Beyond Traditional GWAS

SISG – Module 15
Dr Loic Yengo
l.yengo@imb.uq.edu.au
Institute for Molecular Bioscience
The University of Queensland

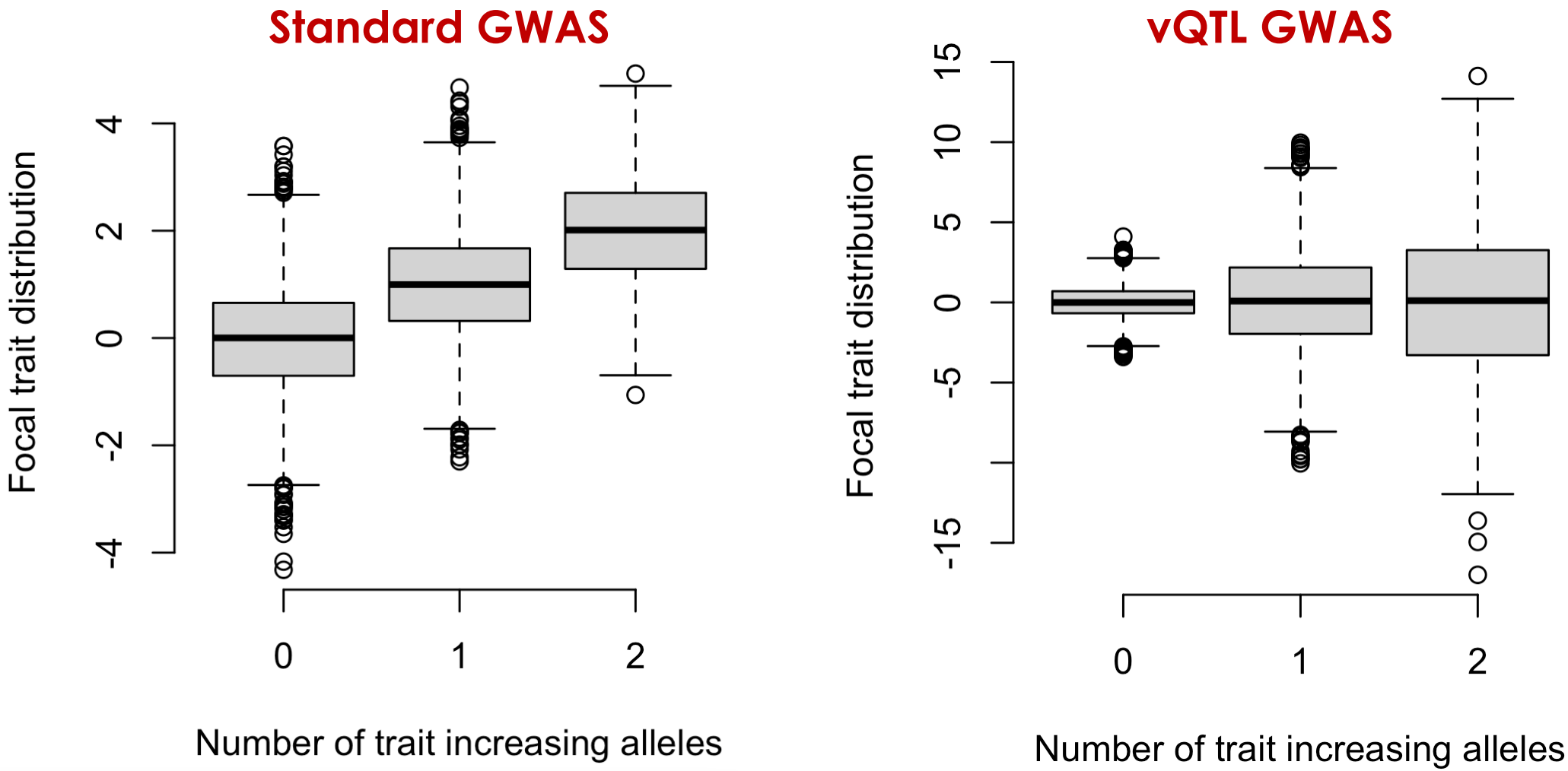
Outline

- Variance QTL
- Interaction Testing
- GWAX
- Time-to-event
- Multi-trait analysis

Outline

- **Variance QTL**
- Interaction Testing
- GWAX
- Time-to-event
- Multi-trait analysis

Variance QTL (vQTL)

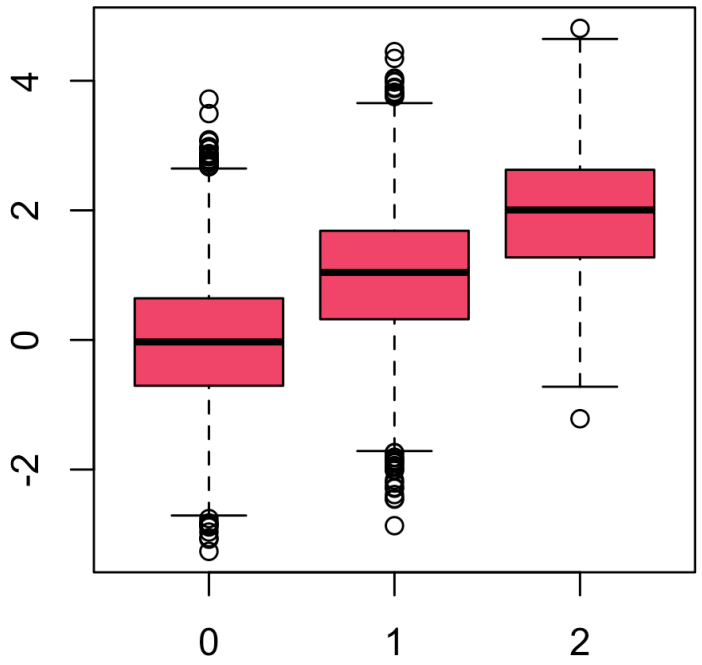


What creates (apparent) vQTL

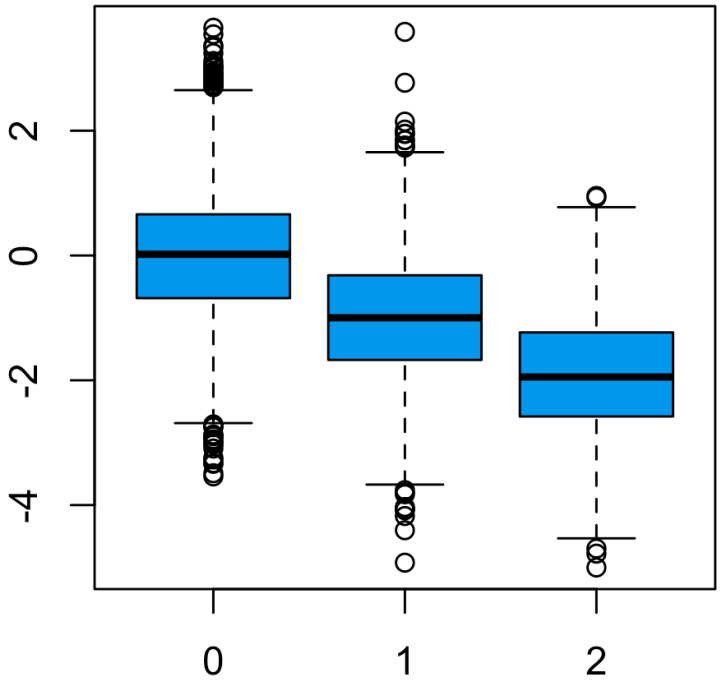
- The scale of the trait (mean-variance) relationship (e.g., skewed distribution)

- Gene-by-Gene or Gene-by-Environment interactions

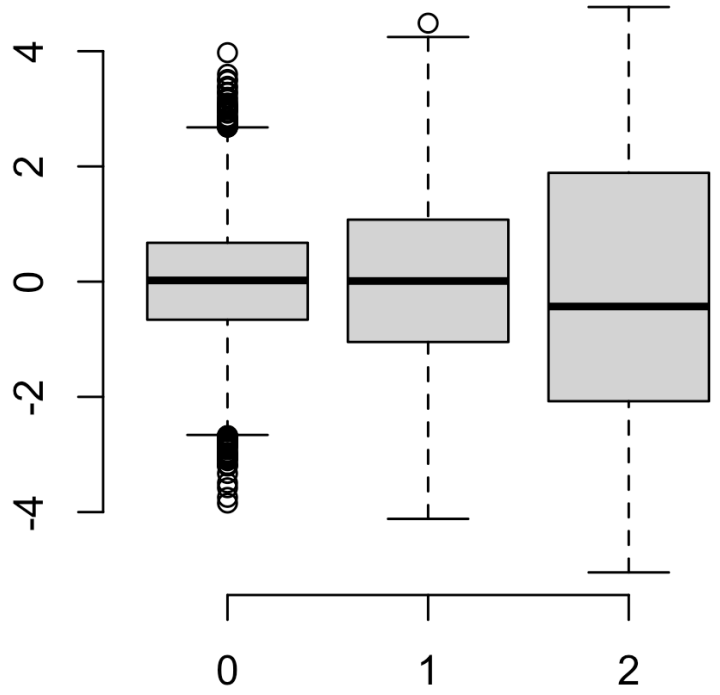
Gene-by-environment interactions



Gene effect in
Environment 1



Gene effect in
Environment 2



Gene effect in
a mixture of individuals
In Environments 1 & 2

Examples

nature

Explore content | About the journal | Publish with us

nature > letters > article

Published: 16 September 2012

FTO genotype is associated with phenotypic variability of body mass index

Jian Yang, Ruth J. F. Loos, ... Peter M. Visscher + Show authors

Nature **490**, 267–272 (2012) | Cite this article

nature genetics

Explore content | About the journal | Publish with us

nature > nature genetics > technical reports > article

Technical Report | Published: 15 October 2018

Identifying loci affecting trait variability and detecting interactions in genome-wide association studies

Alexander I. Young, Fabian L. Wauthier & Peter Donnelly

Nature Genetics **50**, 1608–1614 (2018) | Cite this article

7832 Accesses | **28** Citations | **71** Altmetric | Metrics

Science Advances

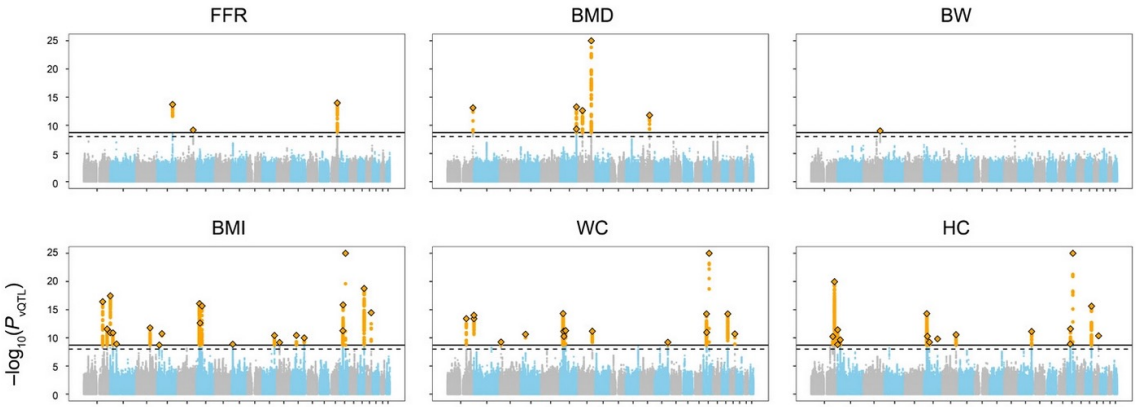
Current Issue | First release papers | Archive | About

HOME > SCIENCE ADVANCES > VOL. 5, NO. 8 > GENOTYPE-BY-ENVIRONMENT INTERACTIONS INFERRED FROM GENETIC EFFECTS ON PHENOTYPIC VARIABILIT…

RESEARCH ARTICLE | HUMAN GENETICS

Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank

HUANWEI WANG, FUTAO ZHANG, JIAN ZENG, YANG WU, KATHRYN E. KEMPER, ANGLI XUE, MIN ZHANG, JOSEPH E. POWELL, MICHAEL E. GODDARD, [...] JIAN YANG +4 authors Authors Info & Affiliations



Testing for vQTL

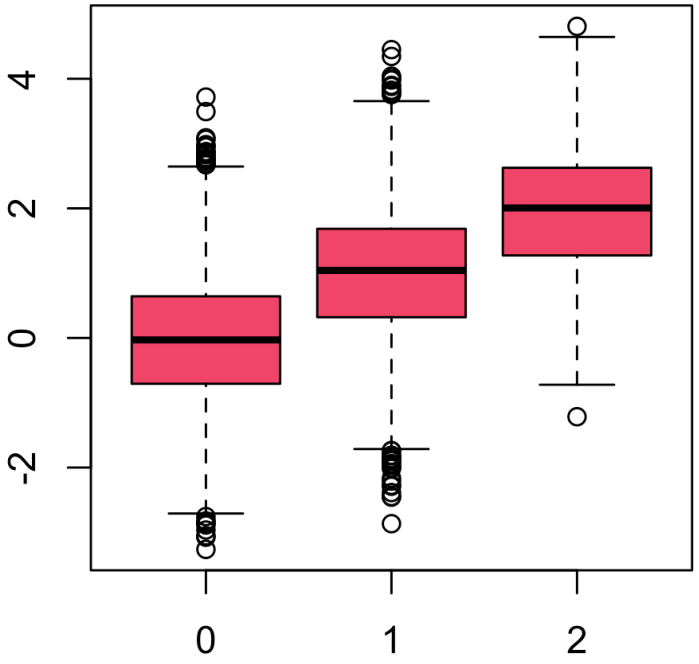
- **Method 1**: Regress $y^2 \sim X$ (Yang 2012) ← No ideal (subject to biases)
- **Method 2**: Levene's Test (Wang 2019) ← Robust! Not flexible for covariates
- **Method 3**: Jointly model mean and variance effect (Young 2018) ← Allows covariates May not be robust to trait distribution assumption

Outline

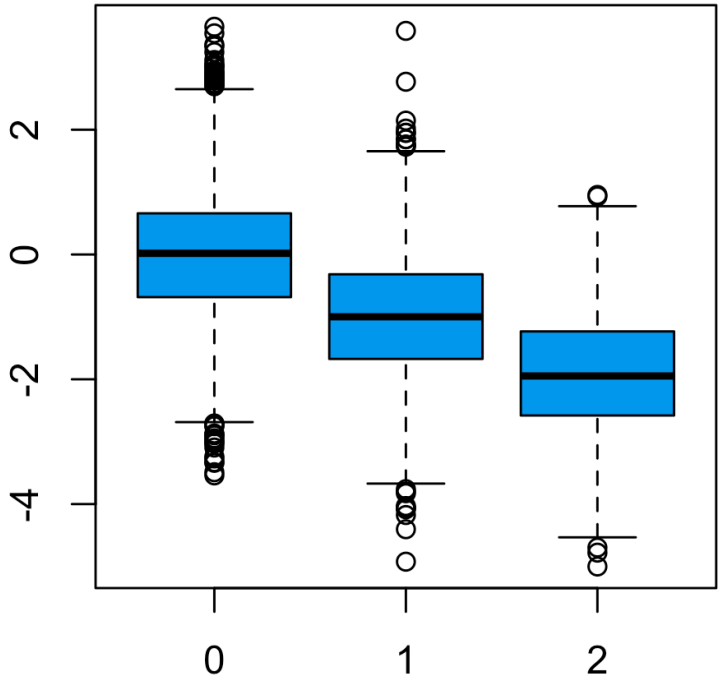
- Variance QTL
- **Interaction Testing**
- GWAX
- Time-to-event
- Multi-trait analysis

Question

Is the effect of SNP1 different in Environment 1 vs Environment 2?



Gene effect in Environment 1



Gene effect in Environment 2

A simple? Interaction test

Y = SNP + Env + **SNP x Env**

True but this can be easily confounded

Recommendation: also adjust for **Covariate x Env**

Review > Biol Psychiatry. 2014 Jan 1;75(1):18-24. doi: 10.1016/j.biopsych.2013.09.006. Epub 2013 Oct 15.

Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution

Matthew C Keller [1]

Affiliations + expand

PMID: 24135711 PMCID: PMC3859520 DOI: 10.1016/j.biopsych.2013.09.006

Free PMC article

Example 1

PLOS GENETICS

Graff et al. (2017)

RESEARCH ARTICLE

Genome-wide physical activity interactions in adiposity – A meta-analysis of 200,452 adults

Mariaelisa Graff[1☯]*, Robert A. Scott[2☯], Anne E. Justice[1☯], Kristin L. Young[1,3☯], Mary

Abstract

Physical activity (PA) may modify the genetic effects that give rise to increased risk of obesity. To identify adiposity loci whose effects are modified by PA, we performed genome-wide interaction meta-analyses of BMI and BMI-adjusted waist circumference and waist-hip ratio from up to 200,452 adults of European (n = 180,423) or other ancestry (n = 20,029). We standardized PA by categorizing it into a dichotomous variable where, on average, 23% of participants were categorized as inactive and 77% as physically active. While we replicate the interaction with PA for the strongest known obesity-risk locus in the *FTO* gene, of which the effect is attenuated by ~30% in physically active individuals compared to inactive individuals, we do not identify additional loci that are sensitive to PA. In additional genome-wide meta-analyses adjusting for PA and interaction with PA, we identify 11 novel adiposity loci, suggesting that accounting for PA or other environmental factors that contribute to variation in adiposity may facilitate gene discovery.

Example 2

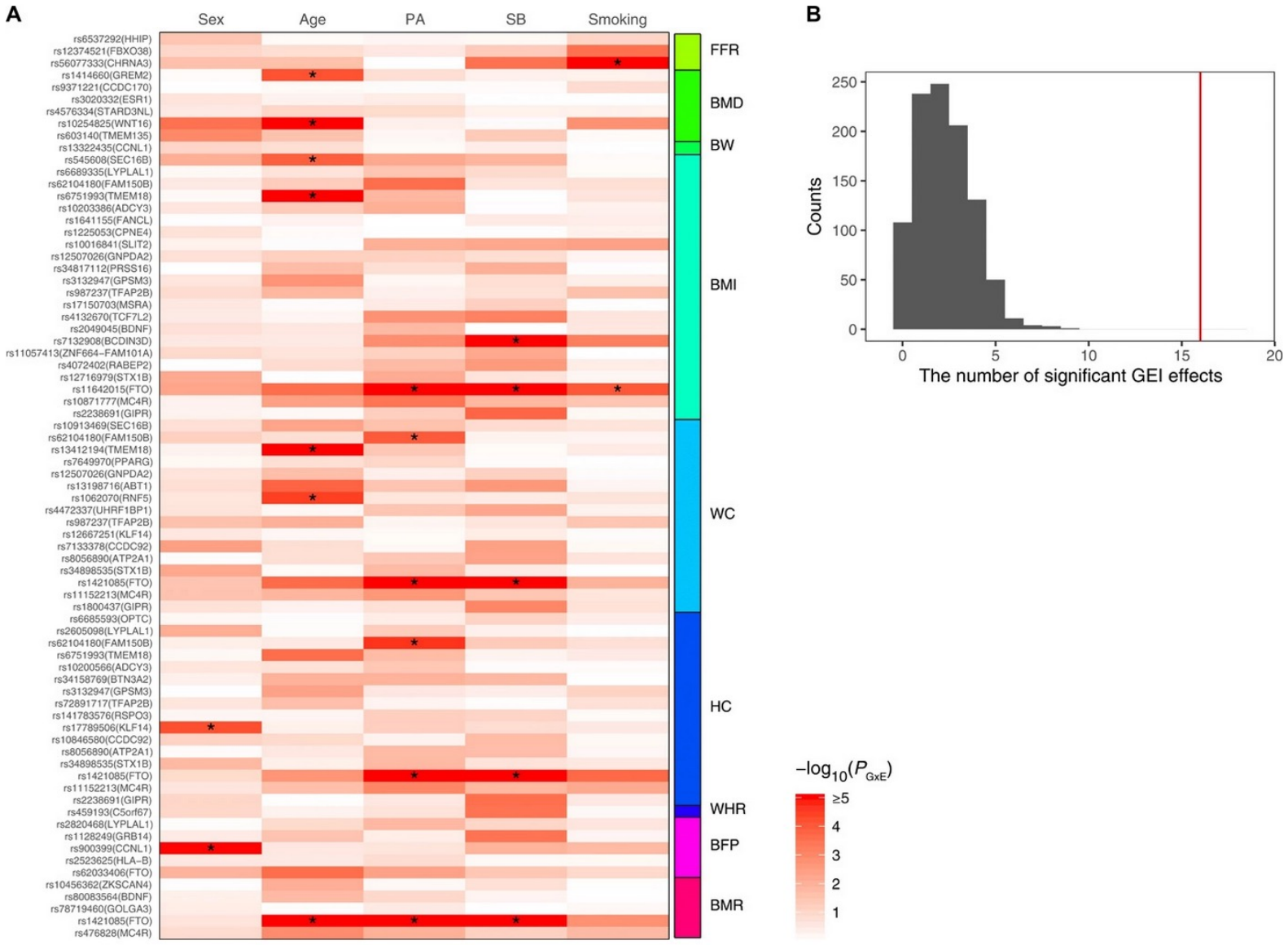


Fig. 5 Enrichment of GEI effects among the 75 vQTLs compared with a random set of QTLs.

Wang et al. (2019)

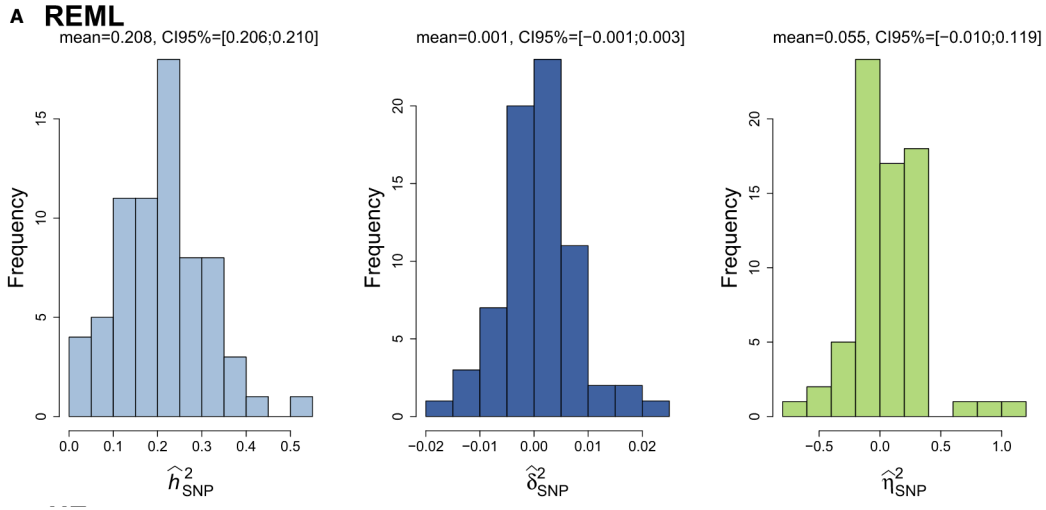
Warning!

Interactions (i.e., nonlinear effects) can be confounded by sample ascertainment (see Practical)

Why is it hard to find GxG interactions?

1) Statistical interaction ≠ Biological/Molecular Interaction

2) Power! GxG (or additive by additive) interaction do not account for a lot of heritability



Hivert et al. (2021)

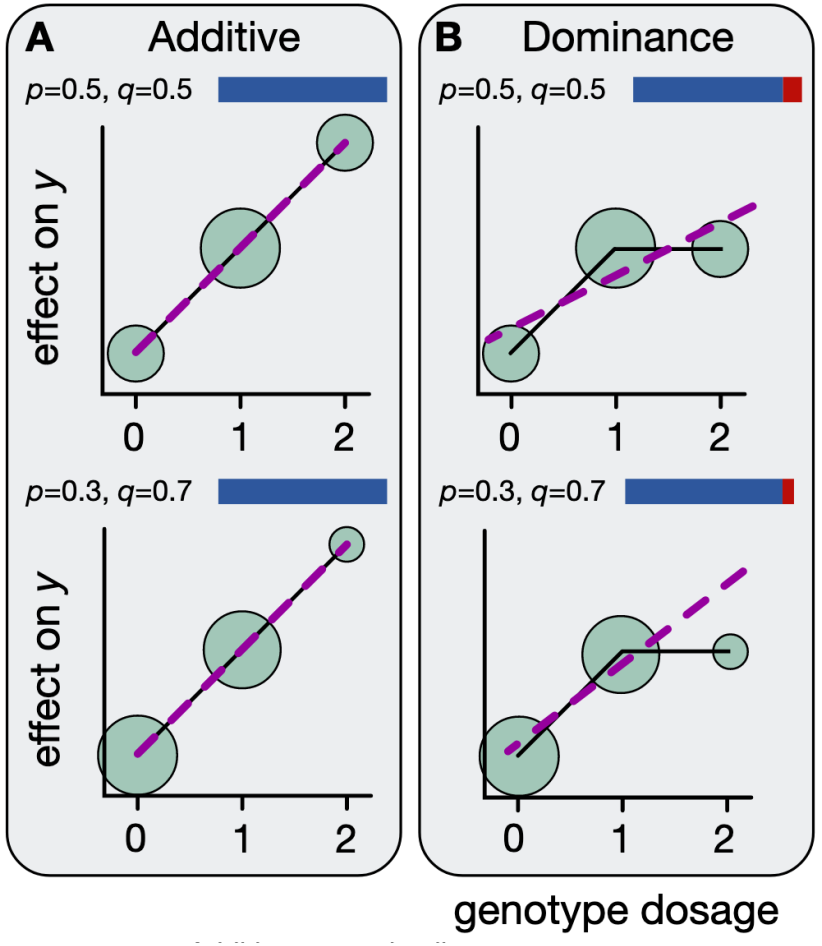
Dominance GWAS

Testing for non-additive effect at a locus

$$y = \mu + \beta_a X + \beta_d X(2 - X) + \varepsilon$$

X={0,1,2}, H=X(2-X) = 1 of heterozygote (0 else)

PLINK: "genotypic" test

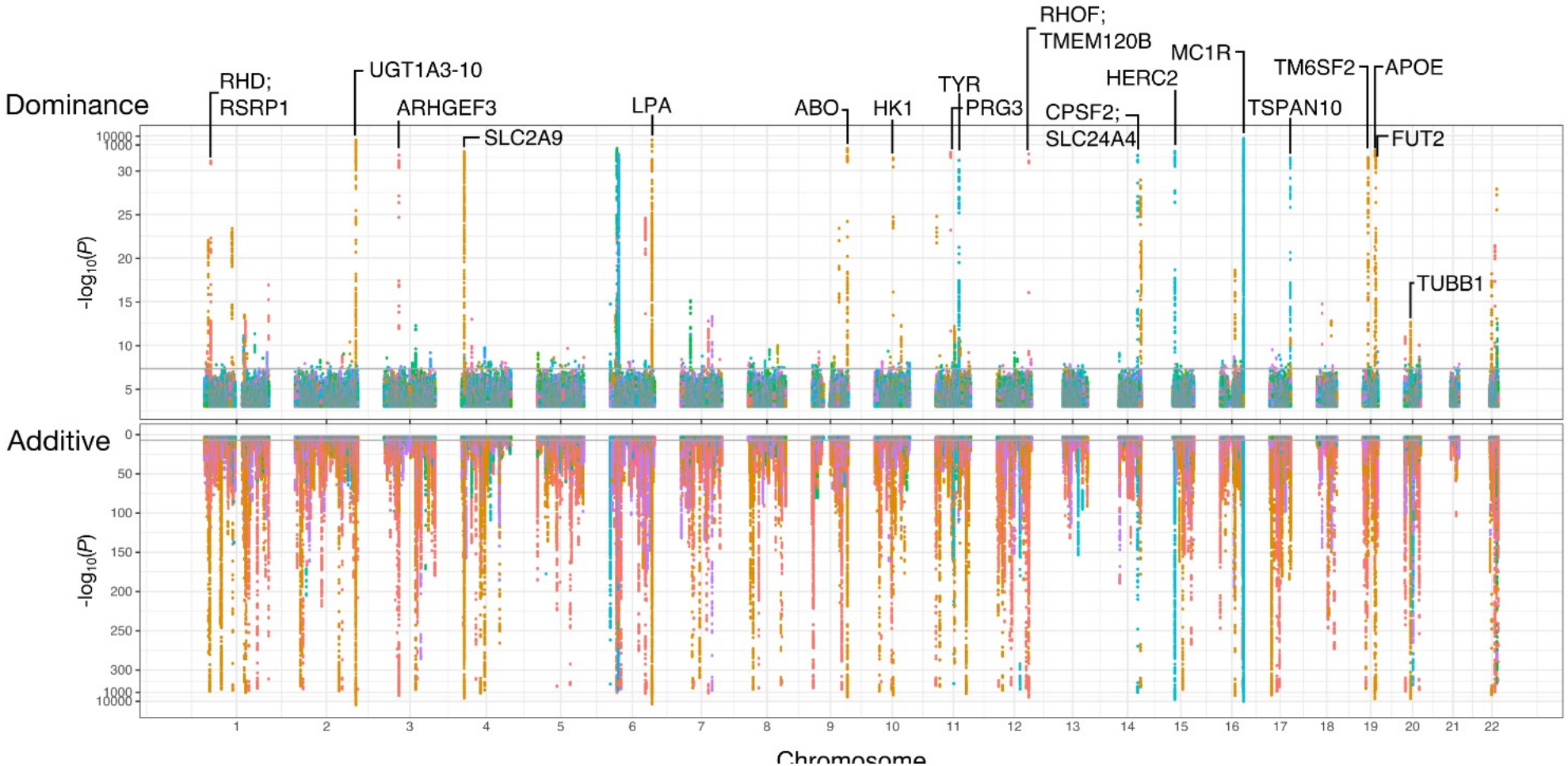


Duncan et al. (2021) - bioRxiv

Analysis of genetic dominance in the UK Biobank

Duncan S. Palmer, Wei Zhou, Liam Abbott, Nikolas Baya, Claire Churchhouse, Cotton Seed, Tim Poterba, Daniel King, Masahiro Kanai, Alex Bloemendal, Benjamin M. Neale

doi: https://doi.org/10.1101/2021.08.15.456387



Outline

- Variance QTL
- Interaction Testing
- **GWAX**
- Time-to-event
- Multi-trait analysis

Rationale

- Test for association using genotypes from on set of individuals and phenotypes from their close relatives (e.g., parents)

- Useful if genotyped individuals are relatively young and interest is in late onset diseases (e.g., Alzheimer's disease)

nature genetics

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › nature genetics › analyses › article

Published: 16 January 2017

Case–control association mapping by proxy using family history of disease

Jimmy Z Liu ✉, Yaniv Erlich & Joseph K Pickrell

Integrated analysis of direct and proxy genome wide association studies highlights polygenicity of Alzheimer's disease outside of the APOE region

Javier de la Fuente ✉, Andrew D. Grotzinger, Riccardo E. Marioni, Michel G. Nivard, Elliot M. Tucker-Drob ✉

A simple idea…but

Effect sizes from GWAX and that of GWAS do NOT have the same expectation!

Therefore, fixed-effect meta-analysis will be biased

Sample-size weighted is a better option.

Outline

- Variance QTL
- Interaction Testing
- GWAX
- **Time-to-event (e.g., age of onset)**
- Multi-trait analysis

Rationale

- Test is genotype is associated with onset
- Very active area of research
- Open question: genetic correlation between onset / severity / susceptibility

Genetics. 2020 May;215(1):41-58. doi: 10.1534/genetics.119.302940. Epub 2020 Mar 4.

Fast Algorithms for Conducting Large-Scale GWAS of Age-at-Onset Traits Using Cox Mixed-Effects Models

Liang He [1], Alexander M Kulminski [1]

nature communications

Article | Open Access | Published: 20 April 2021

Genomic architecture and prediction of censored time-to-event phenotypes with a Bayesian genome-wide analysis

Sven E. Ojavee, Athanasios Kousathanas, Daniel Trejo Banos, Etienne J. Orliac, Marion Patxot, Kristi Läll, Reedik Mägi, Krista Fischer, Zoltan Kutalik & Matthew R. Robinson

An efficient and accurate frailty model approach for genome-wide survival association analysis controlling for population structure and relatedness in large-scale biobanks

Rounak Dey, Wei Zhou, Tuomo Kiiskinen, Aki Havulinna, Amanda Elliott, Juha Karjalainen, Mitja Kurki, Ashley Qin, FinnGen, Seunggeun Lee, Aarno Palotie, Benjamin Neale, Mark Daly, Xihong Lin

doi: https://doi.org/10.1101/2020.10.31.358234

Outline

- Variance QTL
- Interaction Testing
- GWAX
- Time-to-event
- **Multi-trait analysis**

Intuition

If traits share the same causal variants, then analyzing them jointly should improve power (recall COLOC)

nature genetics

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › nature genetics › articles › article

Article | Published: 01 January 2018

Multi-trait analysis of genome-wide association summary statistics using MTAG

Patrick Turley ✉, Raymond K. Walters, Omeed Maghzian, Aysu Okbay, James J. Lee, Mark Alan Fontana, Tuan Anh Nguyen-Viet, Robbee Wedow, Meghan Zacher, Nicholas A. Furlotte, 23andMe Research Team, Social Science Genetic Association Consortium, Patrik Magnusson, Sven Oskarsson, Magnus Johannesson, Peter M. Visscher, David Laibson, David Cesarini ✉, Benjamin M. Neale ✉ & Daniel J. Benjamin ✉

Nature Genetics **50**, 229–237 (2018) | Cite this article

Multi-traits GWAS – a few remarks

- **Clarification**: The (alternative) hypothesis is NOT that a single SNP is associated with at least one trait (just too many degrees of freedom) – Hotelling test

- MTAG works better when traits are highly genetically correlated

- **Caveats**: Genome-wide genetic correlation can be different from local genetic correlation (false positives)

Family-based GWAS

Robust to population stratification

Transmission Disequilibrium Test (TDT)

Spielman, McGinnis and Ewens (1993)

Input: Trio* (Mother + Father + affected Offspring) design

Question: is the "risk" allele more often transmitted to the affected child?

*There are extension beyond trios

Within-family GWAS

Question: are family more with more "risk" allele than their relatives more susceptible to the disease?

$$y = \mu + \beta_{SNP}X + \beta_F\bar{X} + \varepsilon$$

Individual Allele Count

Average allele Count in the family

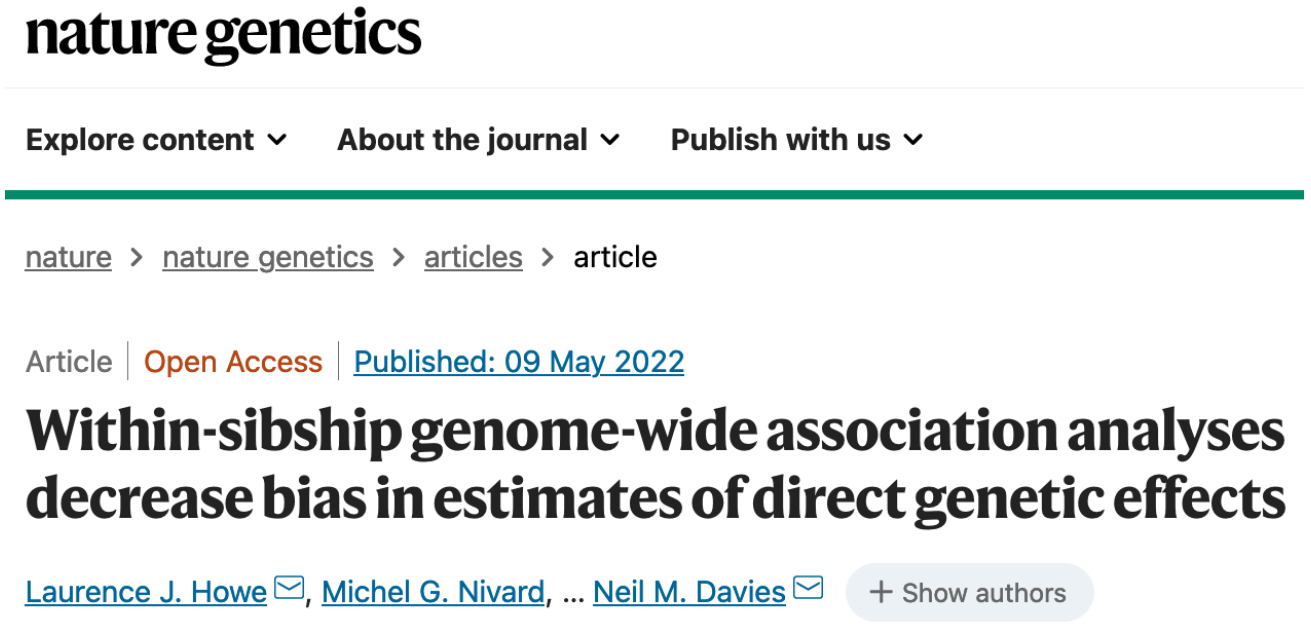
Captures indirect genetic effects

Kong et al.
Young et al.
(Genetic Nurture)

Special case: Sibling-GWAS

Equivalent Model when only sibling-pairs available

$$(y_1 - y_2) = \beta_{SNP}(X_1 - X_2) + \varepsilon$$



Effect sizes are also not inflated by assortative mating

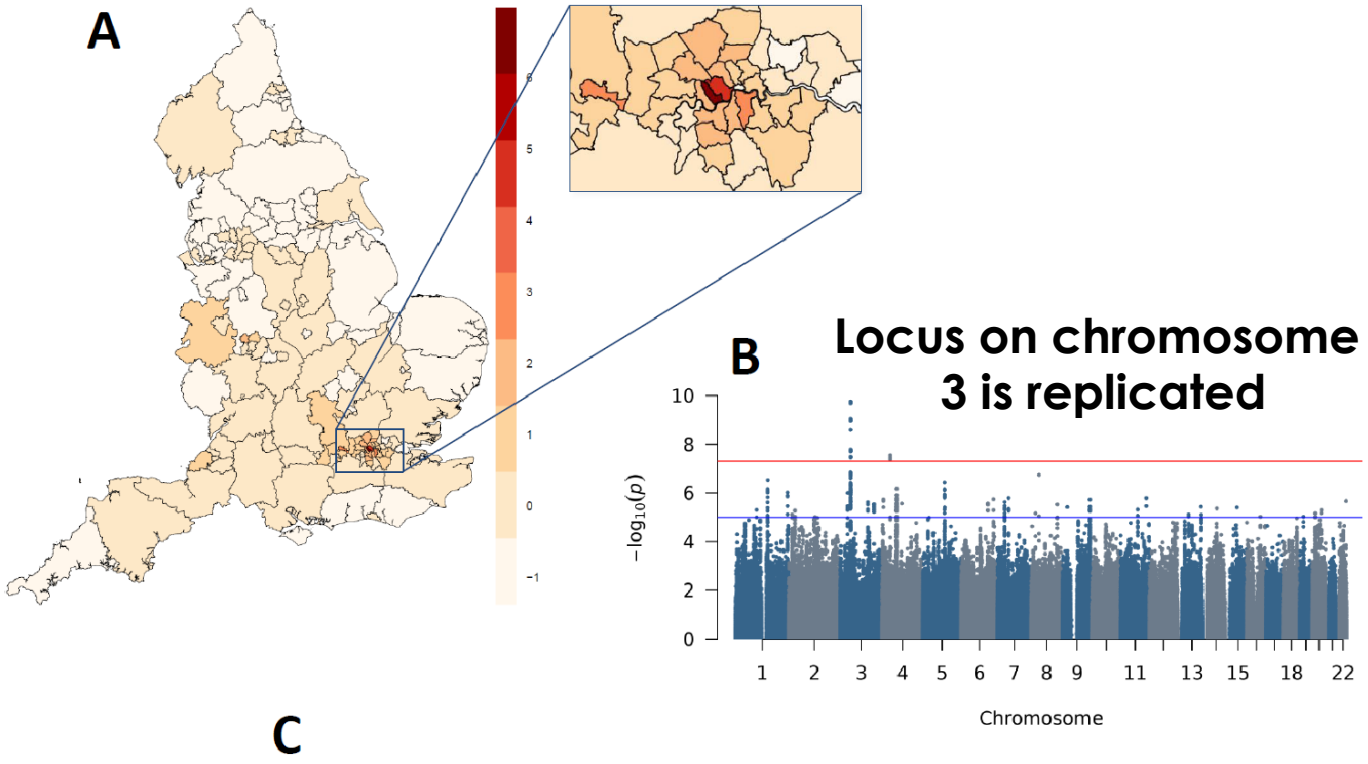
Other *funky* approaches

Regional GWAS?

Association between allele frequency in a region (or a group of individual) and an average/group phenotype

Appealing because many datasets are aggregated (Census / National Stats)

Can we use that for *GWAS'ing* COVID-19?



Abdellaoui 2020 (medRxiv)

Limitations

Causality is hard to establish
(e.g., Migration can change allele frequencies)

Autosomal GWAS of Sex

Is autosomal allele frequency different between males and females?

Other potential explanation
"the autosomal genotype array probe cross-hybridizes with a sex chromosome sequence." – ruled out by Piratsu et al.

Summary and conclusions

Any statistical test can be performed “genome-wide” => may result in an usual GWAS [a) coupling with mixed model – stratification / relatedness; b) computational burden]

Nonlinear effects are hard to detect: very large sample needed + lots of caveats (e.g., scale, ascertainment)

Phenotypes of relatives can be used to improve power (late/early onset diseases)

Future GWAS

- Phenotypes are becoming more complex and multi-dimensional: single cell gene expression, fMRI, ICD10 codes (tree structure)

- Future GWAS should account for time (longitudinal + Cox: onset / severity of symptoms) + shared causal variants across traits)