Response to Bioinformatics Reviewers

Our deep thanks and appreciation to the reviewers for devoting time and attention to our paper and providing valuable comments. Below, we detail our responses.

Reviewer: 1

1. **In the Abstract and Conclusions sections the authors state that pySubnetSB is computationally efficient, reducing computations "from an infeasible 10^78 mapping pairs to a more practical 10^8 mapping pairs." This statement conflicts with Section 4, where the maximum number of mapping pairs is set at 10^12. To ensure consistency, I recommend rephrasing the abstract and conclusion. Additionally, it would be helpful if the authors clarified whether the numbers 10^78 and 10^8 were estimated based on small, average, or large networks. This context is essential for assessing the computational claims.**

   Our thanks for catching this point of confusion. We have done the following to address it.

   a. We clarify the description of computation complexity in the abstract, introduction and conclusion.
      i. Abstract: “We show that pySubnetSB achieves large reductions in computational complexity for subnet discovery. For example, in studies of randomly selected target networks with 100 reactions each with a random reference network with 20 reactions, computations are reduced from an infeasible 10^78 evaluations to a more practical 10^8 evaluations.”
      ii. Introduction: “In our studies of randomly selected target networks with 100 reaction that each embed a random reference networks with 20 reactions, the number of mapping pairs is reduced from a computationally infeasible 10^{78} to a more practical 10^{8}.”
      iii. Conclusions: “In our studies of randomly selected target networks with 100 reactions that each embed a distinct reference networks with 20 reactions, the number of mapping pairs is reduced from a computationally infeasible 10^{78} to a more practical 10^{8}.”

   b. In section 4, we clarify the use of 10**12 as the complexity limit with the text: “To manage computational demands, we set the API parameter `max_num_mapping_pair` so that we evaluate at most 10^12 mapping pairs for a subnet discovery.”

2. **Section 3 appears to primarily describe the methodology rather than present results. Consequently, including such section into the "Methods" section would enhance clarity and structure**.

Our work is entirely computational and so the current work can be viewed as a “methods paper”. Until recently, we have taken the approach that you suggested--- putting the description of the technical approach in the Methods section. Recently, we received feedback from a Bioinformatics reviewer who observed that since the results are really about methodology, the results should be in the Results section. This is the approach we took in the current paper. We think this makes sense, but if you strongly disagree, we are happy to move it to the Methods section.

3. **The statistical method in Section 3.4 relies on fixed constants: 100 reference CRNs and 1000 target CRNs. These constants limit the generalizability of the statistical test. Bootstrapping or another resampling technique could offer more robust and scalable statistical validation, accommodating variable sample sizes and enhancing the reliability of the results**.

   The statistical procedure is stated in a confusing way. The constants 100 and 1000 are used to ensure that entries in Fig 9 have small variances. There is nothing special about 100 and 1000, other than these are sufficiently large to ensure small variances. The procedure has been rewritten as follows:

- Step (1): generate \revision{K^R} reference CRNs with size (M_s^R, M_r^R);
- Step (2): generate \revision{K^T} target CRNs also of size (M_s^R, M_r^R);
- Step (3): for each reference CRN in step (1), count the number of target CRNs in step (2) that are strongly structurally identical, and report the fraction of occurrences of strong structural identity.

K^R, K^R should be sufficiently large so that there is little variability for the statistics calculated in step (3). We find that it is sufficient to use $K^R = 100, K^T=1000$.

4. **The description of the BioModels database and its use in performance assessment is insufficient. The authors should provide greater detail about the dataset, including whether the entire BioModels repository was utilized or if certain filters were applied (e.g., size, completeness, or quality of models). Additionally, providing a URL or reference to the BioModels repository would improve reproducibility and transparency.**

   We appreciate this suggestion. In response, we have created `README.md` in the `data` folder of the project repository. The README identifies the BioModels models used in our study and the reason why some models were not included. The `README.md` also specifies details of data formats to help others make use of the data we created. We have added the following text to the first paragraph of section 4:

   This section studies the occurrence of subnets in the curated branch of BioModels, approximately 1,000 models. (See data/README.md in the github repository for details.)

5. **The experimental section addresses two research questions outlined in the introduction, but these are discussed together in Section 4. Separating the questions into distinct subsections would enhance clarity and focus. Moreover, the conclusions, such as "further work is required to determine if there is a research result," undermine confidence in the practical relevance of the proposed approach. Including validation with examples easily verified in existing literature would strengthen the results and demonstrate the real-world utility of the method.**

   We agree that linking the two use cases to the results is a great idea, and we have restructured the text accordingly. For the first use case, we are able to validate that mechanisms predicted by subnet discovery are present in target models. However, for the other examples (the second example in use case 1 and the example in use case 2), we chose to leave the example as a hypothesis. Our thinking here is twofold. First, generating strong hypotheses is an important part of scientific discovery, even if the hypotheses don't lead to actual results. Second, exploring these hypotheses is a significant undertaking in its own right. We have updated the future work to explore the two hypotheses discovered in this paper.

   The specific change is the last sentence of the paper:

   Longer term, we plan to explore the hypotheses identified in our discovery of subnets in BioModels.

Reviewer: 2

1. Issue in Fig. 2 on page 2 (Major)
The labels for (a) and (b) may be reversed. Please verify both Fig. 2 and the corresponding description in the text (lines 9–12 on page 2, right column).

Great catch. Thanks! The revised figure is at the end of this document.

2. Suggestion for Fig. 2(c) on page 2 (Minor)
Fig. 2(c) could be presented using the same graph format as in Fig. 3 on page 3.

This is an interesting thought, and could provide consistency between Fig. 2 and Fig. 3. However, we were concerned about the internal consistency of Fig. 2 since (a) and (b) are in a text format. In the end, we decided not to use the graph format for (c).

3. Typo (Minor)
Page 4, line 24, right column: should "1031" be "1034"?

Thanks again. It turns out this typo was present a second time as well.

4. Possible Issue (Major)
Page 4, line 36, right column: is the wrong reaction referenced? Should "J1 and J2" be "J1 and J5"?

Your suggestion makes sense. This was written thinking about arbitrary reactions, but the context wasn't clear, and it makes more sense to use the running example. The revised text is:

Consider \revision{the arc from reaction J_1 to J_5 in \fig{fig:bipartite}(b). Here, J_1 is a 1-step predecessor to J_5, and J_5} is a 1-step successor to J_1.

5. Issues and suggestions regarding Fig. 5 and reaction constraints (Major)
The current presentation of Fig. 5 may cause confusion in understanding RC2 and RC3.

> **5.1 Specifically, Fig. 5(c) could be misinterpreted if viewed only alongside Fig. 5(a) and Fig. 5(b), which list only 1-step successor counts. According to the specifications described in lines 41–44 of the right column on page 4, Fig. 5(c) reflects both 1-step and 2-step predecessor counts (RC2), as well as 1-step and 2-step successor counts (RC3).**
> **Suggestion: The authors could either explicitly explain this in the figure/table description, or provide complete information on both 1-step and 2-step counts.**
>
> We have revised the caption for Fig. 5 to address these concerns. The revised text is:
>
> Illustration of calculating} reaction mappings. (a) and (b) are partial data for reaction constraint matrices (reaction type, 1-step successor counts). (c) is the reaction compatibility \revisionb{calculated from (a) and (b)}. (d) displays three reaction mappings (the columns) \revision{of the many reaction mappings; the reaction mappings are based on (c) and are obtained by selecting a unique 351 reaction in the list of compatible reactions for each 1034 reaction.
>
> **5.2 Fig. 5(d) lists 3 examples from all qualified mappings.**
> **Suggestion: The authors could clarify the table subsection with a description such as "3 examples from all." The corresponding text in line 43 of the left column on page 4 may be revised accordingly.**
> **Also, one example could use different mappings such as J1(v11) and J4(v10) to avoid a misimpression that J1 can only map to v10.**
>
> The figure and caption and related text have been changed. The caption is given above.
>
> The new text on page 4 is:

Fig. 5(c) displays the reaction compatibility vector using the constraints in Fig. 5(a) and Fig. 5(b). Fig. 5(d) shows three of the many reaction mappings obtained from Fig. 5(c) by selecting a distinct 351 reaction in the list of compatible reactions for each 1034 reaction.

The revised figure is at the end of this document.

6. Fig. 10 on page 7 (Major)
What is the definition of "induced subnet"? I didn't find it if I didn't miss it.

Another good catch. We had changed my terminology from "induced subnet" to "inferred network" but failed to change a couple of instances of the former. Thanks!

7. pySubnetSB installation and its examples from GitHub (Major)
The availability of the application implementing the methodology, along with its source code, is essential for enabling the community to use and extend this work. The authors mention running the program on a 128 GB, 20-core Mac Studio M1. I do not have a machine with similar specifications, so I attempted to run and test the application on the Ubuntu platform. Following the instructions provided in the manuscript and on the GitHub site, my attempts were unsuccessful.

Attempt 1: Ubuntu 22.04, Python 3.12.7 under Anaconda
*** Installation via pip was successful; pySubnetSB version 1.0.2**
*** Running api_basics.ipynb (downloaded from the GitHub site) triggered a compatibility error: NumPy 2.2.6 was incompatible with other libraries built for NumPy 1.***
*** After downgrading to NumPy 1.25.4, memory_profiler was not found by network.py**
*** Stopped testing**

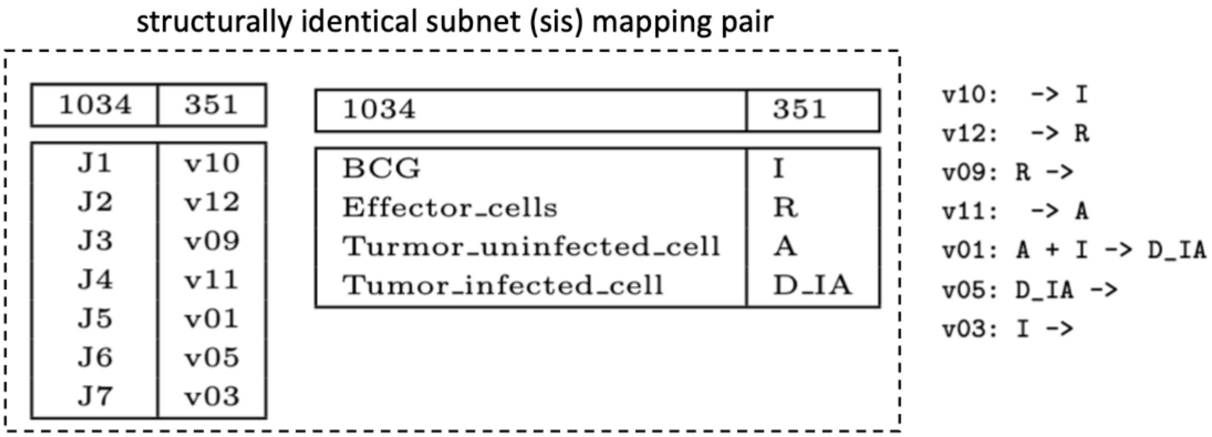
Attempt 2: Ubuntu 20.04, Python 3.8
*** Installation via pip was successful; pySubnetSB version 0.0.1**
*** Running api_basics.ipynb raised an error: no module named "src" in cluster_builder.py**
*** After removing "src", a new error appeared: "type" object is not subscriptable in network.py**
*** Stopped testing**

Strong suggestion: The authors should conduct comprehensive tests on commonly used platforms such as Ubuntu and document all supported environments and configurations clearly on the GitHub repository. Limited platform compatibility will significantly reduce the usability and impact of this work.

Our apologies for not doing a better job of testing the `pySubnetSB` package. We have over 200 unit tests. But we failed to adequately test the package itself. They were simple bugs, and they should have been found before we submitted the paper.

We have taken your recommendations to heart. We used github actions to test on three platforms: Ubuntu (22.04), Windows (10), and Mac OS (14.7.6). For each, tests were run for python 3.9, 3.10, 3.11, and 3.12. You note testing on Ubuntu 20.04 with Python 3.8. These are much older releases. For example, it seems that the installer needed to select a very old version of pySubnetSB (0.0.1) to be compatible with Python 3.8. We have updated our README and the package information to require Python >= 3.9. Last, we have greatly improved the programmatic (non-notebook) example of pySubnetSB in https://github.com/ModelEngineering/pySubnetSB/blob/main/examples/api_basics_programmatic.py.

The revised Fig 2 is:



structurally identical subnet (sis) mapping pair

1034	351
J1	v10
J2	v12
J3	v09
J4	v11
J5	v01
J6	v05
J7	v03

(a) reaction mapping

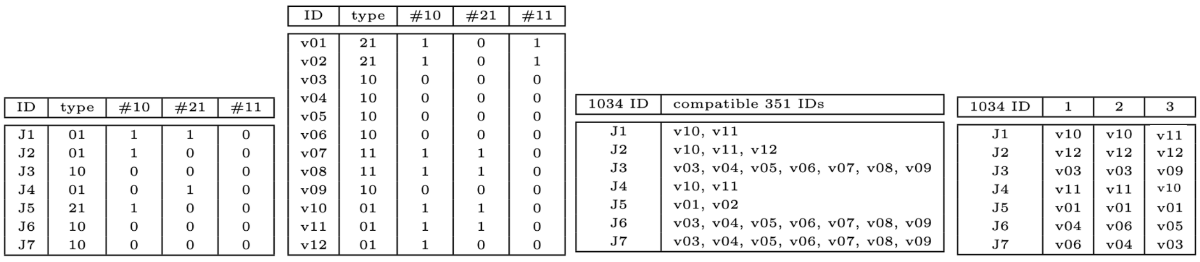
1034	351
BCG	I
Effector_cells	R
Turmor_uninfected_cell	A
Tumor_infected_cell	D_IA

(b) species mapping

```
v10:  -> I
v12:  -> R
v09: R ->
v11:  -> A
v01: A + I -> D_IA
v05: D_IA ->
v03: I ->
```

(c) structurally identical inferred subnet

The revised Fig 5 is



ID	type	#10	#21	#11
J1	01	1	1	0
J2	01	1	0	0
J3	10	0	0	0
J4	01	0	1	0
J5	21	1	0	0
J6	10	0	0	0
J7	10	0	0	0

(a) Model 1034: reference reaction constraint matrix

ID	type	#10	#21	#11
v01	21	1	0	1
v02	21	1	0	1
v03	10	0	0	0
v04	10	0	0	0
v05	10	0	0	0
v06	10	0	0	0
v07	11	1	1	0
v08	11	1	1	0
v09	10	0	0	0
v10	01	1	1	0
v11	01	1	1	0
v12	01	1	0	0

(b) Model 351: target reaction constraint matrix

1034 ID	compatible 351 IDs
J1	v10, v11
J2	v10, v11, v12
J3	v03, v04, v05, v06, v07, v08, v09
J4	v10, v11
J5	v01, v02
J6	v03, v04, v05, v06, v07, v08, v09
J7	v03, v04, v05, v06, v07, v08, v09

(c) reaction compatibility vector

1034 ID	1	2	3
J1	v10	v10	v11
J2	v12	v12	v12
J3	v03	v03	v09
J4	v11	v11	v10
J5	v01	v01	v01
J6	v04	v06	v05
J7	v06	v04	v03

(d) reaction mappings