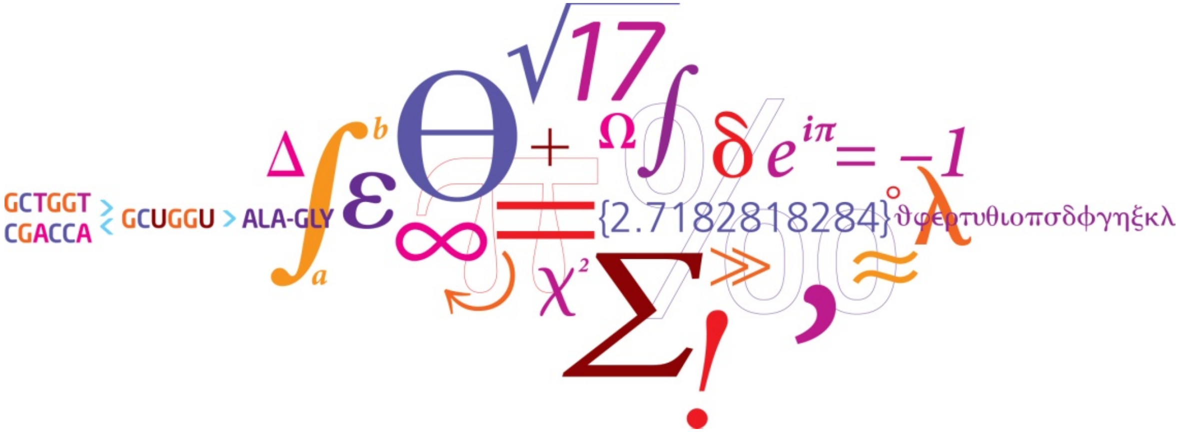
Functional annotation of immunological targets



DTU Biosys
Department of Systems Biology

Selecting the right target

What is a good target
for a vaccine?

What is a good target
for a immunotherapy?

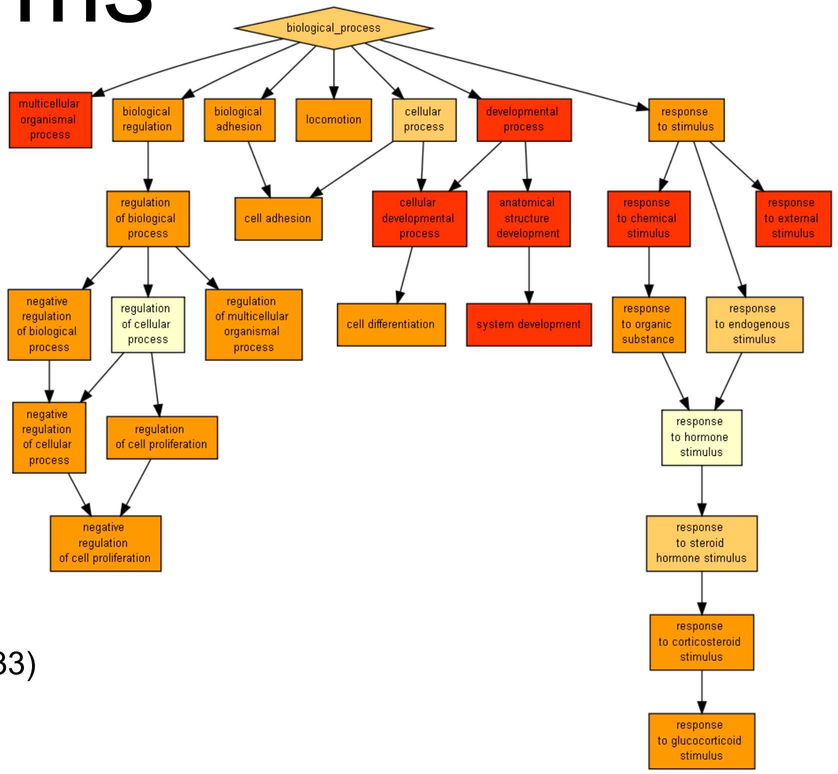
Important features

	Humoral	**Cellular**
Expression level	++	++
Conservation	++	++
Extracellular	++	-
PTMs	++	?
Disorder	++	-
Function	+	+
Redundancy	+	+

Useful tools

Expression level	RNAseq
Conservation	Strains, SNPs, Mutations
Extracellular	TMHMM, SignalP, GO cellular localization
PTMs	NetGlycN, NetPhos
Disorder	DisoPred
Function	Blast, GO, Blast2GO, COG, UniFam, FFPred
Redundancy	Blast, InParanoid, OrthoFinder

GO terms



- Gene Ontology
- 3 main ontologies:
  Process, Function, Localization
- Tree-structure

positive regulation of response to biotic stimulus (GO:0002833)

is a

positive regulation of filamentous growth of a population of unicellular organisms in response to biotic stimulus

Sequence similarity

- Blast (even better, blastx) over an annotated db
- transfer the function if the similarity is good enough

Sequence similarity

- Blast (even better, blastx) over an annotated db

- transfer the function if the similarity is good enough

- Homolog/Paralog problem!

Mutual blast
Close paralogs have similar function

Sequence similarity

- Blast (even better, blastx) over an annotated db

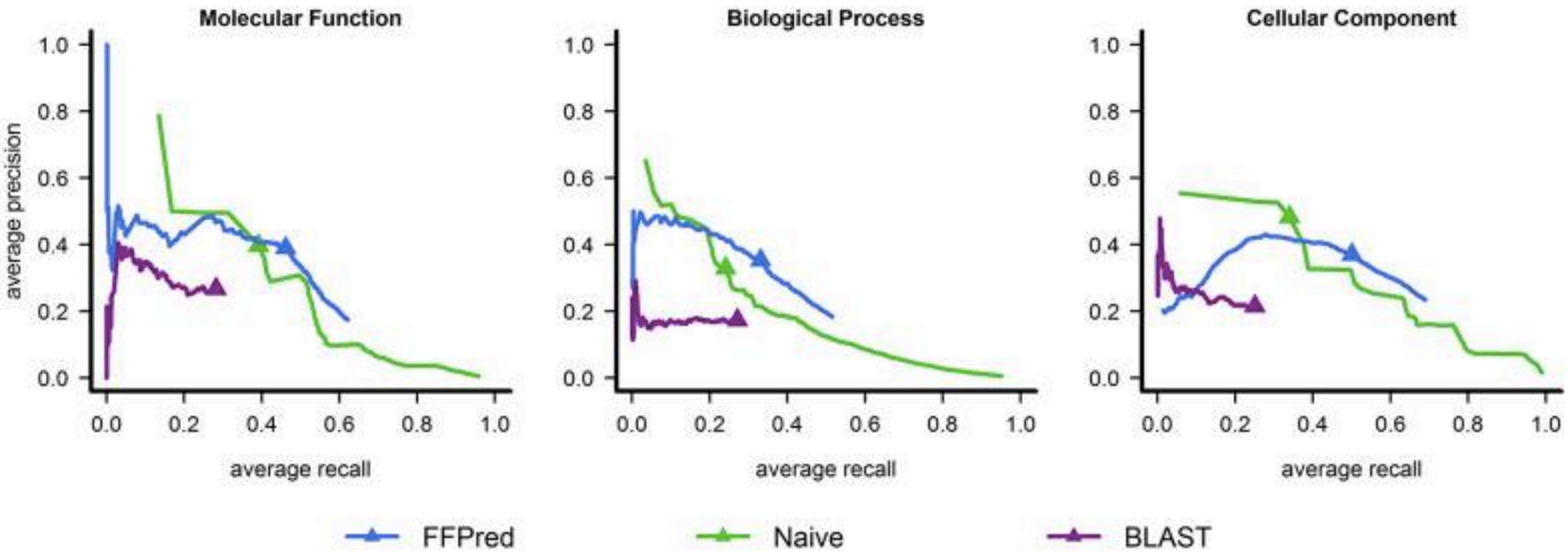
- transfer the function if the similarity is good enough

- Homolog/Paralog problem!

Mutual blast
Close paralogs have similar function

If blast is too slow, try ublast or blat

To blast or not to blast

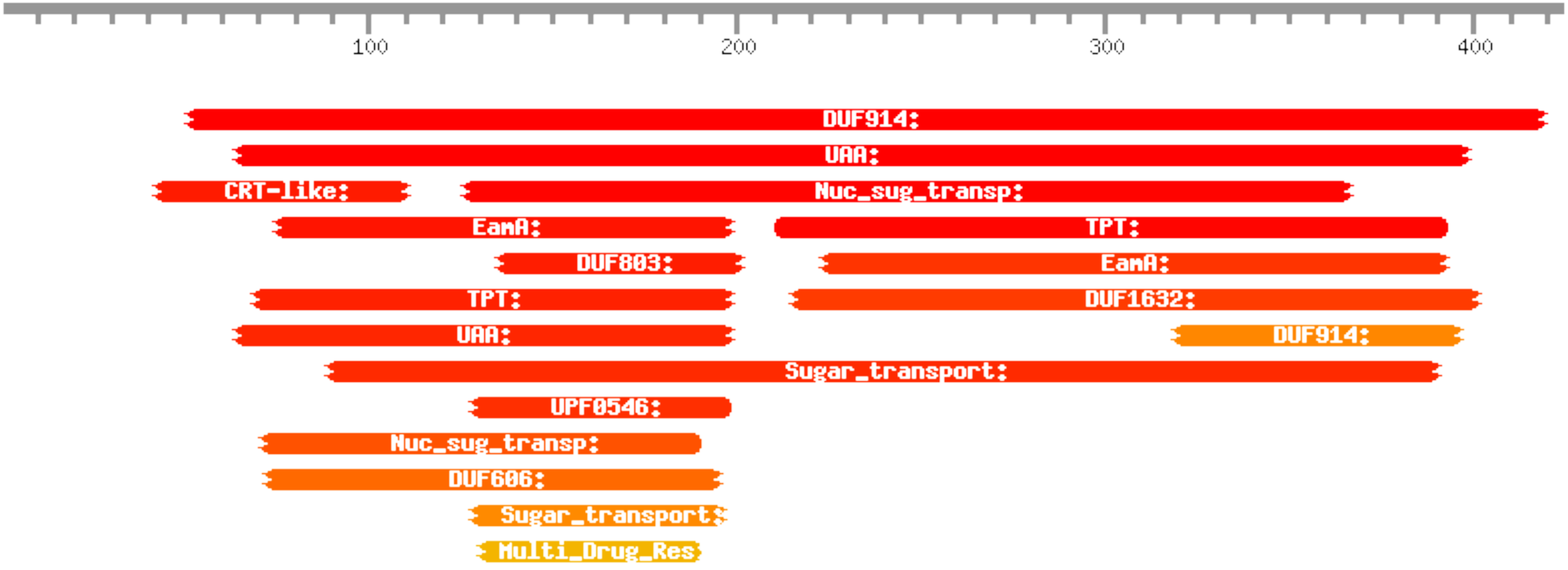


Blast is not very precise AND
Some GO annotations are too shallow to be useful

Pfam

- Protein Families
- HMMs of similar proteins (amino acid sequence)
- Pfam-A and Pfam-B (annotated vs extensive)
- Fast and quite reliable

- Some families are of unknown function



Feature prediction

- B-cell related:
Signal peptide
transmembrane helices
Glycosilation site

- T-cell and APC related
Cleaving sites
MHC transport & presentation

Both:
Similarity to host
hypervariability
non-effective epitopes (hard to predict)

Feature prediction

- B-cell related:
Signal peptide
transmembrane helices
Glycosilation site

- T-cell and APC related
Cleaving sites
MHC transport & presentation

}

CENTERFO
RBIOLOGI
CALSEQU
ENCEANA
LYSIS **CBS**

Both:
Similarity to host
hypervariability
non-effective epitopes (hard to predict)

Case study

Metagenomic sample of a cow disease

Polymicrobial (**treponemas**, fusobacteri, bacteriodetes)

Frequent reinfections

Annotate the proteins
and select which one is the most suitable target for a vaccine

Try and use the following tool:
Pfam (hmmer website)
TMHMM
SignalP
+ other if needed