

Overview: Gene name, protein name, structure,  
expression pattern, function

Picture

Table of Contents

Table of Important Info

#### Gene

- Locus (cytogenetic, base coordinates [+/- strand], neighborhood)
- Common aliases (if not in ToI)
- Number of exons (in total)
- Span of gene (start of txn to polyA site (inc. introns); in bp)

#### Message

- Known isoforms (by name, size, exon usage, abundance)

#### Protein

- Known isoforms (by name, size, domain inclusion, abundance)
- MW, pI, amino acid composition/patterns/repeats
- Domains and motifs by homology
- Secondary structure
- 3° and 4° structure (disulfide bonds [Disulfind], topology within membrane)

#### DNA level regulation

- Promoter
- Transcription factor binding sites
- Expression pattern
  - ubiquitous or tissue-specific?
  - high, moderate, low mRNA abundance
  - NCBI GEO profile across all tissues
  - *in situ* hybridization data
  - protein localization and abundance

#### RNA level regulation

- splice enhancers
- mRNA localization
- predicted stem loops
- translation initiation
- miRNA targeting

#### Protein level regulation

- subcellular localization
- signal peptide? cleavage site?
- acetylation
- glycosylation

- lipid anchor attachment
- SUMOylation
- phosphorylation

#### Homology / Evolution

- Paralogs (with estimated time of divergence); phylogeny of paralogs
- Orthologs (range of organisms possessing orthologs)
  - Table of orthologs
  - Graph showing rate of evolution compared to other genes/proteins
- Distant homologs (most distant homolog detectable)
- Organisms showing no ortholog but proteins with homologous domains

#### Function / Biochemistry

##### Interacting Proteins

- TFs that might bind to regulatory sequence
- Proteins found in Y2H screens (developmental; functional)
- Partners that make sense

##### Clinical Significance

- Pathology
- Disease association
- Mutations (SNPs of interest)

##### References

##### Suggested Reading