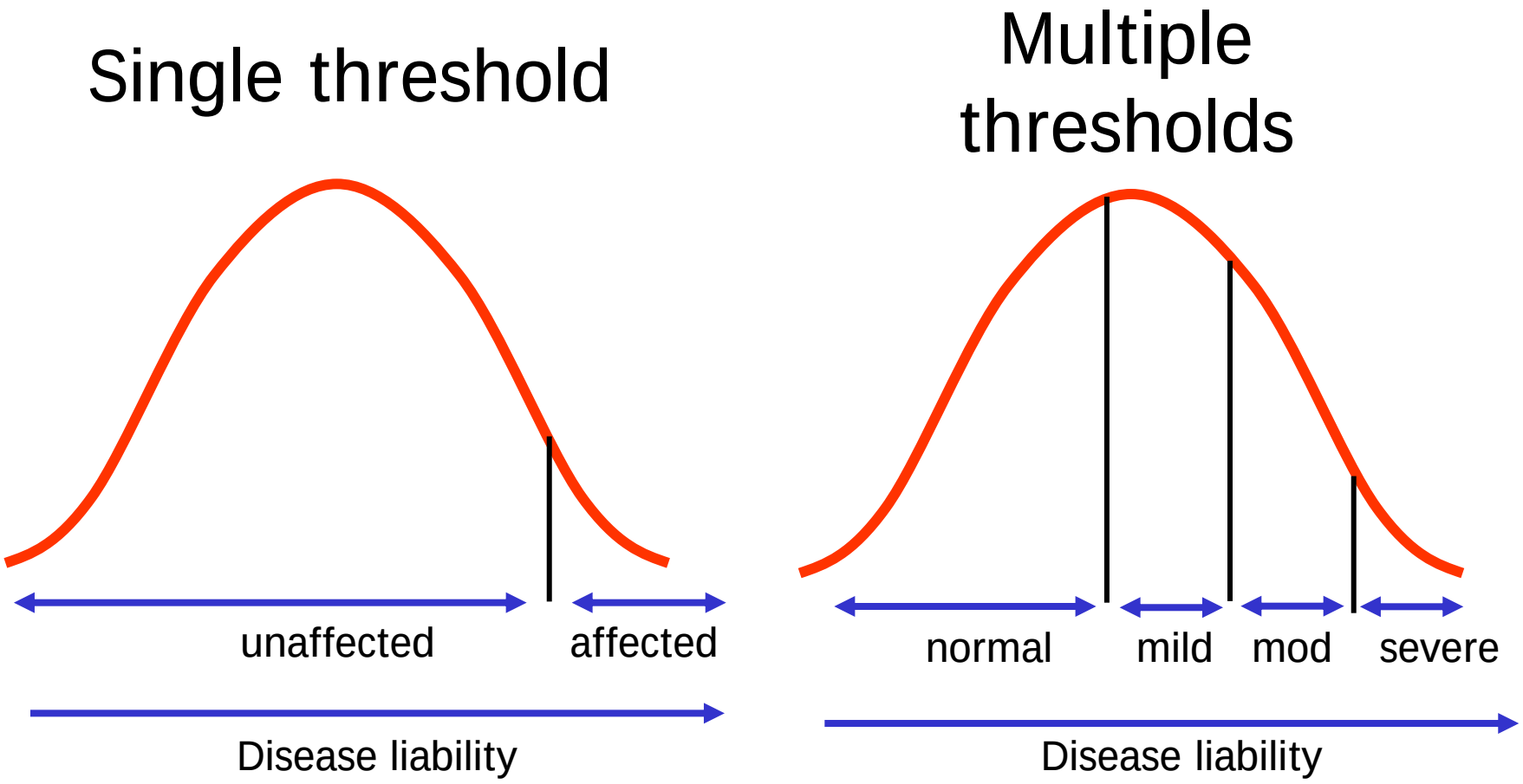
Rationale for QTL analysis

- QTL = quantitative trait locus
- Biology: Understanding genetic variation by dissecting complex traits
  - basic biology
  - applications in agriculture
  - applications in medicine

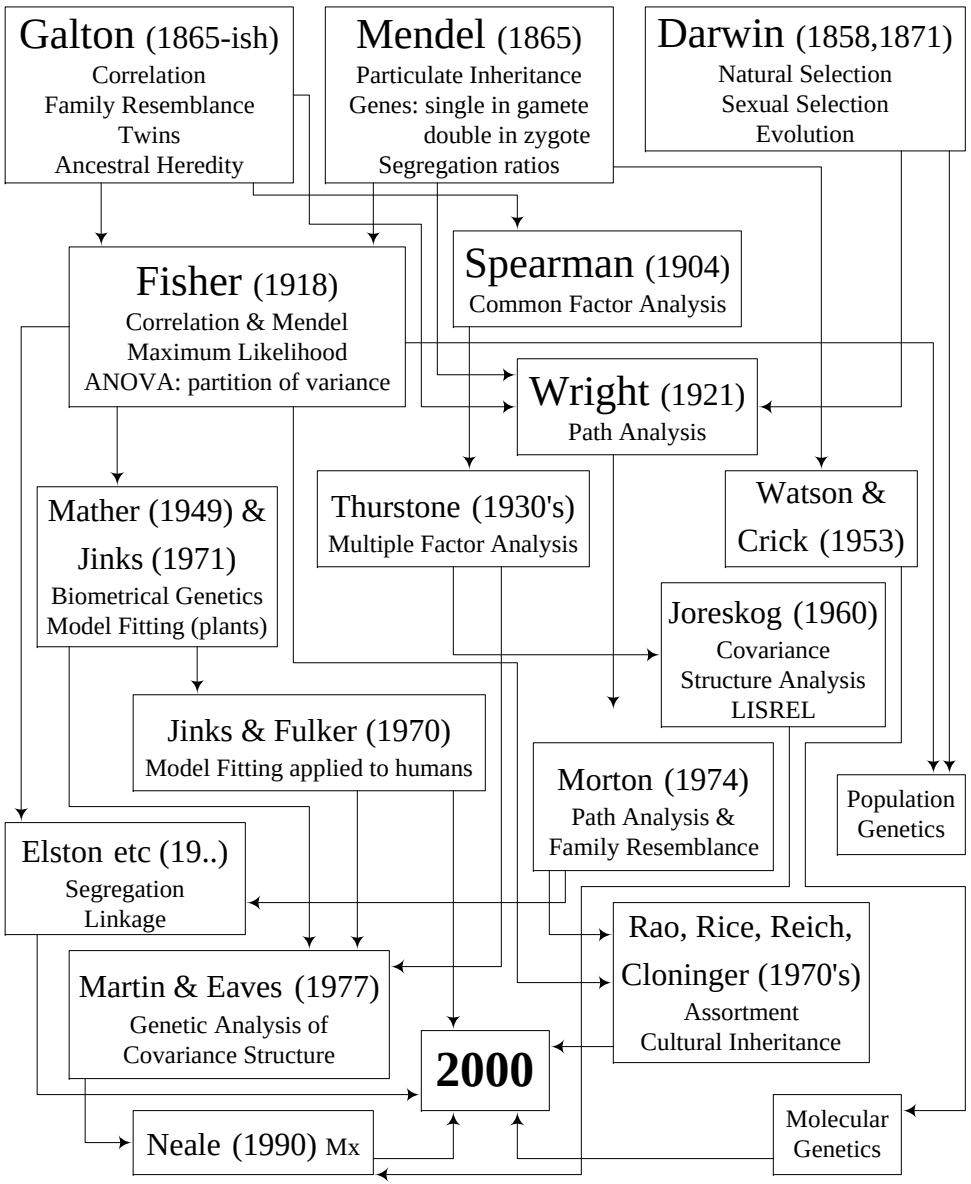




Multifactorial Threshold Model of Disease



People and Ideas



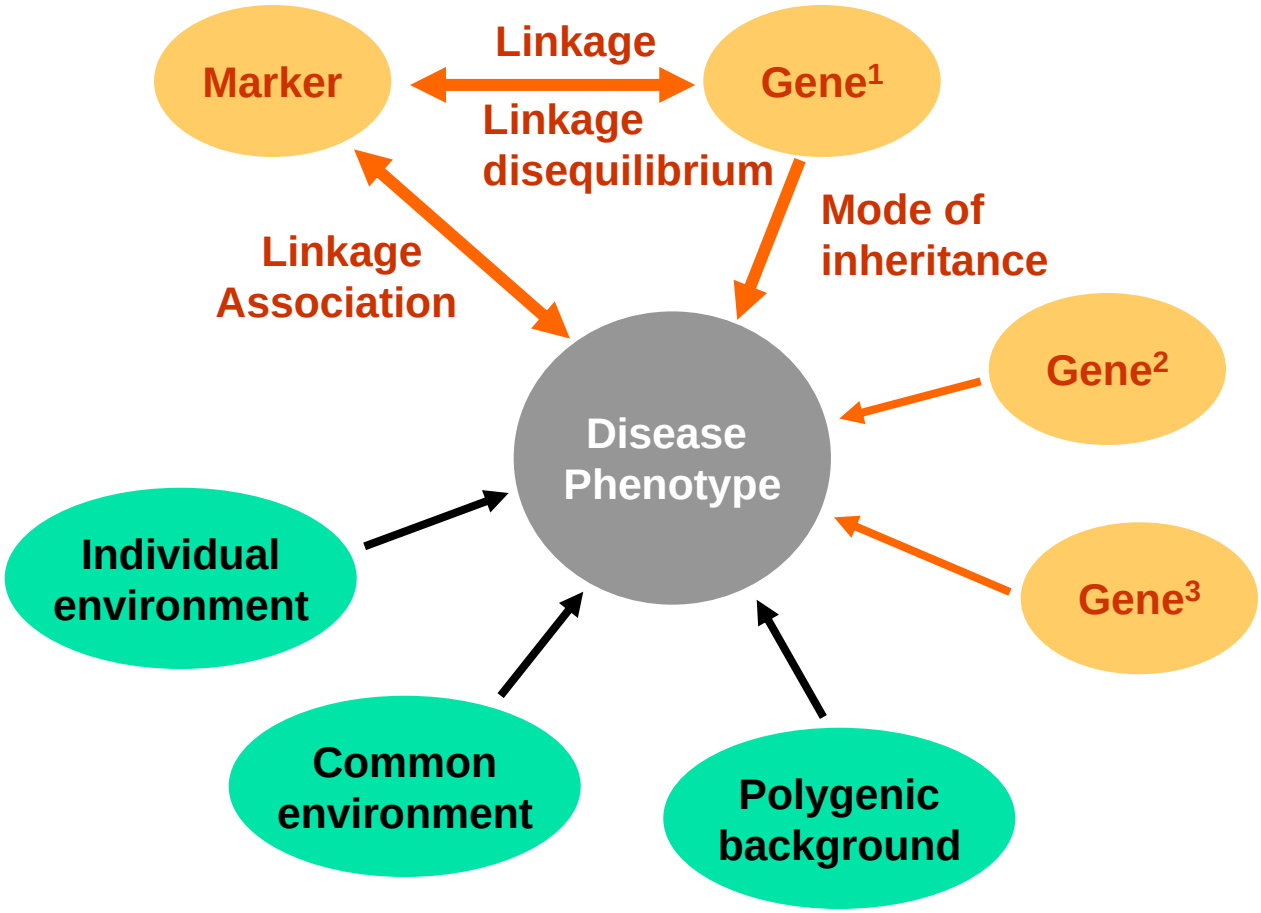
Common diseases

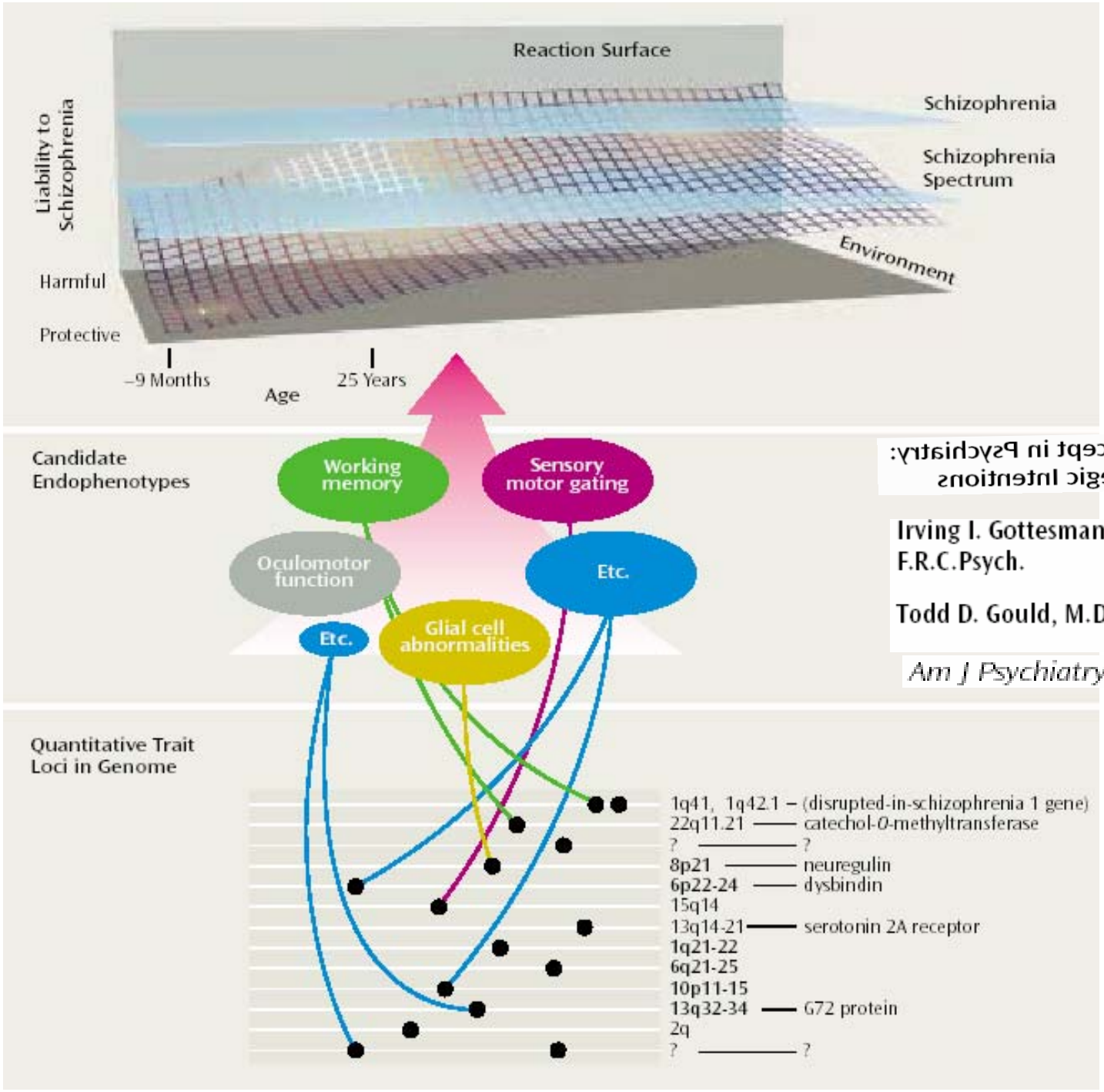
- Estimated life time risk c.60%
- Substantial genetic component
- “Non-Mendelian” inheritance
- Non-genetic risk factors
- Multiple interacting pathways
- Most genes still not mapped

Examples

- Ischaemic heart disease (30-50%, F-M)
- Breast cancer (12%, F)
- Colorectal cancer (5%)
- Recurrent major depression (10%)
- ADHD (5%)
- Non-insulin dependent diabetes (5%)
- Essential hypertension (10-25%)

Complex Trait Model





The Endophenotype Concept in Psychiatry: Etymology and Strategic Intentions

Irving I. Gottesman, Ph.D., Hon. F.R.C.Psych.

Todd D. Gould, M.D.

Am J Psychiatry 160:4, April 2003

Even for “simple” diseases the number of alleles is large

- Ischaemic heart disease (LDR) >190
- Breast cancer (BRAC1) >300
- Colorectal cancer (MLN1) >140

Multiple Rare Alleles Contribute to Low Plasma Levels of HDL Cholesterol

Jonathan C. Cohen,[1,2,3*†] Robert S. Kiss,[5*] Alexander Pertsemlidis,[1] Yves L. Marcel,[5†] Ruth McPherson,[5] Helen H. Hobbs[1,3,4]

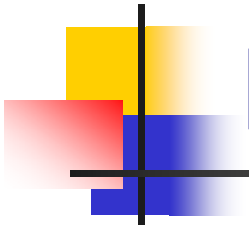
Heritable variation in complex traits is generally considered to be conferred by common DNA sequence polymorphisms. We tested whether rare DNA sequence variants collectively contribute to variation in plasma levels of high-density lipoprotein cholesterol (HDL-C). We sequenced three candidate genes (*ABCA1*, *APOA1*, and *LCAT*) that cause Mendelian forms of low HDL-C levels in individuals from a population-based study. Nonsynonymous sequence variants were significantly more common (16% versus 2%) in individuals with low HDL-C (<fifth percentile) than in those with high HDL-C (>95th percentile). Similar findings were obtained in an independent population, and biochemical studies indicated that most sequence variants in the low HDL-C group were functionally important. Thus, rare alleles with major phenotypic effects contribute significantly to low plasma HDL-C levels in the general population.

Complex disease: common or rare alleles?

Table 1. Sequence variations in the coding regions of *ABCA1*, *APOA1*, and *LCAT*. Values represent the numbers of sequence variants identified in 256 individuals from the Dallas Heart Study (DHS) (128 with low HDL-C and 128 with high HDL-C) and 263 Canadians (155 with low HDL-C and 108 with high HDL-C) (*17*). NS, nonsynonymous (nucleotide substitutions resulting in an amino acid change); S, synonymous (coding sequence substitutions that do not result in an amino acid change). GenBank accession numbers for DHS *ABCA1*, *APOA1*, and *LCAT* sequences are NM_005502, NM_000039, and NM_000229, respectively.

	Sequence variants unique to one group				Sequence variants common to both groups	
	Low HDL-C		High HDL-C			
	NS	S	NS	S	NS	S
			DHS			
ABCA1	14	6	2	5	10	19
APOA1	1	0	0	1	0	1
LCAT	0	1	1	0	1	1
			Canadians			
ABCA1	14	2	2	3	7	5
APOA1	0	1	0	0	2	0
LCAT	6	1	0	0	0	0

[Science 2004]

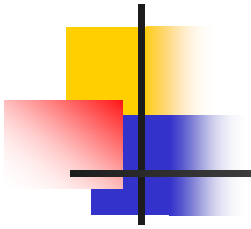


Definitions

- Locus: one of 20-40,000 genes
- Allele: Variant of a specific gene
- Gene: sequence of DNA that codes for a specific function
- Base pair: chemical “letter” of the genome (a gene has many 1000’s of base pairs)
- Genome: all the genes considered together

Defining the Haystack

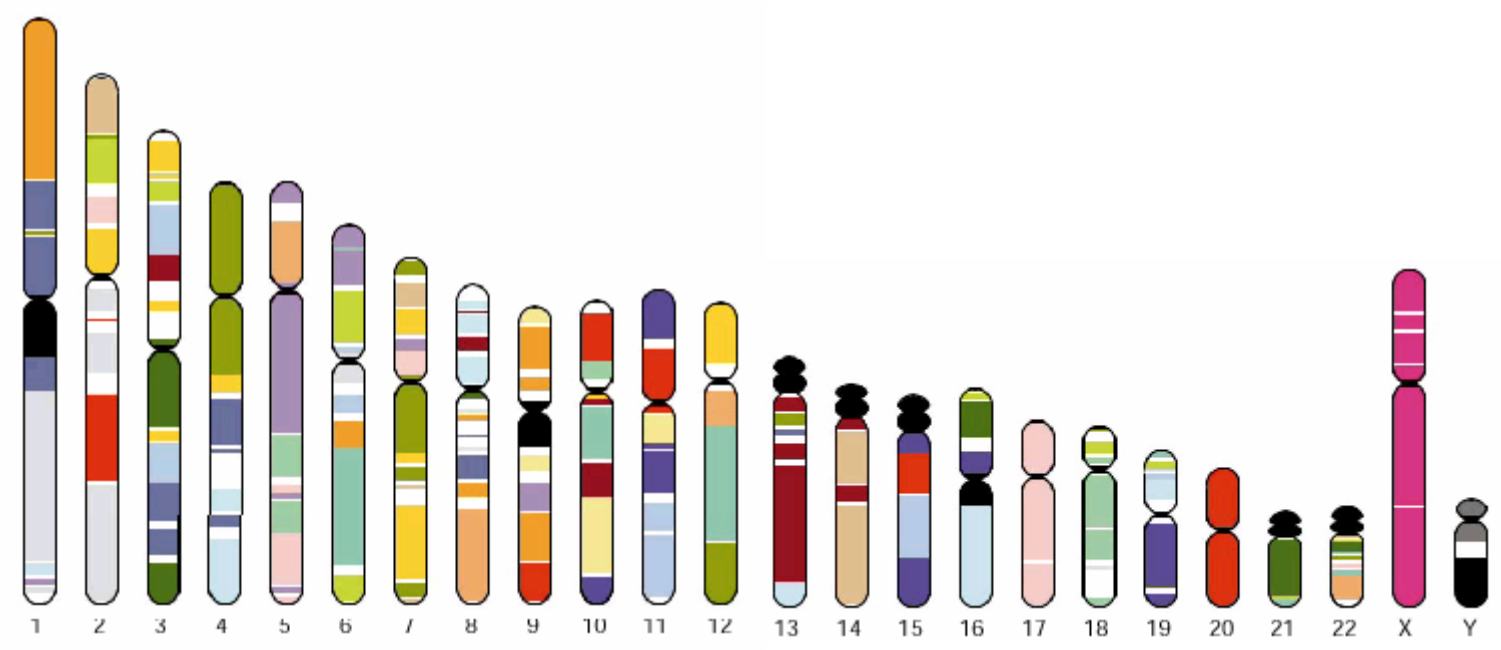
- $3x10^9$ base pairs
- Markers every 6-10kb for association in populations with no recent bottleneck history
- 1 SNPs per 721 b.p. (Wang et al., 1998)
- c.14 SNPs / 10kb = 1000s haplotypes/alleles
- O (10^4 -10^5) genes



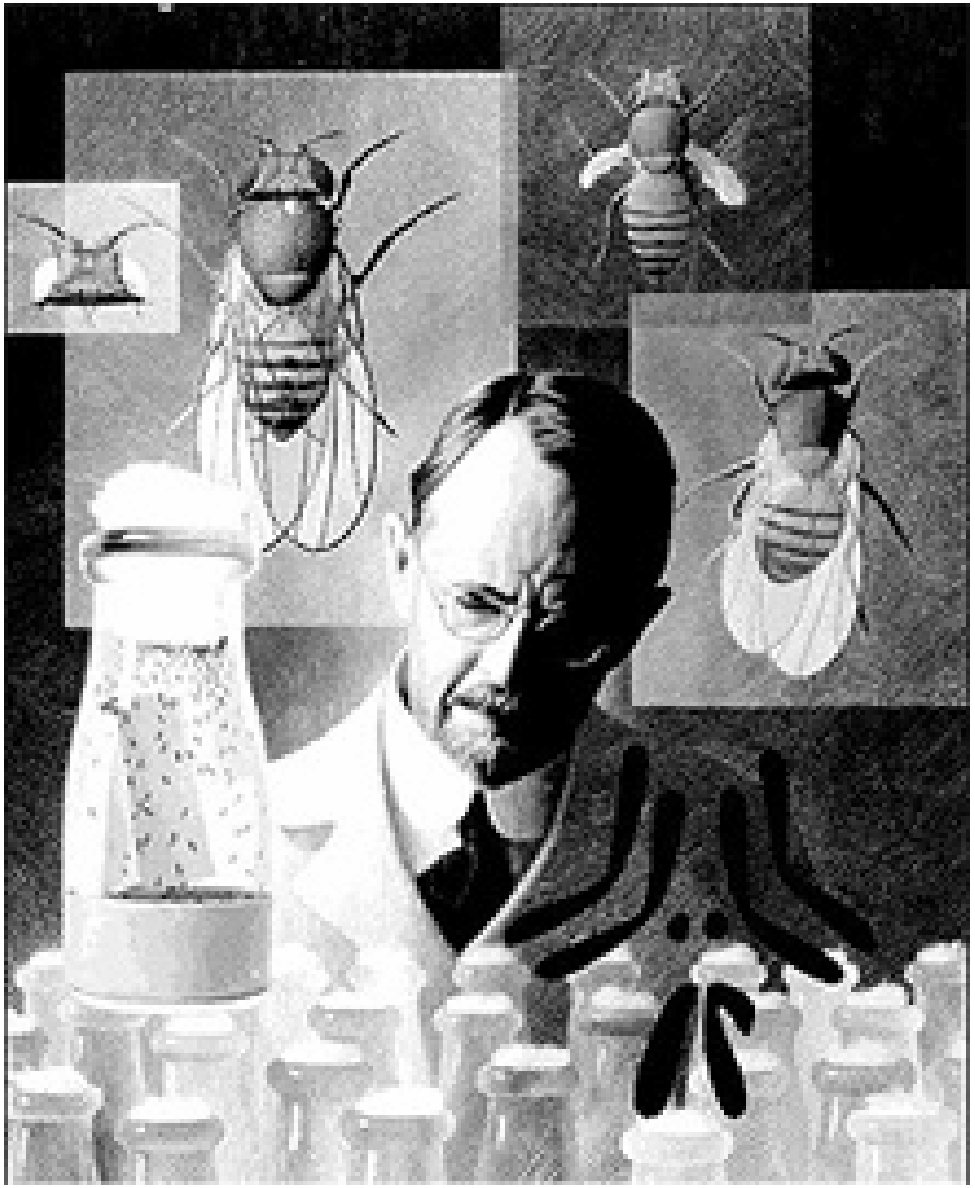
Finding QTLs

- Linkage
- Association

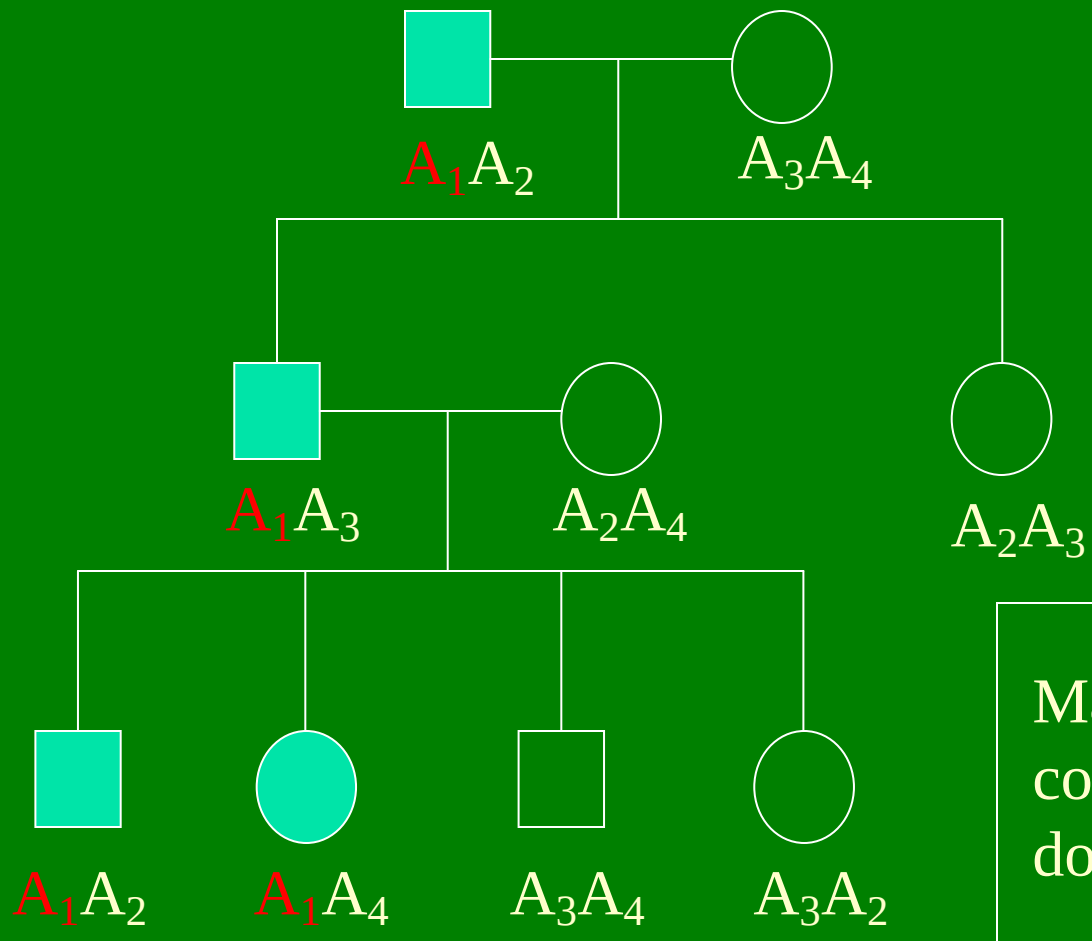
Linkage analysis



Thomas Hunt Morgan – discoverer of linkage



Linkage = Co-segregation



Marker allele A_1 cosegregates with dominant disease

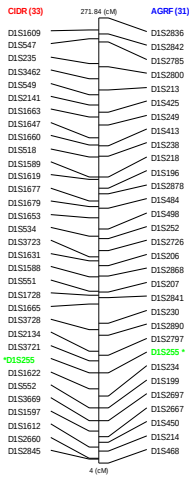
Automated microsatellite genotyping – a major breakthrough (early 90s)

MZ and DZ twins: determining zygosity using ABI Profiler™ genotyping

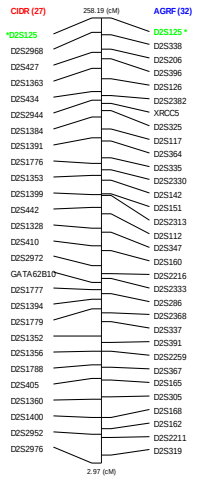
(9 STR markers + sex)

Genotype information for Twin Mole and Twin Maps studies (v23_09)

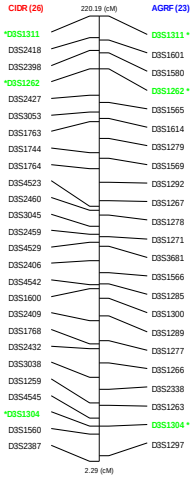
Chromosome 1



Chromosome 2

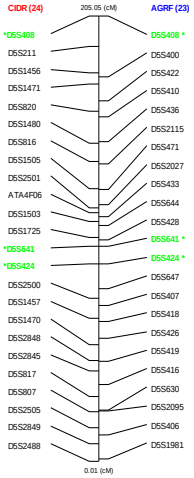


Chromosome 3



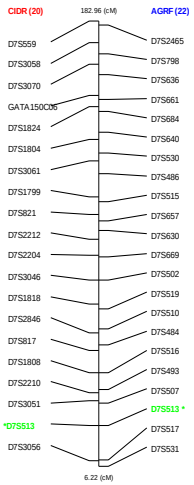
Chromosome 4

Chromosome 5

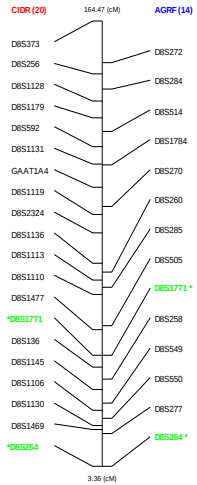


Chromosome 6

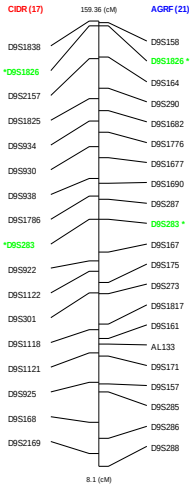
Chromosome 7



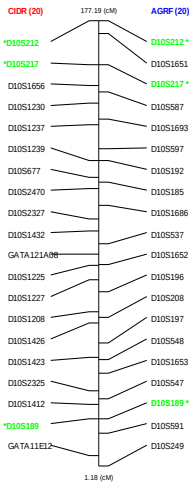
Chromosome 8



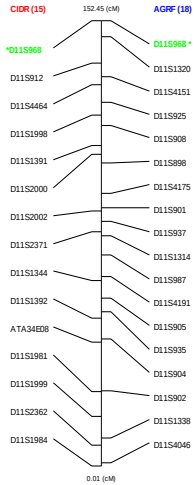
Chromosome 9



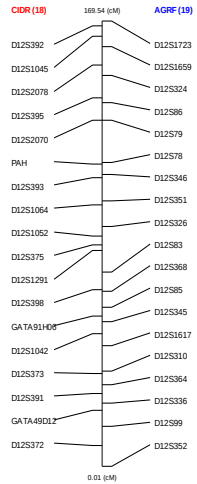
Chromosome 10



Chromosome 11

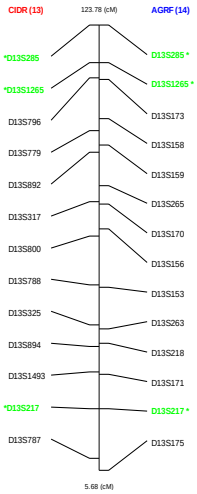


Chromosome 12



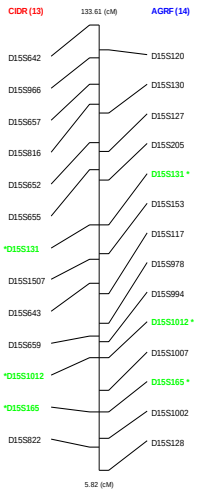
Genotype information for Twin Mole and Twin Maps studies (v23_09)

Chromosome 13

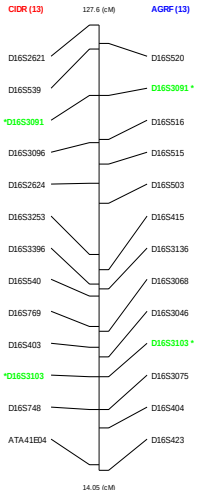


Chromosome 14

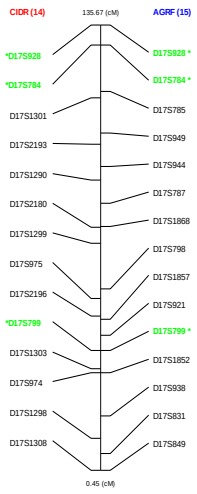
Chromosome 15



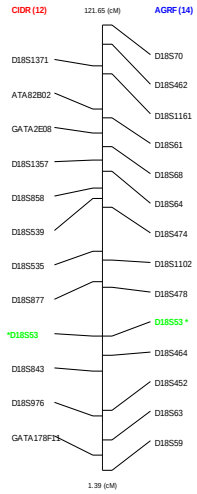
Chromosome 16



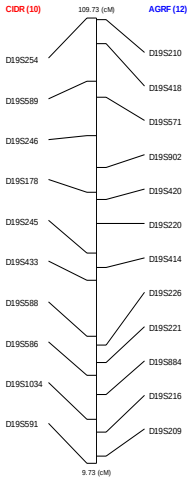
Chromosome 17



Chromosome 18



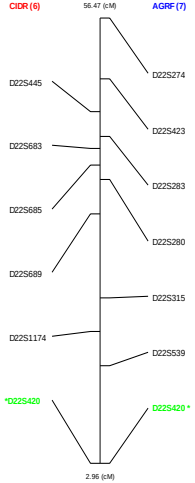
Chromosome 19



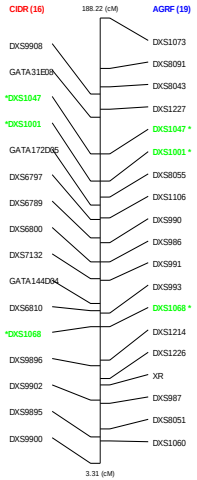
Chromosome 20

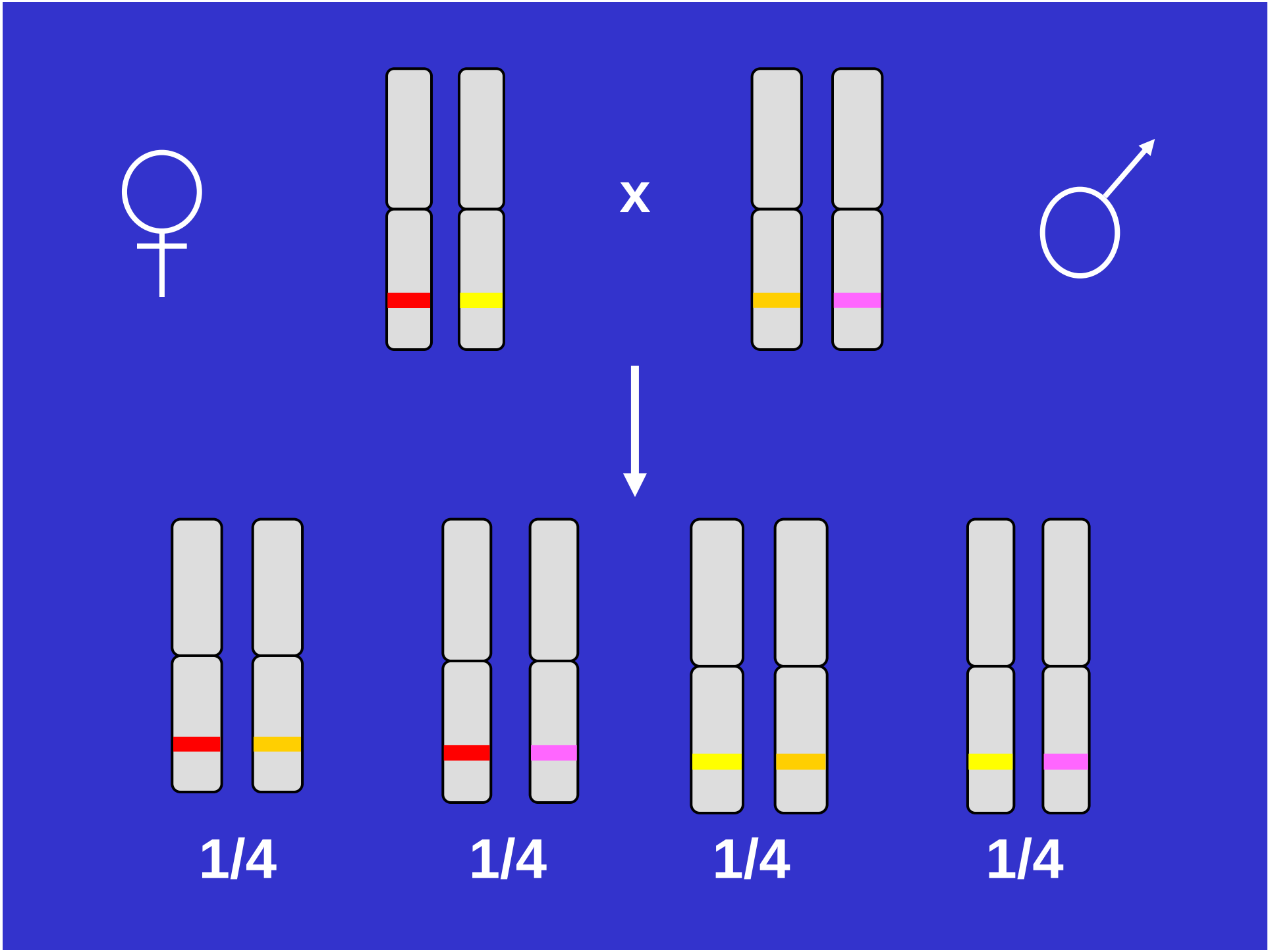
Chromosome 21

Chromosome 22

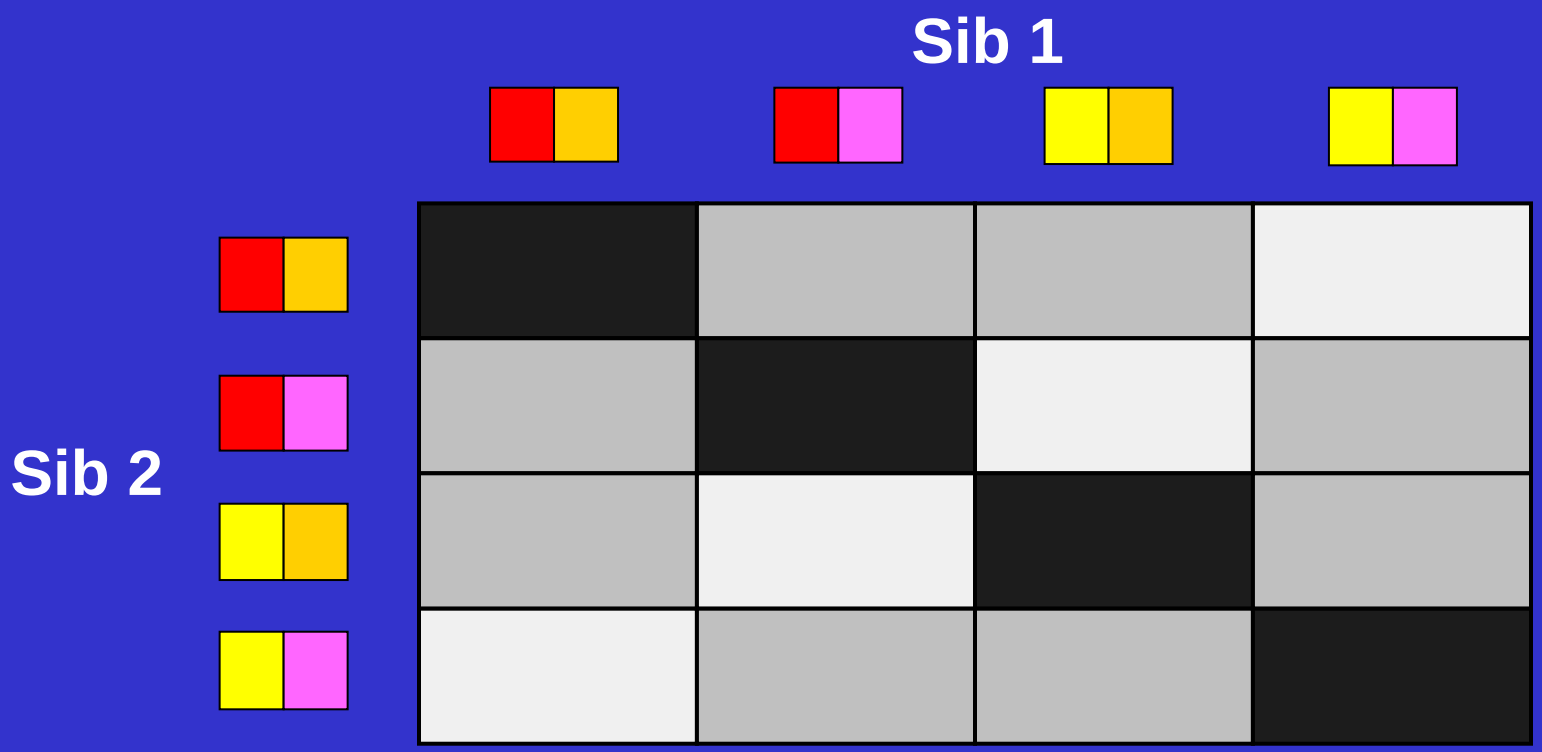


Chromosome X




x
1/4
1/4
1/4
1/4

IDENTITY BY DESCENT



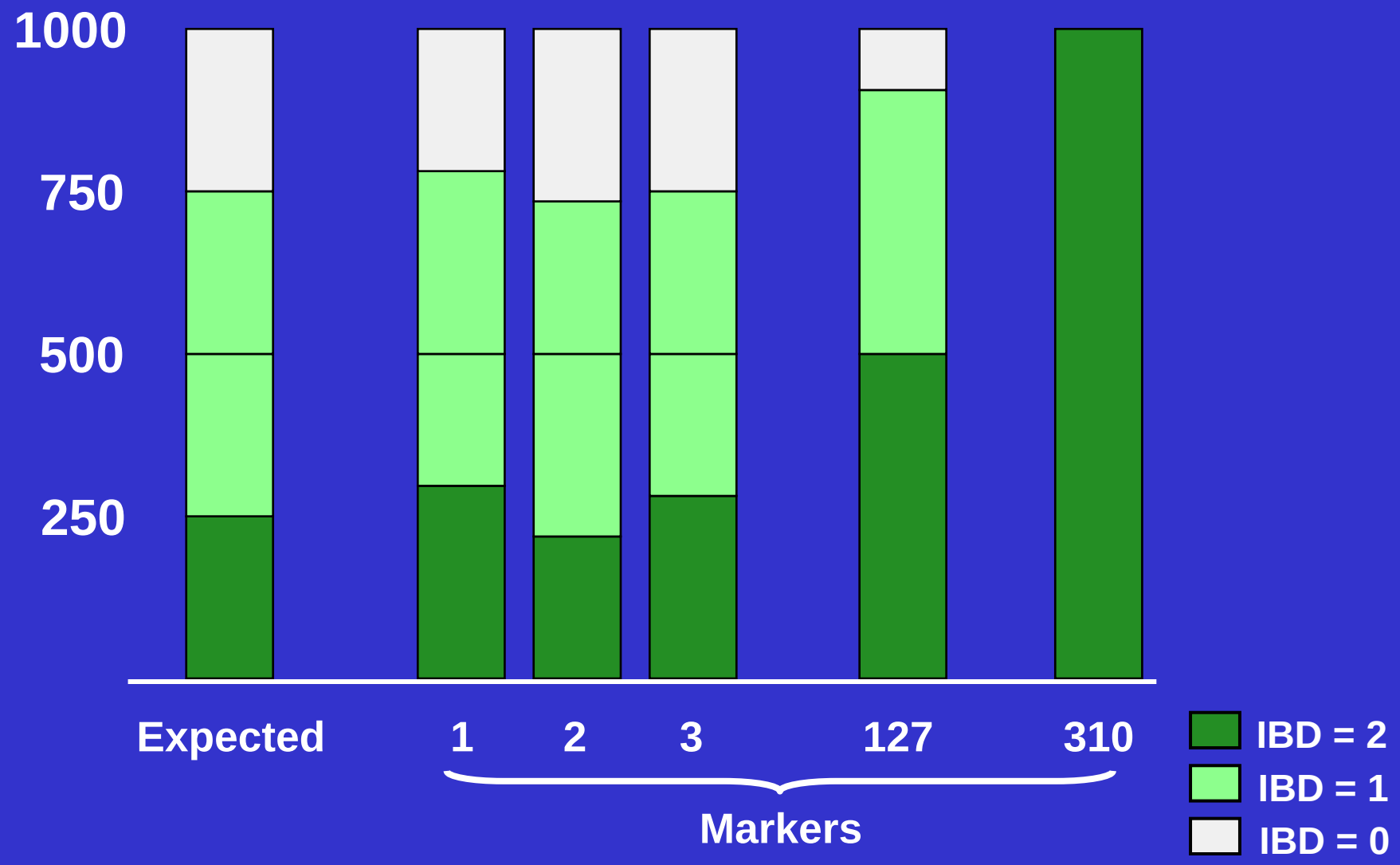
4/16 = 1/4 sibs share BOTH parental alleles IBD = 2

8/16 = 1/2 sibs share ONE parental allele IBD = 1

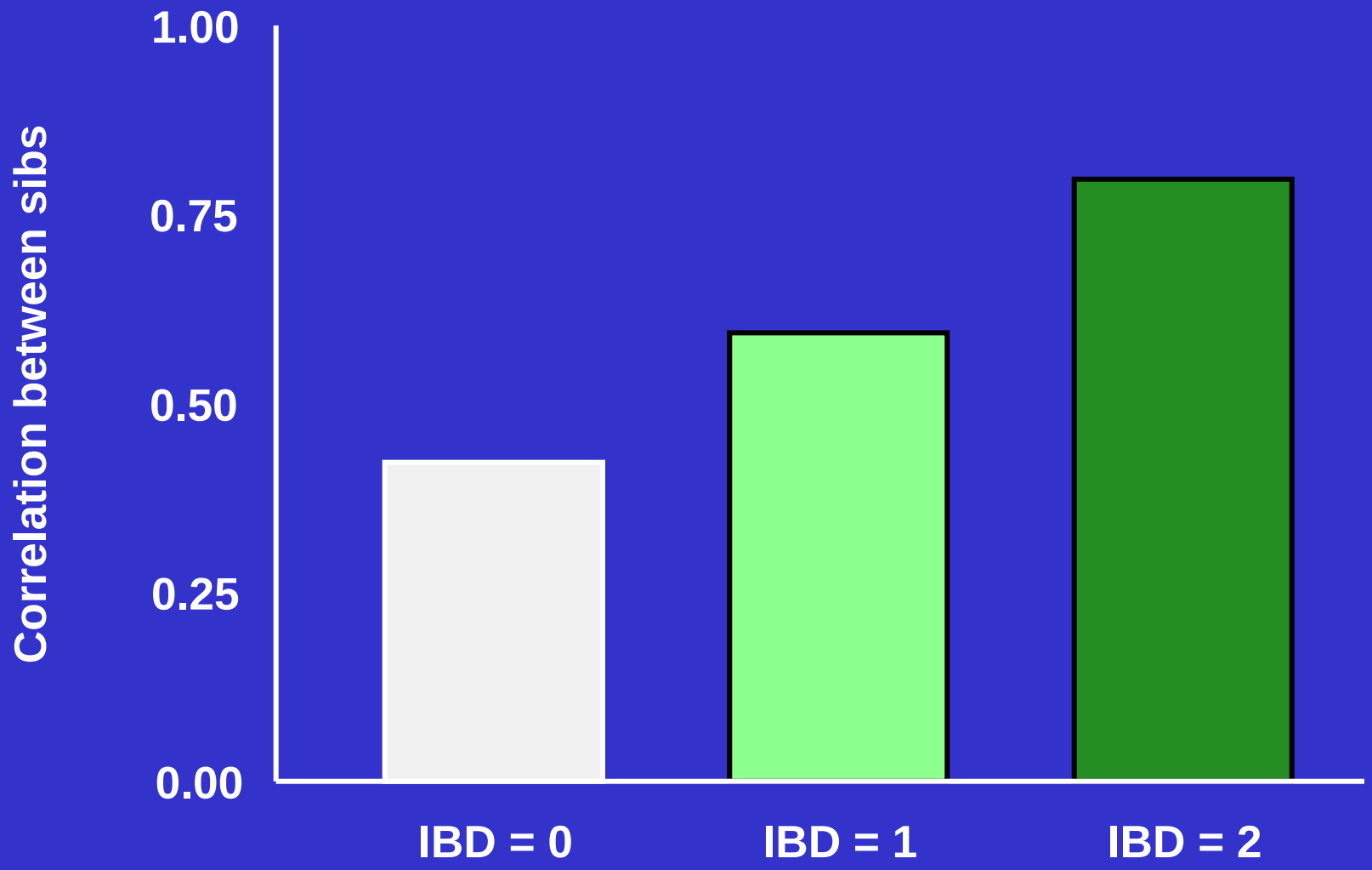
4/16 = 1/4 sibs share NO parental alleles IBD = 0

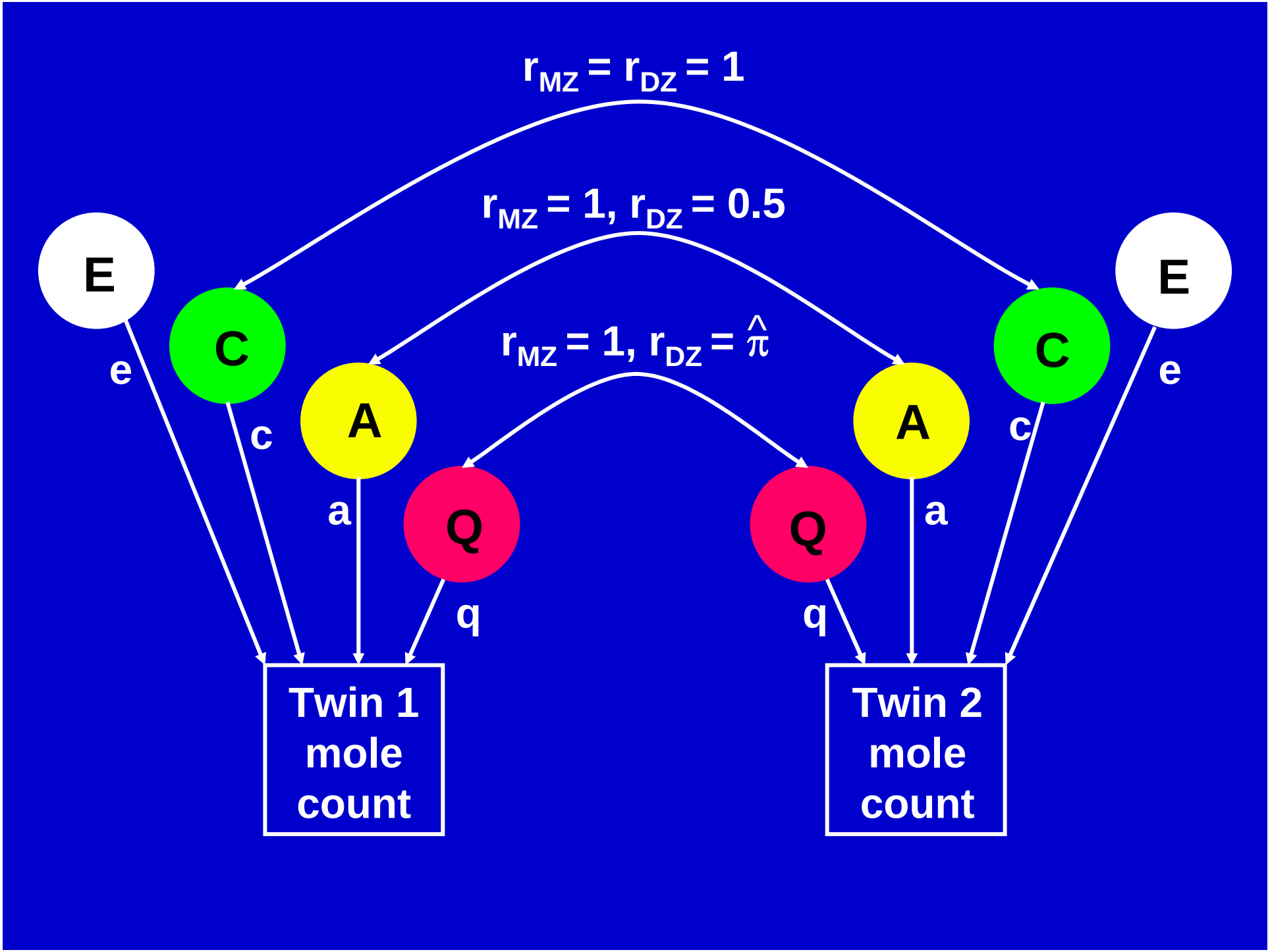
For disease traits (affected/unaffected)

Affected sib pairs selected

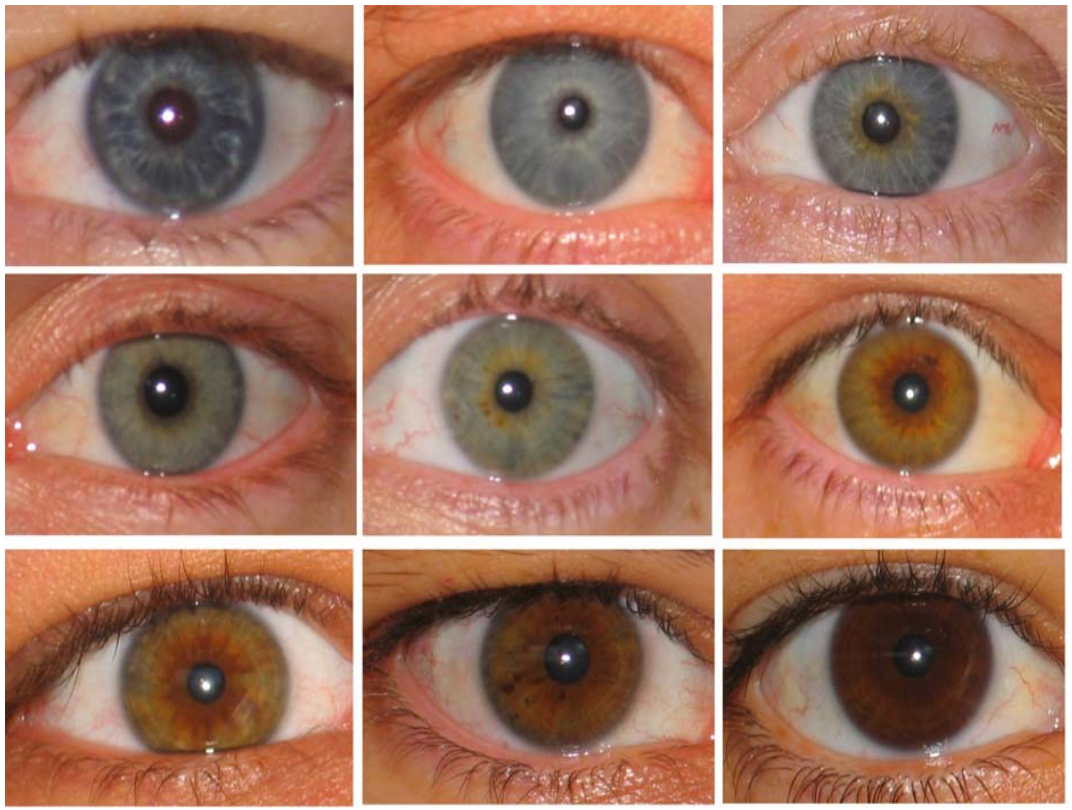


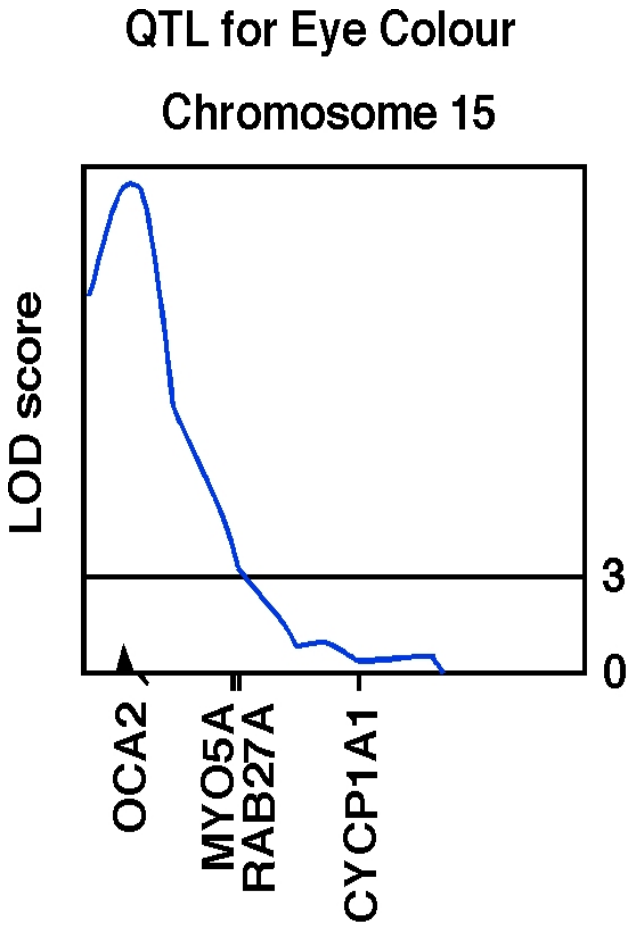
For continuous measures
Unselected sib pairs




$r_{MZ} = r_{DZ} = 1$
$r_{MZ} = 1, r_{DZ} = 0.5$
$r_{MZ} = 1, r_{DZ} = \hat{\pi}$
E
C
A
Q
e
c
a
q
Twin 1 mole count
Twin 2 mole count

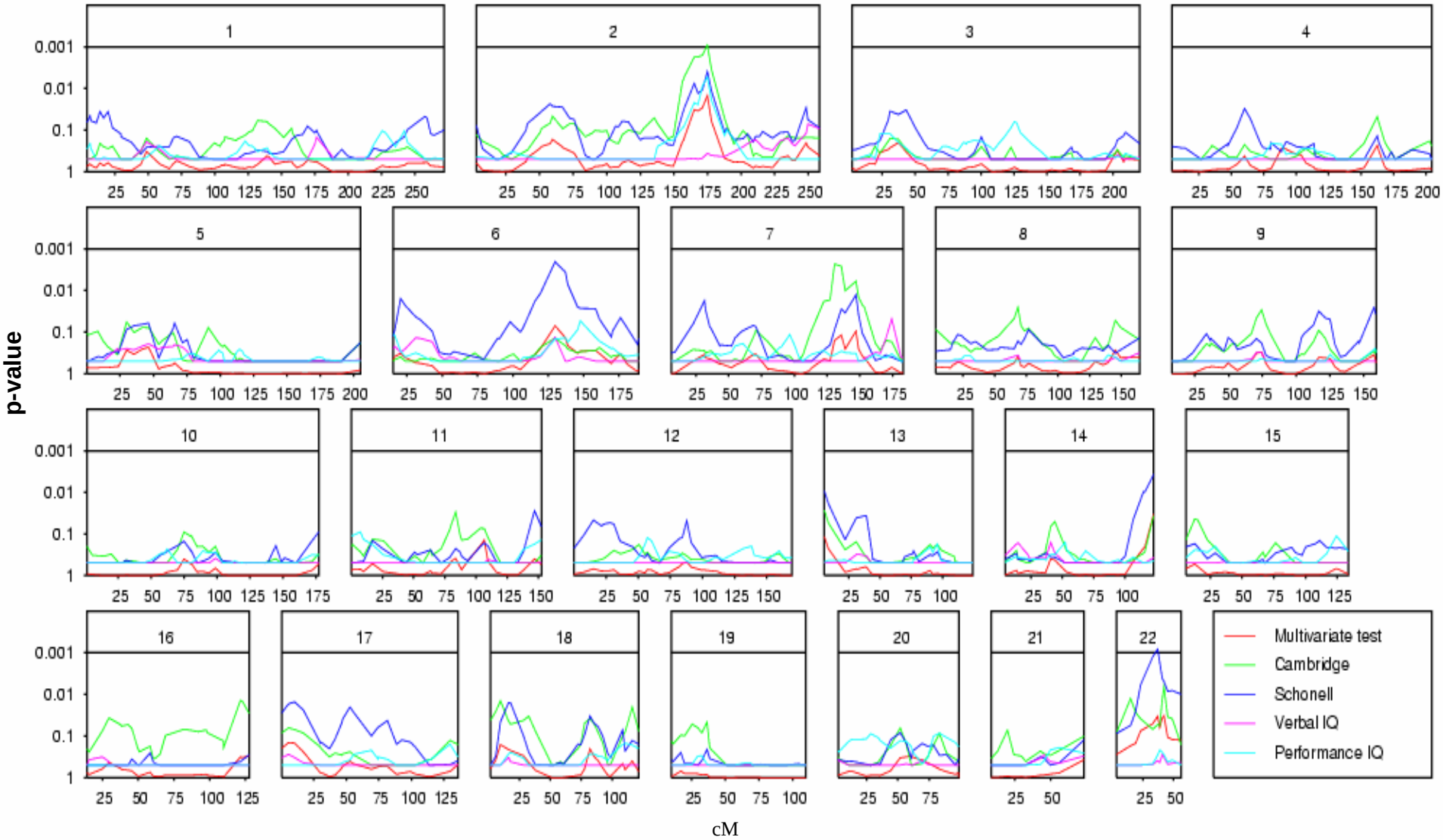
Human OCA2 and eye colour





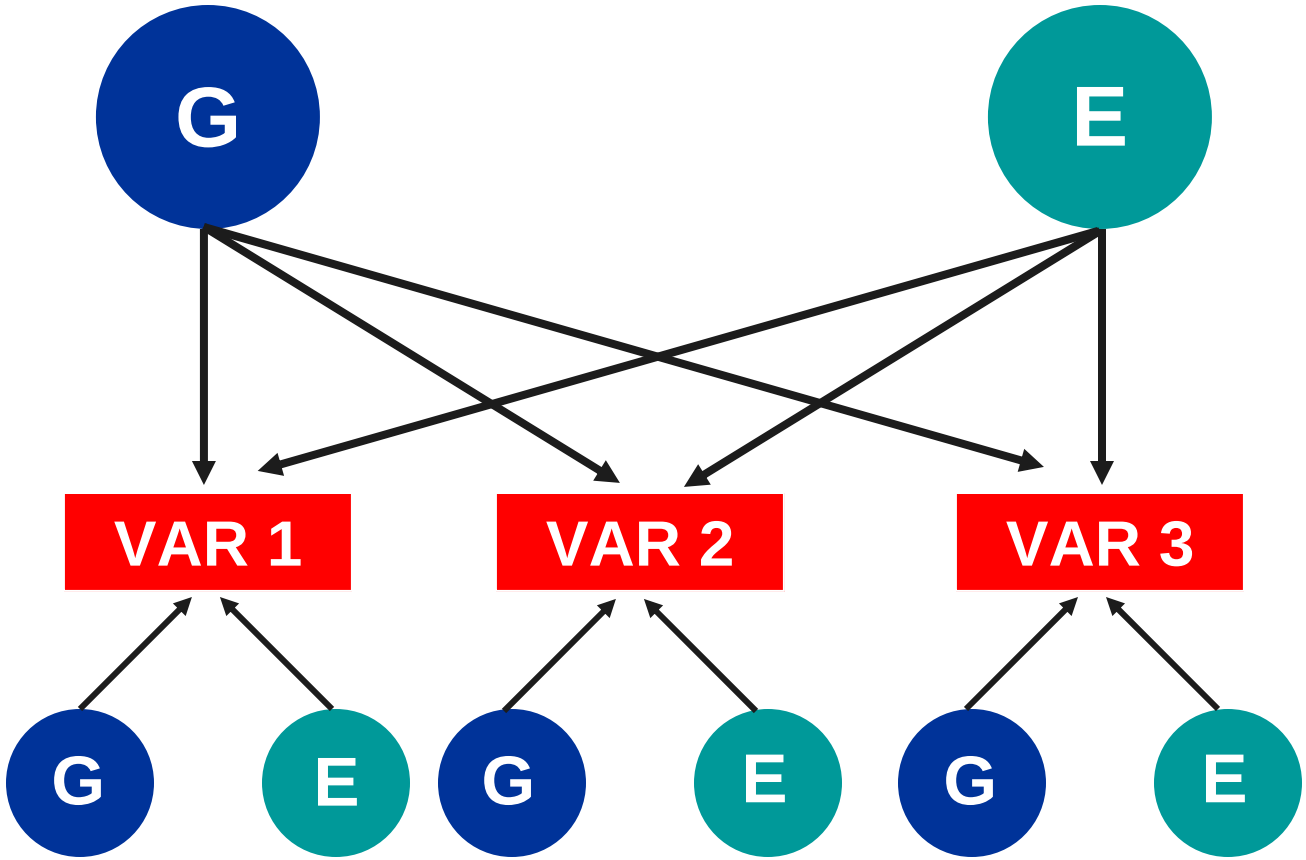
Zhu et al., *Twin Research* 7:197-210 (2004)

Linkage Results for IQ and Reading

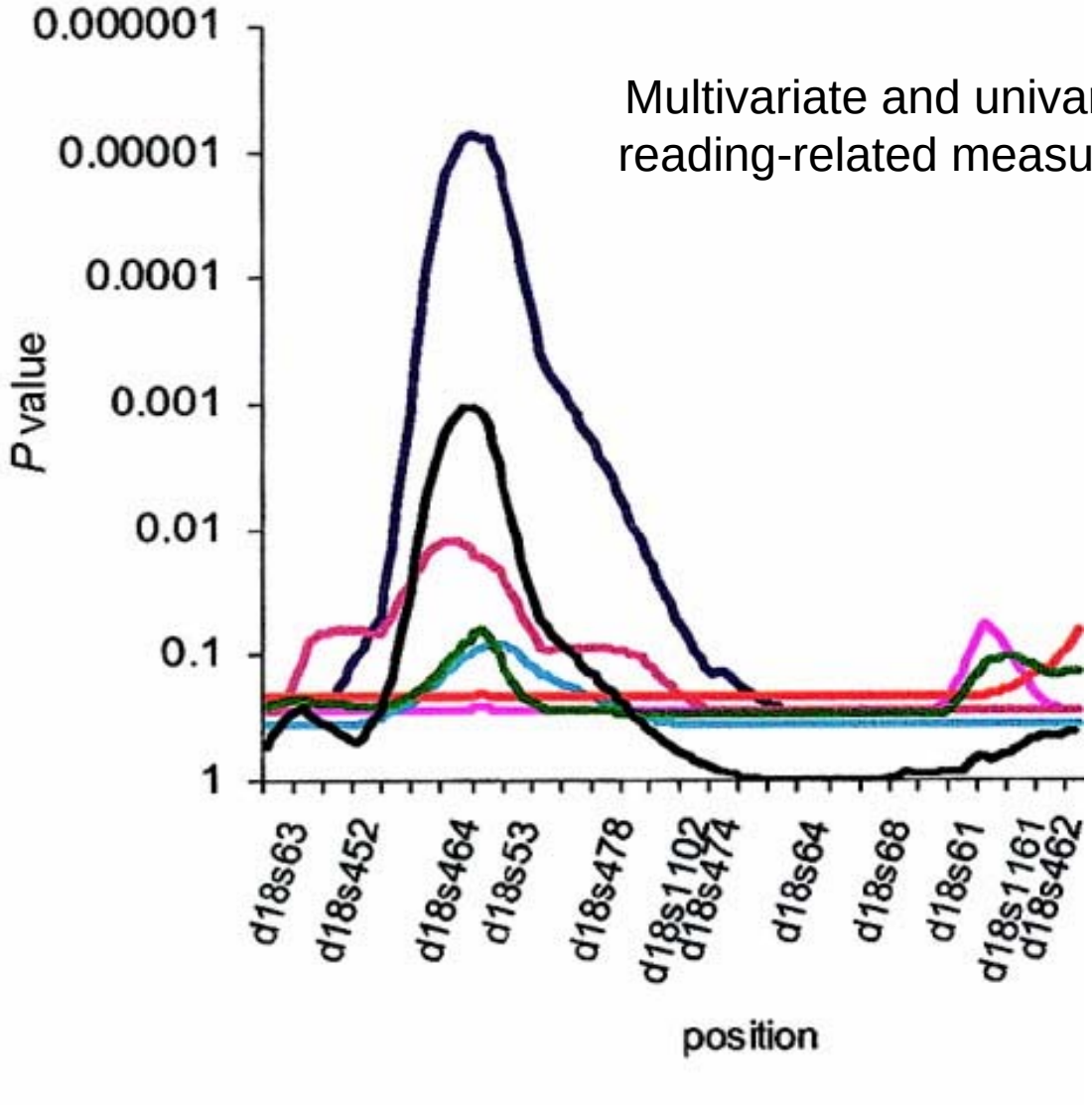


Increasing power for QTL linkage (1)

Multivariate analysis


G
E
VAR 1
VAR 2
VAR 3
G
E
G
E
G
E

Effect of multivariate analysis on linkage power



Multivariate and univariate linkage analysis of six reading-related measures on chromosome 18

Am. J. Hum. Genet., 72:561-570, 2003

Use of Multivariate Linkage Analysis for Dissection of a Complex Cognitive Trait

Angela Marlow, Simon Fisher, Clyde Francks, Laurence MacPhie, Stacey Cherny, Alex Richardson, Joel Talcott, John Stein, Anthony Monaco, and Lon Cardon

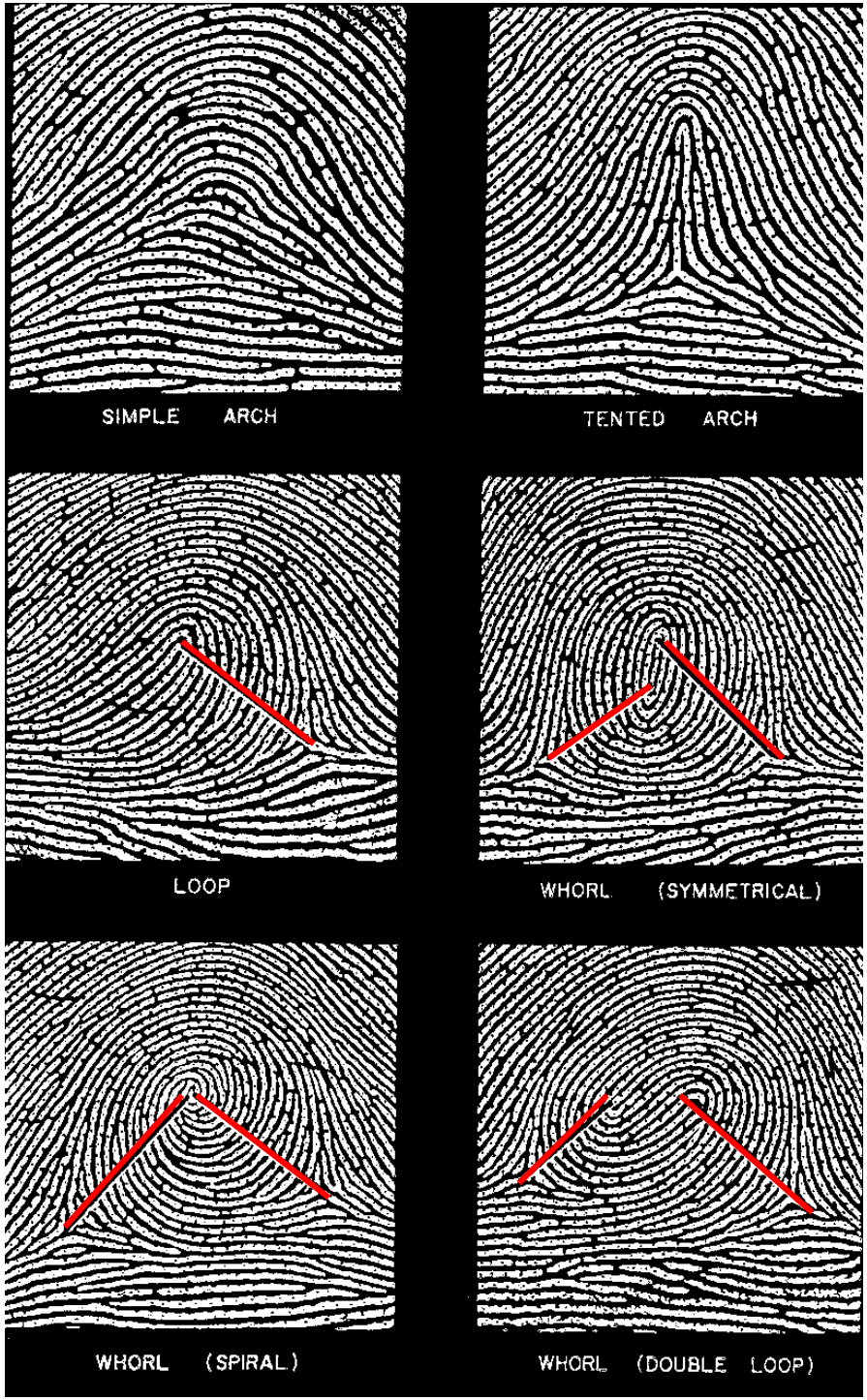
Ridge count

The size of prints can be measured by counting the number of ridges from the triradii to the core

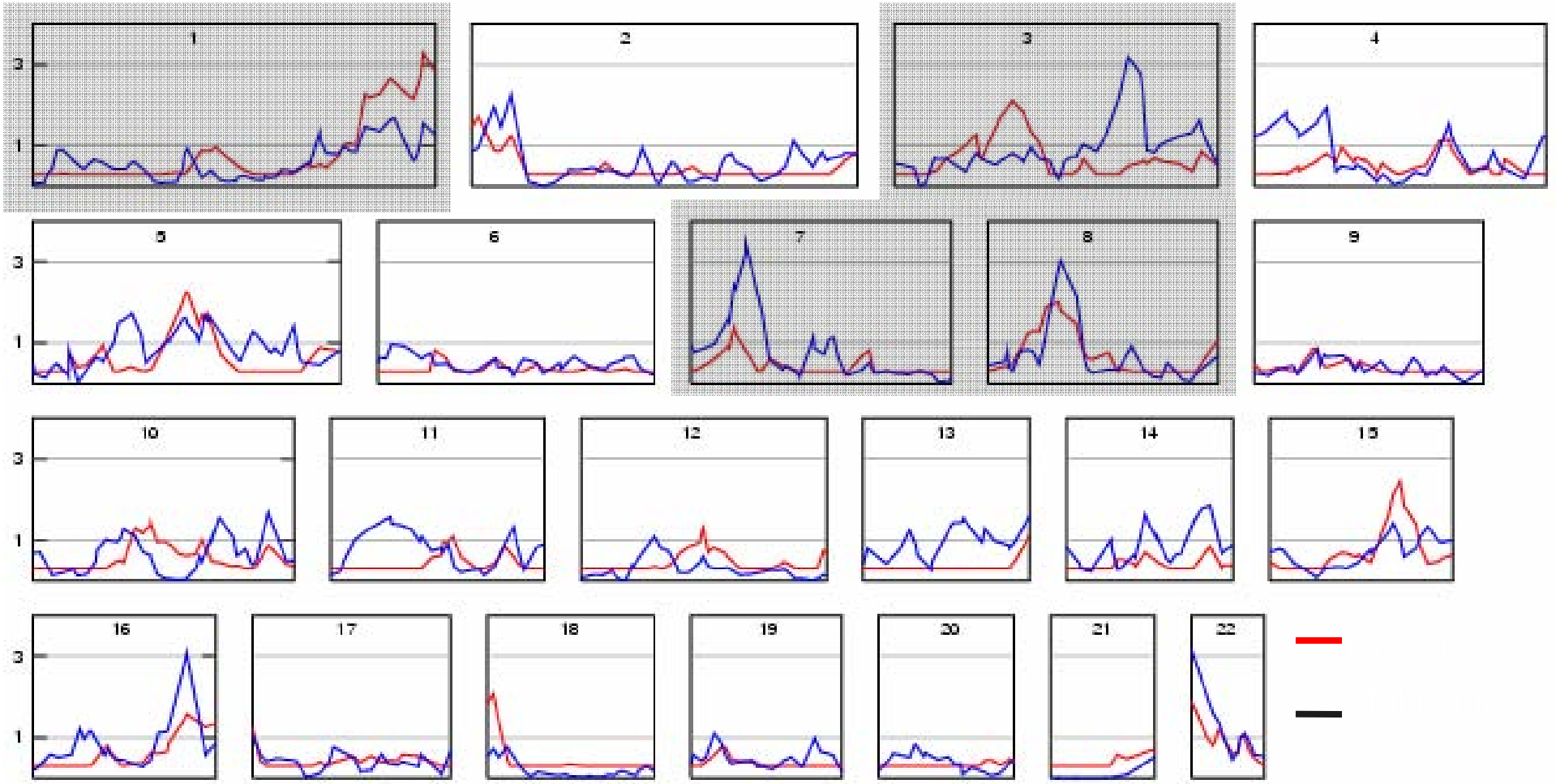
Ridge count can be summed over all fingers to give a total ridge count

Diagram from Holt, 1968

- Highly heritable:
  - MZ r = .94 CI .89 - .96
  - DZ r = .42 CI .34 - .50
  - A .82 CI .56 - .95
  - D .11 CI .00 - .37
  - E .07 CI .05 - .10



TRC vs Multivariate ($-LOG_{10}p$)

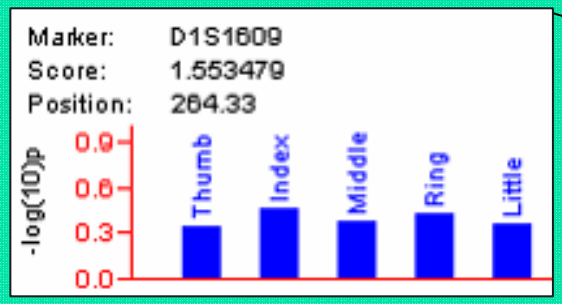


χ_1^2 χ_5^2

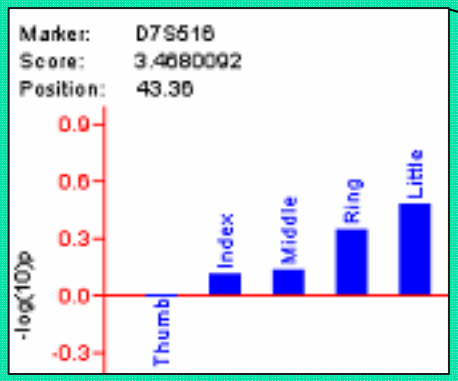
Chromosome 1

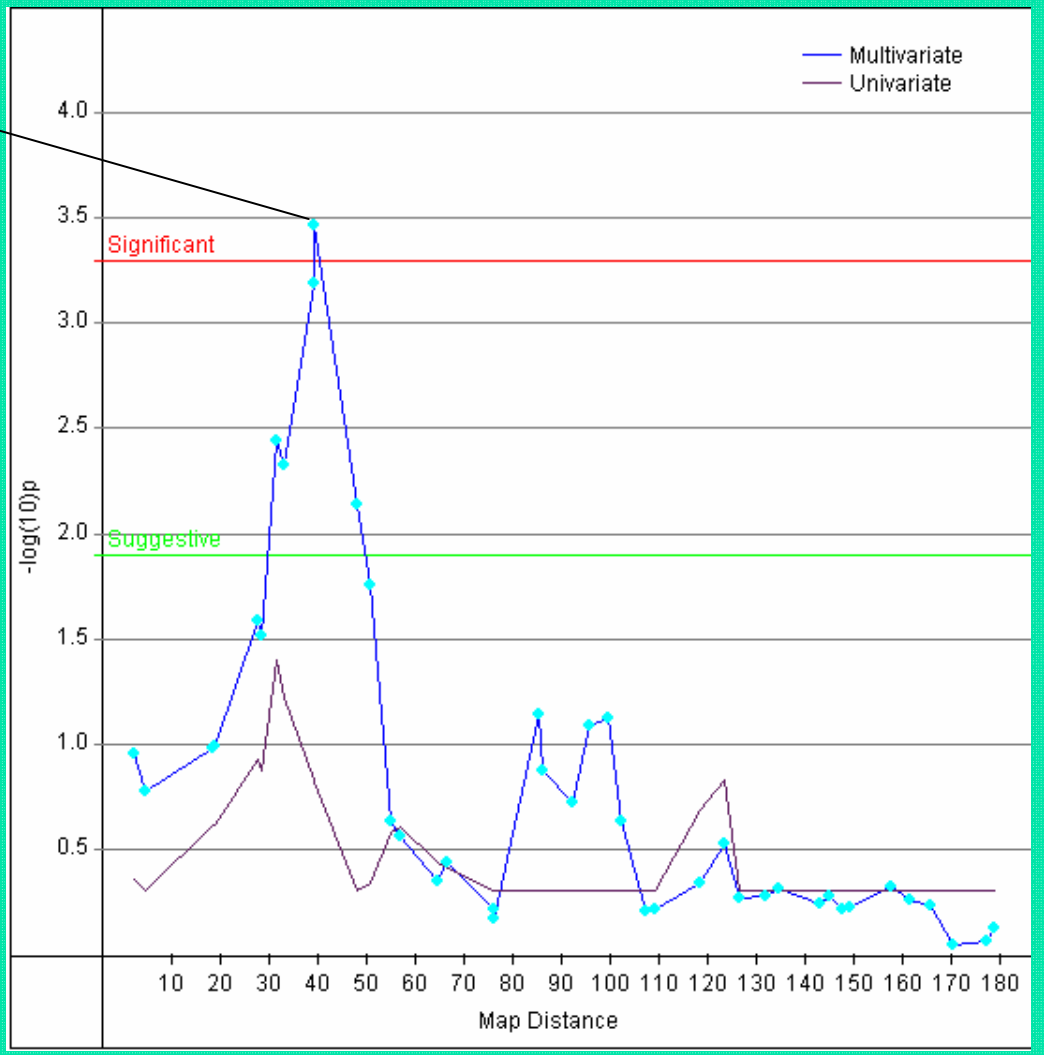
Similar ‘drop chi-squares’ for pleiotropic QTLs

Resulting in a very conservative test



Chromosome 7 ...





Evidence of developmental fields?

[Nature 2004]

articles

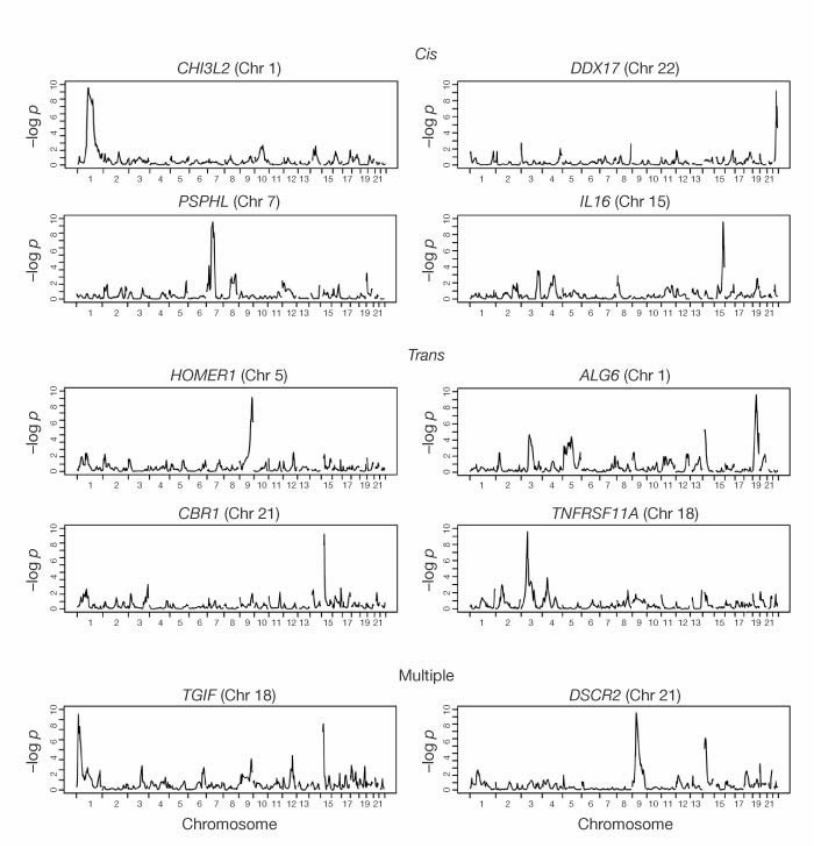
Genetic analysis of genome-wide variation in human gene expression

Michael Morley[1,3*], Cliona M. Molony[2*], Teresa M. Weber[1,3], James L. Devlin[2], Kathryn G. Ewens[2], Richard S. Spielman[2] & Vivian G. Cheung[1,2,3]

[1]Department of Pediatrics and [2]Department of Genetics, University of Pennsylvania, [3]The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA

*These authors contributed equally to this work

Natural variation in gene expression is extensive in humans and other organisms, and variation in the baseline expression level of many genes has a heritable component. To localize the genetic determinants of these quantitative traits (expression phenotypes) in humans, we used microarrays to measure gene expression levels and performed genome-wide linkage analysis for expression levels of 3,554 genes in 14 large families. For approximately 1,000 expression phenotypes, there was significant evidence of linkage to specific chromosomal regions. Both *cis*- and *trans*-acting loci regulate variation in the expression levels of genes, although most act *in trans*. Many gene expression phenotypes are influenced by several genetic determinants. Furthermore, we found hotspots of transcriptional regulation where significant evidence of linkage for several expression phenotypes (up to 31) coincides, and expression levels of many genes that share the same regulatory region are significantly correlated. The combination of microarray techniques for phenotyping and linkage analysis for quantitative traits allows the genetic mapping of determinants that contribute to variation in human gene expression.

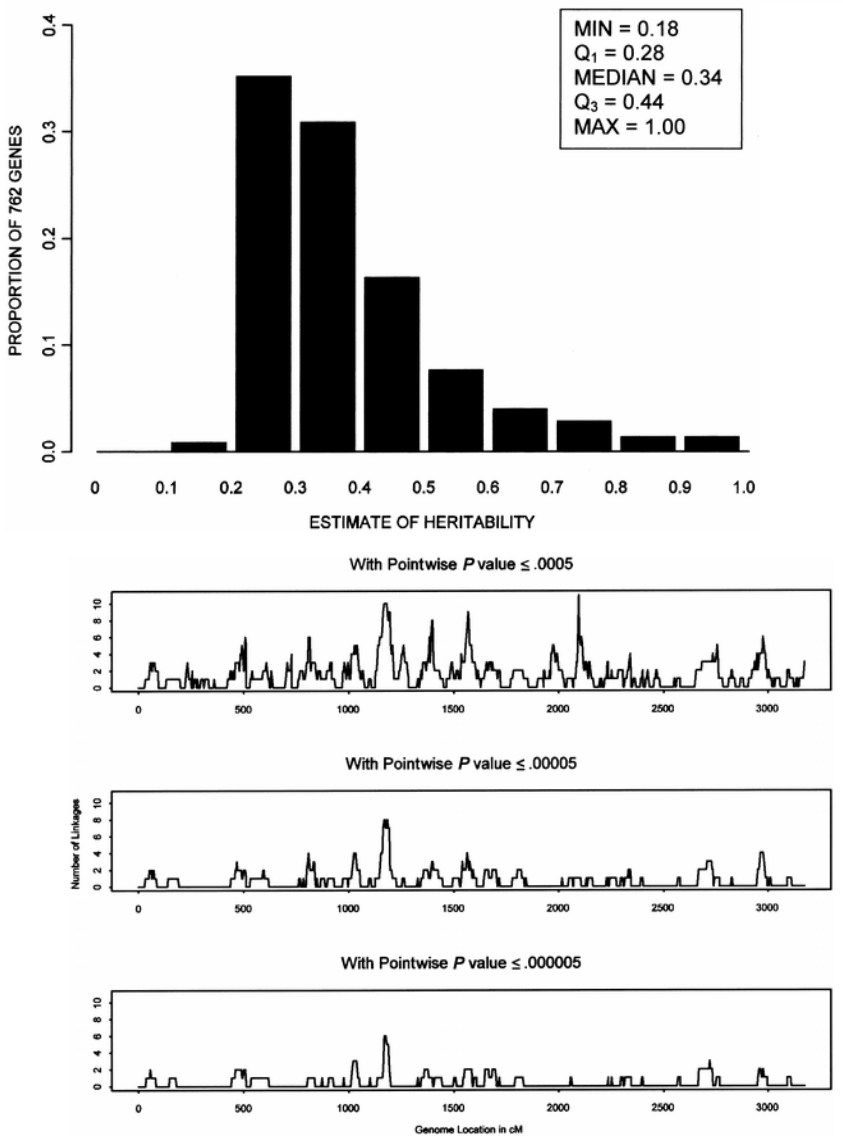


Am. J. Hum. Genet. 75:1094–1105, 2004

Genetic Inheritance of Gene Expression in Human Cell Lines

S. A. Monks,[1,2,4] A. Leonardson,[4] H. Zhu,[4] P. Cundiff,[3] P. Pietrusiak,[5] S. Edwards,[4] J. W. Phillips,[6] A. Sachs,[4] and E. E. Schadt[4]

[1]Department of Statistics, Oklahoma State University, Stillwater, OK; Departments of [2]Biostatistics and [3]Pharmacology, University of Washington, and [4]Rosetta Inpharmatics LLC, Seattle; [5]Department of Epidemiology, Johns Hopkins University, Baltimore; and [6]Merck Research Laboratories, Merck & Co., Rahway, NJ



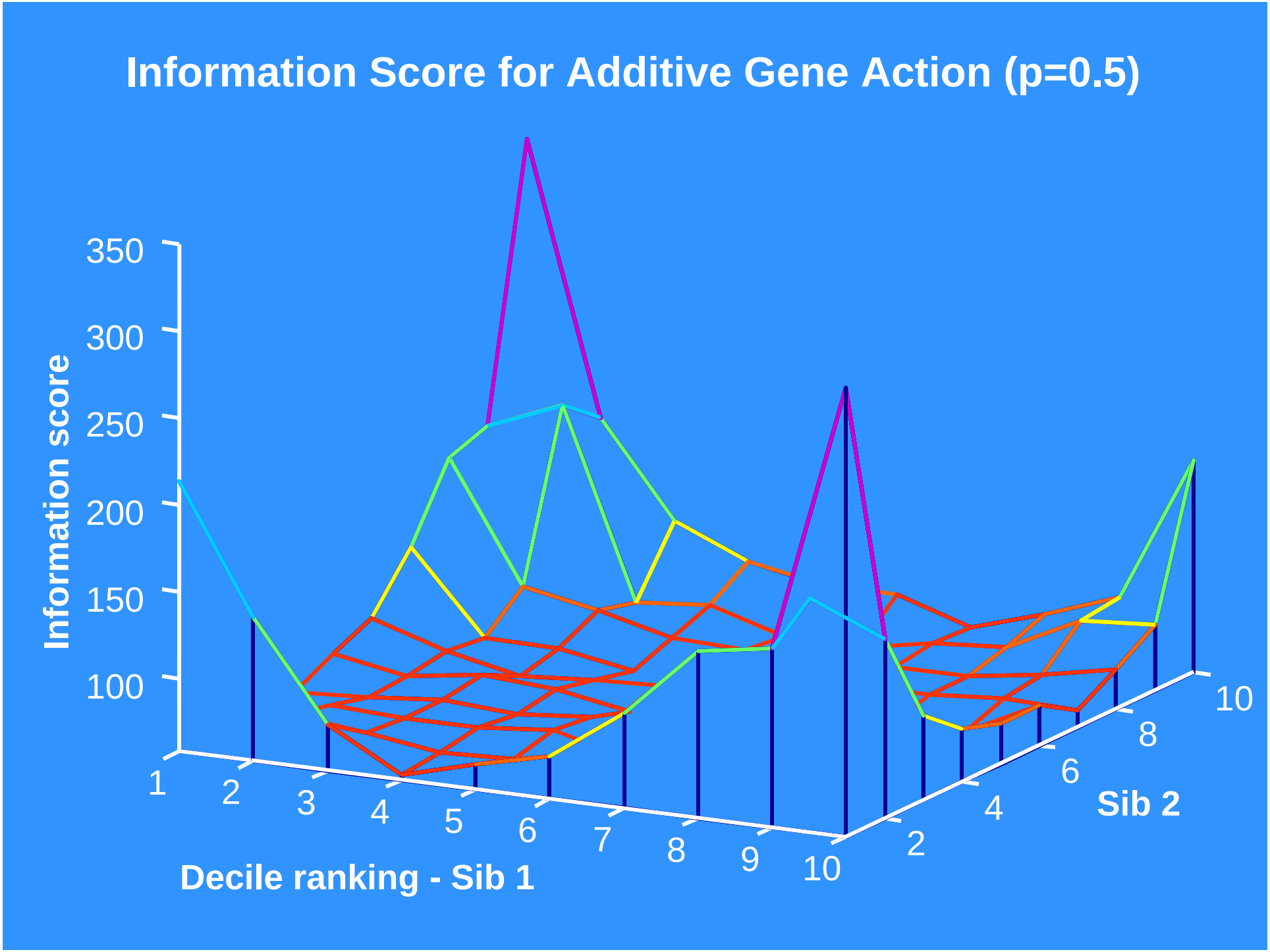
Increasing power for QTL linkage (2)

Selecting extreme pairs

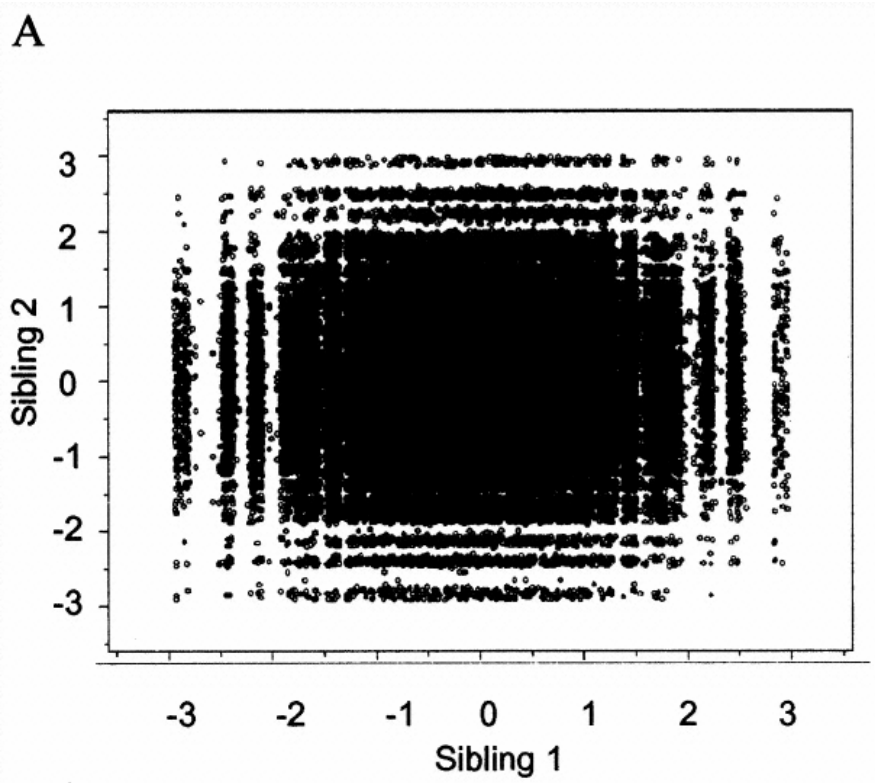
Extreme Discordant Sib Pairs for Mapping Quantitative Trait Loci in Humans

Neil Risch* and Heping Zhang

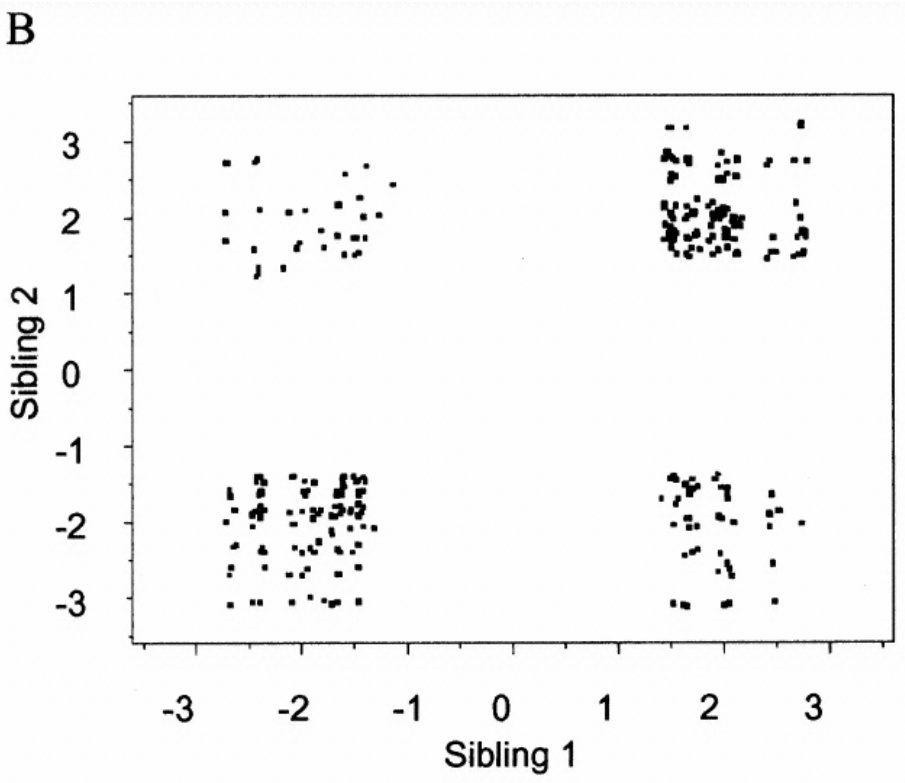
Analysis of differences between siblings (sib pair analysis) is a standard method of genetic linkage analysis for mapping quantitative trait loci, such as those contributing to hypertension and obesity, in humans. In traditional designs, pairs are selected at random or with one sib having an extreme trait value. The majority of such pairs provide little power to detect linkage; only pairs that are concordant for high values, low values, or extremely discordant pairs (for example, one in the top 10 percent and the other in the bottom 10 percent of the distribution) provide substantial power. Focus on discordant pairs can reduce the amount of genotyping necessary over conventional designs by 10- to 40 -fold.


Information Score for Additive Gene Action (p=0.5)
Information score
350
300
250
200
150
100
1
2
3
4
5
6
7
8
9
10
Decile ranking - Sib 1
2
4
6
8
10
Sib 2

Scatterplots of the distribution of neuroticism scores for each sibling pair

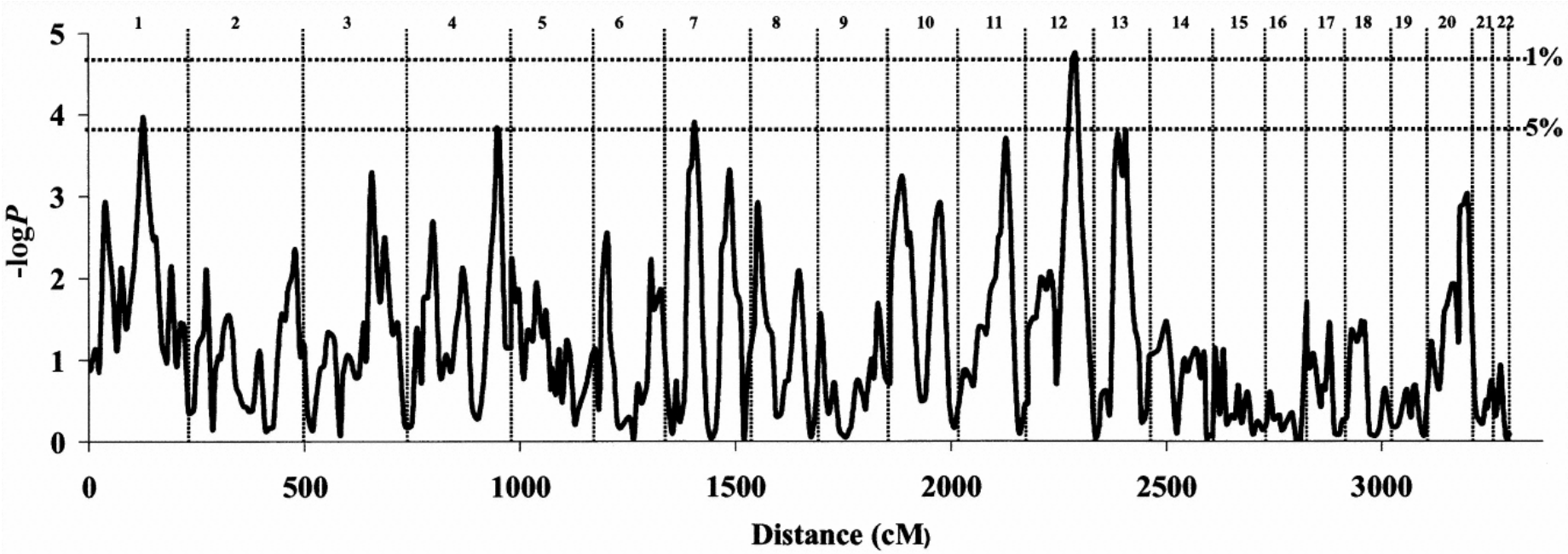


Distribution of entire sample



Distribution of selected sample

Multipoint linkage analysis of the genome for individual variation in neuroticism



The -logP values (vertical axis) for the Visscher-Hopper regression are shown. The cumulative distance is given at the bottom, and chromosome numbers are given at the top. The two dotted, horizontal lines represent the empirically derived genome wide significance thresholds (5% and 1%).

Am. J. Hum. Genet., 72:000, 2003

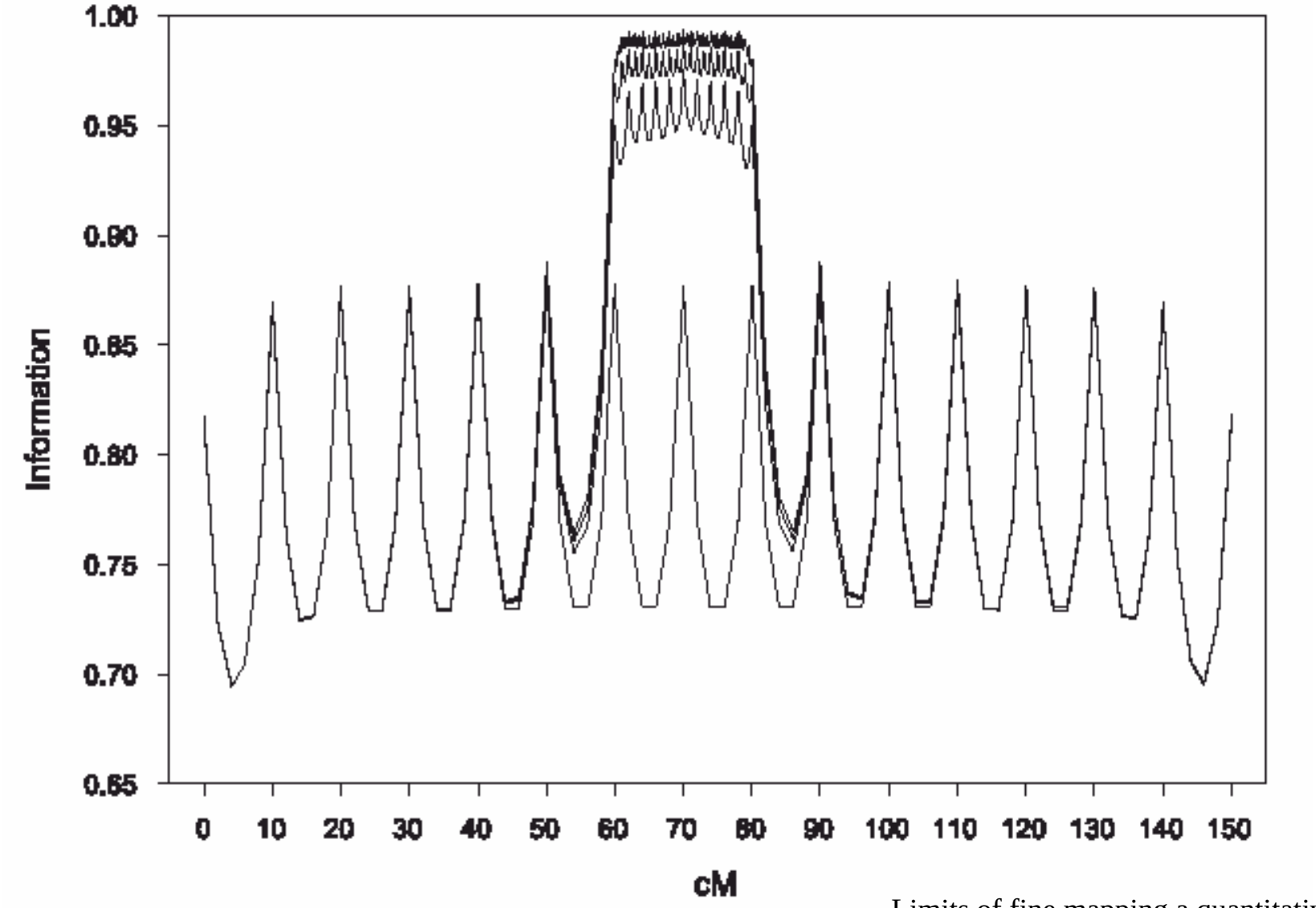
Linkage Analysis of Extremely Discordant and Concordant Sibling Pairs Identifies Quantitative-Trait Loci That Influence Variation in the Human Personality Trait Neuroticism

Jan Fullerton, Matthew Cubin, Hemant Tiwari, Chenxi Wang, Amarjit Bomhra, Stuart Davidson, Sue Miller, Christopher Fairburn, Guy Goodwin, Michael Neale, Simon Fiddy, Richard Mott, David B. Allison, and Jonathan Flint

Increasing power for QTL linkage (3)

Increasing marker density

Information for marker density 0.5, 1, 2, 10cM scan



Limits of fine mapping a quantitative trait
Attwood LD & Heard-Costa NL.
Genetic Epidemiology 24:99-106, 2003

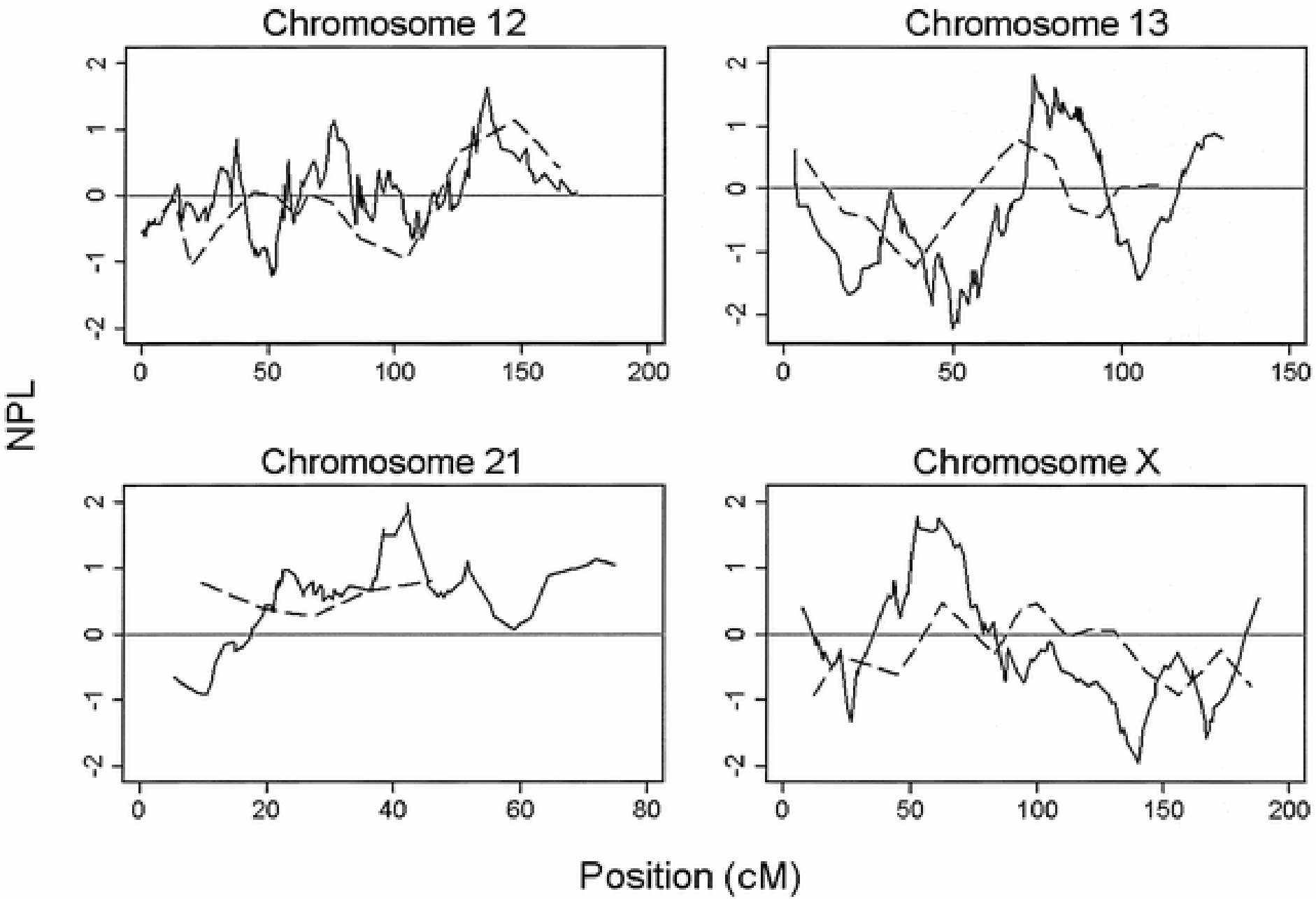
Am. J. Hum. Genet. 75:54–64, 2004

Whole-Genome Scan, in a Complex Disease, Using 11,245 Single-Nucleotide Polymorphisms: Comparison with Microsatellites

Sally John,[1] Neil Shephard,[1] Guoying Liu,[2] Eleftheria Zeggini,[1] Manqiu Cao,[2] Wenwei Chen,[2] Nisha Vasavda,[3] Tracy Mills,[3] Anne Barton,[1] Anne Hinks,[1] Steve Eyre,[1] Keith W. Jones,[2] William Ollier,[1] Alan Silman,[1] Neil Gibson,[3] Jane Worthington,[1] and Giulia C. Kennedy[2]

[1]University of Manchester, Manchester, United Kingdom; [2]Affymetrix, Santa Clara, CA; and [3]AstraZeneca, Macclesfield, United Kingdom

Despite the theoretical evidence of the utility of single-nucleotide polymorphisms (SNPs) for linkage analysis, no whole-genome scans of a complex disease have yet been published to directly compare SNPs with microsatellites. Here, we describe a whole-genome screen of 157 families with multiple cases of rheumatoid arthritis (RA), performed using 11,245 genomewide SNPs. The results were compared with those from a 10-cM microsatellite scan in the same cohort. The SNP analysis detected HLA*DRB1, the major RA susceptibility locus ($P = .00004$), with a linkage interval of 31 cM, compared with a 50-cM linkage interval detected by the microsatellite scan. In addition, four loci were detected at a nominal significance level ($P < .05$) in the SNP linkage analysis; these were not observed in the microsatellite scan. We demonstrate that variation in information content was the main factor contributing to observed differences in the two scans, with the SNPs providing significantly higher information content than the microsatellites. Reducing the number of SNPs in the marker set to 3,300 (1-cM spacing) caused several loci to drop below nominal significance levels, suggesting that decreases in information content can have significant effects on linkage results. In contrast, differences in maps employed in the analysis, the low detectable rate of genotyping error, and the presence of moderate linkage disequilibrium between markers did not significantly affect the results. We have demonstrated the utility of a dense SNP map for performing linkage analysis in a late-age-at-onset disease,


Chromosome 12
Chromosome 13
Chromosome 21
Chromosome X
NPL
Position (cM)

Am. J. Hum. Genet. 75:687–692, 2004

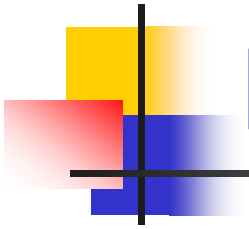
Report

Guidelines for Genotyping in Genomewide Linkage Studies: Single-Nucleotide–Polymorphism Maps Versus Microsatellite Maps

David M. Evans and Lon R. Cardon

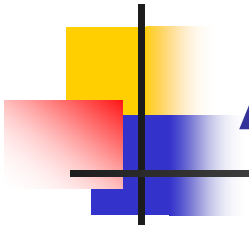
Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

Genomewide linkage scans have traditionally employed panels of microsatellite markers spaced at intervals of ~10 cM across the genome. However, there is a growing realization that a map of closely spaced single-nucleotide polymorphisms (SNPs) may offer equal or superior power to detect linkage, compared with low-density microsatellite maps. We performed a series of simulations to calculate the information content associated with microsatellite and SNP maps across a range of different marker densities and heterozygosities for sib pairs (with and without parental genotypes), sib trios, and sib quads. In the case of microsatellite markers, we varied density across 11 levels (1 marker every 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 cM) and marker heterozygosity across 6 levels (2, 3, 4, 5, 10, or 20 equally frequent alleles), whereas, in the case of SNPs, we varied marker density across 4 levels (1 marker every 0.1, 0.2, 0.5, or 1 cM) and minor-allele frequency across 7 levels (0.5, 0.4, 0.3, 0.2, 0.1, 0.05, and 0.01). When parental genotypes were available, a map consisting of microsatellites spaced every 2 cM or a relatively sparse map of SNPs (i.e., at least 1 SNP/cM) was sufficient to extract most of the inheritance information from the map (>95% in most cases). However, when parental genotypes were unavailable, it was important to use as dense a map of markers as possible to extract the greatest amount of inheritance information. It is important to note that the information content associated with a traditional map of microsatellite markers (i.e., 1 marker every ~10 cM) was significantly lower than the information content associated with a dense map of SNPs or microsatellites. These results strongly suggest that previous linkage studies that employed sparse microsatellite maps could benefit substantially from reanalysis by use of a denser map of markers.



Linkage

- Doesn’t depend on “guessing gene”
- Works over broad regions (good for getting in right ball-park) and whole genome (“genome scan”)
- Only detects large effects (>10%)
- Requires large samples (10,000’s?)
- Can’t guarantee close to gene

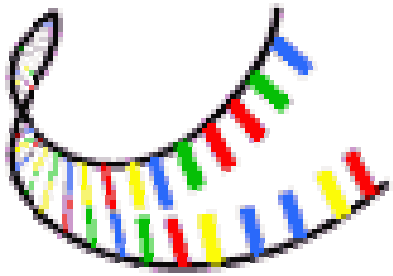


Association

- Looks for correlation between specific alleles and phenotype (trait value, disease risk)

Polymorphism

"Poly" *many* "morphe" *form*

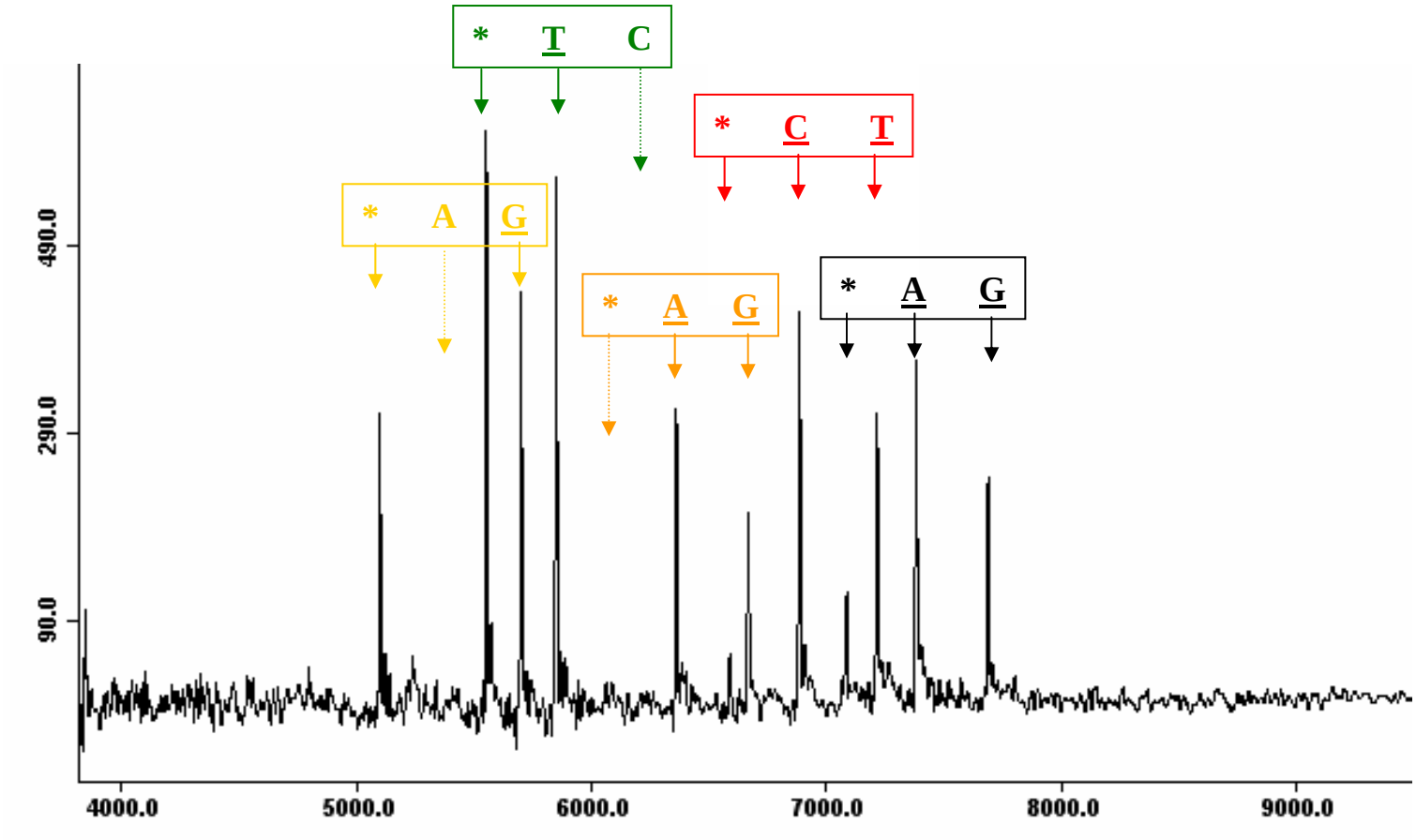


QIMR's Sequenom MassARRAY Installation

SEQUENOM
INDUSTRIAL GENOMICS™

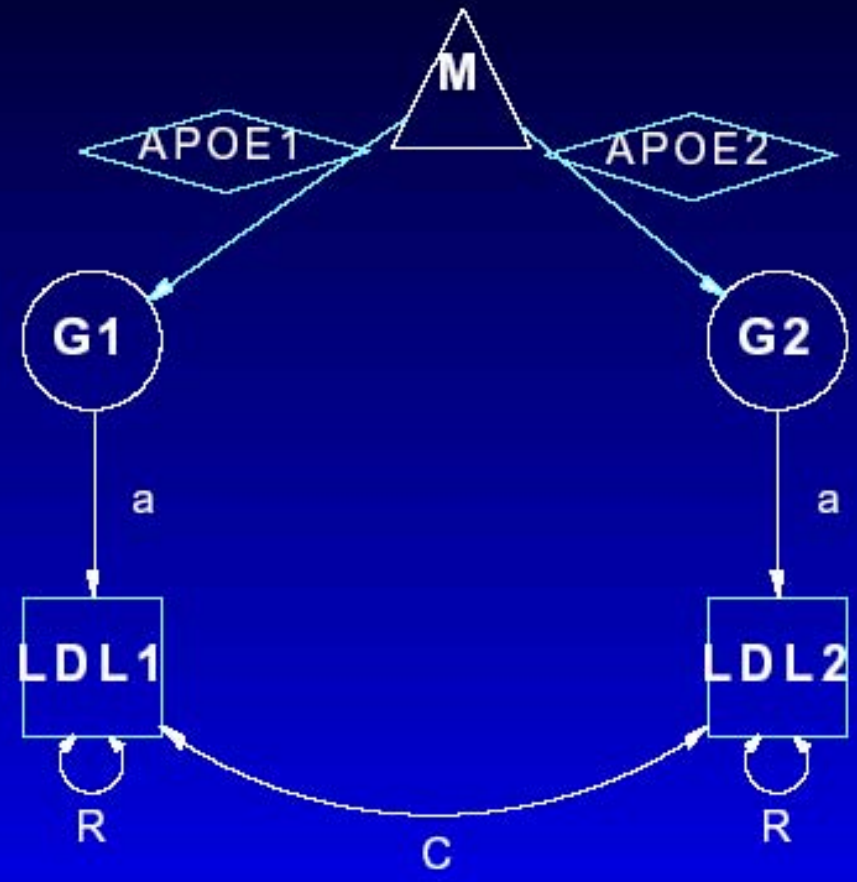
Multiplexing Assays

•SNP 3 •SNP 4 •SNP 1 •SNP 5 •SNP 2



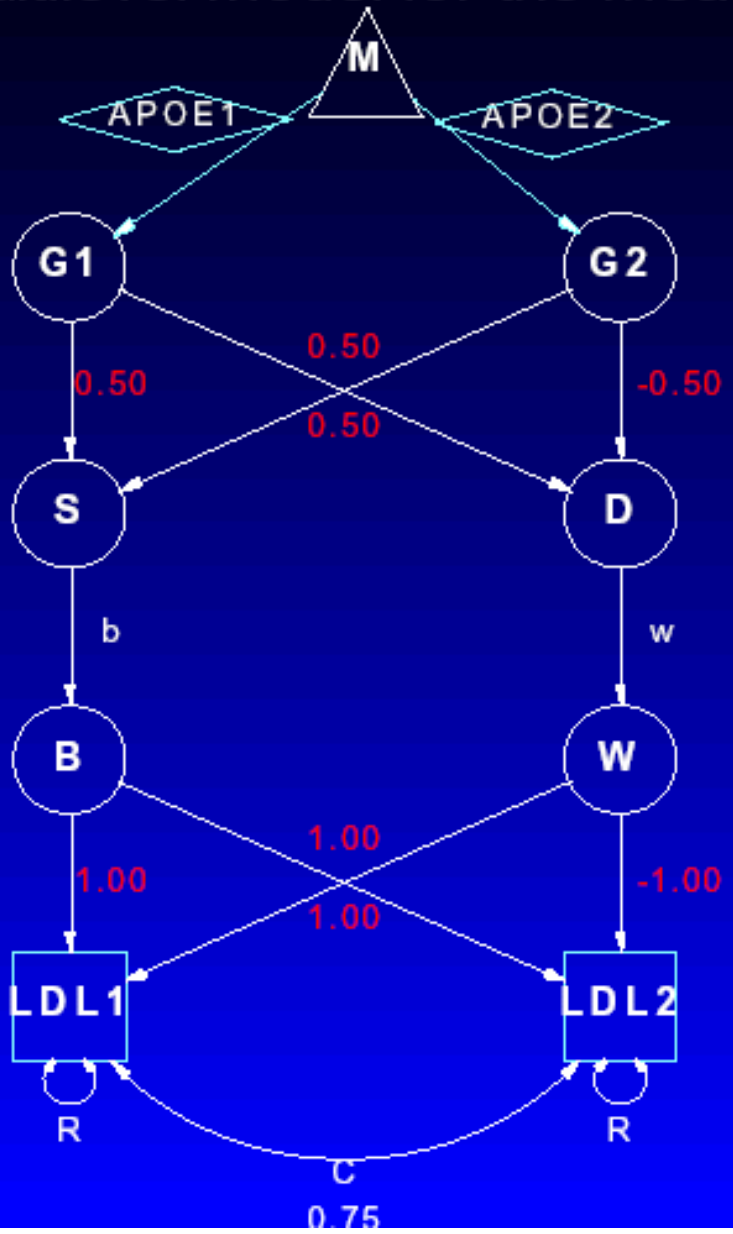
Association Model

Each sib pair i has different MEANS



Fulker Association Model

Multilevel model for the means



Cat's cradle

Diversity and linkage disequilibrium

- HapMap (http://www.hapmap.org/)
  - Samples from four populations

European (CEU): 30 trios

Yoruba (YRI): 30 trios

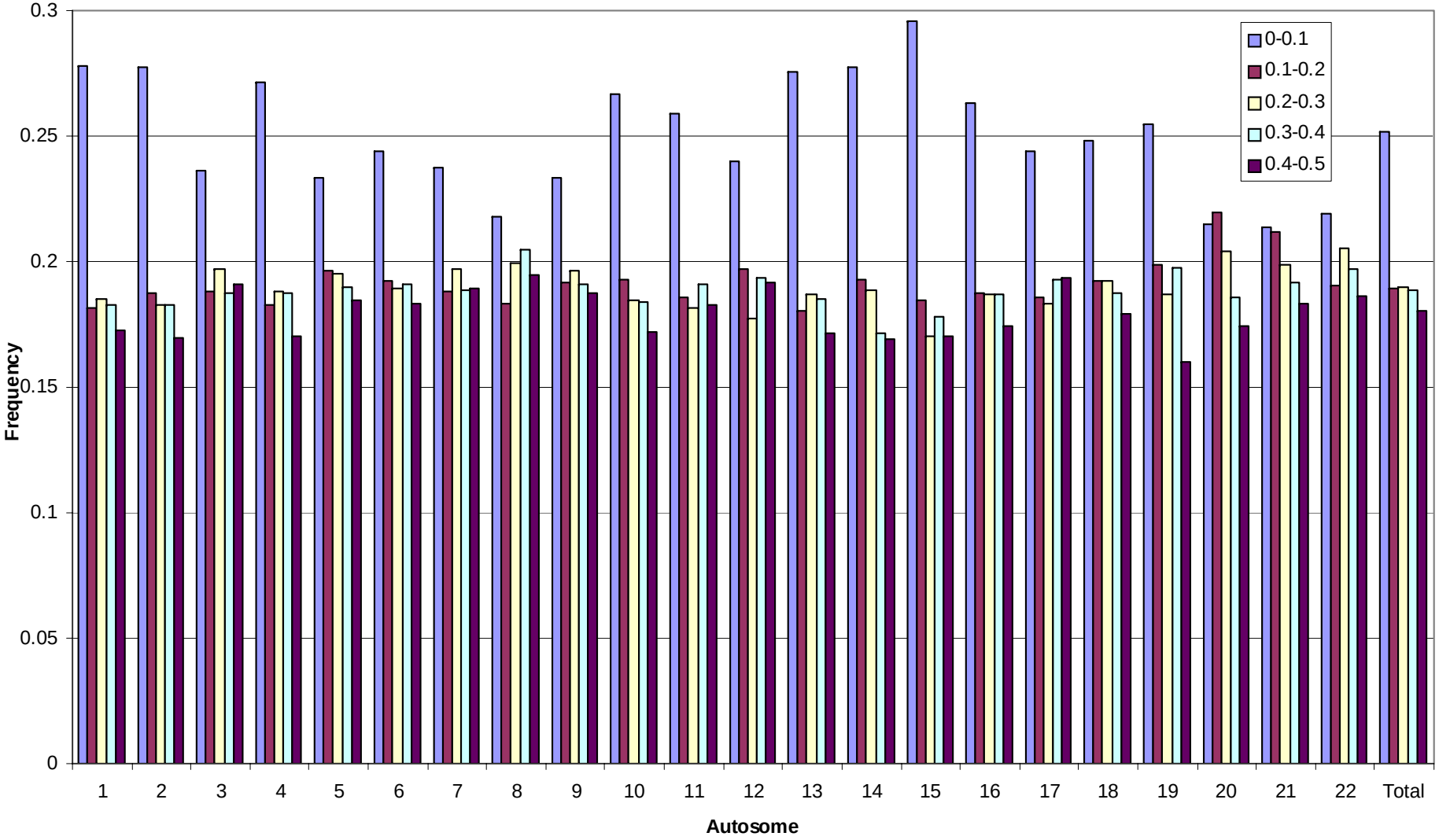
Japanese (JPT): 45 unrelated

Han Chinese (HCB): 45 unrelated

SNP
(Public release 16, March 2005)

Population	Total
Europeans	1,073,663
Africans	1,034,205
Japanese	1,044,416
Han Chinese	1,044,686

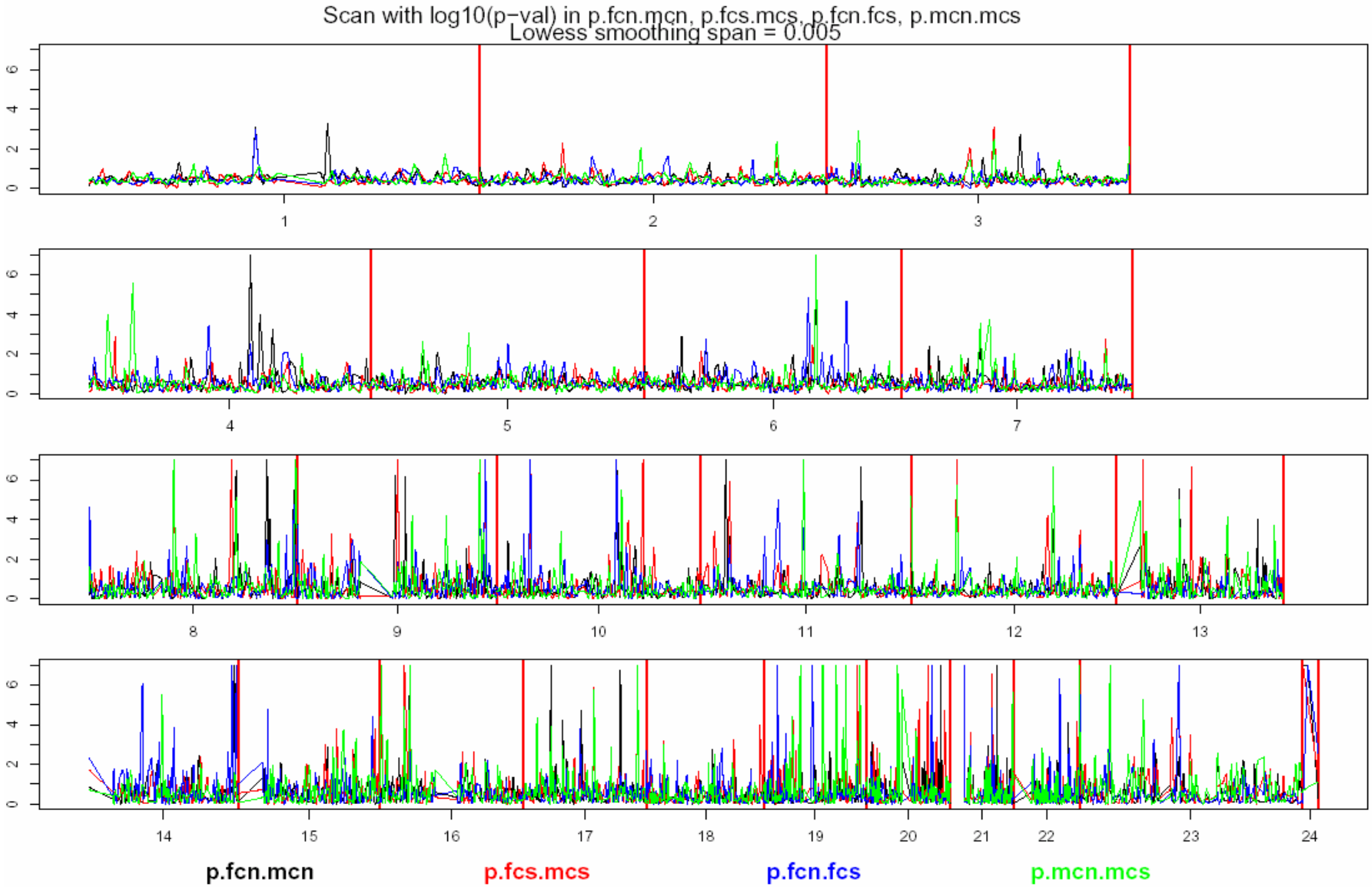
SNP allele frequency distribution for the 22 autosomes obtained from the CEU data



Affy 50k SNP chip (58,960 SNPs)

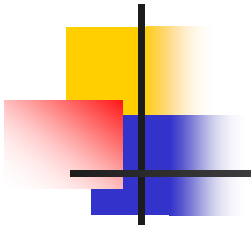
	MZ 1	MZ 2	DZ	Dad	Mum
MZ 1	99.43	3 disc!			Call %
MZ 2	99.99	99.43	sibs		
DZ	75.36	75.35	99.72		Parent-offspring
Dad	70.64	70.64	70.76	99.32	unrelated
Mum	70.48	70.48	70.73	58.02	99.49

Melanoma genome-wide association study



Association

- More sensitive to small effects
- Need to “guess” gene/alleles (“candidate gene”) or be close enough for linkage disequilibrium with nearby loci
- May get spurious association (“stratification”) – need to have genetic controls to be convinced



Problems

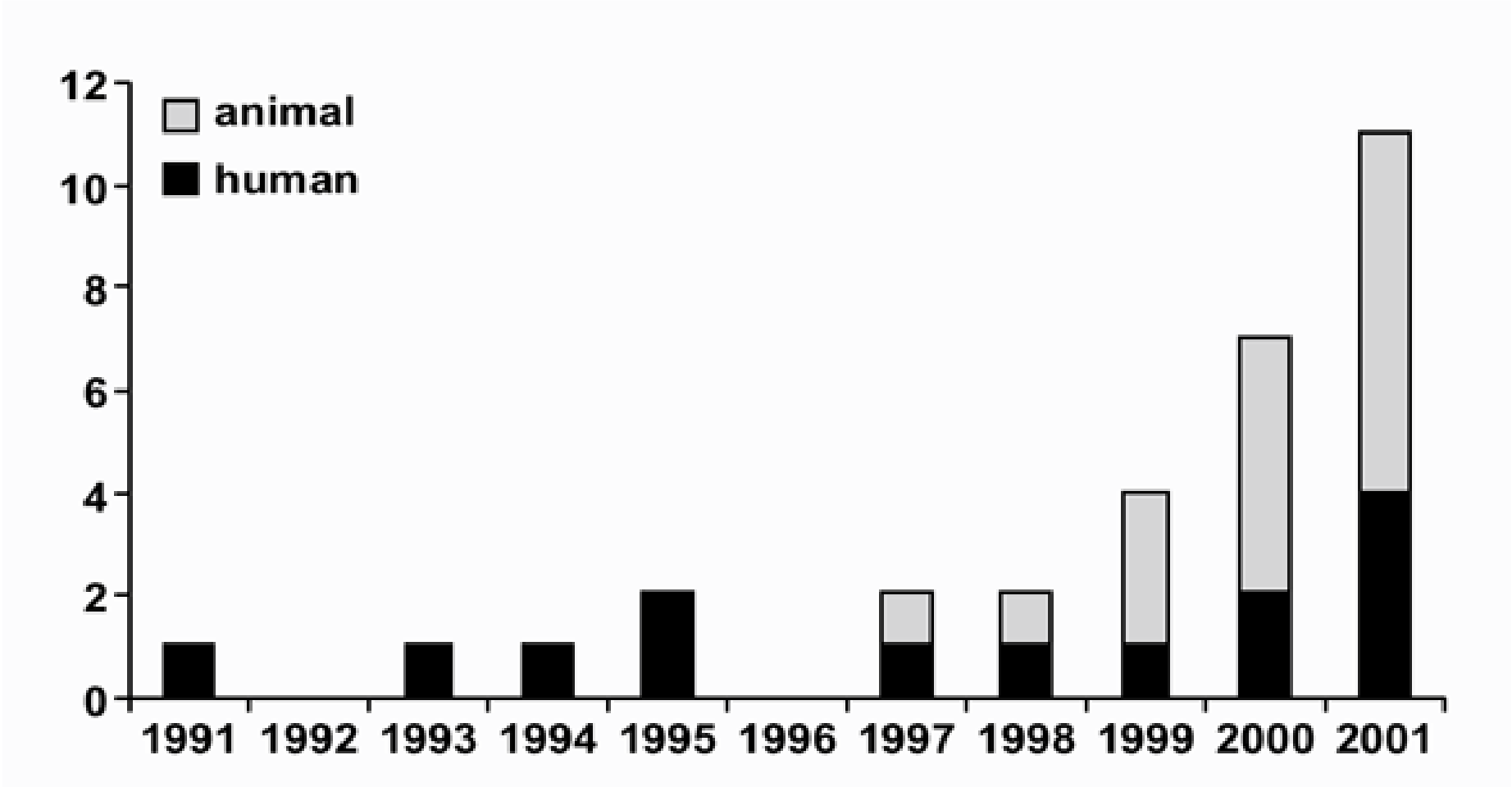
- Large number of loci and alleles/haplotypes
- Possible interactions between genes
- Possible G x E interactions
- Relatively low frequencies of individual risk factors
- Genotype-phenotype relationship unknown
- Signal/noise – minimizing errors within budget
- Scaling of phenotype (continuous, discontinuous)
- Spurious association (stratification)

From QTL to gene: the harvest begins

In the past decade, quantitative trait locus (QTL) mapping has identified hundreds of chromosomal regions containing genes affecting asthma, atherosclerosis, diabetes, hypertension, obesity and other complex phenotypes. The ultimate goal of QTL mapping is to identify the genes underlying these polygenic traits and to gain a better understanding of their physiology and biochemistry. But identifying the QTL genes has been slow and difficult. A commentary in *Nature Genetics* questioned the value of QTL mapping and proposed that mutagenesis strategies offer more promise for identifying the genes determining complex traits[1]. That pessimistic opinion was premature; we herein report that 29 QTL genes have been identified, almost half of them in 2001. We suggest that QTL mapping of complex traits is a promising technique and that the harvest of QTL genes is just beginning.

Ron Korstanje & Beverly Paigen

Number of genes identified from QTL by year



From QTL to gene: the harvest begins: RKorstanje & B Paigen : *Nature Genetics* 31, 235 – 236 (2002)

Table 1 • Genes identified from QTL studies

Polygenic trait	Year	Ref.	Gene	Species	pos	tg	ko	fu
Alzheimer disease	1991	9	*APP*	human				X
Alzheimer disease	1993	10	*APOE*	human				
Ovarian and breast cancer	1994	11	*BRCA1*	human	X			X
Breast cancer	1995	12	*BRCA2*	human	X			X
Insulin resistance	1995	13	*FABP2*	human				
HDL-cholesterol levels	1997	14	*LIPC*	human				
Intestinal cancer	1997	15	*Pla2g2a*	mouse	X	X		
Blood pressure	1998	16	*Atp1a1 / ATP1A1*	rat/human		X		X
Leptin levels	1999	17,18	*POMC*	human			X[a]	X
Asthma	1999	19	*Il4*	mouse	X	X		
Asthma	1999	19	*Il13*	mouse	X	X		
Insulin-mediated glucose uptake	1999	2	*Cd36*	rat		X		
Obesity	2000	20	*Ptpn1/PTPN1*	mouse/human			X[b]	X
Alzheimer disease	2000	21	*PSEN1*	human	X			
Diabetes	2000	22	*Il2*	mouse	X		X[b]	X
Gallstones	2000	23	*Abcc2*	mouse	X			X
Asthma	2000	3	*Hc*	mouse				
Muscle glycogen content	2000	24	*Prkag3*	pig	X		X[c]	X
Crohn disease	2001	25,26	*NOD2*	human	X		X[a]	X
Blood pressure	2001	27	*SCNN1A1*	human			X[a]	
Blood pressure	2001	28	*SCNN1G*	human			X[a]	
Blood pressure	2001	29	*Slc12a1*	rat				
Blood pressure	2001	30	*Cyp11b1*	rat				X
Bone density	2001	5	*COL1A*	human				
Left ventricular mass	2001	31	*Nppa*	rat			X[b]	X
Modifier of tubby hearing	2001	32	*Mtap1a*	mouse	X	X		X
Taste, saccharin response	2001	33	*Tas1r3*	mouse	X	X		X
Tumor susceptibility	2001	34	*Cdkn2a*	mouse	X		X[b]	X
Diabetes	2001	35	*B2m*	mouse		X	X	

pos, found by positional cloning; tg, transgenic insertion of normal gene changes phenotype to normal (for example, transgenic rescue); ko, knockout provides additional evidence ([a]human monogenic syndrome, [b]deletion of gene by homologous recombination produces a mouse with the phenotype typical of the disease, [c]knockout in yeast); fu, functional difference in candidate gene. *APP*, amyloid precursor protein; *APOE*, apolipoprotein E; *BRCA*, breast cancer gene; *FABP2*, fatty acid binding protein 2; *LIPC*, hepatic lipase; *ATP1A1*, α-Na,K-ATPase; *POMC*, pre-pro-opiomelanocortin; *Il*, interleukin; *Cd36*, fatty acid translocase; *PTP1B*, protein tyrosine phosphatase-1B; *PSEN*, presenilin 1; *Abcc2*, ATP-binding cassette, subfamily C2; *Hc*, hemolytic complement (*C5*); *Prkag3*, protein kinase, AMP-activated, γ3; *NOD2*, caspase recruitment domain-containing protein 15 (*CARD15*); *SCNN*, sodium channel, non-voltage gated; *Slc12a1*, Na,K,2Cl-cotransporter; *Cyp11b1*, 11β-hydroxylase; *COL1A*, collagen-1A; *Nppa*, natriuretic peptide precursor A; *Mtap1a*, microtubule-associated protein 1a; *Tas1r3*, taste receptor-3; *Cdkn2a*, cyclin-dependent kinase inhibitor 2a; *B2m*, β2-microglobulin.

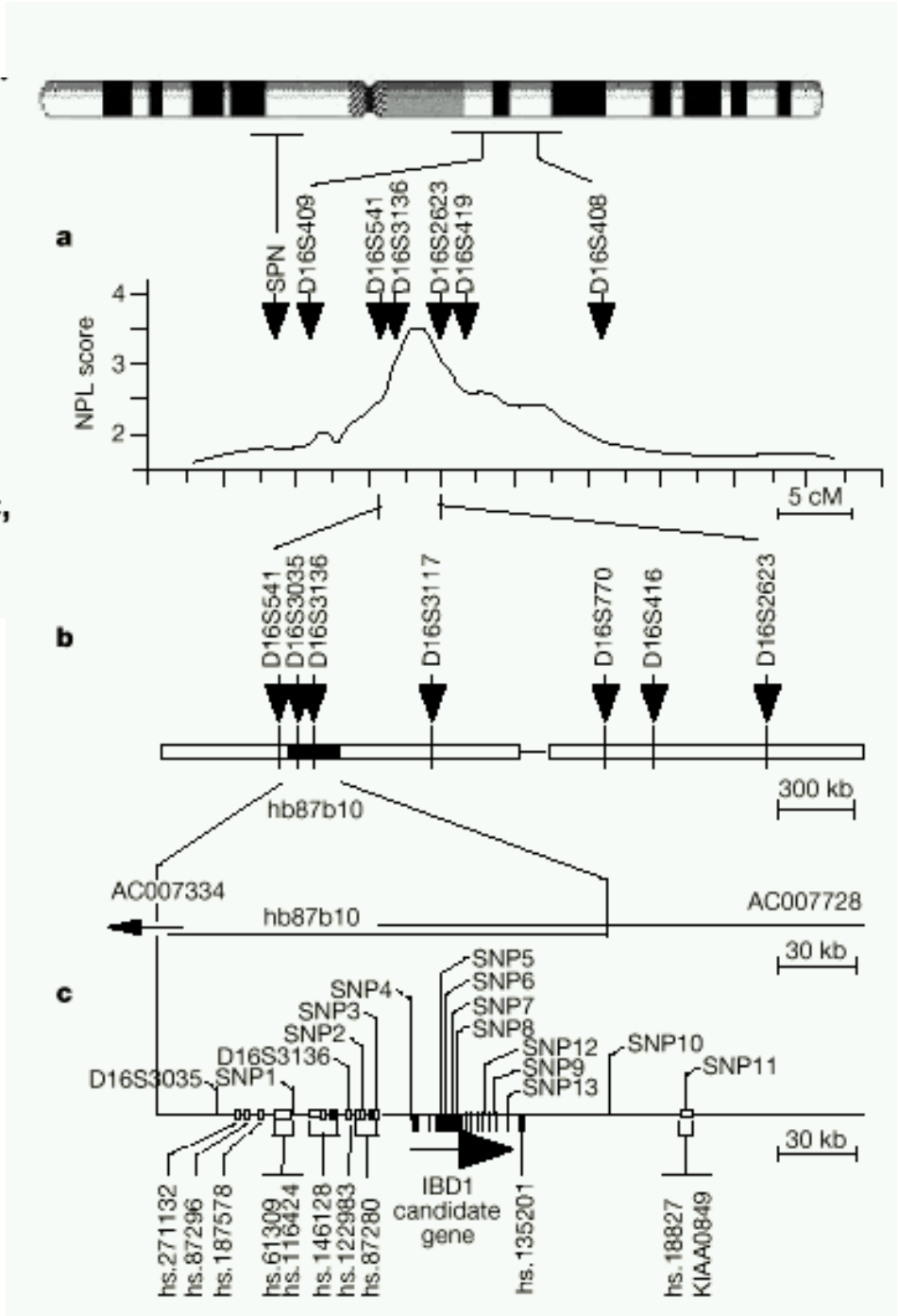
From QTL to gene: the harvest begins: Ron Korstanje & Beverly Paigen : Nature Genetics 31, 235 – 236 (2002)

Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease

Jean-Pierre Hugot*†‡, Mathias Chamaillard*†, Habib Zouali*, Suzanne Lesage*, Jean-Pierre Cézard‡, Jacques Belaiche§, Sven Almer‖, Curt Tysk¶, Colm A. O'Morain#, Miquel Gassull☆, Vibeke Binder**, Yigael Finkel††, Antoine Cortot‡‡, Robert Modigliani§§, Pierre Laurent-Puig†, Corine Gower-Rousseau‡‡, Jeanne Macry‖‖, Jean-Frédéric Colombel‡‡, Mourad Sahbatou* & Gilles Thomas*†¶¶

NATURE | VOL 411 | 31 MAY 2001

First (unequivocal) positional cloning of a complex disease QTL !



EPIGENETIC DISCORDANCE IN IDENTICAL TWINS

The missing “environment” ?

For most heritable complex traits MZ discordance is >50%

- Is this due to exogenous environment?
- Or could it be stochastic epigenetic differences?

- We present a rare case study that may illustrate a more general explanation of MZ discordance (and etiology in the non-twin population)

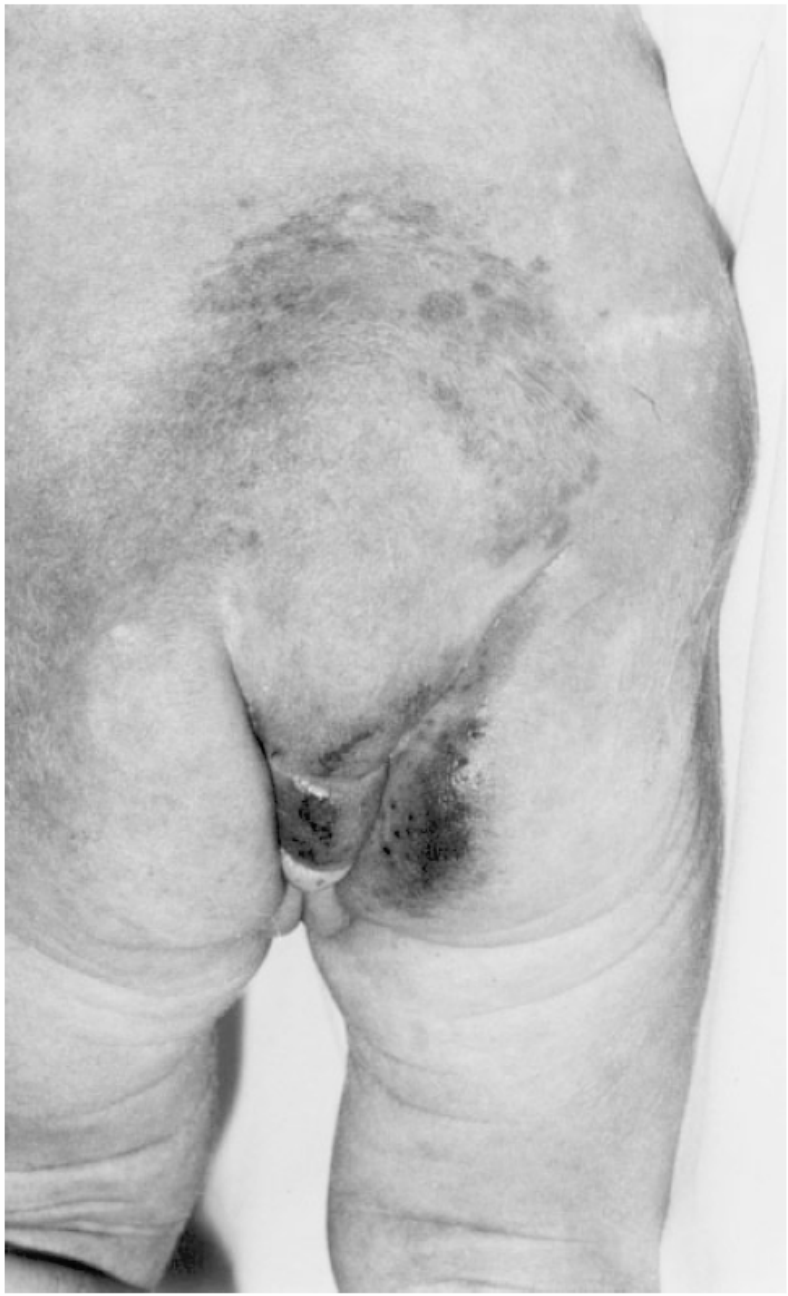


Fig. 1. Patient 1. Soft tumor and abnormal aspect in the lumbosacral area.

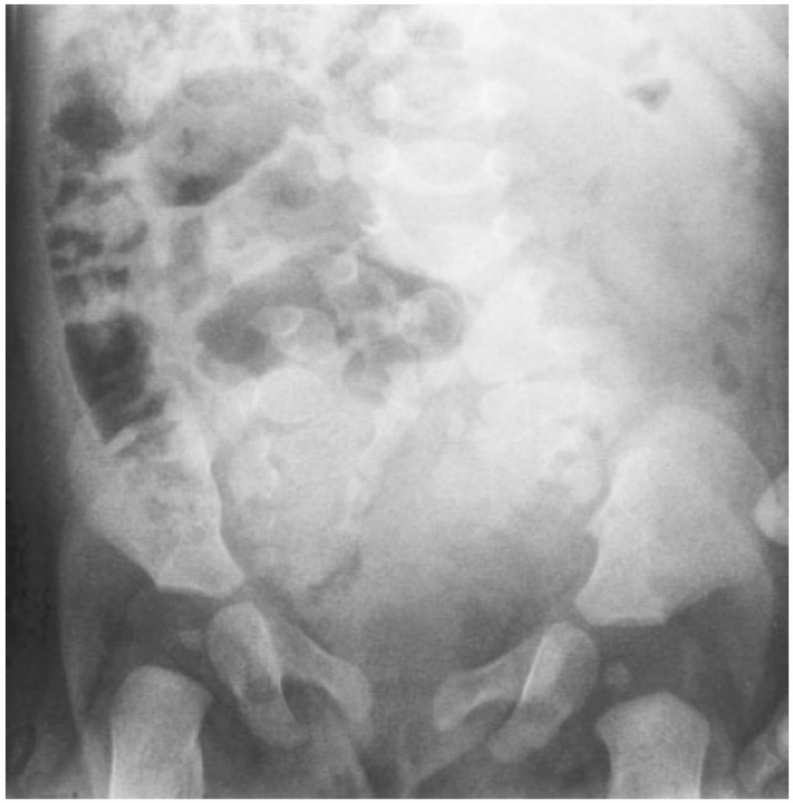
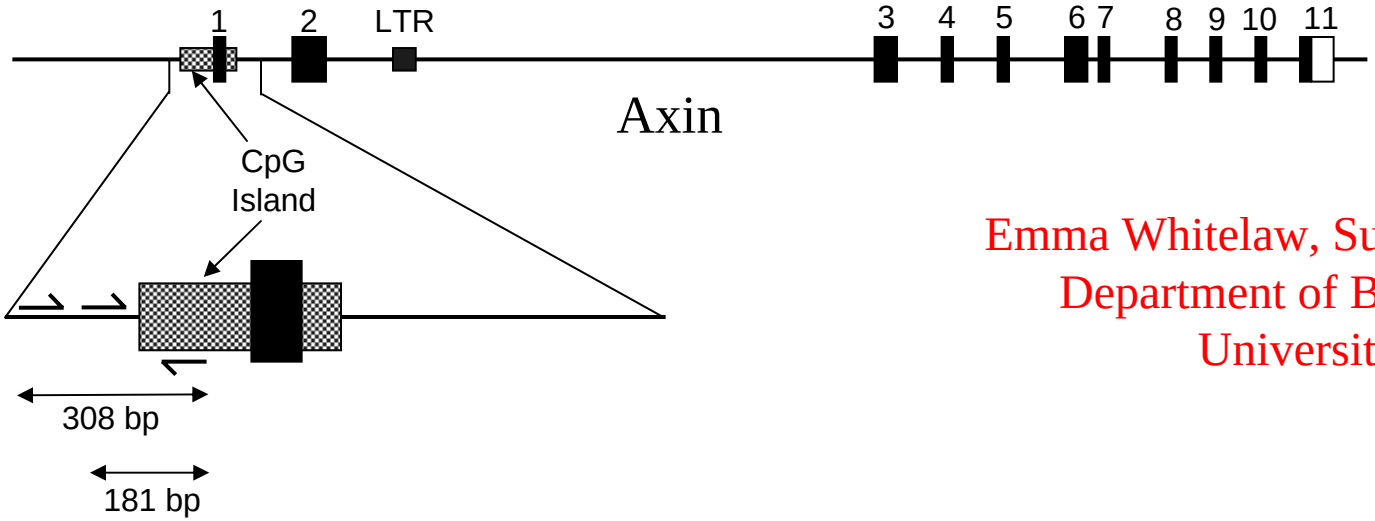


Fig. 2. Patient 1. Radiograph of the vertebral column shows complete duplication of the spine from L4 down.

urethra, a dilated pelvis of the right kidney, bilateral uterus unicornis with normal ovaries, hemivertebrae of thoracic vertebrae 6 and 10, and abnormal curvature of the sacrum. A persistent ductus arteriosus and secundum atrial septum defect was suspected, but results of cardiac investigations at 10 months were normal.

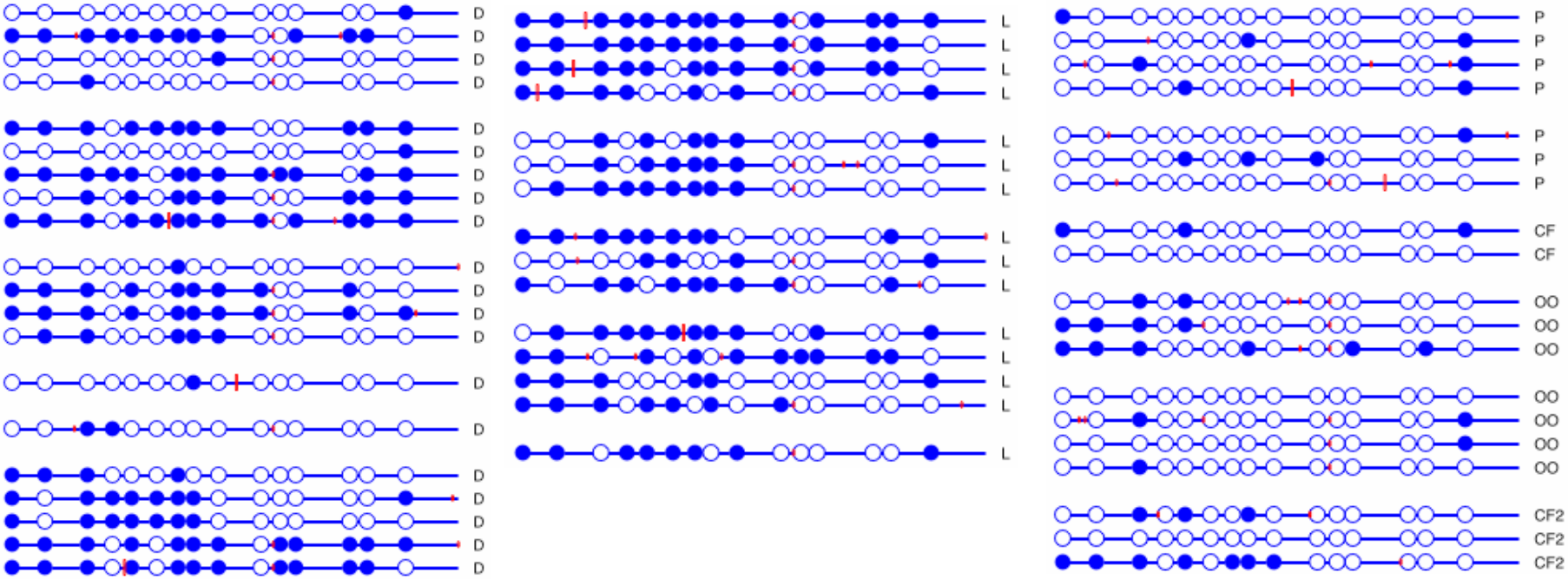
At physical examination for genetic evaluation at 4 months we saw a baby girl with epicanthal folds, but no other minor anomalies. She had a capillary nevus on her left buttock. In the anal region only a dimple was seen. The patient was operated on one day after birth, when a colostomy was made and a fistula connected to the colon

Discordant caudal duplication in MZ twins



Emma Whitelaw, Suyinn Chong
Department of Biochemistry
University of Sydney

Twin 1- unaffected < Twin 2 - affected > Controls [e.g.]



Other studies on MZ discordance

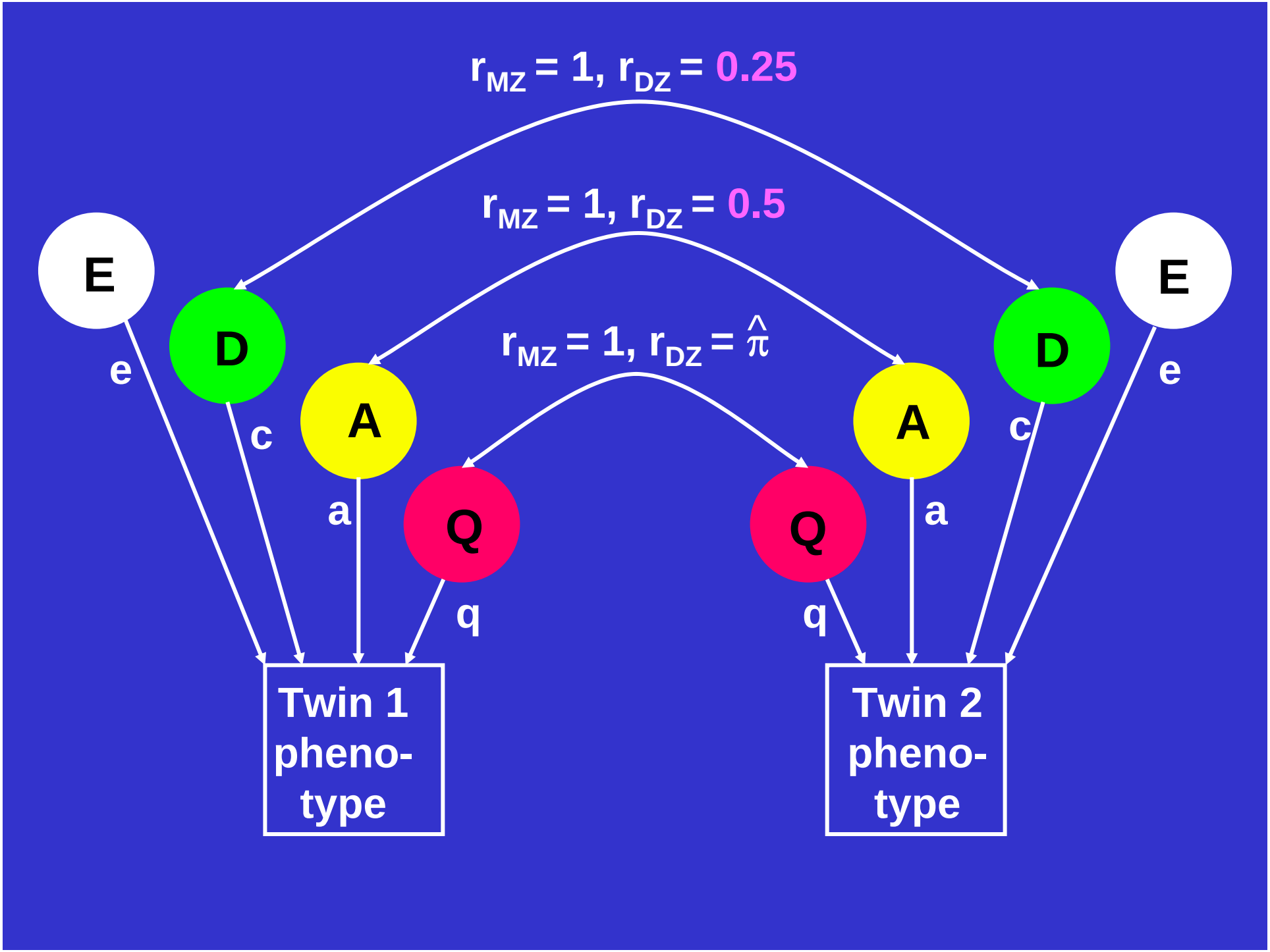
Epilepsy (with Sam Berkovic, Lata Vadlamudi)

Schizophrenia (with B.Mowry, N.Hayward)

Depression (with Art Petronis)

?Male homosexuality (with Vincent Harley)

Will high density genotyping allow us to do genetics without Fisher?


$r_{MZ} = 1, r_{DZ} = 0.25$
$r_{MZ} = 1, r_{DZ} = 0.5$
$r_{MZ} = 1, r_{DZ} = \hat{\pi}$
E
D
A
Q
E
D
A
Q
e
c
a
q
q
a
c
e
Twin 1 pheno-type
Twin 2 pheno-type

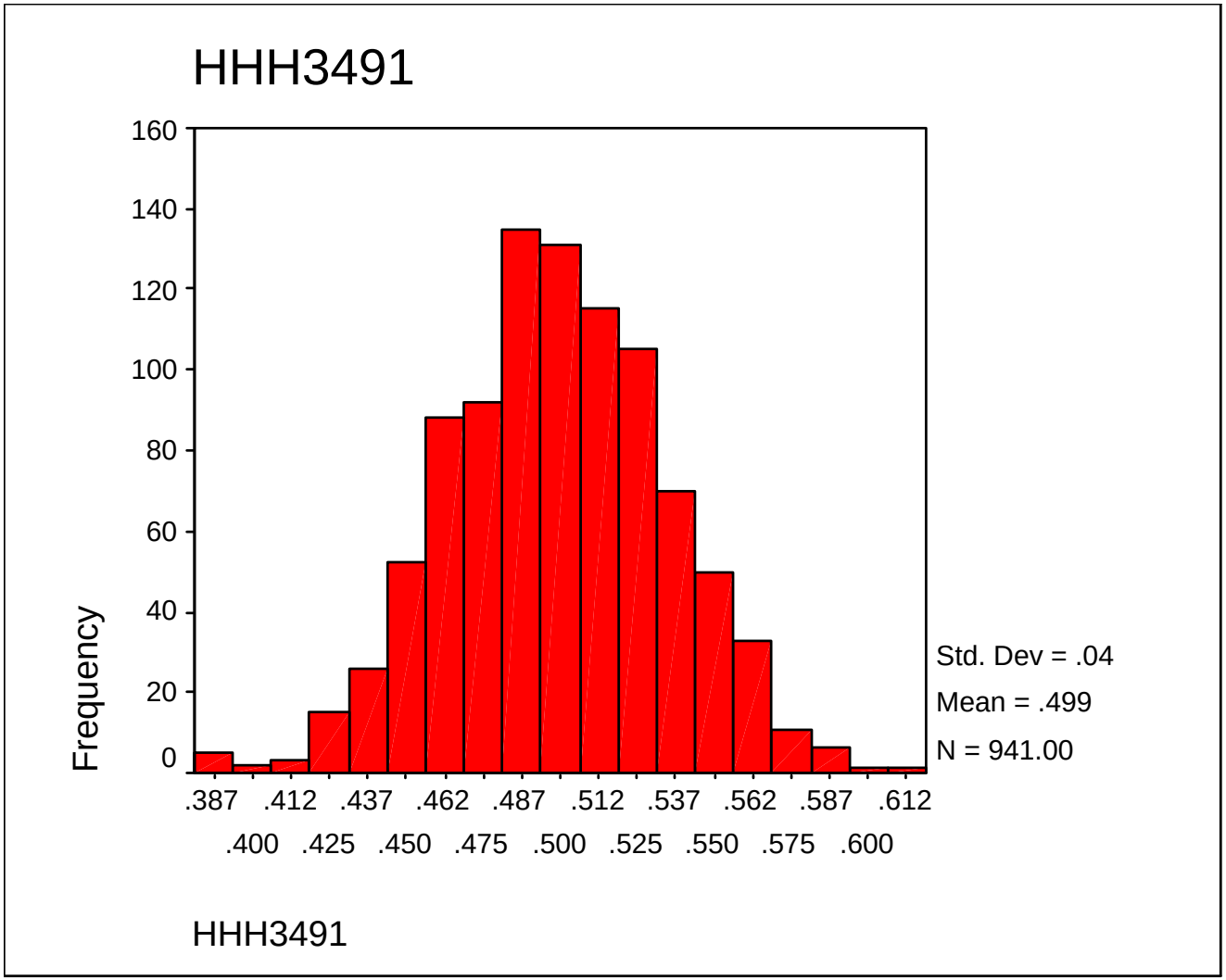
But why do we use the average sib values of

$$r_a = 0.5$$
$$r_d = 0.25$$

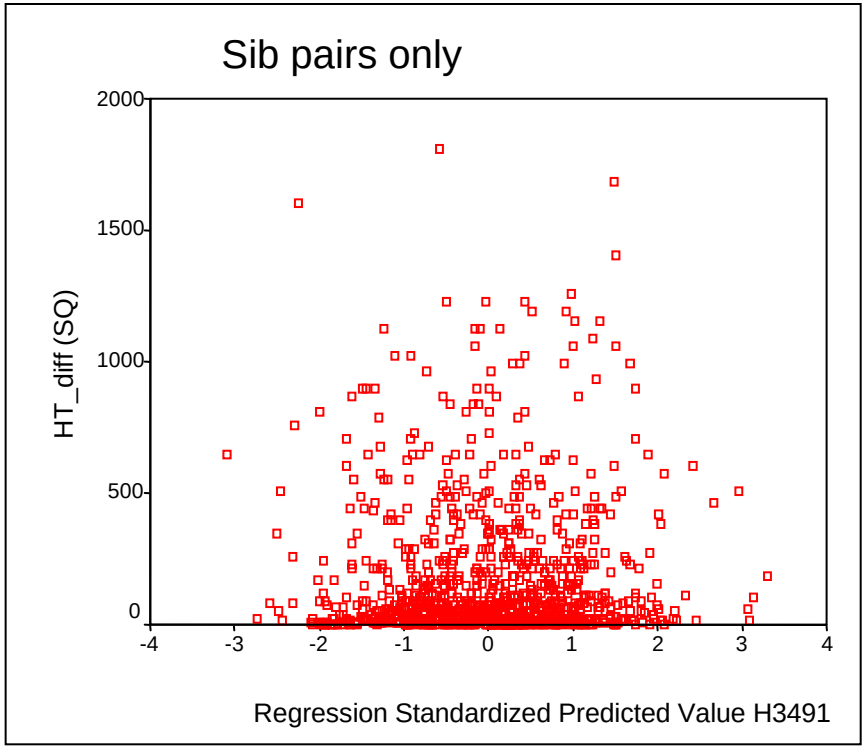
when we can estimate the (almost) exact values for each sib pair from marker data ?

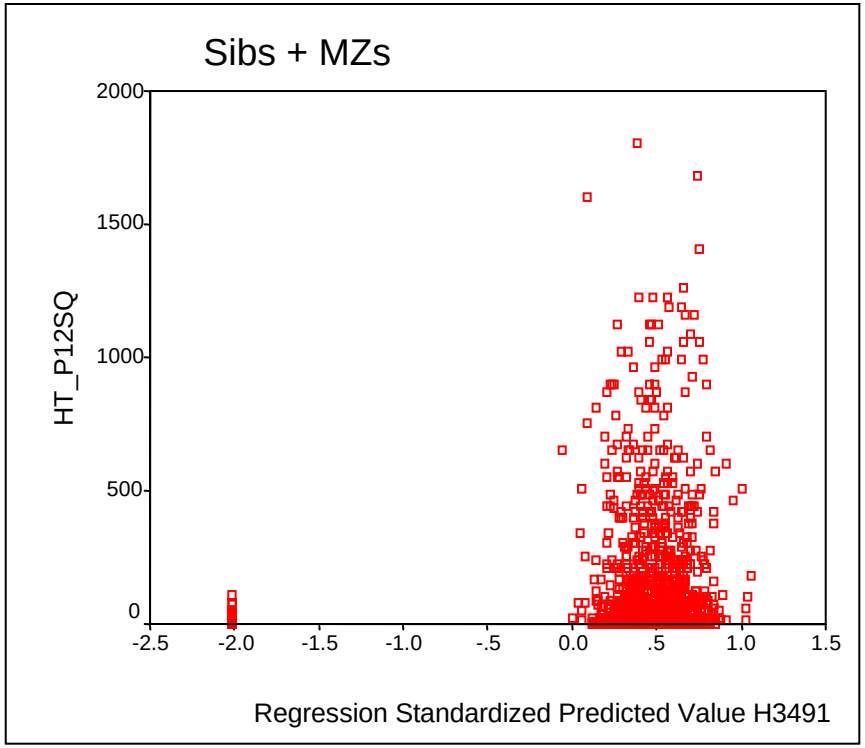
Are there any advantages in doing so ?

The distribution of mean IBD sharing for 941 q.i. sib pairs



Regression statistics for sibling intrapair variance on genomewide mean IBD sharing for height (age, sex corrected) with and without MZ twins





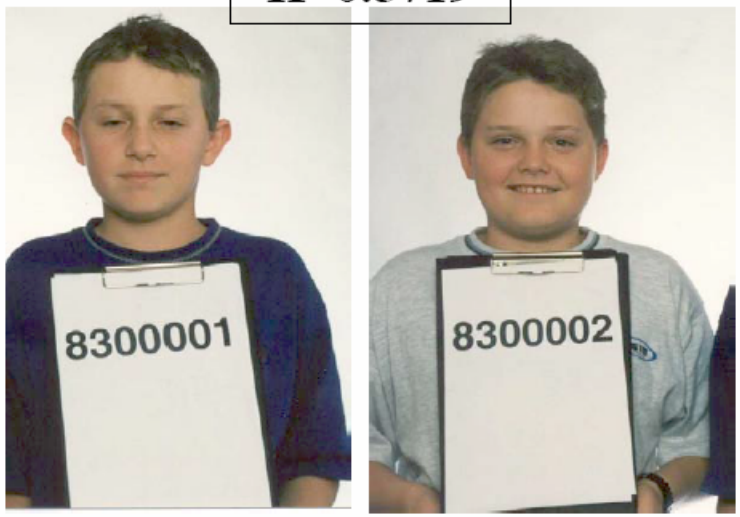
R = -0.025, ns

R = -0.28, p < 0.001

8188001,02
H=0.5677

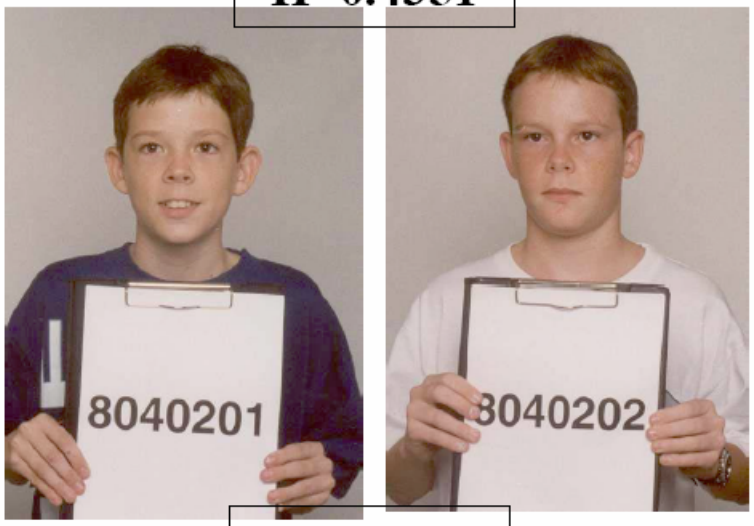
8473001,02
H=0.5577

8300001,02
H=0.5719



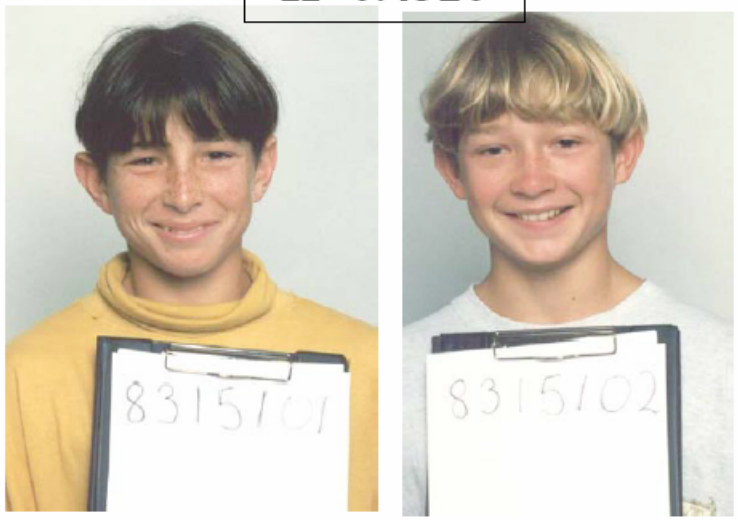
8582601,02
H=0.5640

8040201,02
H=0.4351

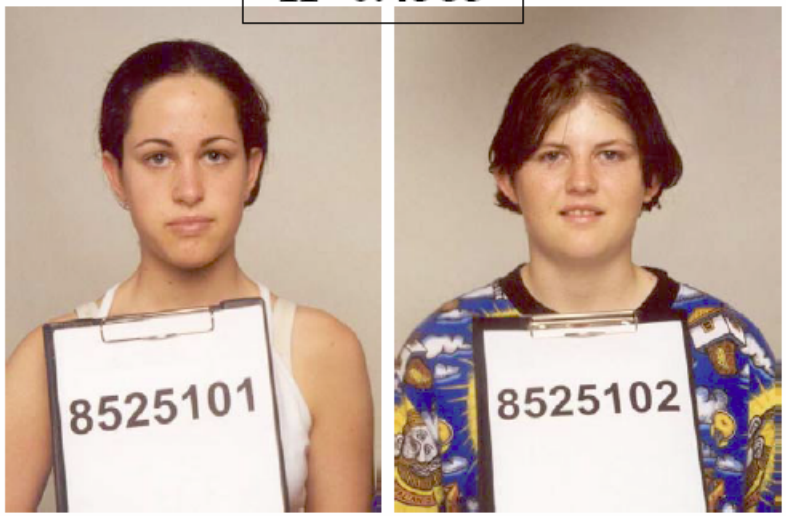


8069101,02
H=0.4291

8315101,02
H=0.4320



8525101,02
H=0.4385



People who should have got the Nobel prize and haven't (yet)...

- Morton for lod scores (1955)
- CAB Smith for linkage heterogeneity (50s)
- Kary Mullis for surfing [he got it for PCR]
- Litt & Weber for microsatellites ~1990
- Mike Neale for Mx
- Lon & Gonzalo for Merlin and QTDT

The official journal of the International Society for Twin Studies

twinresearch

Volume 6 Number 1

February 2003

Editor
Nicholas G. Martin