20th International Workshop on Methodology of Twin and Family Studies: Advanced course

- Lon Cardon (director) 
- Jonathan Flint 
- Jeff Barrett 
- David Evans 
- William Valdar 
- Goncalo Abecasis 
- Mike Neale 
- Hermine Maes 
- Sarah Medland 
- Dorret Boomsma 
- Danielle Posthuma 
- Meike Bartels

- John Hewitt (host) 
- Jeff Lessem 
- Matt McQueen 
- Pak Sham 
- Stacey Cherny 
- Ben Neale 
- Shaun Purcell 
- Manuel Ferreira 
- Nick Martin 
- Kate Morley 
- Marleen de Moor 
- Lannie Ligthart 

Hunting QTLs

Nick Martin

Queensland Institute of Medical Research

Boulder workshop: March 5, 2007

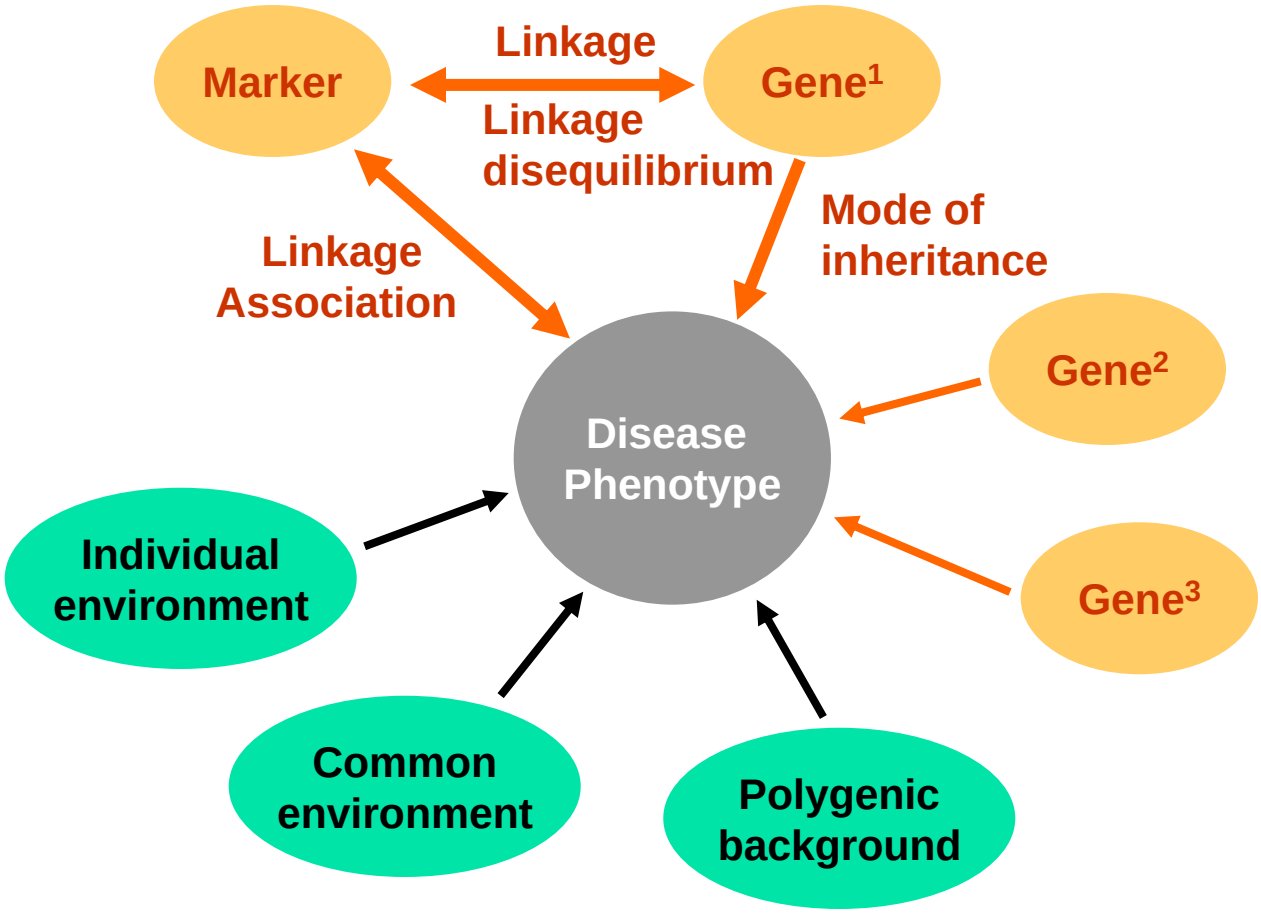
	Year	Location	Type	#Faculty	# Students
TC1	1987	Leuven	Introductory	10	24
TC2	1989	Leuven	Introductory	11	41
TC3	1990	Boulder	Introductory	11	28
TC4	1991	Leuven	Introductory	14	49
			Advanced	12	55
TC5	1993	Boulder	Introductory	13	49
TC6	1994	Boulder	Introductory	16	43
TC7	1995	Helsinki	Introductory	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	82
TCE1	*2003*	*Egmond*	*Introductory*	*15*	*65*
TC17	2004	Boulder	Introductory	18	90
TCE2	*2004*	*Egmond*	*Advanced**	*16*	*64*
TC18	2005	Boulder	Advanced	18	64
TCE3	*2005*	*Egmond*	*Advanced**	*13*	*55*
TC19	2006	Boulder	Introductory	15	93
TCE4	*2006*	*Egmond*	*Advanced*	*12*	*48*
TC20	2007	Boulder	Advanced	21	
TC21	2007	Leuven	Anniversary		

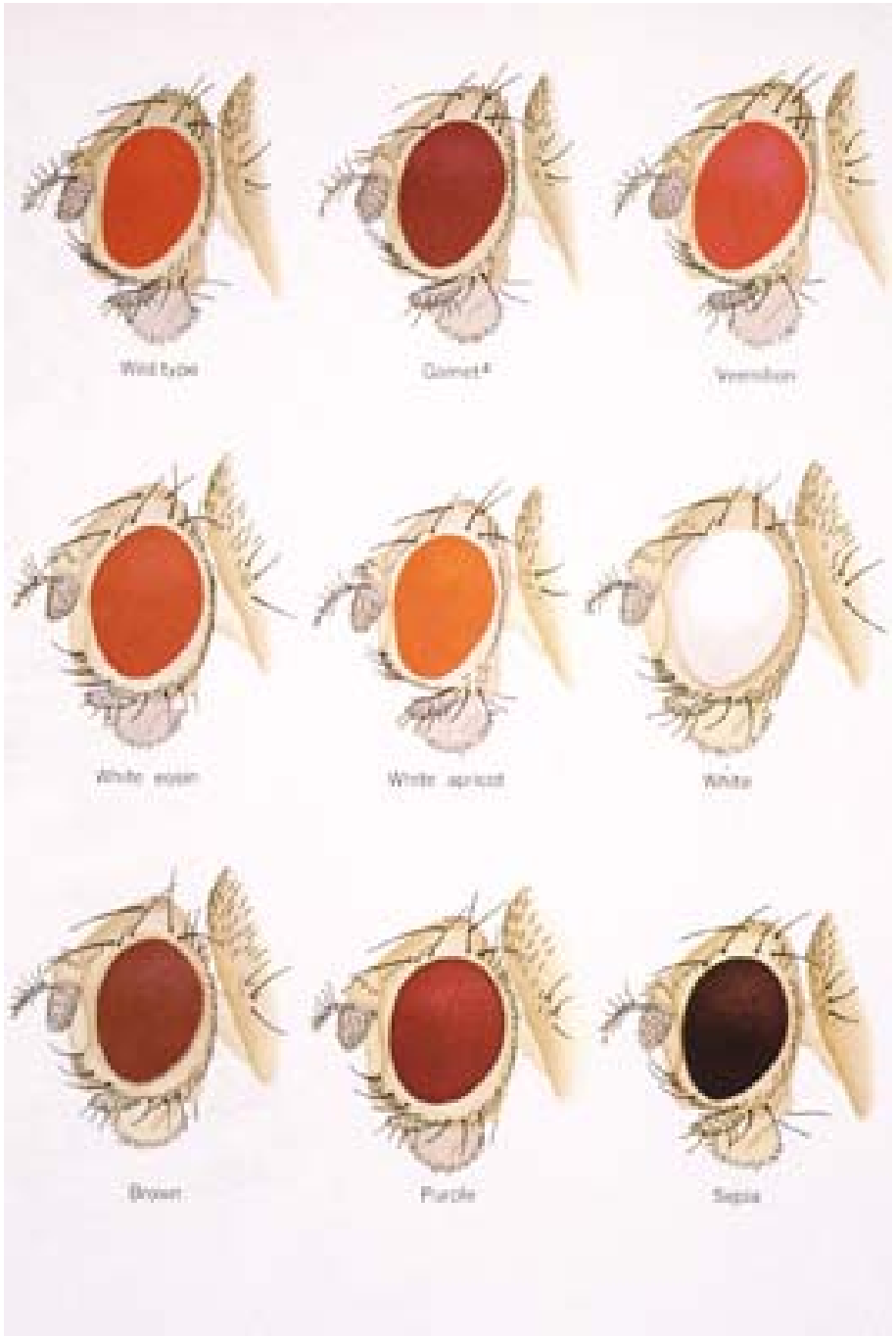
Frequency of attendance of faculty and students

Frequency	1	2	3	4	5	6	7	8	9	10	11	16	18	19	20	21	
Faculty	8	4	4	3	5	2	4	1	2	2	1	1	1	1	2	3	44
Student	585	169	36	14	4	5		1		# of 'Unique' Students							814

Introductory Workshop # of Students	920
Advanced Workshop # of Students	365
Total	1185

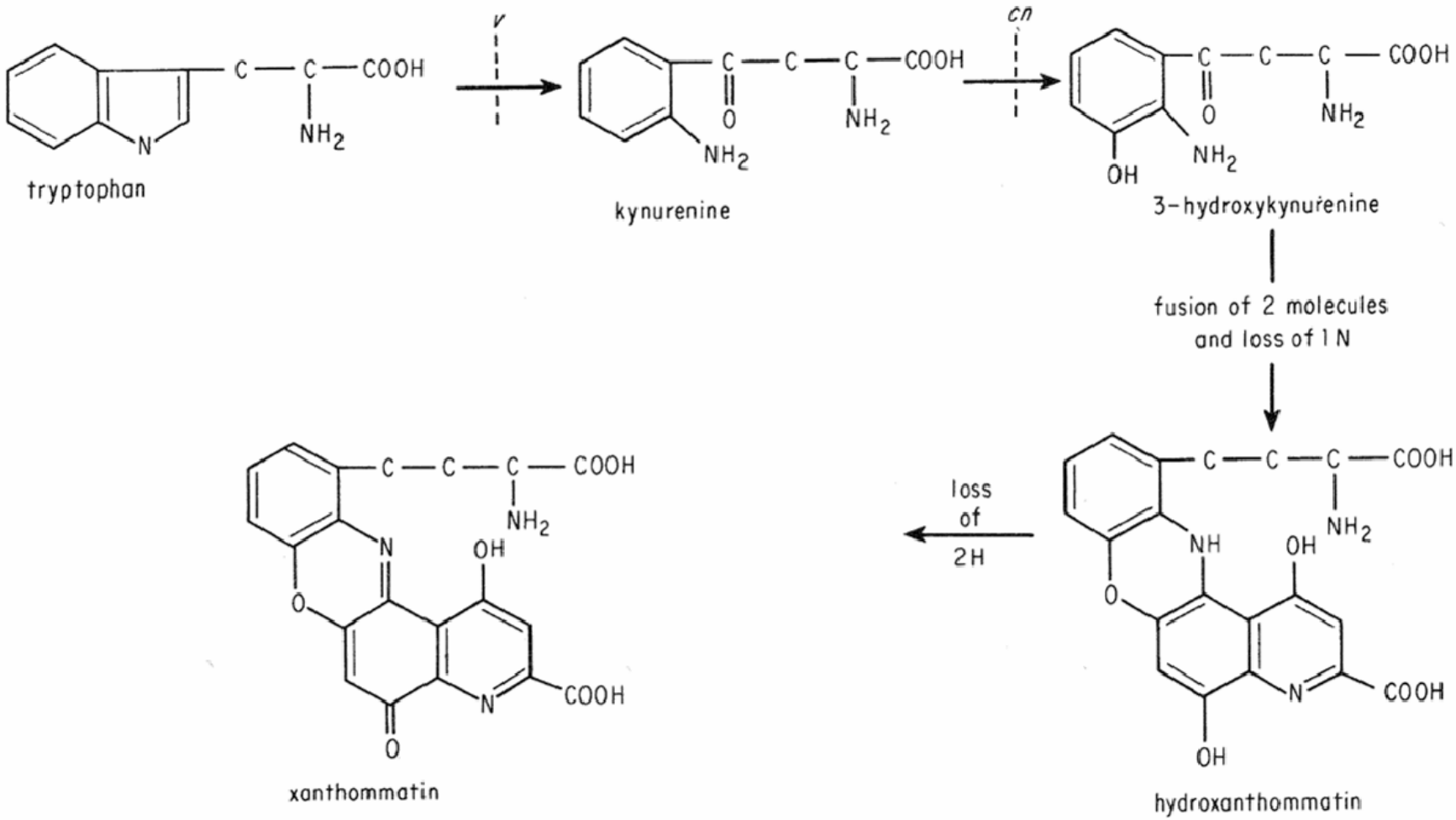
Complex Trait Model



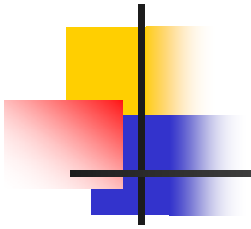


Using genetics to dissect metabolic pathways: Drosophila eye color

Beadle & Ephrussi, 1936



Beadle and Ephrussi, 1936



Finding QTLs

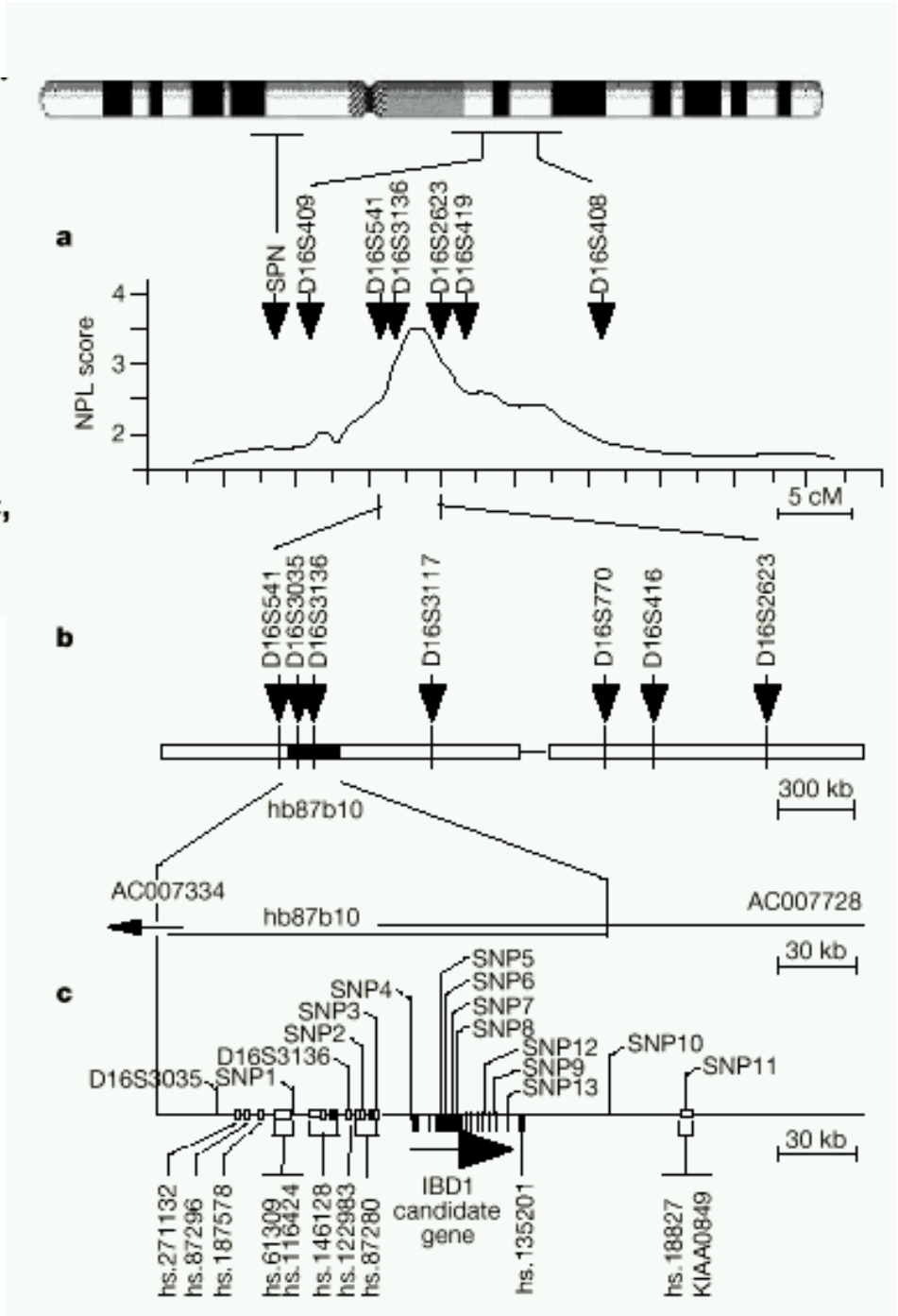
- Linkage
- Association

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

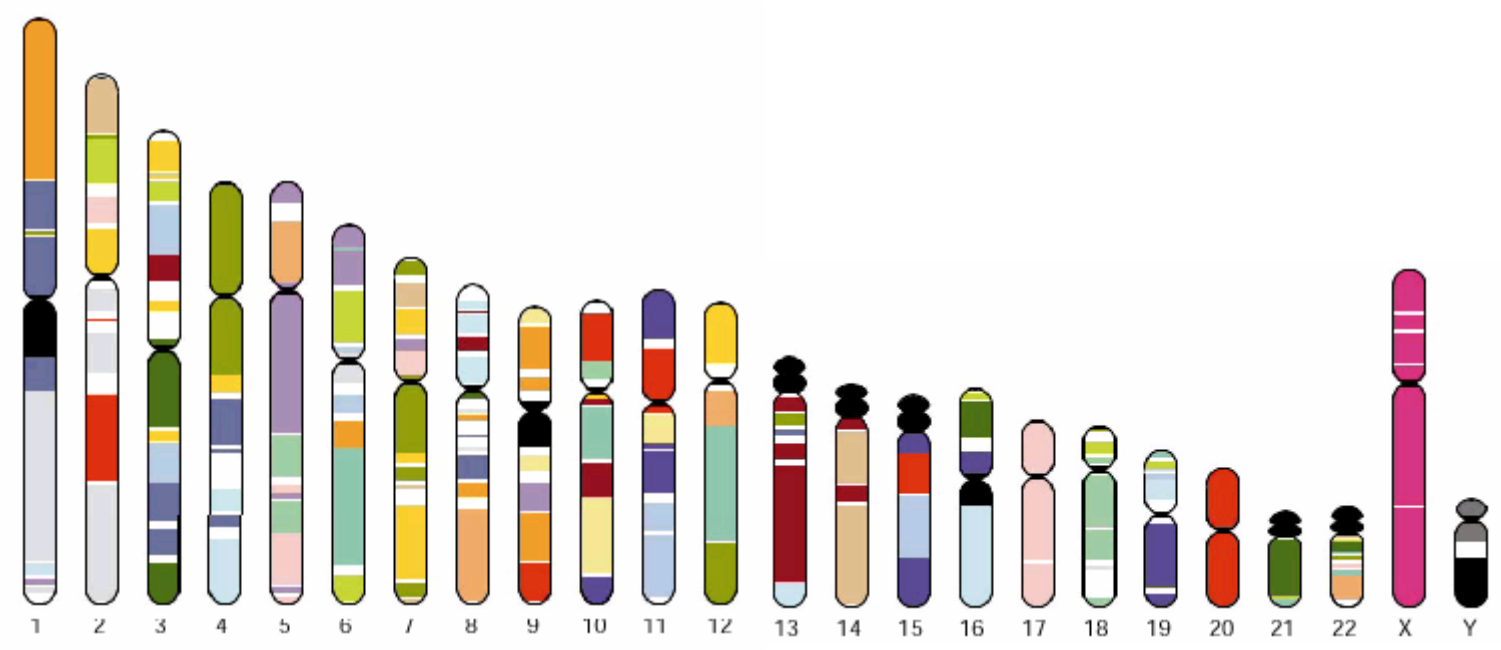
Jean-Pierre Hugot*†‡, Mathias Chamaillard*†, Habib Zouali*, Suzanne Lesage*, Jean-Pierre Cézard‡, Jacques Belaiche§, Sven Almer‖, Curt Tysk¶, Colm A. O'Morain#, Miquel Gassull☆, Vibeke Binder**, Yigael Finkel††, Antoine Cortot‡‡, Robert Modigliani§§, Pierre Laurent-Puig†, Corine Gower-Rousseau‡‡, Jeanne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

NATURE | VOL 411 | 31 MAY 2001

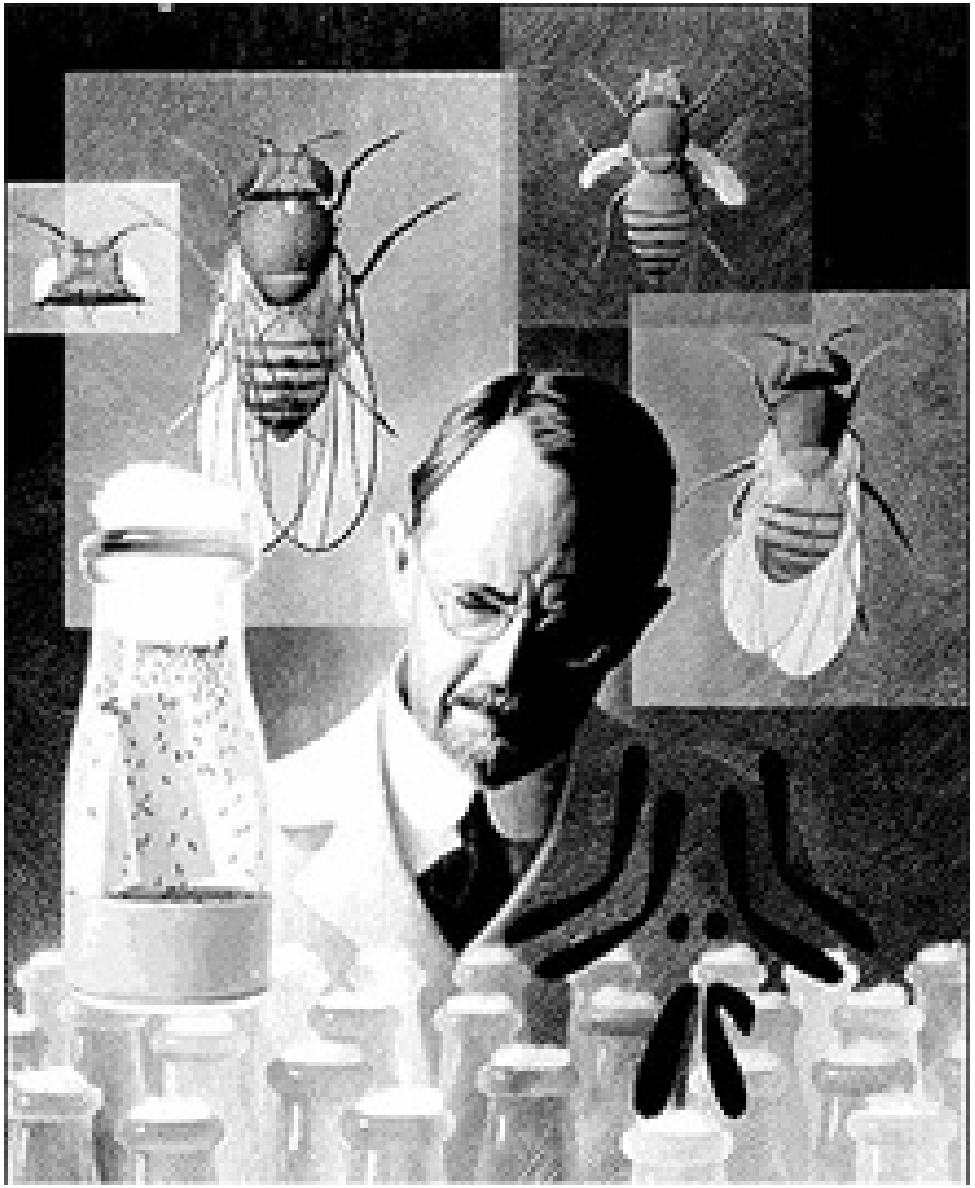
First (unequivocal) positional cloning of a complex disease QTL !



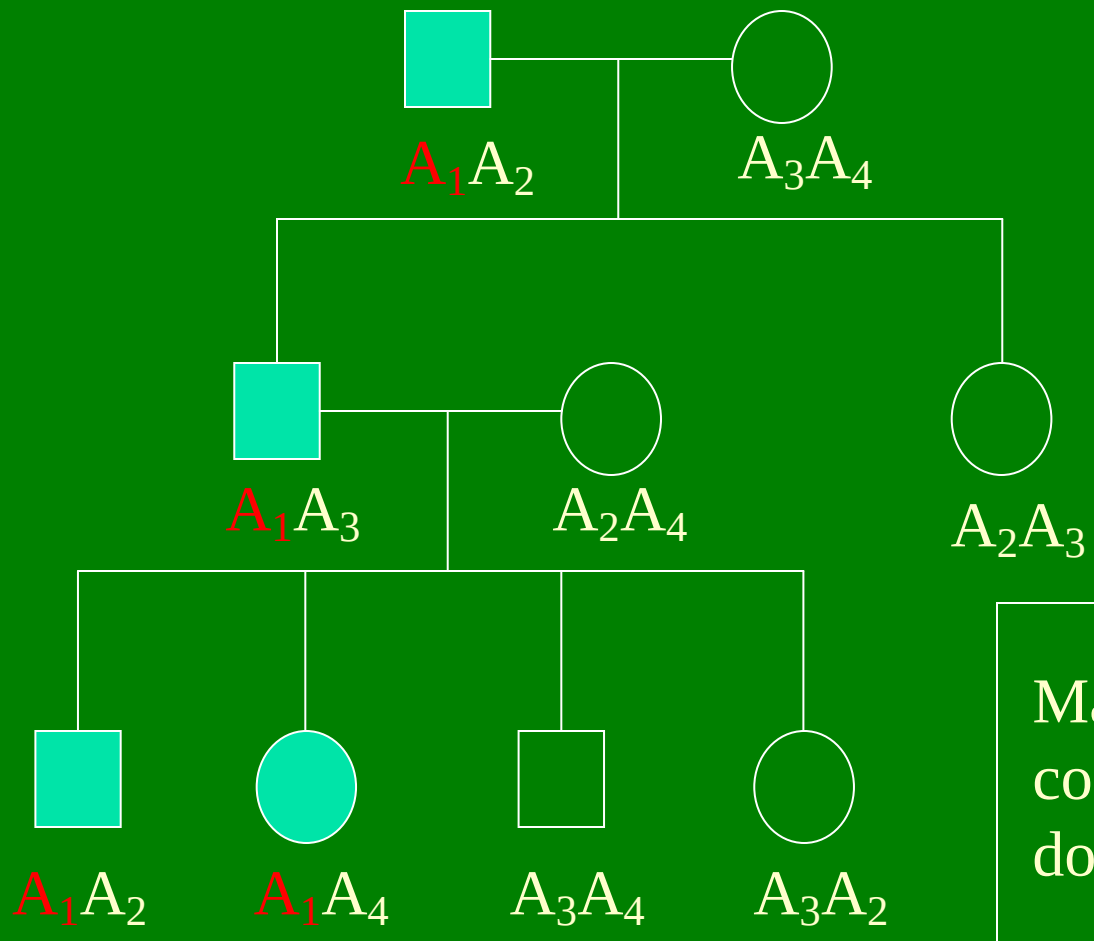
Linkage analysis



Thomas Hunt Morgan – discoverer of linkage

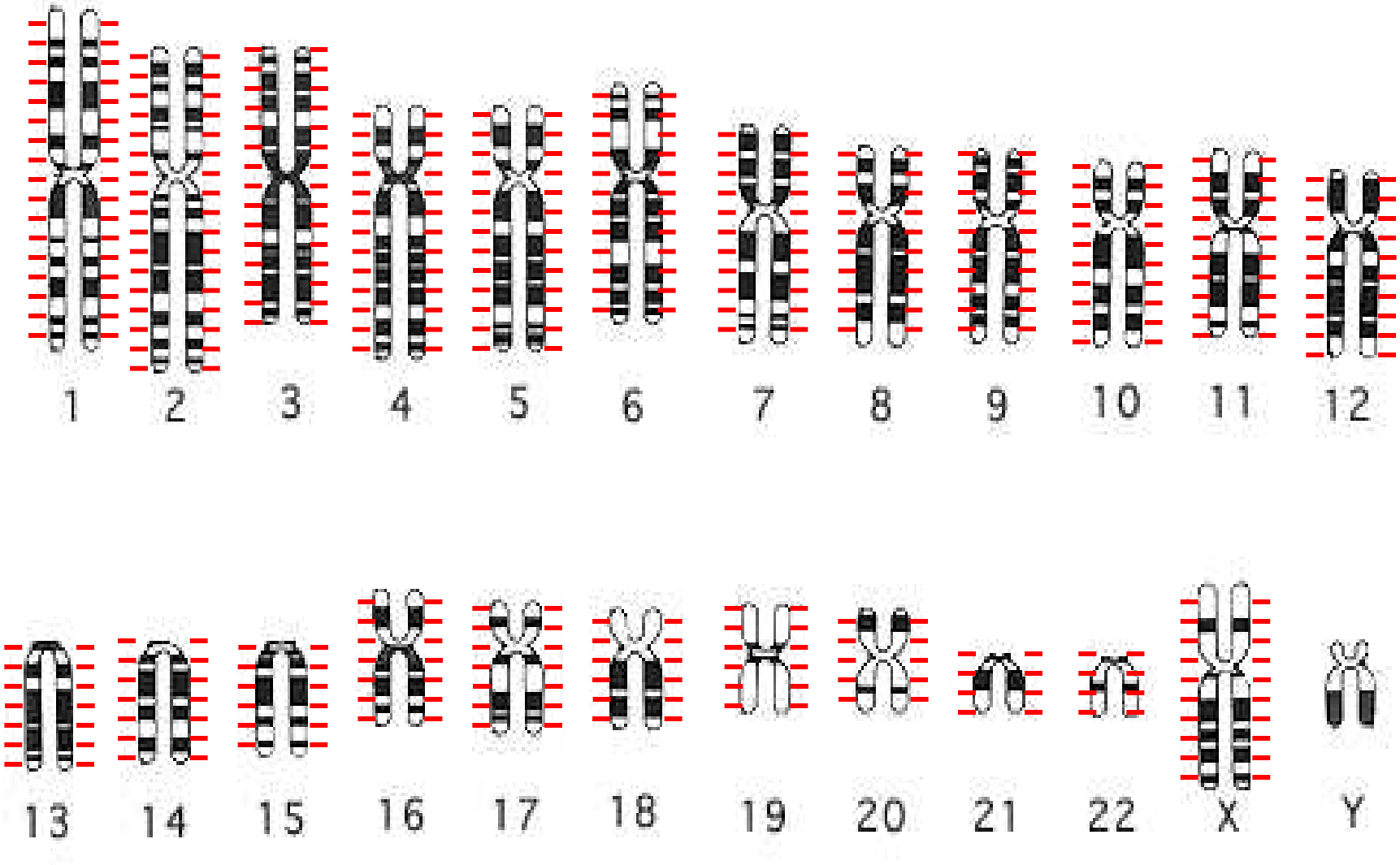


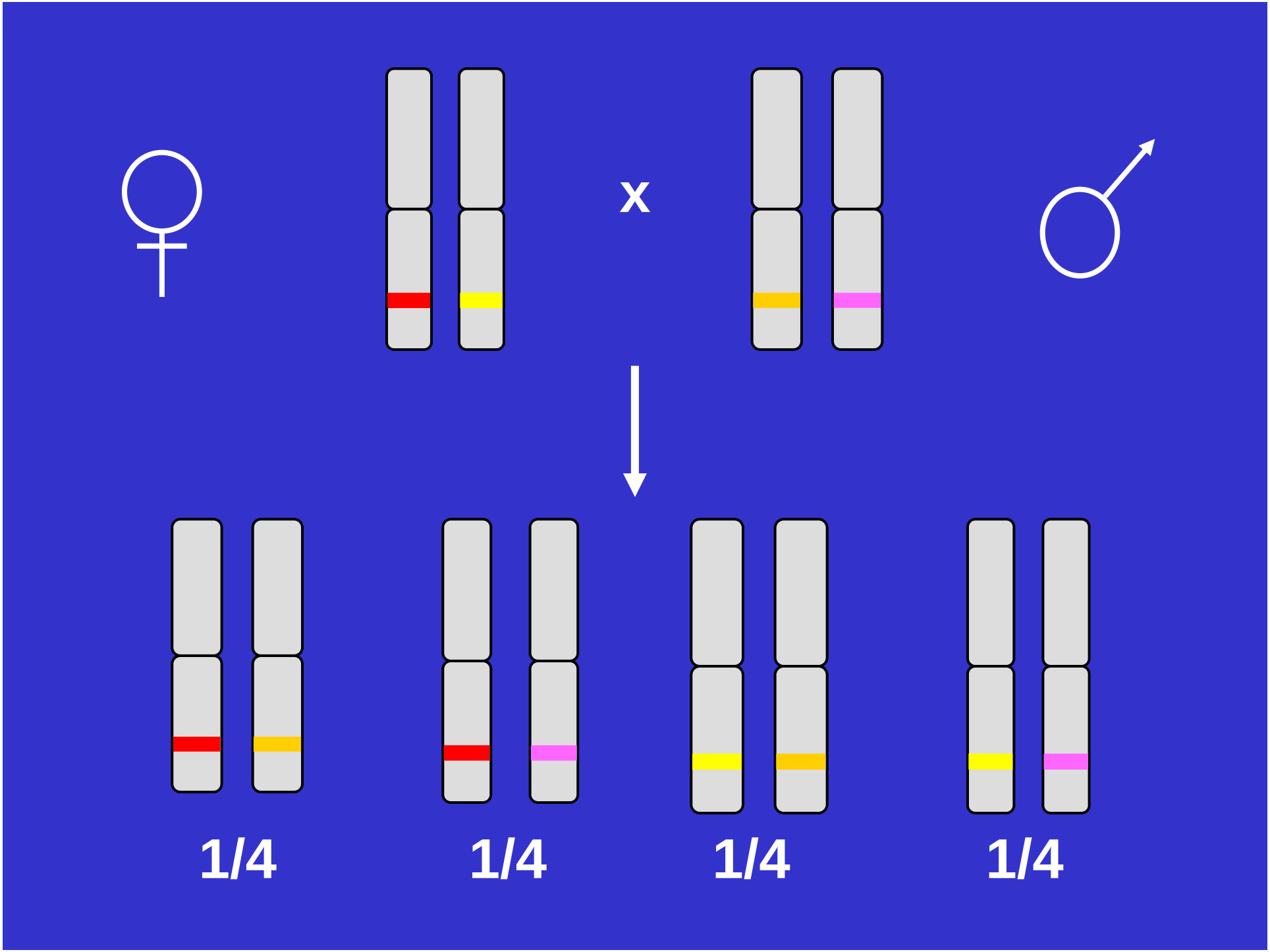
Linkage = Co-segregation



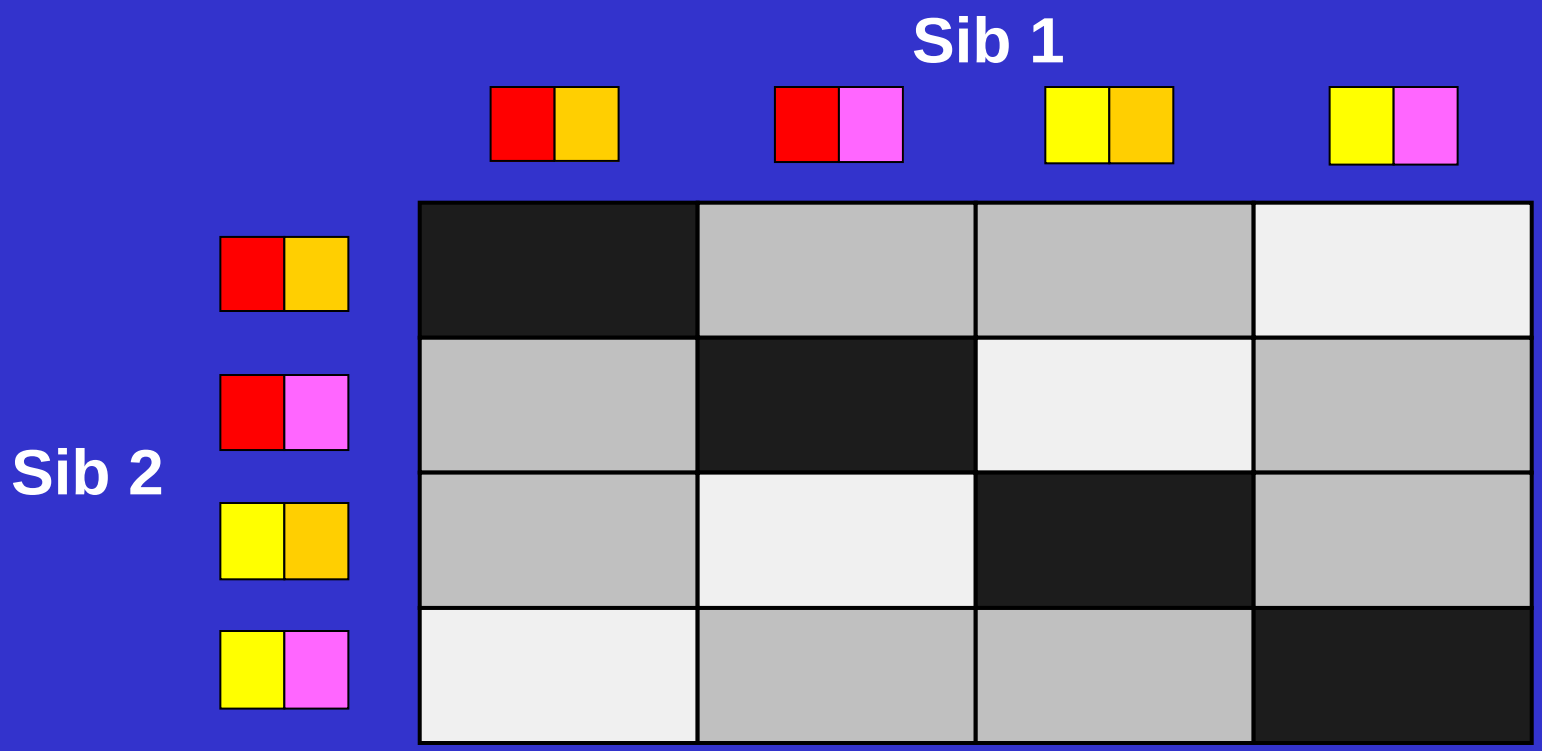
Marker allele A_1 cosegregates with dominant disease

Linkage Markers: microsatellite / SNP/ ...




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

IDENTITY BY DESCENT



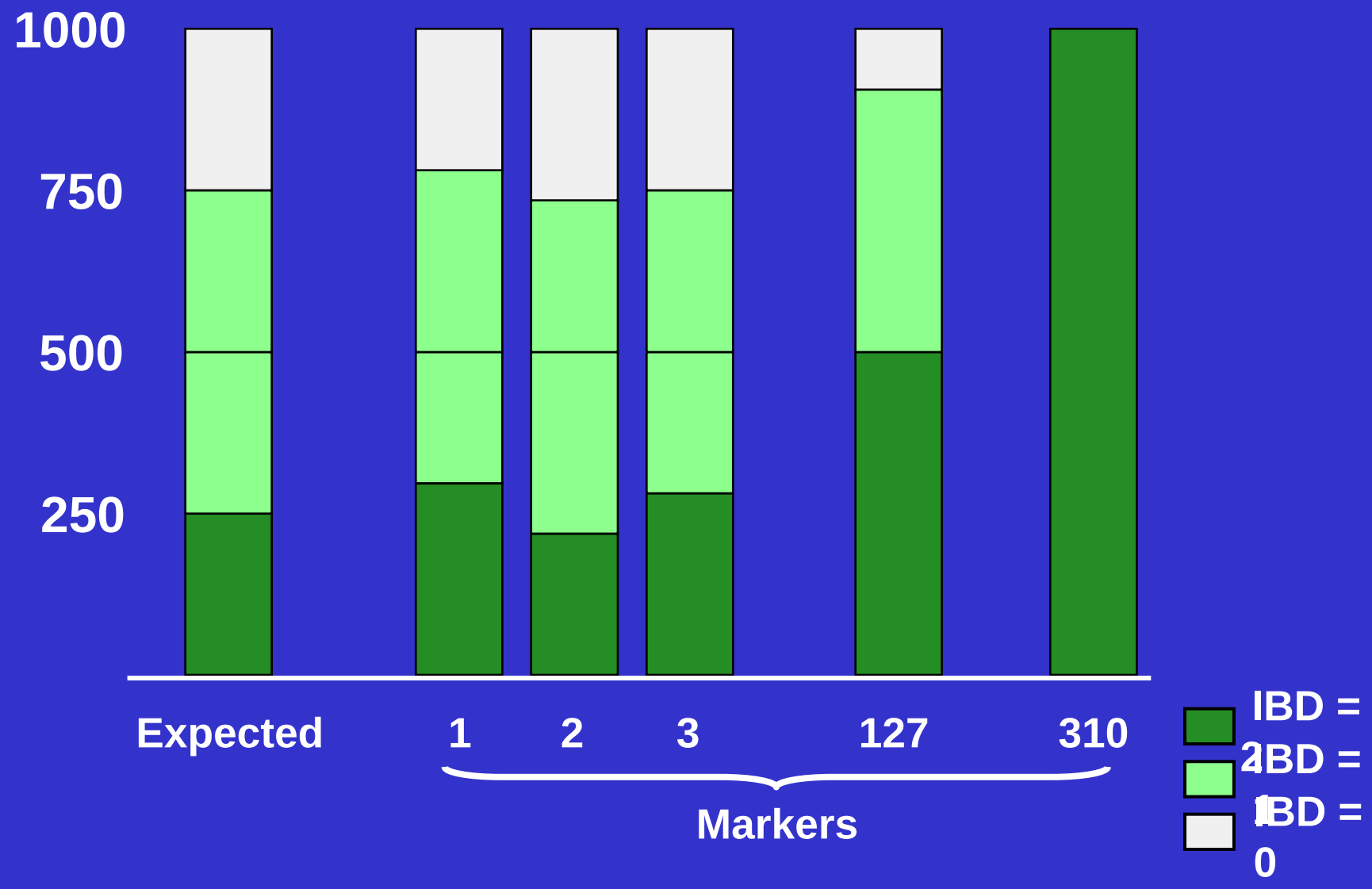
4/16 = 1/4 sibs share BOTH parental alleles IBD = 2

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

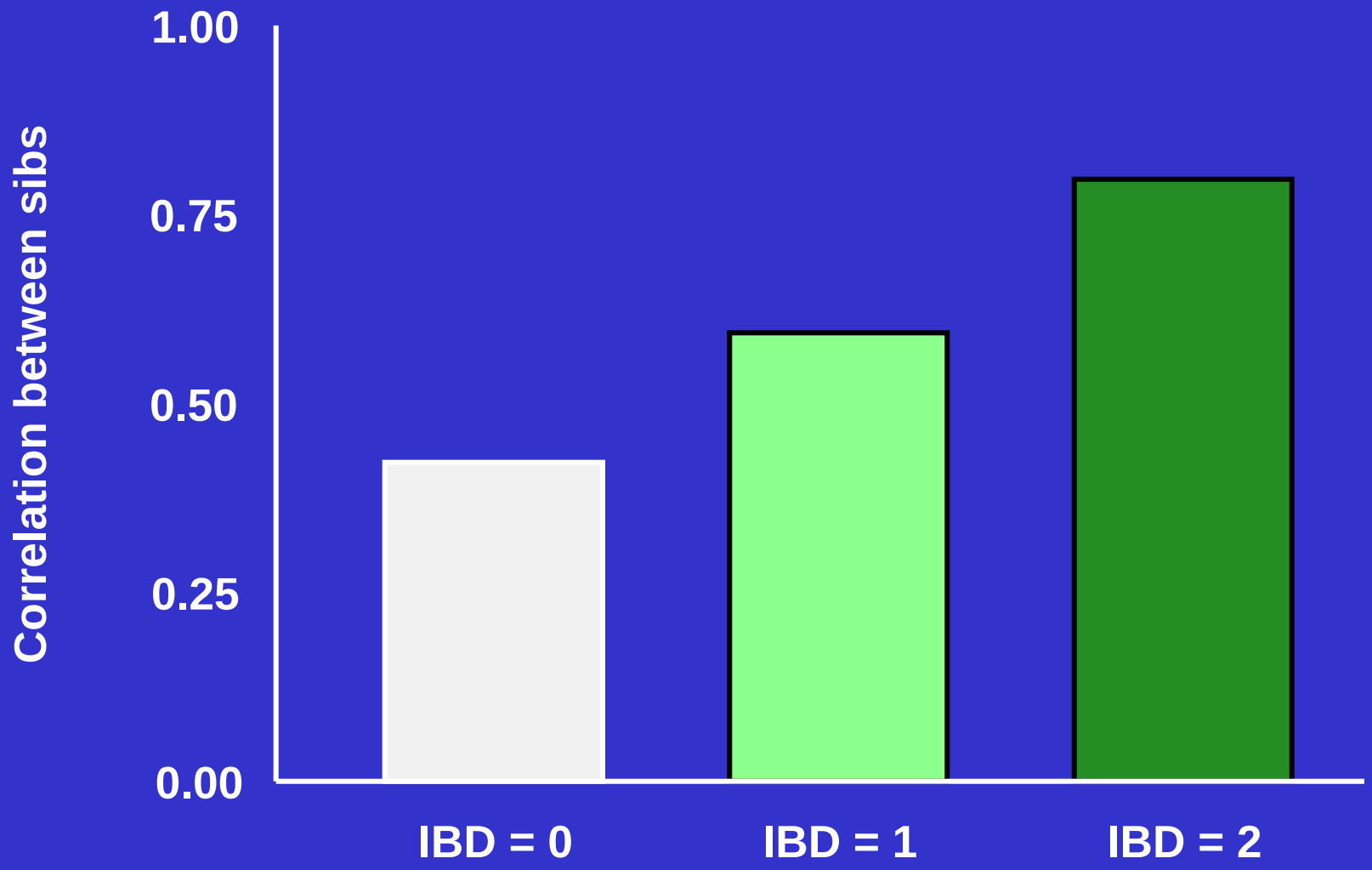
4/16 = 1/4 sibs share NO parental alleles IBD = 0

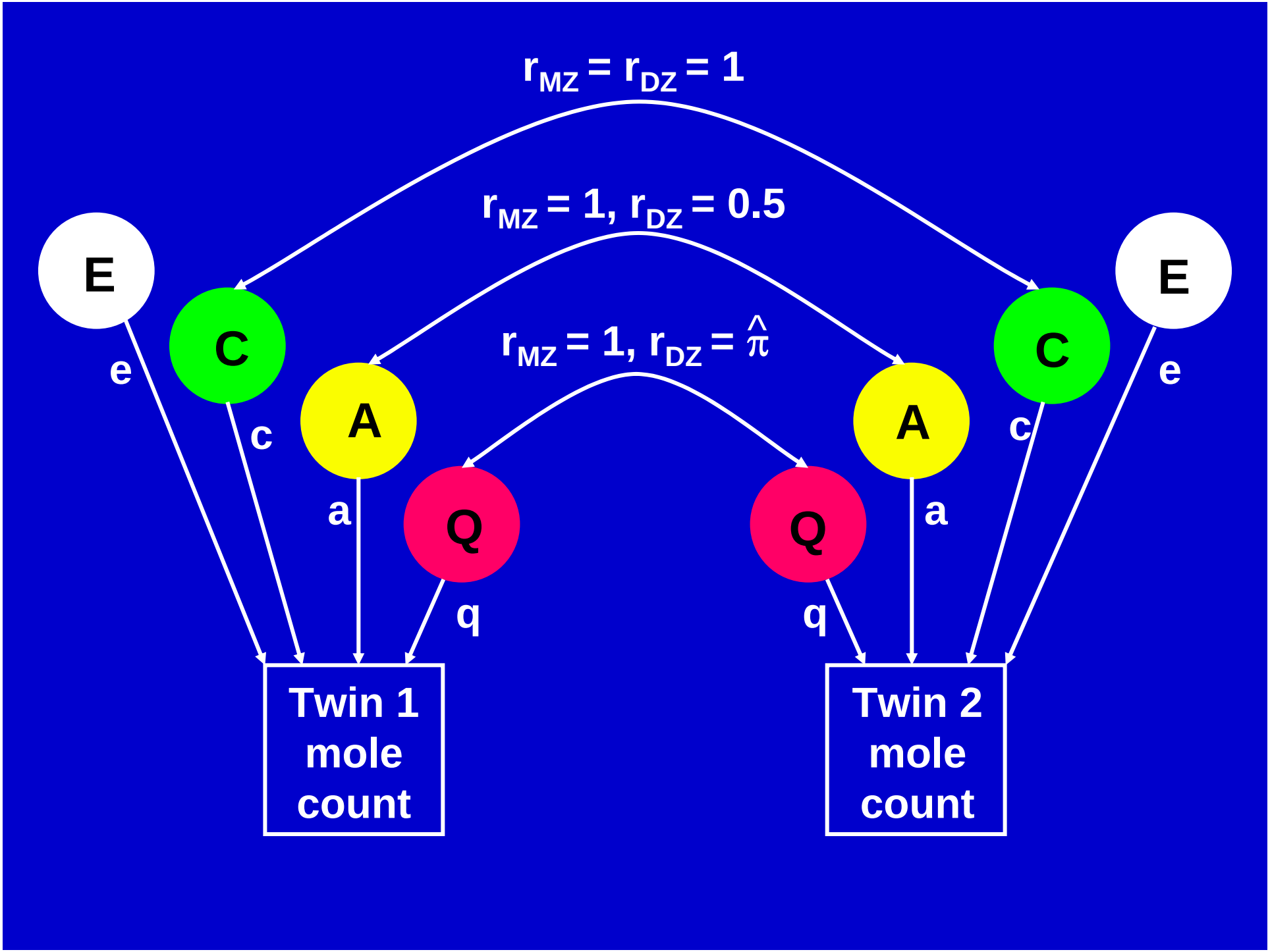
For disease traits (affected/unaffected)

Affected sib pairs selected

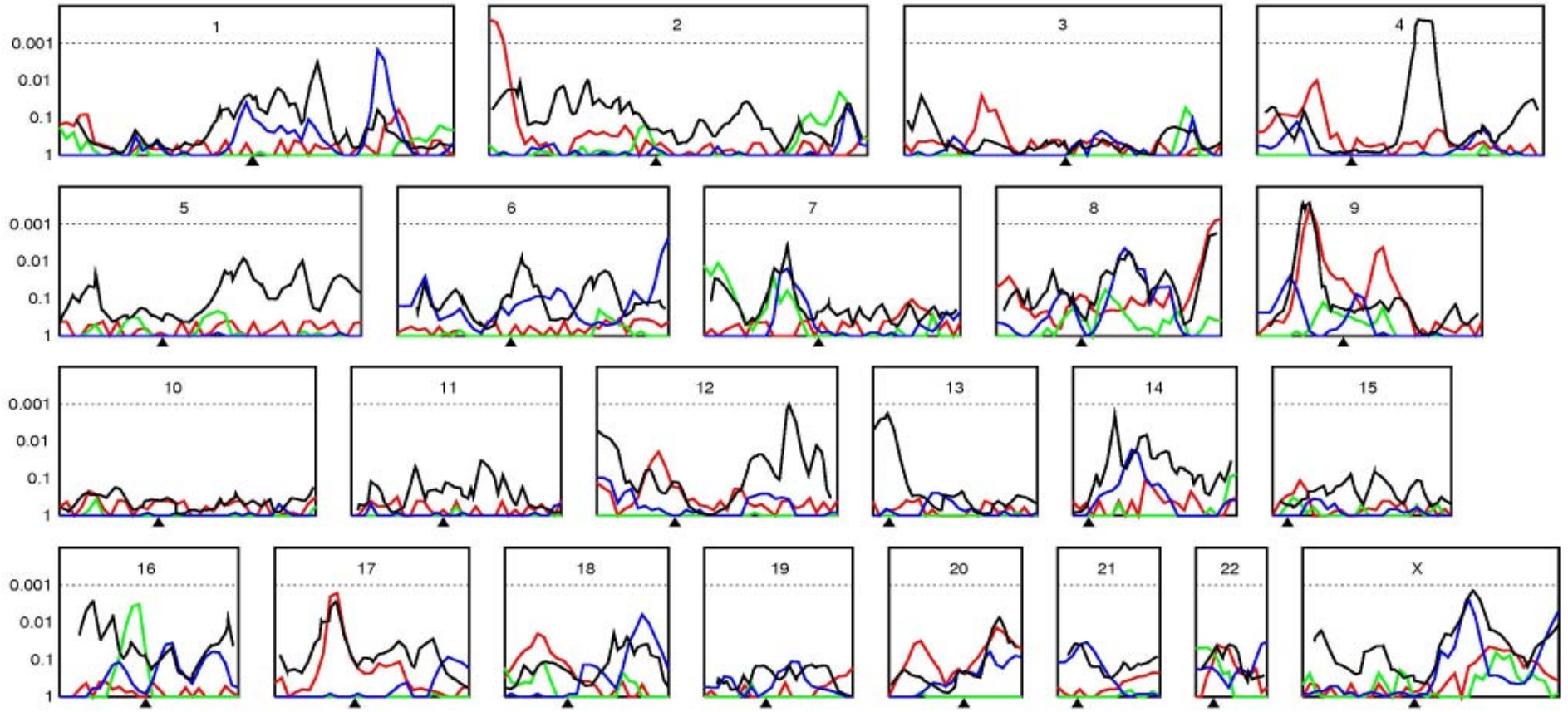


For continuous measures
Unselected sib pairs




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

Linkage for mole counts in Australian twin families

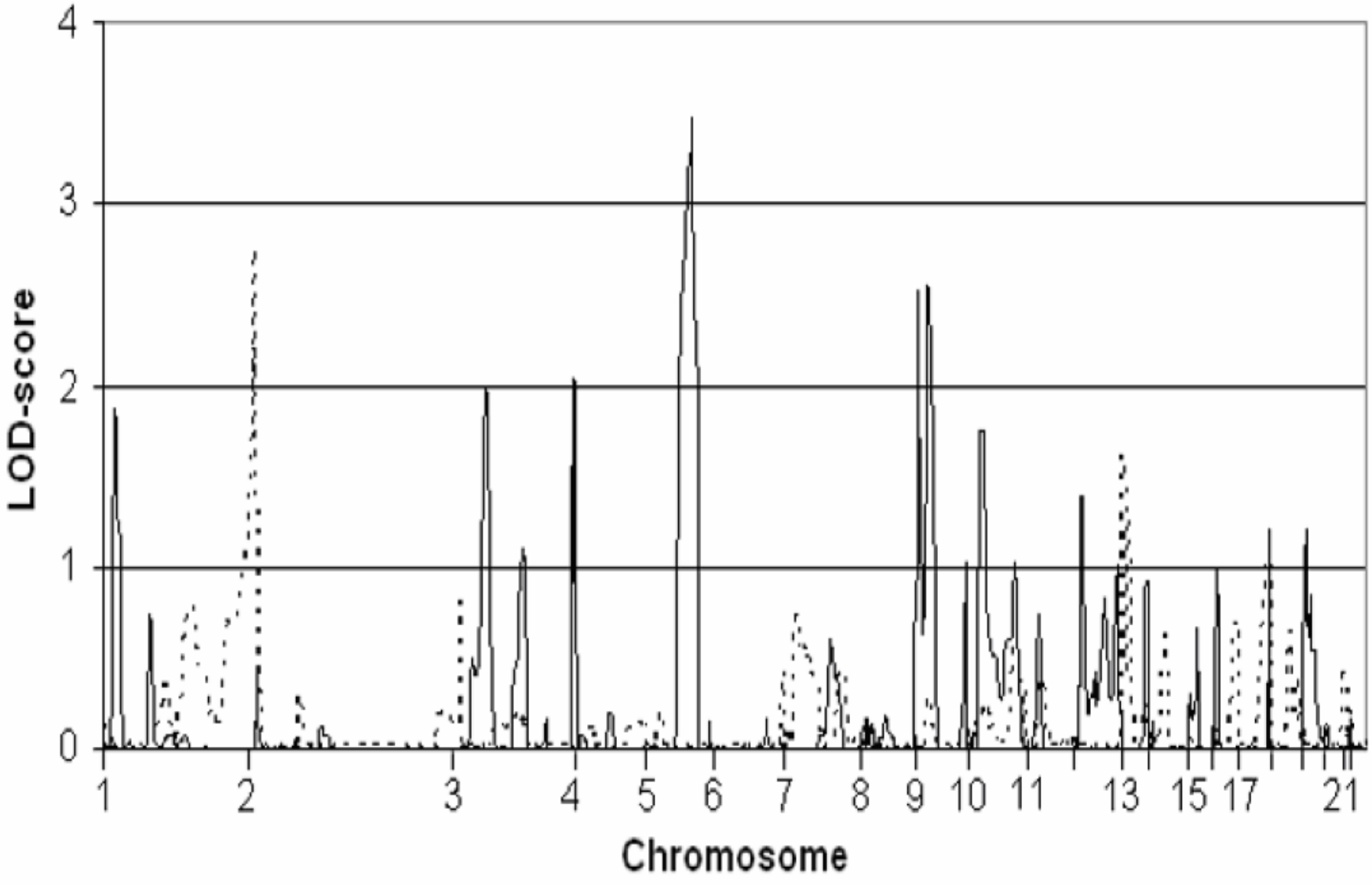


A genome-wide scan for naevus count: linkage to CDKN2A and to other chromosome regions

Gu Zhu[1], Grant W Montgomery[1], Michael R James[1], Jeff M Trent[2], Nicholas K Hayward[1], Nicholas G Martin[1] and David L Duffy*,[1]

European Journal of Human Genetics (2007) 15, 94–102

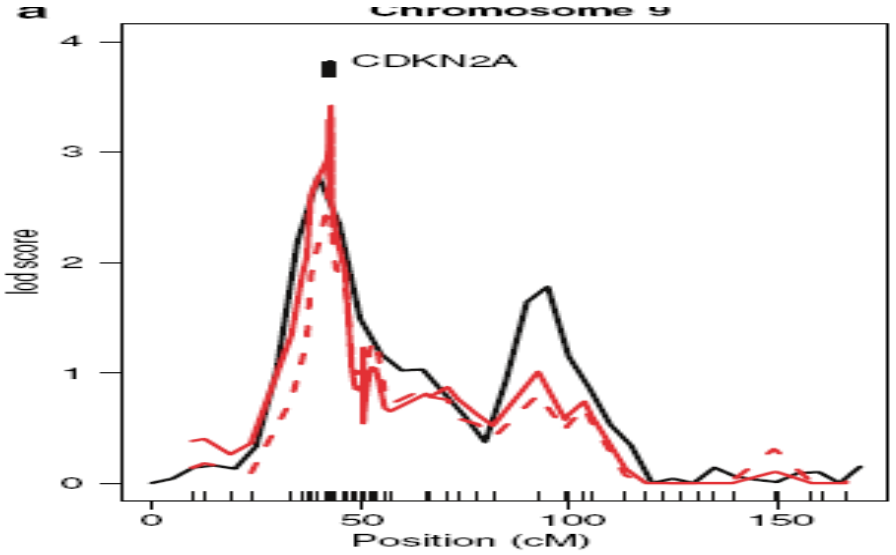
Linkage for mole counts in UK DZ twins



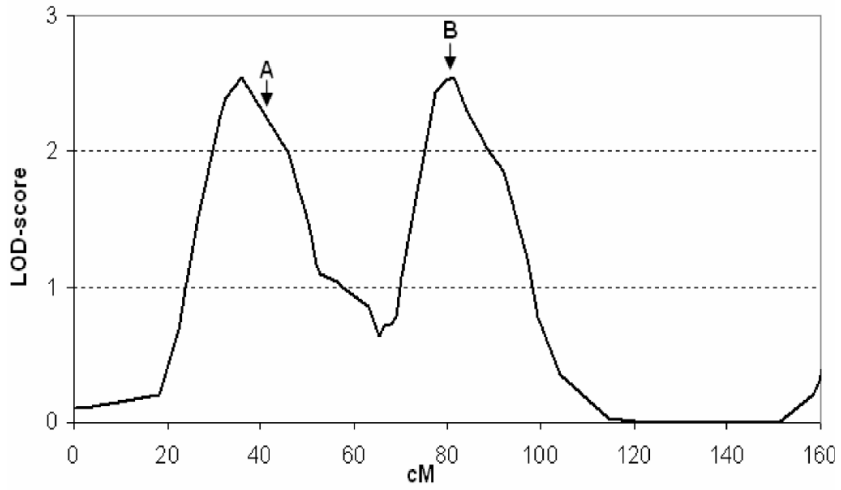
Genome-wide search for nevus density shows linkage to two melanoma loci on chromosome 9 and identifies a new QTL on 5q31 in an adult twin cohort.Falchi M, Spector TD, Perks U, Kato BS, Bataille V. Hum Mol Genet. 2006 Oct 15;15(20):2975-9

Flat mole count: chromosome 9 linkage in Australian and UK twins

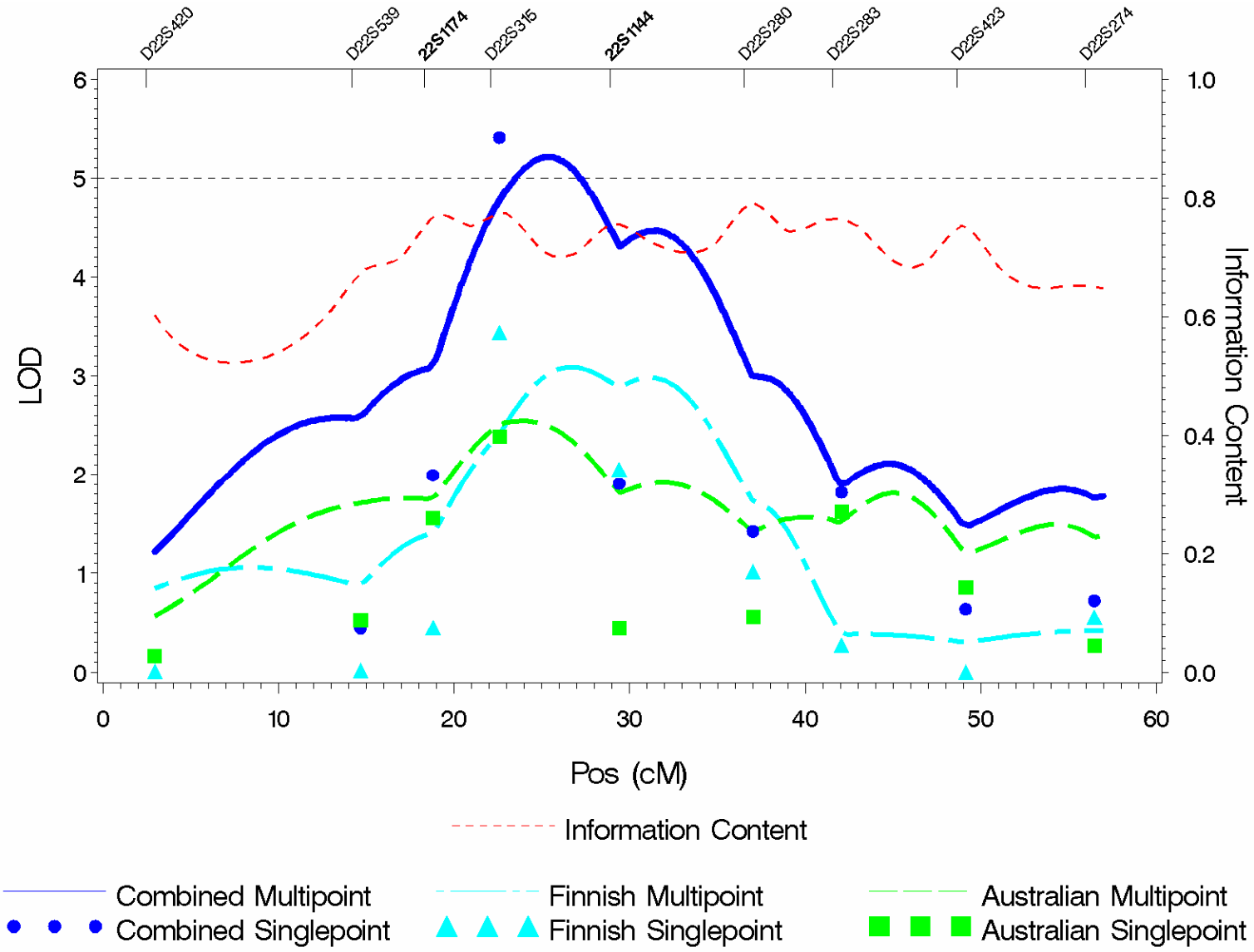
Australia



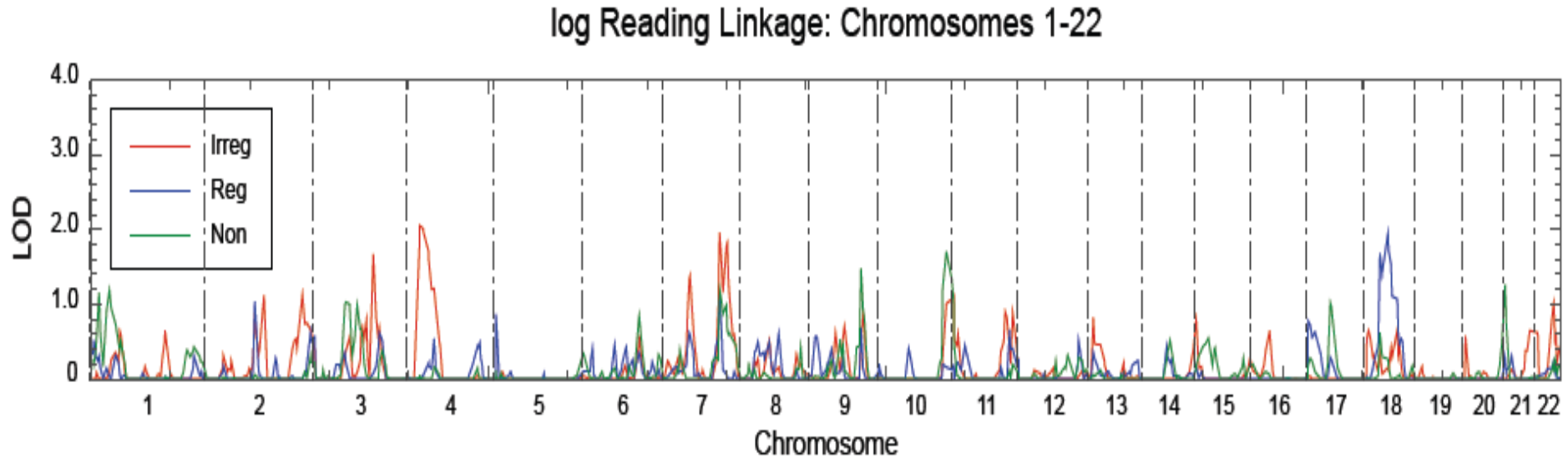
UK

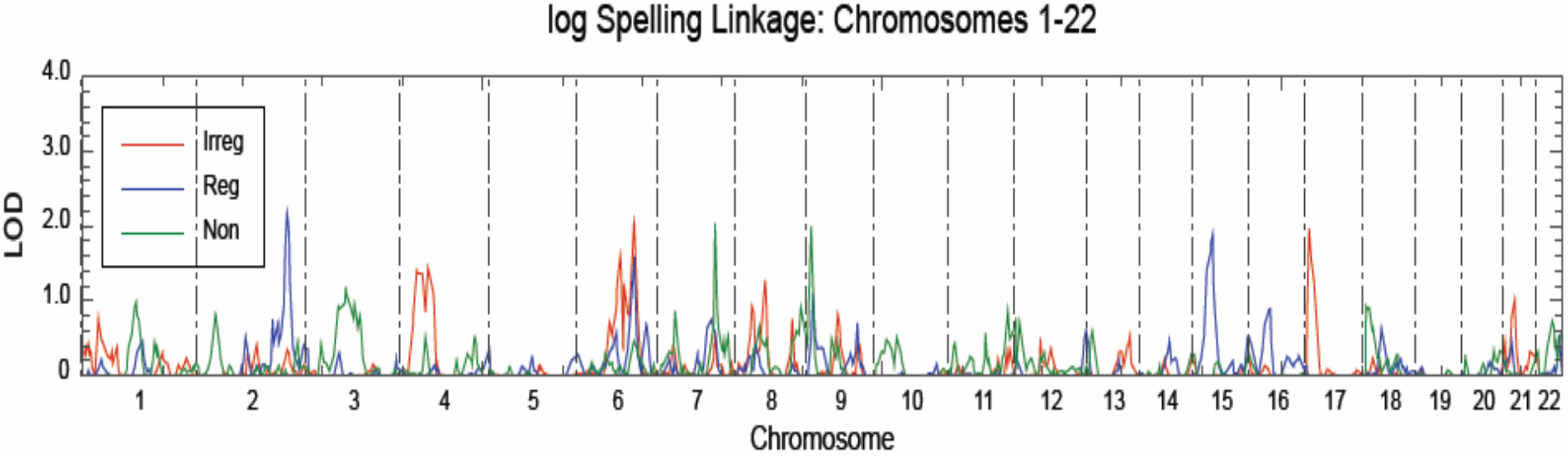


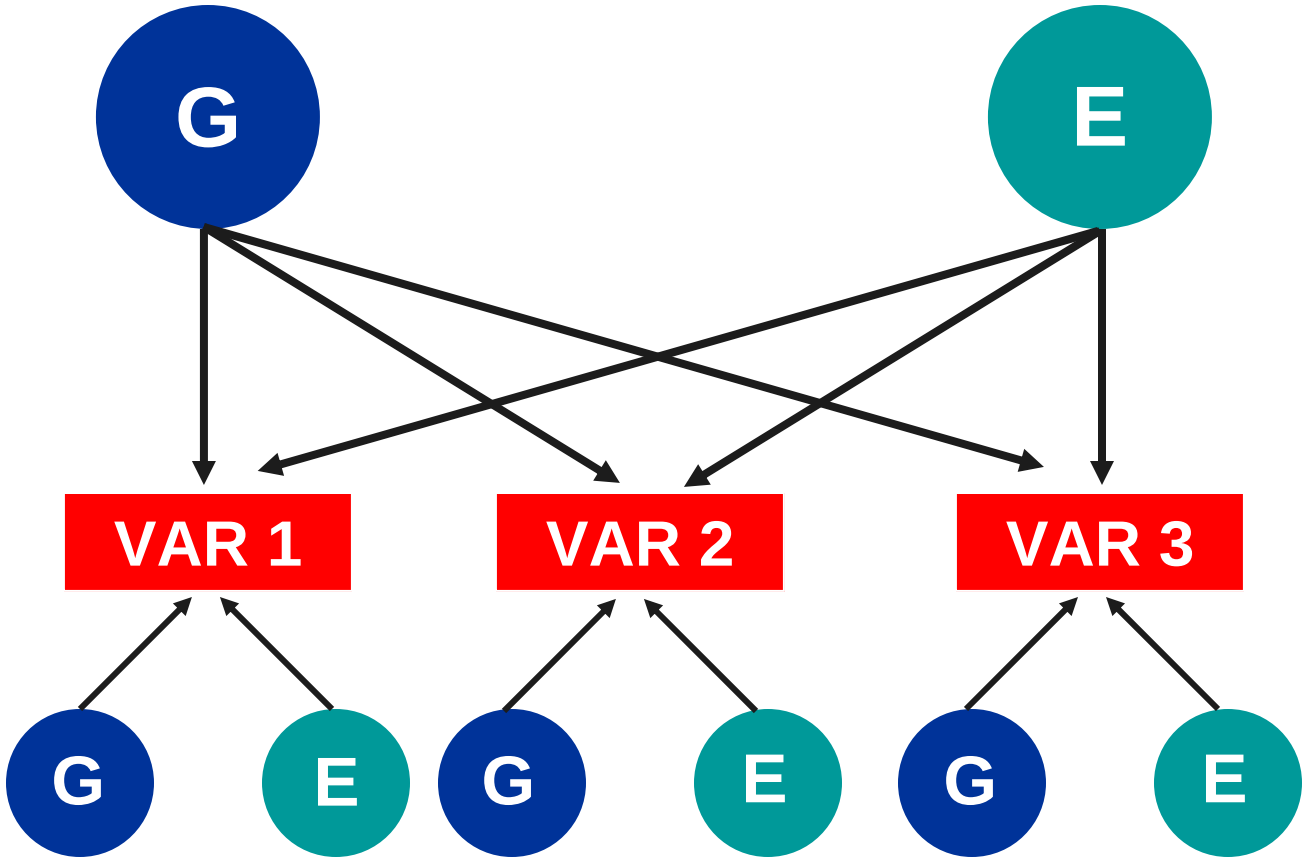
Linkage for MaxCigs24 in Australia and Finland



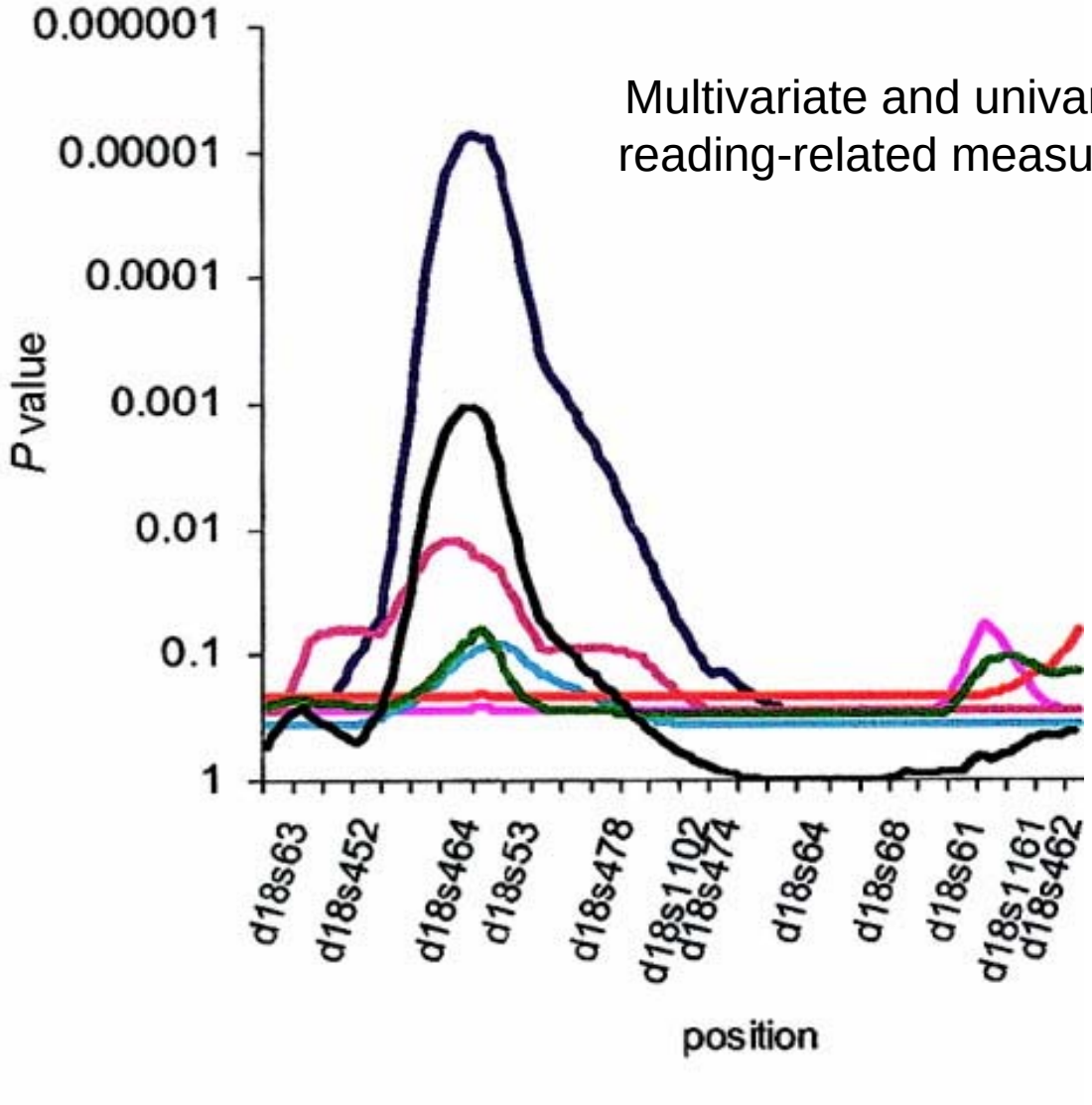
AJHG, in press


log Reading Linkage: Chromosomes 1-22
LOD
4.0
3.0
2.0
1.0
0
Irreg
Reg
Non
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Chromosome


log Spelling Linkage: Chromosomes 1-22
LOD
4.0
3.0
2.0
1.0
0
Irreg
Reg
Non
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Chromosome


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

House dust mite
Atopy
Bronchial Hyperrespons.
FEV_1
Asthma
Airway Obstruction

Robust Estimation of Experimentwise *P* Values Applied to a Genome Scan of Multiple Asthma Traits Identifies a New Region of Significant Linkage on Chromosome 20q13

Manuel A. R. Ferreira,[1] Louise O'Gorman,[1] Peter Le Souëf,[2] Paul R. Burton,[2] Brett G. Toelle,[3] Colin F. Robertson,[4] Peter M. Visscher,[1] Nicholas G. Martin,[1] and David L. Duffy[1]

Am. J. Hum. Genet. 77:1075–1085, 2005

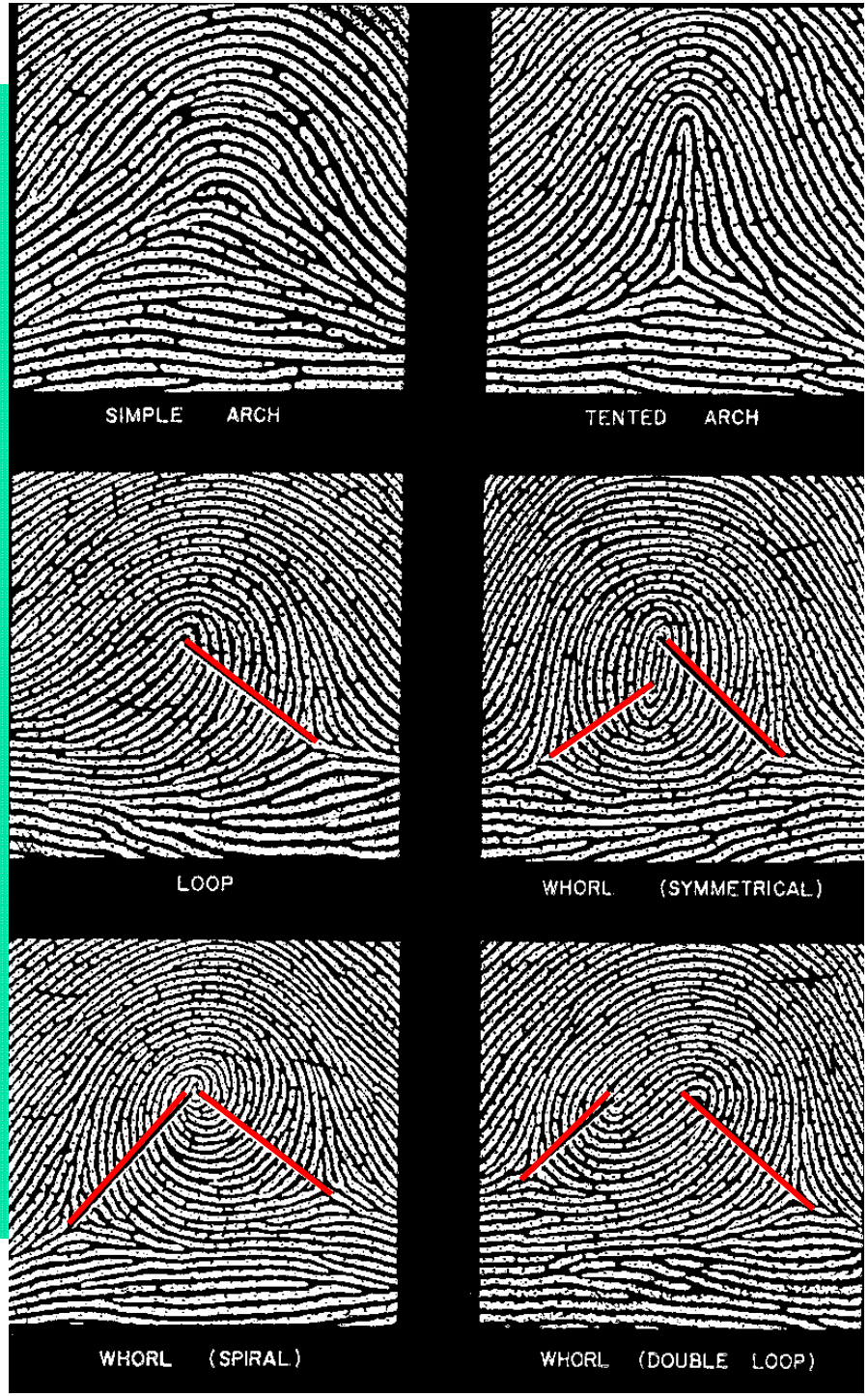
Ridge count

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

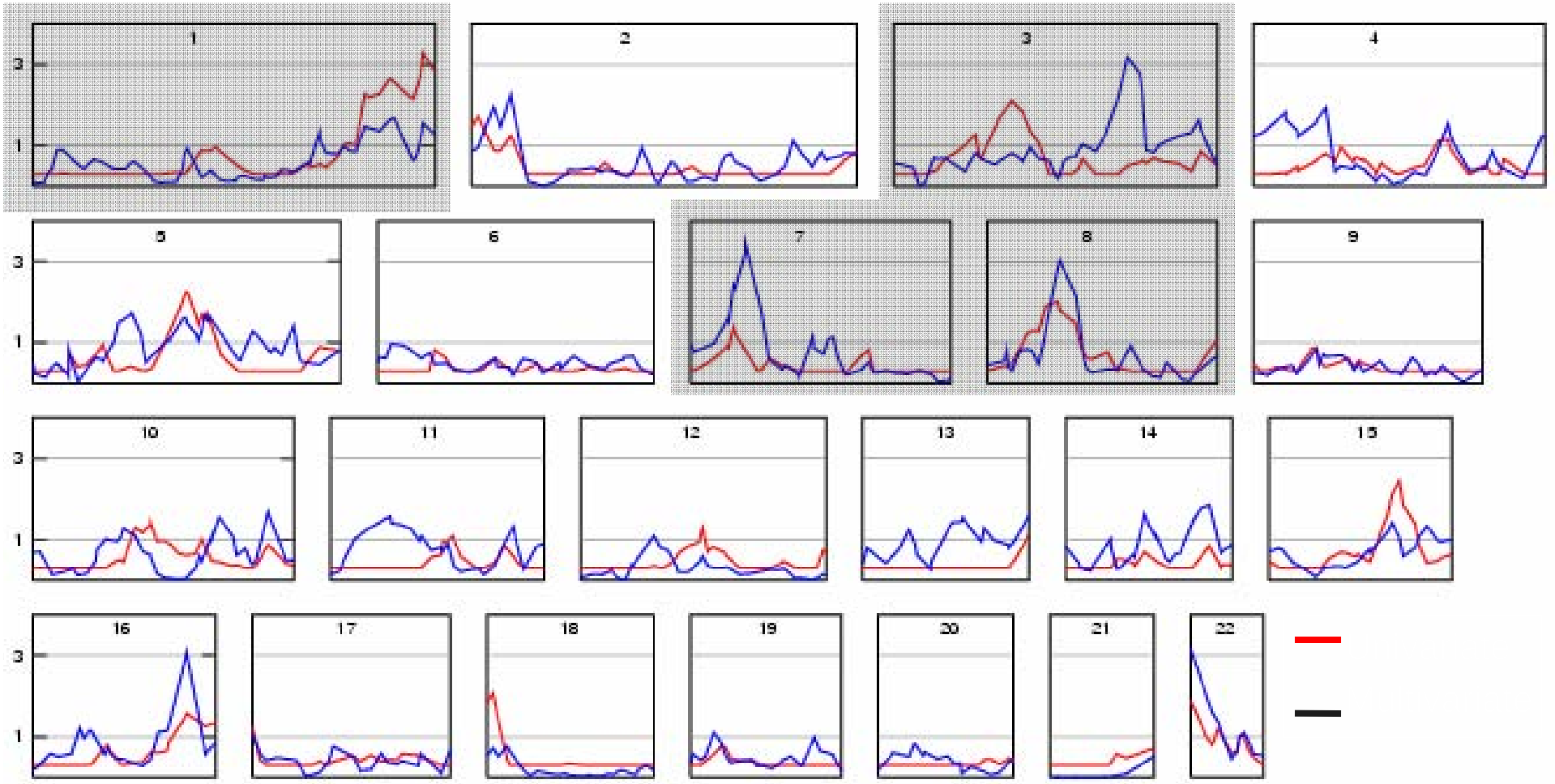
Ridge count can be summed over all fingers to give a total ridge count

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 ($-LOG_{10}p$)

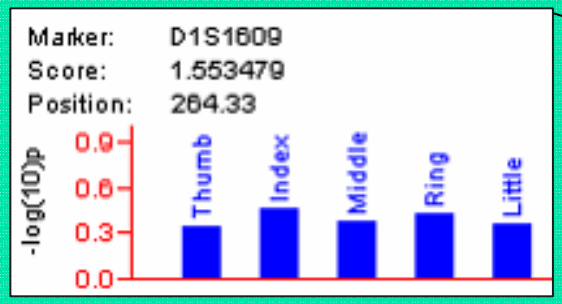


χ_1^2 χ_5^2

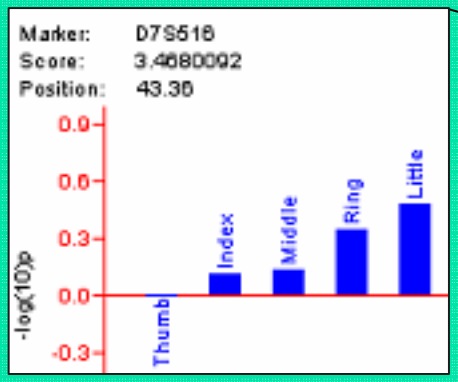
Chromosome 1

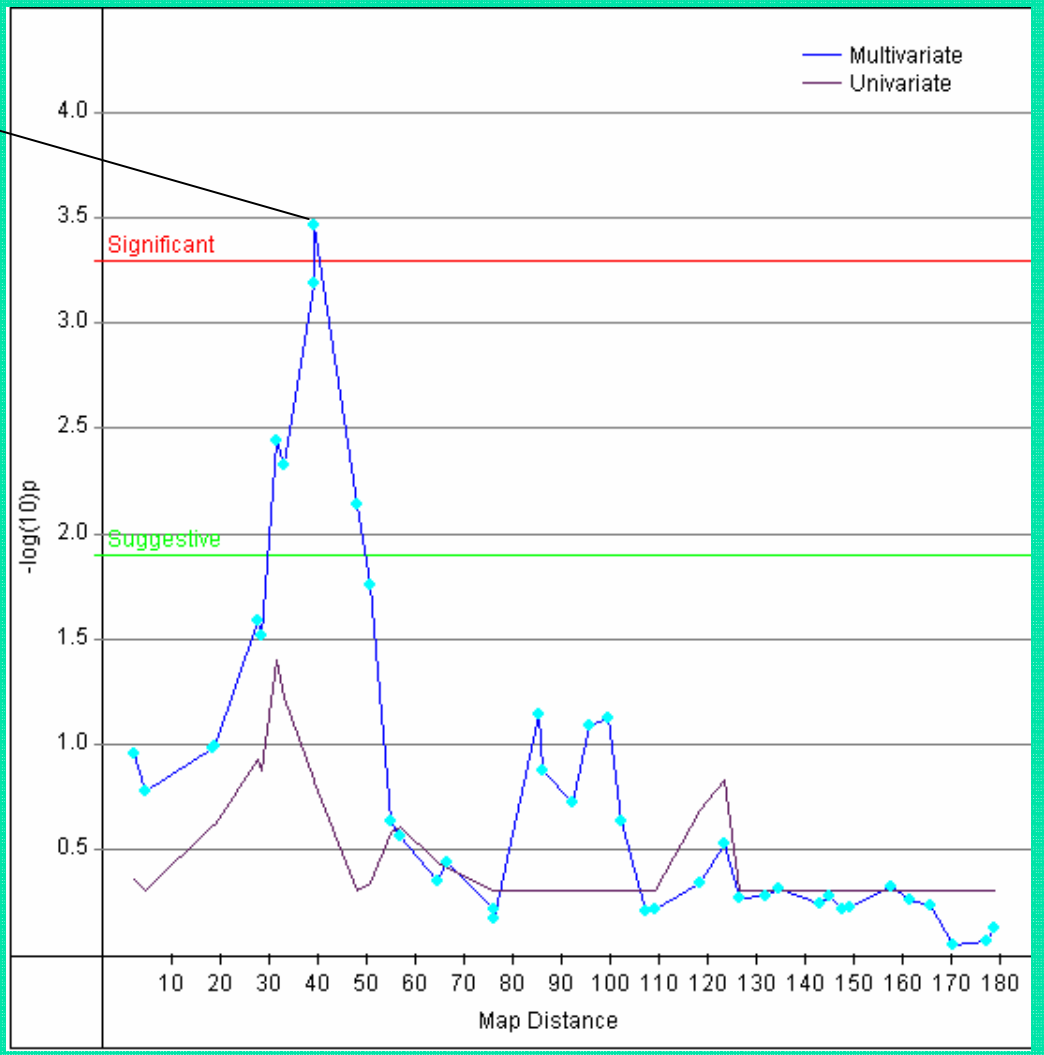
Similar ‘drop chi-squares’ for pleiotropic QTLs

Resulting in a very conservative test



Chromosome 7 ...



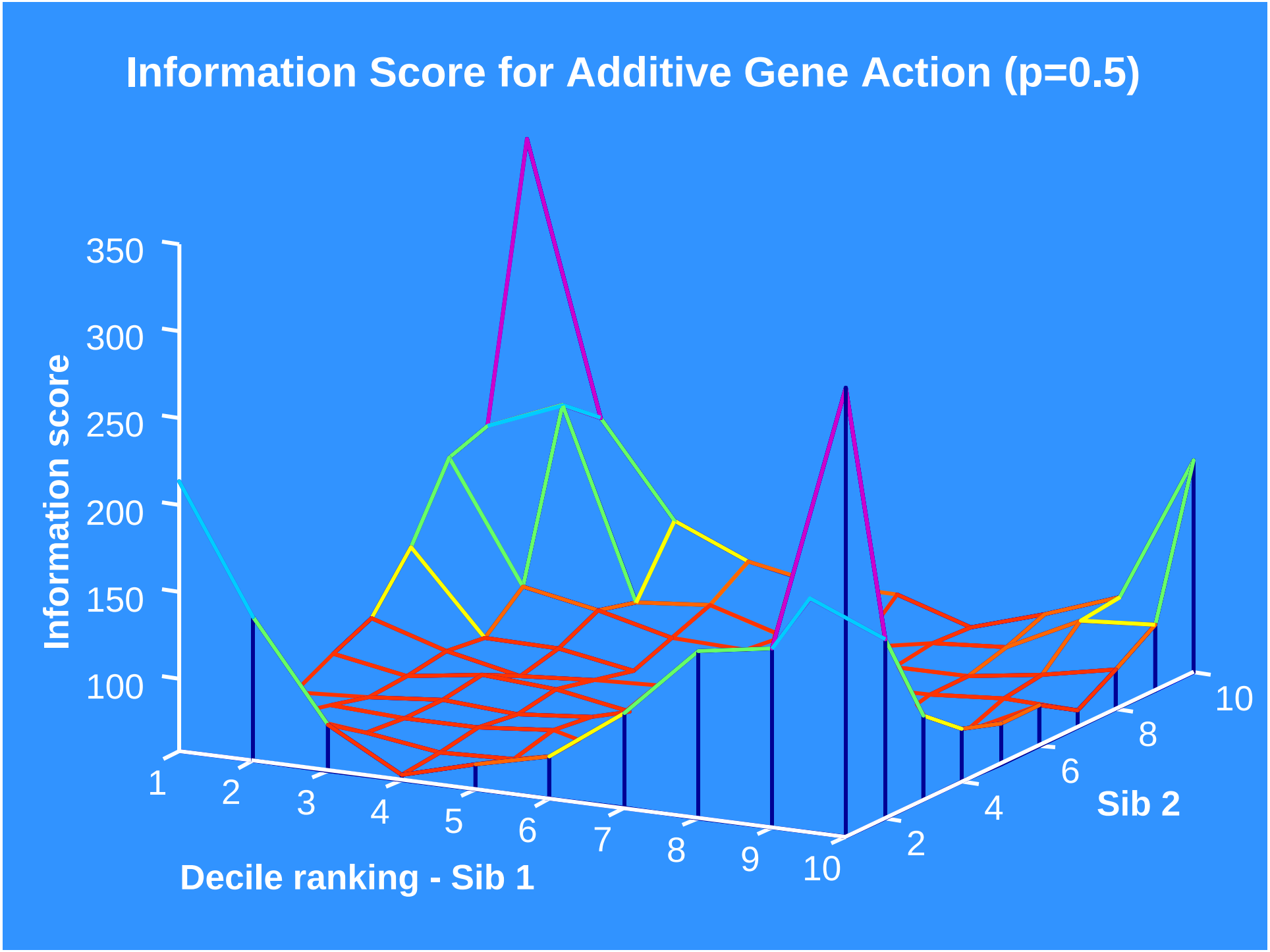


Evidence of developmental fields?

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.

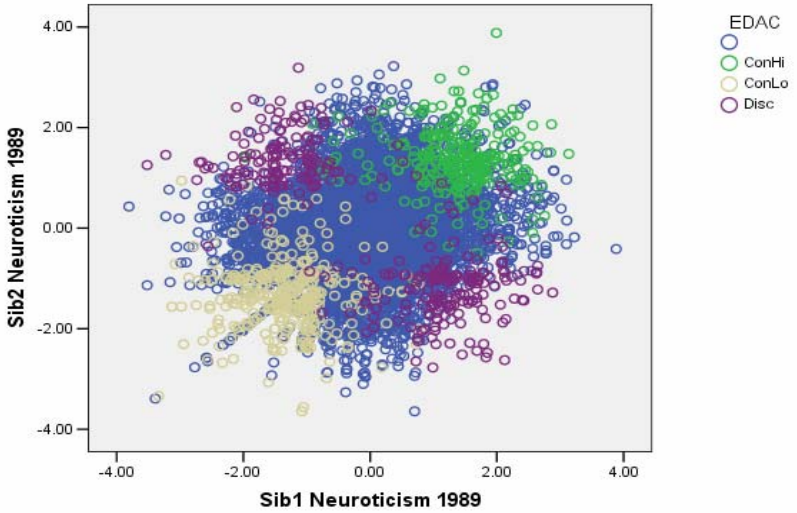

Information Score for Additive Gene Action (p=0.5)
Information score
350
300
250
200
150
100
1
2
3
4
5
6
7
8
9
10
Decile ranking - Sib 1
2
4
6
8
10
Sib 2

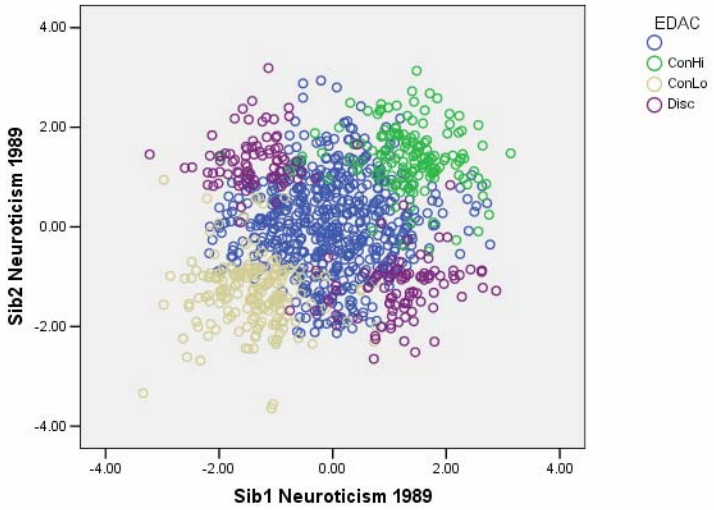
Genotypes available on EDAC plus others

Phenotyped for Neuroticism

Extreme Discordant Concordant Design

Genotyped EDAC plus$_{gn}$

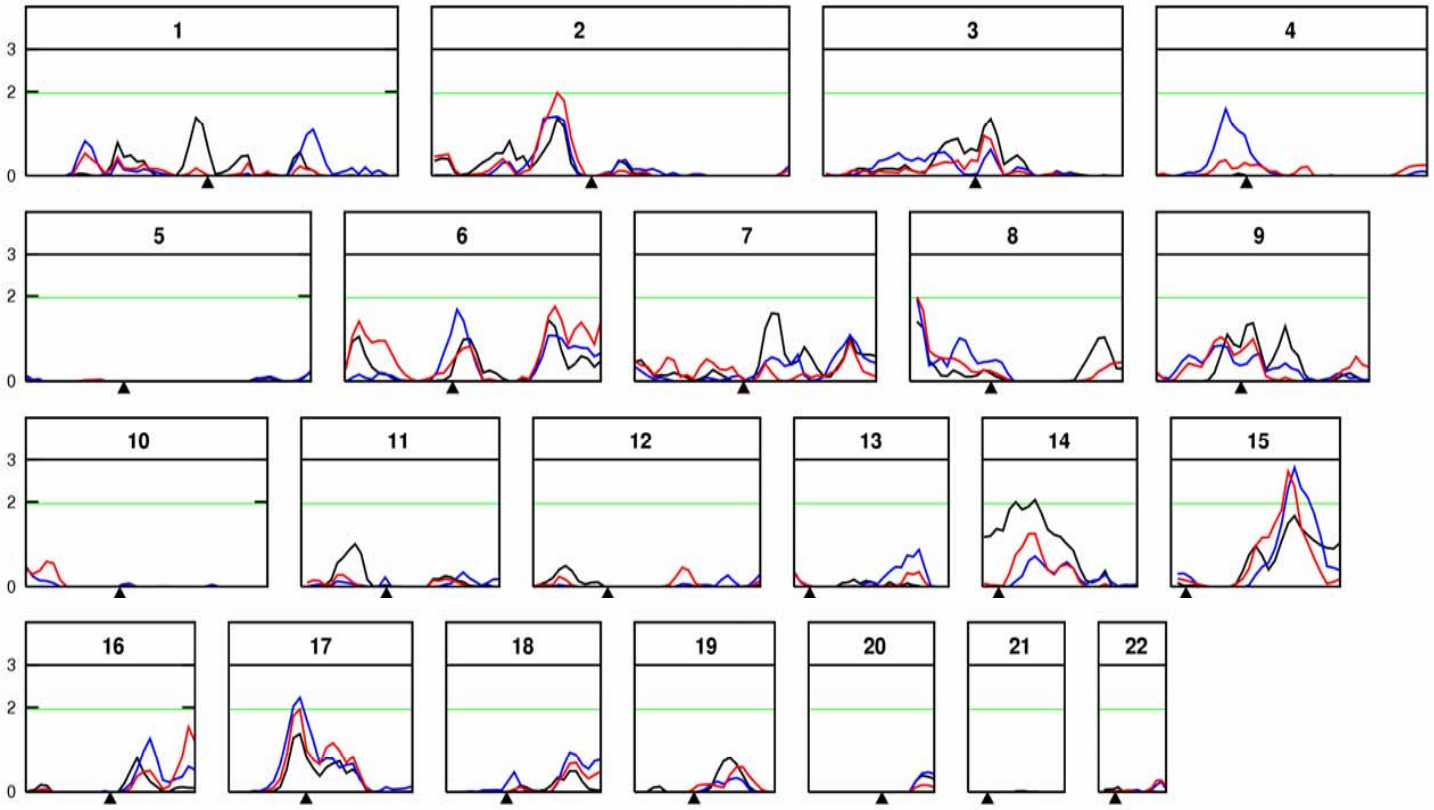




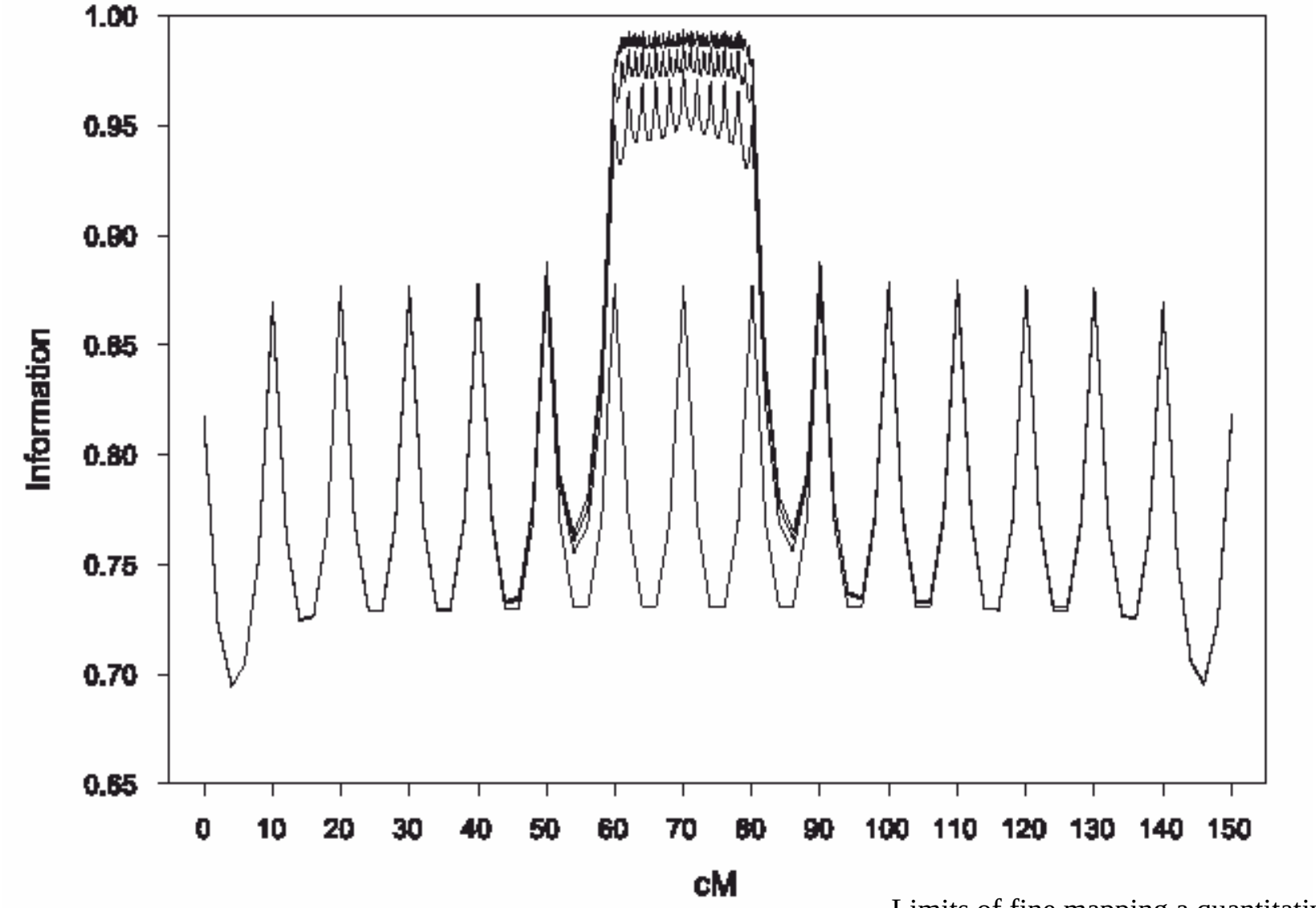
# QISPs	Neuroticism Phenotyped	QISPs share >300 markers	
Concordant Hi	556	343	62%
Concordant Lo	717	497	69%
Discordant	726	463	64%
The rest	8482	858	10%
Total	10481	2161	

Linkage scan EDAC sample – CIDI interview

Depression traits - OZ



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

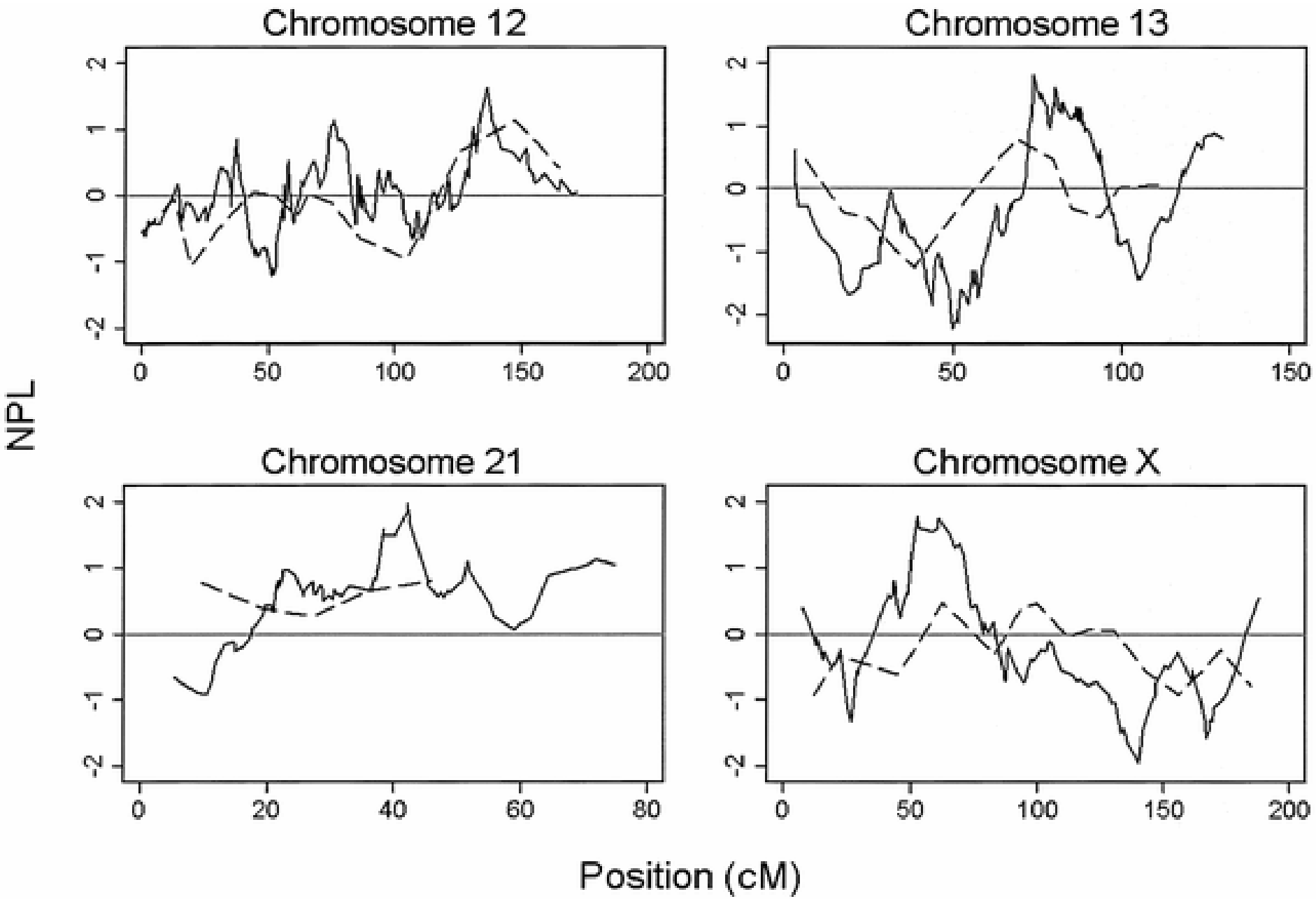
Am. J. Hum. Genet. 75:54–64, 2004

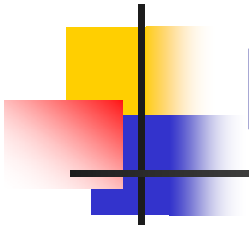
Whole-Genome Scan, in a Complex Disease, Using 11,245 Single-Nucleotide Polymorphisms: Comparison with Microsatellites

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

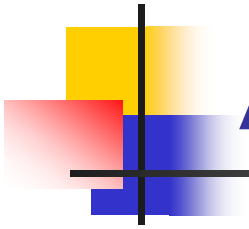
Despite the theoretical evidence of the utility of single-nucleotide polymorphisms (SNPs) for linkage analysis, no whole-genome scans of a complex disease have yet been published to directly compare SNPs with microsatellites. Here, we describe a whole-genome screen of 157 families with multiple cases 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)



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



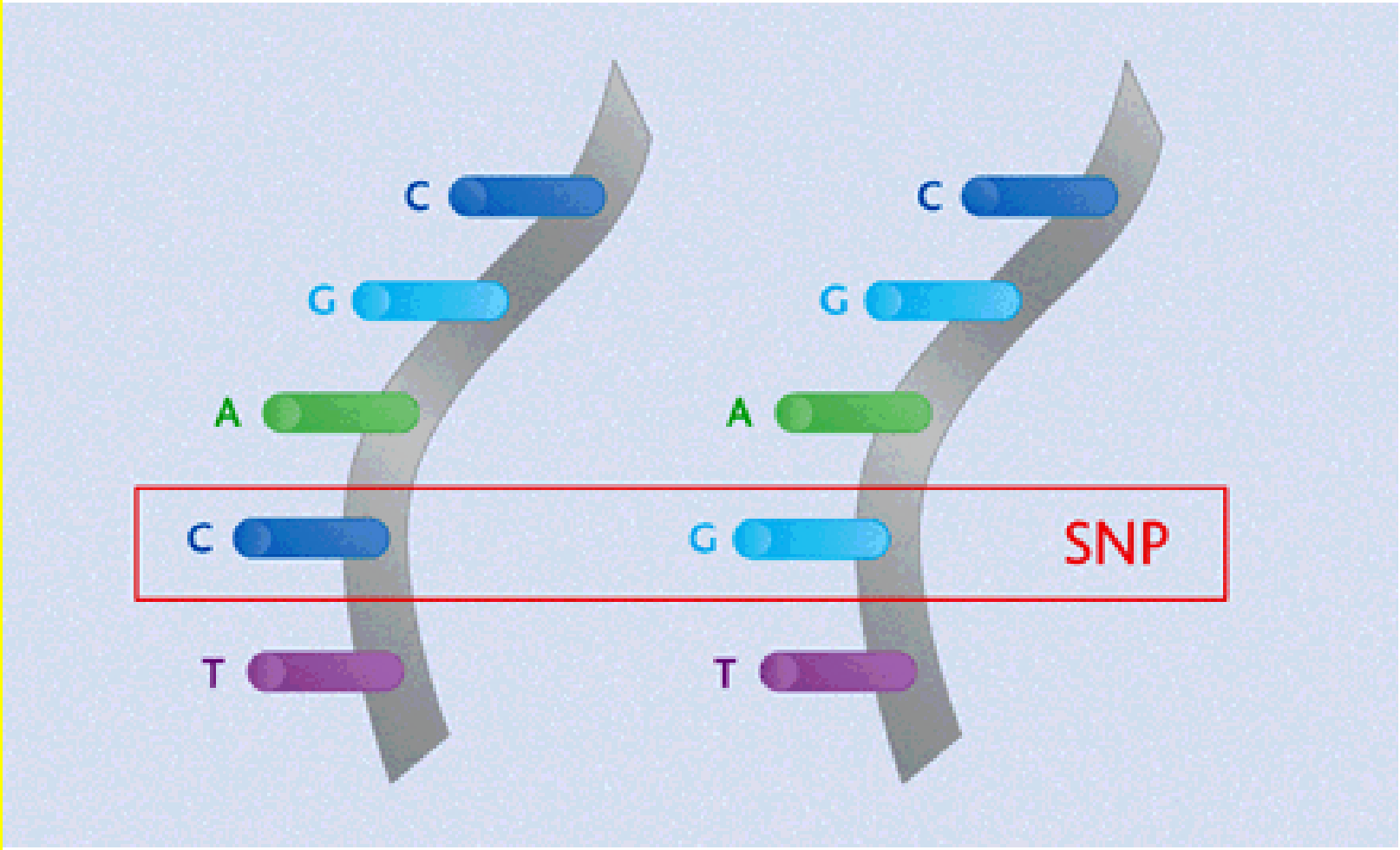
Association

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

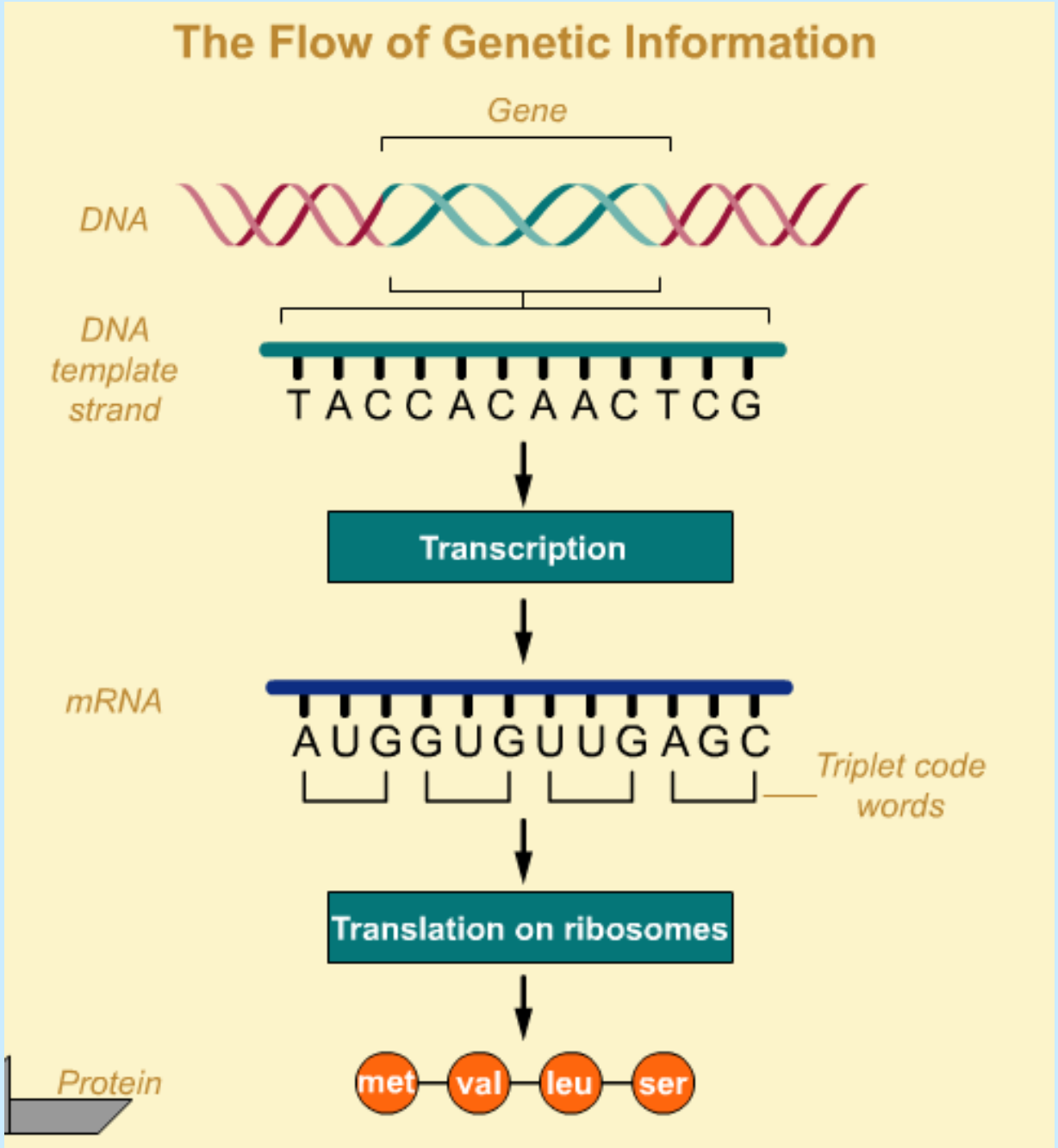
Association

- 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

Variation: Single Nucleotide Polymorphisms

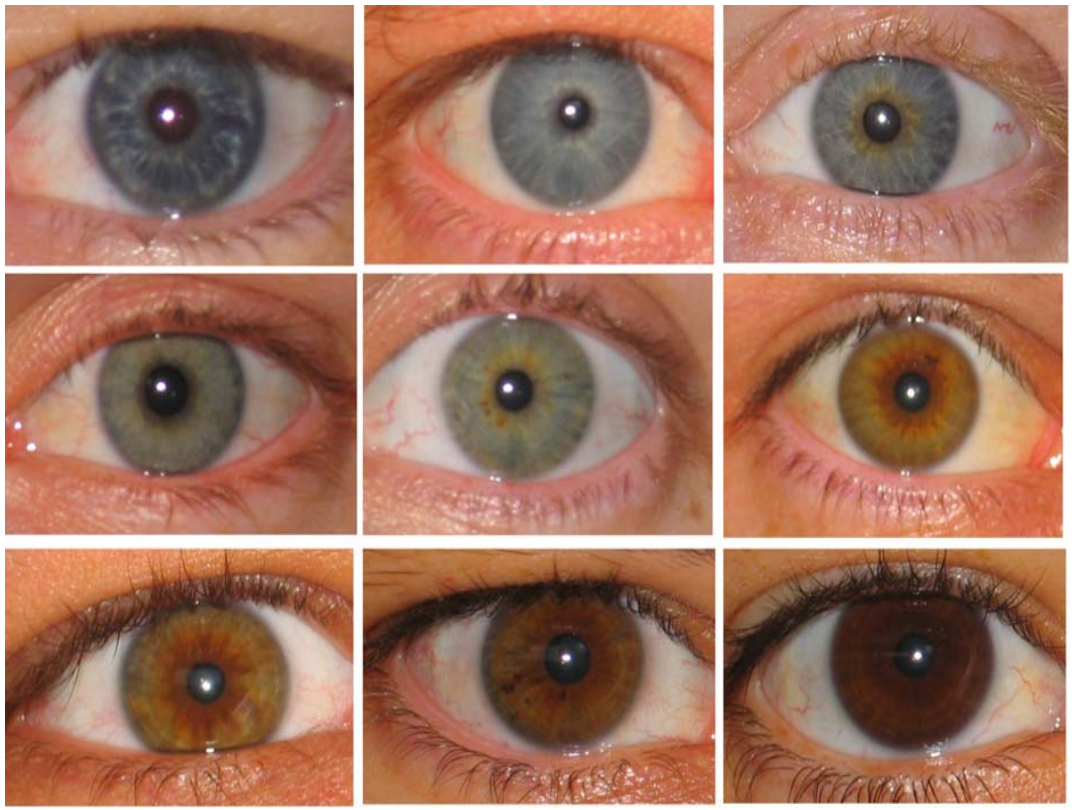


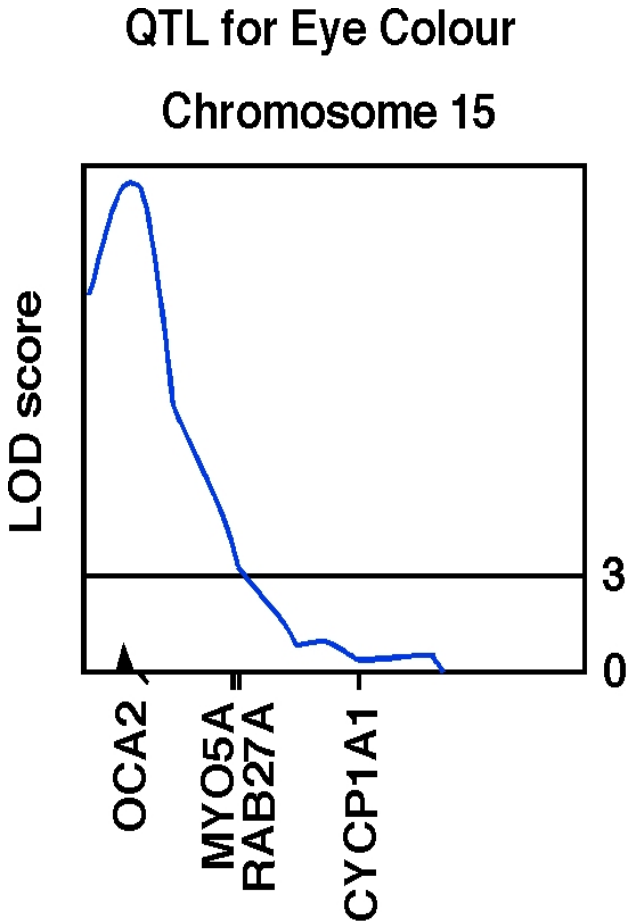
Complex disease marker? SNPs are single-base differences in DNA



Differences (between subjects) in DNA sequence are responsible for (structural) differences in proteins.

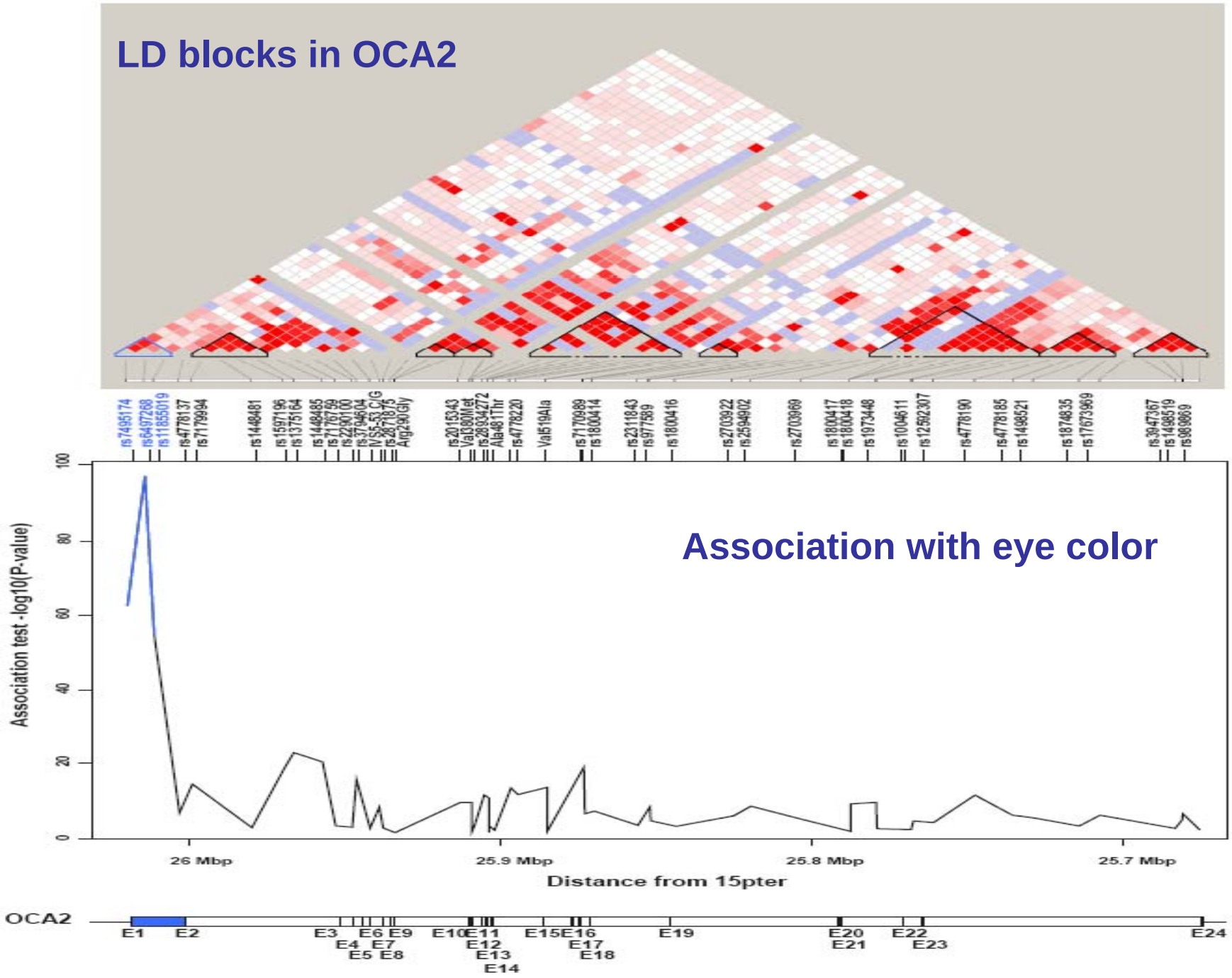
Human OCA2 and eye colour





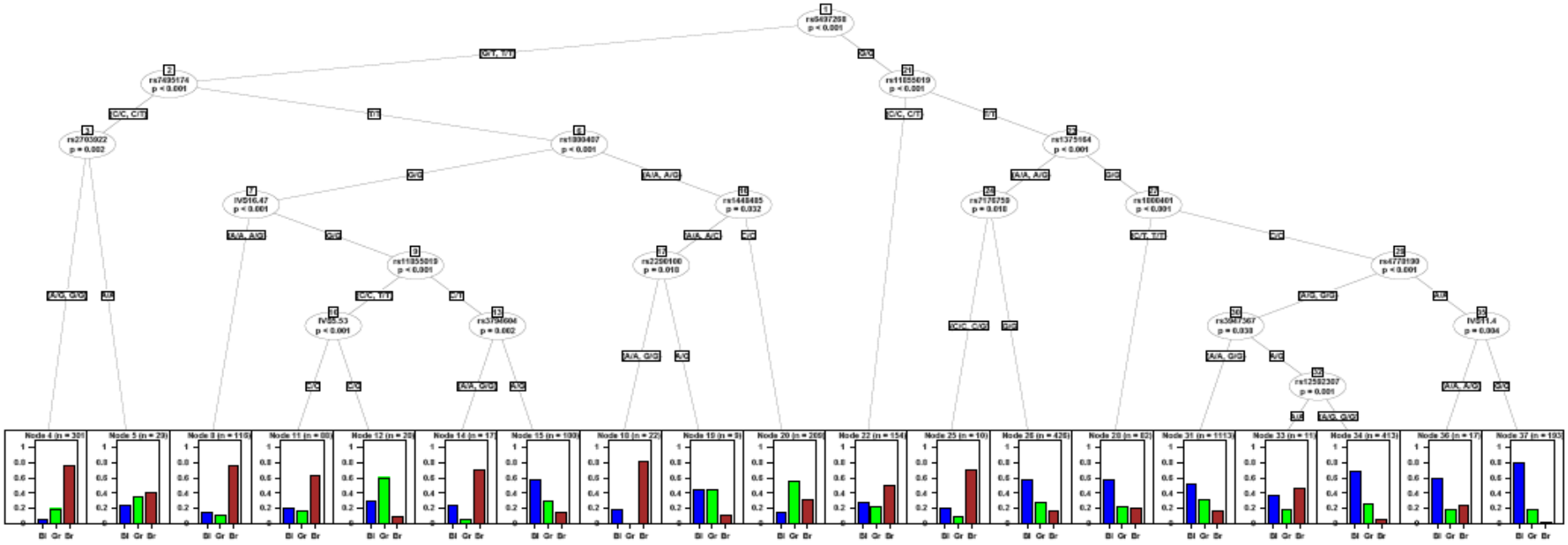
Zhu et al., *Twin Research* 7:197-210 (2004)

LD blocks in OCA2



Association with eye color

Eye colour explained

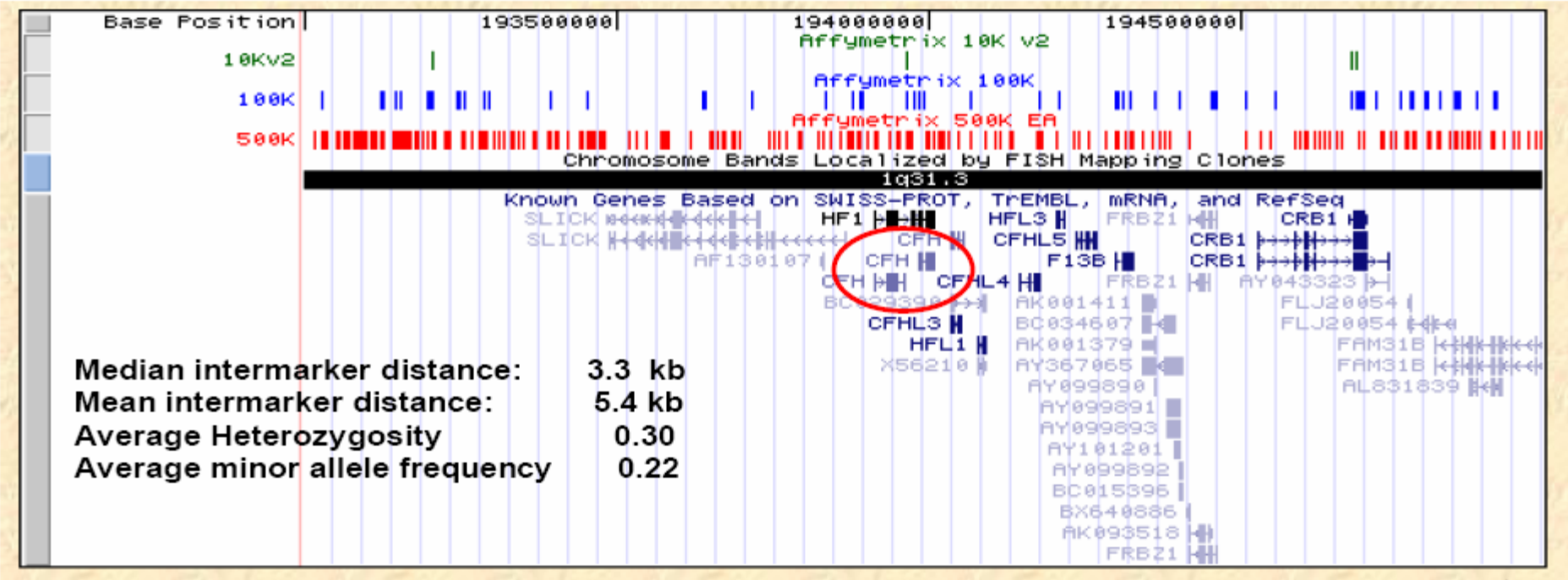


A Three–Single-Nucleotide Polymorphism Haplotype in Intron 1 of *OCA2* Explains Most Human Eye-Color Variation

David L. Duffy,* Grant W. Montgomery,* Wei Chen, Zhen Zhen Zhao, Lien Le, Michael R. James, Nicholas K. Hayward, Nicholas G. Martin, and Richard A. Sturm

American Journal of Human Genetics Volume 80 February 2007

Comparison of Affymetrix 10k, 100k, 500k SNP chips



SNP Genotyping Platforms

Throughput (SNPs Per Assay)

1? **25** **1536**

TaqMan 7900

Sequenom MassARRAY

Illumina BeadStation

Cost Per Assay

Flexibility in Project Design

Unprecedented Call Rates of >99%

FROM FUNG AND SINGLETON ET AL. NEUROLOGY THE LANCET

Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data

Hon-Chung Fung, Sonja Scholz, Mar Matarin, Javier Simón-Sánchez, Dena Hernandez, Angela Britton, J Raphael Gibbs, Carl Langefeld, Matt L Stiegert, Jennifer Schymick, Michael S Okun, Ronald J Mandel, Hubert H Fernandez, Kelly D Foote, Ramón L Rodríguez, Elizabeth Peckham, Fabienne Wavrant De Vrieze, Katrina Gwinn-Hardy, John A Hardy, Andrew Singleton

Summary

Background Several genes underlying rare monogenic forms of Parkinson's disease have been identified over the past decade. Despite evidence for a role for genetics in sporadic Parkinson's disease, few common genetic variants have been unequivocally linked to this disorder. We sought to identify any common genetic variability exerting a large effect in risk for Parkinson's disease in a population cohort and to produce publicly available genome-wide genotype data f
subje

Meth
a coh
SNPs

Findings We have produced around 220 million genotypes in 537 participants. This raw genotype data has been publicly posted and as such is the first publicly accessible high-density SNP data outside of the International HapMap Project. We also provide here the results of genotype and allele association tests.

Interpretation We generated publicly available genotype data for Parkinson's disease patients and controls so that these data can be mined and augmented by other researchers to identify common genetic variability that results in minor and moderate risk for disease.

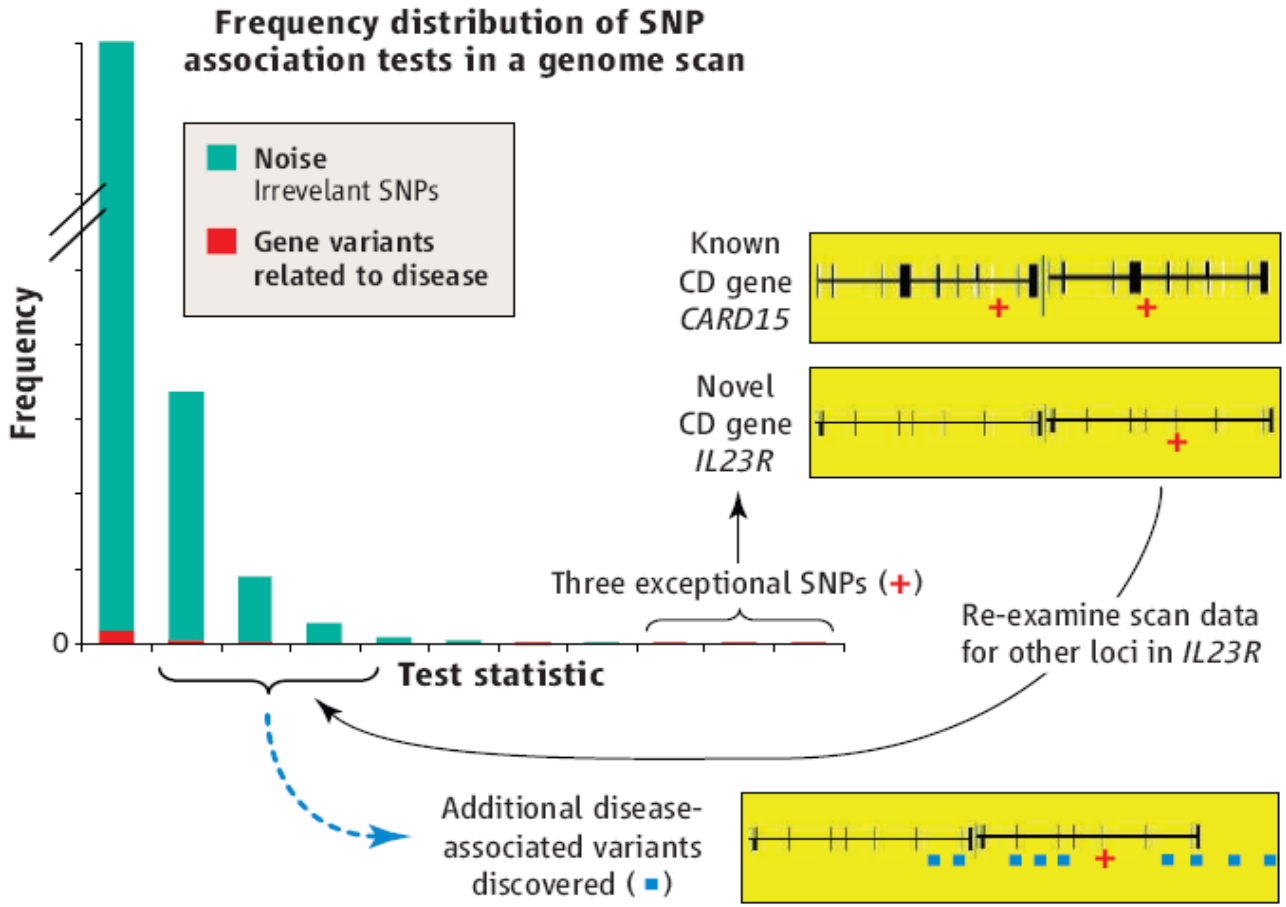
A total of 219,577,497 unique genotype calls were made and the average call rate across all samples was 99.6%.

Published Online
September 27, 2006
DOI:10.1016/S1474-4422(06)70578-6

ogenetics
e Vrieze PhD),
mick BS),
Unit
arin PhD),

ritton MS,
nd
Biology Core
(J R Gibbs BS), National Institute on Aging, National Institutes of Health, Bethesda, MD, USA; Section on Biostatistics, Department of Public Health Sciences, Wake Forest University Health Sciences, Winston-Salem, NC, USA (C Langefeld PhD, M L Stiegert MS); Movement

GWAS for Inflammatory Bowel Disease

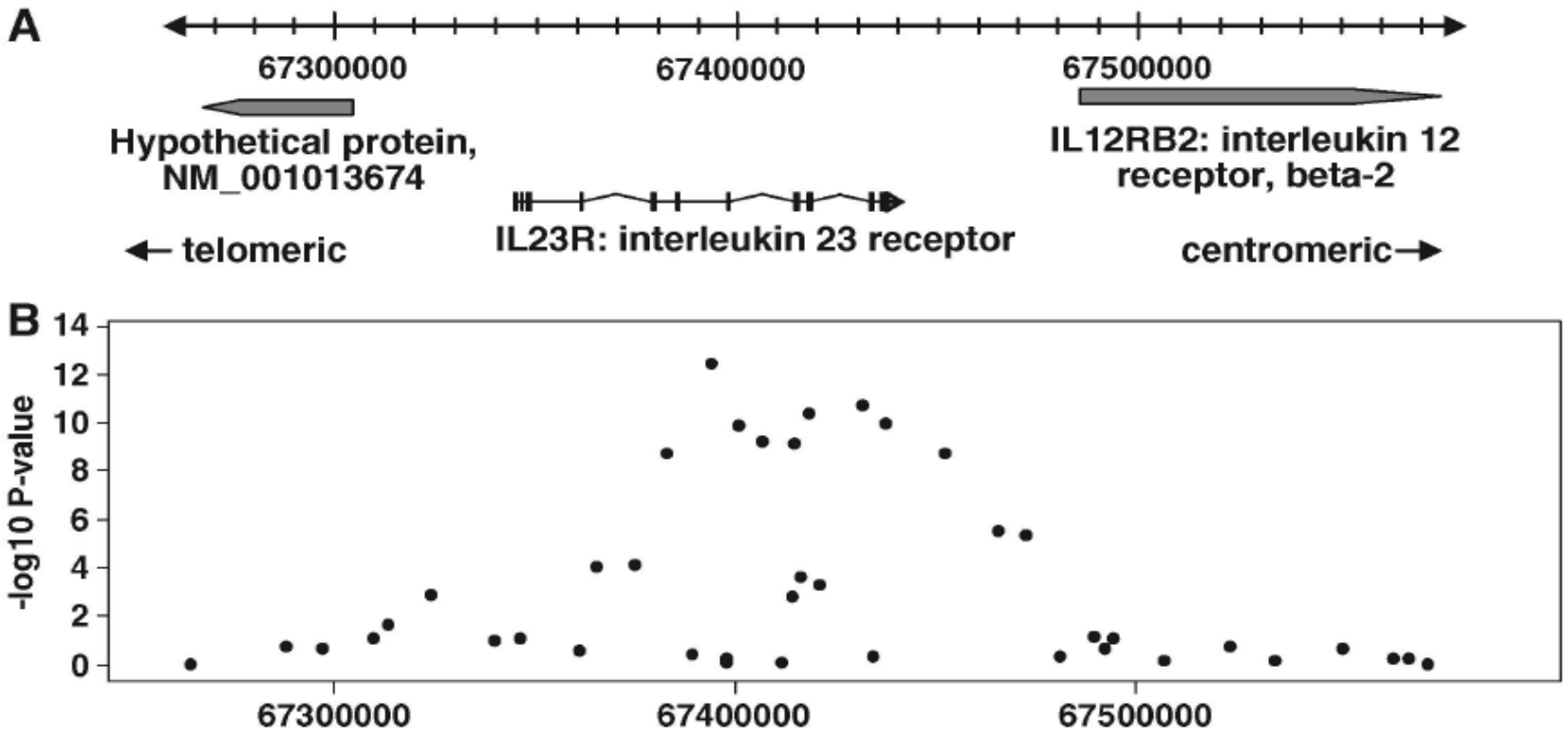


SCIENCE VOL 314 1 DECEMBER 2006

Delivering New Disease Genes

Lon R. Cardon

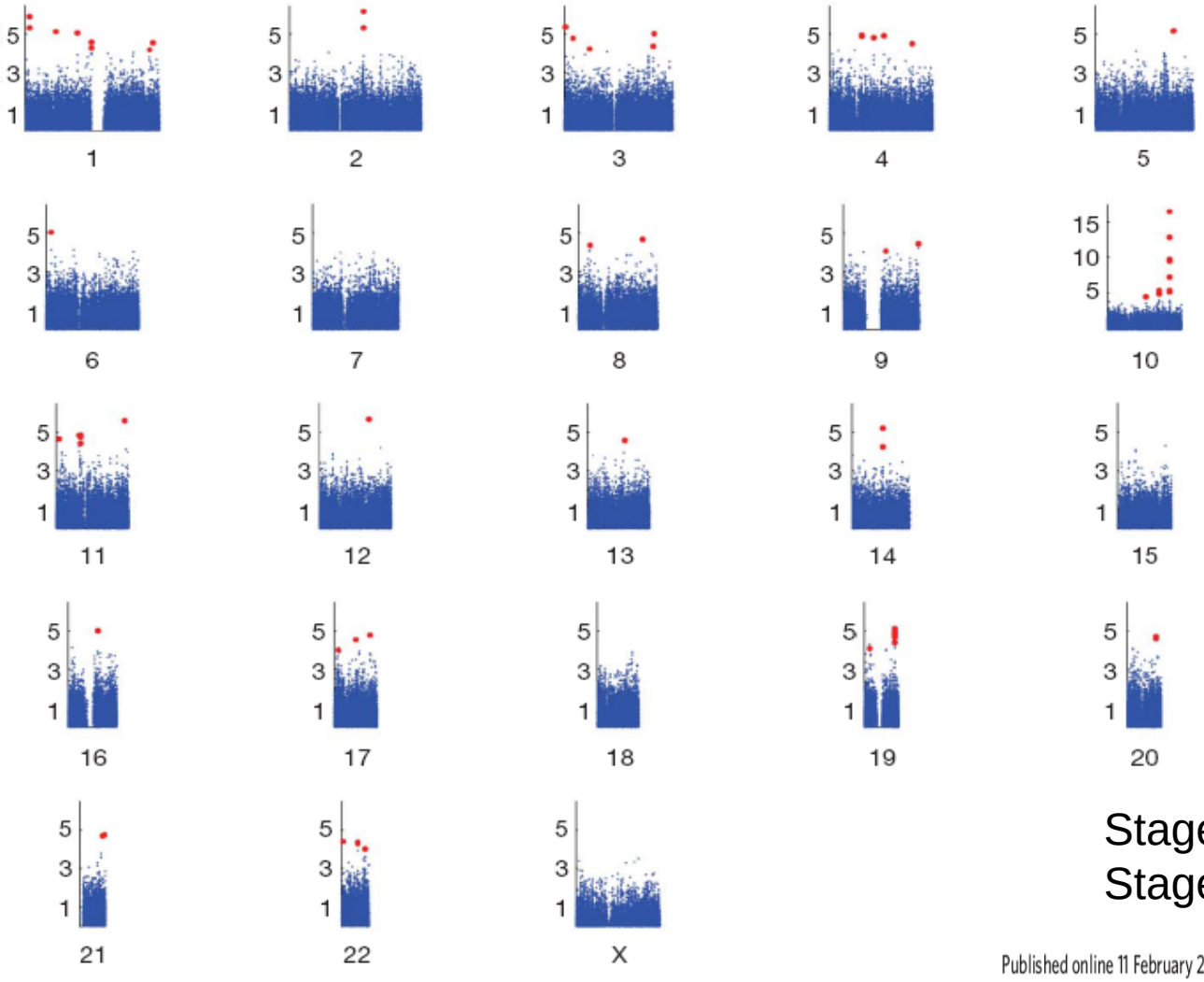
GWAS for Inflammatory Bowel Disease



SCIENCE VOL 314 1 DECEMBER 2006

A Genome-Wide Association Study Identifies *IL23R* as an Inflammatory Bowel Disease Gene

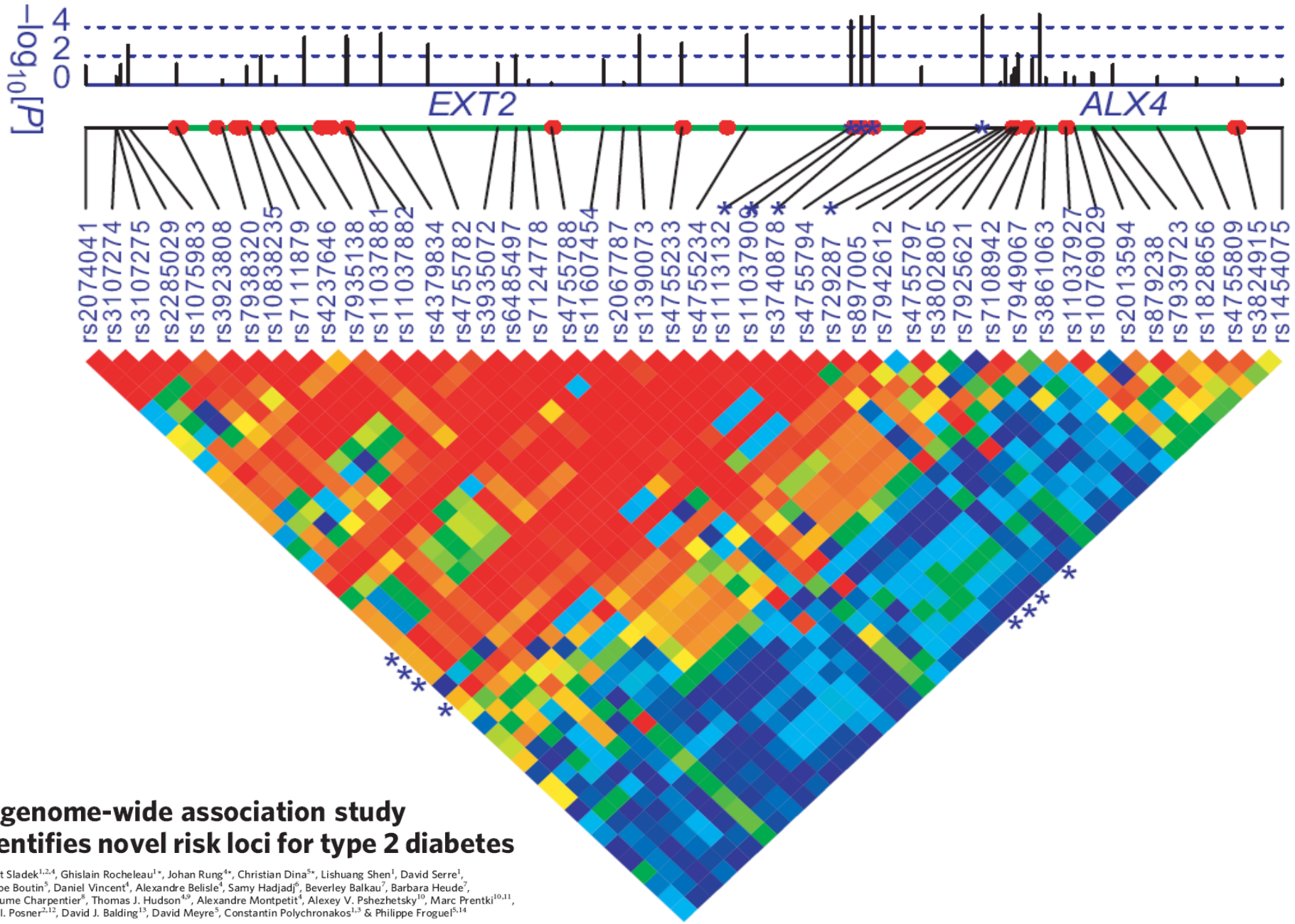
A genome-wide association study identifies novel risk loci for type 2 diabetes



Stage 1: Illumina 100k+300k
Stage 2: Sequenom Iplex

Published online 11 February 2007

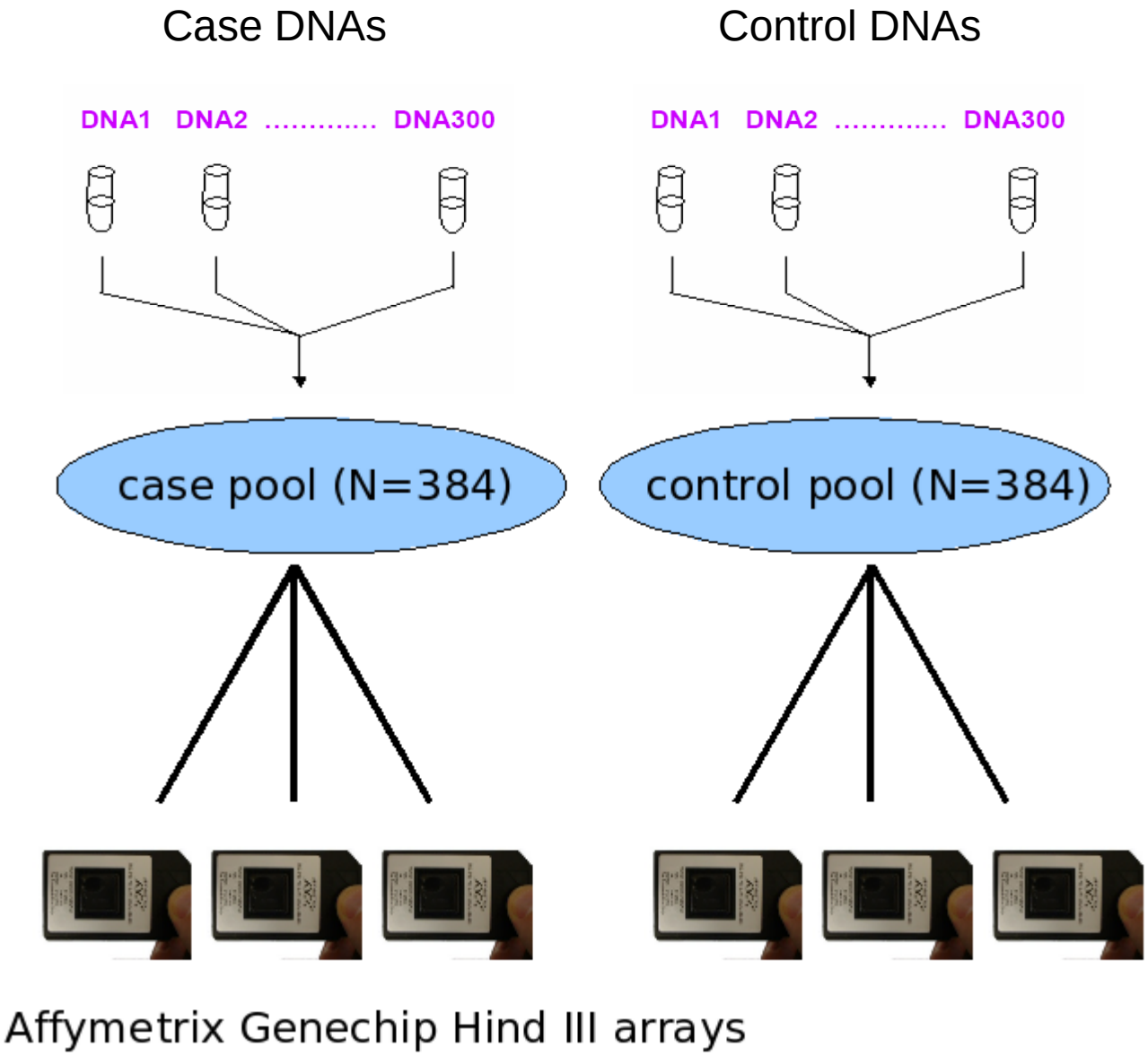
nature



A genome-wide association study identifies novel risk loci for type 2 diabetes

Robert Sladek[1,2,4], Ghislain Rocheleau[1]*, Johan Rung[4]*, Christian Dina[5]*, Lishuang Shen[1], David Serre[1], Philippe Boutin[5], Daniel Vincent[4], Alexandre Belisle[4], Samy Hadjadj[6], Beverley Balkau[7], Barbara Heude[7], Guillaume Charpentier[8], Thomas J. Hudson[4,9], Alexandre Montpetit[4], Alexey V. Pshezhetsky[10], Marc Prentki[10,11], Barry I. Posner[2,12], David J. Balding[13], David Meyre[5], Constantin Polychronakos[1,3] & Philippe Froguel[5,14]

Cutting costs of GWAS by DNA pooling



Affy 500k chip
Pools of Hi/Lo memory Ss

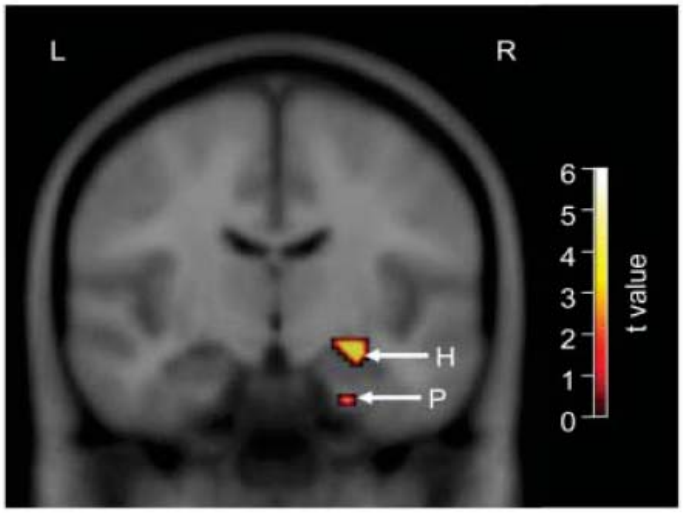
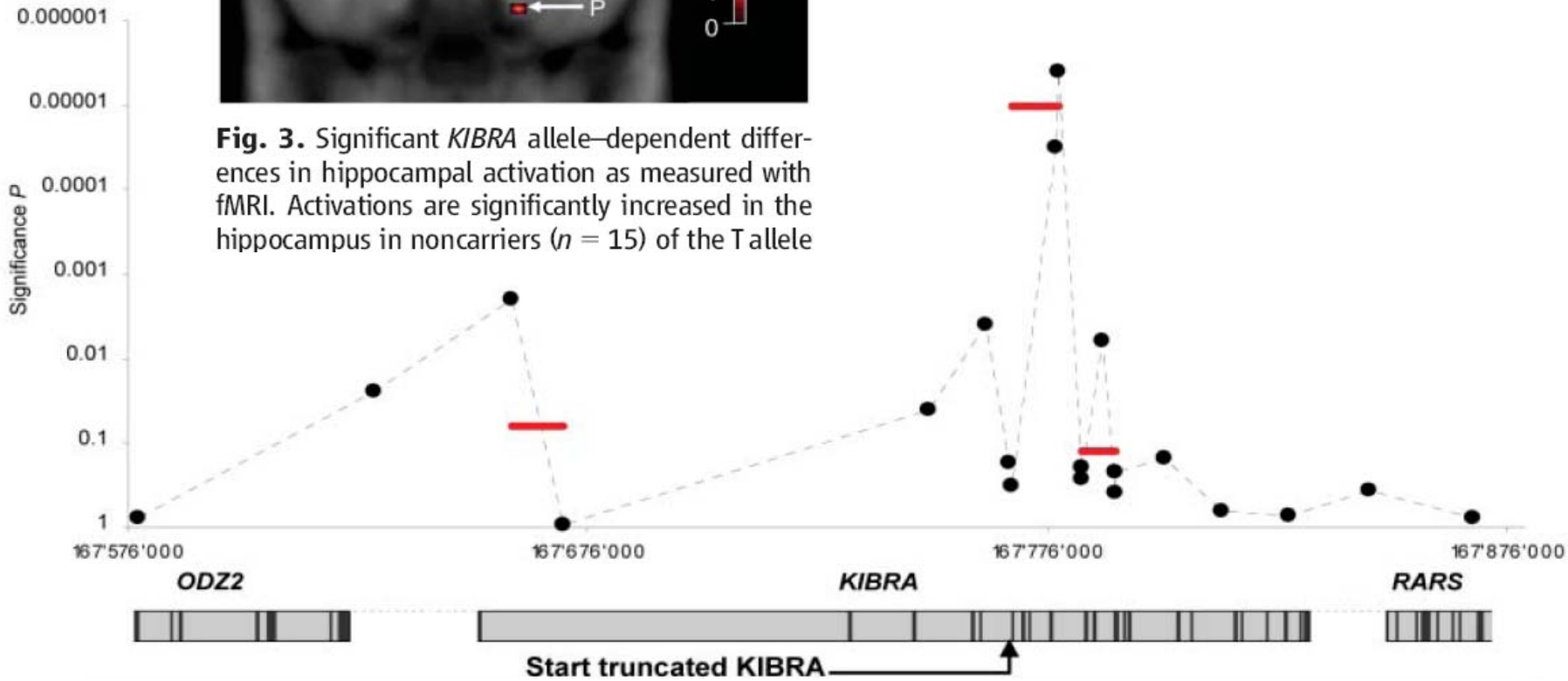


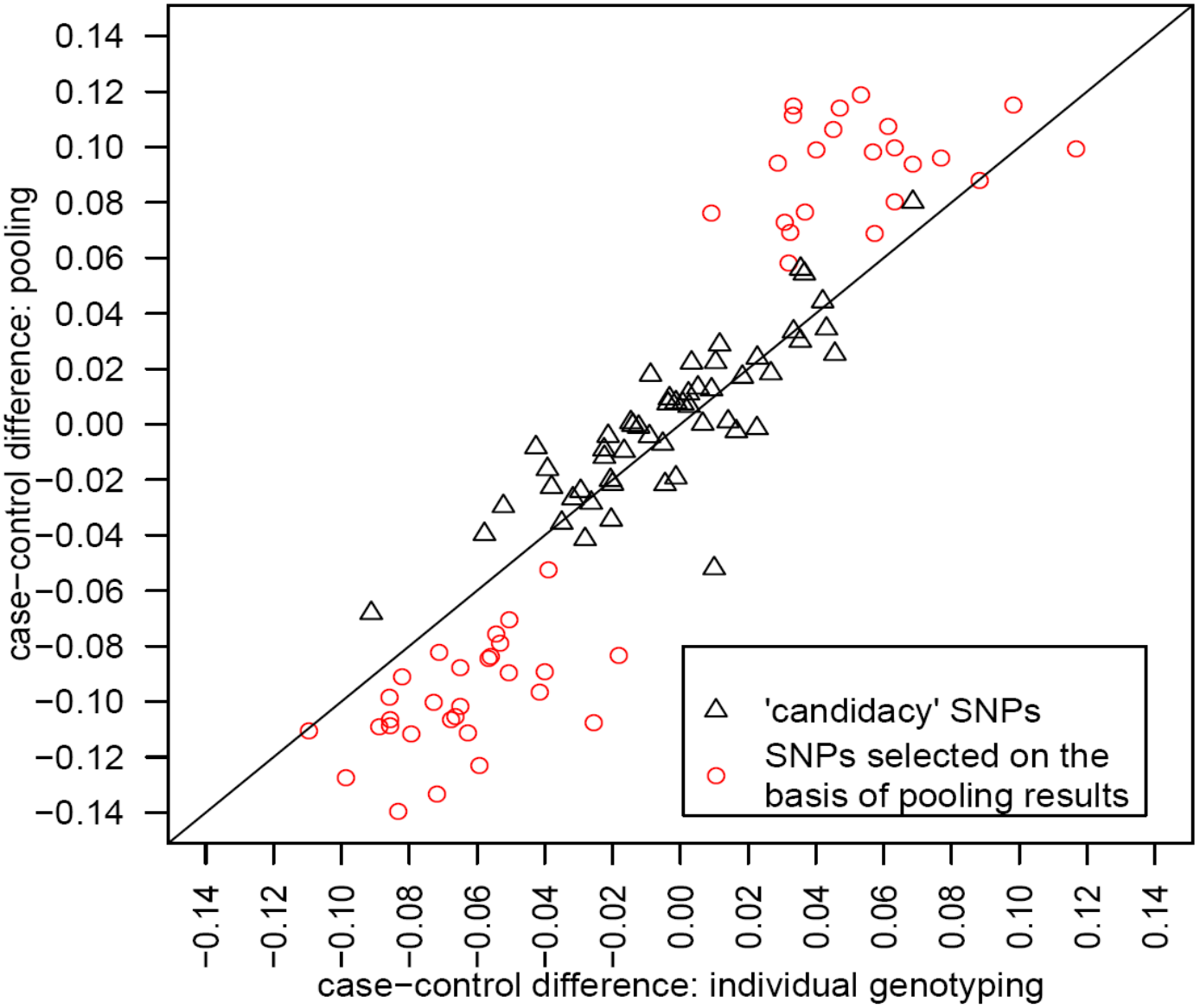
Fig. 3. Significant *KIBRA* allele–dependent differences in hippocampal activation as measured with fMRI. Activations are significantly increased in the hippocampus in noncarriers ($n = 15$) of the T allele



SCIENCE VOL 314 20 OCTOBER 2006

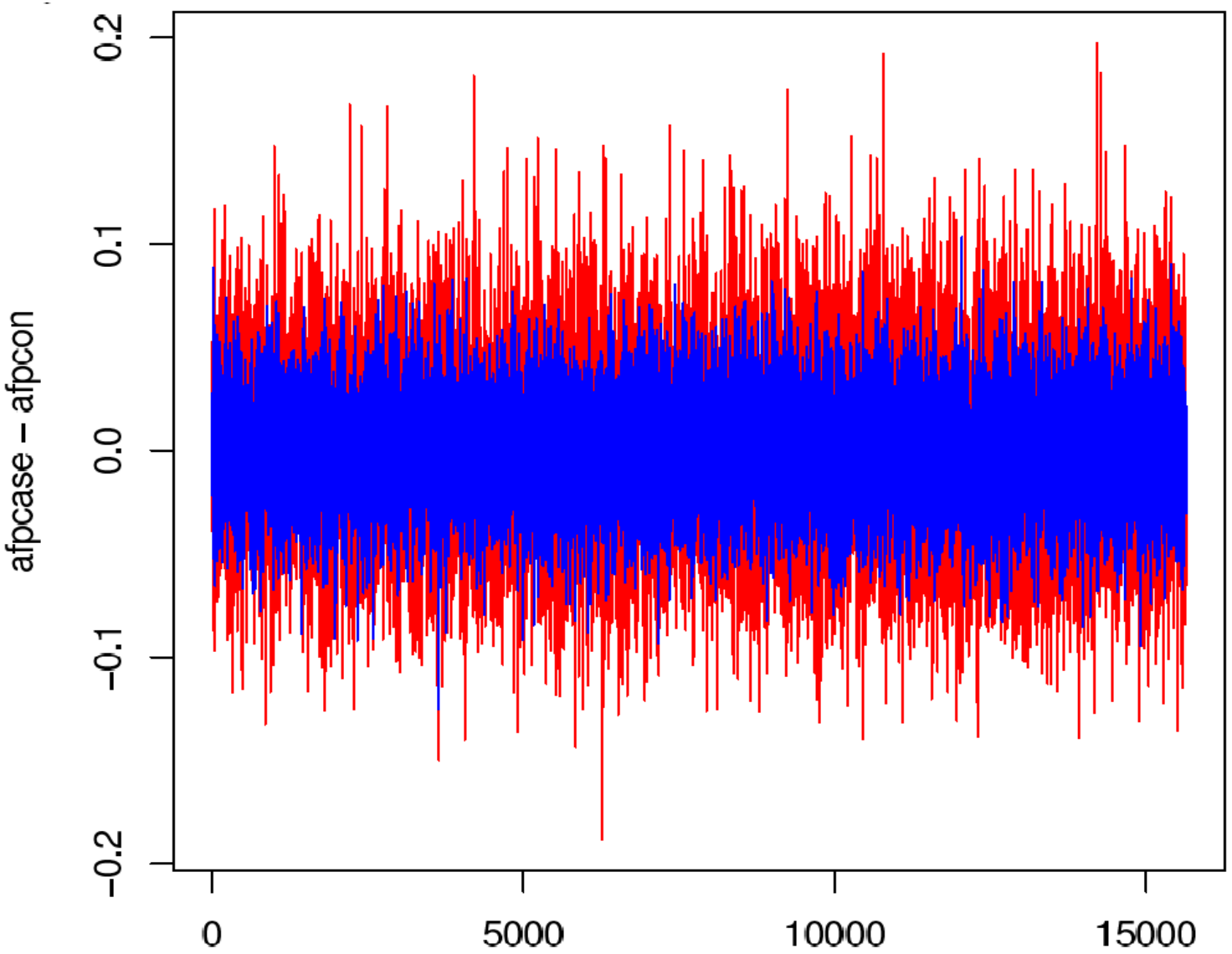
Common *Kibra* Alleles Are Associated with Human Memory Performance

Case-control allele frequency differences: individual genotyping vs pools (Hap300)



Stuart Macgregor, QIMR

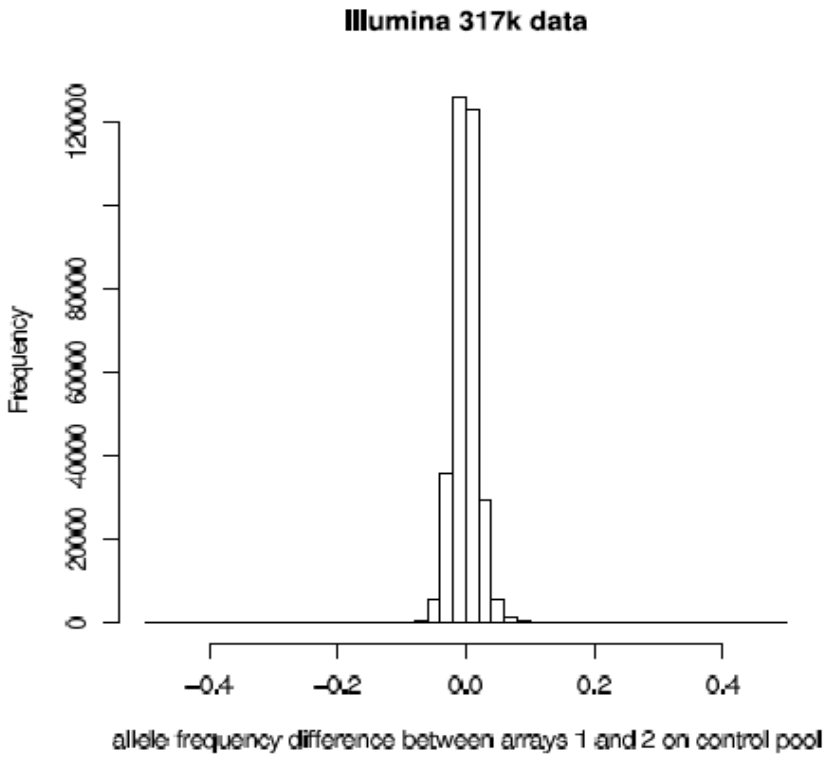
Pooling error for 15,000 SNPs using Illumina Hap300 and Affy 50k arrays

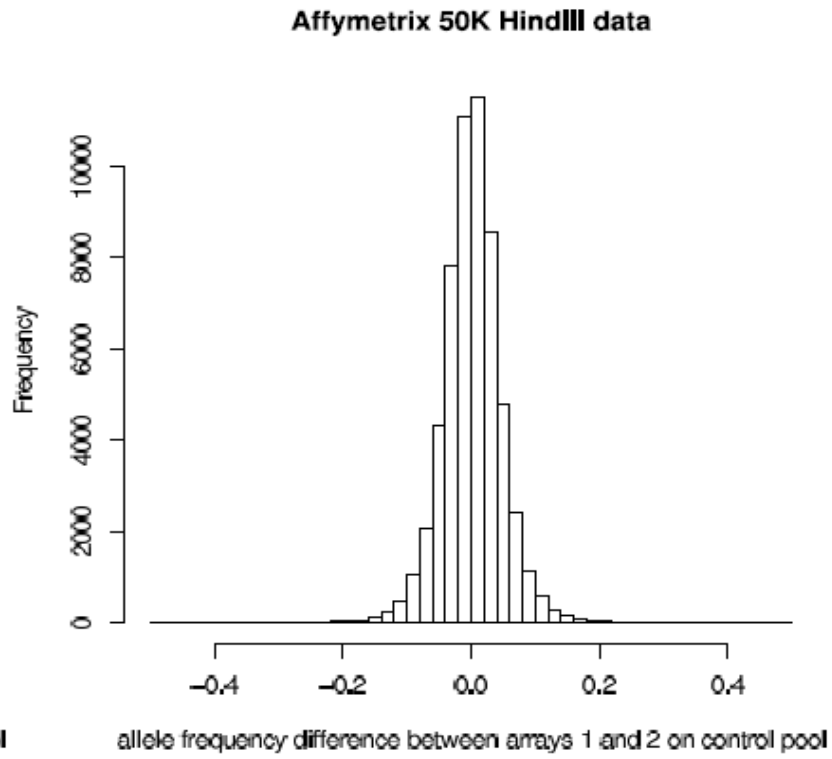


Illumina arrays extract 80% information as IG vs ~30% with Affy: need ~10x Affy arrays

Stuart Macgregor, QIMR

Illumina Hap300 versus Affy 50k array-specific error plots



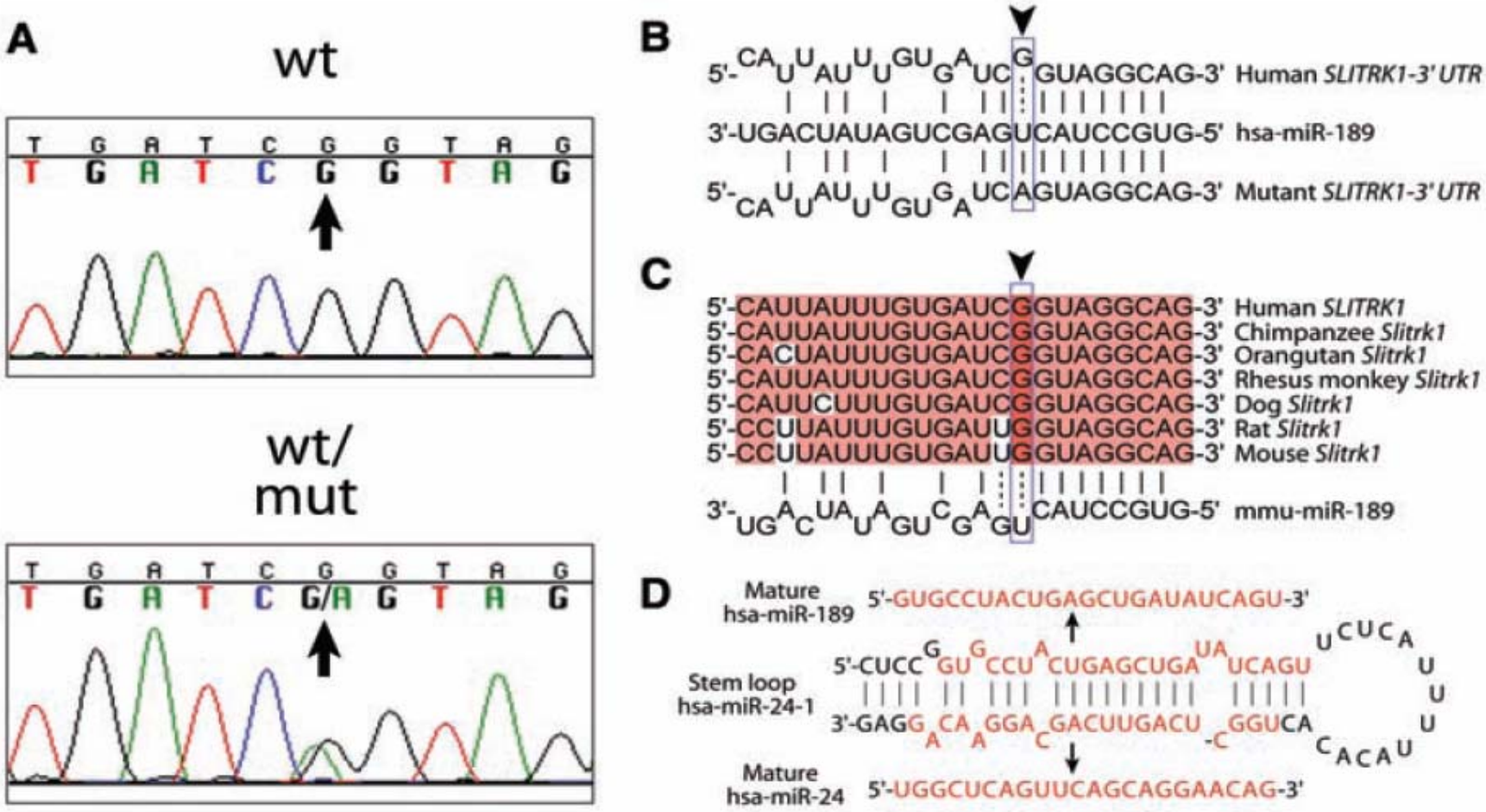


Stuart Macgregor, QIMR

Role of miRNA (binding sites) in disease ?

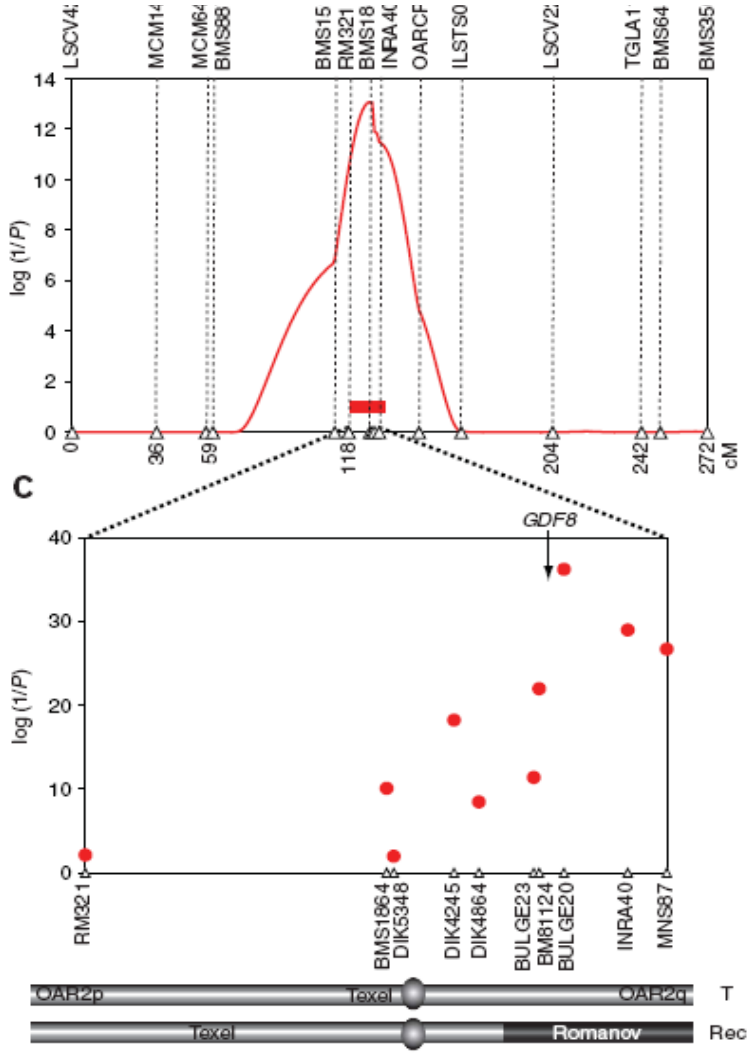
Sequence Variants in *SLITRK1* Are Associated with Tourette's Syndrome

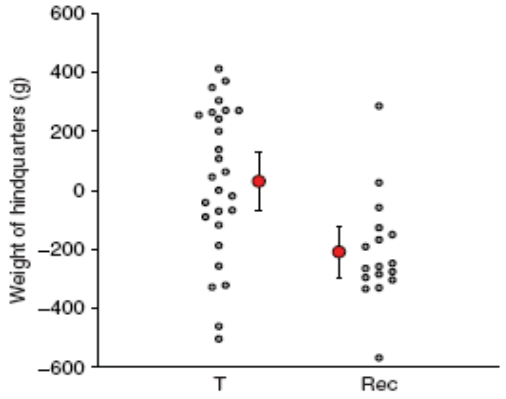
14 OCTOBER 2005 VOL 310 SCIENCE



A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep

and in quantitative traits





Texel sheep are renowned for their exceptional meatiness. To identify the genes underlying this economically important feature, we performed a whole-genome scan in a Romanov × Texel F2 population. We mapped a quantitative trait locus with a major effect on muscle mass to chromosome 2 and subsequently fine-mapped it to a chromosome interval encompassing the myostatin (*GDF8*) gene. We herein demonstrate that the *GDF8* allele of Texel sheep is characterized by a G to A transition in the 3′ UTR that creates a target site for *mir1* and *mir206*, microRNAs (miRNAs) that are highly expressed in skeletal muscle. This causes translational inhibition of the myostatin gene and hence contributes to the muscular hypertrophy of Texel sheep. Analysis of SNP databases for humans and mice demonstrates that mutations creating or destroying putative miRNA target sites are abundant and might be important effectors of phenotypic variation.

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?

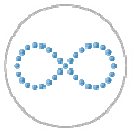
Increasing evidence for Common Disease – Rare Variant hypothesis (CDRV)

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]

Product Portfolio and Application Areas



Human-1	Gene centric association studies
Hap300	WGA with 80% genomic coverage in Caucasians
Hap550	WGA with ~90% genomic coverage in Caucasians and Asians
Hap650Y	WGA with ~90% genomic coverage in Caucasians and Asians AND 67% genomic coverage in Yoruban

Hap450S +

+ Hap240S

1M	The most comprehensive chip that allows whole genome DNA analysis with industry leading SNP coverage in genes, CNV regions and indels

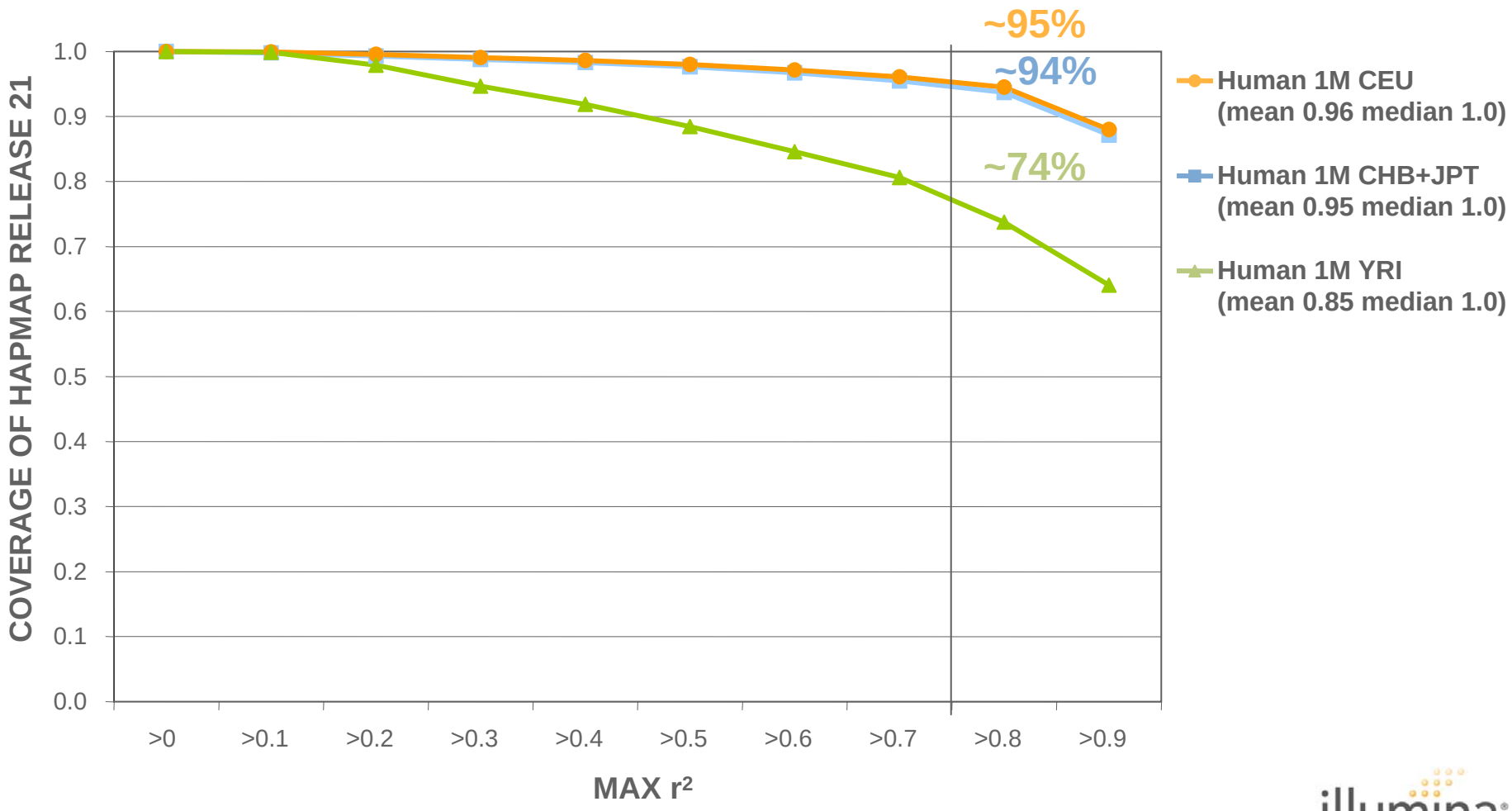
illumina®

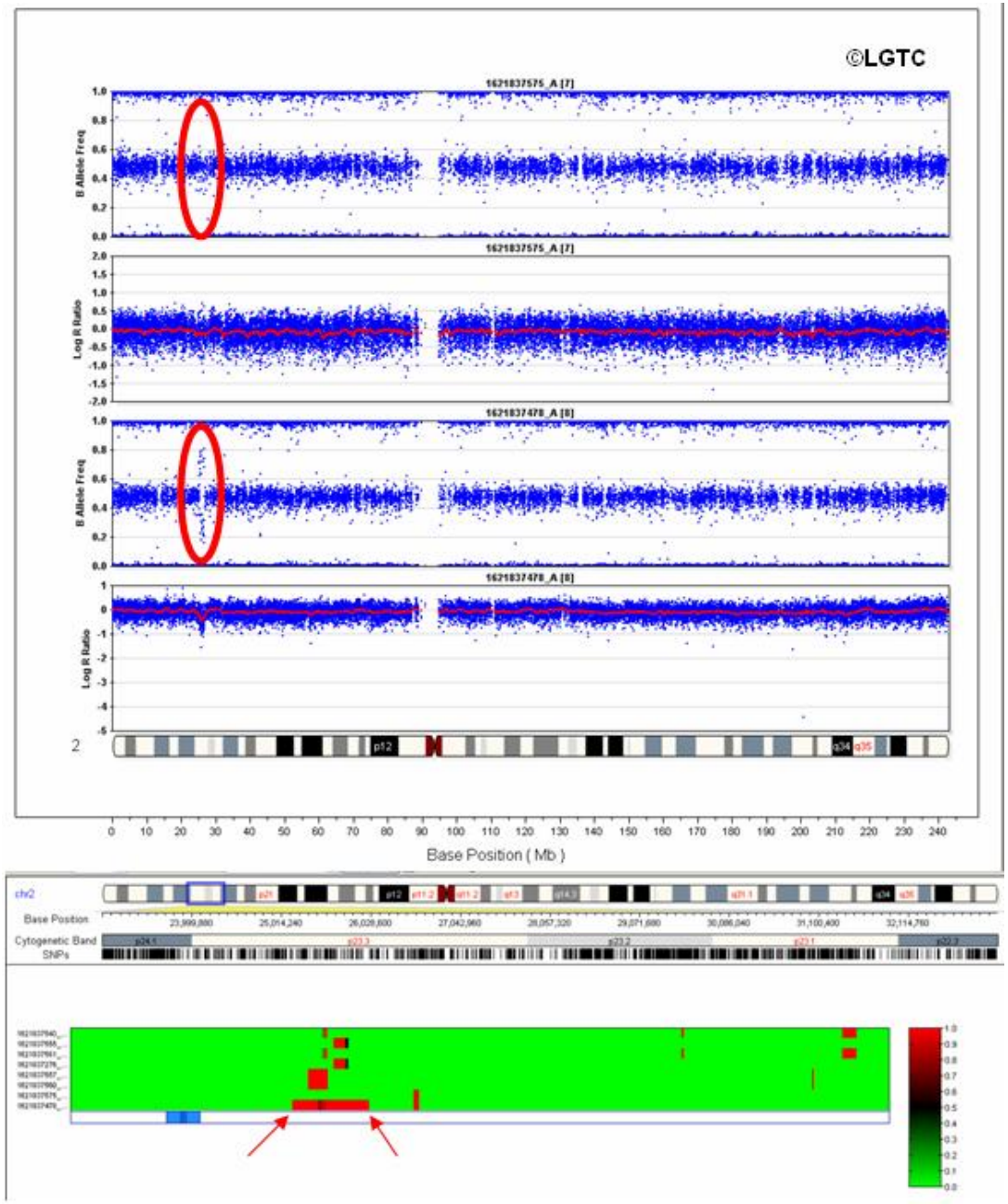
1M Content

CONTENT	NUMBER	VALUE
HumanHap550	555,000	High genomic coverage
SNPs in Genes	400,000	High density of SNPs in coding regions of the genome
SNPs and Probes in both reported and novel Copy Number Variant (CNV) Regions	110,000	High density of SNPs and probes in CNV regions, including “nonSNPable” regions
Additional Caucasian and Asian Tag SNPs	84,000	Higher tag SNP coverage of the genome
Additional African Tag SNPs	100,000	Higher tag SNP coverage of the genome
Even Spacing SNPs	90,000	Ensure complete coverage across the genome, enable new CNV discoveries
ADME/MHC SNPs	17,000	Denser coverage in high value regions/genes
TOTAL	**>1M**	**Unsurpassed power and gene coverage for WGA and CNV studies**

illumina®

Human 1M HapMap Coverage by Population





Copy Number Variation (CNV) in MZ twin pair

The $1,000 Human Genome - Implications for Life Science, Healthcare, and IT

Published by:
IDC

products & services

- overview
- systems & software
- dna analysis solutions
- rna analysis solutions
- solexa applications
- services
- product literature

- print this page

solexa sequencing applications

Illumina's Solexa Sequencing technology offers a powerful new approach to some of today's most important applications for genetic analysis and functional genomics, including:

sequencing and resequencing

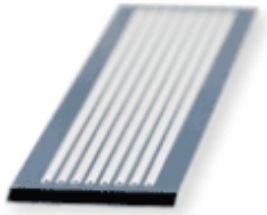
Whether you need to sequence an entire genome or a large candidate region, the Illumina Genome Analyzer System is today's most productive and economical sequencing tool. Solexa sequencing technology and reversable terminator chemistry deliver unprecedented volumes of high quality data, rapidly and economically.

expression profiling

Sequencing millions of short cDNA tags per sample, the Genome Analyzer allows you to generate digital expression profiles at costs comparable to current analog methods. Because our protocol does not require any transcript-specific probes, you can apply the technology to discover and quantitate transcripts in any organisms, irrespective of the annotation available on the organism.

small rna identification and quantification

Solexa sequencing technology also offers a unique and powerful solution for the comprehensive discovery and characterization of small RNAs in a wide range of species. The massively parallel sequencing protocol allows researchers to discover and analyze genome-wide profiles of small RNA in any species. With the potential to generate several million sequence tags economically, the Illumina Genome Analyzer offers investigators the opportunity to uncover global profiles of small RNA at an unprecedented scale.



important information

- product literature
- publications
- faqs
- have a rep contact me

EPIGENETIC DISCORDANCE IN IDENTICAL TWINS

The missing “environment” ?

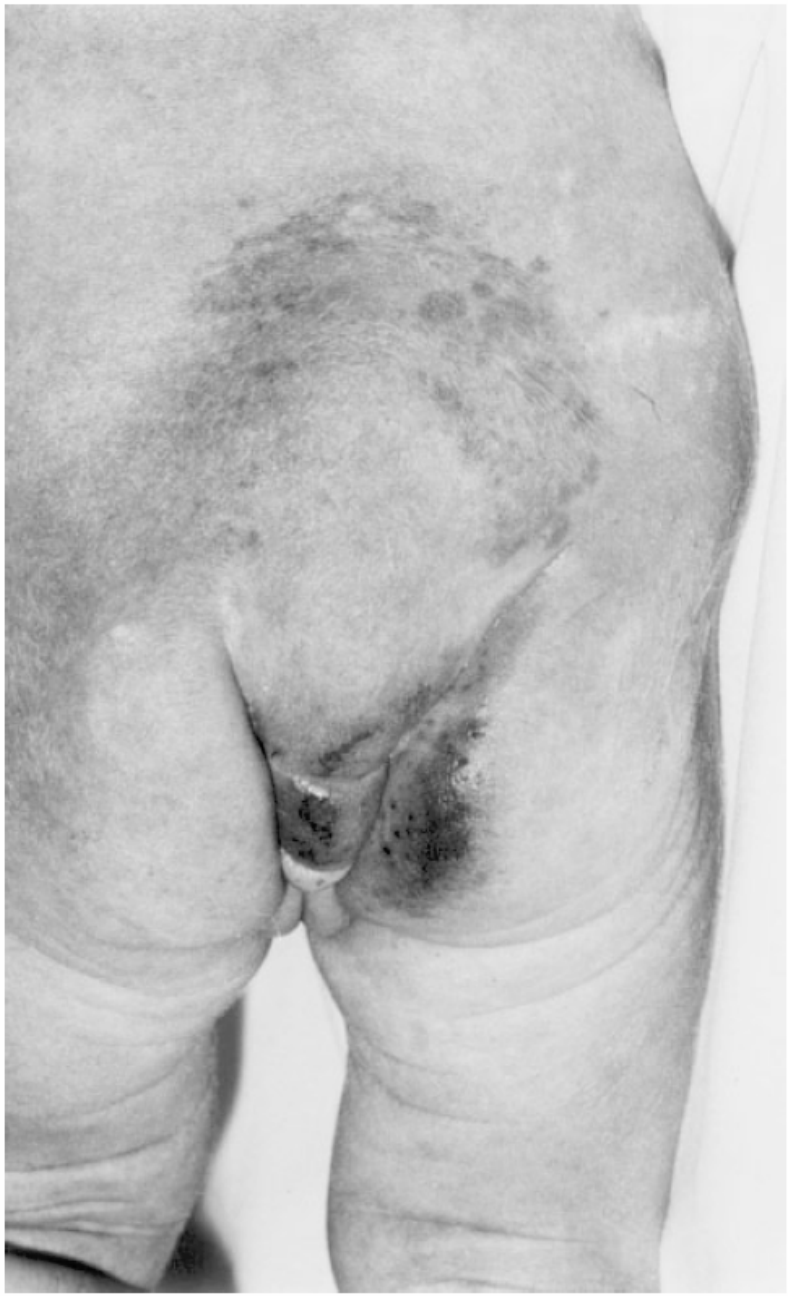


Fig. 1. Patient 1. Soft tumor and abnormal aspect in the lumbosacral area.

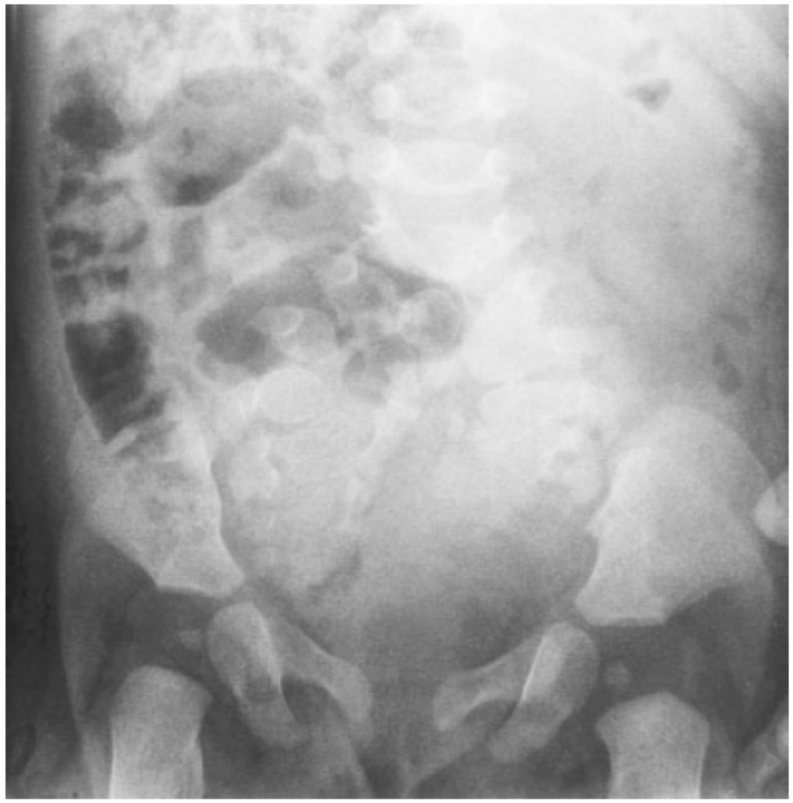
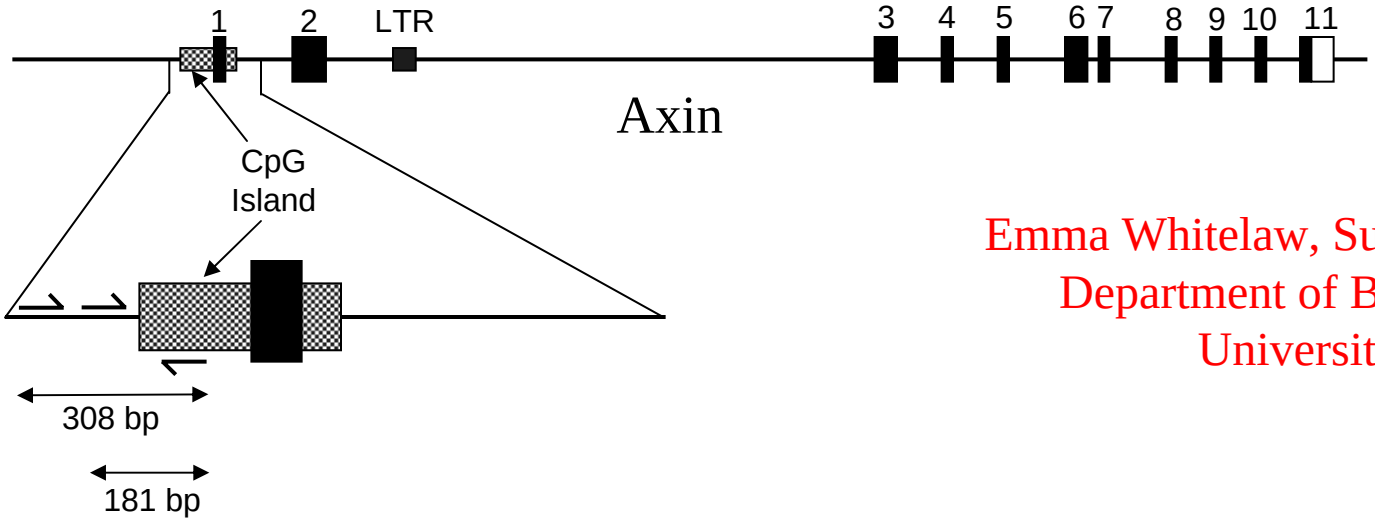


Fig. 2. Patient 1. Radiograph of the vertebral column shows complete duplication of the spine from L4 down.

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.

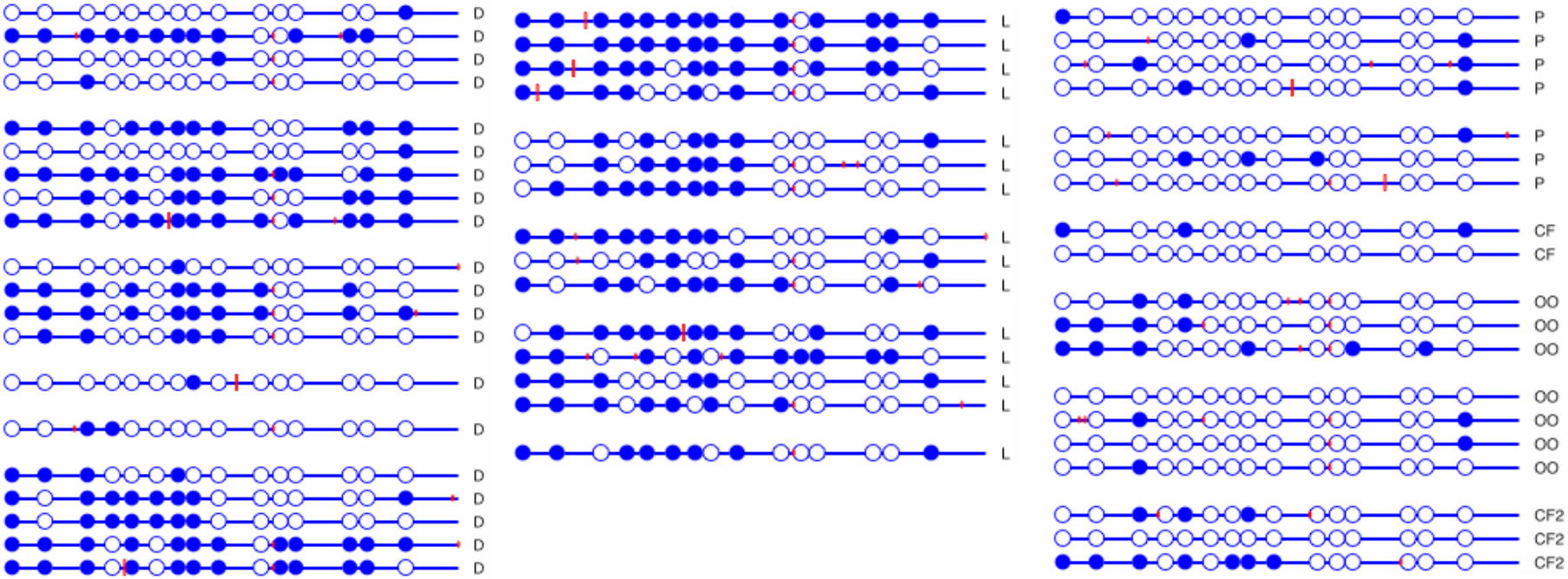
At physical examination for genetic evaluation at 4 months we saw a baby girl with epicanthal folds, but no other minor anomalies. She had a capillary nevus on her left buttock. In the anal region only a dimple was seen. The patient was operated on one day after birth, when a colostomy was made and a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

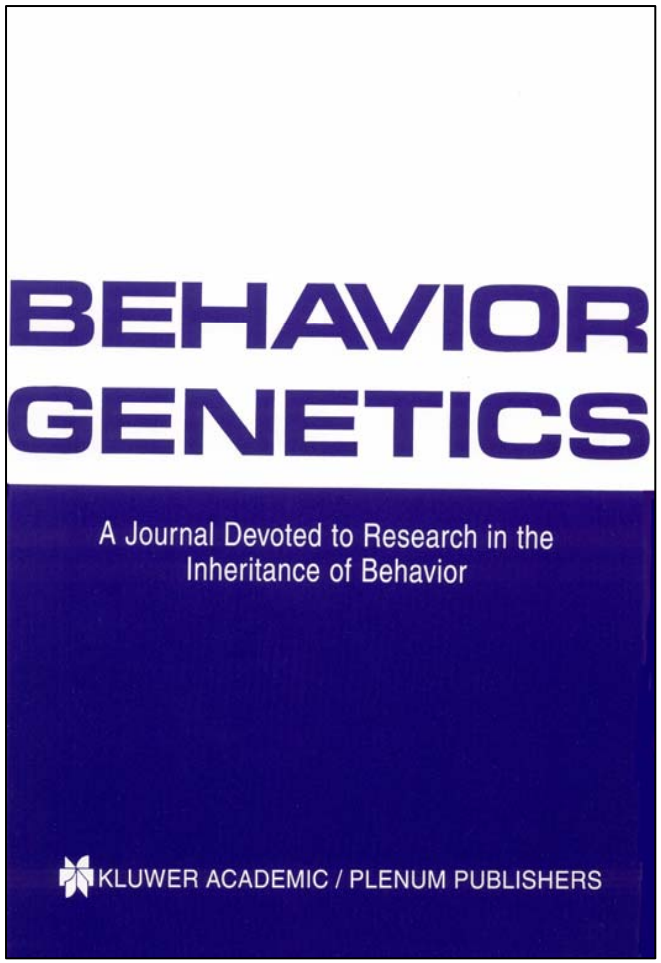
Epilepsy (with S. Berkovic, L. Vadlamudi)

Schizophrenia (with B.Mowry, N.Hayward)

Depression (with A. Petronis, D. Boomsma, P. McGuffin)

Asthma (with M.Ferreira, E.Whitelaw)

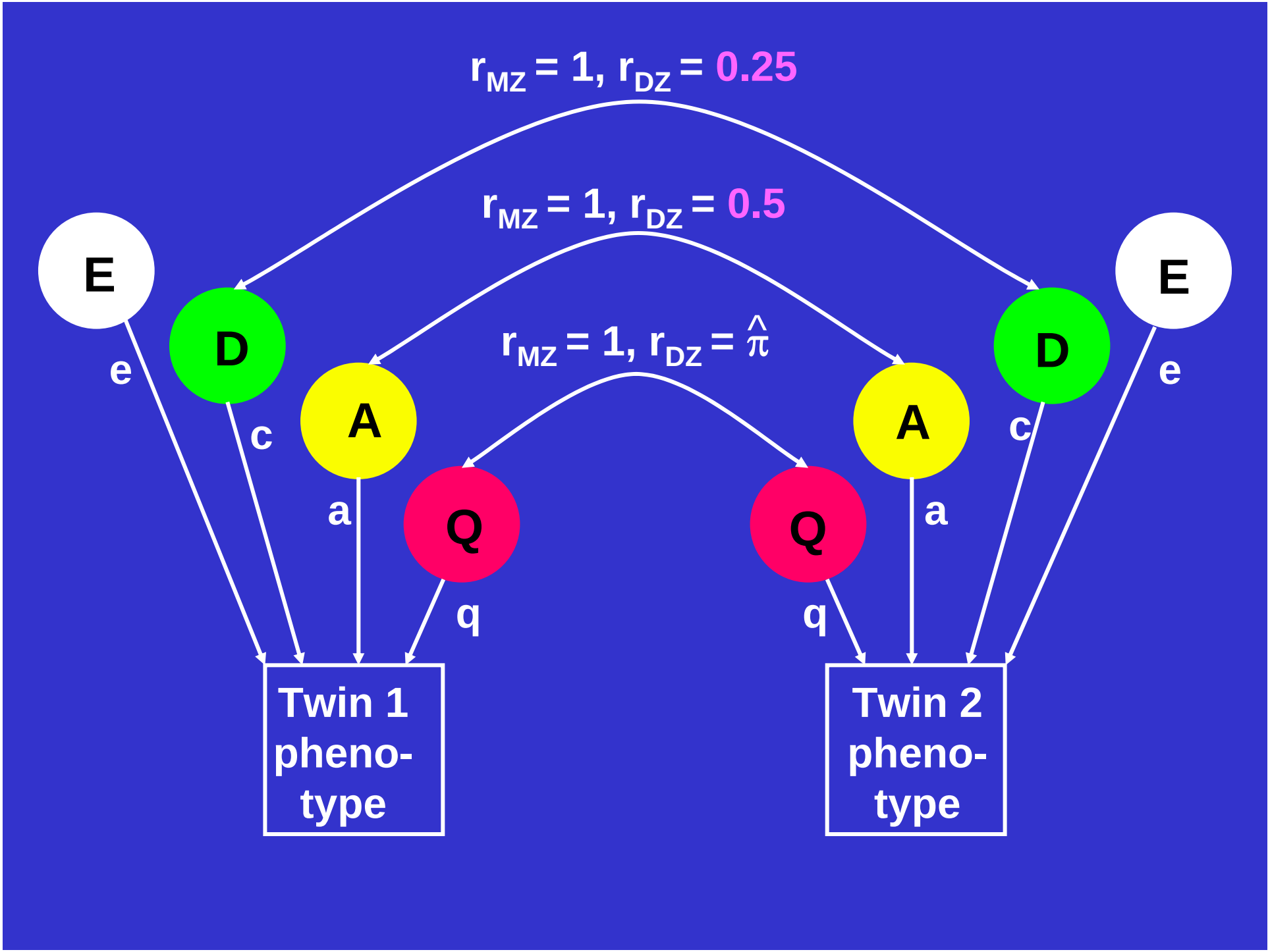
We also run two journals (1)



- Editor: John Hewitt
- Editorial assistant Christina Hewitt
- Publisher: Kluwer /Plenum
- Fully online
- http://www.bga.org

We also run two journals (2)

- Editor: Nick Martin
- Editorial assistant + subscriptions: Marisa Grimmer
- Publisher: Australian Academic Press
- Fully online
- http://www.ists.qimr.edu.au/journal.html


$r_{MZ} = 1, r_{DZ} = 0.25$
$r_{MZ} = 1, r_{DZ} = 0.5$
$r_{MZ} = 1, r_{DZ} = \hat{\pi}$
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 ?

Mean IBD sharing across the genome for the *j*th sib pair was based on IBD estimated from Merlin every centimorgan and averaged at all 3491 points

additive

$$\overline{\hat{\pi}}_{a(j)} = \sum_{i=1}^{3491} \hat{\pi}_{a(ij)} / 3491$$

dominance

$$\overline{\hat{\pi}}_{d(j)} = \sum_{i=1}^{3491} p_{2(ij)} / 3491$$

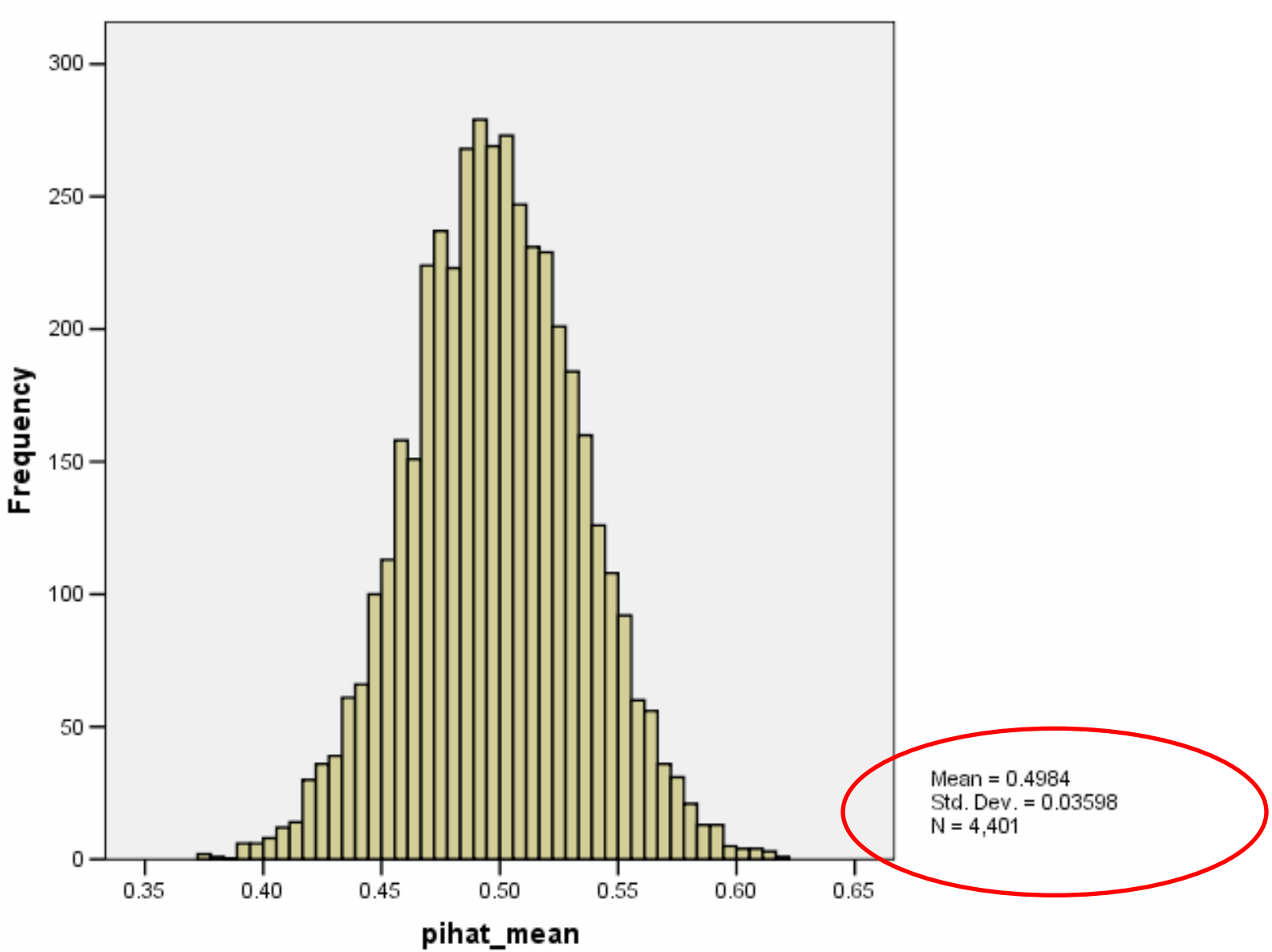
Application

- Phenotype = height

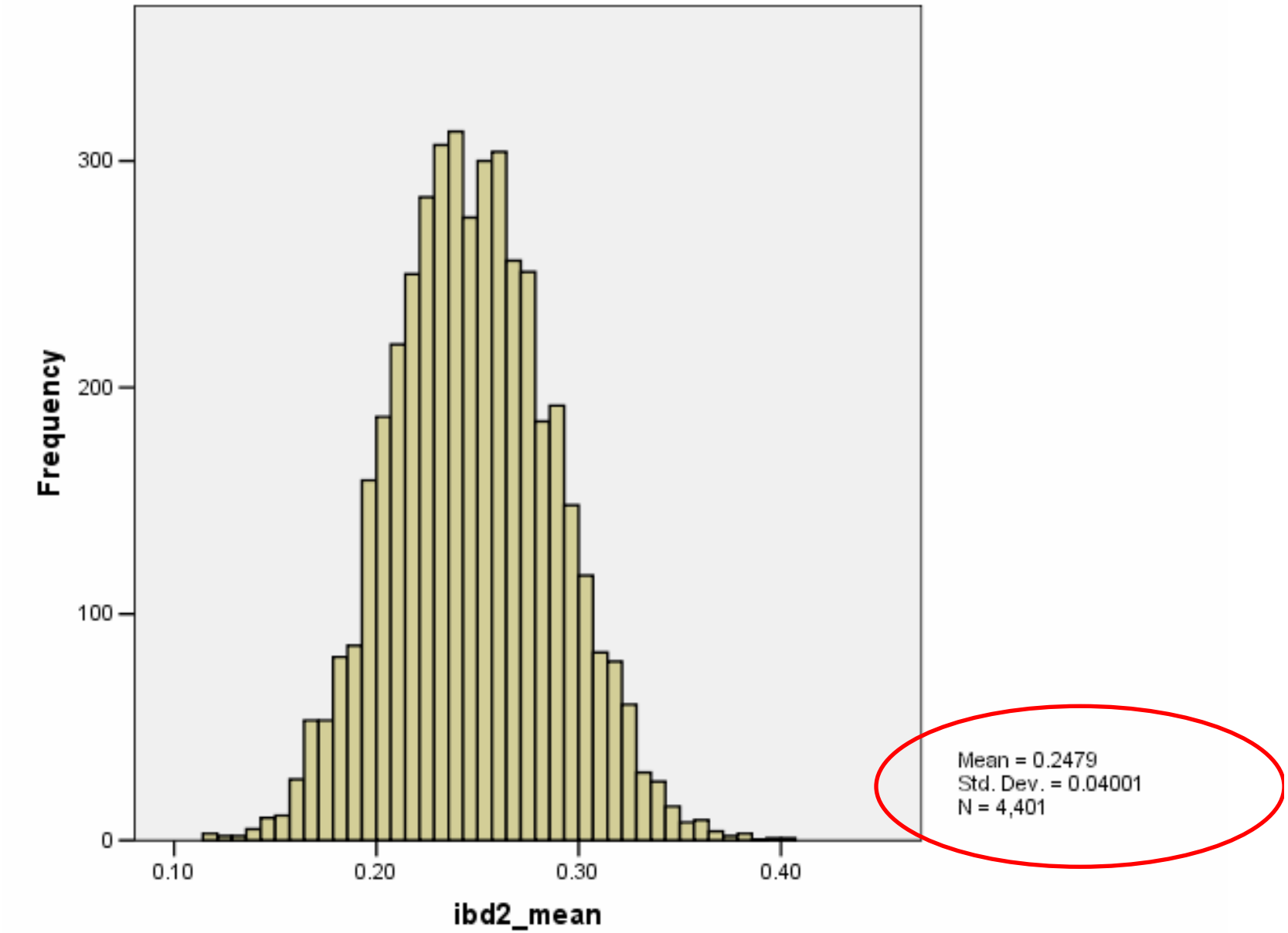
Number of sibpairs with phenotypes and genotypes

Adolescent cohort	931
Adult cohort	2444
Combined	3375

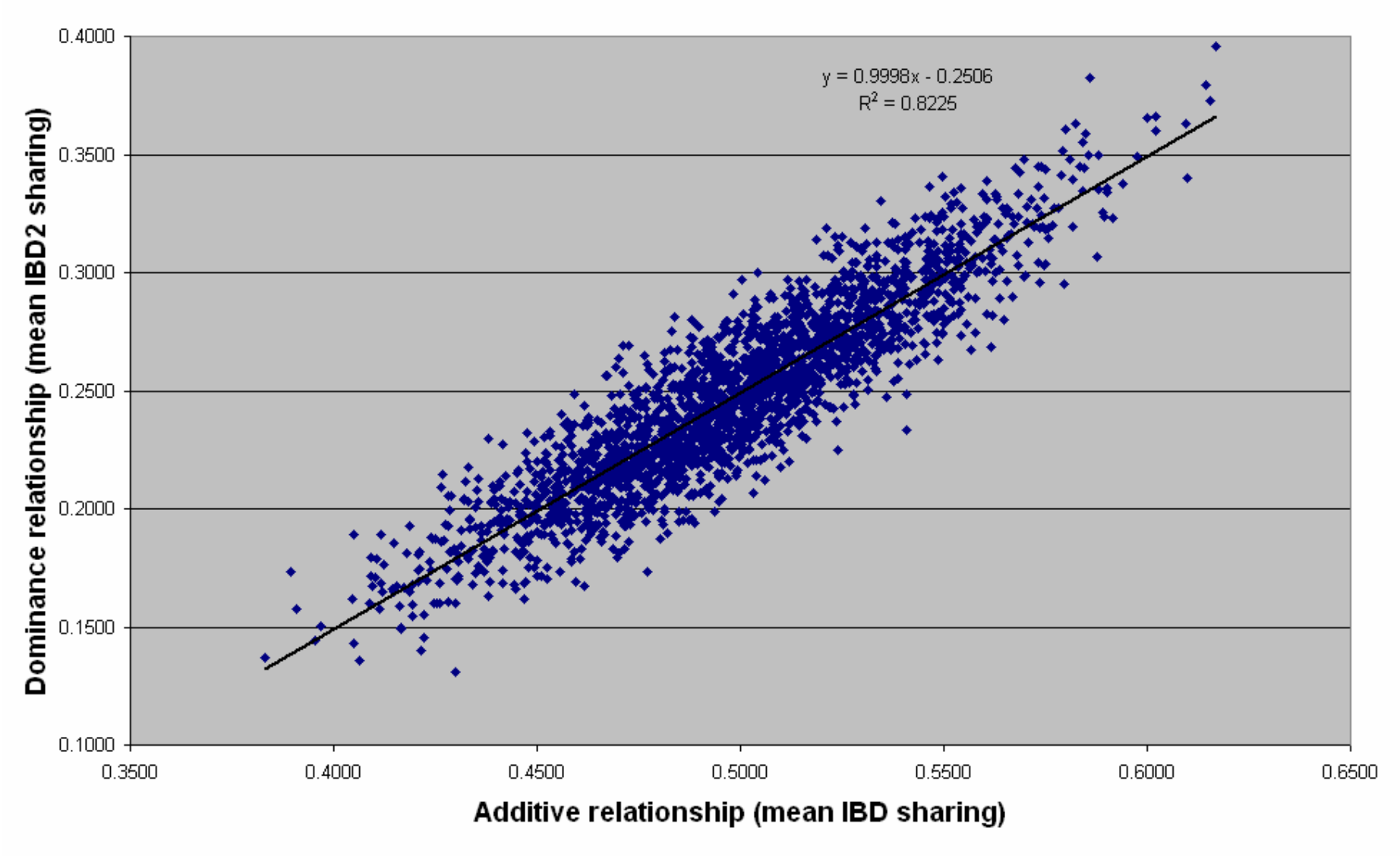
Mean and SD of genome-wide additive relationships



Mean and SD of genome-wide dominance relationships



Additive and dominance relationships correlation = 0.91 (*n*= 4401)



Models

F = Family effect

A = Genome-wide additive genetic

E = Residual

Full model $F + \bar{\hat{\pi}}_{a(j)}A + E$

Reduced model $F + E$

Sampling variances are large

Cohort	F+A (95% CI)
Adolescent	0.80 (0.36 – 0.90)
Adult	0.80 (0.61 – 0.86)
Combined	0.80 (0.62 – 0.85)

►***Estimates of MZ correlation from fullsibs!***

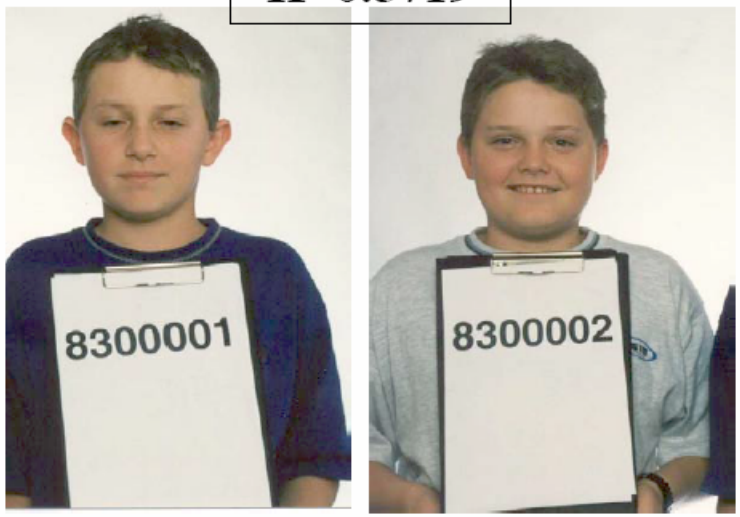
PLOS Genetics, *in press*

And now for IQ! Anyone got sibpairs with IQ + genome scan?

8188001,02
H=0.5677

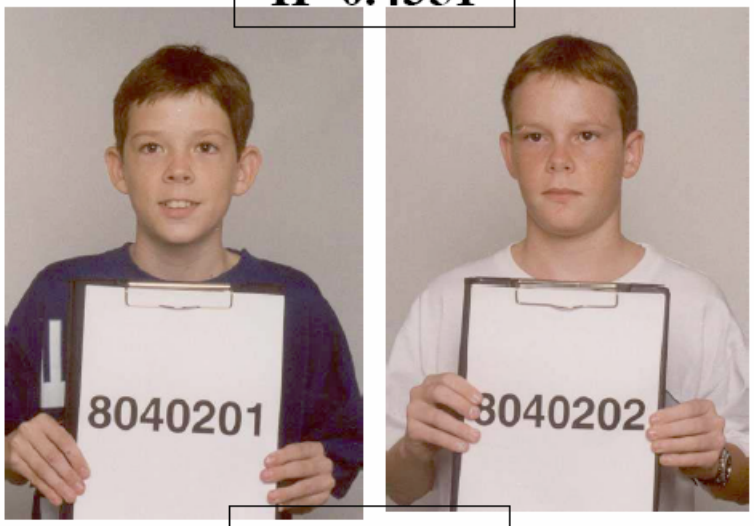
8473001,02
H=0.5577

8300001,02
H=0.5719



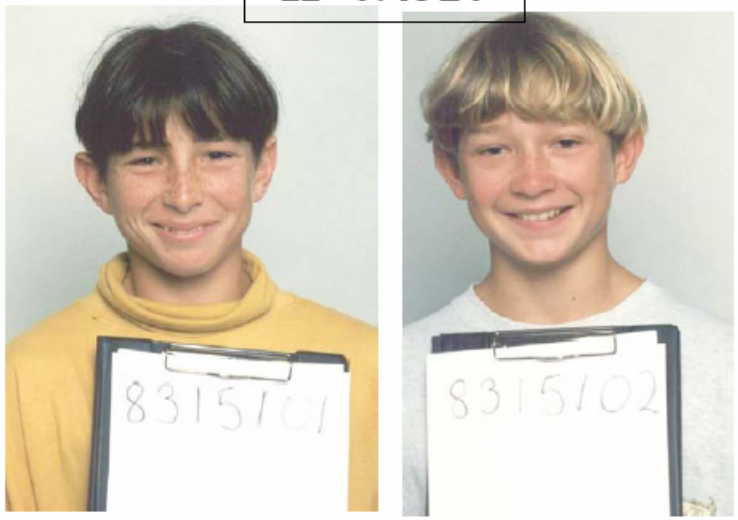
8582601,02
H=0.5640

8040201,02
H=0.4351

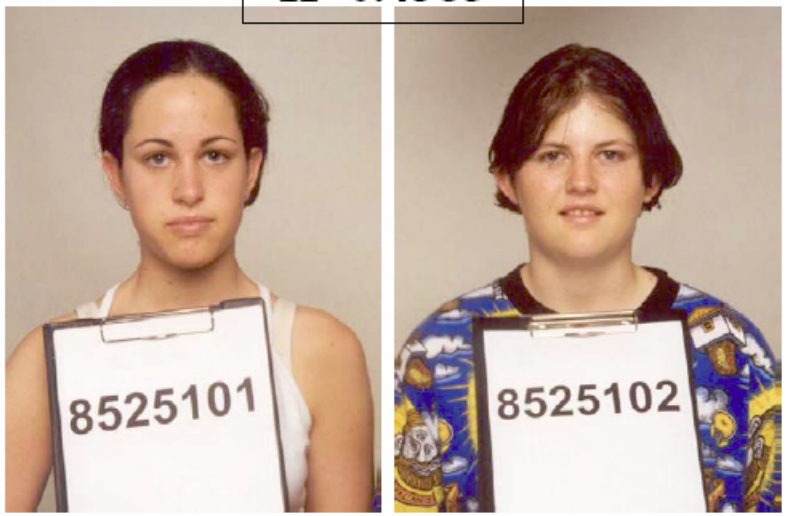


8069101,02
H=0.4291

8315101,02
H=0.4320



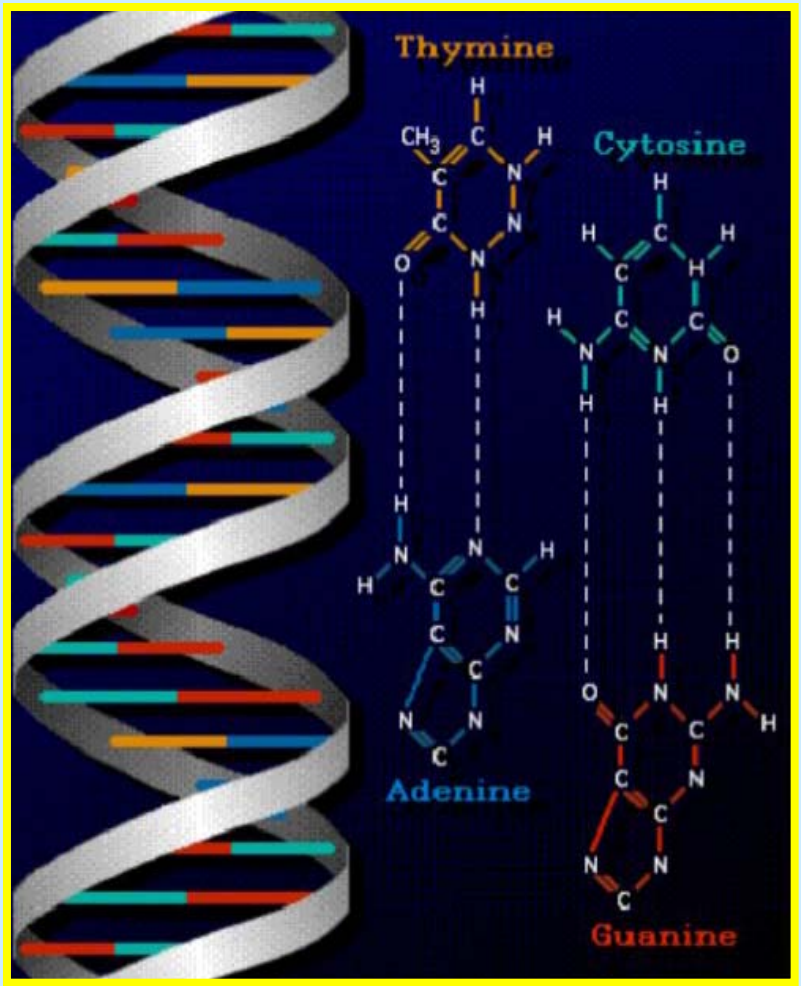
8525101,02
H=0.4385



Comparative Genomics

= differences in DNA sequence

Human-Human 1:1000 = 0.1%



Human-Chimp 1:100 = 1%

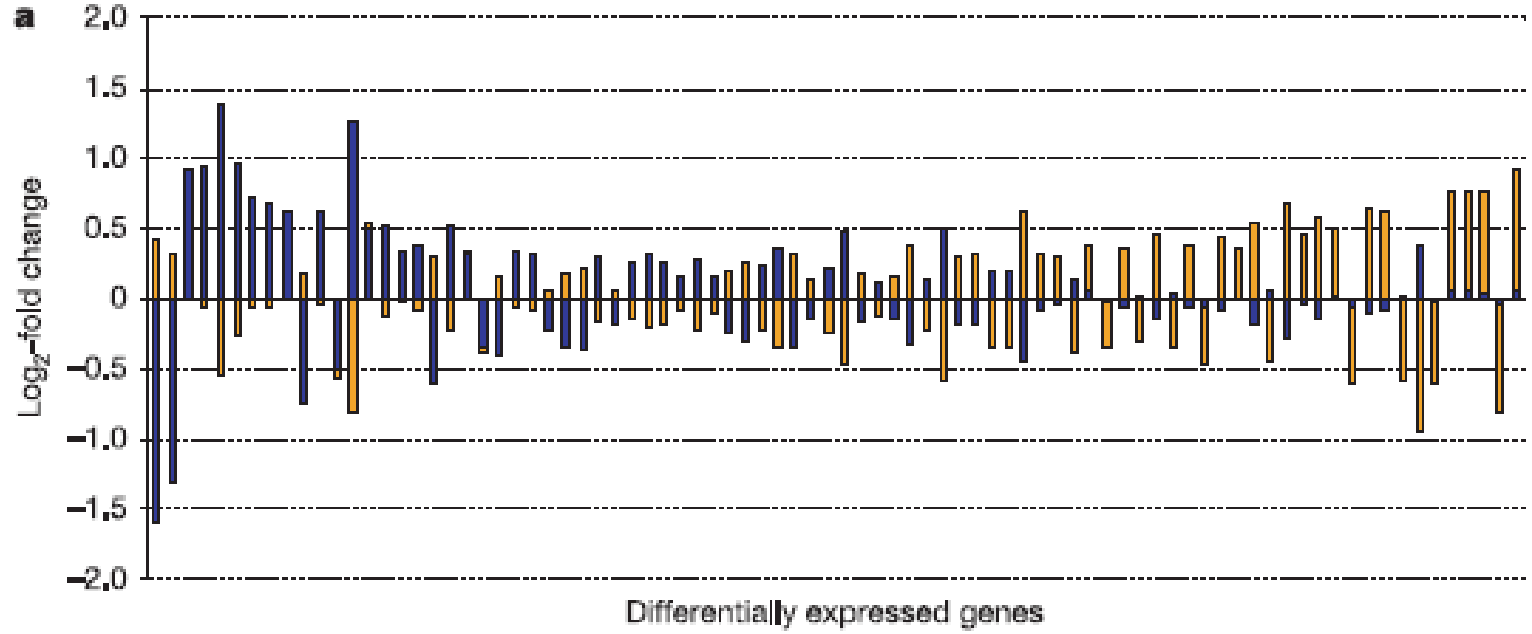
Human-Mouse 1:8 = 15%

nature Vol 440|9 March 2006|doi:10.1038/nature04559

LETTERS

Expression profiling in primates reveals a rapid evolution of human transcription factors

Yoav Gilad[1]†, Alicia Oshlack[2], Gordon K. Smyth[2], Terence P. Speed[2,3] & Kevin P. White[1]



features that point to the action of directional selection. Among the gene set with a human-specific increase in expression, there is an excess of transcription factors; the same is not true for genes with increased expression in chimpanzee.

Which genes have evolved fastest?

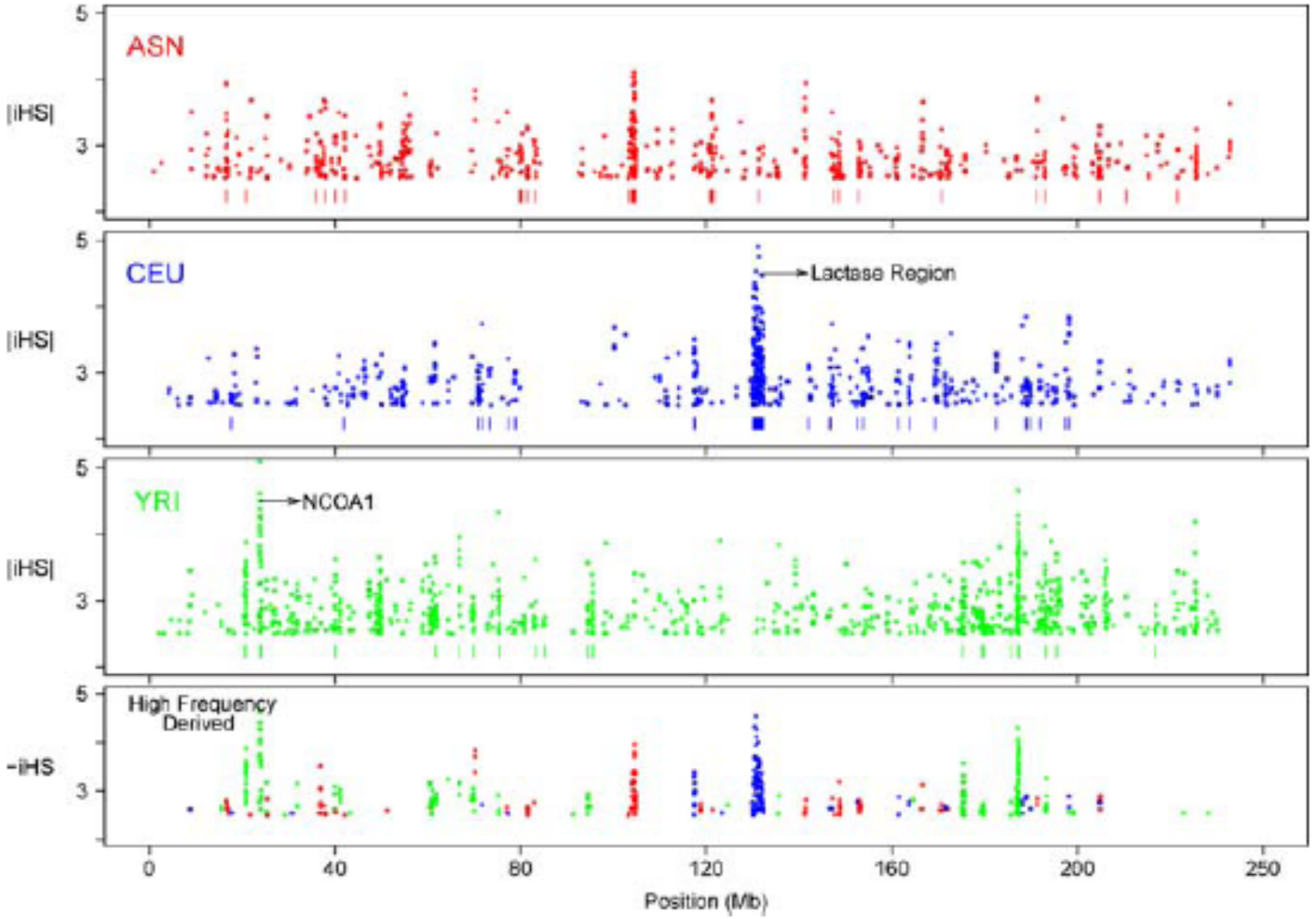


Figure 3. Plots of Chromosome 2 SNPs with Extreme iHS Values Indicate Discrete Clusters of Signals

Table 1. Summary of Some of the Strongest iHS Signals Genome-Wide

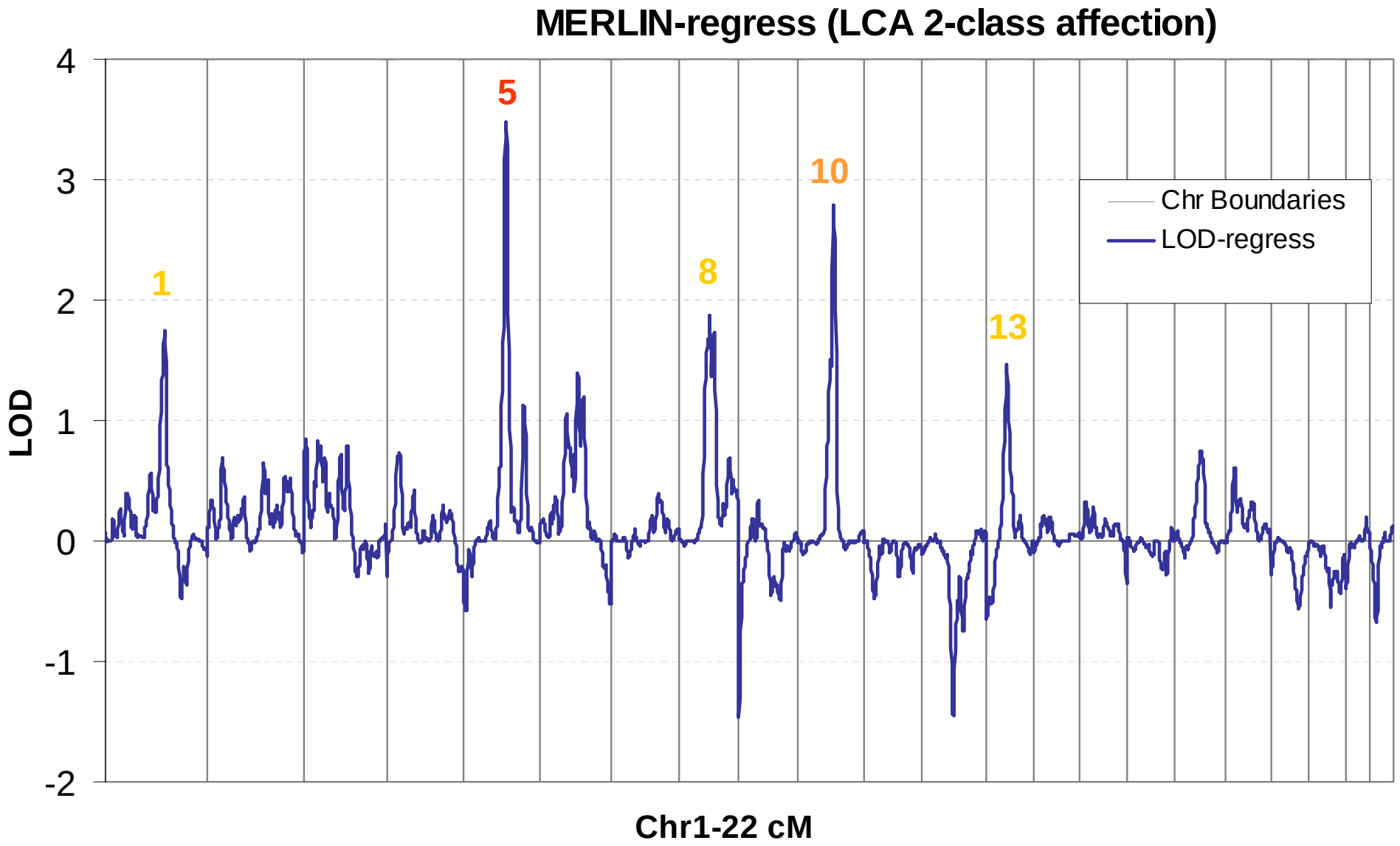
Cytological Position	Genes (Number)	Size (kb)	Pop	Number of SNPs with |iHS| >2.0
1p34.3	*NCDN, TEKT2* (17)	1,200	CEU	74/103
1p31.1	*SLC44A5* (4)	900	ASN	97/150
2p23.3	*NCOA1, ADCY3* (4)	400	YRI	51/76
2q12.3-q13	*SULT1C* cluster (13)	1,100	ASN	108/171
2q21.3-q22.1	*LCT* (15)	2,800	CEU	351/594
2q32.3	None (0)	400	YRI	100/131
4p15.1	None (0)	500	CEU	91/125
			YRI	43/146
4q21–23	*ADH* cluster (8)	100	ASN	21/28
8q11.21–23	*SNTG1* (8)	3,100	ASN	129/1297
			CEU	550/1201
			YRI	212/1451
9p22.3	*C9orf93*(1)	400	ASN	142/204
12q21.2	*SYT1* (3)	700	YRI	108/143
20cen	*ITGB4BP, CEP2,*	800	ASN	101/135
	SPAG4 (24)		CEU	50/153
			YRI	22/154

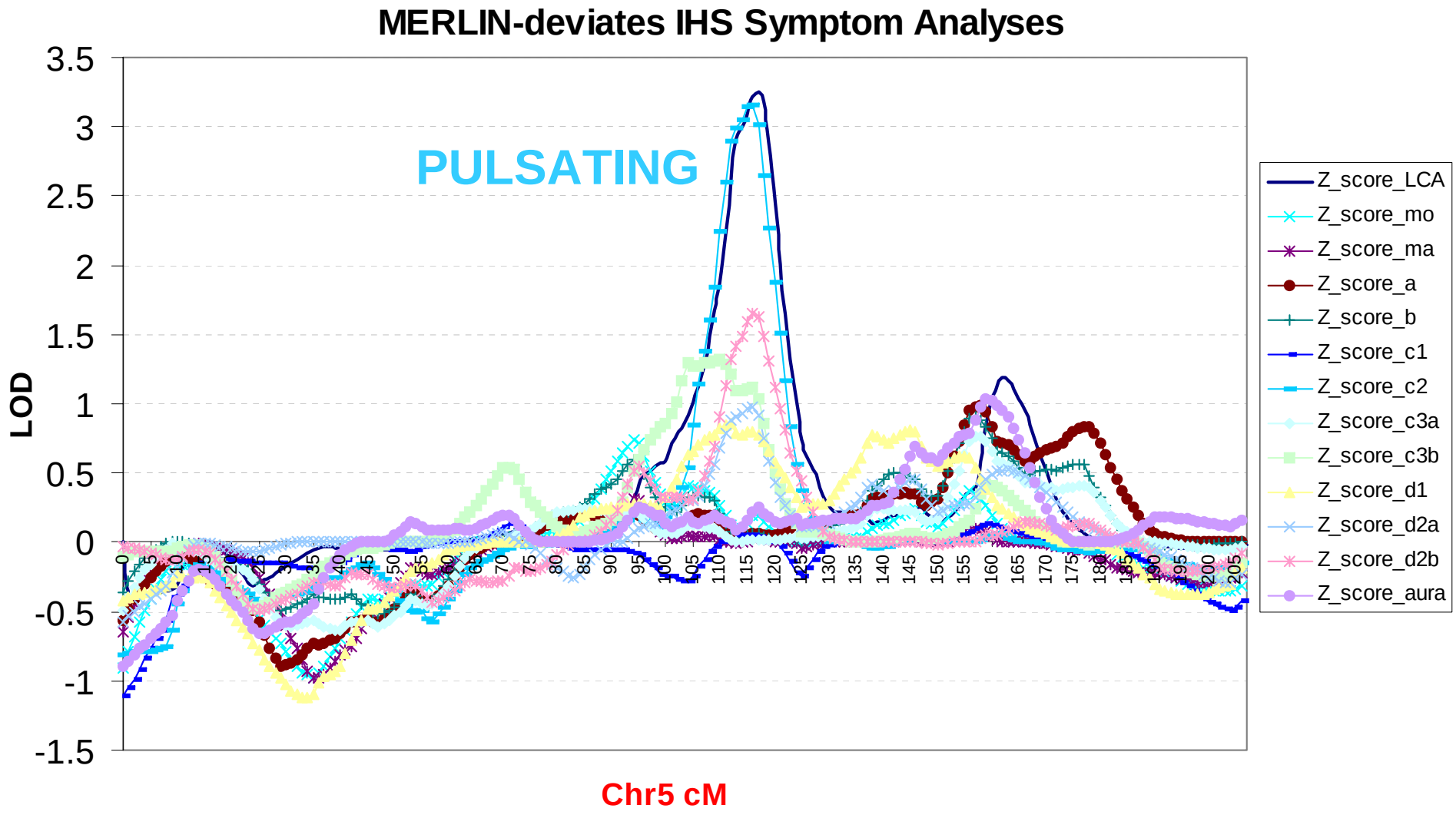
A Map of Recent Positive Selection in the Human Genome

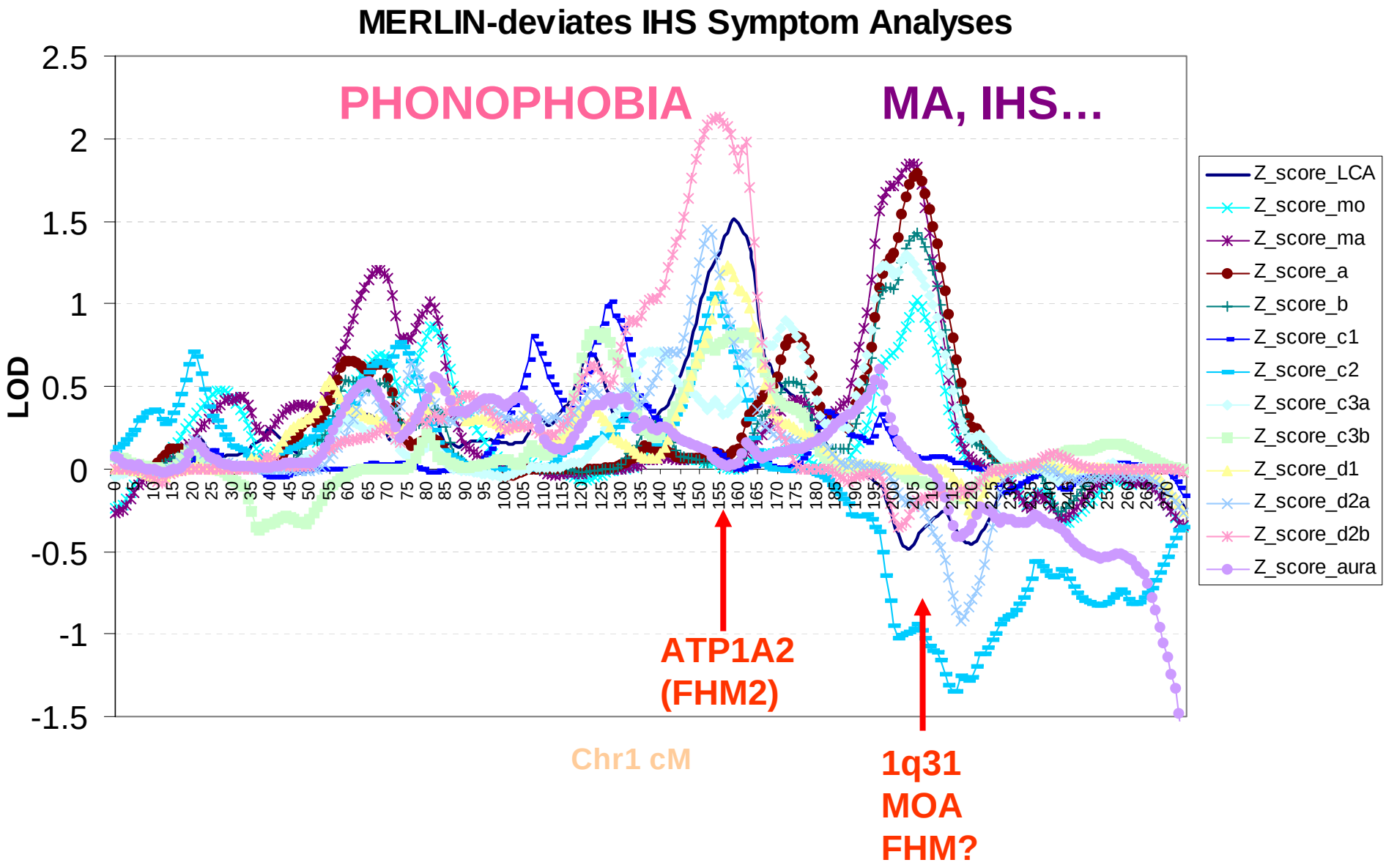
Benjamin F. Voight[☯], Sridhar Kudaravalli[☯], Xiaoquan Wen, Jonathan K. Pritchard[*]

PLoS Biology | www.plosbiology.org
March 2006 | Volume 4 | Issue 3 | e72

Migraine - Genome Scan Results




MERLIN-deviates IHS Symptom Analyses
PULSATING
LOD
Chr5 cM
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205
Z_score_LCA
Z_score_mo
Z_score_ma
Z_score_a
Z_score_b
Z_score_c1
Z_score_c2
Z_score_c3a
Z_score_c3b
Z_score_d1
Z_score_d2a
Z_score_d2b
Z_score_aura


MERLIN-deviates IHS Symptom Analyses
PHONOPHOBIA
MA, IHS…
LOD
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
Z_score_LCA
Z_score_mo
Z_score_ma
Z_score_a
Z_score_b
Z_score_c1
Z_score_c2
Z_score_c3a
Z_score_c3b
Z_score_d1
Z_score_d2a
Z_score_d2b
Z_score_aura
ATP1A2
(FHM2)
Chr1 cM
1q31
MOA
FHM?