Cell Systems

Transparent Peer Review Record

A resource for experimental precision medicine: The BXD family

David G. Ashbrook, Danny Arends, Pjotr Prins, Megan K. Mulligan, Suheeta Roy, Evan G. Williams, Cathleen M. Lutz, Alicia Valenzuela, Casey J. Bohl, Jesse F. Ingels, Melinda S. McCarty, Arthur G. Centeno, Reinmar Hager, Johan Auwerx, Lu Lu, Robert W. Williams

Summary

Initial Submission: Received June 29, 2020
Preprint: https://doi.org/10.1101/672097
Deposited on bioRxiv, June 25, 2019
Scientific editor: Ernesto Andrianantoandro, Ph.D.

First round of review: Number of reviewers: Two
Two confidential, Zero signed
Revision invited August 10, 2020
Major changes anticipated
Revision received August 29, 2020

Second round of review: Number of reviewers: Two
Two original, Zero new
Two confidential, Two signed
Accepted December 21, 2020

Data freely available: Yes
Code freely available: Yes

Cell Systems Transparent Peer Review Record

Editorial decision letter with reviewers' comments, first round of review

Dear Dr. Ashbrook,

I hope this email finds you well. The reviews are back on your manuscript and I've appended them below. You'll see that the reviewers find the manuscript compelling and their comments are intended to strengthen an already strong piece of work. We're happy to invite a revision.

The reviewer recommendations are straightforward, but if you have any questions or concerns about the revision, I'd be happy to talk about them, either over email or by phone. More technical information and advice about resubmission can be found below my signature. Please read it carefully, as it can save substantial time and effort later.

I look forward to seeing your revised manuscript.

All the best,

Ernesto Andrianantoandro, Ph.D.
Scientific Editor, *Cell Systems*

Reviewers' comments:

Reviewer #1: The BXD family of mice and BXD phenome is truly a remarkable resource, which deserves greater recognition and exploitation within the scientific community. Ashbrook and colleagues present a timely update of the BXD resource highlighting the expansion the BDX lines to 150 strains, with genetic maps. The legacy value of the resource has been increased by the associated phenome data of over 7500 classical phenotypes and 100 omic datasets. Although this reviewer is not qualified to assess the statistical genetics presented, the paper communicates a number of important points that should be of wide interest.

The value of this resource for precision medicine is yet to be fully realised and will increase with creative experiments. For example, efficient CRISP engineering of many lines to insert markers, reporters and mutations. I would have liked to have read more about the advantages of the larger panel of mice for matching polygenic risk scores of human disorders to strains that might have predicted susceptibilities and resilence, something that is relevant to the recent 5XFAD paper from Kaczorowski.

Reviewer #2: In this report, Ashbrook and colleagues announce the expansion of the BXD genetic resource to a total size of 150 strains. The larger complement of BXD strains will provide a valuable resource for the mouse modeling community due in part to the large corpus of existing phenotyping that

has been conducted with existing BXD animals, and in part due to the longstanding effort the authors and their collaborators has made to making these phenotype data available through online tools. The authors demonstrate that the expanded cohort increases the power and accuracy of observations made using BXD animals through several methodological demonstrations. This report focuses on presenting the new resource and associated assays (e.g. expanded genotyping) rather than on novel biological observations. The authors are also primary contributors to several web resources related to the BXD strains, including a shiny app devoted to power calculations.

Overall, the quality of writing is very high. The shiny app and tools at https://www.genenetwork.org are well done and highly practical. The introduction reviews the timeline of BXD development in great detail before presenting new data. If that is typical for this journal, it presents no problem, but some of the may be more suitable for a perspective or timeline-style article.

Major comments:
1) It is difficult to extract useful information from Figure 2 because the individual plots are quite small. It's not clear what phenotypes are being represented or how to compare individual results between panels A and B. It would be more useful to choose a few representative examples of phenotypes where evidence for genetic influence was equivocal and is now strengthened by the newly expanded resource, and illustrate them in figure panels, and present these figures at greater resolution in a supplement.

2) Since results for both H-K and LMM methods are overplotted on figure 3, the reader cannot assess the pairwise differences in the statistical results of these two methods from the figure. The values plotted in figure 3 should be reported in a supplementary table.

3) Qualitatively, it appears that the relationship between peak LOD and difference from threshold to peak demonstrated in figure 3 for LMM is closer to linear and less noisy than H-K. Since there is no direct comparison of the peak LOD scores on a pairwise basis, one possible interpretation of the figure is that improved performance of the LMM method shown in figure 3 is meaningfully affected by the results of the permutation approach rather than the LOD score. If permutation testing using H-K tends to generate a nosier and higher significance threshold than LMM, that might account for the observed results. This would still be an interesting and a helpful improvement if the LMM is accurately setting a lower permutation threshold, but it would be helpful for the authors to comment on this variance and indicate whether the peak LOD scores or the size of the candidate region around that LOD peak are improved with LMM vs H-K.

Authors' response to the reviewers' first round comments

Attached.

Cell Systems

Transparent Peer Review Record

Editorial decision letter with reviewers' comments, second round of review

Dear Dr. Ashbrook,

I'm very pleased to let you know that the reviews of your revised manuscript are back, the peer-review process is complete, and only a few minor, editorially-guided changes are needed to move forward towards publication.

In addition to the final comments from the reviewers, I've made some suggestions about your manuscript within the "Editorial Notes" section, below. Please consider my editorial suggestions carefully, ask any questions of me that you need, make all warranted changes, and then upload your final files into Editorial Manager. ***We hope to receive your files within 5 business days, but we recognize that the COVID-19 pandemic may challenge and limit what you can do. Please email me directly if this timing is a problem or you're facing extenuating circumstances.***

I'm looking forward to going through these last steps with you. More technical information can be found below my signature, and please let me know if you have any questions.

All the best,

Ernesto Andrianantoandro, Ph.D.
Scientific Editor, Cell Systems

Editorial Notes

General:

Please reformat as a Report (4 main figures).

Title:

Your title is too generic and doesn't capture the advance in the paper. I suspect it could be more effective. Please include something about expanding the BXD family, the phenome, and the word mouse or murine. As you re-consider your title, note that an effective title is easily found on Pubmed and Google. A trick for thinking about titles is this: ask yourself, "How would I structure a Pubmed search to find this

paper?" Put that search together and see whether it comes up is good "sister literature" for this work. If it does, feature the search terms in your title. You also may wish to consider that PubMed is sensitive to small differences in search terms. For example, “NF-kappaB” returned ~84k hits as of March, 2018, whereas “NFkappaB” only returned ~8200. Please ensure that your title contains the most effective version of the search terms you feature.

Abstract:

The abstract is too long. Please condense to 150 words or less. Please also indicate the original number of BXD strains so readers know how much the family was expanded.

Manuscript Text:

While the text is compelling and clear, it’s rather significantly too long. You’re welcome to go up to 35,000 characters-with-spaces, not including the STAR Methods or the references. (To be explicit, the STAR Methods and references don't “count” towards your manuscript’s length.) Anything longer taxes readers too much; it’s simply too much for the human mind to take in all at once. When you think about what to cut, start with the Discussion. We favor slim Discussions that do not reiterate what's found in the Results beyond a brief transitional summary and are limited to around four medium sized paragraphs. If further cuts are needed, your figures can guide you: paragraphs within the Results section that only pertain to supplemental figures can be slimmed down dramatically (usually into a single sentence that calls out the supplemental figure and states its punchline) or deleted entirely. If it's appropriate, discrete details from those paragraphs can be moved into the Supplemental Figure Legends. After that, please cut any references that aren’t actually necessary.

Too much of the introduction contains interpretations of the results – please remove or move non-redundant material to the Discussion.

Also:

- House style disallows editorializing within the text (e.g. unsurpassed, unrivaled, substantial, marked, striking, surprising, important, etc.), especially the Results section. These terms are a distraction and they aren't needed—your excellent observations are certainly impactful enough to stand on their own. Please remove these words and others like them. “Notably” is suitably neutral to use once or twice if absolutely necessary.
- We don’t allow “priority claims” (e.g. new, novel, etc.). For a discussion of why, read: http://crosstalk.cell.com/blog/getting-priorities-right-with-novelty-claims, http://crosstalk.cell.com/blog/novel-insights-into-priority-claims.
- Please only use the word "significantly" in the statistical sense.

Figures and Legends:

Bar graphs are not acceptable because they obscure important information about the distributions of the underlying data. Please display individual points within your graphs unless their large number obscures the graph's interpretation. In that case, box-and-whisker plots are a good alternative. Also, please ensure that you have defined "n's" specifically and listed statistical tests within your Figure Legends.

STAR Methods:

Cell Press has recently changed the way it approaches "availability" statements for the sake of ease and clarity. Please revise the first section of your STAR Methods as follows, noting that the particular examples used might not pertain to your study.

> RESOURCE AVAILABILITY
>
> **Lead Contact:** Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jane Doe (janedoe@qwerty.com).
>
> **Materials Availability:** This study did not generate new materials. *-OR-* Plasmids generated in this study have been deposited at [Addgene, name and catalog number]. *-OR-* etc.
>
> **Data and Code Availability:**
>
> - **Source data statement** (described below)
> - **Code statement** (described below)
> - **Scripts statement** (described below)
> - Any additional information required to reproduce this work is available from the Lead Contact.

Starting in August of 2020, Cell Systems papers will need to contain a comprehensive and structured "Data and Code Availability" statement. These statements will exceed standard STAR Methods requirements.

Data and Code Availability statements pertain to the source data and original code reported in the study. In this context, **source data** is defined as the collection of individual, unprocessed observations used to generate the figures reported in the paper. Examples include scRNA-seq and proteomic datasets, but also CSV spreadsheets used to generate graphs, and original micrographs in TIFF format. **Code** is defined as any computationally implemented program, algorithm, or pipeline necessary to reproduce the analysis or conclusions reported in a paper. Smaller **scripts** that have been used to visualize data and generate figures should also be included in the statement, as described below.

Data and Code Availability statements are reported in the first section of the STAR Methods**. They have four parts and each part must be present. Each part should be listed as a bullet point, as indicated above. For convenience, a .docx template for Data and Code availability statements can be downloaded here.**

Part 1 pertains to source data. Examples can be used in any number or combination, making sensible modifications as necessary:

- [Data-type] source data have been deposited at [data-type-specific repository] and are publicly available under the accession numbers: [Insert].
- [Data-type] source data have been deposited at [general repository] and are publicly available at [insert DOI].
- [Data-type] source data are available in the paper's Supplemental Information.
- The [data-type] source data reported in this study have not been deposited in a publicly available repository because [reason why data are not public] . They have been archived locally [insert archiving plan]. To request access [insert instructions].
- This paper analyzes existing, publicly available data. These datasets' accession numbers are provided in the Key Resource Table.
- Source data are not provided in this paper but are available from the Lead Contact on request. *(Note: Cell Systems discourages this practice. If you need to make this statement, please discuss it with your editor first.)*

Part 2 pertains to original code. Examples can be used in any number or combination, making sensible modifications as necessary:

- [Adjective] original code is publicly available at [repository name and DOI].
- [Adjective] original code is available in this paper's Supplemental Information.
- The original code reported in this study is not publicly available repository because [reason why data are not public]. Original code has been archived locally [insert archiving plan]. To request access [insert instructions].
- This paper does not report original code.

Part 3 pertains to scripts used to generate figures. Examples to be used in any number or combination:

- The scripts used to generate the figures reported in this paper are available at [repository name and DOI].
- The scripts used to generate the figures reported in this paper are available in this paper's Supplemental Information.
- The scripts used to generate the figures reported in this paper are available in the [name software package, with version, and provide reference or URL] and their use is described in the STAR Methods.
- Scripts were not used to generate the figures reported in this paper.
- Scripts used to generate the figures presented in this paper are not provided in this paper but are available from the Lead Contact on request. *(Note: Cell Systems discourages this practice. If you need to make this statement, please discuss it with your editor first.)*

Part 4 is a statement: “Any additional information required to reproduce this work is available from the Lead Contact.”

Thank you!

Reviewer comments:

Reviewer #1: The authors have addressed my comments. Overall, this is a very timely and useful article.

Reviewer #2: The authors have addressed the comments I raised in the previous review. I recommend their paper be accepted for publication as it is.

We thank the reviewers for their kind and useful comments. We have addressed the comments below, in red italic type.

Reviewers' comments:

Reviewer #1: The BXD family of mice and BXD phenome is truly a remarkable resource, which deserves greater recognition and exploitation within the scientific community. Ashbrook and colleagues present a timely update of the BXD resource highlighting the expansion the BDX lines to 150 strains, with genetic maps. The legacy value of the resource has been increased by the associated phenome data of over 7500 classical phenotypes and 100 omic datasets. Although this reviewer is not qualified to assess the statistical genetics presented, the paper communicates a number of important points that should be of wide interest.

The value of this resource for precision medicine is yet to be fully realised and will increase with creative experiments. For example, efficient CRISP engineering of many lines to insert markers, reporters and mutations. I would have liked to have read more about the advantages of the larger panel of mice for matching polygenic risk scores of human disorders to strains that might have predicted susceptibilities and resilence, something that is relevant to the recent 5XFAD paper from Kaczorowski.

We whole-heartedly agree with Reviewer 1 about the importance of using the BXD family as a panel of background strains on which to test the effects of other mutations.

We agree that the generic risk score approach you mention, as used by Neuner et al., 2019, is of interest, and we have now added the following paragraph on page 28:

> *Neuner and colleauges also evaluated the efficacy of reverse translation from human to mouse (Neuner et al., 2019a). They generated a polygenic genetic risk score using 21 human genes which increase Alzheimer's disease risk, and showed that the allele dosage of was significantly associated with cognitive outcomes in their AD-BXD cohort. This demonstrates that naturally occurring variation in these networks have overlapping effects in mouse and humans. This approach could be applied to many other phenotypes.*

Reviewer #2: In this report, Ashbrook and colleagues announce the expansion of the BXD genetic resource to a total size of 150 strains. The larger complement of BXD strains will provide a valuable resource for the mouse modeling community due in part to the large corpus of existing phenotyping that has been conducted with existing BXD

animals, and in part due to the longstanding effort the authors and their collaborators has made to making these phenotype data available through online tools. The authors demonstrate that the expanded cohort increases the power and accuracy of observations made using BXD animals through several methodological demonstrations. This report focuses on presenting the new resource and associated assays (e.g. expanded genotyping) rather than on novel biological observations. The authors are also primary contributors to several web resources related to the BXD strains, including a shiny app devoted to power calculations.

Overall, the quality of writing is very high. The shiny app and tools at https://www.genenetwork.org are well done and highly practical. The introduction reviews the timeline of BXD development in great detail before presenting new data. If that is typical for this journal, it presents no problem, but some of the may be more suitable for a perspective or timeline-style article.

Major comments:
1) It is difficult to extract useful information from Figure 2 because the individual plots are quite small. It's not clear what phenotypes are being represented or how to compare individual results between panels A and B. It would be more useful to choose a few representative examples of phenotypes where evidence for genetic influence was equivocal and is now strengthened by the newly expanded resource, and illustrate them in figure panels, and present these figures at greater resolution in a supplement.

We have now reformatted the figure to provide a single, large, example plot of a whole-genome QTL map in panels A and B, to make comparison easier. We hope it is now clear that the large peak on chr7 seen with the current genotypes (A), was not significant using the classic phenotypes (B). We have added a supplementary figure (Figure S2) showing the genome-wide plot for several other phenotypes (included in the original Figure 2), to show that the improvement is seen across chromosomes.

2) Since results for both H-K and LMM methods are overplotted on figure 3, the reader cannot assess the pairwise differences in the statistical results of these two methods from the figure. The values plotted in figure 3 should be reported in a supplementary table.

Done, there is now a supplementary excel sheet (Table S2), with all values.

3) Qualitatively, it appears that the relationship between peak LOD and difference from threshold to peak demonstrated in figure 3 for LMM is closer to linear and less noisy than H-K. Since there is no direct comparison of the peak LOD scores on a pairwise basis, one possible interpretation of the figure is that improved performance of the LMM method shown in figure 3 is meaningfully affected by the results of the permutation

approach rather than the LOD score. If permutation testing using H-K tends to generate a nosier and higher significance threshold than LMM, that might account for the observed results. This would still be an interesting and a helpful improvement if the LMM is accurately setting a lower permutation threshold, but it would be helpful for the authors to comment on this variance and indicate whether the peak LOD scores or the size of the candidate region around that LOD peak are improved with LMM vs H-K.

The increased 'noise' in the H-K mapping is an artifact of the plotting: each phenotype is plotted according to its peak LOD calculated by LMM, and this produces an effect where the LMM difference in threshold looks cleaner.

Looking at the new Table S2, we can see that the difference in threshold between LMM and H-K is not substantial, with a mean difference of 0.04 LOD and a median of 0.03 LOD difference. Indeed, when looking at all phenotypes, the increase in peak LOD when using LMM vs H-K is not substantial ether, with LMM having a mean of 0.06 and a median of 0.07 higher peak LOD. Both of these suggest that using the LMM is not having a blanket improvement, but is only improving the 'true positives'.

We have added the below paragraph to page 13:

> *Note that there is not a substantial increase in LOD scores when using an LMM compared to H-K (a mean increase of 0.06 and a median of 0.07), nor is there a substantial change in the mean pgw threshold (LMM 3.51 vs H-K 3.55; Table S2).*