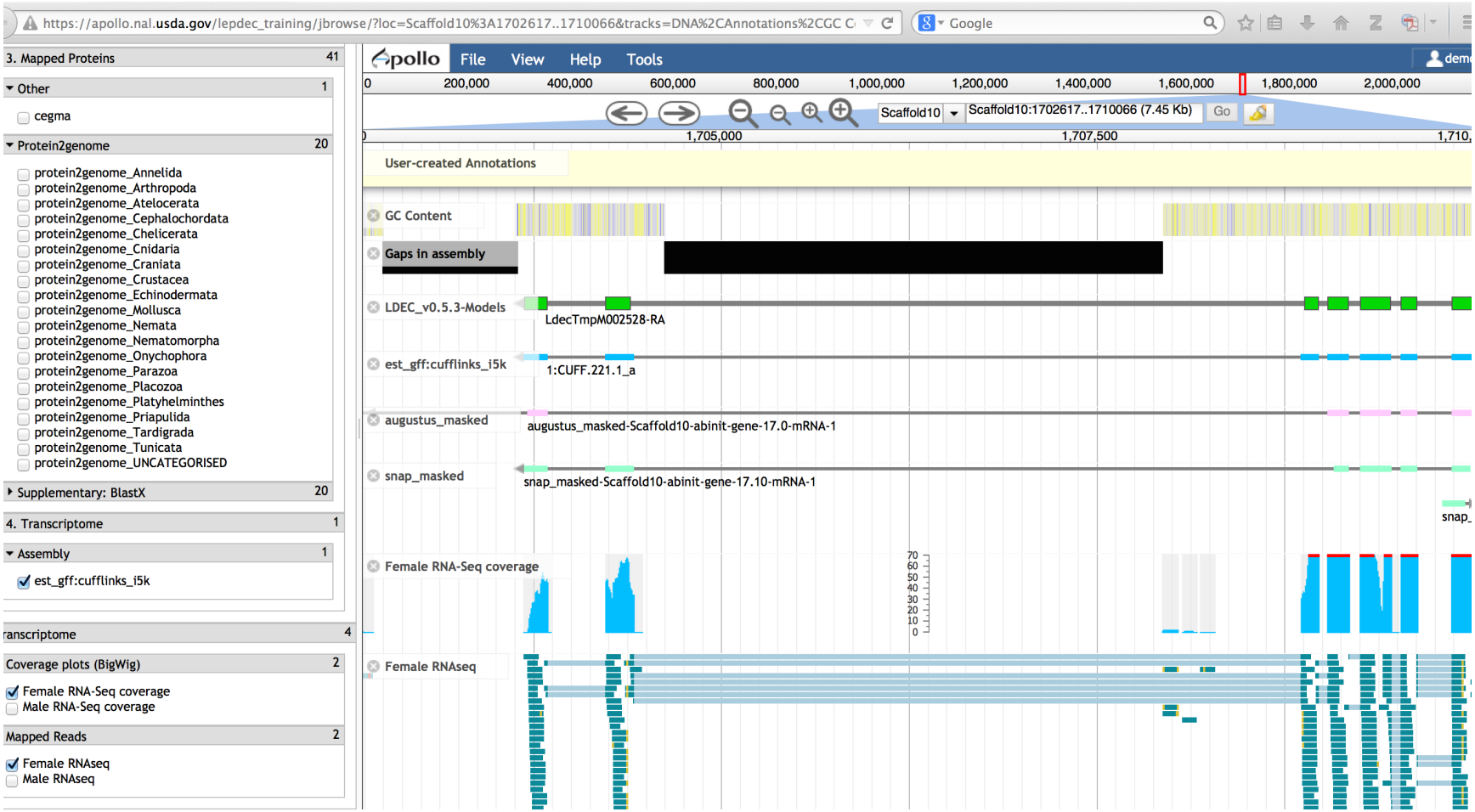
Web Apollo/JBrowse

- JBrowse is a web based genome browser
  - Visualize features that are mapped to a genome
  - These features are displayed as tracks
  - Many different types of data may be
- Web Apollo adds editing functions to JBrowse
  - Manual gene curation
  - Changes automatically saved back to server
  - Edits are visible to other annotators in real-time

Web Apollo

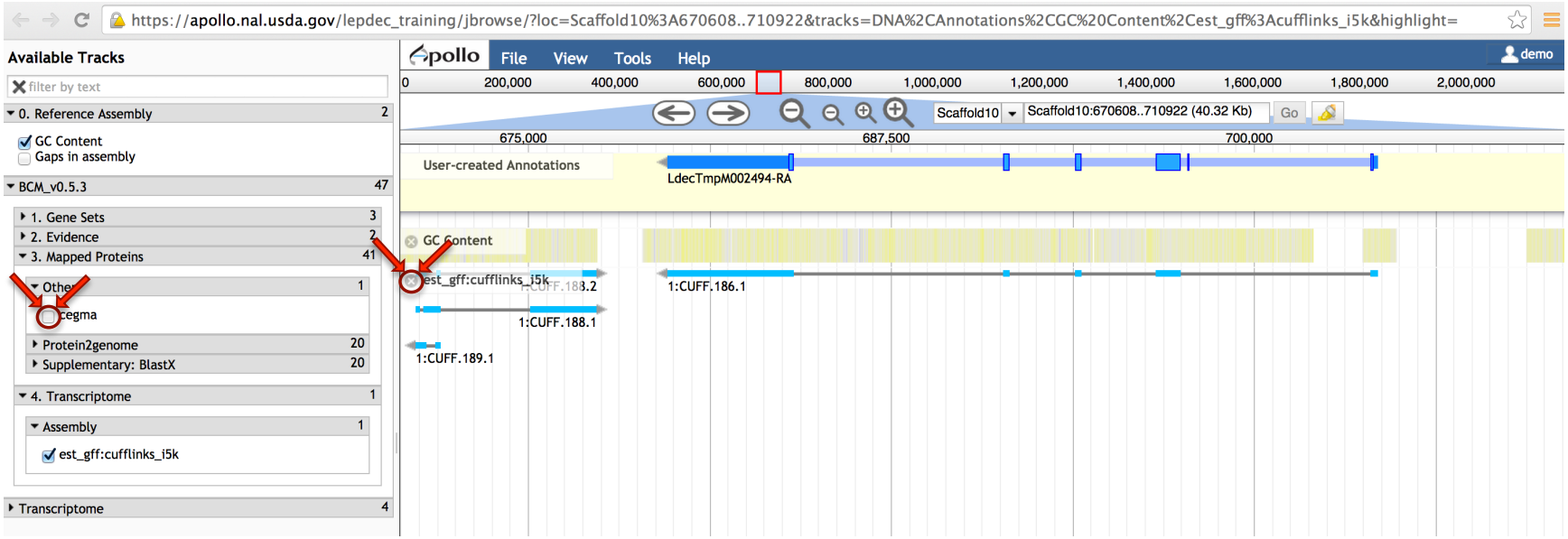

1
2
3
https://apollo.nal.usda.gov/lepdec_training/jbrowse/?loc=Scaffold10%3A1702617..1710066&tracks=DNA%2CAnnotations%2CGC C
Google
3. Mapped Proteins 41
Other 1
cegma
Protein2genome 20
protein2genome_Annelida
protein2genome_Arthropoda
protein2genome_Atelocerata
protein2genome_Cephalochordata
protein2genome_Chelicerata
protein2genome_Cnidaria
protein2genome_Craniata
protein2genome_Crustacea
protein2genome_Echinodermata
protein2genome_Mollusca
protein2genome_Nemata
protein2genome_Nematomorpha
protein2genome_Onychophora
protein2genome_Parazoa
protein2genome_Placozoa
protein2genome_Platyhelminthes
protein2genome_Priapulida
protein2genome_Tardigrada
protein2genome_Tunicata
protein2genome_UNCATEGORISED
Supplementary: BlastX 20
4. Transcriptome 1
Assembly 1
est_gff:cufflinks_i5k
ranscriptome 4
Coverage plots (BigWig) 2
Female RNA-Seq coverage
Male RNA-Seq coverage
Mapped Reads 2
Female RNAseq
Male RNAseq
Apollo File View Help Tools
Scaffold10
Scaffold10:1702617..1710066 (7.45 Kb)
Go
User-created Annotations
GC Content
Gaps in assembly
LDEC_v0.5.3-Models
LdecTmpM002528-RA
est_gff:cufflinks_i5k
1:CUFF.221.1_a
augustus_masked
augustus_masked-Scaffold10-abinit-gene-17.0-mRNA-1
snap_masked
snap_masked-Scaffold10-abinit-gene-17.10-mRNA-1
Female RNA-Seq coverage
Female RNAseq

Viewing evidence

1. The “Available Tracks” panel lists all tracks
   - Tracks are grouped hierarchically
     - Study or analysis
     - Categories based on evidence type
2. The Navigation panel describes the viewed region
   - Coordinate ranges for the entire sequence on top
   - Navigation controls including pan, zoom and search in the middle
   - Coordinate ranges for the current view along the bottom
3. The Evidence panel shows the features in genomic context
   - Each type of data is a separate track
   - Tracks may be displayed in different ways

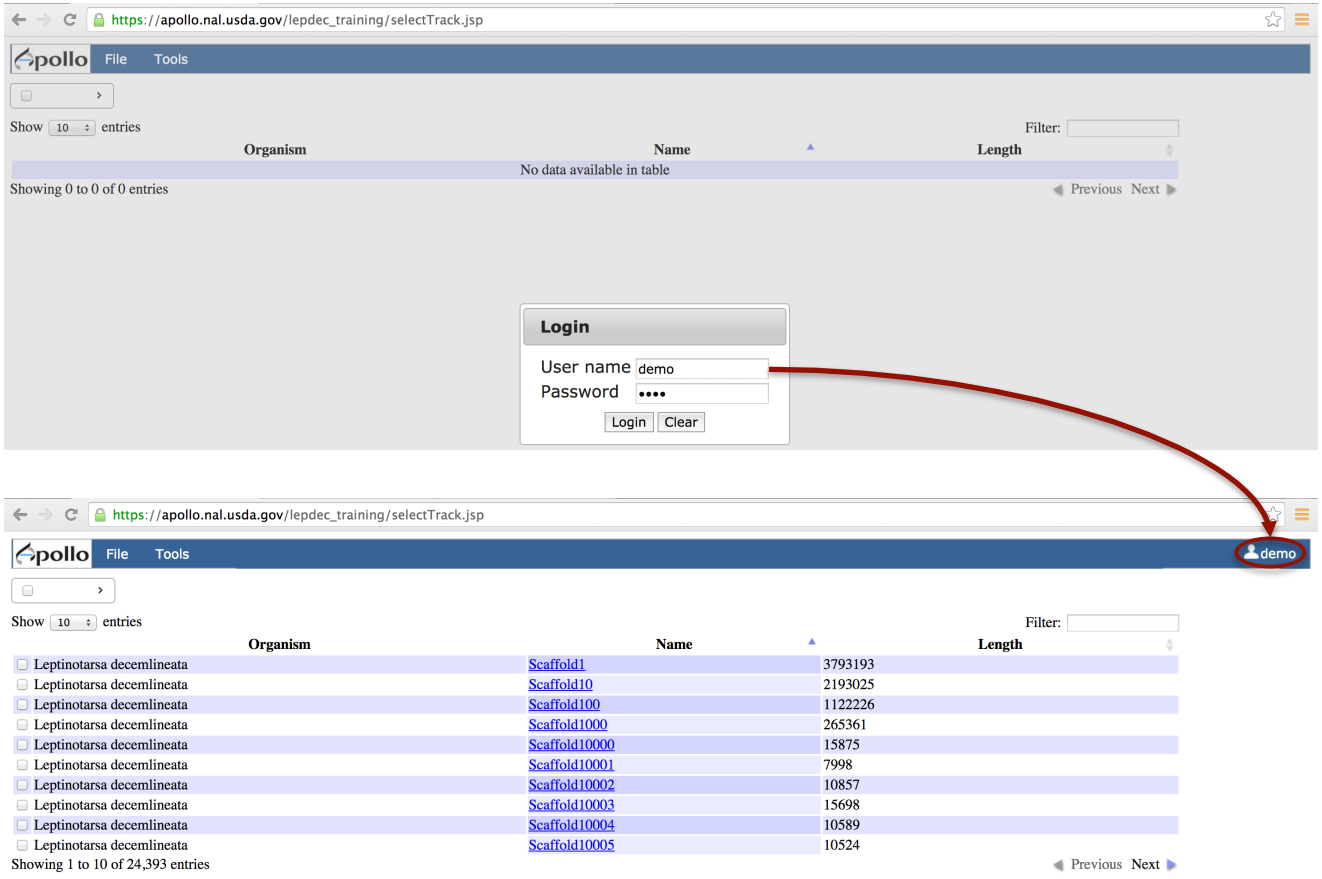
Viewing evidence

- Turning tracks on or off
  - Click a tracks checkbox to turn it on or off
  - Click ‘X’ by the track label also turns the track off



Login

- For scaffold selection view:

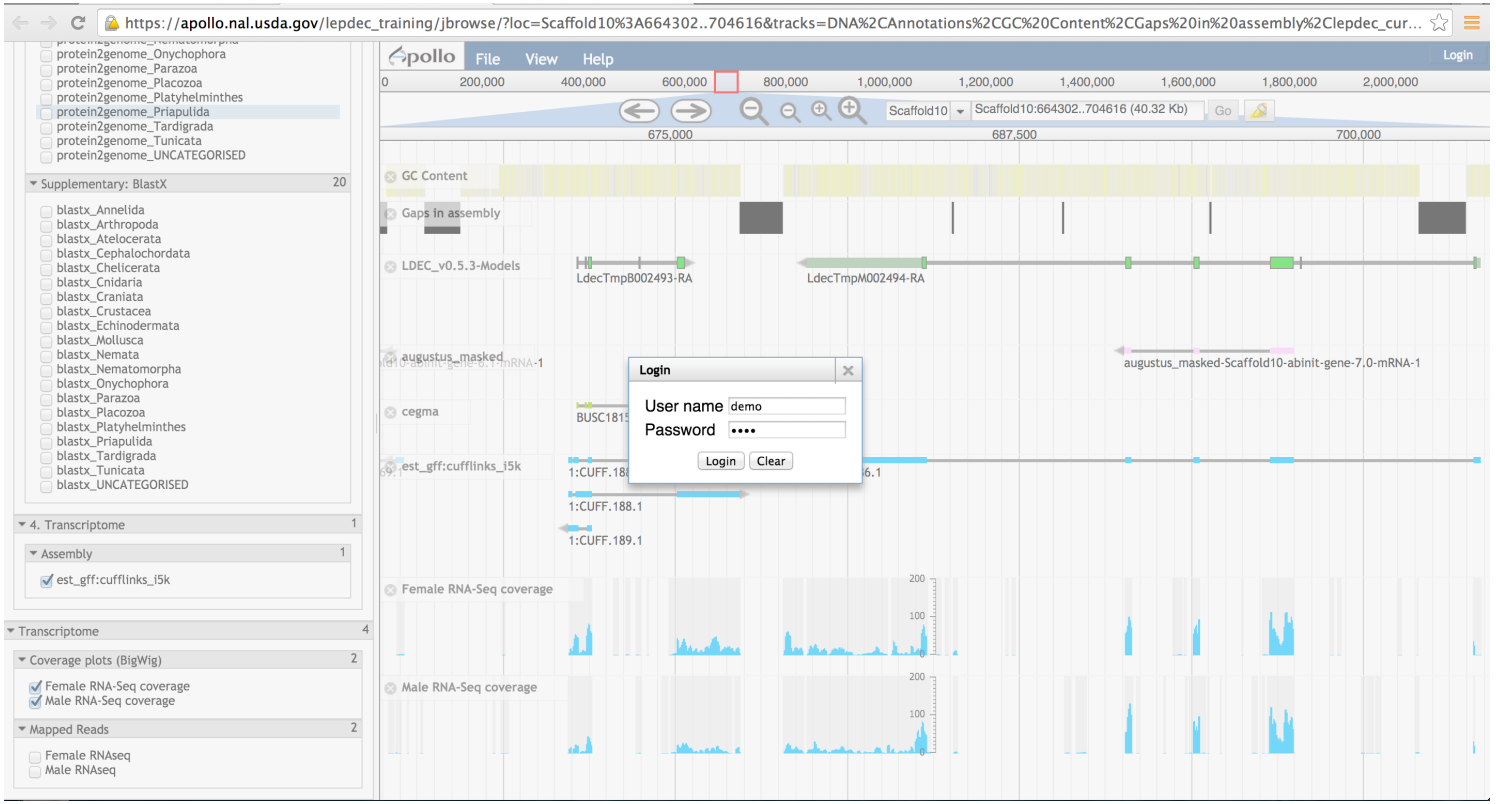


Login

- Directly from JBrowse genome browser

Login

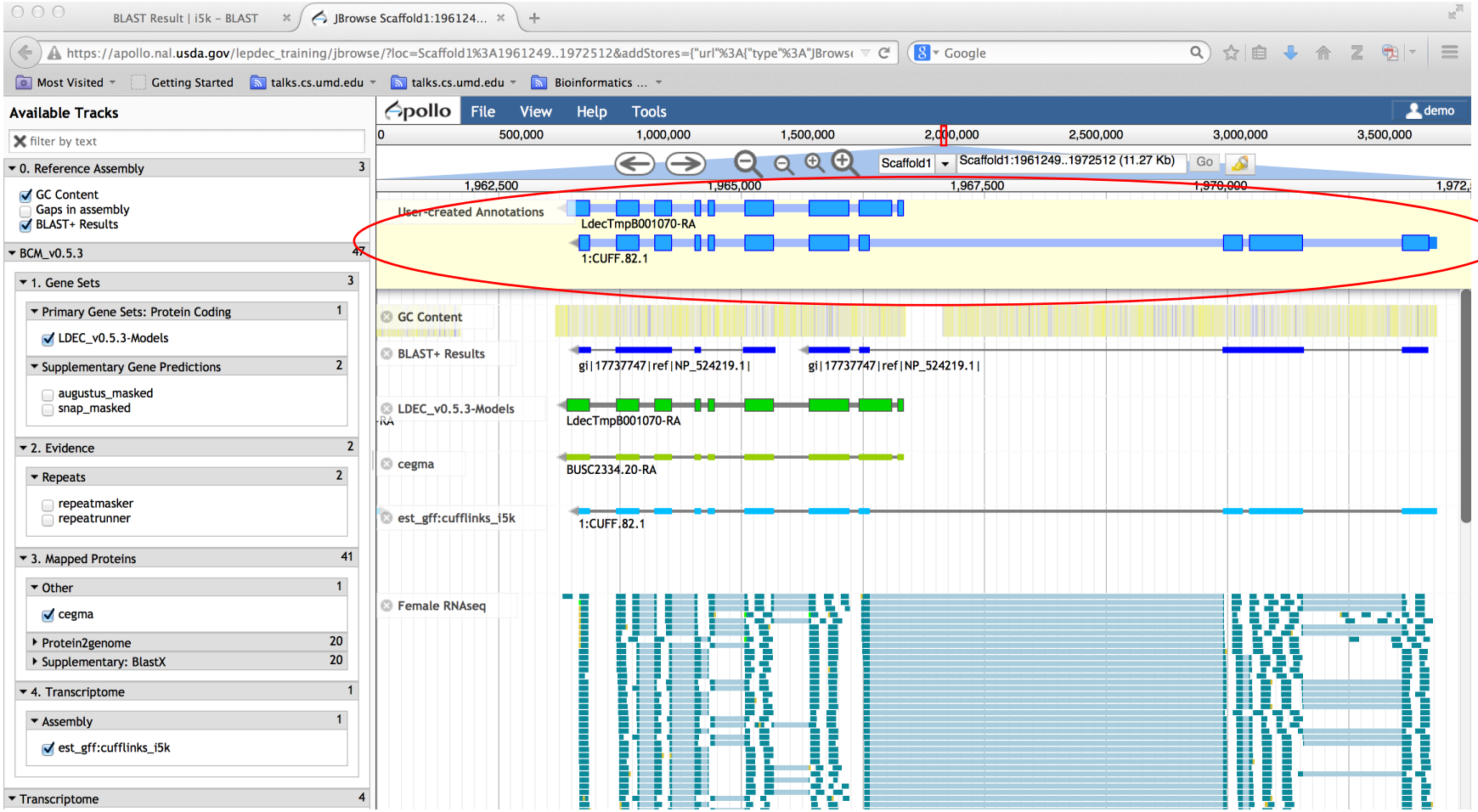
- Directly from JBrowse genome browser



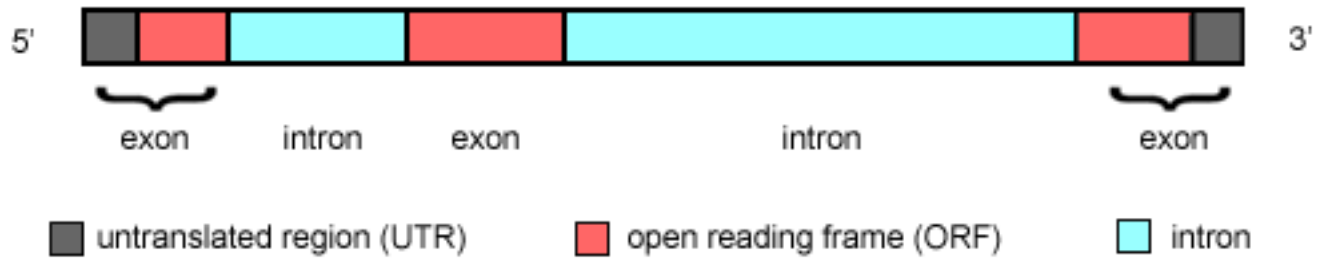
Login

- Directly from JBrowse genome browser

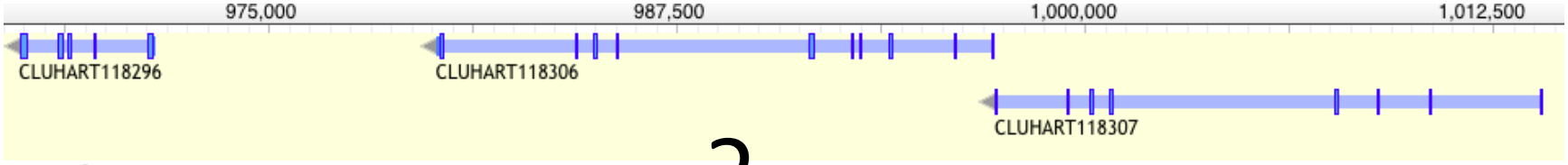
Here is your annotation



Functional annotation



Right, now we have our genes, but what do they do?



?

?

?

Insulin receptor?

Vesicle-trafficking protein?

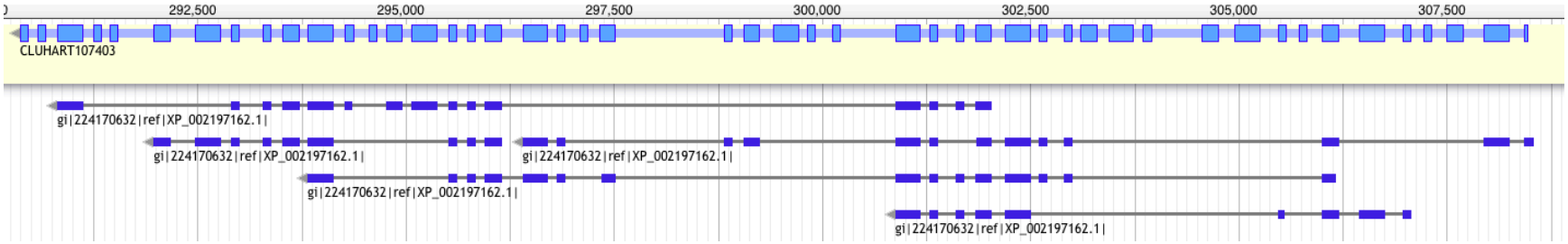
Alcohol dehydrogenase?

Aquaporin?

Transcription factor

MAP kinase kinase kinase?

But we have used proteins in our annotation!



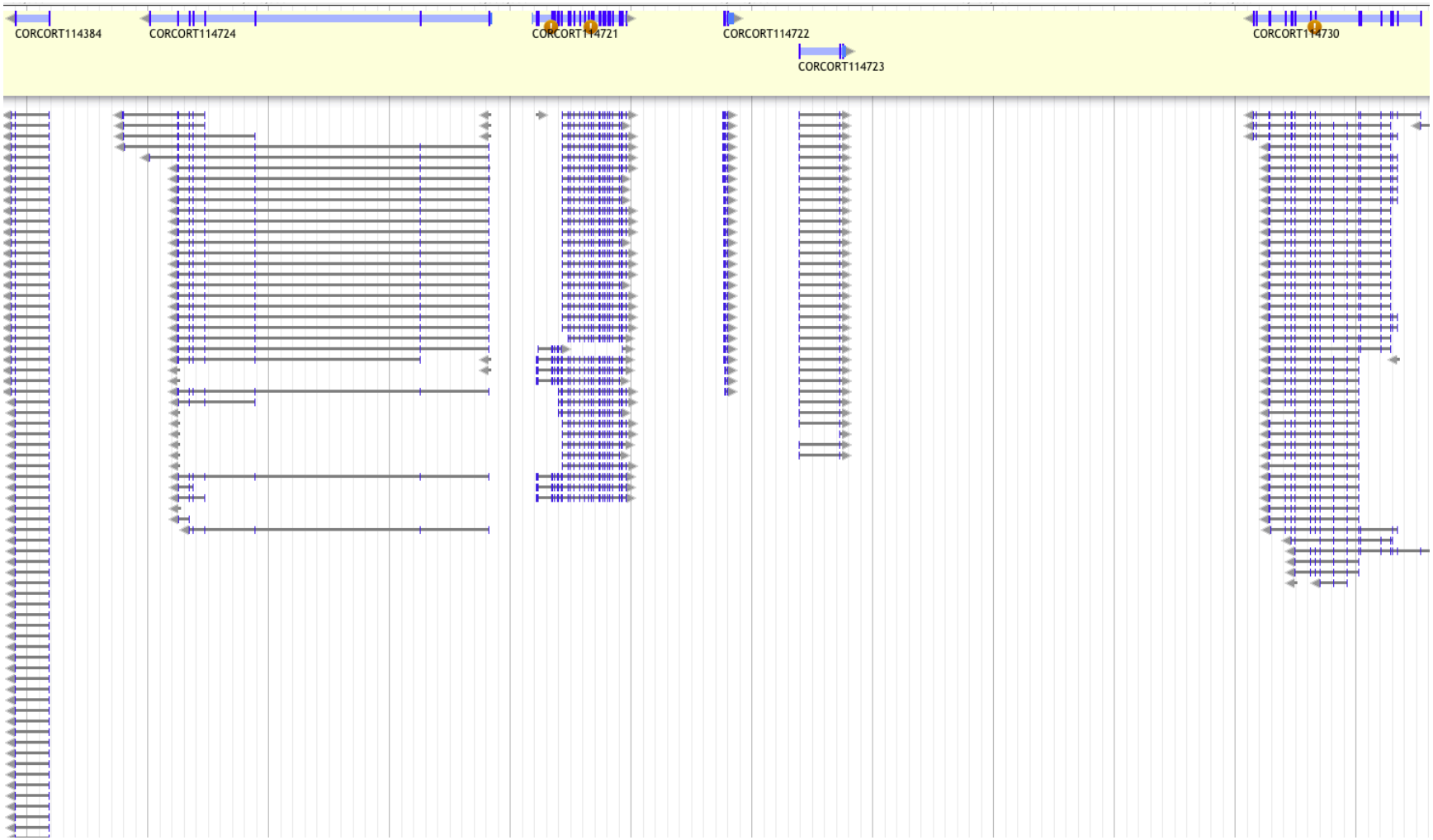
It is actually kind of complex…

… and Maker does not do this for you.

Extract sequences -> functional annotation

- Extract sequences from Webapollo
- Annotate the sequences functionally using Blast and InterProScan

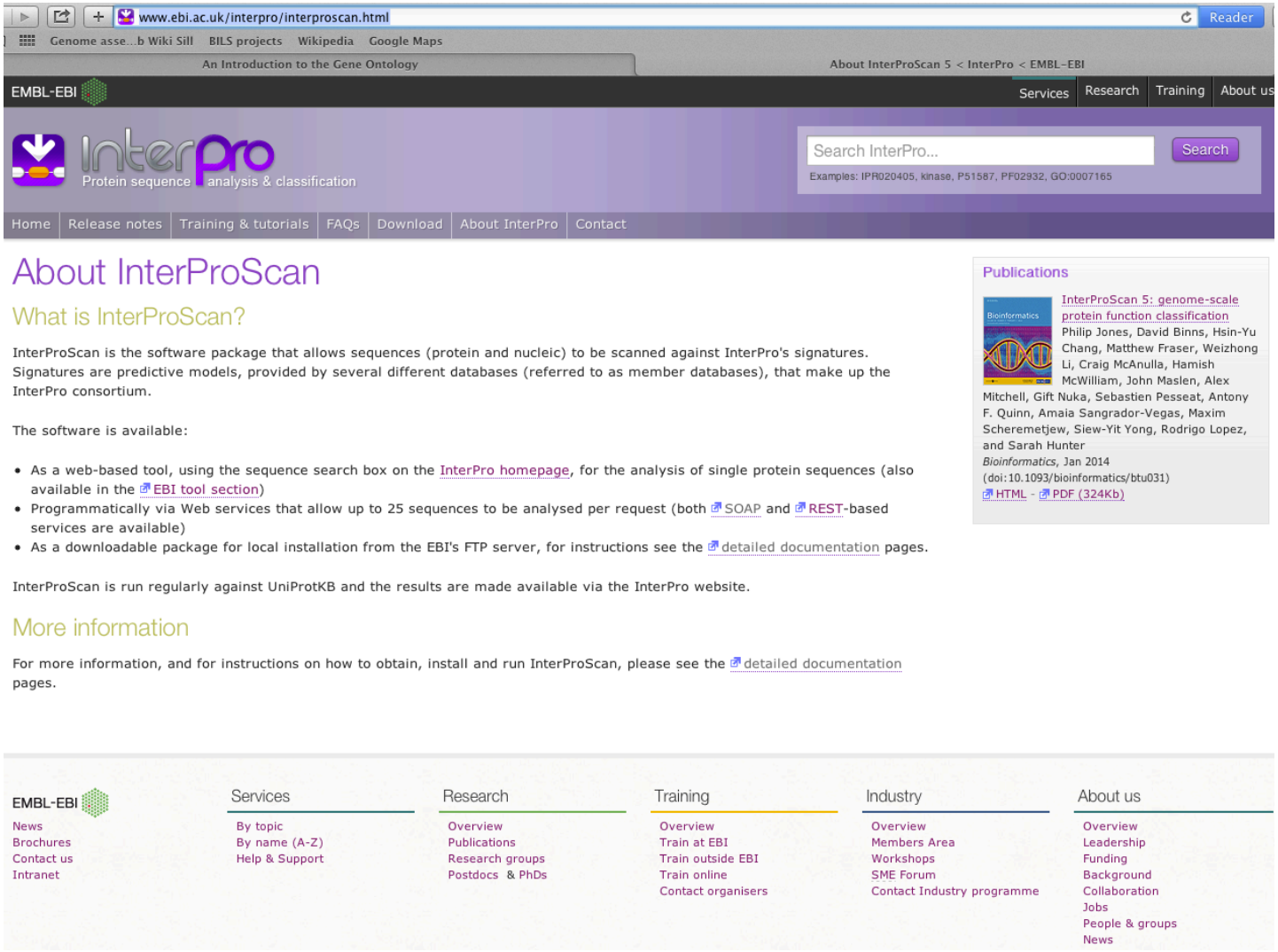
Blast-based annotation - proteins



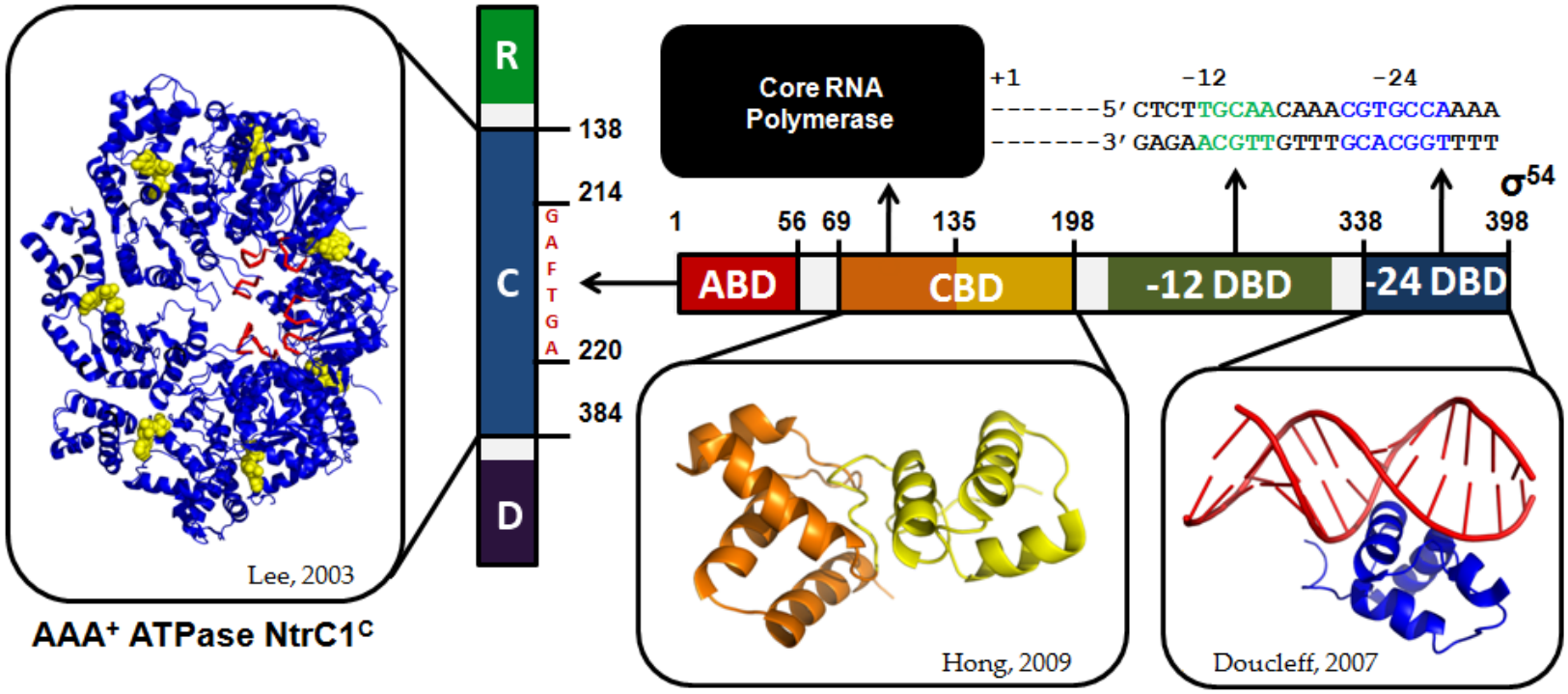
Blast-based approach

- Fairly fast and easy
- Orthology not certain - best blast-hit does not equal orthologous!

Interproscan



Sequence domains

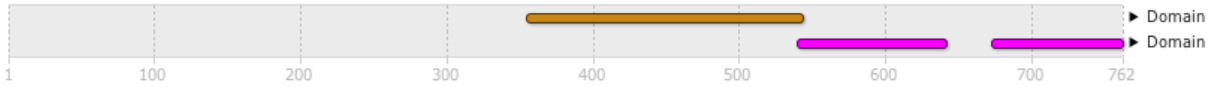


Interproscan results

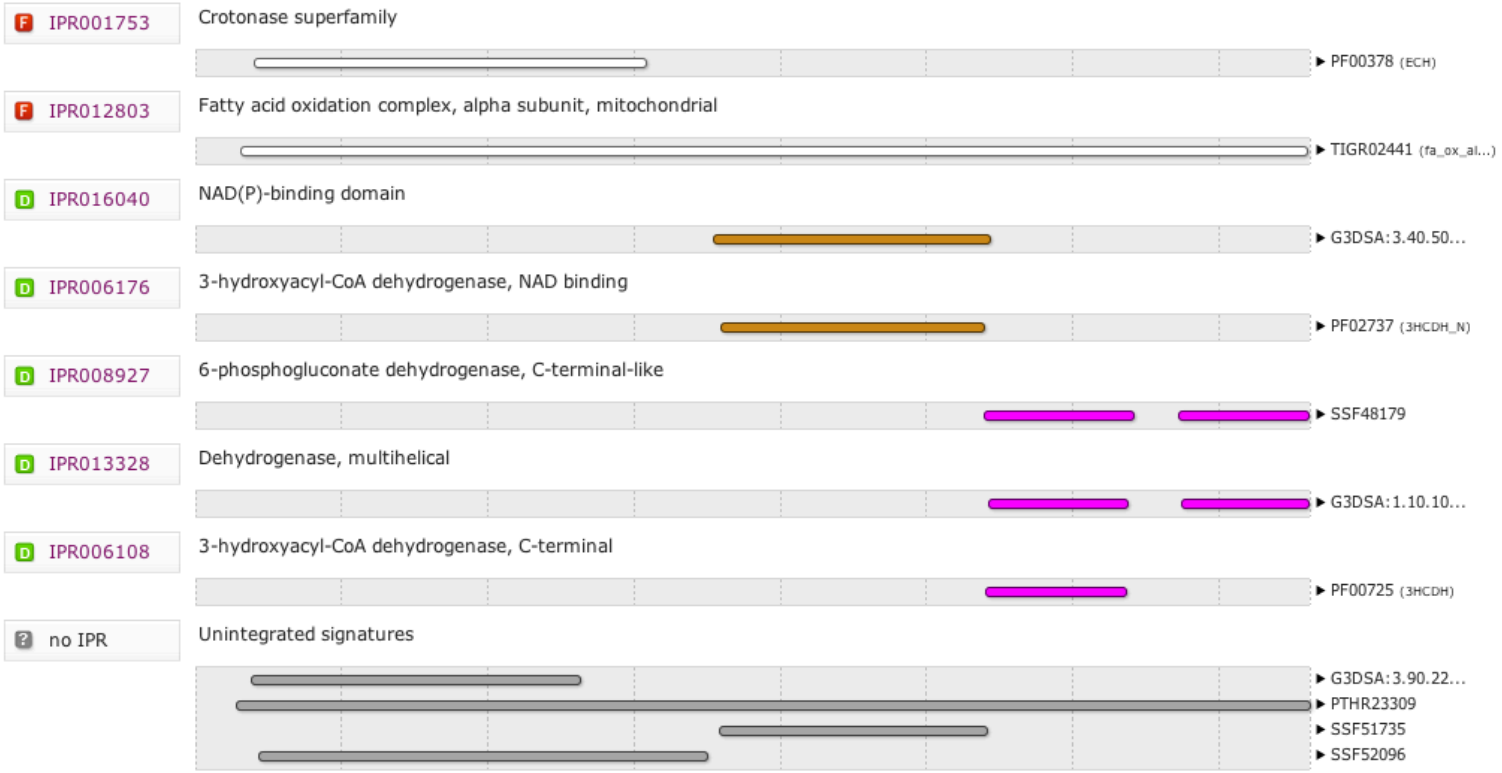
Protein family membership

- Crotonase superfamily (IPR001753)
  - Fatty acid oxidation complex, alpha subunit, mitochondrial (IPR012803)

Domains and repeats



Detailed signature matches



Interproscan results - GO terms

GO term prediction

Biological Process

GO:0006631 fatty acid metabolic process
GO:0006635 fatty acid beta-oxidation
GO:0008152 metabolic process
GO:0055114 oxidation-reduction process

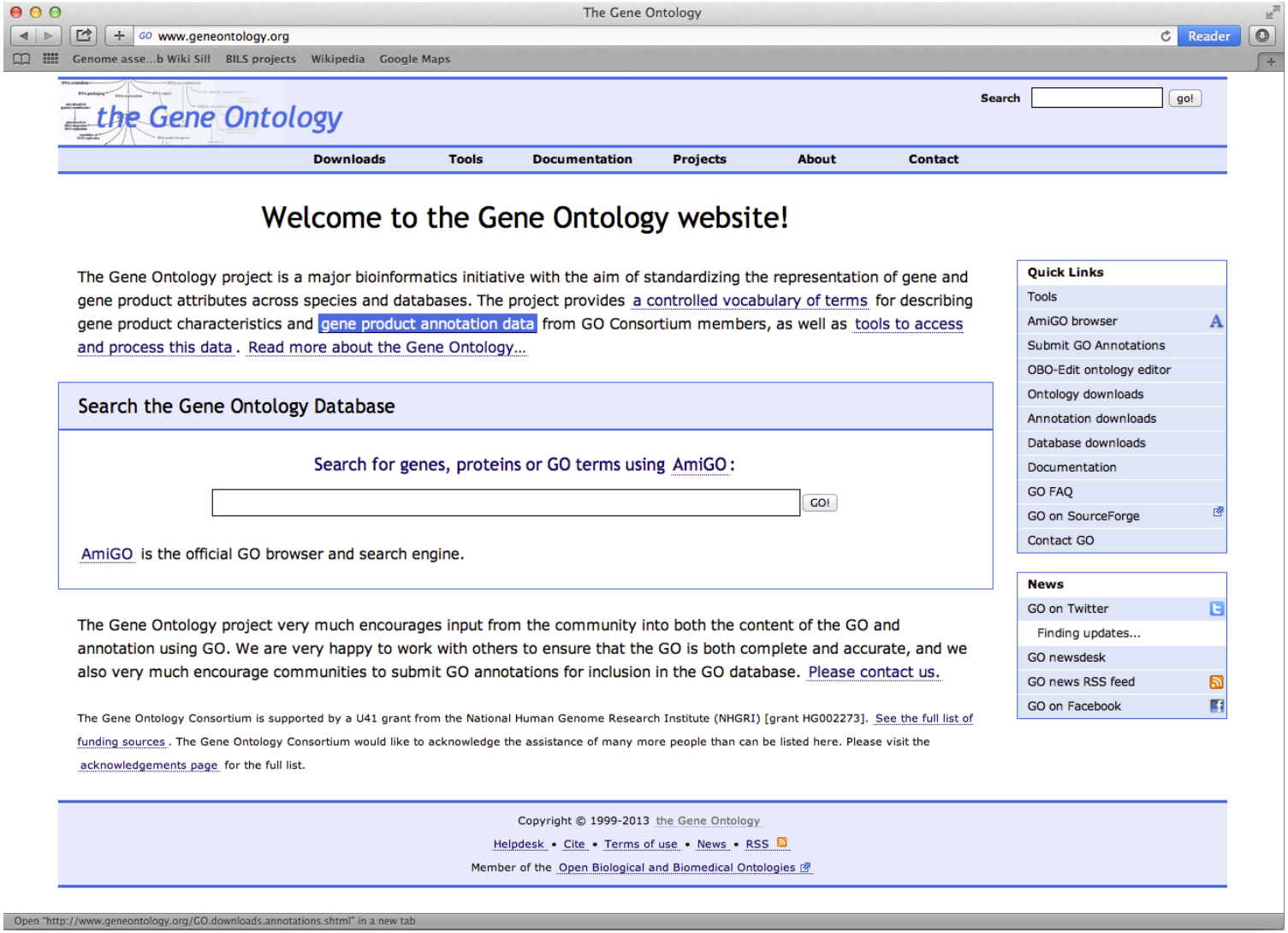
Molecular Function

GO:0003824 catalytic activity
GO:0003857 3-hydroxyacyl-CoA dehydrogenase activity
GO:0004300 enoyl-CoA hydratase activity
GO:0016491 oxidoreductase activity
GO:0016616 oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor
GO:0050662 coenzyme binding

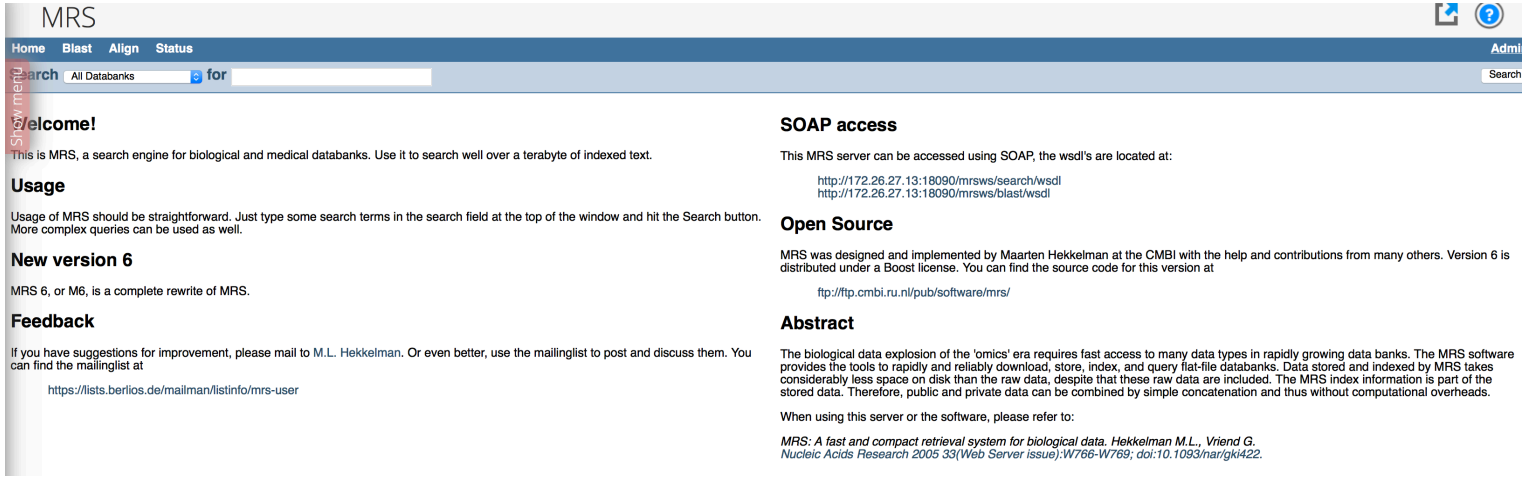
Cellular Component

GO:0005739 mitochondrion
GO:0016507 mitochondrial fatty acid beta-oxidation multienzyme complex

Gene Ontology



Gene Ontology

- C: Cellular Component, e.g., endoplasmatic reticulum, integral to plasma membrane
- P: Biological Process, e.g., pyrimidine metabolic process
- F: Molecular Function, e.g., catalytic acticivity, transporter activity

Use the eBioKit