Canonical Wnt/β-catenin Activity and Differential Epigenetic Marks Direct Sexually Dimorphic Regulation of Irx3 and Irx5 in Developing Mouse Gonads

Megan L. Koth, S. Alexandra Garcia-Moreno, Annie Novak, Kirsten A. Holthusen, Anbarasi Kothandapani, Keer Jiang, Makoto Mark Taketo, Barbara Nicol, Humphry H. C. Yao, Christopher R. Futtner, Danielle M. Maatouk and Joan S. Jorgensen
DOI: 10.1242/dev.183814

Editor: Patrick Tam

Review timeline

Original submission:	10 August 2019
Editorial decision:	18 September 2019
First revision received:	11 December 2019
Editorial decision:	8 January 2020
Second revision received:	8 February 2020
Accepted:	14 February 2020

Original submission

First decision letter

MS ID#: DEVELOP/2019/183814

MS TITLE: Canonical Wnt/β-catenin Activity and Differential Epigenetic Marks Direct Sexually Dimorphic Regulation of Irx3 and Irx5 in Developing Gonads

AUTHORS: Megan L. Koth, Annie Novak, Kirsten A Holthusen, Anbarasi Kothandapani, Keer Jiang, Makoto Taketo, Barbara Nicol, Hung-Chang Humphrey Yao, Sara Alexandra Garcia-Moreno, Christopher R. Futtner, Danielle M. Maatouk, and Joan S. Jorgensen

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 BenchPressand 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 (please also see editor's note) 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.

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 by Koth et al describes the identification of two regulatory elements associated with the Irx3/5 genes that acquire differential epigenetic marks during gonad development due to the activity of WNT signalling in the ovary. A good deal of evidence is supplied supporting the claims made, derived from in vitro ex vivo and in vivo experiments. The manuscript is well written and measured and I think it is an important contribution to our understanding of how WNT/b-catenin regulates somatic cell fate in the ovary because it reveals a plausible mechanism involving positive control of Irx3/5. It is a nice collaborative effort that makes use of epigenomic data and uses that as a basis for drilling down into that mechanism. The comments/questions below are minor and I make them in order to help improve the clarity of the manuscript especially for the reader outside of the sex determination field.

Comments for the author

Perhaps the least convincing data are in Fig. 3 – mainly due to the size of the error bars in some cases and the use of $p \leq 0.1$ as well as $p \leq 0.05$. For example, the specific trend claimed for H3K27Ac enrichment at -305 kb (B) in the testis fails to convince given the general shape of that graph. The authors could perhaps comment on this and make it clear in the text how/why they are distinguishing between 'trend' and 'significance'.

In Fig. 5B, how the authors account for the very high level of reporter gene activity in the testis, which appears to be generally higher than the ovary?

In Fig. 6, the appearance of the H3K27me3 mark as black bars is likely to cause confusion – why does this ChIPseq data not appear as sequence peaks and troughs, like the ATAC-seq data?

When whole gonads are used for expression analysis, could the authors comment on whether cellular heterogeneity affects interpretation of the data?

In supplementary Fig2 it is worth describing the approximate subcellular location of the proteins detected i.e. b-catenin and IRX3. Are these consistent with known function of those proteins?

In the Discussion at line 417, where Cbx2 is introduced, it is worth also mentioning the other epigenetic events reported to be important for sex determination, such as marks deposited by JMJD1A (Kuroki et al 2013 and Kuroki et al 2017) and CBP/p300 (Carre et al 2018).

As a general comment about the Discussion: it is quite long and a little repetitive i.e. it repeats the description of, and data arising from, the experiments in the Results. It would be better if it were more 'synthetic', and offered some speculation on unknowns that remain e.g. the identity and sexually dimorphic control (how?) of the epigenetic writers/erasers (and readers) that deposit and interpret the differential marks reported. Do such marks drive transcription or are they a consequence of it. Does differential epigenomics account entirely for the mutual antagonism between the testis and ovary pathways? A slightly higher-level view, that connects to trends of thought in the field, would be useful.

I apologise if I have missed these details, but it is important that in any data Figures in which statistics is used, the number of samples (n =…) is clearly stated. The same applies to immunostaining data. In addition please include a description of the genetic background of all mouse strains used, since this is important if others attempt to repeat the experiments (and genetic background will strongly influence any phenotypes and their variability).

Reviewer 2

Advance summary and potential significance to field

While it has been known for a decade that activation of beta-catenin in the bipotential gonad somatic cells drives an ovarian fate and suppresses a testicular fate, the downstream mechanism

has remained unclear. Here, Koth et al. place homeobox TFs Irx3 and 5 directly downstream of TCF in ovarian differentiation. They identify TCF/LEF binding motifs in the Irx3/5 locus which have open chromatin, H3K27acetylation, and are bound by TCF7L2 in the fetal ovary but none of these in the fetal testis. A strength of the work is that they test the putative Irx enhancers in cell lines using luciferase and then transfect constructs into cultured gonads. Herein lies the crux of the paper: that the transfected +86kb and -580kb Irx enhancers drive equally robust expression of the reporter in testes and ovaries. This, together with the differential configuration of chromatin and H3K27me3 at these enhancers, leads the authors to conclude that Irx3/5 are activated by TCFL2 exclusively in the fetal ovary because the locus is in an epigenetically favorable configuration. Although the beta-catenin pathway is already specifically activated in female somatic gonad, the epigenetic regulation of Irx3/5 is an additional layer of control in sex differentiation that has been identified.

Comments for the author

The experiments are well designed and rigorously performed. The following should be addressed before it is ready for publication.

1) Does epigenetic regulation of Irx3/5 occur independently of the b-catenin pathway? Given that b-catenin activation can drive ovarian fate in XY gonads, is the downstream activation of Irx3/5 less robust in XY, and are +86bp and -580kb enhancers likewise in a sex-specific epigenetic state in the b-cat GOF testes? The epigenetic regulation of Irx may explain why male to female sex reversal is incomplete in the beta-catenin GOF mouse.
2) While the ChIP experiments are well designed and executed with 3 biologic replicates, the standard deviations are large and significance levels are low. The results may be cleaner with 1-2 more replicates in Figure 3. It would be appropriate to report the exact p values.
3) Figure 1 compares transcript levels in e14.5 gonads from Sf1-cre x active beta-catenin f/f to e11.5 gonads cultured 24 hours in LiCl. This is a difficult comparison, since gonads in culture do not necessarily develop at the same rate as in vivo, however the ATAC-seq first shows opening of chromatin at the Irx3/5 locus at E13.5, so the timing here is somewhat confusing. Have the authors examined Irx gene expression at an earlier timepoint in the beta-catenin GOF, or a later timepoint in the culture? While a minor point, the time of culture should be stated directly in the figure or results (not just in the methods).
4) The readability to a broader audience could be improved in the figures with a few additions, such as cartoons to describe the gonad culture experimental design in Figs 1 and 5 and inclusion of a map of the Irx promoters used in Fig. 5A in the supplement. A summary cartoon of the model of sex-specific epigenetic regulation of the Irx3/5 locus would be helpful.

Reviewer 3

Advance summary and potential significance to field

This paper investigates the possible regulation of Irx3 and 5 by WNT-b-catenin signalling and examines the enrichment of epigenetic marks on potential enhancer elements identified in these genes. The findings have implications for Irx3 and Irx5 regulation in non-gonadal tissues and provide an example of a signalling pathway whose activity upstream of the target genes may alter epigenetic state at the target loci. However, a direct relationship between the levels of Irx3 and 5 observed in XX gonads and altering the epigenetic state in XY gonads was not obvious and how b-catenin interacts to change epigenetic and consequently alter transcriptional outcomes was not demonstrated. The main significance therefore appears to be in the identification of potential Irx enhancers through which b-catenin may regulate these genes. As the same enhancers do not appear to operate in XX germ cells, it is not clear that the same mechanisms will act in other tissues discussed in the study. While very interesting, it is therefore difficult to assess the broader significance of these enhancers with respect to development of other tissues.

Comments for the author

The study investigates the regulation of Irx3 and 5 by WNT-b-catenin signalling and examines the epigenetic state across regions in which enhancer elements were identified in these genes. While the authors provide good evidence that canonical b-catenin regulates Irx3 / 5 in somatic cells of

the ovary (Figure 1), the analysis of the potential enhancer elements through which this activity may be mediated is less convincing (Figures 3-6). This may in part be due to analysis of transfections in whole ovaries rather than in the relevant target somatic cells (developing granulosa cells). This is a difficult task though, and the challenge of getting appropriate construct activity in these cells may hamper the collection of more convincing data. Perhaps this could be overcome by making constructs in vivo using crispR or a similar approach, although this is a significant challenge. The genomic region was analysed in this study was large, but the detail of epigenetic modifications in the specific enhancer regions tested was limited. The study would significantly benefit from more detailed analyses demonstrating the epigenetic state and the activity of +86kb and -580kb enhancer elements identified.

Specific comments:

1. Figure 1A: The iCRT14 drug doses used seem very high. Either this drug is poorly available to the cells or not very potent. It has an IC50 of 40nM in assays for b-catenin responsive transcription, but in BT549 cells used at 10, 20 50uM to inhibit proliferation. Compared to the IC50, the dose used makes me wonder whether the effects in cells are specific. While the data are consistent with the Sf1cre model, arguing in favour that the findings using the drug are reliable, the drug used here does not appear optimal.

2. An ANOVA analysis would seem appropriate in Fig 1A, but a t-test was used.

3. In Figure 1 the authors state that a 0.23-fold change was obtained for Axin2, but the level is actually 23% of control - ie a 75% or 4-fold reduction. This also applies to other genes. Presenting the data as a 0.23-fold change is not accurate. It would be better to represent the fold change for reduced expression as a % decrease. This would actually favor the authors argument that there is a strong reduction in expression.

4. Figure 1C: 50mM LiCl seems a relatively blunt tool for manipulation of b-catenin activity. Is no better drug available? Why was DMSO used as a vehicle – LiCl is solubilized in water or PBS? Why was no dose curve used for LiCl?

5. Figure 3: These data are indicative, but not particularly compelling especially considering a number of differences are measured using a p value of 0.1 and students t test measuring several regions together (ANOVA + post hoc may be more appropriate). Perhaps cleaner and more compelling data would be obtained from purified somatic cells rather than whole ovaries.

6. Figure 5: There is no data to show how efficient the transfection was in this experiment. How did the authors measure transfection efficiency and validate that the experiment has worked? It is surprising that there were no sex specific differences detected here. The full promoter did not appear to confer the same activity as the artificial construct shown in 5B. What is the explanation for this?

7. Figure 5B: Why are the levels in the testis higher than shown in the ovary. Why do the mutations in the reporter not reduce activity in testis?

8. Figure 5: Does this include data from all three experiments combined, or one example (ie one experiment)?

9. The epigenetic differences shown across the region (Figure 2 and Figure 6) are very interesting, but not very detailed and are not validated. As the regions identified are specifically tested in promoter activity assays, validation of their epigenetic state at a more detailed level (ie less than 750kb region) is desirable.

10. Figure 6: No detail is shown for the H3K27me3 enrichment at the loci investigated – just a line in the figure showing regions that are reportedly enriched for H3K27me3 based on data from a separate study. No data is shown for H3K27me3 reductions either at a whole cell level or at specific loci in the TAZ treated samples. The experiment should at least be supported by analysing H3K27me3 levels in the TAZ treated vs untreated cells and preferably across the regions of interest in the samples tested.

11. Figure 6: While there was a significant increase in TAZ/DMSO in the XY samples for Irx3 and Irx5, this response was far below what was observed in Fig 1D for these genes in XX samples. In addition, did not appear to be any differences in the level of transcriptional activation for Irx3, Irx5 for TAZ/DMSO between XY and XX. But the response for p21 was greater in XX than in XY. This is confusing.

12. The description of the changes for Irx3 and Irx5 in Lines 275-278 is not very clear. I am not sure why the transcriptional response for Irx3 and Irx5 should be directly comparable to the positive control p21? Surely this depends on the normal observed transcriptional increase for Irx3 and Irx5 in ovaries vs testes compared to what is observed for p21. From Figure 1, the XX vs XY difference for Irx3 and Irx5 appears to be around 15-20 fold in response to b-catenin activation using the Cre model. I appreciate that TAZ may not result in the same response as b-catenin levels are different in XX vs XY. But it is not clear how much response is really generated by TAZ treatment in XY gonads compared to what can be expected in XX gonads.
This data could be presented two ways 1. As a TAZ/DMSO ratio as provided. 2. As a comparison of Irx3 and Irx5 between untreated XX gonads and TAZ treated XY gonads to show how much of the normal XX levels are reached in TAZ treated XY gonads.

Minor comments:
1. Fig 2B does not add anything new to 2A. If it is to be included, it would be better drawn to match Fig2A with the colored parts shown exactly under their equivalent parts in Fig 2A

First revision

Author response to reviewers' comments

Thank you and the reviewers for your time and consideration of our manuscript entitled Canonical Wnt/β-catenin Activity and Differential Epigenetic Marks Direct Sexually Dimorphic Regulation of *Irx3* and *Irx5* in Developing Gonads, MS ID#: DEVELOP/2019/183814. We especially commend the reviewers' thoughtfulness, as we truly believe that their suggestions added significant value to the manuscript. Accordingly, we have submitted a significantly revised edition of this manuscript for your consideration.

We have carefully considered the comments provided by each reviewer and have addressed each comment below. As a result of additional experiments and data interpretation, our manuscript has significant changes. The revised manuscript shows omitted text via ~~strikethrough~~ of text and new text in red.

Thank you again for considering this manuscript for publication in *Development*. We feel that the suggestions provided by reviewers were extremely helpful and that our revisions have improved the presentation of our data. We hope you and the reviewers agree. We certainly look forward to your reply.

MS# DEVELOP/2019/183814

Thank you to Editor and Reviewers for thoughtful comments that have helped us to improve the presentation of our data.

Reviewer 1 Comments for the Author:
1. Perhaps the least convincing data are in Fig. 3 – mainly due to the size of the error bars in some cases and the use of p≤0.1 as well as p≤0.05. For example, the specific trend claimed for H3K27Ac enrichment at -305kb (B) in the testis fails to convince given the general shape of that graph. The authors could perhaps comment on this and make it clear in the text how/why they are distinguishing between 'trend' and 'significance'.

We acknowledge the variability in the results from whole gonad ChIP experiments shown in Figure 3. While *Irx3* and *Irx5* expression is confined to pregranulosa cells at this timepoint (E13.5-14.5, Fu et al. 2018 PLoS Genetics), we used antibodies for H3K27Ac and TCF7L2, which are present in many cell populations. TCF7L2 is expressed in many somatic cell types of the ovary and testis (Jameson 2012, PLoS Genet, 8, e1002575; GUDMAP database) and H3K27Ac is present in all cells of both testis and ovary; therefore, the variability was expected. To address this comment and to increase clarity of the data, we reported exact statistical p values for each site and highlighted only those with $p < 0.05$ with an asterisk. With this change, we also decided that the distinction between 'trend' and 'significance' was unnecessary and therefore eliminated the descriptions as such. Language has been updated in the Results section starting on page 9, Line 192.

2. In Fig. 5B, how the authors account for the very high level of reporter gene activity in the testis, which appears to be generally higher than the ovary?

This was a common concern among reviewers, and we were also surprised by this finding. Statistical analyses did not indicate any difference between testis and ovary; however, one reason that the testis results might generally be higher is that they are quite a bit easier to inject due to their larger size. We normalized all transfections to SV40 Renilla activity to control for variations in transfection efficiency. In general, luciferase and renilla values were higher in testes, but their ratios—and further normalization to the pBasic luciferase control gonads allowed for direct comparisons.

As noted in later sections of the manuscript, we recognized that the plasmid DNA vectors that we transfected have no epigenetic decorations; therefore, there is essentially unrestricted access to sequences within the promoter in both sexes. Thus, any potential repressor marks would not be present to block activity. To address this concern further, we also examined the proximal promoter sequences of *Irx3* and *Irx5* in more detail. Our H3K27me3-Seq data showed that H3K27me3 marks are highly enriched in XY but were absent in XX cells in the case of both promoters. These data have been included in new **Supplementary Figure S5**. In addition, it has been recognized that β-catenin and Sox factors belong to the same family of transcription factors and can bind similar sequences. We scanned the promoter sequences using the UCSC and JASPAR databases and uncovered several potential binding sites including several Sox binding sequences along with other interesting sites including GATA, EZH2, CEBP, SP1 among others. To increase clarity, we added and reinforced these details in the Discussion section (page 14, starting at Line 356).

3. In Fig. 6, the appearance of the H3K27me3 mark as black bars is likely to cause confusion - why does this ChIP-seq data not appear as sequence peaks and troughs, like the ATAC-seq data?

Thank you for pointing this out to us. We have restructured **Figure 6** to look more like the ATAC-Seq data in Figure 2.

4. When whole gonads are used for expression analysis, could the authors comment on whether cellular heterogeneity affects interpretation of the data?

The expression data in Figure 1 were generated from whole gonads. While *Rps29* and *Axin2* may be present in a number of cell populations, *Fst*, *Irx3* and *Irx5* are only expressed within the pregranulosa cell population during the developmental stages used here. Reports from other groups have shown that β-catenin transcriptional activity is confined to the somatic cell population of the ovary and the impact of *in vivo* β-catenin manipulation (*SF1Cre*$^{+}$*;Ctnnb1*$^{flox/flox}$ and *SF1Cre*$^{+}$*;Ctnnb1*$^{\Delta ex3}$) was targeted specifically to somatic cells using SF1Cre in developing gonads (Chassot et al. 2008 Hum Mol Genet; Tomizuka et al. 2008 Hum Mol Genet.; Maatouk et al. 2008 Hum Mol Genet; Bernard et al. 2012 Endocrinology). In addition, our results indicate that canonical β-catenin activity within female oocytes has no impact on IRX3 expression (**Supplemental Figure S2**). Thus, we feel comfortable that the changes in gene expression shown in Figure 1 using whole gonads represents a relevant change in gene expression to the somatic cell population in particular.

5. In supplementary Fig S2 it is worth describing the approximate subcellular location of the proteins detected i.e. β-catenin and IRX3. Are these consistent with known function of those proteins?

Thank you for your observation regarding subcellular location of β-catenin and IRX3 in **Supplemental Figure S2**. We previously reported the surprising finding that IRX3 was present in the cytoplasm and nucleus of oocytes starting when germline nest breakdown commences (Fu et al. 2018, PLoS Genetics). Further, it has been established that β-catenin can be found in many subcellular locations within the oocyte (Yan et al., 2019 Cell Death Dis; Bothun and Woods, 2019 Histochem Cell Biol; Kumar et al., 2016, Sci Rep; Usongo et al., 2012, Reproduction; Chassot et al., 2011, PLoS One; Jameson et al., 2012, PLoS Genet). Currently, however, the roles for IRX3 and β-catenin within the oocyte are unknown and are under intense investigation. One study in particular showed that transcriptional activity of β-catenin is only present in growing oocytes of postnatal ovaries, starting at the secondary follicle stage. **Figure S2** shows P7 postnatal ovaries that harbor mostly early stage follicles along with some secondary stage follicles. We detected no change in IRX3 in all follicles examined, including secondary stage follicles. We changed the inset to focus on a secondary follicle and have added text within the Figure Legend to address how the timing relates to transcriptional activity.

6. In the Discussion at line 417, where Cbx2 is introduced, it is worth also mentioning the other epigenetic events reported to be important for sex determination, such as marks deposited by JMJD1A (Kuroki et al 2013 and Kuroki et al 2017) and CBP/p300 (Carre et al 2018).

Thank you for your important comments. We are combining our responses to comments 6 and 7 because we decided to incorporate the suggestions from comment 6 into a reconfigured discussion section (comment 7). Specific to the comments in 6, we added text to discuss the overall view of epigenetics within a new paragraph in the discussion (page 15, starting at Line 397). In addition, we believe that we have streamlined the discussion towards a more higher-level view. We hope the reviewers agree.

7. As a general comment about the Discussion: it is quite long and a little repetitive i.e. it repeats the description of, and data arising from, the experiments in the Results. It would be better if it were more 'synthetic', and offered some speculation on unknowns that remain e.g. the identity and sexually dimorphic control (how?) of the epigenetic writers/erasers (and readers) that deposit and interpret the differential marks reported. Do such marks drive transcription or are they a consequence of it. Does differential epigenomics account entirely for the mutual antagonism between the testis and ovary pathways? A slightly higher-level view, that connects to trends of thought in the field, would be useful.

Response combined with number 6 above.

8. I apologise if I have missed these details, but it is important that in any data Figures in which statistics is used, the number of samples (n =…) is clearly stated. The same applies to immunostaining data. In addition, please include a description of the genetic background of all mouse strains used, since this is important if others attempt to repeat the experiments (and genetic background will strongly influence any phenotypes and their variability).

These are important details and we have added them to each figure legend. We also clarified strains of mice used for each experiment within the Methods section.

Reviewer 2 Comments for the Author:
The experiments are well designed and rigorously performed. The following should be addressed before it is ready for publication.
1. Does epigenetic regulation of Irx3/5 occur independently of the β-catenin pathway? Given that β-catenin activation can drive ovarian fate in XY gonads, is the downstream activation of Irx3/5 less robust in XY, and are +86bp and -580kb enhancers likewise in a sex-specific epigenetic state in the β-cat GOF testes? The epigenetic regulation of Irx may explain why male to female sex reversal is incomplete in the beta-catenin GOF mouse.

This is an excellent and very interesting point. We considered several avenues by which we might try to explore this idea. We believe that the best approach would be to isolate gonads from XY *SF1Cre;Ctnnb1*$^{\Delta ex3}$ control and mutant embryos for ChIP-PCR using H3K27Me3 and other epigenetic marks. Unfortunately, current techniques require 100-150 pairs of gonads to perform a single ChIP (for up to 4 antibodies/ChIP, including RNA Pol II and IgG controls), and we are unable to collect the appropriate genotype gonads to perform this experiment within a reasonable timeframe. This paradigm, however, is under serious consideration for future ChIP-Seq assays.

We also do believe that other events besides β-catenin activity is likely to contribute to ovary specific expression of *Irx3/5*. This statement is based on 1) identification of several other potential ovary-specific ATAC-Seq peaks (Figure 2) and 2) the fact that ~40% Irx3 and Irx5 expression remained following somatic cell-specific loss of β-catenin (Figure 1C). We have included text to discuss this within the Discussion section (see page 13, starting Line 339).

2. While the ChIP experiments are well designed and executed with 3 biologic replicates, the standard deviations are large and significance levels are low. The results may be cleaner with 1-2 more replicates in Figure 3. It would be appropriate to report the exact p values.

This was also a concern for reviewer 1 (please see response to Reviewer 1, Question 1). While we agree that 1-2 more replicates will improve the size of the error bars, we did repeat each between 3 - 9 times. Ultimately, we don't believe that the end result will be altered to the extent to warrant the resources required. To increase clarity, we included the n's for each experiment along with exact p values in Figure 3 and believe that this is an important addition that will aid the reader's interpretation of the results.

3. Figure 1 compares transcript levels in e14.5 gonads from Sf1-cre x active beta-catenin f/f to e11.5 gonads cultured 24 hours in LiCl. This is a difficult comparison, since gonads in culture do not necessarily develop at the same rate as in vivo, however the ATAC-seq first shows opening of chromatin at the Irx3/5 locus at E13.5, so the timing here is somewhat confusing. Have the authors examined Irx gene expression at an earlier timepoint in the beta-catenin GOF, or a later timepoint in the culture? While a minor point, the time of culture should be stated directly in the figure or results (not just in the methods).

Thank you for pointing these inconsistencies out to us. Our previous results showed that endogenous Irx3 and Irx5 expression increased at the onset of sex differentiation but did not change much between E13.5 – E15.5 (Kim et al. 2011 Dev. Biol; Kim et al. 2011 Biology of Reprod; Fu at al. 2018 PLoS Genetics). An important feature to the pharmacological cultures was to manipulate the Wnt/□-catenin pathway before the onset of robust sex differentiation activity. To clarify the timing, we included a methodology timeline in new Figure 1A as suggested.

4. The readability to a broader audience could be improved in the figures with a few additions, such as cartoons to describe the gonad culture experimental design in Figs 1 and 5 and inclusion of a map of the Irx promoters used in Fig. 5A in the supplement. A summary cartoon of the model of sex-specific epigenetic regulation of the Irx3/5 locus would be helpful.

Thank you for this helpful suggestion. We included a basic methodology cartoon model in Figure 1 and reference it for the experiment in Figure 5 as suggested. We also developed an overall model to summarize conclusions (see Discussion section, page 12, starting Line 310) and included as a new Figure 7.

Reviewer 3 Comments for the Author:
The study investigates the regulation of Irx3 and 5 by WNT-β-catenin signaling and examines the epigenetic state across regions in which enhancer elements were identified in these genes. While the authors provide good evidence that canonical β-catenin regulates Irx3 / 5 in somatic cells of the ovary (Figure 1), the analysis of the potential enhancer elements through which this activity may be mediated is less convincing (Figures 3-6). This may in part be due to analysis of transfections in whole ovaries rather than in the relevant target somatic cells

(developing granulosa cells). This is a difficult task though, and the challenge of getting appropriate construct activity in these cells may hamper the collection of more convincing data. Perhaps this could be overcome by making constructs in vivo using crispR or a similar approach, although this is a significant challenge. The genomic region was analysed in this study was large, but the detail of epigenetic modifications in the specific enhancer regions tested was limited. The study would significantly benefit from more detailed analyses demonstrating the epigenetic state and the activity of +86kb and -580kb enhancer elements identified.

Specific comments:
1. Figure 1A: The iCRT14 drug doses used seem very high. Either this drug is poorly available to the cells or not very potent. It has an IC50 of 40nM in assays for β-catenin responsive transcription, but in BT549 cells used at 10, 20 50uM to inhibit proliferation. Compared to the IC50, the dose used makes me wonder whether the effects in cells are specific. While the data are consistent with the Sf1cre model, arguing in favour that the findings using the drug are reliable, the drug used here does not appear optimal.

Thank you for this interesting insight. We based our experimental design on the publication that first described the discovery of the iCRT drugs (Gonsalves, Klein, Carson, Katz, Ekas, Evans, Nagourney, Cardozo, Brown, DasGupta PNAS April 12, 2011 108 (15) 5954-5963; https://doi.org/10.1073/pnas.1017496108). iCRT14 was one of few that inhibited the transcriptional properties of ▫-catenin. We were convinced by their data including cell culture experiments that used 25▫M and 50▫M for the purpose of showing transcription inhibition. It may be that differences in responsiveness can also be attributed to differences in drug absorption in monolayer cell culture compared to explant tissue culture. Ultimately, we felt comfortable with our data based on our findings that included both positive and negative controls along with its agreement with the *in vivo* experiments. Similar to cell culture experiments in the Gonsalves, et al. paper, we observed a dose responsive repression of genes, which also suggests selective pharmacological activity of the drug. In sum, to address rigor in the experiments that are outlined in Figure 1, we evaluated loss- and gain- of-function impacts using both *ex vivo* and *in vivo* approaches. Altogether, these suggest a link between canonical β-catenin transcriptional activity and *Irx3* and *Irx5* and gave us the premise to proceed with investigations that would provide a direct regulatory link.

2. An ANOVA analysis would seem appropriate in Fig 1A, but a t-test was used.

Thank you for pointing out this error, we have corrected and reported the new statistical test results in new Figure 1B.

3. In Figure 1 the authors state that a 0.23-fold change was obtained for Axin2, but the level is actually 23% of control - ie a 75% or 4-fold reduction. This also applies to other genes. Presenting the data as a 0.23-fold change is not accurate. It would be better to represent the fold change for reduced expression as a % decrease. This would actually favor the authors argument that there is a strong reduction in expression.

We agree with this interpretation and have therefore altered our text to reflect a percent decrease in activity. Please see new text on page 6, starting Line 114.

4. Figure 1C: 50mM LiCl seems a relatively blunt tool for manipulation of β-catenin activity. Is no better drug available? Why was DMSO used as a vehicle - LiCl is solubilized in water or PBS? Why was no dose curve used for LiCl?

We agree that LiCl is a relatively 'blunt' tool; however, it is also recognized as a long-established tool to evaluate canonical β-catenin activity. We based our *ex vivo* treatment strategy on literature that reported dose ranges between 20-50mM (*please see References 1-5 below*) and our design was based on one particular publication that used the same *ex vivo* paradigm (*Reference 6*). In addition, to maintain rigor in our experimental interrogation, we evaluated the impact of stabilized β-catenin activity using a genetic mouse strategy (*SF1Cre;Ctnnb1$^{\Delta ex3}$*) to complement the *ex vivo* culture experiment.

Thank you for pointing out our error in reporting the LiCl control as DMSO. We had mistakenly used this control as it was the appropriate control for the iCRT14 drugs. To correct this oversight, we repeated experiments using water as the appropriate control for LiCl and observed similar results. These new data are now included in new Figure 1D.

References:
1. Davis, Weakland, Wieland, Farese, West. 1987 Proc. Natl Acad Sci 84:3728-3732
2. Tanwar, Kaneko-Tarui, Zhang, Rani, Taketo, Teixeira. 2010 Biol Reprod 82:422-432
https://doi.org/10.1095/biolreprod.109.079335
3. Usongo, Li, Farookhi. 2012 Dev Dyn 242:291-300.
https://doi.org/10.1002/dvdy.23919
4. Bai, Chang, Cheng, Chu, Leung, Yang. 2017 Endocrinology 158:2813-2825.
https://doi.org/10.1210/en.2017-00287
5. Bernard, Ryan, Sim Czech, Sinclair, Koopman, Harley. 2012 Endocrinology 153:901- 912.
6. Maatouk, DiNapoli, Alvers, Parker, Taketo, Capel. 2008 Hum Mol Genet 17:2949-2955.

5. Figure 3: These data are indicative, but not particularly compelling especially considering a number of differences are measured using a p value of 0.1 and students t test measuring several regions together (ANOVA + post hoc may be more appropriate). Perhaps cleaner and more compelling data would be obtained from purified somatic cells rather than whole ovaries.

All three reviewers commented on this particular Figure (see responses to Reviewer 1, comment 1 and Reviewer 2, comment 2). We have added text to the results section to explain the variability in data and have eliminated the confusing language regarding trend vs significance. We acknowledge that the homogeneous population within whole gonad ChIP contributes to data variability and agree that we would likely achieve cleaner data from purified cells; however, the approach using isolated somatic cells took over a year to obtain sufficient cells numbers for the genomic-Seq data. Even the whole gonad ChIP requires accumulation of 100-150 pairs of gonads for each experiment, which takes approximately 3- 4 months to acquire. Therefore, we consider somatic cell ChIP to be beyond the scope of this manuscript at this time. To address the variability in the results, we updated the Results section (see page 9, starting Line 192).

6. Figure 5: There is no data to show how efficient the transfection was in this experiment. How did the authors measure transfection efficiency and validate that the experiment has worked? It is surprising that there were no sex specific differences detected here. The full promoter did not appear to confer the same activity as the artificial construct shown in 5B. What is the explanation for this?

The details for transfection assays were not included in this manuscript. We did however, include language within the text to point readers to a previous paper from our laboratory where we described this technique in detail (Gao L., et al. 2011, Biol Reprod 84: 422). We acknowledge that transfection efficiency is typically not achieved to the same level as would be expected in monolayer cell culture; however, there are no cell lines that represent fetal gonad somatic cells available at this time and we developed this technique to provide an instrument that would allow acquisition of more physiologically relevant significance. To address this potential issue, we employ rigorous means to normalize data, first, gonads are cotransfected with SV40-Renilla, which is used as an internal transfection efficiency control and then second, each experiment is normalized to gonads that were injected with the promoterless control (pBasic/SV40-Renilla).

We were also surprised that our data indicated no sexually dimorphic differences in activity. We interpreted this data to indicate that there was additional information required that was not present in our plasmid DNA constructs. In addition, it is recognized that plasmid DNA is devoid of epigenetic decorations, which would allow potentially unrestricted access to all sequences in ovary and testis. To address this, we evaluated our H3K27me3 ChIP-Seq data and found that there was substantial enrichment on promoter sequences in XY, but not XX cells. This is shown in new **Supplementary Figure S5**. This was also commented upon by Reviewer 1 (please see response to their comment # 2). We included new text within the Discussion section to address these important caveats (see page 14, starting line 356).

Thank you for your question regarding the difference in activity for the full-length promoter and the artificial construct that contains enhancer + full-length promoter in 5B. We have also wondered about this. The obvious difference is the addition of the enhancers; therefore, we searched those sequences for potential binding sites of factors that may inhibit activity in gonads of either sex. One particularly intriguing possibility was the identification of several SOX9 binding sites. Studies to further evaluate this possibility are underway.

7. Figure 5B: Why are the levels in the test is higher than shown in the ovary. Why do the mutations in the reporter not reduce activity in testis?

This was a common concern among reviewers, and we were also surprised by this finding. Statistical analyses did not indicate any difference between testis and ovary; however, one reason that the testis results might generally be higher is that they are quite a bit easier to inject due to their larger size. We normalized all transfections to SV40 Renilla activity to control for variations in transfection efficiency. In general, luciferase and renilla values were higher in testes, but their ratios—and further normalization to the pBasic luciferase control gonads allowed for direct comparisons.

The data showing ovary-specific sensitivity to the mutations in the enhancers was very exciting to us. There are two mutations, one in each enhancer. Notably, the mutations are the same—a single base pair mutation that changes the TCF/LEF binding site from canonical to non-binding. This is the exact mutation that differentiates the widely used TOP- flash vs FOP-flash reporters. Thus, we interpreted this finding to show exquisite sensitivity to β-catenin/TCF/LEF DNA binding and therefore, support our hypothesis that canonical β- catenin directly regulates the Irx3/5 locus, which can only be achieved within the ovary where β-catenin transcriptional regulation occurs. Regarding testis expression, we recognize that the plasmid DNA lacks epigenetic marks, thus making all sequences accessible to male-specific factors. For example, SOX factors can bind the same sequences as TCF/LEF and there are likely others that are exposed. These factors are similar to the responses detailed above (Reviewer 1, comment #2, and the response immediately above, #6).

8. Figure 5: Does this include data from all three experiments combined, or one example (ie one experiment)?
These data include all experiments combined. To clarify, we report the n for each experiment.

9. The epigenetic differences shown across the region (Figure 2 and Figure 6) are very interesting, but not very detailed and are not validated. As the regions identified are specifically tested in promoter activity assays, validation of their epigenetic state at a more detailed level (ie less than 750kb region) is desirable.

We agree that the data shown for ATAC-Seq (in addition to DNAse-Seq, data not shown) were not very detailed and indicated only that chromatin in specific regions was accessible. We included new information regarding ATAC-Seq and H3K27me3 ChIP-Seq data for the *Irx3* and *Irx5* proximal promoters in XX vs XY cells in new **Supplementary Figure S5**. Of note, there is abundant enrichment of the repressor mark on both promoters only in XY cells. As pointed out, we did clone specific regions into reporter constructs. To evaluate the ATAC-Seq data, we used an antibody against the active mark, H3K27Ac, for each region using ChIP-PCR. The primers were designed to focus specifically to the region of the canonical ▫-catenin/TCF/LEF binding site with amplicons no larger than 100bp. We did not repeat the H3K27Me3 ChIP-Seq data as we considered that additional experiments using whole gonads would not add significant new knowledge to an already robust evaluation using isolated cells of interest. As suggested, we incorporated new insets of our ChIP-Seq data into **Figure 6** and focused on ~5kb of +86kb and -580kb enhancer sites in addition to *Irx3* and *Irx5* promoter sequences mentioned above.

10. Figure 6: No detail is shown for the H3K27me3 enrichment at the loci investigated - just a line in the figure showing regions that are reportedly enriched for H3K27me3 based on data from a separate study. No data is shown for H3K27me3 reductions either at a whole cell level or at specific loci in the TAZ treated samples. The experiment should at least be supported by

analysing H3K27me3 levels in the TAZ treated vs untreated cells and preferably across the regions of interest in the samples tested.

This is an interesting point and is similar to comment 1 from Reviewer 2. We considered several avenues by which we might try to explore this idea. We believe that the best approach would be to isolate gonads from XY *SF1Cre;Ctnnb1*$^{\Delta ex3}$ control and mutant embryos for ChIP-PCR using H3K27Me3 and other epigenetic marks. Unfortunately, current techniques require 100-150 pairs of gonads to perform a single ChIP (for up to 4 antibodies/ChIP, including RNA Pol II and IgG controls), and we are unable to collect the appropriate genotype gonads to perform this experiment within a reasonable timeframe. This paradigm, however, is under serious consideration for future ChIP-Seq assays.

11. Figure 6: While there was a significant increase in TAZ/DMSO in the XY samples for Irx3 and Irx5, this response was far below what was observed in Fig 1D for these genes in XX samples. In addition, did not appear to be any differences in the level of transcriptional activation for Irx3, Irx5 for TAZ/DMSO between XY and XX. But the response for p21 was greater in XX than in XY. This is confusing.

Thank you for this important question. While it is tempting to compare expression levels between treatment paradigms (Figure 1 vs Figure 6), we believe each treatment paradigm tests very different questions and each drug will have unique effects. To address the specific comparison in Figure 1D; these data are from XY testes from mice carrying a stabilized form of ▫-catenin. This model is very specific to a transcription factor and is expected to be extremely robust based on its phenotype (Maatouk DM et al. 2008 Hum Mol Genet 17:2949). The goal for the TAZ-treated gonads was to inhibit a single DNA methylation pathway related to H3K27Me3. We recognize that this experiment has limitations as it interferes with a pathway that likely impacts other factors. A statement to this has been included in the discussion section (see page 14, starting Line 378).

As noted, the transcript levels for *Irx3* and *Irx5* are not different between treated ovaries and testes, but the response for p21 is quite different. Our treatment time course (E11.5 onset and 48 hours culture) falls during the most profound period of sex differentiation; thus, we expect sex-specific transcriptomes that will respond in different ways to drug treatment. We rephrased our results section to clarify how we interpreted the data (page 12, starting Line 283).

12. The description of the changes for Irx3 and Irx5 in Lines 275-278 is not very clear. I am not sure why the transcriptional response for Irx3 and Irx5 should be directly comparable to the positive control p21? Surely this depends on the normal observed transcriptional increase for Irx3 and Irx5 in ovaries vs testes compared to what is observed for p21. From Figure 1, the XX vs XY difference for Irx3 and Irx5 appears to be around 15-20 fold in response to b- catenin activation using the Cre model. I appreciate that TAZ may not result in the same response as b-catenin levels are different in XX vs XY. But it is not clear how much response is really generated by TAZ treatment in XY gonads compared to what can be expected in XX gonads. This data could be presented two ways 1. As a TAZ/DMSO ratio as provided. 2. As a comparison of Irx3 and Irx5 between untreated XX gonads and TAZ treated XY gonads to show how much of the normal XX levels are reached in TAZ treated XY gonads.

We appreciate the importance of this question and evaluated different means of reporting the results. In the end, we opted to go with the original report (option 1) because we believe that it is difficult to compare to untreated samples and across sexes, especially at this particular time in development (see response for comment 11 above). We included additional text to increase clarity in describing the experiment and its interpretation (see page 12, starting Line 283).

Minor comments:
1. Fig 2B does not add anything new to 2A. If it is to be included, it would be better drawn to match Fig2A with the colored parts shown exactly under their equivalent parts in Fig 2A

Thank you for this helpful suggestion. We included the chromatin map within Figure 2A (now just figure 2) as suggested.

Editor's note:

1. Refine the statistical reporting of quantitative data (reviewer 1 and3)

These have been refined as suggested.

2. Clarify the co-localization of IRX3/5 and beta-catenin

This was clarified, see response to Reviewer 1, comment 5.

3. Consider examining the Irx3/5 locus in sex reverse mutants

We have considered this suggestion and do believe it is an excellent idea; however, the ChIP experiments require 100-150 gonads/ChIP, which takes over 3 months to acquire on a wild type background. Accumulation of sufficient gonads for a sex-specific mutant is likely to take over a year and thus, not feasible for this particular manuscript. This is, however, under serious consideration for ChIP-Seq experiments for future analysis.

4. Please streamline the Discussion (reviewer 1)

We altered the Discussion considerably.

Second decision letter

MS ID#: DEVELOP/2019/183814

MS TITLE: Canonical Wnt/β-catenin Activity and Differential Epigenetic Marks Direct Sexually Dimorphic Regulation of Irx3 and Irx5 in Developing Gonads

AUTHORS: Megan L Koth, Annie Novak, Kirsten A Holthusen, Anbarasi Kothandapani, Keer Jiang, Makoto Taketo, Barbara Nicol, Hung-Chang Humphrey Yao, Sara Alexandra Garcia-Moreno, Christopher R Futtner, Danielle M Maatouk, and Joan S Jorgensen

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 BenchPressand click on the 'Manuscripts with Decisions' queue in the Author Area.

As you will see, the referees recognised that some of the concerns raised in the previous review have been addressed satisfactorily but there are issues that need clarification or further attention (please see Editor's note appended to this letter). If you are able to address the remaining issues to your best ability and revise the manuscript along the lines suggested, which may involve further experiments, I will be happy receive another 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' concerns.

Please 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 paper offers new and interesting data concerning the role of IRX factors in sex determination using a variety of approaches.

Comments for the author

The authors have responded to my comments and have done so to my satisfaction. They have worked hard to nuance their interpretations and the manuscript has been improved by acknowledging some of the drawbacks and difficulties in their data (mostly based on technical issues). They argue convincingly that generation of additional data, based on refined ChIP experiments, whilst desirable, are unrealistic given the time-scales involved when working on a small developing organ.

Reviewer 2

Advance summary and potential significance to field

The authors have satisfactorily addressed my comments and concerns from the prior submission. The accessibility of the manuscript is improved with Figure 1A, however the reader appreciation would be increased in figure 5 if a schematic were included to explain the ex vivo transfection experiment.

Comments for the author

The authors may consider Cut and Run as a low input alternative to ChIP in their future studies of gonad somatic cells.

Reviewer 3

Advance summary and potential significance to field

As per original review

Comments for the author

Reviewer 3 Comments for the author
The study investigates the regulation of Irx3 and 5 by WNT-β-catenin signaling and examines the epigenetic state across regions in which enhancer elements were identified in these genes. While the authors provide good evidence that canonical β-catenin regulates Irx3 / 5 in somatic cells of the ovary (Figure 1), the analysis of the potential enhancer elements through which this activity may be mediated is less convincing (Figures 3-6). This may in part be due to analysis of transfections in whole ovaries rather than in the relevant target somatic cells (developing granulosa cells). This is a difficult task though, and the challenge of getting appropriate construct activity in these cells may hamper the collection of more convincing data. Perhaps this could be overcome by making constructs in vivo using crispR or a similar approach, although this is a significant challenge. The genomic region was analysed in this study was large, but the detail of epigenetic modifications in the specific enhancer regions tested was limited. The study would significantly benefit from more detailed analyses demonstrating the epigenetic state and the activity of +86kb and -580kb enhancer elements identified.

Specific comments:
1. Figure 1A: The iCRT14 drug doses used seem very high. Either this drug is poorly available to the cells or not very potent. It has an IC50 of 40nM in assays for β-catenin responsive transcription, but in BT549 cells used at 10, 20 50uM to inhibit proliferation.
Compared to the IC50, the dose used makes me wonder whether the effects in cells are specific. While the data are consistent with the Sf1cre model, arguing in favour that the findings using the drug are reliable, the drug used here does not appear optimal.

Thank you for this interesting insight. We based our experimental design on the publication that first described the discovery of the iCRT drugs (Gonsalves, Klein, Carson, Katz, Ekas Evans,

Nagourney, Cardozo, Brown, DasGupta PNAS April 12, 2011 108 (15) 5954-5963; https://doi.org/10.1073/pnas.1017496108). iCRT14 was one of few that inhibited the transcriptional properties of β-catenin. We were convinced by their data including cell culture experiments that used 25μM and 50μM for the purpose of showing transcription inhibition. It may be that differences in responsiveness can also be attributed to differences in drug absorption in monolayer cell culture compared to explant tissue culture. Ultimately, we felt comfortable with our data based on our findings that included both positive and negative controls along with its agreement with the in vivo experiments. Similar to cell culture experiments in the Gonsalves, et al. paper, we observed a dose responsive repression of genes, which also suggests selective pharmacological activity of the drug. In sum, to address rigor in the experiments that are outlined in Figure 1, we evaluated loss- and gain-of-function impacts using both ex vivo and in vivo approaches. Altogether, these suggest a link between canonical β-catenin transcriptional activity and Irx3 and Irx5 and gave us the premise to proceed with investigations that would provide a direct regulatory link.

Response 1 Reviewer comment: ok.

2. An ANOVA analysis would seem appropriate in Fig 1A, but a t-test was used.

Thank you for pointing out this error, we have corrected and reported the new statistical test results in new Figure 1B.

Response 2 Reviewer comment:: The authors have addressed this concern

3. In Figure 1 the authors state that a 0.23-fold change was obtained for Axin2, but the level is actually 23% of control - ie a 75% or 4-fold reduction. This also applies to other genes. Presenting the data as a 0.23-fold change is not accurate. It would be better to represent the fold change for reduced expression as a % decrease. This would actually favor the authors argument that there is a strong reduction in expression.

We agree with this interpretation and have therefore altered our text to reflect a percent decrease in activity. Please see new text on page 6, starting Line 114.

Response 3 Reviewer comment: The authors have addressed this concern

4. Figure 1C: 50mM LiCl seems a relatively blunt tool for manipulation of β-catenin activity. Is no better drug available? Why was DMSO used as a vehicle – LiCl is solubilized in water or PBS? Why was no dose curve used for LiCl?

We agree that LiCl is a relatively 'blunt' tool; however, it is also recognized as a long-established tool to evaluate canonical β-catenin activity. We based our ex vivo treatment strategy on literature that reported dose ranges between 20-50mM (please see References 1-5 below) and our design was based on one particular publication that used the same ex vivo paradigm (Reference 6). In addition, to maintain rigor in our experimental interrogation, we evaluated the impact of stabilized β-catenin activity using a genetic mouse strategy (SF1Cre;Ctnnb1Δex3) to complement the ex vivo culture experiment.

Thank you for pointing out our error in reporting the LiCl control as DMSO. We had mistakenly used this control as it was the appropriate control for the iCRT14 drugs. To correct this oversight, we repeated experiments using water as the appropriate control for LiCl and observed similar results. These new data are now included in new Figure 1D.

References:
1. Davis, Weakland, Wieland, Farese, West. 1987 Proc. Natl Acad Sci 84:3728-3732
2. Tanwar, Kaneko-Tarui, Zhang, Rani, Taketo, Teixeira. 2010 Biol Reprod 82:422-432 https://doi.org/10.1095/biolreprod.109.079335
3. Usongo, Li, Farookhi. 2012 Dev Dyn 242:291-300. https://doi.org/10.1002/dvdy.23919
4. Bai, Chang, Cheng, Chu, Leung, Yang. 2017 Endocrinology 158:2813-2825. https://doi.org/10.1210/en.2017-00287

5. Bernard, Ryan, Sim Czech, Sinclair, Koopman, Harley. 2012 Endocrinology 153:901-912.
6. Maatouk, DiNapoli, Alvers, Parker, Taketo, Capel. 2008 Hum Mol Genet 17:2949-2955.

Response 4 Reviewer comment: This is a common response - ie compound A has been used before. However, the use of LiCl by other groups does not mean that LiCl is the best or most specific compound to use for this pathway – it is probably more indicative that subsequent studies have followed suit and other (more specific and more potent) agonists have not been investigated in the field. More sophisticated and specific compounds are available and less likely to have off target effects. Notwithstanding this reservation, the in vivo analysis supports the LiCl analysis so it seems appropriate.

5. Figure 3: These data are indicative, but not particularly compelling especially considering a number of differences are measured using a p value of 0.1 and students t test measuring several regions together (ANOVA + post hoc may be more appropriate). Perhaps cleaner and more compelling data would be obtained from purified somatic cells rather than whole ovaries.

All three reviewers commented on this particular Figure (see responses to Reviewer 1 comment 1 and Reviewer 2, comment 2). We have added text to the results section to explain the variability in data and have eliminated the confusing language regarding trend vs significance. We acknowledge that the homogeneous population within whole gonad ChIP contributes to data variability and agree that we would likely achieve cleaner data from purified cells; however, the approach using isolated somatic cells took over a year to obtain sufficient cells numbers for the genomic-Seq data. Even the whole gonad ChIP requires accumulation of 100-150 pairs of gonads for each experiment, which takes approximately 3-4 months to acquire. Therefore, we consider somatic cell ChIP to be beyond the scope of this manuscript at this time. To address the variability in the results, we updated the Results section (see page 8, starting Line 189).

Response 5 Reviewer comment: The data here is important for the rest of the study as it defines the +86 and -580 elements as important. However, the data do not appear to be not particularly robust – for example in A (XX) there appears to be no difference in the data for -305 and -580 for TCF/Lef despite that observation that one yielded a significant difference but the other did not. This is also the case for the H3K27ac for +86 and -305. Similar criticisms can be made of other comparisons in the figure. Considering this concern was raised by all reviewers, I am not convinced that the response sufficiently addresses the reviewer's concerns. While it is appreciated that sample size is a substantial impediment robust ChIP approaches using 100,000+ cells are now routine. In female mice, at least 50,000 somatic cells can be retrieved from a gonad pair at E14.5. 100 gonad pairs would yield >5 million cells. Figure 4 is more convincing than Figures 3 and 5, but alone is limited.

6. Figure 5: There is no data to show how efficient the transfection was in this experiment.
How did the authors measure transfection efficiency and validate that the experiment has worked?
It is surprising that there were no sex specific differences detected here. The full promoter did not appear to confer the same activity as the artificial construct shown in 5B.
What is the explanation for this?

The details for transfection assays were not included in this manuscript. We did however include language within the text to point readers to a previous paper from our laboratory where we described this technique in detail (Gao L., et al. 2011, Biol Reprod 84: 422). We acknowledge that transfection efficiency is typically not achieved to the same level as would be expected in monolayer cell culture; however, there are no cell lines that represent fetal gonad somatic cells available at this time and we developed this technique to provide an instrument that would allow acquisition of more physiologically relevant significance. To address this potential issue, we employ rigorous means to normalize data, first, gonads are cotransfected with SV40-Renilla, which is used as an internal transfection efficiency control and then second, each experiment is normalized to gonads that were injected with the promoterless control (pBasic/SV40-Renilla).

We were also surprised that our data indicated no sexually dimorphic differences in activity.
We interpreted this data to indicate that there was additional information required that was not present in our plasmid DNA constructs. In addition, it is recognized that plasmid DNA is devoid of epigenetic decorations, which would allow potentially unrestricted access to all sequences in ovary

and testis. To address this, we evaluated our H3K27me3 ChIP-Seq data and found that there was substantial enrichment on promoter sequences in XY, but not XX cells. This is shown in new Supplementary Figure S5. This was also commented upon by Reviewer 1 (please see response to their comment # 2). We included new text within the Discussion section to address these important caveats (see page 14, starting line 353).

Thank you for your question regarding the difference in activity for the full-length promoter and the artificial construct that contains enhancer + full-length promoter in 5B. We have also wondered about this. The obvious difference is the addition of the enhancers; therefore, we searched those sequences for potential binding sites of factors that may inhibit activity in gonads of either sex. One particularly intriguing possibility was the identification of several SOX9 binding sites. Studies to further evaluate this possibility are underway.

Reviewer comment for points 6-11 is below (after point 11)

7. Figure 5B: Why are the levels in the testis higher than shown in the ovary. Why do the mutations in the reporter not reduce activity in testis?

This was a common concern among reviewers, and we were also surprised by this finding. Statistical analyses did not indicate any difference between testis and ovary; however, one reason that the testis results might generally be higher is that they are quite a bit easier to inject due to their larger size. We normalized all transfections to SV40 Renilla activity to control for variations in transfection efficiency. In general, luciferase and renilla values were higher in testes, but their ratios—and further normalization to the pBasic luciferase control gonads allowed for direct comparisons.

The data showing ovary-specific sensitivity to the mutations in the enhancers was very exciting to us. There are two mutations, one in each enhancer. Notably, the mutations are the same—a single base pair mutation that changes the TCF/LEF binding site from canonical to non-binding. This is the exact mutation that differentiates the widely used TOP-flash vs FOP-flash reporters. Thus, we interpreted this finding to show exquisite sensitivity to β-catenin/TCF/LEF DNA binding and therefore, support our hypothesis that canonical β-catenin directly regulates the Irx3/5 locus, which can only be achieved within the ovary where β-catenin transcriptional regulation occurs. Regarding testis expression, we recognize that the plasmid DNA lacks epigenetic marks, thus making all sequences accessible to male-specific factors. For example, SOX factors can bind the same sequences as TCF/LEF and there are likely others that are exposed. These factors are similar to the responses detailed above (Reviewer 1, comment #2, and the response immediately above, #6).

8. Figure 5: Does this include data from all three experiments combined, or one example (ie one experiment)?

These data include all experiments combined. To clarify, we report the n for each experiment.

9. The epigenetic differences shown across the region (Figure 2 and Figure 6) are very interesting, but not very detailed and are not validated. As the regions identified are specifically tested in promoter activity assays, validation of their epigenetic state at a more detailed level (ie less than 750kb region) is desirable.

We agree that the data shown for ATAC-Seq (in addition to DNAse-Seq, data not shown) were not very detailed and indicated only that chromatin in specific regions was accessible. We included new information regarding ATAC-Seq and H3K27me3 ChIP-Seq data for the Irx3 and Irx5 proximal promoters in XX vs XY cells in new Supplementary Figure S5. Of note, there is abundant enrichment of the repressor mark on both promoters only in XY cells. As pointed out, we did clone specific regions into reporter constructs. To evaluate the ATAC-Seq data, we used an antibody against the active mark, H3K27Ac, for each region using ChIP-PCR. The primers were designed to focus specifically to the region of the canonical β-catenin/TCF/LEF binding site with amplicons no larger than 100bp. We did not repeat the H3K27Me3 ChIP-Seq data as we considered that additional experiments using whole gonads would not add significant new knowledge to an already robust

evaluation using isolated cells of interest. As suggested, we incorporated new insets of our ChIP-Seq data into Figure 6 and focused on ~5kb of +86kb and -580kb enhancer sites in addition to Irx3 and Irx5 promoter sequences mentioned above.

10. Figure 6: No detail is shown for the H3K27me3 enrichment at the loci investigated – just a line in the figure showing regions that are reportedly enriched for H3K27me3 based on data from a separate study. No data is shown for H3K27me3 reductions either at a whole cell level or at specific loci in the TAZ treated samples. The experiment should at least be supported by analysing H3K27me3 levels in the TAZ treated vs untreated cells and preferably across the regions of interest in the samples tested.

This is an interesting point and is similar to comment 1 from Reviewer 2. We considered several avenues by which we might try to explore this idea. We believe that the best approach would be to isolate gonads from XY SF1Cre;Ctnnb1Δex3 control and mutant embryos for ChIP-PCR using H3K27Me3 and other epigenetic marks. Unfortunately, current techniques require 100-150 pairs of gonads to perform a single ChIP (for up to 4 antibodies/ChIP, including RNA Pol II and IgG controls), and we are unable to collect the appropriate genotype gonads to perform this experiment within a reasonable timeframe. This paradigm, however, is under serious consideration for future ChIP-Seq assays.

11. Figure 6: While there was a significant increase in TAZ/DMSO in the XY samples for Irx3 and Irx5, this response was far below what was observed in Fig 1D for these genes in XX samples. In addition, did not appear to be any differences in the level of transcriptional activation for Irx3, Irx5 for TAZ/DMSO between XY and XX. But the response for p21 was greater in XX than in XY. This is confusing.

Thank you for this important question. While it is tempting to compare expression levels between treatment paradigms (Figure 1 vs Figure 6), we believe each treatment paradigm tests very different questions and each drug will have unique effects. To address the specific comparison in Figure 1D; these data are from XY testes from mice carrying a stabilized form of β-catenin. This model is very specific to a transcription factor and is expected to be extremely robust based on its phenotype (Maatouk DM et al. 2008 Hum Mol Genet 17:2949). The goal for the TAZ-treated gonads was to inhibit a single DNA methylation pathway related to H3K27Me3. We recognize that this experiment has limitations as it interferes with a pathway that likely impacts other factors. A statement to this has been included in the discussion section (see page 15, starting Line 375).

As noted, the transcript levels for Irx3 and Irx5 are not different between treated ovaries and testes, but the response for p21 is quite different. Our treatment time course (E11.5 onset and 48 hours culture) falls during the most profound period of sex differentiation; thus, we expect sex-specific transcriptomes that will respond in different ways to drug treatment. We rephrased our results section to clarify how we interpreted the data (page 12, starting Line 281).

Response 6-11 Reviewer comment: The points above have only been partially addressed.
However, I still think that this data is variable in quality and could be strengthened. It would be preferable to validate the transfection efficiency in this study rather than to rely on what has happened in previous studies. The explanation that the vector does not have H3K27me3 does not really address this issue. It appears that despite the lack of H3K27me3 on the vector in XX and XY, the expression in XY is higher than in XX and the difference generated by deleting the +86 and -580 elements is modest. The data shown in tracks for looks interesting, but as I understand was not validated in this study. Figure 6D demonstrated a modest, but significant difference in XX for Irx3 and 5 over DMSO in response to EZH2 inhibition, but this difference is ~1.7 fold in XY and 2.4 fold in XX - ie the difference in response is less than 1 fold between the sexes over a general increase of ~2 fold. Moreover, the difference in response between the sexes was not significant – ie there was not a significantly greater response in XX than in XY.
This does not seem to be a particularly robust and would benefit from further analysis. A direct analysis of epigenetic state in conjunction with the qRTPCR analysis rather than correlating data from past studies with the qRTPCR data (Fig 6D).

12. The description of the changes for Irx3 and Irx5 in Lines 275-278 is not very clear. I am not sure why the transcriptional response for Irx3 and Irx5 should be directly comparable to the positive control p21? Surely this depends on the normal observed transcriptional increase for Irx3 and Irx5 in ovaries vs testes compared to what is observed for p21. From Figure 1, the XX vs XY difference for Irx3 and Irx5 appears to be around 15-20 fold in response to b-catenin activation using the Cre model. I appreciate that TAZ may not result in the same response as b-catenin levels are different in XX vs XY. But it is not clear how much response is really generated by TAZ treatment in XY gonads compared to what can be expected in XX gonads.
This data could be presented two ways 1. As a TAZ/DMSO ratio as provided. 2. As a comparison of Irx3 and Irx5 between untreated XX gonads and TAZ treated XY gonads to show how much of the normal XX levels are reached in TAZ treated XY gonads.

We appreciate the importance of this question and evaluated different means of reporting the results. In the end, we opted to go with the original report (option 1) because we believe that it is difficult to compare to untreated samples and across sexes, especially at this particular time in development (see response for comment 11 above). We included additional text to increase clarity in describing the experiment and its interpretation (see page 12, starting Line 281).

Response 12 reviewer comment: I think a greater response would be expected in XX than in XY gonads (ie option 2 is important). This was not evident. This issue has yet to be convincingly addressed.

Minor comments:
1. Fig 2B does not add anything new to 2A. If it is to be included, it would be better drawn to match Fig2A with the colored parts shown exactly under their equivalent parts in Fig 2A

Thank you for this helpful suggestion. We included the chromatin map within Figure 2A (now just figure 2) as suggested.

Figure 2A is now easier to understand

Second revision

Author response to reviewers' comments

MS# DEVELOP/2019/183814

Thank you again for your consideration with our manuscript, we truly value this process and strongly believe that our reviewers have helped us to improve the presentation of our complicated, but important findings. In response to reviewer comments, we are excited to present new data and text that provide considerable substance towards our story. We sincerely hope that you agree. Page/line references match the manuscript that contains all track changes. New text is in red, omitted text is with ~~strikethrough~~ font.

Reviewer 1 Comments for the Author:
The authors have responded to my comments and have done so to my satisfaction. They have worked hard to nuance their interpretations and the manuscript has been improved by acknowledging some of the drawbacks and difficulties in their data (mostly based on technical issues). They argue convincingly that generation of additional data, based on refined ChIP experiments, whilst desirable, are unrealistic given the time-scales involved when working on a small developing organ.

Thank you for these encouraging comments.

Reviewer 2 Advance Summary and Potential Significance to Field:
The authors have satisfactorily addressed my comments and concerns from the prior submission. The accessibility of the manuscript is improved with Figure 1A, however the reader appreciation would be increased in figure 5 if a schematic were included to explain the ex vivo transfection experiment.

We added a schematic of the transfection experimental paradigm to Figure 5 as suggested (Figure 5A).

Reviewer 2 Comments for the Author:
The authors may consider Cut and Run as a low input alternative to ChIP in their future studies of gonad somatic cells.

We appreciate the alternative method suggestion and recognize that this approach is not trivial. Success would require FACS for somatic cells from genetically modified animals that would facilitate identification of this specific cell population. We also have colleagues (at NIEHS) that have been troubleshooting the Cut and Run technique for over a year using isolated cells from mice and have not yet obtained publishable results. We do intend to pursue methods for low input alternatives for future experiments as suggested. Of note, we did use a low-input protocol for all of our genomic-Seq experiments: Multiplexed Indexed T7 ChIP- seq (MINT-ChIP) by Van Galen et al., 2015 *Molecular Cell*. This method has been validated for as low as 500 cells. In our studies, MINT-ChIP was performed on 30-150K FAC-purified gonadal cells at two developmental time points. We have included this new information in our materials and methods section.

Reviewer 3 Comments for the Author:

Note: all of these statements were originally presented in one paragraph. Here we separated each section so that we could address each comment directly with the goal to increase clarity.

Response 6-11 Reviewer comment: The points above have only been partially addressed. However, I still think that this data is variable in quality and could be strengthened. It would be preferable to validate the transfection efficiency in this study rather than to rely on what has happened in previous studies.

As noted in our previous response, we acknowledge that there will be transfection variability from gonad to gonad. This reality is not different from cell line transfection analysis. We noted that we control for transfection variability by the co-transfection of SV40 Renilla in addition to all Irx3-promoter-luciferase constructs (pGL3). Transfection data is reported as a fold change of Luciferase/Renilla ratios that are also normalized to the no-promoter pGL3 basic control for each sex gonad. This is standard practice for transfection analyses.

We have published this technique twice (Gao L., et al. 2011, Biol Reprod 84: 422 and Wainwright E.N. et al. 2013, Biol Reprod 89: 34) and have used it for another manuscript that is under review (see adjacent figure). We have also routinely used this technique to interrogate specific binding elements for sexually dimorphic regulation for countless genes of interest within our own and in collaboration with other laboratories. Our first publication goes into tremendous detail to show that the technique is robust and includes evaluation of reporter expression by whole mount analysis along with section double label IHC to validate reporter and cell type specificity (Gao L., et al. 2011, Biol Reprod 84: 422). While the promoter of interest is not the Irx3 promoter, the premise is the same.

To illustrate transfection efficiency, we present copy number qPCR results from a different manuscript that is currently under review. [Details provided confidentially to referees have been removed as the data in question are included in a manuscript under review elsewhere.]

The explanation that the vector does not have H3K27me3 does not really address this issue. It appears that despite the lack of H3K27me3 on the vector in XX and XY, the expression in XY is higher than in XX and the difference generated by deleting the +86 and -580 elements is modest.

After much consideration and discussion, we respectfully disagree. It is possible that we do not fully

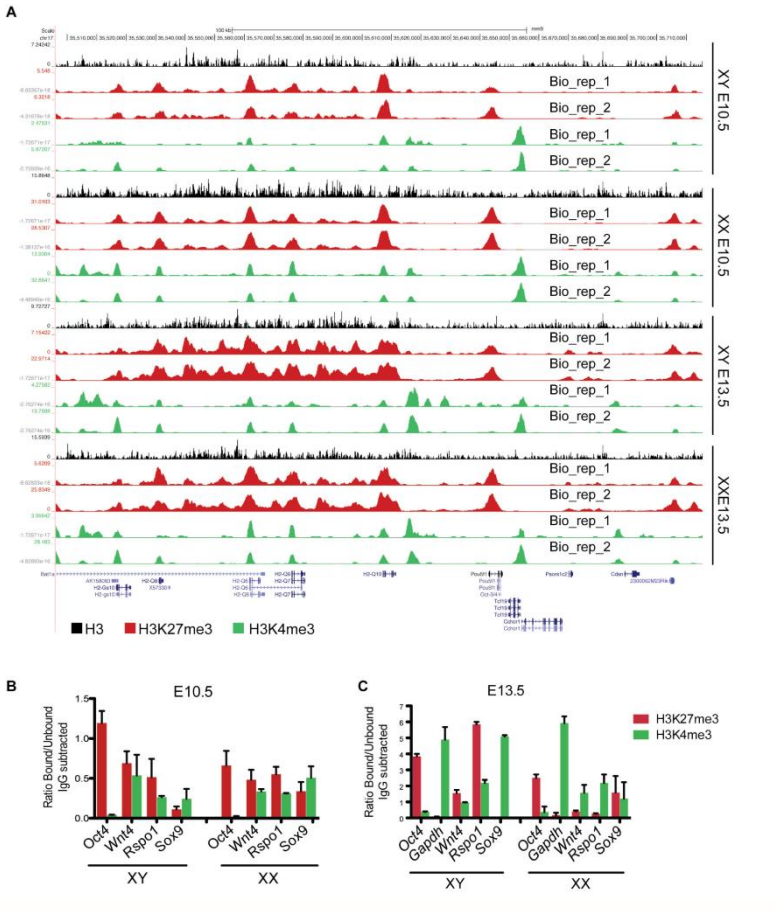
understand the intent of this comment. H3K27me3 is a repressor mark. More importantly, plasmid DNA is completely void of histones, histone markers or even DNA methylation. In other words, there is absolutely no epigenetic impact on plasmid DNA regulation. This configuration allows access to ALL sequences and thus, any available transcription factors will be allowed access. Notably, ▫catenin/TCF/LEF factors and Sox factors have been shown to bind the same sequences. Other sequences will also be available in both gonads that may otherwise be unavailable depending on sex. Thus, it is not surprising that there is no difference between male and female data. We consider that the most important information is that SINGLE POINT MUTATIONS within the TCF/LEF binding sites causes ~60% decrease only in ovaries. We interpret this as a powerful visualization that these enhancers are indeed direct targets for ▫catenin/TCF/LEF transcriptional regulation.

The data shown in tracks for looks interesting, but as I understand was not validated in this study. Figure 6D demonstrated a modest, but significant difference in XX for Irx3 and 5 over DMSO in response to EZH2 inhibition, but this difference is ~1.7 fold in XY and 2.4 fold in XX - ie the difference in response is less than 1 fold between the sexes over a general increase of ~2 fold. Moreover, the difference in response between the sexes was not significant - ie there was not a significantly greater response in XX than in XY. This does not seem to be a particularly robust and would benefit from further analysis. A direct analysis of epigenetic state in conjunction with the qRTPCR analysis rather than correlating data from past studies with the qRTPCR data (Fig 6D).

Here we made significant changes and completely replaced Figure 6. We appreciate the reviewer's point regarding the results from the EZH2 inhibition studies. After careful consideration and evaluation of the biology of ovary versus testis development at this time point, we recognized that it is almost impossible to evaluate the impact of EZH2 in specific cell populations at this time. Most important is the differential in cell cycle progression between the testis and ovary at this time in development. The testis undergoes profound growth via cell proliferation starting around E11 and is one of the first and most important sexually dimorphic distinctions during sex determination (2003 Schmahl et al. Dev Biol 258:264). This difference in cell cycle progression will have an impact on epigenetic markers that is also likely to incorporate significant individual cell-to-cell variability. Therefore, we recognized that our EZH2 inhibition studies are unlikely to provide a clear picture and felt the best course was to remove the graph altogether (Figure 6D).

With respect to data validation, we present to our reviewers Supplemental Fig 1 from our previously published paper (see below), which includes validation PCR for the ChIP-Seq experiments (see figure below, Garcia-Moreno 2019 PLoS Genetics 15(5):e1007895.) The ChIP-seq was validated by ChIP-qPCR on 3 biological replicates, each replicate consisted of pooled FACS-purified cells from multiple embryos; the same protocol used to obtain samples for ChIP-seq. Promoters from genes known to drive sex determination and differentiation were chosen for validation. We referenced these data within the results sections, see Page 11, starting LINE 264.

Finally, we constructed a new Figure 6, which is designed to reinforce the H3K27me3 data, as suggested by reviewers 2 and 3. In Figure 6A, we now show data from 4 independent biological replicates (2 additional replicates that have not been previously published) of the H3K27me3-seq data for both male and female somatic cell populations at E13.5 for the enhancers within the Irx3/5 locus. Each biological replicate was performed on 30-150k FAC-purified cells pooled from gonads of multiple embryos. Of note, the ChIP-seq protocol we used (MINT-ChIP by van Galen et al., 2015 *Molecular Cell*) has been validated on as low as 500 cells. These new data are in the process of being deposited into the NCBI GEO database. We also include data to show the time-dependent progression of H3K27me3 + ATAC-Seq events at each enhancer site between E10.5 (pre-sex determination) and E13.5 (post sex-determination) across the *Irx3/5* locus and a magnified view of each locus. We believe that 4 independent biological replicates show that our data are robust and repeatable. These new data are described in the results and discussion sections, see Page 10, starting LINE 248.



Supplemental Figure 1 from Garcia-Moreno et al., 2019 *PLoS Genetics*. (A) ChIP-seq tracks for H3 (black), H3K27me3 (red) and H3K4me3 (green) for both biological replicates are shown side by side in XY and XX, E10.5 and E13.5 purified supporting cells. (B&C) ChIP-qPCR validation of ChIP-seq for H3K27me3 (red) and H3K4me3 (green) in FACS-purified E10.5 XY and XX cells (B) and E13.5 XY and XX (C). Each ChIP-qPCR was performed on 3 biological replicates, each replicate contained pooled cells from several gonads. ChIP-seq tracks for depicted genes are in Figs 2 & 3. Values represent mean ± SEM.

Response 12 reviewer comment: I think a greater response would be expected in XX than in XY gonads (ie option 2 is important- As a comparison of Irx3 and Irx5 between untreated XX gonads and TAZ treated XY gonads to show how much of the normal XX levels are reached in TAZ treated XY gonads.). This was not evident. This issue has yet to be convincingly addressed.

We have removed these data (see comment above).

Editor's note:
-Consider performing a Low-(cell number) input profiling experiment as an alternative to ChIP seq analysis to substantiate the results on the gonadal somatic cells (Review 2 and Review 3 comment on your response #5)

Please see response to Reviewer #2 above.

-Validity of the results of profiling chromatin accessibility and H3K27me3 marks (R3 comment to your response #6-11)

Please see response to Reviewer #3 above.

Third decision letter

MS ID#: DEVELOP/2019/183814

MS TITLE: Canonical Wnt/β-catenin Activity and Differential Epigenetic Marks Direct Sexually Dimorphic Regulation of Irx3 and Irx5 in Developing Gonads

AUTHORS: MEGAN L KOTH, S. Alexandra Garcia-Moreno, Annie Novak, Kirsten A Holthusen, Anbarasi Kothandapani, Keer Jiang, Makoto Taketo, Barbara Nicol, Hung-Chang Humphrey Yao, Christopher R Futtner, Danielle M Maatouk, and Joan S Jorgensen
ARTICLE TYPE: Research Article

I am satisfied with your response to the review and the revision. I am happy to tell you that your manuscript is accepted for publication in Development, pending our standard ethics check.