EMBO reports

DnaQ exonuclease-like domain of Cas2 promotes spacer integration in a type I-E CRISPR-Cas system

Gediminas Drabavicius, Tomas Sinkunas, Arunas Silanskas, Giedrius Gasiunas, Česlovas Venclovas, and Virginijus Siksnys

Review timeline:

Submission date:	24 November 2017
Editorial Decision:	9 January 2018
Revision received:	6 April 2018
Editorial Decision:	17 April 2018
Revision received:	4 May 2018
Accepted:	8 May 2018

Editor: Esther Schnapp

Transaction Report:

(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 9 January 2018

Thank you for your patience while your manuscript was peer-reviewed at EMBO reports. I am sorry for the delay in getting back to you; we have only now received the full set of enclosed referee reports as well as referee cross-comments.

As you will see, all referees acknowledge that the data are interesting. While referee 2 feels that in vivo data would be required to strengthen the study, the other 2 referees indicate in their cross-comments that in vivo data would go beyond the scope of this study, but that referee 2's concerns should be addressed in the discussion of the manuscript. All other points raised by the referees will need to be addressed.

Given these constructive comments, we would thus 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. You can either publish the study as a short report or as a full article. For short reports, the revised manuscript should not exceed 27,000 characters (including spaces but excluding materials & methods and references) and 5 main plus 5 expanded view (EV) figures. The results and discussion sections must further be combined, which will help to shorten the manuscript text by eliminating some redundancy that is inevitable when discussing the same experiments twice. For a normal article there are no length limitations, but it should have more than 5 main figures and the results and discussion sections must be separate. In

both cases, the entire materials and methods must be included in the main manuscript file, and we cannot offer more than 5 EV figures. Additional supplementary information will need to be moved to an Appendix file. Alternatively, all EV figures can be part of the Appendix file, but in this case they will not be integrated and clickable in the html version of the manuscript. You can find more information about our file types in our guide to authors online.

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 must be provided in the figure legends.

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.

When submitting your revised manuscript, we will require:

- a complete author checklist, which you can download from our author guidelines (http://embor.embopress.org/authorguide#revision). Please insert page numbers in the checklist to indicate where in the manuscript the requested information can be found. The completed author checklist will also be part of the RPF (see below).
- a letter detailing your responses to the referee comments in Word format (.doc)
- a Microsoft Word file (.doc) of the revised manuscript text
- editable TIFF or EPS-formatted figure files in high resolution

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

For the preparation of manuscript figures please check our figure guidelines at http://www.embopress.org/sites/default/files/EMBOPress_Figure_Guidelines_061115.pdf

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File (RPF) 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:

In this paper, Siknsys and colleagues describe biochemical analysis of an unusual CRISPR adaptation module from S. thermophiles where Cas2 is fused to a DnaQ exonuclease domain. The authors show that Cas1 and Cas2 form a complex as expected and support integration (probably half site integration) into a supercoiled plasmid. They also shown that the DnaQ domain has the expected 3' to 5' exonuclease activity. They then show that DnaQ is important to trim extended 3' overhangs from protospacers to a standard length (typically 5 nt) for integration, and demonstrate this both in vitro and by sequencing integration products.

This work is interesting as it is one of very few studies that have coupled protospacer processing to integration. The finding that the DnaQ exonuclease domain trims the 3' ends of protospacer to around 5 nt is a strong signal that other exonucleases will fulfil the same function in other systems, broadening the likely interest in this paper. The authors should cite and discuss the just-published work on prespacer processing and integration in the type I-A system, which is highly relevant to the work described here (https://doi.org/10.1093/nar/gkx1232).

There are some important issues that need to be addressed, as follows:

1. In the figure showing the integration site mapping, the authors show half-site integration at the leader-distal ends of repeat 1 and repeat 2. This is of course determined by their choice of pcr primers, but there is an emerging consensus that the first half integration reaction occurs at the leader-proximal end, and generally integration is much more efficient at this half-site. The choice of leader-distal site presented here is therefore surprising given the state of the art. The authors should repeat this experiment with PCR primers that amplify the products of integration at the leader-proximal site. This will allow them to compare the integration efficiency at the two half-sites and check if there is more specificity for the leader-repeat junction. These experiments are trivial - if they have not been carried out already they just need a new set of PCR primers.
2. In the figure showing metal dependence of integration (EV2), the authors show integration of a radioactive protospacer is supported by magnesium but not manganese - a very unexpected finding. Calcium also supports integration, which is also surprising. Given that other Cas1 enzymes favour Manganese, and indeed Manganese is almost always suitable for magnesium-requiring enzymes, the authors should consider carefully: a) were the lanes mixed up? Or b) is there any precedent for such a metal ion dependence in the literature ? Please discuss.
3. In the figure showing protospacer integration into plasmids as a function of 3' flap length, the quantitation is expressed as "% of relaxed plasmid" and seems to have a maximum of 1.0. Is this a mistake? By eye, the percentage of relaxed plasmid often seems to be around 50% and varies from 10-80 %. Also, although error bars are shown in this figure I can find no information on the statistical treatment of data.
4. Since this paper was submitted, a highly relevant paper on integration in a type I-A system has been published (https://doi.org/10.1093/nar/gkx1232). In the revision, it would be sensible to discuss the related work, which is broadly consistent with the findings and interpretation of this manuscript.

Referee #2:

This is a technically sound paper describing a potentially novel function of a domain of a naturally fused Cas2 protein in generating adaptation intermediates. The in vitro data are of high quality typical for the PI and his group. The paper lacks in one important aspect - in vivo demonstration of relevance of the findings. If such data are provided the paper will become an important contribution to the field.

Referee #3:

This study examines CRISPR adaptation by a Type I-E CRISPR system containing a Cas2-DnaQ fusion protein. DnaQ is the editing subunit of DNA Polymerase III having 3'-5' exonuclease activity. Using a co-expression and purification strategy the authors establish that Cas1 and Cas2-DnaQ form a stable complex, with Cas1x4 and Cas2-DnaQx2 stoichiometry. They then showed that this complex could integrate 3'-overhang protospacers into a plasmid containing a CRISPR array and that the DnaQ domain did indeed have 3'-5' exonuclease activity against single-stranded and double-stranded DNA. Finally, the authors examined the role of DnaQ activity on integration finding that the DnaQ domain trimmed long 3' overhangs to an optimal length for integration.

This is a very timely and well-executed study. The conclusions are clearly supported by the data and the manuscript is well written and presented. I do have a few minor suggestions for improvement.

Minor Comments:

Abstract: as CRISPR systems vary substantially between even closely related strains please indicate the strain (DGCC7710) of S. thermophilus used in these studies.

Considering the importance of Figure 3B, it might be helpful to include a quantification of the gel.

All sequencing data should be included as supplemental information with the manuscript. Also, the number of sequencing reactions performed should be included in the methods, simply stating the percentage is not sufficient.

Cross-comments from referee 2:

There are several known instances of well-documented and meticulously researched in vitro nuclease activities shown to be not relevant in in vivo studies (Cas2 of Type I-E being a prominent and obvious one). Given the purely in vitro character of this (very fine) work, it looks to me that in vivo studies are needed to show that in vitro findings are relevant (or not).
If the issue of (possible) spurious nuclease activity is raised in discussion to keep expectations at check, I'd support publication.

1st Revision - authors' response 6 April 2018

Referee #1:

There are some important issues that need to be addressed, as follows:

1. In the figure showing the integration site mapping, the authors show half-site integration at the leader-distal ends of repeat 1 and repeat 2. This is of course determined by their choice of pcr primers, but there is an emerging consensus that the first half integration reaction occurs at the leader-proximal end, and generally integration is much more efficient at this half-site. The choice of leader-distal site presented here is therefore surprising given the state of the art. The authors should repeat this experiment with PCR primers that amplify the products of integration at the leader-proximal site. This will allow them to compare the integration efficiency at the two half-sites and check if there is more specificity for the leader-repeat junction. These experiments are trivial - if they have not been carried out already they just need a new set of PCR primers.
Reply: Experiments suggested by reviewer were performed and results are included in Figure EV5. Taken together data show that the protospacers that are integrated have 3‘ end overhang distribution centered on 5nt, independent on which side of the first repeat the integration takes place. We updated text on page 8 to reflect this.

2. In the figure showing metal dependence of integration (EV2), the authors show integration of a radioactive protospacer is supported by magnesium but not manganese - a very unexpected finding. Calcium also supports integration, which is also surprising. Given that other Cas1 enzymes favour Manganese, and indeed Manganese is almost always suitable for magnesium-requiring enzymes, the authors should consider carefully: a) were the lanes mixed up? Or b) is there any precedent for such a metal ion dependence in the literature ? Please discuss.
Reply: We re-analysed metal ion requirements for integration from scratch using integration complexes with a catalytically active and dead DnaQ. The results show, that magnesium, calcium and manganese all support integration. It turned out that our previous observation that manganese ions did not support integration was due to the fact that Mn^{2+} activates DnaQ domain so strongly, that protospacers were degraded before integration could take place. Integration experiment with inactive DnaQ, shows that manganese ions indeed support integration. Calcium seems to support integration as well, however integration effciency in the presence of calcium is lower than in the case of magnesium.

3. In the figure showing protospacer integration into plasmids as a function of 3' flap length, the quantitation is expressed as "% of relaxed plasmid" and seems to have a maximum of 1.0. Is this a mistake? By eye, the percentage of relaxed plasmid often seems to be around 50% and varies from

10-80 %. Also, although error bars are shown in this figure I can find no information on the statistical treatment of data.

Reply: We have changed the plot to show absolute values of relaxed plasmid, rather than relative values to the maximum relaxation observed. Following reviewers suggestion, we included a sentence about statistical treatment of the data in the legend of the figure EV4 on page 28. The error bars represent standard errors of the mean values of the samples.

4. Since this paper was submitted, a highly relevant paper on integration in a type I-A system has been published (https://doi.org/10.1093/nar/gkx1232). In the revision, it would be sensible to discuss the related work, which is broadly consistent with the findings and interpretation of this manuscript.

Reply: The findings of the relevant paper were cited and briefly discussed in relation to our findings on page 10.

Referee #2:

This is a technically sound paper describing a potentially novel function of a domain of a naturally fused Cas2 protein in generating adaptation intermediates. The in vitro data are of high quality typical for the PI and his group. The paper lacks in one important aspect - in vivo demonstration of relevance of the findings. If such data are provided the paper will become an important contribution to the field.

Cross-comments from referee 2:
There are several known instances of well-documented and meticulously researched in vitro nuclease activities shown to be not relevant in in vivo studies (Cas2 of Type I-E being a prominent and obvious one). Given the purely in vitro character of this (very fine) work, it looks to me that in vivo studies are needed to show that in vitro findings are relevant (or not).
If the issue of (possible) spurious nuclease activity is raised in discussion to keep expectations at check, I'd support publication.

Reply: Following reviewer‘s comment we have included a sentence in the main text (Pages 8 and 9), that findings of this study may not be extraploatable to *in vivo* systems, as per reviewer‘s observation.

Referee #3:

This study examines CRISPR adaptation by a Type I-E CRISPR system containing a Cas2-DnaQ fusion protein. DnaQ is the editing subunit of DNA Polymerase III having 3'-5' exonuclease activity. Using a co-expression and purification strategy the authors establish that Cas1 and Cas2-DnaQ form a stable complex, with Cas1x4 and Cas2-DnaQx2 stoichiometry. They then showed that this complex could integrate 3'-overhang protospacers into a plasmid containing a CRISPR array and that the DnaQ domain did indeed have 3'-5' exonuclease activity against single-stranded and double-stranded DNA. Finally, the authors examined the role of DnaQ activity on integration finding that the DnaQ domain trimmed long 3' overhangs to an optimal length for integration.

This is a very timely and well-executed study. The conclusions are clearly supported by the data and the manuscript is well written and presented. I do have a few minor suggestions for improvement.

Minor Comments:

Abstract: as CRISPR systems vary substantially between even closely related strains please indicate the strain (DGCC7710) of S. thermophilus used in these studies.

Reply: We indicated the strain in the abstract.

Considering the importance of Figure 3B, it might be helpful to include a quantification of the gel.

Reply: We included quantification of the gel, as per reviewer's suggestion in the Figure 3B, just below timepoints. We also included a note in the legend of the figure to reflect this addition on page 23.

All sequencing data should be included as supplemental information with the manuscript. Also, the number of sequencing reactions performed should be included in the methods, simply stating the percentage is not sufficient.

Reply: Following referee's suggestion, all sequencing data will be uploaded as source data for Figure EV5. Number of sequencing reactions is indicated in the figure EV5 legend on page 29.

2nd Editorial Decision 17 April 2018

Referee 1 has reviewed all your responses to the referee comments and supports the publication of your study now. S/he only suggests one more change that I would like you to incorporate before we proceed with the official acceptance.

A few other changes are also needed:

- please send us a short running title and up to 5 keywords
- the callout to Fig 2C comes after Fig 3A, please correct
- please upload table EV1 in either word or excel format, and please change the callout in the manuscript text to table EV1
- figure EV3 needs to fit on one single page
- please upload the Source Data for fig 2B + 2C in one single file

I also attach a word file with comments from our data editors on the figure legends. Please make the necessary changes in the word file (and full manuscript file) using the "track changes" option and send us back the corrected file for us to check.

EMBO press papers are accompanied online by 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. Please send us this information along with the revised manuscript.

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

REFEREE REPORT

Referee #1:

The authors have considered the comments of all referees and made helpful changes to the data and the text. The work as it now stands looks suitable for publication, and is an interesting addition to the field. To bring the paper absolutely up to date, the authors could add a brief mention of the two papers recently published in Molecular Cell, and Cell Reports which report on the activity of Cas4 in prespacer processing in type I-C and I-D systems.

2nd Revision - authors' response 4 May 2018

Following issues were addressed in response to your request:

- We have cited papers suggested by the reviewer as per recommendation. Changes are indicated as a comment in Manuscript file.

- Text file 'Running Title and Keywords' is attached.
- We have included a short sentence in the draft that calls out Fig 2C before Fig 3A. This change is indicated as a comment in the manuscript file.
- Table EV1 has been uploaded to the manuscript central. Callouts have been changed in the text as well.
- Figure EV3 has been downscaled and now fits into a single page. File uploaded to the manuscript central.
- Source Data for Fig 2B and 2C has been uploaded to the manuscript central as a single image file.
- We made changes and approvals to figure legends as requested.
- Summary and Highlights file is attached
- Synopsis image has been uploaded to the manuscript central and attached to this mail.

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: Virginijus Siksnys
Journal Submitted to: EMBO Reports
Manuscript Number: EMBOR-2017-45543V2

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?	No statistical methods were used to determine sample sizes as the effect sizes could not be pre-specified.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
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.	NA
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.	NA
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?	NA
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	NA
Is there an estimate of variation within each group of data?	Yes
Is the variance similar between the groups that are being statistically compared?	No groups of data were statistically compared.

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).	NA
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.	NA
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	NA
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).	We have included Sanger sequencing data used to generate Figure EV5 as a source data file for that figure.
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