

# Biological Pathway Exchange (BioPAX) Format and Pathway Commons Update

COMBINE  
10/16/2015

Augustin Luna  
[lunaa@cbio.mskcc.org](mailto:lunaa@cbio.mskcc.org)

# BioPAX Background

# Pathguide: The Pathway Resource List

## Pathguide» the pathway resource list

### Navigation

Protein-Protein Interactions  
Metabolic Pathways  
Signaling Pathways  
Pathway Diagrams  
Transcription Factors / Gene Regulatory Networks  
Protein-Compound Interactions  
Genetic Interaction Networks  
Protein Sequence Focused  
Other

### Search

Organisms  
All  
Availability  
All  
Standards  
All  
Reset Search

### Analysis

Statistics  
Database Interactions

### Contact

Comments, Questions, Suggestions are Always Welcome!

### Complete Listing of All Pathguide Resources

Pathguide contains information about **547** biological pathway related resources and molecular interaction related resources. Click on a link to go to the resource home page or 'Details' for a description page. Databases that are free and those supporting BioPAX, CellML, PSI-MI or SBML standards are respectively indicated.

If you know of a pathway resource that is not listed here, or have other questions or comments, please [send us an e-mail](#).

### News

**Major new update of Pathguide**  
August 2013  
We now have information about ~550 resources!.

**Visual navigation added May 2010**  
Click the 'Database interactions' link on the left menu to access.

### Protein-Protein Interactions

Database Name (Order: alphabetically | [by web popularity](#) )

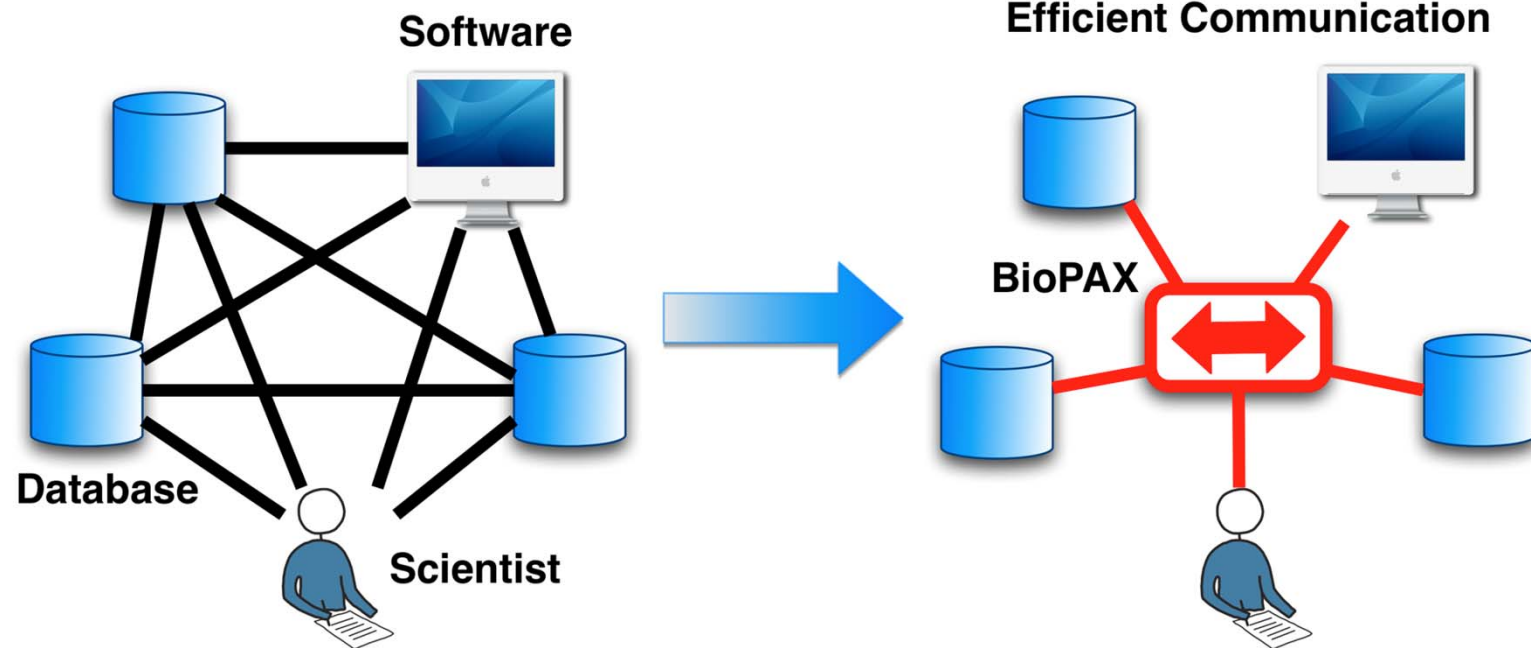
	Full Record	Availability	Standards
2P2ldb - The Protein-Protein Interaction Inhibition Database	<a href="#">Details</a>	Free	
3D-Interologs - 3D-Interologs	<a href="#">Details</a>	Free	
3DID - 3D interacting domains	<a href="#">Details</a>	Free	
ADAN - Prediction of protein-protein interaction of modular domains	<a href="#">Details</a>	X	
AHD2.0 - Arabidopsis Hormone Database 2.0	<a href="#">Details</a>	Free	
AllFuse - Functional Associations of Proteins in Complete Genomes	<a href="#">Details</a>	X	
aMAZE - Protein Function and Biochemical Pathways Project	<a href="#">Details</a>	X	
ANAP - Arabidopsis Network Analysis Pipeline	<a href="#">Details</a>	Free	
AnimalTFDB - Animal Transcription Factor Database	<a href="#">Details</a>	Free	
AntiJen - AntiJen a Kinetic, Thermodynamic and Cellular Database	<a href="#">Details</a>	Free	
APID - Agile Protein Interaction DataAnalyzer	<a href="#">Details</a>	Free	
AS-ALPS - Alternative Splicing - induced ALteration of Protein Structure	<a href="#">Details</a>	Free	
ASD - Allosteric Database	<a href="#">Details</a>	Free	
ASEdb - Alanine Scanning Energetics Database	<a href="#">Details</a>	Free	
ASPD - Artificial Selected Proteins/Peptides Database	<a href="#">Details</a>	Free	
ATDB - Animal Toxin Database	<a href="#">Details</a>	Free	
AtPID - Arabidopsis thaliana Protein Interactome Database	<a href="#">Details</a>	Free	
AtPIN - Arabidopsis thaliana Protein Interactome Network	<a href="#">Details</a>	Free	
Bacteriome.org - Bacterial Protein Interaction Database for Escherichia Coli	<a href="#">Details</a>	Free	
BIANA - Biologic Interaction and Network Analysis	<a href="#">Details</a>	Free	

<http://www.pathguide.org/>

[lunaa@cbio.mskcc.org](mailto:lunaa@cbio.mskcc.org)

10/16/2015

# BioPAX Goal: Unification of Pathway Data



# BioPAX Scope

- To represent biological interactions and pathways
  - Pathways are collections of interactions that biologists have found useful to group together for organizational, historic, biophysical or other reasons
  - Types
    - Metabolic pathways
    - Signaling pathways
    - Protein-protein, molecular interactions
    - Gene regulatory pathways
  - Tends to capture interaction information described in experimental biology literature
- Encourage a community-wide effort to distribute pathway data in standard format

# Getting started with BioPAX and Paxtools API

- BioPAX Website: <http://biopax.org/>
- BioPAX Tutorial: <https://github.com/cannin/biopaxTutorial>
- Paxtools Web Console: <http://pk0iye.appspot.com/>

The screenshot displays the Paxtools Web Console interface, which is a web-based tool for working with BioPAX data. The interface is divided into four main panels:

- Code Editor:** Contains Java code for loading and processing BioPAX data. The code includes imports for `FileInputStream`, `SimpleIOHandler`, and `Model`, and methods for loading a BioPAX file and printing participants in biochemical reactions.
- BioPAX OWL File:** Displays the BioPAX OWL file content, which is an XML document representing a biochemical pathway. The file includes namespaces for `xsd`, `owl`, `rdf`, and `bp`, and defines a `ProteinReference` class.
- Output:** Displays the output of the code execution, showing the BioPAX OWL file content and the results of the `printParticipants` method.
- Code Examples:** Provides a list of code snippets for various BioPAX operations, including `listFiles_0.groovy`, `firstModel_1.groovy`, `simpleIO_2.groovy`, `basicTraversal_3.groovy`, and `merging_7.groovy`.

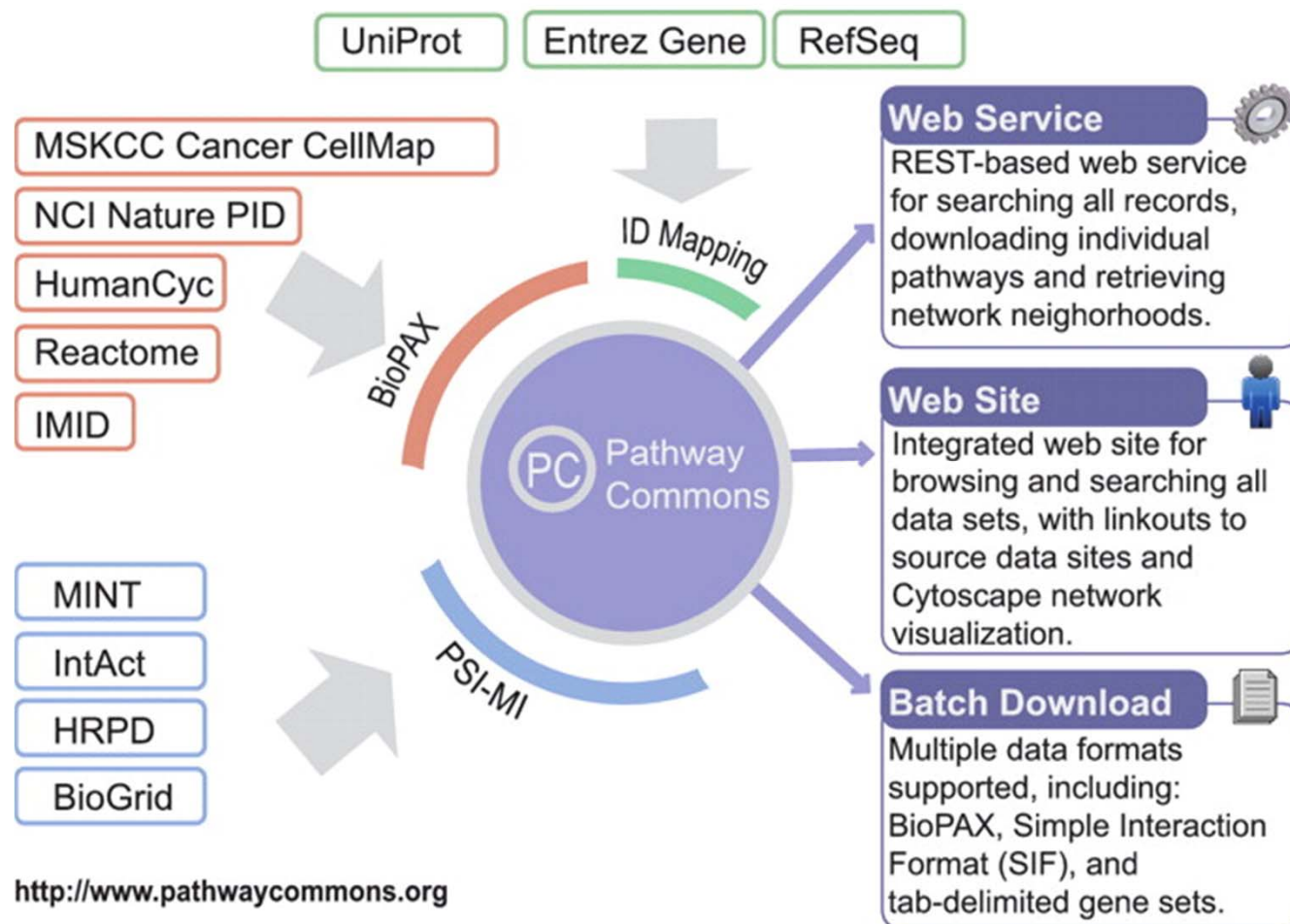
# Latest Developments Involving BioPAX

- BioPAX Level 3 (2010) was the last major release
- Pathway Commons released (2011)
  - <http://www.pathwaycommons.org/>
  - An aggregated database of BioPAX data
  - Has been the major focus for promotion and usage of BioPAX
- BioPAX Validator (2013)
  - <http://biopax.baderlab.org/validator/>

# Pathway Commons



# Pathway Commons Database



Cerami et al (2011)

# Interactions and Databases in Pathway Commons Version 8 (Upcoming)

- Pathway Commons includes a growing set of databases across a spectrum of databases
- Includes 20 databases with ~1.2 million interactions
- New versions of Pathway Commons are released roughly twice a year

Reactome	11924
NCI PID	16017
PhosphoSitePlus	13642
HumanCyc	7024
HPRD	40618
PantherDB	5282
DIP	7102
BioGRID	244843
InAct	98347
BIND	35566
TRANSFAC	261624
mirTarBase	51214
DrugBank	19159
Recon X	10910
CTD	313174
KEGG	4472
SMPDB	29693
INOH	6127
NetPath	7214
WikiPathways	9829
Total	1193781

# BioPAX/Pathway Commons Tools

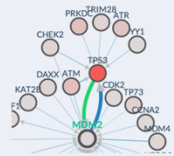
# BioPAX/Pathway Commons Tools

Pathway Commons

Download F.A.Q. Publications Contact

## For biologists

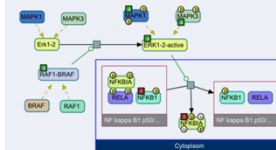
Search, visualize and download Pathway Commons pathways as part of an integrated network analysis ([more](#))



### Simple

See genes in pathway context

PCViz



### Advanced

See detailed processes

ChiBE



### Analyze

Search and analyze pathway relationships

CyPath2

## For computational biologists and software developers

Download all pathways in BioPAX, SIF and other formats for pathway and network analysis. Build software on top of Pathway Commons using our web service API ([more](#))



### PC2: Web service

BioPAX Level 3. Advanced graph queries. Normalized and up-to-date data. Programmatic access. Batch downloads.

Pathway Commons 2



### BioPAX & Paxtools

Standard language for Biological Pathway Exchange and a software library for handling data in BioPAX.

BioPAX & Paxtools



### PaxtoolsR

An R interface for Paxtools software and Pathway Commons webservice.

PaxtoolsR



### PC: Previous web service

Obsolete, last updated 2011

Pathway Commons

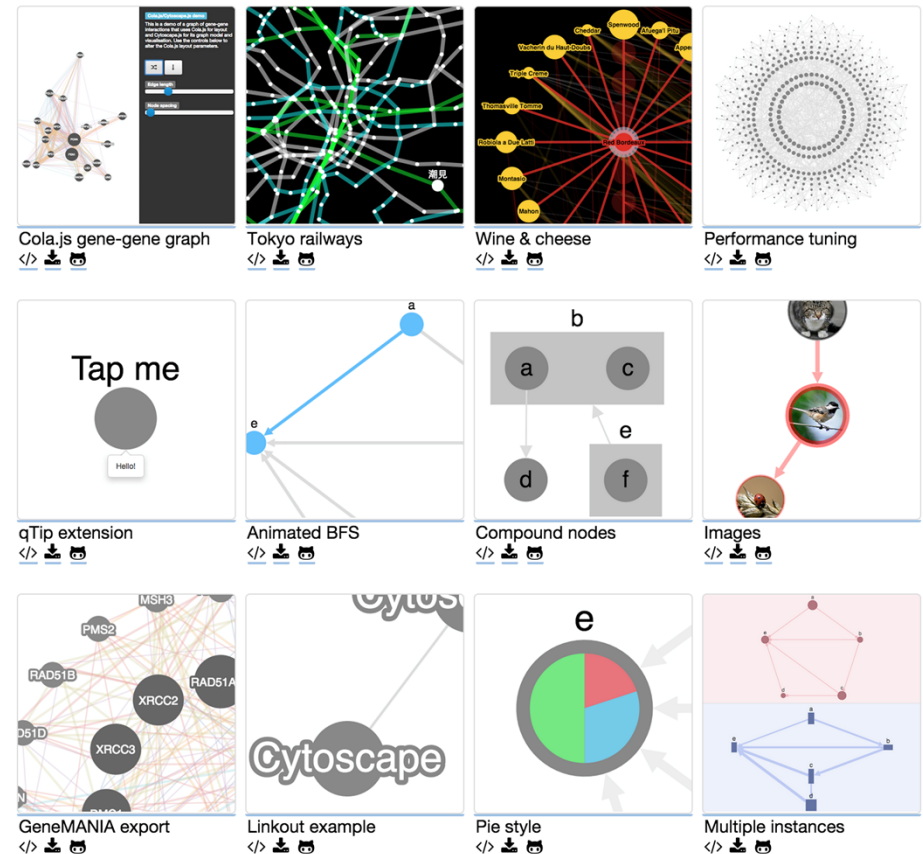
# Focus on Outreach

- Pathway Commons (PC) received renewed funding this year
- Several grant reviewer comments highlighted the need for “biologist” usage of the data
- We are hiring outreach coordinator to interact with biology communities outside of computational biology
- Focus on web-based tools to focus on interacting with PC and PC-based analyses

# Outreach Using Web-Based Tools

# Cytoscape.js: A Javascript library for analysis and visualization of networks

- Can handle large networks
- Provides analysis algorithms
- Provides various layouts
- Customizable features
  - Colors
  - Tooltips
  - Linkouts
  - Node images
  - Compound nodes
  - ...
- Basis for other projects
  - PCViz
  - SBGNViz



<http://js.cytoscape.org/>

[lunaa@cbio.mskcc.org](mailto:lunaa@cbio.mskcc.org)

10/16/2015

# UniProt Mockup

None

# h

- |   |                      |
|---|----------------------|
| ✓ | FUNCTION             |
| ✓ | NAMES & TAXONOMY     |
| ✓ | SUBCELLULAR LOCATION |
| ✓ | PATHOLOGY & BIOTECH  |
| ✓ | PTM / PROCESSING     |
| ✓ | EXPRESSION           |
| ✓ | INTERACTION          |
| ✓ | STRUCTURE            |
| ✓ | FAMILY & DOMAINS     |
| ✓ | SEQUENCES (11)       |
| ✓ | CROSS-REFERENCES     |
| ✓ | PUBLICATIONS         |
| ✓ | ENTRY INFORMATION    |
| ✓ | MISCELLANEOUS        |
| ✓ | SIMILAR PROTEINS     |

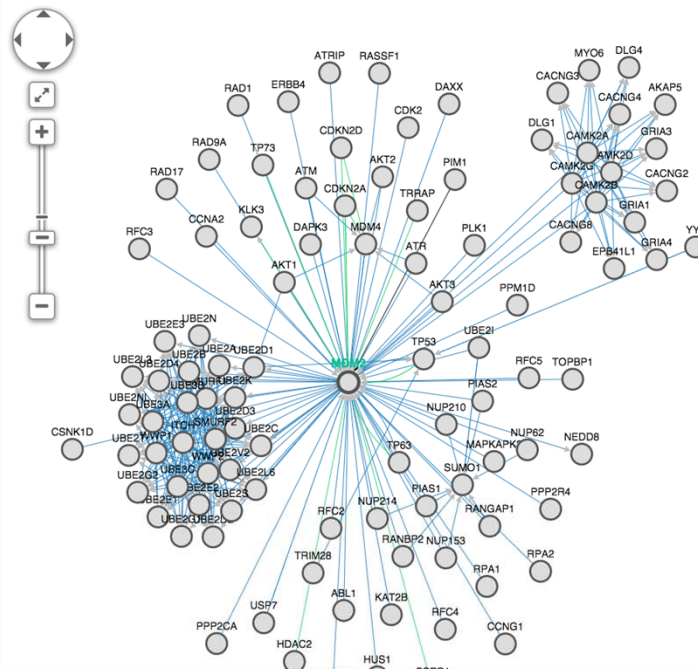
Entry version 188 (14 May 2014)  
Sequence version 1 (01 Apr 1993)

## Previous versions

**V** Help video

<sup>i</sup>

### Pathway View



PCViz Pathway Commons Network Visualizer

### Description

## Interaction

## PIM1 controls state change of MDM2

11504915, 11715018,  
11850850, 11923280,  
11960368, 11983168,  
15001356, 15024078,  
15169778, 15527798,  
15625077, 15688025,  
16774943, 17107963,  
17157787, 17297446,  
17386267, 18209096,  
18332867, 18359851,  
18467333, 19008219,  
19203586, 19630168,  
19664995, 19690332,  
19920202, 20657550,  
20708156, 21445305,  
22322096, 22349818,  
23132866, 23186163,  
25159151

## Participant #1

PIM1

## Participant #2

MDM2

## Co-Citations

0

Mockup by Arman Aksoy

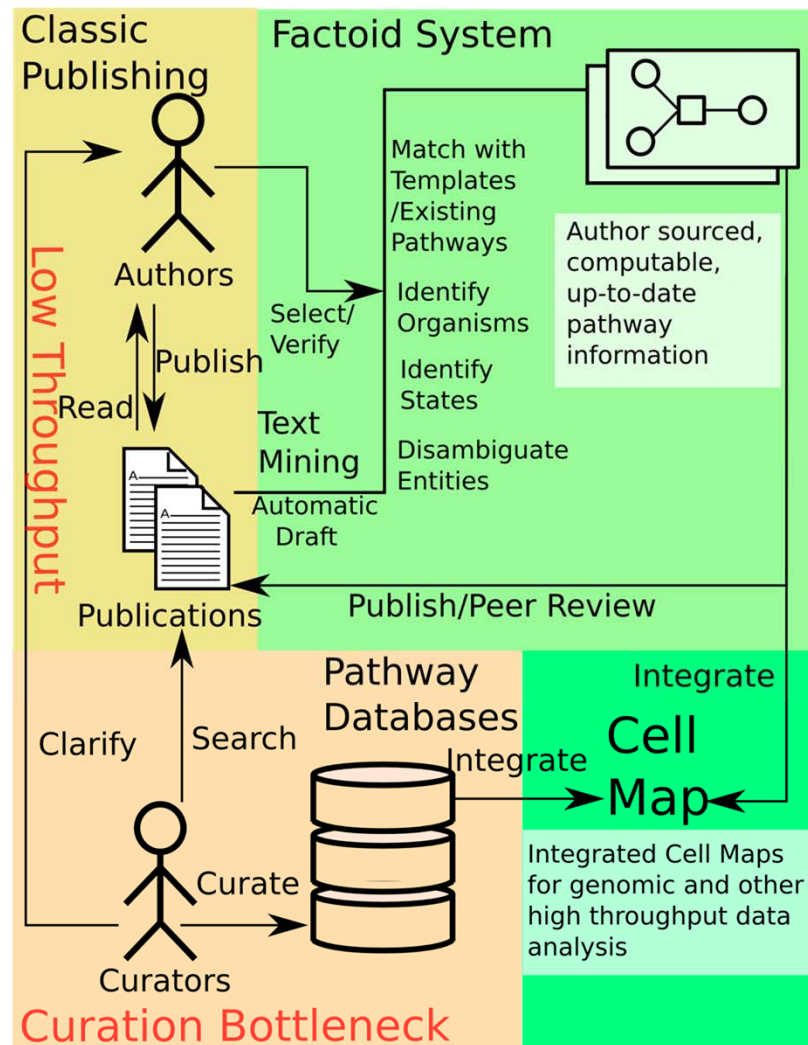
[lunaa@cbio.mskcc.org](mailto:lunaa@cbio.mskcc.org)

10/16/2015



# Expanding Pathway Commons

# Factoid: Expanding Pathway Commons During Manuscript Submission Process

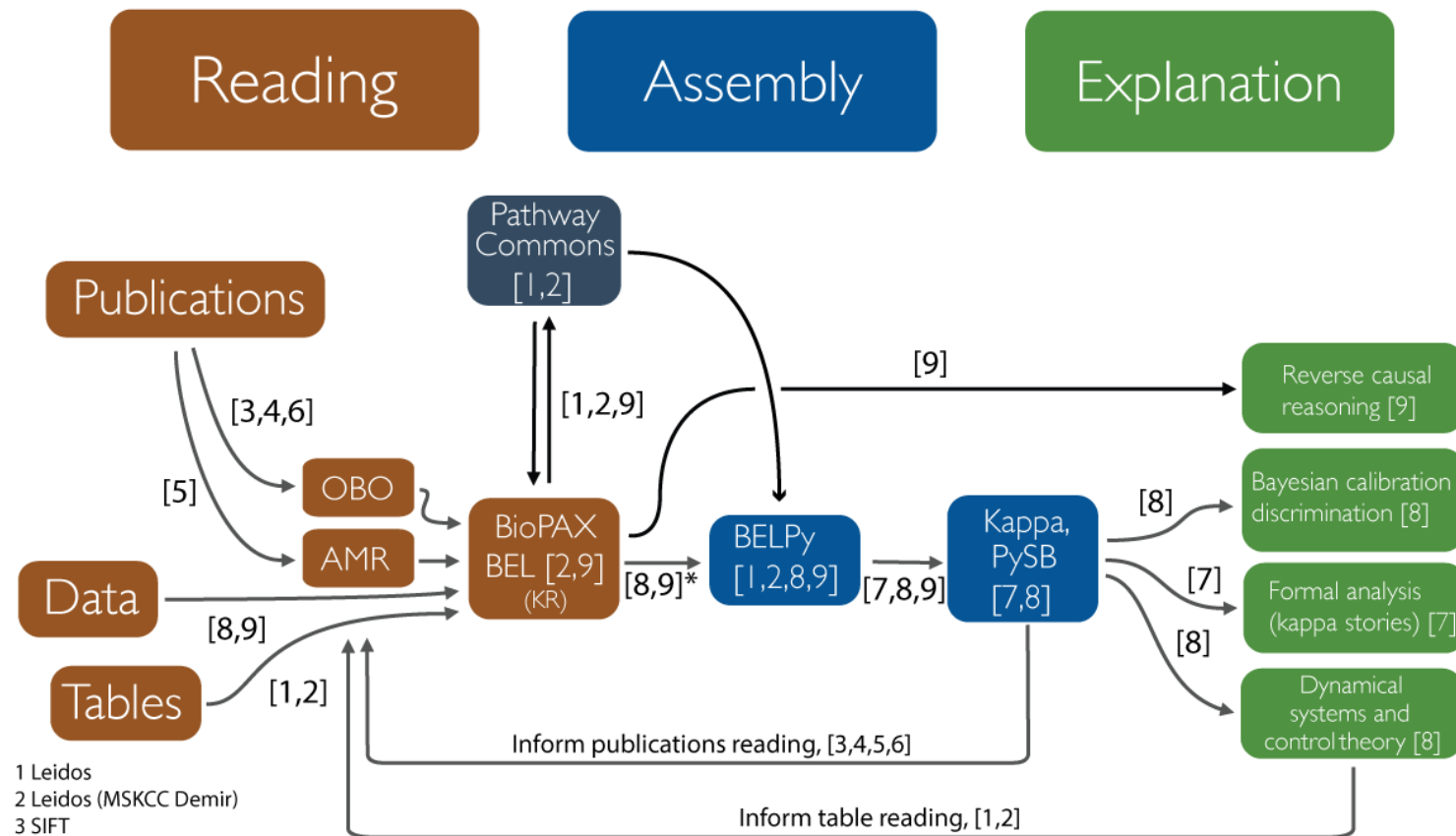


<http://factoid.baderlab.org>

[lunaa@cbio.mskcc.org](mailto:lunaa@cbio.mskcc.org)

10/16/2015

# Novel Data Acquisition Through the DARPA Big Mechanism Project



1 Leidos  
2 Leidos (MSKCC Demir)  
3 SIFT  
4 SIFT (UC Hunter)  
5 ISI  
6 CMU (USC Burns)  
7 HMS\_Fontana  
8 HMS\_Sorger  
9 HMS\_Sorger (UCSD Pratt)

\* OBO - Open Biomedical Ontology, AMR - Abstract Meaning Representation

<http://www.darpa.mil/our-research>

[lunaa@cbio.mskcc.org](mailto:lunaa@cbio.mskcc.org)

10/16/2015

# Future Updates to BioPAX Specification

# BioPAX Level 4

- Last update to BioPAX specification
  - BioPAX Level 3 (2010)
- Missing from BioPAX 3
  - Negative Information
  - Polymerization
  - Disease pathways
  - Multicellular pathways
  - Improved representation of evidence and interaction confidence
- Where are the databases with these types of data?
  - Continued work on BioPAX Level 4 will occur as new databases emerge
  - Text-mining projects offer opportunities to extract new data for existing interactions

# Acknowledgements

- BioPAX Community
- Current Pathway Commons Team
  - Ozgun Babur
  - Emek Demir (OHSU)
  - Onur Sumer
  - Igor Rodchenkov (U of Toronto)
  - Gary Bader (U of Toronto)
  - Chris Sander
- COMBINE Organizers
- NHGRI/DARPA Funding
- Database Maintainers and Developers
  - Reactome
  - NCI-PID
  - BioCyc
  - PantherDB
  - And many others
- Developers of Exporters
  - Andreas Draeger (KEGGTranslator)
  - Huaiyu Mi (CellDesigner)
  - Mark Woon (PharmGKB)
  - And many others

If you know anyone or are interested in the outreach coordinator position, let us know!