PathRings User Guide

PathRings is a web-based application for visualization and analysis of biological pathways. Users can visualize and search pathways from Reactome, as well as perform ortholog and gene expression analysis by uploading orthologous and gene expression data. This project is part of the PathBubbles project: http://sites.google.com/a/umbc.edu/pathbbubles/. The recommended web browser to use is Chrome.

The web-based PathRings is available for data analysis at http://raven.anr.udel.edu/~sunliang/PathBubbles/. All source code is available at Github here: https://github.com/ivcl/PathBubbles.

The PathRings manuscript is under review:

> Yongnan Zhu, Liang Sun, Alexander Garbarino, Carl Schmidt, Jinglong Fang, J. Chen. PathRings: an web-based tool for Exploration of Ortholog and Expression Data in Biological Pathways[J] (submitted to BMC Bioinfomatics)
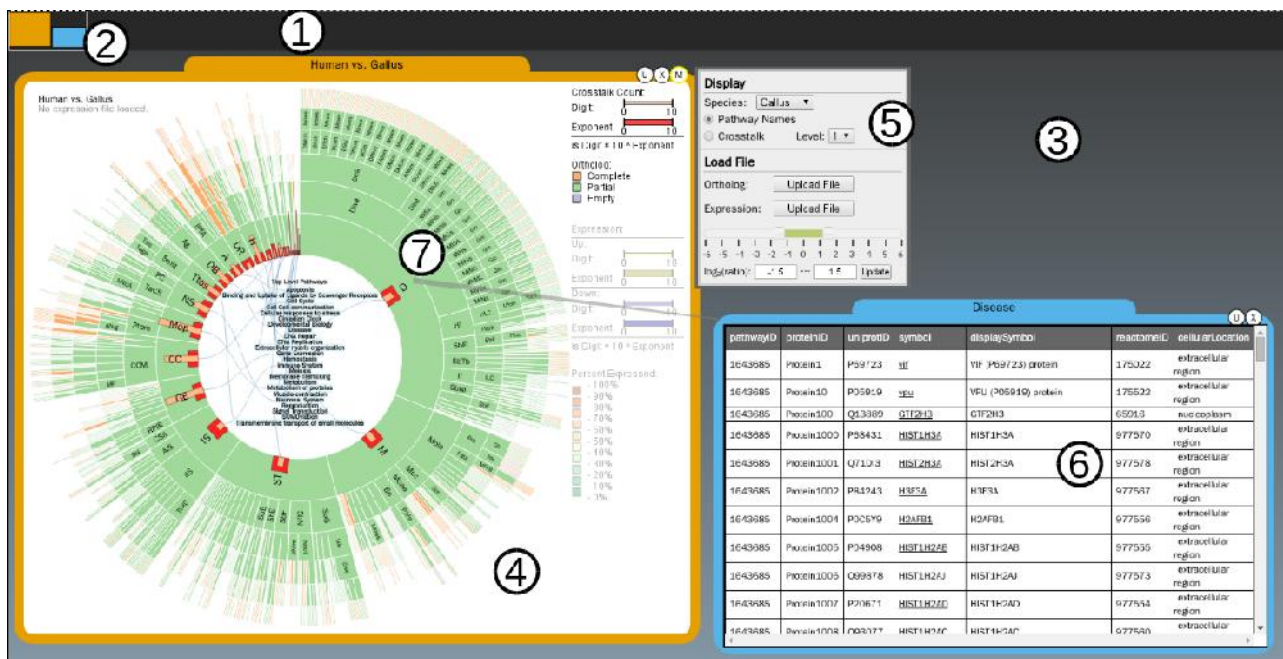
Figure 1: An overview of the PathRings interface.   1) The navigation bar. 2) The current view box. 3) The main workspace. 4) A sunburst (or treering) pathway visualization bubble. 5) A sub-menu. 6) A table bubble. 7) A link between tables.

1. Introduction

PathRings supports data exploration of four main types of pathway relationships: hierarchical gene pathways, cross-talking, orthologous, and gene expression. Each functional view is presented in a "bubble". These bubbles may be grouped together. Figure 1 provides an overview of the PathRings interface. At the top is a navigation bar (1) which provides an overview of the current layout, and along with (2) which can be used to quickly scroll to different sections of the exploration canvas. The main workspace (3) is where the various bubbles are laid out. (4) and (6) are examples of two different bubble types. (5) is a menu for data configuration. (7) is a link from (4) to (6), showing the relationship between the two.

2. Navigation Bar

The navigation bar spans the top of the PathRings interface. Each currently open bubble is displayed on the bar in a simplified form. The current view is represented with a white box, which may be dragged to change the current view.
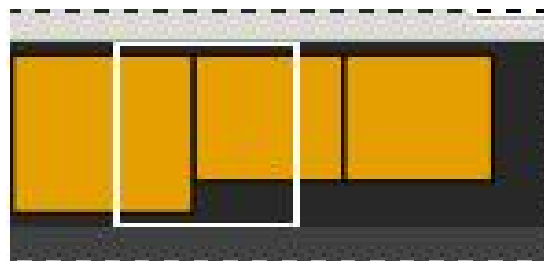

Figure 2: The Navigation Bar

3. Main Workspace

Below the navigation bar is the main workspace, which is the main display area. Right-clicking on any open spot in the main workspace will bring up a context menu with the following options:

- Open Entire Pathway: Creates a new, default pathway bubble.

- Delete All: Will delete all bubbles.

- Open Help: Opens this document.

4. General Bubble Mechanics

All bubbles share a few common traits, and are manipulated in a similar manner. While the mouse is over an outside edge of a bubble, the cursor will change. Clicking and dragging will allow the user to resize the bubble. Clicking and dragging anywhere else on the bubble's border, including the title, will allow the user to move the bubble. Each bubble also has several buttons in the upper-right corner. All will have the 'X' button, for deleting the bubble, and the 'U' button, for ungrouping it. The 'M' button will open a menu for further data queries.

Bubbles may be grouped and moved together. Grouping is done by dragging and dropping one bubble with the cursor placed over another bubble, or done implicitly when a bubble is created from a parent bubble. Bubbles may then be ungrouped by clicking and grabbing the bubble, starting from the 'U' button in the top right corner.
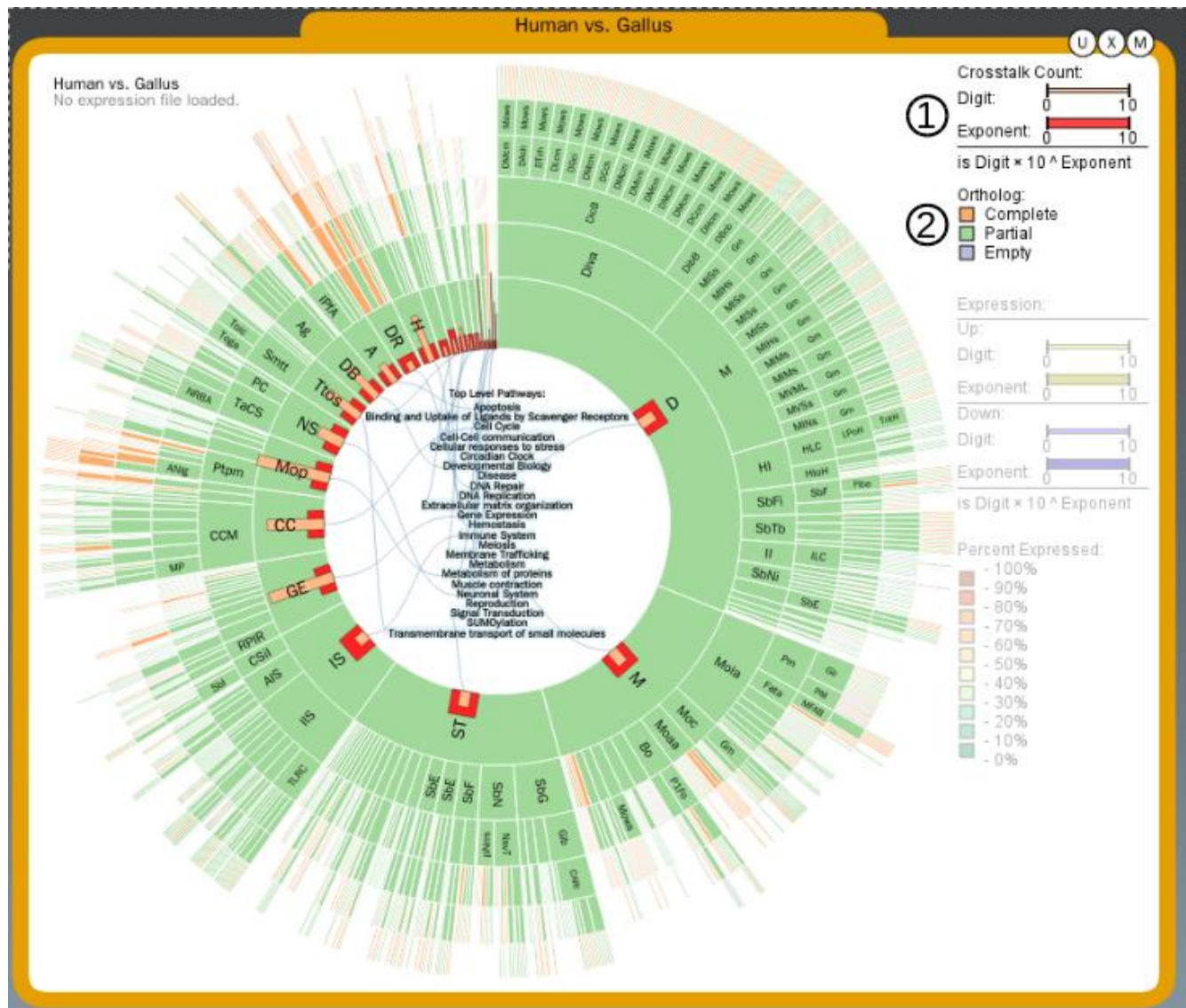
5. Pathway Bubbles

5.1 Main Bubble

Figure 3: A Pathway Bubble showing crosstalk counts.

The Pathway Bubble visualizes up to the entire set of Reactome pathways. In the center of the bubble is a list of the top-level pathways being shown. There is a line drawn from each of these pathways to its actual representation in the graph. Hovering over the pathway name with the mouse will highlight this linking line, so that the actual pathway may easily be distinguished. Using the menu, you may opt to hide this information and instead display lines showing crosstalk between the different pathways.

The main visualization has a different meaning depending on whether or not expression data has been uploaded. On the right of the bubble is a legend describing both depictions. The parts of the legend not currently in use will be faded from view.

Before it has been uploaded, each pathways is colored according to how closely the selected species or

uploaded ortholog data matches that of humans. The key is labeled as (2) in Figure 3. It distinguishes between *Complete*, *Partial*, and *Empty* pathways. A Complete pathway has all of its genes also found in humans, and an Empty one has none.

The actual bars in the chosen pathway level show how many crosstalk genes are present in the given pathway. This is presented using a split vector in order to improve readability. The length of the thicker bar represents the order of magnitude of the value, and the thinner bar is the value of its leading digit. The key for this is labeled as (1) in Figure 3.
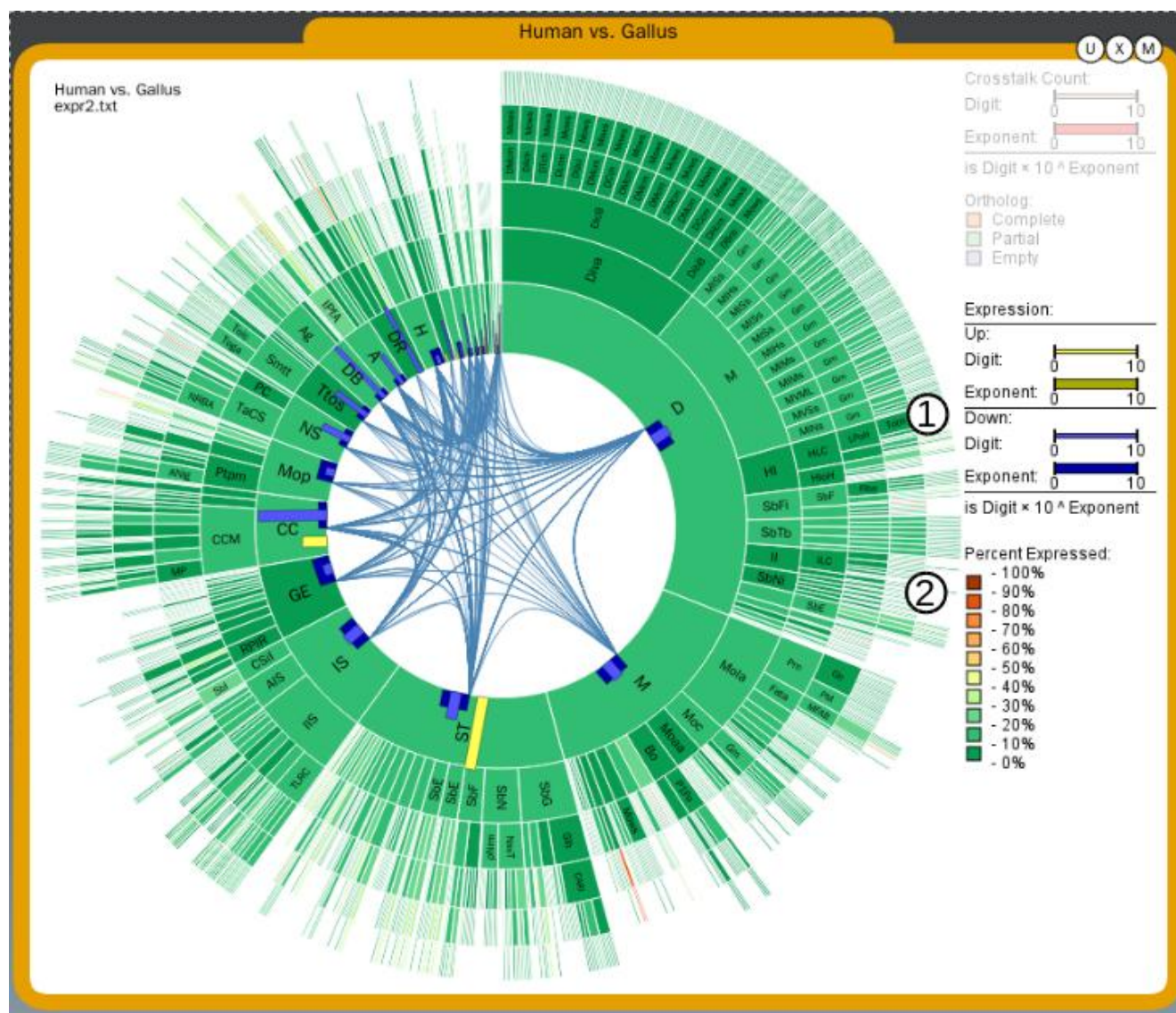

Figure 4: A Pathway Bubble showing gene expression.

Once an expression file is uploaded, the visualization will be updated. Each pathway is color coded by what percentage of genes in that pathway is up or down expressed, as explained by (2) in Figure 4. Each pathway in the selected level will show up to two different split bars – one showing the number of up expressed genes and one showing the number of down expressed. These are displayed in the same split format: The length of the thicker bar represents the order of magnitude of the value, and the

thinner bar is the value of its leading digit. The key for this is labeled as (1) in Figure 4.

The upper-left corner displays the currently selected species, as well as any loaded ortholog or expression files.

The actual diagram itself can be clicked on to bring up a new bubble. Left clicking on a pathway section will bring that pathway section up in a new pathway bubble. Right clicking on a pathway section will bring up a new table bubble listing all of the components of that pathway. The crosstalk or expression bars may also be right clicked – in this instance, the resulting table bubble will only have those components which are part of the displayed bar. If crosstalk links are currently visible, moving the mouse over an expression or crosstalk bar will highlight those links. Any of these newly created bubbles will be visibly linked to the original bubble.

5.2 Bubble Menu

The menu opened with the circular 'M' button in the top right corner is used to load data files and otherwise manipulate the display:

1. This selects a species to *compare* vs. *humans*.

2. This controls what is displayed in the center of the diagram. Either the list of pathway names, or links showing crosstalk between pathways.

3. This selects at which level of the graph the bars showing crosstalk or expression are displayed, as well as what level the crosstalk links connect to if those are selected. Level 1 is the innermost and highest ring displayed. Higher levels move towards the edge.

4. This checkbox is only visible once expression data has been loaded. While checked, this will shrink the "Percent Expressed" scale to have a smaller maximum (See Figure 6), based on the highest expressed pathway in that bubble.

5. This lets you choose an orthology or expression file to load. See section 7 for the file formats.

6. The bottom portion lets you select exactly what constitutes up or down expressed. Using the $\log_2$ of the ratio, anything below the first value is considered down expressed, and anything above the second is considered up expressed. The values may be changed either with the sliders or by typing them in directly. A new expression file must be loaded, or the Update button pushed, before these changes will take effect.
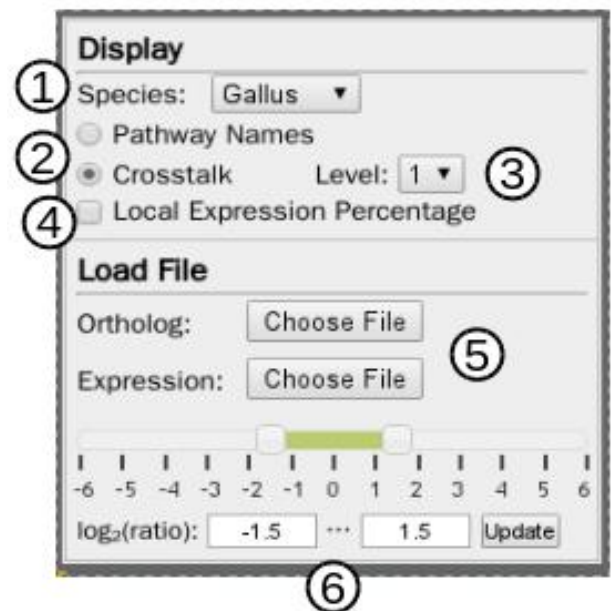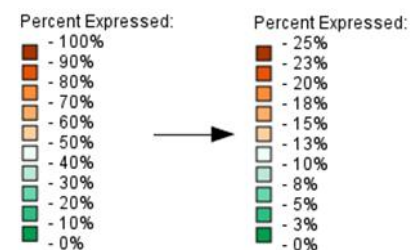

Figure 5: The Pathway Bubble's Menu.

Figure 6: Local Expression Percentage

In addition, any changes to species, ortholog, or expression data or display will be propagated to any linked bubbles as well.

6. Table Bubble

The table bubble is comparatively simple. The columns may be left clicked to sort the table by that column. Clicking again will cycle through ascending and descending orders. Elements in the symbol column may be left clicked to bring up an NCBI information pane on them. Right clicking on a symbol element will bring up a new table bubble composed of just that symbol. If the crosstalk column is present, its entries may be clicked to highlight those crosstalking events in an attached Pathway Bubble, if possible.

7. Data Format

This section describes the file formats used in the uploading option for the Pathway Bubble.

7.1. Ortholog Gene Format

Your own ortholog gene data may be uploaded. It should be a tab-delimited, two column text file. The first column is the gene symbol, and the second is that gene's orthologous relationship between this species and human. If this gene cannot be found in any human pathway, \N should be used. Otherwise, it should be the NCBI Entrez gene ID.

There is an example ortholog file for use available at:

http://raven.anr.udel.edu/~sunliang/PathBubbles/documents/Orthology.example1.txt.

Symbol	DbId
ADA	\N
CDH2	414745
AK13	421497
MED6	426282
NRE3	39528

Table 1: Ortholog Data Format Example

7.2. Expression Data Format

Your own gene expression data may be uploaded as well. It should be a tab-delimited, three column text file. The first column is the Entrez gene identification number, the second is the gene symbol, and the third is the expression ratio.

There is an example expression file for use available at:

http://raven.anr.udel.edu/~sunliang/PathBubbles/documents/GeneExpression.TGF.txt

Gene Id	Symbol	Ratio
374096	SMAD6	0.6892992938620315
395132	SMAD3	7.128911138923654
395247	SMAD2	3.370212765957447
395543	SMAD9	Infinity
395679	SMAD5	3.013579576317219

Table 2: Expression Data Format Example