MSDA tutorial: How to read output Excel files obtained from GO ontologies extractions.

In the output Excel file you get after GO ontologies extraction, multiple columns will have been added to the initial input Excel file. These columns include general information about genes/proteins, and specific ontology annotations.

Global information

Identifier used	The identifier used to extract GO ontologies is given here. If more than one accession number per cell was present in the input Excel file, the first one will have been selected. If it returned no result, the next one will then have been selected, etc.
NCBI protein accession number	Corresponding NCBInr accession number
GeneID	Corresponding GeneID number
Symbol	Name of the gene coding for the protein
UniProtKB accession number	Corresponding UniProtKB accession number
General description	Description of the protein in the UniProtKB database
Recommended Name	Recommended name of the protein in the UniProtKB database
Alternative Names	Alternative names of the protein in the UniProtKB database
Taxonomy	Taxonomy ID of the studied species for which GO terms were extracted. For more details on TaxID, please refer to the NCBI website (http://www.ncbi.nlm.nih.gov/taxonomy)
Reference database	Name of the reference database for the studied organism
Reference accession number	Corresponding accession numbers in the reference database

Ontology annotations

For each selected ontology category, three separate columns are added in the output Excel file:

Ontology	GO terms **directly** assigned to the used identifier in the GO database. The evidence code for each GO term is given between brackets to indicate the extent a given annotation to a particular term is supported. (see details on evidence codes at http://www.geneontology.org/GO.evidence.shtml) **Warning: there is more than one line in most Excel cells.**
Ancestors	List of all GO terms in the GO database that are related to the GO terms directly assigned to a given entry, but only considering ancestries (parental lineages). (see details at http://www.geneontology.org/GO.ontology.structure.shtml)
Graphic tree view	A graphic representation of extracted GO terms, including ancestries, is provided through a reconstructed AmiGO URL link. (see details at http://www.geneontology.org/GO.ontology.relations.shtml) **Warning: The links are not "clickable" due to their extreme length. You will have to copy and paste them in your Web browser.**