

BioGRID REST Service, BiogridPlugin2 and BioGRID WebGraph: new tools for access to interaction data at BioGRID

Andrew G. Winter*, Jan Wildenhain and Mike Tyers

Wellcome Trust Centre for Cell Biology and School of Biological Sciences, The University of Edinburgh, Edinburgh, EH9 3JR, UK

Associate Editor: Jonathan Wren

ABSTRACT

Summary: The Biological General Repository for Interaction Datasets (BioGRID) representational state transfer (REST) service allows full URL-based access to curated protein and genetic interaction data at the BioGRID database. Appending URL parameters allows filtering of data by various attributes including gene names and identifiers, PubMed ID and evidence type. We also describe two visualization tools that interface with the REST service, the BiogridPlugin2 for Cytoscape and the BioGRID WebGraph.

Availability and implementation: BioGRID data and applications are completely free for commercial and non-commercial use. <http://webservice.thebiogrid.org/resources/interactions> (REST Service), <http://wiki.thebiogrid.org/doku.php/biogridrest> (REST Service parameter list and help), <http://webservice.thebiogrid.org/resources/application.wadl> (REST Service WADL), <http://thebiogrid.org/download.php> (BiogridPlugin2, v2.1 download), <http://wiki.thebiogrid.org/doku.php/biogridplugin2> (BiogridPlugin2 help) and http://tyerslab.bio.ed.ac.uk/tools/BioGRID_webgraph.php (BioGRID WebGraph).

Contact: andrew.winter@ed.ac.uk, m.tyers@ed.ac.uk

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on December 10, 2010; revised on January 24, 2011; accepted on January 30, 2011

1 INTRODUCTION

The aim of the Biological General Repository for Interaction Datasets (BioGRID, <http://thebiogrid.org>) is to curate and distribute comprehensive collections of physical and genetic interactions from all major model organisms, including human (Breitkreutz *et al.*, 2008; Reguly *et al.*, 2006). BioGRID currently houses 362 355 interactions, categorized by evidence code and supplemented with annotation such as modifications and phenotypes. BioGRID data are accessible through a searchable web interface, the Osprey network visualization tool (Breitkreutz *et al.*, 2003) and partner databases such as *Saccharomyces* Genome Database (Engel *et al.*, 2010). Downloadable data files in tab-delimited or PSI-MI 2.5 XML (Kerrien *et al.*, 2007) formats are available for off-line data analysis. However, since these files are often large and/or complex, full utilization of the data can require specialized scripts to be written.

From release 3.0.64 (May 2010), BioGRID has been made available in MITAB format (Kerrien *et al.*, 2007), a tab-delimited format, with BioGRID evidence codes mapped to PSI-MI ontology

terms. The PSICQUIC web service has been established by the IMEx consortium (<http://imex.sourceforge.net>) providing federated access to multiple interaction databases, including BioGRID, in MITAB format (Orchard *et al.*, 2009). Although access via PSICQUIC provides all available interactions in BioGRID, it is not ideal for provision of BioGRID data as it does not contain original BioGRID evidence codes, post-translational modifications, phenotypes or free-text qualifiers. Here, we describe a new representational state transfer (REST) web service allowing access to full BioGRID data, and two visualization tools that programmatically access it. Example use cases for all components are provided in Supplementary File 1.

2 BIOGRID REST SERVICE

BioGRID data have been made available through a REST-style web service at a dedicated URL, <http://webservice.thebiogrid.org/resources/interactions>, hereafter abbreviated to 'rest_url'. Parameters can be appended to the rest_url to filter the results returned, for example, 'rest_url/?searchNames=true&geneList=mdm2|tp53&includeHeader=true' will return all interactions between any genes with official name 'mdm2' or 'tp53'. Other parameters allow filtering by various identifiers, PubMed ID, experimental evidence and other data-associated features. A full list of parameters, help manual and a formal description in Web Application Description Language (WADL) for the BioGRID REST service are available.

In addition to BioGRID interaction data, interactions from the IMEx databases: IntAct (Aranda *et al.*, 2010), MINT (Ceol *et al.*, 2010) and DIP (Salwinski *et al.*, 2004) have been parsed, mapped to BioGRID evidence codes according to established rules (http://wiki.thebiogrid.org/doku.php/imex_import), and included as an option for retrieval in the REST service data (including originating database interaction identifiers), using the rest_url parameter 'sourceDatabaseList'. IMEx interaction files are downloaded from their respective web sites and integrated into the REST service to coincide with each monthly BioGRID release, currently (December 2010) providing >300 000 additional interactions and >90 000 new (i.e. not curated in BioGRID) non-redundant interactions. These additional interactions are currently only available through the BioGRID REST service.

3 BIOGRID PLUGIN2

Interaction data are often most intuitively viewed and manipulated from within a graphical network visualization tool such as Cytoscape (Shannon *et al.*, 2003) that has features that are continuously

*To whom correspondence should be addressed.

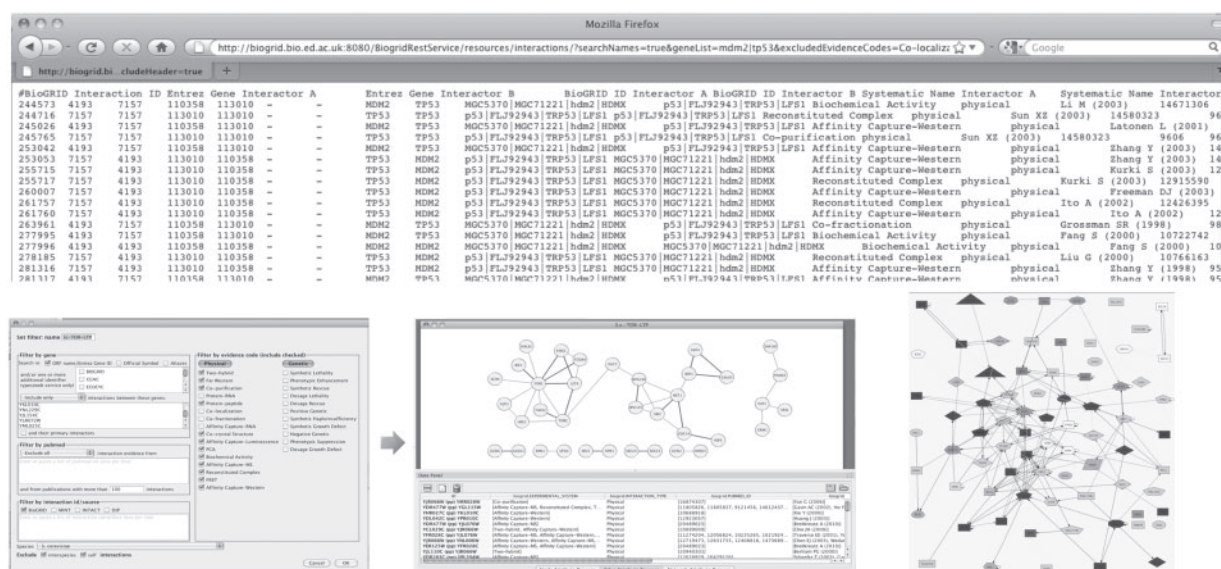


Fig. 1. BioGRID REST service application. Top panel: REST service query results viewed within a browser window. Bottom panel: Cytoscape BiogridPlugin2 custom filter interface (left) specifying import of interactions for a user-defined target of rapamycin (TOR)-related gene list from the REST service, and the resulting network view and edge attributes (centre); BioGRID WebGraph displaying results of a search on an extended TOR-related gene list (right).

enhanced by development of user-contributed plugins. Although Cytoscape has several plugins that act as web service clients allowing interactions to be imported from various sources across the web, none provide easy import of full BioGRID data. Osprey, another popular network visualization tool developed in conjunction with BioGRID, represents interactions in a flexible and rapidly expandable graphical format, allows functional comparisons between datasets and can retrieve data from the BioGRID database across the Internet. However, Osprey does not currently have an open source plugin-style architecture that enables users to apply different functions created by outside developers to network analysis. We have developed an enhanced plugin for Cytoscape called BioGRIDPlugin2 that can import data from the BioGRID REST service to allow the complete functionality of Cytoscape and its available plugins to be applied to BioGRID data. During import, data can be filtered according to various attributes such as gene names and identifiers, PubMed IDs and evidence codes, via a user-friendly custom filter dialog window (Fig. 1). Multiple filters can be created and saved for later use. Other useful functions include node expansion and export of filtered networks allowing creation of custom tab files based on the filters.

4 BIOGRID WEBGRAPH

BioGRID webgraph is a web application, which interacts with the BioGRID REST service to provide simple gene network graphs for user-provided gene lists. Graphs are created using scalable vector graphics, and can be saved as XGMML files for visualization in Cytoscape.

5 IMPLEMENTATION

The BioGRID REST service is implemented in Java, runs on a Redhat Linux 8 server, in Glassfish Application Server 3.1 and

connects to a MySQL 5 database running on a separate Redhat Linux 8 server. The Cytoscape plugin has been tested with Cytoscape 2.6 on Linux, MacOSX and Windows XP SP2. BioGRID webgraph is written in PHP5 and PEAR.

ACKNOWLEDGEMENTS

We thank Bobby-Joe Breitkreutz and Chris Stark for reading the manuscript and testing.

Funding: Biotechnology and Biological Sciences Research Council of the UK (grant number BB/F010486/1); the National Institutes of Health National Center for Research Resources (1R01RR024031); the European Commission FP7 Program (2007-223411).

Conflict of Interest: none declared.

REFERENCES

- Aranda, B. *et al.* (2010) The IntAct molecular interaction database in 2010. *Nucleic Acids Res.*, **38**, D525–D531.
- Breitkreutz, B.J. *et al.* (2003) Osprey: a network visualization system. *Genome Biol.*, **4**, R22.
- Breitkreutz, B.J. *et al.* (2008) The BioGRID Interaction Database: 2008 update. *Nucleic Acids Res.*, **36**, D637–D640.
- Ceol, A. *et al.* (2010) MINT, the molecular interaction database: 2009 update. *Nucleic Acids Res.*, **38**, D532–D539.
- Engel, S.R. *et al.* (2010) Saccharomyces Genome Database provides mutant phenotype data. *Nucleic Acids Res.*, **38**, D433–D436.
- Kerrien, S. *et al.* (2007) Broadening the horizon—level 2.5 of the HUPO-PSI format for molecular interactions. *BMC Biol.*, **5**, 44.
- Orchard, S. *et al.* (2009) Implementing data standards: a report on the HUPO-PSI workshop September 2009, Toronto, Canada. *Proteomics*, **10**, 1895–1898.
- Reguly, T. *et al.* (2006) Comprehensive curation and analysis of global interaction networks in *Saccharomyces cerevisiae*. *J. Biol.*, **5**, 11.
- Salwinski, L. *et al.* (2004) The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res.*, **32**, D449–D451.
- Shannon, P. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.