Peer Community In Genomics

DefenseFinder update advances prokaryotic antiviral system research

Sishuo Wang *based on peer reviews by* ***Pierre Pontarotti*** *,* ***Pedro Leão*** *and 1 anonymous reviewer*

F. Tesson, R. Planel, A. Egorov, H. Georjon, H. Vaysset, B. Brancotte, B. Néron, E. Mordret, A Bernheim, G. Atkinson, J. Cury (2024) A Comprehensive Resource for Exploring Antiphage Defense: DefenseFinder Webservice, Wiki and Databases. BioRxiv, ver. 4, peer-reviewed and recommended by Peer Community in Genomics.
https://doi.org/10.1101/2024.01.25.577194

Submitted: 18 April 2024, Recommended: 12 August 2024

Cite this recommendation as:
Wang, S. (2024) DefenseFinder update advances prokaryotic antiviral system research. *Peer Community in Genomics*, 100373. 10.24072/pci.genomics.100373

Published: 12 August 2024

Prokaryotic antiviral systems, such as CRISPR-Cas and restriction-modification systems, provide defense against viruses through diverse mechanisms including intracellular signaling, chemical defense, and nucleotide depletion. However, bioinformatic tools and resources for identifying and cataloging these systems are still in development. The work by Tesson and colleagues (2024) presents a significant advancement in understanding the defense systems of prokaryotes. The authors have provided an update of their previously developed online service DefenseFinder, which helps to detect known antiviral systems in prokaryotes genomes (Tesson et al. 2022), plus three new databases: one serving as a wiki for defense systems, one housing experimentally determined and AlphaFold2-predicted structures, and a third one consisting of precomputed results from DefenseFinder. Users can analyze their own data through the user-friendly interface. This initiative will help promote a community-driven approach to sharing knowledge on antiphage systems, which is very useful given their complexity and diversity. The authors' commitment to maintaining an up-to-date platform and encouraging community contributions makes this resource accessible to both newcomers and experienced researchers in the rapidly growing field of defense system research. Experienced researchers will find that there are ways to contribute to the future expansion of these databases, while new users can easily access and use the platform. Overall, the updated DefenseFinder, as well as the other databases introduced in the manuscript, are well-suited for researchers (both dry- and wet-lab ones) interested in antiphage defense. I am hopeful that the efforts by the authors will collectively create valuable online resources for researchers in this field and will foster an environment of open science and accessible bioinformatics tools.

References:

Tesson F, Hervé A, Mordret E, Touchon M, d'Humières C, Cury J, Bernheim A (2022) Systematic and quantitative view of the antiviral arsenal of prokaryotes. Nature Communications, 13, 2561. https://doi.org/10.1038/s41467-022-30269-9

Tesson F, Planel R, Egorov A, Georjon H, Vaysset H, Brancotte B, Néron B, Mordret E, Atkinson G, Bernheim A, Cury J (2024) A comprehensive resource for exploring antiphage defense: DefenseFinder webservice, wiki and databases. bioRxiv, ver. 4 peer-reviewed and recommended by Peer Community in Genomics. https://doi.org/10.1101/2024.01.25.577194

Reviews

Evaluation round #2

DOI or URL of the preprint: https://doi.org/10.1101/2024.01.25.577194
Version of the preprint: 3

Authors' reply, 24 July 2024

Dear Dr. Wang,

Thank you for the writing suggestions to improve the quality of the paper. We uploaded a new version of the paper (https://www.biorxiv.org/content/10.1101/2024.01.25.577194v4) with those suggestions, as well as with the figures in higher quality.

Best regards

Jean Cury, on behalf of the authors.

Decision by Sishuo Wang, posted 15 July 2024, validated 16 July 2024

Minor revision (some remaining grammar issues)

Dear Dr. Cury,

The revised manuscript greatly addressed the concerns from the reviewers. I am here providing a few additional comments on the writing. Since they are easy to accommodate, I'm hopeful that you will be able to resubmit shortly. Thanks for your efforts.

- In the 1st paragraph, consider adding "and" between "disruption, production of antiviral molecules"
- change "or with a domain replacement" to "or genes with ....."
- p6: phylums -> phyla
- p9: ..... -> etc.
- p10: "1,500 predictions of homodimers or dimers". Do you mean "homo- and heterodimers"?
- p10: "2 proteins, complexes with up to 4 proteins were also computed in 1:1 stoichiometry", where 2 -> two, 4 -> four
- Postgresql -> PostgreSQL
- p13: Refseq -> RefSeq

- Some sentences should be merged with the previous paragraph. As an example, the sentence starting with "Importantly for the community".
- The resolution of Figs. 2-3 should be improved, as they are screenshots. However, please ignore this comment if you have prepared independent higher-resolution figures in pdf or svg format (the current file I am reviewing may come from a word file)

Evaluation round #1

DOI or URL of the preprint: https://doi.org/10.1101/2024.01.25.577194
Version of the preprint: 2

Authors' reply, 10 July 2024

We thank the reviewers' and the editor's comments. You'll find attached our responses to the reviewers' remarks.

Download author's reply
Download tracked changes file

Decision by Sishuo Wang, posted 06 June 2024, validated 06 June 2024

Dear Dr. Cury.

Thank you for submitting your work to PCI. I've received 3 reviewer reports, all of which agree on the point that the study is generally well done and represents an important contribution to the field. However, they also raised some questions. While most seem minor, it is important to note that they need to be addressed before the work can be accepted. To this end, I'd like to invite you to prepare a revised manuscript, taking into account the suggestions by the reviewers. I look forward to receiving the revision.

Best,
Sishuo

Reviewed by Pierre Pontarotti , 06 May 2024

This is an updated version of the highly successful 'Defense Finder,' with more phage defense systems described and the availability of new web services. The new version of Defense Finder will be an important tool for scientists interested in anti-phage defense systems

Reviewed by anonymous reviewer 1, 06 June 2024

REVIEW on
A Comprehensive Resource for Exploring Antiphage Defense: DefenseFinder Webservice, Wiki and Databases.
doi: https://doi.org/10.1101/2024.01.25.577194

Title and abstract
Does the title clearly reflect the content of the article?
Mostly, Yes. I suggest the following change since the systems do not only target phages, but also other mobile genetic elements.
"Comprehensive Resource for Exploring Prokaryotic Defense Systems: DefenseFinder Webservice, Wiki and Databases."
Does the abstract present the main findings of the study?

Yes. Slight improvements are suggested in the text below.

Introduction

Are the research questions/hypotheses/predictions clearly presented?

Yes

Does the introduction build on relevant research in the field?

Yes

Materials and methods

Are the methods and analyses sufficiently detailed to allow replication by other researchers?

Yes.

Are the methods and statistical analyses appropriate and well described?

Yes (mostly). See comments below.

Results

In the case of negative results, is there a statistical power analysis (or an adequate Bayesian analysis or equivalence testing)?

I don't know

Are the results described and interpreted correctly?

Yes. But since this work is introducing a tool, databases and an encyclopedia, only some first (broad) insights are shown.

Discussion

Have the authors appropriately emphasized the strengths and limitations of their study/theory/methods/argument?

Yes.

Are the conclusions adequately supported by the results (without overstating the implications of the findings)?

Yes.

In the here presented work, done by Tesson et al., a website has been created to improve the bioinformatic detection of antiphage systems. This includes an updated version of DefenseFinder with a web service, plus three databases (a wiki on defense systems, a structure database with experimentally determined and AlphaFold2 predicted structures, and a precomputed DefenseFinder results database).

Overall, this is a great work and contribution for the community!

The arsenal and complexity of defense systems are huge and are expanding very fast, making them hard to track. This work facilitates transparency and a fast acquisition of knowledge on the systems by providing a clear overview. The creation of wiki pages was a lot of work and is highly appreciated. The experimental validation part is very nicely presented.

I have only a few comments and minor concerns:

· How are orphans (HMM-only hits) treated or should be treating. These orphans are genes involved in systems but do not make complete ones? This was not addressed in this or the previous article(s). Please add a paragraph what would be a recommend approach to deal with them.

· What about interactions or even synergy between defense systems? And, also, how is the concept of layers of defense systems considered and could be represented?

These are just thoughts and suggestions that could be added in future versions. Please add some clarifications on these points since these topics are addressed in the community (as reflected in current literature).

Minor concerns

I tested the web service and it works well in my point of view. Is it possible to download figures of the genomic organizations (in good quality) without the need to make screenshots?

Moreover, hits on several contigs detected in a multi-FASTA file are all displayed on 'one genome arrow'. Can this be displayed on separate contigs (which visualization should be limited to 10 or an adjustable number)?

Abstract

o L.19-21: "all known antiphages defense mechanism" is in my point of view (slightly) overstated. There are also other tools that predict the presence of antiviral systems such as PADLOC (PMC9252829), which won't fully overlap with DefenseFinder, indicating that one tool cannot find all systems. I suggest to remove 'all known'.

o L. 26-29 I suggest the following change for a better readability:

"To overcome these challenges, we present a hub of resources on defense systems, including: 1) an updated version of DefenseFinder with a web-service search function, 2) a community-curated repository of knowledge on the systems, and 3) precomputed databases, which include annotations done on RefSeq genomes and structure predictions generated by AlphaFold."

Methods

o L39-42. Sentence is unclear. Verb is missing (starting from L41).

o Creation of a homepage is quite specific and I do not have the know-how to give any comments this part. The only question that I have is, if the pages will be maintained on the long-term? Is there a funding behind?

o I suggest combining protein selection and structure prediction (L.304-319) together. What does "best hit" mean? Is the selection based on the 30% identity and 70% coverage? This threshold is not very high and may be biased by different structures. In other words, what are the maximum differences between sequences, and does it make sense to take just one sequence as representative?

o Pfam annotation (L.326-333): What is meaning of superimposed in this context? Per protein or domain/sequence position?

Conclusion

o I recommend rephrasing the first section to avoid frequent repetition of 'antiphage.'

o Additionally, I suggest to add a small section describing future plans for updating and improving this platform/hub. This could include plans for exploring interactions between defense systems, conducting docking simulations using the predicted structures, providing tutorials, and even organizing workshops or events to foster collaborations in computational work on defense mechanisms.

Download the review

Reviewed by Pedro Leão, 22 May 2024

Florian Tesson and colleagues make a significant contribution to the field of defense systems by developing and providing access to three comprehensive databases. This initiative can greatly impact researchers at all levels, offering experienced individuals the opportunity to contribute to the databases' future expansion. It also serves as an accessible platform for newcomers and enthusiasts in the rapidly growing field of defense system research. Below are my minor comments and suggestions:

Line 74-75: "The information of the web server is integrated into the rest of the website." It would be interesting to explain a bit more what this means for the general public, and potential user of the website.

Line 82-83: We suggest the authors to add this sentence to the beginning of the next paragraph to improve readability.

Line 115: "We use pyrodigal v3.0.131 to identify and annotate the coding regions..." . I believe this process identify the coding regions, and translate them, not annotate. Please double check.

Line 116-117: It's not clear how these tests are making "the development of future features more robust". can you elaborate a bit more on this?

Line 222-223: "For systems and subsystems where protein accessions were impossible to retrieve, we selected another representative." Could you please make this process more clear? What protein accessions were not possible to retrieve? The experimentally validade proteins structures? proteins with experimentally

validade anti-phage function?

The same on the methods session regarding this process (Line 306-307): "For systems with no accession available". We would recommend the authors to be more precise. For systems with no experimental validation?

Line 322-323: We believe the number (n) of archaeal and bacterial genomes are inverted. Please double check it.