Additional File 5. List of primers and specific PCR conditions used to amplify the indicated regions in red yeast species. Two PCR protocols (**A** and **B**) shown below were used to amplify the different genomic regions under study. For each region, it is indicated which of two protocols was used. Primer annealing sites are shown at the end for each studied region.

A. For PCR screening, reactions were performed in a final volume of 10 μl with the following components: 1X DreamTaq Buffer (Fermentas, Canada), 0.20 mM of each of the four dNTPs (GE Healthcare), 1% DMSO, 1.0 μM of each primer, 100 ng of genomic DNA, and 0.5 U DreamTaq DNA polymerase (Fermentas, Canada). Thermal cycling consisted of a 5-minute denaturation step at 95°C, followed by 35 cycles of denaturation at 95°C for 30 s, 30 s at the annealing temperature (variable), and extension at 72°C (variable time). The annealing temperatures, extension times and primer sequences are given for each case. A final extension of 7 min at 72°C was performed at the end of each reaction. PCRs reactions were adjusted to 50 μl final volume when sequencing was required.

B. For the long range PCR (LR-PCR), reactions were performed in a final volume of 50 μl and contained the following components: 1X Long PCR buffer, 2 mM of $MgCl_2$, 0.2 mM of each of the four dNTPs (GE healthcare), 1.0% DMSO, 1.0 μM of primer, 150 ng of genomic DNA and 2.5 U Long PCR enzyme mix (Fermentas, Canada). Thermal cycling consisted of a 3-minute denaturation step at 94ºC, followed by an initial round of 10 cycles of denaturation at 95ºC for 20 s, annealing for 30 s (variable temperature), extension at 68ºC (variable time), and a second round of 25 cycles increasing the extension time in each cycle (variable). A final extension of 10 minutes at 68ºC was performed. The annealing temperatures, extension times and primer sequences are given for each case.

Additional File 5. Continued.

Amplification of *STE3. A1* allele – Protocol A			
Species	**Primer Name**	**Sequence (5' → 3')**	**Annealing/ Extension**
R. babjevae, *Rh. glutinis*, *R. diobovatum*, *Rh. araucariae*, *R. kratochvilovae*, "*Rh. hamamotoiana"*, *R. sphaerocarpum*, *R. azoricum*, *"R. oreadum"*, *Rh. colostri*,	**MC159** **MC164**	ACGYTGCTSTTCATYTTYTGGTG GAAGATSGCRCASGAGATKGCGTA	60.0ºC / 60 s
"Rh. sesimbrana", *R. paludigenum*, *S. ruineniae*	**MC190** **MC191**	ACCYTGCTYCTCATYTTYTGG ATGGCGCTYTTKGACGAYTC	59.0ºC / 60 s
R. toruloides	**MC057** **MC058**	CATCTTCTGGTGTKTGGTCAMCGT GGGASGSTGTAGGCGATTTG	60.0ºC / 60 s
S. salmonicolor, *S. johnsonii*	**MC053** **MC054**	CATCTTCTTCCTCCTCCTGCC GGGTAGGGAGGAAGCGAGAT	59.5ºC / 90 s
Amplification of the intervening region between *RibL18ae* and *RNAPOL* genes (encompasses the *STE3. A1*) – Protocol B			
Species	**Primer Name**	**Sequence (5' → 3')**	**Annealing/ Extension**
S. metaroseus , *S. pararoseus*	**MC135** **MC138**	GAGAAGGAGCCKCARCCCAA CCTGTCCKCCKCGCATCTT	60ºC / 7 min – 10 cy. 60ºC / 7 min + 5 s/ cy. – 25 cy.
	MC188 **MC189**	ACCGWCCMTCGACYTTCTACT ACTCACRCTCTCVACACGCTC	60ºC / 4 min – 10 cy. 60ºC / 4 min + 2 s/ cy. – 25 cy.
	MC160 **MC192**	ACTYTCMTCTACATYTTYTGGTG CACCCTCACYTATATTTTCTGG	Primer-walking sequencing (*S. metaroseus*) Primer-walking sequencing (*S. pararoseus*)
Sp. ruberrimus, *Sp. blumeae*	**MC188** **MC189**	ACCGWCCMTCGACYTTCTACT ACTCACRCTCTCVACACGCTC	60ºC / 4 min – 10 cy. 60ºC / 4 min + 2 s/ cy. – 25 cy.
	MC160	ACTYTCMTCTACATYTTYTGGTG	Primer-walking sequencing
Sp. jilinensis	**MC188** **MC189**	ACCGWCCMTCGACYTTCTACT ACTCACRCTCTCVACACGCTC	60ºC / 4 min – 10 cy. 60ºC / 4 min + 2 s/ cy. – 25 cy.
	MC193	CACCTTGATCTACGTTTTCTGG	Primer-walking sequencing
No amplification of the *STE3.A1* allele was obtained in the following species using any combination of the above mentioned primers			
Rh. graminis, *"Rh. pinicortis"*, *R. dairenensis, Rh. mucilaginosa, Rh. pacifica, "Rh. pyritica", "R. feracious", R. fluviale*, *S. microsporus*, *Sp. odoratus*, *R. lusitaniae*, *S. longiusculus*			

Additional File 5. Continued.

Amplification of *STE3. A2* allele – Protocol A			
Species	**Primer Name**	**Sequence (5' → 3')**	**Annealing/ Extension**
R. babjevae, *Rh. graminis*, *R. kratochvilovae*,	**MC124** **MC125**	AACAKCGSGCCWCTGTACTGGCAG GCSACSACCCGAYSAAAGTTGCG	63.0ºC / 60 s
S. salmonicolor, *S. johnsonii*	**MC126** **MC127**	AATRCCGCTCCGCTYTACTGGCA TGTTCGGGTCTACAGGTTGGAGAG	62.0ºC / 90 s
R. diobovatum, *Rh. araucariae*, *"Rh. hamamotoiana"*, *R. paludigenum*, *R. toruloides*, *Rh. dairenensis*, *R. azoricum*, *R. fluviale*, *S. microsporus*, *R. lusitaniae*	**MC166** **MC169**	AKCGTKCCKCTSTACTGGCA CCGAAMGCGGCRAARAARAT	58.0ºC / 90 s
Rh. colostri	**MC166** **MC168**	AKCGTKCCKCTSTACTGGCA CYGAAMGCSGCGAAGAARAT	59.0ºC / 60 s
Amplification of the intervening region between *LSm7* and *RNAPOL* genes (encompasses the *STE3. A2*) – Protocol B			
Sp. ruberrimus	**MC123** **MC189**	AAAGGSTACGACCAGYTSCTCAAYCT ACTCACRCTCTCVACACGCTC	62ºC / 4 min – 10 cy. 62ºC / 4 min + 2 s/ cy. – 25 cy.
	MC195 **MC198**	TACAGGGTAARGGTATCTT CTGACAGGACGCATAGATACAT	54ºC / 2 min – 10 cy. 54ºC / 2 min + 1 s/ cy. – 25 cy.
Sp. pararoseus	**MC123** **MC189**	AAAGGSTACGACCAGYTSCTCAAYCT ACTCACRCTCTCVACACGCTC	62ºC / 4 min – 10 cy. 62ºC / 4 min + 2 s/ cy. – 25 cy.
	MC196 **MC199**	GCGAAGTHGCWGCTCGAATC CCGATAGCCGTCCAACAGG	60ºC / 2 min – 10 cy. 60ºC / 2 min + 1 s/ cy. – 25 cy.
S. longiusculus	**MC123** **MC189**	AAAGGSTACGACCAGYTSCTCAAYCT ACTCACRCTCTCVACACGCTC	62ºC / 4 min – 10 cy. 62ºC / 4 min + 2 s/ cy. – 25 cy.
	MC194 **MC197**	CCGTGAAGGACGCCTCGAT CTCTCGCTCAATGACTACCTT	59ºC / 2 min – 10 cy. 59ºC / 2 min + 1 s/ cy. – 25 cy.
	MC167	AATRCCGCKCCKCTYTACTGGCA	Primer-walking sequencing
No amplification of the *STE3.A2* allele was obtained in the following species using any combination of the abovementioned primers			
Rh. glutinis, *"Rh. sesimbrana"*, *"Rh. pinicortis", Rh. mucilaginosa, Rh. pacifica, R. sphaerocarpum*, *"Rh. pyritica", "R. feracious", Sp. odoratus*, *"R. oreadum"*, *S. longiusculus*, *Sp. blumeae*, *Sp. jilinensis*			

Additional File 5. Continued.

Confirm synteny in the vicinity of the *STE3.A1* allele in selected red yeast species (see Figure 3 and Additional File 2) **– Protocols A** and **B** Amplicons were obtained for the following red species but only fully sequenced in those for which a GenBank Accession number is given					
Species	**Primer Name**	**Sequence (5' → 3')**	**Annealing/ Extension**	**Amplified region**	**GenBank Accession Number**
R. babjevae, R. diobovatum	**MC109** **MC122**	CGTCCGCCGCCCGAACATCAAG TGGCGMGTYCCCGAACCYTACAAC	65ºC / 4 min – 10 cy. 65ºC / 4 min + 2 s/ cy. – 25 cy.	*Rib L18ae – RibL6* (includes the *STE3.A1*)	
R. kratochvilovae					JN246653
Rh. araucariae					JN246654
R. azoricum					JN246655
R. toruloides	**MC109** **MC161**	CGTCCGCCGCCCGAACATCAAG AAGAYSGCCGGCAKGCTCCARTA	65ºC / 4 min – 10 cy. 65ºC / 4 min + 2 s/ cy. – 25 cy.	*Rib L18ae – RtSTE3.A1*	
	MC057 **MC022**	CATCTTCTGGTGTKTGGTCAMCGT GGCGGATGGATTGGTTGAGC	60ºC / 6 min – 10 cy. 60ºC / 6 min + 2 s/ cy. – 25 cy.	*RtSTE3.A1 – RHA1*	[ref. 36]
R. sphaerocarpum	**MC109** **MC122**	CGTCCGCCGCCCGAACATCAAG TGGCGMGTYCCCGAACCYTACAAC	65ºC / 4 min – 10 cy. 65ºC / 4 min + 2 s/ cy. – 25 cy.	*Rib L18ae – RibL6* (includes the *STE3.A1*)	JN246664
	MC159 **MC164**	ACGYTGCTSTTCATYTTYTGGTG GAAGATSGCRCASGAGATKGCGTA	60.0ºC / 60 s		
Sp. ruberrimus, *Sp. blumeae*	**MC188** **MC189**	ACCGWCCMTCGACYTTCTACT ACTCACRCTCTCVACACGCTC	60ºC / 4 min – 10 cy. 60ºC / 4 min + 2 s/ cy. – 25 cy.	*Rib L18ae – RNAPOL* (includes the *STE3.A1* and *RibL6* genes)	
	MC160	ACTYTCMTCTACATYTTYTGGTG	Primer-walking sequencing		
S. pararoseus	**MC135** **MC138**	GAGAAGGAGCCKCARCCCAA CCTGTCCKCCKCGCATCTT	60ºC / 4 min – 10 cy. 60ºC / 4 min + 2 s/ cy. – 25 cy.	*Rib L18ae – RNAPOL* (includes the *STE3.A1* and *RibL6* genes)	JN246670
	MC188 **MC189**	ACCGWCCMTCGACYTTCTACT ACTCACRCTCTCVACACGCTC	Primer-walking sequencing		
	MC192	CACCCTCACYTATATTTTCTGG			

Additional File 5. Continued.

Species	Primer Name	Sequence (5' → 3')	Annealing/ Extension	Amplified region	GenBank Accession Number
S. metaroseus	**MC135** **MC138**	GAGAAGGAGCCKCARCCCAA CCTGTCCKCCKCGCATCTT	60ºC / 4 min – 10 cy. 60ºC / 4 min + 2 s/ cy. – 25 cy.	*Rib L18ae – RNAPOL* (includes the *STE3.A1* and *RibL6* genes)	
	MC188 **MC189**	ACCGWCCMTCGACYTTCTACT ACTCACRCTCTCVACACGCTC	Primer-walking sequencing		
	MC160	ACTYTCMTCTACATYTTYTGGTG			
Sp. jilinensis	**MC188** **MC189**	ACCGWCCMTCGACYTTCTACT ACTCACRCTCTCVACACGCTC	60ºC / 4 min – 10 cy. 60ºC / 4 min + 2 s/ cy. – 25 cy.	*Rib L18ae – RNAPOL* (includes the *STE3.A1* and *RibL6* genes)	
	MC193	CACCTTGATCTACGTTTTCTGG	Primer-walking sequencing		
S. ruineniae	**MC109** **MC122**	CGTCCGCCGCCCGAACATCAAG TGGCGMGTYCCCGAACCYTACAAC	65ºC / 4 min – 10 cy. 65ºC / 4 min + 2 s/ cy. – 25 cy.	*Rib L18ae – RibL6* (includes the *SruRHA1.A2*)	

Confirm synteny in the vicinity of the *STE3.A2* allele in selected red yeast species (see Figure 3 and Additional File 2) **– Protocols A** and **B** Amplicons were obtained for the following red species but only fully sequenced in those for which a GenBank Accession number is given					
Species	**Primer Name**	**Sequence (5' → 3')**	**Annealing/ Extension**	**Amplified region**	**GenBank Accession Number**
R. babjevae, Rh. graminis, *R. lusitaniae*	**MC122** **MC123**	TGGCGMGTYCCCGAACCYTACAAC AAAGGSTACGACCAGYTSCTCAAYCT	(LR-PCR) 63ºC / 4 min – 10 cy. 63ºC / 4 min + 2 s/ cy. – 25 cy.	*LSm7 – RibL6* (includes the *STE3.A2*)	
R. diobovatum					JN246579
R. kratochvilovae					JN246585
Rh. araucariae					JN246588
Rh. colostri					JN246603
S. johnsonii					JN246613
R. paludigenum, R. toruloides	**MC123** **MC124**	AAAGGSTACGACCAGYTSCTCAAYCT AACAKCGSGCCWCTGTACTGGCAG	62.0ºC / 2 min	*STE3.A2 – LSm7*	

Additional File 5. Continued.

Species	Primer Name	Sequence (5' → 3')	Annealing/ Extension	Amplified region	GenBank Accession Number
Sp. ruberrimus	**MC123** **MC189**	AAAGGSTACGACCAGYTSCTCAAYCT ACTCACRCTCTCVACACGCTC	62ºC / 4 min – 10 cy. 62ºC / 4 min + 2 s/ cy. – 25 cy.	*LSm7 – RNAPOL* (includes the *STE3.A2 and RibL6* genes)	
	MC195 **MC198**	TACAGGGTAARGGTATCTT CTGACAGGACGCATAGATACAT	54ºC / 2 min – 10 cy. 54ºC / 2 min + 1 s/ cy. – 25 cy.		
Sp. pararoseus	**MC123** **MC189**	AAAGGSTACGACCAGYTSCTCAAYCT ACTCACRCTCTCVACACGCTC	62ºC / 4 min – 10 cy. 62ºC / 4 min + 2 s/ cy. – 25 cy.	*LSm7 – RNAPOL* (includes the *STE3.A2 and RibL6* genes)	JN246610 JN246611
	MC196 **MC199**	GCGAAGTHGCWGCTCGAATC CCGATAGCCGTCCAACAGG	60ºC / 2 min – 10 cy. 60ºC / 2 min + 1 s/ cy. – 25 cy.		
S. longiusculus	**MC123** **MC189**	AAAGGSTACGACCAGYTSCTCAAYCT ACTCACRCTCTCVACACGCTC	(LR-PCR) 62ºC / 4 min – 10 cy. 62ºC / 4 min + 2 s/ cy. – 25 cy.	*LSm7 – RNAPOL* (includes the *STE3.A2 and RibL6* genes)	JN246612
	MC194 **MC197**	CCGTGAAGGACGCCTCGAT CTCTCGCTCAATGACTACCTT	(LR-PCR) 59ºC / 2 min – 10 cy. 59ºC / 2 min + 1 s/ cy. – 25 cy.		
	MC167	AATRCCGCKCCKCTYTACTGGCA	Primer-walking sequencing		
S. salmonicolor	**MC123** **MC189**	AAAGGSTACGACCAGYTSCTCAAYCT ACTCACRCTCTCVACACGCTC	62ºC / 4 min – 10 cy. 62ºC / 4 min + 2 s/ cy. – 25 cy.	*LSm7 – RNAPOL* (includes the *STE3.A2 and RibL6* genes)	
	MC126 **MC122**	AATRCCGCTCCGCTYTACTGGCA TGGCGMGTYCCCGAACCYTACAAC	62ºC / 2 min – 10 cy. 62ºC / 2 min + 1 s/ cy. – 25 cy.		

Amplification of *HD1/HD2* alleles – Protocol B			
Species	**Primer Name**	**Sequence (5' → 3')**	**Annealing/ Extension**
R. babjevae, *Rh. glutinis, Rh. graminis*, *R. diobovatum*,	**MC118** **MC120**	TGTTGRYGAACCAGGTRTCGAKCTG AGGATGKCGAGGACSTCGKGSGTGAA	65ºC / 1 min 30 s – 10 cy. 65ºC / 1 min 30 s + 1 s/ cy. – 25 cy.
S. salmonicolor, *S. johnsonii*	See Reference [21] of the manuscript		

Additional File 5. Continued.

Amplification intervening region between *STE20* and *KAP95* genes (encompasses the *RHA2*) – Protocol A					
Species	**Primer Name**	**Sequence (5' → 3')**	**Annealing/ Extension**	**Amplified region**	**GenBank Accession Number**
S. salmonicolor (ML 2241, NRRL Y-17498) *S. johnsonii* (CBS 1522)	**MC040** **MC073**	GCCCGAAGTCGTCAAGCAGAAGGA TTCCCSAAYGGCCARCTCAAGGAGCC	63.0ºC / 3 min	*STE20 – KAP95* (includes the *RHA2* genes)	JN246569 JN246568 JN246571
	MC113 **MC115**	AGCCRGCGCTTSGGATCAGTTAC AAGGAAGGATGCTYACTTGCGG	Primer-walking sequencing		
S. salmonicolor (PYCC 4558)	**MC040** **MC073**	GCCCGAAGTCGTCAAGCAGAAGGA TTCCCSAAYGGCCARCTCAAGGAGCC	63.0ºC / 3 min	*STE20 – KAP95* (includes the *RHA2* genes)	JN246570
	MC113 **MC115** **MC130**	AGCCRGCGCTTSGGATCAGTTAC AAGGAAGGATGCTYACTTGCGG TCCGAACAATGGATCAACAGGC	Primer-walking sequencing		
S. jonhsonii (strains: PYCC 4351)	**MC040** **MC073**	GCCCGAAGTCGTCAAGCAGAAGGA TTCCCSAAYGGCCARCTCAAGGAGCC	63.0ºC / 3 min	*STE20 – KAP95* (includes the *RHA2* genes)	JN246572
	MC114 **MC115**	GGAATCGGGACCCCAAGAAAC AAGGAAGGATGCTYACTTGCGG	Primer-walking sequencing		
R. toruloides	See Reference [21] of the manuscript				

Additional File 5. Continued.