Embryonic requirements for *Tcf12* in the development of the mouse coronal suture

Man-chun Ting, D'Juan T. Farmer, Camilla S. Teng, Jinzhi He, Yang Chai, J. Gage Crump and Robert E. Maxson
DOI: 10.1242/dev.199575

Editor: Patrick Tam

Review timeline

Original submission:	2 March 2021
Editorial decision:	6 April 2021
First revision received:	14 October 2021
Editorial decision:	10 November 2021
Second revision received:	18 November 2021
Accepted:	22 November 2021

Original submission

First decision letter

MS ID#: DEVELOP/2021/199575

MS TITLE: Embryonic requirements for Tcf12 in the development of the mouse coronal suture

AUTHORS: Man-chun Ting, D'Juan Farmer, Camilla S Teng, Jinzhi He, Yang Chai, Gage Crump, and Robert E Maxson

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 recommend a substantial revision of your manuscript before we can consider publication. If you are able to revise the manuscript along the lines suggested, which may involve further experiments, I will be happy receive a revised version of the manuscript. Your revised paper will be re-reviewed by one or more of the original referees, and acceptance of your manuscript will depend on your addressing satisfactorily the reviewers' major concerns. 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 manuscript reports the novel characterization of the phenotype craniosynostosis phenotypes based on the complete and tissue-specific deletion of Tcf12 in mice. The findings reported in the manuscript help to further understand the basis for the phenotype.

Comments for the author

The manuscript entitled ‚Embryonic requirements for Tcf12 in the development of the mouse coronal suture' describes the phenotypic changes upon total or conditional loss of the bHLH transcription factor Tcf12. Conditional deletion was performed in two different compartments neural crest (Wnt1-Cre) and mesenchyme (Mesp1-Cre), or in combination in both compartments. Furthermore, the authors used the Doxycyclin-inducible Gli1-CreERT2 line to delete.
TCF12 can interact with TWIST1, and heterozygous mutations in both genes account for most Saethre-Chotzen syndrome cases. Yet, in mice heterozygosity of Tcf12 does not recapitulate the coronal suture defect. Here, the individual role of Tcf12 in mice was explored.
Overall the manuscript is well written and conclusive. The results are very well documented and presented in the figures, which are of high quality.
There are only a few minor concerns/comments:

Introduction:
The introduction may be restructured, as currently results are intermingled with the state of the art. It would be better to present them only in the last paragraph. E.g. 'Whereas Twist1 and Tcf12 are expressed in postnatal…, we show that conditional deletion…in the postnatal…does not affect maintenance of the sutures, pointing to largely embryonic roles of Twist1 and Tcf12 in suture regulation.' The conclusion, however, that Twist 1 also has a largely embryonic role is not supported by any data in the manuscript or by citations from the literature.

Material & Methods:
Sufficiently described, but catalog order numbers should be provided for all reagents.
In the paragraph on quantification and statistical analysis, the software used for the statistical analysis should be mentioned. Regarding the Student's T-Test please specify if this was a paired, two-sample equal variance, or two-sample unequal variance type of test.

Results:
The number of individual samples analyzed should always be mentioned in the figure legends.
What was the overall percentage of homozygous mice recovered? The current phrase 'recovered at low frequency' is a bit vague. What is the percentage of perinatal lethality of the Tcf12 homozygous specimens?
It is a bit puzzling that in the text the authors state that in 8 out of 22 Tcf12 mutant mice partial fusions of the coronal suture were observed, yet, in Table 1 the total number is 33, and the penetrance of coronal craniosynostosis in 58% here versus 36% in the main text. Then on the other hand in Table 2 only 22 Tcf12-/- specimens were analyzed. This needs to be clarified.
Oram & Gridley defined in 2005 a craniosynostosis index, does this differ from the coronal synostosis index mentioned here, and if so what are the differences. It would have been good to include here in the result part where the severity of the coronal suture defects was determined the respective reference again, or briefly mention how this was done.
Noticeably, the coronal synostosis index in the mutants has a high variance, as the SEM +/- 1.8 is bigger than the index number, which is 1.7. Table 2 should be referenced here.
Please state the penetrance of the heterotopic bones observed upon neural crest-specific deletion of Tcf12 and the number of samples analyzed.
Can the authors comment on the differences in the penetrance of the Twist1+/- craniosynostosis phenotype reported here (100%) versus Oram & Gridley 2005 (83%).
Sp7 is also regarded as a marker for osteoblast precursors that are not yet 100% committed, as such the statement made based on the Sp7 immunostaining results that Tcf12 and Twist 1 function to inhibit premature osteoblast differentiation is not 100% correct. Maybe the authors could further elaborate on this explaining if they consider the BrdU+;Sp7+ population as being the precursors and the BrdU-;Sp7+ population as mature osteoblasts.

Please provide a reference for the statement that Tcf12 overlaps postnatally with the Gli1-CreERT2 activity/expression. The postnatal deletion produced no phenotype. This negative result could also be due to inefficient deletion, can the authors provide data on the postnatal deletion efficiency to support their finding.

Figure legends:
- Number of individual samples analyzed should be mentioned.
- The quantitative data seem to be displayed as box-and-whisker plots with the median, minimum to maximum (whiskers) displayed, and not error bars (SEM)
- Figure 2: Please indicate that the quantification was performed using P21 samples
- Figure 4: Typo mesencnyme instead of mesenchyme
- Figure S1: The figure legend is placed at an odd position within the figure legends.

Reviewer 2

Advance summary and potential significance to field

The authors of this paper seek to understand the individual role of Tcf12 in producing the phenotypes associated with Saethre-Chotzen syndrome particularly fusion of the frontal and parietal bones at the coronal suture.
It is known that heterozygous loss-of-function mutations in the basic HLH transcription factors TWIST1 and TCF12 cause Saethre-Chotzen syndrome in human as do their orthologues in mouse, but it is unclear exactly what role Tcf12 plays in the process. Tcf12 is a binding partner for Twist and so the authors wish to investigate whether Tcf12 has the same ability in changing growth rates and closing sutures as does Twist, or if it is merely associative. They find that homozygous loss of Tcf12 has some of the same effects as heterozygous Twist and compound Twist1;Tcf12 heterozygous mice.
The authors conclude that Twist1-Tcf12 heterodimers are critical for coronal suture formation.
The discussion of results is lacking resolution. It ends without a summary statement or a suggestion for the next set of experiments and none of what is presented is tied back to Saethre-Chotzen syndrome. The last sentence is a jump that leaves the reader wondering what the point is.

Comments for the author

Although it appears that most of the findings are sound, data are presented with little explanation for the reader and there seem to be minor problems with statistics as presented. The authors assume the reader's knowledge of all markers and cells so that this paper is really only understandable by a small group of specialists. Even the disease, Saethre-Chotzen, the driving force for the research is not given much attention.
Introduction In summarizing broad expression of Tc3 and Tc4 reported by Wang et al, what exactly is meant? Broad expression of what? patterns, intensity? or are they broadly expressed as in expressed across a broad range of cell types?
Last sentence of the first paragraph does not follow from the previous sentence so the introduction is hard to follow. Consider adding something focusing on Tcf12 or choosing a different transition between these two concepts.
In paragraph 3, the term ES is used but with no introduction of what ES stands for or exactly what these cells are. Likewise, in the next paragraph Grem1 is not introduced or defined. Grem1 is later discussed in the Results section and perhaps this could be moved to the first time Grem1 is mentioned. There should be some introduction to the relationship between mesenchymal cells and sutures.

Methods Histology and immunohistochemistry: There is no reasoning given for the specific ages/ stages used in analysis.
Is there a citation for the BrdU time schema of dosing the dam and dissecting the pups?
There are many undefined abbreviations in this section Quantitation and statistical analysis: The suture assessment of Oram and Gridley was designed to evaluate suture fusion across several sutures, not a single suture. In the original analysis each animal received a composite score. It might be better to state that the authors adapted this method to specifically look at the coronal suture. Also, how were the sutures "looked at". Oram and Gridley observed specimens stained for

alizarin red, but this has been shown to be inadequate. How were the sections selected? how far apart were they? In what plane were all of the sections made? Further, not all the statistical analyses used are reported here. Explanation of analysis for figure four is completely missing. Additionally, the methods as stated in the results/ figure are inappropriate.
Results The term “wild-type siblings” is an imprecise descriptor. They cannot be wildtype if they were part of the experiment. Perhaps use “unaffected” or provide their genotype. Later on the term “controls” is used – again, the genotype of the mice should be used rather than descriptors like this as the “controls” change for each experiment. Lack of specificity makes it hard to follow the argument.

Using only the homozygous Tcf12 pups that survive birth introduces a bias – these are likely the less affected mice, so that the percentages reported do not represent the impact of homozygosity on suture closure. The authors acknowledge this in the discussion but do not do so in the reported findings. Why is there no figure or discussion of the Tcf12 homozygous embryos at earlier ages? It seems much could be learned from them. Suture phenotypes seem noncritical compared to the other phenotypes reported.
The authors state that loss of Tcf12 in both neural crest and mesoderm prevented coronal suture formation. Stating this requires a summary or figure defining the cells or morphology recorded when a suture is judged “not formed” and “formed”. The terms suture “formation” is not clearly defined. Figure 2 only shows the mice at P6 and P21 that cannot possibly be proof that a suture never existed as sutures form much earlier.

Reduction of Tcf12 function affects osteoblast dynamics and asymmetric distribution of Grem1+ mesenchyme : Since symmetry is usually discussed with reference to the midline, this section requires a specific definition of the boundary across which symmetry is being judged. The authors state: ”Interestingly, we observed that Grem1+ cells were asymmetrically distributed above (lateral to) the frontal bone and below (medial to) the parietal bone in wild-type E14.5 controls (Fig 4D, F, H and I).” and “the frontal bone were selectively lost (Fig 4E, G and H), resulting in a more symmetric arrangement of Grem1+ cells. Grem1+ cells both above and below the parietal bone were also reduced in mutants.”

What does “lateral to/above” and “medial to/below” mean in this context? Are the authors referring to the endocranial and ectocranial surfaces of the bones? Also, the nature of the lack of symmetry is not described or explained. Were there more equal numbers of Grem1+ cells above the frontal bone and below the parietal bone?
Neural crest-mesoderm boundary defects in Tcf12 null mice: The significance of Eph-ephrin signaling should at least be briefly introduced earlier so the reader understands the significance here.
The shifts in coronal suture placement reported for Figure 2 are hard to see and I would want a measurement to validate that. Smaller frontal or parietal could be measured, or the distance from some standard anatomical location (not affected by suture closure) could be used. It is not clear what we are looking at in these figures as there is no clear suture in Fig 2 F or H.
Also in Figure 2J-O, how was the red line established?
Are the authors assuming that a closed suture means that one never formed?
Lack of formation versus premature closure are different pehnomena. This should be demonstrated in younger mice (embryos).
Tcf12 is dispensable in postnatal suture mesenchyme for suture patency: The authors state that: “Fourteen weeks after induction, histology revealed no defects in the coronal suture when compared with untreated wild-type controls (Fig 6B,C).” What does ‘no defects’ mean? Was it patent? Were these really wild-type controls as in C57B6 mice of the same age, or were they littermates from this experiment?

Discussion In the first paragraph, the statements given about Twist1 and Tcf12 do not specifically tie them to coronal suture development. Maybe this is just a lack of the proper citation but it is left to the reader to connect the dots. Equally, the statement “Whereas heterozygous loss-of-function mutations of TCF12 in humans cause coronal synostosis, Tcf12 heterozygosity does not in mice.” needs a citation The authors state: “Rather than controlling proliferation, our lineage analysis of neural crest-derived cells in Tcf12 mutants suggests that Tcf12 functions primarily to restrict osteoblast differentiation at the growing bone fronts.” What is the evidence for this?

Why would islands of bone form with excess osteoblast production? Why would the osteoblasts not coalesce with already forming osteoblasts and bone? Are the spaces between these “extra” bones acting as sutures? Does this have something to do with nc-derived osteoblasts differentiating over mesoderm derived dura? The evidence for Tcf12 in maintaining suture patency postnatally is weak given the unknowns regarding periosteum and dura.
The discussion is lacking resolution. It ends without a summary statement or a suggestion for the next set of experiments and none of what is presented is tied back to Saethre-Chotzen syndrome. The last sentence is a jump that leaves the reader wondering what the point is. How would Tcf12 control rate of growth? How do you know that rate is what being change and not timing?
Figures Figure 1- define all abbreviations in the figure... af, pb, fb? Providing anatomical context for orientation is also needed.

Figure 2- would make a greater impact if labeled with mesoderm deletion and neural crest deletion and both. Normally p values are reported as $p<0.001$ if they are that low. Why is the scale bar in K so much larger than the other images? Since there are two different stains in J-O but they initially look similar, it would be good to put the stain target or methodology on the image for example adding ALP to J-N and Lac-Z to O. For Figure 2.I. : are these ratios based on suture length? how was that measured? This should be in the methods section.
Are the parietal bones overlapping at the sagittal suture or are they fused at P6 in the Tcf12flox/flox-;Mesp1-Cre and the double mutants? And if they are overlapping, why are they overlapping? This is the morphology of the coronal suture (one bone growing over the other) which the authors explain by the asymmetric presence of Grem1+ mesenchymal cells. Since parietal bones are mesoderm derived, this would pertain to the effect on mesoderm-derived cells only.

Figure 3- Lack of consistency with abbreviations makes it really hard for the reader to follow any argument or even read the figure. e.g., fs vs ms for the same area.
What counter stain was used for these images? if ALP signal is purple, it is difficult to be convinced of staining especially with the LacZ (pink and blue make purple)
Figure S1- What is the n (number of sections quantified from what number of individuals)?
Figure 4 H-I What is the n (number of sections quantified from what number of individuals of what genotype)? The histogram for this data is confusing... are the blue and orange areas a percentage of the total number of cells counted? what are the error bars on each section indicating? It seems that these data would be better analyzed using anova rather than a series of t-tests.
Figure 5 Abbreviations are not consistent making it extremely hard to follow. All abbreviations need to be defined and be consistent throughout the entire manuscript.
Figure 6. Why was 14 weeks chosen as the age? Is it possible that these are just H&E rather than Aliazarin Red and hematoxylin/eosin?
Table 1. data for littermates with the genotypes that are not the focus of the research should be shown for comparison.

Reviewer 3

Advance summary and potential significance to field

Ting et al. examines the mechanism of craniosynostosis (CS) in mouse models of Seathre-Chotzen syndrome, focusing on Tcf12 mutants. The study documents some important findings, and reinforces the idea of common mechanisms across species i.e. the clear demonstration of genetic interaction between Twist and Tcf12 seen in humans, zebrafish, and mice. It also offers additional support for the hypothesis that accelerated expansion of the frontal and parietal bones is a major driver in coronal CS. Based on distribution of Grem+ cells in normal and mutant mice, they also offer an intriguing possible explanation for the consistent direction in overlap of bones at the mouse coronal suture. Overall, the manuscript makes an important contribution and has high quality data, but could be improved by addressing a couple major and some minor concerns.

Comments for the author

While TCF12 haploinsufficiency in humans is associated with coronal CS, both copies must be mutated in mouse, where the phenotype is still only partially penetrant. Also, while some Tcf12

mutants do not survive the perinatal period, the CS in those who do is mild. One significant point that is still unclear is the extent and underlying cause of death of Tcf12 mutants. First, it is not clearly stated what the survival rate is (observed/expected at specific time points). This seems like data that the authors should already have and could easily present.
Importantly, they speculate that the increased severity and penetrance of CS in the Tcf12 conditional mutants shown in Table 2 is because the more severely affected Tcf12 mutants are dying earlier, an explanation difficult to judge without supporting numbers on survival rates. The authors' arguments around this point are confusing. They seem to suggest that mutants with more severe CS are dying prenatally, although normally CS is not lethal.

The developmental specimens shown in Fig. 2J-O are intended to provide important data supporting their model of early increased growth in the skull bones. However it is not easy to judge the size of the bones from those images. The argument would be greatly strengthened by more/better images, or preferably measurements of bone size in multiple samples with statistics. In conjunction, they could compare these conditional mutants to Tcf12 mutants at the same prenatal stages, to support their explanation of the very mild CS observed postnatally in Tcf12 mutants.
While a minor point in the paper, the asymmetric distribution of Grem+ cells at the suture is interesting and could point to a mechanism not just for CS in the mutants, but asymmetric bone overlap in WT mice. If available, information about when the asymmetric distribution first arises (and when it first looks different in mutants) could shed additional light on the pathogenesis of CS. In other words if the bones are growing faster in the mutants, why do they fuse, instead of just forming the normal overlap at an earlier stage?
A final minor point is that for readers less familiar with CS, and with the grading of severity, it might be helpful to provide examples of the index summarized in Table 2.

First revision

Author response to reviewers' comments

Response to Reviews

REVIEWER #1:

1.1 - The introduction may be restructured, as currently results are intermingled with the state of the art. It would be better to present them only in the last paragraph. E.g. 'Whereas Twist1 and Tcf12 are expressed in postnatal…, we show that conditional deletion…in the postnatal…does not affect maintenance of the sutures, pointing to largely embryonic roles of Twist1 and Tcf12 in suture regulation.' The conclusion, however, that Twist 1 also has a largely embryonic role is not supported by any data in the manuscript or by citations from the literature.

We have modified the sentence about postnatal function to refer only to Tcf12, as that is what is studied in this manuscript. We have also moved this sentence to the last paragraph of the Introduction as requested.

> Lines 97-101: "In addition, we show that conditional deletion of *Tcf12* in the postnatal *Gli1+* sutural stem cell domain does not affect maintenance of sutures, further pointing to a largely embryonic role of Tcf12 in suture regulation."

1.2 - Material & Methods: Sufficiently described, but catalog order numbers should be provided for all reagents. In the paragraph on quantification and statistical analysis, the software used for the statistical analysis should be mentioned. Regarding the Student's T-Test please specify if this was a paired, two-sample equal variance, or two-sample unequal variance type of test.

We have added catalog numbers to the Methods. We clarify that "Pairwise comparisons among genotypes were analyzed by a two-tailed Student's t-test".

1.3 - The number of individual samples analyzed should always be mentioned in the figure legends.

We have now added the number of individual embryos and sections analyzed where appropriate in each figure legend.

1.4 - What was the overall percentage of homozygous mice recovered? The current phrase 'recovered at low frequency' is a bit vague. What is the percentage of perinatal lethality of the Tcf12 homozygous specimens?

We did not have precise records of genotypes surviving birth, which is likely compounded by some perinatally lethal pups being cannibalized by the mother shortly after birth. Instead, we now cite the original study on the *Tcf12* null showing that they are born at the expected Mendelian frequency but largely die off in the first two weeks.

> Lines 209-211: "It had previously been reported that homozygous *Tcf12*$^{-/-}$ mice are born at the expected frequency from a *Tcf12*$^{+/-}$ incross, with the vast majority dying by 2 weeks after birth (Zhuang et al., 1996). We similarly observed that most homozygous *Tcf12*$^{-/-}$ mice died within the first 2 weeks."

1.5 - It is a bit puzzling that in the text the authors state that in 8 out of 22 Tcf12 mutant mice partial fusions of the coronal suture were observed, yet in Table 1 the total number is 33, and the penetrance of coronal craniosynostosis in 58% here versus 36% in the main text. Then on the other hand in Table 2 only 22 Tcf12-/- specimens were analyzed. This needs to be clarified.

The original Table 1 had included both embryonic and postnatal mutant mice analyzed, whereas Table 2 had referred to only the postnatal mice. We have now restructured the tables such that Table 1 focuses on non-synostosis Tcf12 mutant phenotypes, and Table 2 on synostosis. We also repeated the analysis of Tcf12 mutant synostosis at embryonic stages (E17-E18) and now include these numbers in Table 2. We find that the synostosis penetrance is 50% and the index 0.7 at embryonic stages, versus penetrance of 36% and index of 0.5 at postnatal stages (P0-P21). These differences are not statistically significant, suggesting that there is not a large difference in synostosis frequency or severity in Tcf12 mutant embryos that die before birth.

1.6 - Oram & Gridley defined in 2005 a craniosynostosis index, does this differ from the coronal synostosis index mentioned here, and if so what are the differences. It would have been good to include here in the result part where the severity of the coronal suture defects was determined the respective reference again, or briefly mention how this was done. Noticeably, the coronal synostosis index in the mutants has a high variance, as the SEM +/- 1.8 is bigger than the index number, which is 1.7. Table 2 should be referenced here.

We have added details of the craniosynostosis index scoring to the Methods, including references to two papers from our lab (Yen et al., 2010; Teng et al., 2018).

> ***Quantitation of craniosynostosis and statistical analyses***
> The severity of coronal synostosis was quantified using a scoring method adapted from Oram and Gridley's craniosynostosis index (Oram and Gridley, 2005; Yen et al., 2010) applied specifically to the coronal suture (Teng et al., 2018). For each skull, the two coronal sutures were scored. The extent of fusion was assessed microscopically (0: unfused; 1: <50% fused; 2: ≥50% fused and 3: 100% fused.) Scores for left and right sutures were added. Thus, for example, the maximum score given to two coronal sutures is 6. Phenotypic scoring was performed blinded to genotype. Pairwise comparisons among genotypes were analyzed by a two-tailed Student's t-test.

1.7 - Please state the penetrance of the heterotopic bones observed upon neural crest-specific deletion of Tcf12 and the number of samples analyzed.

In the legend for Figure 3, we now state that the phenotype is 100% penetrant and that 4 conditional mutants were analyzed at both E18.5 and P0.

1.8 - Can the authors comment on the differences in the penetrance of the Twist1+/- craniosynostosis phenotype reported here (100%) versus Oram & Gridley 2005 (83%).

100% was an editing mistake on our part. We have updated Table 2 with the correct number of 44%. We note that the Twist1+/- analysis was not the prime focus of this paper on Tcf12. While 44% (4/9) is lower than 83%, it is not statistically significant due to low sample size. We also note that genetic penetrance frequently varies based on genetic drift.

1.9 - Sp7 is also regarded as a marker for osteoblast precursors that are not yet 100% committed, as such the statement made based on the Sp7 immunostaining results that Tcf12 and Twist 1 function to inhibit premature osteoblast differentiation is not 100% correct. Maybe the authors could further elaborate on this explaining if they consider the BrdU+;Sp7+ population as being the precursors and the BrdU-;Sp7+ population as mature osteoblasts.

We agree that Sp7 labels both osteoblast precursors and osteoblasts. For example, recent work from the Maes lab showing that Sp7+ cells are an abundant source of new osteoblasts in mouse long bones. We have therefore changed "Sp7+ osteoblasts" to "Sp7+ osteolineage cells" throughout the manuscript. We do not consider BrdU-;Sp7+ to be a good criterion for mature osteoblasts, however, as there are likely Sp7+ progenitors that fail to incorporate BrdU during the time window of application due to the variable nature of replication.

1.10 - Please provide a reference for the statement that Tcf12 overlaps postnatally with the Gli1-CreERT2 activity/expression. The postnatal deletion produced no phenotype. This negative result could also be due to inefficient deletion, can the authors provide data on the postnatal deletion efficiency to support their finding.

We clarify that Tcf12 expression in coronal suture mesenchyme as presented in Fig 6A matches the suture mesenchyme domain labeled by Gli1-CreERT2 in Zhao et al, 2015. We have also performed new experiments that now show a 60% reduction in *Tcf12* expression in Gli1-CreERT2; Tcf12-f/f mice after tamoxifen administration (Fig 6D,E). We have added a paragraph to the Discussion to discuss the interpretations and caveats of these results.

> Lines 381-386: "However, expression of *Tcf12* in the coronal suture and surrounding tissues was not completely eliminated in *Gli1-CreERT2; Tcf12-f/f* mice following postnatal tamoxifen administration, likely due to incomplete activity of CreERT2 and/or expression of *Tcf12* in some *Gli1-CreERT2*-negative lineage cells. Thus, future experiments using different tools to completely eliminate postnatal Tcf12 function will be needed to better assess roles for Tcf12 postnatally, especially as we observe continued *Tcf12* expression at the coronal suture until at least 4 months after birth."

1.11 - The quantitative data seem to be displayed as box-and-whisker plots with the median, minimum to maximum (whiskers) displayed, and not error bars (SEM).

Graphs in Figures 2,4,6 now show error bars +/- SEM , as well as p values for comparisons.

1.12 - Figure 2: Please indicate that the quantification was performed using P21 samples

We now indicate this in legend to Figure 2.

1.13 - Typo mesencnyme instead of mesenchyme.

Corrected.

1.14 - Figure S1: The figure legend is placed at an odd position within the figure legends.

Fixed.

REVIEWER #2:

2.1 - Although it appears that most of the findings are sound, data are presented with little explanation for the reader and there seem to be minor problems with statistics as presented. The authors assume the reader's knowledge of all markers and cells so that this paper is really only understandable by a small group of specialists. Even the disease, Saethre-Chotzen, the driving force for the research is not given much attention.

We have added further introduction of Saethre-Chotzen syndrome in the first paragraph of the Introduction, and then return to put our work in the context of this disease in the last paragraph of the Discussion. We address the minor problems to statistics elsewhere in this response letter.

2.2 - In summarizing broad expression of Tc3 and Tc4 reported by Wang et al, what exactly is meant? Broad expression of what? patterns, intensity? or are they broadly expressed as in expressed across a broad range of cell types?

In a previous iteration of this paper, we had included analysis of *tcf3a* and *tcf3b* zebrafish mutants and *Tcf3* mouse mutants. Later, however, we removed these data as we felt the analyses were incomplete and inconclusive. We have therefore deleted references to Tcf3 and Tcf4 in the current manuscript.

2.3 - Last sentence of the first paragraph does not follow from the previous sentence so the introduction is hard to follow. Consider adding something focusing on Tcf12 or choosing a different transition between these two concepts.

We are unclear which paragraph the reviewer is referring to. In the first paragraph of the Introduction, the last sentence states that the individual role of Tcf12 in skull development has not been reported. The previous two sentences refer to the fact that Tcf12 function has only been studied in mouse and zebrafish as an enhancer of Twist1 loss, thus setting up in the last sentence the motivation of the current manuscript to study Tcf12 requirements on their own.

2.4 - In paragraph 3, the term ES is used but with no introduction of what ES stands for or exactly what these cells are. Likewise, in the next paragraph, Grem1 is not introduced or defined. Grem1 is later discussed in the Results section and perhaps this could be moved to the first time Grem1 is mentioned. There should be some introduction to the relationship between mesenchymal cells and sutures.

We now define ES as "embryonic stem cell" in the Introduction. We also move introduction of Grem1 from the Results to Introduction, as well as define the relationship between the suture mesenchymal cells and sutures.

> Lines 92-95: "The suture mesenchyme between the apposing frontal and parietal bones has been shown to house skeletal stem cells (Zhao et al, 2015). In the appendicular skeleton, Grem1 marks mesenchymal progenitors responsible for bone formation (Worthley et al., 2015)."

2.5 – Methods: Histology and immunohistochemistry: There is no reasoning given for the specific ages/ stages used in analysis.

We have added a section to the Methods entitled "Rationale for choice of developmental stages for analysis" in which we describe our reasoning for using specific ages/stages.

2.6 - Is there a citation for the BrdU time schema of dosing the dam and dissecting the pups?

We now cite (Teng et al., 2018).

2.7 - Methods: There are many undefined abbreviations in this section.

We now define abbreviations throughout the Methods.

2.8 - Quantitation and statistical analysis: The suture assessment of Oram and Gridley was designed to evaluate suture fusion across several sutures, not a single suture. In the original analysis each animal received a composite score. It might be better to state that the authors adapted this method to specifically look at the coronal suture. Also, how were the sutures "looked at". Oram and Gridley observed specimens stained for alizarin red, but this has been shown to be inadequate. How were the sections selected? how far apart were they? In what plane were all of the sections made?

We have added a section in the Methods entitled "Quantitation of craniosynostosis and statistical analyses" where we cite our recent study adapting the Oram and Gridley method to the coronal suture in Alizarin-stained heads (Teng et al., 2018). We also clarify in the Methods that we "sectioned at 8 µm intervals" in a "defined area perpendicular to the bone fronts", and that "five sections from the medial to lateral aspects of the coronal suture were quantified per animal and averaged."

2.9 - Methods: Explanation of analysis for figure four is completely missing. Additionally, the methods as stated in the results/ figure are inappropriate.

We have modified this figure to better show in the diagrams where Grem1+ cells above and below the frontal and parietal bones were counted. In the Methods (Lines 197-202) we describe how Grem1+ cells were counted. In the bar graphs in Figure 4, we state the p values for each comparison, and in the Methods we state that "pairwise comparisons were analyzed by a two-tailed Student's t-test, and the p-values were corrected for multiple testing with Bonferroni correction." We have also substantially modified the text in the Results to better describe the comparison of ectocranial and endocranial situated cells.

2.10 - The term "wild-type siblings" is an imprecise descriptor. They cannot be wildtype if they were part of the experiment. Perhaps use "unaffected" or provide their genotype. Later on the term "controls" is used – again, the genotype of the mice should be used rather than descriptors like this as the "controls" change for each experiment.

For each experiment, we now list the precise genotype used for controls. In some cases, wild types were used from similarly staged but independent crosses. In other cases, sibling were used as controls and for these we now describe the exact genotype.

2.11 - Using only the homozygous Tcf12 pups that survive birth introduces a bias – these are likely the less affected mice, so that the percentages reported do not represent the impact of homozygosity on suture closure. The authors acknowledge this in the discussion but do not do so in the reported findings. Why is there no figure or discussion of the Tcf12 homozygous embryos at earlier ages? It seems much could be learned from them. Suture phenotypes seem noncritical compared to the other phenotypes reported. The authors state that loss of Tcf12 in both neural crest and mesoderm prevented coronal suture formation. Stating this requires a summary or figure defining the cells or morphology recorded when a suture is judged "not formed" and "formed". The terms suture "formation" is not clearly defined. Figure 2 only shows the mice at P6 and P21 that cannot possibly be proof that a suture never existed as sutures form much earlier. Are the authors assuming that a closed suture means that one never formed? Lack of formation versus premature closure are different phenomena. This should be demonstrated in younger mice (embryos).

We note that based on alkaline phosphatase staining in Fig. 5A,B, we can detect loss of the coronal suture in Tcf12 mutants as early as embryonic stage E13.5. In addition, we have conducted new analysis at E17 and E18 in Figure 1. This shows that based on Alizarin Red fluorescence staining of calvarial bones, Tcf12 mutants display partial fusions of the coronal suture before birth. We have also included videos of confocal images of regions of coronal sutures of wild type and Tcf12 mutant skulls (Movie 1 and Movie 2). These videos confirm the synostosis phenotype apparent in Alizarin-stained skulls. Quantification of the embryonic suture fusions have been added to Table 2. By comparing to our previous postnatal analysis, we find that the penetrance of synostosis in Tcf12-/- mice at embryonic stages (50%, n=10) is not statistically significantly different from the penetrance at postnatal stages (36%, n=17). Combined with the E13.5 analysis of mutant sections, this now

better supports our conclusion that suture formation rather than maintenance is perturbed in Tcf12-/- mice.

> Lines 344-347: "The observation of coronal suture fusion by Alizarin Red staining as early as E17.5, as well as the similar penetrance and severity of suture defects at late embryonic and postnatal stages, points to defects in the formation as opposed to the maintenance of the coronal suture in *Tcf12*$^{-/-}$ mice."

2.12 - Grem1+ mesenchyme : Since symmetry is usually discussed with reference to the midline, this section requires a specific definition of the boundary across which symmetry is being judged. The authors state: "Interestingly, we observed that Grem1+ cells were asymmetrically distributed above (lateral to) the frontal bone and below (medial to) the parietal bone in wild-type E14.5 controls (Fig 4D, F, H and I)." and "the frontal bone were selectively lost (Fig 4E, G and H), resulting in a more symmetric arrangement of Grem1+ cells. Grem1+ cells both above and below the parietal bone were also reduced in mutants."
What does "lateral to/above" and "medial to/below" mean in this context? Are the authors referring to the endocranial and ectocranial surfaces of the bones? Also, the nature of the lack of symmetry is not described or explained. Were there more equal numbers of Grem1+ cells above the frontal bone and below the parietal bone?

We have significantly modified this section to make it clearer we are comparing the ectocranial and endocranial sides of the calvarial bones. We have also added new data using Six2 in situs to show that mesenchymal asymmetry is present as early as E12.5 (see also response to Reviewer 3).

> Lines 281-286: "Interestingly, we observed that Grem1+ cells were asymmetrically distributed around the bones at the forming coronal suture, with more Grem1+ cells above (closer to skin) versus below (closer to brain) the frontal bone and conversely more Grem1+ cells below versus above the parietal bone in wild types at E14.5 (Fig 4A-F). This asymmetric distribution of Grem1+ cells in the ectocranial versus endocranial layers correlates with the parietal bone reproducibly overlapping above the frontal bone at the mature coronal suture."

In the Discussion, we also now better explain the nature of lack of asymmetry and in the Legend to Fig. 4 we provide statistical analysis that there are now more equal numbers of Grem1+ cells above the frontal bone of Tcf12 mutants.

2.13 - The significance of Eph-ephrin signaling should at least be briefly introduced earlier so the reader understands the significance here.

We now introduce Eph-ephrin signaling in the Introduction.

> Lines 64-66: "Loss of function of the ephrin receptor, EphA4, also causes cell mixing at the neural crest-mesoderm boundary and also results in synostosis of the frontal and parietal bones (Merrill et al., 2006)."

2.14 - The shifts in coronal suture placement reported for Figure 2 are hard to see and I would want a measurement to validate that. Smaller frontal or parietal could be measured, or the distance from some standard anatomical location (not affected by suture closure) could be used. It is not clear what we are looking at in these figures as there is no clear suture in Fig 2 F or H. Also in Figure 2J-O, how was the red line established?

In Fig. 2E-G, we have assessed the position of the coronal suture at P21 by measuring the length of the sagittal suture (white line, ss) and metopic suture (black line, ms). In Fig. 2I, we then display the ratio of these two measurements. In Fig. 2F (Wnt1-Cre; Tcf12-f/f knockout), we have indicated the patent coronal suture used for measurements, as well as some discoloration of the bone more anteriorly which is not a fused suture (*). We agree that in Fig. 2H (Wnt1-Cre; Mesp1-Cre; Tcf12-f/f) where the coronal suture is lost, it is difficult to determine where the suture would have been as there are two distinct faint lines – either of which might reflect the fused suture. We have therefore chosen to remove measurements in this genotype from Fig. 2I.

2.15 - Tcf12 is dispensable in postnatal suture mesenchyme for suture patency: The authors state that: "Fourteen weeks after induction, histology revealed no defects in the coronal suture when

compared with untreated wild-type controls (Fig 6B,C).” What does ‘no defects’ mean? Was it patent? Were these really wild-type controls as in C57B6 mice of the same age, or were they littermates from this experiment?

We thank the reviewer for this comment as we have discussed in further detail with our co-authors to get more accurate details of this experiment. As now stated in the manuscript, RNAscope in situ hybridization of *Tcf12* was performed at fourteen weeks, but for conditional deletion experiments tamoxifen was added for three consecutive days at one month after birth, with suture defects analyzed at two months. We also clarify that *Tcf12-fl/+* littermate controls were used.

2.16 – Discussion: In the first paragraph, the statements given about Twist1 and Tcf12 do not specifically tie them to coronal suture development. Maybe this is just a lack of the proper citation but it is left to the reader to connect the dots. Equally, the statement “Whereas heterozygous loss-of-function mutations of TCF12 in humans cause coronal synostosis, Tcf12 heterozygosity does not in mice.” needs a citation.

At the end of the first paragraph, we have modified the concluding sentence to link Twist1 and Tcf12 to coronal suture development.

> Lines 53-56: “While these experiments show combined requirements of Twist1 and Tcf12 for early development of the coronal suture (Teng et al., 2018), the individual role of Tcf12 in skull development had remained unexplored in model organisms, leaving it unclear whether Twist1 and Tcf12 control similar or distinct processes during suture development.”

In the Introduction, we cite (Sharma et al.,2013) for lack of synostosis in Tcf12 heterozygous mice. Restructuring of the Discussion had removed a similar statement.

2.17 - Discussion: The authors state: “Rather than controlling proliferation, our lineage analysis of neural crest-derived cells in Tcf12 mutants suggests that Tcf12 functions primarily to restrict osteoblast differentiation at the growing bone fronts.” What is the evidence for this?

We have deleted this sentence as we discuss effects on the Grem1+/Six2+ mesenchyme populations more thoroughly in the following paragraph.

2.18 – Discussion: Why would islands of bone form with excess osteoblast production? Why would the osteoblasts not coalesce with already forming osteoblasts and bone? Are the spaces between these “extra” bones acting as sutures? Does this have something to do with nc-derived osteoblasts differentiating over mesoderm derived dura?

These ectopic bone islands, or Wormian bones, have been commonly reported in human patients and a number of mouse models. We now cite one study from our group – (Roybal et al., 2010). We agree that it is interesting that they fail to fuse into a larger bone, yet this is what is commonly reported and we do not have additional data to address possible mechanisms keeping these bone islands separate.

> Lines 361-364: “These bony islands arise from neural crest-derived cells of the frontal foramen, reminiscent of our previous findings that inactivation of *Msx1* and *Msx2* in neural crest causes inappropriate conversion of non-osteogenic neural crest-derived cells to osteoblasts (Roybal et al., 2010).”

2.19 - The evidence for Tcf12 in maintaining suture patency postnatally is weak given the unknowns regarding periosteum and dura.

We agree and now more fully acknowledge potential caveats to this experiment.

> Lines 380-386: “However, expression of *Tcf12* in the coronal suture and surrounding tissues was not completely eliminated in *Gli1-CreERT2; Tcf12-f/f* mice following postnatal tamoxifen administration, likely due to incomplete activity of CreERT2 and/or expression of *Tcf12* in some *Gli1-CreERT2*-negative lineage cells. Thus, future experiments using different tools to completely eliminate postnatal Tcf12 function will be needed to rule out roles for Tcf12

postnatally, especially as we observe continued *Tcf12* expression at the coronal suture until at least 4 months after birth."

2.20 - The discussion is lacking resolution. It ends without a summary statement or a suggestion for the next set of experiments and none of what is presented is tied back to Saethre-Chotzen syndrome. The last sentence is a jump that leaves the reader wondering what the point is. How would Tcf12 control rate
of growth? How do you know that rate is what being change and not timing?

We have added an additional paragraph at the end of the Discussion that ties our findings back to Saethre-Chotzen syndrome and highlights important future experiments.

Lines 409-416: "Our study points to similar roles for Twist1 and Tcf12 in controlling embryonic processes essential for coronal suture formation, thus explaining why Saethre-Chotzen Syndrome can be caused by heterozygous inactivating mutations in either *TWIST1* or *TCF12*. The similarities of defects in heterozygous *Twist1* and homozygous *Tcf12* mutant mice further support the model that Twist1 and Tcf12 function as heterodimers to control unique sets of genes important for regulation of osteoblast progenitors. Future experiments will be required to identify genomic regions specifically bound by such heterodimers, as well as the molecular and cellular mechanisms by which these transcription factors regulate the timing and rate of calvarial bone addition to ensure proper coronal suture formation."

2.21 - Figure 1- define all abbreviations in the figure... af, pb, fb? Providing anatomical context for orientation is also needed.

We have provided a diagram of the anatomy of the skullcap in a new Fig. 1A and have made sure that all abbreviations are defined in the legend.

2.22 - Figure 2- would make a greater impact if labeled with mesoderm deletion and neural crest deletion and both. Normally p values are reported as $p<0.001$ if they are that low. Why is the scale bar in K so much larger than the other images? Since there are two different stains in J-O but they initially look similar, it would be good to put the stain target or methodology on the image for example adding ALP to J-N and Lac-Z to O. For Figure 2.I. : are these ratios based on suture length? how was that measured? This should be in the methods section.

We have added "neural crest deletion" and "mesoderm deletion" and both to the figure. We prefer to use exact p values as we feel they are more informative (e.g. in physics only $p<0.00001$ is considered significant). The scale bar in K was a mistake that we have corrected. We explain the LacZ in the legend to panel O and note that all panels in J-O have ALP. We also added a paragraph to the "Quantitation of craniosynostosis and statistical analyses" sub-section of the Methods describing how we measured the ratio of the sagittal to metopic suture.

Lines 186-191: "For the measurement of coronal suture position, and hence the relative proportions of the parietal and frontal bones, we took the ratio of the sagittal suture to the metopic suture. The sagittal suture was defined as the distance between the juncture of the sagittal and lambdoid suture and the juncture of the sagittal and coronal suture. The metopic suture was defined as the distance between the juncture of the metopic and coronal suture and the anterior end of the frontal bones."

2.23 - Are the parietal bones overlapping at the sagittal suture or are they fused at P6 in the Tcf12flox/flox-;Mesp1-Cre and the double mutants? And if they are overlapping, why are they overlapping? This is the morphology of the coronal suture (one bone growing over the other) which the authors explain by the asymmetric presence of Grem1+ mesenchymal cells. Since
parietal bones are mesoderm derived, this would pertain to the effect on mesoderm-derived cells only.

We thank the reviewer for this observation and now note in the text that both *Tcf12* null mice at E18 and P0, as well as mesoderm-deletion mice at P6, display this abnormal overlap of the parietal bones at the sagittal suture. The observation of this abnormal overlap when *Tcf12* is deleted from mesoderm but not neural crest is consistent with these parietal bones being mesoderm-derived.

2.24 - Figure 3- Lack of consistency with abbreviations makes it really hard for the reader to follow any argument or even read the figure. e.g., fs vs ms for the same area. What counter stain was used for these images? if ALP signal is purple, it is difficult to be convinced of staining especially with the LacZ (pink and blue make purple)

We now note that the counter stain is eosin and have carefully gone through the manuscript to make sure all abbreviations are consistent (in particular only using metopic, instead of frontal and metopic interchangeably). While we understand that the blue LacZ stain in Fig. 2O makes the dark ALP stain less prominent, in our hands it was not a problem scoring the size of the nascent parietal bone for new Fig. 2P. Unfortunately, we did not have double deletion E14.5 samples without LacZ to replace the current ones, and to generate new samples would have taken many months of breeding.

2.25 - Figure S1- What is the n (number of sections quantified from what number of individuals)?

In legend to Fig. S1, we now state the number of embryos and number of sections per embryo.

2.26 - Figure 4 H-I What is the n (number of sections quantified from what number of individuals of what genotype)? The histogram for this data is confusing... are the blue and orange areas a percentage of the total number of cells counted? What are the error bars on each section indicating? It seems that these data would be better analyzed using anova rather than a series of t-tests.

We now state the n number for embryos and sections per embryo in the legend. The blue and orange areas show the standard error of the mean for cell counts in the respective boxes for each embryo (i.e. the values in the histograms represent mean of cell counts per box for each of the 5 wild type or mutant embryos used). We used Bonferroni correction for multiple comparisons.

The intent of the comparisons was to determine how Grem1+mesenchyme patterns change in specific domains. Thus, rather than using an ANOVA to determine if all areas between control and mutant are statistically different, we chose to test how specific regions differ between control and mutants, and to correct for multiple testing.

> Legend to Fig. 4: “Error bars represent the standard error of the mean. P values were calculated using a two-tailed Student's t-test with Bonferroni correction for multiple comparisons.”

2.27 - Figure 5 Abbreviations are not consistent making it extremely hard to follow. All abbreviations need to be defined and be consistent throughout the entire manuscript.

We have carefully revised the manuscript to ensure that abbreviations are consistent throughout.

2.28 - Figure 6. Why was 14 weeks chosen as the age? Is it possible that these are just H&E rather than Aliazarin Red and hematoxylin/eosin?

We chose 14 weeks for *Tcf12* in situ hybridization in Fig. 6A as we wanted an adult stage. We see similar expression of *Tcf12* in the coronal suture at 2 months of age in new Fig. 6D. We also thank the reviewer for pointing out the error in describing the stain – this is just H&E and we have corrected it in the legend.

2.29 - Table 1: Data for littermates with the genotypes that are not the focus of the research should be shown for comparison.

We have added data for *Tcf12* and *Twist1* heterozygous littermates to Table 2. We do not have data for individual heterozygotes related to Table 1, though we note that the compound *Tcf12*; *Twist1* heterozygotes have 0% penetrant curly tail, exencephaly, and open ventral body wall phenotypes.

REVIEWER #3:

3.1 - While TCF12 haploinsufficiency in humans is associated with coronal CS, both copies must be mutated in mouse, where the phenotype is still only partially penetrant. Also, while some Tcf12 mutants do not survive the perinatal period, the CS in those who do is mild. One significant point that is still unclear is the extent and underlying cause of death of Tcf12 mutants. First, it is not clearly stated what the survival rate is (observed/expected at specific time points). This seems like data that the authors should already have and could easily present. Importantly, they speculate that the increased severity and penetrance of CS in the Tcf12 conditional mutants shown in Table 2 is because the more severely affected Tcf12 mutants are dying earlier, an explanation difficult to judge without supporting numbers on survival rates. The authors' arguments around this point are confusing. They seem to suggest that mutants with more severe CS are dying prenatally, although normally CS is not lethal.

Unfortunately we did not have precise records of genotypes surviving birth, which is likely compounded by some perinatally lethal pups being cannibalized by the mother shortly after birth. Instead, we now cite the original study on the *Tcf12* null showing that they are born at the expected Mendelian frequency but largely die off in the first two weeks. We agree that lethality is unlikely to be due to the partially penetrant fusion of the coronal suture and instead likely reflects defects in blood, nervous system, or some other tissue that was not the focus of our study. We also performed new analysis at embryonic stages (E17 and E18) and found a similar penetrance and severity of coronal synostosis as at postnatal stages (new Fig. 1B and Table 1). We have therefore removed discussion of the more severely affected *Tcf12* mutants dying earlier as our new data no longer support this model.

> Lines 209-212: "It had previously been reported that homozygous $Tcf12^{-/-}$ mice are born at the expected frequency from a $Tcf12^{+/-}$ incross, with the vast majority dying by 2 weeks after birth (Zhuang et al., 1996). We similarly observed that most homozygous $Tcf12^{-/-}$ mice died within the first 2 weeks."

3.2 - The developmental specimens shown in Fig. 2J-O are intended to provide important data supporting their model of early increased growth in the skull bones. However, it is not easy to judge the size of the bones from those images. The argument would be greatly strengthened by more/better images, or preferably measurements of bone size in multiple samples with statistics. In conjunction, they could compare these conditional mutants to Tcf12 mutants at the same prenatal stages, to support their explanation of the very mild CS observed postnatally in Tcf12 mutants.

We have now quantified the altered frontal and parietal bone growth at embryonic stages in conditional *Tcf12* deletion mutants in new Fig. 2P. This reveals that frontal bone growth is accelerated compared to the parietal bone in neural crest deletion mutants, and parietal bone growth accelerated compared to the frontal bone in mesoderm deletion mutants, although altered bone growth only reaches statistical significance for neural crest *Tcf12* deletion.

3.3 - While a minor point in the paper, the asymmetric distribution of Grem+ cells at the suture is interesting and could point to a mechanism not just for CS in the mutants, but asymmetric bone overlap in WT mice. If available, information about when the asymmetric distribution first arises (and when it first looks different in mutants) could shed additional light on the pathogenesis of CS. In other words, if the bones are growing faster in the mutants, why do they fuse, instead of just forming the normal overlap at an earlier stage?

While we did not detect Grem1 expression at an earlier stage (E13.5), we instead examined expression of Six2 which we had recently reported labeled the coronal suture mesenchyme and was lost in Twist1+/-; Tcf12+/- mice at E15.5. Interestingly, analysis of Six2 at E12.5 and E13.5 reveals that mesenchyme asymmetry is a feature of the frontal and parietal bones well before they meet. We also explain that loss of this asymmetry in mutants may contribute to fusions versus overlap due to bones meeting end-on-end.

> Lines 286-295: "We also attempted to examine whether Grem1+ cell asymmetry was established earlier at E13.5 before the parietal and frontal bones meet but did not observe

significant Grem1 expression at this stage. Instead, we examined the expression of *Six2*, which we had recently shown broadly labels coronal suture mesenchyme (Farmer et al., 2021). We find that, similar to Grem1, *Six2* labels more cells above the frontal relative to above the parietal bone at E14.5, though additional *Six2* expression in the dura below the bones prevented us from assessing relative expression below the frontal and parietal bone. At E13.5 and as early as E12.5, when the frontal and parietal bones are well separated, we also observed more *Six2* expression above the frontal than above the parietal bone. These results indicate that mesenchyme asymmetry above versus below the bones precedes coronal suture formation and requires *Tcf12*, with loss of asymmetry in mutants correlating with the frontal and parietal bones meeting end-on-end and often fusing."

Lines 394-397: "We find that enrichment of *Six2* above the frontal but not above the parietal bone begins as early as E12.5 when the bones are still distant from one another. One possibility is that these asymmetries skew the growth of the two bones off their central axes such that that the parietal reproducibly overlaps above the frontal bone."

3.4 - A final minor point is that for readers less familiar with CS, and with the grading of severity, it might be helpful to provide examples of the index summarized in Table 2.

As the grading of severity has already been reported multiple times, we have chosen to add citations in the Methods to some of our previous papers containing examples along the spectrum. Lines 179-181: "The severity of coronal synostosis was quantified using a scoring method adapted from Oram and Gridley's craniosynostosis index (Oram and Gridley, 2005; Yen et al., 2010) applied specifically to the coronal suture (Teng et al., 2018)."

Second decision letter

MS ID#: DEVELOP/2021/199575

MS TITLE: Embryonic requirements for Tcf12 in the development of the mouse coronal suture

AUTHORS: Man-chun Ting, D'Juan Farmer, Camilla S Teng, Jinzhi He, Yang Chai, Gage Crump, and Robert E Maxson

I have now received 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 overall evaluation is positive and we would like to publish a revised manuscript in Development, provided that the comments of referees (also see Editor's note) can be satisfactorily addressed . Please attend to the comments in your revised manuscript. If you do not agree with any of the suggestions explain clearly why this is so.

Reviewer 1

Advance summary and potential significance to field

The manuscript entitled, Embryonic requirements for Tcf12 in the development of the mouse coronal suture' describes the phenotypic changes upon total or conditional loss of the bHLH transcription factor Tcf12. The conditional deletion was performed in two different compartments neural crest (Wnt1-Cre) and mesenchyme (Mesp1-Cre), and in combination in both compartments. Furthermore, the authors used the Doxycyclin-inducible Gli1-CreERT2 line to delete.
TCF12 can interact with TWIST1, and heterozygous mutations in both genes account for most Saethre-Chotzen syndrome cases. Yet, in mice heterozygosity of Tcf12 does not recapitulate the coronal suture defect. Here, the individual role of Tcf12 in mice was explored.

Comments for the author

As already pointed out last time, the manuscript is well written and conclusive. The results are very well documented and presented in the figures, which are of high quality.
The authors have addressed most of the raised concerns in a satisfying way.
Yet, I am concerned about the statistical method applied by the authors.
Pairwise comparison in form of a paired Student's T-test cannot be used for the statistical analysis of different mouse specimens (based on their different genotypes) originating from various litters.
An unpaired T-test needs to be performed instead.

Second revision

Author response to reviewers' comments

Response to reviewers

Reviewer 1

“As already pointed out last time, the manuscript is well written and conclusive. The results are very well documented and presented in the figures, which are of high quality.
The authors have addressed most of the raised concerns in a satisfying way.

Yet, I am concerned about the statistical method applied by the authors.
Pairwise comparison in form of a paired Student's T-test cannot be used for the statistical analysis of different mouse specimens (based on their different genotypes) originating from various litters.
An unpaired T-test needs to be performed instead.”

Response:

We did actually perform unpaired T-tests but failed to state this clearly. Our intent was to communicate that the analysis was performed between pairs of conditions. However, we now see that our description could be read as indicating that we used paired Student T-tests--when in fact we performed unpaired T-tests. We have now clarified language to make clear that we used unpaired T-tests. For complete transparency, we have also included source files which are now referenced in the Methods section and are included in the Supplementary Information section.

Reviewer 3

“In their revision, the authors adequately addressed my previous concerns. In particular, the genetics data is less confusing now, and a couple of the figures improved with added data. I still had a few minor comments about the figures, partly dealing with newly added data.

1. While the alizarin red staining in Fig 1 is a good addition, the panels are very difficult to see, and would show up better in grayscale.”

Response: We now include grayscale images for fluorescent Alizarin Red images in Figure 1B- G.

2. “I did not catch this in the first submission, but it would be better to change the colors in 5G-J from red and green to magenta and green (as they are in Fig 4), to make them color-blind friendly.”

Response:

We replaced green-red images in Figure 4D,E and Figure 5G,H with green-magenta to facilitate visualization for all readers.
3. “In 6F and G, the purple for Tcf12 doesn’t contrast well with the blue DAPI, and they should use a different color (magenta?).”

Response:

We exchanged the magenta images in Figure 6F, G with images in green to increase contrast and visibility of the RNAscope signal.

Third decision letter

MS ID#: DEVELOP/2021/199575

MS TITLE: Embryonic requirements for Tcf12 in the development of the mouse coronal suture

AUTHORS: Man-chun Ting, D'Juan Farmer, Camilla S Teng, Jinzhi He, Yang Chai, Gage Crump, and Robert E Maxson
ARTICLE TYPE: Research Article

I am satisfied with the response to review and the revision of the figures. Your manuscript has been accepted for publication in Development, pending our standard ethics checks.