

INOH pathway database

Ken Ichiro FUKUDA¹, Takao ASANUMA², Satoko YAMAMOTO², Tatsuya KUSHIDA², Emi HATTORI³, Yuki YAMAGATA², and Toshihisa TAKAGI⁴
fukuda-cbrc@aist.go.jp, asanuma-takao@aist.go.jp, yamamoto-satoko@aist.go.jp,
kushida-tatsuya@aist.go.jp, hattori-emi@aist.go.jp, yamagata-yuki@aist.go.jp, tt@k.u-tokyo.ac.jp

¹Computational Biology Research Center, National Institute of Advanced Industrial Science and Technology, Tokyo JAPAN

²Institute for Bioinformatics Research and Development, Japan Science and Technology Agency,

³Information and Mathematical Science Laboratory, Inc., Tokyo, JAPAN,

⁴Graduate School of Frontier Sciences, University of Tokyo.

INOH is a human curated signal transduction pathway database of model organisms including human, mouse, rat and others.

A pathway has a hierarchical and recursive structure, which means a pathway consists from a set of pathways (sub-pathways) and the pathway itself can be a sub-pathway of other pathways. A sub-pathway is defined by a process or reaction and its input/output/controller material, e.g., a protein and its phosphorylated form and an enzyme. In INOH, pathways are called “events” and each event consists of other events (subpathways) or a process, input/output materials and controllers.

The aim of INOH project is to provide pathway knowledge that resides in the scientific literature in a machine-accessible form. To that end, not only the topological structure of pathways but also the background knowledge that biologists use to interpret pathway descriptions in the literature, has to be made explicit and available to computers.

Hence every pathway component in INOH (events, proteins, chemicals, etc.) is annotated by a biological ontology.

The Event Ontology provides a DAG structured controlled vocabulary for events. It has four top categories for biological events: (1) “Molecular event” includes molecular level phenomena such as “molecular interactions,” (2) “Cellular event” includes phenomena such as “cell growth,” (3) “Organism event” includes phenomena such as “immune response,” and (4) “Physiological event” includes phenomena such as “homeostasis”. And the background knowledge that an event is involved in a specific biological phenomenon is encoded as a “part-of” relationship: e.g., “B cell receptor signaling (through IKK-NF-kappaB cascade)” is part-of “Cell proliferation.” Background knowledge about positive and negative regulations of sub-pathways/pathways is also encoded in this ontology. For example, “Sumoylation of STAT1/3/4 by PIASx/y” is-a “Negative regulation of transcription in the JAK STAT pathway by (STAT sumoylation by PIAS)”.

The MoleculeRole ontology[3] is a bio-ontology of protein names and protein family names and captures the relations among (1) molecule function names (e.g. protein serine/threonine kinase) (2) abstract molecule names (e.g. R-smad), (3) concrete molecule names (e.g. Smad2 of human), (4) sequence accession numbers and (5) complexes and their subunits.

These ontological knowledge can be utilized during a pathway query. For example, if one finds a protein entity annotated as “MAP-kinase”, she knows from the MoleculeRole Ontology that it may indicate ERK1 of a human, JNK1 of a mouse, p38alpha of a rat, etc. If she finds a protein entity annotated as “Smurf”, then the ontology tells her that it is a ubiquitin ligase.

To represent pathways in multiple abstraction levels, the INOH pathway model is based on a compound-graph structure[1]. The “INOH Curation Tool” is a JAVA based platform independent graphical pathway navigation tool to edit and search pathways or ontologies in the INOH database. This tool provides automatic layouting of compound graph pathways. INOH ontology can be accessed through a web application, called Ontology Viewer. Ontology Viewer allows the user to search the ontology by names, synonyms, and ontology IDs. The ontologies and the system are freely available at <http://www.inoh.org>.

[1] Fukuda K, Takagi T. 2001. Knowledge representation of signal transduction pathways. *Bioinformatics* 17(9): 829-837.

[2] Fukuda K, Yamagata Y, Takagi T. 2003. FREX: a query interface for biological processes with a hierarchical and recursive structures. *In Silico Biol* 4: 0007.

[3] Yamamoto, S., Asanuma, T., Takagi, T., Fukuda, K. 2004. The Molecule Role Ontology: an ontology for annotation of signal transduction pathway molecules in the scientific literature, *Comparative and Functional Genomics* 5, 528-536.