

Detection and replication of epistasis influencing transcription in humans

Joseph Powell and Gibran Hemani

The University of Queensland Diamantina Institute
and
Queensland Brain Institute

11th Oct 2013

What is epistasis?

Definition

The effect on the phenotype caused by locus A depends on the genotype at locus B....

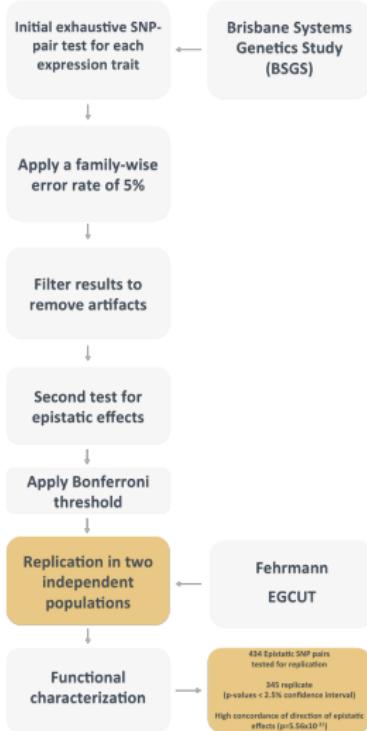
Epistasis has been reported in model organisms through artificial gene knockouts, artificial line crosses and hybridisation.

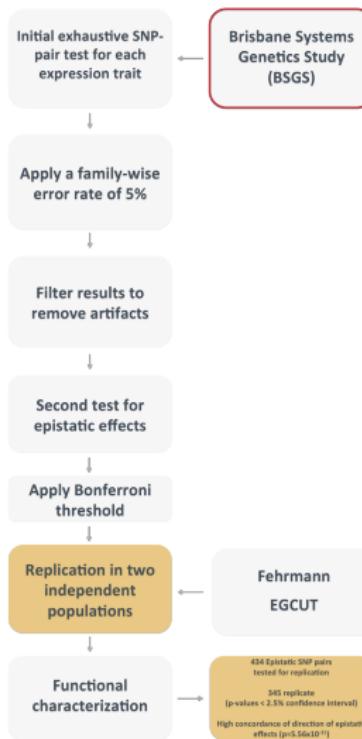
Our aim was to systematically search for instances of epistasis amongst common variants for genetic variation that has arisen in natural populations.

Gene expression

- Transcription measured for thousands of genes
- Typically heritable
- Loci (eQTLs) commonly have very large effect sizes
- Good candidates to search for epistasis

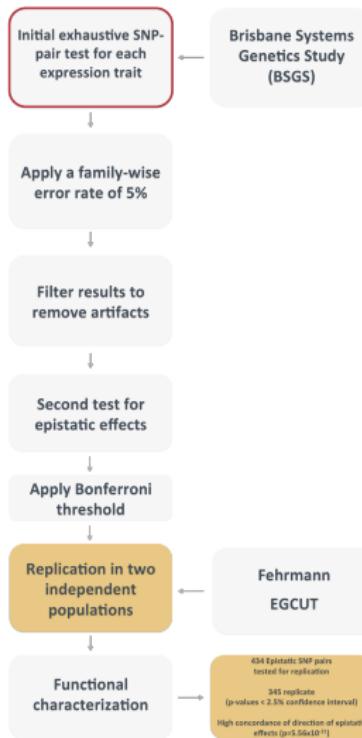






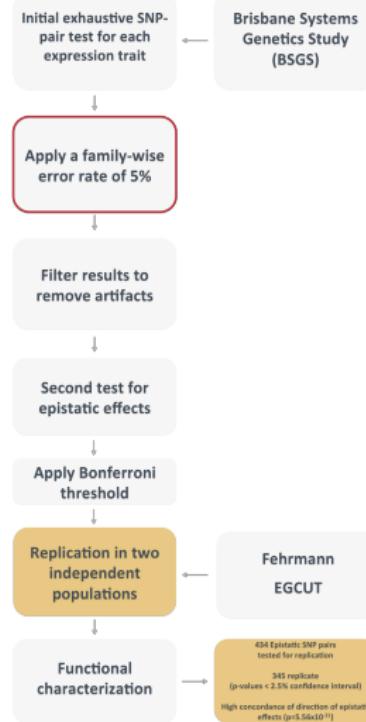
Brisbane Systems Genetics Study (BSGS)

- 846 healthy individuals
 - 528,509 autosomal SNPs
 - RNA measured for peripheral blood (Illumina HT-12v4.0)
 - Expression levels for 7,339 probes



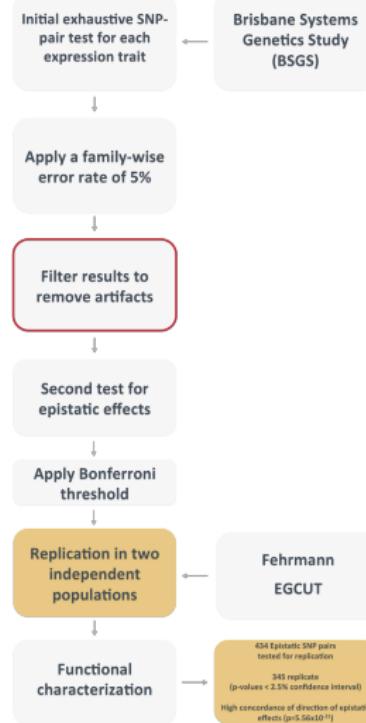
Initial exhaustive SNP-pair test for each expression trait

- Exhaustive SNP x SNP testing for each probe
- epiGPU software and GPU clusters
- 8 d.f. F-test
- Over quadrillion tests



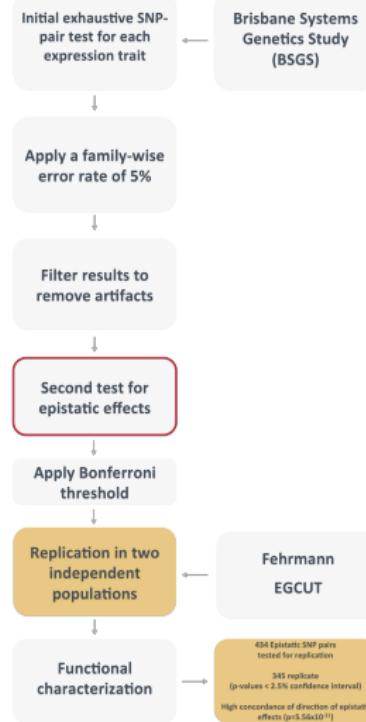
Apply a family-wise
error rate of 5%

- Permutation and bonferroni used to give 5% FWER
- Permutation: Single probe FWER; $T^* = 2.13 \times 10^{-12}$
- Correct for 7,339 probes
- Experiment wide FWER;
 $T^* / 7,339 = 2.91 \times 10^{-16}$



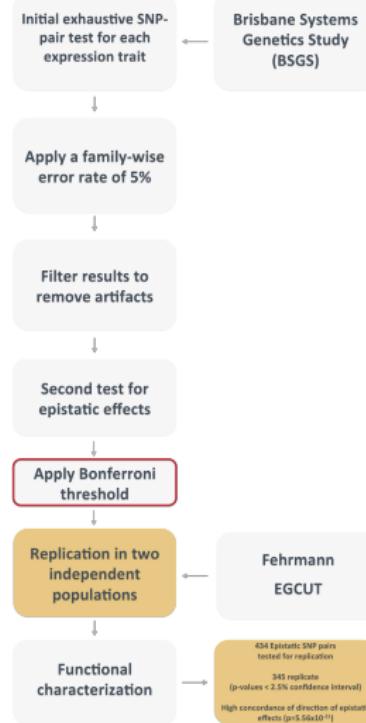
Filter results to remove artifacts

- For SNP pairs that passed the 2.91×10^{-16} threshold
- All 9 genotypes classes present
- Minimum class size of 5
- No LD between SNP pairs ($r^2 < 0.1$ and $D'^2 < 0.1$)
- No single loci additive or dominance effects
- 11,155 pairs carried forward



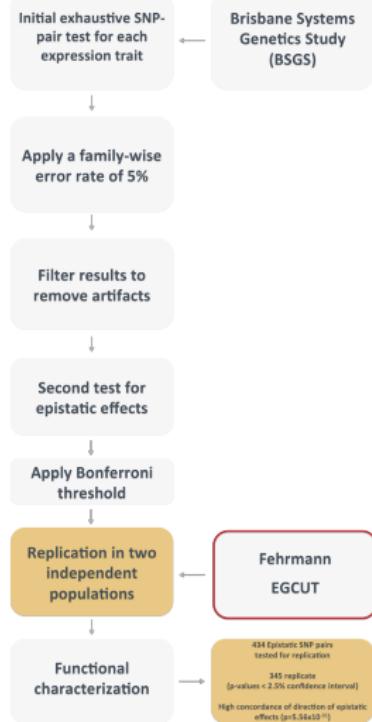
Second test for epistatic effects

- Nested ANOVA for 11,155 pairs
- Contrasts full genetic (8 d.f.) vs marginal effects (4 d.f)
- Thus, testing for the contribution of epistatic variance



Apply Bonferroni threshold

- Epistatic effects significant at $p < 0.05/11,155 = 4.48 \times 10^{-6}$
- 501 SNP pairs with significant interaction terms



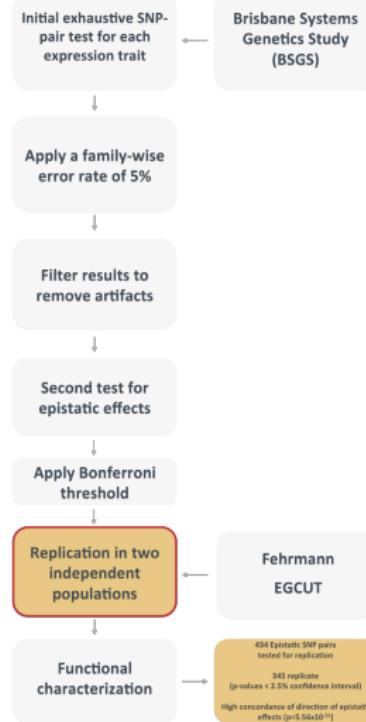
Fehrman EGCUT

Fehrman

- 1,240 individuals
(Netherlands)

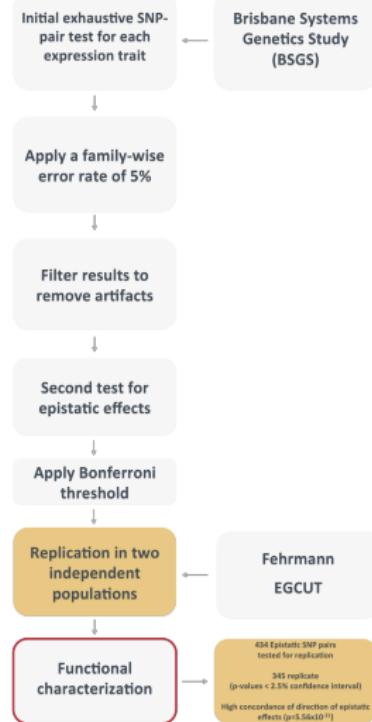
EGCUT

- 891 individuals (Estonian)
- RNA measured for peripheral blood (Illumina HT-12v3.0)
- Genotyped with Illumina arrays



Replication in two independent populations

- Replication of 501 epistatic hits using two independent cohorts
- Same filtering criteria
- 434 pairs which could be tested in both cohorts



Functional characterization

Analyses to elucidate possible functional mechanisms

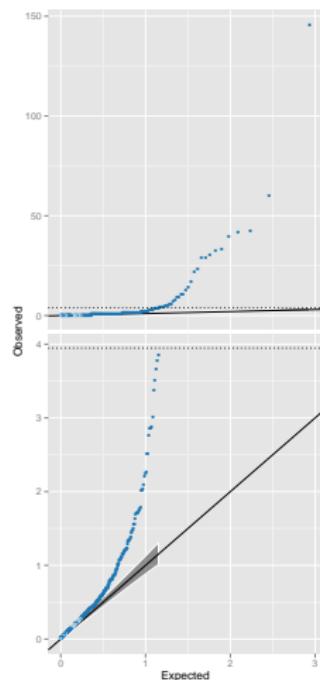
- Non-synonymous mutations
- GWAS and known eQTL overlap
- Genome segmentation
- Chromosome interactions
- Tissue specific transcription regions

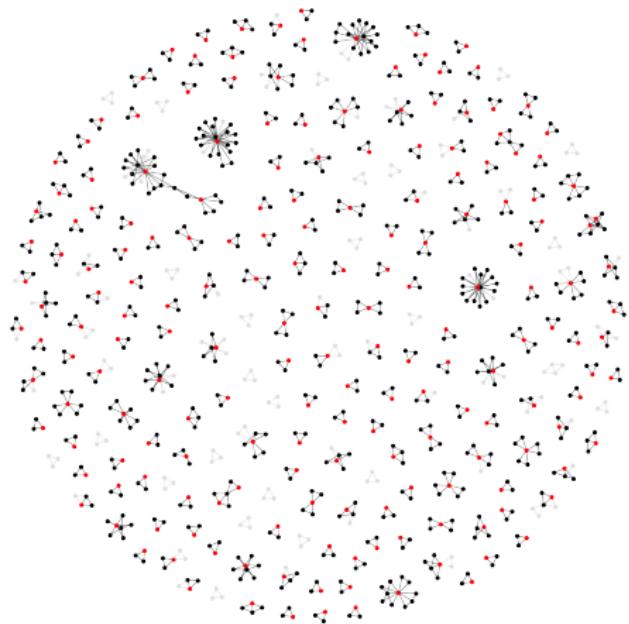
Q-Q plots of interaction p-values from replication

The top panel shows all 434 discovery SNPs that were tested for interactions.

The bottom panel shows the same data as the top panel but excluding the 30 most significant interactions

Matched direction of epistatic effects $p = 5.56 \times 10^{-31}$





Discovery and replication of epistatic networks. All 434 putative genetic interactions.

Black nodes represent SNPs and red nodes represent expression probes.

345 (80%) interactions had p -values exceeding the 2.5% confidence interval from replication data. The remaining interactions are depicted in grey.

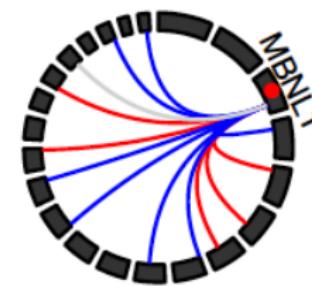
MBNL1

Involved in RNA modification
and regulation of splicing

Has one cis-effect, controlled by
13-trans SNPs

Each pair is independent

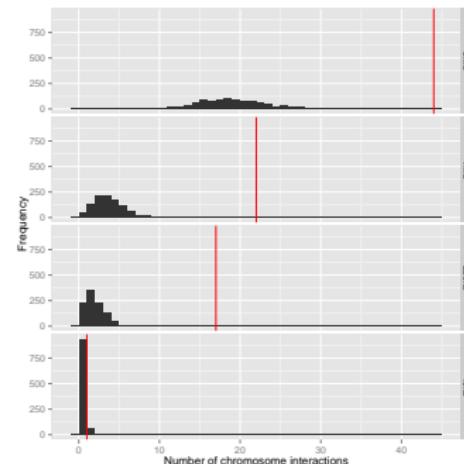
Masking pattern: when the
trans-SNP is homozygous for the
masking allele the decreasing
allele of the cis-SNP no longer
has an effect



Chromosome interactions

Are epistatic SNPs located in positions known to interact?

- Chromosome interactions identified in K562 cell lines using Hi-C and ChIP-seq
 - red lines = N SNPs within 20kb, 500kb, 2Mb and 10Mb of known interacting regions
 - Histograms = null distributions



Conclusions

This study presents the first evidence for multiple instances of segregating common polymorphisms interacting to influence human traits.

Computational and statistical frameworks can identify and replicate epistasis in complex traits.

Functional characterization of epistatic loci suggests a large number of possible mechanisms that can lead to non-additive genetic variation.

Acknowledgements

Complex Trait Genomics Group

- Gib Hemani
- Kostya Shakhbazov
- Jian Yang
- Allan Mcrae
- Peter Visscher

University of Groningen

- Harm-Jan Westra
- Lude Franke

QIMR Berghofer

- Nick Martin
- Anjali Henders
- Grant Montgomery

Georgia Institute of Technology

- Greg Gibson

University of Tartu

- Andres Metspalu
- Tonus Esko

Visit us at www.complextraitgenomics.com