Life Science Alliance

Recursive splicing discovery using lariats in total RNA sequencing

Emma Hoppe, Dylan Udy, and Robert Bradley
DOI: https://doi.org/10.26508/lsa.202201889

Corresponding author(s): Robert Bradley, Fred Hutchinson Cancer Center

Transaction Report:

February 2, 2023

Re: Life Science Alliance manuscript #LSA-2022-01889

Dr. Robert K. Bradley
Fred Hutchinson Cancer Center
Computational Biology Program
FHCRC
1100 Fairview Avenue North
Seattle, WA 98109

Dear Dr. Bradley,

Thank you for submitting your manuscript entitled "Recursive splicing discovery using lariats in total RNA sequencing" to Life Science Alliance. The manuscript was assessed by expert reviewers, whose comments are appended to this letter. We invite you to submit a revised manuscript addressing the Reviewer comments.

To upload the revised version of your manuscript, please log in to your account: https://lsa.msubmit.net/cgi-bin/main.plex

You will be guided to complete the submission of your revised manuscript and to fill in all necessary information. Please get in touch in case you do not know or remember your login name.

While you are revising your manuscript, please also attend to the below editorial points to help expedite the publication of your manuscript. Please direct any editorial questions to the journal office.

The typical timeframe for revisions is three months. Please note that papers are generally considered through only one revision cycle, so strong support from the referees on the revised version is needed for acceptance.

When submitting the revision, please include a letter addressing the reviewers' comments point by point.

We hope that the comments below will prove constructive as your work progresses.

Thank you for this interesting contribution to Life Science Alliance. We are looking forward to receiving your revised manuscript.

Sincerely,

Eric Sawey, PhD
Executive Editor
Life Science Alliance
http://www.lsajournal.org

A. THESE ITEMS ARE REQUIRED FOR REVISIONS

-- A letter addressing the reviewers' comments point by point.

-- An editable version of the final text (.DOC or .DOCX) is needed for copyediting (no PDFs).

-- High-resolution figure, supplementary figure and video files uploaded as individual files: See our detailed guidelines for preparing your production-ready images, https://www.life-science-alliance.org/authors

-- Summary blurb (enter in submission system): A short text summarizing in a single sentence the study (max. 200 characters including spaces). This text is used in conjunction with the titles of papers, hence should be informative and complementary to the title and running title. It should describe the context and significance of the findings for a general readership; it should be written in the present tense and refer to the work in the third person. Author names should not be mentioned.

-- By submitting a revision, you attest that you are aware of our payment policies found here: https://www.life-science-alliance.org/copyright-license-fee

B. MANUSCRIPT ORGANIZATION AND FORMATTING:

Full guidelines are available on our Instructions for Authors page, https://www.life-science-alliance.org/authors

We encourage our authors to provide original source data, particularly uncropped/-processed electrophoretic blots and spreadsheets for the main figures of the manuscript. If you would like to add source data, we would welcome one PDF/Excel-file per figure for this information. These files will be linked online as supplementary "Source Data" files.

IMPORTANT: It is Life Science Alliance policy that if requested, original data images must be made available. Failure to provide original images upon request will result in unavoidable delays in publication. Please ensure that you have access to all original microscopy and blot data images before submitting your revision.

Reviewer #1 (Comments to the Authors (Required)):

Here the authors use computational approach to identify recursive splicing sites within introns. They used total RNA sequencing data and identify a number of new recursive splice sites not previously known. They show that long introns are not necessarily a requirement for recursive splicing, as previously published. They provide evidence that all sizes of introns, including small ones can use this kind of specialized splicing event. They find that while choice of such splice sites is stochastic, some may be conserved and regulated. The authors also provide examples where recursive splicing is used for exon exclusion. Overall, this clear and well-written manuscript provides support for more a broader use of recursive splicing in the human transcriptome. I support its publication.

Reviewer #2 (Comments to the Authors (Required)):

This study by Hoppe et al. aims to expand our knowledge of recursive splicing in human cells by using unbiased methods to identify and characterize sites of recursive splicing. To do so, they search for high-throughput sequencing reads that support the presence of intron lariats from recursive splicing. Using these sites, they are able to show a broader range of introns with recursive splicing, higher than expected conservation across vertebrates, and roles for these sites in regulating alternative splicing. Overall, this study is well written and presents solid analyses in support of an expanded role for recursive splicing in mammalian cells. However, there a few analysis details and points that I think need to be clarified before publication.
Specific Comments:
(1) I could not find a place where the authors detail the samples used for these analyses. In the methods, they say that they search the SRA to find human nascent RNA-seq datasets but give no context for which samples were ultimately chosen. I understand that they likely used a lot of samples, but think that there needs to be a mention somewhere in the Results or Methods section about the actual samples used to give context for how to compare their results with other studies. The authors similarly do not mention how many datasets or total reads were used, which provides important context for their power to identify recursive sites relative to other studies.
(2) On a related note, could the lack of overlap (Figure 1C) between the sites identified here and other studies be due to differences in cell types? The authors also do not show overlap with the Wan et al. 2021 paper, which identified recursive sites and uses nascent RNA-seq data.
(3) Can the authors comment at all on the probably that these recursive sites are being used consistently to aid in splicing of recursive introns or stochastically, such that any given recursive site is being used a minority of the time at random (as suggested by Wan et al. 2021)? This might be a useful discussion in the context of the suggestion that these recursive sites in shorter introns are aiding in the choice of alternative splice sites for those introns.

REVIEWER 1

Here the authors use computational approach to identify recursive splicing sites within introns. They used total RNA sequencing data and identify a number of new recursive splice sites not previously known. They show that long introns are not necessarily a requirement for recursive splicing, as previously published. They provide evidence that all sizes of introns, including small ones can use this kind of specialized splicing event. They find that while choice of such splice sites is stochastic, some may be conserved and regulated. The authors also provide examples where recursive splicing is used for exon exclusion. Overall, this clear and well-written manuscript provides support for more a broader use of recursive splicing in the human transcriptome. I support its publication.

We thank the reviewer for the kind and encouraging comments on our manuscript.

REVIEWER 2

This study by Hoppe et al. aims to expand our knowledge of recursive splicing in human cells by using unbiased methods to identify and characterize sites of recursive splicing. To do so, they search for high-throughput sequencing reads that support the presence of intron lariats from recursive splicing. Using these sites, they are able to show a broader range of introns with recursive splicing, higher than expected conservation across vertebrates, and roles for these sites in regulating alternative splicing. Overall, this study is well written and presents solid analyses in support of an expanded role for recursive splicing in mammalian cells. However, there a few analysis details and points that I think need to be clarified before publication.

We thank the reviewer for their appreciation of our study's contributions to the literature on recursive splicing.

Specific Comments:
I could not find a place where the authors detail the samples used for these analyses. In the methods, they say that they search the SRA to find human nascent RNA-seq datasets but give no context for which samples were ultimately chosen. I understand that they likely used a lot of samples, but think that there needs to be a mention somewhere in the Results or Methods section about the actual samples used to give context for how to compare their results with other studies. The authors similarly do not mention how many datasets or total reads were used, which provides important context for their power to identify recursive sites relative to other studies.

We thank the reviewer for pointing out this oversight. We edited **Table S1** so that it clearly states the relevant details and statistics for each sample that we analyzed in our study and additionally states the total number of reads that were included in our analysis over all samples. We additionally state these important statistics in the revised text. We edited the Methods ("Dataset selection and annotation") and Data Availability sections to clearly state that the relevant details and statistics for the samples that we analyzed are found in **Table S1**.

On a related note, could the lack of overlap (Figure 1C) between the sites identified here and other studies be due to differences in cell types?

We thank the reviewer for highlighting this important point. As the reviewer notes, the impact of cell type differences is an intriguing and important question to highlight that could contribute to the relatively low overlap that we observed. We edited the manuscript to emphasize this important point:

> The limited overlap that we observed could additionally arise in part from differences in the cell types that were analyzed in each study. Our analysis utilized a diverse selection of cell lines and patient-derived fibroblasts (**Table S1**) which were distinct from the cell types analyzed in other works (Duff et al. (2015): selected adult tissue samples; Sibley et al. (2015): adult brain tissue; Zhang et al. (2018): ovarian, embryonic stem cell, and embryonic stem cell-derived forebrain cell lines; Wan et al. (2021): bronchial epithelial cell line).

We thank the reviewer for this excellent suggestion. We revised the manuscript to include the recursive sites identified by Wan et al, 2021 in our analysis of overlap between different studies of recursive splicing. This new analysis is found in the revised **Table 1** and **Figure S2**.

Comment 3: *Can the authors comment at all on the probability that these recursive sites are being used consistently to aid in splicing of recursive introns or stochastically, such that any given recursive site is being used a minority of the time at random (as suggested by Wan et al. 2021)? This might be a useful discussion in the context of the suggestion that these recursive sites in shorter introns are aiding in the choice of alternative splice sites for those introns.*

We thank the reviewer for raising this interesting point. To tackle this question, we mapped lariats for conventional splicing across the genome and compared conventional versus recursive splicing rates in the introns for which recursive splicing had been identified. Although there was not enough data to definitively assess whether each intron appears to be regulated or stochastic (and this may depend on factors that the reviewer mentioned above, like cell type), we sought to identify whether there was a difference between shorter and longer introns in the usage of recursive versus conventional splicing. As suggested by Wan et al, 2021 and highlighted by the reviewer, this analysis supported a preference for conventional splicing among shorter introns, with a relatively strong association in the expected direction. We describe this new analysis in the main text in the section "Lariat sequencing identifies recursive splicing in diverse introns":

> Finally, we compared the relative rates of identification of conventional and recursive lariats for the introns in which we identified recursive splicing to determine whether recursive splicing appeared to be occurring infrequently or stochastically (as suggested by Wan et al., 2021) versus as a more common or dominant mode of splicing. Approximately half of the introns (46 of 100) with recursive lariats had at least one conventional lariat that we identified. For short introns (<1,000 nt), we observed a preference for lariats arising from conventional versus recursive splicing (**Fig S4A and S4B**). A Spearman's rank correlation performed on the relative number of unique recursive lariats with mismatches compared to the intron length in those introns with at least one lariat of each type revealed a positive association between intron length and the relative number of recursive lariats (ρ = 0.4912, p = 0.0005272). These data suggest that certain introns identified here, especially smaller introns, may be primarily subject to recursive splicing as a stochastic and less frequent mode of splicing, while others show a preference for recursive over conventional splicing.

We also highlighted this point in the revised Discussion:

> Our work here indicates that the frequency of use of recursive versus conventional splicing may be a function of intron length, with stochasticity playing a stronger role in smaller introns. However, further work will be needed to confirm this observation and elucidate the role, if any, recursive splicing plays in regulating alternative introns, as the work here primarily focuses on constitutive introns.

April 19, 2023

RE: Life Science Alliance Manuscript #LSA-2022-01889R

Dr. Robert K. Bradley
Fred Hutchinson Cancer Center
Computational Biology Program
FHCRC
1100 Fairview Avenue North
Seattle, WA 98109

Dear Dr. Bradley,

Thank you for submitting your revised manuscript entitled "Recursive splicing discovery using lariats in total RNA sequencing". We would be happy to publish your paper in Life Science Alliance pending final revisions necessary to meet our formatting guidelines.

Along with points mentioned below, please tend to the following:
-please upload your manuscript text as an editable doc file
-please upload your main and supplementary figures as single files and add a separate figure legend section to the main manuscript text
-please upload your table files as excel or doc files or make sure that they're included in the doc file of your main manuscript text
-please add the Twitter handle of your host institute/organization as well as your own or/and one of the authors in our system
-please use the [10 author names, et al.] format in your references (i.e. limit the author names to the first 10)

If you are planning a press release on your work, please inform us immediately to allow informing our production team and scheduling a release date.

LSA now encourages authors to provide a 30-60 second video where the study is briefly explained. We will use these videos on social media to promote the published paper and the presenting author (for examples, see https://twitter.com/LSAjournal/timelines/1437405065917124608). Corresponding or first-authors are welcome to submit the video. Please submit only one video per manuscript. The video can be emailed to contact@life-science-alliance.org

To upload the final version of your manuscript, please log in to your account: https://lsa.msubmit.net/cgi-bin/main.plex
You will be guided to complete the submission of your revised manuscript and to fill in all necessary information. Please get in touch in case you do not know or remember your login name.

To avoid unnecessary delays in the acceptance and publication of your paper, please read the following information carefully.

A. FINAL FILES:

These items are required for acceptance.

-- An editable version of the final text (.DOC or .DOCX) is needed for copyediting (no PDFs).

-- High-resolution figure, supplementary figure and video files uploaded as individual files: See our detailed guidelines for preparing your production-ready images, https://www.life-science-alliance.org/authors

-- Summary blurb (enter in submission system): A short text summarizing in a single sentence the study (max. 200 characters including spaces). This text is used in conjunction with the titles of papers, hence should be informative and complementary to the title. It should describe the context and significance of the findings for a general readership; it should be written in the present tense and refer to the work in the third person. Author names should not be mentioned.

B. MANUSCRIPT ORGANIZATION AND FORMATTING:

Full guidelines are available on our Instructions for Authors page, https://www.life-science-alliance.org/authors

We encourage our authors to provide original source data, particularly uncropped/-processed electrophoretic blots and spreadsheets for the main figures of the manuscript. If you would like to add source data, we would welcome one PDF/Excel-file per figure for this information. These files will be linked online as supplementary "Source Data" files.

Submission of a paper that does not conform to Life Science Alliance guidelines will delay the acceptance of your manuscript.

It is Life Science Alliance policy that if requested, original data images must be made available to the editors. Failure to provide original images upon request will result in unavoidable delays in publication. Please ensure that you have access to all original data images prior to final submission.

The license to publish form must be signed before your manuscript can be sent to production. A link to the electronic license to publish form will be sent to the corresponding author only. Please take a moment to check your funder requirements.

Reviews, decision letters, and point-by-point responses associated with peer-review at Life Science Alliance will be published online, alongside the manuscript. If you do want to opt out of having the reviewer reports and your point-by-point responses displayed, please let us know immediately.

Thank you for your attention to these final processing requirements. Please revise and format the manuscript and upload materials within 3 days.

Thank you for this interesting contribution, we look forward to publishing your paper in Life Science Alliance.

Sincerely,

Eric Sawey, PhD
Executive Editor
Life Science Alliance
http://www.lsajournal.org

April 21, 2023

RE: Life Science Alliance Manuscript #LSA-2022-01889RR

Dr. Robert K. Bradley
Fred Hutchinson Cancer Center
Computational Biology Program
FHCRC
1100 Fairview Avenue North
Seattle, WA 98109

Dear Dr. Bradley,

Thank you for submitting your Research Article entitled "Recursive splicing discovery using lariats in total RNA sequencing". It is a pleasure to let you know that your manuscript is now accepted for publication in Life Science Alliance. Congratulations on this interesting work.

The final published version of your manuscript will be deposited by us to PubMed Central upon online publication.

Your manuscript will now progress through copyediting and proofing. It is journal policy that authors provide original data upon request.

Reviews, decision letters, and point-by-point responses associated with peer-review at Life Science Alliance will be published online, alongside the manuscript. If you do want to opt out of having the reviewer reports and your point-by-point responses displayed, please let us know immediately.

IMPORTANT: If you will be unreachable at any time, please provide us with the email address of an alternate author. Failure to respond to routine queries may lead to unavoidable delays in publication.

Scheduling details will be available from our production department. You will receive proofs shortly before the publication date. Only essential corrections can be made at the proof stage so if there are any minor final changes you wish to make to the manuscript, please let the journal office know now.

DISTRIBUTION OF MATERIALS:
Authors are required to distribute freely any materials used in experiments published in Life Science Alliance. Authors are encouraged to deposit materials used in their studies to the appropriate repositories for distribution to researchers.

You can contact the journal office with any questions, contact@life-science-alliance.org

Again, congratulations on a very nice paper. I hope you found the review process to be constructive and are pleased with how the manuscript was handled editorially. We look forward to future exciting submissions from your lab.

Sincerely,

Eric Sawey, PhD
Executive Editor
Life Science Alliance
http://www.lsajournal.org