

Supplementary material

Mating-type locus organization and mating-type chromosomes in the bipolar edible button mushroom *Agaricus bisporus*

Marie Foulongne-Oriol¹, Ozgur Taskent², Ursula Kües³, Anton S. M. Sonnenberg⁴, Arend van Peer⁴ and Tatiana Giraud²

¹ INRAE, MycSA, Mycologie et Sécurité des Aliments, 33882 Villenave d'Ornon, France

² Ecologie Systématique Evolution, Bâtiment 360, CNRS, AgroParisTech, Université Paris-Saclay, 91400 Orsay, France

³ Molecular Wood Biotechnology and Technical Mycology, Büsgen-Institute and Goettingen Center for Molecular Biosciences (GZMB), University of Goettingen, Büsgenweg 2, 37077, Goettingen, Germany

⁴ Plant Breeding, Wageningen University and Research, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

A

a1-1	MTVSSIRK-----RLIQAEESFLAASNER--FMEDFEQTWNILAVDIVKATSANQMSPEDKSLADIV	60
a1-2	MTISSIRK-----RLIHAEEESFLAAKNDH--LMEDFEYAWTSLAVDIMKATSANQMTQEDIALADAV	60
b1-1	MSTRSSRKTVKQSLCRRFSDAESAFLDALDVGGGAMTAFEFEWHLRDDFEAACIQGDVDPETASYARLV	70
a1-1	SSRITSLCKTFIDLTYPCCDISLSSEQELERILEEHESGNHGSQSRSEETLHPAPALYIKQAYSWLMQN	130
a1-2	SSRITSLCKTFVEFTNACDISLSFEQDLERILEEHDHGNHCSNHYSSEEMPRPAPAPYIKQAYSWLQVN	130
b1-1	TYRIAELISSSFLKEATCKQLTDSLMDLASIPELRQS----QVLDTPQETISSR--RDVSLAAEWLSKN	134
a1-1	LSNPYPKPKEVREAIARKAGSDPKHVENWFGDVRKRMGWNGIRKRHFENKRHLTVAAATAFFGGQSSELLT	200
a1-2	LSNPYPKPKEVREAIARKAGSDPKHVENWFGDARKRMGWNGIRKRHFENKRRLTVAAATAFFGGQSSELLT	200
b1-1	FFNPYPSSSTVDRISHQSNWNRKDVDAWFEARRRIGWNDIRKQYFONKRAGAVERATEFFNG-SPTCFD	203
a1-1	DLVRQDFAFMSADAEDMYSKKFMESKLARHIAGVTGSPRKKPSQSAS-VENETSSSYNHP-----	259
a1-2	DPVRQDFAFMSADAEDMYSKKFMESKLARHIAGVTESPRKKRSQSAS-VEGGSSSSFNYS-----	259
b1-1	AVLSQAFIDMESRVQDLYVEPYKPSKLAVVLG-----KGHSDEAAFMOKGCFSTPEKSPRTLSPSRSL	267
a1-1	-----LFHPSLPHTOCNA-RKRFRSLDSTDEVLDINNFOKRPRFSYSGDSSPAASHSPLPSTSPQSPVP	322
a1-2	-----LFHPSLPHTOCNP-RKRFRSLDLDTDVLDK---KKRRLSYSGDSSPAALHSPLPTTSPQSPVP	319
b1-1	TSSRSASPIDSSVSTCQTNKRKPRS-DKTDQYHSQHP-RKKLR-----EEEEPTTIM	317
	<u>NLS</u> <u>NLS</u>	
a1-1	HAIPATFQMNROTSEFE---DHLPPSHPSADSRSV-SLSQVLELNRTSCNPSPSSSTLPSPDVSDEPITSQ	388
a1-2	HAIPATFQMNCOTSEFE---DHLSPRHPSADSRPA-SLSQVQDLNRTSCNASPSSYTLPDFVSDEPITSQ	385
b1-1	LQAQPTCSSVSQSSLEEVSDAIT-KGSSAVARPARSPRLSPFLDPQRKQLHEEALDLPNNAFGGSRITIR	386
a1-1	YIPNEATKRSDSVSMASFQOSP-TLHSSALSILHNTLPSSLSRKENTQFPTEFQAPVTYSEPVTQGF	457
a1-2	HIPNEDTKRSDLVSMASIROSS-ILHSPALSLPRNILPSSLSPRKENTQFPTEFQMPVPCSESVVTQGF	454
b1-1	SQTTSLGHQSNACTPMAPDLGSINPTLSLNFPSASSTIMPDLQVDFVHDISHAEI--SELSYTFPGP	454
a1-1	LAPSASDITHD-----TIFSGSTANDRFLAISSDFAPDLSSAQDVNATSWPATIANTQT-----	513
a1-2	LAPLASDVTHE-----TIFSGSTVNDRLFLSTSHLDFASRSPSSVQDVNATSWPGTIANTQI-----	510
b1-1	-----LTHDQPTLMQPVDPQWS-NSATS AHLPTYDLSMSLLELASQAMP PNGNTDVVPSSDIFNNCS	516
a1-1	--DGQFFGYLDPLPTEFTTFTAQVDALPDFADVDLTFGNLFDLDASAAFPSPDDEFDIT---ILPANRAENL	578
a1-2	--DGQFFGYLDPLPTEFTTFTAQVDALPDFADVDLSFGNLFDLDAASAAFPSPDDEFDIS---MLPVNRAENL	575
b1-1	FYDDQLFELL-----RPFGRGAPAVPPMQNFQGLNTEVDISLAAPY---LDFESCTSQIPQPIDDI	576
a1-1	ATNDLFTLTP----AVESSVSKFLQLEQLKKHRNRIOEYINTLEIQIATTLAA	627
a1-2	ATNDLLTLTP----AAENSVSKFLQLEQLKKHRNRIOEYINTLEIQIATTLAA	624
b1-1	LADFIHNFSAAPPVDDAEK-LKRVAVHVEMLKREKEHLQERLOSLEYVSLPLVINKRLT	631

B

a2-1	MPAFRISDECSQRNMGVAS-VLSVLRKALCTKTPLPHSATLPSLSNIQLPQAPSIAIVLKTHCVGQSLAE	69
------	--	----

```

b2-2  MSPKIQETQAILSSRVPSFNISRPALQFGGLSNSRRNNWTPPLQFPPTSPAAQLVQLGCAFEESQ 69
b2-1  MPLIDETP---ASIPHQRISDALSK 23

a2-1  ELTRTYEQHVSDLKCTEAPIQGLLQHYRDN-----VSSDIIEKQIARYISLYESRIKHWNLL 129
b2-2  EAANIFASRVKDLQSTYERHYEKILFRRLSKYSSNTELRNLAKHINRVLDARSRNITSWQEAILQAASIR 139
b2-1  SLSQSYQSKVNHVEVNFQKPRQVQYRLRQAPGESEQ--TPQATTVISGAYMIQLKTWESHTLKKIARR 91

a2-1  SNEIKREKS-----EENTERIPILMTYFKHNAYPSIDDRNWLARETENTPQOIGLWFONRRRRSRKAGV 193
b2-2  VQERLKQASSDERPVEKPEWTPFLKKYFAQNAYPTHQDKVILATKTGMTVROIYDWFQNHRTREKRKGI 209
b2-1  LRK-NRCIEPTCKPSEKYEVVPLLNSYFKRNAYPSTSDRASLAKKTNMSROIEVWFQNHNRRAHKNQ 160
-----

a2-1  ELKRAGPDEISPNLFADDEEPYSPOPGTRPRVSSPEVT-TVEEVPLSCSSVDATLRSHOHLRTHDFTN 262
b2-2  PLKRPDIEGSPEFSSEDSPTSRVTTPEDDDMLPKKRR-VATDEPLALSPVESTSTQDFDFFCLPPDAVW 278
b2-1  ILNRVKDQQLPSDALPDEPPERSSTPPKNSHSSRKCLSHSPSSSIQAPASPLSPTQCQSPPSIFSATCV 230

a2-1  FELPP-YAFPTYNHEATPSCWSRPPETPFPEPNWDRK-----PSSHRKNRHARSTATSVDD 318
b2-2  PRSYSEKDMKAFWECFVEPNLGSFTWPRELQPDSRSK-----RKALDSNIDEQIAESFASK 335
b2-1  QTISSTSKSSPLPEPLEYAFPPAFPRACDDDTFAIKGKDPWFPFAPQWIRLPAQTINVRKRKRRTISIDE 300

a2-1  LSLMLAVKLRIFCNHFSSSAVADTPWHAT-LYHTSPSPLPPALVRSKQFVIITIGPTSQSIRSSLKRKRS 387
b2-2  LHFRSSLCKRTSLECEETSDIFEPWSYHRLYPKRHPASVYNQLTEPSKCAQPARPTELFPKYRRTLNS 405
b2-1  FAVEFAQHLSCRHLCSEPSASKKKRSSWRDATHVYPCRGHPADLGQYQQKASLESSHHSPRHHSRPS 370

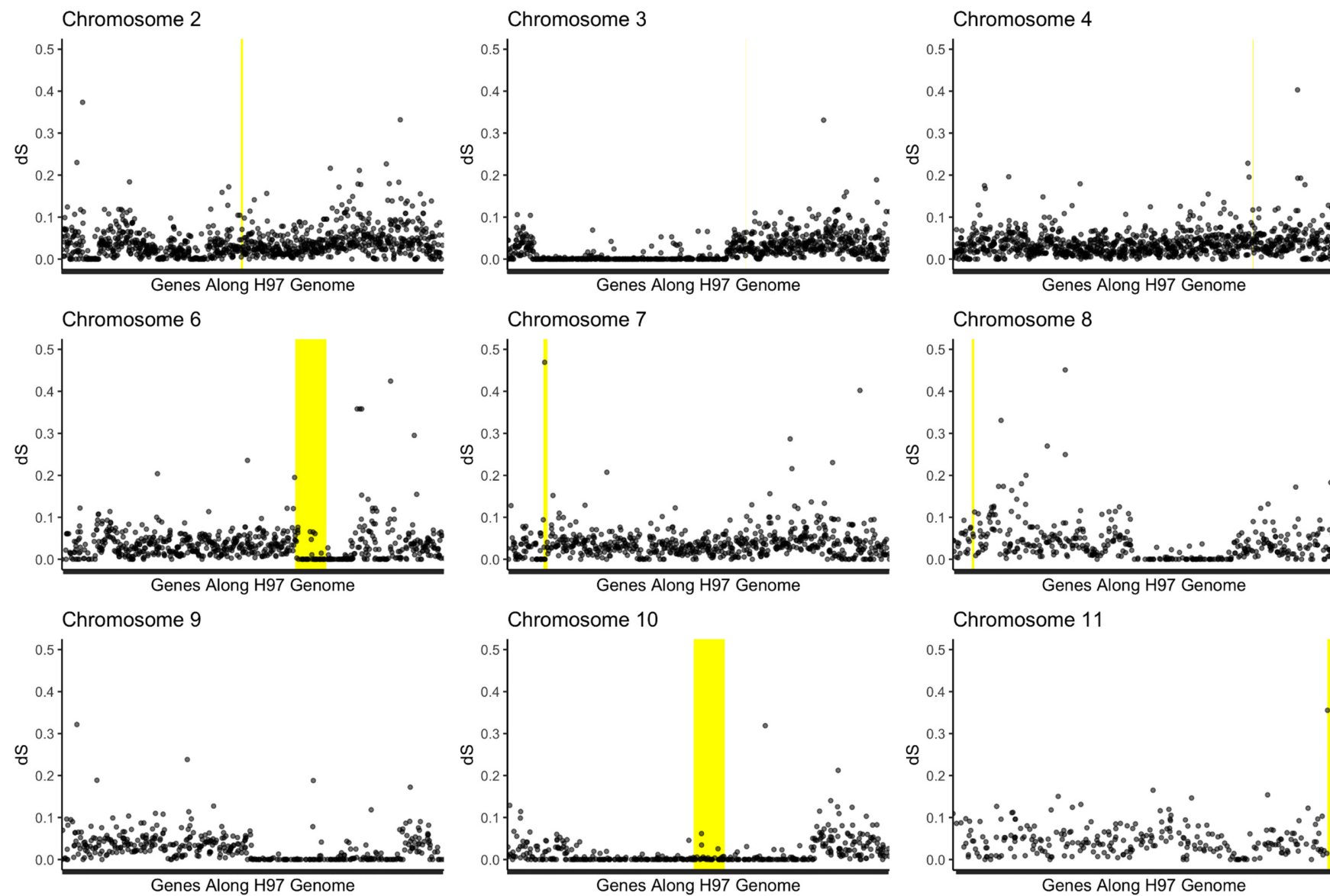
a2-1  EINEEQCTKRFRFSASSCSSFIGGDFYSSDDSSCTSSSPGPQTTPPSAHLSLPTPPRFVDWSKFSIPTV 457
b2-2  SPSSRRRASSSSSHSTCSISSSTSPSSSIASSPPSMIRCSWSSLDTFTPGSSPRRPLQSLGISGLKGF 475
b2-1  ISSRRSSHTCMPSGSESSVSTVSSPSSFDSTPLTSLDASPLLLPQNDKAYASESSLSVQLQSLMLPGF 440

a2-1  PQAYGSC 464
b2-2  EVYAGPFVGLNMSISS 491
b2-1  LQYAGSSVKFGTAVS 455

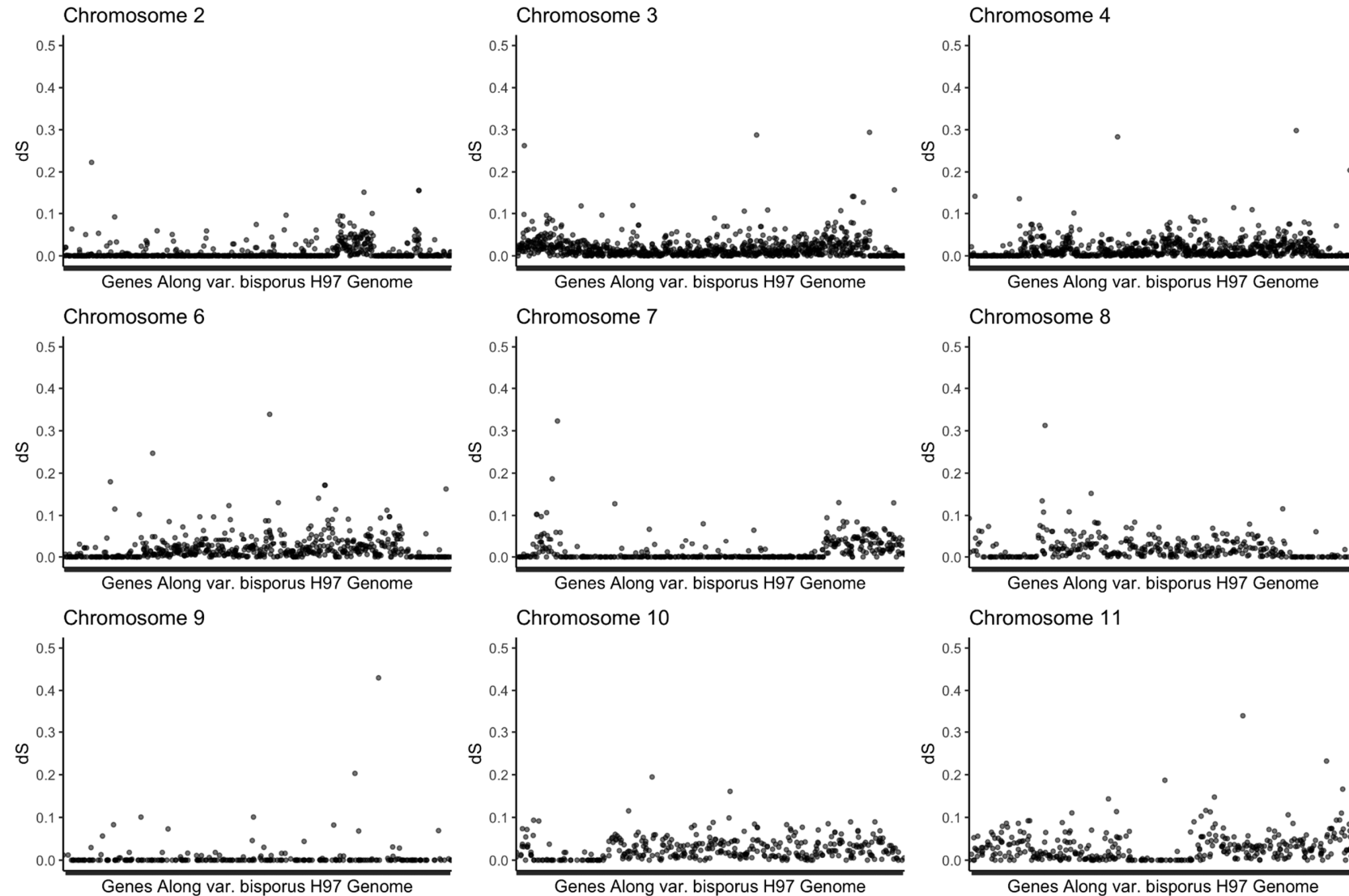
```

Supplementary Figure S1 Alignment of **A. HD1** and **B. HD2** proteins from *Agaricus bisporus* var. *bisporus* H97 and *A. bisporus* var. *burnettii* JB137-S8. The positions of the homeodomains are marked by a dashed line. NLS marks motifs of alike structure of the bipartite nuclear localisation signals in *Coprinopsis cinerea* (Spit *et al.*, 1998) and *Laccaria bicolor* (Niculita-Hirzel *et al.*, 2008).

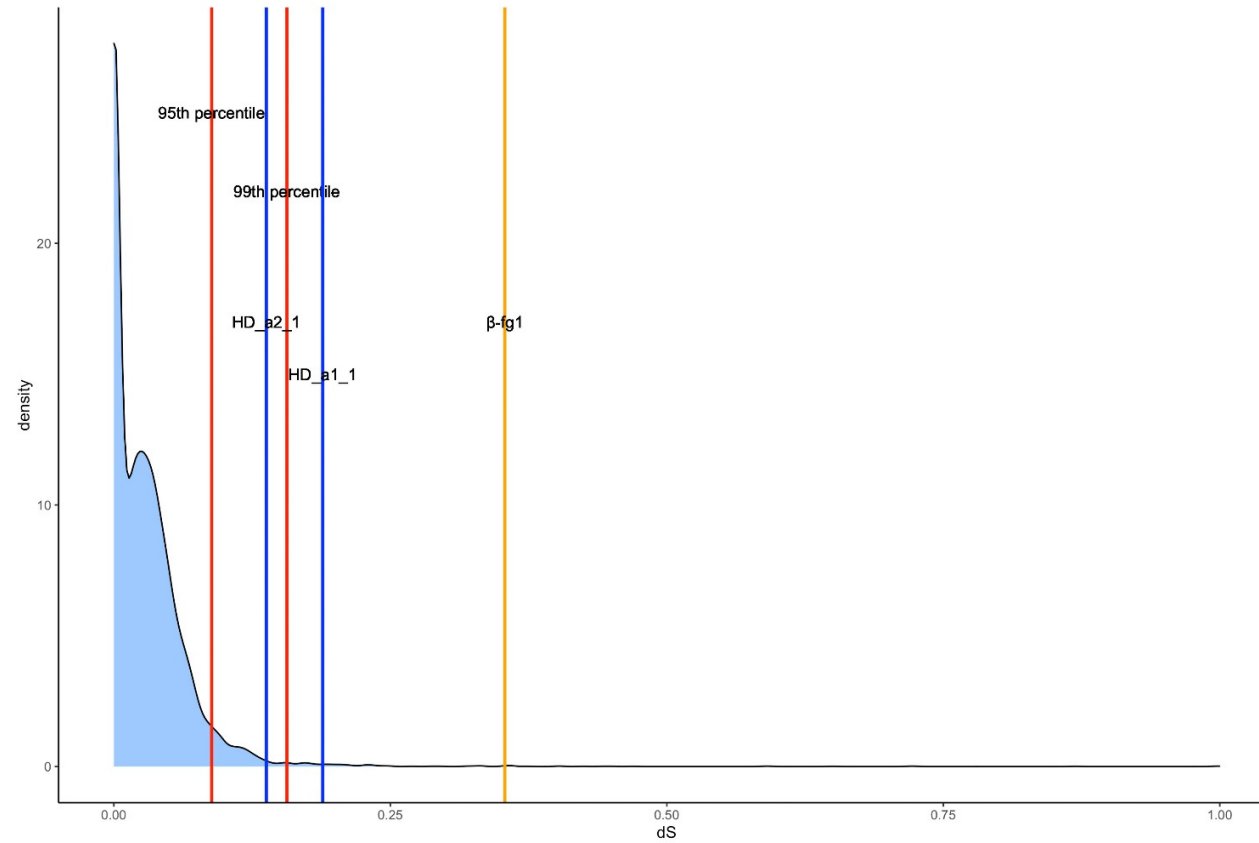
A



B



Supplementary Figure S2. Synonymous substitution rates between the homokaryotic genomes of *Agaricus bisporus* var. *bisporus* H97 and H39 (A) and between *A. bisporus* var. *burnettii* H119p1 and H119p4 (B). Y-axis shows the synonymous divergence levels between the alleles found in the homokaryotic genomes of H97 and H39 of alternative mating types, issued from the U1 cultivar. X-axis shows the gene order along the H97 genome. Yellow highlights on the background shows the putative centromeric regions. The other chromosomes are shown in Figure 8.



Supplementary Figure S3. Distribution of the rate of synonymous substitutions (dS) between orthologous gene pairs found in the constitutive homokaryon genomes of H97 and H39 of *Agaricus bisporus* var. *bisporus* Hurst U1. Density plot shows the relative frequencies of genes with different dS values. The 95th and 99th percentile values of the distribution are highlighted with vertical red lines. The dS values for the two *HD* genes and *β -fg1* gene are shown with blue and orange vertical lines, respectively.

Supplementary Table S1 : Gene comparison between *Agaricus bisporus* var *bisporus* H97 v3.0 and *Agaricus bisporus* var *burnettii* JB137-S8: genomic position in the two strains, DNA identity and protein identity

<i>Agaricus bisporus</i> var <i>bisporus</i> H97 v3.0		<i>Agaricus bisporus</i> var <i>burnettii</i> JB137-S8		DNA identity (%)		Protein identity (%)
Gene ID/name	Position (start to stop codon)	Gene ID/name	Position (start to stop codon)	Genomic DNA	Coding sequence	
239372/ <i>mip1</i>	scaffold_1:897420-899760 (+)	67360/ <i>mip1</i>	scaffold_1:2040314-2042755 (-)	98	98	99
113833/ <i>a1-1</i>	scaffold_1:895062-897100 (+)	124448/ <i>a1-2</i>	scaffold_1:2043142-2045173 (-)	88	89	82
239365/ <i>a2-1</i>	scaffold_1:893233-894863 (-)	144391/remnant of <i>a2-2</i>	scaffold_1:2045374-2045559 (+)	77	-*	-
-	-	144392/footprint of a <i>mip1</i> gene	scaffold_1:2045600-2045806 (-)	(77 to <i>mip1</i>)	-*	-
-	-	144356/ <i>b1-1</i>	scaffold_1:2080256-2082318 (-)	-	-	-
239364/ <i>b2-2</i>	scaffold_1:890959-892499 (-)	124456/ <i>b2-1</i>	scaffold_1:2082520-2084040 (+)	-	-	26
1181595/ β - <i>fg2</i>	scaffold_1:887425-889480 (+)	144375/ β - <i>fg2</i>	scaffold_1:2084717-2086770 (-)	98	99	98
1137415/ β - <i>fg1</i>	scaffold_1:883864-884895 (+)	81839/ β - <i>fg1</i>	scaffold_1:2087937-2088968 (-)	94	96	98
239575/ <i>ph11</i>	scaffold_1:2918372-2918506 (+)	144336/ <i>ph11</i>	scaffold_1:59978-60141 (-)	99	99	98
239576/ <i>ph12</i>	scaffold_1:3231293-3231622 (+)	144361/ <i>ph12</i>	scaffold_14:29026-29157 (+)	100	100	100
239578/ <i>ph13</i>	scaffold_1:3236723-3236857 (-)	144359/ <i>ph13</i>	scaffold_14:34681-34815 (-)	99	99	98
239573/ <i>ph14</i>	scaffold_5:200825-200999 (-)	86158/ <i>ph14</i>	scaffold_9:175788-175616 (-)	97	97	96
239579/ <i>ste3.1</i>	scaffold_5:2047209-2048830 (-)	121275/ <i>ste3.1</i>	scaffold_7:898975-900596 (-)	98	99	97
239590/ <i>ste3.2</i>	scaffold_13:276835-278479 (+)	144367/ <i>ste3.2</i>	scaffold_17:267571-269202 (+)	98	98	98
239584/ <i>ste3.3</i>	scaffold_13:392785-394687 (+)	144363/ <i>ste3.3</i>	scaffold_17:379554-381458 (+)	98	98	98
195352/ <i>ste3.4</i>	scaffold_12:1236632-1238485 (-)	123419/ <i>ste3.4</i>	scaffold_38:95442-97029 (+)	96	97	98

*contains frameshift(s) and stop codon(s)

Supplementary Table S2: Microsynteny analysis performed by tblastn searches on JGI page <http://genome.jgi-psf.org/Copci1/Copci1.home.html> (settings: default) with genes of the regions with *HD* mating type genes on the chromosomes 1 of *Agaricus bisporus* var. *bisporus* H97 (scaffold 1) compared to the genome of *Coprinopsis cinerea* Okayama 7 (no hit: -)

<i>Agaricus bisporus</i> var. <i>bisporus</i> H97		<i>Coprinopsis cinerea</i> Okayama 7	
Gene	Position on chromosome I	Gene	Position on chromosome I
189405 (<i>pab1</i>)	932956-935538 (-)	CC1G_01849T0 (<i>pab1</i>)	2699003-2701362 (-)
1165947	931452-932725 (-)	-	-
1120398 (<i>up1</i>)	930119-931431 (+)	CC1G_01848T0	2697506-2698930 (+)
1165946 (<i>ubc12</i>)	928622-930005 (-)	CC1G_01847T0	2696166-2697285 (-)
1072572	928377-928937 (+)	CC1G_13707T0	2695644-2696187 (+)
1153365 (<i>up3</i>)	927132-928291 (-)	CC1G_01846T0	2694335-2695507 (-)
1095010 (<i>up4</i>)	924999-927074 (-)	CC1G_01845T0	2692021-2694125 (-)
1072437 (<i>up5</i>)	922378-924450 (-)	CC1G_13706T0	2689754-2691849 (-)
1181608 (<i>up6</i>)	919573-922217 (-)	CC1G_13705T0	2686968-2688467 (-)
1145681 (<i>glydh</i>)	915514-919100 (-)	CC1G_01843T0	2683012-2686558 (-)
1153361 (<i>up7</i>)	912956-915306 (+)	CC1G_01842T0	2680968-2682946 (+)
1031411 (<i>sec61</i>)	912405-912835 (-)	CC1G_01841T0	2680073-2680505 (-)
1095004	910403-912306 (+)	CC1G_01840T0	2678200-2679927 (+)
1103173 (<i>up8</i>)	906900-909825 (+)	CC1G_01839T0	2674086-2677215 (+)
1165935 (<i>up9</i>)	905234-906831 (+)	CC1G_01838T0	2672530-2673857 (+)

1137419 (<i>up10</i>)	904507-905196 (-)	CC1G_01837T0	2671654-2672349 (-)
1165933	902316-904458 (-)	-	-
1072561 (<i>up11</i>)	900667-902087 (-)	CC1G_01836T0	2669318-2671452 (-)
239372 (<i>mip1</i>)	897420-899938 (-)	CC1G_01835T0 (<i>mip1</i>)	2666417-2668995 (-)
113833 (<i>a1-1</i>)	895062-897100 (+)	CC1G_01834T0 (<i>a1-2</i>)	2666140-2664189 (+)
239365 (<i>a2-1</i>)	893233-894463 (-)	CC1G_01833T0 (<i>a2-2</i>)	2663799-2662251 (-)
-	-	CC1G_01832T0	2660867-2660048 (+)
-	-	CC1G_01831T0 (<i>b1-2</i>)	2655981-2653848 (+)
239364 (<i>b2-2</i>)	890959-892499 (-)	Modified CC1G_13704T0 (<i>b2-2</i>)	2653400-2651473 (-)
-	-	CC1G_01828T0 (<i>d1-1</i>)	2653400-2651473 (-)
1181595 (<i>β-fg2</i>)	887425-889480 (+)	-	-
1137415 (<i>β-fg1</i>)	883864-884895 (+)	CC1G_01827T0 (<i>β-fg</i>)	2646406-2647545 (+)
1181614	882720-883653 (+)	CC1G_01826T0	2645288-2646234 (+)
1094996 (<i>glgen</i>)	879120-882345 (+)	CC1G_01825T0	2641261-2644545 (+)
1094995	876725-878971 (-)	CC1G_01824T0	2638742-2641057 (-)
1103148 (<i>blac2</i>)	874323-876620 (+)	-	-
1160308 (<i>mad2</i>)	873029-874228 (-)	CC1G_01823T0	2637048-2637945 (-)
1165929 (<i>snx4</i>)	871257-873344 (+)	CC1G_01822T0	2634949-2636783 (+)

1031396 (<i>up12</i>)	868800-871138 (-)	CC1G_01821T0	2632303-2634830 (-)
1031395 (<i>rpb2</i>)	864573-868430 (+)	CC1G_13703T0 (two genes fused to one model)	2626696-2631959 (+)
1120344 (<i>up13</i>)	862870-864414 (+)		
1165927 (<i>fmnor1</i>)	860988-862678 (-)	CC1G_01818T0	2624916-2626437 (-)
1181620 (<i>qdr1</i>)	858946-860793 (+)	-	-
1103133 (<i>ndse</i>)	857528-858720 (+)	CC1G_01817T0	2623644-2624843 (+)
1153345 (<i>upa4</i>)	855520-857395 (+)	CC1G_01816T0	2621439-2623236 (+)
1024634	853801-855409 (-)	CC1G_01815T0	2619103-2621082 (-)
213701	850192-853206 (+)	CC1G_01814T0	2615859-2618928 (+)
1165924	849031-850133 (+)	CC1G_01813T0	2614502-2615625 (+)
1153341	845123-848866 (-)	CC1G_01812T0	2610634-2614222 (-)
1094983	840285-843205 (+)	CC1G_01811T0	607625-2610066 (+)
1024630	838934-839727 (+)	CC1G_01810T0	2606242-2607178 (+)
1120321 (<i>cwc25</i>)	837108-838693 (-)	CC1G_01809T0	2604610-2606051 (-)
1137404(<i>lim15/dmc1</i>)	836248-837091 (-)	CC1G_01808T0	2602748-2604371 (-)
1165918 (<i>RNAb30</i>)	833488-834784 (-)	-	-
1165917 (<i>hst3</i>)	830976-833064 (-)	-	-
61363 (<i>rp6</i>)	829899-830250 (+)	CC1G_01807T0	2601740-2602693 (+)

1094976	824356-828040 (-)	CC1G_01806T0	2597092-2600714 (-)
-	-	CC1G_01805T0	2595167-2596736 (+)
1153333	821738-823496 (+)	CC1G_01804T0	2592521-2594140 (+)
1072421	820710-821631 (+)	CC1G_13702T0	2591415-2592320 (+)
1120304	819167-820691 (+)	CC1G_01803T0	2589970-2591279 (+)
1145648	817392-818862 (-)	CC1G_01802T0	2588157-2589877 (-)
1094972	815356-817579 (+)	CC1G_01801T0	2586076-2588027 (+)
1044354	813756-814924 (+)	CC1G_01800T0	2584273-2585762 (+)
1120294	812245-813487 (-)	CC1G_01799T0	2582928-2584111 (-)
-	-	CC1G_01798T0	2581293-2582126 (+)
1072934	810638-812175 (+)	CC1G_01797T0	2578317-2579859 (+)
1165910	809176-810262 (-)	CC1G_01796T0	2576916-2577993 (-)
1103082	807160-808228 (-)	CC1G_01795T0	2575851-2576750 (-)
1165908	802076-804569 (-)	CC1G_01794T0	2572983-2575510 (-)
1153321	797672-802037 (+)	CC1G_13701T0	2568876-2572693 (-)
1153320	795950-797603 (-)	CC1G_01791T0	2567390-2568674 (-)
213648	794661-796492 (+)	CC1G_01790T0	2564978-2566678 (+)
1165904	793411-794430 (-)	CC1G_01789T0	2562478-2563426 (-)

		CC1G_01788T0	2561151-2562077 (-)
1120267	791196-793007 (-)	CC1G_01787T0	2559457-2561006 (-)
1165902	788993-790798 (-)	CC1G_01786T0	2557814-2558980 (-)
1103060	787718-789588 (+)	CC1G_01785T0	2556370-2557695 (+)
1165901 (<i>gpd</i>)	785731-787605 (-)	CC1G_01784T0	2554397-2556162 (-)
147399	783211-785503 (-)	CC1G_01783T0	2551954-2554178 (-)
1103054	780120-782052 (-)	-	-
1103058 (<i>ade8</i>)	776958-779907 (+)	CC1G_01782T0 (<i>ade8</i>)	2548109-2551563 (+)

References

- Niculita-Hirzel H, Labbe J, Kohler A, le Tacon F, Martin F, Sanders IR, Kues U. 2008.** Gene organization of the mating type regions in the ectomycorrhizal fungus *Laccaria bicolor* reveals distinct evolution between the two mating type loci. *New Phytologist* **180**(2): 329-342.
- Spit A, Mellor EJC, Casselton LA. 1998.** A role for heterodimerization in nuclear localization of a homeodomain protein. *Proc. Natl. Acad. Sci, USA* **95**: 6228-6233.