

WebLOG update:

Agaricus bisporus 8x Release [project ID 4084801 / August 2009]

JGI have issued a revision of the Agaricus bisporus 8X genome assembly for Agaricus bisporus (H97, lead homokaryon. Jeremy Schmutz of the JGI has provided an assembly report/update <PDF>.

This is update to the initial draft release (2009 03) and incorporates two additional breaks based on linkage mapping data provided by Anton Sonnenberg. Many thanks to Anton and Jeremy Schmutz for their work on this.

This draft genome will now go into genome improvement, additional joins may yet be made to complete the chromosomes.

**[asm_release] Agaricus bisporus 8x v1.1 Release [project ID 4084801]
JGI update -- 2009 08 07**

Note:

This is update to the intial draft release that makes two additional breaks based on the map integration by Anton Sonnenberg.

As this genome will be going into genome improvement additional joins may be made to complete the chromosomes.

Assembly Stats:

Main genome scaffold total: 29
 Main genome contig total: 254
 Main genome scaffold sequence total: 30.2 MB
 Main genome contig sequence total: 30.0 MB (-> 0.7% gap)
 Main genome scaffold N/L50: 6/2.3 MB
 Main genome contig N/L50: 35/262.5 KB
 Number of scaffolds > 50 KB: 19
 % main genome in scaffolds > 50 KB: 99.7%

Minimum Scaffold Length	Number of Scaffolds	Number of Contigs	Total Scaffold Length	Total Contig Length	Scaffold Contig Coverage
All	29	254	30,233,745	30,028,778	99.32%
1 kb	29	254	30,233,745	30,028,778	99.32%
2.5 kb	29	254	30,233,745	30,028,778	99.32%
5 kb	24	247	30,214,694	30,010,776	99.33%
10 kb	21	243	30,195,633	29,991,890	99.33%
25 kb	19	240	30,156,271	29,952,770	99.33%
50 kb	19	240	30,156,271	29,952,770	99.33%
100 kb	19	240	30,156,271	29,952,770	99.33%
250 kb	17	234	29,709,594	29,509,530	99.33%
500 kb	17	234	29,709,594	29,509,530	99.33%
1 mb	13	208	27,060,522	26,878,775	99.33%
2.5 mb	3	58	9,008,119	8,986,541	99.76%
5 mb	0	0	0	0	0.00%

Coverage Stats:

LIB	COV.	INSERT	STDDEV
FWBC	3.84x	2904	+/- 405
FWBB	3.92x	6367	+/- 572
FWBA	0.74x	39552	+/- 2828
Total	8.50x		

Completeness Stats:

To evaluate the completeness of the genomic sequence we placed the clustered ESTs. We used all clusters > 400bp.

Of the 12,621 clusters 12,050 (97.54%) are well placed (at 90% ID, 85% Coverage) in the genome.

267 clusters are identified as artifacts.

78 (0.63%) are found at > 50% coverage and 226 (1.83%) cluster sequences are not found.

The 226 not found clusters are mostly not found in the reads either and appear to probably be poorly sequenced clusters built from 454 reads that disagree with the genome.

It doesn't appear that these indicate a problem in the genome sequence.

The scaffolds are divided into the following sets:

- 1) Main genome
- 2) Mitochondrion
- 3) Repeat
- 4) Excluded scaffolds

Production Stats:

LIB	INSERT	READS	FAILED(%)	VECTOR(%)
FWBC	2904	191136	5634(2.9)	7242(3.8)
FWBB	6367	179424	6821(3.8)	2035(1.1)
FWBA	39552	36192	2416(6.7)	167(0.5)
Total	N/A	406752	14871(3.7)	9444(2.3)

LIB	UNPAIRED(%)	PAIRED(%)	GOODP20	STDDEV
FWBC	3846(2.0)	174414(91.3)	772.60	+/-99.35
FWBB	4060(2.3)	166508(92.8)	772.51	+/-102.36
FWBA	1535(4.2)	32074(88.6)	724.65	+/-106.89
Total	9441(2.3)	372996(91.7)	768.42	+/-102.24