

# Introduction to Pathway Modeling

Slides by: Kristina Hanspers, PhD

Martina Kutmon, PhD

Friederike Ehrhart, PhD

Lauren J. Dupuis, PhD

Maastricht University Department of Bioinformatics – BiGCaT

**GLADSTONE INSTITUTES**  
*SCIENCE OVERCOMING DISEASE*



**Helis Academy**

**Interreg**   
**Vlaanderen-Nederland**  
Europees Fonds voor Regionale Ontwikkeling



# What are biological pathways?

According to the NIH, “A biological pathway is a series of **interactions** among **molecules** in a cell that leads to a certain product or a change in the cell.”



# What are biological pathways?

molecules

DNA



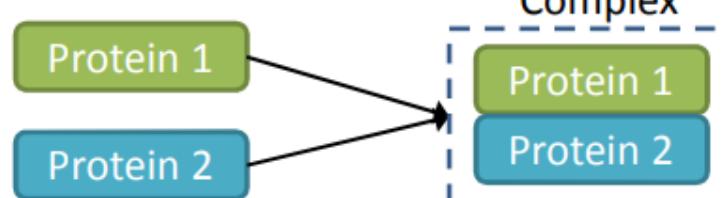
Genomics-25,000 genes



# What are biological pathways?

## interactions

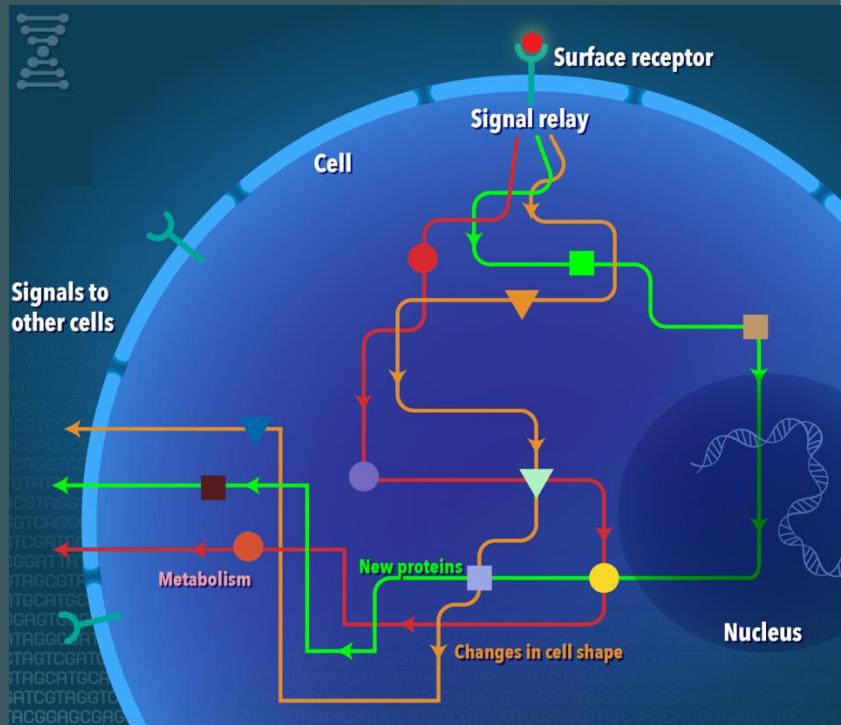
*Complex Assembly:*





# What are biological pathways?

- Signaling pathways
- Metabolic pathways
- Gene regulation pathways



Images:  
[https://www.genome.gov/images/content/biologic\\_pathways.jpg](https://www.genome.gov/images/content/biologic_pathways.jpg)

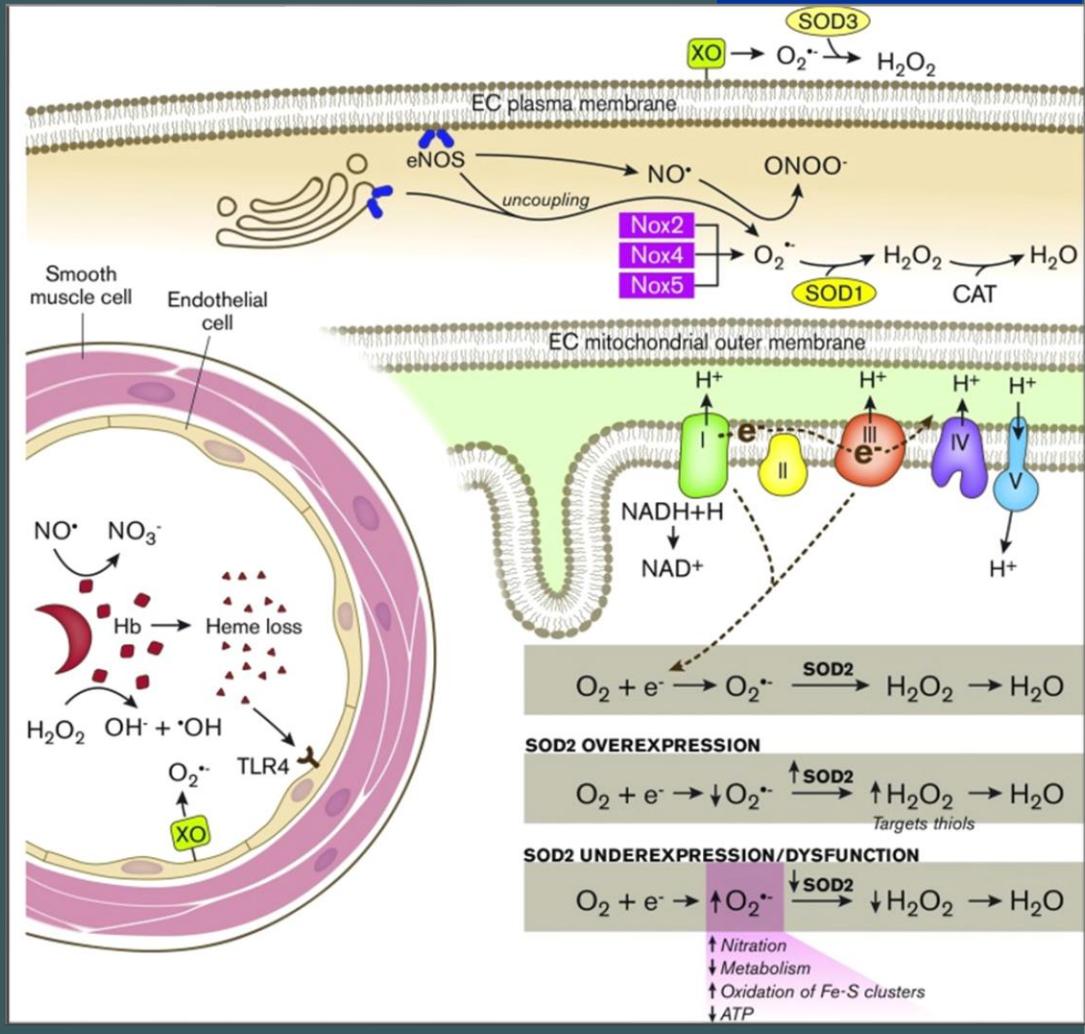
Images:  
[http://www.genome.gov/multimedia/illustrations/Biological\\_Pathways.pdf/](http://www.genome.gov/multimedia/illustrations/Biological_Pathways.pdf/)



# Why pathways?

- Intuitive representation of complex information  
→ Over 1000 published each month
- Proteins / genes / metabolites
- Interactions / reactions
- Subcellular location

Dosunmu-Ogunbi et al, *Blood Advances*, 2019



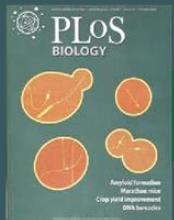
# Why pathway models?



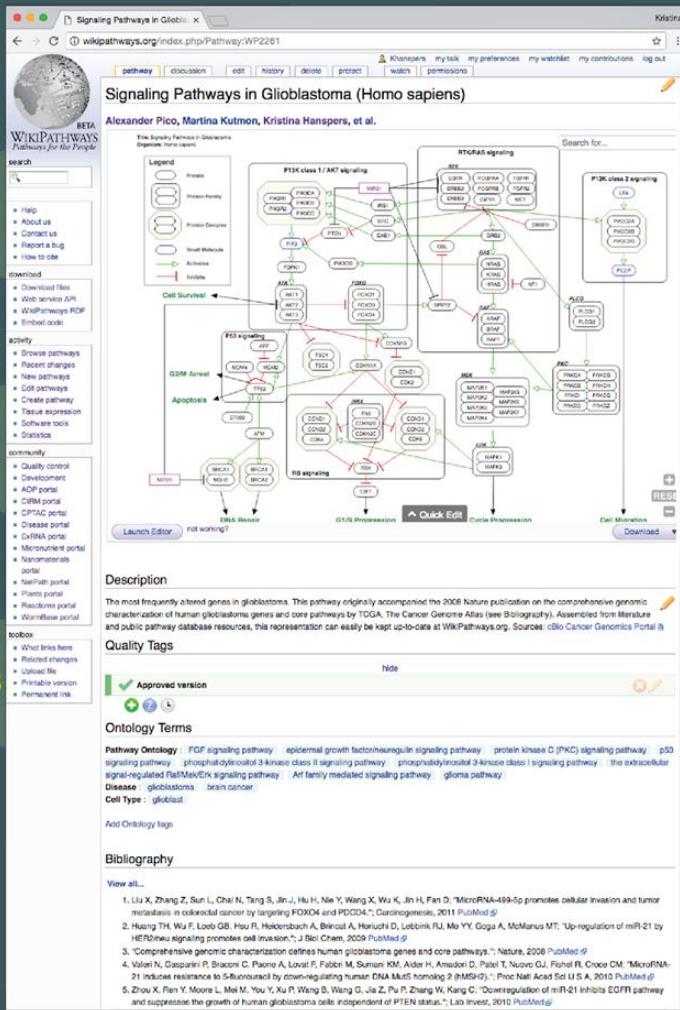
Literature



Experimental data



Static figures



Authorship

Figure images

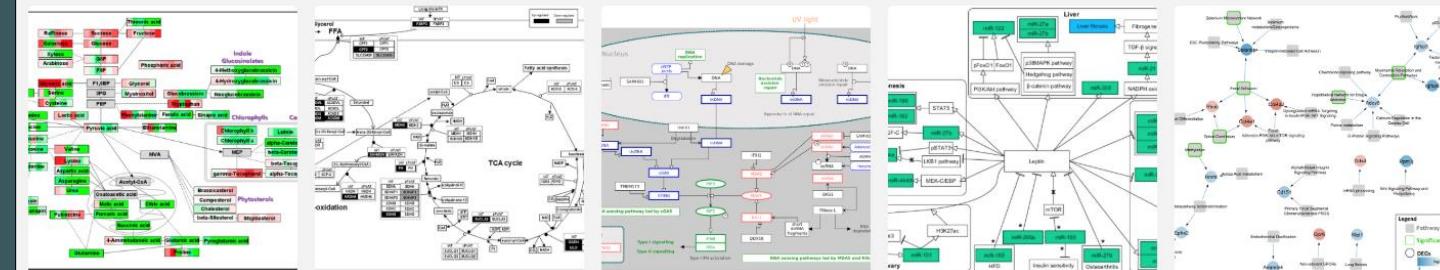
Annotated models

Visualization & analysis

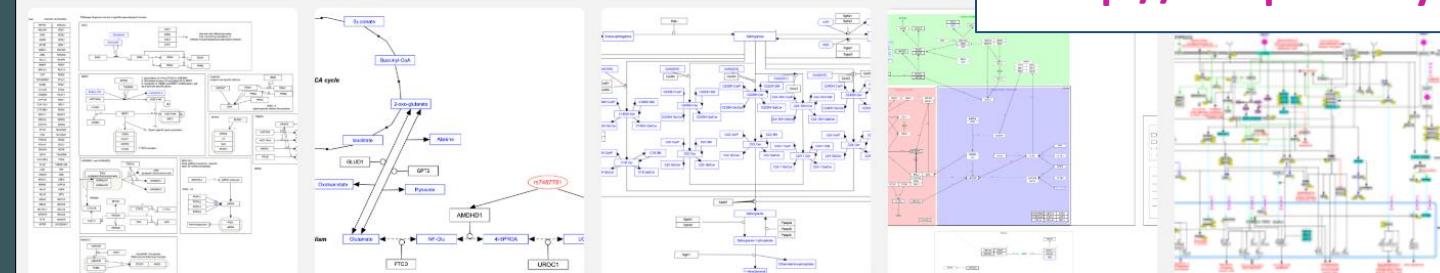
Collaboration

Distributed resource





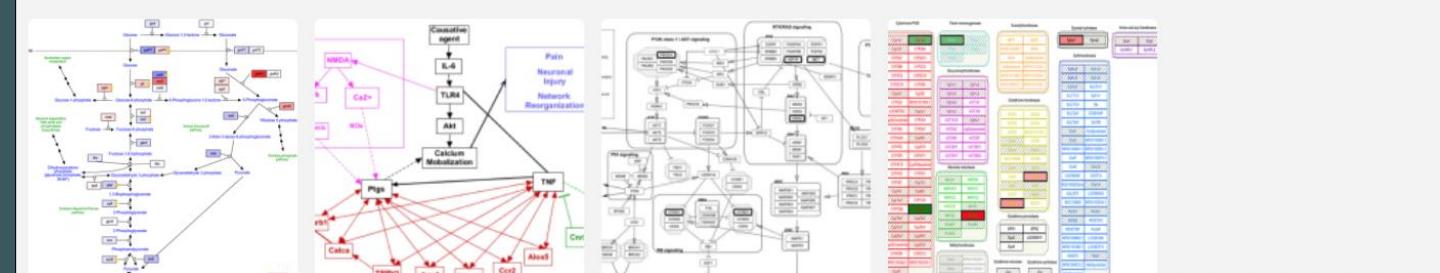
September 2019



## WikiPathways Publications

<http://wikopathways.tumblr.com/>

August 2019





# What pathway tools will we use?

# Editing Tools: WikiPathways and PathVisio



For drawing and editing pathways

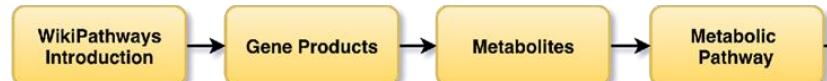


Pathway database



# WikiPathways Academy

## Biology 101



## Walk along a Pathway



## Pathway Building Blocks



## My First Metabolic Pathway



## WikiPathways



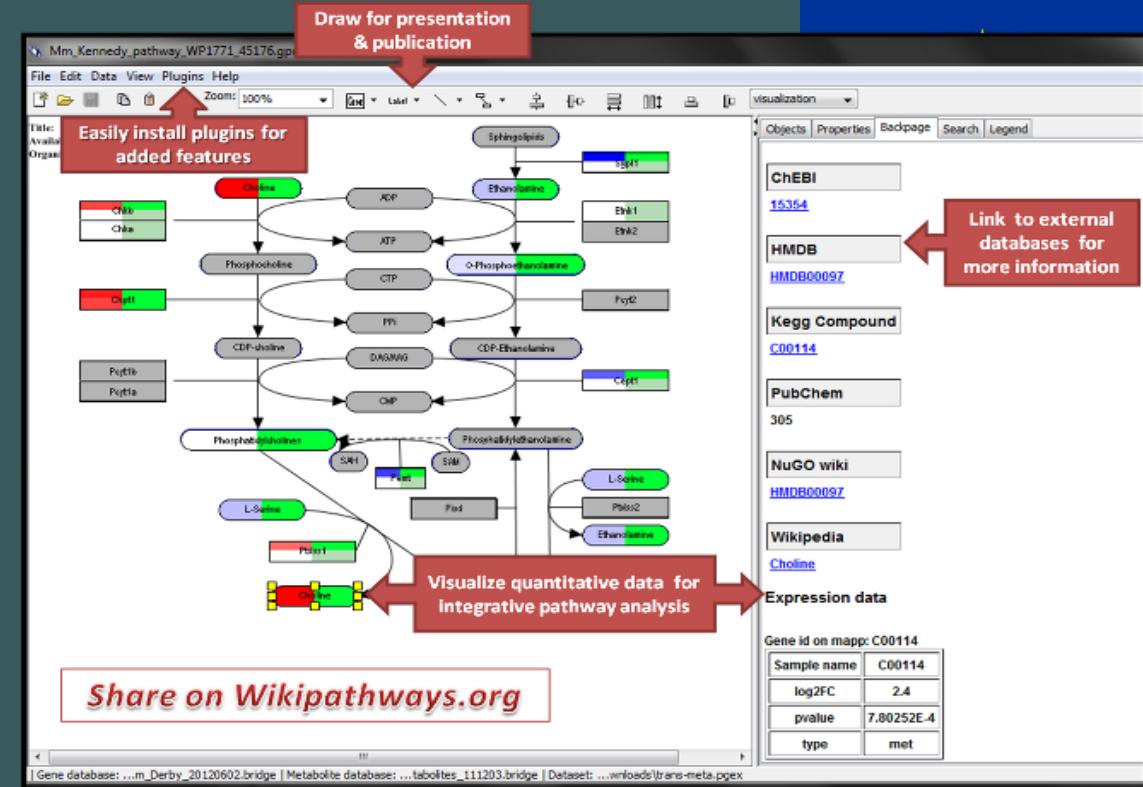
[academy.wikipathways.org](http://academy.wikipathways.org)

## Setting Pathway Information Free



# PathVisio

PathVisio 3.3.0  
[www.pathvisio.org](http://www.pathvisio.org)

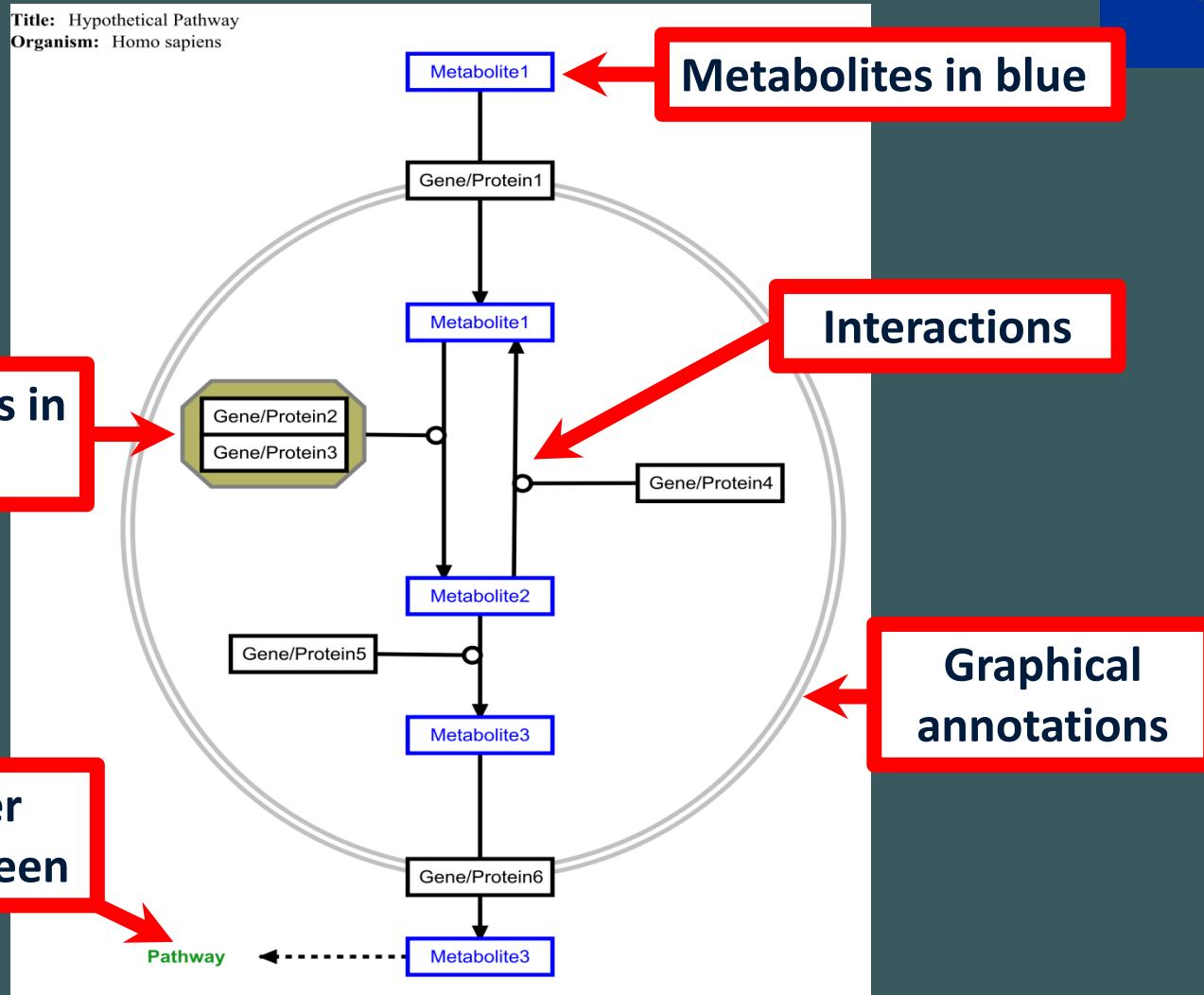


- Java standalone desktop application
- Draw and annotate biological pathways
- Visualize your experimental data on pathways
- Find pathways regulated in your dataset

# Pathway Models in PathVisio



Title: Hypothetical Pathway  
Organism: Homo sapiens



# Pathway Models in PathVisio



More than just images - **fully annotated models**

DataNode properties

Comments Literature Annotation

Search

Manual entry

Text label Adenine

Biological Type Metabolite

Identifier HMDB00034

Database HMDB

Accession Number HMDB00034

Common Name Adenine

Description

Adenine is a purine base. Adenine is found in both DNA and RNA forms adenosine, a nucleoside, when attached to ribose triphosphate (ATP), a nucleotide, when three phosphate groups are added. Adenine is involved in metabolism as one of the basic methods of transferring energy in the body. Adenine is also involved in serious hereditary disorders, which should be considered in the differential diagnosis of neurological deficit, renal disease, self-mutilation and other symptoms. (OMIM 300322, 229600, 603027, 17052198, 17520339)

DataNode properties

Comments Literature Annotation

Search

Manual entry

Text label p53

Biological Type GeneProduct

Identifier ENSG00000141510

Database Ensembl

Cancel Ok

links to other databases

e!Ensembl

Home > Human

Location: 17:7,512,445-7,531,642 Gene: TP53

Gene: TP53 (ENSG00000141510)

Gene summary

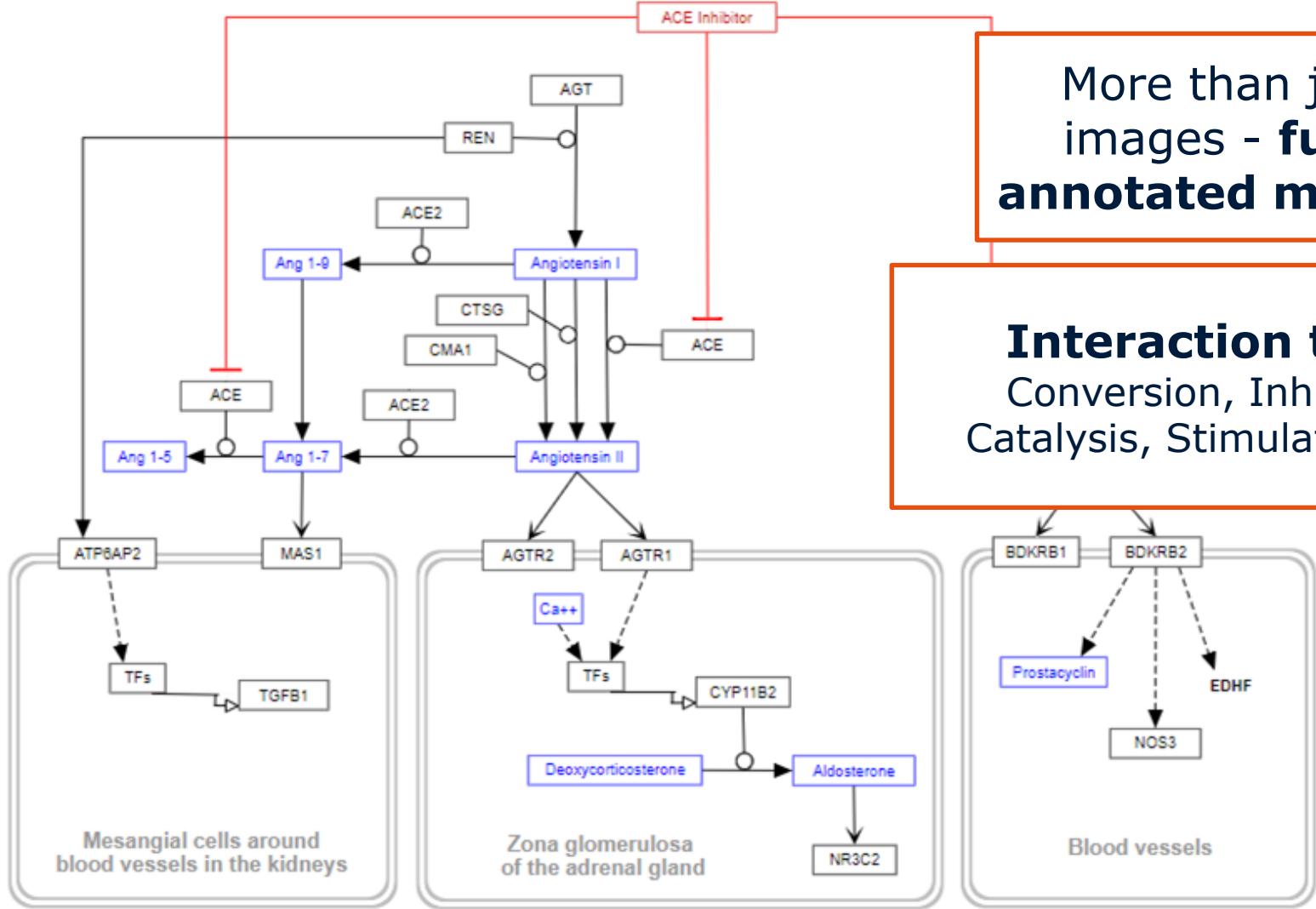
- Splice variants (3)
- Supporting evidence
- Sequence
- External references (7)
- Regulation
- Comparative Genomics
  - Genomic alignments (36)

Transcripts

There are 3 transcripts in this gene: TP53-201, TP53-202, TP53-203

TP53-201	ENST00000434410
TP53-202	ENST00000434411
TP53-203	ENST00000434412

# Pathway Models in PathVisio



More than just images - **fully annotated models**

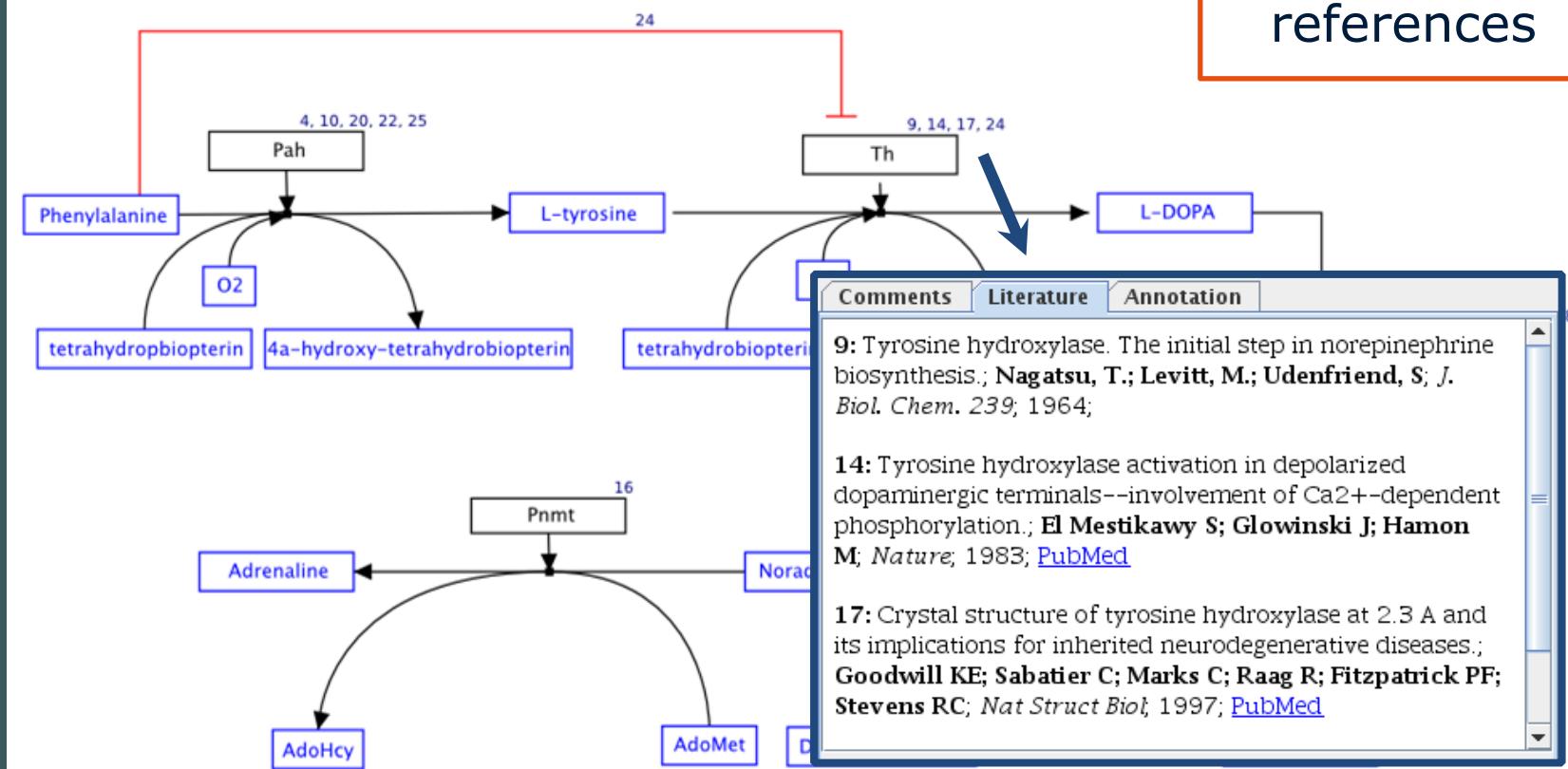
**Interaction types**  
Conversion, Inhibition,  
Catalysis, Stimulation, etc.



# Pathway Models in PathVisio

More than just  
images - **fully  
annotated models**

literature  
references





# Upload to WikiPathways

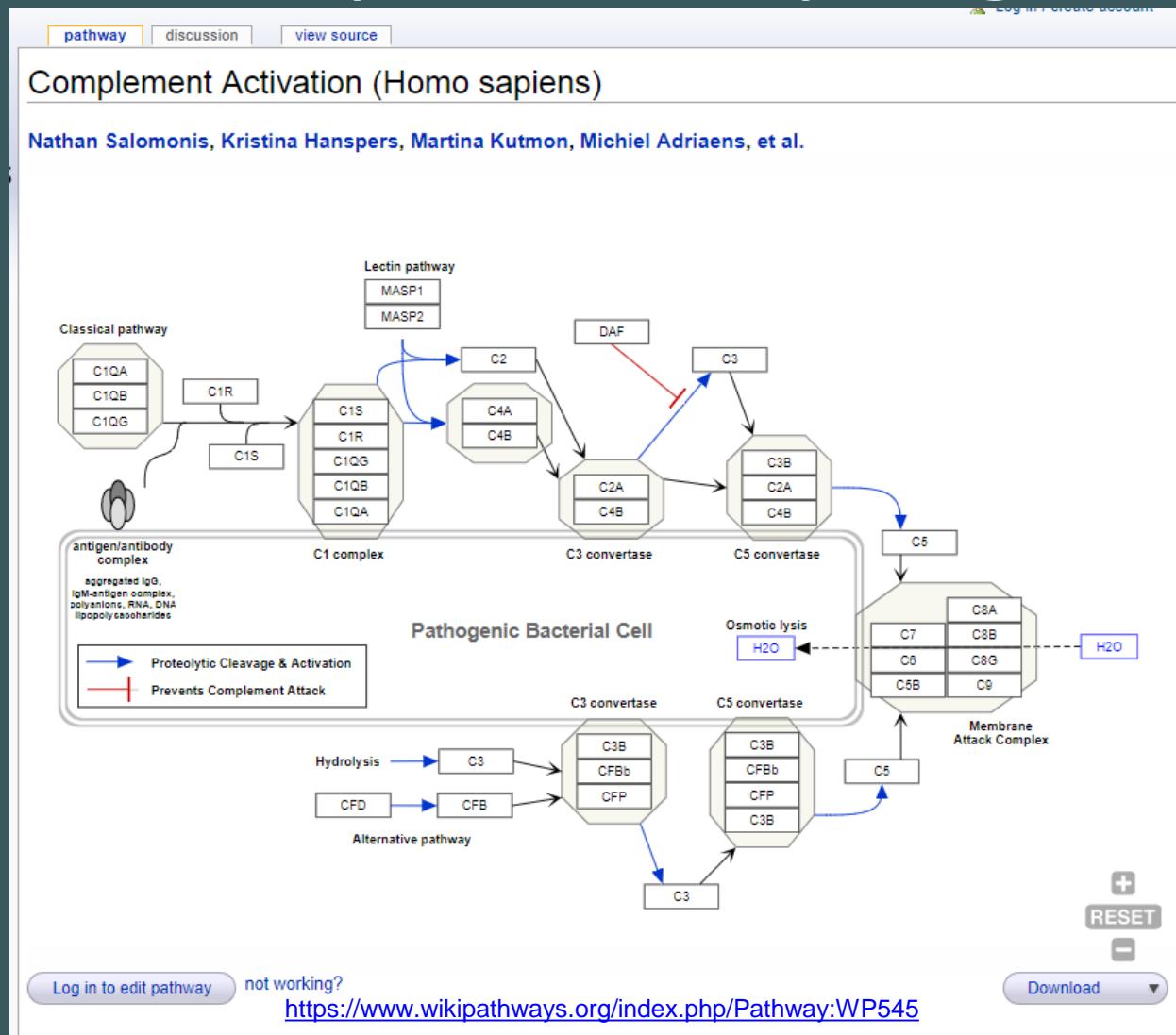
- New pathways can be directly uploaded to WikiPathways through the WikiPathways plugin
- Plugin also allows curators to search for, load and update existing pathways



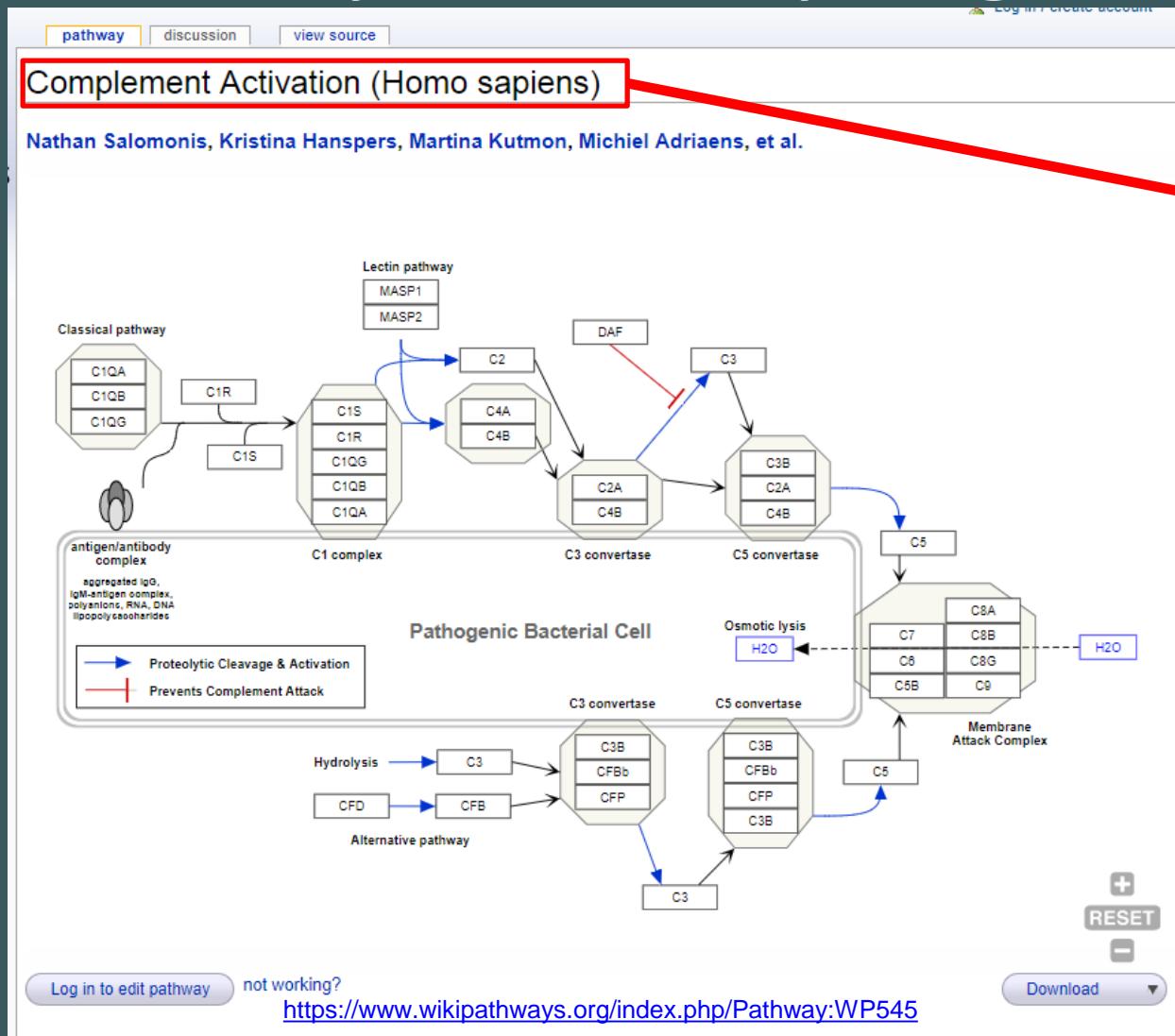
developed by Sravanthi Sinah



# WikiPathways Pathway Pages



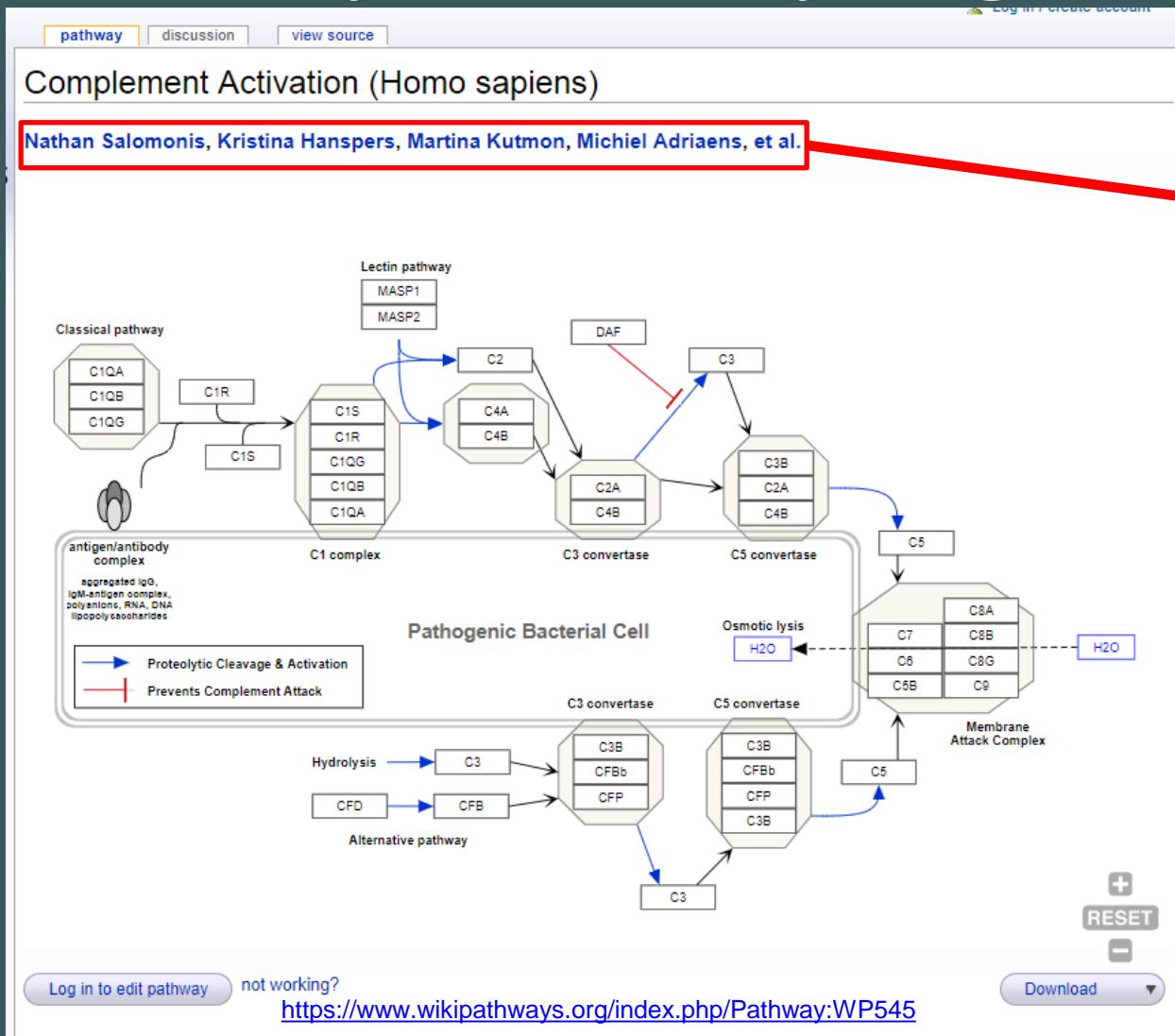
# WikiPathways Pathway Pages



Title /  
Organism

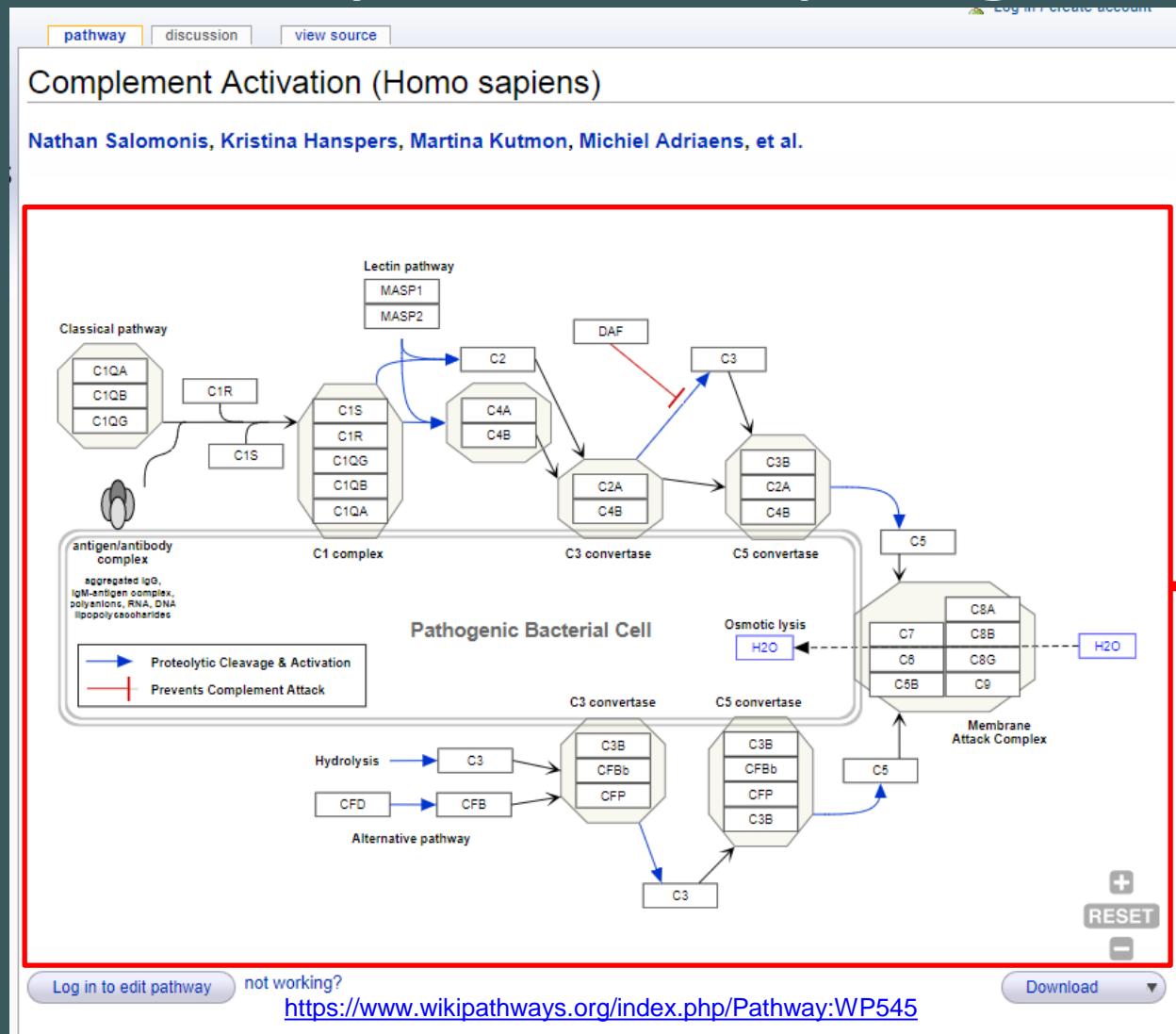


# WikiPathways Pathway Pages



List of Curators

# WikiPathways Pathway Pages



Interactive  
Pathway  
Viewer



# WikiPathways Pathway Pages

## Description

The complement system is a biochemical cascade that helps, or complements, the ability of antibodies to clear pathogens from an organism. It is part of the innate immune system that is not adaptable and does not change over the course of an individual's lifetime. However, it can be recruited and brought into action by antibodies.

Classical pathway of activation of the complement system is a group of blood proteins that mediate the specific antibody response. [source: Wikipedia]

The Classical pathway begins with circulating C1Q binding to an antigen on the surface of a pathogen, which goes on to activate and recruit 2 copies of C1. The activated C1 complex cleaves C2 and C4. Activated cleavage products C2A and C4B combine to form C3 convertase, which cleaves C3. The cleaved C3 fragments form C5 convertase, which cleaves C5. The cleavage product C5B joins C6, C7, C8 and multiple copies of C9 to form the Membrane Attack Complex, which inserts into the target cell, leading to osmotic lysis. The Decay accelerating factor (DAF) inhibits C3 convertase.

The Lectin pathway involves mannose-binding lectin (MBL) binding the surface of the pathogen instead of C1Q. MBL-associated serine proteases MASP1 and MASP2 substitute for the C1 complex, leading to the formation of C3 convertase and the subsequent cascade.

The Alternative pathway relies on the spontaneous hydrolysis of C3 and the cleavage of factor B (CFB) by factor D (CFD), which form an alternative C3 convertase. Additional copies of the cleavage product C3B are recruited to the complex, resulting in an alternative C5 convertase, which cleaves C5 and contributes to the formation of the Membrane Attack Complex.

**Pathway description**

## Quality Tags

[hide](#)

- Featured version**
- Approved version**
- 

## Ontology Terms

Pathway Ontology : [innate immune response pathway](#) [immune response pathway](#) [classical complement pathway](#)



# WikiPathways Pathway Pages

## Description

The complement system is a biochemical cascade that helps, or complements, the ability of antibodies to clear pathogens from an organism. It is part of the innate immune system that is not adaptable and does not change over the course of an individual's lifetime. However, it can be recruited and brought into action by antibodies.

Classical pathway of activation of the complement system is a group of blood proteins that mediate the specific antibody response. [source: Wikipedia]

The Classical pathway begins with circulating C1Q binding to an antigen on the surface of a pathogen, which goes on to activate and recruit 2 copies of C1. The activated C1 complex cleaves C2 and C4. Activated cleavage products C2A and C4B combine to form C3 convertase, which cleaves C3. The cleaved C3 fragments are C3a and C3b. C3b joins C5 to form C5 convertase, which cleaves C5. The cleavage product C5B joins C6, C7, C8 and multiple copies of C9 to form the Membrane Attack Complex (MAC). MAC embeds into the target cell, leading to osmotic lysis. The Decay accelerating factor (DAF) inhibits C3 convertase.

The Lectin pathway involves mannose-binding lectin (MBL) binding the surface of the pathogen instead of C1Q. MBL-associated serine proteases MASP1 and MASP2 act as the equivalents of C1Q, in place of the C1 complex, leading to the formation of C3 convertase and the subsequent cascade.

The Alternative pathway relies on the spontaneous hydrolysis of C3 and the cleavage of factor B (CFB) by factor D (CFD), which form an alternative C3 convertase. Additional copies of the cleavage product C3B are recruited to the complex, resulting in an alternative C5 convertase, which cleaves C5 and contributes to the formation of the Membrane Attack Complex.

## Quality Tags

[hide](#)

- Featured version**
- Approved version**
- 

**Quality  
Tags**

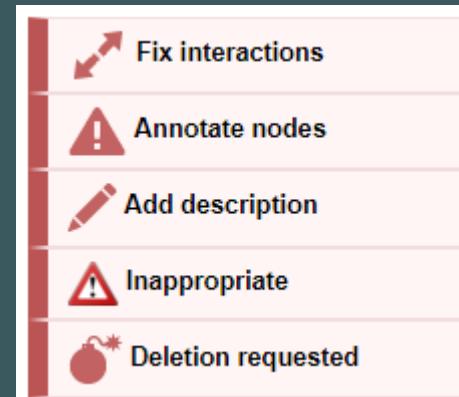
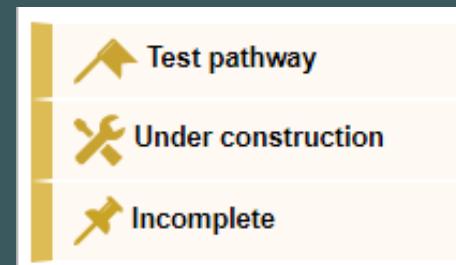
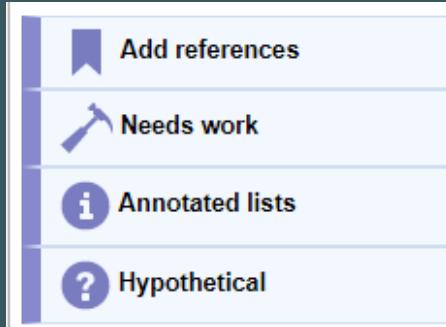
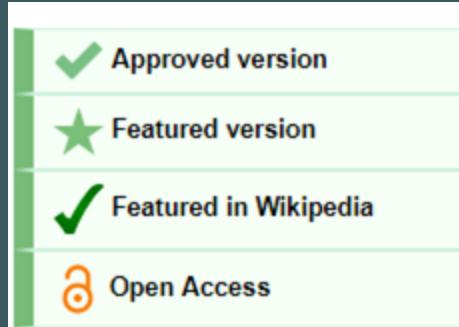
## Ontology Terms

Pathway Ontology : [innate immune response pathway](#) [immune response pathway](#) [classical complement pathway](#)



# Quality Tags

- Information about the state of the pathway
  - Is the pathway included in any collections?
  - How complete is the pathway?





# WikiPathways Pathway Pages

## Description

The complement system is a biochemical cascade that helps, or complements, the ability of antibodies to clear pathogens from an organism. It is part of the immune system that is not adaptable and does not change over the course of an individual's lifetime. However, it can be recruited and brought into action.

Classical pathway of activation of the complement system is a group of blood proteins that mediate the specific antibody response. [source: Wikipedia]

The Classical pathway begins with circulating C1Q binding to an antigen on the surface of a pathogen, which goes on to activate and recruit 2 copies of C1. The activated C1 complex cleaves C2 and C4. Activated cleavage products C2A and C4B combine to form C3 convertase, which cleaves C3. The cleaved C3 fragments are joined by C4B to form C5 convertase, which cleaves C5. The cleavage product C5B joins C6, C7, C8 and multiple copies of C9 to form the Membrane Attack Complex, which inserts into the target cell, leading to osmotic lysis. The Decay accelerating factor (DAF) inhibits C3 convertase.

The Lectin pathway involves mannose-binding lectin (MBL) binding the surface of the pathogen instead of C1Q. MBL-associated serine proteases MASP1 and MASP2 act in place of the C1 complex, leading to the formation of C3 convertase and the subsequent cascade.

The Alternative pathway relies on the spontaneous hydrolysis of C3 and the cleavage of factor B (CFB) by factor D (CFD), which form an alternative C3 convertase. Additional copies of the cleavage product C3B are recruited to the complex, resulting in an alternative C5 convertase, which cleaves C5 and contributes to the formation of the Membrane Attack Complex.

## Quality Tags

[hide](#)

- Featured version
- Approved version
- 

## Ontology Terms

Pathway Ontology : [innate immune response pathway](#) [immune response pathway](#) [classical complement pathway](#)

Ontology  
Terms



# Ontology Terms

- Controlled vocabularies to categorize the pathway
- Support of three ontologies
  - <http://bioportal.bioontology.org/>
  - Pathway, disease and cell type ontology



# Ontology Terms

## Ontology Terms

Pathway Ontology : innate immune response pathway immune response pathway classical complement pathway

[Hide Ontology Options](#)

Type Ontology term..

To add a tag, either select from the available ontology trees below or type a search term in the search box.

### Pathway Ontology

- + classic metabolic pathway
- + disease pathway
- + drug pathway
- + regulatory pathway
- + signaling pathway

### Disease

- + disease by infectious agent
- + disease of anatomical entity
- + disease of cellular proliferation
- + disease of mental health
- + disease of metabolism
- + genetic disease
- + physical disorder
- + syndrome

### Cell Type

- + abnormal cell
- + cell in vitro
- + compound eye retinal cell
- + native cell
- + neural cell
- + neuron



# WikiPathways Pathway Pages

## Bibliography

1. Brook E, Herbert AP, Jenkins HT, Soares DC, Barlow PN; "Opportunities for new therapies based on the natural history of C1Q nephropathy". *Nephrol Dial Transplant*. 2016;31(10):1713-1718.

## History

[View all...](#)

[Compare selected versions](#)

Compare	Revision	Action	Time	User	Comment
<input checked="" type="radio"/>	89636 <small>Featured Approved</small>	<a href="#">view</a>	06:10, 22 September 2016	Egonw	Connected lines.
<input checked="" type="radio"/> <input type="radio"/>	87260	<a href="#">view</a>	14:39, 20 July 2016	MaintBot	missing graphids
<input type="radio"/> <input checked="" type="radio"/>	82142	<a href="#">view</a>	01:23, 9 September 2015	AlexanderPico	fixed C1Q complex
<input type="radio"/> <input checked="" type="radio"/>	82141	<a href="#">view</a>	01:19, 9 September 2015	AlexanderPico	Modified description
<input type="radio"/> <input checked="" type="radio"/>	82140	<a href="#">view</a>	01:13, 9 September 2015	AlexanderPico	Modified title

[Compare selected versions](#)

## External references

## DataNodes

[View all...](#)

Name	Type	Database reference	Comment
C1QA	GeneProduct	<a href="#">712 (Entrez Gene)</a> ⓘ	update: C1QA HUMAN
C1QB	GeneProduct	<a href="#">713 (Entrez Gene)</a> ⓘ	
C1QG	GeneProduct	<a href="#">714 (Entrez Gene)</a> ⓘ	update C1QC HUMAN
C1R	GeneProduct	<a href="#">ENSG00000159403 (Ensembl)</a> ⓘ	
C1S	GeneProduct	<a href="#">716 (Entrez Gene)</a> ⓘ	

## Annotated Interactions

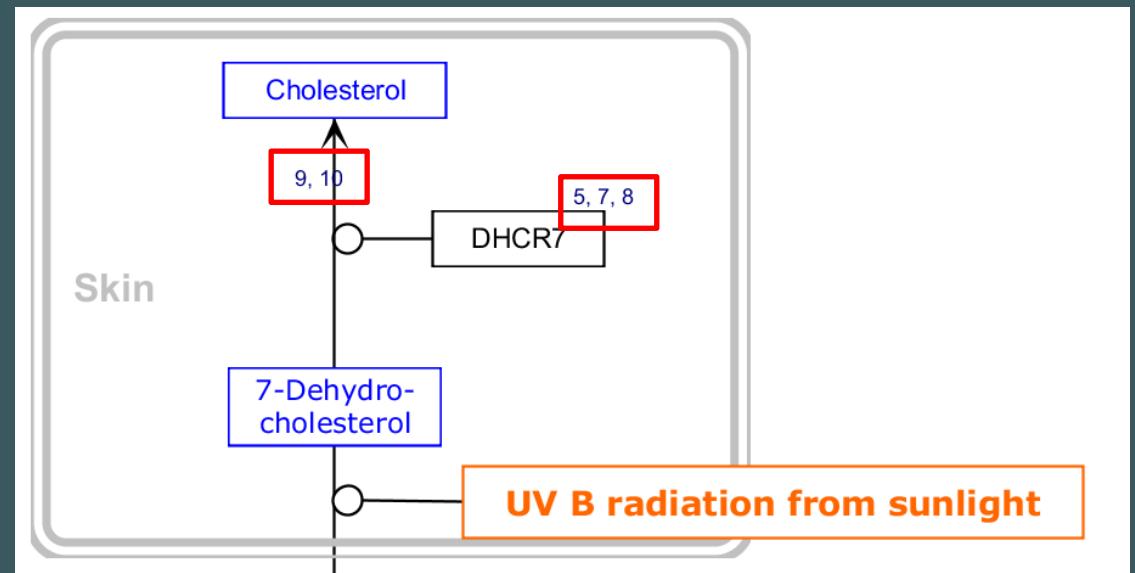
No annotated interactions

**Bibliography**



# Bibliography

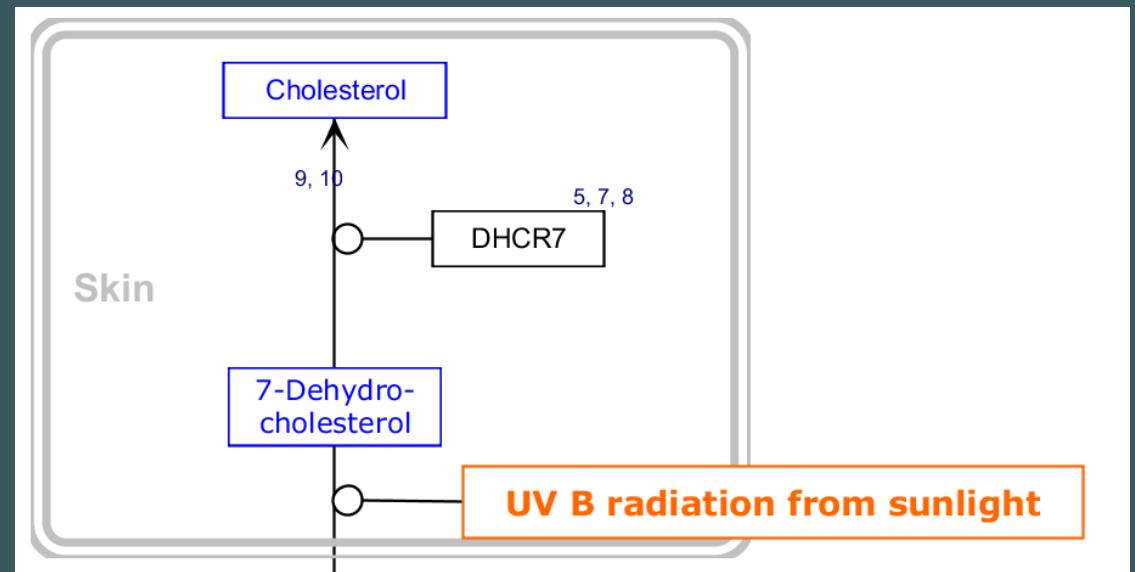
- Every element of the pathway can be linked to literature references including





# Bibliography

- Every element of the pathway can be linked to literature references including
  - Data nodes
  - Interactions
  - Pathway





# WikiPathways Pathway Pages

## Bibliography

- Brook E, Herbert AP, Jenkins HT, Soares DC, Barlow PN; "Opportunities for new therapies based on the natural history of rare diseases."

## History

[View all...](#)

[Compare selected versions](#)

Compare	Revision	Action	Time	User	Comment
<input checked="" type="radio"/>	89636 Featured Approved	<a href="#">view</a>	06:10, 22 September 2016	Egonw	Connected lines.
<input checked="" type="radio"/> <input type="radio"/>	87260	<a href="#">view</a>	14:39, 20 July 2016	MaintBot	missing graphids
<input type="radio"/> <input checked="" type="radio"/>	82142	<a href="#">view</a>	01:23, 9 September 2015	AlexanderPico	fixed C1Q complex
<input type="radio"/> <input checked="" type="radio"/>	82141	<a href="#">view</a>	01:19, 9 September 2015	AlexanderPico	Modified description
<input type="radio"/> <input checked="" type="radio"/>	82140	<a href="#">view</a>	01:13, 9 September 2015	AlexanderPico	Modified title

[Compare selected versions](#)

## External references

## DataNodes

[View all...](#)

Name	Type	Database reference	Comment
C1QA	GeneProduct	<a href="#">712 (Entrez Gene)</a>	update: C1QA HUMAN
C1QB	GeneProduct	<a href="#">713 (Entrez Gene)</a>	
C1QG	GeneProduct	<a href="#">714 (Entrez Gene)</a>	update C1QC HUMAN
C1R	GeneProduct	<a href="#">ENSG00000159403 (Ensembl)</a>	
C1S	GeneProduct	<a href="#">716 (Entrez Gene)</a>	

## Annotated Interactions

No annotated interactions

History



# WikiPathways Pathway Pages

## Bibliography

- Brook E, Herbert AP, Jenkins HT, Soares DC, Barlow PN; "Opportunities for new therapies based on the natural history of rare diseases".

## History

[View all...](#)

[Compare selected versions](#)

Compare	Revision	Action	Time	User	Comment
<input checked="" type="radio"/>	89636 Featured Approved	view	06:10, 22 September 2016	Egonw	Connected lines.
<input checked="" type="radio"/> <input type="radio"/>	87260	view	14:39, 20 July 2016	MaintBot	missing graphids
<input type="radio"/> <input checked="" type="radio"/>	82142	view	01:23, 9 September 2015	AlexanderPico	fixed C1Q complex
<input type="radio"/> <input checked="" type="radio"/>	82141	view	01:19, 9 September 2015	AlexanderPico	Modified description
<input type="radio"/> <input checked="" type="radio"/>	82140	view	01:13, 9 September 2015	AlexanderPico	Modified title

[Compare selected versions](#)

## External references

## DataNodes

[View all...](#)

Name	Type	Database reference	Comment
C1QA	GeneProduct	712 (Entrez Gene)	update: C1QA HUMAN
C1QB	GeneProduct	713 (Entrez Gene)	
C1QG	GeneProduct	714 (Entrez Gene)	update C1QC HUMAN
C1R	GeneProduct	ENSG00000159403 (Ensembl)	
C1S	GeneProduct	716 (Entrez Gene)	

## Annotated Interactions

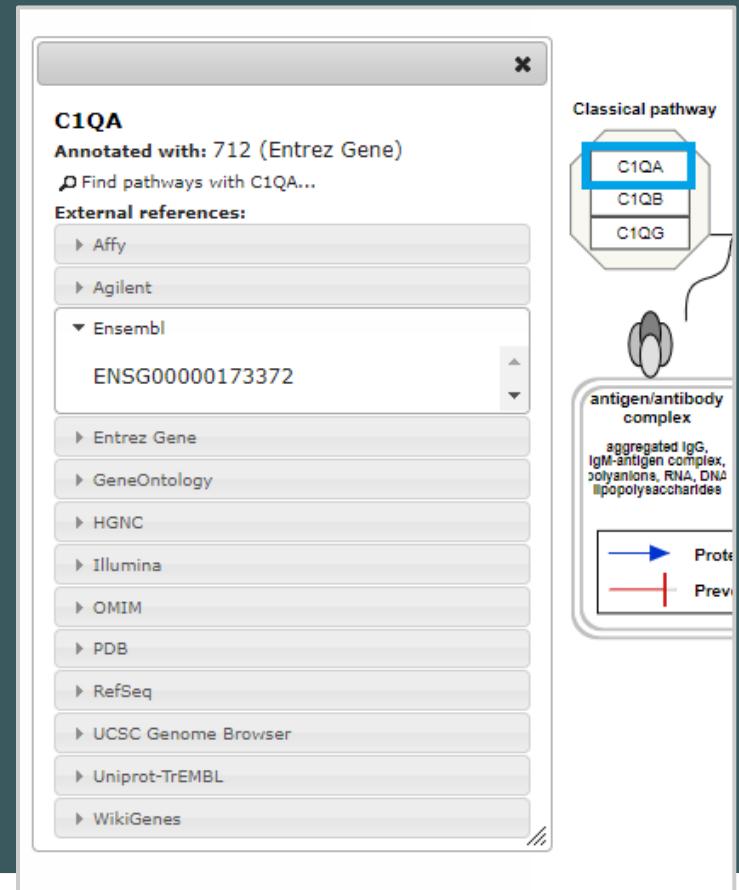
No annotated interactions

External  
References

# External References

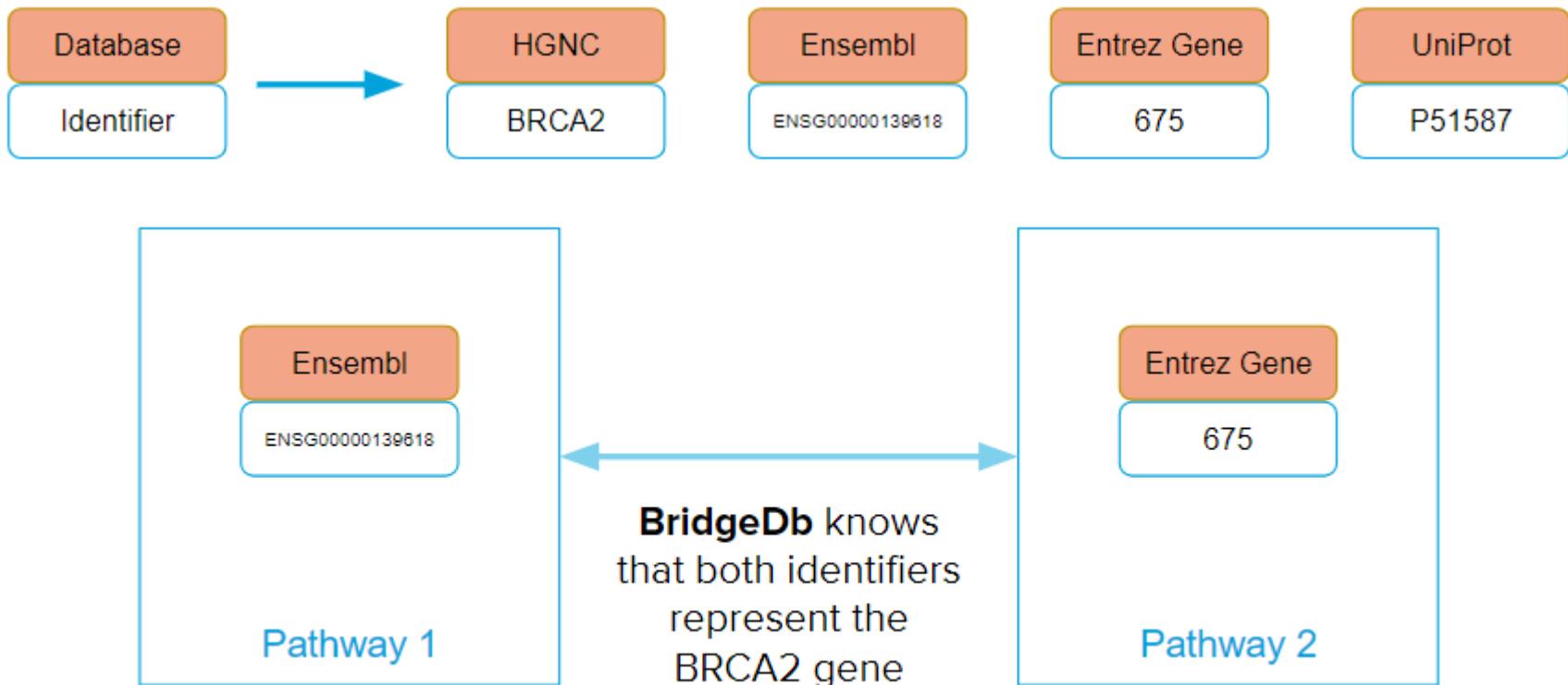


- Powered by BridgeDb
  - Mapping databases available for gene products, metabolites and interactions
  - Identifier mapping framework
  - [www.bridgedb.org](http://www.bridgedb.org)
- Pathway curators can use their identifier system of choice





# External References





# Community Portals

WikiPathways BETA  
Pathways for the People

page discussion view source history

Welcome to WikiPathways BETA

WikiPathways is a database of biological pathways maintained by a community of experts.

Find Pathways

Search

You can search by:

- Pathway name (*Apoptosis*)
- Gene or protein name (*p53*)
- Any page content (*cancer*)

Get Pathways

Download

Multiple formats and methods

Edit Pathways

Learn

Play with the editor tools

community

- Quality control
- Development
- WikiPathways Blog
- AOP portal
- CIRM portal
- CPTAC portal
- Renal Genomics portal
- Disease portal
- ExRNA portal
- Lipids portal
- Micronutrient portal
- Nanomaterials portal
- NetPath portal
- Plants portal
- Rare Disease portal
- Reactome portal
- WormBase portal

Today's Featured Pathway

Synaptic Vesicle Pathway (*Homo sapiens*)

Curator of the Week

Susan Coort (Maastricht University)

Updates

- November 2019 Release: 172 edits by 19 contributors and 15 new pathways.
- October 2019 Release: 728 edits by 16 contributors and 5 new pathways
- September 2019 Release: 905 edits by 15 contributors and 5 new pathways
- August 2019 Release: 416 edits by 13 contributors and 8 new pathways
- July 2019 Release: 284 edits by 21 contributors and 18 new pathways
- June 2019 Release: 270 edits by 29 contributors and 4 new pathways
- May 2019 Release: 243 edits by 23 contributors and 14 new pathways
- April 2019 Release: 213 edits by 20 contributors and 17 new pathways

more...

Discussions

9 November 2019

Welcome to Wikopathways! (12) by Daniela Digles

more...

Forum

Re: [wp-discuss] I kindly request to help me in importing and mapping ID

INTERREG VLAANDEREN-NEDERLAND

Maastricht Universiteit

EUROPEES FONDS VOOR REGIONALE ONTWIKKELING

COINT PROGRAMME

DISEASES

Helis Academy



# Community Portals

portal discussion view source history Log in / create account

## Portal:RareDisease

### Rare Disease pathways at WikiPathways

This portal highlights WikiPathways content related to rare genetic human diseases, and is designed as a central organizing point for exploring, curating and expanding the collection of rare disease pathways.

A rare or orphan disease is a disease which affects relatively few people. The exact definition varies between 1:1000 and 1:200.000. E.g. it can be 1:1500 (USA) , 1:2000 (EU) or 1:2500 (Japan). Rare diseases are mostly caused by genetic variation making them chronic and hard to cure. The severity depends on the affected gene and its physiological implications.

This portal is funded by:

**EUROPEAN JOINT PROGRAMME  
RARE DISEASES**

**elixir**

**Netherlands  
RETT  
EXPERTISE CENTRE**

**Rare Disease Pathways**

**Featured Pathway**

**Neurodegeneration with brain iron accumulation (NBIA) subtypes pathway (Homo sapiens)**

**Explore rare disease pathways on WikiPathways**

- MECP2 and Associated Rett Syndrome
- Amyotrophic lateral sclerosis (ALS)
- Fanconi Anemia Pathway
- The effect of progerin on the involved genes in Hutchinson-Gilford Progeria Syndrome
- Lamin A-processing pathway
- Viral Acute Myocarditis

edit

beta

WIKIPATHWAYS  
Pathways for the People

search

Help  
About us  
Contact us  
Report a bug  
How to cite

download

Download files  
Web service API  
WikiPathways RDF  
Embed code

activity

Browse pathways  
Recent changes  
New pathways  
Edit pathways  
Create pathway  
Statistics

tools

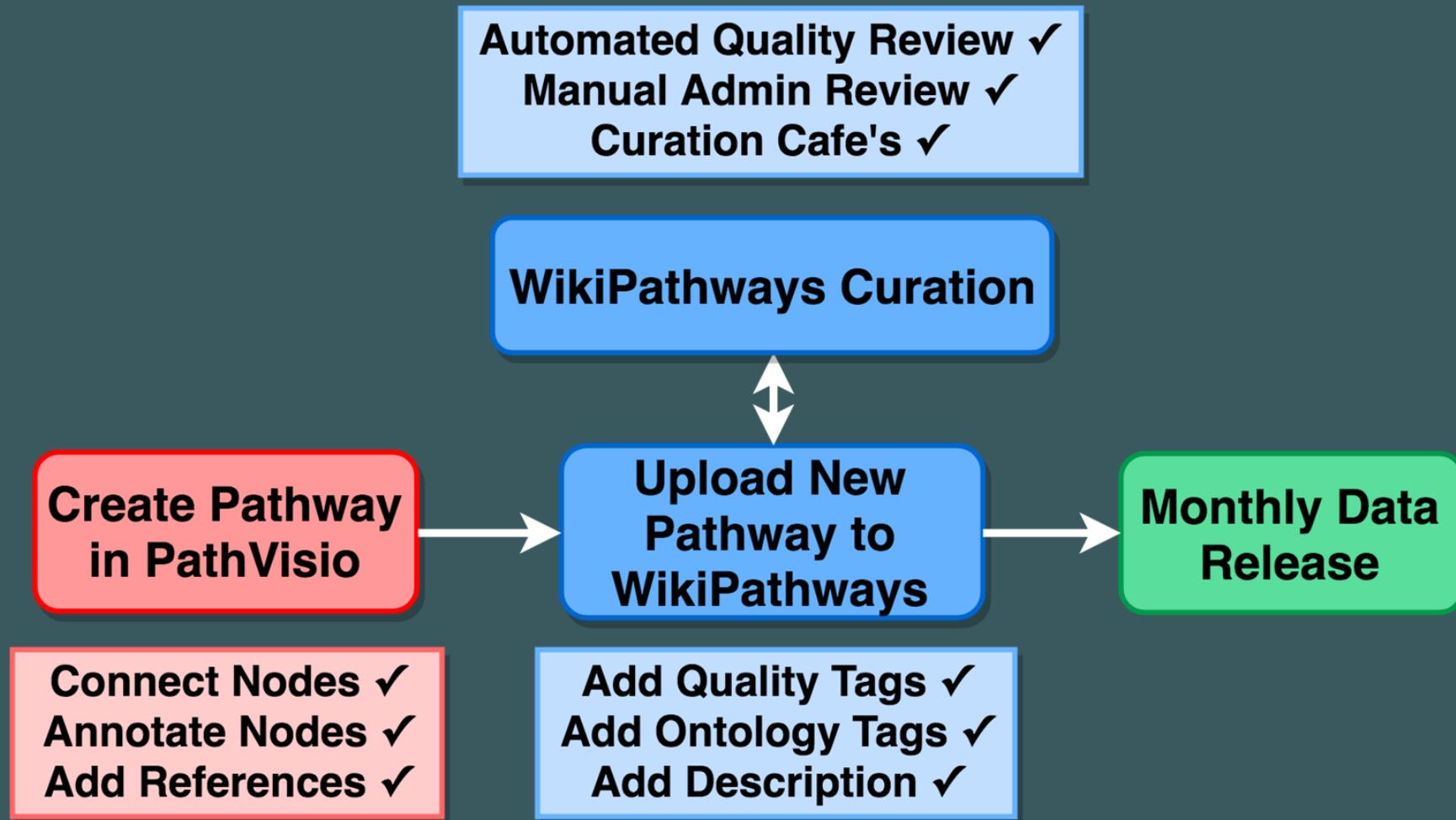
PathwayWidget  
Pathway Finder  
Software tools

community

Quality control  
Development  
WikiPathways Blog  
AOP portal



# Life-Cycle of a Pathway Model at WikiPathways



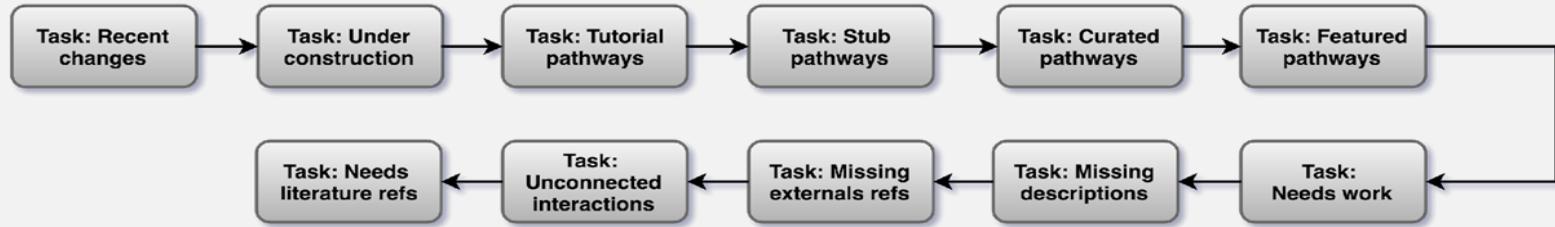


# Quality Control Protocol

## QA Team



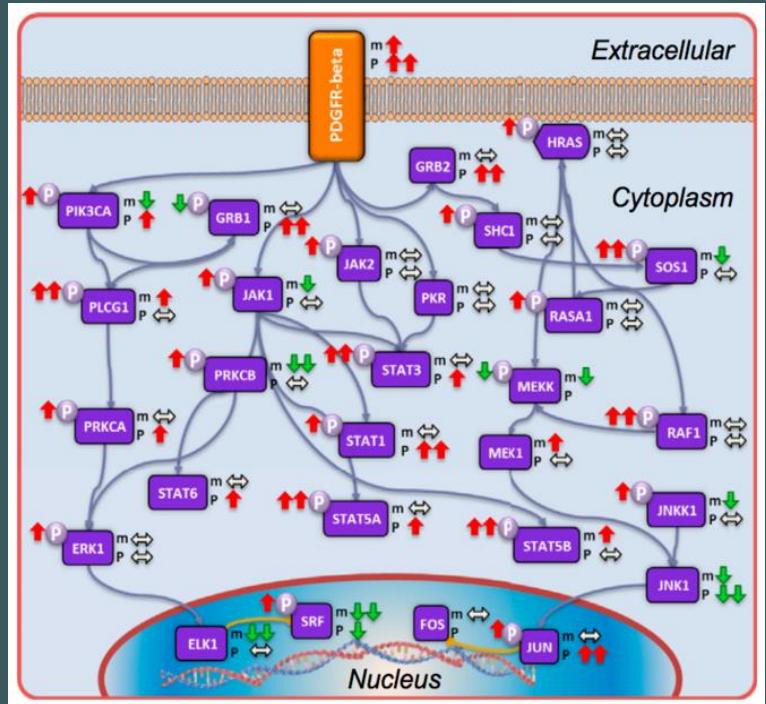
### QA Protocol



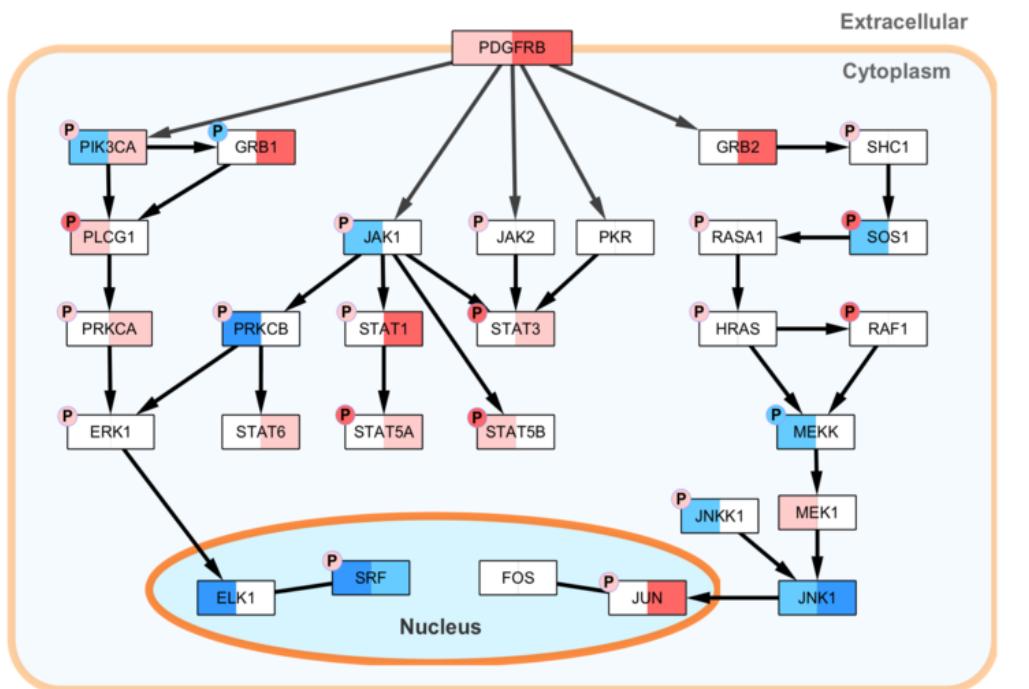
<https://wikopathways.github.io/academy/qaprotocol.html>



# Machine readability



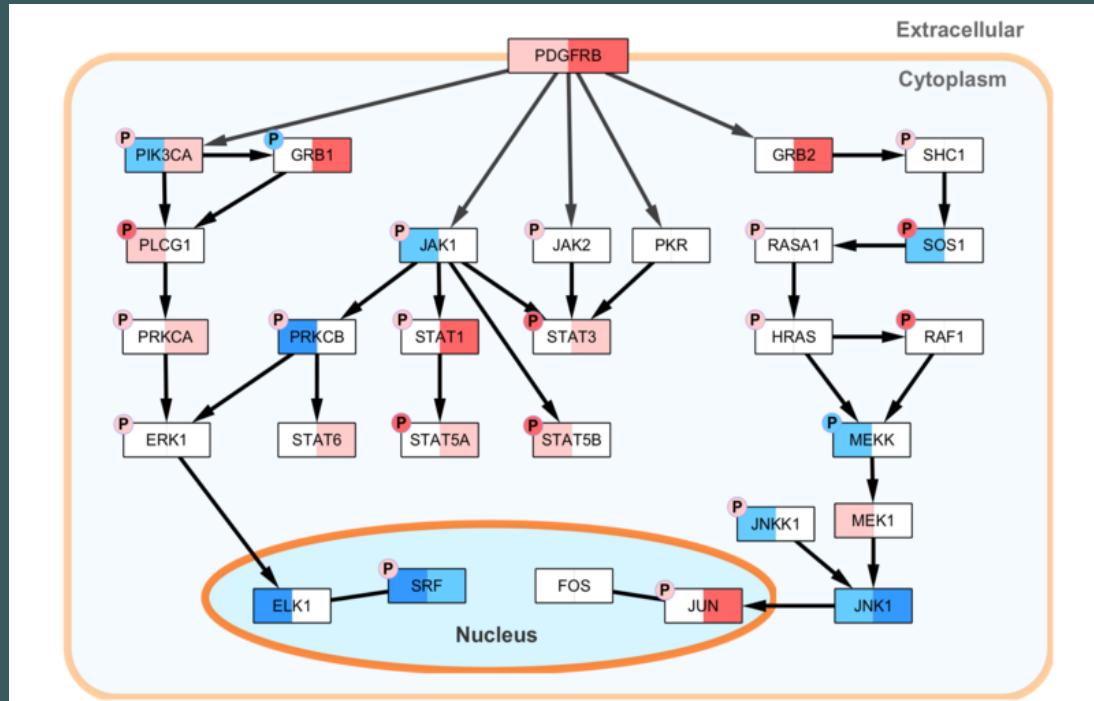
Static image  
Zhang et al, Cell 2016



PDGFR-beta pathway with  
transcriptomic/phosphoproteomic data  
Zhang et al, Cell 2016



# Machine readability



PDGFR-beta pathway with  
transcriptomic/phosphoproteomic data  
Zhang et al, *Cell* 2016



# Machine readability



[www.wikipathways.org](http://www.wikipathways.org)

How do we use  
them for analysis?



# Data Import

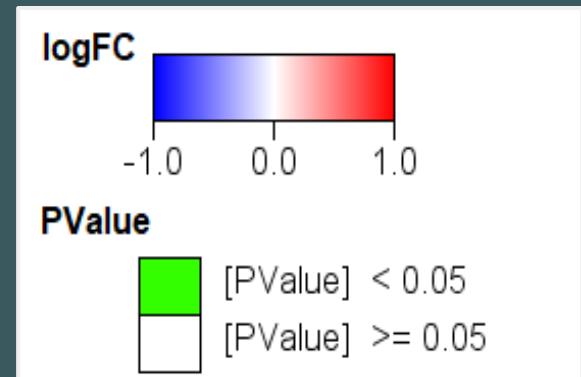


- Wizard guiding you through data import
- Multiple identifier systems supported
  - Ensembl, NCBI Gene, Affymetrix, Illumina, HGNC
  - Focus on gene expression for practical
  - PathVisio can handle other types of omics data
- BridgeDb takes care of ID mapping (.bridge file)
- Database (.pgex), log file and default visualization are created for each imported dataset

# Data Visualizations

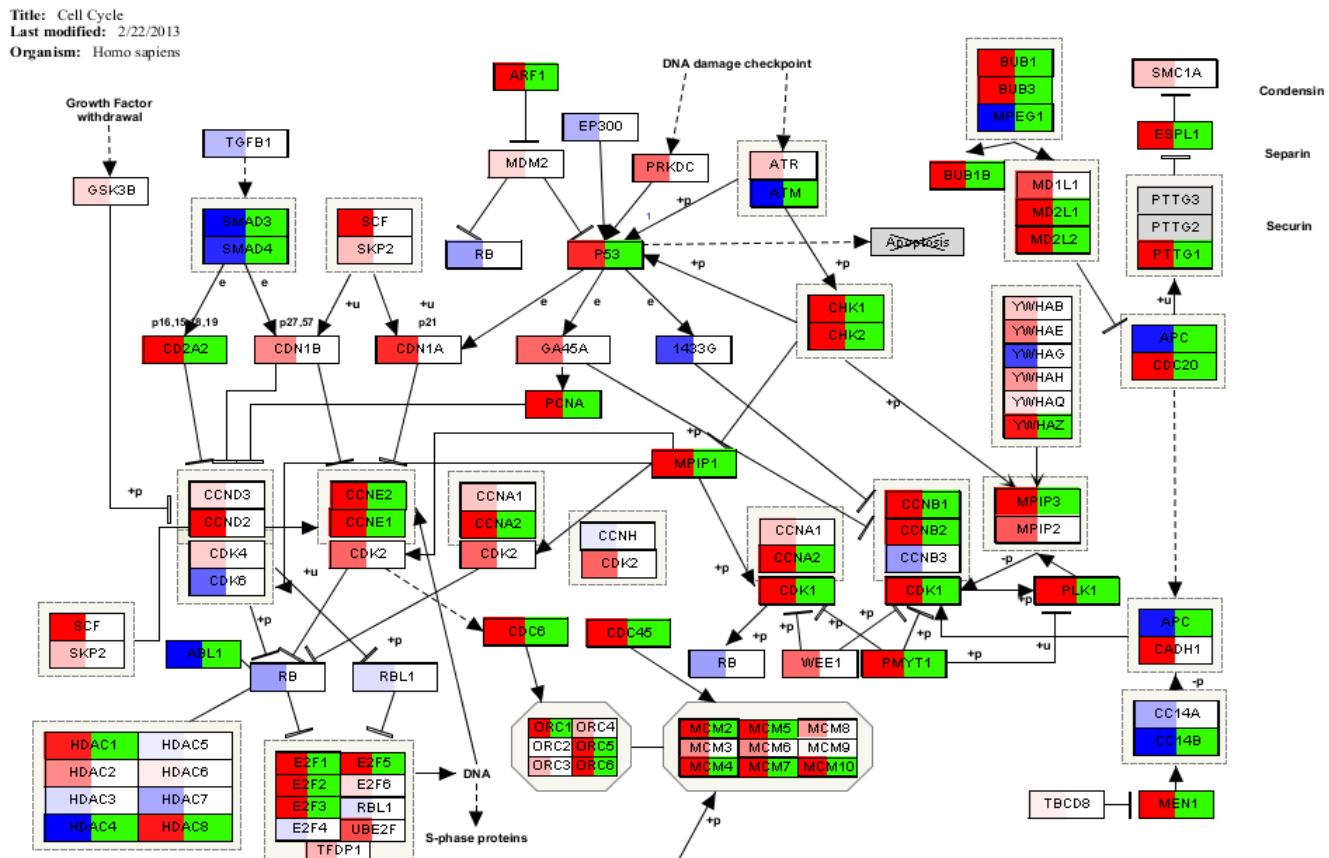
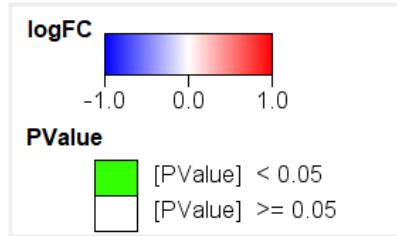


- Data visualization on data nodes and interactions
- Visualization styles
  - Color gradient (continuous data like log2FC)
  - Color rules (distinct categories like pval cutoff)
- Multi-omics visualization
- Time-series visualization





# Data Visualizations

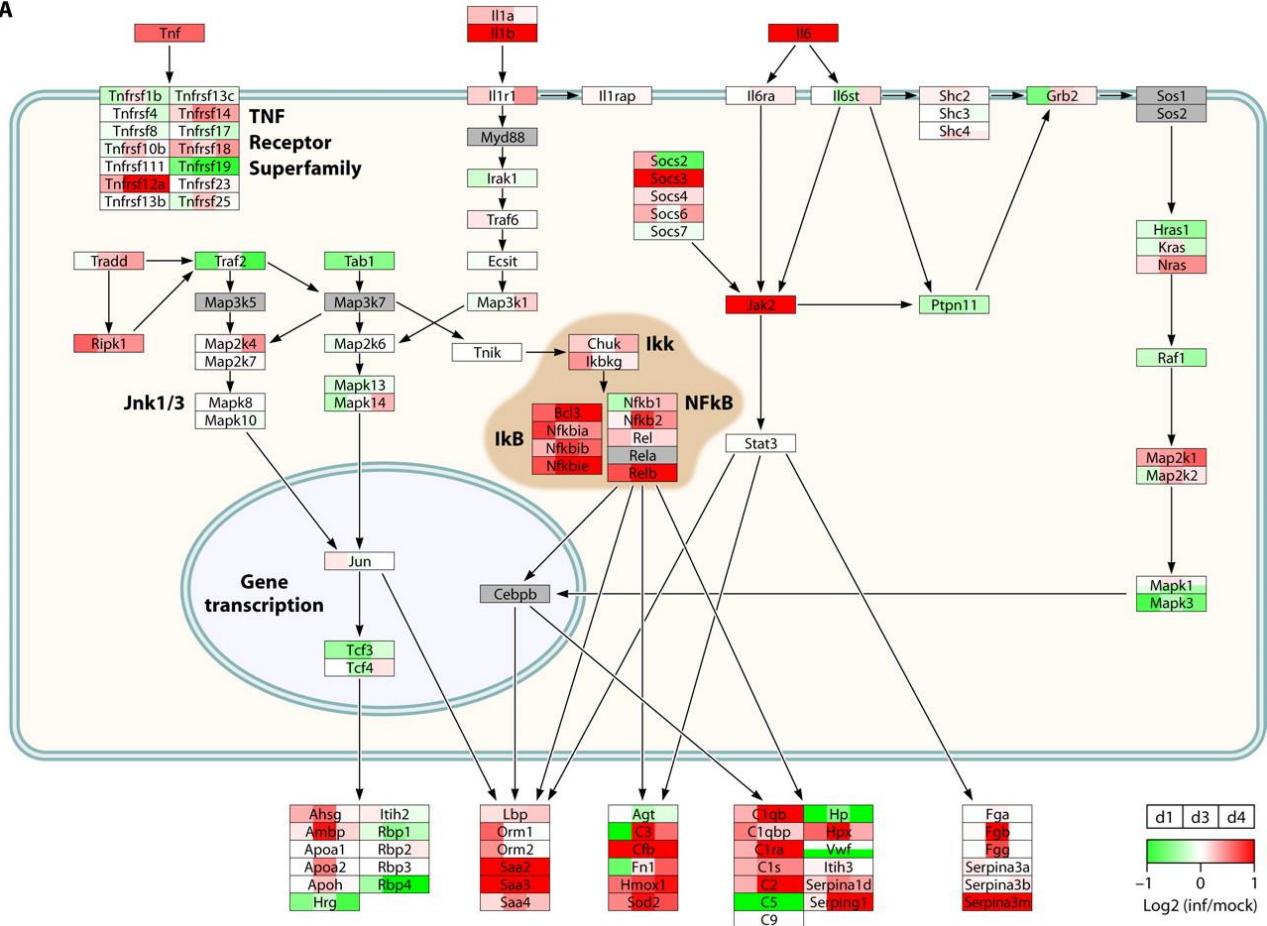


Transcriptomics cancer dataset visualized on the  
human Cell Cycle Pathway:  
<https://www.wikipathways.org/instance/WP179>



# Data Visualizations

A

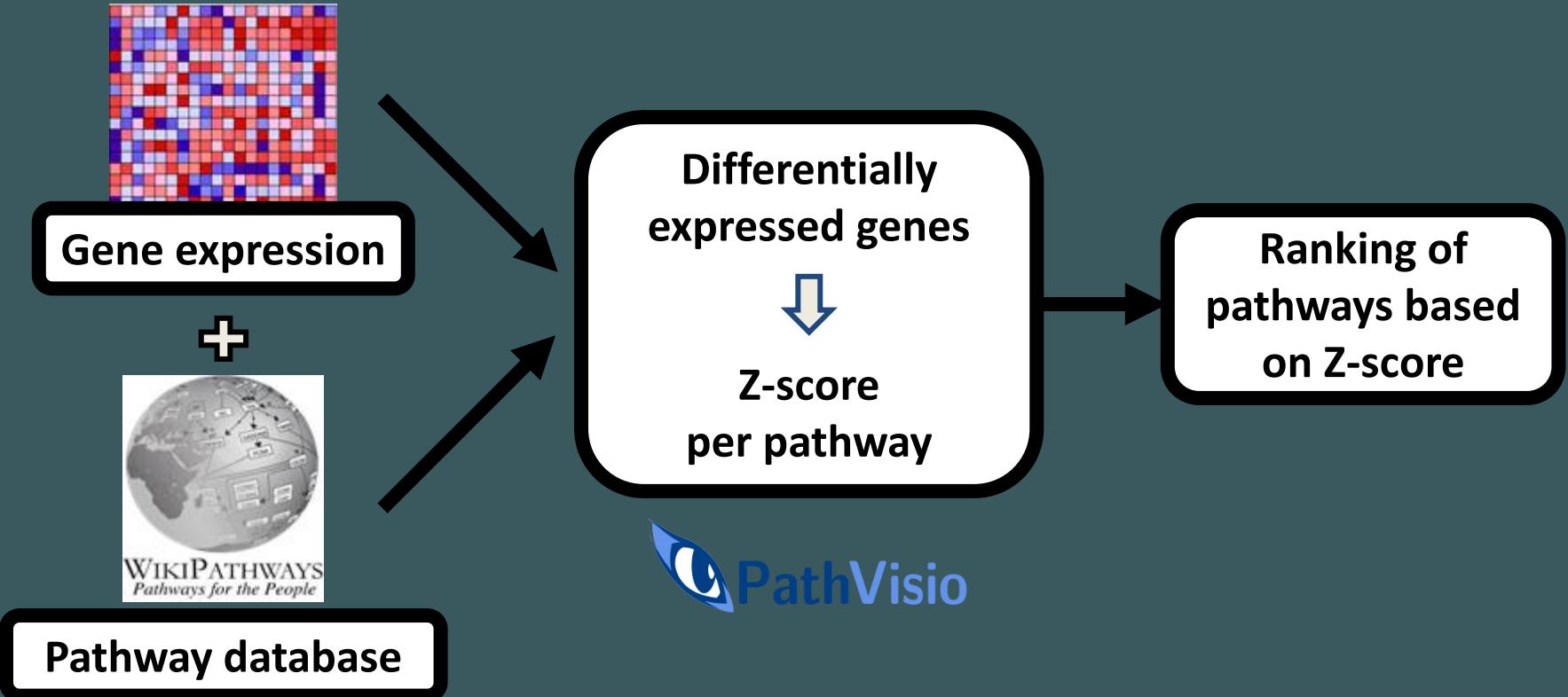


**Time-series data visualization**  
(virus infection over multiple days)  
Into the eye of the cytokine storm.

<https://doi.org/10.1128/MMBR.05015-11>



# Pathway Statistics





# Pathway Analysis Methods

- **Overrepresentation analysis:**
  - Input list → e.g. significantly up- or down-regulated genes
  - Background list → e.g. all measured genes
  - Statistical test → e.g. Fisher's exact test (hypergeometric test)

$$Z\text{-score} = \frac{(r - n \frac{R}{N})}{\sqrt{n \frac{R}{N} (1 - \frac{R}{N}) (1 - \frac{n-1}{N-1})}}$$

- Z-Score is calculated for each pathway
  - Results in ranked list of pathways
- Four variables in the formula: N, R, n, r

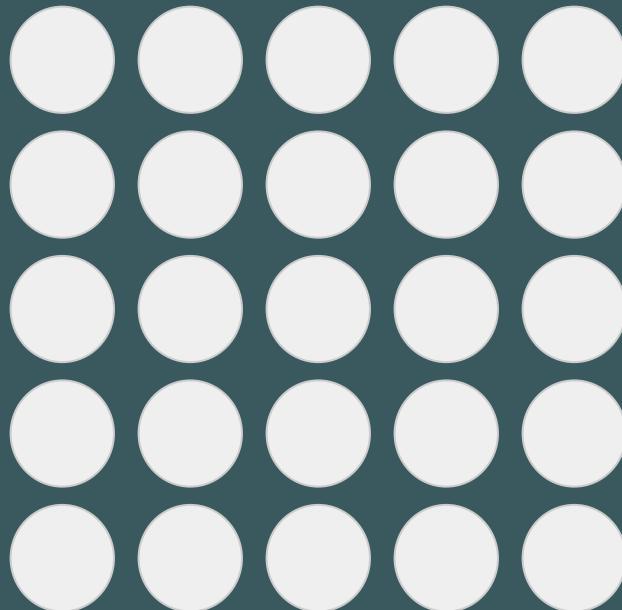


# Z-Score ORA

$$Z\text{-score} = \frac{(r - n \frac{R}{N})}{\sqrt{n \frac{R}{N} (1 - \frac{R}{N}) (1 - \frac{n-1}{N-1})}}$$

**N = 25**

background list (total number of measured genes in experiment)





# Z-Score ORA

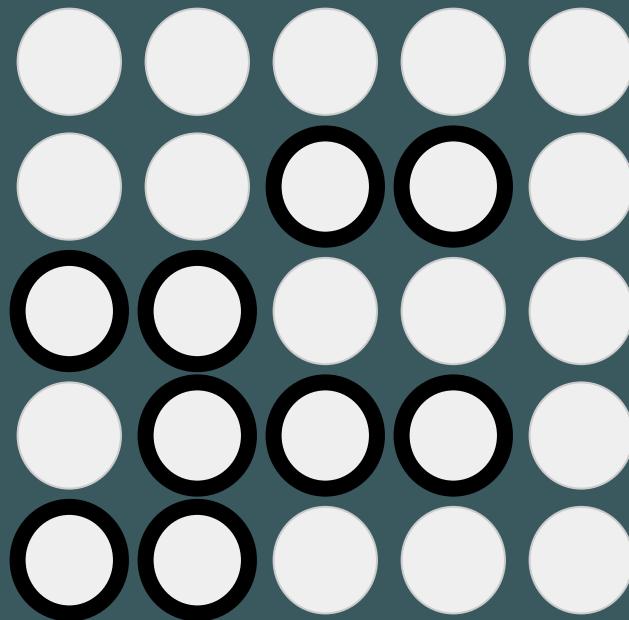
$$Z\text{-score} = \frac{(r - n \frac{R}{N})}{\sqrt{n \frac{R}{N} (1 - \frac{R}{N}) (1 - \frac{n-1}{N-1})}}$$

**N = 25**

background list (total number of measured genes in experiment)

**R = 9**

input list (number of changed genes in experiment)





# Z-Score ORA

$$Z\text{-score} = \frac{(r - n \frac{R}{N})}{\sqrt{n \frac{R}{N} (1 - \frac{R}{N}) (1 - \frac{n-1}{N-1})}}$$

**N = 25**

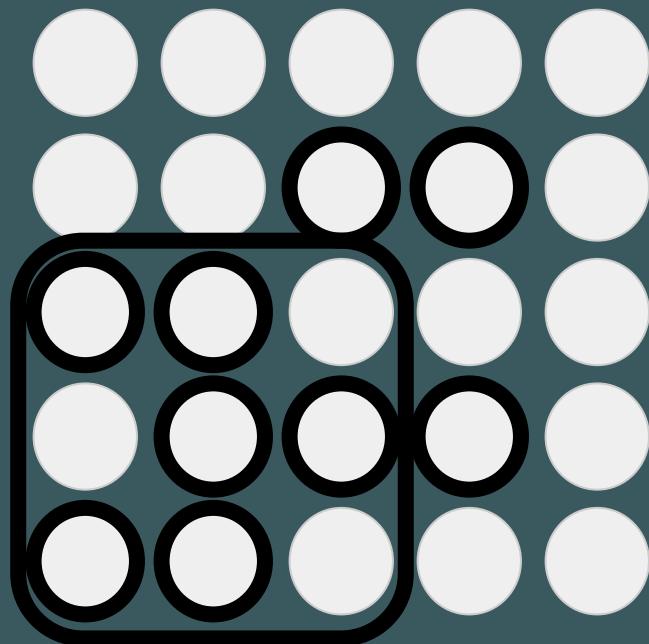
background list (total number of measured genes in experiment)

**R = 9**

input list (number of changed genes in experiment)

**n = 9**

total number of genes in pathway



Pathway X



# Z-Score ORA

$$Z\text{-score} = \frac{(r - n \frac{R}{N})}{\sqrt{n \frac{R}{N} (1 - \frac{R}{N}) (1 - \frac{n-1}{N-1})}}$$

**N = 25**

background list (total number of measured genes in experiment)

**R = 9**

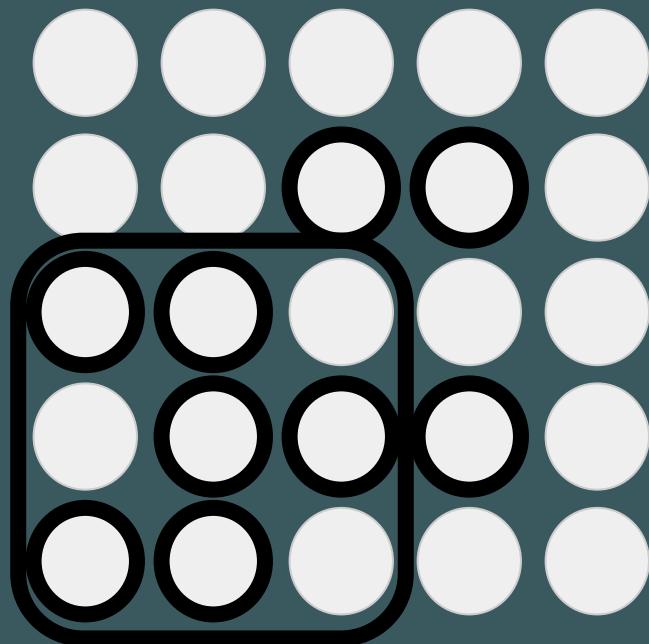
input list (number of changed genes in experiment)

**n = 9**

total number of genes in pathway

**r = 6**

number of changed genes in pathway



Pathway X



# Z-Score ORA

$$Z\text{-score} = \frac{(r - n \frac{R}{N})}{\sqrt{n \frac{R}{N} (1 - \frac{R}{N}) (1 - \frac{n-1}{N-1})}}$$

**N = 25**

background list (total number of measured genes in experiment)

**R = 9**

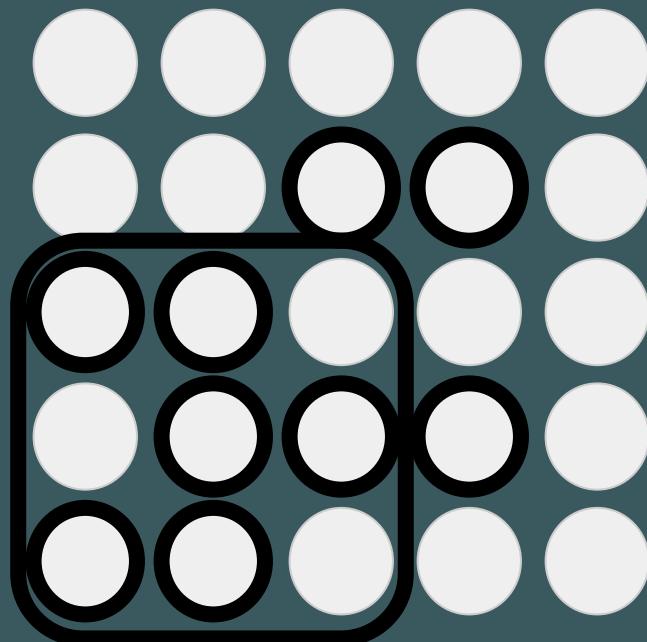
input list (number of changed genes in experiment)

**n = 9**

total number of genes in pathway

**r = 6**

number of changed genes in pathway



Pathway X

**Z-Score for pathway X = 2.347**



# Z-Score ORA

- What does the Z-Score tell you?
  - Z-Score > 1.96
    - Significantly more genes than expected are changed in the pathway → altered pathways in the experiment (different between the groups)
  - Z-Score = 0
    - Distribution of changed genes in the pathway is the same as in the complete dataset
  - Z-Score = < -1.96
    - Significantly less genes than expected are changed in the pathway → very stable pathway (not affected in experiment)



# Pathway Statistics

- Overrepresentation analysis to identify altered pathways in a dataset of interest
- Hypergeometric test (Fisher's exact test)

<b>Up-regulated pathways</b> (log2FC > 2, p-value < 0.05)	<b>Z-score</b>	<b>Perm. p-value</b>
Cell Cycle	6.12	0.001
G1 to S cell cycle control	4.26	0.002
Synaptic Vesicle Pathway	3.89	0.001
DNA Damage Response	3.88	0.002
ATM Signaling Pathway	3.80	0.001

<b>Down-regulated pathways</b> (log2FC < -2, p-value < 0.05)	<b>Z-score</b>	<b>Perm. p-value</b>
Complement and Coagulation Cascades	5.87	0.001
Complement Activation	5.84	0.001
Adipogenesis	5.49	0.001
Differentiation of white and brown adipocyte	5.44	0.001
Triacylglyceride Synthesis	4.53	0.001



# Acknowledgements

**GLADSTONE  
INSTITUTES**



Maastricht University

core development and curation team  
and many contributors and curators  
around the world



# On to the practical!