Response to review comments

1 Reviewer 1

Hemani et al. have performed a complete genome-scan for pair-wise epistatic eQTL effects using expression data for 7400 transcripts assayed with microarrays and RNA extracted from PBMCs of 850 individuals genotyped for 500,000 SNPs. After filtering on LD (within and between SNP pairs) the authors report 501 SNPs pairs that are significantly (after Bonferroni correction) influencing the expression level of a target gene. The authors show that - for 30 of 434/501 replicatable SNP pairs - the epistatic component of the effect remains significant (after Bonferroni correction) in a confirmation cohort comprising expression data obtained with the same array on PBMCs of > 2,000 individuals. Moreover, the ranked p-values of the top 316/404 exceed the 97.5% confidence surface occupied by QQ-plots under the null hypothesis, supporting the veracity of the majority of the corresponding epistatic effects. The majority of interacting SNP pairs include a cis- and a trans-acting SNPs. Several instances are described where multiple trans-SNPs modify the effect of a cis-SNP. The authors show that SNPs in high LD with the interacting cis-SNPs (but not trans-SNPs), tend to be located in transcriptionally active chromatin regions, particularly in haematopoeitic cell lines. Moreover, the authors show that interacting SNPs tend to be located in the vicinity of chromosome segments shown to physically interact in the nucleus of K562 blood cells. The authors make an attempt to quantify the relative importance of additive vs epistatic effects in determining gene expression level. In conclusion, this work provides strong evidence that pair-wise epistasis contributes significantly to variation in transcript levels, encouraging similar efforts applied to other complex traits including diseases.

The topic of epistasis is of great interest to the field of complex trait genetics. The paper is in general clearly and succinctly written. The methodology is sound. The results are convincing.

I only have minor comments:

1. Abstract: 2.5% confidence interval of the distribution under the null hypothesis ... I dont think that this is what the authors mean. Should be rephrased here and throughput the text.

2. Abstract: ... within 2MB of regions ... Some contradictions here and throughout the text. Is it 2 or 2.5? According to Suppl. Fig. 11, its should be 2.5Mb

3. Main text, page 3: ... 316 of the remaining 404 discovered SNPs ... come

confusiuon here and in other places. These are 404 SNP pairs not SNPs.

4. I am not sure that trimming SNPs based on LD within pairs completely avoided haplotype effects. Yet, as the majority of interactions are cistrans this is not a measure issue. It would be nice to provide an idea of the average distance (+ range) for cis-cis interactions.

5. Suppl. Methods: Equ. 2: Is it correct that the diagonal includes both $\sigma^2 A + \sigma^2 E$?

6. Suppl. Methods: Page 4, line 1: How do the authors explain that the epistatic component was significant for a much larger proportion of SNP pairs for which one or both SNP had a highly significant marginal effect?

7. Suppl. Fig. 1: I would have liked to see the right panel (null) generated not with random SNPs but with the reshuffled 434 or 404 SNP pairs. The legend refers to 2.5% FDR ... Is this truly FDR?

Finally, despite the inclusion and use of Hi-C data, the paper really does not elaborate about possible molecular mechanisms that might underlie the observed epistatic effects. Were the trans-SNPs more often causing cis-eQTL effects than expected by chance alone? Were they (and SNP in LD) enriched in coding SNP? Were they located in the vicinity of specific types of genes (transcription factors)? Etc.

2 Reviewer 2

Authors carry out epistasis analyses on human eQTL GWAS data. Across 7339 gene expression levels in blood in a cohort of 846 individuals, they detect 501 pairwise SNP interactions, some of which replicate in at least one of two replication data sets. Authors perform some bioinformatic enrichment analyses of interaction SNPs.

Abstract and introduction background: epistasis has been reported in many mapping studies of natural trait variation in multiple species, including for gene expression levels.

What is the evidence that transcription levels are less polygenic than higher level phenotypes?

Page 3, "remarkable similarity in GP maps" needs to be quantified.

Page 4, "cis-cis" interactions are defined as both SNPs on same chromosome as expression gene: These can be very far away and unlikely to be cis, especially if filter of any SNPs in LD is applied here. Interaction results between SNPs on same chromosome are frequently artefactual due to small sample size of "recombinant" haplotype classes because of LD.

Page 4, genes and SNPs involved in very many interactions are not expected given the sparseness of interactions detected, and are likely to reflect technical artefacts.

Bottom of page 5 and top of page 6, enrichement analyses are weakly informative at best. From weak enrichment of cis-acting SNPs vs trans-acting SNPs for transcriptionally active regions in haematopoietic cells it seems unreasonable to draw conclusions about their biological relevance.

Page 5, there is no justification for applying interaction threshold to additive effects. Should match false discovery rates or effect sizes but not thresholds for classes with very different statistical properties in regards to multiple testing and power. Leads to huge underestimate of additive effects.

Methods to identify epistatic QTL are confusing. Test of full vs null model should capture significant additive, epistatic, and/or dominance effects and post hoc methods could be used to disentangle which terms are contributing, but significance after post hoc filters hard to evaluate. Significance threshold for the full vs null model is cited in main text on page 3 before stating that 501 interactions were discovered. This is not exactly appropriate, as this was threshold for full vs null, not the criteria used to determine if there was significant epistasis. It is not clear how many of the 501 interactions are actually significant, nor what the false discovery rate is for this set. It is unclear what filters 1 and 2 äre on (methods page 4).

Should at least use additive-by-additive epistasis model alongside the full model, to increase statistical power and generate better context via comparison to previous work, where this is what is standardly done.

It is not clear how filtering out SNPs with significant additive or dominant effects (methods page 3) is consistent with results in the first full paragraph of page 4 (main text), which notes many interaction SNP pairs with significant main effects.

An appropriate FDR threshold for the tests of the (full) model vs the (additive and dominance) model would be more informative than the Bonferroni threshold used for the post hoc determination of epistatic pairs.

How is the ̈null distribution of no epistatic effects(bottom of page 3) determined?

(top of page 4) Is the dependence on LD between observed SNPs and causal variants the most noteworthy explanation for the lack of replication between discovery and replication samples.

What is the relevance of the statement that ̈patterns of epistasis used for statistical decomposition are not designed to resemble biological function ̈in the context of that paragraph (end of first full paragraph page 4).