Genetic Association Studies and GWASes

John F Pearson

Biostatistics and Computational Biology unit
Public Health and General Practice
University Otago Christchurch 18/3/2013

Menu

- Why do genetic association studies and what do we get from them?
- Some work from Christchurch

Heritability

Heritability allows a comparison of the relative importance of genes and environment to the variation of traits within and across populations.

Nature Reviews Genetics **9**, 255-266 (April 2008) | doi:10.1038/nrg2322

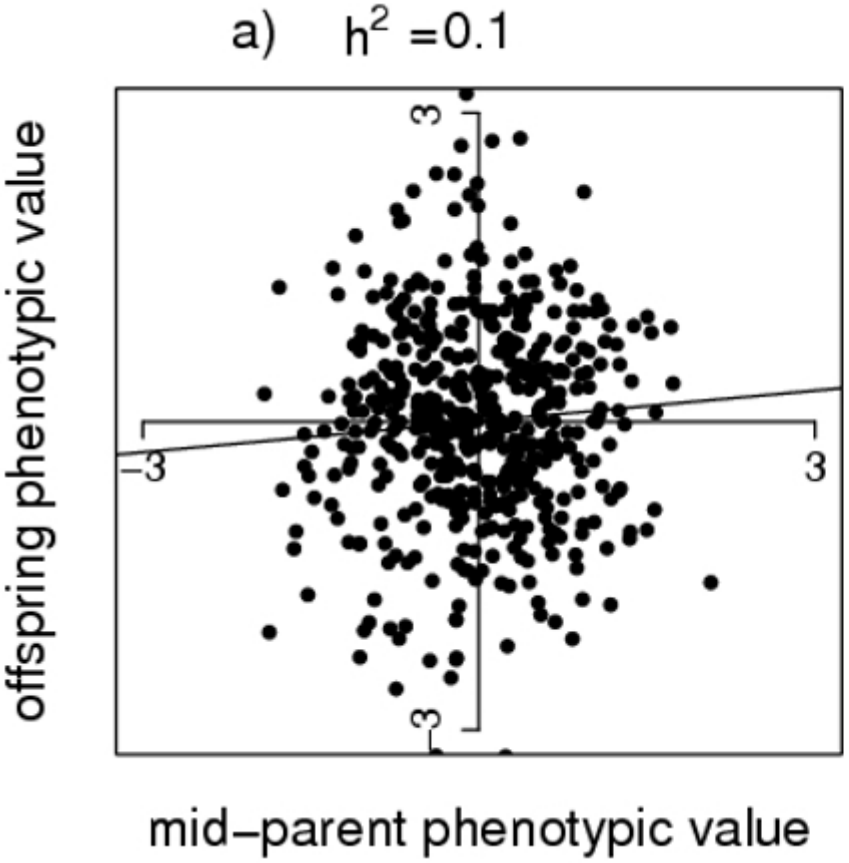
Heritability (Genetics 101)

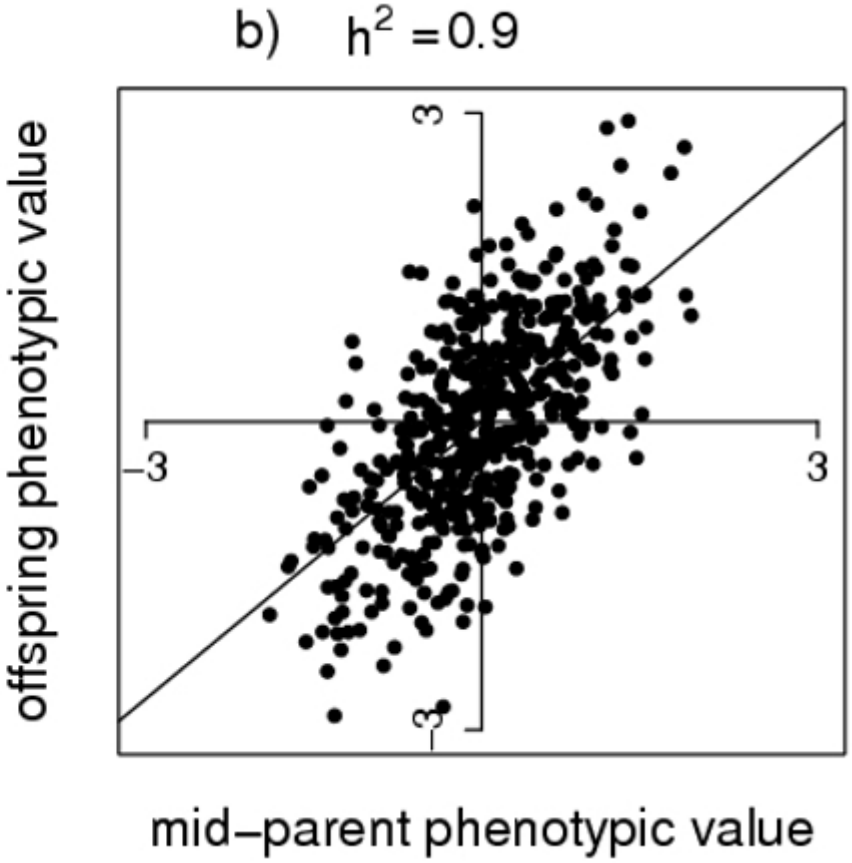
Phenotype (P) = Genotype (G) + Environment (E)

\+ ——

heritability

\+ + ——





heritability

Wray, N. & Visscher, P. (2008) Estimating trait heritability. *Nature Education* 1(1)

\+ \+ ——

Heritability

Transactions of the Royal Society of Edinburgh vol 52: 399-433 (1918)

–**The Correlation between Relatives on the Supposition of Mendelian Inheritance.** By **R. A. Fisher, B.A.** *Communicated by* Professor J. Arthur Thomson. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

"… the square of the standard deviation as the measure of variability. We shall term this quantity the Variance of the Normal population to which it refers,…"

http://digital.library.adelaide.edu.au/dspace/bitstream/2440/15097/1/9.pdf

Heritability

Transactions of the Royal Society of Edinburgh vol 52: 399-433 (1918)

–**The Correlation between Relatives on the Supposition of Mendelian Inheritance.** By **R. A. Fisher, B.A.** *Communicated by* Professor J. Arthur Thomson. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

I do not feel that this kind of work affects us biologists much at present. It is too much of the order of problem that deals with weightless elephants upon frictionless surfaces, where at the same time we are largely ignorant of the other properties of the said elephants and surfaces.

Yrs sincerely
R. C. Punnett

Notes and Records of the Royal Society of London, Vol. 31, No. 1 (Jul.,1976), pp. 151-162
http://www.jstor.org/stable/531555

Heritability

Transactions of the Royal Society of Edinburgh vol 52: 399-433 (1918)

The Correlation between Relatives on the Supposition of Mendelian Inheritance. By **R. A. Fisher, B.A.** *Communicated by* Professor J. Arthur Thomson. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

(1) Pearson's report on Fisher's paper (1916)

I have not examined in detail the results of this paper, as I am overfussed with other work, but submit the following remarks: The keynote of the memoir

Notes and Records of the Royal Society of London, Vol. 31, No. 1 (Jul.,1976), pp. 151-162
http://www.jstor.org/stable/531555

Heritability

Transactions of the Royal Society of Edinburgh vol 52: 399-433 (1918)

–The Correlation between Relatives on the Supposition of Mendelian Inheritance. By **R. A. Fisher, B.A.** *Communicated by* Professor J. Arthur Thomson. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

… the paper was inevitably refused because it had been sent for refereeing to ‘a “mathematician who knew no biology” and to a “biologist who knew no mathematics” `.

Attributed to R. A. Fisher

Notes and Records of the Royal Society of London, Vol. 31, No. 1 (Jul.,1976), pp. 151-162
http://www.jstor.org/stable/531555

Heritability

Proc Natl Acad Sci U S A. 1920 June; 6(6): 320–332.

THE RELATIVE IMPORTANCE OF HEREDITY AND ENVIRONMENT IN DETERMINING THE PIEBALD PATTERN OF GUINEA-PIGS

By Sewall Wright

Bureau of Animal Industry, United States Department of Agriculture

Communicated by R. Pearl, March 17, 1920

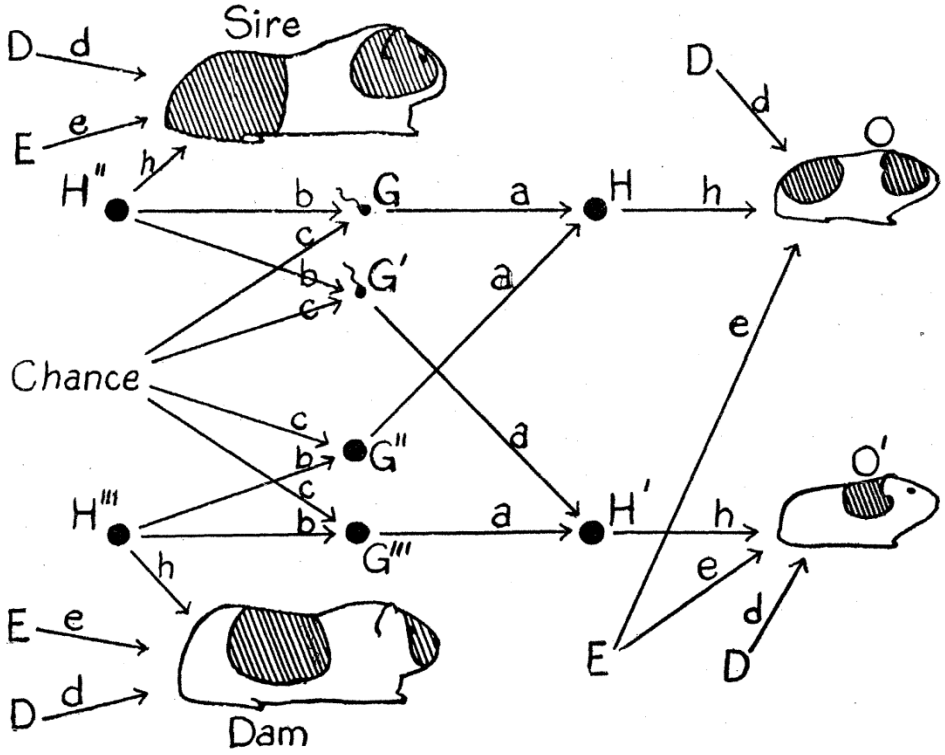
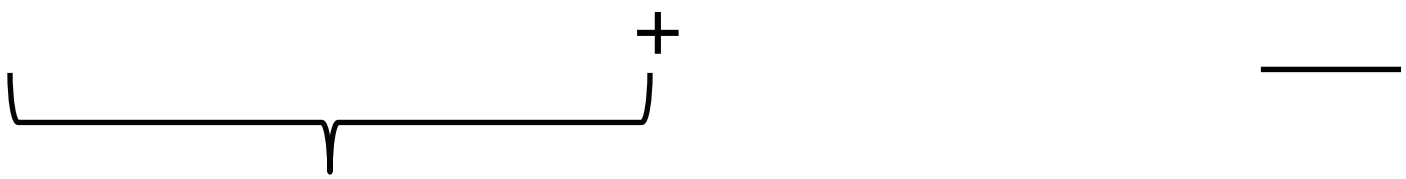


FIG. 5.

Heritability (Genetics 101)

Phenotype (P) = Genotype (G) + Environment (E)



Gene environment covariance

+ + 2

Heritability: human height

Annals of Human Biology, March–April 2007; 34(2): 206–215

The mysterious trend in American heights in the 20th century

JOHN KOMLOS[1] & BENJAMIN E. LAUDERDALE[2]

[1]*Chair of Economic History, University of Munich, Munich, Germany and* [2]*Department of Politics, Princeton University, Princeton, New Jersey, USA*

Results: After being the tallest population in the world ever since colonial times, Americans are now shorter than most Western and Northern Europeans and as much as 4.7–5.7 cm shorter than the Dutch, who are the tallest in world today.

Conclusion: it was hypothesized that either American diets are sub-optimal or that the universal health care systems and social safety net of the European welfare states are providing a more favorable early-life health environment than does the American health care system.

Heritability: human height

Standard deviation of height in pop is about 7cm

Reported heritability's are about 0.8

Nature Reviews Genetics **9**, 255-266 (April 2008) | doi:10.1038/nrg2322

Heritability: human height

Standard deviation of height in pop is about 7cm

Reported heritability's are about 0.8

Nature Reviews Genetics **9**, 255-266 (April 2008) | doi:10.1038/nrg2322

Heritability: human height

Standard deviation of height in pop is about 7cm

Reported heritability's are about 0.8

Standard deviation of adult offspring around parental mean height is about 5.4cm.

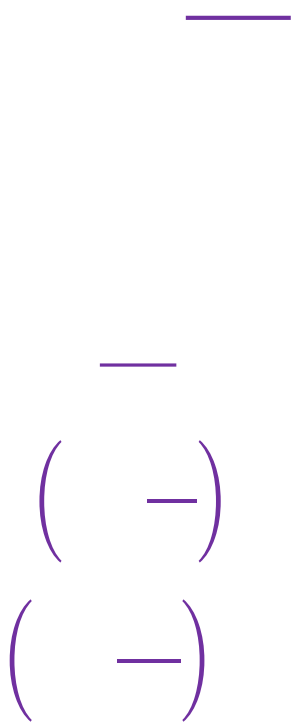
Nature Reviews Genetics **9**, 255-266 (April 2008) | doi:10.1038/nrg2322

Heritability: human height

Standard deviation of height in pop is about 7cm

Reported heritability's are about 0.8

Standard deviation of adult offspring around parental mean height is about 5.4cm.



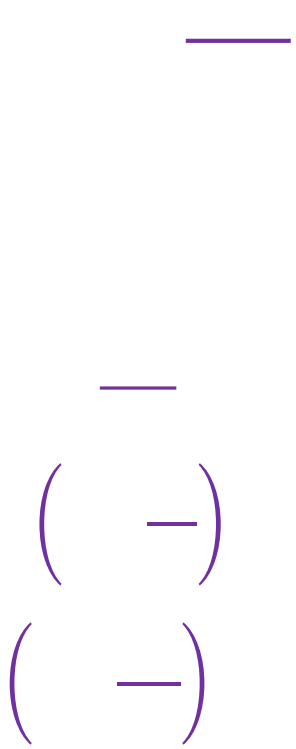
Heritability: human height

Standard deviation of height in pop is about 7cm

Reported heritability's are about 0.8

Standard deviation of adult offspring around parental mean height is about 5.4cm.

Not much smaller than sd in the population.



Heritability: human height

Nature Genetics **39**, 1245 - 1250 (2007)
Published online: 2 September 2007 | doi:10.1038/ng2121

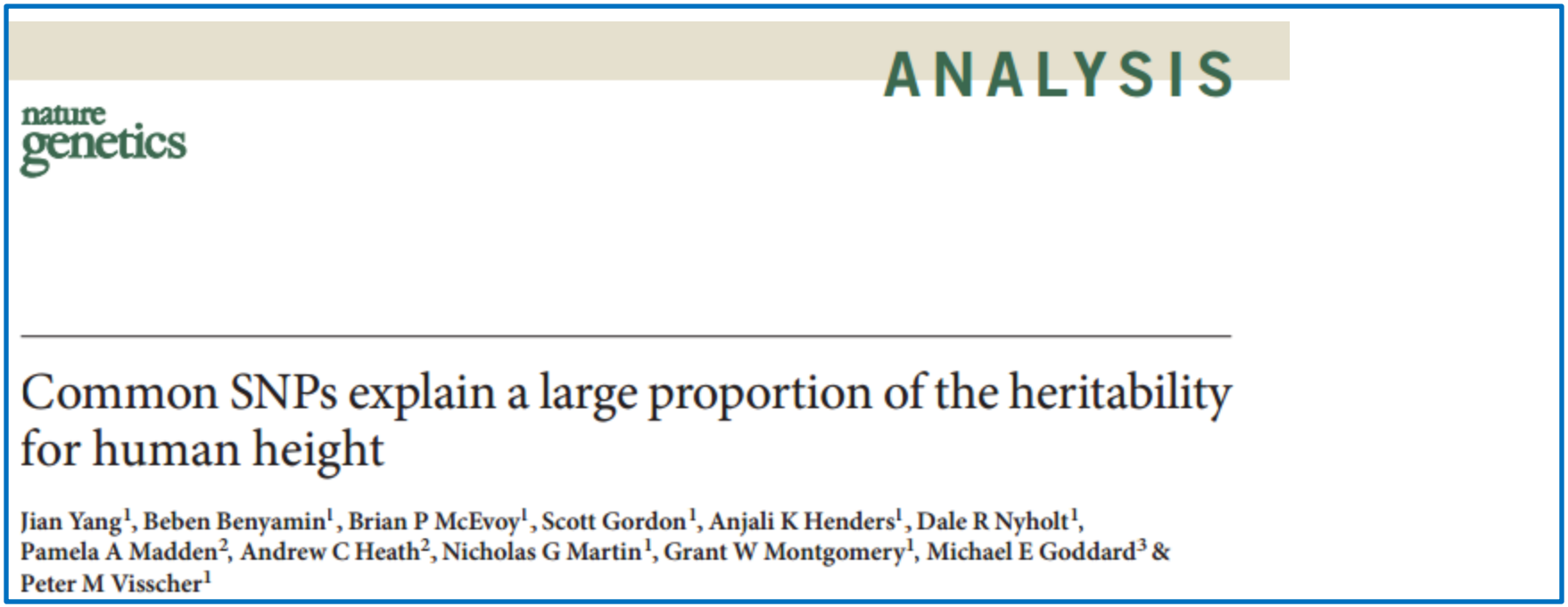
A common variant of *HMGA2* is associated with adult and childhood height in the general population

Michael N Weedon[1,2,21], Guillaume Lettre[3,4,21], Rachel M Freathy[1,2,21],

"We estimate that rs1042725 explains 0.3% of population variation in height (0.4 cm increased adult height per C allele)."

High Heritability does not imply high effect genes.

Heritability: human height



ANALYSIS

nature genetics

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

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

We show that 45% of variance can be explained by considering all SNPs simultaneously. Thus, most of the heritability is not missing but has not previously been detected because the individual effects are too small to pass stringent significance tests.

No high effect genes does not imply low heritability.

Missing heritability

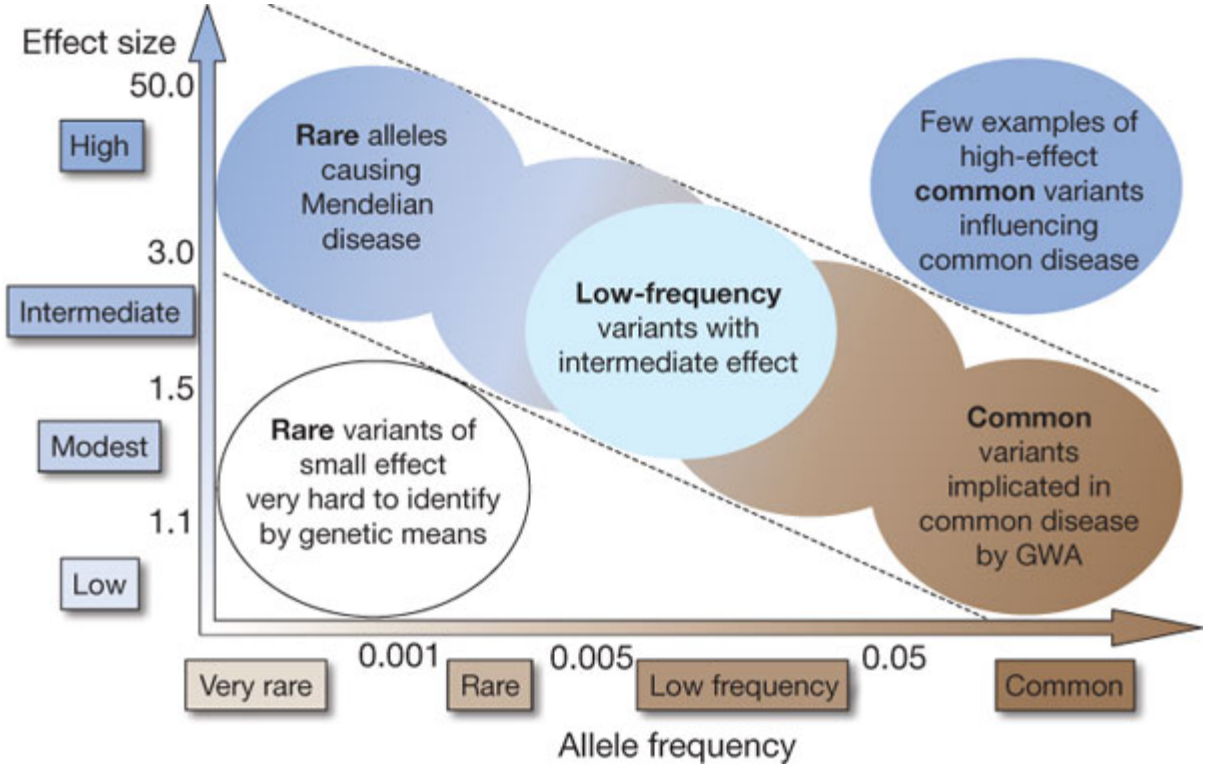
TA Manolio *et al. Nature* **461**, 747-753 (2009) doi:10.1038/nature08494

The case of the missing heritability

Missing heritability

TA Manolio *et al. Nature* **461**, 747-753 (2009) doi:10.1038/nature08494

Feasibility of identifying genetic variants by risk allele frequency and strength of genetic effect (odds ratio).



Heritability: human height

Heritability:

- Is a ratio of the additive genetic variance to the phenotypic variance
- Is specific to a population

High Heritability:

- Does not imply genetic determination
- Does not imply high genetic effects
- Is not informative about between group means

Hereditability is not constant

Genetic association studies

What's being done at UOC (partial view)

Candidate gene

Expression association

GxE

Genome wide association

CNV

Candidate gene studies

Replication

Refined phenotype

Candidate gene studies

Am J Gastroenterol 2012; 107:589–596; doi:10.1038/ajg.2011.437; published online 13 December 2011

Clinical and Genetic Risk Factors for Perianal Crohn's Disease in a Population-Based Cohort

T W Eglinton MBChB, FRACS[1,2], R Roberts BSc (Hons) I, PhD[3], J Pearson BSc, MSc, PhD[4], M Barclay MBChB, MD, FRACP[5,6], T R Merriman PhD[3], F A Frizelle MBChB, M Med Sc, FACS, FRACS[1,2] and R B Gearry MBChB, PhD, FRACP[5,6]

Candidate gene studies: Replication

Gene	Locus	alleles	Additive		
			OR	CI	P
CARD8	rs2043211	A/T	0.83	(0.69-1.00)	0.056
NALP3	rs3582941	C/A	0.97	(0.62-1.51)	0.903
IL2_21	rs6822844	T/G	12.14	(9.14-16.46)	$<10^{-15}$
IRGM	rs1336118	T/C	1.42	(1.01-2.00)	0.042
IRGM	rs4958847	G/A	1.40	(1.04-1.89)	0.025
NCF4	rs4821544	C/T	1.11	(0.91-1.35)	0.311
ATG16L1	rs2241880	T/C	0.71	(0.59-0.84)	0.0001
NOD2	rs2066842	C/T	1.22	(0.97-1.52)	0.085
NOD2	rs2066844	C/T	1.56	(1.10-2.25)	0.013
NOD2	rs2066845	G/C	1.76	(0.96-3.32)	0.035
NOD2	rs2066847	0/C	2.21	(1.25-4.07)	0.011
IL23R	rs1120902	G/A	0.51	(0.32-0.78)	0.004

"…The Bonferroni correction for multiple testing of 13 SNPs (ten genes) adjusts the significance level to 0.0038 for these results…"

Am J Gastroenterol 2012; 107:589–596; doi:10.1038/ajg.2011.437

Candidate gene studies: Refined phenotype

Genotypic associations for Perianal Disease within Crohn's Disease cohort.

Gene	Locus	alleles	Additive			
			OR	(CI)	P	P adjusted
CARD8	rs2043211	A/T	0.98	(0.71-1.33)	0.88	0.84
NALP3	rs35829419	C/A	1.15	(0.54-2.30)	0.70	0.48
IL2_21	rs6822844	T/G	1.02	(0.73-1.42)	0.92	0.50
IRGM	rs13361189	T/C	1.40	(0.85-2.27)	0.18	0.23
IRGM	rs4958847	G/A	1.22	(0.79-1.85)	0.37	0.35
NCF4	rs4821544	C/T	1.46	(1.08-1.98)	0.01	0.03
ATG16L1	rs2241880	T/C	0.93	(0.69-1.24)	0.61	0.42
NOD2	rs2066842	C/T	1.26	(0.91-1.75)	0.16	0.59
NOD2	rs2066844	C/T	1.32	(0.80-2.13)	0.26	0.84
NOD2	rs2066845	G/C	1.53	(0.63-3.49)	0.33	0.75
NOD2	rs2066847	0/C	1.19	(0.55-2.41)	0.64	0.74
IL23R	rs11209026	G/A	0.97	(0.40-2.15)	0.94	0.94

Am J Gastroenterol 2012; 107:589–596; doi:10.1038/ajg.2011.437

Multiple testing

Family wise error rate (FWER):

Control probability of making 1 false positives.

Bonferroni: declare significance at α/n (0K … 0.05/n)

- Tests that no genes are associated
- High probability of rejecting true results (type II error)

False discovery rate (FDR)

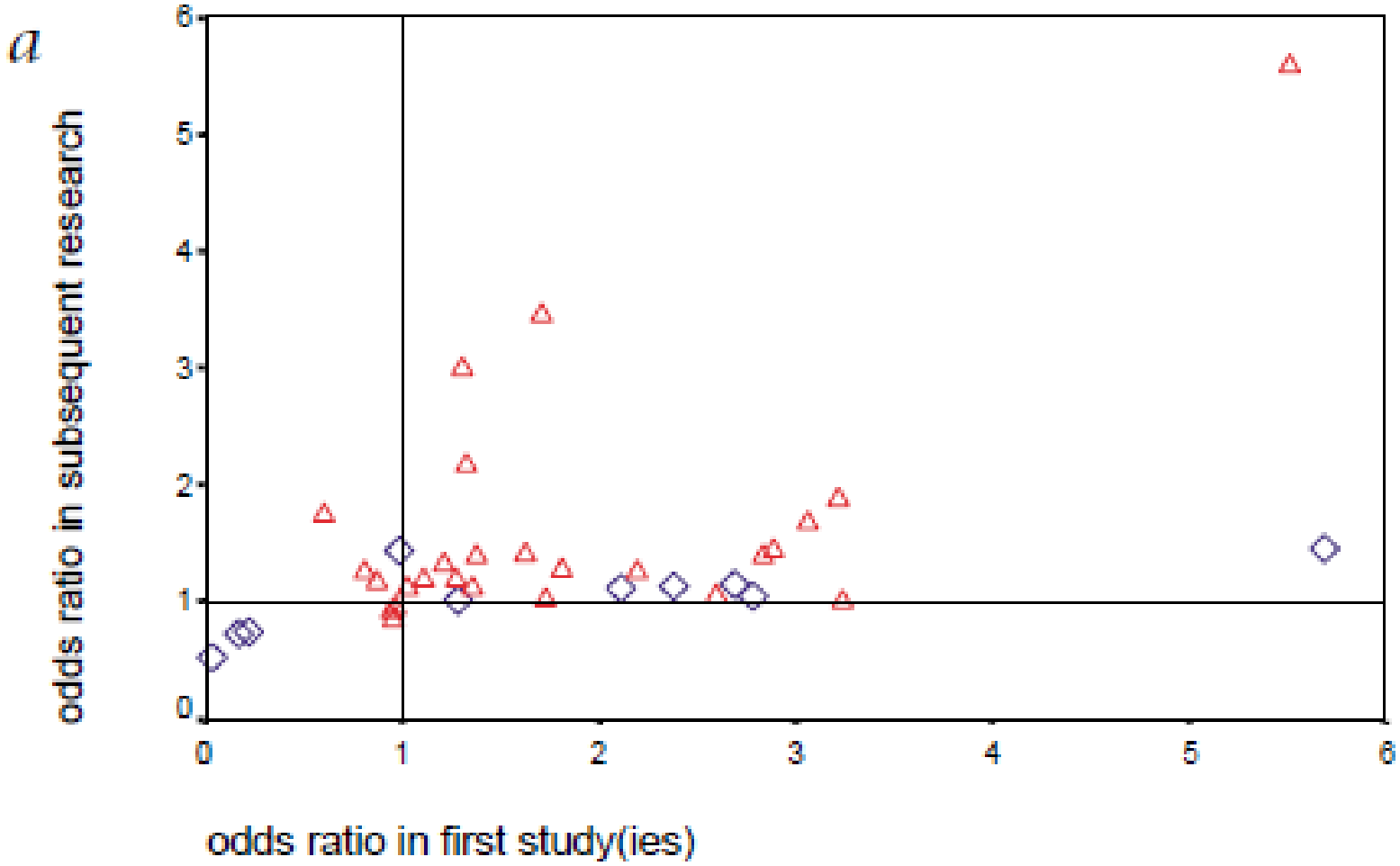
- We can usually cope with a few false discoveries, just control the rate at which we make them.
- BH method adapted to BKY procedure which is also robust to departures from independence .

Benjamini, Y. and Hochberg, Y. (1995) Journal of the Royal Statistical Society, Series B, 57, 289-300.
Benjamini, Y. and Yekutieli, D. (2001).. Annals of Statistics, 29, 1165-1188.

Nature Genetics **29**, 306 - 309 (2001)
Published online: 15 October 2001; | doi:10.1038/ng749

Replication validity of genetic association studies

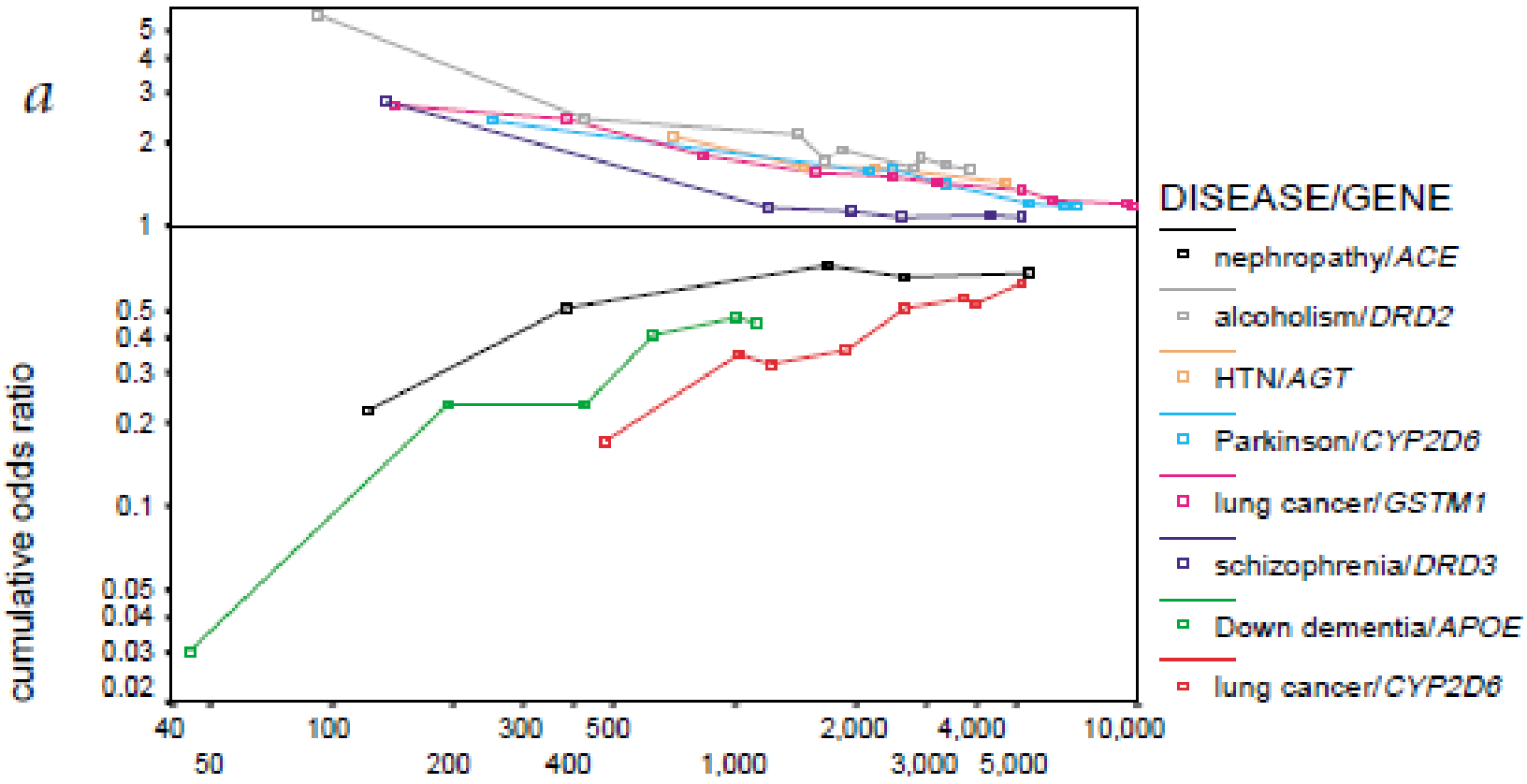
John P.A. Ioannidis[1, 2, 3], Evangelia E. Ntzani[1], Thomas A. Trikalinos[1] & Despina G. Contopoulos-Ioannidis[1, 4]



Nature Genetics **29**, 306 - 309 (2001)
Published online: 15 October 2001; | doi:10.1038/ng749

Replication validity of genetic association studies

John P.A. Ioannidis[1, 2, 3], Evangelia E. Ntzani[1], Thomas A. Trikalinos[1] & Despina G. Contopoulos-Ioannidis[1, 4]

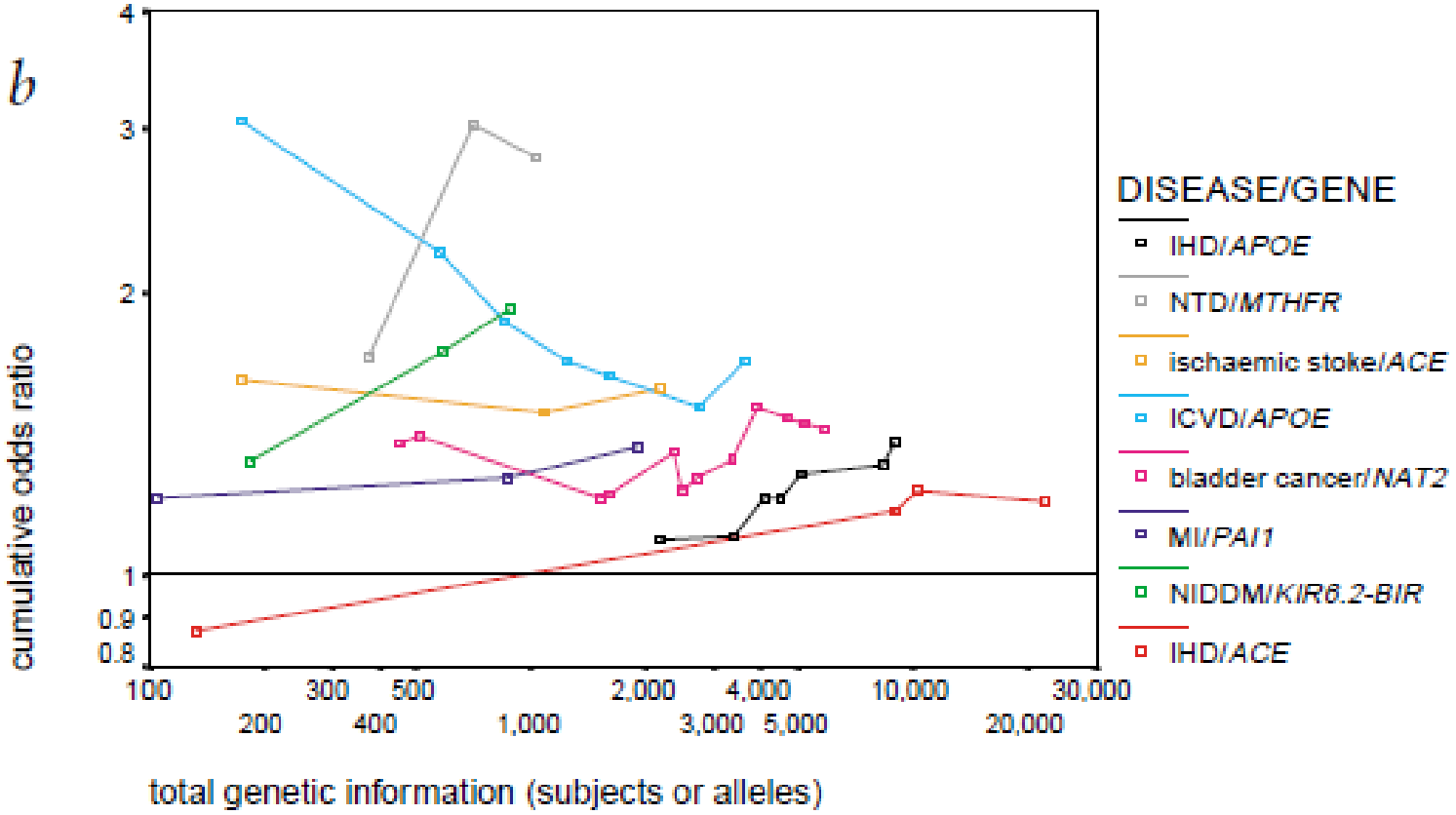


a, Eight topics in which the results of the first study or studies differed beyond chance (*P*<0.05)

Nature Genetics **29**, 306 - 309 (2001)
Published online: 15 October 2001; | doi:10.1038/ng749

Replication validity of genetic association studies

John P.A. Ioannidis[1, 2, 3], Evangelia E. Ntzani[1], Thomas A. Trikalinos[1] & Despina G. Contopoulos-Ioannidis[1, 4]

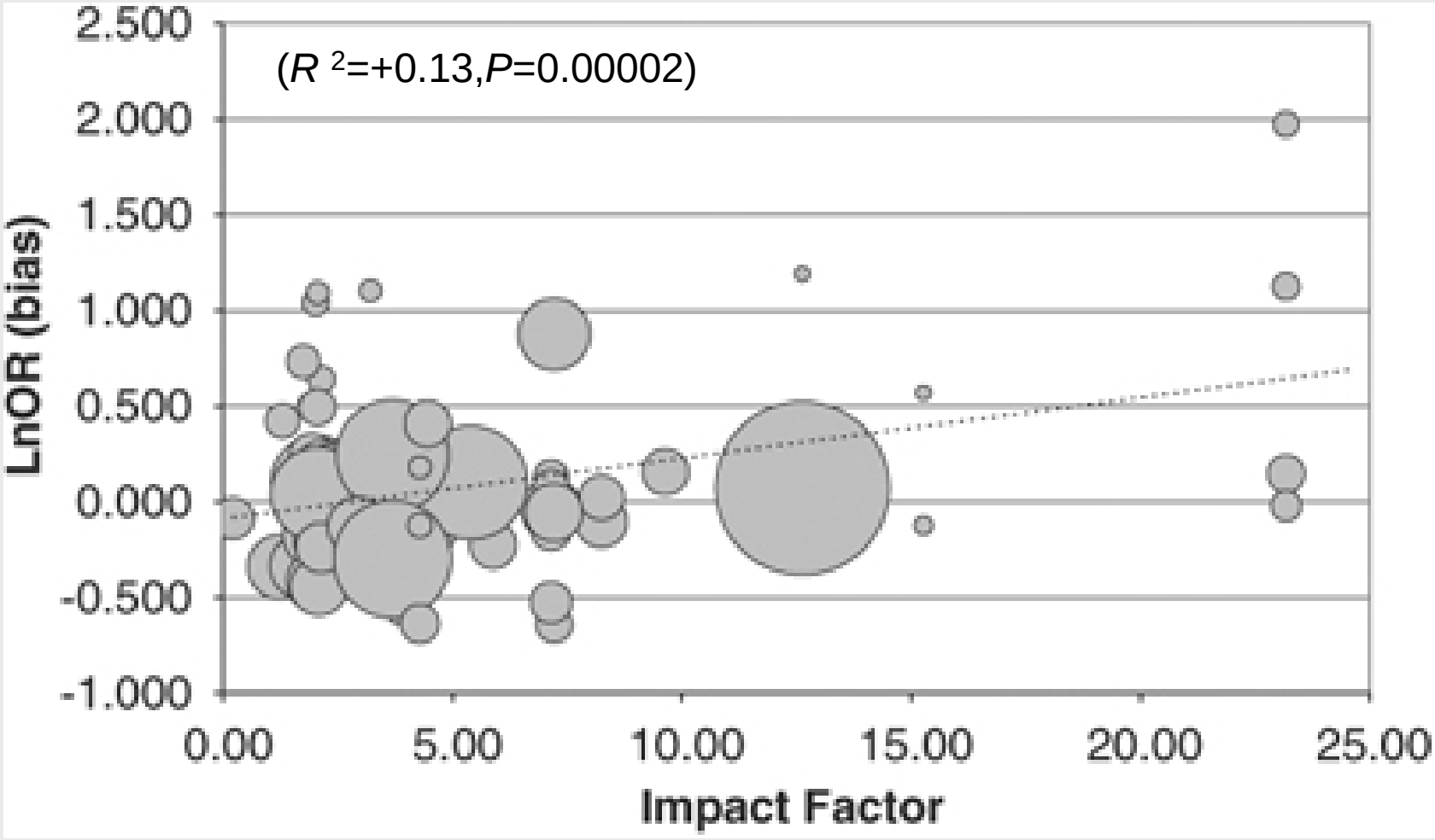


b, Eight topics in which the first study or studies did not claim formal statistical significance for the genetic association

Molecular Psychiatry (2009) **14,** 119–120; doi:10.1038/ mp.2008.77

Bias in genetic association studies and impact factor

M R Munafò[1], G Stothart[1] and J Flint[2]

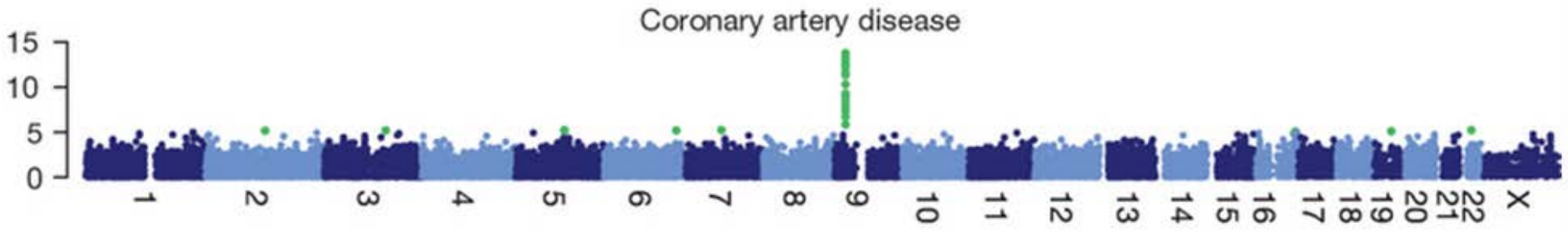


Genome wide association study

Nature **447**, 661-678 (7 June 2007) | doi:10.1038/nature05911; Received 26 March 2007; Accepted 11 May 2007

Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls

- The Wellcome Trust Case Control Consortium



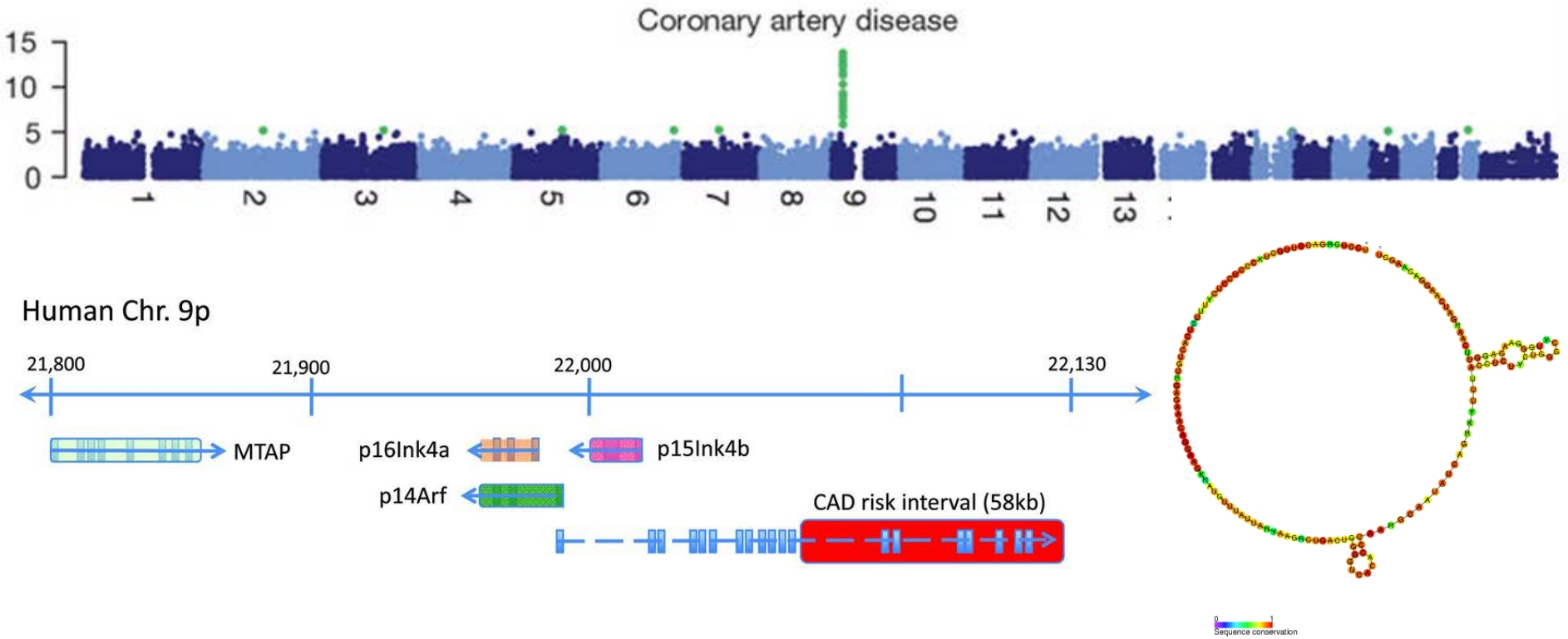
Circulation.2012; 126: 1896–1906

Genome wide association study

Nature **447**, 661-678 (7 June 2007) | doi:10.1038/nature05911; Received 26 March 2007; Accepted 11 May 2007

Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls

- The Wellcome Trust Case Control Consortium



Circulation. 2012; 126: 1896-1906

Gene Expression

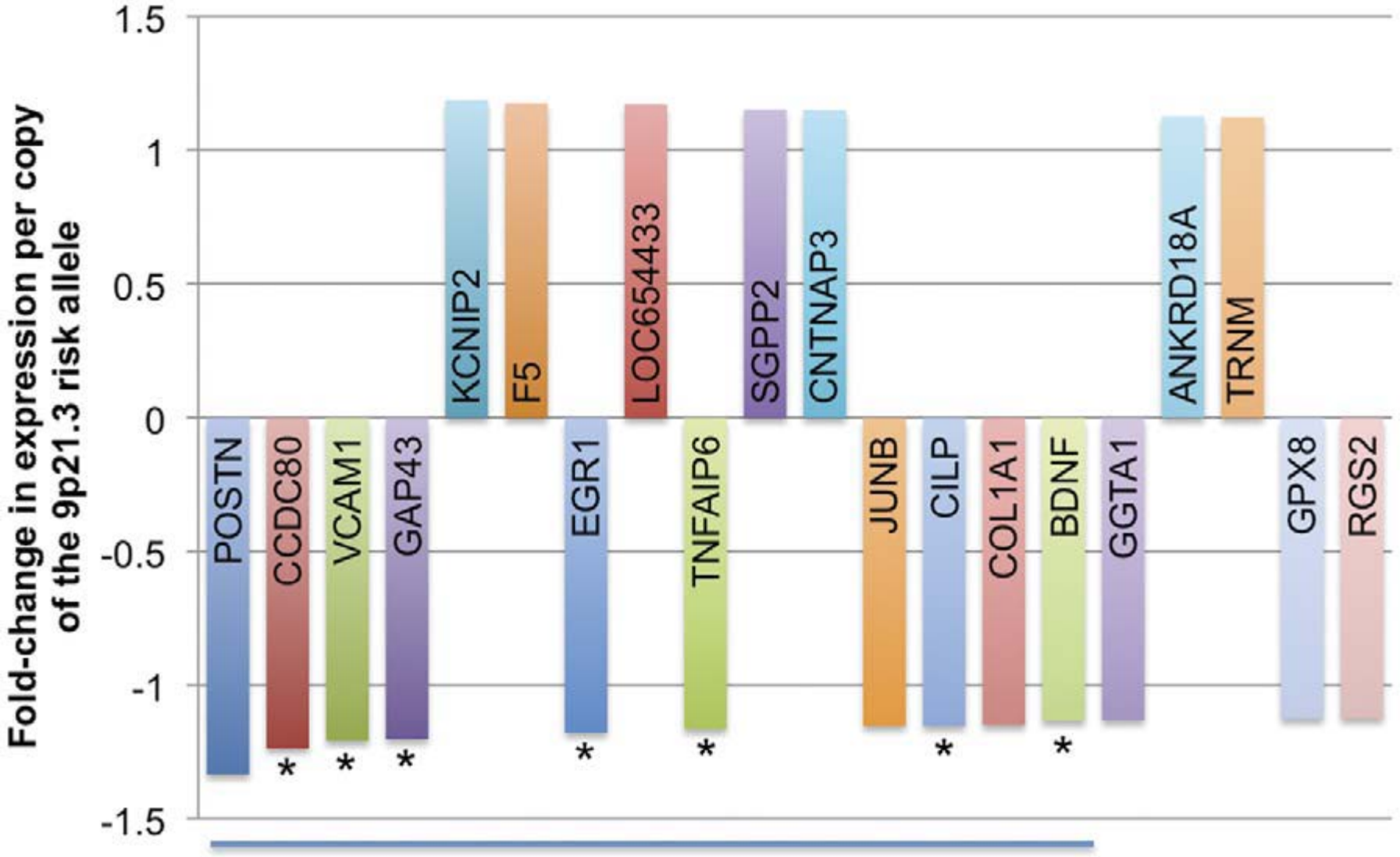
OPEN ACCESS Freely available online

PLoS one

The Chromosome 9p21.3 Coronary Heart Disease Risk Allele Is Associated with Altered Gene Expression in Normal Heart and Vascular Tissues

Anna P. Pilbrow[1]*, Lasse Folkersen[2], John F. Pearson[3], Chris M. Brown[4], Les McNoe[4], Nancy M. Wang, Wendy E. Sweet[5], W. H. Wilson Tang[5], Michael A. Black[4], Richard W. Troughton[1], A. Mark Richards[1], Anders Franco-Cereceda[6], Anders Gabrielsen[2], Per Eriksson[2], Christine S. Moravec[5], Vicky A. Cameron[1]

Gene Expression

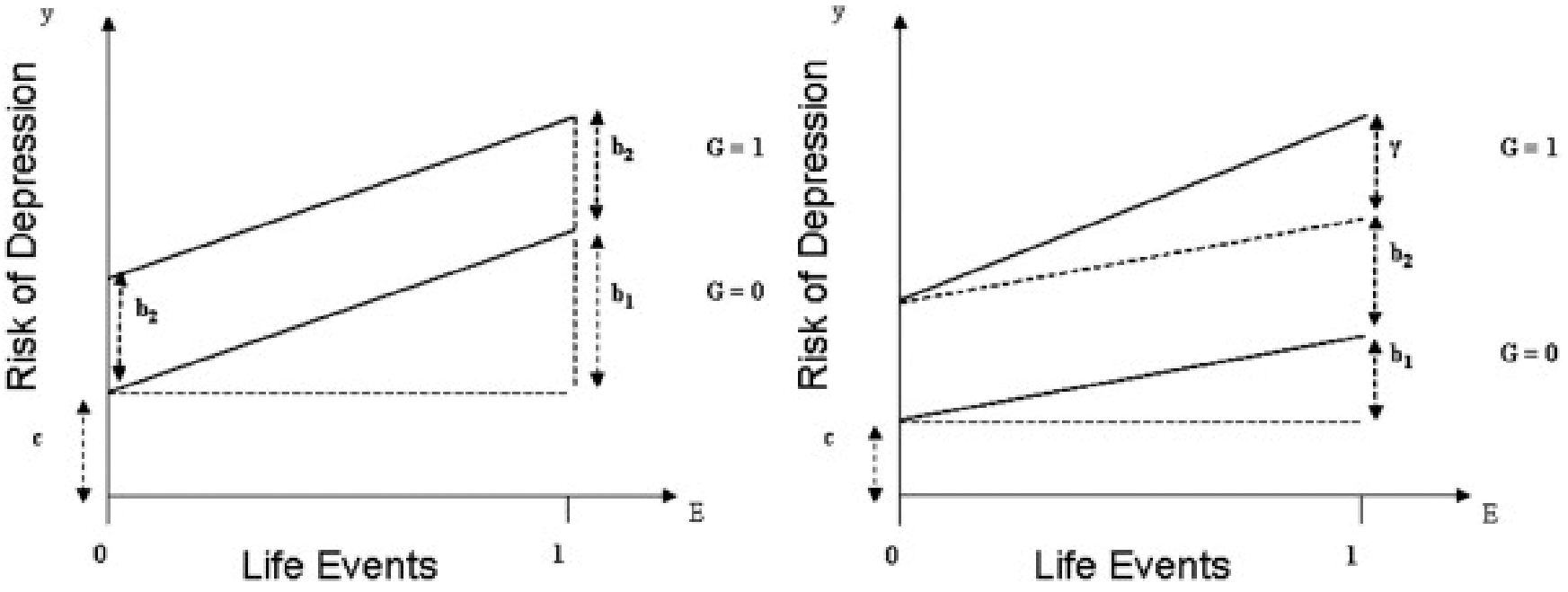


Gene x Environment

Biological Psychiatry, Volume 65, Issue 3, 1 February 2009, Pages 211–219

Gene × Environment Interactions at the Serotonin Transporter Locus

Marcus R. Munafò, Caroline Durrant, Glyn Lewis, and Jonathan Flint



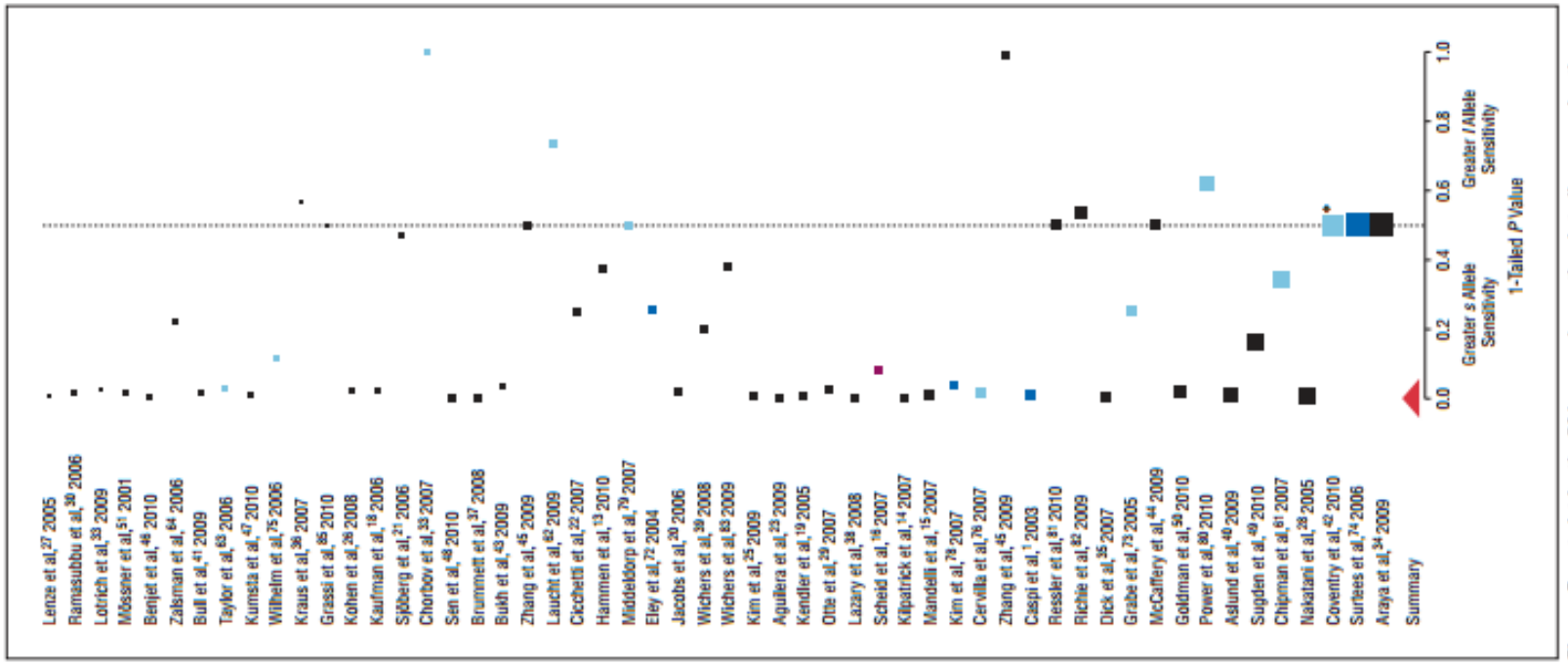
Gene x Environment

Arch Gen Psychiatry. 2011;68(5)

The Serotonin Transporter Promoter Variant (5-HTTLPR), Stress, and Depression Meta-analysis Revisited

Evidence of Genetic Moderation

Katja Karg, BSc; Margit Burmeister, PhD; Kerby Shedden, PhD; Srijan Sen, MD, PhD



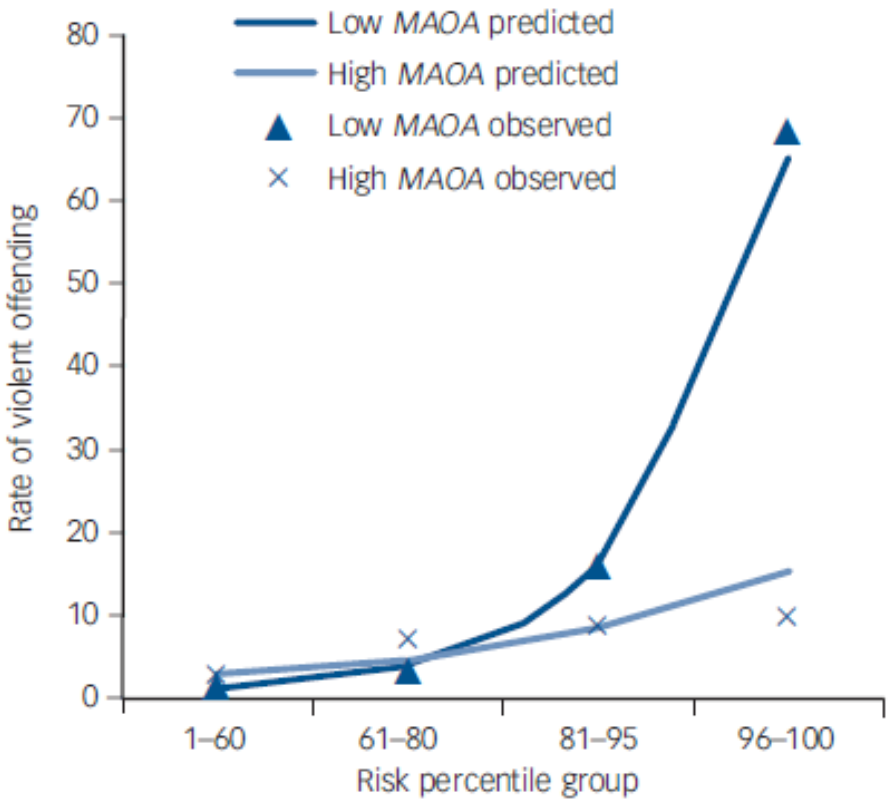
Gene x Environment: CHDS

The British Journal of Psychiatry (2012) 200, 116–123

Moderating role of the MAOA genotype in antisocial behaviour

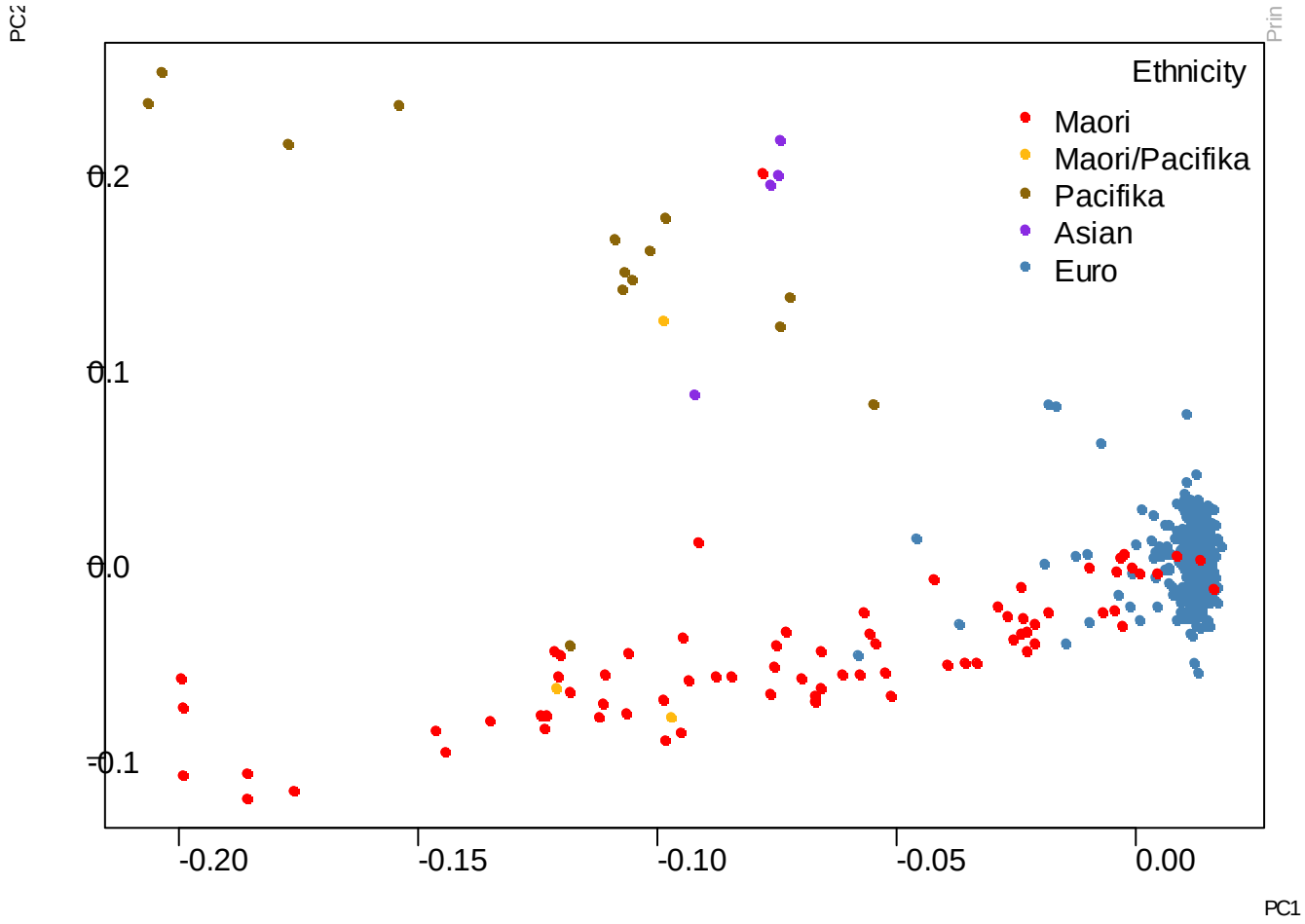
David M. Fergusson, Joseph M. Boden, L. John Horwood, Allison Miller and Martin A. Kennedy

399 Males
Genotyped with a VNTR
In the 5' promoter of MAOA
Gene on Chr X.



Gene x Environment: CHDS

739 participants genotyped for 657,000 SNPS
Illumina Human 660W Quad v1

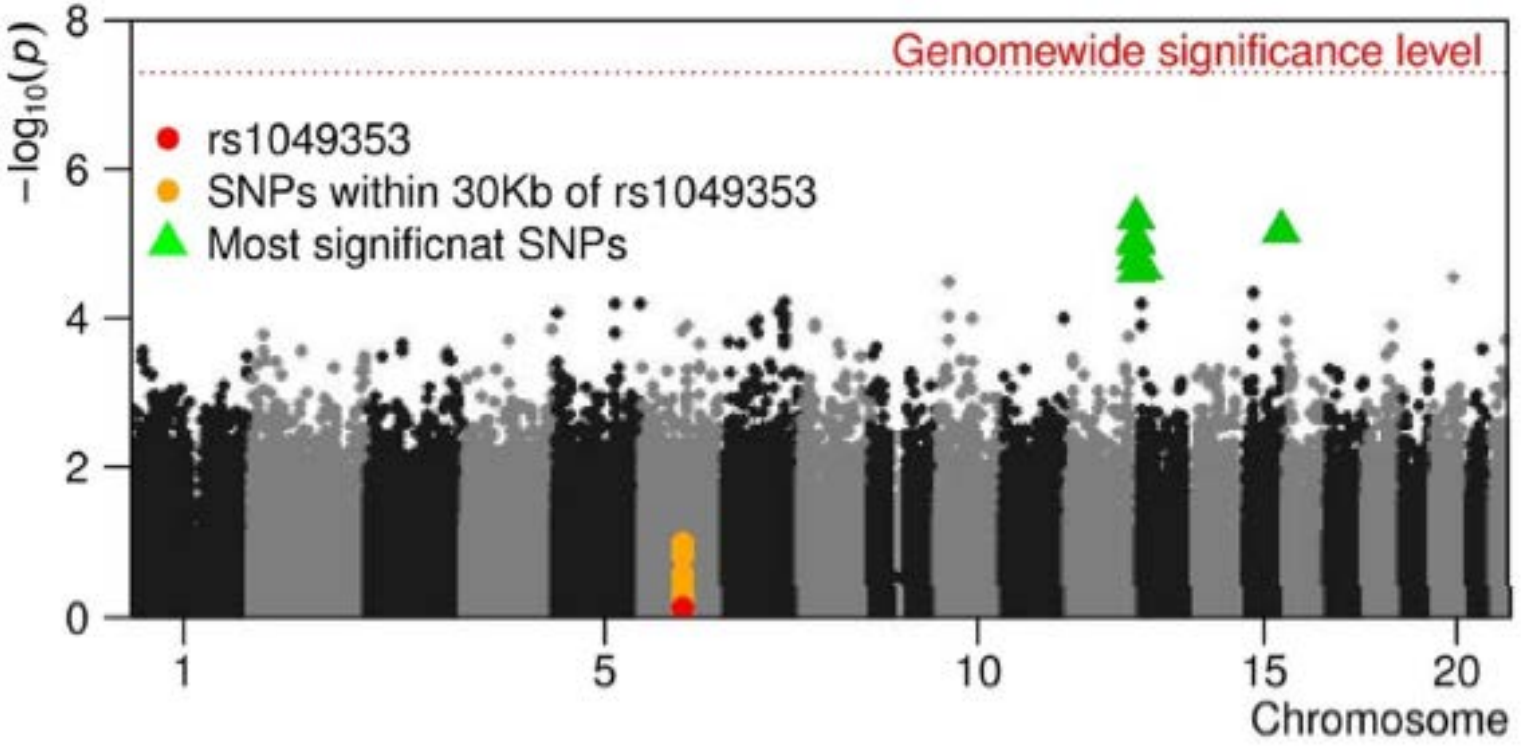


Gene x Environment: CHDS

Am J Med Genet Part B 162B:224–226.

Increased Risk of Major Depression by Childhood Abuse Is Not Modified by *CNR1* Genotype

John F. Pearson,[1*] David M. Fergusson,[2] L. John Horwood,[2] Allison L. Miller,[3] Patrick F. Sullivan,[4] Liu E. Youfang,[4] and Martin A. Kennedy[3]

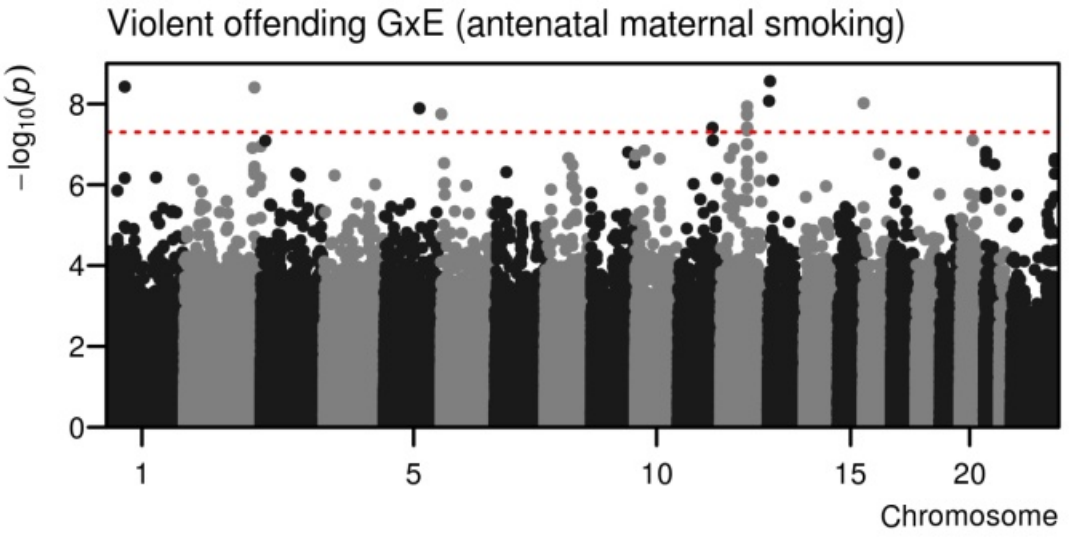


Gene x Environment: CHDS

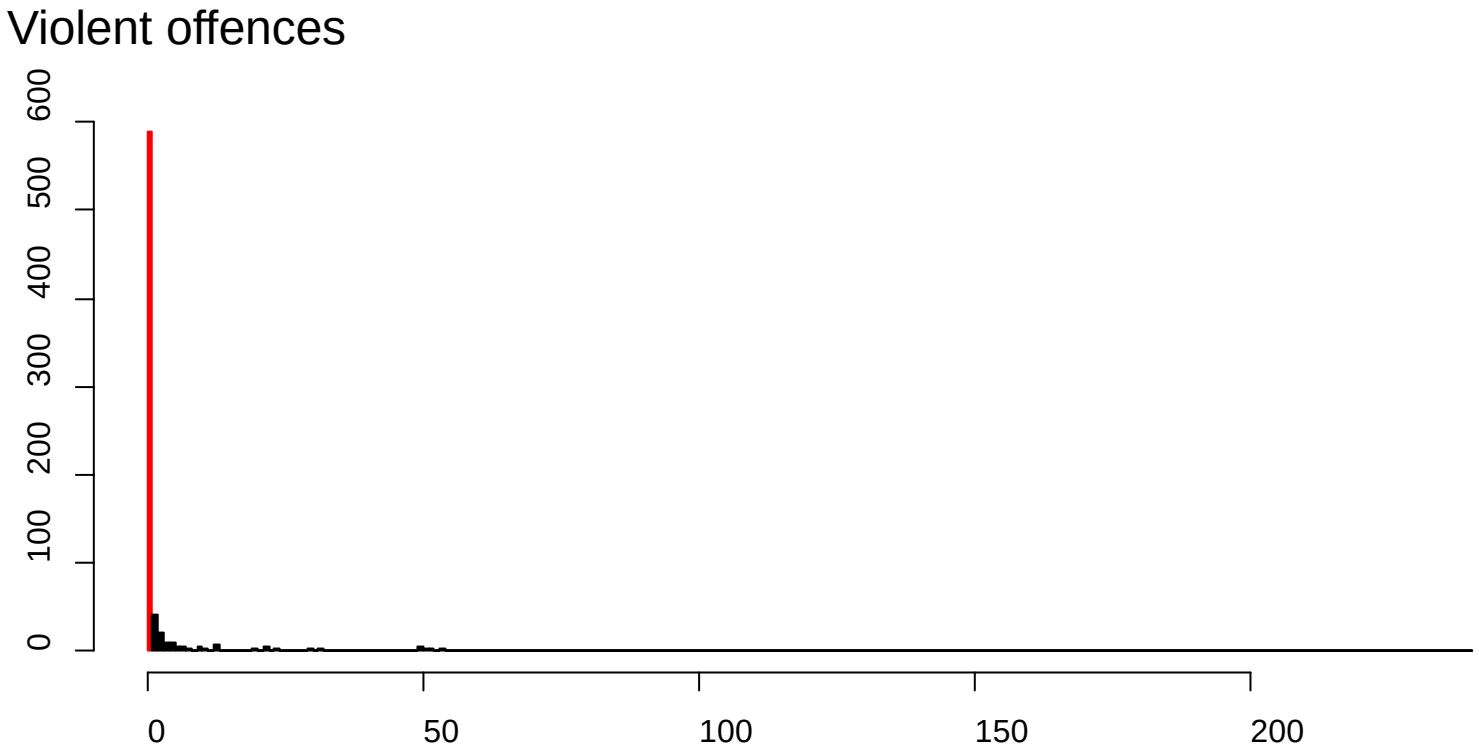
What's next?

- MAOA modifies the effects of early life stressors, which other genes do?
- Can we generate a score combining susceptibility loci?

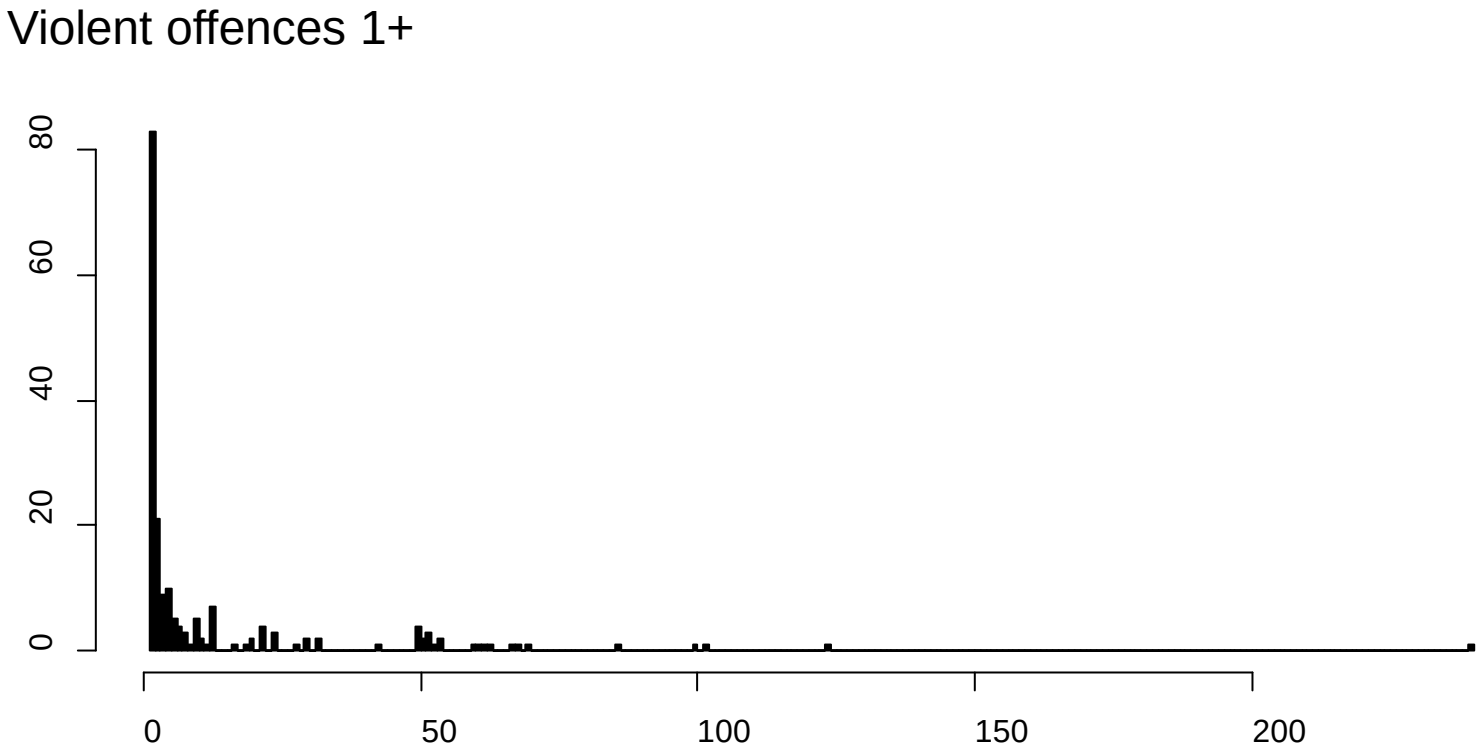
The data is difficult:
Fit with negative binomial or over-dispersed Poisson models



Count data



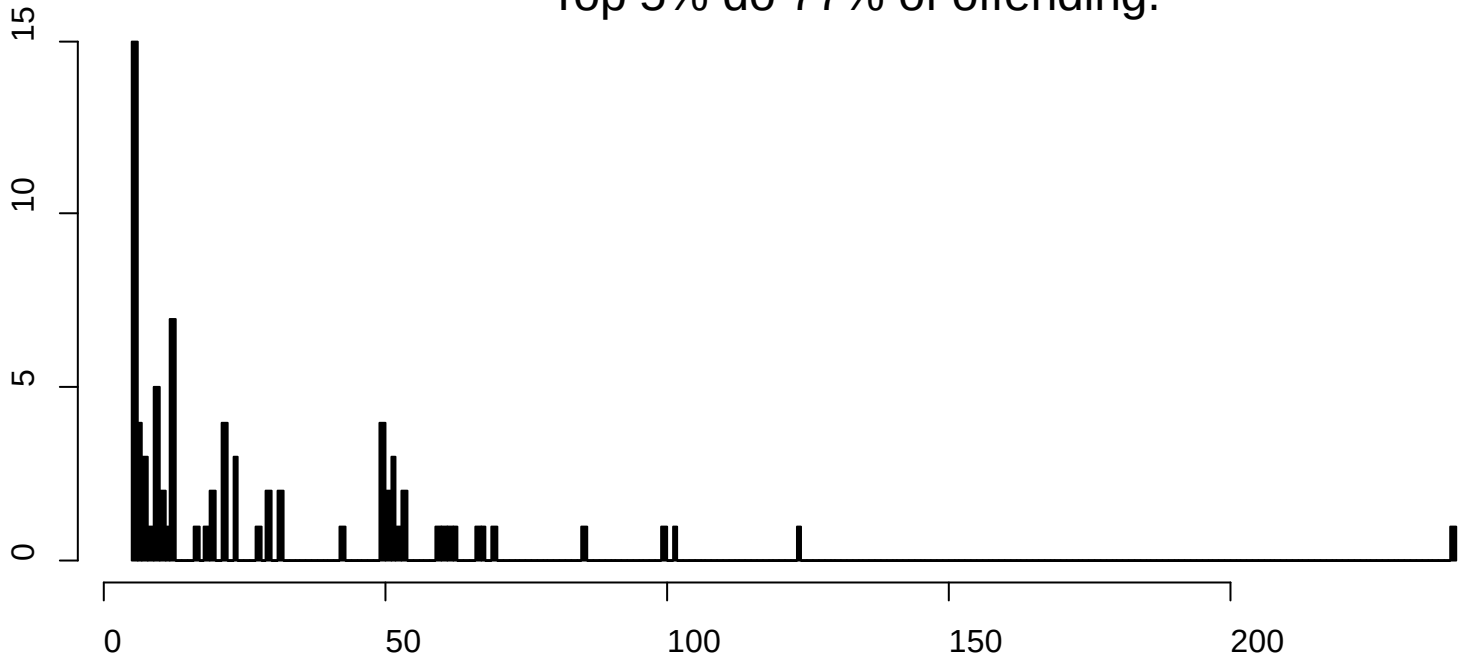
Count data



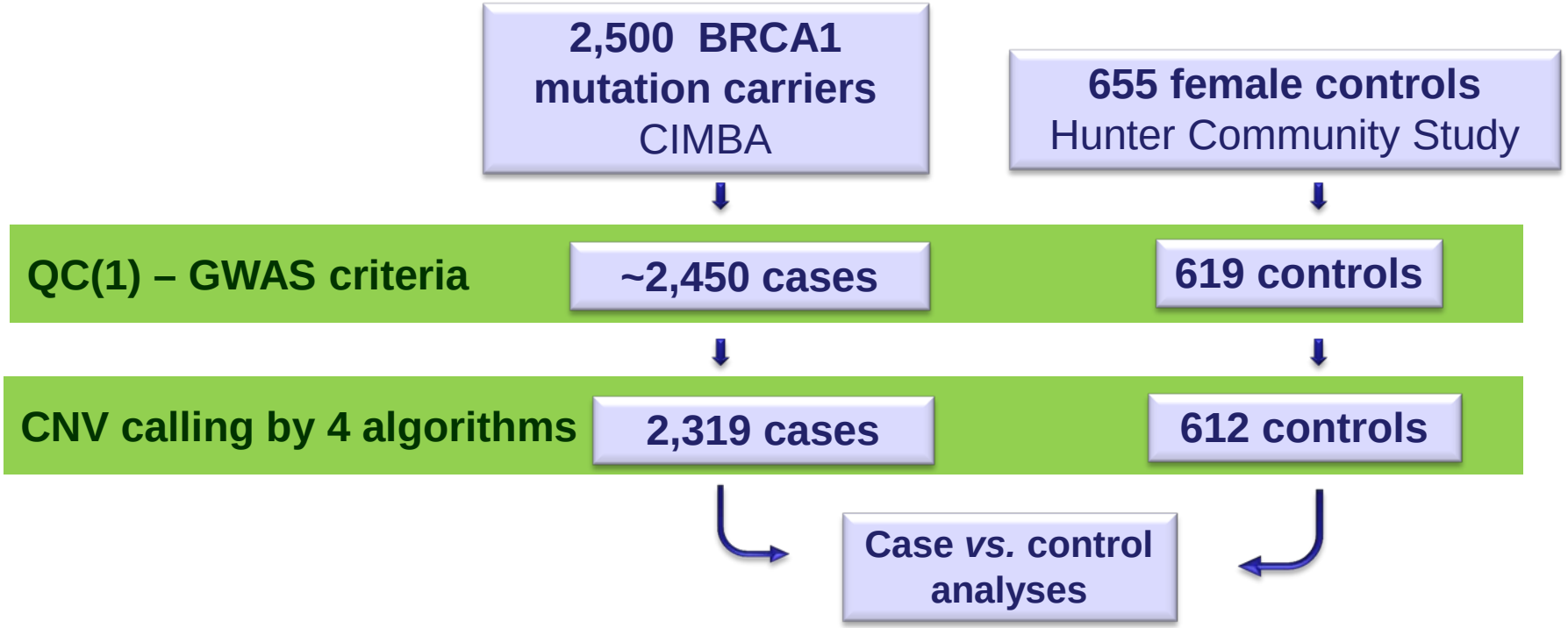
Count data

Violent offences 5+

Top 10% do 91% of offending.
Top 5% do 77% of offending.



Copy Number Variations



Data is censored, but not at random: ovarian cancer preceding breast cancer were Considered controls for breast cancer. Test a score statistic based on the likelihood of the observed genotype conditional on the disease phenotype.

Barnes et al. (2012) *Evaluation of association methods for analysing modifiers of disease risk in carriers of high-risk mutations.* Genetic Epidemiology 36: 274-291

Take home

- Reviews are not always private
- Heritability does not imply determinism
- Initial large effects reduce on replication
- GWAS need more than a big n.

Acknowledgements

- Anna Pilbrow
- Vicky Cameron

- Logan Walker
- Mackenzie Cancer Research Group
- Amanda Spurdle
- Louise Marquart

- David Fergusson
- John Horwood
- Joe Boden
- Martin Kennedy

- Tim Eglinton
- Rebecca Roberts

… Apologies if your name isn’t here but should be.