



This site provides online material for the www.cosmoss.org **2nd Physcomitrella genome workshop** that was held

on

July 1st-3rd 2009

at

University of Freiburg, Germany



[Here](#) are more photos...

Contents

- [1 Venue](#)
- [2 Speakers](#)
- [3 Cosmoss accounts](#)
- [4 First day](#)

- ◆ 4.1 Morning session
 - ◇ 4.1.1 The P. patens genome
 - ◇ 4.1.2 Overview of cosmoss.org resources
 - ◇ 4.1.3 Cosmoss sequence retrieval
 - ◇ 4.1.4 The cosmoss.org databases
 - ◇ 4.1.5 Hands-on experience I
 - 4.1.5.1 Menu, documentation, mailing list and wiki
 - 4.1.5.2 Sequence Retrieval
 - 4.1.5.3 Keyword search
- ◆ 4.2 Afternoon session
 - ◇ 4.2.1 BLAST, homology and hit filtering
 - ◇ 4.2.2 Hands-on experience II
 - 4.2.2.1 (batch) BLAST & hit filtering
 - ◇ 4.2.3 Reciprocal BLAST
 - ◇ 4.2.4 CSV BLAST
 - ◇ 4.2.5 Hands-on experience III
- 5 Second day
 - ◆ 5.1 Morning session
 - ◇ 5.1.1 The cosmoss.org genome browser
 - 5.1.1.1 other gbrowse instances for additional genomes
 - ◇ 5.1.2 Hands-on experience IV
 - ◇ 5.1.3 Genome browser: hidden treasures
 - ◇ 5.1.4 Hands-on experience V
 - 5.1.4.1 BLAST gbrowse integration
 - 5.1.4.2 Advanced navigation and zooming
 - 5.1.4.3 Exporting sequence annotations and publication quality images
 - 5.1.4.4 Highlighting
 - 5.1.4.5 Displaying custom annotation
 - ◆ 5.2 Afternoon session
 - ◇ 5.2.1 V1.0., 1.1, 1.2, 1.5, 2.0 ?!
 - ◇ 5.2.2 Annotation interface
 - ◇ 5.2.3 Hands-on experience VI
 - ◇ 5.2.4 Hands-on experience VII

Venue

Biology Campus, Hauptstr. 1, Computer Pool & Lounge

Speakers

- Stefan Rensing
- Andreas Zimmer

- [Daniel Lang](#)

Cosmoss accounts

You have to be a registered cosmoss user in order to access some of linked material.

For the workshop participants, accounts have been created (Unless you already had one) with an initial password that has been sent by e-mail. Please change it as soon you [log in](#).

Everyone else is also invited to register! Please contact helpdesk-cosmoss@uhura.biologie.uni-freiburg.de for an account!

First day

Morning session

The *P. patens* genome

speaker

Stefan

slides

[Genome.pdf](#)

Overview of cosmoss.org resources

speaker

Daniel

slides

[cosmoss_overview.pdf](#)

Cosmoss sequence retrieval

speaker

Andreas

slides

[Sequence Retrieval 09.pdf](#)

The cosmoss.org databases

speaker

Daniel

slides

[cosmoss_databases.pdf](#)

Hands-on experience I

Using cosmoss.org (help, documentation), sequence retrieval, databases

prepared by

Andreas and Daniel

Menu, documentation, mailing list and wiki

Hint: Use STRG + left mouse click to open a link in a new window



Open

<http://www.cosmoss.org>

Tasks

1. Explore the the menu system
2. Find the FAQ
3. Find the BLAST documentation - What are the e-value threshold defaults?
4. Explore the wiki - Find out Daniel's ICQ number
5. How many predicted genes are in Genome Annotation V1.0?
6. Continents of the World: Where can *Physcomitrella* be found?

Sequence Retrieval

documentation

[Sequence Retrieval](#)

Familiarize yourself with the Cosmoss Retrieval system:

Tasks

Try to retrieve the following sequences:

BJ172647

BJ179866

PP015013150R

Cosmoss_workshop_2009

You can access the sequence retrieval via the transcriptome and genome menu.

Hint: These accession numbers above are from the [pp0304](#) annotated virtual transcript database. You can retrieve multiple sequences by providing their accession numbers as a comma- or space-separated

For a vast number of sequences you could upload your request in a text file.

File format: text
Provide accession number per line.

Select the following accession numbers and copy them into a new text file (e.g. notepad) and save it.

```
Pp1s459_1V2.1|cosmoss|Phypa_173430  
Pp1s7_181V2.1|cosmoss|Phypa_201973  
Pp1s204_88V2.1|cosmoss|Phypa_221004  
Pp1s204_91V2.1|cosmoss|Phypa_59935  
Pp1s352_13V2.1|cosmoss|Phypa_152025  
Pp1s312_46V2.1|cosmoss|Phypa_225236  
Pp1s66_47V2.1|cosmoss|Phypa_106210  
Pp1s545_4V2.1|cosmoss|Phypa_109367  
Pp1s459_9V2.1|cosmoss|Phypa_110675  
Pp1s56_169V2.1|cosmoss|Phypa_125839  
Pp1s56_166V2.1|cosmoss|Phypa_125903  
Pp1s188_40V2.1|cosmoss|Phypa_168764  
Pp1s545_4V2.2|cosmoss|Phypa_201189  
Pp1s109_143V2.1|cosmoss|Phypa_233894  
Pp1s352_57V2.1|cosmoss|Phypa_61317  
Pp1s66_46V2.1|cosmoss|Phypa_77574  
Pp1s12_223V2.1|cosmoss|Phypa_115069  
Pp1s46_41V2.1|cosmoss|Phypa_123666  
Pp1s251_42V2.1|cosmoss|Phypa_146969  
Pp1s161_16V2.1|cosmoss|Phypa_87740  
Pp1s161_32V2.1|cosmoss|Phypa_87752  
Pp1s339_14V2.1|cosmoss|Phypa_151552  
Pp1s459_12V2.1|cosmoss|Phypa_8310
```

Goto the [cosmoss.org](#) sequence retrieval:

Select database: *P.patens.V1.2* proteins

This [database](#) contains all *P.patens* released proteins. In comparison to the release V1.1 all gene models overlapping with transposons, non-protein-coding genes (e.g. tRNA genes) were removed. The *Physcomitrella patens* genome accession numbers work for both transcripts and proteins databases. Just change the database to change to your favored sequence type.

Browse for the previously created file and submit your request

Save the sequences in a new file in FASTA format

Select only a subset of the sequences and save it to a new file

Keyword search

documentation

Keyword search

How many *geranylgeranyl pyrophosphate synth(et)ases* (GGPS) are in the virtual transcriptome?

Try to find the corresponding pp0304 virtual transcripts by keyword search!

Tasks

1. Read the documentation
2. Use the simple search menu to find the GGPS's in pp0304
3. Play around with the advanced search option.

This is the advanced query that works:

```
"geranylgeranyl"[DESC] AND "phosphate"[DESC] AND "synth"[DESC]
```

Finally, here are the two loci in the genome:

- Pp1s3_331V2.1
- Pp1s8_136V2.1

The two initial pp0304 transcripts are highlighted with a yellow box.

Afternoon session

BLAST, homology and hit filtering

speaker

Stefan

slides

BLAST.pdf

Hands-on experience II

(batch) BLAST & hit filtering

prepared by

Stefan

material

BLAST_hands_on.pdf

Reciprocal BLAST

speaker

Andreas

presentation

[Reciprocal BLAST searches 09.pdf](#)

CSV BLAST

speaker

Stefan

presentation

[CSV-BLAST.pdf](#)

Hands-on experience III

CSV-BLAST and reciprocal BLAST

prepared by

Stefan and Andreas

material

- [CSV-BLAST hands on.pdf](#)
- [Hands on reciprocal blast.pdf](#)

Second day

Morning session

The cosmoss.org genome browser

speaker

Andreas

presentation

[Basic GenomeBrowser 09.pdf](#)

other gbrowse instances for additional genomes

- [Physcomitrella plastid genome](#)
- [Physcomitrella mitochondrial genome](#)

- [Arabidopsisgenome TAIR7](#)
- [Selaginella genome FM2](#)
- [Vitis genome V1.0](#)
- [organellar genomes](#)

Hands-on experience IV

Genome browser basics

prepared by

Andreas and Daniel

presentation

[Hands on gbrowse basics.pdf](#)

Genome browser: hidden treasures

speaker

Daniel

presentation

[cosmoss_gbrowse_treasures.pdf](#)

links

- [Scalable Vector Graphics \(SVG\)](#)
 - [Inkscape](#) An Open Source vector graphics editor, with capabilities similar to Illustrator, CorelDraw, or Xara X, using the W3C standard Scalable Vector Graphics (SVG) file format.
-

Hands-on experience V

Genome browser: customization and special features

prepared by

Daniel

documentation gbrowse

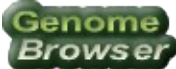
[general gbrowse help](#)

BLAST gbrowse integration

documentation

[BLAST gbrowse integration](#)

Tasks

1. BLAST with an arbitrary transcript/EST vs the scaffolds and follow the  link.
2. Compare results when BLASTing w/o low complexity filtering!
3. Compare the BLAST to the spliced-alignment results. Are there lonely exons?

other gbrowse instances for additional genomes

4. BLAST with the *Arabidopsis* protein AT5G13930 .1 vs the v1.2 gene models and follow the



link.

5. Compare one of the hit *Physcomitrella* loci vs the *Arabidopsis* locus [AT5G13930](#)

Advanced navigation and zooming

Tasks

1. Find out the definition of a *gene* in [SQ](#) by using the *Ontology_term* cross-link in a gene feature's mouse-over window
2. Zoom into a CDS exon and back again to the region of its *mommy* mRNA or gene
3. Play around with the zoom function to inspect EST and cDNA spliced alignments!

Exporting sequence annotations and publication quality images

Tasks

1. Export the upstream 10kbp your most favorite region to [FASTA](#) format and reverse it when necessary to reflect the gene's orientation.
2. *OPTIONAL*: Get another upstream region and try to find shared putative promoter elements using e.g. [AlignACE](#)
3. Save your most favorite region as a [png](#) image.
4. Check out the exemplary PDF created with Inkscape from the SVG of a region: www.cosmoss.org/inkscape_example.pdf

Highlighting

example locus

[scaffold_29:573541..579040](#)

documentation

[available colors](#) (scroll down to *Colors*)

Tasks

1. Let gbrowse highlight a feature and zoom out again, in order to see whether it overlaps with another feature in another track ([highlight and zoom to the region of this feature](#) link in the mouse-over)
2. Highlight your favorite gene model for the locus using the *Highlight feature* box in the *Display panel*.
3. Visualize an PCR experiment on the genomic locus! You've just amplified and sequenced an genomic PCR product for the locus using the primer coordinates below. Highlight the genomic region using the *Highlight regions* box in the *Display panel*.

forward primer mapping

[scaffold_29:575397..575417](#)

reverse primer mapping

[scaffold_29:577964..577950](#)

at	<input checked="" type="checkbox"/> MIPS ANGELA long_terminal_repeat	<input checked="" type="checkbox"/> Helitrons Bennetzen lab
oson	<input checked="" type="checkbox"/> TE related PFAM domains	
		<input type="button" value="Configure tracks..."/> <input type="button" value="Update"/>
Display Settings		
	Highlight feature(s) (feature1 feature2...)	
	<input type="text" value="all_Phypa_205678@yellow"/>	
	Highlight regions (region1:start..end region2:start..end)	
	<input type="text" value="scaffold_29:577964-577950"/>	
	Region Size (bp)	
	<input type="text" value="10 kbp"/>	
		<input type="button" value="Update"/>
Add your own tracks		

Displaying custom annotation

You can draw your own features and have your own custom track in the browser!

documentation

[gbrowse custom annotation help](#)

example locus

[scaffold_29:573541..579040](#)

Tasks

1. Go through the documentation
2. Visualize an RT-PCR experiment on the genomic locus! You've just amplified and sequenced an RT-PCR product for the locus below. Display it as custom annotation using the coordinates:

```
RT-exon1 scaffold_29:575397..575570
RT-exon1 scaffold_29:576849..576875
RT-exon1 scaffold_29:577042..577335
RT-exon1 scaffold_29:577629..577692
```

See the [example file](#) which combines everything. Adjust it, if you like!

Afternoon session

V1.0., 1.1, 1.2, 1.5, 2.0 ?!

speaker

Stefan

Highlighting

slides

[versions.pdf](#)

Annotation interface

speaker

Daniel

slides

[genonaut.pdf](#)

Hands-on experience VI

Annotation interface

prepared by

Daniel

Hands-on experience VII

?Putting the pieces together - The Summary HowTo? & (on your own/Q&A)

The Summary HowTo

prepared by

Stefan (caution: this is using v1.2!)

slides

[HowToSummary.pdf](#)

material (in the order you might need it)

[the paper](#)

[A.t. MS1 query as FASTA](#)

[P.p. MS1 homologs as FASTA](#)

[A.t. MS1 homologs as FASTA](#)

[the BLAST csv result against Swissprot](#)

[the Excel table from above - three tabs](#)

[the n.r. acc. nos. from above](#)

[the combined protein sequences in FASTA format](#)

V1.0., 1.1, 1.2, 1.5, 2.0 ?!

the alignment of above in FASTA format

the alignment in PHYLIP format

sneak preview

the two P.p. loci are

```
scaffold_314:310561..315560  
scaffold_271:262092..267091
```

an ML tree:

```
((A9TNA4_PHY:0.065768,A9TTIO_PHY:0.028371):0.304285,(((A5AL84_VIT:0.048466,A7P414_VIT:0.000000)
```

bootstrapped:

```
((A9TNA4_PHY:0.065768,A9TTIO_PHY:0.028371)100:0.304285,(((A5AL84_VIT:0.048466,A7P414_VIT:0.0000
```