26th Fungal Genetics Conference (March 15-20, 2011)

hskwan@eservices.cuhk.edu.hk

Genome sequence and genetic linkage analysis of Shiitake mushroom *Lentinula edodes*

H.S. Kwan, C.H. Au, M.C. Wong, J. Qin, I.S.W. Kwok, W.W.Y. Chum, P.Y. Yip, K.S. Wong, L. Li, Q.L. Huang, W.Y. Nong

The Chinese University of Hong Kong, Hong Kong SAR, PRC

ABSTRACT

Lentinula edodes (Shiitake/Xianggu) is an important cultivated mushroom. Understanding the genomics and functional genomics of *L. edodes* allows us to improve its cultivation and quality. Genome sequence is a key to develop molecular genetic markers for breeding and genetic manipulation. We sequenced the genome of ***L. edodes* monokaryon L54A** using **Roche 454 and ABI SOLiD** genome sequencing. Sequencing reads of about 1400Mb were *de novo* assembled into a **40.2 Mb** genome sequence. We compiled the genome sequence into a searchable database with which we have been **annotating the genes** and **analyzing the metabolic pathways**. In addition, we have been using many molecular techniques to analyze genes differentially expressed during development. Gene ortholog groups of *L. edodes* genome sequence compared across genomes of several fungi including mushrooms identified gene families unique to mushroom-forming fungi. We used a mapping population of haploid basidiospores of dikaryon L54 for genetic linkage analysis. High-quality variations such as single nucleotide polymorphisms, insertions, and deletions of the mapping population formed **a high-density genetic linkage map**. We compared the linkage map to the *L. edodes* L54A genome sequence and located selected quantitative trait loci. The Shiitake community will benefit from these resources for genetic studies and breeding.

INTRODUCTION

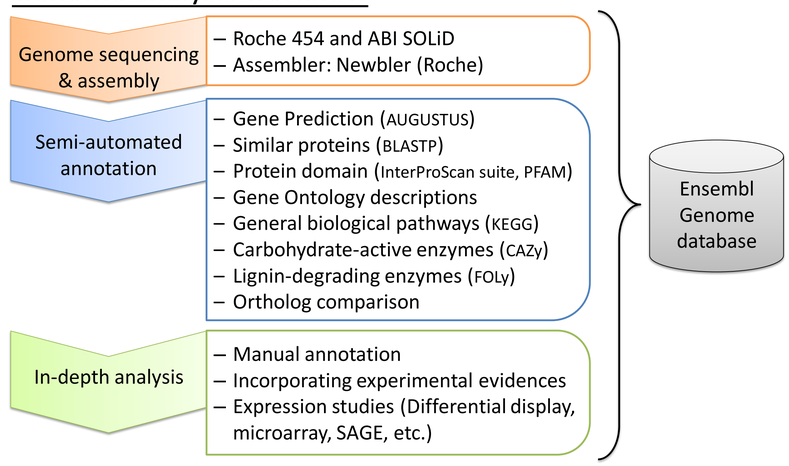
Lentinula edodes

- One of the most cultivated mushroom
- White rot basidiomycetes

Objective:

- To construct *L. edodes* reference genome sequence
- To annotate genes and analyze metabolic pathways
- To build mushroom genome analysis platform
- To elucidate the molecular mechanism of fruiting body development
- To develop markers for genetics studies and assist breeding program

Genome analysis workflow



RESULTS AND DISCUSSION

Reference genome sequence and gene annotation

Table 1. Genome assembly and annotation of *L. edodes* L54A.

Total length of sequencing reads (Mb)	1400
No. of scaffolds	767
Total length of assembly (Mb)	40.2 (92.4% in scaffolds)
N50 sequence length (kb)	110
Mode coverage	11X
No. of protein-coding gene prediction*	13382
With GO term(s)	4885
With Pfam domain(s)	6877
With KEGG orthology	2535

*AUGUSTUS *L. edodes* model (trained with *S. commune* proteome, assisted with *L. edodes* ESTs)

Analysis of protein families

Table 2. Pfam protein families with highest number of predicted proteins of *L. edodes*.

Rank	Pfam ID	No. of predicted proteins *Le*	*Cc*	*Lb*	*Sc*	Family	Description
1	PF00400	479	575	778	573	*WD40*	WD domain, G-beta repeat
2	PF00069	187	232	213	201	*Pkinase*	Protein kinase domain
5	PF00665	133	1	4	1	*rve*	Integrase core domain
6	PF00078	118	0	7	3	*RVT_1*	Reverse transcriptase (RNA-dependent DNA polymerase)
8	PF00106	99	60	63	129	*adh_short*	short chain dehydrogenase
9	PF00076	98	96	105	106	*RRM_1*	RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)
11	PF00023	78	215	38	36	*Ank*	Ankyrin repeat
12	PF03732	73	0	1	4	*Retrotrans_gag*	Retrotransposon gag protein
13	PF04082	70	45	50	61	*Fungal_trans*	Fungal specific transcription factor domain
14	PF00005	64	71	68	69	*ABC_tran*	ABC transporter
16	PF07727	58	0	1	1	*RVT_2*	Reverse transcriptase (RNA-dependent DNA polymerase)
17	PF00083	55	30	28	49	*Sugar_tr*	Sugar (and other) transporter
18	PF00004	54	54	46	50	*AAA*	ATPase family associated with various cellular activities (AAA)
19	PF00107	48	29	26	42	*ADH_zinc_N*	Zinc-binding dehydrogenase
20	PF00385	47	9	10	9	*Chromo*	'chromo' (CHRromatin Organisation MOdifier) domain

Le, Lentinula edodes; Cc, Coprinopsis cinerea; Lb, Laccaria bicolor; Sc, Schizophyllum commune.

Analysis of biological pathways

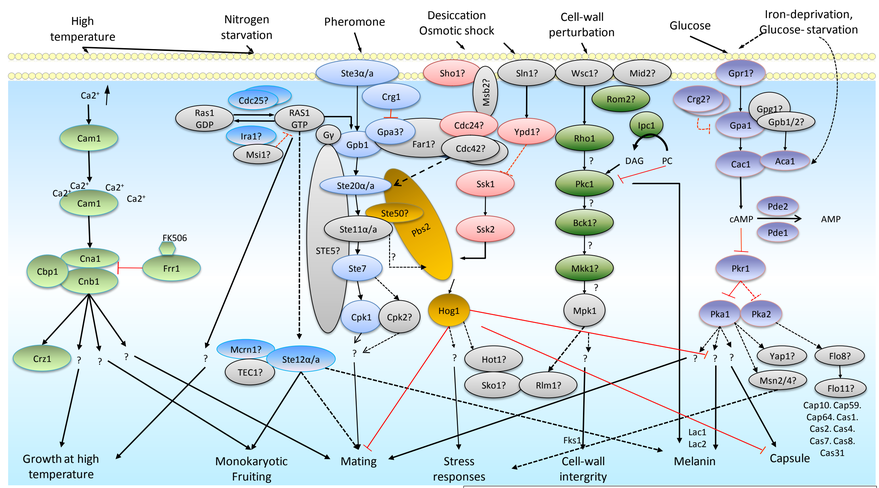


Fig 1. Component of signal transduction pathways of *L. edodes* with reference to *S. cerevisiae*.

Identifying lignocellulolytic enzymes

Table 3. Comparison of the number of candidate CAZymes and FOLymes among basidiomycetes.

	Carbohydrate-Active EnZymes database (CAZy)			Fungal Oxidative Lignin enzymes database (FOLy)		
Family	CBM1, GH5, GH6, GH7, GH9, GH61	GH10, GH11, GH43, GH51, GH74	PL1, PL3, PL4, PL9, GH28, CE1	Lignin oxidase		Lignin degrading auxiliary enzyme
Enzyme	Cellulase	Hemicellulase	Pectinase	Laccase	Peroxidase	Glyoxal oxidase
Le	34	11	24	13	8	5
Pc	78	17	6	0	16	1
Cc	114	17	8	17	1	0
Activities	Plant cell wall polysaccharides degradation			Lignin degradation		H_2O_2 generation

Le, Lentinula edodes; Cc, Coprinopsis cinerea; Pc, Phanerochaete chrysosporium

Mushroom comparative genomics platform

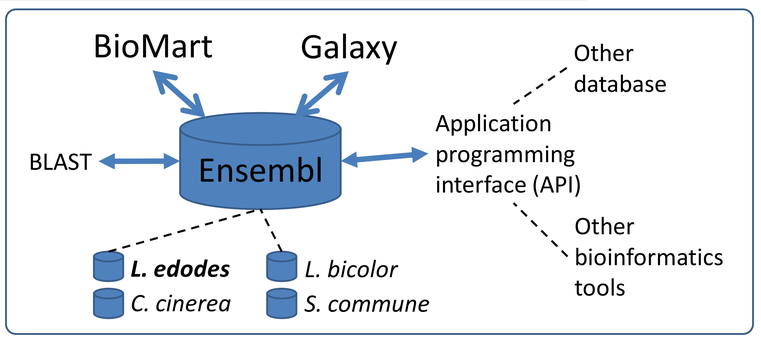


Fig 4. The architecture of mushroom genome database and analysis tool.

Functions:

- General genome browser of visualization of genome sequence and detailed gene annotation
- Multiple-species comparison
  - Gene family,
  - Genome sequence alignment
- Multiple-strain comparison
  - Genetic variations, eg. SNPs, indels

Construction of a high-density genetic linkage map

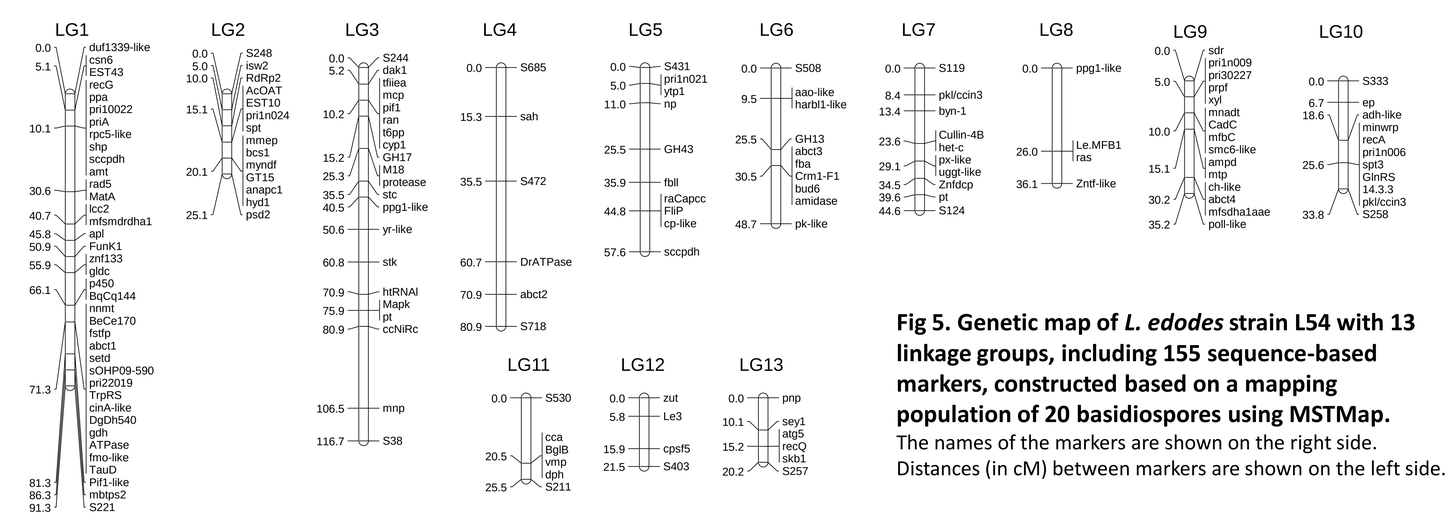


Fig 5. Genetic map of *L. edodes* strain L54 with 13 linkage groups, including 155 sequence-based markers, constructed based on a mapping population of 20 basidiospores using MSTMap.
The names of the markers are shown on the right side. Distances (in cM) between markers are shown on the left side.

CONCLUSION

Our works provided

- sequence information for revealing genes of important biological process
- a better understanding of the molecular basis of the growth and development of mushrooms
- valuable resources for large amount sequence-based markers
- genetic map allowing mapping of Quantitative Trait Loci (QTLs)
- resources for genetic studies and breeding of Shiitake mushroom

REFERENCES

- Chum et al. 2011. *Fungal Genet Biol* 48: 359-369.
- Goffeau et al. 1996. *Science* 274: 546-567.
- Kämper et al. 2006. *Nature* 444: 97-101.
- Loftus et al. 2005. *Science* 307: 1321-1324.
- Martin et al. 2008. *Nature* 452: 88-92.
- Martinez et al. 2004. *Nat Biotech* 22: 695-700.
- Martinez et al. 2008. *Nat Biotech* 26: 553-560.
- Nierman et al. 2005. *Nature* 438: 1151-1156.
- Ohm et al. 2010. *Nat Biotech* 28: 957-963.
- Stajich et al. 2010. *PNAS* 107: 11889-11894.

ACKNOWLEDGEMENT

This project was supported by the RGC General Research Fund CUHK467810 from the Research Grants Council of the HKSAR, PR China.

Nature Precedings - doi:10.1038/npre.2012.6855.1 : Posted 31 Jan 2012