

# Software for Graphical Notations in PathVisio and Lessons from its Usage

SBGN

4/29/2013

Augustin Luna

augustin@mail.nih.gov

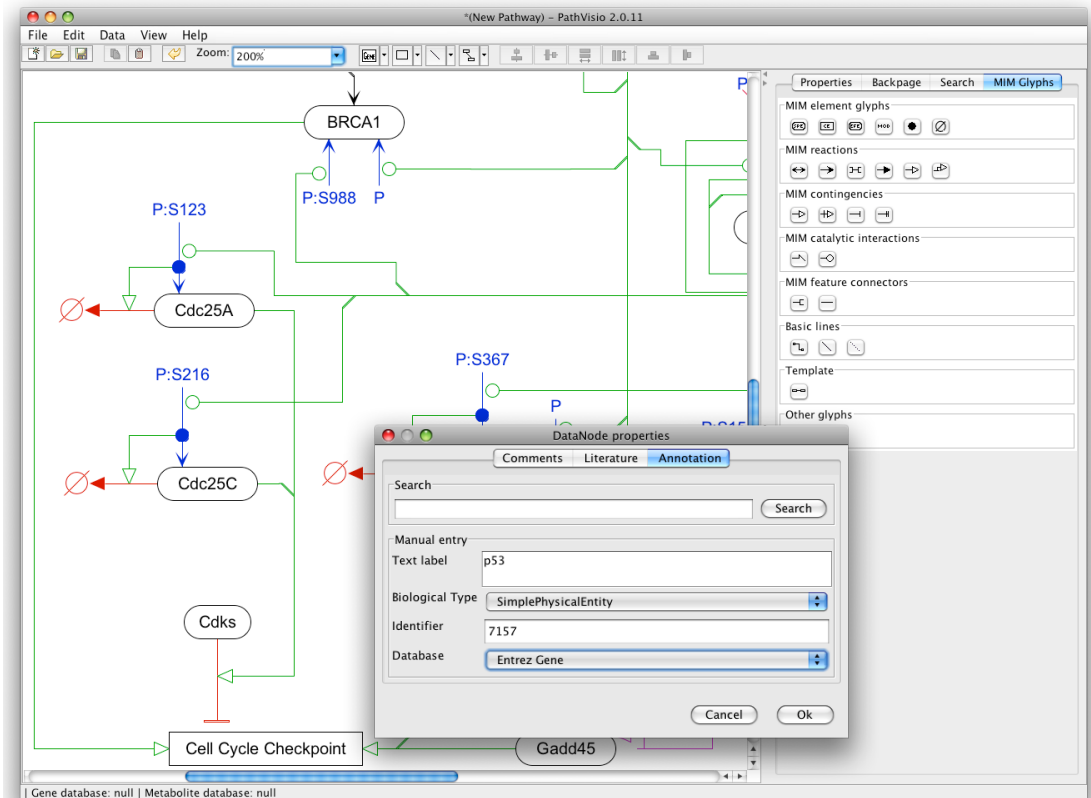
Bioinformatics Program, Boston University,  
Boston, MA, USA

Laboratory of Molecular Pharmacology,  
National Cancer Institute, Bethesda, MD, USA



# PathVisio

- Java-based
- Possesses a flexible plugin architecture
- Entities and interactions can be annotated with comments and literature references
- Associated pathway database, WikiPathways



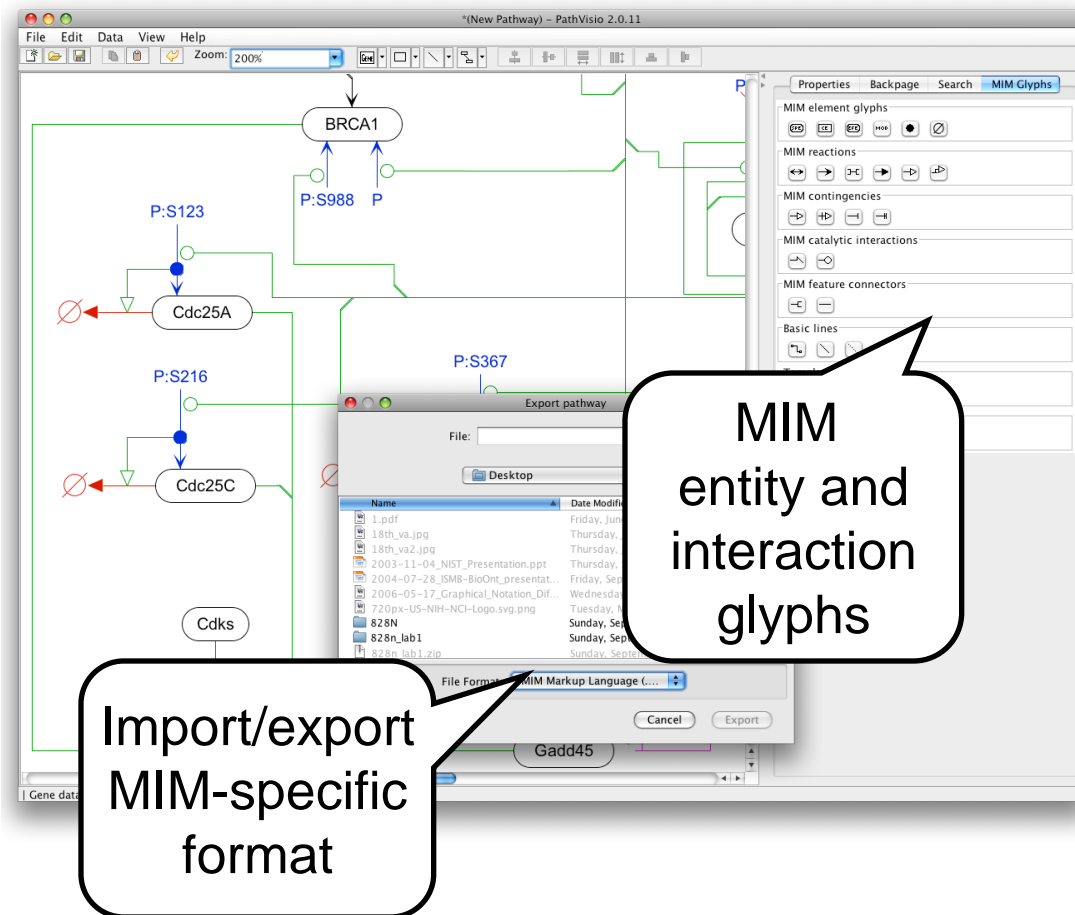
van Iersel, MP et al. (BMC Bioinformatics) 2008

augustin@mail.nih.gov

4/29/2013

# PathVisio-MIM Plugin

- Supports Molecular Interactions Map (MIM) notation created Kurt Kohn in 1999
  - Easy access to MIM glyphs
  - Import/export MIM-specific format meant for validation
- SBGN plugin has also been created for PathVisio



Luna A. et al. (Bioinformatics) 2011

# PathVisio-Validator Plugin

- Works with rulesets created using Schematron or Groovy
- Schematron: Generic XML validation framework
  - Language independent; stylesheets applied to XML datasets
- Groovy: Java-compatible scripting language
- MIM and SBGN rulesets created
- Other best practices rulesets created based on WikiPathways and E-MIMs curation rules

## Schematron-Based Ruleset

```
<!-- Validate Branching Interaction Arrowhead Prohibited Sets -->  
<iso:pattern name="validate-branching-interaction-prohibited-sets" id="validate-br  
  <iso:rule context="mimVis:InteractionGlyph/mimVis:Point[starts-with(@arrowHead
```

```
<iso:let name="vis-id" value="../@visId"/>  
<iso:let name="inter" value="//mimVis:InteractionGlyph[@visId=$vis-id]"/>
```

Highlighted Errors

Error: Unattached Line

Rule Groups

Validation Messages

Ruleset Selection

Errors: 3

Title: mimmi\_validation

Group: All

1.) Error - The end of an interaction should be possess a visRef attribute.

2.) Error - Non-branched interactions possessing a 'CovalentBondCleavage' arrowhead should be terminated with a 'Line' arrowhead.

3.) Error - A covalent bond cleavage interaction should be connected to one of the following entity types: entity feature, simple physical entity, implicit complex, conceptual entity, or explicit complex.

Highlight All

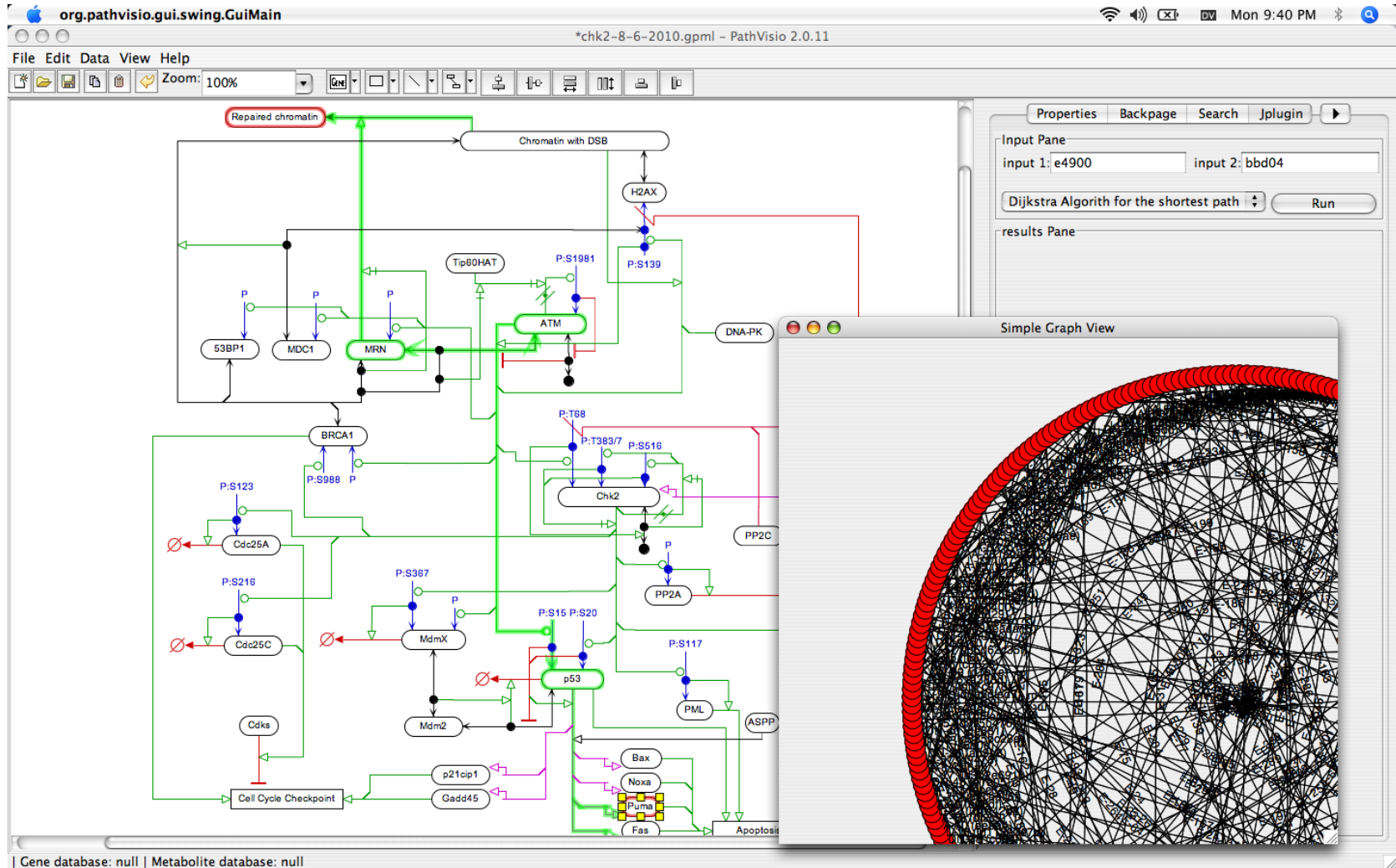
Errors & Warnings

Validate

Choose Ruleset

Luna A. et al. (BMC Bioinformatics) 2011; Chandan K. et al., (Bioinformatics) 2012; van Iersel, MP et al. (Bioinformatics) 2012

# Network Analysis: Are Entities in a MIM Connected?



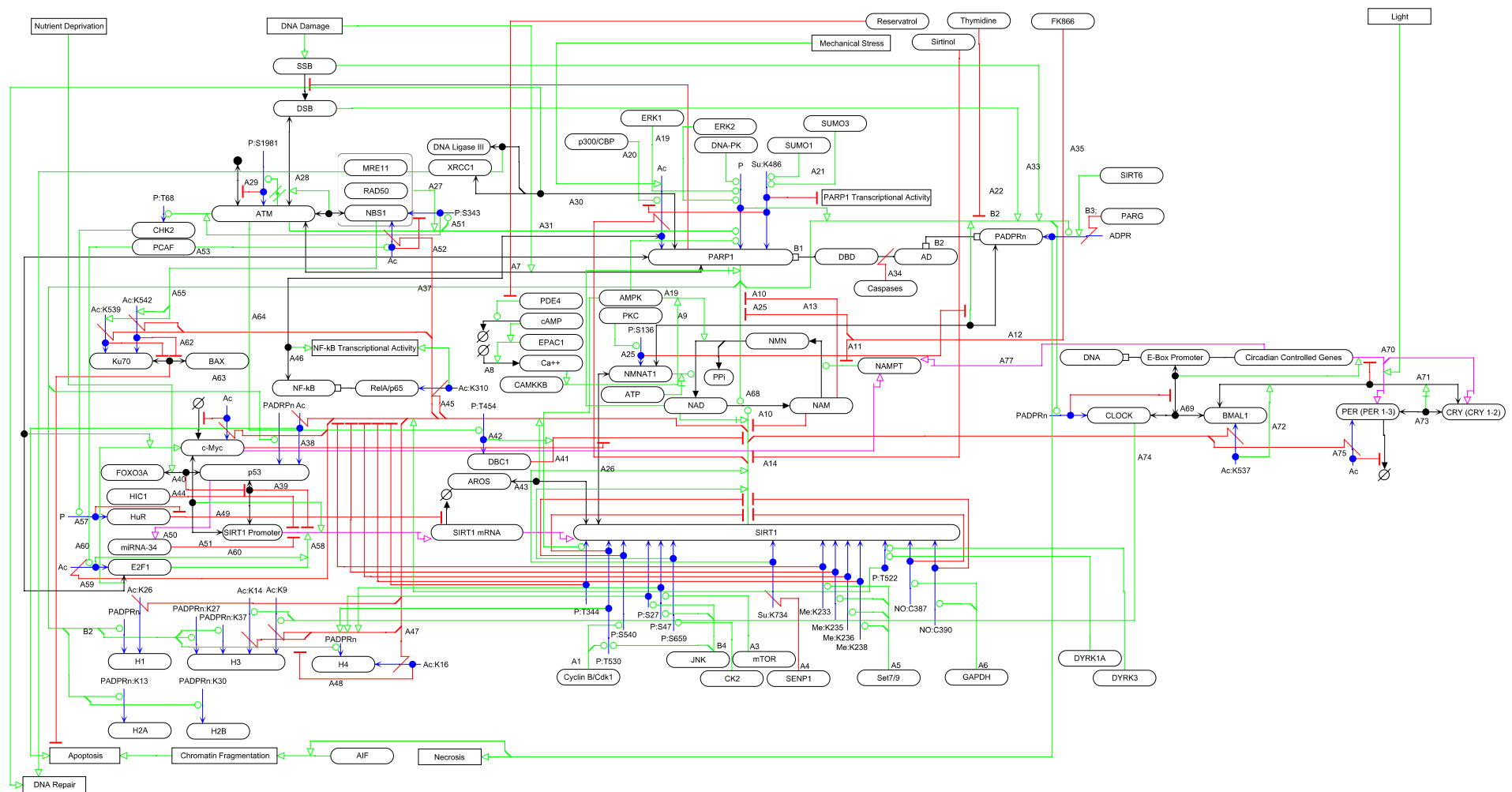
Chandan, K. and Luna, A.

# WikiPathways Database

- Search for pathways, genes, proteins, etc.
- Need to extend support for SBGN and MIM

The screenshot shows the WikiPathways website in a web browser. The browser's address bar displays 'http://wikipathways.org/index.php/WikiPathways'. The page has a navigation sidebar on the left with sections: 'search', 'navigation' (Home, Help), 'pathway' (Create, Browse, Wish List, Download), 'overview' (Recent Changes, Most Viewed, Most Edited, New Pathways), 'community' (About us, Contact us, How to cite, Curation events, GenMAPP portal, BiGCaT portal, Micronutrient portal, Development), and 'toolbox' (What links here). The main content area is titled 'Welcome to WikiPathways BETA' and includes a description: 'In the new tradition of Wikipedia, WikiPathways is an open, public platform dedicated to the curation of biological pathways by and for the scientific community. More about WikiPathways...'. Below this is a 'Finding Pathways' section with a 'Search' box (Google Custom Search) and a 'Browse' section (Browse Pathways). The 'Search' section lists search criteria: 'You can search by: Pathway name (Apoptosis), Gene or protein name (p53), Any page content (cancer)'. The 'Browse' section lists 'Browse by species and category'. There is a 'Contributing New Pathways' section with 'Create' (Create a new pathway page) and 'Suggest' (Add a pathway to the wish list) options. A 'Sample Pathway Pages' section is also visible. On the right, there is a 'Today's Featured Pathway' section showing the 'p38 MAPK Signaling Pathway (BioCarta) (Mus musculus)' with a diagram. Below this is a 'Latest edits' section listing recent updates, such as 'TNF-alpha and mucus production in lung epithelium (Rattus norvegicus)' by Kristina Hanspers.

<http://wikipathways.org>

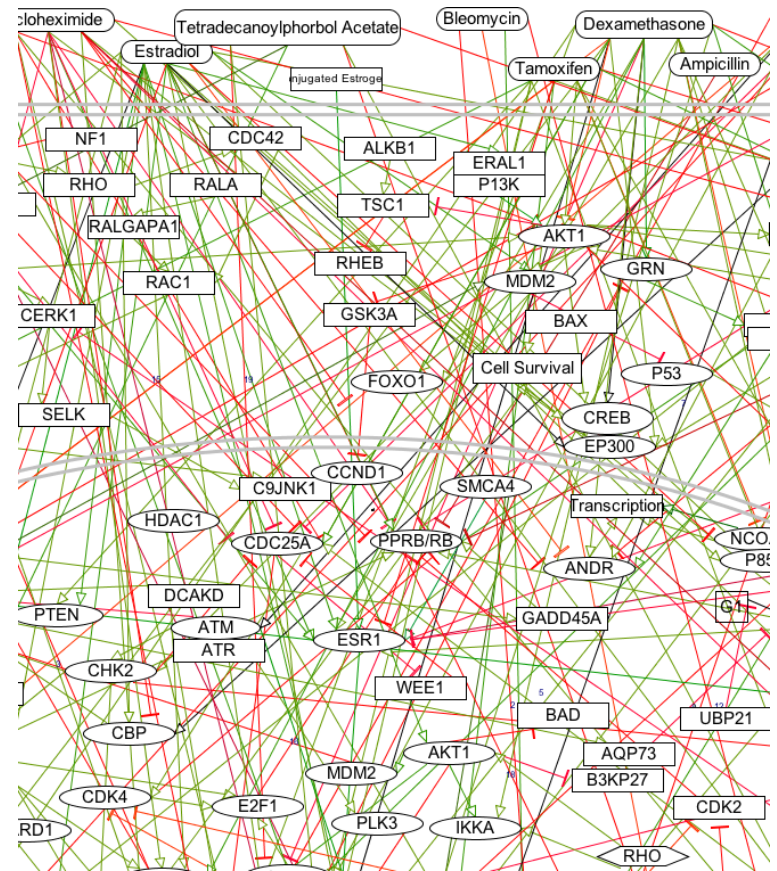


Review of SIRT1 and PARP1 Interactions: ~50 proteins, ~200 interactions; ~100 papers



# Layout, layout, layout...

- Layout is the biggest challenge in large diagrams
- Orthogonal layouts help keep diagrams readable
  - PathVisio has limited interaction routing capabilities; can keep lines orthogonal
- Difficult to extract smaller diagrams with entity selection or expand diagrams with new entities
- PathVisio-MIM allows more connection points on entities

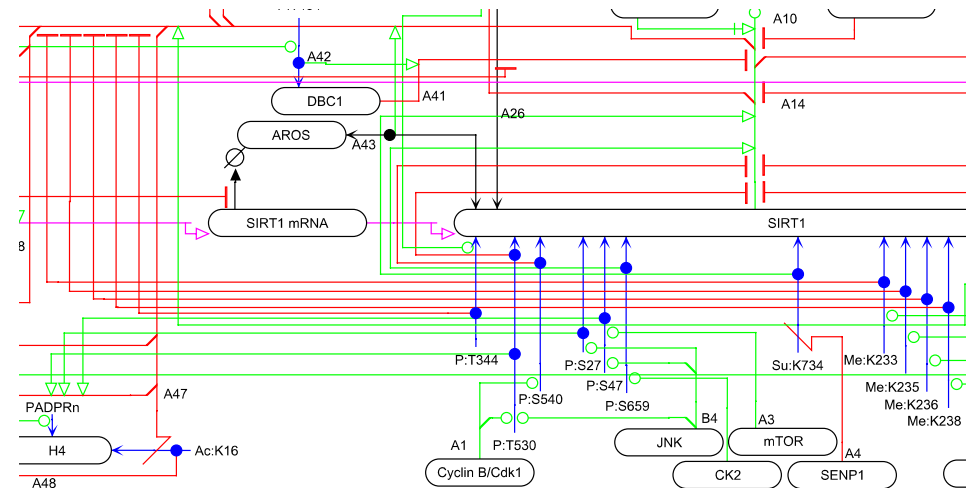


Breast Cancer Pathway  
WikiPathways



# Repetitive Interactions

- Repetitive interaction types
  - Add to the visual complexity
  - Make large diagrams difficult manage and layout



Many inhibitions stemming  
from entity modifications

# Entity Identification Problems

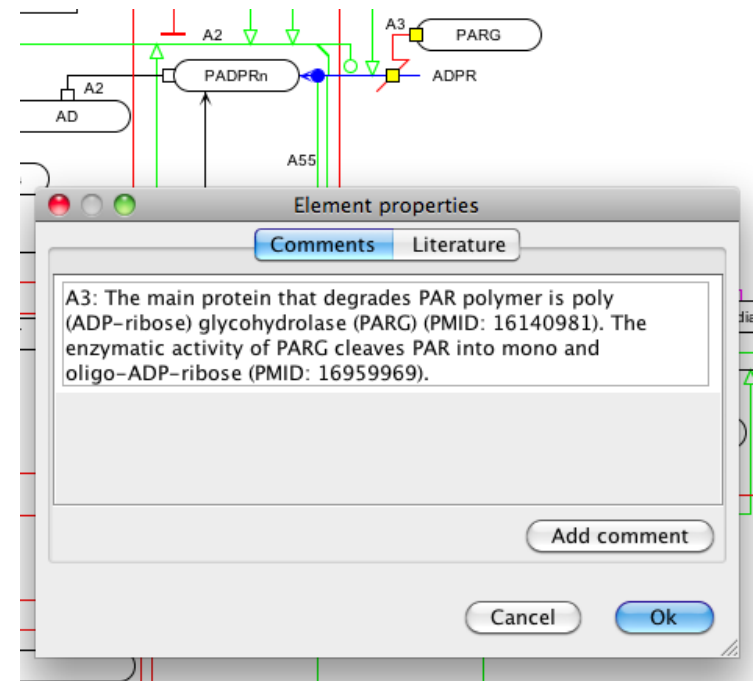
- Lack of information from publications for referencing some entities
  - Histones
  - Protein Kinase C (PKC); rarely are specific isomers referred to in publications

<a href="#">HIST1H1A</a>	histone cluster 1, H1a
<a href="#">HIST1H1B</a>	histone cluster 1, H1b
<a href="#">HIST1H1C</a>	histone cluster 1, H1c
<a href="#">HIST1H1D</a>	histone cluster 1, H1d
<a href="#">HIST1H1E</a>	histone cluster 1, H1e
<a href="#">HIST1H1PS1</a>	histone cluster 1, H1, pseudogene 1
<a href="#">HIST1H1PS2</a>	histone cluster 1, H1, pseudogene 2
<a href="#">HIST1H1T</a>	histone cluster 1, H1t
<a href="#">HIST1H2AA</a>	histone cluster 1, H2aa
<a href="#">HIST1H2AB</a>	histone cluster 1, H2ab
<a href="#">HIST1H2AC</a>	histone cluster 1, H2ac
<a href="#">HIST1H2AD</a>	histone cluster 1, H2ad
<a href="#">HIST1H2AE</a>	histone cluster 1, H2ae
<a href="#">HIST1H2AG</a>	histone cluster 1, H2ag
<a href="#">HIST1H2AH</a>	histone cluster 1, H2ah
<a href="#">HIST1H2AI</a>	histone cluster 1, H2ai
<a href="#">HIST1H2AJ</a>	histone cluster 1, H2aj
<a href="#">HIST1H2AK</a>	histone cluster 1, H2ak
<a href="#">HIST1H2AL</a>	histone cluster 1, H2al

HGNC Identifiers for Histones

# Linking Diagrams, Annotations, and Publication Text

- All interactions possess annotation labels
- Best practices MIM ruleset checks for properly annotated MIMs and annotation format
- Annotations eventually extracted to table
- Manuscript/diagram syncing tedious
  - PathVisio issue no easy way to renumber annotations



Annotation with multiple references

Desnoyers et al. 1999; Hassa, Haenni et al. 2006; Messner, Altmeyer et al. 2010). PARP1 is a transcriptional co-activator where PAR acts as a signal helping to regulate transcription (Luo and Kraus 2012). PAR is quickly cleaved by poly (ADP-ribose) glycohydrolase (PARG) [[A3]] (Hassa, Haenni et al. 2006). PARP1 becomes highly activated by DNA strand breaks;

Publication Manuscript

# Notation/Tool Instructions

- Instruction overload for many users
  - How do you use the notation?
  - How do I use the tool?
  - How do I use the analysis plugins?

## Contents

### [The PathVisio-Faceted Search plugin](#)

[Q: Who are the intended users of the PathVisio-Faceted Search plugin?](#)

[Q: How do I install/run the PathVisio-Faceted Search plugin?](#)

[Q: What are the requirements to use this plugin?](#)

[Q: How do I use the PathVisio-Faceted Search plugin ?](#)

[Q: How do I create a facets from local experimental data](#)

[Managing Dataservices](#)

[Batch Adding Facets](#)

[Facet Selection Notes](#)

[Q: How do I access webservices](#)

[Accessing Other Webservices](#)

[Tips for Biomart Webservice Facets](#)

[Q: How do I add facets using Groovy?](#)

[Running Groovy Facets](#)

[Return Values for Groovy Facets](#)

[Special Variables/Methods for Groovy Facet Code](#)

[Complete List of Biomart Attributes and Organism Databases](#)

[Tips for using Groovy Facets](#)

[Groovy Facet Code Examples](#)

[Basic Example: Create Facet from DataNode ID](#)

[Basic Example: Create Facet from Biomart Data](#)

[Complex Example: Linking User Data and Biomart Data](#)

[Complex Example: Returning Data from Other Webservices and URLs](#)

[Downloading and Parsing Text Files \(JSON, XML, etc.\) with Groovy](#)

[Complex Example: Working with Gene Ontology Information](#)

### [Troubleshooting](#)

#### [Plugin Performance](#)

[Facet Adding \(Time in Seconds\)](#)

[Performance Testing Conditions and Notes](#)

[Facet Adding \(Memory in MB\)](#)

[Performance Testing Conditions and Notes](#)

[Other Performance Testing Conditions and Notes](#)

### [Downloads](#)

#### [Links](#)

## Faceted-Search Plugin Help

# Acknowledgements

- NCI LMP
  - Mirit Aladjem
  - Kurt Kohn
  - Margot Sunshine
  - Sohyoung Kim
  - Lucas Chang (Cornell)
  - Noah Zweben (Columbia)
  - Katya Malison
- Bogazici University
  - Turkan Haliloglu
  - Can Ozturan
  - Itir Karac
  - Mine Edes
- NCI-Frederick
  - Ruth Nussinov
- Pathvisio Community
  - Martjin van Iersel (EBI)
  - Martina Kutmon (U of Maastricht)
- Google Summer of Code
  - Alexander Pico (UCSF)
  - Kumar Chandan (RIT)
  - Jake Fried (UMD)
- BioPAX Community
- SBGN Community