

Molekularbiologische Datenbanken

Übungen

Aufgabe 5

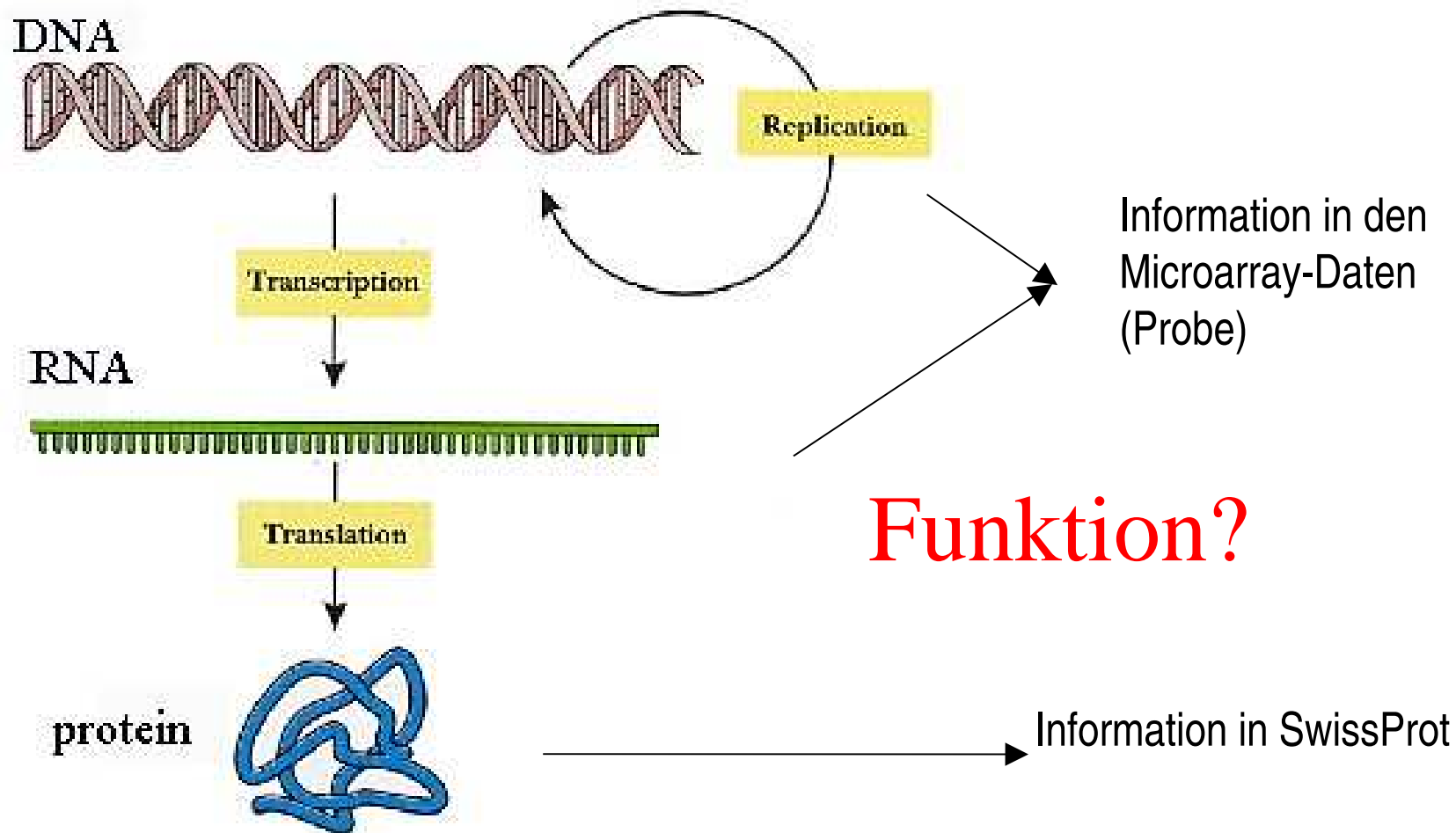
Silke Trißl
Ulf Leser



Wissensmanagement in der
Bioinformatik



Aufgabe 5 – Integrieren von GeneOntology



GeneOntology - Annotationen

- „The goal of the Gene Ontology™ Consortium is to produce a **dynamic controlled vocabulary** that can be applied to all organisms even as knowledge of gene and protein roles in cells is accumulating and changing.“

Annotation in EMBL-Einträgen

Funktion der Gene / Protein beschrieben durch die Description, Keywords, Features

```
DE   Chlamydomonas reinhardtii retinal binding protein (cop4) gene, complete cds.  
...  
KW   .  
...  
FT           /product="retinal binding protein"  
FT           ...  
FT           /note="chlamyopsin-4; microbial type rhodopsin;  
FT           hypothetical light gated ion channel; channelopsin 2"  
FT           /gene="cop4"  
FT           /product="retinal binding protein"
```

Annotation über GeneOntology

- GO-Annotation von cop4

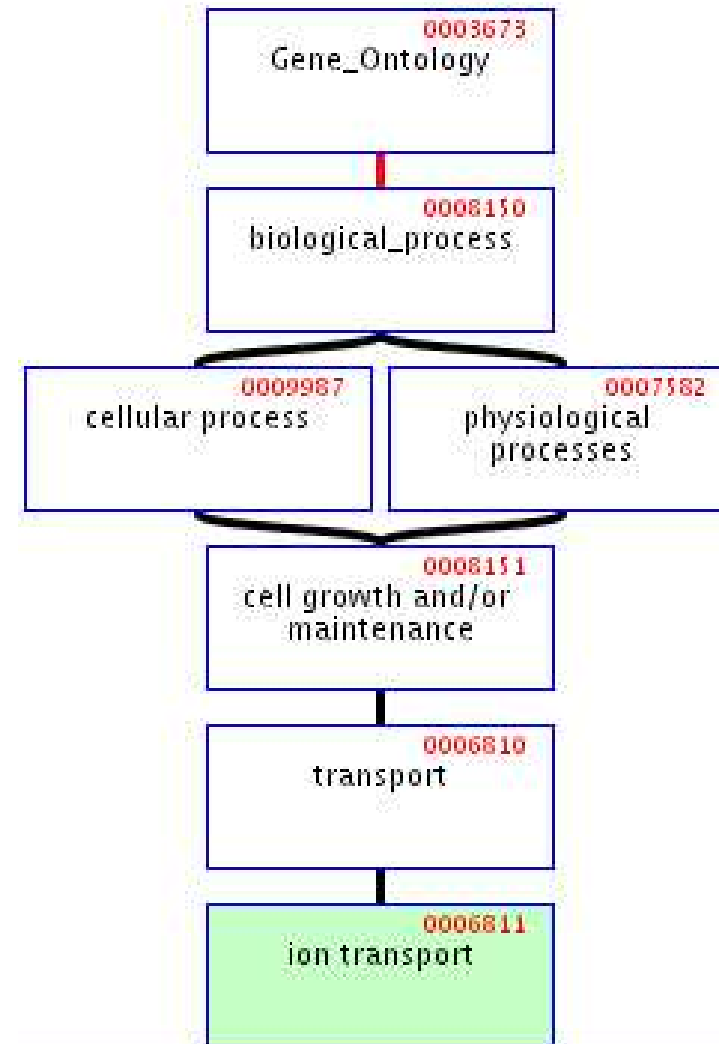
<u>Select</u>	Name	GO ID
process (2)		
<input type="checkbox"/>	ion transport	<u>GO:0006811</u>
<input type="checkbox"/>	protein folding	<u>GO:0006457</u>
function (1)		
<input type="checkbox"/>	ion channel activity	<u>GO:0005216</u>
component (1)		
<input type="checkbox"/>	membrane	<u>GO:0016020</u>

Aufteilung von GO

- molecular function
 - Aufgabe, die von einzelnen Genprodukten ausgeführt wird
- biological process
 - breiteres biologisches Ziel das durch eine Aneinander-reihung von molecular functions erreicht wird
- cellular component
 - subzelluläre Strukturen, Komponenten, und makromolekulare Komplexe

Struktur innerhalb von GO

- Multi-hierarchisch
 - Ein parent-node kann mehrere children haben
 - Ein child-node kann mehrere partents haben

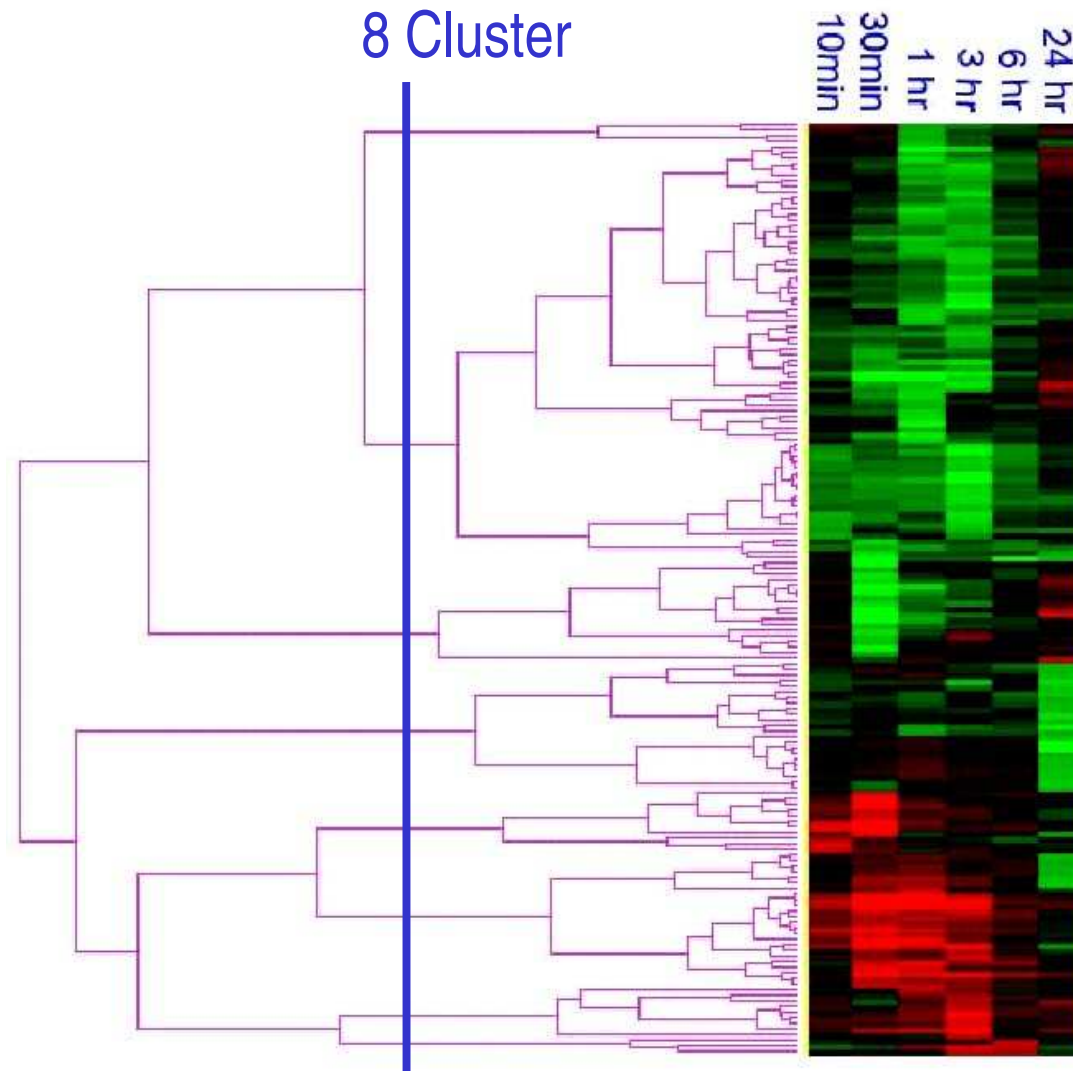


GO und Microarray

- Clustern von Microarray-Experimenten

Frage:

- Was haben Cluster gemeinsam?



GO und Microarray

- GO Term Finder¹:
„The GO Term Finder searches for significant shared GO terms, or parents of the GO terms, used to annotate gene products in a given list.“
- Automatische Analyse von Clustern nach Gemeinsamkeiten mit GO terms möglich

¹ <http://genome-www4.stanford.edu/cgi-bin/SGD/GO/goTermFinder>

neu: GO und Microarray

- Preprocessing von Microarray-Daten mit GO:
- Bildung von Gruppen mit gleicher/ähnlicher GeneOntology Accession
- Clustering innerhalb dieser Gruppen nach Average Difference Intensity

Aufgabe 5 - Daten

- Daten über einen GO-Eintrag
 - term.txt
 - internal_id, GO-name, term_type, acc, is_obsolete, is_root
- Beziehung zwischen einem parent und einem child
 - parent_child.txt
 - relationship_id, relationship_type_id, parent_id, child_id
- Die Werte sind durch tabs getrennt

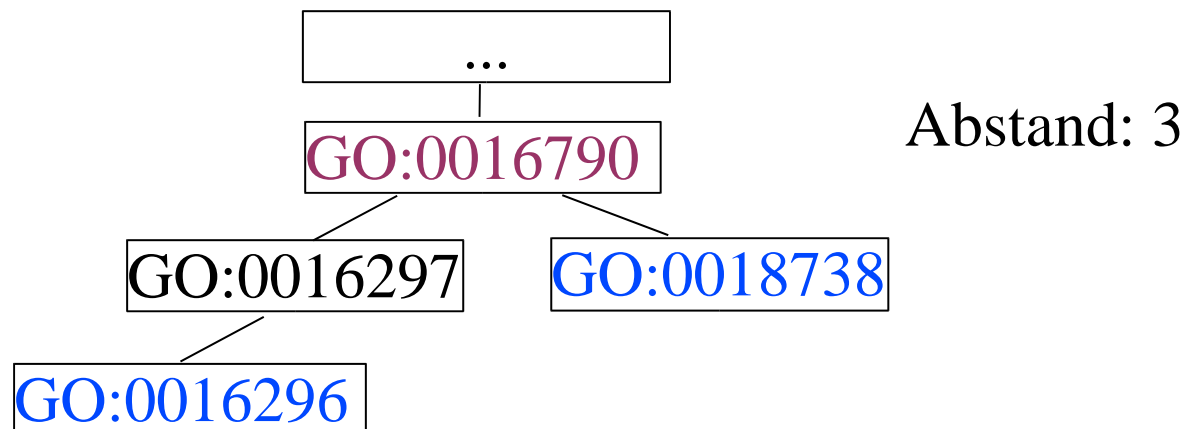
Aufgaben:

- Finden des kleinsten gemeinsamen Knotens

- Beispiel:

- GO:0030371 -> GO:0045182 -> GO:0003674 -> GO:0003673
- GO:0030533 -> GO:0003674 -> GO:0003673
- GO:0003674

Finden des kleinsten gemeinsamen Knotens mit Abstand



Fragen?

- Aufgabe 5 über Goya oder auf Web-page
- Folien auf der Web-page
- Lösung bis 16.07. , 17 Uhr per e-mail oder in RUD25 IV.102
- Mi, 02.07. entfällt (bin die Woche nicht da)

