

MAIN DATA

- SlimFinder results (per dataset)
- Occurrences
- Random PPI
- Domain PPI
- ID Mapping (Sequence file)

ComparMotif

GO : Gene

GO : GO

SLIM (Pattern)

- Gene Hubs
- Domain Hubs
- Random Hubs
- Spokes
- GO x 3
- Hubs Parent → Results
- Spoke Gene → Occurrences
- ComparMotif → ComparMotif (?)
- Unifake (?)
- Unifake (?)

Lists linking to pages

Hub & Spoke Genes connected (DATASET3) → Results = Net

GENE

- GO (SF & Occ)
- Hub
- Spoke
- Interactions
- Domains
- Interdomain
- UPC (Net)
- Unifake (Net)
- Rer
- RurC

DATASET

- GO
- Hub (SF & Occ)
- Gene (Net)
- Interdomain

RANDOM (DATASET) x 2

- Gene/Patterns
- Dataset (Hub) → Results
- ComparMotif → Results

SlimFinder - (Main) Results table = Dataset-level results

then

- RunID
- Dataset
- Type
- Pattern
- parent run
- dataset name
- type of "hub" used to make dataset
- returned SLIM prediction

(Main) Occurrence table = Protein-level results

- RunID
- Dataset
- Pattern
- StartPos
- EndPos

Pairwise PPI Table

- Hub
- Spoke
- Domain PPI
- Domain Spoke

NEW

- Sequence Gene
- Species

10 Mapping Table = linking to external sites via Gene

EnLoc

- Gene
- GOID

GO Relationship

- GOID 1
- GOID 2
- Relationship

- GO
- GOID

- ComparMotif
- Pattern

- DATASETS
- Hub Dataset
- Spoke Gene
- Type
- Species

- PAPER
- Paper
- RunID