Dear Dr. Daetwyler and Dr. Long,

Response to reviews

Specifics of response follow:

Sincerely,

Alice

GENETICS/2019/302613   
**Genetic Associations in Four Decades of Multi-Environment Trials Reveal Agronomic Trait Evolution in Common Bean**   
  
Dear Dr. MacQueen:   
  
Three experts in the field have reviewed your manuscript, and I have read it as well, and I am pleased to tell you that it is potentially suitable for publication in GENETICS. However, the reviewers have suggested additional analyses and have provided in-depth comments and concerns that need to be addressed in a revised manuscript. You can read their reviews at the end of this email.   
  
If you submit a revised manuscript please include a response to each of the reviewers' comments. I will highlight several here, but please consider all comments made by the reviewers. R1 suggest investigating the use of factor analytic models to generate the BLUPs for downstream analysis. My own experience has also been that FA models may be more appropriate in data structure such as yours. R2 ask that you discuss why a key regions was not detected in your analyses. R3 suggest more work is needed to be able to claim that pleiotropy and not linkage has been observed. They also suggest, as the use of Mash is emphasised in the manuscript, that you may want to compare Mash to another method with similar aims.   
  
I look forward to receiving a revised manuscript. I expect it could be submitted within 90 days, but please let me know if you think you will need more time to complete the revision. A revised manuscript will be considered a resubmission, and may be sent out for review.   
  
Follow this link to submit the revised manuscript: <https://genetics.msubmit.net/cgi-bin/main.plex?el=A3NR5CLP7A7JPq4I7A9ftdmmOpcr6vwXLMGsQygHL0xQZ>   
  
If you have questions about the reviews or this message, please contact me.   
  
  
Sincerely,   
  
Hans Daetwyler   
Associate Editor   
GENETICS   
  
  
Anthony Long   
Senior Editor   
GENETICS   
  
***look at these editors – ask what software they have to get BLUPs in incomplete datasets – anything that might internally fit a VCOV model for me to get BLUPs***

***Ask the editor/reviewer #1 if they have suggestions for software***

***Tom thinks ASREML would do it… most flexible mixed model that quantitative genetic breeders use. Wouldn’t have an automatic way to compare VCOV matrices… you’d have to use likelihood to do this.***

--------------------------------------------------------------------   
  
  
Reviewer #1 (Comments for the Authors (Required)):   
  
The main objective of the manuscript is to identify genomic regions regulating multiple traits in common bean. To identify these regions, the authors apply the mash method, proposed by a recent paper by Urbut et al. (2019, Nat. Genetics) to the genotypic BLUPs (genetic main effects) for multiple traits obtained across environments. By calculating the genotypic main effects across environments, the authors circumvent the issue of high imbalance in the multi-environment data. That allows them to concentrate on the multi-trait analysis. The main question addressed in this paper is relevant to many crops and although the paper implements already existing methodology, they are the first ones (to my knowledge) to implement this method in crop data coming from historical multi-environment data. The manuscript is well written and the supporting data and code available in GitHub are clear and complete. That certainly increases the impact of the paper, as they facilitate the implementation in other crops. However, I do have a few comments:

- My main concern is about the calculation of the BLUPs for the genotypic main effect. The authors used rrBLUP to fit a model with locations, years and their interactions as fixed and with genotype main effects as random. The variance-covariance between genotypes was modelled with a compressed Kinship matrix calculated with GAPIT. The model that was used looks rather incomplete to me. I suggest to model the genotype by environment interaction explicitly (with a factor analytic structure, for example) and then produce the BLUPs for the genotypic main effects. Or, at least, to decompose the GxE into GxL, GxY and GxLxY and then produce the predictions for the genotypic main effect. That would also probably improve the heritability estimates, which are rather low now as part of the GxE variance is accounted as residual variance.

*The reviewer is correct that we did not model the GxE explicitly to calculate the BLUPs. This was due to the extremely sparse nature of the GxE matrices – or Genotype by Location by Year matrices – for each phenotype. However, a decomposition of GxLxY matrix and the use of a factor analytic structure could both reduce the sparseness of the matrix and make inclusion of GxE variance possible.*

*As discussed, the GxE matrices for each phenotype were extremely sparse – the genetic correlation matrices constructed from GxLxY data never have more than 15% of cells with correlations. Since most genotypes were only grown in sequential years, for 1-4 years, the GxY matrix was even sparser – no genetic correlation matrix for any phenotype from GxY data had more than 6% of cells with correlations. Given that factor analysis requires matrices without missing values, we judged that these matrices did not have enough complete cases for factor analysis. We are unaware of methods to model such sparse matrices. However, 9 phenotypes had between 40% and 99% of genetic correlations constructed from GxL data: seed yield, seed weight, days to maturity, days to flowering, seedfill duration, biomass, plant height, harvest index, and lodging. For these nine phenotypes, we used factor analysis to produce covariates for the GxL matrix that we then included in our calculation of BLUPs for the genotypic main effects. The script is available at \_link\_to\_R\_script\_.*

*As I understand it, you would like me to model the GxE explicitly using factor analysis, and use fa() or similar to produce a variance and covariance matrix to put in to rrBLUP, in addition to using the compressed Kinship matrix calculated with GAPIT. I believe the genetic variance matrix can still be modeled as a kinship matrix, but I will need to add covariates (using covariate= in kin.blup()) for each genotype from a factor analysis of all genotypes across all locations and years.*

*I looked at Boer et al 2007 which had a nice breakdown of the types of variance and covariance structures that could be included in a model to generate BLUPs. I had a model where genetic variance and environmental variance did not covary. – I did not specify heterogeneity of genetic variances across individual environments, nor heterogeneity of genetic correlations between pairs of environments. I specified genetic variances by the kinship matrix, and a common set of fixed effects for location and the interaction between location and year, which were assumed to be uniform across all genotypes.*

Minor comments   
- Line 273: the address should be <https://github.com/Alice-MacQueen/CDBNgenomics/tree/master/data-raw>, instead of <https://github.com/Alice-MacQueen/CDBNgenomics/tree/master/analysis-paper>

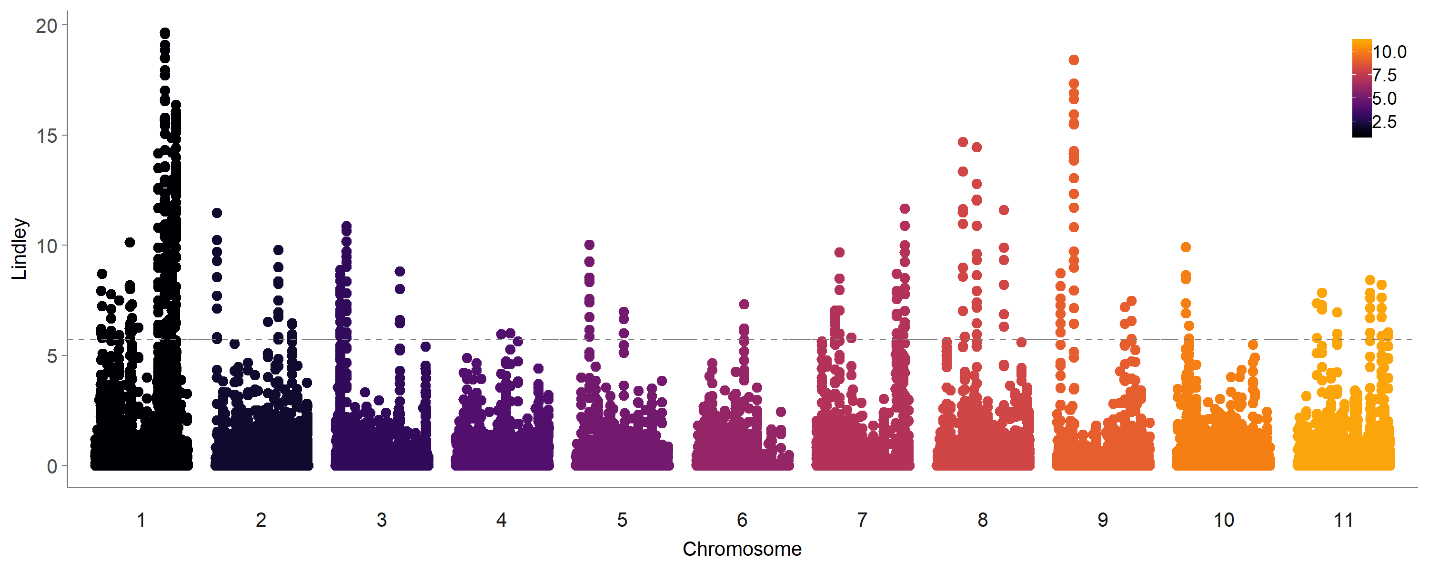
*Edited.*

- Figure 3. The point estimates in sub-figure D and E appear very small. Please use the same size as in sub-figure F  *Edited. Will remake figure once I redo the analysis with local score/ new GxE covariance matrix*  
  
  
  
Reviewer #2 (Comments for the Authors (Required)):   
  
Review of manuscript GENETICS/2019/302613, "Genetic Associations in Four Decades of Multi-Environment Trials Reveal Agronomic Trait Evolution in Common Bean "   
Authors: Alice H. MacQueen, Jeffrey W. White, Rian Lee, Juan M. Osorno, Jeremy Schmutz, Phillip N. Miklas, Jim Myers, Phillip E. McClean, Thomas E. Juenger   
  
The authors mine phenotypic data collected from bean trials conducted over 70 years across 10-20 locations in North America to detect significant genetic effects that define genomic regions associated with traits of agronomic importance. Much of the findings align with genomic regions previously identified in QTL and GWAS studies. One of the surprising results were the genomic regions not detected by the MET analysis. The fundamental reason to conduct the CDBN trials was to measure performance (yield) of new bean lines across locations and years. The MET analysis detected genomic regions associated with yield on Pv01 adjacent to the determinacy allele which dramatically affects yield but the analysis did not detect a critical region on Pv03 (~38.27Mb) that has been shown to be associated with yield in 10 independent QTL and GWAS studies (Kelly, 2018). Interestingly, a candidate gene for days to maturity was detected in this same region of Pv03 (36.8Mb) and maturity is a major factor in plant yield. This brings into question the value of the MET analysis as a tool to detect genomic regions associated with the major economic driver in bean breeding.

*We anticipated we might have seen this region, but we didn’t: here are two possible reasons for why that might be.*

*One possibility is that the additional noise inherent in a GWAS on data from multiple locations and multiple years – even though partly accounted for in our model – gave us very low power to detect associations of minor, or even moderate effect (phenotypic variation explained by the SNP of, say, 0.05 to 0.15. With two exceptions, studies in Kelly 2018 had the R2 of the most significant association on Pv03 of less than 0.15. In addition, the additive effect in these studies of this locus on yield ranged between 0.7 kg/ha and 219 kg/ha, with the average effect 110-140 kg/ha. Our only significant association was correlated with a difference in seed yield of 104 kg/ha. If this region had an effect of less than 104 kg/ha, as it did in five out of 14 studies (36%), we would not have detected an effect of this region.*

*When we used the local score approach, which has increased power to detect associations of smaller effect, we detected a significant region for seed yield on Pv03 at 36.988 Mb (and a very strong hit for days to maturity only 55kb away), confirming that this region does have an effect on seed yield in this panel, but that our original approach did not have power to detect it.*

**

*There are many caveats to an approach like the one we used, and thus we considered some overlaps with previous studies necessary to demonstrate the usefulness of this method, but no particular single overlap or lack of overlap, even a robust one of moderate effect, either sufficient or damning evidence for or against the efficacy of this method. We now mention that we detect this region for days to maturity, but not seed yield, and cite the paper.*

Lines 516 and 707: The authors state an alternate allele of Ur-3 first 'appeared' in pinto Sierra and Starlight great northern. 'CDBN until 1988, when it appeared in the pinto Sierra and the great northern Starlight". This sentence needs to be rewritten as new resistant alleles do NOT just 'appear' in new varieties they are bred into the variety over generations of selection. Replace 'appeared' with 'was first observed'

*True. Edited.*  
  
Line 522: The determinacy factor (growth habit) in navy beans was first reported on Pv07 by Kolkman and Kelly (2003). No actual physical mapping was possible in common bean in that era but the same genomic region was confirmed later by Kwak et al. (2012). All the factors listed as associated with the alternate allele on Pv07 (line 601) would be associated with determinacy.

*We edited the discussion (lines 661-664) to include this work.*  
  
Lines 523-524: The authors state: "Phvul.001221100, recently identified as the photoperiod sensitivity locus Ppd, or PHYTOCHROME A3 (WELLER et al. 2019)". The association of days to flower and the Ppd gene and the candidate phytochrome A3 gene was first reported by Kamfwa et al. (2015) on Pv01 at 48.3Mb. Please acknowledge the first report.

*Edited.*

Line 545: 'early on' replace with 'in the 1980s'

*Edited.*

Line 686: suggest adding '........delayed flowering to prevent pod set on lower nodes (Vandermark et al. 2014)'.

*Edited.*  
  
[Citation: Kelly, J.D. 2018. Developing improved varieties of common bean, 2:3-17. In: Achieving sustainable cultivation of grain legumes. Vol. 2: Improving cultivation of particular grain legumes (eds. Sivasankar S., Bergvinson D., Gaur P., Agrawal S.K., Beebe S., Tamò M.) Burleigh Dodds Science Publishing, Cambridge UK, pp. 376.]   
  
[Citation. Kolkman, J.M. and J.D. Kelly. 2003. QTL conferring resistance and avoidance to white mold (Sclerotinia sclerotiorum) in common bean (Phaseolus vulgaris). Crop Sci 43:539-548.]   
  
  
  
  
Reviewer #3 (Comments for the Authors (Required)):   
  
This manuscript describes a genome-wide association study of agronomic traits and yield components in common bean using several decades of breeder phenotyping in a cooperative trial system (the CDBN). The study uses single trait GWAS and a multi-trait method (mash) to identify associations. It argues that the colocalization of association signals reflect shifts in the ideotypes over several decades of breeding. This is an interesting issue, of interest from both a plant breeding and evolutionary genetics perspective. The common bean nurseries provide a good system to investigate it.   
  
Strengths: The study rigorously develops a genotype-phenotype data set including (i) the digitization and curation of multi-decade agronomic phenotype resource and (ii) genotyping of unarchived entries from CDBN to generate a more complete data set. These analyses are well described in the text and the supplemental material, which should facilitate reuse of this resource. The main hypothesis of the study, that shifts in bean ideotypes in the past several decades have led to selection on pleiotropic loci, is potentially interesting to a wide range of geneticists.   
  
Weaknesses: The main hypothesis of the study is not consistently addressed throughout the manuscript. Instead much of the manuscript is exploratory (see objectives in line 166-168). The hypothesis could be further tested by genome scans. The prediction would be a significant enrichment of outlier SNPs within the mash peak regions. Evidence of pleiotropy from the single trait GWAS is also not clear. A systematic analysis of colocalization (compared to an appropriate null model) would be needed to show evidence of pleiotropy.

*This reviewer appears to want clarification of the main message of the paper.*   
  
Specific points to address:   
  
Line 147-149: "Selection for a common bean crop ideotype ... is known to have led to pleiotropic effects...". This claim is not supported by existing evidence. A small phenotypic association study is cited to support the claim (Soltani et al. 2016), but this study does not distinguish pleiotropy vs. linkage. Instead this claim appears to be the main hypothesis to be tested in the current study. (Line 712-713 makes the same claim: "Selection for the common bean ideotype is known to have led to pleiotropic effects")   
*Neither our study nor the previous study we cite provide conclusive evidence that pleiotropy, rather than linked effects, are producing … Genetic manipulation is necessary to distinguish between pleiotropy and linkage. This kind of genetic manipulation is outside of the scope of this paper. In the paper, we are careful to say*

Line 159-161: The text highlights the novelty and value of using mash for crops, here and elsewhere in the manuscript. If this is an important point to make, then this method should be compared to other methods, e.g. MTMM (Korte et al. 2012 Nat Gen) or similar. Otherwise it doesn't seem relevant to emphasize this point.

*This is not the point of our paper – we are not methods developers, and we’re not astute enough in the methods development and comparison approaches to be able to do this comprehensively or well.*

Line 272: For location and year the variance is the parameter of interest so a random effect seems more appropriate.   
  
Line 280-282: I know it's common to use the term "control" to describe the population structure (Q) and kinship terms (K), but it's not clear that they "control" anything and I'd argue it's more appropriate to say the Q and K terms "account" for genetic background effects (Vilhjálmsson and Nordborg 2013 Nat Rev Genet).   
  
Line 286, 307-309: Explain these distance cutoffs relative to the extent of LD in the germplasm. Overall, the lack of a linkage disequilibrium analysis in the study makes it difficult to interpret the distance cutoffs (e.g. 20 or 200kb) used to identify the candidate genes.

*We agree that distance cutoffs are arbitrary and certainly meaningless without knowledge of LD. During our revisions, we became aware of a new method to better account for LD structure effects on p-values and the localization of effects in the genome.*   
  
Line 320-324: This text is repeated from the introduction.   
  
Line 364-386: This first section of the Results has no results, but a mixture introductory and discussion points. These points should be covered elsewhere if needed, but not in Results.   
  
Line 398: "used to determine the narrow-sense heritability, h2". The Material and Methods don't describe the method. Since the previous sentence mentions kinship, maybe a SNP-based estimate of h2 was used? Provide details.   
  
Line 825: "h2 is narrow sense heritability, defined as Vg / (Vg + Ve)". This is the definition of broad-sense heritability, not narrow-sense heritability (which would have Va, the additive heritability in the numerator).