nature portfolio

Peer Review File

Functional analysis of ESRP1/2 gene variants and CTNND1 isoforms in orofacial cleft pathogenesis

Corresponding Author: Dr Eric Liao

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript, “Functional analysis of ESRP1/2 gene variants and identification of Esrp-regulated CTNND1 isoforms in orofacial cleft pathogenesis” demonstrates the variable function of known ESRP variants and highlights key mRNA targets in cleft palate development. The authors used two systems to test the role of ESRP1 and ESRP2 variants found in humans —zebrafish model with esrp1 knocked out and esrp2 knocked down with translation-blocking morpholino, and murine cell line with Esrp1 and Esrp2 doubly knocked out. These models had clear phenotypes they could utilize for their assays; the zebrafish model showed defective morphology of the anterior neurocranium, and the double KO mouse cell line was assayed for splicing defects in many Esrp1/2 targets. The authors next performed "rescue" experiments with the human variants of ESRP1 and ESRP2, and assessed which variants are functional and which are pathogenic. They tested 7 ESRP1 and 12 ESRP2 variants and found that most are benign mutations—when expressed in the models they can rescue the phenotypes. Out of the 19, only 1 frameshift mutation in ESRP1 and 1 missense and 1 nonsense mutation in ESRP2 were found to be pathogenic (1 missense mutation in ESRP2 was only pathogenic in the murine cell line). The two models were mostly in agreement. The authors show that the in vivo and in vitro testing is critical for determination of pathogenicity of genetic variants because many in silico methods including AlphaMissense did not accurately predict the results.
Next part of the paper was slightly hard to follow in terms of flow. The authors shifted focus on Ctnnd1 and its splicing, and localization of isoforms. They showed that overexpression of Ctnnd1 is able to rescue the morphological phenotype in zebrafish model. It seemed like there is a logical jump from the previous section. Perhaps they are looking for downstream targets of ESRP1/2 that are associated with the OFC phenotype? It would be good if the authors can explain why they focused on Ctnnd1.

Overall, the manuscript provides critical, new information about ESRPs and I recommend it for publication. I have a few comments that can improve the manuscript further.

1. The authors use Arghef11 splicing as a functional readout for ESRP1/2 mutations. Authors should consider adding 2-3 more known targets to further validate this result, preferably with targets related to the phenotype such as CTNND1.

2. Ctnnd1 is a known ESRP target. This should be emphasized in the section “Alternative splicing to generate epithelial isoform of Ctnnd1 requires Esrp 1/2 function”.

3. The last section in the results feels disjointed with the rest of the manuscript. This section should be included prior to the data in Figure 6 and the data should be represented in a figure.

Minor:
1. Typo – Lines 406-407. “As because”.

2. Figure 1A – Arrows to indicate the fused otoliths.

3. Figure 4 needs letters to indicate panels for clarity. Image quality is not great.

4. In the discussion, the emphasis on the issues with bioinformatics is too heavy. It is important to acknowledge these issues so other researchers are aware, however, the authors should more strongly emphasize the importance of the results from their physiological experiments. As written, these results feel underwhelming.

Reviewer #2

(Remarks to the Author)
This is a very well-written and organized study that used zebrafish and Py2T in vitro systems to uncover the pathological function of coding mutations in ESRP1/2 and CTNND1 genes. The study emphasized the importance of experimental assays to determine functional gene variants compared to in silico computational predictive models, including SIFT, PolyPhen-2, and AlphaMissense. The manuscript also identified CTNND1 as a cleft lip and palate gene that harbors a large number of gene mutations which account for almost 1.5% of cleft lip/palate (CL/P) cases, making this gene the second most important OFC after IRF6. The study tested a good number of gene variants in ESRP1/2 using a zebrafish system to evaluate the prediction efficiency of the predictive computational models and concluded that the in silico models overestimated the pathological impact of coding variants.

There are a few concerns that need to be addressed to improve the quality of the manuscript. The statements used to compare the predictive accuracy of computer models were very broad without sufficient evaluation and analysis of previous studies that conducted similar assays to test the function of coding variants in vitro and in vivo compared to predictive models. The study cited a few publications that were consistent with the study's findings but failed to include a sufficient number of studies to justify the generalized statements.

The limitations of the current study are the use of genetic knockout of Esrp1 and morpholinos to knockdown Esrp2 for creating an efficient in vivo system to test a good number of coding variants, while testing a smaller number of variants in vitro. Using this type of in vivo system introduces several layers of complexity which might add noises that could impact the efficiency of the pathogenicity test. One discrepancy between the zebrafish and mouse model is the gene duplications and paralogues in zebrafish which required the knockout of both Esrp1 and 2 to create OFC phenotype, while knockout of Esrp1 alone was sufficient to cause CLP in mice. Also, the study didn't test all the gene variants in Py2T cells, which were tested in zebrafish. It was not clear why they decided to test less than 50% of the variants in vitro and still made general statements about the efficiency of in vivo and in vitro systems to identify functional variants.

The justification for studying CTNND1 as one of the target genes of ESRP1/2 was not satisfactory and came out of the blue in the manuscript. More justification for focusing on this target gene is required in the abstract and introduction.

The variant V540H was tested in Py2T but not in the zebrafish system, and no information or prediction was provided. Please confirm if this is a typo.

Some gene variants showed different effects when tested in zebrafish compared to Py2T cells like N643S, V540H, and R667C. If all coding variants of Esrp1 were tested in vitro, it would have provided a stronger comparison between zebrafish and the in vitro system. Hence, additional clarification for the differences is required.

Modify the statement "Targeted disruption of Esrp1 and Esrp2 in the mouse resulted in bilateral cleft lip and palate (21, 46)." to targeted disruption of Esrp1 alone.

- It wasn't clear why a murine breast cancer cell line (Py2T) from a breast tumor of an MMTV-PyMT transgenic mouse was used as an in vitro model instead of normal keratinocytes or oral epithelial cells.
- Line 108, modify heterozygous
- Line 122, this statement is not clear "We also showed that Esrp-regulated Ctnnd1 to generate an epithelial isoform that co-localized with Esrp1 and Esrp2 in the mouse and zebrafish embryonic oral epithelium."
- Line 626, What does SVI stand for?

Reviewer #3

(Remarks to the Author)
Da Silva et al use zebrafish and cell-based assays in parallel to test potentially pathogenic rare ESRP1 and ESRP2 variants and show that in silico, in vitro and in vivo assays do not always agree, highlighting the importance of functional testing of patient variants. They further observe that overexpression of Ctnnd1 rescues the zebrafish phenotype, and that Espr1/2 mutants had differential expression of Ctnnd1 isoforms in the epithelium, supporting an Esrp-Ctndd1 regulatory pathway. Lastly, they identified novel CTNND1 variants in a cohort of OFC patients. These findings are important and provide novel insight on epithelial contributors to the pathogenicity of OFCs. The conclusions and data interpretation are robust, valid and reliable and the manuscript is well written.

There are some areas of the manuscript that would benefit from clarification and justification, listed below.

Major:
1. It is unclear what dataset is being used for the CTNND1 variant discovery in OFC cohorts, neither the methods nor the results clearly explain this portion of the study. The authors say that this was a recent analysis, is this from the GMKF dataset or a separate dataset? It’s also unclear what the 500 genes implicated in OFCs that were analyzed are and what kind of analysis is being done (it’s also not explained well in the methods).
2. The experiments using Py2T need justification, why is a mammary tumor cell line being used? How does it compare to

oral epithelial cells and how are findings in this cell line relevant to facial development?

Author Rebuttal letter:

Manuscript COMMSBIO-24-0943-T

Functional analysis of ESRP1/2 gene variants and identification of Esrp-regulated CTNND1 isoforms in orofacial cleft pathogenesis

Thank you very much for expert reviews of this work. We are grateful that there was a consensus between reviewers on the merit of this manuscript to recommend publication pending revisions. We have carefully made all the recommended changes in accordance with the reviewer's comments. We hope that you will find this revised manuscript is now appropriate for publication.

Additionally, the title, abstract and the manuscript were revised to observe the word limits of the journal for these sections.

Thank you very much for your consideration.

Reviewer #1 (Remarks to the Author):
The manuscript, "Functional analysis of ESRP1/2 gene variants and identification of Esrp-regulated CTNND1 isoforms in orofacial cleft pathogenesis" demonstrates the variable function of known ESRP variants and highlights key mRNA targets in cleft palate development. The authors used two systems to test the role of ESRP1 and ESRP2 variants found in humans—zebrafish model with esrp1 knocked out and esrp2 knocked down with translation-blocking morpholino, and murine cell line with Esrp1 and Esrp2 doubly knocked out. These models had clear phenotypes they could utilize for their assays; the zebrafish model showed defective morphology of the anterior neurocranium, and the double KO mouse cell line was assayed for splicing defects in many Esrp1/2 targets. The authors next performed "rescue" experiments with the human variants of ESRP1 and ESRP2, and assessed which variants are functional and which are pathogenic. They tested 7 ESRP1 and 12 ESRP2 variants and found that most are benign mutations—when expressed in the models they can rescue the phenotypes. Out of the 19, only 1 frameshift mutation in ESRP1 and 1 missense and 1 nonsense mutation in ESRP2 were found to be pathogenic (1 missense mutation in ESRP2 was only pathogenic in the murine cell line). The two models were mostly in agreement. The authors show that the in vivo and in vitro testing is critical for determination of pathogenicity of genetic variants because many in silico methods including AlphaMissense did not accurately predict the results.
Next part of the paper was slightly hard to follow in terms of flow. The authors shifted focus on Ctnnd1 and its splicing, and localization of isoforms. They showed that overexpression of Ctnnd1 is able to rescue the morphological phenotype in zebrafish model. It seemed like there is a logical jump from the previous section. Perhaps they are looking for downstream targets of ESRP1/2 that are associated with the OFC phenotype? It would be good if the authors can explain why they focused on Ctnnd1.
Response: The reviewer is correct that the project did utilize the esrp1/2 mutant model as a functional screen to analyze potential targets of Esrp splicing. We found that over-expression of ctnnd1 rescues the phenotype in the esrp1/2 mutant fish. We revised the Introduction (page 5, lines 113-119) to clarify the transition in the paper from ESRP1/2 gene variant testing to CTNND1: "We previously showed that Esrp1/2 regulated splicing of Ctnnd1 (60). Using RNAscope, we found that Ctnnd1 transcripts co-localized with Esrp1 and Esrp2 in the mouse and zebrafish embryonic oral epithelium. The esrp1/2 zebrafish model also presented a functional assay to test the function of Esrp-regulated genes such as Ctnnd1. In fact, exogenous expression of ctnnd1 mRNA in zebrafish esrp1/2 mutants partially rescued the cleft ANC, foreshortened pectoral fin and fused otolith phenotypes."

Overall, the manuscript provides critical, new information about ESRPs and I recommend it for publication. I have a few comments that can improve the manuscript further.

1. The authors use Arghef11 splicing as a functional readout for ESRP1/2 mutations. Authors should consider adding 2-3 more known targets to further validate this result, preferably with targets related to the phenotype such as CTNND1.

Response: Thank you very much for your suggestion, we considered using different

controls, but we chose Arhgef11 because Esrp1/2 regulation of Arhgef11 and its epithelial vs. mesenchymal isoforms has been published previously and primers could be designed to easily discern the difference. Early in this work, we examined the splicing pattern of several other Esrp-regulated genes that we identified (Enah and Ctnnd1, for example) but the splicing patterns of these genes were more complex. We believe to redo all the transfections using another marker would not change the conclusions of the paper. To enhance the reproducibility and robustness of the gene variant functional tests, we believe using an independent model such as the zebrafish was effective and corroborated the Py2T results.

2. Ctnnd1 is a known ESRP target. This should be emphasized in the section "Alternative splicing to generate epithelial isoform of Ctnnd1 requires Esrp 1/2 function".

Response: Thank you for your comment, we emphasized this prior evidence with "We and others demonstrated that Esrp1 and Esrp2 regulate the alternative splicing of Ctnnd1, generating isoforms that differ between epithelial and mesenchymal cell types (20, 60, 66, 73), making Ctnnd1 an interesting Esrp1/2 target that has also been implicated in CL/P (61)." page 28, lines 523-526.

3. The last section in the results feels disjointed with the rest of the manuscript. This section should be included prior to the data in Figure 6 and the data should be represented in a figure.

Response: We have moved this section earlier as recommended, page 32, lines 602-618. We have also included a figure indicating the newly identified CTNND1 gene variants as recommended in Supplementary Figure 3.

Minor:

1. Typo – Lines 406-407. "As because".

Response: Revised, we removed "as" from the phrase

2. Figure 1A – Arrows to indicate the fused otoliths.

Response: White arrows were added indicating the otoliths

3. Figure 4 needs letters to indicate panels for clarity. Image quality is not great.

Response: We improved the image quality and added letters for clarity in the figure and the subtitles.

4. In the discussion, the emphasis on the issues with bioinformatics is too heavy. It is important to acknowledge these issues so other researchers are aware, however, the authors should more strongly emphasize the importance of the results from their physiological experiments. As written, these results feel underwhelming.

Response: We agreed with the reviewer, and we re-wrote the bioinformatics part. "These results highlight the need for experimental models to enhance the validity of in silico predictions of protein function. We found that while the SIFT and PolyPhen-2 algorithms have a positive predictive value when they align in predicting benign variants, they tend to overestimate the prevalence of pathogenic variants". Page 36, lines 694-698. While AlphaMissense provided slightly better predictions for ESRP1 than SIFT and PolyPhen-2, in the case of ESRP2, AlphaMissense over-interpreted benign variants as pathogenic. " Page 37, Line 699-701.

Reviewer #2 (Remarks to the Author):

This is a very well-written and organized study that used zebrafish and Py2T in vitro systems to uncover the pathological function of coding mutations in ESRP1/2 and CTNND1 genes. The study emphasized the importance of experimental assays to determine functional gene variants compared to in silico computational predictive models, including SIFT, PolyPhen-2, and AlphaMissense. The manuscript also identified CTNND1 as a cleft lip and palate gene that harbors a large number of gene mutations which account for almost 1.5% of cleft lip/palate (CL/P) cases, making this gene the second most important OFC after IRF6. The study tested a good number of gene variants in ESRP1/2 using a zebrafish system to evaluate the prediction efficiency of the predictive computational models and concluded that the in silico models overestimated the pathological impact of coding variants.

There are a few concerns that need to be addressed to improve the quality of the manuscript. The statements used to compare the predictive accuracy of computer models were very broad without sufficient evaluation and analysis of previous studies that conducted similar assays to test the function of coding variants in vitro and in vivo compared to predictive models. The study cited a few publications that were consistent with the study's findings but failed to include a sufficient number of studies to justify the generalized statements.

Response: Thank you very much for the comments. We completely agree that we should not draw a general statement from the analysis of a couple of genes. We believe that while bioinformatics tools have been very helpful and widely used to predict the functionality of missense variants, the performance may vary across genes, and, in the case of ESRP2, the algorithms tend to suffer from false positives (over-calling benign variants as pathogenic). This justifies the use of gene-specific models for variant interpretation. In the revised manuscript, we have cited additional papers reviewing bioinformatics tools for variant effect prediction (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4896183/, https://pubmed.ncbi.nlm.nih.gov/33261662/, https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2022.981005/full), Page 3, line 73-74 and we have cited several additional work, such as Mutation Significance Cutoffs (https://www.nature.com/articles/nmeth.3739) and AnnotBoost (https://www.nature.com/articles/s41467-020-20087-2), to strengthen the discussion. Page 4, line 82-84.

The limitations of the current study are the use of genetic knockout of Esrp1 and morpholinos to knockdown Esrp2 for creating an efficient in vivo system to test a good number of coding variants, while testing a smaller number of variants in vitro. Using this type of in vivo system introduces several layers of complexity which might add noises that could impact the efficiency of the pathogenicity test. One discrepancy between the zebrafish and mouse model is the gene duplications and paralogues in zebrafish which required the knockout of both Esrp1 and 2 to create OFC phenotype, while knockout of Esrp1 alone was sufficient to cause CLP in mice. Also, the study didn't test all the gene variants in Py2T cells, which were tested in zebrafish. It was not clear why they decided to test less than 50% of the variants in vitro and still made general statements about the efficiency of in vivo and in vitro systems to identify functional variants.

Response: Thank you very much for your comment. We clarified in the text that we decided to test the variants that had different results compared to the in silico analysis. “We aimed to obtain an additional functional assessment for those gene variants testing results that contradicted in silico prediction. We performed site-directed mutagenesis to introduce the 11 gene variants, that were electroporated into Esrp1/2 DKO PY2T cells and performed the RT-PCR assay 24 hours post-electroporation.” Page 22-23, lines 443-446.

The justification for studying CTNND1 as one of the target genes of ESRP1/2 was not satisfactory and came out of the blue in the manuscript. More justification for focusing on this target gene is required in the abstract and introduction.

Response: Thank you for your comment, we added clarifications in the abstract, introduction, and in the CTNND1 sections.

The variant V540H was tested in Py2T but not in the zebrafish system, and no information or prediction was provided. Please confirm if this is a typo.

Response: We did test the V540M variant in the Py2T cells but it was not tested in the zebrafish model, we have removed it from the data presented in Figure 3.

Some gene variants showed different effects when tested in zebrafish compared to Py2T cells like N643S, V540H, and R667C. If all coding variants of Esrp1 were tested in vitro, it would have provided a stronger comparison between zebrafish and the in vitro system. Hence, additional clarification for the differences is required.

Response: Thank you very much for your comment, we cannot be sure why we see a difference in the variant N643S, we hypothesize that this variant may be hypomorphic, or that Arhgef11 is just one readout of Esrp1 mRNA splicing activity. For the R667C variant, we got the same result between fish and cells where we see a rescue in both models, suggesting that this variant is benign. V540M was removed.

Modify the statement "Targeted disruption of Esrp1 and Esrp2 in the mouse resulted in bilateral cleft lip and palate (21, 46)." to targeted disruption of Esrp1 alone.

Response: Thank you, we modified the text to “Targeted disruption of Esrp1 in the mouse resulted in bilateral cleft lip and palate (21).” Page 4, Line 94-95.

- It wasn't clear why a murine breast cancer cell line (Py2T) from a breast tumor of an MMTV-PyMT transgenic mouse was used as an in vitro model instead of normal keratinocytes or oral epithelial cells.

Response: We used Py2T cell line because it was a model previously used for the Esrp1 and Esrp2 studies and was shown to be a suitable epithelial cell model. We added that to the text. “The Py2T cell line has been used effectively to study epithelial mesenchymal transition and we have previously generated and characterized Esrp1 and Esrp2 double knock-out Py2T lines (23, 53).” Page 5, line 108-111.

- Line 108, modify heterozygous

Response: Revised

- Line 122, this statement is not clear "We also showed that Esrp-regulated Ctnnd1 to generate an epithelial isoform that co-localized with Esrp1 and Esrp2 in the mouse and zebrafish embryonic oral epithelium."

Response: We modified the phrase “We previously showed that Esrp1/2 regulated splicing of Ctnnd1 (60). Using RNAscope, we found that Ctnnd1 transcripts co-localized with Esrp1 and Esrp2 in the mouse and zebrafish embryonic oral epithelium”. Page 5, line 113-115.

- Line 626, What does SVI stand for?

Response: SVI stands for sequence variant interpretation, we added this to the text

Reviewer #3 (Remarks to the Author):

Da Silva et al use zebrafish and cell-based assays in parallel to test potentially pathogenic rare ESRP1 and ESRP2 variants and show that in silico, in vitro and in vivo assays do not always agree, highlighting the importance of functional testing of patient variants. They further observe that overexpression of Ctnnd1 rescues the zebrafish phenotype, and that Espr1/2 mutants had differential expression of Ctnnd1 isoforms in the epithelium, supporting an Esrp-Ctndd1 regulatory pathway. Lastly, they identified novel CTNND1 variants in a cohort of OFC patients. These findings are important and provide novel insight on epithelial contributors to the pathogenicity of OFCs. The conclusions and data interpretation are robust, valid and reliable and the manuscript is well written.

There are some areas of the manuscript that would benefit from clarification and justification, listed below.

Major:
1. It is unclear what dataset is being used for the CTNND1 variant discovery in OFC cohorts, neither the methods nor the results clearly explain this portion of the study. The authors say that this was a recent analysis, is this from the GMKF dataset or a separate dataset? It’s also unclear what the 500 genes implicated in OFCs that were analyzed are and what kind of analysis is being done (it’s also not explained well in the methods).

Response: This data is the cumulative analysis from CLP and CPO cohorts, we added to the text “OFC associated genes were based on a previously published study that curated a list of approximately 500 genes based on known clinical syndromes and association results from GWAS (61)" page 6, line 143-145

2. The experiments using Py2T need justification, why is a mammary tumor cell line being used? How does it compare to oral epithelial cells and how are findings in this cell line relevant to facial development?

Response: We used Py2T cell line because it was a model previously used for the Esrp1 and Esrp2 studies and was shown to be a suitable epithelial cell model. We added that to the text. “The Py2T cell line has been used effectively to study epithelial mesenchymal transition and we have previously generated and characterized Esrp1 and Esrp2 double knock-out Py2T lines (23, 53).” Page 5, line 108-111.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I'd like to thank the authors for their thoughtful and complete responses to my comments. In their revision, the authors have adequately addressed all concerns included in my previous review. For this reason, I believe the article is now acceptable for publication.

Reviewer #2

(Remarks to the Author)
The authors did a great job addressing this reviewer's comments and suggestions, and all requested modifications were adapted in the revised manuscript. I have no additional comments.

Reviewer #3

(Remarks to the Author)
The changes made by the authors have addressed the clarification issues that were brought up.