F1000Research

Check for updates

SOFTWARE TOOL ARTICLE

REVISED CellMap visualizes protein-protein interactions and subcellular localization [version 2; peer review: 2 approved]

Christian Dallago [iD] [1], Tatyana Goldberg[1,2], Miguel Angel Andrade-Navarro[3], Gregorio Alanis-Lobato[3], Burkhard Rost[1,4-6]

[1]Department of Informatics, Bioinformatics & Computational Biology , TUM (Technical University of Munich), Munich, Germany
[2]Center for Doctoral Studies in Informatics and its Applications (CeDoSIA), TUM (Technical University of Munich) Graduate School, Munich, Germany
[3]Faculty of Biology, Johannes Gutenberg University Mainz, Mainz, Germany
[4]Department of Biochemistry and Molecular Biophysics, Columbia University, New York City, NY, USA
[5]Institute for Food and Plant Sciences WZW, Freising, Germany
[6]Institute for Advanced Study (TUM-IAS), Munich, Germany

Abstract
Many tools visualize protein-protein interaction (PPI) networks. The tool introduced here, CellMap, adds one crucial novelty by visualizing PPI networks in the context of subcellular localization, i.e. the location in the cell or cellular component in which a PPI happens. Users can upload images of cells and define areas of interest against which PPIs for selected proteins are displayed (by default on a cartoon of a cell). Annotations of localization are provided by the user or through our in-house database. The visualizer and server are written in JavaScript, making CellMap easy to customize and to extend by researchers and developers.

Keywords
subcellular location, biological visualization, protein-protein interaction

iSCB
INTERNATIONAL SOCIETY FOR COMPUTATIONAL BIOLOGY

This article is included in the International Society for Computational Biology Community Journal gateway.

Open Peer Review

Reviewer Status ✔ ✔

	Invited Reviewers	
	1	**2**
REVISED version 2 published 01 Feb 2018	✔ report	
	↑	
version 1 published 11 Oct 2017	? report	✔ report

1 **Sandra Orchard** [iD] , European Bioinformatics Institute (EMBL-EBI), European Molecular Biology Laboratory, Cambridge, UK

2 **Augustin Luna** [iD] , Harvard Medical School, Boston, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Christian Dallago (christian.dallago@tum.de)

Author roles: Dallago C: Resources, Software, Visualization, Writing – Original Draft Preparation; **Goldberg T**: Conceptualization, Data Curation, Supervision, Validation, Visualization, Writing – Review & Editing; **Andrade-Navarro MA**: Data Curation, Writing – Review & Editing; **Alanis-Lobato G**: Data Curation, Writing – Review & Editing; **Rost B**: Conceptualization, Data Curation, Funding Acquisition, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

How to cite this article: Dallago C, Goldberg T, Andrade-Navarro MA *et al.* **CellMap visualizes protein-protein interactions and subcellular localization [version 2; peer review: 2 approved]** F1000Research 2018, **6**:1824 (https://doi.org/10.12688/f1000research.12707.2)

First published: 11 Oct 2017, **6**:1824 (https://doi.org/10.12688/f1000research.12707.1)

Introduction

Many tools visualize different aspects of protein-protein interaction (PPI) networks; the most prominent might be Cytoscape[1]. Existing visualizations of large PPI networks continue to be difficult to use. Some proteins interact with many hundreds or thousands of others. Often referred to as 'PPI hairballs', such hubs are in the way of understanding large data sets. Many ways have been proposed to resolve such hairballs through the addition of biologically meaningful dimensions such as pathways[2] or time[3].

Another dimension was first introduced a decade ago, namely the overlay of PPIs with subcellular localization[4]. Combining PPI networks with protein location provide an intuitive way of laying out PPI networks on a graphical representation of the cell, and might reduce the clutter from PPI hairballs. This decade-old solution[4] no longer copes with today's data, in terms of scalability nor of customizability and in terms of ease-of-use.

CellMap, the prototype introduced here, takes up on the idea of PPI visualization constrained by protein location, and provides a simple visual interface for users to explore protein location inside a cell. It presents this information in a graphically pleasant way and offers several customization features. The framework has been optimized to simplify future developments, such as the addition of further data dimensions (e.g. inclusion of protein trafficking). An instance of the tool with localization data from a previous publication that includes protein localizations of the human proteome[5] and PPI data from the Human Integrated Protein-Protein Interaction rEference (HIPPIE) resource[6] is available at http://cell.dallago.us.

Methods

Implementation

The CellMap prototype is an integrated portal that exposes API calls to retrieve images (representing cells) and protein information, as well as a frontend to visualize protein location and PPI data. The portal is fully written in JavaScript, namely in the JavaScript interpreter node.js (https://nodejs.org) for the backend and vanilla JavaScript for the frontend. The portal is deployed to the public through a Docker container. Docker is a technology that allows shipping of packaged services such as web applications to customers and users without the need to install dependencies other than the Docker engine (available through: https://www.docker.com). For the representation of cell images as maps, the Leaflet framework is used. Leaflet is a JavaScript-based tool used to represent maps (http://leafletjs.com).

Data about proteins are stored as JSON documents in a Mongo (http://mongodb.com) database. All information about the interaction partners and the subcellular localization of a protein is stored in a single JSON document, making the data structure simple to understand for non-experts and enabling them to deploy prototypes using their own data. Figure 1 schematically represents a protein data model (for a specific example for a protein object: http://cell.dallago.us/api/proteins/search/Q99943).

Operation

In CellMap, users can choose to upload new maps (images of cells). They can modify the location of regions of interest (ROIs) for a selected map (Figure 2), and visualize the locations of selected proteins on a map or render protein-protein interaction networks from a set of selected proteins.

To maintain a consistent coloring scheme for different cellular compartments throughout a set of different images, each compartment is assigned a unique color through the hash of the compartment's name (e.g. light blue = vacuole, Figure 3B). Using this coloring approach, users might eventually learn to associate color with compartment. When proteins are loaded into the map, they are assigned pseudo-random coordinates representing a point that lies within the boundaries of the ROI in which they are localized (Figure 3D). A circle of a given radius is placed on the randomly generated point (Figure 3E-F), and the circle will be filled with the same color as the compartment in which the protein is located in (Figure 3B and 3F).

Users can choose between two visualization options: the subcellular location in the context of the protein-protein interaction
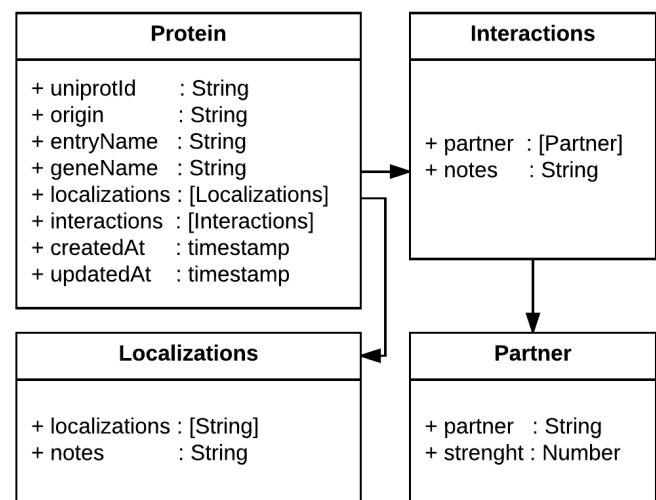

Figure 1. Diagram of the data representation in CellMap. In the figure we present a diagram of the Protein class, which contains several attributes of type *String*, two fields of type *timestamp* and two arrays (in square brackets) that reference the *Interactions* and *Localizations* classes. The arrows highlight the referenced models. This simple representation of information about a protein, its protein-protein interaction partners and its localizations enables the tool to be reused with one's own datasets.

Figure 2. Section of a screenshot of the CellMap editing tool on a private instance of the portal. In the screenshot, an authorized user with editing capabilities draws a polygon (dark green) representing a new cellular compartment or region of interest (ROI). The user has a set of tools on the left side that can be used to draw polygons, lines, squares or circles. Once the new region has been drawn, the user can associate a cellular compartment through the dropdown input on the top-right and submit the new information to the server. The image used for this screenshot was taken from Wikimedia's user Royroydeb, under CC BY-SA 4.0 (http://bit.ly/2fuYRiE) and is used in this figure for demonstrative purposes only, as using it on the online version of CellMap would infringe copyright.
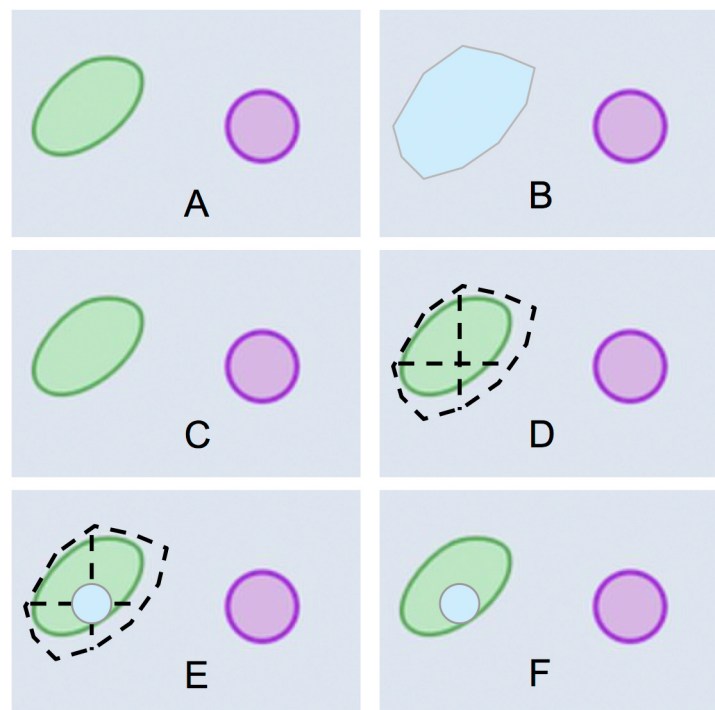

Figure 3. Definition of an area and drawing of protein circle. (**A**) Section of a cartoon image of a cell; (**B**) user-drawn polygon representing the area occupied by a vacuole; (**C**) how the section of the cartoon image is displayed on the PPI/map viewer; (**D**) random point calculation inside vacuole-polygon-defined area; (**E**) drawing of a protein circle located inside the vacuole, (**F**) result of loading a protein localized in the vacuole as shown by PPI/map viewer.

viewer (PPI viewer, Figure 4A, http://cell.dallago.us/ppi), and the protein subcellular location viewer (Map viewer, Figure 4B, http://cell.dallago.us/map). The two viewers can load the same images of cells (maps) and collect localization data from the same source, in the publicly available instance by 5. The PPI viewer offers the possibility to overlay networks between proteins being visualized. The map viewer displays all locations reported for a given protein simultaneously, while the PPI viewer only displays only one location at a time (by default: the first localization in the array of localizations as described in the protein data model, Figure 1); users can manually change the location by clicking on the protein circle and selecting a new location from the information box (Figure 5). Both the PPI and the map viewer are enriched by several controls (Figure 6): The top-left controls enable actions including: the navigation to the home of CellMap (Figure 6, panels 1 and 2, A), switching from the map viewer to the PPI viewer and vice versa, keeping the proteins currently loaded in the view (Figure 6, panels 1 and 2, B), reducing the opacity of the cell map, highlighting the protein circles (Figure 6, panels 1 and 2, C), zooming in- and out of the map and PPI viewers (Figure 6, panels 1 and 2, D), and visualization of the global network among all proteins loaded in the visualizer (Figure 6, panel 1, E). The top-right control allows to temporarily hide loaded proteins or activate an overlay of the user-drawn localizations (Figure 6, panel 4). The top-center search panel allows users to load new proteins by searching for their UniProt identifier, primary gene or primary protein name[7] into the viewer (Figure 6, panel 3).

To facilitate the retrieval of proteins and their interacting partners, CellMap provides basic search functionalities. Users can search for proteins based on their UniProt identifiers, by their gene identifiers or by their protein names. When performing the search, the page renders a grid containing boxes, each representing a different protein (Figure 7). Inside the boxes, the UniProt identifier
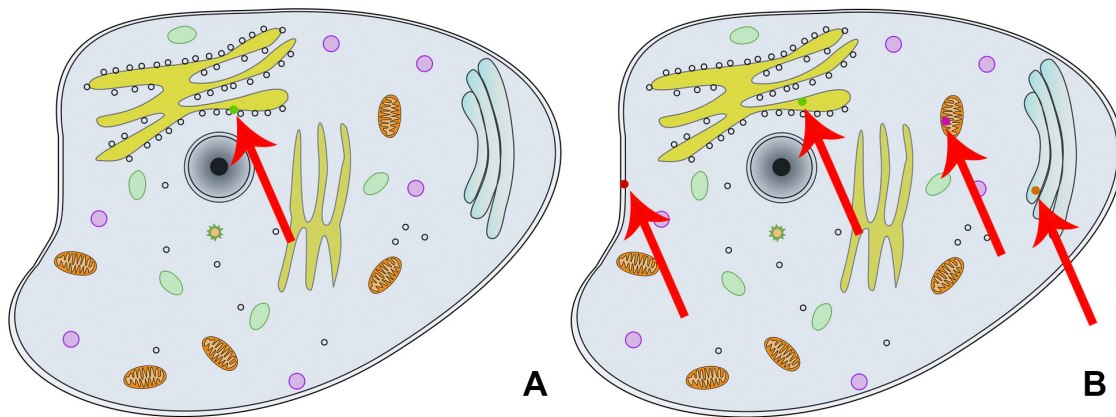

Figure 4. Comparison between PPI viewer and map viewer. The left view (**A**) shows the PPI viewer, which depicts the result of loading protein Q9NR71 and displays a circle for the first localization found in the array of locations (http://cell.dallago.us/ppi?p=Q9NR71); The right panel (**B**) shows the Map viewer, which depicts the result of loading the same protein Q9NR71 and displays a circle for the protein in each of its reported location (http://cell.dallago.us/map?p=Q9NR71). The red arrows are overlaid on top of the screenshots to highlight where the protein circles have been drawn in the viewers, since fitting the screenshot on the page reduces the overall size of the images.
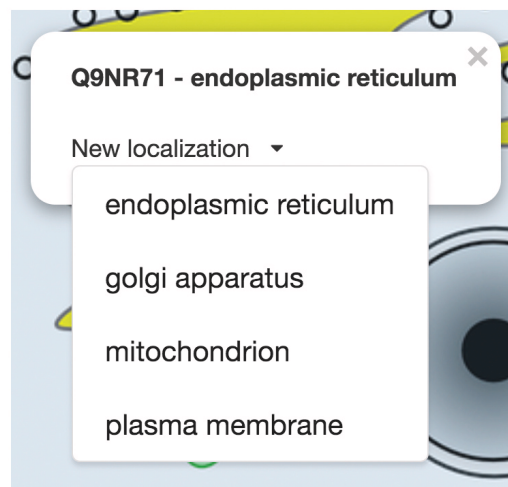

Figure 5. Protein information box. Top: information about the selected protein. Bottom: new localization selection box rendered in the PPI viewer when clicking on the protein circle (http://cell.dallago.us/ppi?p=Q9NR71).

Figure 6. Controls used in the different viewers. (**1**) Top-left controls of PPI viewer; (**2**) top-left controls of map viewer; (**3**) top-center search panel of PPI/map viewer; (**4**) top-right layer control on PPI/map viewer.
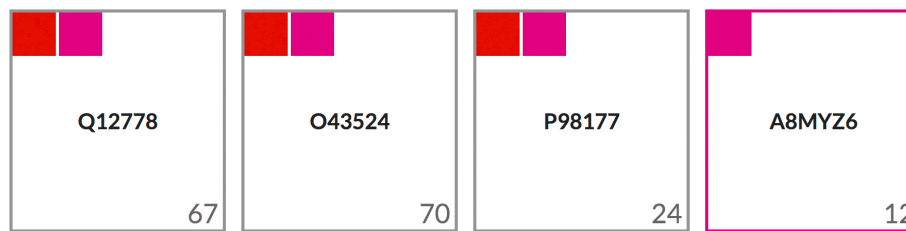

Figure 7. Results of searching for protein "*foxo*". The screenshot of this section of the home page shows four proteins that match the search criterion "*foxo*" either by their UniProt identifier, primary gene name or primary protein name. The protein boxes contain the UniProt identifiers of the matched proteins (center) and display the number of interaction partners (bottom-right) and several color-filled boxes graphically representing the localizations reported for the matched proteins (top-left).

for the protein that matched the search criterion is displayed. Starting on the top-right of every box a smaller colored square for each compartment is displayed in which that protein is localized. For proteins annotated to be in a single compartment, the border of the outer box (representing one protein as indicated by the UniProt ID in the center of the box) will get the color of that compartment (2nd box in Figure 7). Clicking on one of the colored squares will filter results based on the compartment represented by that color. In the bottom-right of each box, the total number of PPI partners are annotated.

Discussion

Some CellMap functionality is exemplified by a heat shock protein (HSPA4; Heat shock 70 kDa protein 4, UniProt identifier P34932) with many interaction partners (338, according to HIPPIE, http://cbdm-01.zdv.uni-mainz.de/~mschaefer/hippie/query.php?s=HSPA4) in different compartments. The objective was to showcase how CellMap can simplify PPI hairballs. We visualize the same PPI network using CellMap (Figure 8A) and Cytoscape[1] in the form of the Cytoscape.js version used by HIPPIE (Figure 8B) and the Cytoscape desktop version (Figure 8C).
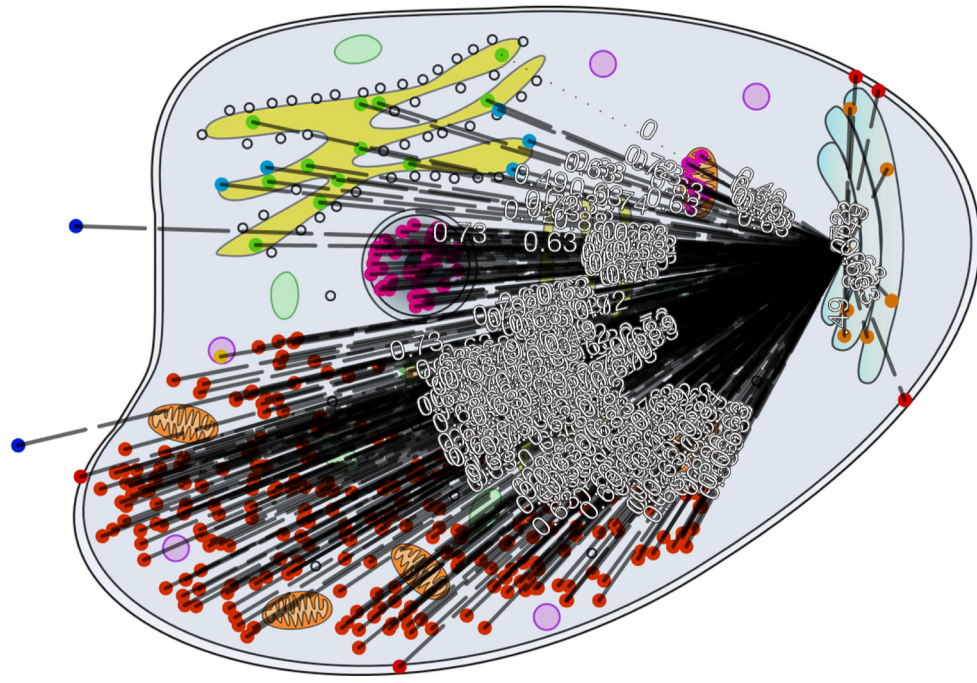
None of the three viewers solves the PPI hairball problem completely. Without zooming in, the information density for

338 protein pairs is too high to be helpful. HIPPIE's layout for Cytoscape.js (Figure 8B) clearly improves over the standard Cytoscape desktop version (Figure 8C) by centering the view around HSPA4, the protein of interest. In CellMap (Figure 8A) the biologically relevant differences between pairs from the same and from different compartments remain visible.
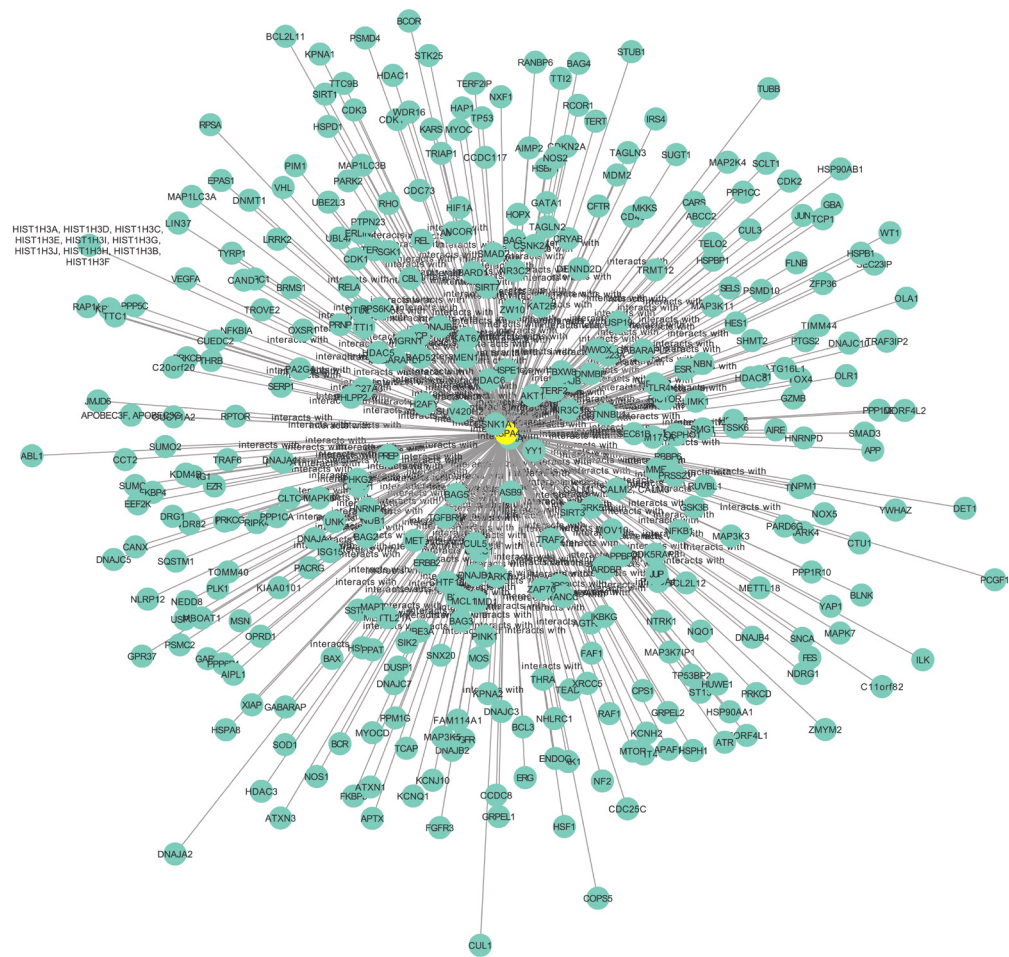
By using a biologically relevant dimension (protein localization), instead of drawing nodes in positions based on edge weight (force layout of Cytoscape), some aspects of the protein and its partners become obvious at first glance, e.g. that HSPA4 interacts with many nuclear and cytoplasmic proteins, as well as with proteins that are secreted (extra-cellular) and located in the Endoplasmic Reticulum (ER, Figure 8). This may suggest the hypothesis HSPA4 to be an important hub involved in process spanning across compartments. Such a hypothesis is presented in our supplementary material (Figure SOM_1), where we analyze the visualization of the FOXO3 protein through CellMap.

One disadvantage of CellMap over the Cytoscape.js view is that the protein identifiers are not visible at all on the static image (protein identifiers become visible through mouse-over events in CellMap). However, in the image shown (Figure 8) the Cytoscape.js names also remain unreadable. Another problem with CellMap are the numbers displayed on edges (experimental
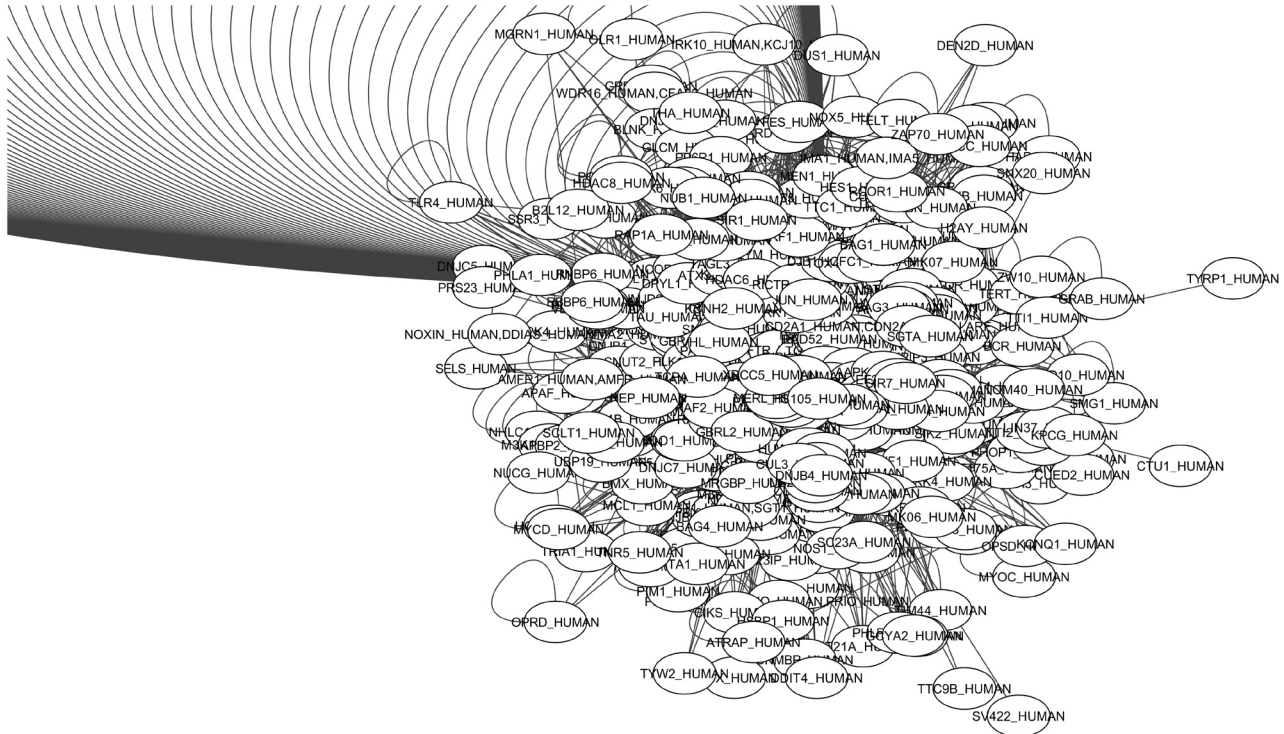
A

B

C

Figure 8. PPI hub in CellMap (**A**), Cytoscape.js (**B**) and Cytoscape desktop (**C**). For HSPA4 (Heat shock 70 kDa protein 4, UniProt identifier P34932), we show some of the PPIs known (according to HIPPIE HSPA4 has 338 interaction partners). We chose this as one example of a protein with many more PPIs than the average protein ("PPI hub"). The figure compares how three different PPI viewers cope with the HSPA4 network: (**A**) CellMap (http://cell.dallago.us/protein/P34932), (**B**) HIPPIE's Cytoscape.js visualizer and (**C**) the desktop version of Cytoscape. Proteins in CellMap are represented as colored dots on the map (image) of the cell, and upon selecting the protein of interest an overlay of edges is drawn. In Cytoscape and Cytoscape.js, proteins are represented as nodes containing a label (protein name as UniProt identifier), and edges are directly inferred from the data. The Cytoscape.js visualization was taken directly from HIPPIE. The Cytoscape network was automatically drawn upon loading the HIPPIE dataset and selecting the protein of interest and it's direct neighbors.

reliability of the PPI as given by HIPPIE). In our view, this information is extremely important to look at interactions, but we are still lacking a more sophisticated mechanism to visualize these numbers.

CellNetVis[8] is a recent tool that also connects localization with PPI networks. It emphasizes the way PPI networks are laid out through the adaptation of a so-called force-directed layout (using the tool While). Although CellMap and CellNetVis are founded on a similar idea, user experience and focus differ importantly. For instance, CellMap can be driven by data from users that define the number of compartments on a map, and provide localizations. In contrast, CellNetVis uses a fixed subset of compartments and an ad hoc diagram for the cell. Additionally, CellMap comes with out of the box data for the human proteome and allows the community to grow the tool by enriching datasets (images and localizations), whereas CellNetVis has a per-use approach,

allowing to visualize networks stored in specialized XGMML files. Another unique aspect of CellMap is the openness to introduce further biologically meaningful dimensions (beyond location such as time or pathways) that increase the usefulness of PPI visualization tools to create new testable hypotheses.

Conclusions

CellMap is a prototype providing a portal exploring the idea of using protein subcellular location as the basis to construct more complete visualizations of biological data, such as protein-protein interactions (PPIs). Using this paradigm, we claim that additional information, such as pathways, can be layered on top of the current visualization of subcellular location to potentially generate meaningful biological insights. The source code for the portal is publicly available and an instance of the portal with location data from a previous publication about the subcellular localization of the human proteome[5] and protein-protein interaction data from

HIPPIE[6] (http://cbdm-01.zdv.uni-mainz.de/~mschaefer/hippie) is running at http://cell.dallago.us. The visualization tool is written in JavaScript, thereby tapping into a very large user base for customized extensions and modifications. With the release of the prototype, we aim at creating a user base and awareness of the tool, ultimately collecting precious feedback from experimentalists and technical users alike.

Abbreviations

2D: two dimensions, API: Application Program Interface, ID: identifier, JSON: JavaScript Object Notation, PPI: protein-protein interaction, ROI: region of interest.

Software availability

The CellMap prototype is released as open source software under the GNU General Public License v3.0. Documentation, source code and viewer are available at https://github.com/sacdallago/cellmap. Archived source code as at the time of publication is available at https://doi.org/10.5281/zenodo.904324[9]. An example of use with protein localization data from a recent publication[5] and from the HIPPIE database of protein-protein interactions[6] is available at http://cell.dallago.us.

Competing interests

No competing interests were disclosed.

References

1. Shannon P, Markiel A, Ozier O, *et al.*: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res.* 2003; **13**(11): 2498–2504.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

2. Chaurasia G, Malhotra S, Russ J, *et al.*: **UniHI 4: new tools for query, analysis and visualization of the human protein-protein interactome.** *Nucleic Acids Res.* 2009; **37**(Database issue): D657–D660.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

3. Ma DK, Stolte C, Krycer JR, *et al.*: **SnapShot: Insulin/IGF1 Signaling.** *Cell.* 2015; **161**(4): 948–948.e1.
    PubMed Abstract | **Publisher Full Text**

4. Ofran Y, Yachdav G, Mozes E, *et al.*: **Create and assess protein networks through molecular characteristics of individual proteins.** *Bioinformatics.* 2006; **22**(14): e402–e407.
    PubMed Abstract | **Publisher Full Text**

5. Ramilowski JA, Goldberg T, Harshbarger J, *et al.*: **A draft network of** ligand-receptor-mediated multicellular signalling in human. *Nat Commun.* 2015; **6**: 7866.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

6. Alanis-Lobato G, Andrade-Navarro MA, Schaefer MH: **HIPPIE v2.0: enhancing meaningfulness and reliability of protein-protein interaction networks.** *Nucleic Acids Res.* 2017; **45**(D1): D408–D414.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7. The UniProt Consortium: **UniProt: the universal protein knowledgebase.** *Nucleic Acids Res.* 2017; **45**(D1): D158–D169.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

8. Heberle H, Carazzolle MF, Telles GP, *et al.*: **CellNetVis: a web tool for visualization of biological networks using force-directed layout constrained by cellular components.** *bioRxiv.* 2017.
    Publisher Full Text

9. Dallago C: **CellMap: open software for PPI and protein localization visualization in JavaScript.** *Zenodo.* 2017.
    Data Source

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 12 February 2018

https://doi.org/10.5256/f1000research.15085.r30439

✔ **Sandra Orchard** (iD)

European Bioinformatics Institute (EMBL-EBI), European Molecular Biology Laboratory, Cambridge, UK

The authors have addressed my concerns and I have no further comments to add.

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 18 January 2018

https://doi.org/10.5256/f1000research.13762.r29754

✔ **Augustin Luna** (iD)

Memorial Sloan-Kettering Cancer Center, New York, NY, USA

The tool provides an interesting feature to help declutter visualizations of biological networks using localization information. Some comments:

- It would be good if the names of the used databases was stated in the last paragraph of the introduction.

- The tool would be more intuitive for new users, if it provided descriptions the various colors used on the site with the same explanation as in the paper. For example, the colored boxes that

represent localizations in the search results and the dot colors used for the protein visualization on the cell map.

- It is unclear from the paper all the types of interactions might be shown in the represented networks.

- Also, it is unclear from the paper, what happens to the network visualization in the cases where the identified proteins are present in multiple locations.

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?
Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?
Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 29 Jan 2018

Christian Dallago, TUM (Technical University of Munich), Munich, Germany

Dear Dr. Luna,

Thank you very much for your input on our work.

We have submitted a new version of the manuscript, which should address points one and four of your comment.

As to the second point: we have created a new feature item for our next release that displays a button on the map viewer to display a modal with the legend. As of now: a legend is available by scrolling down to the second half of the page in the map or ppi viewers (e.g. http://cell.dallago.us/map/573c87c182a9e1ae1e37d08e?p=P04637) and expanding the "Legend" tab. We understand that this can be overseen and improved, therefore we thank you for the input.

As to the third point: in this manuscript, we focus on discussing the software implementation and visualization abilities of CellMap, rather than the data sources used in the example deployment hosted on http://cell.dallago.us. More information about the types of interactions reported by the HIPPIE data source can be found in the latest paper describing HIPPIE (http://nar.oxfordjournals.org/content/early/2016/10/28/nar.gkw985) and directly on the HIPPIE information page (http://cbdm-01.zdv.uni-mainz.de/~mschaefer/hippie/information.php#sources).

Please, feel free to suggest any other changes to both our manuscript and tool.

Best regards,
Christian Dallago.

Competing Interests: No competing interests were disclosed.

Reviewer Report 13 November 2017

https://doi.org/10.5256/f1000research.13762.r27544

?   **Sandra Orchard** 🆔
European Bioinformatics Institute (EMBL-EBI), European Molecular Biology Laboratory, Cambridge, UK

The authors describe a tool for visualising PPI networks in the context of the subcellular localisation of the searched protein.

1. I have twice tried to review this paper and both times the http://cell.dallago.us link gave me a MongoDB error. I have therefore had to review the paper without being able to view the tool. This is not satisfactory. I was unable to test the conclusions about the tool and its findings.

2. The tool uses a static interaction compilation database (HIPPIE) as the source of PPIs. Did the authors not consider using the PSICQUIC web service, which gives the users considerably more options as to where to source their PPI data from, and also allows the visualisation of protein-small molecule interactions and also potentially the site of action of protein-drug interactions, also available via PSICQUIC. It would also allow the data to be as up to date as the latest release of each database, which will be more frequent than releases of HIPPIE.

3. I am not clear where the subcellular location data comes from. This may be obvious to regular users of CellMap but not to me, an should be stated in the paper for other user who do not know this.

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?
Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?
Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?
Partly

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 25 Nov 2017

Christian Dallago, TUM (Technical University of Munich), Munich, Germany

Dear Dr. Orchard,

Thank you for your valuable input on the tool!

With regard to point 1: we apologize for the broken database connection, unfortunately, the deployment system missed that flag and thus didn't restart the service. We have fixed the issue and the website is now running. Up until now, I have not identified any other issues that could prevent the web server to run properly.

With regard to point 2: the software presented in this paper has a dual-purpose. On the one hand, we want to give the ability to discover protein localization and protein-protein interaction from two known sources (HIPPIE for PPI, and subcellular localization from a publication, which describes localization for the human proteome based on a consensus of experimental data and state-of-the-art prediction models (http://doi.org/10.1038/ncomms8866)). On the other hand, we want to propose a system that can be reused on user-defined data (as long as it complies with the format the visualization tool digest, as from Figure 1) and be integrated as JavaScript visualization tool in different portals. For now, we would like to avoid having a direct integration of the portal with external tools via, for example, API calls. In an upcoming version of the portal, we will offer scripts to populate the database from different sources for the two data entities (protein localization and interaction).

PSICQUIC generates interaction data on-demand, which can later be downloaded. Obtaining the data requires some time: a user input one specific protein identifier, selects the databases to use to collect interaction data, submits a cluster job and finally gets access to the data. Searching for protein P45381 identified 80 interactions in all online databases. After several hours, the job was

not finished, so we decided to lower the number of databases to fetch information from. Reducing the number of databases produced results quickly. The results page of PSICQUIC presents a table of interactions and visualizes a graph, which we could not load due to lack of compatibility with the Chrome browser. We believe it would be interesting to present CellMap at the level of this resource and will contact the authors of the tool to discuss what the best idea in this regard would be. Fetching the data from PSICQUIC as it is now and putting it into the portal requires to also normalize the PSICQUIC data and map it to protein localization data. Writing a parser for the PSI-MITAB tables is straightforward, the normalization and mapping of identifiers should occur externally to CellMap. We will create a guide on how this can be done in the next days and put it on the landing page of CellMap.

Integrating protein-molecule data and displaying these entities meaningfully is an interesting idea for the future development of the CellMap tool.

With regard to point 3: the data about protein localization stems from a publication of our group (http://doi.org/10.1038/ncomms8866). The data on protein subcellular localization for humans published through this paper was the starting point for the development of CellMap. In the current manuscript, we focused more on describing the visualization tool, rather than going into detail about how the localization data was retrieved (which in this case is by building a consensus over experimental (where available) and predicted localisations for 6 subcellular compartments). This is again because we didn't want to develop a tool around this specific data source, but rather offer the possibility to change the origin for the localization data in the future.

We appreciate the suggestions for further data sources and data entities that can be used and integrated into CellMap. In upcoming releases, we will make sure to offer a bigger variety of data sources and scripts to populate and update the information on protein subcellular localization, and protein-protein interaction data used by the visualization tool. Additionally, we will contact the authors of PSICQUIC to discuss if it would be possible to integrate CellMap in the results page of a cluster job.

Best regards,
Christian Dallago, Tatyana Goldberg & Burkhard Rost.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias

- You can publish traditional articles, null/negative results, case reports, data notes and more

- The peer review process is transparent and collaborative

- Your article is indexed in PubMed after passing peer review

- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research