Cytoscape 2.8: New Features for Data Integration and Network Visualization

Michael E. Smoot*, Keiichiro Ono, Johannes Ruscheinski, Peng-Liang Wang, Trey Ideker
Departments of Medicine and Bioengineering, University of California San Diego, 9500 Gilman Dr. La Jolla, CA 92093.

ABSTRACT
Summary: Cytoscape is a popular bioinformatics package for biological network visualization and data integration. Version 2.8 introduces two powerful new features – Custom Node Graphics and Attribute Equations – which can be used jointly to greatly enhance Cytoscape’s data integration and visualization capabilities. Custom Node Graphics allow an image to be projected onto a node, including images generated dynamically or at remote locations. Attribute Equations provide Cytoscape with spreadsheet-like functionality in which the value of an attribute is computed dynamically as a function of other attributes and network properties.

Availability and Implementation: Cytoscape is a desktop Java application released under the Library Gnu Public License (LGPL). Binary install bundles and source code for Cytoscape 2.8 are available for download from http://cytoscape.org.

Contact: msmoot@ucsd.edu

1 INTRODUCTION

Networks are pervasive in biology. Vast data sets are being gathered to delineate networks of various types and levels, including networks of genetic and protein-protein interactions networks of transcriptional and post-transcriptional regulation, social networks, and multiscale networks that span several of these domains (Fowler, et al., 2009; Tan, et al., 2007). These network data extend and compliment a great deal of other information available in the biomedical sciences. Although various data sets can appear quite different in quality and quantity, they all are reflections of the same underlying biological system and its responses. Thus data integration, along with integrated visualization, is key.

Cytoscape, now in its eighth year of development, has become a standard tool for integrated analysis and visualization of biological networks (Cline, et al., 2007; Shannon, et al., 2003). Its central organizing principle is a network graph, with biological entities (e.g. genes, proteins, cells, patients) represented as nodes and biological interactions represented as edges between nodes. Data are integrated with the network using attributes, which map nodes or edges to specific data values such as gene expression levels or protein functions. Attribute values can be used to control visual aspects of nodes and edges (e.g. shape, color, size) as well as to perform complex network searches, filtering operations, and other analysis.

Version 2.8 of Cytoscape has introduced two significant new features that improve its ability to integrate and visualize complex data sets. The first feature allows non-programmers to map graphical images onto nodes, which greatly increases the power and flexibility with which integrated data can be visualized. The second feature is the introduction of spreadsheet-like equations into Cytoscape’s Attribute Browser to enable advanced transformation and combination of data sets directly within Cytoscape. Separately, each of these features provides useful new capabilities to Cytoscape. Taken together, however, these features provide a mechanism for expressing relationships between sets of data while simultaneously visualizing the integrated results.

2 INTEGRATED VISUALIZATION

2.1 Custom Node Graphics

A key function of Cytoscape is to allow diverse types of attribute data to be visualized on the nodes and edges of a biological network. Scalar data can be linked to simple visualization properties such as node color, shape, or size as node and edge attributes. To represent multivariate data associated with a node, Cytoscape can control the visualization of each node using a custom graphical image (since version 2.3) through a programming API. Version 2.8 introduces the ability for non-programmers to specify images through the Cytoscape GUI and to map these images to nodes using Cytoscape’s standard VizMapper interface.

Using this new feature, images are loaded using a standard file browser or by dragging and dropping them into a pool of available images maintained by Cytoscape. For cases involving large numbers of custom images, Cytoscape also allows images to be loaded as node attributes by providing a Uniform Resource Locator (URL) address for each image. Given a URL, Cytoscape will try to read that URL and generate an image for display. Images can be mapped to nodes using Cytoscape’s VizMapper framework, which connects the visualization properties of the network to attribute data. Placement and sizing of images can also be directly controlled within Cytoscape.

In addition to rich, multivariate visualizations (Figs. 1A, B) images enable shading, highlights, and other aesthetic effects to be applied to nodes (Fig. 1C) to further enhance network visualization.

2.2 Equations

The Cytoscape Attribute Browser provides a tabular view of data that have been loaded into Cytoscape and attached to nodes, edges, or networks. To improve the ability of users to manipulate attribute data within Cytoscape we have added Attribute Equations to version 2.8. Attribute Equations extend the existing spreadsheet-like behavior of the Cytoscape Attribute Browser (rows and columns of
The following equation for the `imageURL` attribute:

```
String attribute, which we name "imageURL". We then write the
Protein Data Bank (http://www.pdb.org). We then create a new
populate the value of this attribute with a valid protein name in the
creating a data attribute called "pdbID" and, for each node, we
then mapping that image to a node in the network. We begin by
URL that points to an image of the 3D structure of a protein and
This can be accomplished in Cytoscape v2.8 by constructing a
ics and Equations, we integrate networks of protein interactions
as a first example illustrating the joint use of Custom Node Graph-
version formula, which refers to the element in the data attribute.
variables that are available in the network. For those nodes where
formula for the visualization of this node, we write the following:

The following equation is copied to all nodes via an option in the right-click
menu of the Attribute Browser. For each node, Cytoscape interprets
the `pdbID` variable using the corresponding `pdbID` attribute value and concatenates the resulting strings into a valid URL. We then use the Cytoscape VizMapper to map the `imageURL` attribute to a Custom Graphics visual property. As described in section 2.1, this visual property interprets the URL and loads the image it points to. Following this procedure produces rich network visualizations such as those seen in Figures 1D and 1E.

A second example that demonstrates the use of both new features is the GLINECHART function available from the GoogleChart-Functions plugin (available in the Plugin Manager). This function uses the Google Visualization API (http://code.google.com/apis/visualization) to dynamically generate a URL to an image. For instance, if a user has time course data (values \{1.1, 2.3, 1.7, 0.6\}) for an experiment stored in an attribute named "timeCourse", the following function call will produce a URL of a line chart image of the data with a Y-axis range of 0 to 3:

```
= GLINECHART(0,3,$timeCourse)
```

This function evaluates to a URL, which in turn is interpreted by the VizMapper as a node image, resulting in integrated network visualizations such as those shown in Figs. 1A, B.

### 3 CONCLUSION

Custom node graphics provide a new tool for non-programmers that allows rich new network visualizations to be created that integrate large and complex data sets. Equations provide a powerful mechanism for data transformation within Cytoscape. Together, these features can be used to embed rich visualizations of data within the nodes of large networks.

### ACKNOWLEDGEMENTS

Our team would like to acknowledge the the Ideker Laboratory and the core Cytoscape development team.

**Funding** This work was supported by the National Institutes of Health [Grant Numbers GM070743, P01 HG005062].

### REFERENCES


