# Lichtheimia - Basal fungal gene expression analysis completed by ncRNA annotation

## Konstantin Riege<sup>1</sup>

<sup>1</sup>Faculty of Mathematics und Computer Science RNA Bioinformatics and High Throughput Analysis Friedrich-Schiller University Jena

TBI Winterseminar, 16.02. 2015

伺下 イヨト イヨト

Phylogeny Pathogenicity

- L. corymbifera, L. ramosa and L. hyalospora
- Basal lineage fungi between the fork of animals and plants
   L. corymbifera is located at the base of Mucorales



Figure : Phylome analysis of single copy genes Outgroup: Microsporidia (spore forming eukaryotic parasites)

Phylogeny Pathogenicity

- L. corymbifera, L. ramosa
- $\rightarrow$  Pathogen for immunodeficient people
  - Formation of blood clots
  - Blood vessel destruction Mortality rates in case of
    - HIV/Aids: 100%
    - Leukemia: 1%
    - Diabetes: 0%
- L. hyalospora
- $\rightarrow \, \mathsf{Non-pathogen}$
- $\rightarrow$  Common fermenting agent in Asian food industry

DNA Sequencing RNA Sequencing

- L. corymbifera, L. ramosa
  - Illumina HiSeq, assembled by Velvet
  - 454, assembled by Newbler
    - $\rightarrow$  Merged by Minimus2
    - $\rightarrow$  Finalized with <code>GapCloser</code> and <code>SEQeul</code>
- L. hyalospora assembly downloaded from JGI



DNA Sequencing RNA Sequencing

```
Ion Torrent sequencing of 4 conditions with 2 biological replicates Coverage: 60 \rightarrow 15-22 mio. reads per sample
```

- WT
- Endoplasmic reticulum stress by dithiothreitol (DTT)
- Protein maturation stress by guanine deaminase (GDA)
- Heat stress

・ロト ・同ト ・ヨト ・ヨト

- Trimming with sliding window 3'  $\rightarrow$  5', l: 10, mean quality: 20
  - Sickle
  - PrinSeq
  - Trimmomatic
  - QTrim
  - FastX Clipper (no sliding window)
- Mapping
  - Segemehl
  - Tophat 2
- Gene prediction
  - Gorap for ncRNAs
  - snoStrip for snoRNAs
  - AUGUSTUS for proteins
  - Cufflinks for de novo transcipts
  - Cuffmerge

A D



- Blast to generate interspecies gene map
- InterProscan for functional annotation
- DESeq for differential gene expression analysis
  - WT vs. Stress
  - Species vs. Species
  - Pathogen vs. Non-Pathogen
- PANTHER GO term enrichment

A (10) < (10) </p>

Trimming and Mapping Gene prediction





RAW Read count: 21.360.689 Read length: 8-363

Sickle Read count: 16.418.297 (76,8%) Read length: 20-270



FastX Read count: 7.207.081 (33,7%) Read length: 8-251

PrinSeq Read count: 20.412.499 (95,6%) Read length: 20-363

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

Trimming and Mapping Gene prediction



45-49 70-79 100-109 140-149 100-189 220-229

202-260







PrinSeq Read count: 20.412.499 (95,6%) Read length: 20-363

QTrim Read count: 19.916.956 (93.2%) Read length: 20-362

15-19 25-29 35-39 45-49 00-69 00-69 100-149 250-299 Epution in read (bp)

Trimming and Mapping Gene prediction



Mapping Statistics Sample 2 GDA

High sensitivity of Segemehl due to handling of

- Unreliable prediction of Homopolymers during Ion Torrent sequencing
- Multiple splice junctions in long reads

< A >

- ∢ ≣ ▶

-

Trimming and Mapping Gene prediction

	L. corymbifera	L. ramosa	L. hyalospora	
Denovo transcripts	731	1213	1465	
Proteins	12379	11510	12062	
tRNAs	174	171	130	
5S rRNA	1	0	0	
5.8S rRNA	0	1	0	
18S rRNA	1	1	0	
28S rRNA	0	1	0	
U1	2	2	2	
U2	2	3	3	
U4	1 (Reads)	1	1	
U5	2	3	3	
U6	4	5	3	
U11	1	1	1	
U12	1	1	0	
U4atac	1	1	1	
U6atac	1	1	1	
RNase P	1	1	1	
RNase MRP	1	1	1	
TPP	1	1	1	
SRP	1	1	1	
U3	3	3	4	
snoRNA	22+(9)+3	24+(8)+4	□ 22+ <b>(</b> 9)+4 ≡	

Konstantin Riege

Lichtheimia - Basal fungal gene expression analysis completed by

GO and miRNA like gene annotation miRNA biogenesis Thank you

- Functional analysis and GO term enrichment
- Prediction of milRNAs, due to presence of related proteins

Dicer	Hairpin cleavage		
ADAR	Destabilizing dsRNA		
TUT4	sRNA degradation by uridylation		
Importin 8	Nuclear import		
KIF17B	binds and guides PIWI		
Gemin3	snRNP assembly		
Armitage	RISC assembly		
Argonaut	RISC assembly		
elF4E	Cap binding and translation inhibiting		
FIERY1	Hairpin loop degradation		
HEN1	sRNA degradation inhibition		
TNRC6B	repression of miRNA translation		

(日) (同) (三) (三)

э

GO and miRNA like gene annotation miRNA biogenesis Thank you





Konstantin Riege Lichtheimia - Basal fungal gene expression analysis completed by

GO and miRNA like gene annotation miRNA biogenesis Thank you

Showed in silico and experimentally for Penicillium marneffei

- sRNA sequencing and mapping
- 2 Loci with 17 to 30 nt perfect matches
- Ilongate to precursor RNA
- Check for hairpin structure and hybridization temperatures  $\rightarrow$  22 milRNAs
- Target prediction and qPCR
  - $\bullet\,$  Dicer, Argonaute knockdown  $\rightarrow\,$  milRNA expression decreases
  - $\bullet\,$  milRNA knockdown  $\rightarrow$  target protein expression increases

・ロト ・同ト ・ヨト ・ヨト

GO and miRNA like gene annotation miRNA biogenesis Thank you







### Shrestha et.al.

 $\ensuremath{\mathsf{QTrim}}$  : a novel tool for the quality trimming of sequence reads generated using the Roche/454 sequencing platform

BMC Bioinformatics, 15:33, January, 2014.



#### Lau et.al.

Identification of microRNA-like RNAs in mycelial and yeast phases of the thermal dimorphic fungus Penicillium marneffei

PLoS Negl Trop Dis, 7(8), August, 2013.

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

GO and miRNA like gene annotation miRNA biogenesis Thank you



< 🗗 ▶

< 注 → < 注

æ

GO and miRNA like gene annotation miRNA biogenesis Thank you







< 🗗 >

< ≣⇒

3

æ

GO and miRNA like gene annotation miRNA biogenesis Thank you



< 17 >

2

Э

æ