chem-bla-ics

PlantMetWiki: a linked open data service for querying and analyzing plant pathway knowledge

Egon Willighagen

Published January 5, 2026

Citation
Willighagen, E. (n.d.). In *chem-bla-ics*. chem-bla-ics. https://doi.org/10.59350/wyz5v-vts33

Keywords
Wikipathways, Gpml, Rdf

Copyright
Copyright © Egon Willighagen 2026. Distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Back on October I presented *Everything you always wanted to know: plant pathway modelling in WikiPathways* (doi:10.5281/zenodo.18149988) at the *Knowledge Graphs for Plant and Microbiome Multiomics* symposium (see this archived LinkedIn post) on 14th October 2025 (youtube recording). I had not found time yet to post about this meeting, but it was an awesome list of speakers, regrettable absense of some others, but resulting in new contacts and some slowly evolving collaborations.

Previously, plant pathways were somewhat negatively prioritized at our BiGCaT research group. Something with Dutch academic politics. But that was 10 years ago, and with the notion that human health very much involves the exposome, which includes live around humans, I think the plant pathway science is important to human health. Even just the human health impacts of drops in biodiversity. Or the impact on our nutrition supply chain of climate change.

Anyway, I am happy that Elena and Denise pulled me into a collaboration to create an RDF-based knowledge graph about plant pathways. Their idea was to PlantCyc pathways (their license seems to allow that; doi:10.1093/nar/gkae991), convert that to GPML (by Max) and then to RDF. That last step is where I come in. The details will follow later, but Elena announced the project on LinkedIn (archived link), so time to blog about it myself too.

I am happy with this effort, not just because we now have pathways in RDF form for more than 500 species, but also because it requires continued development of the WikiPathways solutions, like GPML and libGPML and the RDF generation code, but also BridgeDb (doi:10.1186/1471-2105-11-5). The latter provides the identifier mapping infrastructure, but needed to be extended for the new species (something I had to do earlier this year for several caffeine synthesis pathways developed at the DBCLS BioHackathon 2025).

Lars gave me a tip on how to scale this up (after a manual addition), verifier.globalnames.org (doi:10.5281/zenodo.17245658, which greatly helped me out. It translates species names into identifiers, and their JSON is very rich in that process as well as easy to process. So, a custom script allowed me to update BridgeDb more efficiently. Highly recommended!

So, the resulting knowledge base is available at plantmetwiki.bioinformatics.nl and looks like this (also big thanks to Marvin for support in setting this up!):

chem-bla-ics

PlantMetWiki
Plant Metabolic Pathways Wiki
SPARQL Explorer
Tutorial
SPARQL Query:
1
Query
Reset
Export CSV
Export JSON
Export XML
Get Permalink
Fullscreen Mode
Show Prefixes
SPARQL Endpoint:
https://plantmetwiki.bioinformatics.nl/sparql
SPARQL Examples:
https://github.com/pathway-lod/SPARQLQue
Type part of the query file name to search for...
Search
Clear
1.(Meta)Data
2.PlantMetWiki
3.Biosynthetic gene cluster (BGC) Crosslinks
4.Federated
Powered by WikiPathways (doi:10.1093/nar/gkad960) and Snorql - Extended Edition — Cookie Policy — PlantMetWiki SPARQL Explorer is licensed under — GitHub Organisation — PlantMetWiki includes modified data derived from the PlantCyc / PMN databases. Original data © Carnegie Institution for Science. Licensed under the PMN Open Database License.