Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Isak Sylvin

MAKER Annotation Suite

What, Why and How?

Yandell Lab
Department of Human Genetics - University of Utah

Why (is MAKER useful)?

Sequencing project

Sequence without annotations

```
GTGAAGGAAGGGGCAAGGGTTTTTCTGGCGCGGCTTGAGGACCTTAAAGGGCGCAAGACCTGGACCCCGG
CCCGGGCCGCCCTGAGGCGGAGCCTGGCCTTTGAGCCCGGGACCTATCCCCCGGCCATCGGGATGGTGGA
GCCCCTTCTACCCGGGGAAGCGGAGGGATGGATCCGGGAGGCCTACTACCTGGTGGCGGCCCTCTACGCC
CTCAAGGATGGGGAGCACCGGGAAGGGCGTACCCTGGCCCAGGCCCTAAGGGAAAAGGCGGGGGATTCCG
CCAGTGTGGAGCATCGGTTTCTGGCCCTTTTGGATGCGGACCGGGACCAGATCGCTTTCCGACTAAGGCA
AAGCGTGGTCCTGGTAGAAGGGGGTTTGGACTTCGCGCGGCTTTTGGCAGACCTGCTGGCCTGGTTTAGC
CCGGACCGGTGGGTGCAGGCCCGTTGGGCGAGGGAGTTTTACGGCGTGAAGGAAGAGGAGAAGGAGGAGG
TGAAGGCATGA
```

It's just letters!

What (is MAKER)?

Annotation ways

Manual:
Naming features by hand
SLOWEST

With a program:
Need to install it and know about all the options for it
FASTER & EASIER

With a pipeline/suite:
A single program that runs many programs. Need very little tinkering.
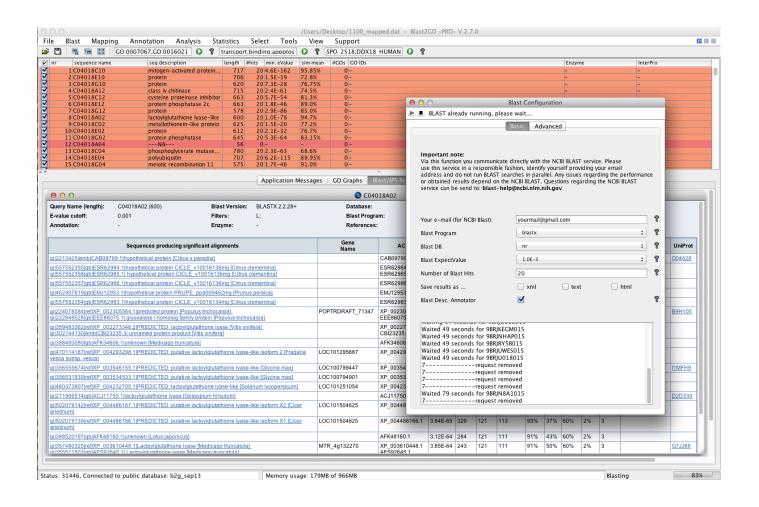*FASTEST & EASIEST*

Pipelines

Pipeline = A single program that runs many programs.

Many exist: BLAST2GO, ANNOVAR, MAKER, DIYA etc.

Most are hard to use, but easier than figuring out all components

Blast2Go screen example

MAKER Description

Ab initio gene prediction	EST alignment	Protein alignment	Repeats	Small RNA
SNAP	Exonerate	Exonerate	Repeatmasker	tRNAscan-SE
Augustus	BlastN	BlastX	Repeatrunner	snoscan
FGENESH*		tBlastX		
GeneMark-ES				
GeneMark-S				

*Commercial product

How (do we use MAKER)?

Overview

1) *Log onto mobaXterminal
2) *Load pre-installed modules
3) Create configuration file
4) Edit configuration file
5) Run a short MAKER test
6) Check the results
7) Run MAKER for real
8) Wait….
9) Wait….
10) Check the results
11) Ponder the results!

Running MAKER is easy

```
module load maker
maker –CTL
maker maker_bopts.ctl maker_exe.ctl maker_opts.ctl
```

About the files

Maker_exe.ctl

> Installation paths for all software; only relevant if installing

Maker_bopts.ctl

> Settings for all prediction programs like BLAST; feel free to tinker around

Maker_opts.ctl

> Input files (like genome) and analysis to run. ***MUST*** be tinkered with.

What we won't cover

Downloading and installing MAKER:
http://yandell.topaz.genetics.utah.edu/cgi-bin/maker_license.cgi
Linux experience necessary!

Extra resoures:

http://gmod.org/wiki/MAKER_Tutorial

Don't just click and go

Problems are NOT obvious:

    Did we compare it with enough data?

    Was the criteria for matches too strict or too loose?

    Was there something related to the type of species?

        Prokaryotic/eukaryotic/archea?

        GC-content?

        Repetitive segment content?

        Etc.

Take-away message

Individual programs
    Harder to start, error checking is fast
    Gives very specific predictions

Pipeline

    Easy to start, takes long to check for errors
    Solution: Test on a portion of the data!

    Gives many types of predictions
    Solution: Crop the output, or make sure you know what everything
            is.

HPC Maker

Very few prediction programs
Data is just as an example
Runs 5-10 minutes

If you finish early

1. Play with the .ctl files (bopts, opts)
2. See if you can run MAKER with very small files (<20 MB) from NCBI uploaded to HPC
3. If you sit on a LINUX pc and have time, try installing MAKER – It's relatively easy