

# TermBrain

the HPO term predictor

Jonathan Boidol, Rene Schoeffel, Yann Spöri

December 7, 2013

# Homology based function prediction

General approach:

- ▶ Search for annotated similar sequences with blast and hhblits (hits)
- ▶ Build subgraph of HPO containing the found annotations
- ▶ Calculate confidence for every annotation from some distance measure to the hits

# Preparations

- ▶ Prepare databases for annotated sequences
- ▶ Represent HPO Graph in predictor
- ▶ Merge trees corresponding to hits

# Features

- ▶ use neural network to calculate confidence per node
- ▶ each node is assigned features derived from the merged tree
  - ▶ number of hits
  - ▶ min. E-value
  - ▶ avg. E-value      [3, 0.0074, 0.45, 84, ...]
  - ▶ longest hit
  - ▶ ...