Dear all,

You are receiving this email because you are a named author on the Hemani et al. (2014) paper published in Nature. This is a very important email. Please read the remainder.

You last heard from us about this paper in October 2015, when we submitted, with everyone’s permission, a proposed Addendum to the Hemani et al. 2014 paper and its subsequent Brief Communications Arising. The text of that Addendum is copied further below. Our submitted Addendum was not dealt with for a long time, presumably related to the change-over in July 2015 from Magdalena Skipper to Orli Bahcall.

We finally heard from Nature with a ‘Decision’ about our submitted Addendum in June 2017, which is to propose that we voluntarily retract the paper. The full email, including referee comments, is given below.

The reason for the proposed voluntary retraction is that Nature believes that the conclusions from Heman et al. 2014 are no longer supported by the data and therefore Nature wishes to remove the claim of detection and replication of epistasis from the literature. We have had 2 calls with Orli to talk about how the decision was reached and what the process was (is). We apologise for not contacting you earlier – this is because we wanted to get all the facts first and it proved difficult to set up a phone call between us and Orli (which took place on July 27th).

As you can read in the decision letter, Nature proposes that a modified version of the proposed Addendum is published as a new Brief Communications Arising (BCA) that will be linked to the paper. This appears to be an unusual offer.

In terms of process, we have been told that there is no formal appeal process, but we are allowed to get back to Nature to argue if/why we disagree with the decision. The alternative to voluntary retraction is involuntary retraction by the journal.

We believe that we should argue our case with Nature, but would like your opinion. We can set up a Skype or Zoom call to discuss. A response by email is fine as well of course.

Best wishes,

Peter, Grant, Jian, Joseph & Gib

**Decision email from Nature**

Received by email to PMV on June 92017.

Dear Peter,  
  
This email follows to summarize our discussion by phone yesterday. Thank you again for submitting your proposed Addendum on your earlier publication at Nature, “Detection and replication of epistasis influencing transcription in humans.”  
  
We have sent your manuscript to two of the referees who reviewed your original publication, as well as a new referee with expertise in human, statistical genetics and eQTL. These 3 referee reports are attached below.   
  
I have now discussed this matter in detail with my colleagues at Nature together with our research integrity and reproducibility committee at Nature Research.   
  
First, we all share in appreciation for your continued efforts to reanalyze the methods used in detection of epistasis in your original publication, and for sharing your findings in the form of this proposed Addendum. We recognize that this further analysis provides an important contribution that is useful in interpreting the results of the original publication, and for researchers planning further studies in detection of epistasis.  
  
We have also considered how the new analyses influences the conclusions from the original publication, and have discussed this with the referees. The Referees feel that the new analyses, combined with the concerns raised in the earlier Brief Communications Arising by Frayling and colleagues, together raise concerns on the main conclusions of the original publication.   
  
After discussion and review of these concerns, and in particular the inflated false positive rate now shown for this method, we also feel that the main conclusions of the original publication in terms of detection and replication of epistasis, and the frequency of epistasis in these datasets, are not sufficiently supported.   
  
We therefore believe that voluntary retraction of the original research publication would be an appropriate course of action.  
  
As mentioned above, we recognize the contribution provided by your additional analyses, and that this will be useful for our readers. We therefore would like to publish this new analysis in the format of a Brief Communications Arising (BCA), linked to the retraction. We would also like to make it clear by means of this BCA and the retraction text the nature of the progress in this work since your original publication, and that you have initiated these additional analyses.  
  
Please use the following link to submit your revised manuscript as a Brief Communications Arising:  
  
<http://mts-nature.nature.com/cgi-bin/main.plex?el=A3K4BTv4A6BReb6J7A9ftduVzKxWCddj21i9Sotu6fqAZ>  
**NOTE:** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to your co-authors, please delete this link first.  
  
We ask you to ensure that your revised manuscript is in the correct format for Brief Communications Arising (for details, see the guidelines at www.nature.com/nature/authors/gta/commsarising.html). The formatting for Brief Communications Arising are about 700 words, with two figures and up to 15 references, attaching any other supporting data as supplementary material (this is for reviewing purposes only).   
  
I have also included below an initial proposed draft for the retraction text, as a guide for the format used. Please let me know of any questions, and I will be available to work with you in reformatting this manuscript and the retraction text.  
  
Once again, we appreciate and commend you for your work to continue to analyze the methods and results from your original publication.  
  
Yours sincerely  
  
Orli Bahcall, PhD  
Senior Editor   
Nature  
Email: [o.bahcall@us.nature.com](mailto:o.bahcall@us.nature.com)  
  
Proposed draft for retraction text:   
  
In this Article, we reported the detection and replication of epistatic interactions between common single nucleotide polymorphisms (SNPs) influencing gene expression in peripheral blood, including both cis and trans interactions. Wood et al. (Frayling BCA 2014) attempted to replicate these findings in an additional whole-genome sequencing dataset and found that a large fraction of these epistatic effects could be explained by tagging causative SNPs that were not genotyped. In our response (Visscher 2014), we acknowledged that this may explain some of the reported cis-cis interactions, but it remained unclear whether this influenced cis-trans interactions. In an accompanying BCA (Visscher 2017), we now report additional analyses on the same datasets used in our original publication and find that in the presence of imperfectly tagged cis-eQTLs with large additive effects, the F-test statistic we used to test for epistasis can result in an inflated false positive rate. Based on  
these analyses, we now conclude that there is not statistical support for the majority of the reported epistatic interactions in this publication. The authors [list authors that agree] are therefore retracting this paper.   
  
  
  
  
  
Referees' comments:  
  
Referee #1 (Remarks to the Author):  
  
The proposed amendment, following up on the BCA from Wood et al., validates the concern that the original findings of epistasis in human eQTL data were not statistically robust by showing that the putative findings of epistasis are subject to an inflated false positive rate, which is impossible to estimate accurately, but which is consistent with the null hypothesis that none of the reported interactions are real. Given that the title of the paper is "Detection and replication of epistasis influencing transcription in humans", and that the proposed amendment clearly shows that there is in fact no valid evidence for either detection or replication of such epistasis, the central conclusions of the paper should be deemed to be invalid.  
  
  
Referee #2 (Remarks to the Author):  
  
In 2014, Hemani et al. reported the common occurrence of epistatic eQTL effects typically involving linked (cis-cis) or unlinked (cis-trans) variants. Wood et al. (2014) subsequently suggested that a large fraction of these epistatic effects were actually “haplotype effects” tagging ungenotyped causative SNPs. While this explanation may satisfactorily explain at least some of the cis-cis effects, it is less obvious that it might also account for some of the cis-trans effects. Hemani et al. now present an addendum to their 2014 paper that indicates that the test-statistic for cis-trans interactions is inflated in the presence of large additive cis-effects. They suggest that this inflation is due to non-normality of the residuals, and that this probably accounts for a large proportion of the (hence false positive) trans-cis effects reported in their 2014 paper.   
Both the Wood et al. (2014) paper combined with this addendum in essence indicate that the vast majority if not all epistatic interactions reported in the original paper are false positives. The two reports also provide at least partial solutions on how to avoid this artifact in future studies. Hence, publishing this paper is essential, whether as an addendum or as a retraction. As “the central conclusions of the paper are deemed to be invalid “, the retraction unfortunately seems the most appropriate format.   
The authors mention “For the alternative explanation involving cis-cis interactions, both SNPs in cis are correlated with an unknown causal variant, whereas for cis-trans interactions, the SNP in trans is not correlated with either the SNP in cis or the unknown causal variant.” I actually wonder how much of the inflation of the test statistic could be due – with limited sample size – to LD between the cis-trans SNPs and the causative variant. This would mean that both cis-cis and cis-trans false positive epistatic effects would be due to the same problem.   
I was reviewer for the original manuscript and stated: “I am not sure that trimming SNPs based on LD within pairs completely avoided haplotype effects. Yet, as the majority of interactions are cis-trans this is not a measure issue.” I apologize for not having helped the authors more effectively in picking up this problem.   
  
  
  
Referee #3 (Remarks to the Author):  
  
  
I believe that this corrigendum is a valuable contribution to the interpretation of the original paper and the previous corrections. It is also of broad value to the field of association studies in general, since it provides further insight into the technical issues involved in testing for epistasis.  
  
Furthermore, I commend the authors for taking the responsibility of following up on their earlier analyses, and publishing the results. This honesty is unfortunately all-too-rare, and I greatly respect them for doing this.  
  
  
  
  
  
This email has been sent through the Springer Nature Manuscript Tracking System NY-610A-SN&MTS  
  
Confidentiality Statement:This e-mail is confidential and subject to copyright. Any unauthorised use or disclosure of its contents is prohibited. If you have received this email in error please notify our Manuscript Tracking System Helpdesk team at[*http://platformsupport.nature.com*](http://platformsupport.nature.com).

Details of the confidentiality and pre-publicity policy may be found here[*http://www.nature.com/authors/policies/confidentiality.html*](http://www.nature.com/authors/policies/confidentiality.html)

[Privacy Policy](http://www.nature.com/info/privacy.html) | [Update Profile](http://mts-nature.nature.com)

**Addendum for Hemani *et al*. (2014), Nature 508:249-53**

(submitted by email to editor Orli Bahcall on 22 October 2015)

We reported the detection and replication of epistatic interactions for gene expression (Hemani et al. 2014). Subsequently, in a Brief Communications Arising (BCA), Wood et al. (2014) proposed an alternative explanation for the reported interactions, namely the existence of an imperfectly tagged causal variant in *cis* which is in linkage disequilibrium (LD) with both genotyped SNPs. Using an independent data set with gene expression and genome sequences, Wood et al. 2014 reported a substantial reduction of signal for the interaction term by adjusting the expression trait by the top associated genetic variant from the sequence data. We acknowledged that such “haplotype effects” could explain a number of the reported *cis-cis* interactions but it was not clear whether this alternative explanation applied to all reported interactions, in particular the *cis-trans* interactions (Hemani et al. 2014 BCA reply). For the alternative explanation involving *cis-cis* interactions, both SNPs in *cis* are correlated with an unknown causal variant, whereas for *cis*-trans interactions, the SNP in trans is not correlated with either the SNP in *cis* or the unknown causal variant.

We have now performed additional empirical analyses on exactly the same data used in the original publication (Hemani et al. 2014) to attempt to resolve the alternative explanations for the results given by Hemani et al. (2014) and Wood et al. (2014 BCA). We used additional genome-wide association testing, simulation studies and empirical derivations to study the behaviour of the test statistic for epistasis when epistatic interactions are not present. These investigations have revealed that the F-test we used to detect epistasis is not robust to large additive effects of causal variants in *cis* that are not perfectly tagged by genotyped SNPs. A large imperfectly tagged additive effect in *cis* violates the assumed null distribution of no epistatic interaction because residuals are not normally distributed (they are a mixture of binomial and normal errors). This violation of assumption can create widespread genome-wide inflation of the test statistic, both for *cis-cis* and *cis-trans* effects, which can cause an inflated false-positive rate if the theoretical F-distribution is used to calculate a p-value (as was done in Hemani et al. 2014).

In Figure 1 below we report the genome-wide inflation factor of the F-test for epistasis for the 501 SNPs pairs that passed our significance filters (Hemani 2014). This factor was calculated by transforming the p-value from the observed F–test of each test for epistasis to a 2 statistic with 1 degree of freedom and then taking the median 2 statistic across all tests divided by its value under the null hypothesis, 0.455. The inflation factor was calculated from a genome-wide analysis that held the SNP with the strongest additive effect in *cis* constant and then performing an interaction test with SNPs on other chromosomes. The genome-wide analysis excluded SNPs in a 5 Mb region around the 2nd SNPs of the reported pair and SNPs that were on the same chromosome as the 1st SNP. In Table 1 we report the average inflation of the test statistic per probe, for probes that had at least 5 SNP pairs that contributed to the 501 selected pairs. Figure 1 and Table 1 show that for many probe *cis*-SNP pairs, in particular for the probes for which we reported *cis-trans* effects in Hemani et al. (2014), there is a large inflation of the test statistic. Note that since we removed 5Mb regions around the 2nd SNPs the actual inflation could be slightly larger than those shown in Figure 1 and Table 1. The two probes for which replication was reported in Table 1 of Hemani (2014), TMEM149 and MBNL1, have median genome-wide test statistics that are, respectively, more than 5 and 3 times the value under the null hypothesis of no epistasis.

The consequence of these findings is that in the presence of imperfectly tagged eQTLs with large additive effects, the false positive rate from the test statistic for epistasis we employed is larger than assumed. A larger false positive rate is consistent with finding apparent replication (since the same applies to the replication data) and with having a much lower replication rate than expected given the stringency in the discovery sample. It is also the likely explanation why the statistical evidence for *cis-trans* effects is reduced after an adjustment for the additive effect of the best-tagging sequence variant in the *cis* region (Wood et al. 2014).

Whilst we cannot resolve with certainty alternative explanations for all reported epistatic interactions (either the 501 pairs discovered or the 30 pairs replicated) with these additional analyses, they show that in the presence of imperfectly tagged large *cis*-eQTLs, tests for epistasis such as that employed by us can result in an inflated false positive rate. This conclusion is general and alternative explanations may also apply to other studies that have reported pairwise interactions in the presence of large additive effects using SNP-array based genotype data in particular. Unfortunately, there is, to our knowledge, no existing or proposed analysis solution to this problem. The problem could be attenuated using whole genome sequencing data or very well imputed data. Creating an empirical null distribution by correction for genome-wide inflation for any pair of SNPs on different chromosomes in a full genome-wide scan for epistasis is at present computationally too demanding, and would not resolve *cis-cis* interactions.

Figure 1: Distribution of the ratio of the median F-statistic for epistasis and its value under the null hypothesis from a genome-wide empirical pairwise epistasis analysis that held the cis-SNP constant, for each of the 501 SNP pairs that passed the significance filters in Hemani et al. (2014).

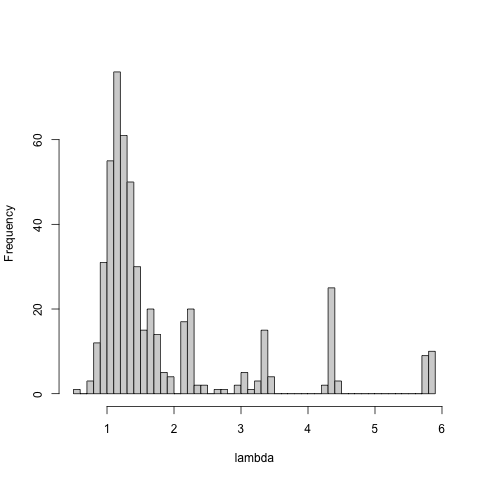


Table 1: Probes with multiple cis-trans interactions have genome-wide inflation of the test statistic for epistasis.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Probe Name | SNP 1 | SNP 2 | Gene | Mean Lambda | N pairs |
| ILMN\_1704730 | rs1884655 | rs10255470 | CD93 | 2.88 | 10 |
| ILMN\_1710752 | rs2123758 | rs2786014 | NAPRT1 | 2.15 | 8 |
| ILMN\_1717234 | rs1157079 | rs7733671 | CAST | 4.31 | 17 |
| ILMN\_1720059 | rs12435486 | rs7837237 | HMBOX1 | 2.29 | 7 |
| ILMN\_1738784 | rs10930170 | rs12120009 | PPP2R5A | 2.24 | 6 |
| ILMN\_1755589 | rs11080134 | rs11169322 | DIP2B | 1.16 | 6 |
| ILMN\_1786426 | rs2839013 | rs8106959 | TMEM149 | 5.65 | 20 |
| ILMN\_1804396 | rs1293455 | rs2655991 | C14ORF4 | 1.38 | 7 |
| ILMN\_2313158 | rs10869600 | rs13069559 | MBNL1 | 3.15 | 15 |
| ILMN\_2372639 | rs17159840 | rs10059004 | TRAPPC5 | 4.17 | 17 |
| ILMN\_3231952 | rs12947580 | rs8079215 | ARL17B | 2.16 | 6 |