

Systems biology

# Web-based network analysis and visualization using CellMaps

Francisco Salavert<sup>1,2</sup>, Luz García-Alonso<sup>1,†</sup>, Rubén Sánchez<sup>1</sup>, Roberto Alonso<sup>1,3</sup>, Marta Bleda<sup>4</sup>, Ignacio Medina<sup>5</sup> and Joaquín Dopazo<sup>1,2,6,\*</sup>

<sup>1</sup>Computational Genomics Department, Centro de Investigación Príncipe Felipe (CIPF), Valencia 46012, Spain, <sup>2</sup>Bioinformatics of Rare Diseases (BIER), CIBER de Enfermedades Raras (CIBERER), Valencia, Spain, <sup>3</sup>Bull-CIPF, Computational Genomics Chair, Valencia 46012, Spain, <sup>4</sup>Department of Medicine, University of Cambridge, School of Clinical Medicine, Addenbrooke's Hospital, Hills Road, Cambridge CB2 0QQ, UK, <sup>5</sup>HPC Service, UIS, University of Cambridge, Cambridge, UK and <sup>6</sup>Functional Genomics Node, (INB, PRB2, ISCIII) at CIPF, Valencia 46012, Spain

\*To whom correspondence should be addressed.

<sup>†</sup>Present address: European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Cambridge, UK  
Associate Editor: Janet Kelso

Received on November 16, 2015; revised on May 18, 2016; accepted on May 23, 2016

## Abstract

**Summary:** CellMaps is an HTML5 open-source web tool that allows displaying, editing, exploring and analyzing biological networks as well as integrating metadata into them. Computations and analyses are remotely executed in high-end servers, and all the functionalities are available through RESTful web services. CellMaps can easily be integrated in any web page by using an available JavaScript API.

**Availability and Implementation:** The application is available at: <http://cellmaps.babelomics.org/> and the code can be found in: <https://github.com/openncb/cell-maps>. The client is implemented in JavaScript and the server in C and Java.

**Contact:** [jdopazo@cipf.es](mailto:jdopazo@cipf.es)

**Supplementary information:** [Supplementary data](#) are available at *Bioinformatics* online.

## 1 Introduction

Networks are becoming increasingly important to understand the relationships between genotype and phenotype, specifically in the context of disease. Consequently, the demand for efficient network visualization and analysis tools, which can interoperate with modern data analysis environments, such as the web, is increasing in parallel. Most of the available web-based network analysis portals often offer limited functionality, typically restricted to network visualization and, in some cases, simple analysis options, such as GeneMania (Warde-Farley *et al.*, 2010) or STRING (Franceschini *et al.*, 2013). Other web tools are available, but based on Java, such as Ondex Web (Taubert *et al.*, 2014) or Visant (Hu *et al.*, 2013). Very recently, a JavaScript library with Cytoscape functionalities has been made available (Franz *et al.*, 2015).

Here, we present CellMaps, an open-source HTML5 web-based application that brings advanced visualization and network analysis to end-users. Additionally, a JavaScript API was developed to facilitate the integration of CellMaps in any web page. CellMaps offers the interactivity and the graphical advanced features typical for desktop applications and the performance of client-server applications.

## 2 Implementation

CellMaps server code is written in Java and C. RESTful web services and database queries have been implemented in Java. CellMaps client has been implemented in JavaScript using the HTML5 and SVG standards, providing a rich and user-friendly interface built with Polymer web components (<https://www.polymer-project.org/1.0/>).

A JavaScript API facilitates an easy integration of the CellMaps client by using plain HTML tags. Computations and analyses are executed in remote high-end servers and all the functionalities are available through RESTful web services. Data and session results can be stored and recovered locally or in remote servers (for registered users), where computation is performed. Relevant biological network information, including IntAct interactions, is available through CellBase (Bleda *et al.*, 2012b) web services (<http://bioinfo.hpc.cam.ac.uk/cellbase/webservices/rest>). KEGG and Reactome pathways are directly taken from their respective web services. CellMaps runs in the latest versions of ‘evergreen’ web browsers (Chrome, Firefox, Safari and Edge) and has no browser plugins or other libraries requirement.

### 3 Features

#### 3.1 Data input/output

Networks can be imported from local files in SIF, XLSX or text formats and exported in SVG or SIF formats. Pre-defined networks can also be imported from Reactome, KEGG or IntAct using CellBase (Bleda *et al.*, 2012b) web services. Attribute files containing meta-data information about nodes and edges (e.g. gene expression, GO terms, etc.) can also be imported as simple tab-delimited matrices. Sessions are maintained in the server for registered users but can also be saved in JSON format. See Additional Information for detail.

#### 3.2 Network editing and management

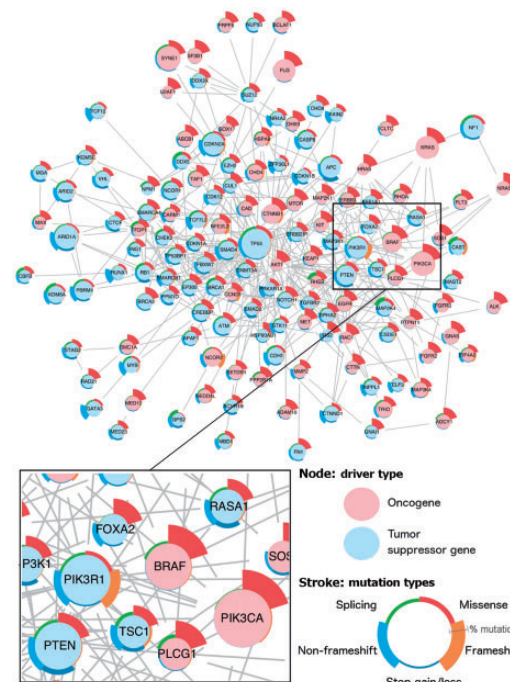
Once the network and attribute data are loaded, options for adding and deleting nodes as well as defining the number and type of relationships (edges) among them are available. Based on edge and node adjacency the user can perform different convenient ways of node and edge selection options, in a convenient way. To facilitate the visualization of the network CellMaps provides different layouts, among others taken from Cytoscape.js (Franz *et al.*, 2015), several shapes, colors, sizes and transparency settings for nodes and edges as well as the upload of background images.

Node and edge attributes can be visualized using simple visual properties as color, size, etc. and more sophisticated features as donut and pie charts (see an example in Fig. 1). They can be imported from files or from CellBase, which provides information for a total of 76 species. Of remarkable interest is gene-specific information on diseases (from ClinVar and UNIPROT), drugs (<http://dgidb.org>), Gene Ontology and tissue-specific gene expression (<http://www.proteinatlas.org>).

#### 3.3 Network analysis

Different options for network analysis and interpretation are available (described in more detail in the [Supplementary information](#)):

- *Find interactions*: locates a list of proteins within the interactome and returns the subnetwork that connects them, along with their first degree interactors.
- *Community Structure Definition*: finds the organization substructure of the network using different popular algorithms (*FastGreedy Walktrap* or *Infomap*) implemented in the *igraph* (Csardi and Nepusz, 2006) R package.
- *Topological Study*: allows calculating basic network topological parameters such as the connection degree, the betweenness centrality and the clustering coefficient.



**Fig. 1.** The network of interactions among cancer drivers identified in the ICGC cancer genomes through mutations in their interacting interfaces (Porta-Pardo *et al.*, 2015), completed with known drivers. The bottom left of the figure shows an example of nodes with donut charts in a magnification of the square in the network. Donut charts allows the simultaneous representation of several features in the nodes (connected by the gray lines)

- *Network Enrichment*: finds all the existing connections between pairs of proteins in a list and then, it applies a test to check if the level of connectivity of the network inferred is significantly higher than it can be expected by chance (Minguez *et al.*, 2009), a testing strategy also adopted by other resources, such as STRING (Franceschini *et al.*, 2013).
- *Gene Set Network Enrichment*: applies the gene set version of the previous test, by looking for significant subnetworks in a ranked or prioritized list of genes or proteins (Garcia-Alonso *et al.*, 2012).
- *Single Enrichment*: carries out a conventional Gene Ontology enrichment test (Al-Shahrour *et al.*, 2004) of the network.

#### 3.4 Example applications

As an example, CellMaps is embedded in RENATO (Bleda *et al.*, 2012a) providing an interactive graphical interface for the representation of regulatory networks (<http://renato.bioinfo.cipf.es>). Another example is Babelomics (Alonso *et al.*, 2015) (<http://www.babelomics.org>), which implements it in several options. It is also used in two pathway analysis programs, PathiWays (Sebastian-Leon *et al.*, 2013) and PathiPred (Amadoz *et al.*, 2015) (<http://pathiways.babelomics.org/>).

#### 3.5 Future directions

In addition to KEGG and Reactome, we are planning to integrate other web services such as Pathway Commons (<http://www.pathwaycommons.org/>) or the recently released Ndex (Pratt *et al.*, 2015). We are also planning to include more options, from Cytoscape.js and other useful graphic libraries.

## 4 Discussion

Despite the functionalities of desktop applications are increasingly being ported to the web, there are no efficient, user-friendly web-based solutions for the representation and analysis of biological networks available yet. CellMaps is an open-source web-based application oriented to end-users for advanced interactive visualization and network analysis that offers the performance of a web-based client/server application.

## Funding

This work is supported by grants BIO2014-57291-R from the Spanish Ministry of Economy and Competitiveness (MINECO), and Plataforma de Recursos Biomoleculares y Bioinformáticos PT 13/0001/0030 from the ISCIII, both co-funded with European Regional Development Funds (ERDF), PROMETEOII/2014/025 from the Conselleria de Educació of the Valencian Community, and the Fundació la Marató TV3 [20133134].

*Conflict of Interest:* none declared.

## References

- Al-Shahrour, F. *et al.* (2004) FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics*, **20**, 578–580.
- Alonso, R. *et al.* (2015) Babelomics 5.0: functional interpretation for new generations of genomic data. *Nucleic Acids Res.*, **43**, W117–W121.
- Amadoz, A. *et al.* (2015) Using activation status of signaling pathways as mechanism-based biomarkers to predict drug sensitivity. *Sci. Rep.*, **5**, 18494.
- Bleda, M. *et al.* (2012a) Inferring the regulatory network behind a gene expression experiment. *Nucleic Acids Res.*, **40**, W168–W172.
- Bleda, M. *et al.* (2012b) CellBase, a comprehensive collection of RESTful web services for retrieving relevant biological information from heterogeneous sources. *Nucleic Acids Res.*, **40**, W609–W614.
- Csardi, G. and Nepusz, T. (2006) The igraph software package for complex network research. *Int. J. Complex Syst.*, 1695. <http://igraph.org/c/doc/igraph-introduction.html#idm470952874400>.
- Franceschini, A. *et al.* (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.*, **41**, D808–D815.
- Franz, M. *et al.* (2015) Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics*, **32**, 309–311.
- Garcia-Alonso, L. *et al.* (2012) Discovering the hidden sub-network component in a ranked list of genes or proteins derived from genomic experiments. *Nucleic Acids Res.*, **40**, e158.
- Hu, Z. *et al.* (2013) VisANT 4.0: Integrative network platform to connect genes, drugs, diseases and therapies. *Nucleic Acids Res.*, **41**, W225–W231.
- Mínguez, P. *et al.* (2009) SNOW, a web-based tool for the statistical analysis of protein-protein interaction networks. *Nucleic Acids Res.*, **37**, W109–W114.
- Porta-Pardo, E. *et al.* (2015) A pan-cancer catalogue of cancer driver protein interaction interfaces. *PLoS Comput. Biol.*, **11**, e1004518.
- Pratt, D. *et al.* (2015) NDEx: the network data exchange. *Cell Syst.*, **1**, 302–305.
- Sebastian-Leon, P. *et al.* (2013) Inferring the functional effect of gene expression changes in signaling pathways. *Nucleic Acids Res.*, **41**, W213–W217.
- Taubert, J. *et al.* (2014) Ondex Web: web-based visualization and exploration of heterogeneous biological networks. *Bioinformatics*, **30**, 1034–1035.
- Warde-Farley, D. *et al.* (2010) The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.*, **38**, W214–W220.