Benno Schwikowski · Frederik Gwinner · Oriol Guitart
 Systems Biology Lab · Institut Pasteur, Paris
Assembling Networks for Use in Biomedicine

Pathway assembly via integration of networks

Network based study of disease

Projection of molecular profiles on protein networks to reveal active modules

Network evolutionary comparison

Inference of networks from gene expression and other data

Integration of genetic and physical interactions

Advanced Network Visualizations

Network-based disease diagnosis, prognosis, and stratification

Models of cell dynamics

Network-based rational drug design

Moving from genome-wide association studies (GWAS) to network-wide “pathway” association (NWAS)
Cytoscape Overview

• Cytoscape is a Java application developed for the **visualization and analysis of biological networks**.

• Cytoscape is **open source** and has a plugin architecture that allows external developers to easily extend the capabilities for the core platform.

• Cytoscape is downloaded ~**2500** times per month.

• There are nearly **100 plugins** available through our website: [http://cytoscape.org](http://cytoscape.org)

• Very popular in the Systems Biology community, but also used in other domains like the Semantic Web.
Visualizing dense information

King et. al., Physiol Genomics. 2005 Sep 21;23(1):103-18.
Hypercube® rules and their overlap

J.-F. Bureau, F. Gwinner
The Collective Dynamics of Smoking in a Large Social Network

Nicholas A. Christakis, M.D., Ph.D., M.P.H., and James H. Fowler, Ph.D.

NEJM 358;21  WWW.NEJM.ORG  MAY 22, 2008
Cytoscape collaborators

Benno Schwikowski, Pasteur
Gary Bader, University of Toronto
Chris Sander, MSKCC
Trey Ideker, UCSD
James Fowler, UCSD
Bruce Conklin, Gladstone, UCSF
Mark Mercola, Sanford-Burnham
Margaret Wrensch, UCSF
James Adler, Agilent
Philippe Noirot, INRA, France
Nicolas Christakis, Harvard
Charles Sawaya, MSKCC
Marc Ladanyi, MSKCC
Andrew Emili, University of Toronto
Brenda Andrews and Charlie Boone, University of Toronto
Philippe Sansonetti, Pasteur

RAVIN site
Driving Biological Project
Collaboration and Service
Development

- The Cytoscape **core** is developed by a group of developers provided by Cytoscape Consortium members and volunteers.

- Cytoscape is **extended** by external plugin writers otherwise unaffiliated with the project.
Proliferation of Cytoscape Plugins
Dissemination

• Cytoscape is released under the LGPL software license - it is **free** software available for download from our website.

• The hub of our dissemination efforts is the project Website: [http://cytoscape.org](http://cytoscape.org)

• The site includes:
  • Manual, tutorials, and other documentation.
  • Subversion source code repository.
  • Discussion/Help mailing lists.
  • Bug Tracker.
Training and Education

- Numerous tutorials available on our website.
- Video lectures available on our website.
- Annual Symposium and Developer’s retreat.
  - Held each year.
  - Features invited talks, a plugin expo, tutorials, demonstrations, and user feedback forums.
- Participation in the Google Summer of Code.
- Taught in graduate, undergraduate and other classes (e.g. UCSD, UCSF, Lund University, DTU, ISB, ...).
What’s the problem?

- The Cytoscape codebase is very feature-rich, but is becoming increasingly **difficult to maintain**, and more importantly **difficult to extend**.
What’s next?

Modular Network Biology Toolkit
(aka Cytoscape 3.0)
What does this mean?

• A set of independent Java modules (jar files).

• A well defined and principled API - culmination of our team’s collective experience with Cytoscape.

• Use of OSGi to support and enforce modularity.

• Use of Spring-DM to abstract away the complexities of OSGi.

• Use of Maven to facilitate distribution and integration of modules.

• Follow accepted best programming practices (information hiding, code to interfaces, dependency injection, extensive unit testing, scrum, code quality metrics, semantic versioning, thread safe, ...).
Overall Goal?

• Make things **easier**!
• Easier to **use**
  • Simple programming model = more consistent user interface.
• Easier to **understand**
  • Well defined APIs, well defined dependencies, Maven archetypes.
• Easier to **maintain**
  • Clear APIs, separate API and implementation, semantic versioning, well understood dependencies.
• Easier to **extend**
  • (ditto)
Capabilities

• Do everything the current desktop version of Cytoscape does.

• Run in headless mode to support batch operations.

• Run in daemon mode to support backend web services.

• Take advantage of multithreaded and/or clustered environments.

• Many new features (scripting in different languages, 3D rendering, custom graphics, ...).
Improved Infrastructure

• Organize around a website with mailing lists, bug tracker, source repository but expand with a Maven repository and code quality analysis tools.

• Enhance our web tutorials with screencasts and movies.

• Enhance developer documentation and provide Maven archetypes to get plugin writers up and running quickly.

• Continue the successful Symposium and Retreat.
1A. Modular Layouts and Views
2A. Semantic Zooming: from genes to exons
### Information Layering Control Panel

<table>
<thead>
<tr>
<th>On/Off</th>
<th>Data Set</th>
<th>Visual</th>
</tr>
</thead>
<tbody>
<tr>
<td>🌟</td>
<td>Gene Expression A</td>
<td>Node color</td>
</tr>
<tr>
<td></td>
<td>Exon Array B</td>
<td>Exon color</td>
</tr>
<tr>
<td>🌟</td>
<td>Protein interactions</td>
<td>Edge color</td>
</tr>
<tr>
<td></td>
<td>Interaction confidence</td>
<td>Edge width</td>
</tr>
<tr>
<td></td>
<td>Load data set...</td>
<td>Choose...</td>
</tr>
</tbody>
</table>

**Diagram:**
- **Kinetochore**
- **Replication Fork**
Cytoscape Network Inference (CYNI)

Iterative network modeling process using Cytoscape
Questions we’d like biologists to ask

• What is the best network explaining the data?

• Which *parts* of this network are well-supported?

• Is there a well-supported subnetwork?

• Which experiment could be done to better distinguish different possible models?

• Given a model, which parts are consistent/inconsistent with the data?

• Which interactions could be added (removed) to make the data compatible with the model?
Questions for you

- In which (other) ways can such a platform can be useful?
- Major challenges?
- Required features to be useful for you?
- Most important features for biologists?

Thoughts and collaboration welcome.