THE EMBO JOURNAL

GTSF-1 is required for formation of a functional RNA-dependent RNA Polymerase complex in *C. elegans*

Miguel Vasconcelos Almeida, Sabrina Dietz, Stefan Redl, Emil Karaulanov, Andrea Hildebrandt, Christian Renz, Helle D. Ulrich, Julian König, Falk Butter and René F. Ketting.

Review timeline:

Submission date:	26th Febuary 2018
Editorial Decision:	16th March 2018
Revision received:	22nd March 2018
Accepted:	24th March 2018

Editor: Anne Nielsen

Transaction Report:

(Transcription note: This manuscript was reviewed at another journal before being transferred to The EMBO Journal. The original referee reports are not included in this Review Process File as they are not covered by the EMBO Press transparency policy.)

1st Editorial Decision 16th March 2018

Thank you for submitting your manuscript to The EMBO Journal along with the referee reports from another journal as well as your point-by-point response to the concerns raised. These files have now all been seen by an arbitrating advisor whose comments are shown below.

As you will see, our advisor finds that the changes requested in the previous round of review have been well incorporated and that the study offers sufficient advance to merit publication in The EMBO Journal. The advisor does make a couple of suggestions for how the manuscript presentation could be further improved but these are all rather minor points that will not require the generation of additional data. I would therefore like to invite you to submit a revised manuscript in which you address the four minor points mentioned by the advisor. Regarding point 2, I'll leave it up to you to decide how you structure the narrative in the final manuscript.

Arbitrating advisor:

The manuscript by Ketting and colleagues on the role of C. elegans Gtsf1 in the 26G small RNA pathway is a comprehensive characterization of this small Zinc finger protein in worms, which yields several novel and important findings. The experiments are well executed, appropriate controls are included, the interpretation of the data is sound, and the text is overall well written.
The authors have improved their manuscript also in light of the previous review rounds, and have addressed most points raised in a satisfactory manner. I further agree with the authors responses towards the more challenging questions as of how Gtsf1 mechnistically works. This is certainly a separate and exciting follow-up question that is beyond the scope of this paper.

I have only four suggestions:
1. For IF images, it is in my opinion advantageous to show the individual channels as black and

white or as inverted black and white images as this gives superior contrast
2. I understand the historical development of this work. The authors wanted to probe for the function of Gtsf1 in elegans. I do wonder, however, whether there would not be a much more exciting way to actually motivate this study, which would be to start with ERIC and the fact that this complex is evidently regulated in its assembly. That must mean something very important and would give the story an entirely different spin. I realize that this would require major restructuring of the work and figures. But the authors simply might want to think about this.
3. The thing that in my opinion does not get entirely clear is the point about whether Gtsf1 leaves the mature complex or not. If it is still present (even at lower levels), why would not all complex members be identified by mass-spec? The authors might want to add clarity here.
4. Finally, there are several Gtsf1 genes in flies and I think also in mammals. The function of these is entirely unclear. It could be that C. elegans Gtsf1 is in fact close to these sister genes rather than to the Gtsf1 that is described in the fly and mouse literature.
The authors should include this as a discussion point.

1st Revision - authors' response 22nd March 2018

We thank the arbitrator for the time and effort to evaluate our manuscript. Below we will provide a point-by-point reply to each individual comment that was made.

1. For IF images, it is in my opinion advantageous to show the individual channels as black and white or as inverted black and white images as this gives superior contrast

Reply: We did so, and updated Figure 1 with GTSF-1::mCherry and GFP::ALG-3 fluorescence signal as black and white.

2. I understand the historical development of this work. The authors wanted to probe for the function of Gtsf1 in elegans. I do wonder, however, whether there would not be a much more exciting way to actually motivate this study, which would be to start with ERIC and the fact that this complex is evidently regulated in its assembly. That must mean something very important and would give the story an entirely different spin. I realize that this would require major restructuring of the work and figures. But the authors simply might want to think about this.

Reply: The way that the manuscript was set-up was to also emphasize the difference we see with the role of GTSF1 in piRNA biology in other organisms. That's why we believe that starting from that angle brings very relevant information. In addition, we believe we do not make sufficient progress on understanding why RRF-3 is regulated as it is to take the perspective suggested by the reviewer. Hence, we maintain the original flow of the manuscript.

3. The thing that in my opinion does not get entirely clear is the point about whether Gtsf1 leaves the mature complex or not. If it is still present (even at lower levels), why would not all complex members be identified by mass-spec? The authors might want to add clarity here.

Reply: That is a very good point and is currently simply not completely resolved. To further clarify this, we added the following sentence in the first paragraph of our Discussion section "What is the exact molecular function of GTSF-1?":

"We would like to point out that it is unclear why we do not observe all ERIC components in GTSF-1 IP-MS in embryos (**Fig 7B**), given that GTSF-1 does co-fractionate with mature ERIC (**Fig 7E, H**). There may be several reasons for this. For example, it may be that the epitope of GTSF-1 is inaccessible within ERIC in embryos, or that GTSF-1 more easily dissociates from mature ERIC than from pre-ERIC. We do observe some enrichment of PIR-1 and ERI-1 in GTSF-1 IP-MS in embryos, suggesting the latter scenario may indeed apply."

Also in the legend of Figure 8, where we show our model, we added: "It is unclear how stable the association between GTSF-1 and the mature complex is."

4. Finally, there are several Gtsf1 genes in flies and I think also in mammals. The function of these is entirely unclear. It could be that C. elegans Gtsf1 is in fact close to these sister genes rather than to the Gtsf1 that is described in the fly and mouse literature.
The authors should include this as a discussion point.

Reply: We now added the following sentence in the third paragraph of our discussion section "The double CHHC zinc finger as a protein-protein interaction module":
"Of note, the function of Gtsf1 paralogs in flies has not yet been determined. It may be that these paralog CHHC zinc finger proteins may interact with other proteins via their zinc fingers, and thus have a more similar role to CeGTSF-1 in sRNA biology."

2nd editorial decision 24th March 2018

Thank you for submitting the revised version of your manuscript along with a point-by-point response to the comments made by our arbitrating advisor. I have looked at both and I am happy to inform you that the manuscript has now been officially accepted for publication in The EMBO Journal.

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: Rene F Ketting
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2018-99325

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 $\chi 2$ 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?	Triplicates were assumed as providing sufficient statistical power for most experiments.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	N/A
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	For example for fertility experiments where single animals were used, a particular animal was excluded from the analysis if the animal could not be maintained during the considered egg laying period - e.g. if the nematode crawled to the wall of the plate, subsequently dying.
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.	When assaying for fertility, for example, where single animals were isolated, animals were isolated during early stages of development, before the L4 stage, to randomize for potential germline defects.
For animal studies, include a statement about randomization even if no randomization was used.	N/A
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.	No blinding was performed.
4.b. For animal studies, include a statement about blinding even if no blinding was done	No blinding was performed.
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.	Mostly we used non-parametric Mann-Whitney-Wilcoxon, so normal distribution was not assumed.
Is there an estimate of variation within each group of data?	No.
Is the variance similar between the groups that are being statistically compared?	Not addressed.

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/

C- Reagents

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).	All antibodies used are commercially available, except for the DCR-1 antibody see Duchaine et al, 2006 and Thivierge et al, 2012. We validated our polyclonal GTSF-1 antibodies by the absence of a band in gtsf-1 mutants which lack most of the coding sequence.
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.	Several C. elegans strains were used and created for this study - see supplementary information.
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
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.	NA

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	Sequencing data have been deposited to the NCBI Gene Expression Omnibus (GEO) and proteomics data are available at the ProteomeXchange Consortium via PRIDE. GEO: GSE103432 PRIDE: PXD007665.
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).	See above.
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