Computational Biology LU 2014

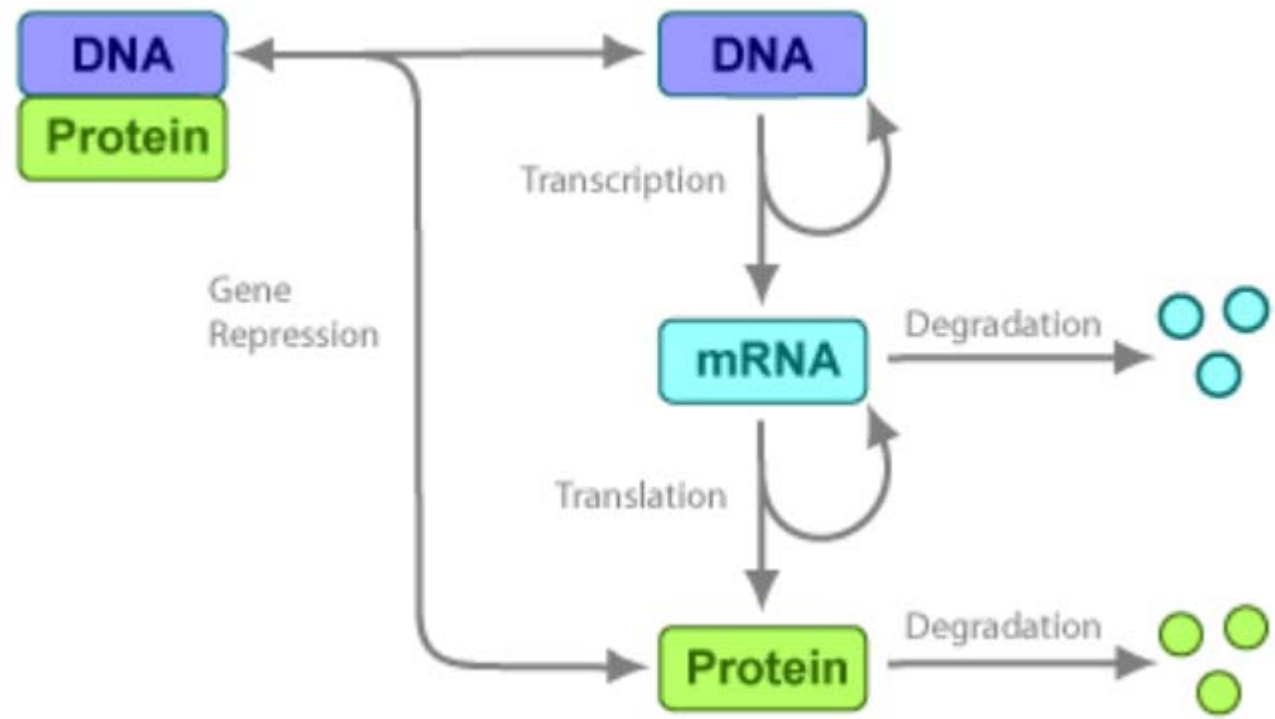
miRNA Analysis

Bioinformatics Group
Institute for Knowledge Discovery
Graz University of Technology
http://genome.tugraz.at
Petersgasse 14, A-8010 Graz

Outline

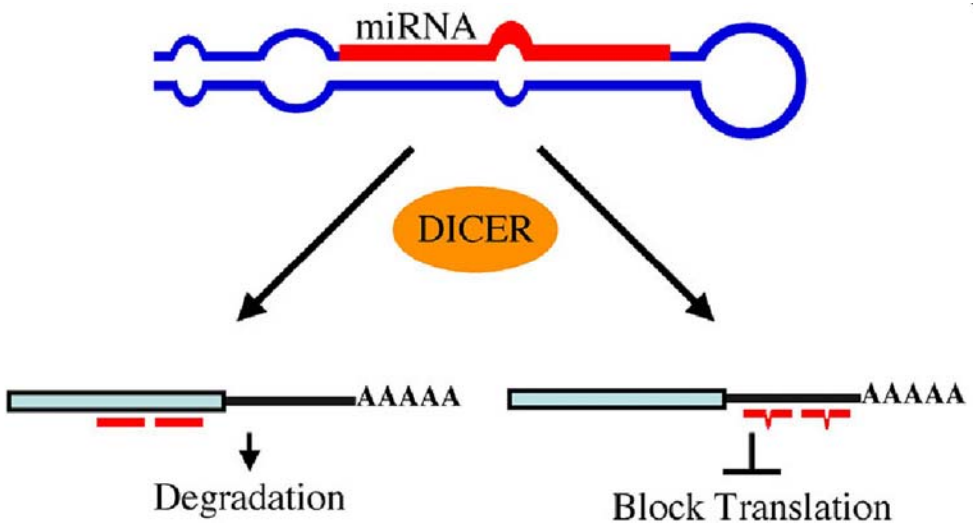
- miRNAs
  - Introduction
  - Function
- Gene Ontology Analysis
- Exercises

Introduction – Gene Regulation



miRNAs are post-transcriptional regulators

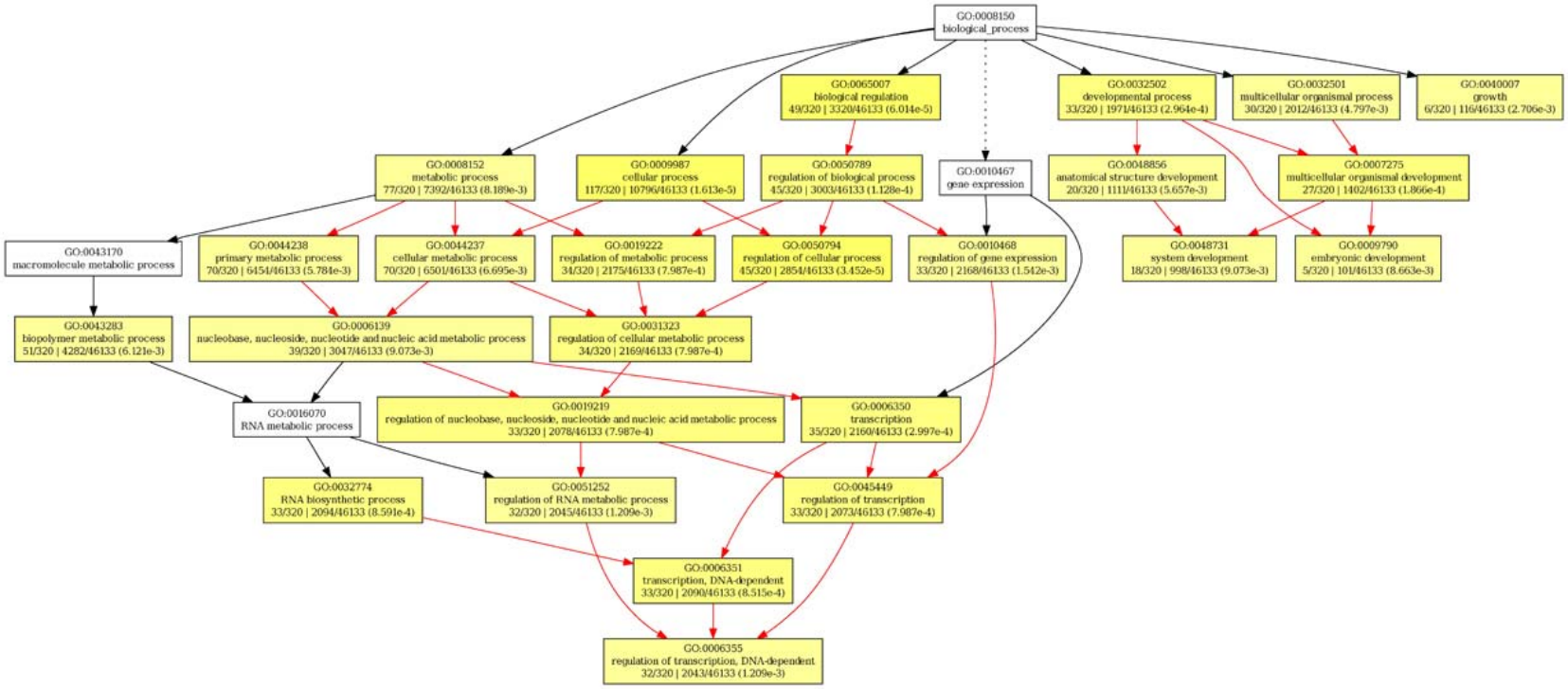
- Were discovered in the 1990s
- Acceptance as distinct biological class in 2000
- Are on average 22 nucleotides long
- Bind on mRNAs
- Indirect gene regulation through mRNA degradation or mRNA translation blocking



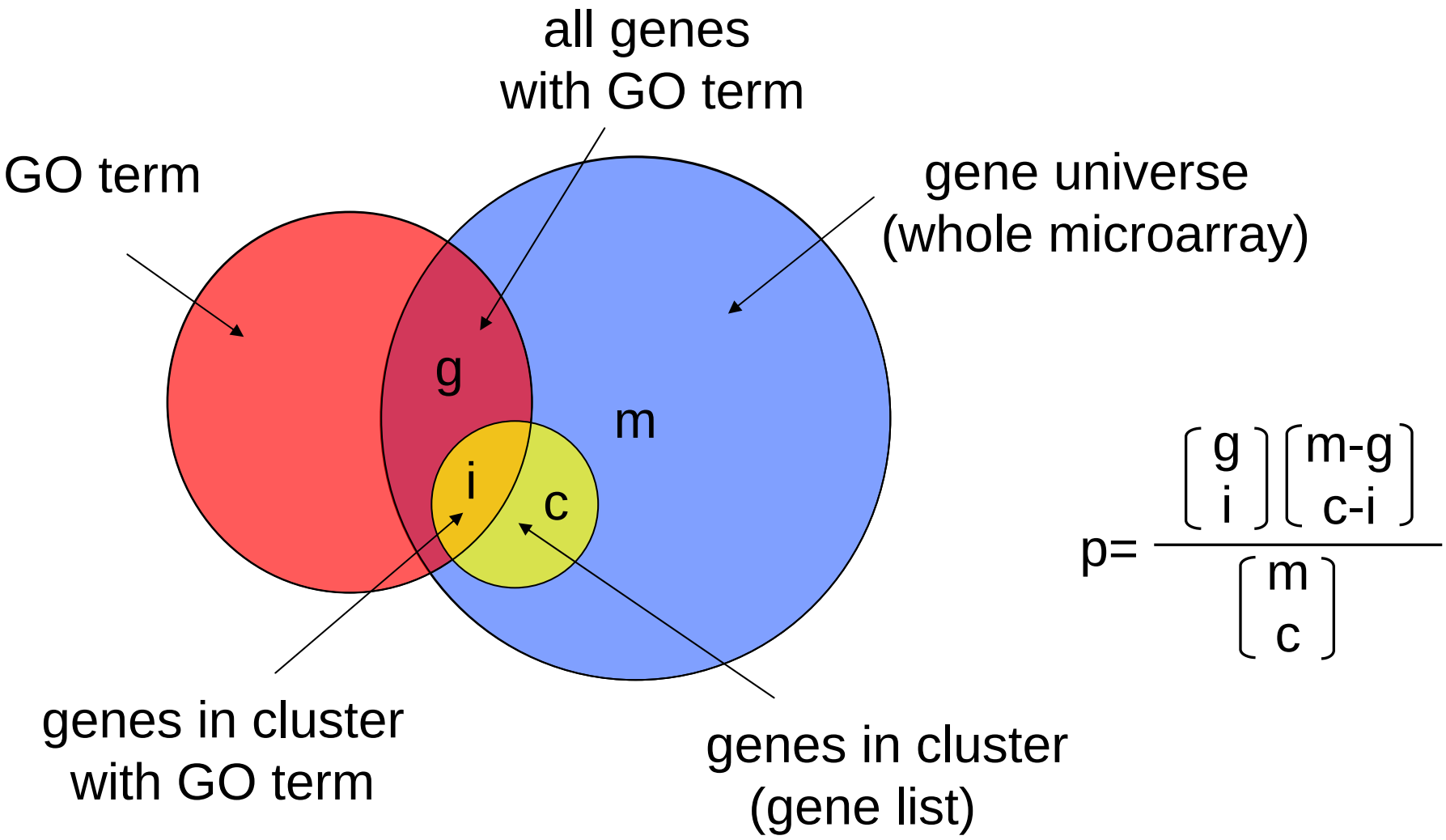
Gene Ontology

- http://www.geneontology.org/
- The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism.
- The three organizing domains of GO are:
  - biological processes (lipid metabolism)
  - molecular function (hydrolase activity)
  - cellular compartment (mitochondria).
- Each entry has a unique numerical identifier of the form GO:nnnnnnn (GO:0022008) and a term name (neurogenesis)

Directed acyclic graph (DAG)



Overrepresentation analysis



$$p = \frac{\binom{g}{i}\binom{m-g}{c-i}}{\binom{m}{c}}$$

Exercise

- We start with a list of miRNAs.
- Obtain a list of validated target genes of the miRNAs
- Perform Gene Ontology analysis on the validated target genes.

Exercises – R hints

- Install packages from bioconductor:
  GOstats, hgu133a.db
- Install CRAN R packages:
  igraph, RCurl, XML.
  Use: install.packages(c(“”,…))
- Do not forget to load the packages before working with them: library(“…”)

Exercises – R hints

- unique(), duplicated()
- hyperGTest() – GO analysis
- getURLContent() – read data from a URL
- htmlParse() – parse text from URL
- readHTMLTable() – read the parsed html text