Identification of enhancer regulatory elements that direct epicardial gene expression during zebrafish heart regeneration

Yingxi Cao, Yu Xia, Joseph J. Balowski, Jianhong Ou, Lingyun Song, Alexias Safi, Timothy Curtis, Gregory E. Crawford, Kenneth D. Poss and Jingli Cao
DOI: 10.1242/dev.200133

Editor: Steve Wilson

Review timeline

Original submission:	23 August 2021
Editorial decision:	23 September 2021
First revision received:	23 November 2021
Editorial decision:	5 January 2022
Second revision received:	6 January 2022
Accepted:	11 January 2022

Original submission

First decision letter

MS ID#: DEVELOP/2021/200133

MS TITLE: Identification of enhancer regulatory elements that direct epicardial gene expression during zebrafish heart regeneration

AUTHORS: Yingxi Cao, Yu Xia, Joseph J Balowski, jianhong ou, Lingyun Song, Alexias Safi, Timothy Curtis, Gregory E Crawford, Kenneth D Poss, and Jingli Cao

I have now received all the referees' reports on the above manuscript, and have reached a decision. The referees' comments are appended below, or you can access them online: please go to BenchPress and click on the 'Manuscripts with Decisions' queue in the Author Area.

As you will see, the referees express considerable interest in your work, but have some significant criticisms and suggestions for improving your manuscript. If you are able to revise the manuscript along the lines suggested, I will be happy receive a revised version of the manuscript. Please also note that Development will normally permit only one round of major revision.

We are aware that you may be experiencing disruption to the normal running of your lab that make experimental revisions challenging. If it would be helpful, we encourage you to contact us to discuss your revision in greater detail. Please send us a point-by-point response indicating where you are able to address concerns raised (either experimentally or by changes to the text) and where you will not be able to do so within the normal timeframe of a revision. We will then provide further guidance. Please also note that we are happy to extend revision timeframes as necessary.

Please attend to all of the reviewers' comments and ensure that you clearly highlight all changes made in the revised manuscript. Please avoid using 'Tracked changes' in Word files as these are lost in PDF conversion. I should be grateful if you would also provide a point-by-point response detailing how you have dealt with the points raised by the reviewers in the 'Response to Reviewers' box. If you do not agree with any of their criticisms or suggestions please explain clearly why this is so.

Reviewer 1

Advance summary and potential significance to field

The group previously reported that regulatory elements termed tissue regeneration enhancer elements (TREEs) are activated by injury and regulate gene expression programs for regenerative responses (Kang et al. Nature 2016). The present work is a follow-up study to identify TREEs in epicardial cells. The authors performed ATAC-seq and H3K27ac ChIP-seq with epicardial cells purified from regenerating hearts of tcf21 reporter fish. They thoroughly characterized genomic regions undergoing an increase or decrease in accessibility, identifying many candidate TREEs in regenerating epicardial cells. They focused on the elements found in the novel epicardial marker genes rgmb, gnai3, and ncam1a, proving that TREEs regulate regeneration-specific gene programs in epicardial cells.

I liked the paper and support for publication in Development. The experiments were carefully performed, as evident in the validation of the isolated TREEs with several stable transgenic reporter lines. The data presented are of high quality and support the authors' conclusions well. The resource provided in this study will be helpful in future research in the field of developmental and regeneration biology.

Comments for the author

I have one comment that should be considered addressing. The authors characterized the ncam1a TREEs in sections (Figs. 6I, 7H) but not the rgmb and gnai3 TREEs (Figs. 4, 5). These TREEs may need to be also analyzed in sections to make the characterizations more convincing. The whole-mount confocal images indicate their activation in epicardial and perivascular cells (Figs. 4G and 5H). However, in situ hybridization signal for rgmb seems also positive in the endocardium at a similar level to that in the epicardium (Fig. 4E arrows). This may suggest the activation of the element in non-epicardial cells during regeneration.

Reviewer 2

Advance summary and potential significance to field

This study uses ATAC-seq and RNA-seq to map regulatory regions that are activated by injury in epicardial cells. To do this authors have used a transgenic zebrafish line expressing a nuclear localized EGF in epicardial cells (driven by the tcf21 promoter) and FACSorting was performed on uninjured and injured (3dpi and 7dpi) hearts. In the ATAC-seq data peaks with increased and decreased accessibility are identified and a motif search was performed to identify enriched motifs and their corresponding transcription factors that can bind to these motifs. In addition, the peaks with increased and decreased accessibility were linked to their nearest gene and correlated with changes in expression using the RNAseq data. To validate their findings the authors clone identified regulatory elements and test whether these in combination with a minimal promoter can regulate GFP expression in the injured heart by generating transgenic lines.

The epicardium is immediately activated upon injury and is an important cell type during heart regeneration. While its importance have been recognized, very little is known about the regulation of epicardial activation and how epicardial cells undergo EMT, invade the injury area and regulate pro-regenerative processes. This work provides a helpful resource by identifying genome-wide activation of regulatory elements and corresponding genes in epicardial cells during zebrafish heart regeneration.
However, several issues need to be addressed by the authors.

Comments for the author

Major comments:
1. The authors need to demonstrate that they have sorted a clean population of epicardial cells on which they performed the ATAC-seq and RNAseq analysis. While they do show some RNA-seq and ATACseq peaks of know epicardial genes in figure S1 (which they seem not to refer to), a better

characterization is needed. For example, is there any contamination from other cell types (cardiomyocytes, endothelial cells)?
Some more in situ hybridizations on genes that are induced by the injury in their data would be helpful to validate the dataset. In addition, epicardial cells form a heterogeneous population and tcf21 expression labels a subset of them (PMID: 32084358). This should be recognized.
2, The authors use Homer to look for enriched binding motifs in the ATACseq data and identify a list of putative transcription factors. Are the identified transcription factors expressed in epicardial cells and more importantly, induced by injury?
3. To validate their data the author selected putative regulatory elements and made transgenic lines to test their activity in vivo. The authors do not explain very well how they selected the regulatory elements that they used to generate the transgenic lines. In figure 1A they show the top 20 distal regions with the highest fold increase in accessibility. Why they continue only with regulatory elements close to ncam rgmb, gnai3 and not any of the other is not clearly described. It is also not clear whether the here described regulatory elements were the only elements that have been tested in vivo or whether more elements were initially included. To get a good overview of how predictable the ATACseq peaks are for biological activity the authors should provide a clear description of their selection procedure and give a complete overview of all the elements that have been tested.
4. To show the activity of the enhancers the authors performed whole mount imaging in most cases. These interpretation of these images can be hard sometimes (for example 4F). The red signal from the tcf21:H2A-mcherry is weak which may be due to the red color (another color like magenta may work better). Are the authors only showing one section of a confocal image or is this a stack? Is all expression that is visible only from the outer (epicardial layer) or is signal from deeper layers also shown? It would be helpful to complement the whole mount pictures with cross sections (such as in Fig 6I and 7H). In addition, the GFP expression is often very patchy (such as in 4F). What could be the reason for this patchy expression?
5. The authors in some cases make conclusions about the cell types that express the transgene (for example perivascular cells). Authors should either provide more evidence (e.g. colocalization with markers) or should be careful to make such statements.
6. The authors show that they can identify regulatory elements that are active in epicardial cells during heart regeneration. The example genes and elements are OK to demonstrate the application of the data the relevance of the selected genes for heart regeneration is not obvious (yet). The epicardium is important for several processes during heart regeneration including the regulation of cardiomyocyte proliferation by the expression of the growth factor neuregulin1. As very little is known about the regulation of Nrg1, potentially this dataset could help to stimulate future research into the mechanism of Nrg1 regulation. Therefore showing the RNA-seq data and ATAC-seq data in the Nrg1 locus could make a strong argument for the importance and relevance of this study.
7. The authors conclude that the identified injury-specific enhancers that are not active during development. This is not accurate as they show that some of these enhancers are active during development albeit in other cell types.

Reviewer 3

Advance summary and potential significance to field

Summary of content:
Using an ATAC-seq strategy based on sorted epicardial cells from injured vs. control ventricles in zebrafish the authors identify genome-wide signatures of epicardial-specific tissue regeneration enhancer elements (TREEs) expected to contribute to regenerative epicardial gene expression responses. The authors find that the enhancer mark H3K27ac is enriched predominantly in the category of regions showing gain of chromatin accessibility during regeneration which indicates TREE identity supported by motif and pathways enrichment analysis. RNA-seq analysis further demonstrates that the top 20 distal candidate TREEs (with highest fold increase in chromatin accessibility) are found nearby genes upregulated upon injury. For functional validation of such TREEs the authors then select elements in the vicinity of three of the genes associated to top enriched candidate TREEs and with likely relevant functions during heart regeneration (ncam1a, gnai3 and rgmb). Comparing fluorescent reporter transgenesis in injured vs. control ventricles the authors confirm TREE-specific activity (absent in the larval epicardium but activated upon injury) in the majority of elements tested (n=5/6). Together, these results add to the perception of TREEs as

a wide-spread regulatory unit in distinct regenerative processes and indicate that enhancer landscapes near genes with roles in regeneration are commonly harboring multiple TREEs with partially overlapping activities that likely contribute to transcriptional robustness.

Strengths and weaknesses:
Overall, while the concept of TREEs has been established and corroborated previously using alternative tissue sources, the authors convincingly demonstrate the validity of their strategy for identification of candidate TREEs in epicardial cells during cardiac regeneration and present a useful resource for TREE identification in future studies. The fact that a subset of these cis-regulatory elements appears evolutionarily conserved yields an interesting perspective for investigation of such elements in mammalian model systems.
As a primary weakness of the study, direct evidence that the TREEs validated by transgenic reporter analysis are functionally contributing to upregulation (or robust expression) of the associated target genes is missing and based on correlation of gene transcript signatures.

Novelty and significance:
Recent studies in zebrafish have elaborated on TREEs in other tissue types (e.g. PMID: 32883834) or have dissected individual TREE functionality (PMID: 33246928). However, to my knowledge the genomic enhancer landscapes specific to epicardial cells in the regenerating adult zebrafish epicardium have not yet been identified and published. The results presented in this study contribute to our understanding of the regulatory landscapes driving heart regeneration in zebrafish and lead to interesting predictions in terms of their regulatory architecture. Therefore, I consider the resources and conclusions offered by this study of general interest for the developmental biology community.

Quality:
The study is well written and follows a general logic. At a technical level the study appears solid by using appropriate genome-wide and statistical analysis tools as well as a convincing number of biological replicates for stage-specific analyses in ATAC-seq (n=4) and RNA-seq experiments (n=2), as well as transgenic reporter elements (2-4 transgenic lines have been established per element). However, in some cases the enhancer activities observed and described are rather correlative and it is difficult to conclude whether an enhancer-enhancer or gene-enhancer activity is overlapping or not (see comments below).

Comments for the author

I encourage revision of the manuscript by considering the following points:

Major points:
1) I consider the main weakness of this study that the validation of TREEs is restricted to correlation of transgenic enhancer-reporter activity and gene expression dynamics. While the proposition of in vivo enhancer deletion experiments to define the functional impact of the reporter-validated TREES on regulation of target gene expression might be too bold, the manuscript would gain from improved functional evidence that the identified enhancers are active specifically in cells upregulating expression levels of the associated target genes upon injury.
In addition, motif analysis in direct comparison with evolutionary conservation would be preferential to functionally link these enhancers to possible upstream regulatory pathways and lead to a better mechanistic understanding of the GRNs promoting cardiac regeneration.
2) In Fig. 5 how do the authors explain the differences of gnai3-E2:EGFP transgenic reporter activities seen in lines 1 and 4? While transgenic line 1 drives EGFP reporter activity around the injury side at 3 dpa, line 4 exhibits patches of positive cells in a broader area (including the remote zone), and generally more distant from the injury site.
The authors also should point out better the limitations of the use of transgenic reporters, that can lead to different intensities in the observed enhancer activity profiles, as seen in other lines (e.g. rgmb-E1:EGFP line1 and line2), and that can be related to effects of transgenic integration or activity (e.g. copy number).
3) Related to comment 2, while for rgbm (Fig.4E) and ncam1a (Fig. 6D) the authors provide in situ hybridization results to verify upregulation of these genes and localization of their transcripts at 3 dpa, such a result is lacking for gnai3, and it remains unclear in which cells or domain(s) a functional effect of the enhancer activity is to be expected.

It would be beneficial if the authors could add data to clarify the local distribution of gnai3 transcripts and comment on the correlated overlap with enhancer activities observed in the different transgenic lines.
4) It would be preferential if the authors could provide more depth about the large subset of chromatin regions identified that lose accessibility upon injury (e.g. motif context, other histone marks) and during the regeneration process (n=5588 at 3 dpa), and speculate on their function in relation to associated gene expression signatures.

Minor points:
1) The fact that the ncam1a-E4 enhancer activities are presented in context of evolutionary conservation in Fig. 3 (and the corresponding main text) but the related results from transgenic analyses are shown only in Fig. 6 appears somewhat confusing. The authors should restructure this context to keep the display of the ncam1a open chromatin/enhancer landscape next to the findings of the correspondin transgenic reporter (as for rgmb and gnai3 -related figures).
2) The authors might want to discuss their findings in relation to the recently published preprint "Distinct epicardial gene regulatory programmes drive development and regeneration of the zebrafish heart" (Weinberger et al., Biorxiv, 2021, doi: https://doi.org/10.1101/2021.06.29.450229) which uses a similar approach to define the epicardial enhancer landscapes during zebrafish heart regeneration.

First revision

Author response to reviewers' comments

Here, we list each suggestion of the reviewers and describe how we have addressed the suggestions in our revision.

Reviewer 1 Advance summary and potential significance to field
The group previously reported that regulatory elements termed tissue regeneration enhancer elements (TREEs) are activated by injury and regulate gene expression programs for regenerative responses (Kang et al. Nature 2016). The present work is a follow-up study to identify TREEs in epicardial cells. The authors performed ATAC-seq and H3K27ac ChIP-seq with epicardial cells purified from regenerating hearts of tcf21 reporter fish. They thoroughly characterized genomic regions undergoing an increase or decrease in accessibility, identifying many candidate TREEs in regenerating epicardial cells. They focused on the elements found in the novel epicardial marker genes rgmb, gnai3, and ncam1a, proving that TREEs regulate regeneration-specific gene programs in epicardial cells. I liked the paper and support for publication in Development. The experiments were carefully performed, as evident in the validation of the isolated TREEs with several stable transgenic reporter lines. The data presented are of high quality and support the authors' conclusions well. The resource provided in this study will be helpful in future research in the field of developmental and regeneration biology.

A: We thank the reviewer for the effort to review our manuscript and the supportive comments.

Reviewer 1 Comments for the author
I have one comment that should be considered addressing. The authors characterized the ncam1a TREEs in sections (Figs. 6I, 7H) but not the rgmb and gnai3 TREEs (Figs. 4, 5). These TREEs may need to be also analyzed in sections to make the characterizations more convincing. The whole-mount confocal images indicate their activation in epicardial and perivascular cells (Figs. 4G and 5H). However, in situ hybridization signal for rgmb seems also positive in the endocardium at a similar level to that in the epicardium (Fig. 4E, arrows). This may suggest the activation of the element in non-epicardial cells during regeneration.

A: Thank you for this suggestion. We have now included section images for the rgmb and gnai3 TREEs in the revised Figures 6I, 7F, and 8F. As you can see in Figure 6I, rgmb-E1:EGFP is expressed in the surface layers of the ventricle but not in the endocardium. Thus, the rgmb-E1 enhancer does not have endocardial activity. As we noted in the manuscript, rgmb-E1 activity may contribute

partially to rgmb expression. The following sentence is now included on Page 14: “However, numerous rgmb+EGFP- cells were observed in the epicardium, indicating that rgmb-E1 only contributes partially to the gene activity.”

Reviewer 2 Advance summary and potential significance to field
This study uses ATAC-seq and RNA-seq to map regulatory regions that are activated by injury in epicardial cells. To do this, authors have used a transgenic zebrafish line expressing a nuclear localized EGFP in epicardial cells (driven by the tcf21 promoter) and FACSorting was performed on uninjured and injured (3dpi and 7dpi) hearts. In the ATAC-seq data peaks with increased and decreased accessibility are identified and a motif search was performed to identify enriched motifs and their corresponding transcription factors that can bind to these motifs. In addition, the peaks with increased and decreased accessibility were linked to their nearest gene and correlated with changes in expression using the RNAseq data. To validate their findings the authors clone identified regulatory elements and test whether these in combination with a minimal promoter can regulate GFP expression in the injured heart by generating transgenic lines.

The epicardium is immediately activated upon injury and is an important cell type during heart regeneration. While its importance has been recognized, very little is known about the regulation of epicardial activation and how epicardial cells undergo EMT, invade the injury area and regulate pro-regenerative processes. This work provides a helpful resource by identifying genome-wide activation of regulatory elements and corresponding genes in epicardial cells during zebrafish heart regeneration. However, several issues need to be addressed by the authors.

A: We thank the reviewer for the effort to review our manuscript and the insightful critiques to improve the work.

Reviewer 2 Comments for the author

Major comments:
1. The authors need to demonstrate that they have sorted a clean population of epicardial cells on which they performed the ATAC-seq and RNAseq analysis. While they do show some RNA-seq and ATACseq peaks of know epicardial genes in figure S1 (which they seem not to refer to), a better characterization is needed. For example, is there any contamination from other cell types (cardiomyocytes, endothelial cells)? Some more in situ hybridizations on genes that are induced by the injury in their data would be helpful to validate the dataset. In addition, epicardial cells form a heterogeneous population and tcf21 expression labels a subset of them (PMID: 32084358). This should be recognized.

A: Thank you for this comment. We have previously validated our method for efficient and high-purity isolation of live tcf21:nucEGFP+ cells from adult zebrafish hearts (Cao et al., Development, 2016). A published Figure panel is shown below for your reference.

After each FACS sorting, we plated the isolated cells in dishes or on a coverslip and observed > 95% purity of EGFP+ cells. We have included an image as new Figure 1B. We believe this is a better method than q-PCR of cell type markers to validate the purity. It is not possible to rule out trace amount contamination from cardiomyocytes, endothelial cells, and others. However, our purity ensures data integrity. We have added the following sentence in the method section on Page 18: “After FACS sorting, we plated the isolated cells in dishes or on a coverslip and observed > 95% purity of EGFP+ cells.”

As suggested, we have now included in situ hybridization images of a few upregulated genes, including fstl1a, fstl1b, plcxd3, triqk, plod2, arhgap4a, and gnai3 in the new Figures S3 and 7D. As you can see, these genes except arhgap4a are induced in the surface layer of the ventricle enriched in epicardial cells. In additional, many other upregulated genes (Table S3) identified from our datasets have been validated previously for injury-induced epicardial cell expression in our lab, such as mdka, tmsb4x, p4hb, hspa5, timp2b, fn1a, and fn1b (Cao et al., Development, 2016; Wang et al., Developmental Biology, 2013). We have now included the track images of fn1a in Figure S1. These results suggest that our datasets are reliable for the discovery of epicardial regulatory programs.

Lastly, we have recognized the limitation of using the tcf21 reporter and cited the reference (PMID: 32084358). We also noted that the tcf21 reporter is the best available reagent so far for isolating zebrafish epicardial cells in Paragraph 1 on Page 5: “To isolate epicardial cells, we used an EGFP reporter driven by the regulatory sequences of tcf21. Although epicardial cells are a heterogeneous population and tcf21 conceivably does not label the entire population, it is the best available pan-epicardial marker that labels both quiescent and injury-responding epicardial cells in zebrafish (Cao et al., 2016; Kikuchi et al., 2011a; Weinberger et al., 2020).”

2, The authors use Homer to look for enriched binding motifs in the ATACseq data and identify a list of putative transcription factors. Are the identified transcription factors expressed in epicardial cells and more importantly, induced by injury?

A: Thank you for the comment. As shown below in the MA plot and Table S3 (3 dpa vs. Ctrl), these TFs are expressed (or even highly expressed like junba and tcf21) in epicardial cells but not always induced by injury.

Characterizing expression of each of these many TFs is not a goal of the current manuscript. However, it was previously reported that epicardial TGF-beta signaling contributes to zebrafish heart regeneration (Chablais et al., Development, 2012), consistent with our observation of enrichment of Smad2/3/4 binding sites in epicardial enhancers. This also applies to other TFs with reported epicardial functions, including Tcf21 (Hu et al., 2020), Runx1 (Koth et al., 2020), TEADs (Xiao et al., 2018), C/EBPb (Huang et al., 2012), and Gli2 (Choi et al., 2013; Sugimoto et al., 2017; Wang et al., 2015). In addition, we have now included an in situ hybridization result in Figure 2F to show expression of junba in the presumed epicardium both before and after heart injury.

3. To validate their data the author selected putative regulatory elements and made transgenic lines to test their activity in vivo. The authors do not explain very well how they selected the regulatory elements that they used to generate the transgenic lines. In figure 1A they show the top 20 distal regions with the highest fold increase in accessibility. Why they continue only with regulatory elements close to ncam, rgmb, gnai3 and not any of the other is not clearly described. It is also not clear whether the here described regulatory elements were the only elements that have been tested in vivo or whether more elements were initially included. To get a good overview of how predictable the ATACseq peaks are for biological activity the authors should provide a clear description of their selection procedure and give a complete overview of all the elements that have been tested.

A: Thank you for this critique. We have clarified our selection criteria in the revised manuscript on Pages 9-10. For the top 20 distal regions (Figure 3A), only 9 are linked to upregulated genes or contain conserved sequences (Figure 3B, D). The linked genes are arhgap4a, plcxd3, fn1a, gnai3, triqk, rgmb, plod2, fam98b, and ncam1a. Further in situ hybridization analysis (Figure S3) narrowed this list to 7 genes with confirmed epicardial expression upon heart injury (plcxd3, fn1a, gnai3, triqk, rgmb, plod2, and ncam1a). With the goal to find novel regulatory programs of epicardial regeneration, we further prioritized ncam1a and rgmb for their role in neuronal development (Table S5, enrichment annotation results) and gnai3 for its G protein property. We had then successfully cloned 6 candidate regions for these 3 genes and established 21 stable lines for these regions. As shown in the current study, 5 of 6 candidate enhancers (~83%) have injury-induced epicardial activity, suggesting that our dataset and the prioritization strategy are promising for discovering new epicardial regulators. Using these 3 genes as an example, we expect this work will provide a resource to foster further studies in the field.

4. To show the activity of the enhancers the authors performed whole mount imaging in most cases. These interpretation of these images can be hard sometimes (for example 4F). The red signal from the tcf21:H2A-mcherry is weak which may be due to the red color (another color like magenta may work better). Are the authors only showing one section of a confocal image or is this a stack? Is all expression that is visible only from the outer (epicardial layer) or is signal from deeper layers also shown? It would be helpful to complement the whole mount pictures with cross sections (such as in Fig 6I and 7H). In addition, the GFP expression is often very patchy (such as in 4F). What could be the reason for this patchy expression?

A: Thank you for pointing out these issues. We have switched these mCherry signals to magenta color in all figures. The weak red channel is likely caused by PDF conversion and low magnification. These are whole-mount confocal images, and all EGFP and mCherry signals were captured through multiple z stacks from the ventricular surface to the trabecular muscle layer until no signal was detectable. The low-magnification images are maximum projections of z stacks to show the big picture of expression patterns, while the high-mag view panels (including the single-channel images) are from single optical sections to assess colocalization. We have clarified this in Figure Legends.

We have now included section images for rgmb and gani3 TREEs in the revised Figures 6I, 7F, and 8F, which indicate restricted EGFP expression to the ventricular surface.

We also noticed the patchy expression pattern of rgmb-E1:EGFP (in the revised Figure 6F), which is consistent across hearts. The EGFP signals are often close to large vessels (see below). This may suggest pro-angiogenic or related functions of rgmb, which could be addressed in a follow-up study. The following language is included in Paragraph 1 on Page 14: “We noticed patchy EGFP expression that was often adjacent to vessels (Figure 6F-H). This may suggest pro-angiogenic or related functions of rgmb.”

5. The authors in some cases make conclusions about the cell types that express the transgene (for example perivascular cells). Authors should either provide more evidence (e.g. colocalization with markers) or should be careful to make such statements.

A: Thank you for the suggestion. In the revised manuscript, we included Hybridization Chain Reaction (HCR) staining image of pdgfrb, a marker of cardiac mural cells (PMID 34310924 and bioRxiv 2021.04.27.441161; doi: https://doi.org/10.1101/2021.04.27.441161). As shown in the new Figures 4J, 6H, and S6G, the EGFP+ cells that aligned in parallel are often pdgfrb+. We have revised our statements to include parallel alignment and pdgfrb expression as criteria for potential perivascular cell properties in Paragraph 1 on Page 12: “In addition, we observed mCherry+EGFP+ cells aligned in parallel and these EGFP+ cells expressed the mural cell marker pdgfrb (Ando et al., 2021) (Figure 4G, arrowheads; Figure 4J, arrows), suggesting ncam1a-E2 activity in tcf21+ perivascular cells.”

6. The authors show that they can identify regulatory elements that are active in epicardial cells during heart regeneration. The example genes and elements are OK to demonstrate the application of the data, the relevance of the selected genes for heart regeneration is not obvious (yet). The epicardium is important for several processes during heart regeneration including the regulation of cardiomyocyte proliferation by the expression of the growth factor neuregulin1. As very little is known about the regulation of Nrg1, potentially this dataset could help to stimulate future research into the mechanism of Nrg1 regulation. Therefore showing the RNA-seq data and ATAC-seq data in the Nrg1 locus could make a strong argument for the importance and relevance of this study.

A: Thank you for the suggestion. We have now shown RNA-seq and ATAC-seq data at the nrg1 locus in new Figure S1 and Table S3. We also included a prediction of potential TF binding sites in the promoter regions and putative enhancer regions in new Figure S2 and Table S4. We detected at least two distinct nrg1 transcripts (nrg1-202, nrg1-205) in our samples and identified at least 3 nrg1-linked putative enhancer regions. We did not observe injury-induced nrg1 expression (3 dpa vs. ctrl, FDR = 0.296; 7 dpa vs ctrl, FDR = 0.08; Table S3). This is possibly due to the low expression level of nrg1 (only detectable by using the RNAscope technique) and the limited expression in a small subpopulation of tcf21+ cells (Gemberling et al., eLife, 2015). However, ATAC-seq and Chip-seq results indicated promising promoter and enhancer regions. Further motif analysis of these 5 regions indicated the presence of numerous binding sites for AP-1 subunits, Retinoid X Receptors (RXRA, RXRG), and Retinoic Acid Receptors (RARA, RARB, RARG) (Figure S2 and Table S4). This result suggests that the AP-1 complex and RA signaling may regulate Nrg1 expression during heart regeneration, which warrants further studies.

In addition, Follistatin-like 1 (Fstl1) is another epicardium-derived growth factor that supports CM proliferation in mammalian heart injury models (Wei et al., Nature, 2015). We have included profiling and in situ hybridization results of the zebrafish orthologs – fstl1a and fstl1b in Figure S3. As you can see, we identified transcript level increases and several ATAC-seq peaks with increased

accessibility for both genes (Figure S3A, B) during regeneration. In situ hybridization results demonstrated injury-induced expression in apparent epicardial cells for both genes. These results are now included in Paragraph 2 on Page 7. They further support the importance and relevance of our datasets in detecting epicardial genes relevant to regeneration and as candidate TREEs.

7. The authors conclude that they identified injury-specific enhancers that are not active during development. This is not accurate as they show that some of these enhancers are active during development albeit in other cell types.

A: Thank you for this comment. We have revised our statement in the Conclusions section to clarify that these enhancers are not active in the developing epicardium: “By contrast, the epicardial TREEs we identified direct injury-induced but not developmental expression in the epicardium, suggesting they are customized to the epicardial regeneration machinery.”

Reviewer 3 Advance summary and potential significance to field

Summary of content:
Using an ATAC-seq strategy based on sorted epicardial cells from injured vs. control ventricles in zebrafish the authors identify genome-wide signatures of epicardial-specific tissue regeneration enhancer elements (TREEs) expected to contribute to regenerative epicardial gene expression responses. The authors find that the enhancer mark H3K27ac is enriched predominantly in the category of regions showing gain of chromatin accessibility during regeneration which indicates TREE identity supported by motif and pathways enrichment analysis. RNA-seq analysis further demonstrates that the top 20 distal candidate TREEs (with highest fold increase in chromatin accessibility) are found nearby genes upregulated upon injury. For functional validation of such TREEs the authors then select elements in the vicinity of three of the genes associated to top enriched candidate TREEs and with likely relevant functions during heart regeneration (ncam1a, gnai3 and rgmb). Comparing fluorescent reporter transgenesis in injured vs. control ventricles the authors confirm TREE-specific activity (absent in the larval epicardium but activated upon injury) in the majority of elements tested (n=5/6). Together, these results add to the perception of TREEs as a wide-spread regulatory unit in distinct regenerative processes and indicate that enhancer landscapes near genes with roles in regeneration are commonly harboring multiple TREEs with partially overlapping activities that likely contribute to transcriptional robustness.

Strengths and weaknesses:
Overall, while the concept of TREEs has been established and corroborated previously using alternative tissue sources, the authors convincingly demonstrate the validity of their strategy for identification of candidate TREEs in epicardial cells during cardiac regeneration and present a useful resource for TREE identification in future studies. The fact that a subset of these cis-regulatory elements appears evolutionarily conserved yields an interesting perspective for investigation of such elements in mammalian model systems. As a primary weakness of the study, direct evidence that the TREEs validated by transgenic reporter analysis are functionally contributing to upregulation (or robust expression) of the associated target genes is missing and based on correlation of gene transcript signatures.

Novelty and significance:
Recent studies in zebrafish have elaborated on TREEs in other tissue types (e.g. PMID: 32883834) or have dissected individual TREE functionality (PMID: 33246928). However, to my knowledge the genomic enhancer landscapes specific to epicardial cells in the regenerating adult zebrafish epicardium have not yet been identified and published. The results presented in this study contribute to our understanding of the regulatory landscapes driving heart regeneration in zebrafish and lead to interesting predictions in terms of their regulatory architecture. Therefore, I consider the resources and conclusions offered by this study of general interest for the developmental biology community.

Quality:
The study is well written and follows a general logic. At a technical level the study appears solid by using appropriate genome-wide and statistical analysis tools as well as a convincing number of biological replicates for stage-specific analyses in ATAC-seq (n=4) and RNA-seq experiments (n=2),

as well as transgenic reporter elements (2-4 transgenic lines have been established per element). However, in some cases the enhancer activities observed and described are rather correlative and it is difficult to conclude whether an enhancer-enhancer or gene-enhancer activity is overlapping or not (see comments below).

A: We thank the reviewer for the effort to review our manuscript and the insightful critiques to improve the work.

Reviewer 3 Comments for the author

I encourage revision of the manuscript by considering the following points:

Major points:
1) I consider the main weakness of this study that the validation of TREEs is restricted to correlation of transgenic enhancer-reporter activity and gene expression dynamics. While the proposition of in vivo enhancer deletion experiments to define the functional impact of the reporter-validated TREES on regulation of target gene expression might be too bold, the manuscript would gain from improved functional evidence that the identified enhancers are active specifically in cells upregulating expression levels of the associated target genes upon injury. In addition, motif analysis in direct comparison with evolutionary conservation would be preferential to functionally link these enhancers to possible upstream regulatory pathways and lead to a better mechanistic understanding of the GRNs promoting cardiac regeneration.

A: Thank you for this comment. We have now included Hybridization Chain Reaction (HCR) staining results for genes ncam1a, rgmb, and gnai3. As shown in the revised Figures 4L, 5I, 5J, 6J, 7G, and 8G, EGFP signals of the TREE reporters are colocalized with expressions of the associated target genes. No definitive epicardial expression of these genes was observed in the Ctrl samples. Some nonspecific background staining is evident in the muscle (noted in the Figure legend). In addition, we noticed that the gene expression domains are broader than the EGFP signals of the linked enhancers for ncam1a-E2, rgmb-E1, gnai3-E1, and gnai3-E2, suggesting that these enhancers may contribute partially to activities of the linked genes. This comment has been added in the relevant paragraphs.

As suggested, we have included motif analysis of the differentially regulated ATAC-seq peaks (3 dpa vs. Ctrl) that contain conserved sequences as new Figure 3E. Interestingly, the enriched motifs are very similar to those of all differentially regulated ATAC-seq peaks (3 dpa vs. Ctrl, Figure 2E), with AP-1 motifs topping the positive regulators and Tcf21 and WT1 motifs leading the negative regulators. This result may suggest conserved epicardial machinery that mediates heart regeneration and imply translational potentials in activating similar programs to promote heart repair.

2) In Fig. 5 how do the authors explain the differences of gnai3-E2:EGFP transgenic reporter activities seen in lines 1 and 4? While transgenic line 1 drives EGFP reporter activity around the injury side at 3 dpa, line 4 exhibits patches of positive cells in a broader area (including the remote zone), and generally more distant from the injury site. The authors also should point out better the limitations of the use of transgenic reporters, that can lead to different intensities in the observed enhancer activity profiles, as seen in other lines (e.g. rgmb-E1:EGFP line1 and line2), and that can be related to effects of transgenic integration or activity (e.g. copy number).

A: Thank you for pointing this out. For this particular gnai3-E2:EGFP reporter line 1, the original image was from a heart with a relatively small injury. We have examined numerous hearts and confirmed a similar expression pattern to that of line 4. To avoid misleading, we have replaced that image with one from a heart bearing a standard wound size (revised Figure 8C). We agree that the insertion regions in the genome and the copy numbers might affect the reporter activities. We have included the following statement to recognize the limitation in the Conclusions section: “In our study, we noticed variations between stable lines of the same enhancer, likely explained by different genome insertion sites and copy numbers among lines. Establishing and analyzing multiple lines for each candidate enhancer is critical to define the activity patterns with this transgenic strategy.”

3) Related to comment 2, while for rgbm (Fig.4E) and ncam1a (Fig. 6D) the authors provide in situ hybridization results to verify upregulation of these genes and localization of their transcripts at 3 dpa, such a result is lacking for gnai3, and it remains unclear in which cells or domain(s) a functional effect of the enhancer activity is to be expected. It would be beneficial if the authors could add data to clarify the local distribution of gnai3 transcripts and comment on the correlated overlap with enhancer activities observed in the different transgenic lines.

A: We apologize for this oversight. The in situ hybridization result of gnai3 has now been included as new Figure 7D, which demonstrated injury-induced gene expression in the presumed epicardial cells. In addition, as noted above, we have now included HCR staining results for genes ncam1a, rgmb, and gnai3. As shown in the revised Figures 4L, 5I, 5J, 6J, 7G, and 8G, EGFP signals of the TREE reporters are colocalized with the expressions of the associated target genes. In addition, we noticed that the gene expression domains are broader than the EGFP signals of the linked enhancers for ncam1a-E2, rgmb-E1, gnai3-E1, and gnai3-E2, suggesting that these enhancers may contribute partially to activities of the linked genes.

4) It would be preferential if the authors could provide more depth about the large subset of chromatin regions identified that lose accessibility upon injury (e.g. motif context, other histone marks) and during the regeneration process (n=5588 at 3 dpa), and speculate on their function in relation to associated gene expression signatures.

A: Thank you for the suggestion. In the revised manuscript, we have included motif analysis results of these downregulated ATAC-seq regions in Figure 3E and revised the text to reflect the new results on Pages 7-8: “To identify candidate transcriptional regulators active in epicardial cells during heart regeneration, we assayed for enriched nucleotide motifs within regions with differential accessibility at 3 dpa (vs. Ctrl) using HOMER (Heinz et al., 2010)… Interestingly, the top hits of regions with decreased accessibility are binding motifs of Tcf21 and WT1, signature TFs of the epicardium. This may suggest a transition in cell state, which may warrant further investigation. Other enriched top hits include motifs belonging to Gata6 (Kolander et al., 2014), ERG (ETS Transcription Factor ERG), Meis1 (Crespillo et al., 2021; Huang et al., 2012), and Foxo3 (Figure 2E).”

For the histone marks, although the decreased ATAC-seq regions also show H3K27Ac marks (Figure 2A), they do not show changes in the H3K27Ac signature (Figure 2B, bottom, ratio = 1). This is mentioned in Paragraph 2 on Page 6.

Minor points:
1)The fact that the ncam1a-E4 enhancer activities are presented in context of evolutionary conservation in Fig. 3 (and the corresponding main text) but the related results from transgenic analyses are shown only in Fig. 6 appears somewhat confusing. The authors should restructure this context to keep the display of the ncam1a open chromatin/enhancer landscape next to the findings of the corresponding transgenic reporter (as for rgmb and gnai3 -related figures).

A: Thank you for the suggestion. We have rearranged the figures accordingly.

2)The authors might want to discuss their findings in relation to the recently published preprint “Distinct epicardial gene regulatory programmes drive development and regeneration of the zebrafish heart” (Weinberger et al., Biorxiv, 2021, doi: https://doi.org/10.1101/2021.06.29.450229) which uses a similar approach to define the epicardial enhancer landscapes during zebrafish heart regeneration.

A: We have now cited this preprint and echoed the finding of AP-1 as a major regulator of epicardial development and regeneration response in Paragraph 1 on Page 8: “In agreement with our finding, a recent preprint also reported the increased presence of the AP-1 complex subunit binding motifs in ATAC-seq peaks preferentially accessible in the injured epicardium (Weinberger et al., 2021).”

We also recognized the complementary nature of both studies in the Conclusions section: “Notably, a recent preprint used a similar approach to identify epicardial enhancers, reporting distinct regulatory programs during epicardial development and regeneration (Weinberger et al., 2021).

Enhancers linked to genes loxa, ppfibp1a, col12a1a, and mdka were found to be sufficient to direct gene expression in the embryonic epicardium, although whether they have enhancer activity during regeneration was not addressed."

Second decision letter

MS ID#: DEVELOP/2021/200133

MS TITLE: Identification of enhancer regulatory elements that direct epicardial gene expression during zebrafish heart regeneration

AUTHORS: Yingxi Cao, Yu Xia, Joseph J Balowski, Jianhong Ou, Lingyun Song, Alexias Safi, Timothy Curtis, Gregory E Crawford, Kenneth D Poss, and Jingli Cao

I have now received all the referees reports on the above manuscript, and have reached a decision. The referees' comments are appended below, or you can access them online: please go to BenchPress and click on the 'Manuscripts with Decisions' queue in the Author Area.

The referees are happy with yourrevisionsand we will be happy to publish your manuscript in Development. However, as a research article, Development requires separate Results and Discussion sections and so you will need to alter the text to conform to our formatting guidelines prior to acceptance.

Reviewer 2

Advance summary and potential significance to field

The epicardium is immediately activated upon injury and is an important cell type during heart regeneration.
While its importance have been recognized, very little is known about the regulation of epicardial activation and how epicardial cells undergo EMT, invade the injury area and regulate pro-regenerative processes. This work provides a helpful resource by identifying genome-wide activation of regulatory elements and corresponding genes in epicardial cells during zebrafish heart regeneration.

Comments for the author

In this revised version of their manuscript the authors have adequately addressed all questions. The new data that is included in the figures and the revised display of the data has improved the quality of the manuscript and I want to congratulate the authors with their results.

Reviewer 3

Advance summary and potential significance to field

The current study provides a valuable resource for relevant cis-regulatory elements involved in the control of injury-induced epicardial gene expression during zebrafish heart regeneration. These results extend our understanding of the epicardial gene regulatory networks underlying cardiac regeneration by uncovering the dynamic units of critical enhancer landscapes and elucidating relevant underlying TF motifs. Such transcriptional enhancers (and related motif grammar) can be utilized to direct customized injury-induced expression of (multipurpose) transgenic cassettes and/or pro-regenerative factors, with the potential to serve as genetic driver sequences in therapeutic applications to promote regeneration of the mammalian heart.
Comments for the author

The authors have addressed all my comments and concerns in a detailed and satisfactory manner and overall significantly improved their manuscript during this revision process. I have no further comments and support publication.

Second revision

Author response to reviewers' comments

1. We reduced the Abstract to less than 180 words.
2. We separated the Results and Discussion sections. The former Conclusions section is now the Discussion section.
3. We made a few other minor edits for clarifications, which are marked in red.

Third decision letter

MS ID#: DEVELOP/2021/200133

MS TITLE: Identification of enhancer regulatory elements that direct epicardial gene expression during zebrafish heart regeneration

AUTHORS: Yingxi Cao, Yu Xia, Joseph J Balowski, Jianhong Ou, Lingyun Song, Alexias Safi, Timothy Curtis, Gregory E Crawford, Kenneth D Poss, and Jingli Cao
ARTICLE TYPE: Techniques and Resources Article

I am happy to tell you that your manuscript has been accepted for publication in Development, pending our standard ethics checks.