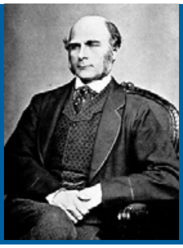
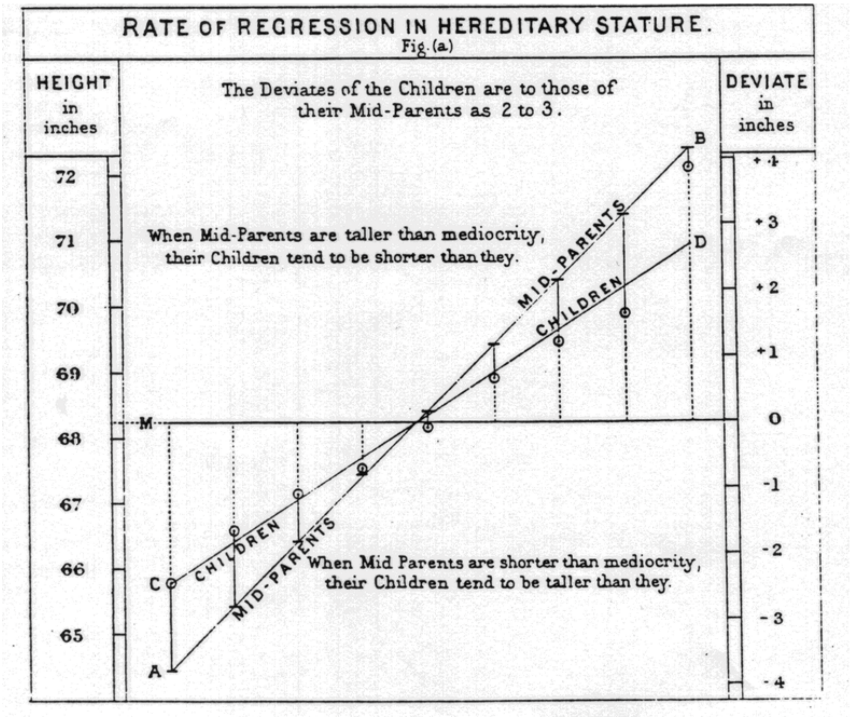
← Found on IBG website

LD Score – new methods for heritability

Benjamin Neale, Ph.D.
Analytic and Translational Genetics Unit, MGH
Stanley Center for Psychiatric Research & Program in Medical and Population Genetics, Broad Institute

Galton





- Children resemble parents
- Origin of term “regression”

East 1915: Inheritance of corolla lemgth in Nicotiana Longiflora

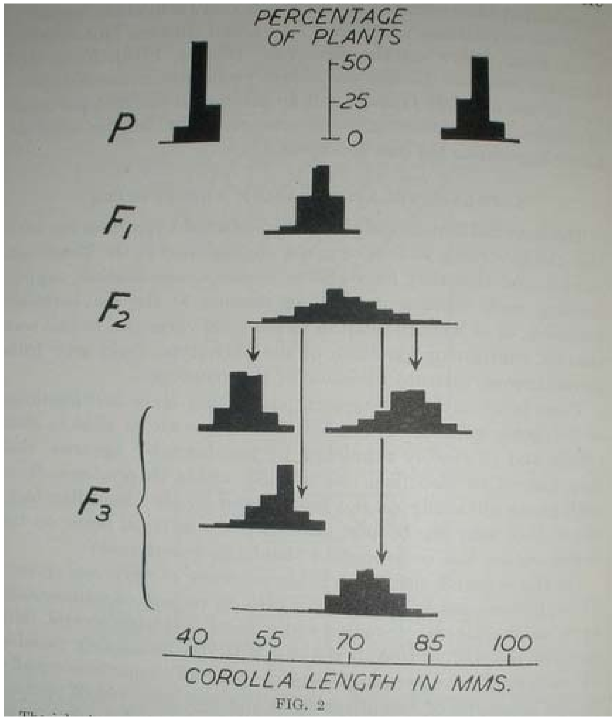
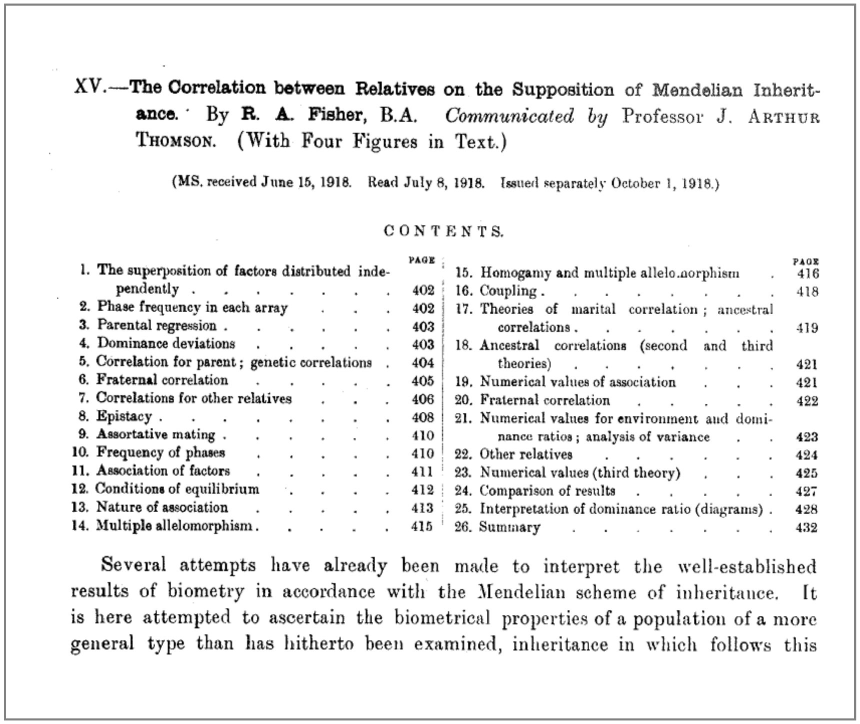


FIG. 2

- Breeding experiments in plants clearly demonstrate polygenic inheritance

Neo-Darwinist reconciliation
RA Fisher



XV.—**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.)

CONTENTS.

Several attempts have already been made to interpret the well-established results of biometry in accordance with the Mendelian scheme of inheritance. It is here attempted to ascertain the biometrical properties of a population of a more general type than has hitherto been examined, inheritance in which follows this

LD Score regression

With thanks


Brendan Bulik-Sullivan


Hilary Finucane


Po-Ru Loh


Mark Daly


Alkes Price

How does LD shape association?

LD Score regression distinguishes confounding from polygenicity in genome-wide association studies

Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Nick Patterson, Mark J Daly, Alkes L Price & Benjamin M Neale

Affiliations | Contributions | Corresponding author

Nature Genetics **47**, 291–295 (2015) | doi:10.1038/ng.3211
Received 07 March 2014 | Accepted 07 January 2015 | Published online 02 February 2015

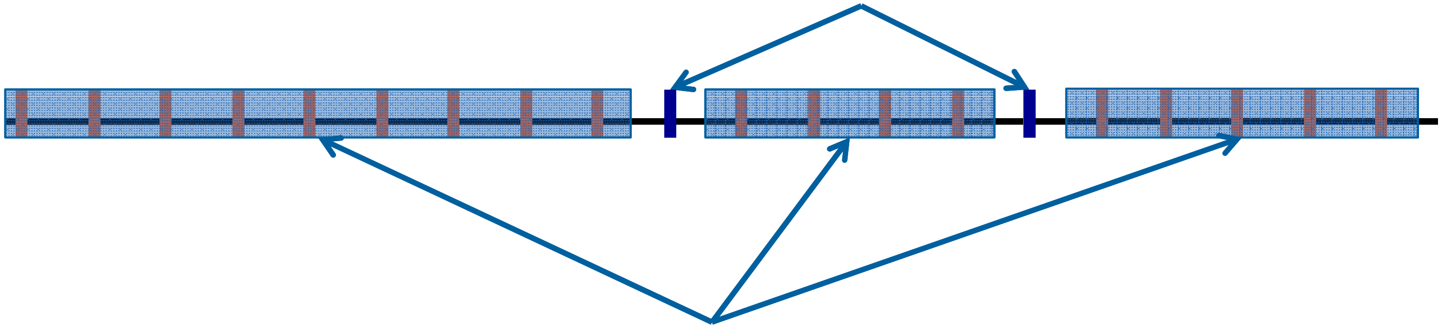
How does LD shape association?

LD Score regression distinguishes confounding from polygenicity in genome-wide association studies

Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Nick Patterson, Mark J Daly, Alkes L Price & Benjamin M Neale

Affiliations | Contributions | Corresponding author

Nature Genetics **47**, 291–295 (2015) | doi:10.1038/ng.3211
Received 07 March 2014 | Accepted 07 January 2015 | Published online 02 February 2015



Lonely SNPs [no LD]

LD blocks

How does LD shape association?

▌ Lonely SNPs [no LD]

▨ LD blocks

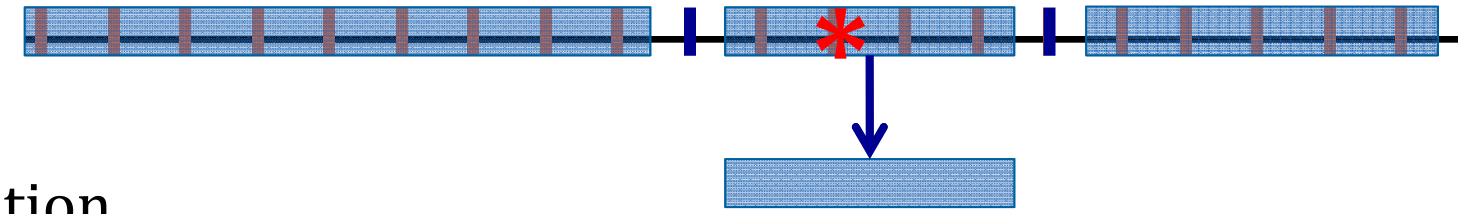
* Causal variants

LD Score regression distinguishes confounding from polygenicity in genome-wide association studies

Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Nick Patterson, Mark J Daly, Alkes L Price & Benjamin M Neale

Affiliations | Contributions | Corresponding author

Nature Genetics **47**, 291–295 (2015) | doi:10.1038/ng.3211
Received 07 March 2014 | Accepted 07 January 2015 | Published online 02 February 2015



Association

All markers correlated with a causal variant show association

How does LD shape association?

▌ Lonely SNPs [no LD]

▨ LD blocks

* Causal variants

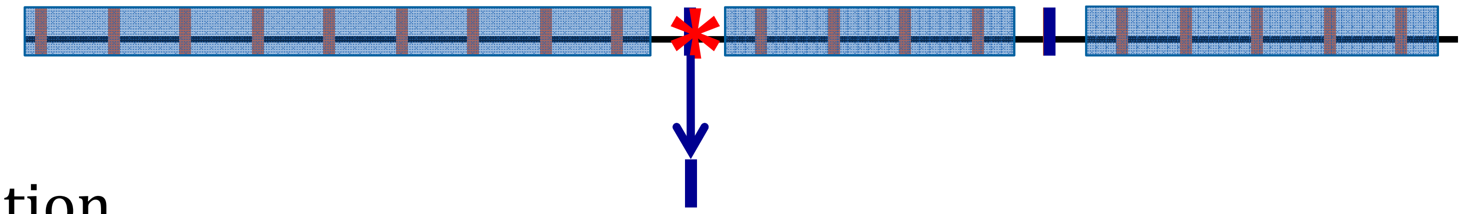
LD Score regression distinguishes confounding from polygenicity in genome-wide association studies

Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Nick Patterson, Mark J Daly, Alkes L Price & Benjamin M Neale

Affiliations | Contributions | Corresponding author

Nature Genetics **47**, 291–295 (2015) | doi:10.1038/ng.3211

Received 07 March 2014 | Accepted 07 January 2015 | Published online 02 February 2015



Association

Lonely SNPs only show association if they are causal

What happens under polygenicity?

▮ Lonely SNPs [no LD]

▨ LD blocks

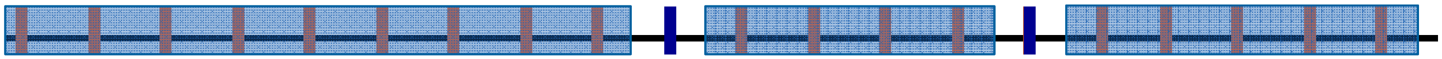
* Causal variants

LD Score regression distinguishes confounding from polygenicity in genome-wide association studies

Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Nick Patterson, Mark J Daly, Alkes L Price & Benjamin M Neale

Affiliations | Contributions | Corresponding author

Nature Genetics 47, 291–295 (2015) | doi:10.1038/ng.3211
Received 07 March 2014 | Accepted 07 January 2015 | Published online 02 February 2015

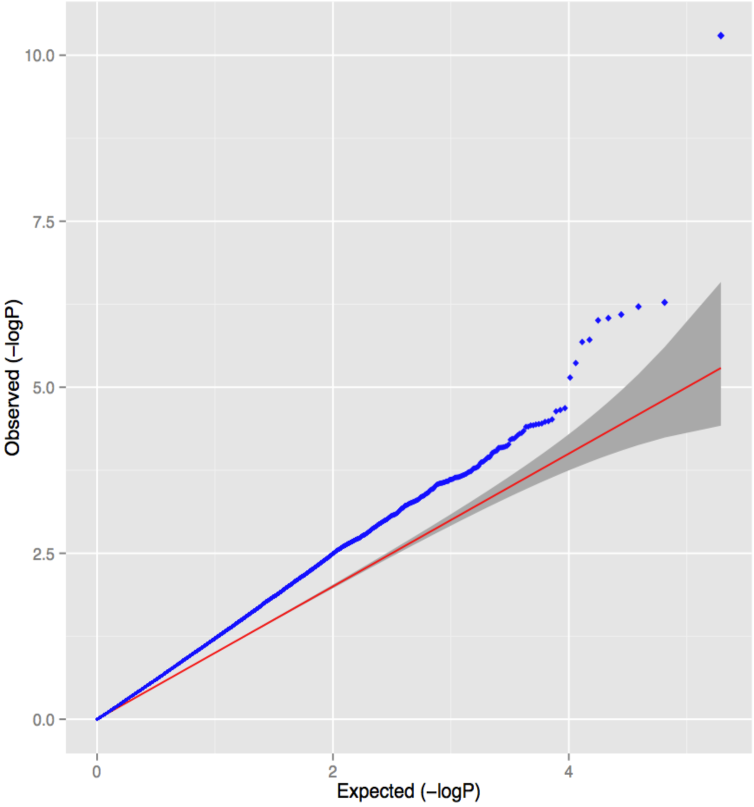


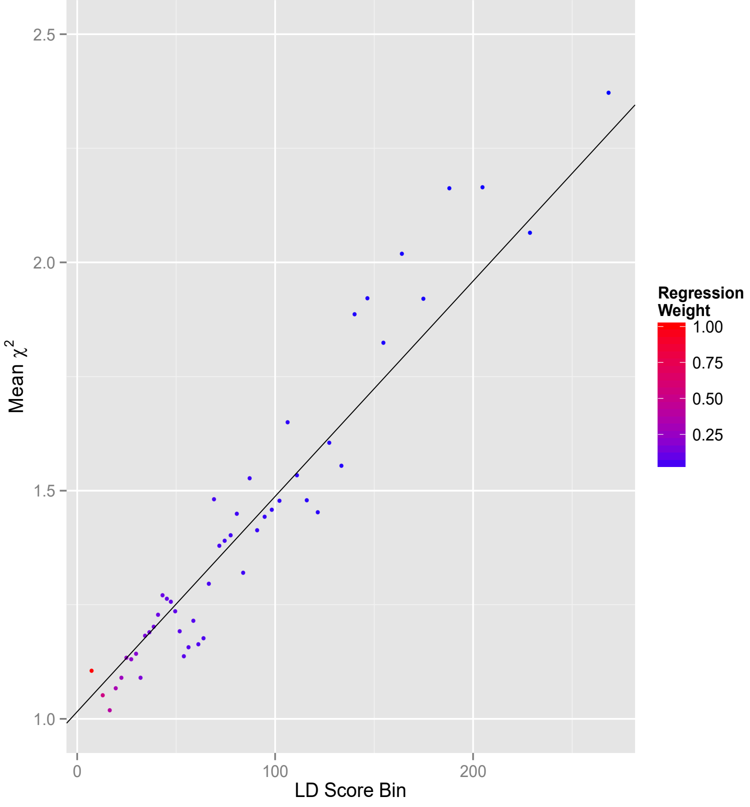
Assuming a uniform prior, we see SNPs with more LD friends showing more association

The more you tag, the more likely you are to tag a causal variant

Simulated polygenic architecture

Lambda = 1.30 LD score intercept = 1.02





What happens under stratification?

▌ Lonely SNPs [no LD]

▨ LD blocks

* Causal variants

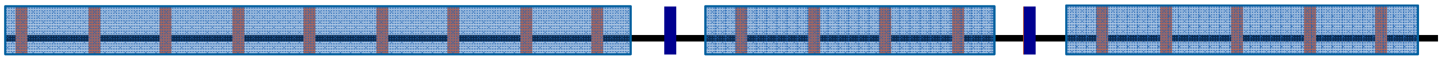
LD Score regression distinguishes confounding from polygenicity in genome-wide association studies

Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Nick Patterson, Mark J Daly, Alkes L Price & Benjamin M Neale

Affiliations | Contributions | Corresponding author

Nature Genetics **47**, 291–295 (2015) | doi:10.1038/ng.3211

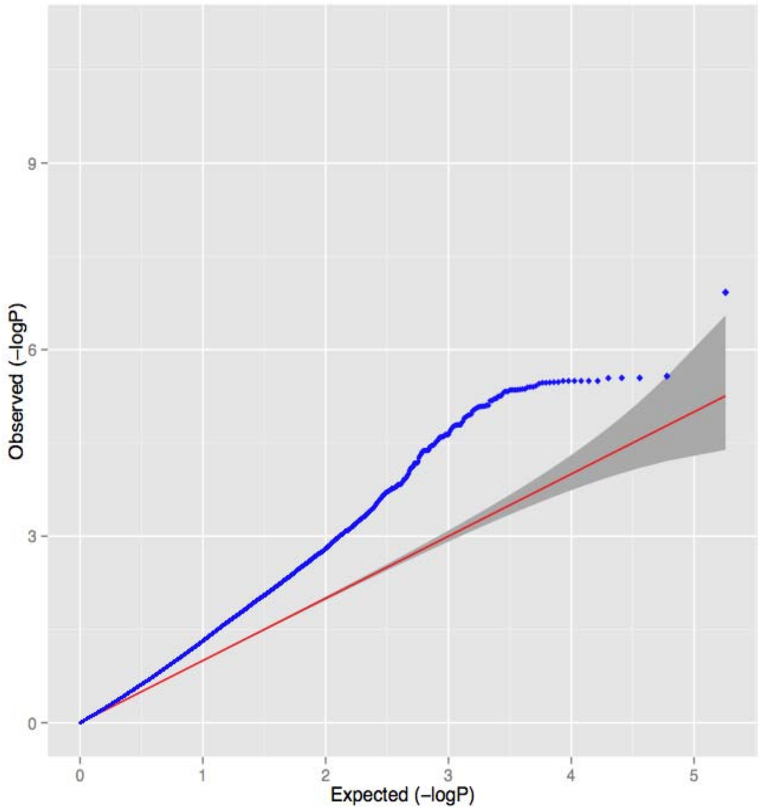
Received 07 March 2014 | Accepted 07 January 2015 | Published online 02 February 2015

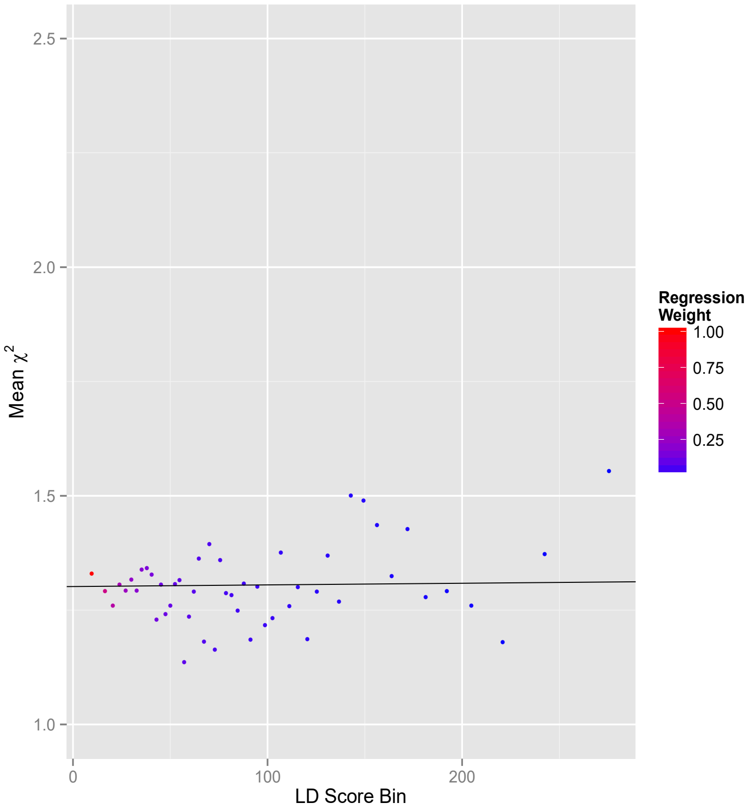


Under pure drift we expect LD to have no relationship to differences in allele frequencies between populations

UK controls versus Sweden controls

Lambda = 1.30 LD score intercept = 1.32





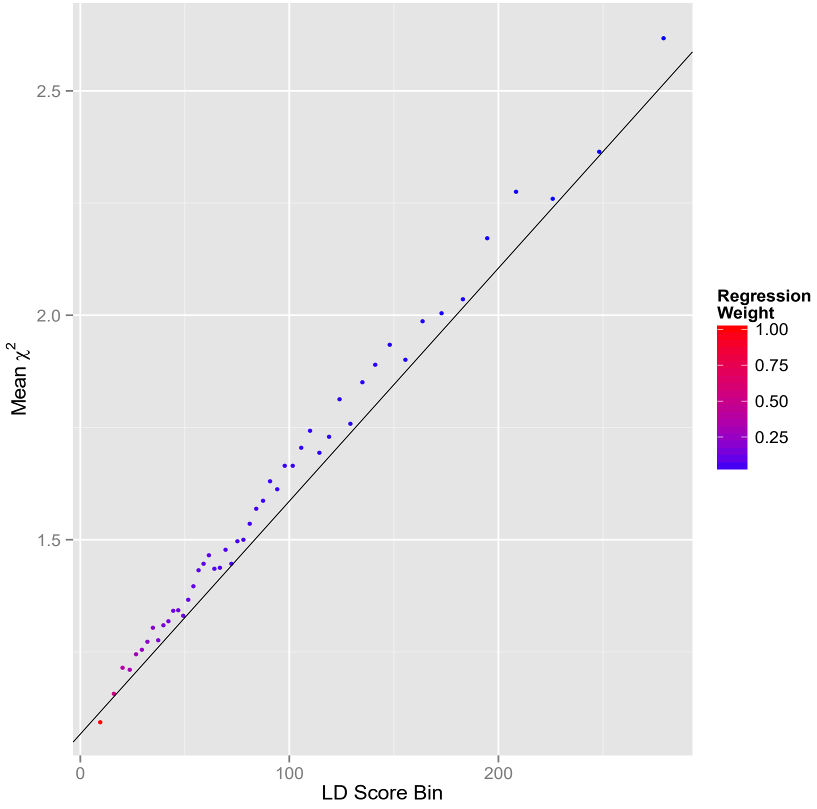
PGC Schizophrenia

Lambda = 1.48

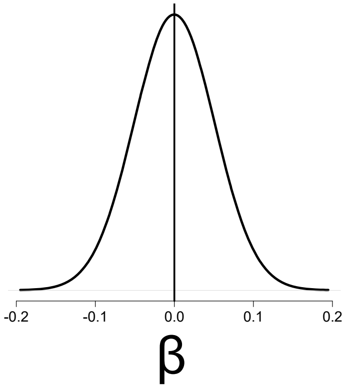
Intercept = 1.06

Slope *p*-value < 10^{-300}

Overwhelming majority of inflation is consistent with polygenic architecture



LD Score regression



Draw polygenic effects from N (0, h^2/m), var =

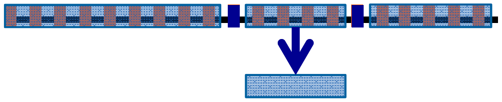
What is the E[χ^2] for variant *j*?

$$E[\chi_j^2] = 1 + Na + \frac{\boxed{h_g^2} N}{M} l_j$$

New estimator of heritability

where N=sample size, M=# of SNPs, a=inflation due to confounding, h^2g is heritability (total obs.) and l_j is the *LD Score*

$$l_j = \sum_{k=j} r_{jk}^2$$



Bulik-Sullivan et al. Nature Genetics 2015
Yang et al. EJHG 2011

Brainstorm Project

Verneri Anttila

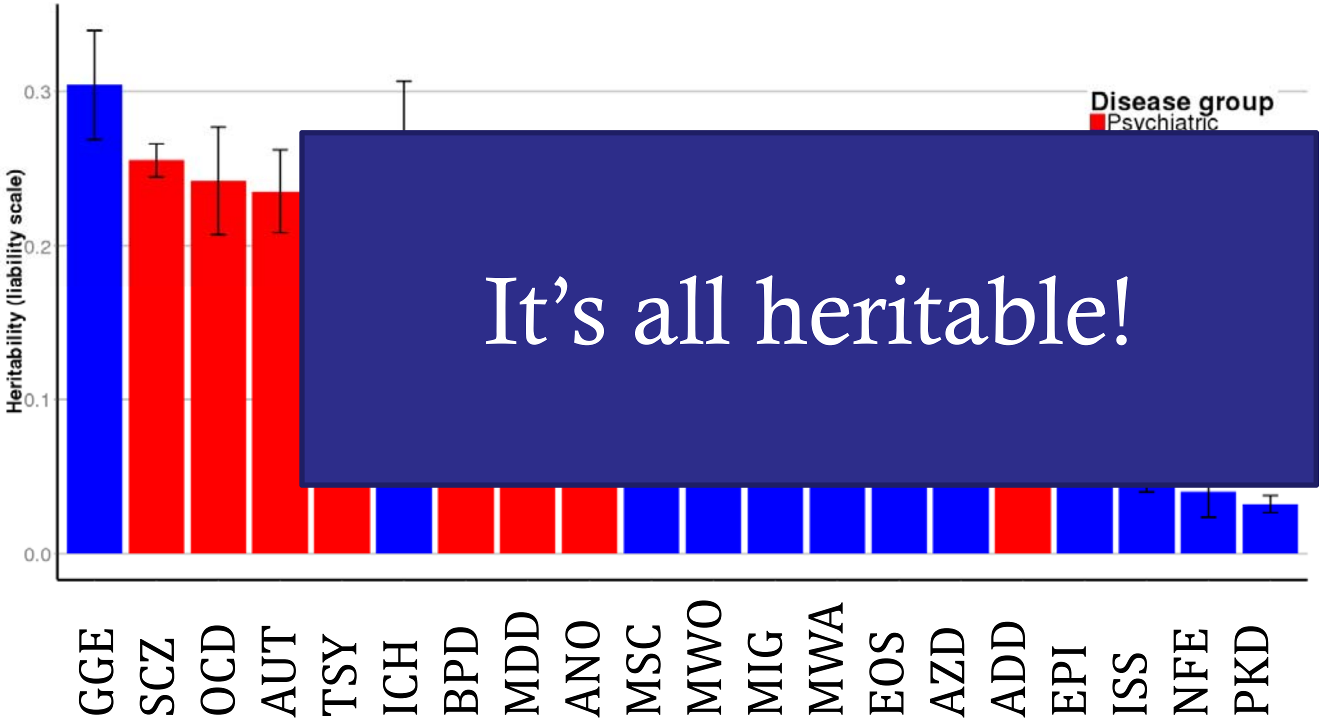
Aiden Corvin

Brendan Bulik-Sullivan
Hilary Finucane
Jonathan Rosand
Aarno Palotie
Mark Daly
Patrick Sullivan
Bobby Koeleman
Nick Wood
Julie Williams
Alessandro Biffi

Jeremiah Scharf
Kenneth Kendler
Stephan Ripke
Alkes Price
Chris Cotsapas
Padhraig Gormley
Zhi Wei
Rainer Malik
Hailiang Huang
Andrea Byrnes

Dongmei Yu
Laramie Duncan
Kai-How Farh
Namrata Gupta
Miriam Raffeld
...and many, many others
in their respective study groups

Univariate heritability from common variation



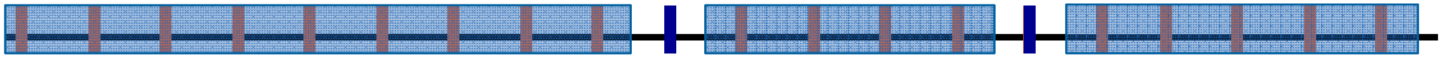
It’s all heritable!

GGE = Generalized Epilepsy
SCZ = Schizophrenia
OCD = Obsessive Compulsive Disorder
AUT = Autism
TSY = Tourette’s Syndrome
ICH = Intracerebral Hemorrhage
BPD = Bipolar Disorder
MDD = Major Depressive Disorder
ANO = Anorexia Nervosa
MSC = Multiple Sclerosis
MWO = Migraine without Aura
MIG = Migraine
MWA = Migraine with Aura
EOS = Early Onset Stroke
AZD = Alzheimer’s Disease
ADD = Attention Deficit/Hyperactivity
EPI = Epilepsy (all)
ISS = Ischemic Stroke
NFE = Non-acquired focal epilepsy
PKD = Parkinson’s Disease

Key extension 1
Functional partitioning

- ▮ Lonely SNPs [no LD]
- ▨ LD blocks
- * Causal variants

- ▮ DHS
- ▮ Coding



LD Score 9 1 4 1 5

Finucane et al. 2015 Nat Gen

$$l_j = \sum_{k \in C} r_{jk}^2$$

Key extension 1
Functional partitioning

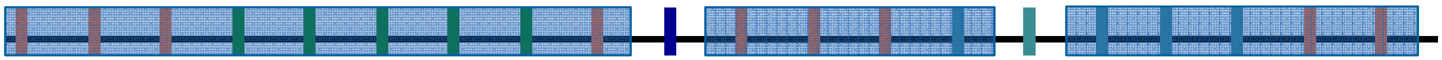
▮ Lonely SNPs [no LD]

▨ LD blocks

* Causal variants

▮ DHS

▮ Coding



LD Score	9	1	4	1	5
DHS Score	5	0	0	0	0
Coding Score	0	0	1	1	3

Finucane et al. 2015 Nat Gen

$$l_j = \sum_{k \in C} r_{jk}^2$$

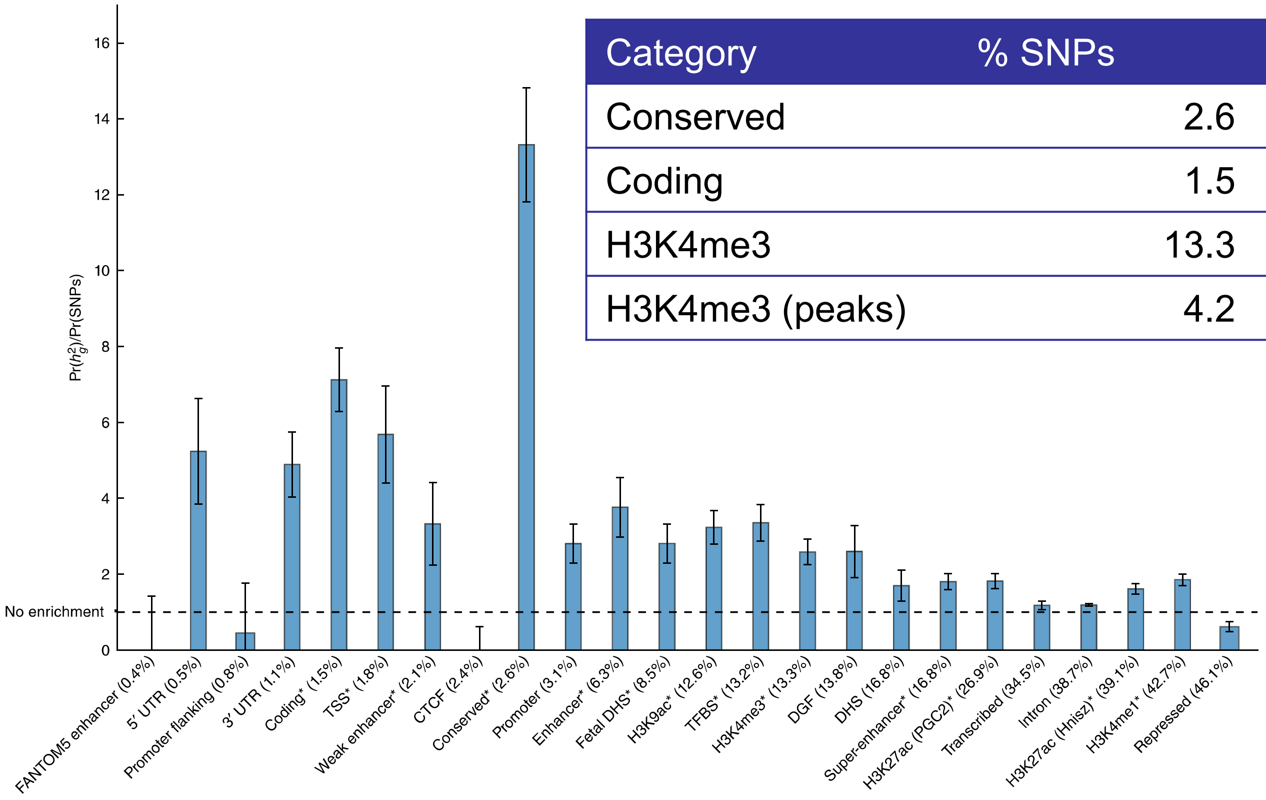
Mark	Source/reference
Coding, 3’ UTR, 5’ UTR, Promoter, Intron	UCSC; Gusev et al., in press AJHG
Digital Genomic Footprint, TFBS	ENCODE; Gusev et al., in press AJHG
CTCF binding site, Promoter Flanking, Repressed, Transcribed, TSS, Enhancer, Weak Enhancer	ENCODE; Hoffman et al., 2012 Nucleic Acids Research
DHS, fetal DHS, H3K4me1, H3K4me3, H3K9ac	Trynka et al., 2013 Nature Genetics.*
Conserved	Lindblad-Toh et al., 2011 Nature
FANTOM5 Enhancer	Andersson et al., 2014 Nature
lincRNAs	Cabili et al., 2011 Genes Dev
DHS and DHS promoter	Maurano et al., 2012 Science
H3K27ac	Roadmap; PGC2 2014 Nature

*Post-processed from ENCODE and Roadmap data by S. Raychaudhuri and X. Liu labs

Datasets for GWAS Selected for a Z>7 for h^2

Phenotype	Reference	Phenotype	Reference
Height	Lango Allen, 2010	Schizophrenia	PGC, 2014
BMI	Speliotes, 2010	Bipolar	Sklar, 2011
Age of menarche	Perry, 2014	Anorexia	Boraska, 2014
LDL	Teslovich, 2010	Education years	Rietveld, 2013
HDL	Teslovich, 2010	Ever smoked	TAG, 2010
Triglycerides	Teslovich, 2010	Rheumatoid Arth	Okada, 2014
CAD	Schunkert, 2011	Crohn's Disease	Jostins, 2012
T2D	Morris, 2012	Ulcerative Colitis	Jostins, 2012
Fasting Glucose	Manning, 2012		

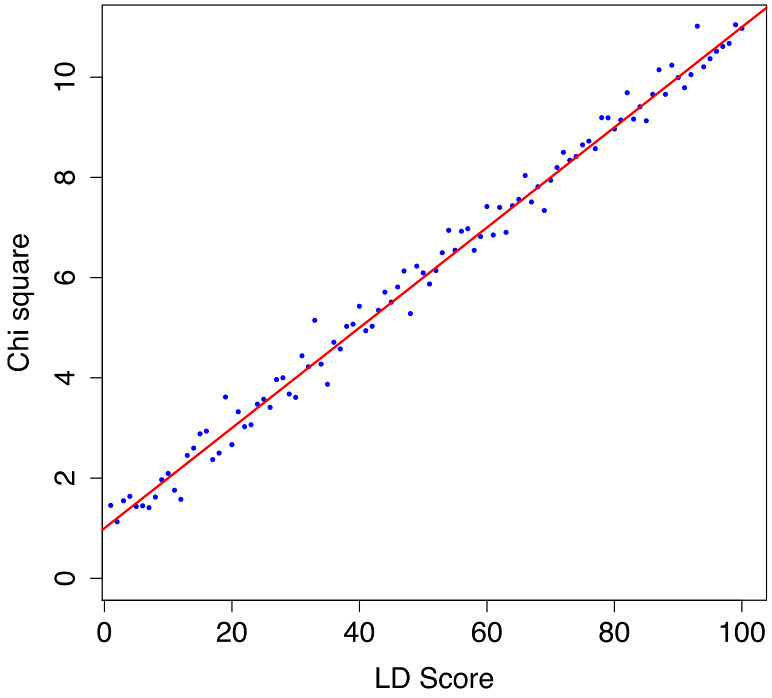
Average enrichments per class
Collapsed results across 17 traits



Category	% SNPs	% h^2	Enrichment
Conserved	2.6	34.7	13.4x
Coding	1.5	10.4	7.0x
H3K4me3	13.3	34.4	2.6x
H3K4me3 (peaks)	4.2	15.8	3.8x

LD Score regression
Genetic correlation

| Trait 1

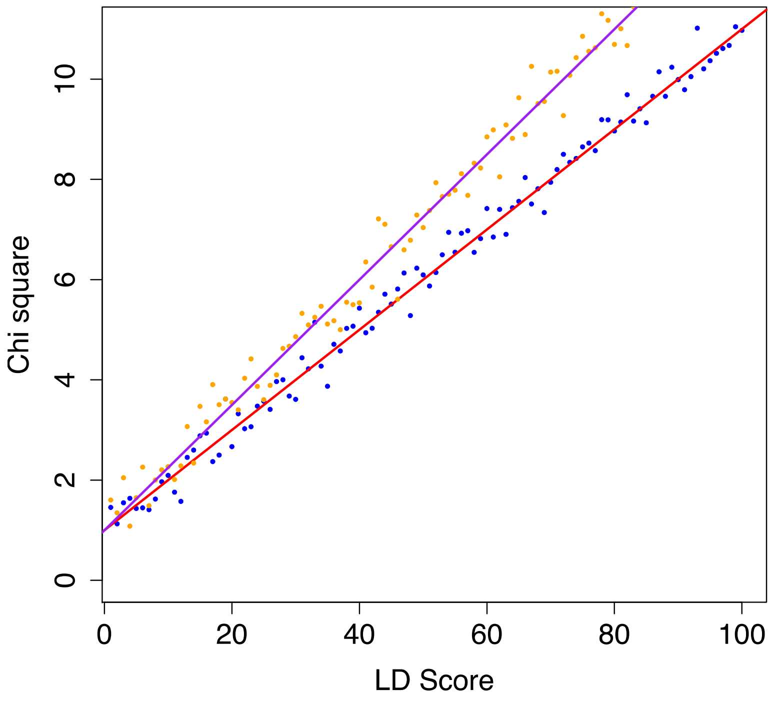


Slope estimates heritability

LD Score regression
Genetic correlation

Trait 1

Trait 2

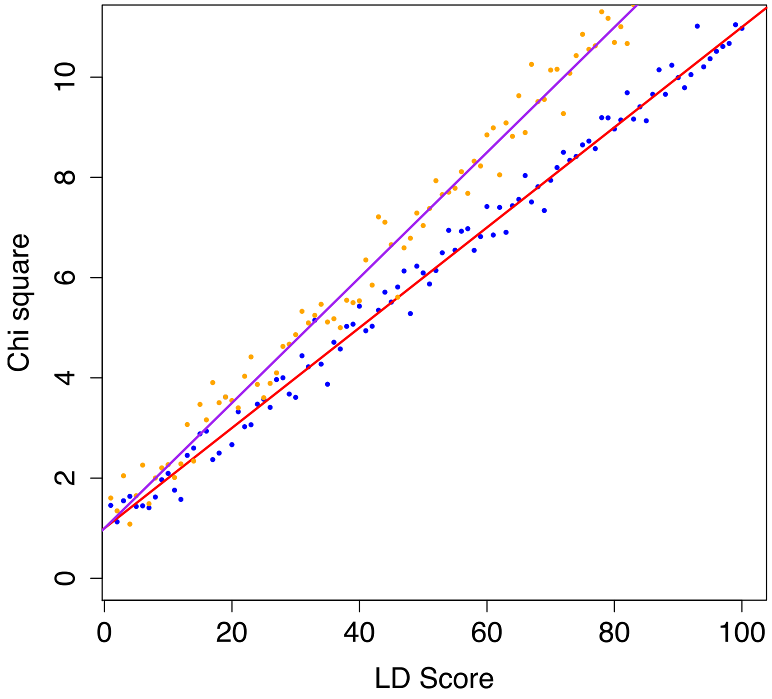


We can a second trait and obtain two heritability estimates

LD Score regression
Genetic correlation

| Trait 1

| Trait 2



$Z*Z = \chi^2$

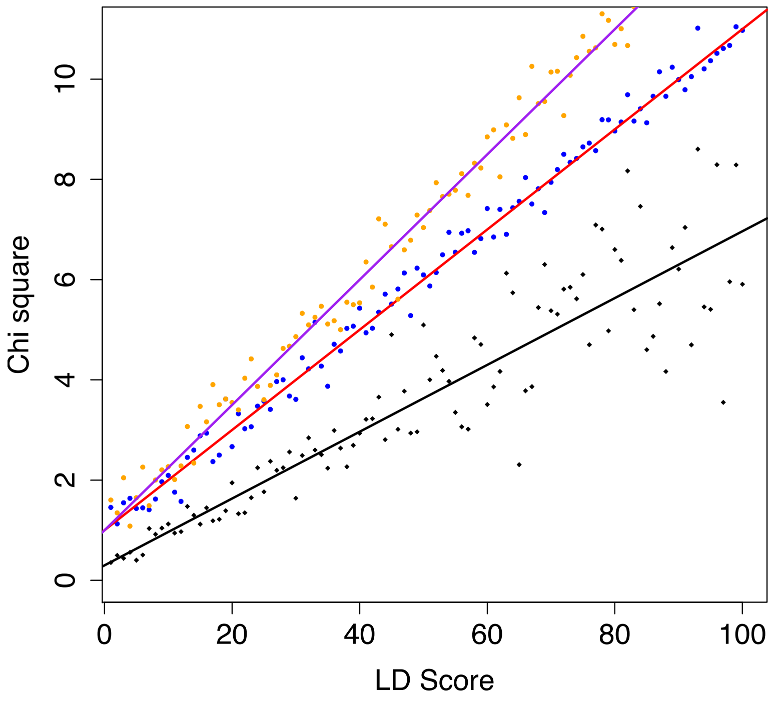
So we can estimate genetic covariance from the product of the Z-scores

LD Score regression
Genetic correlation

Trait 1

Trait 2

R_G



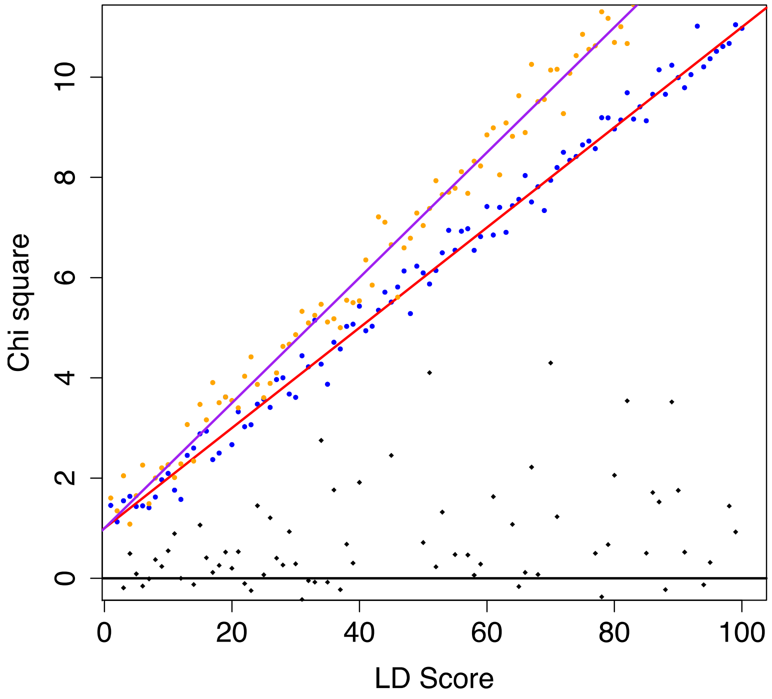
$Z*Z = \chi^2$

So we can estimate genetic covariance from the product of the Z-scores for the two traits

$R_G = 0.5$

LD Score regression
Genetic correlation

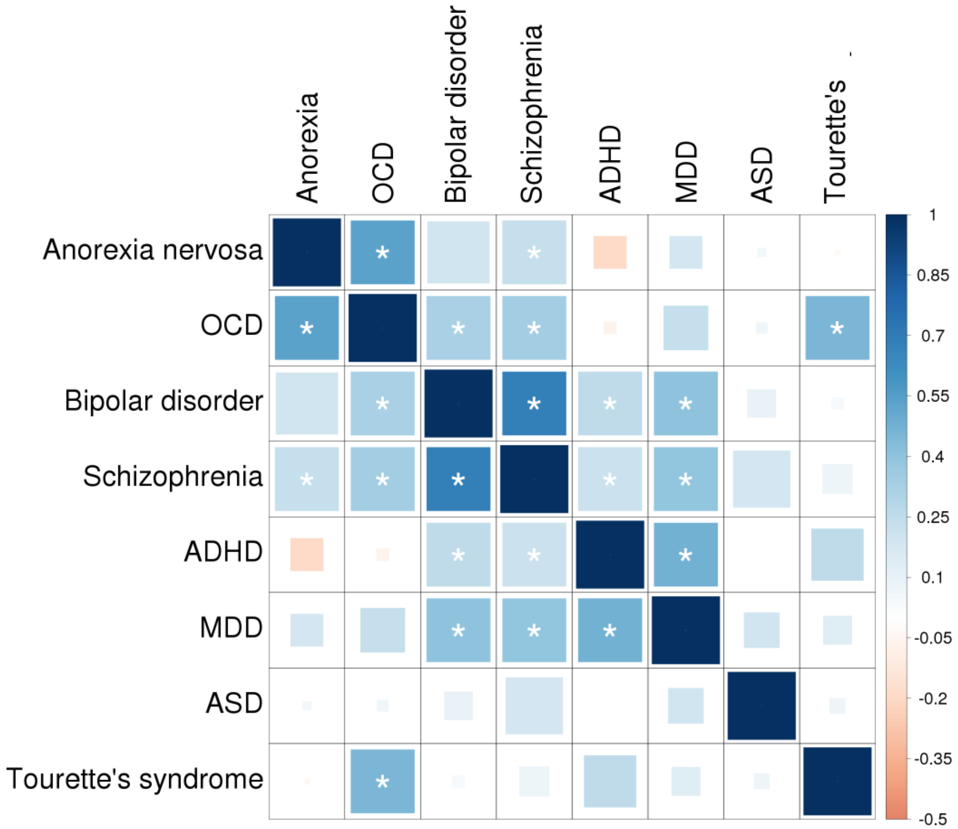
▌Trait 1
▌Trait 2
▌R_G



Here $R_G = 0$

This approach is robust to sample overlap as all variants are equally inflated

Within psychiatry



Broad sharing of common genetic risk factors across psychiatric illness

Risk factors do not respect traditional boundaries

* denotes significant correlation

Conclusions

- Direct estimation of heritability from GWAS is possible
- Heritability is an arbiter of relative importance for classes of variation
- Heritability can be used to elucidate the relationships between different phenotypes