EMBO reports

Mouse GTSF1 is an essential factor for secondary piRNA biogenesis

Takuji Yoshimura, Toshiaki Watanabe, Satomi Kuramochi-Miyagawa, Noriaki Takemoto, Yusuke Shiromoto, Akihiko Kudo, Masami Kanai-Azuma, Fumi Tashiro, Satsuki Miyazaki, Ami Katanaya, Shinichiro Chuma, Jun-ichi Miyazaki

Review timeline:

Submission date:	20 January 2016
Editorial Decision:	1 February 2016
Revision received:	17 November 2017
Editorial Decision:	21 December 2017
Revision received:	11 November 2017
Accepted:	16 January 2018

Editor: Esther Schnapp

Transaction Report: This manuscript was transferred to *EMBO reports* following review at *The EMBO Journal*

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 1 February 2016

Thank you for the transfer of your manuscript with referee reports to EMBO reports. I have sent it and the reports to an advisor, and given the set of comments we have (all pasted below), we decided that we can offer to publish a revised manuscript that successfully addresses all referee and the advisor's concerns.

The advisor notes that the main concern is the single replicate of the deep sequencing data. S/he also agrees with referee 1 that IP/MS needs to be performed in testis extract, that the 16-35 ntRNA needs to be deep sequenced, that piRNA sequence composition by position needs to be globally analyzed, and that Z-scores for piRNA ping-pong need to be included. All minor concerns also need to be addressed.

Given these constructive comments, we would like to invite you to revise your manuscript with the understanding that the referee concerns must be fully addressed and their suggestions taken on board. Please address all referee concerns in a complete point-by-point response. Acceptance of the manuscript will depend on a positive outcome of a second round of review. It is EMBO reports policy to allow a single round of revision only and acceptance or rejection of the manuscript will therefore depend on the completeness of your responses included in the next, final version of the manuscript.

Revised manuscripts should be submitted within three months of a request for revision; they will otherwise be treated as new submissions. Please contact us if a 3-months time frame is not sufficient for the revisions so that we can discuss this further. Given the 10 main figures, we would publish your study as a full article, for which there are no length limitations. Please note that the EMBO reports reference style is numbered (this needs to be changed) and that supplementary figures and tables are called expanded view figure/table now (EV1, 2, etc). Please add the figure legends for EV figures to the end of the main manuscript file and upload the EV figures separately upon submission

of your revised manuscript. The advance of our expanded view content is that it is integrated into the main text online and expands when clicked.

Regarding data quantification, please specify the number "n" for how many experiments were performed, the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values in the respective figure legends. This information is currently incomplete and must be provided in the figure legends. Please also include scale bars in all microscopy images.

We now strongly encourage the publication of original source data with the aim of making primary data more accessible and transparent to the reader. The source data will be published in a separate source data file online along with the accepted manuscript and will be linked to the relevant figure. If you would like to use this opportunity, please submit the source data (for example scans of entire gels or blots, data points of graphs in an excel sheet, additional images, etc.) of your key experiments together with the revised manuscript. Please include size markers for scans of entire gels, label the scans with figure and panel number, and send one PDF file per figure or per figure panel.

We would also welcome the submission of cover suggestions, or motifs to be used by our Graphics Illustrator in designing a cover.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File to accompany accepted manuscripts. This File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

You are able to opt out of this by letting the editorial office know (emboreports@embo.org). If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

I look forward to seeing a revised version of your manuscript when it is ready. Please let me know if you have questions or comments regarding the revision.

REFEREE REPORTS

Referee 1:

The authors previously described mouse Gtsf-1 as a factor required for transposon silencing (Yoshimura et al., 2009). In that study mouse Gtsf-1 was shown to be a cytoplasmic protein. Recently, we and others showed that fly Gtsf-1 is a nuclear protein, interacting with the nuclear Piwi in fly ovaries. The fly study implicated Gtsf-1 as an effector protein for nuclear silencing by Piwi, but not as a piRNA biogenesis factor.

In this manuscript, the authors show that mouse Gtsf-1 is both cytosolic and nuclear in embryonic germ cells. Using recombinant GST-fusion proteins they show that Gtsf-1 interacts with a number of proteins in mouse testes extracts, including all the three mouse Piwi proteins. The most interesting aspect of this work is the demonstration that nuclear Miwi2 is not loaded with piRNAs in the Gtsf-1 KO mouse.

So in addition to a number of other factors, Gtsf-1 is essential for loading Miwi2. This implicates Gtsf-1 as a piRNA biogenesis factor and not as a nuclear effector. So fly and mouse Gtsf-1 proteins function differently, how? is not clear. This study will be of great importance to researchers in the piRNA field.

Comments
1) Can the authors speculate on how Gtsf-1 might work? The interaction data, the cytoplasmic localization etc, how does that help us understand how the protein might facilitate Miwi2 loading?
2) Some discussion on how the same protein has different function in flies and mice will be useful.
3) The in vitro GST-Gtsf1 and piRNA pathway factors interactions using mouse testes lysates (Fig.

6): A mass spectrometry analysis of the pull downs, if available, should be provided. One problem with such experiments is that incubation of any GST-fusion protein with cell lysates can bring down almost any protein one is searching for. The GST control is always clean in such experiments.
4) The quality of images can be improved. It is currently hard to see anything.

Referee 2:

Yoshimura et al follow up on their earlier studies of GTSF1 protein, which is known to be required for spermatogenesis and transposon repression in male germ cells of mice. Here, they evaluate localization of GTSF1 with respect to murine PIWI proteins (MILI and MIWI2); the requirement of GTSF1 for localization of PIWI and other proteins in the piRNA pathway, and for the production of piRNAs. The authors observe that GTSF1 frequently colocalizes with with MIWI2 and TDRD9 but only occasionally with MILI. GTSF1 also accumulates in the nuclei of fetal prospermatogonia. Pull down experiments using GST fusion proteins suggest that GTSF1 interacts with PIWI complexes in adult and embryonic testes and W107 residue could play an important role in these interactions. GTSF1 interaction with MIWI2 appear to be largely due to an RNA component of the complex. However, only Tdrd9 showed direct interaction with GTSF1 following protein co-expression in cultured cells. Consistent with defects in MIWI2 localization to cytoplasmic granules and the nucleus, no piRNAs co-immunoprecipitate with the protein and piRNA amplification is abolished.

Overall, this is a carefully executed study and the data will be of interest to the piRNA field. The new data complement results of prior studies on other components of the piRNA pathway in mice as well as several paper that examined GTSF1 homolog in the fruit fly. The only open question is whether the paper offers sufficiently novel insights into piRNA biogenesis at the level anticipated at this particular journal but I would leave this up the editors.

Minor comments:

- MIWI2 and TDRD9 cytoplasmic granules are not just processing bodies (P-bodies). They share some components with somatic cells' P-bodies but are distinct from those both by the presence of numerous piRNA pathway proteins and their appearance (See and cite Aravin et al 2009). This point needs to be clarified.

- Co-localization of GSTF1 to MIWI2 and MILI cytoplasmic bodies should be quantified - this information would be valuable for subsequent studies dissecting piRNA pathway proteins' functions.

Advisor's comments:

This manuscript provides compelling evidence that GSTF1 acts in piRNA
biogenesis in mice, unlike its role in carrying out silencing in flies. I suspect that the experiments presented here will prompt a reexamination of GSTf1 in flies.
There are quite a few problems that need to be fixed in a revised manuscript-
most importantly, replicates of the deep sequencing data. While the
bioinformatics is generally below the standard in the field, I suspect that the
authors have gotten very little if anything wrong. More likely, they have simply
missed insights that a more complete analysis would have revealed.

Specific Comments
(1) Figure 2a: a two-tailed, unpaired t-test or other appropriate statistical test needs to be used to determine whether these data are significant. The effect on UTR sequences is much larger than on the L1 ORF2 or IAP Gag ORF sequence. Why?
(2) What are the intense, L1 ORF1p-containing foci in Gstf1−/−
prospermatogonia?
(3) Figure 4b: A statistical test is required to determine whether the intermitochondrial distance is longer in Gstf1−/− than in Gstf1+/−.
(4) I agree with Reviewer 1 that GST-pull downs provide weak evidence at best.
Immunoprecipitation/mass spectrometry needs to be performed in testis extract.
This is particularly important given the discordant results from the experiments

using GST-pull down and those using ectopic expression in cultured somatic cells.
(5)The small RNA sequencing data appears to be from a single biological replicate. The experiment needs to be repeated so that they authors have at least two biological replicates; three would be optimal.
(6) Figure 7c: The authors appear not to have sequenced the MILI- and MIWI2-bound piRNA populations. Instead they analyze them solely by end labeling. Unfortunately these data is quite unconvincing. It appears that full-length tRNAs are the main RNA recovered in the immunoprecipitates, even in the Gstf1 heterozygotes. The immunoprecipitated 16-35 nt RNA needs to be deep sequenced and fully analyzed as is standard for the field.
(7)Figure 8b: The author's analysis of piRNA sequence composition is helpful, but cannot substitute for a global analysis of piRNA sequence composition by position.
(8) Figure 8c: Z-scores for piRNA Ping-Pong need to be included. While it is useful to analyze Line1- and IAP-mapping piRNAs, Ping-Pong analysis for all genome-mapping piRNAs should also be presented. Moreover, the authors have missed a key finding indicated by their analysis of IAP piRNAs. The authors determine the 5′-5′ distance for IAP-mapping piRNAs on opposite genomic strands for GSTf1 heterozygous and homozygous. For heterozygotes, they observe the expected 10 nt peak indicative of Ping-Pong. In contrast, in E17.5 testes from the Gstf1 homozygous mutant testes, there is a prominent peak at 17 nt (Z-score is needed to determine its statistical significance). What is this peak?

Minor Points
(1) I don't agree with the minor point of Reviewer 2:
"MIWI2 and TDRD9 cytoplasmic granules are not just processing bodies (P-bodies). They share some components with somatic cells' P-bodies but are distinct from those both by the presence of numerous piRNA pathway proteins and their appearance (See and cite Aravin et al 2009). This point needs to be clarified."
Since 2009, the authors have used the term "processing bodies" to describe the perinuclear foci containing TDRD9 and MIWI2. Nowhere in the manuscript do they abbreviate processing body to "P body." The nomenclature is confusing, and indeed it confused Reviewer 2.

1st Revision - authors' response 17 November 2017

To referee 1:

Comment (1): *Can the authors speculate on how Gtsf-1 might work? The interaction data, the cytoplasmic localization etc, how does that help us understand how the protein might facilitate Miwi2 loading?*
Response (1):
Our study strongly suggested that mouse GTSF1 is involved in the MILI-piRNA guiding mechanism and thus is essential for secondary piRNA biogenesis (see Figs. 6B and 6C; p.14 line 5-17 in the revised manuscript). We propose that mouse GTSF1 is involved in the mechanism(s) for stabilizing and/or grasping the target RNA so that MILI-piRNA can slice it at a specific position (see Fig. 6D). We have added some discussion on this issue to the revised manuscript (p.15 line 3-8).

Comment (2): *Some discussion on how the same protein has different function in flies and mice will be useful.*
Response (2):
DmGTSF1 interacts with nuclear Piwi complex and coordinately functions in the transcriptional silencing of transposons (Donertas et al, 2013; Ohtani et al, 2013). Therefore, it has been predicted that mouse GTSF1 is an effector of transcriptional gene silencing, but is not crucial for ping-pong cycle (Xiol et al, 2014; Yu et al, 2015). DmGTSF1 is unlikely to be involved in the ping-pong cycle because i) DmGTSF1 is not localized to perinuclear nuages, in which the ping-pong cycle by Aub and Ago takes place, ii) the absence of DmGTSF1 has no impact on the

perinuclear localization of Aub and Ago3, and iii) lack of DmGTSF1 results in increased sense piRNA level, probably because an intact ping-pong cycle processes the accumulated sense transposon transcripts (Donertas et al, 2013). On the other hand, our study showed that mouse GTSF1 is probably involved in the MILI-piRNA guiding mechanism and thus is essential for secondary piRNA biogenesis. This evidence lets us to speculate that mouse GTSF1 is probably required for the MIWI2-piRNA silencing complex to stabilize and/or grasp the nascent RNA transcribed from the target genomic locus depending on the guide sequence of piRNA (Fig. 6D). We propose this model because it can account for not only the common molecular function of GTSF1 in both MILI-piRNA and MIWI2-piRNA complexes, but also the evolutionarily conserved function of GTSF1 in transcriptional retrotransposon silencing and therefore, is in line with the previous reports that DmGTSF1 functions with Piwi for the establishment of H3K9me3 at transposon loci (Donertas et al, 2013; Muerdter et al, 2013; Ohtani et al, 2013). Based on the above consideration, we have added some discussion on this point to our revised manuscript (p.15 line 9-22 and p.16 line 16-26).

Comment (3): *The in vitro GST-Gtsf1 and piRNA pathway factors interactions using mouse testes lysates (Fig. 6): A mass spectrometry analysis of the pull downs, if available, should be provided. One problem with such experiments is that incubation of any GST-fusion protein with cell lysates can bring down almost any protein one is searching for. The GST control is always clean in such experiments.*
Response (3):
We understand the Reviewer's and Advisor's concern whether the in vitro interactions between GTSF1 and complexes involved in piRNA pathway truly occur in vivo. We thus performed immunoprecipitation of testis lysates with anti-GTSF1 antibody instead of GST-pull down. The immunoprecipitates were subjected to MS analysis. As summarized in Table EV1, the result strongly supported that mouse GTSF1 binds to and is a component of PIWI complexes in vivo (p.9 line 21-25 in the revised manuscript).

Comment (4): *The quality of images can be improved. It is currently hard to see anything.*
Response (4):
We tried to produce improved figures from the original images by signal level adjustment. The image processing was equally applied across the entire image and also to controls. We are willing to disclose the source data, if requested. We believe that stained foci can be easily recognized in the revised figures (see Figs. 1A, 1B, 1D-1P, and EV1C).

To referee 2:

Comment (1): *MIWI2 and TDRD9 cytoplasmic granules are not just processing bodies (P-bodies). They share some components with somatic cells' P-bodies but are distinct from those both by the presence of numerous piRNA pathway proteins and their appearance (See and cite Aravin et al 2009). This point needs to be clarified.*
Response (1):
According to the Referee 2's comment (1) and Advisor's minor comment (1), we have changed a term 'processing body' to 'pi-P body' to avoid readers' possible confusion (Aravin et al, 2009) (p.4 line 7, p.6 line 12 and 17, p.7 line 6, 10, 12, and 13, p.8 line 2, 16, 19, and 20, p16 line 12, and 14, and p.33 line 2, 7, 15, 18, and 21 in the revised manuscript).

Comment (2): *Co-localization of GSTF1 to MIWI2 and MILI cytoplasmic bodies should be quantified - this information would be valuable for subsequent studies dissecting piRNA pathway proteins' functions.*
Response (2):
We agree with the reviewer's comment. As shown in Fig. 1, all MIWI2 foci appeared to correspond to large GTSF1 foci (see Fig. 1F), and most of GTSF1 foci appeared to be co-stained or overlapped with some of MILI foci (see Fig. 1G). However, quantification of these observations seemed very difficult because anti-GTSF1 antibody has relatively high background staining especially in double staining condition (see Fig. 1D and 1G) and thus it is hard to detect small GTSF1 foci. We would like to pursue this point in the future studies.

To Advisor:

Responses to Specific Comments

Comment (1): *Figure 2a: a two-tailed, unpaired t-test or other appropriate statistical test needs to be used to determine whether these data are significant. The effect on UTR sequences is much larger than on the L1 ORF2 or IAP Gag ORF sequence. Why?*

Response (1):

The experiment shown in Figure 2a was intended to show that GTSF1 is essential for repressing LINE-1 and IAP expression during the development from prospermatogonia to spermatogonia. To better clarify this point, we performed immunofluorescent staining for P0, P4, and P8 testes instead of RT-qPCR. The result clearly showed derepression of LINE-1 and IAP expression in GTSF1-/- prospermatogonia (see Fig. EV1A).

Comment (2): *What are the intense, L1 ORF1p-containing foci in Gstf1-/- prospermatogonia?*

Response (2):

These foci are probably Line-1 ribonucleoproteins (L1 RNPs), which were described at least in two papers (Martin & Branciforte, 1993; Soper et al, 2008).

Comment (3): *Figure 4b: A statistical test is required to determine whether the intermitochondrial distance is longer in Gstf1-/- than in Gstf1+/-.*

Response (3):

We found that our data were not sufficient for statistical analysis. We have removed these data (Figure 4b) and relevant description from the revised manuscript.

Comment (4): *I agree with Reviewer 1 that GST-pull downs provide weak evidence at best. Immunoprecipitation/mass spectrometry needs to be performed in testis extract. This is particularly important given the discordant results from the experiments using GST-pull down and those using ectopic expression in cultured somatic cells.*

Response (4):

We understand the Reviewer's and Advisor's concern whether the in vitro interactions between GTSF1 and complexes involved in piRNA pathway truly occur in vivo. We thus performed immunoprecipitation of testis lysates with anti-GTSF1 antibody instead of GST-pull down. The immunoprecipitates were subjected to MS analysis. As summarized in Table EV1, the result strongly supported that mouse GTSF1 binds to and is a component of PIWI complexes in vivo (p.9 line 21-25 in the revised manuscript).

Comment (5): *The small RNA sequencing data appears to be from a single biological replicate. The experiment needs to be repeated so that they authors have at least two biological replicates; three would be optimal.*

Response (5):

According to the Advisor's comment, we confirmed our results and conclusion based on the data from deep-sequencing of two biological replicates of total small RNAs (see Figs. 3, 4, and EV3)

Comment (6): *Figure 7c: The authors appear not to have sequenced the MILI- and MIWI2-bound piRNA populations. Instead they analyze them solely by end labeling. Unfortunately these data is quite unconvincing. It appears that full-length tRNAs are the main*

RNA recovered in the immunoprecipitates, even in the Gstf1 heterozygotes. The immunoprecipitated 16-35 nt RNA needs to be deep sequenced and fully analyzed as is standard for the field.

Response (6):

According to the Advisor's comment, we performed additional experiments and obtained reliable data for piRNA detection by end-labeling (see Fig. 3D; p.10 line 24 – p.11 line 2 in the revised manuscript). We also performed deep-sequencing of MILI- and MIWI2-bound piRNAs (see Figs. 3, 4, EV3, and EV4; p.11 line 3-11 and 13-20 in the revised manuscript).

Comment (7): *Figure 8b: The author's analysis of piRNA sequence composition is helpful, but cannot substitute for a global analysis of piRNA sequence composition by position.*

Response (7):

We prepared reference figures showing global analysis of piRNA sequence composition by position (see Ref. Figs. 1-4).

Comment (8): *Figure 8c: Z-scores for piRNA Ping-Pong need to be included. While it is useful to analyze Line1- and IAP-mapping piRNAs, Ping-Pong analysis for all genome-mapping piRNAs should also be presented. Moreover, the authors have missed a key finding indicated by their analysis of IAP piRNAs. The authors determine the 5′-5′ distance for IAP-mapping piRNAs on opposite genomic strands for GSTf1 heterozygous and homozygous. For heterozygotes, they observe the expected 10 nt peak indicative of Ping-Pong. In contrast, in E17.5 testes from the Gstf1 homozygous mutant testes, there is a prominent peak at 17 nt (Z-score is needed to determine its statistical significance). What is this peak?*

Response (8):

We have changed "relative frequency" in Y-axis to "Z-score" (Zhang et al, 2011) for 5′-5′ distance analysis of piRNAs. Further, we have prepared reference figures showing that i) 17-nt peak was not observed when 5′-5′ distance was analyzed within uniquely mapped piRNAs by Z-scores (see Ref. Fig. 5A) and ii) piRNA pairs corresponding to the 17-nt peak were mapped to limited positions in IAP genomic sequences (see Ref. Fig. 5B). This implies that the 17-nt peak is probably artifacts.

Response to Minor Points

Comment (1): *I don't agree with the minor point of Reviewer 2:*
"MIWI2 and TDRD9 cytoplasmic granules are not just processing bodies (P-bodies). They share some components with somatic cells' P-bodies but are distinct from those both by the presence of numerous piRNA pathway proteins and their appearance (See and cite Aravin et al 2009).

This point needs to be clarified."
Since 2009, the authors have used the term "processing bodies" to describe the perinuclear foci containing TDRD9 and MIWI2. Nowhere in the manuscript do they abbreviate processing body to "P body." The nomenclature is confusing, and indeed it confused Reviewer 2.

Response (1):

According to the Referee 2's comment (1) and Advisor's comment (1), we have changed a term 'processing body' to 'pi-P body' to avoid readers' possible confusion (Aravin et al, 2009; p.4 line 7, p.6 line 12 and 17, p.7 line 6, 10, 12, and 13, p.8 line 2, 16, 19, and 20, p16 line 12, and 14, and p.33 line 2, 7, 15, 18, and 21 in the revised manuscript).

References for responses

Aravin AA, van der Heijden GW, Castaneda J, Vagin VV, Hannon GJ, Bortvin A (2009) Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. PLoS Genet 5: e1000764

Donertas D, Sienski G, Brennecke J (2013) Drosophila Gtsf1 is an essential component of the Piwi-mediated transcriptional silencing complex. Genes Dev 27: 1693-705

Martin SL, Branciforte D (1993) Synchronous expression of LINE-1 RNA and protein in mouse embryonal carcinoma cells. Mol Cell Biol 13: 5383-92

Muerdter F, Guzzardo PM, Gillis J, Luo Y, Yu Y, Chen C, Fekete R, Hannon GJ (2013) A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila. Mol Cell 50: 736-48

Ohtani H, Iwasaki YW, Shibuya A, Siomi H, Siomi MC, Saito K (2013) DmGTSF1 is necessary for Piwi-piRISC-mediated transcriptional transposon silencing in the Drosophila ovary. Genes Dev 27: 1656-61

Soper SFC, van der Heijden GW, Hardiman TC, Goodheart M, Martin SL, de Boer P, Bortvin A (2008) Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. Dev Cell 15: 285-297

Xiol J, Spinelli P, Laussmann MA, Homolka D, Yang Z, Cora E, Coute Y, Conn S, Kadlec J, Sachidanandam R, Kaksonen M, Cusack S, Ephrussi A, Pillai RS (2014) RNA clamping by Vasa assembles a piRNA amplifier complex on transposon transcripts. Cell 157: 1698-711

Yu Y, Gu J, Jin Y, Luo Y, Preall JB, Ma J, Czech B, Hannon GJ (2015) Panoramix enforces piRNA-dependent cotranscriptional silencing. Science 350: 339-42

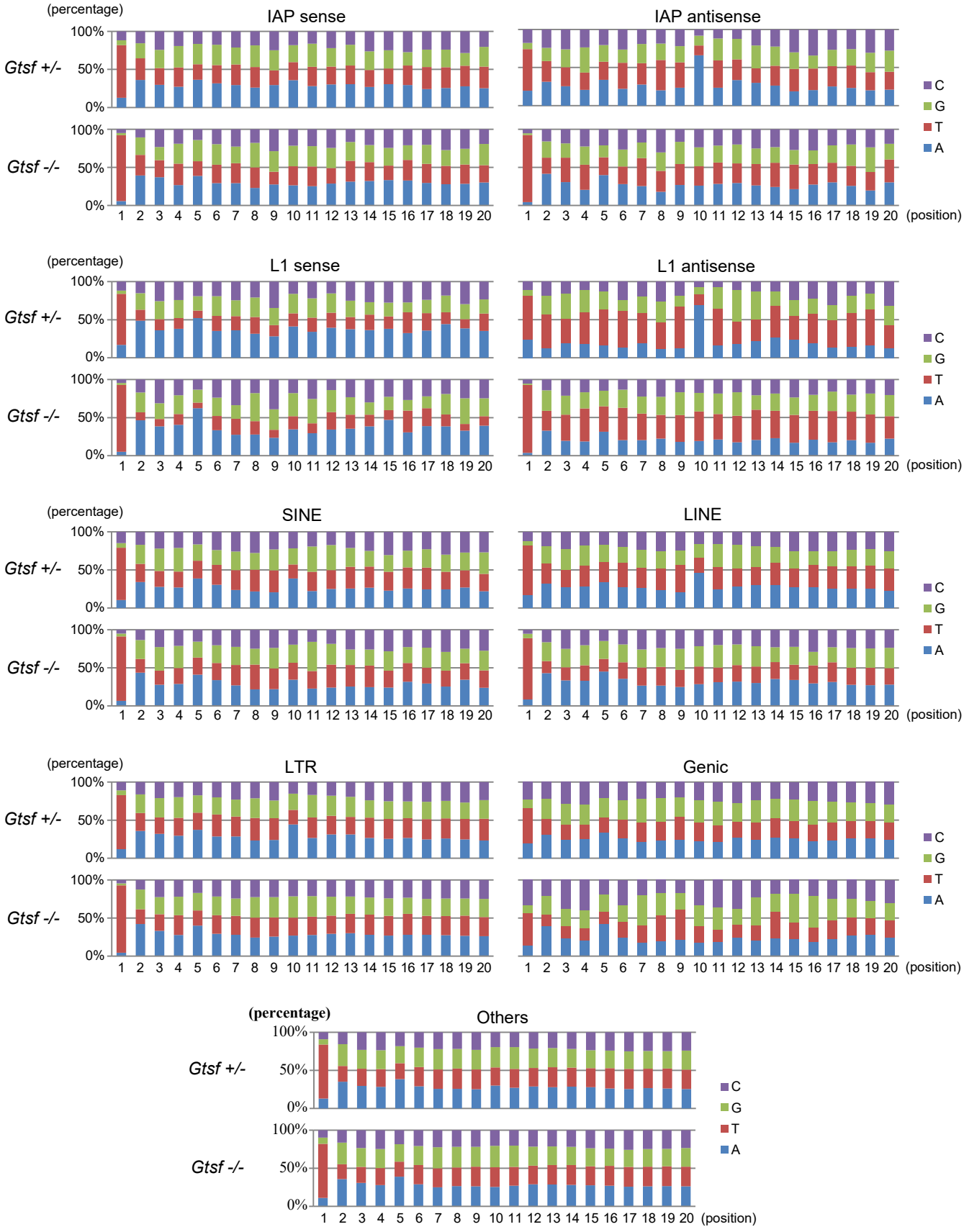
Zhang Z, Xu J, Koppetsch BS, Wang J, Tipping C, Ma SM, Weng ZP, Theurkauf WE, Zamore PD (2011) Heterotypic piRNA Ping-Pong Requires Qin, a Protein with Both E3 Ligase and Tudor Domains. Mol Cell 44: 572-584

Corresponding table of the figures between the first and revised manuscripts

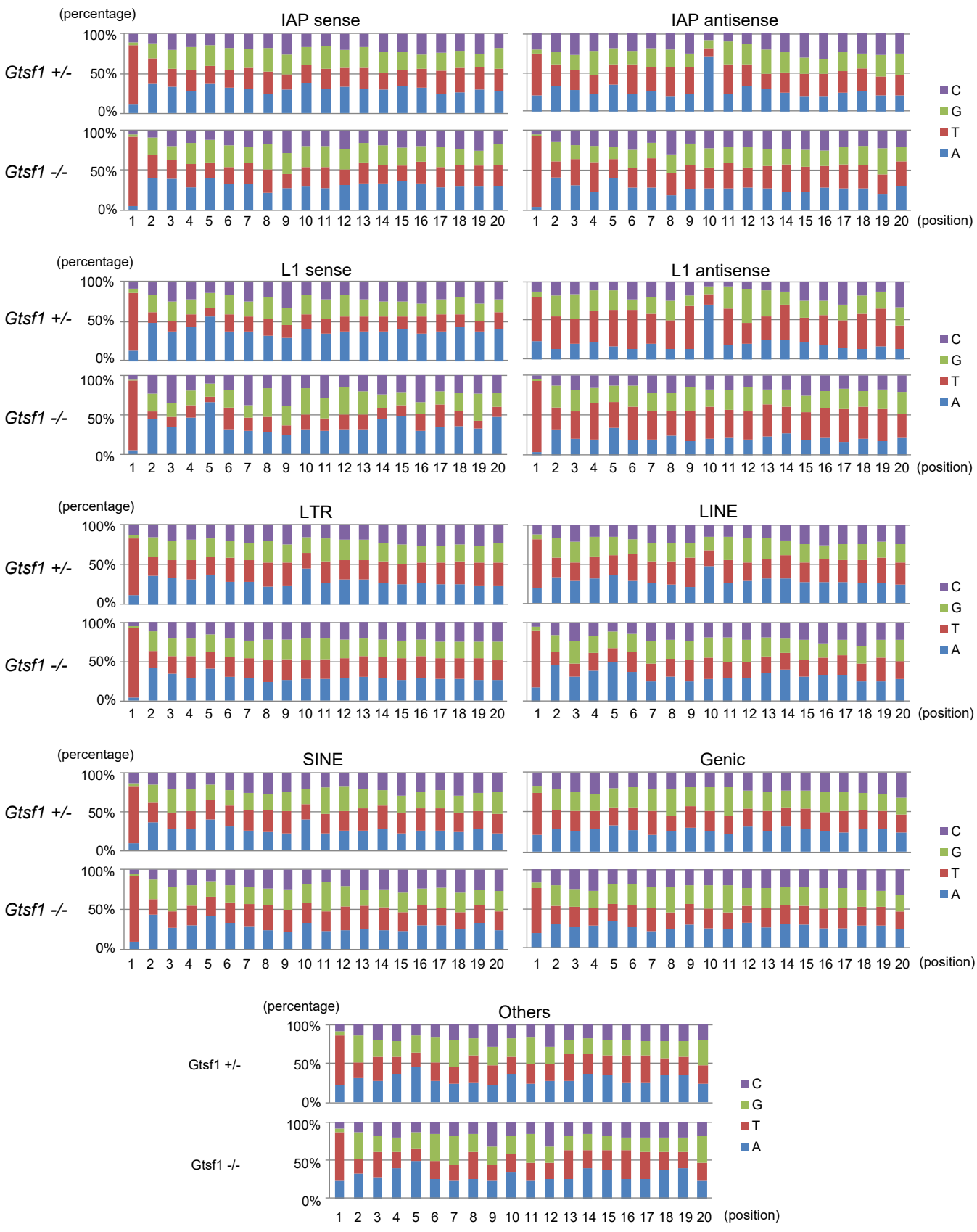
First submission	Revised submission
Fig.1A	Fig.1D
Fig.1B	Fig.1E
Fig.1C	Fig.1F
Fig.1D	Fig.1G
Fig.2A	retracted
Fig.2B	Fig.1A
Fig.2C	Fig.1B
Fig.3A	Fig.1H
Fig.3B	Fig.1I
Fig.3C	Fig.1J
Fig.3D	Fig.1K
Fig.3E	Fig.1L
Fig.3F	Fig.1M
Fig.4A	retracted
Fig.4B	retracted
Fig.5A	Fig.1N
Fig.5B	Fig.1O
Fig.5C	Fig.1P
Fig.6A	Fig.2A
Fig.6B	Fig.EV2A
Fig.6C	Fig.2B
Fig.6D	Fig.EV2B

Fig.6E	Fig.2C
Fig.6F	Fig.2D
Fig.7A	Fig.EV3B
Fig.7B	Fig.EV3A
Fig.7C	retracted
Fig.7D	Fig.3B
Fig.7E	retracted
Fig.8A	retracted
Fig.8B	Fig.4C and 4D
Fig.8C	Fig.4E
Fig.9	retracted
Fig.S1A	Fig.EV1C
Fig.S1B	Fig.EV1C
Fig.S2A	Fig.EV1D
Fig.S2B	Fig.EV1D
Fig.S3	Fig.EV2D
Fig.S4A	retracted
Fig.S4B	retracted
Fig.S5A	Fig.EV6E and EV6F
Fig.S5B	Fig.EV7E and EV7F
Supplemental Table S1	Table EV2

Base composition in Total small RNA sequences rep#1

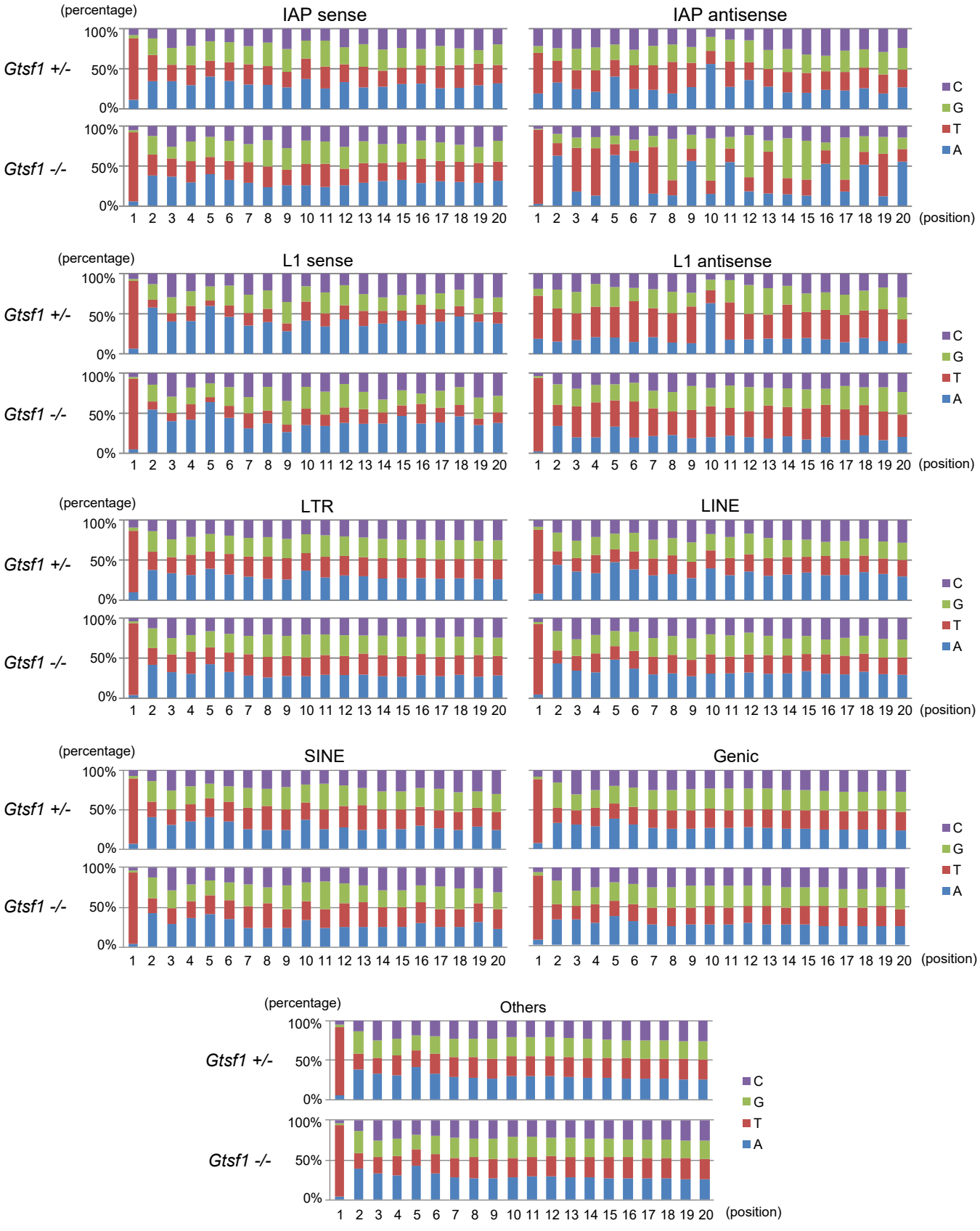


Reference Figure 2



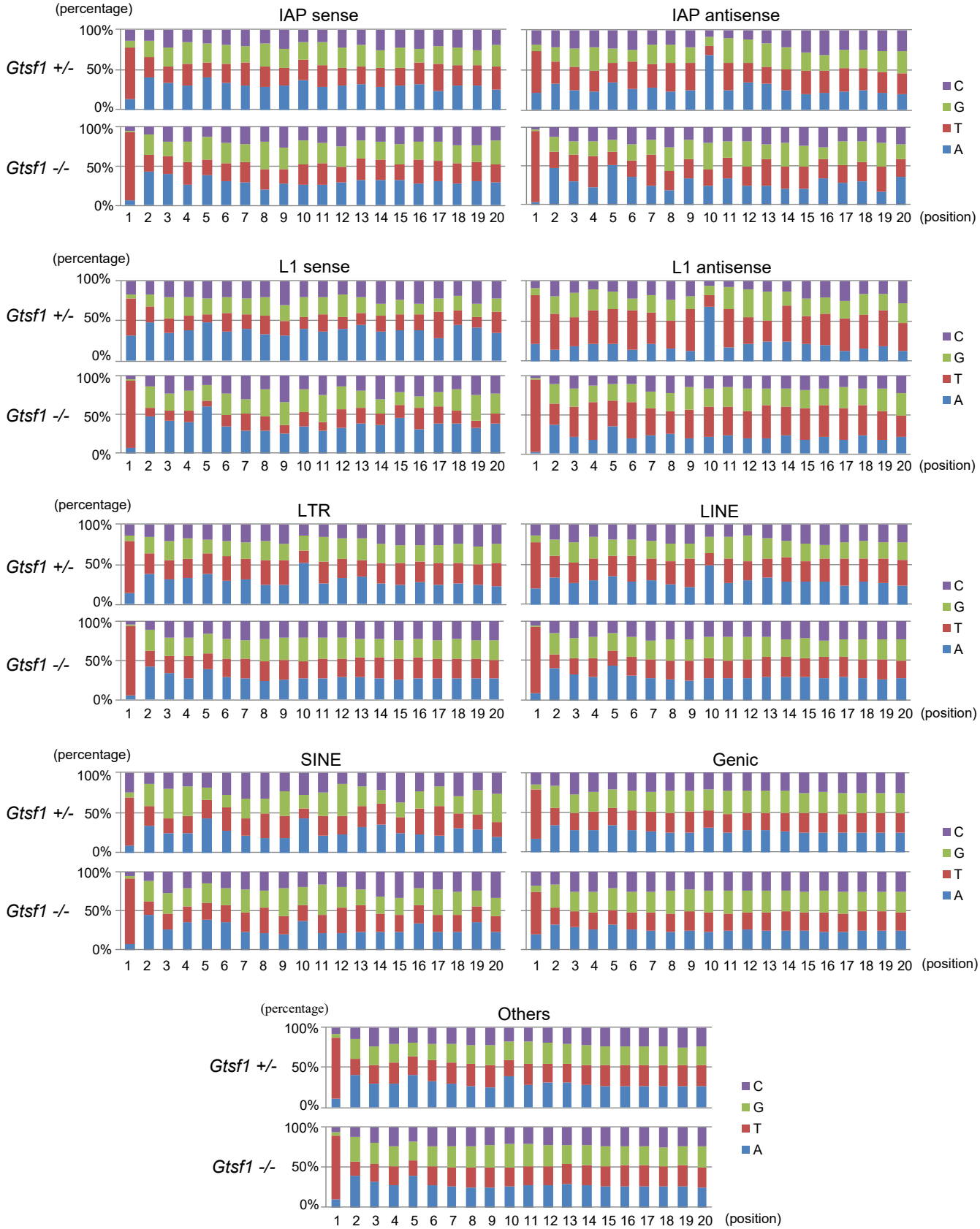
Reference Figure 3

Base composition in MILI-bound piRNA sequences



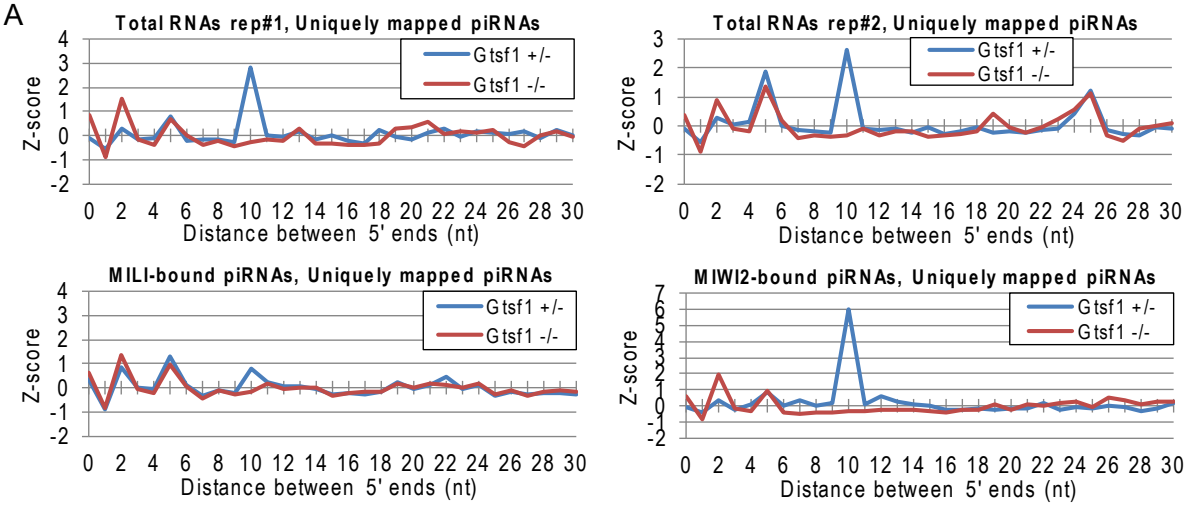
Base composition in MIWI2-bound piRNA sequences

Referece Figure 4



Reference Figure 5

A



B

2nd Editorial Decision 21 December 2017

Thank you for the submission of your revised manuscript to EMBO reports. We have now received the comments from the referees that were asked to assess it, and I am happy to say that both supports its publication now.

I think that referee 1's suggestion to change the title is good. A few other changes are also needed:

- EMBO reports reference style is numbered and can be found in EndNote, please correct. A maximum of 10 author names may be listed.

- please send us a completed author checklist that can be found here: http://embor.embopress.org/authorguide#revision. The completed author checklist will also be part of the review process file (RPF) published online along with accepted manuscripts at EMBO press (transparent peer-review).

- Figure EV1D is missing a sale bar.

- Figure EV4D states n=2, in this case no error bars can be shown. The experiment either needs to be repeated one more time, or the error bars removed. If n=2 all data points from both experiments can be shown along with their mean.

- Tables EV2 and EV3 could be called Dataset 1 and 2 instead. Table EV4 could be a regular table in the methods section.

- All EV figures and tables need to be uploaded as individual files please.

- We normally only allow 5 EV figures, exceptionally 6. Can you either combine some of the EV figures, or move one figure to the main manuscript file? Alternatively, you can move all or extra EV figures to an Appendix file, which is, however, not embedded in the main manuscript text online, as the EV figures are. You can find more information about EV figures and Appendix in our guide to authors.

- We need an ORCID ID for all corresponding authors. Please add this ID to your personal profile page in our online manuscript system. We can unfortunately not do this for you.

- please send us A) a short (1-2 sentences) summary of the findings and their significance, B) 2-3 bullet points highlighting key results and C) a synopsis image that is 550x200-400 pixels large (the height is variable). You can either show a model or key data in the synopsis image. Please note that text needs to be readable at the final size.

- The abstract needs to be written in present tense when it comes to the findings reported in the manuscript, please amend.

I look forward to seeing a final version of your manuscript as soon as possible. Please let me know if you have any questions or comments.

REFEREE REPORTS

Referee #1:

I am happy with the revised version of the manuscript and it is now ready for publication. An excellent study that will clarify the role of GTSF1 in the mouse system and draw a contrasting picture to that in flies. What this study shows clearly is that effect on the nuclear function of the piRNA pathway is due to cytoplasmic role of GTSF1 in biogenesis of piRNAs loaded into the nuclear MIWI2.

Perhaps the authors may consider to make a minor change as listed below. It is only a suggestion.

The title says that GTSF1 is crucial for piRNA-guided cleavage of target RNAs. This is a very strong statement. If one believes that Argonautes (Piwis) are small RNA-guided nucleases, they do not need any help from additional proteins to find targets and cleave. Title of the paper could be "Mouse GTSF1 is a secondary piRNA biogenesis factor" or "Mouse GTSF1 is essential for biogenesis of secondary piRNAs associating with MIWI2".

Referee #2:

The authors are to be commended for responding so thoroughly to the original critiques. They have produced a truly beautiful paper that is ready for immediate publication in EMBO Reports.

2nd Revision - authors' response 11 November 2017

Thank you very much for your letter dated December 21, 2017, together with the referees' comments on our revised manuscript. We are glad to hear that our manuscript is acceptable for publication in *EMBO Reports*.
According to Referee 1's comment, we have changed the title to "Mouse GTSF1 is an essential factor for secondary piRNA biogenesis". Further, according to the required points for publication, we have revised the text and figures. Our responses to each point are described below. Additionally, we checked the whole manuscript and made minor revisions (font and format) in accordance with the author guidelines.

Point 1: *EMBO reports reference style is numbered and can be found in EndNote, please correct. A maximum of 10 author names may be listed.*
Response 1:
We changed to EMBO reports reference style.

Point 2: *please send us a completed author checklist that can be found here: http://embor.embopress.org/authorguide#revision. The completed author checklist will also be part of the review process file (RPF) published online along with accepted manuscripts at EMBO press (transparent peer-review).*
Response 2:
We filled the checklist and attached it.

Point 3: *Figure EV1D is missing a sale bar.*
Response 3:
We have added scale bars in Fig EV1D.

Point 4: *Figure EV4D states n=2, in this case no error bars can be shown. The experiment either needs to be repeated one more time, or the error bars removed. If n=2 all data points from both experiments can be shown along with their mean.*
Response 4:
We removed the error bars. Instead, we showed the data points with their mean in Figs 4C, 4D, and EV4D. We have added a sentence explaining these graphs in each figure legend.

Point 5: *Tables EV2 and EV3 could be called Dataset 1 and 2 instead. Table EV4 could be a regular table in the methods section.*
Response 5:
We moved Table EV2 to Dataset 1, Table EV3 to Dataset 2, Table EV4 to Table in Materials and Methods section, and Table EV5 to Table EV2.

Point 6: *All EV figures and tables need to be uploaded as individual files please.*
Response 6:
We attached individual PDF files.

Point 7: *We normally only allow 5 EV figures, exceptionally 6. Can you either combine some of the EV figures, or move one figure to the main manuscript file? Alternatively, you can move all or extra EV figures to an Appendix file, which is, however, not embedded in the main manuscript text online,*

as the EV figures are. You can find more information about EV figures and Appendix in our guide to authors.
Response 7:
We moved Fig EV5 to Fig 6 and gave a new figure title for this figure. Accordingly, we moved Fig 6 to Fig 7, Fig EV6 to Fig EV5, and Fig EV7 to Fig EV6.

Point 8: *We need an ORCID ID for all corresponding authors. Please add this ID to your personal profile page in our online manuscript system. We can unfortunately not do this for you.*
Response 8:
We added the ID https://orcid.org/0000-0003-2475-589X to our corresponding author's personal profile page.

Point 9: *please send us A) a short (1-2 sentences) summary of the findings and their significance, B) 2-3 bullet points highlighting key results and C) a synopsis image that is 550x200-400 pixels large (the height is variable). You can either show a model or key data in the synopsis image. Please note that text needs to be readable at the final size.*
Response 9:
We prepared a word document and a synopsis image (550x350 pixels large) for synopsis section and attached them.

Point 10: *The abstract needs to be written in present tense when it comes to the findings reported in the manuscript, please amend.*
Response 10:
We changed sentences in past tense to in present one in the Abstract.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Jun-ichi Miyazaki
Journal Submitted to: EMBO Reports
Manuscript Number: EMBOR-2016-42054V2

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	For animal experiments, we chose minimal sample size that allowed us to analyse the data statistically, according to the principle of the three Rs. The statictical data we analysed between two groups were obtained based on technical and biological replication.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	Sample size was described in "Materials and Methods" or Figure legends.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	NA
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	Yes, animals or samples subjected to treatment were chosen at random.
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	Yes. Whenever possible, experiments were performed in a blinded manner.
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes. Normality test and F test.
Is there an estimate of variation within each group of data?	Yes
Is the variance similar between the groups that are being statistically compared?	Yes.

C- Reagents

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	We described these informations in "Materials and Methods" - Antibodies.
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	We previously described these informations in our research article "Yoshimura T, et al. (2009) Gtsf1/Cue110, a gene encoding a protein with two copies of a CHHC Zn-finger motif, is involved in spermatogenesis and retrotransposon suppression in murine testes. Dev Biol 335: 216-27."
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	All experiments involving animals were carried out in accordance with the institutional guidelines under the protocols (No. 21-089 and 25-061), which were approved by the Animal Care and Use Committee of the Osaka University Graduate School of Medicine. We described these in "Materials and Methods" - Animals.
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	Mice were housed and maintained in a controlled environment according to the institutional guidelines. Our animal institute has obtained the Animal Welfare Assurance (#A5950-01) from the Office of Laboratory Animal Welfare (OLAW) of National Institutes of Health, USA.

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	We described accession numbers to the source data of deep sequencing at the end of "Materials and Methods" - Accession numbers.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	NA
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA