

What is Gene Ontology?

- The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism (<http://www.geneontology.org>)

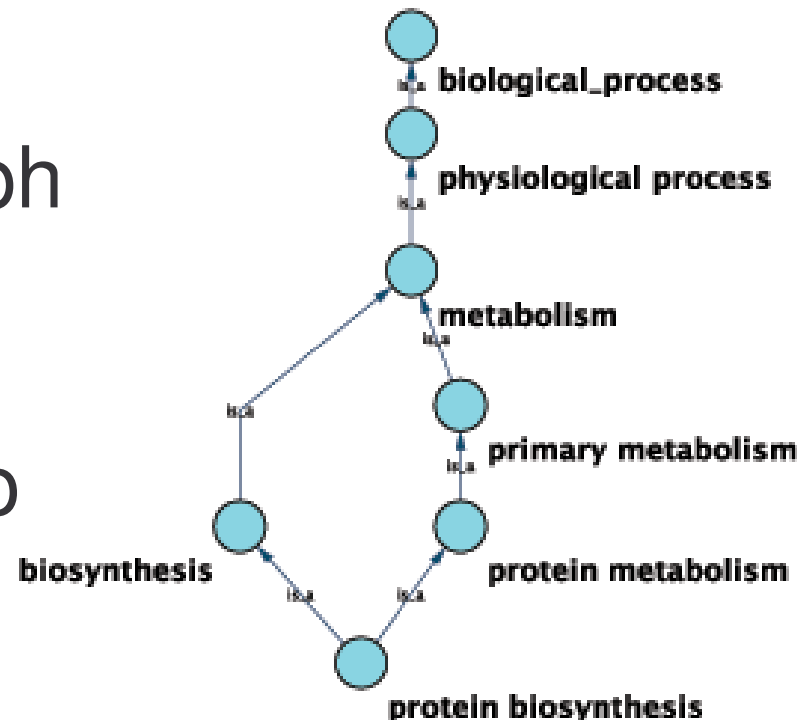
What are the ontologies

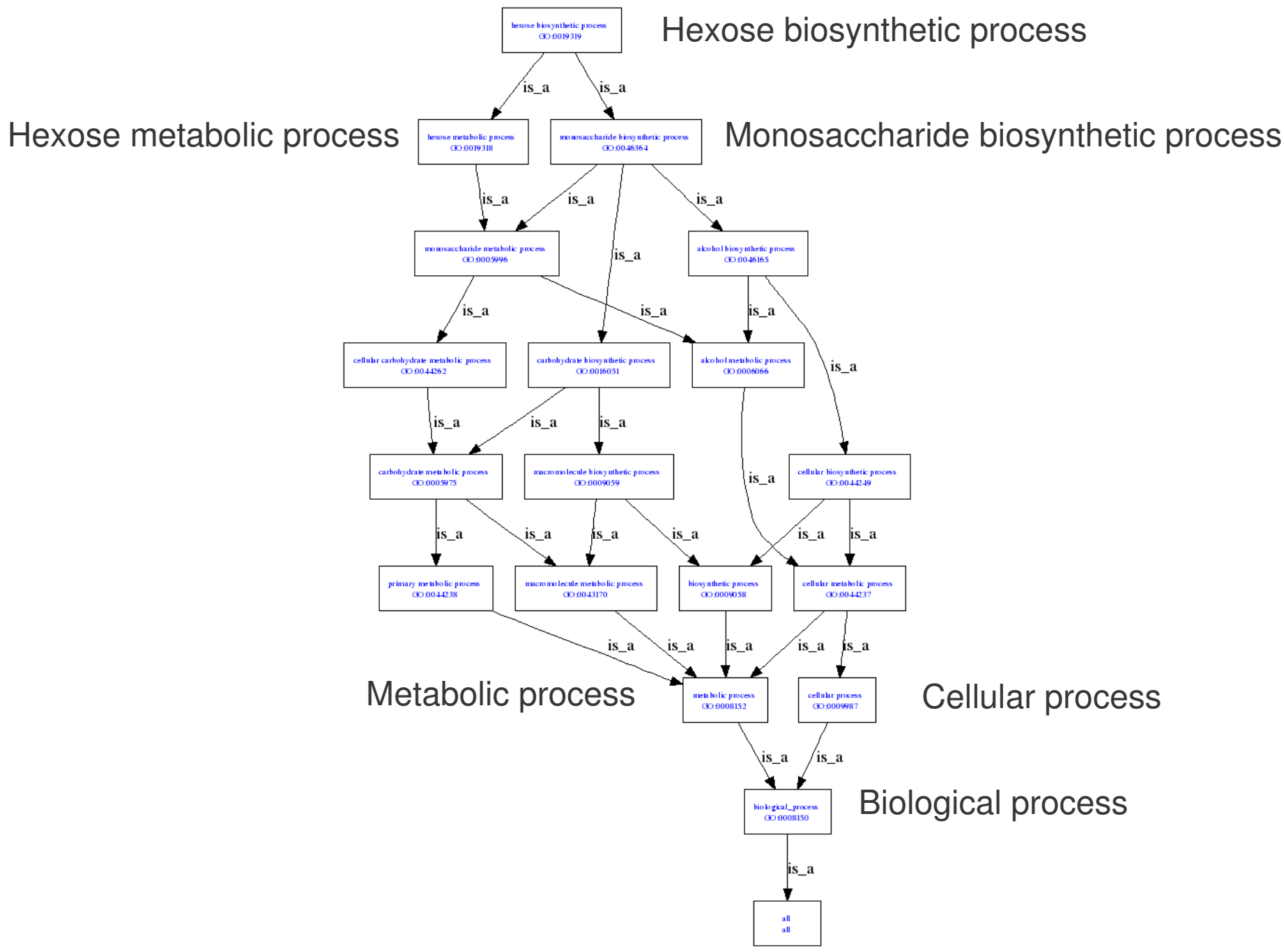
- Organised as three separate ontologies
 - Molecular functions
 - Biological processes
 - Cellular components
- Each gene or gene product may
 - have more than one molecular function
 - take part in more than one biological process
 - act in more than one cellular component

Structure of ontologies

- Shows the relation between different terms
 - One term may be a more specific description of another more general term

- Directed Acyclic Graph (DAG)
- Similar to hierarchies
- Allows a child node to have more than one parent





Hexose biosynthetic process

Hexose metabolic process

Monosaccharide biosynthetic process

Metabolic process

Cellular process

Biological process

Browsing ontologies

The image shows two overlapping software windows from the SpotPix Data application. The top window, titled "GO DAG SpotPix Data", displays a hierarchical ontology tree. The selected term is "mitotic cell cycle" (GO:0000278), which is highlighted in blue. Other terms in the tree include "chromosome segregation", "cell cycle", "meiotic cell cycle", "endomitotic cell cycle", "abortive mitotic cell cycle", "regulation of ubiquitin ligase activity during mitotic cell cycle", and "interphase of mitotic cell cycle".

The bottom window, titled "Gene Graph - SpotPix Data", shows a line chart of gene expression data. The chart has a y-axis ranging from -5.0 to 5.0 and an x-axis with labels for various patients (e.g., Patient 1 ff, Patient 2 ff, ..., Patient 10 etter). A table of gene data is visible on the left side of this window.

Gene Graph Data Table:

Ind...	Na...	ID	Re...	Us...	Gr...
41...	ce...	16...	1	1...	█
41...	U...	46...	1	1...	█
41...	F...	50...	1	1...	█
42...	pr...	42...	2	2...	█
42...	W...	37...	1	1...	█
42...	ad...	24...	1	1...	█
42...	co...	70...	1	1...	█
42...	Tr...	70...	1	1...	█
42...	de...	42...	1	1...	█
42...	ce...	51...	1	1...	█
42...	nu...	24...	1	1...	█
42...	pr...	50...	1	1...	█
42...	Tr...	50...	1	1...	█
42...	cy...	84...	1	1...	█
42...	ln...	42...	1	1...	█
42...	22...	22...	1	1...	█
42...	hy...	37...	1	1...	█
42...	tr...	82...	1	1...	█
42...	U...	H...	5	5...	█
42...	ne...	13...	1	1...	█
42...	ch...	51...	1	1...	█
42...	R...	28...	1	1...	█

Gene Ontology Analysis

- We have a list of differentially expressed genes from e.g. SAM
- Which biological processes do these genes take part in?
- Is there an over-representation of the number of genes belonging to a particular biological process, compared to what could be expected?

Question

- If we look at the dataset containing all of our genes and see that 10% of these belong to cell cycle. We then do a differentially expressed genes analysis and get a list of genes we believe are significantly changed. How many of the genes in the gene list do you expect belong to cell cycle?

Gene Ontology Analysis

Reference list

Statistical comparison between the two GO components

The image displays a software interface for Gene Ontology (GO) analysis. It includes several windows:

- Project:** Shows a tree view with 'Breast cancer' and 'SAM' selected.
- GO DAG SAM:** A hierarchical tree of GO terms for the SAM component, with 'mitotic cell cycle' highlighted.
- GO DAG Breast cancer:** A hierarchical tree of GO terms for the Breast cancer component, with 'mitotic cell cycle' highlighted.
- GO Comparison result:** A table comparing the two components.

GO term	Exact test p value	Genes in selecti...	Genes in reference	Enrichment
mitotic cell cycle	4.782E-13	15	135	14.305
mitosis	1.182E-10	11	82	17.27
M phase of mitotic cell c...	1.677E-10	11	85	16.661
cytokinesis	1.88E-10	11	86	16.467
cell division	1.88E-10	11	86	16.467
cell cycle	2.375E-9	18	397	5.837
			29	26.636
			69	13.061
			3	128.743
			5	77.246
			23	22.39
			26	19.807
			10	38.623
			13	29.71
			16	24.139
			3	85.828
			21	18.392
			21	18.392
			56	9.196
DNA replication initiation	0.001	2	5	51.497
chromosome	0.001	5	101	6.373
cytoskeleton	0.002	8	276	3.732
kinesin complex	0.002	2	7	36.784
biological_process	0.003	50	4601	1.399
nucleus	0.006	20	1419	1.815
blood coagulation	0.008	3	50	7.725
mitotic chromosome co...	0.008	2	16	16.093
cell proliferation	0.009	5	159	4.049
regulation of mitosis	0.009	2	17	15.146
wound healing	0.011	3	56	6.887
- Gene Graph - SAM Column rearranged:** A line graph showing gene expression levels across different patients. The y-axis ranges from -5.0 to 5.0. The x-axis lists patients from Patient 1 to Patient 12 (with 'Patient 12 error').