



WP5

LiceBase Introductory Course

Part 4 Advanced steps

Michael Dondrup
Bergen 03.10.2014
Oslo 6.10.2014



Outline

- Overview on LiceBase, getting help
- Terms and definitions
- Basic functions: using Tripal
- Basic functions: using GBrowse
- Basic functions: BLAST
- Advanced topics: BLAST, GO, KEGG, and friends
- Exercise: searching genes by functional annotation
- Exercise: finding G O I's
- Annotation guide: RNAi experiment
- Exercise: annotate an RNAi experiment
- Annotation guide, novel genes and gene function
- Q & A

Script: Advanced annotation tasks

- There is some pre-computed automatic functional annotation: BLAST vs several dbs, KAAS, InterProScan
- So we don't need to run these methods each time
- Open: pre-computed Blast search
- Search: 'shaven', 'chitin synthase'
- Restrict search to NR database, check additional evidence for this gene opening the protein too.

Script: precomputed Blast

- **->Search BLAST results**
- **This is often the first choice to find genes**
- Search mask has various filtering fields, pre-filled with sensible values
- Search 'cuticle', restrict result further: e-value = 0, 1e-180, 1e-10
- restrict organism to 'tribolium', 'salmonis', 'ixodes', 'droso'
- sorting is by score, grouping is by gene, try out different sorting of the list

Script: GO

- -> organism -> salmon louse -> GO report
- GO - very fine grained vs. GO slims more coarse categories
- Use full GO for now
- Search term using AmiGO <http://amigo.geneontology.org>
- -> response to stimulus -> .. -> signal transduction
- export fasta file

Script: KEGG

- explain KEGG if not already known
- -> organism -> salmon louse -> KEGG report
- there is the BRITE hierarchy, open terms
- you have to use KEGG page for searching terms for now
- search 'chitin synthase'
- the commonly known KEGG pathway maps are at the bottom: Kegg Orthology (KO)
- open TCA cycle, show annotated genes
- you can open the reconstructed pathway

Exercise searching and finding genes by functional annotation

- Does the salmon louse contain a putative Major vault protein? Is MVP found in other hexapoda/crustaceans (hint: check blast vs nr for drosophila, check *Calanus finmarchicus* homologs); write down the EMLSAG00..... id(s) of the genes. Check all available evidence and download an image of the expression pattern. Check also if the gene model is on the correct strand (genome browser) and if the coverage is good. Hint: use precomputed blast.
- Check all the evidence for MVP, and read the link-outs on domains. Is the evidence convincing?
- Make a little overview over 'Calpain'-like genes and potential isoforms found in the salmon louse
- Export a gene set for the GO term 'symbiosis, encompassing ... parasitism' (hint: use AmiGO browser to search location in GO)
- Some people are interested in ftz-F1 like receptors, check the annotation in the KEGG reports. Check all the evidence, if possible, check gene models for correctness. (hint: ftz-F1 = Fushi tarazu-F1)
- Locate the chitin synthesis sub-pathway (hint: KO Amino sugar and nucleotide sugar metabolism (map00520)) and open the reconstructed pathway in KEGG, is it complete? Compare to *Drosophila*
- Find your gene of interest (GOI) by whatever input you have, e.g. from its DNA sequence and annotate it (adding synonym and Description if it has an EMLSAG ID), we will treat the case of not found GOIs later