The genetics of complex traits
Amazing progress (much by ppl in this room)

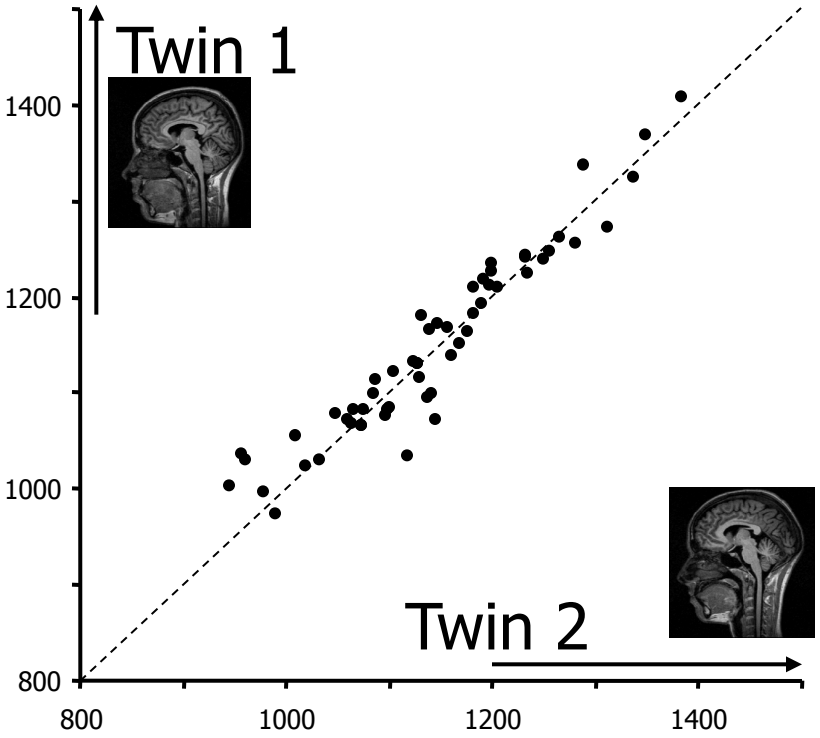
Nick Martin
Queensland Institute
of Medical Research
Brisbane

Boulder workshop
March 11, 2016

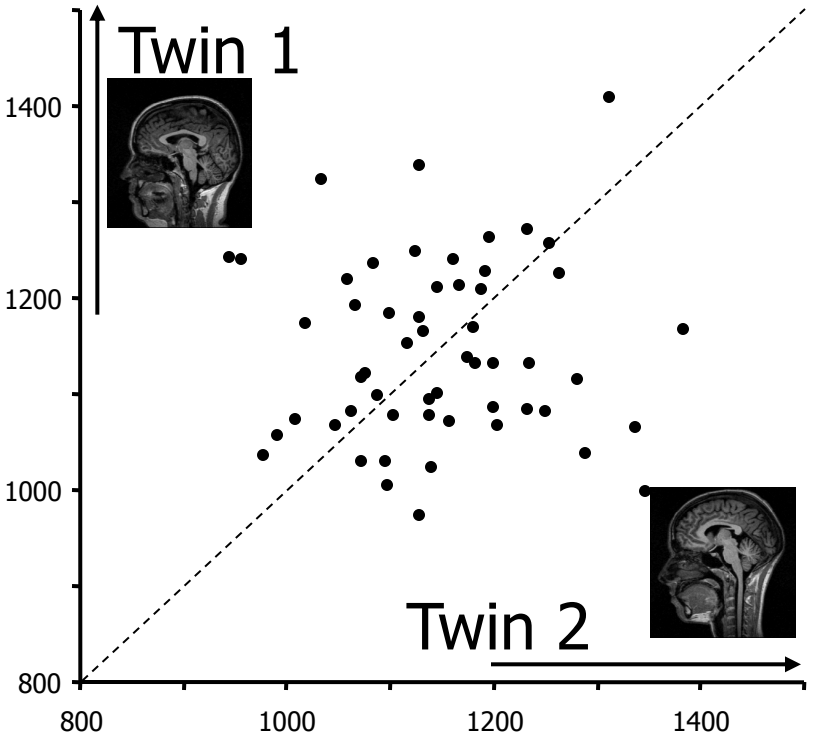
Genetic Epidemiology: Stages of Genetic Mapping

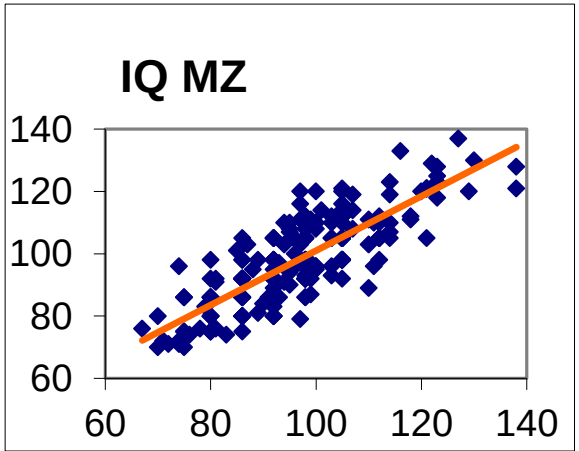
- Are there genes influencing this trait?
  - Genetic epidemiological (twin / family) studies OR heritability based on measured genetic variants
- What are those genes?
  - Association analysis (meta-analysis / pathway)
- How do they work beyond the sequence?
  - Epigenetics, transcriptomics, proteomics
- What can we do with them ?
  - Translational medicine

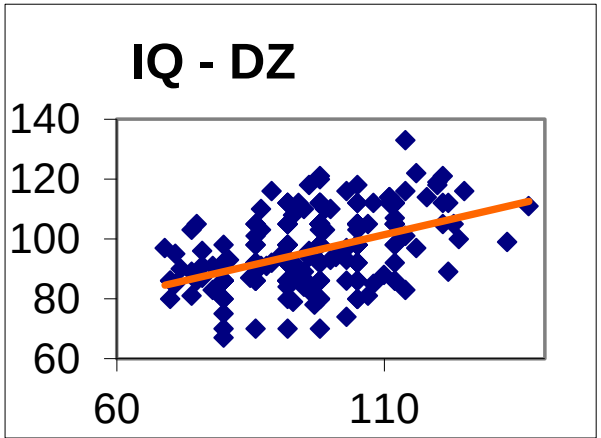
Brain volume (ml) in MZ twin pairs



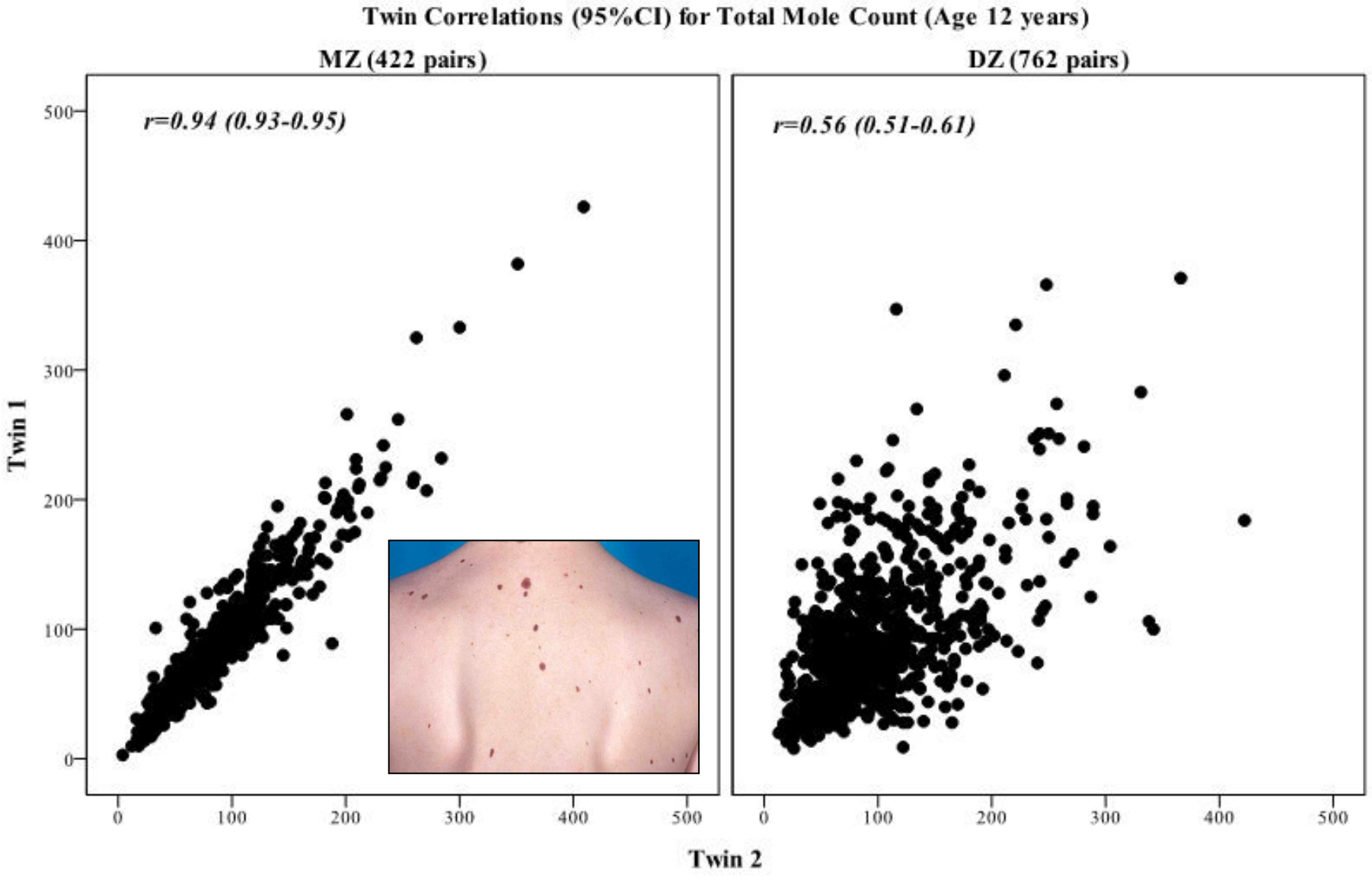
Brain volume in DZ twin pairs





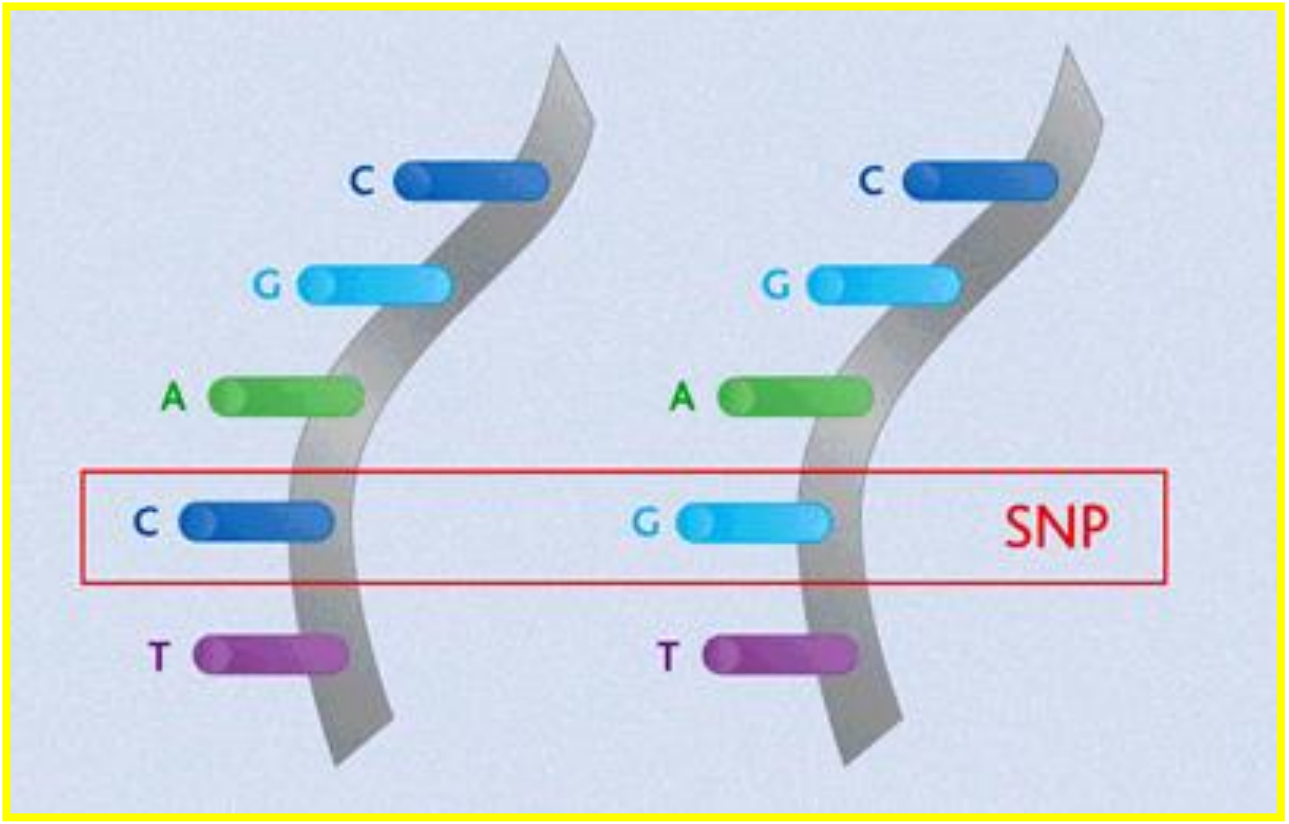


Twin correlations for total mole count

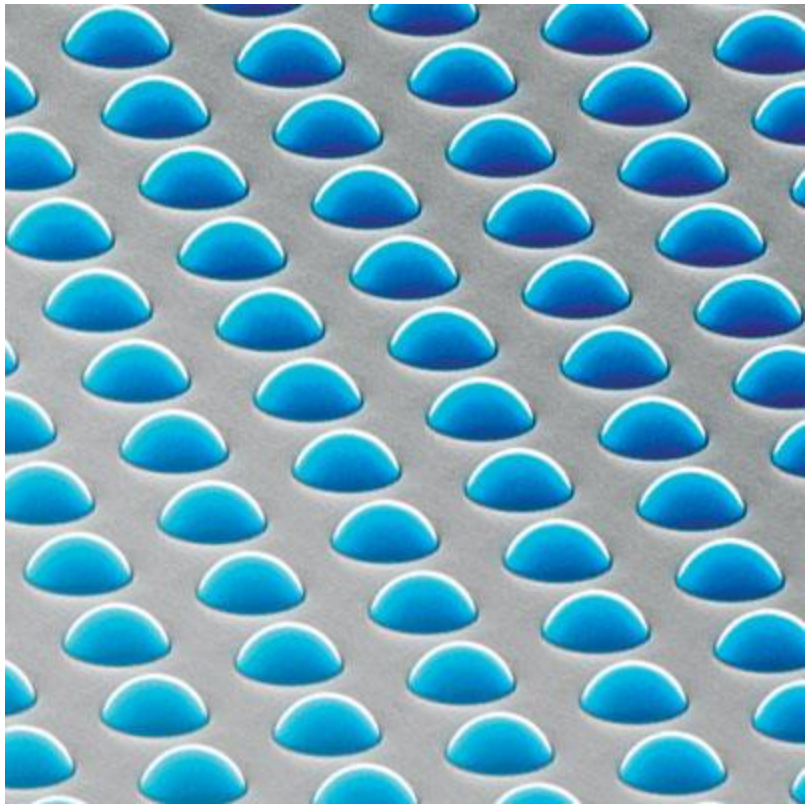


Finding the genes - association

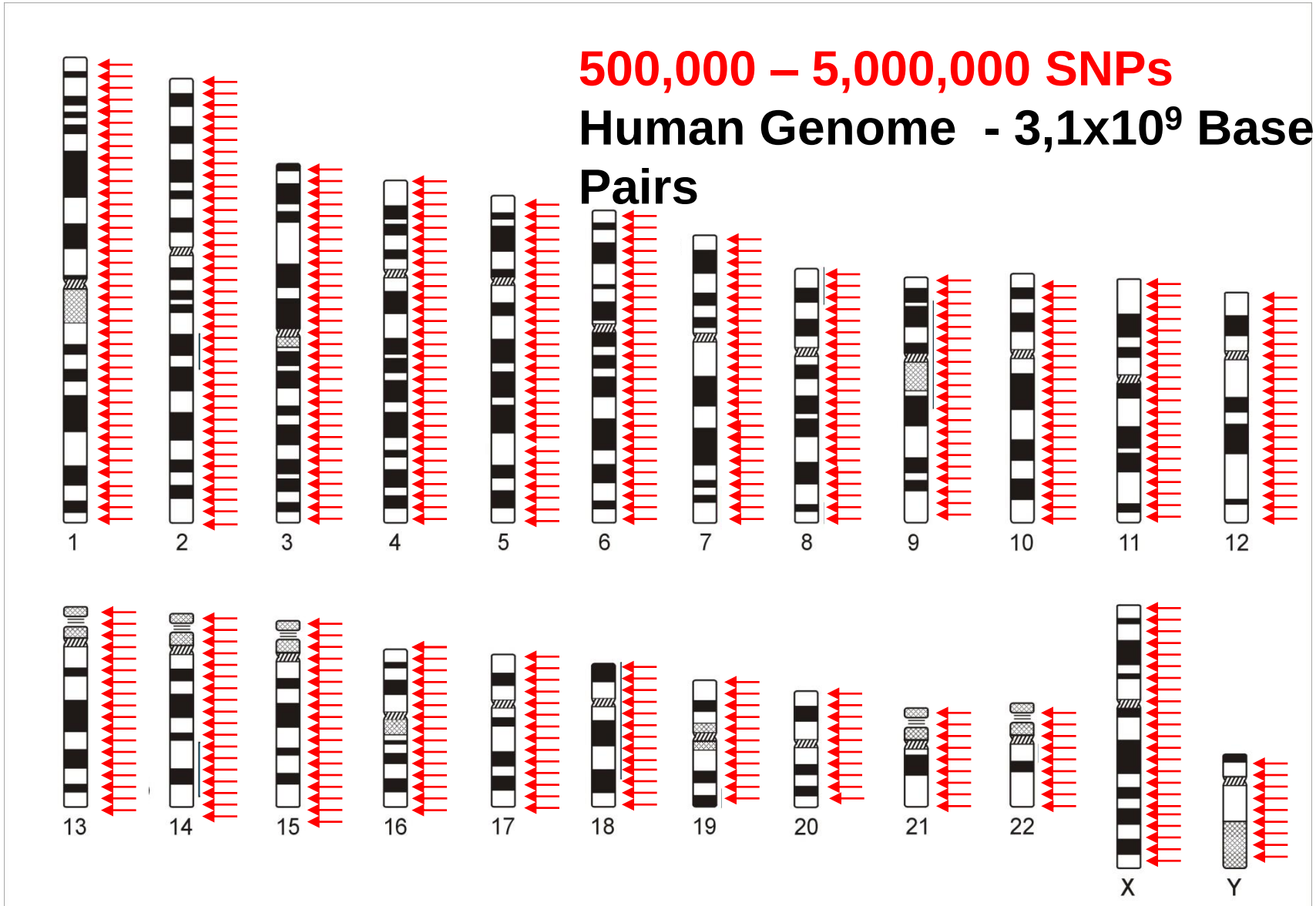
Looks for correlation between specific alleles and phenotype (trait value, disease risk) using single nucleotide polymorphisms (SNPs)



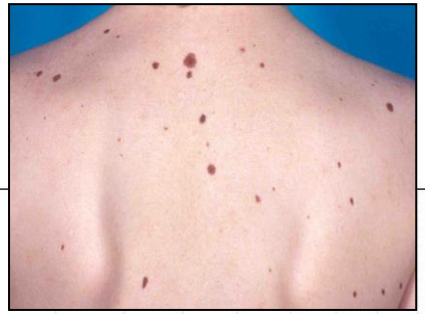
High density SNP arrays – up to 1 million SNPs

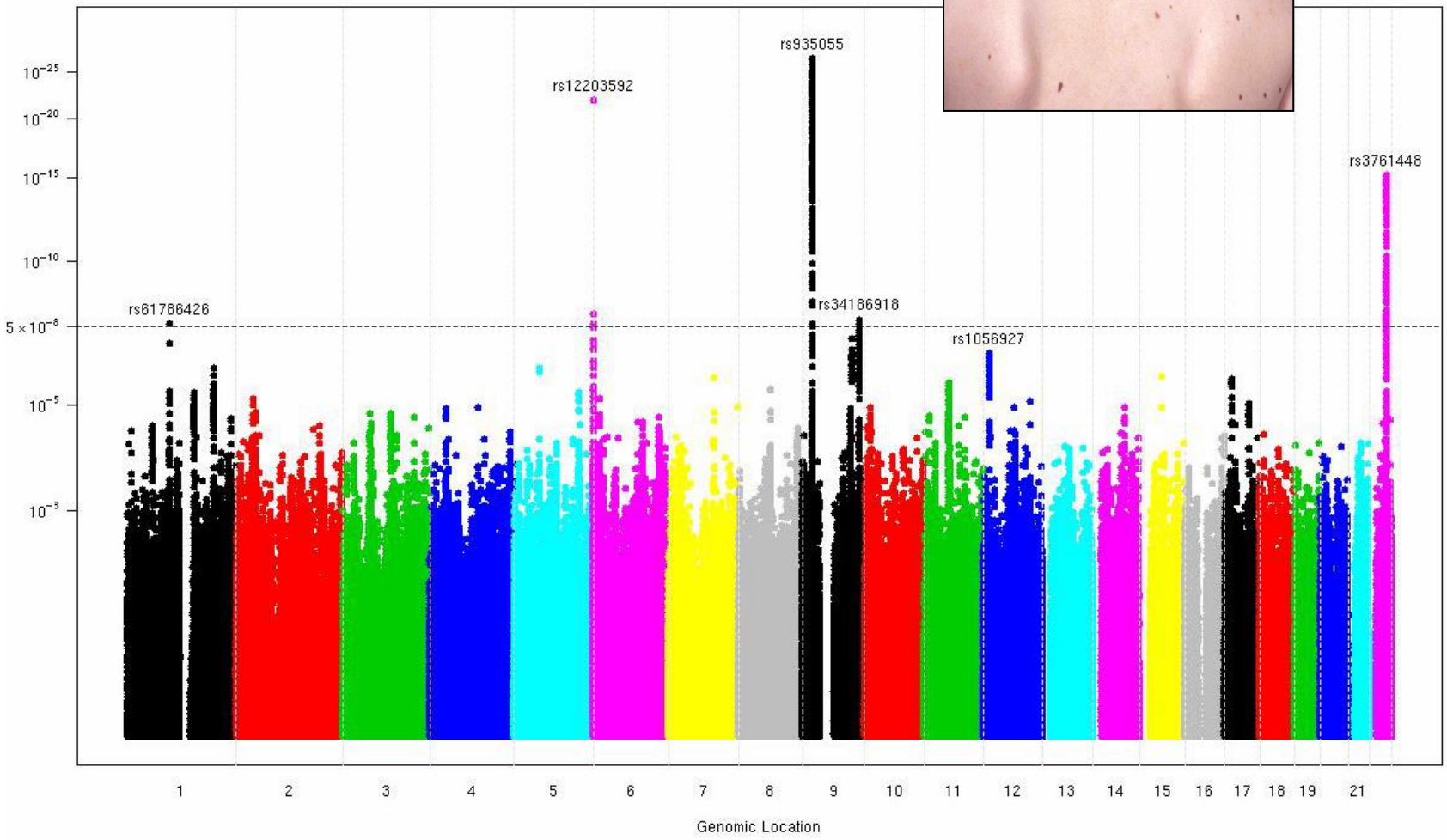


Genome-Wide Association Studies



GWAS for moliness(11 studies, 23,000 subjects, 5 hits)

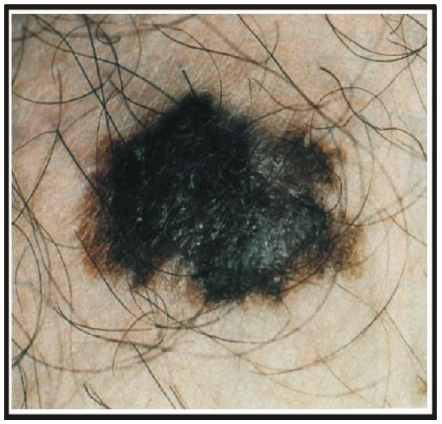


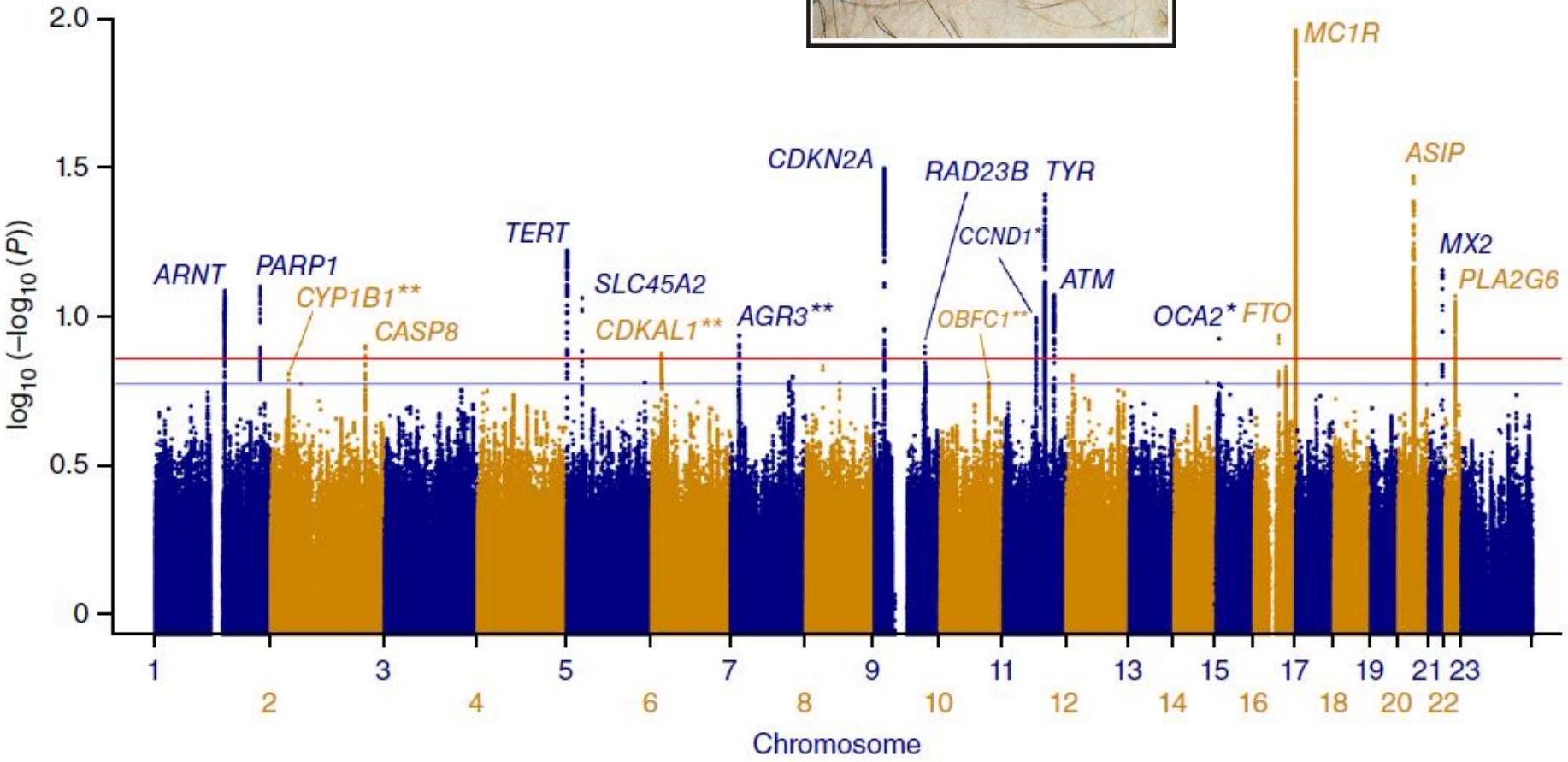


David Duffy, Gu Zhu

Melanoma

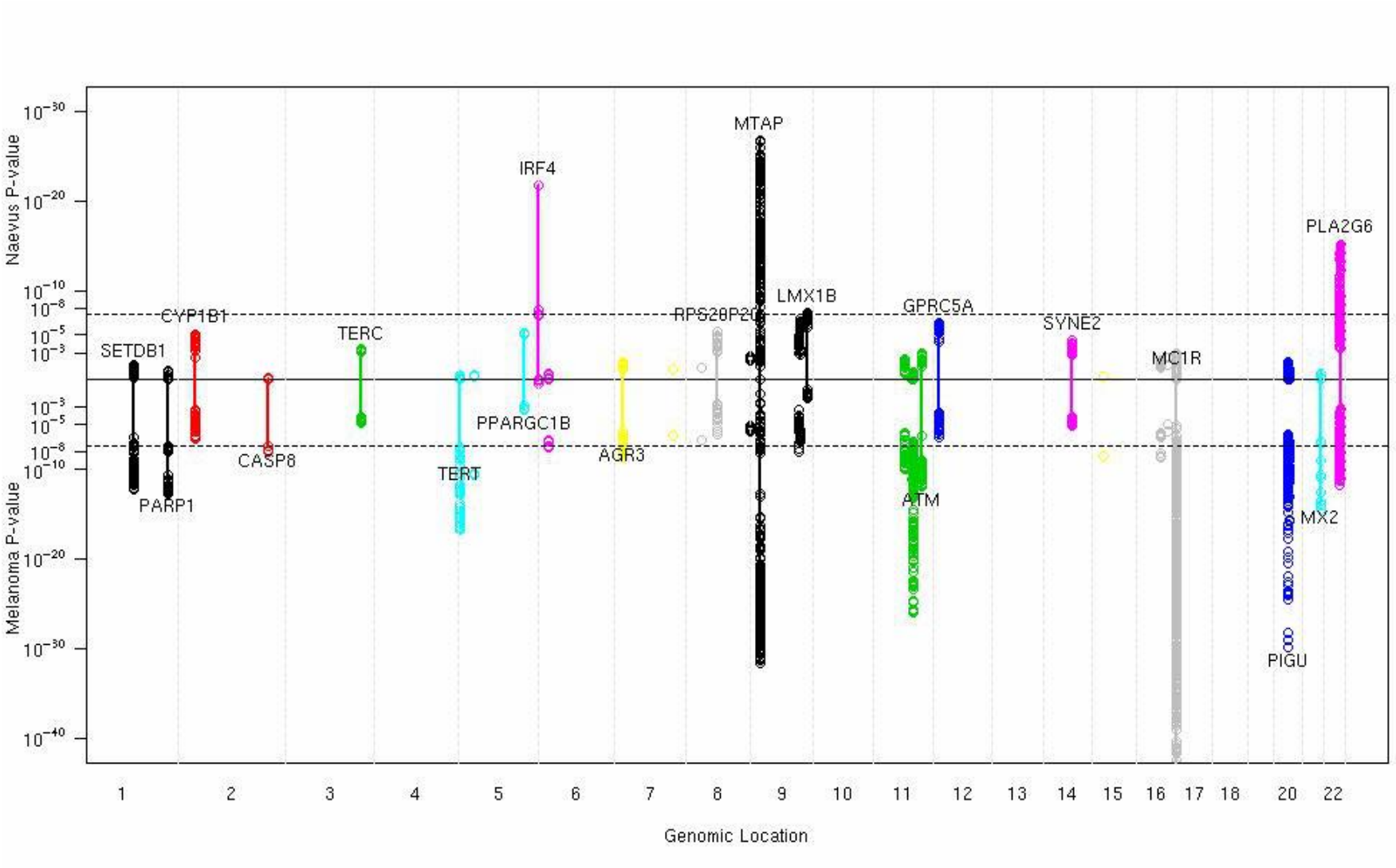
11 studies,
15,990 cases, 26,409 controls
20 loci (5 new)





Matthew Law

GWAS meta-analysis for moliness + melanoma

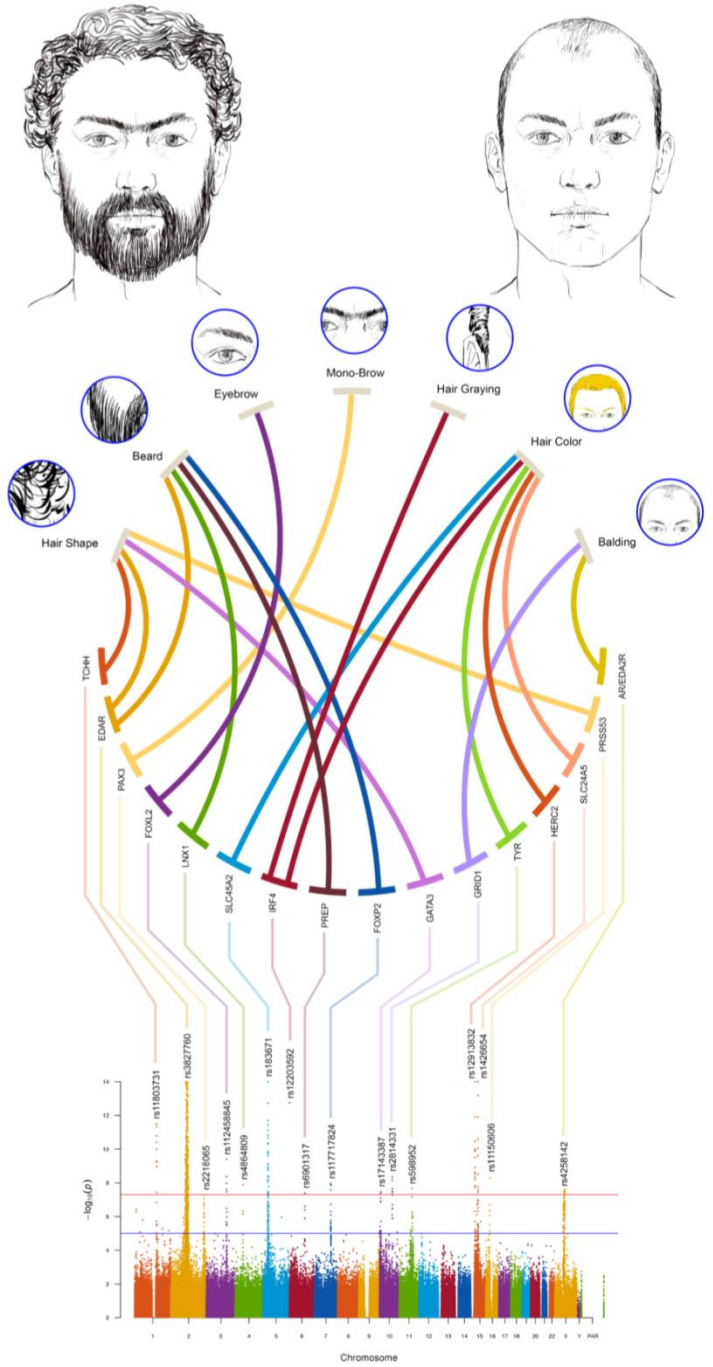


David Duffy

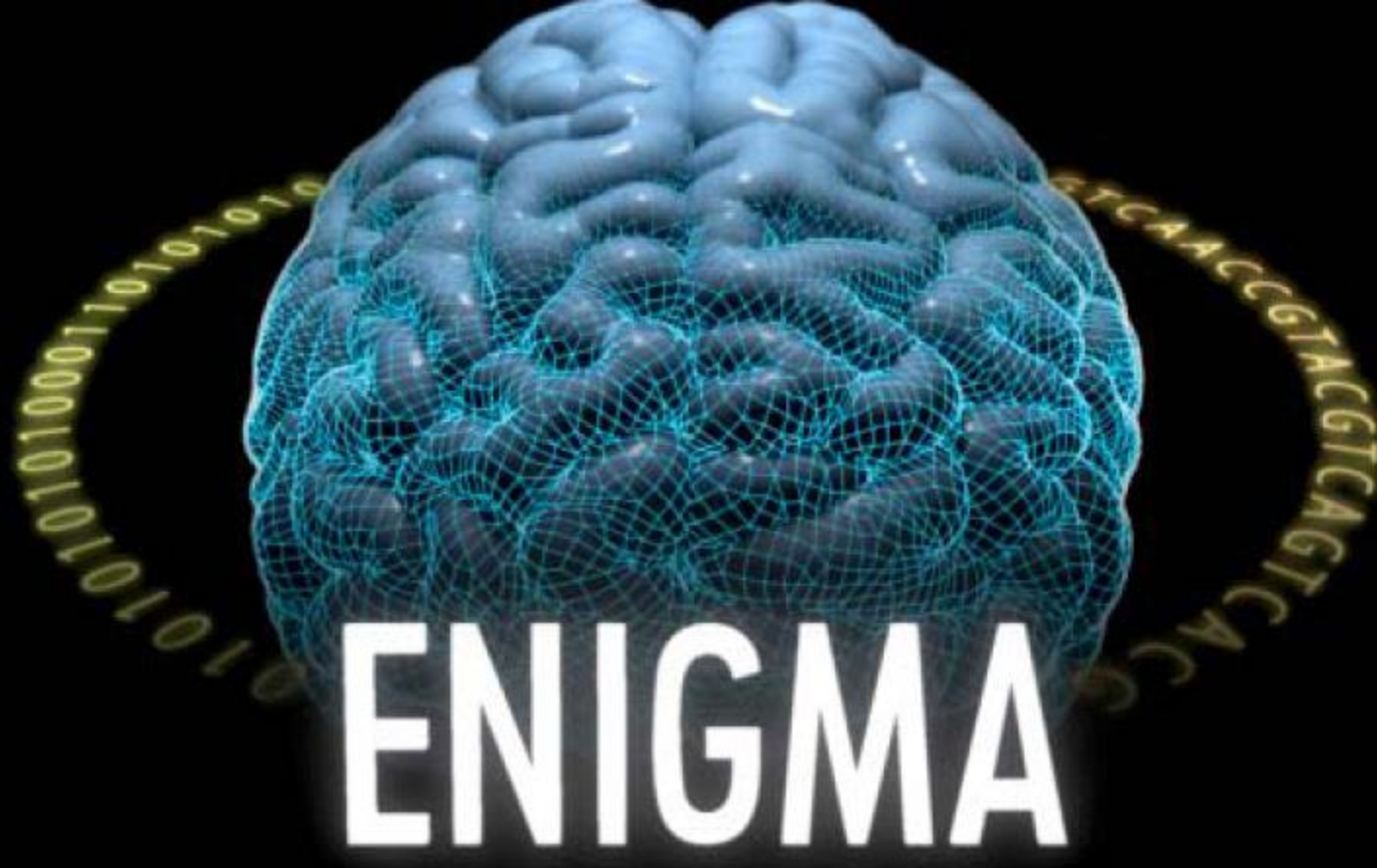
THE GENETIC BASIS OF VARIATION IN FACIAL AND SCALP HAIR: A GENOME-WIDE ASSOCIATION STUDY IN ADMIXED LATIN AMERICANS

Kaustubh Adhikari…Carla Gallo… Andrés RuizLinares

- GWAS in **> 6,000 Latin Americans** for features of scalp hair (shape, color, graying, baldness) and facial hair (beard thickness, mono brow, eye-brow thickness).
- GW significant association at 14 genomic regions with at least one of the hair features examined (p-values 5×10^{-8} to 3×10^{-119}).
- Two traits (scalp hair shape and beard thickness) are associated with SNPs in the ectodysplasin A receptor gene region, a key regulator of embryonic skin appendage development.
- The other associated regions include novel loci for scalp hair shape and baldness, and the **first reported loci for hair graying, mono-brow, and eye-brow and beard thickness**.
- The novel variant associated with scalp hair shape is a Q30R substitution in the Protease Serine S1 family member 53 (PRSS53). We demonstrate that this enzyme is highly expressed in the hair follicle, especially the inner root sheath, and that the Q30R substitution affects enzyme processing and secretion.
- **The genome regions associated with hair features are enriched for signals of selection, consistent with proposals regarding the evolution of human hair.**



Nat Commun. 2016 Mar;7:10815.

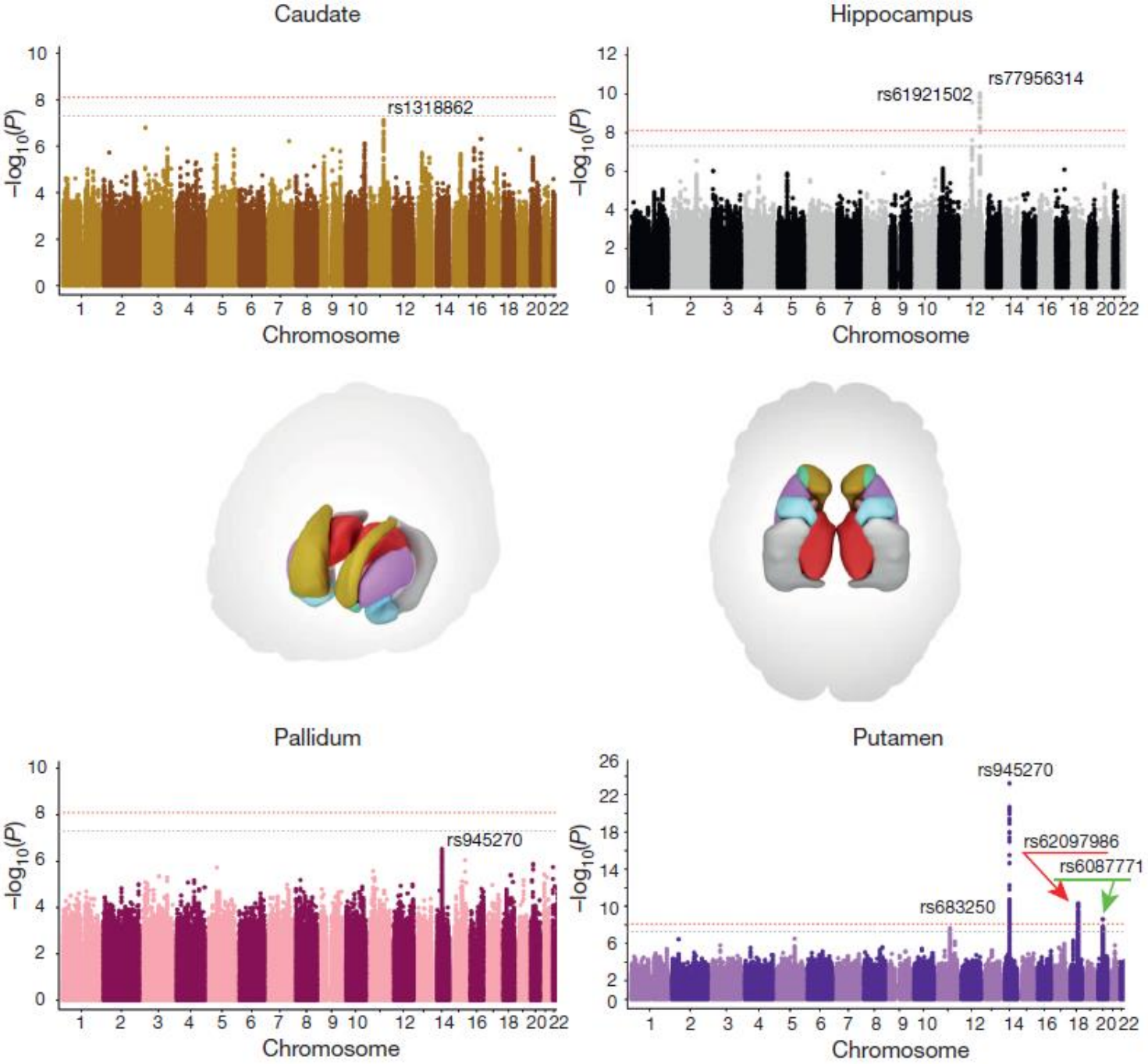


Enhancing Neuro Imaging Genetics Through Meta-Analys

The ENIGMA Network brings together researchers in imaging genomics, to understand brain structure and function, based on MRI, DTI, fMRI and genomewide association scan (GWAS) data. The ENIGMA Network has several goals:

About ENIGMA

http://enigma.loni.ucla.edu/

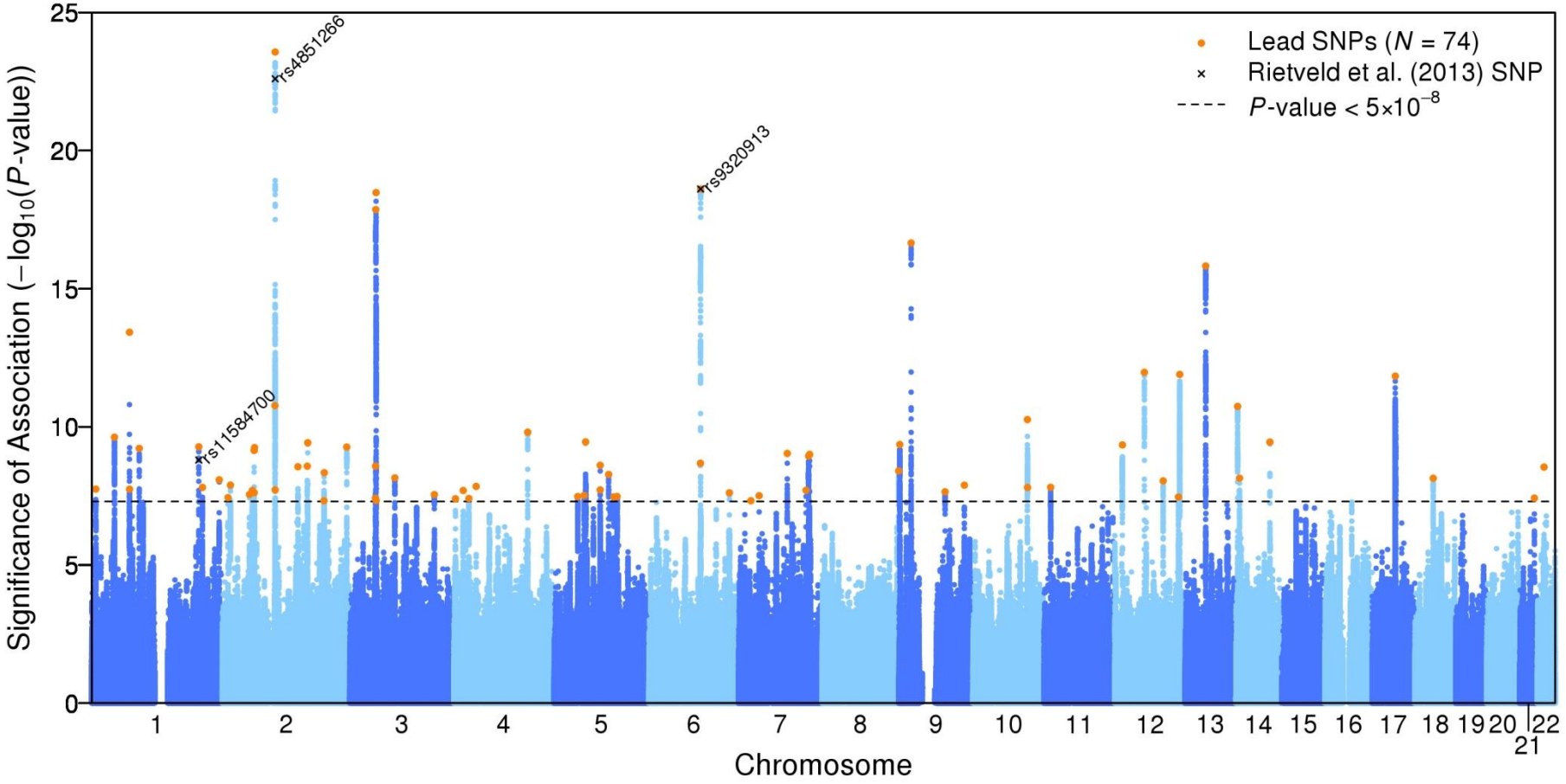

Caudate
rs1318862
−log10(P)
Chromosome
Hippocampus
rs61921502
rs77956314
−log10(P)
Chromosome
Pallidum
rs945270
−log10(P)
Chromosome
Putamen
rs945270
rs62097986
rs6087771
rs683250
−log10(P)
Chromosome

GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment

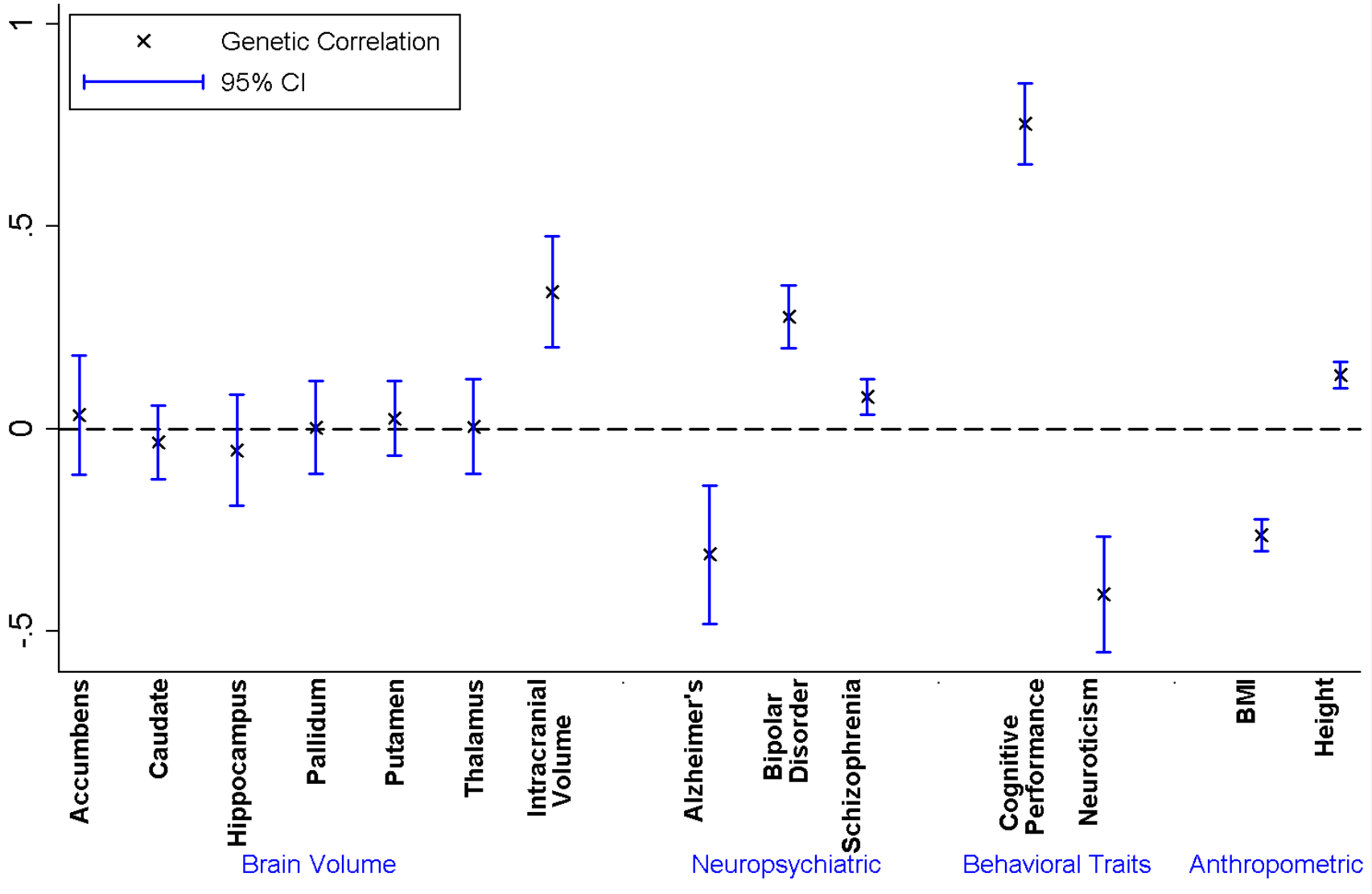
All authors with their affiliations appear at the end of this paper.

A genome-wide association study (GWAS) of educational attainment was conducted in a discovery sample of 101,069 individuals and a replication sample of 25,490. Three independent single-nucleotide polymorphisms (SNPs) are genome-wide significant (rs9320913, rs11584700, rs4851266), and all three replicate. Estimated effects sizes are small (coefficient of determination $R^2 \approx 0.02\%$), approximately 1 month of schooling per allele. A linear polygenic score from all measured SNPs accounts for ≈2% of the variance in both educational attainment and cognitive function. Genes in the region of the loci have previously been

EduYears (*N* = 293,723)



SNPs in orange are the approximately 74 independent genome-wide significant associations ("lead SNPs"). The black x's labeled with rs numbers are the 3 Rietveld et al.[14] SNPs.



Genetic overlap between *EduYears* and other traits.
Results from Linkage-Disequilibrium (LD) Score regressions: estimates of genetic correlation with brain volume, neuropsychiatric, behavioral, and anthropometric traits for which GWAS summary statistics were available in the public domain.

Personality Genomics Consortium

Meta-analysis for IRT-based Neuroticism and Extraversion: 28 cohorts

Marleen de Moor

Stephanie van den Berg

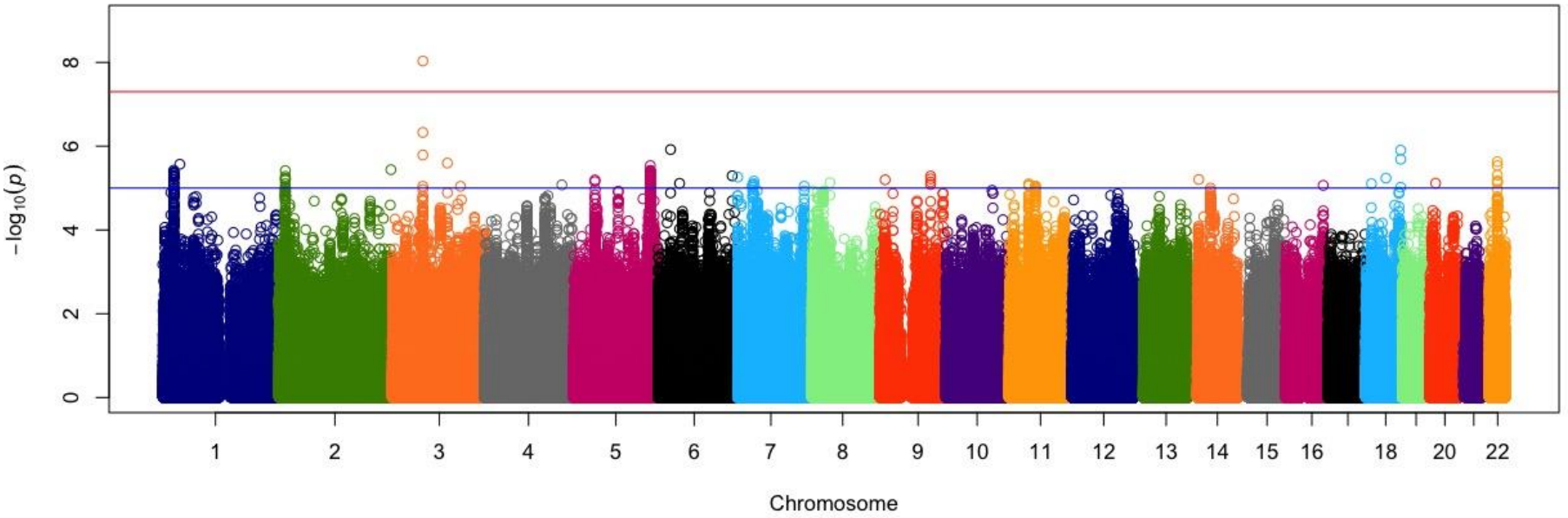
Behav Genet (2014) 44:295–313
DOI 10.1007/s10519-014-9654-x

ORIGINAL RESEARCH

Harmonization of Neuroticism and Extraversion phenotypes across inventories and cohorts in the Genetics of Personality Consortium: an application of Item Response Theory

Stéphanie M. van den Berg · Marleen H. M. de Moor · Matt McGue · Erik Pettersson · Antonio Terracciano · Karin J. H. Verweij · Najaf Amin · Jaime Derringer · Tõnu Esko · Gerard van Grootheest · Narelle K. Hansell · Jennifer Huffman · Bettina Konte · Jari Lahti ·

Neuroticism: N ~ 63,000
Meta-analysis with 1000G imputation for all 28 studies

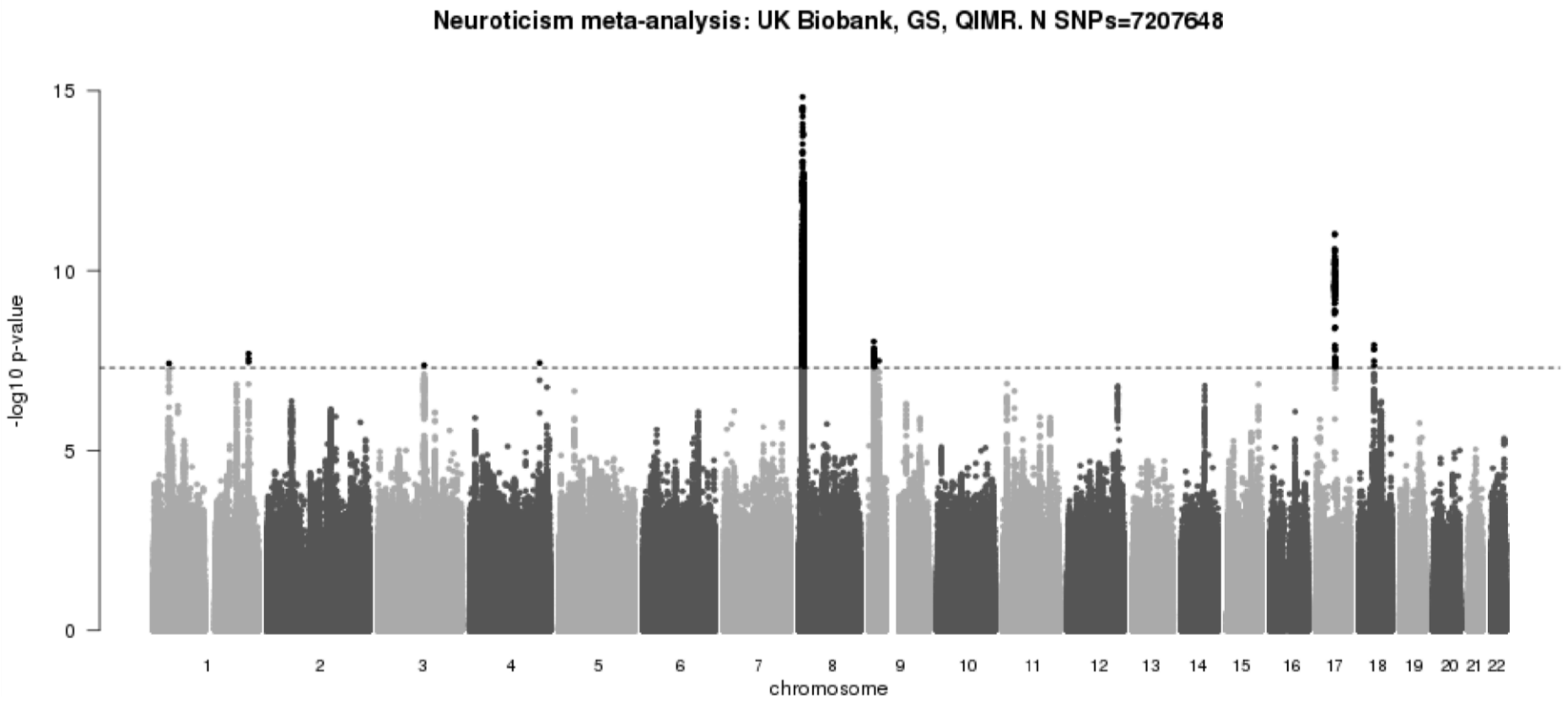


A genome-wide significant SNP was found on 3p14 in the *MAGI1* gene (rs35855737; P=9.26 x 10^{-9}).

JAMA Psychiatry. 2015 Jul;72(7):642-50.

Neuroticism: UK Biobank, n > 106,000 identifies 9 neuroticism-associated loci

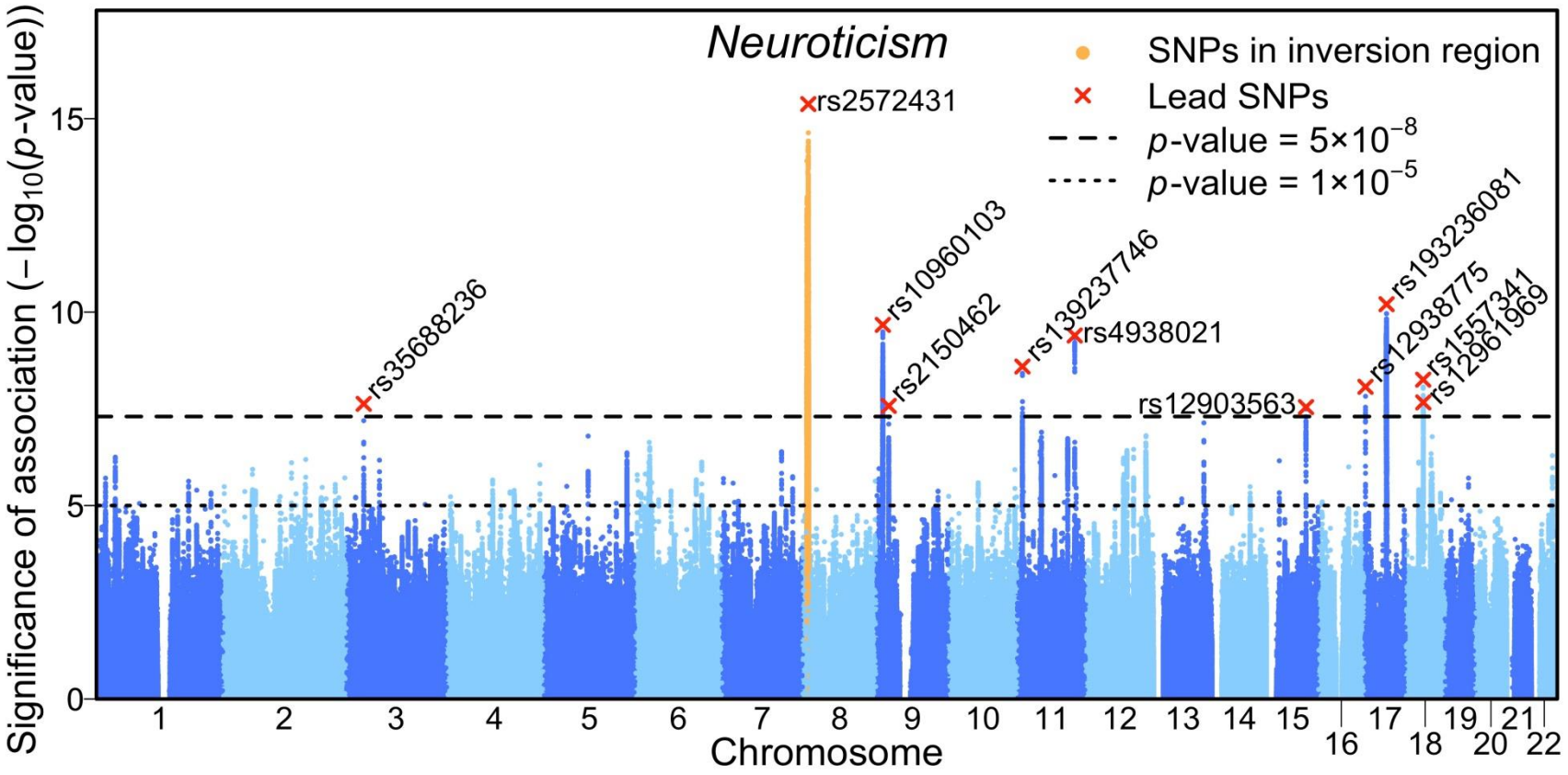
Daniel J Smith, Valentina Escott-Price, Gail Davies, Mark Bailey, **Lucía Colodro-Conde**, …, Ian Deary, Michael O'Donovan



Molecular Psychiatry, accepted 3/3/16

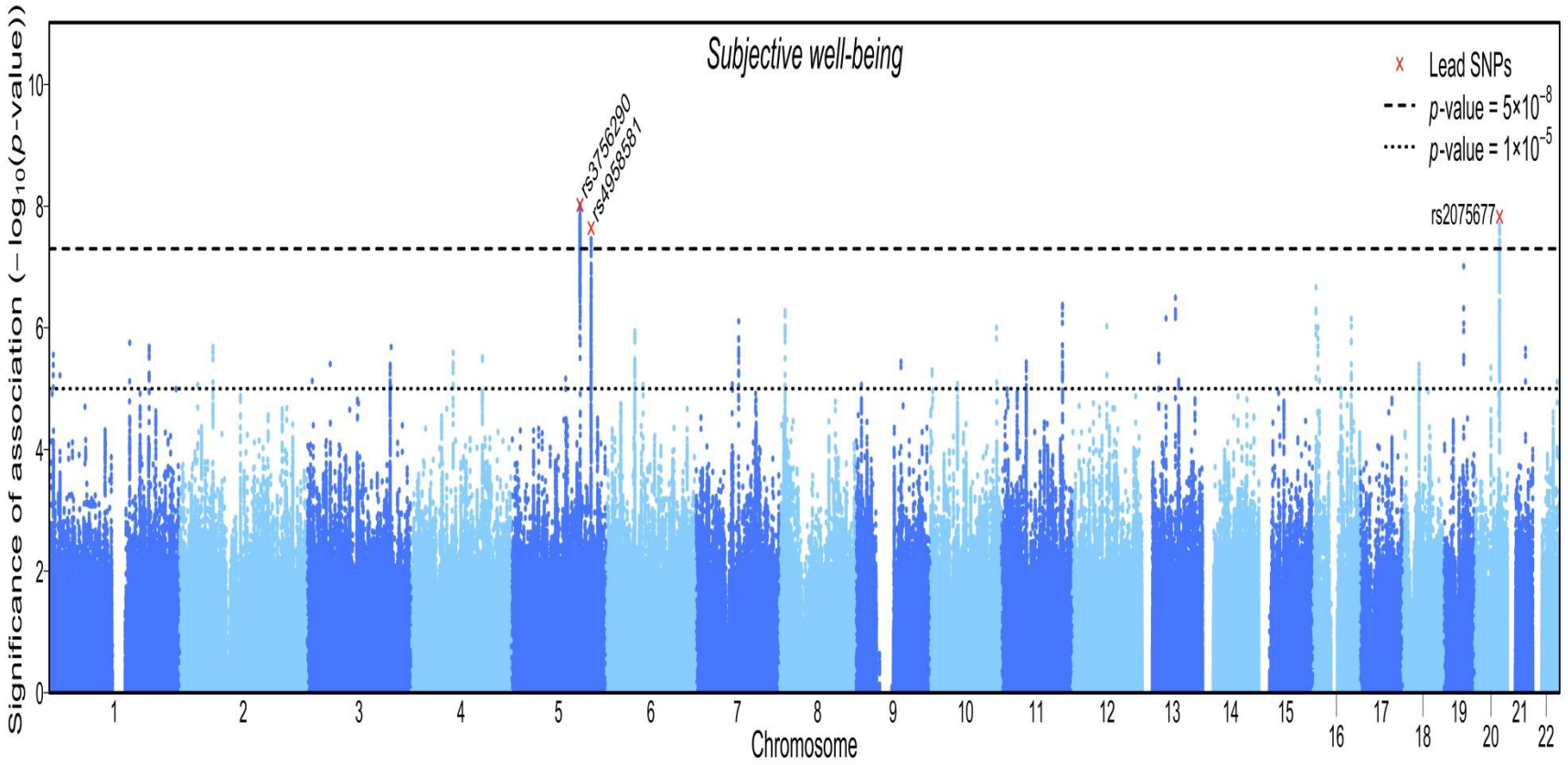
Neuroticism: UK Biobank + Personality Genomics Consortium m-a (N = 170,910) identifies 11 loci (2 more than UKB alone)

ALARM! power of meta-analysis versus a single large uniformly phenotyped and genotyped resource?

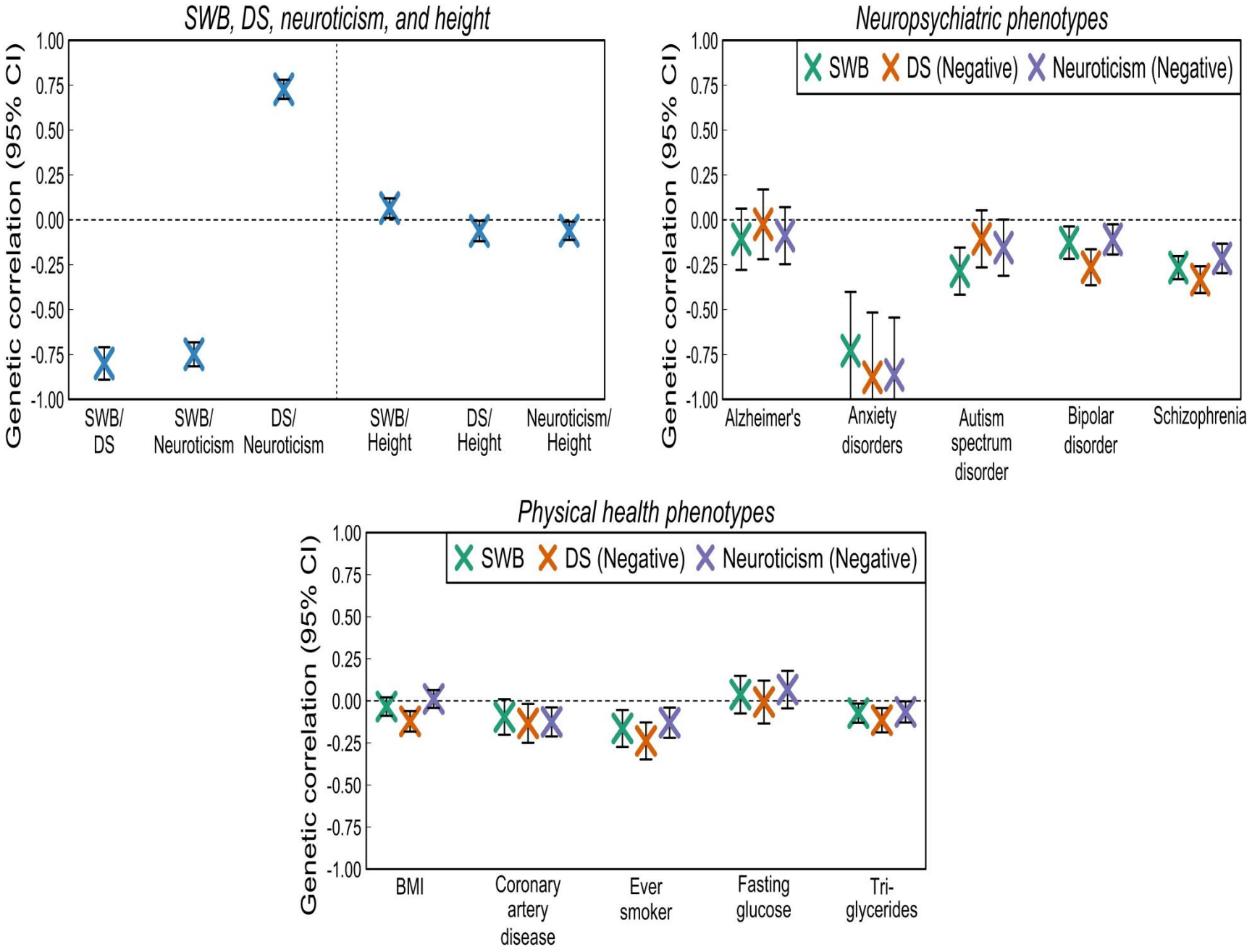


Nature Genetics, accepted 2/3/16

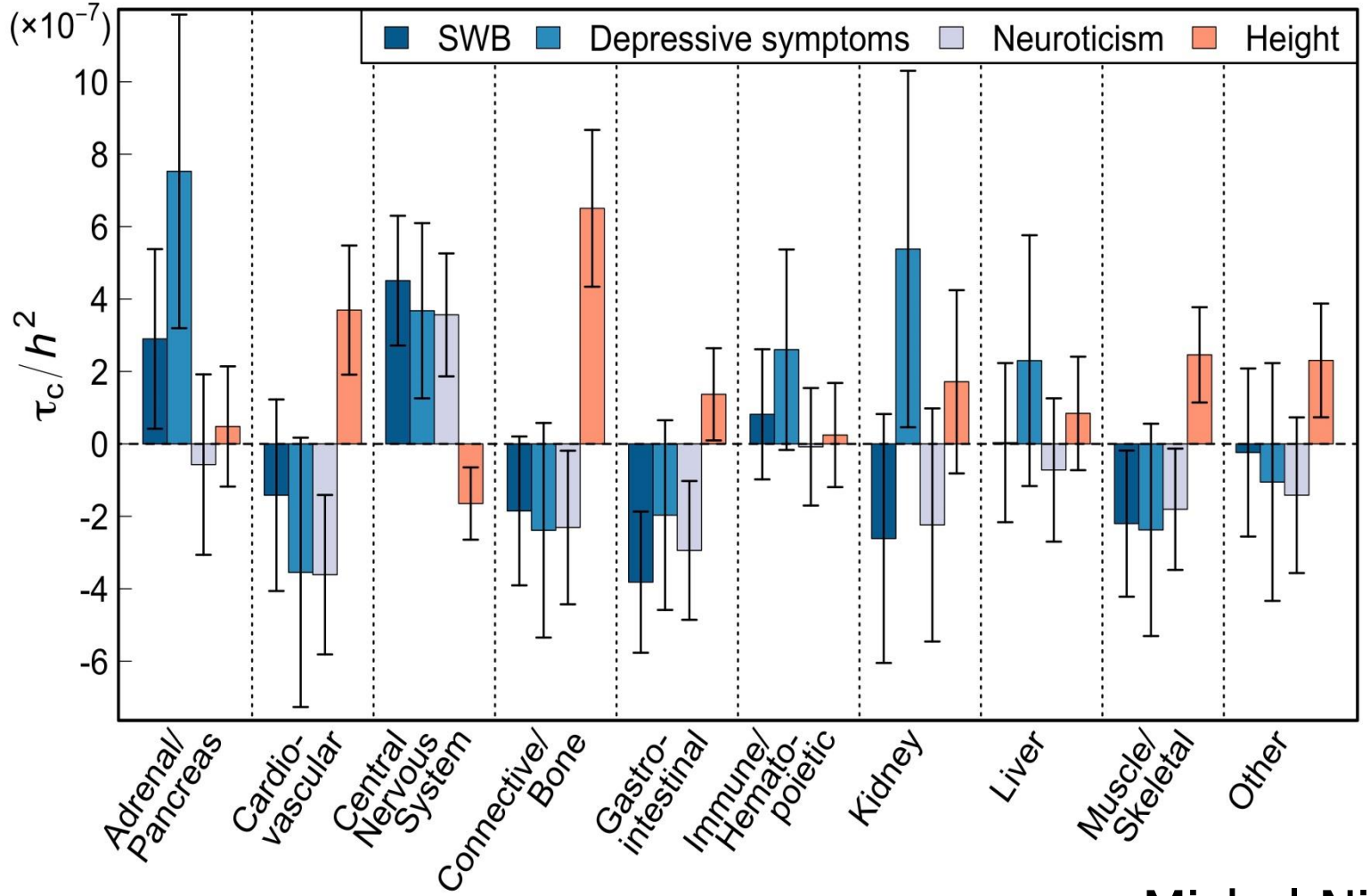
Subjective well-being : N = 298,420



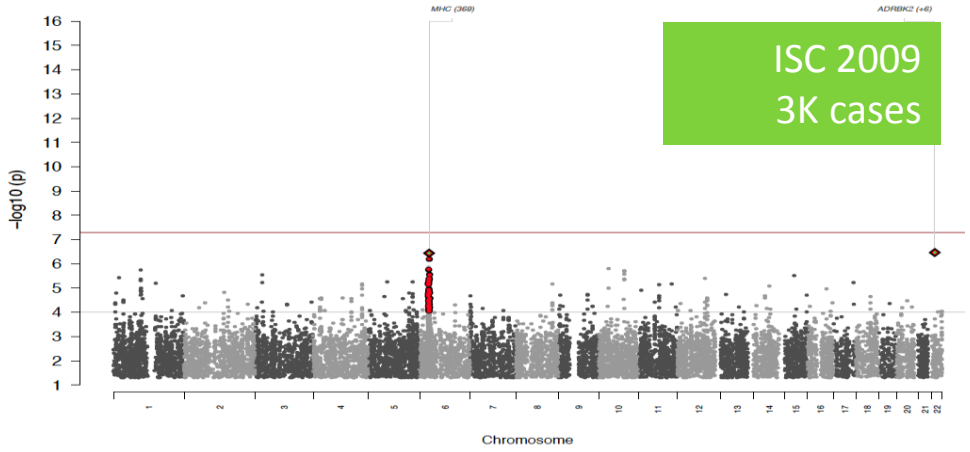
Bartels et al: *Nature Genetics,* accepted 2/3/16


SWB, DS, neuroticism, and height
Genetic correlation (95% CI)
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
SWB/ DS
SWB/ Neuroticism
DS/ Neuroticism
SWB/ Height
DS/ Height
Neuroticism/ Height
Neuropsychiatric phenotypes
SWB
DS (Negative)
Neuroticism (Negative)
Alzheimer's
Anxiety disorders
Autism spectrum disorder
Bipolar disorder
Schizophrenia
Physical health phenotypes
BMI
Coronary artery disease
Ever smoker
Fasting glucose
Tri-glycerides

Across our phenotypes, loci regulating expression in central nervous system and adrenal/pancreas tissues are strongly enriched for association and contrast with Height

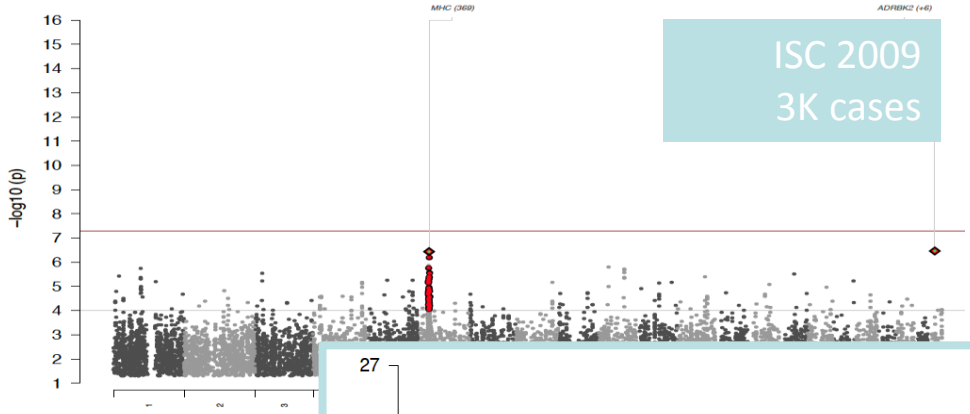


Michel Nivard

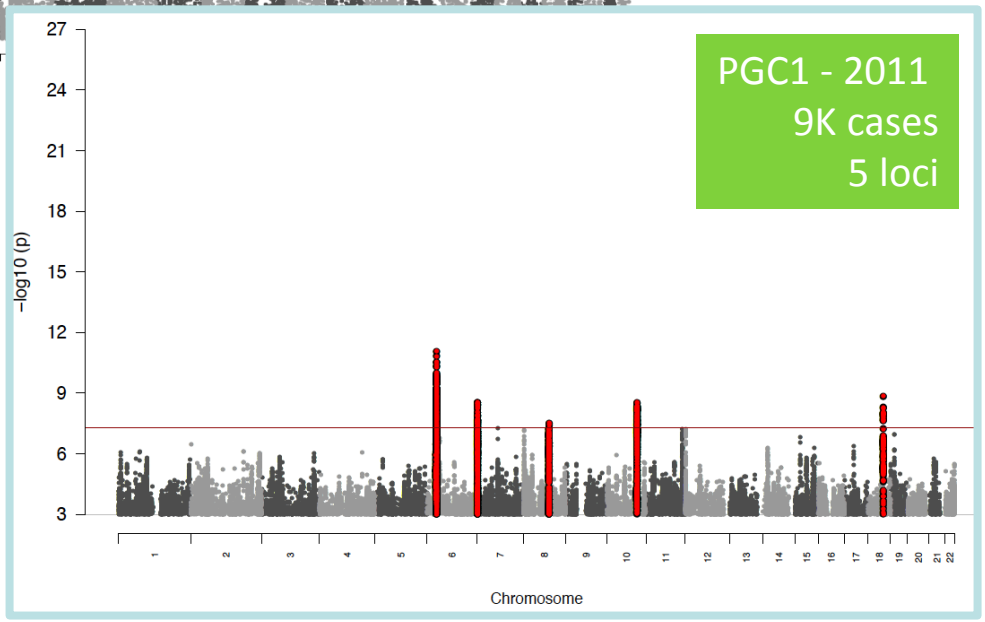


Schizophrenia 2009

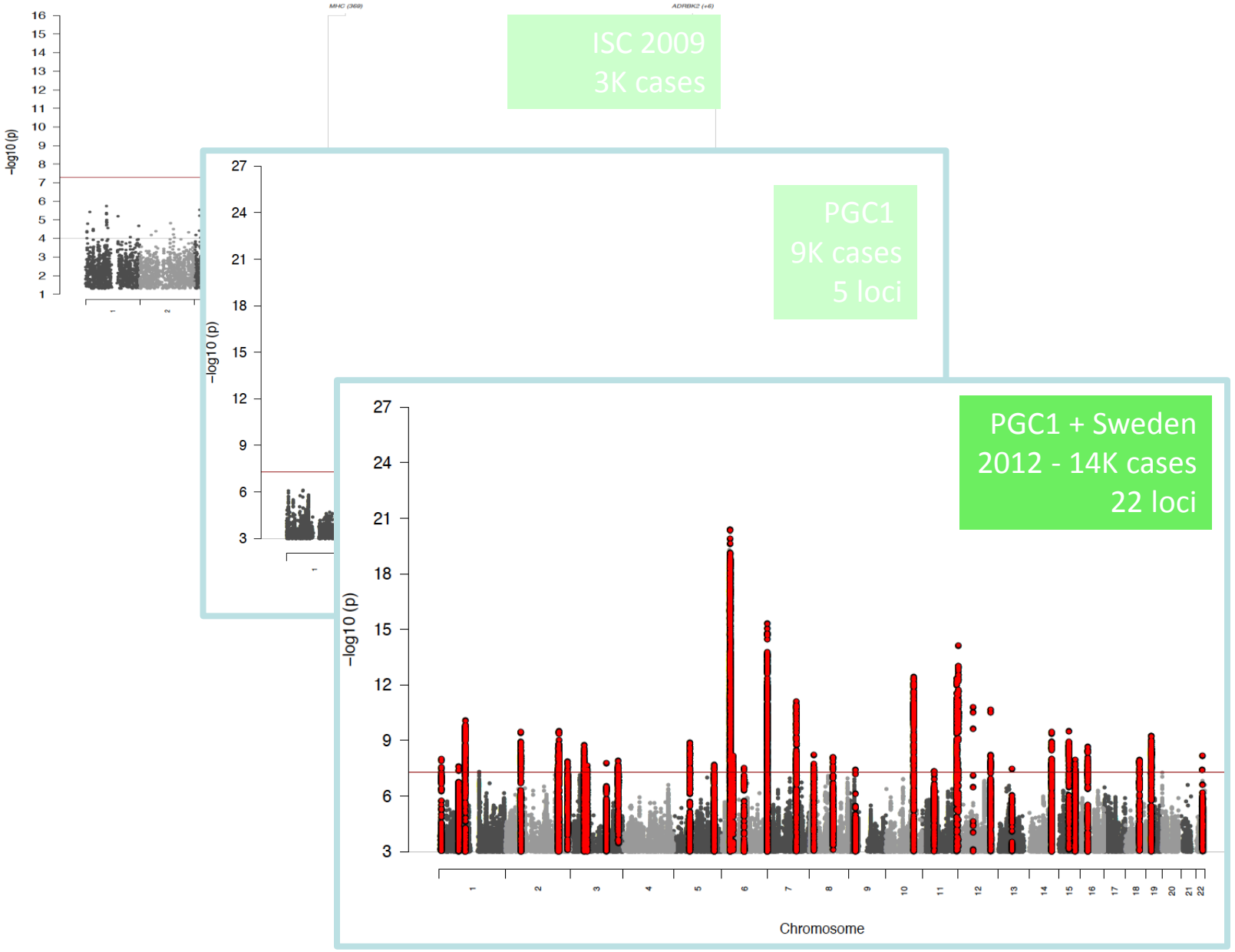
Dramatic progress in GWAS for Schizophrenia

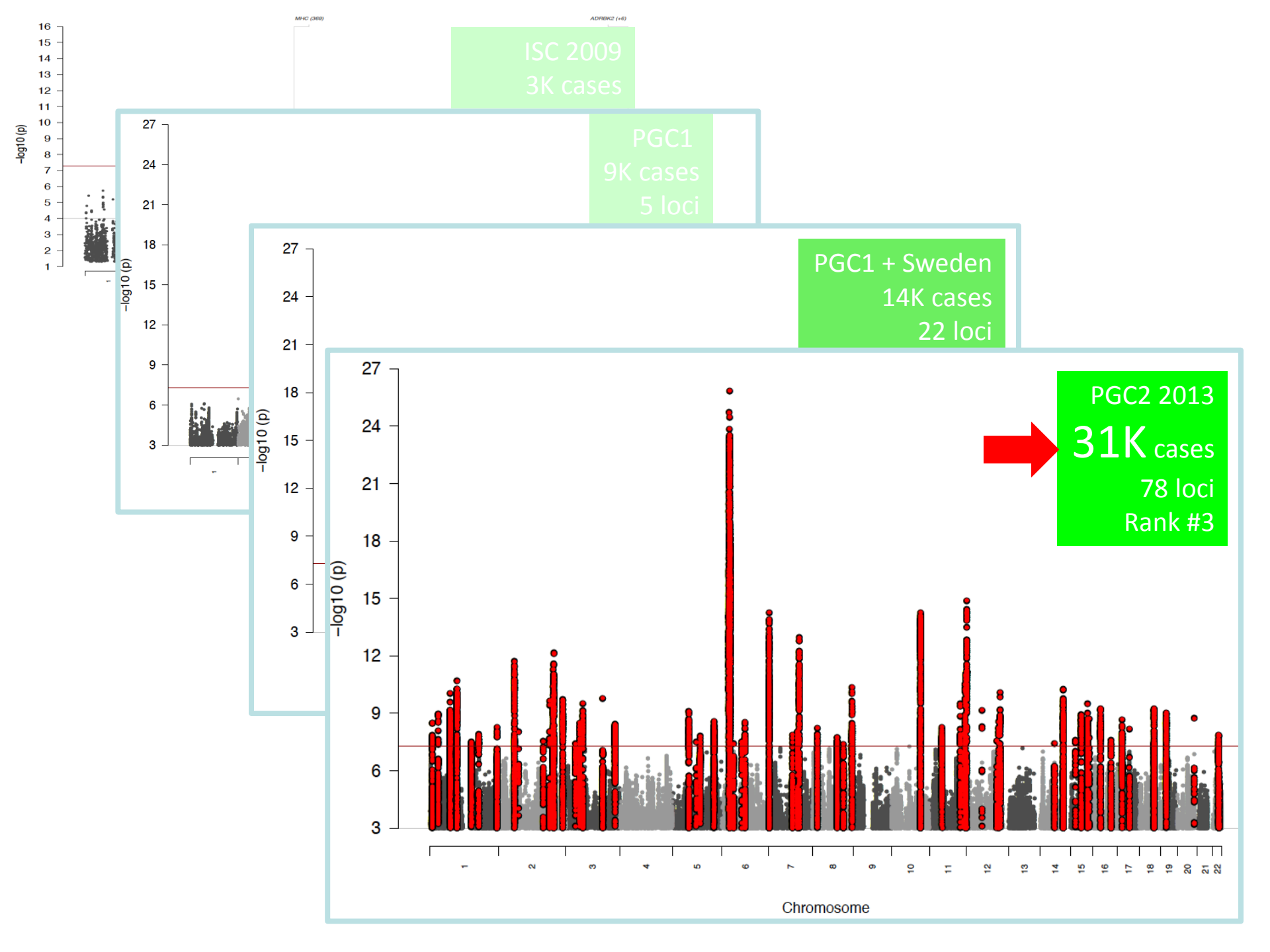


ISC 2009
3K cases

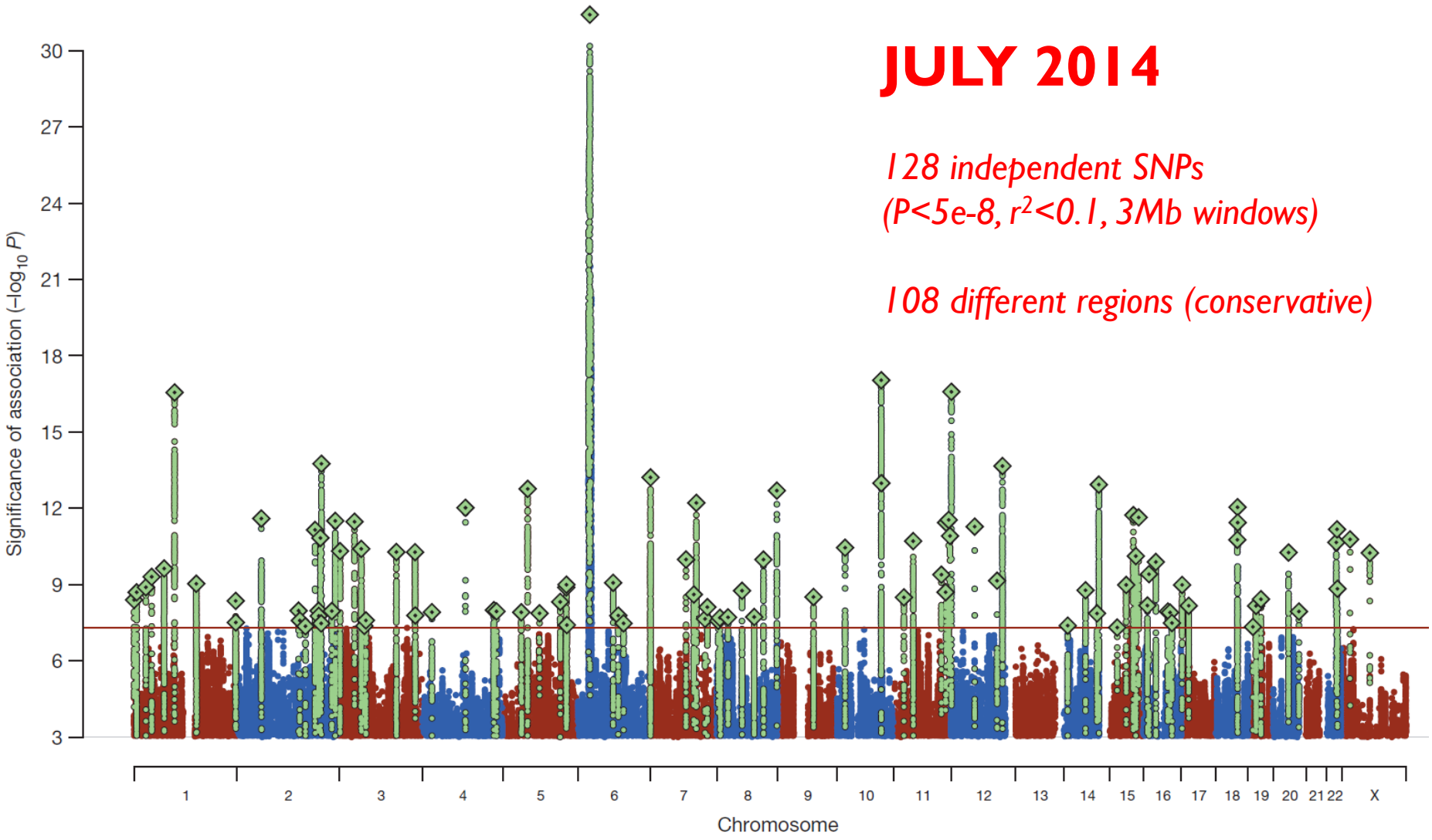


PGC1 - 2011
9K cases
5 loci


ISC 2009
3K cases
PGC1
9K cases
5 loci
PGC1 + Sweden
2012 - 14K cases
22 loci
−log10 (p)
Chromosome

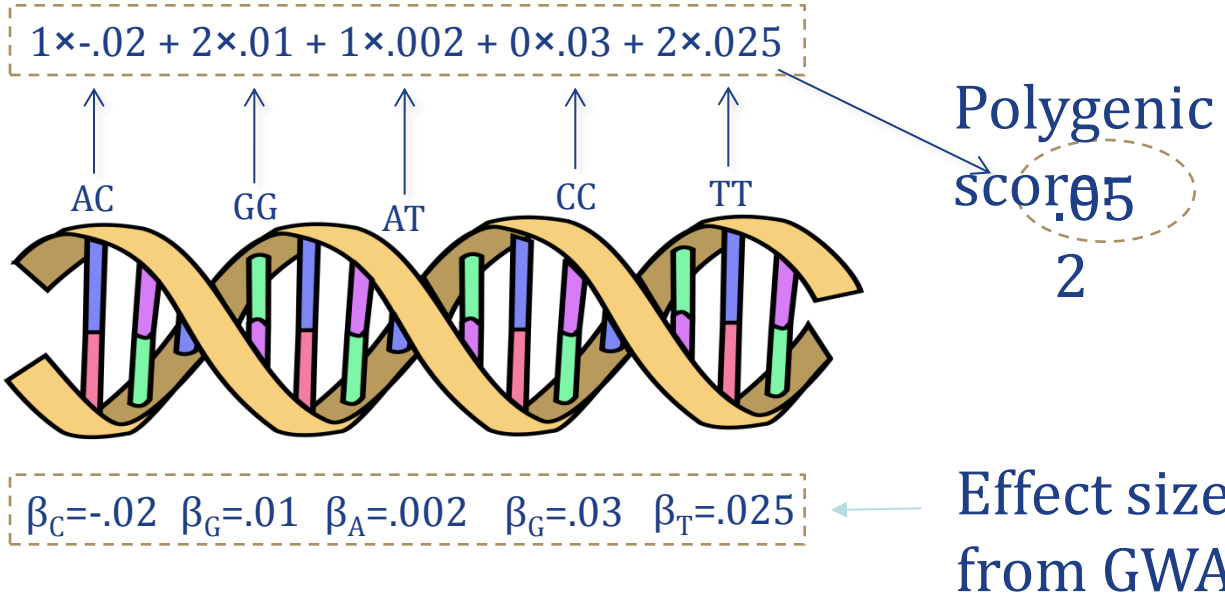

ISC 2009
3K cases
PGC1
9K cases
5 loci
PGC1 + Sweden
14K cases
22 loci
PGC2 2013
31K cases
78 loci
Rank #3
−log10 (p)
Chromosome

Schizophrenia: meta-analysis of 49 case control samples (34,241 cases and 45,604 controls)



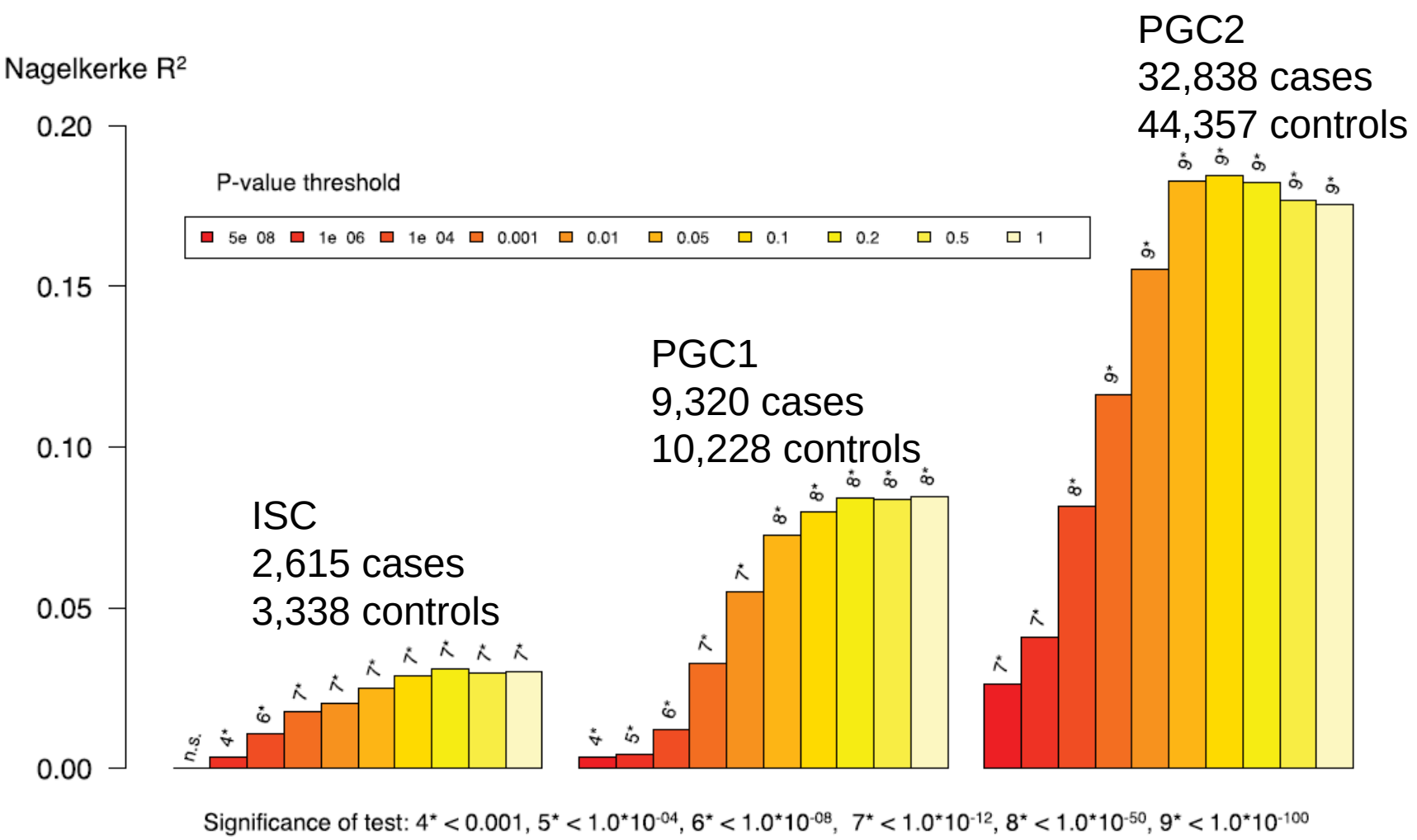
Polygenic Risk Scores

Polygenic Risk Scores capture (part of) someone's genetic "risk" by summing all risk alleles weighted by the effect sizes estimated in a Genome-Wide Association Study (GWAS)

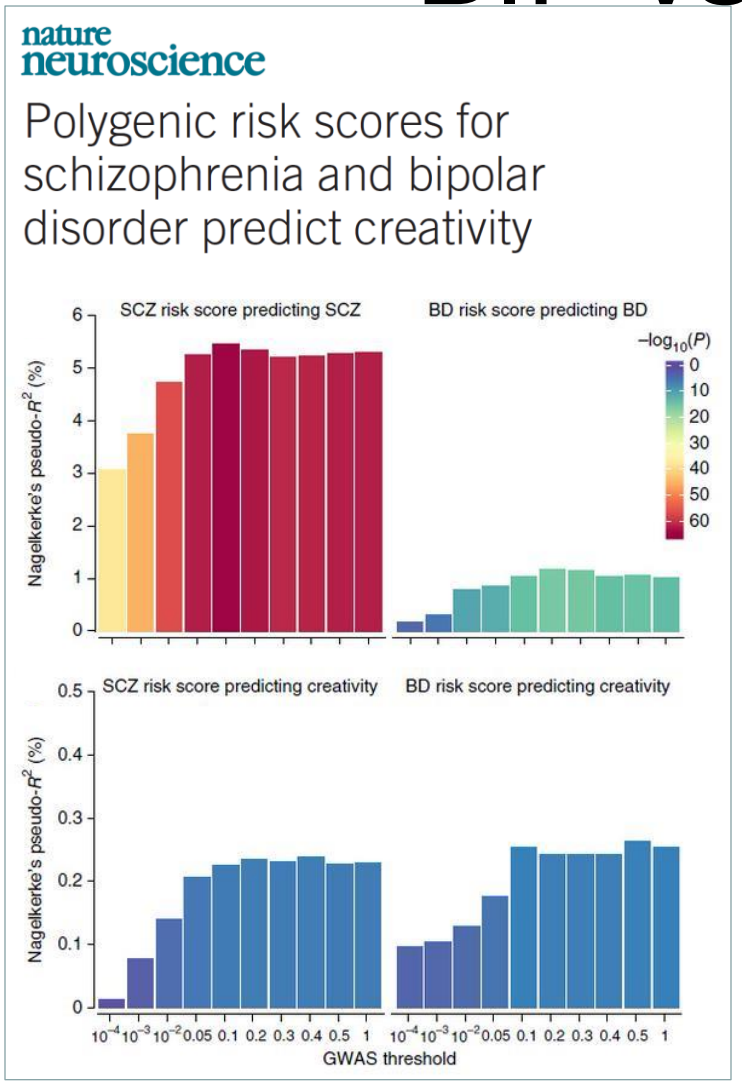


Abdel Abdellaoui

Polygenic risk score for schizophrenia using MGS sample as target and ISC, PGC1, and PGC2 as prediction samples (minus MGS!)



Polygenic Risk Scores: SCZ & BIP vs Creativity



Higher polygenic scores for schizophrenia and bipolar disorder associated with artistic society membership or creative profession in Iceland & replication cohorts.

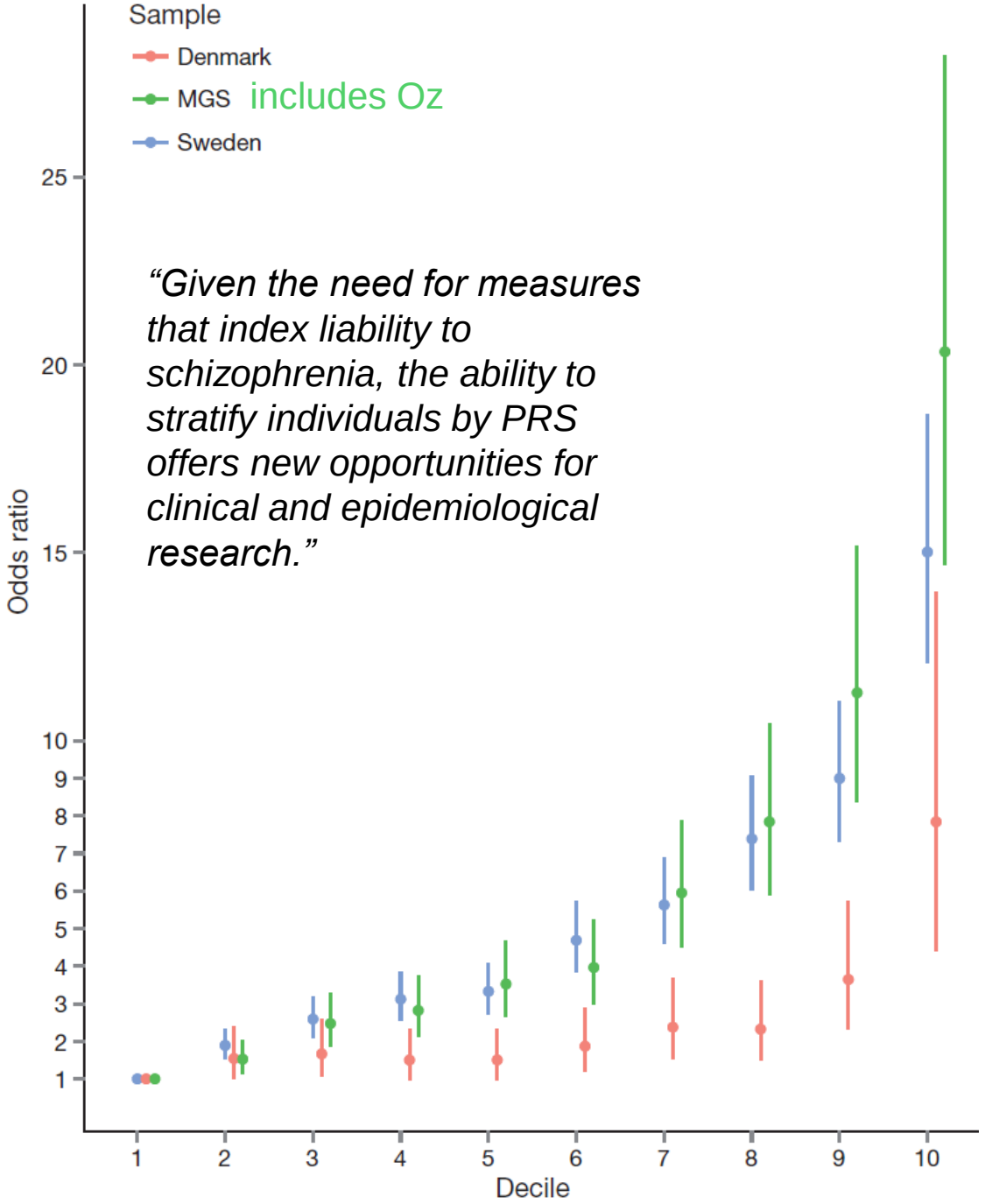
Table 1 Prediction of artist status by schizophrenia and bipolar disorder polygenic risk scores in four replication cohorts

Phenotype and study	*n*, artist	*n*, non-artist	Schizophrenia score OR	Schizophrenia score *P*	Bipolar disorder score OR	Bipolar disorder score *P*
Artistic profession						
NTR	118	6,160	1.31	0.022	1.43	0.00015
RS	62	7,819	1.18	0.19	1.10	0.46
STR	65	8,828	1.25	0.096	1.16	0.23
NBS	28	4,265	1.12	0.54	0.95	0.79
Combined	273	27,072	1.23	0.0021	1.23	0.00086
CAQ-Arts score	*n*					
STR	804		1.11	0.11	1.06	0.34

n values for artist and non-artist are shown where applicable. Polygenic risk scores were calculated using a *P*-value threshold of 0.2 and were standardized to have a mean of 0 and a s.d. of 1. Results are from logistic regression except in the case of the CAQ-Arts score, where ordered logistic regression was used. NTR, Netherlands Twin Registry; RS, Rotterdam Study; STR Swedish Twin Registry; NBS, Nijmegen Biomedical Study.

Abdel Abdellaoui

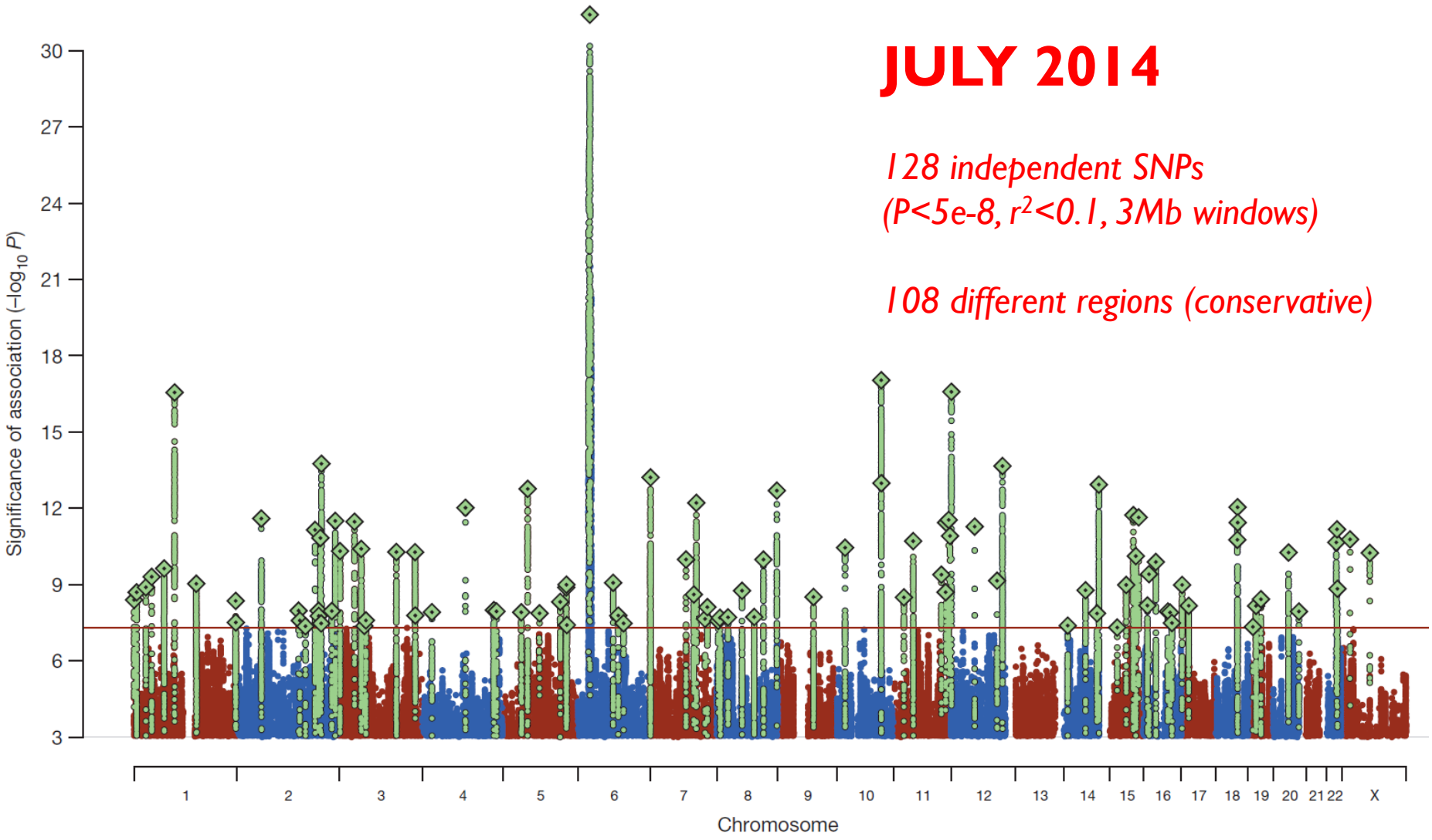
Odds ratio for schizophrenia by polygenic risk score (PRS) decile in the Sweden, Denmark, and Molecular Genetics of Schizophrenia (US+Oz) studies.



"Given the need for measures that index liability to schizophrenia, the ability to stratify individuals by PRS offers new opportunities for clinical and epidemiological research."

Risk alleles and weights were derived from 'leave one out' analyses in which those samples were excluded from the GWAS meta-analysis. The threshold for selecting risk alleles was $P<0.05$.

Schizophrenia: meta-analysis of 49 case control samples (34,241 cases and 45,604 controls)



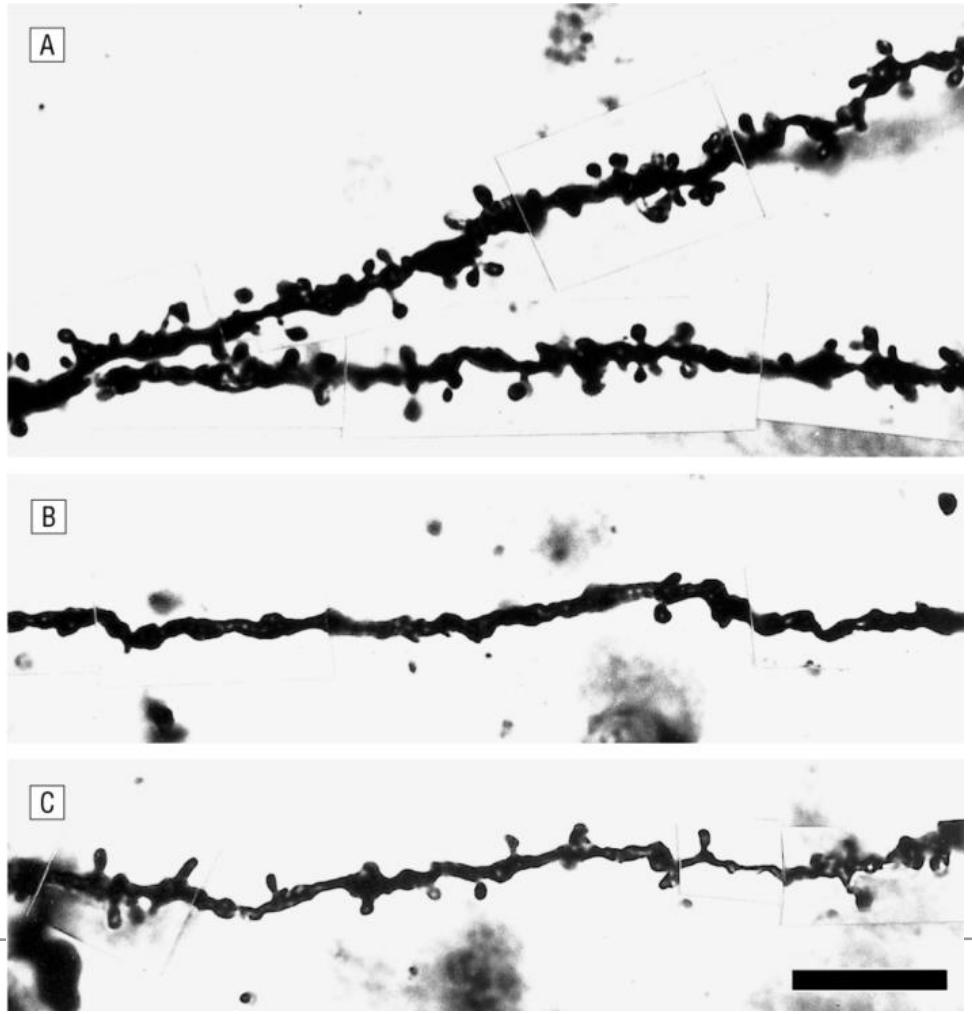
Schizophrenia risk from complex variation of complement component 4

Aswin Sekar[1,2,3], Allison R. Bialas[4,5], Heather de Rivera[1,2], Avery Davis[1,2], Timothy R. Hammond[4], Nolan Kamitaki[1,2], Katherine Tooley[1,2], Jessy Presumey[5], Matthew Baum[1,2,3,4], Vanessa Van Doren[1], Giulio Genovese[1,2], Samuel A. Rose[2], Robert E. Handsaker[1,2], Schizophrenia Working Group of the Psychiatric Genomics Consortium*, Mark J. Daly[2,6], Michael C. Carroll[5], Beth Stevens[2,4] & Steven A. McCarroll[1,2]

Schizophrenia is a heritable brain illness with unknown pathogenic mechanisms. Schizophrenia's strongest genetic association at a population level involves variation in the major histocompatibility complex (MHC) locus, but the genes and molecular mechanisms accounting for this have been challenging to identify. Here we show that this association arises in part from many structurally diverse alleles of the complement component 4 (*C4*) genes. We found that these alleles generated widely varying levels of *C4A* and *C4B* expression in the brain, with each common *C4* allele associating with schizophrenia in proportion to its tendency to generate greater expression of *C4A*. Human C4 protein localized to neuronal synapses, dendrites, axons, and cell bodies. In mice, C4 mediated synapse elimination during postnatal development. These results implicate excessive complement activity in the development of schizophrenia and may help explain the reduced numbers of synapses in the brains of individuals with schizophrenia.

From: **Decreased Dendritic Spine Density on Prefrontal Cortical Pyramidal Neurons in Schizophrenia**

Arch Gen Psychiatry. 2000;57(1):65-73. doi:10.1001/archpsyc.57.1.65



Basilar dendrites and spines on dorsolateral prefrontal cortex layer 3 pyramidal neurons from normal control subject (A) and 2 subjects with schizophrenia (B and C). The calibration bar equals 10 µm.

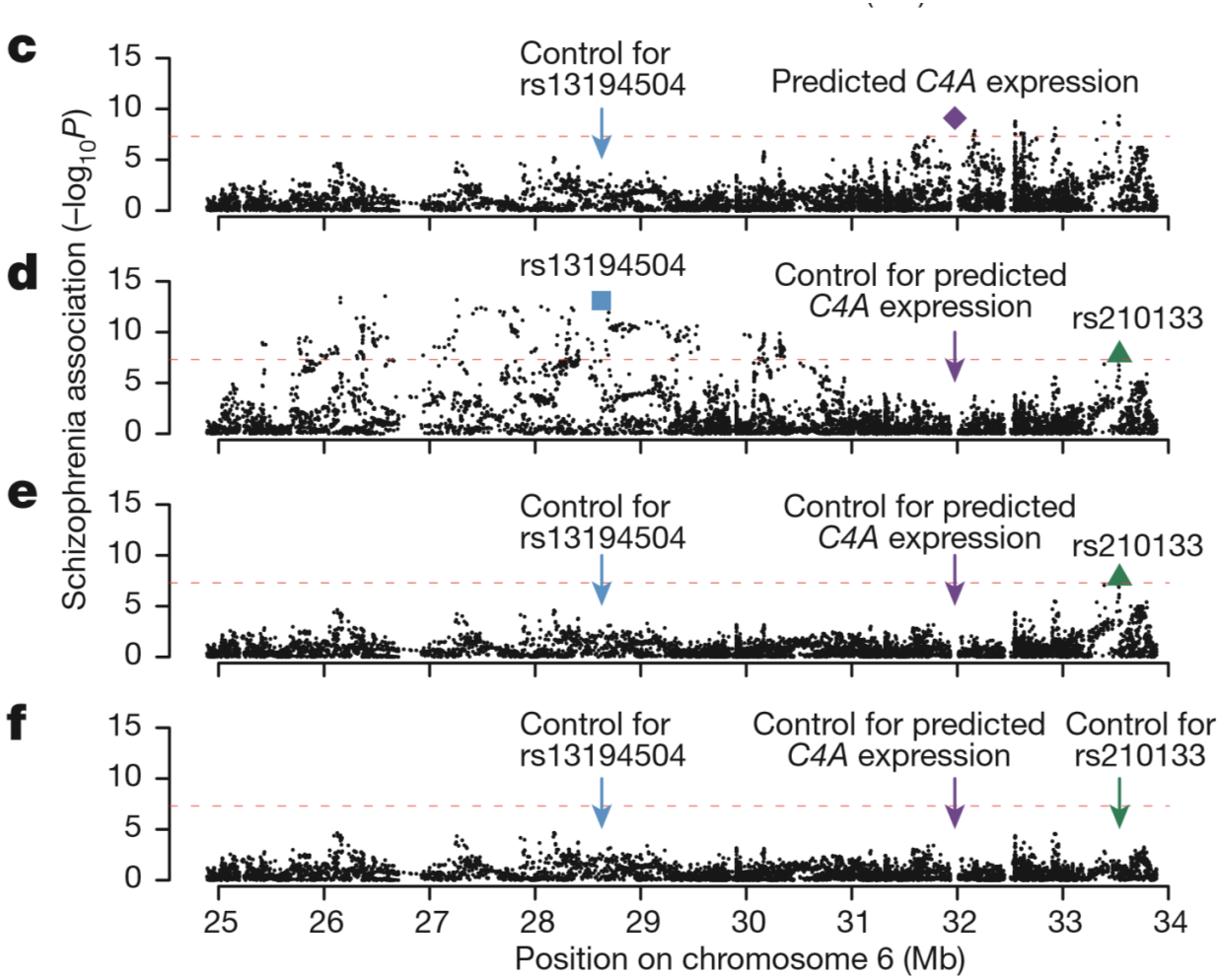


Figure 4 | Association of schizophrenia to *C4* and the extended MHC locus. Association of schizophrenia to 7,751 SNPs across the MHC locus and to genetically predicted expression levels of *C4A* and *C4B* in the brain (represented in the genomic location of the *C4* gene). The data shown are based on analysis of 28,799 schizophrenia cases and 35,986 controls of

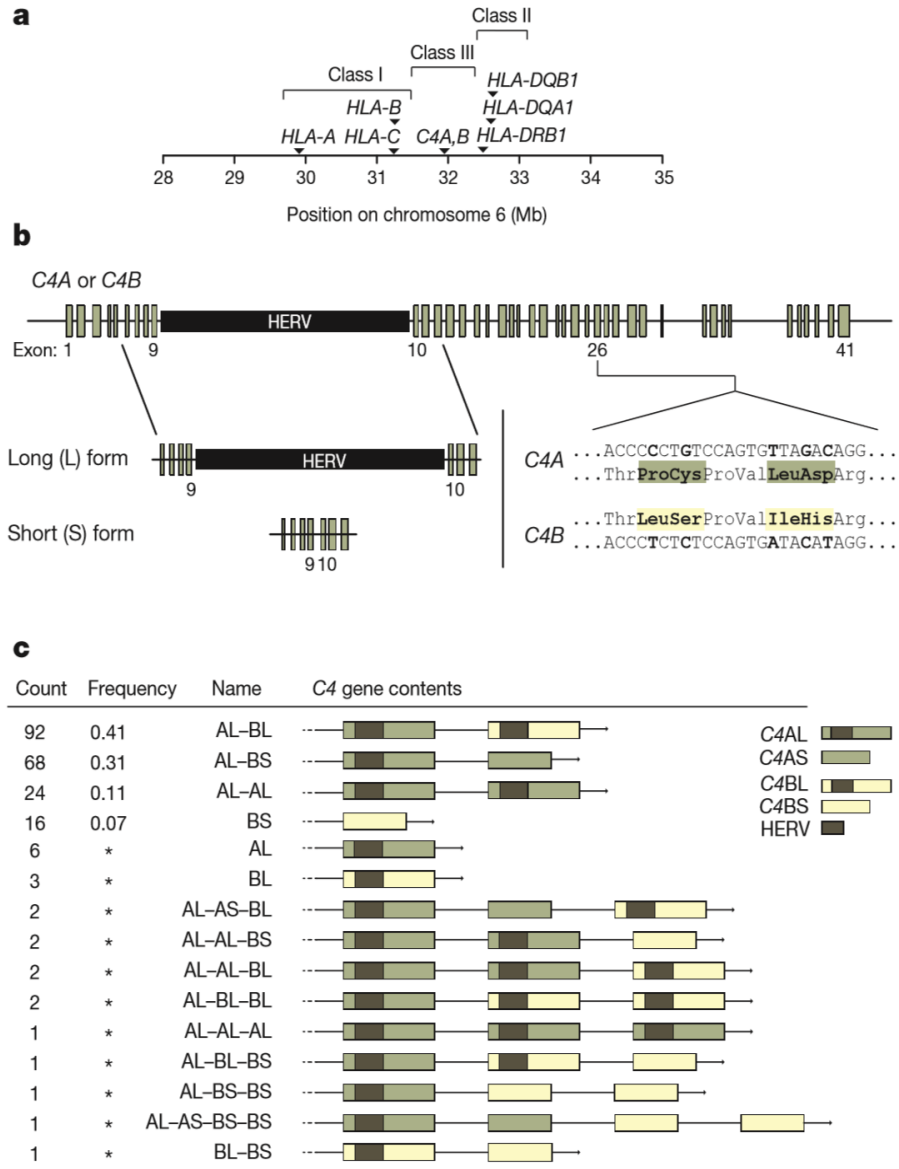


Figure 1 | Structural variation of the complement component 4 (*C4*) gene. a, Location of the *C4* genes within the major histocompatibility complex (MHC) locus on human chromosome 6. **b**, Human *C4* exists

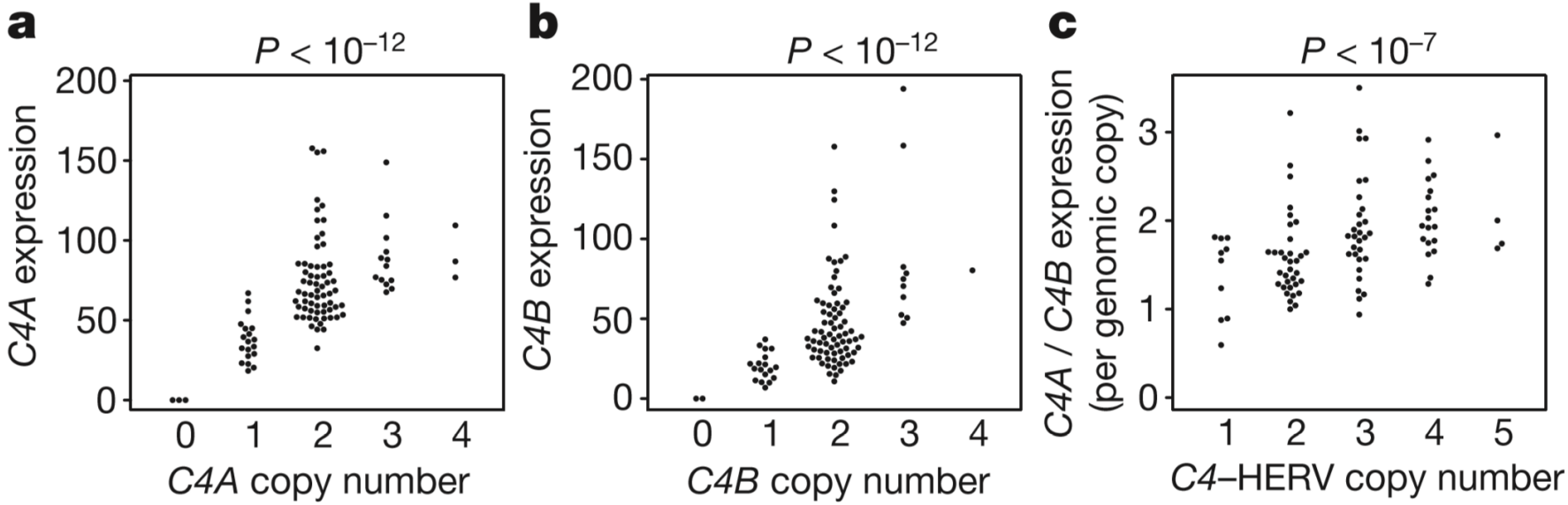


Figure 3 | Brain RNA expression of *C4A* and *C4B* in relation to copy numbers of *C4A*, *C4B*, and the *C4*–HERV. a, **b**, mRNA expression of *C4A* (**a**) and *C4B* (**b**) was measured (by ddPCR) in brain tissue from 244 individuals. The copy numbers of *C4A*, *C4B*, and the *C4*–HERV were measured (by ddPCR) in genomic DNA from the brain donors. The

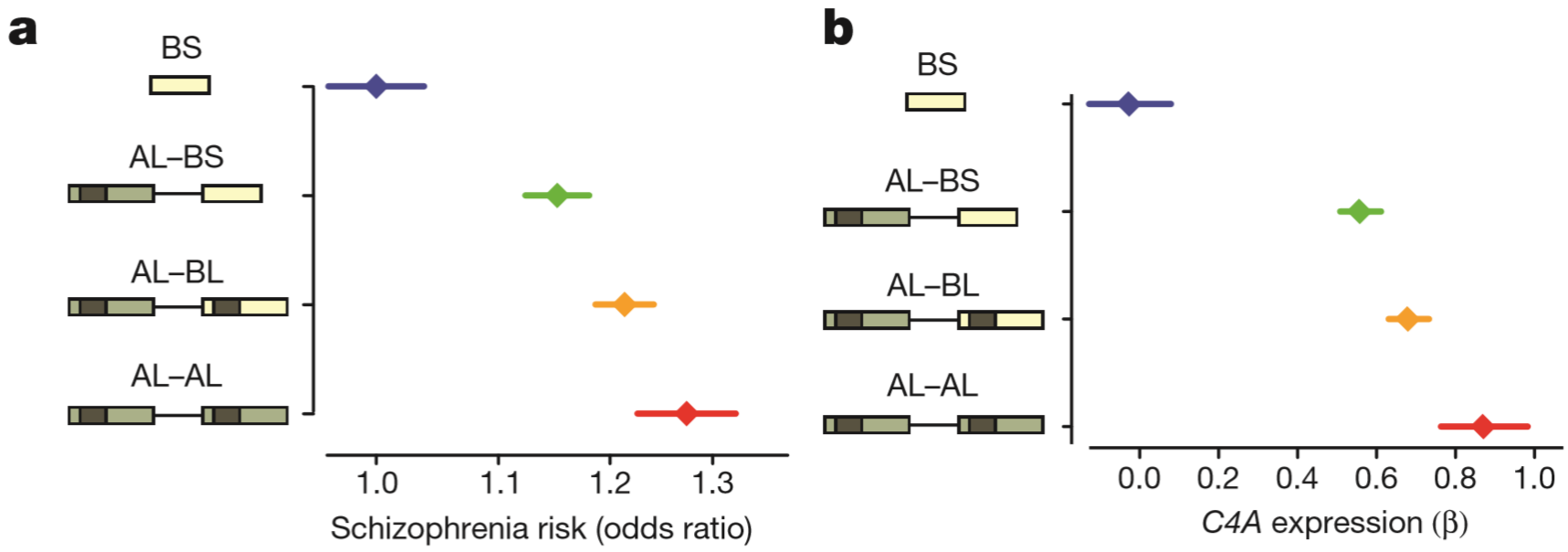
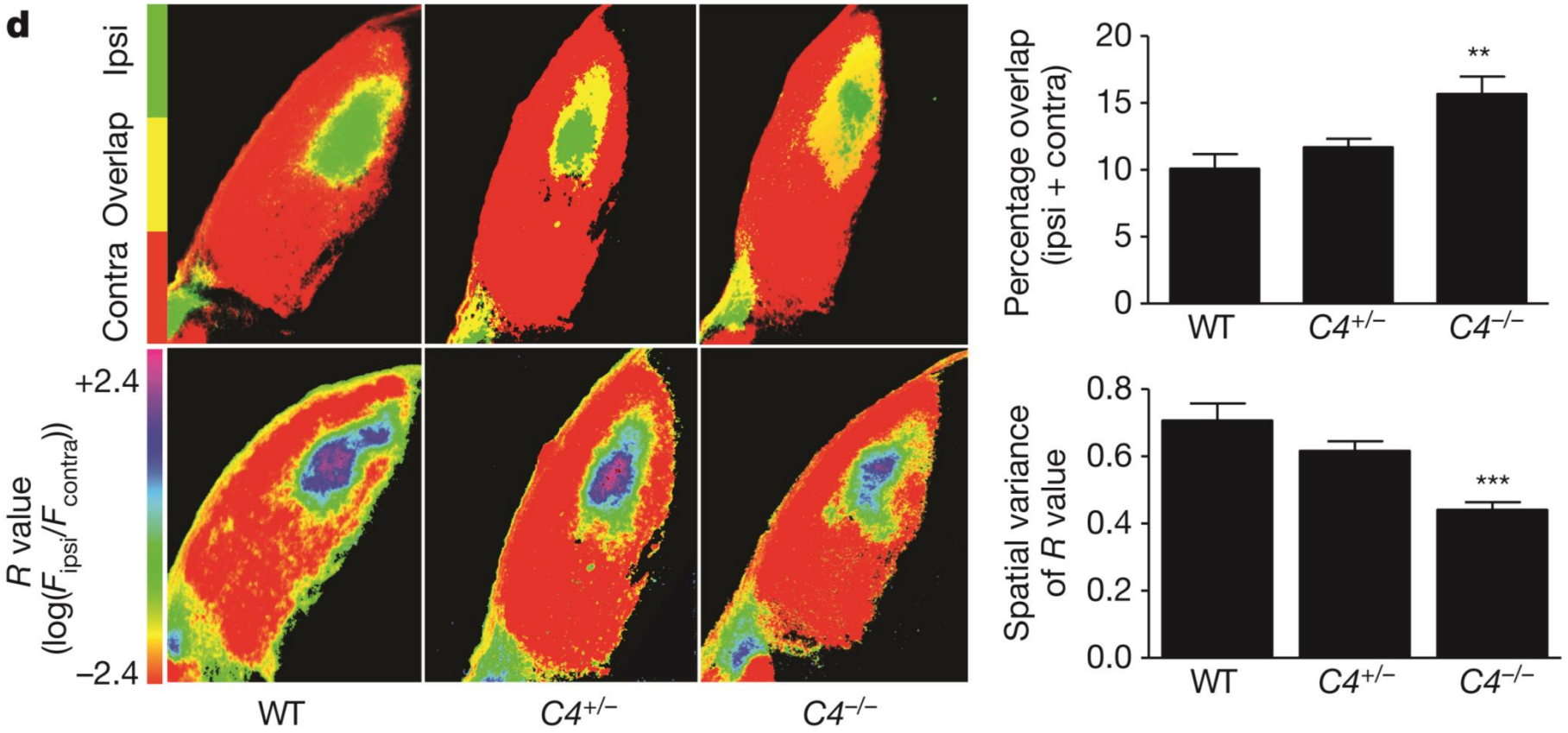


Figure 5 | *C4* structures, *C4A* expression, and schizophrenia risk. a, Schizophrenia risk associated with four common structural forms of *C4* in analysis of 28,799 schizophrenia cases and 35,986 controls. **b**, Brain *C4A* RNA expression levels associated with four common structural forms of *C4*. β was calculated from fitting *C4A* RNA expression (in

Impairments in schizophrenia tend to affect higher cognitive functions and recently expanded brain regions. Waves of postnatal synapse elimination occur in many brain regions. Synaptic projections from retinal ganglion cells (RGCs) onto thalamic relay neurons within the dorsal lateral geniculate nucleus (dLGN) of the visual thalamus undergo activity-dependent synaptic refinement. We found that C4 RNA was expressed in the LGN and in RGCs purified from the retina during the period of synaptic pruning.



Synaptic refinement in mice with 0, 1, or 2 copies of C4. These images represent the segregation of ipsilateral and contralateral RGC projections to the dLGN

3 Stages of Genetic Mapping

- Are there genes influencing this trait?
  - Genetic epidemiological studies
- What are those genes?
  - Association analysis
- What can you do with them?
  - Translational research

NATURE GENETICS ADVANCE ONLINE PUBLICATION

The support of human genetic evidence for approved drug indications

Matthew R Nelson[1], Hannah Tipney[2], Jeffery L Painter[1], Judong Shen[1], Paola Nicoletti[3], Yufeng Shen[3,4], Aris Floratos[3,4], Pak Chung Sham[5,6], Mulin Jun Li[6,7], Junwen Wang[6,7], Lon R Cardon[8], John C Whittaker[2] & Philippe Sanseau[2]

Table 1 The relative value of genetic support for the probability that a target-indication pair progresses along the drug development pipeline, based on historical drug trial information

	p(progress|genetic support)/(progress|no genetic support)		
Progression	GWASdb and OMIM	GWASdb	OMIM
Phase I to phase II	1.2 (1.1–1.3)	1.2 (1.1–1.3)	1.2 (1.1–1.3)
Phase II to phase III	1.5 (1.3–1.7)	1.4 (1.2–1.7)	1.6 (1.3–1.9)
Phase III to approval	1.1 (1.0–1.2)	1.0 (0.8–1.2)	1.1 (0.9–1.3)
Phase I to phase III	1.8 (1.5–2.1)	1.8 (1.4–2.1)	1.9 (1.5–2.3)
Phase I to approval	2.0 (1.6–2.4)	1.8 (1.3–2.3)	2.2 (1.6–2.8)

"We estimate that selecting genetically supported targets could double the success rate in clinical development. Therefore, using the growing wealth of human genetic data to select the best targets and indications should have a measurable impact on the successful development of new drugs."

The End