

## **The Apoptosis Gene Ontology project and the IntAct-Apoptosis dataset: updating the conceptual representation of apoptosis**

Pablo Porras Millán, Paola Roncaglia, Emily Dimmer, Jane Lomax, Rachael Huntley, Sandra Orchard, Marine Dumousseau, Samuel Kerrien, Henning Hermjakob and the GO and IntAct teams at the EBI.

EBI-EMBL, Wellcome Trust Genome Campus, CB10 1SD, Hinxton, UK

`pst@ebi.ac.uk`

The comprehensive and detailed analysis and depiction of biological pathways and reactions requires the use of resources and databanks in which previous knowledge is stored in an organized and standardized way and made available for the scientific community. Resources such as protein and gene databanks, molecular interactions databases or annotation ontologies used to characterize genes and gene products help to provide researchers with a reference from which they can infer current scientific knowledge, expanding and annotating their own datasets.

As participants in the APO-SYS Consortium ([www.apo-sys.eu/](http://www.apo-sys.eu/)), a project funded by the European Community to focus on “Apoptosis Systems Biology Applied to Cancer and AIDS”, the molecular interactions database IntAct ([www.ebi.ac.uk/intact](http://www.ebi.ac.uk/intact)) and the Gene Ontology (GO, [www.geneontology.org](http://www.geneontology.org)) have taken a significant effort to improve annotation and representation of apoptotic cell death in their respective domains. Both resources are hosted at the European Bioinformatics Institute (EBI-EMBL). IntAct is currently one of the biggest primary repositories of molecular interactions, hosting more than 300.000 binary interaction evidences represented with a high level of detail and extracted from the manual curation of over 5.700 publications. The GO project is an international effort with the aim to consistently and systematically annotate gene products by using a controlled vocabulary. Annotations performed with this ontology help representing and interpreting biological data both in detailed, small-scale setups and in the comprehensive, ‘omics’-sized studies that have revolutionized our understanding of biology in the last 10-15 years. GO term annotation and GO enrichment analysis have become fundamental tools in the study of protein-protein interaction networks.

The research groups involved in the APO-SYS Consortium highlighted the need for a thorough revision of apoptosis-related terms in GO in order to properly depict current knowledge in the field. In a collaborative effort, members of the IntAct and the GO teams at the EBI have undertaken a guided curation effort of literature recommended by the consortium experts, creating more granular terms that can accurately represent the scientific consensus in apoptosis. Old, inaccurate terms have been updated or obsoleted and the relationships between both updated and newly created terms have been thoroughly revised. A major re-annotation effort with the goal to increase the granularity and extension of current annotations has already

started, and it will continue until the major players in the apoptotic process are properly annotated.

In addition, previously existing entries in IntAct involving gene products with a role in apoptosis have been identified and integrated into an “IntAct-Apoptosis” group in the database, providing a dataset with almost 30.000 interactions involved in the apoptotic process and in other types of cell death in different species, mostly human and mouse, but also other organisms such as *C. elegans*, *D. melanogaster* or even *S. cerevisiae*.

In our communication, we give an overview of the updated apoptosis ontology and we show the differences in representation and analysis as it is applied to the IntAct-Apoptosis dataset. We will also describe the workflow of the project, in which the integration of the input of the apoptosis experts, along with the effort of IntAct curators and developers and GO curators and editors, has been key to the success of this endeavor.

The IntAct apoptosis dataset is available at [http://www.ebi.ac.uk/intact/query/annot:"dataset:apoptosis"](http://www.ebi.ac.uk/intact/query/annot:).

Acknowledgements:

**Supported by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° HEALTH-F4-2007-200767 for APO-SYS.**