Maastricht University

Department of Bioinformatics - BiGCaT

“Where do you go from here?” Semantics of directions in biological pathways

Ryan A. Miller[1], Martina Kutmon[1,2], Egon L. Willighagen[1], Andra Waagmeester[1,4], Jonathan Mélius[1], Anwesha Bohler[1], Nuno Nunes[1], Peter Woollard[3], Chris T. Evelo[1,2]

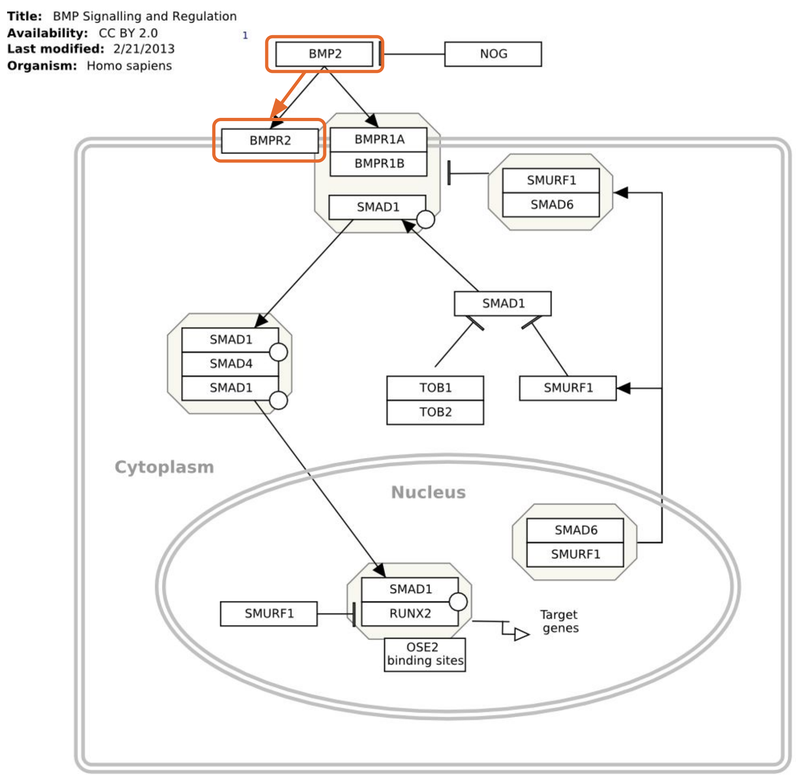
[1] Department of Bioinformatics - BiGCaT, NUTRIM, Maastricht University, The Netherlands
[2] Maastricht Centre for Systems Biology (MaCSBio), Maastricht University, The Netherlands
[3] Computational Biology, GlaxoSmithKline, Stevenage, UK
[4] Micelio, Antwerp, Belgium

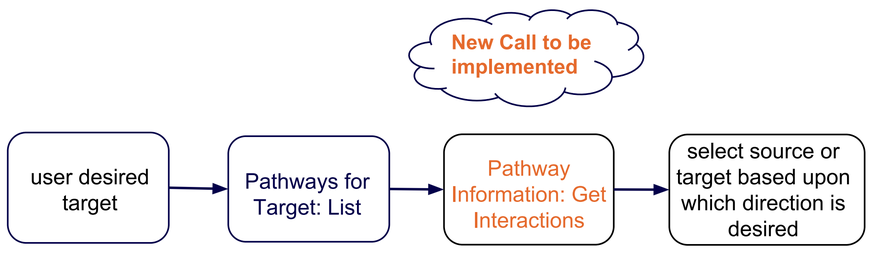
Interactions in the Open PHACTS Discovery platform

The edges between biomolecular graph nodes can have information that is important to describe the interactions. An arrow connecting two molecules in a pathway for instance implies a directed interaction. In the pathway diagram this offers a visual cue for the reader. We described the interactions and their directionality in the new WikiPathways RDF. It is now possible to query the WikiPathways RDF to find information about gene products that are either upstream or downstream of the gene product of interest, in one or more pathways. This, for example, exposes related potential drug targets within the same pathway in a computer readable way. The same process can also be used for metabolites to find products or substrates, or for any other type of data node that is present in a pathway. Now that all this has been done, it is possible to use this new RDF information to traverse a WikiPathways pathway step by step.

WikiPathways and Open PHACTS

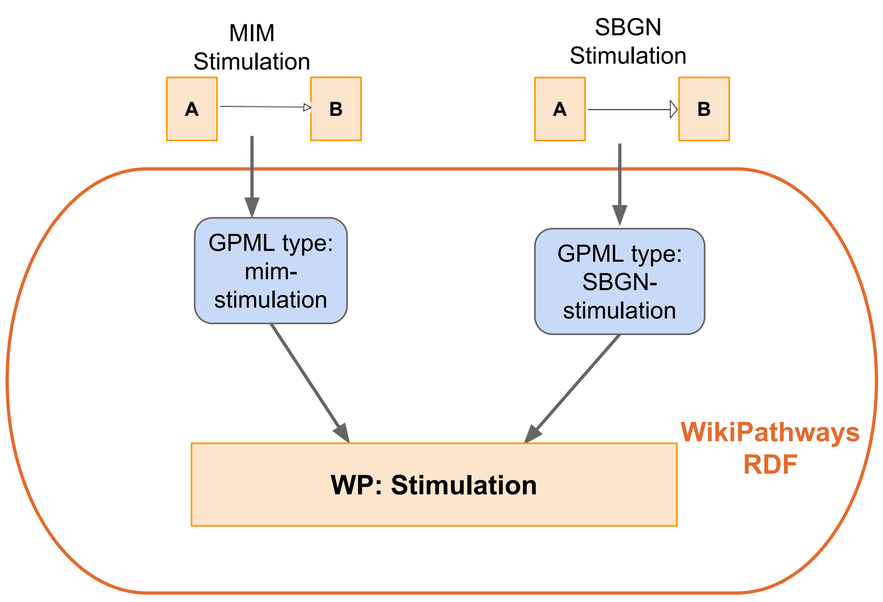
WikiPathways is a core public resource for curating biological pathways. It currently contains 2,236 pathways from more than twenty different species. WikiPathways RDF has proven to be useful for curation purposes and to rapidly find pathway elements. WikiPathways is also useful for finding new routes across the pathway. The WikiPathways RDF is currently loaded into the Open PHACTS platform. Information pertaining to pathways can be obtained through Open PHACTS API calls that can, for instance, find genes of interest and isolate in which pathways they are present. The Open PHACTS platform contains data from many other resources. Which, for instance, allows searching for pathways related to a disease first, then select the proteins in these pathways and finally select drugs targeting these proteins.





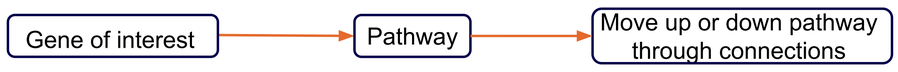
WikiPathways RDF and vocabularies used

The WikiPathways RDF is converted from the GPML (WikiPathways XML) files of the pathways on WikiPathways. The GPML and WP vocabularies are the two definition sets for pathways in WikiPathways that we used. GPML is the graphical vocabulary for WikiPathways RDF (what it looks like), WP is the semantic vocabulary for the RDF (what it means). The RDF is able to handle storage of biomolecular data identified in different resources and recognize that they are related (e. g. a UniProt protein and a Refseq gene). The key addition to WikiPathways RDF that we describe here is the inclusion of information about interactions connecting nodes.



Interaction representation in RDF using WP vocabulary

```
<http://rdf.wikipathways.org/Pathway/WP1425_r74390/WP/Interaction/e077e>
a                   wp:DirectedInteraction , wp:Interaction  ;
dcterms:isPartOf    <http://identifiers.org/wikipathways/WP1425_r74390>  ;
wp:isAbout          <http://rdf.wikipathways.org/Pathway/WP1425_r74390/Interaction/e077e>  ;
wp:participants     <http://identifiers.org/ncbigene/659> , <http://identifiers.org/ensembl/ENSG00000125845> ;
wp:source           <http://identifiers.org/ensembl/ENSG00000125845> ;
wp:target           <http://identifiers.org/ncbigene/659> .
```



From Interaction to Directionality

Having semantic information about how two nodes are connected in a Pathway diagram is already useful information. This means that we know that node A and node B are connected to each other and their interaction has a specific type. This interaction type is a representation of the biological process that it takes to get from one product to another. Directionality is another layer of this information that allows the navigation across the pathway while knowing which product is affecting others. In WikiPathways convention, any sort of directed interaction in RDF has a source and a target. The source being the start of the line and the target being the side that has the arrowhead attached. Knowing the order in which products affect each other is directly applicable in instances such as drug repurposing and target repositioning.

Acknowledgements

The Open PHACTS project is supported by the Innovative Medicines Initiative Joint Undertaking under grant agreement no. 115191, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies’ in kind contribution.

Correspondence to:

✉ ryan.miller@maastrichtuniversity.nl

Department of Bioinformatics - BiGCaT
Maastricht University
P.O. Box 616
6200 MD Maastricht, The Netherlands

Links:

WikiPathways	http://www.wikipathways.org
WikiPathways RDF	http://rdf.wikipathways.org
WikiPathways SPARQL	http://sparql.wikipathways.org
Open PHACTS API	https://dev.openphacts.org