

Supplementary data

Figure S1: Example of surperimposed forward chromatograms resulting from a mixed of ITSa and ITSb alleles present in strain CBS7103

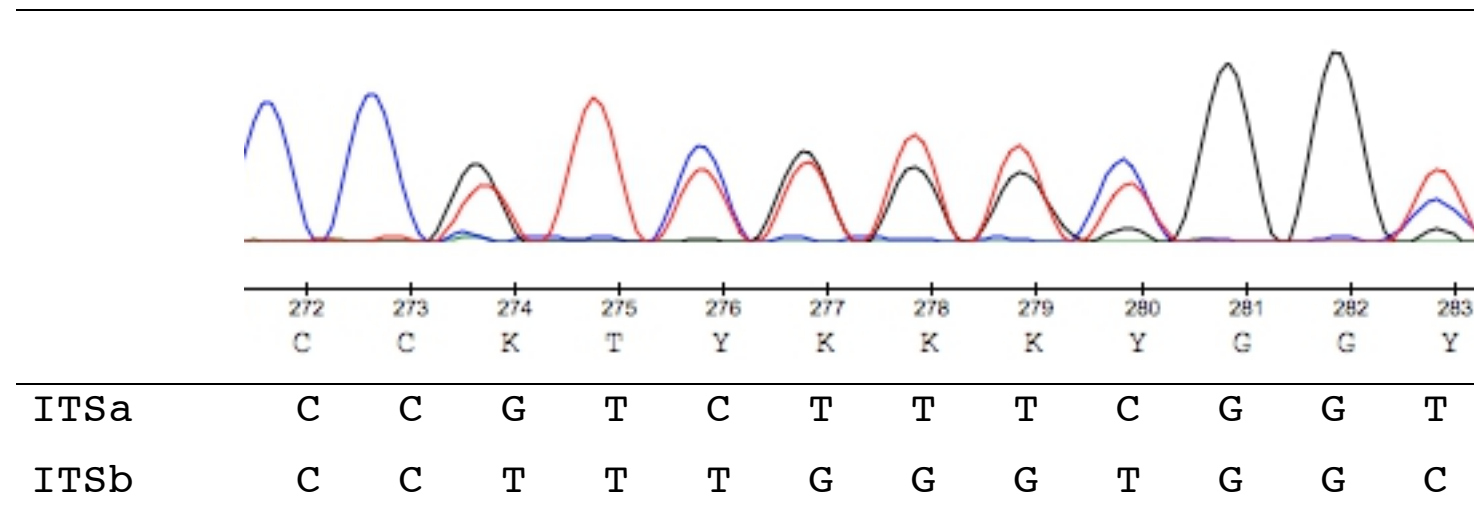


Table S1: Yeasts strains used in this study and their corresponding nucleotide sequence accession number in GenBank

Species	Strain designation at CBS and other culture collections		GenBank accession number		
			EF1	D1D2	ITS
<i>C. inconspicua</i>	CBS180	NRRLY2029	EF552464 This study KM252819	EF550240	AB179767
<i>C. inconspicua</i>		CIU71062		U71062	
<i>C. inconspicua</i>		UL325		HQ641283	
<i>C. inconspicua</i>		EXF-5648		JF766626	
<i>C. inconspicua</i>		1.5		JN417608	
<i>C. inconspicua</i>		CNRMA9.403			KP131716
<i>C. inconspicua</i>		PMM10-1744L			KP131718
<i>C. inconspicua</i>		CNRMA7.1481			KP131714
<i>C. inconspicua</i>					AB179767
<i>C. inconspicua</i>		WM11			AJ853766
<i>C. inconspicua</i>		WM02.462			KP131719
<i>C. inconspicua</i>		CNRMA9.723			KP131717
<i>C. inconspicua</i>		CNRMA7.1482			KP131715
<i>P. cactophila</i>		PMM08-858L			KP132499
<i>P. cactophila</i>		PMM09-2821L			KP132500
<i>P. cactophila</i>	CBS6926	NRRLY-10963	EF552465 This study 252820	EF550241	This study KM252828 KM252830 KM252831
<i>P. cactophila</i>	CBS7059		this study KM252801		this study KM252832 KM252833
<i>P. cactophila</i>	CBS2155		this study KM252803		this study KM252827
<i>P. cactophila</i>	CBS6928		this study KM252804	this study KM252821	this study KM252832
<i>P. cactophila</i>	CBS7103				This study KM252828 KM252829 KM252830
<i>P. cactophila</i>		IMAU5Y035		GU460181	
<i>P. cactophila</i>		IMAU5Y040		GU460184	
<i>P. cactophila</i>		IMAU5Y029		GU460177	
<i>P. cactophila</i>		IMAU5Y023		GU460174	
<i>P. cactophila</i>		YOH		GQ179989	
<i>P. cactophila</i>		PCU75731		U75731	
<i>P. cactophila</i>		YB2		KC512909	
<i>P. norvegensis</i>		AB179768.1			AB179768
<i>P. norvegensis</i>		F2fp15		JN110384	
<i>P. norvegensis</i>		F2fp21		JN110386	
<i>P. norvegensis</i>	CBS1922	IFO1020		AJ508574	AB278161
<i>P. norvegensis</i>		TY27		FJ972223	
<i>P. norvegensis</i>		ATCC20686		JQ070172	
<i>P. norvegensis</i>		PNU75730		U75730	
<i>P. norvegensis</i>		PUMY020		JN391355	
<i>P. norvegensis</i>		PUMY021		JN391356	
<i>P. norvegensis</i>		L6058			KC408998
<i>P. norvegensis</i>	CBS1911				AB278162

<i>P. norvegensis</i>		ATCC20686			JQ070096
<i>P. norvegensis</i>	CBS2327				AB278170
<i>P. norvegensis</i>		NCPF3861			AM74530
<i>P. norvegensis</i>		NYSDOH83MR 168			AY139789
<i>P. norvegensis</i>	AB179768.1				AB179768
<i>P. norvegensis</i>	CBS2145				AB278169
<i>P. norvegensis</i>		ST3481-03			AY939799
<i>P. norvegensis</i>	CBS2144				AB278168
<i>P. norvegensis</i>	CBS2128				AB278167
<i>P. norvegensis</i>	CBS2126				AB278166
<i>P. norvegensis</i>	CBS2125				AB278165
<i>P. norvegensis</i>	CBS1921				AB278163
<i>P. norvegensis</i>	CBS1953				AB278164
<i>P. norvegensis</i>	CBS6564	NRRLY7687	EF552463	EF550239	This study KM252826
<i>P. norvegensis</i>		PMM08-1253L			KP132518
<i>P. norvegensis</i>		PMM09-1823L			KP132519
<i>P. norvegensis</i>		PMM09-768L			KP132520
<i>P. norvegensis</i>		WM881			KP132523
<i>P. norvegensis</i>		WM887			KP132528
<i>P. norvegensis</i>		WM877			KP068971
<i>P. norvegensis</i>		WM875			KP132521
<i>P. norvegensis</i>		WM885			KP132526
<i>P. norvegensis</i>		WM884			KP132525
<i>P. norvegensis</i>		WM883			KP132524
<i>P. norvegensis</i>		WM879			KP132522
<i>P. norvegensis</i>		WM886			KP132527
<i>P. norvegensis</i>		CNRMA7.1213			KP132517
<i>P. norvegensis</i>		WM999			KP132529
<i>P. insulana</i>	CBS11169	TSU00-106.5	This study KM252818	EU747339.1 / this study KM252825	This study KM252834
<i>P. pseudocactophila</i>	CBS6929	NRRLY17239	EF552466	EF550242 this study KM252822	
<i>P. pseudocactophila</i>		PPU75732		U75732	
<i>P. kudriavzevii</i>	ATCC24210	NRRLY-5396	EF552446	EF550222	AY939808

Table S2: Blastn results of the complete or uncomplete ITS1/4 sequences tested in this study.

Strain	5'-end chromatogram readable size	3'-end chromatogram readable size	Genbank best hits of consensus sequence species	Genbank best hits of accession number	Maximum % identity	E value
Strains identified as <i>C. inconspicua</i> or <i>P. cactophila</i> using MALDI-TOF						
SA1461, SA1463, CBS6928, CBS2155, CBS180	369- 373	369- 373	<i>C. inconspicua</i>	AB179761	99% or 99% <sup>1</sup> or 100%	0.0
SA952, SA1635, SA1459, CBS7059	68	29-68	<i>C. ernobii</i>	EU343827	100%	2e-26
			<i>C. inconspicua</i>	AM745303	100%	2e-26
SA1729, SA1730, SA1457, SA1458, SA1462, SA2185, CBS7103	265	68-74	<i>C. ernobii</i>	EU343827	100% or 99% <sup>1</sup>	4e-135
			<i>C. inconspicua</i>	AB179767	100% or 99% <sup>1</sup>	4e-135
CBS6926	0 <sup>2</sup>	0 <sup>2</sup>	NA	NA	NA	NA
Strains identified as <i>P. norvegensis</i> using MALDI-TOF						
CBS6564	404	404	<i>P. norvegensis</i>	AB278164	99% <sup>1</sup>	0.0
SA1460	88	0 <sup>2</sup>	<i>P. norvegensis</i>	KC408998	98% <sup>1</sup>	4e-35
SA2546	0 <sup>2</sup>	0 <sup>2</sup>	NA	NA	NA	NA
Strains not identified using MALDI-TOF						
CBS11169	0 <sup>2</sup>	306	Uncultured fungus clone	KJ173550	95%	2e-132
			<i>C. inconspicua</i>	AM711386	95%	3e-131
				NR111116	94%	2e-128
CBS6929	51	101	<i>C. ernobii</i>	EU343827	84%	3e-18
			<i>C. inconspicua</i>	AB179767	85%	6e-20
			<i>P. norvegensis</i>	AB278164	80%	1e-15

<sup>1</sup> percentage could be 100% when considering point mutation of heterozygoty

<sup>2</sup> fully uninterpretable

Table S3: Blastn results of complete or incomplete D1/D2 domain sequences tested in this study.

Strains	5'-end chromatogram readable size	3'-end chromatogram readable size	GenBank Best hits of consensus sequence species	GenBank Best hits accession number	Maximum identity	% E value
Strains identified as <i>C. inconspicua</i> or <i>P. cactophila</i> using MALDI-TOF						
SA1730, CBS6926, 573-574		573-574	<i>P. cactophila</i>	GU460184	99% <sup>1</sup>	0.0
CBS6928, CBS180			<i>C. inconspicua</i>	EF550240		
SA1729, SA2185, 55-68-71-108		0	<i>C. inconspicua</i>	EF550240	99% <sup>1</sup>	6e-46
CBS7103, CBS7059						
CBS2155	51	0	<i>C. inconspicua</i>	EF550240	100%	3e-17
			<i>P. norvegensis</i>	KF268295	100%	3e-17
			<i>P. cactophila</i>	GU460184	100%	3e-17
SA952, SA1635, SA1457, SA1458, SA1459, SA1461, SA1462, SA1463	0 <sup>2</sup> -12	0 <sup>2</sup>	NA	NA	NA	NA
Strains identified as <i>P. norvegensis</i> using MALDI-TOF						
CBS6564	571	571	<i>P. norvegensis</i>	EF550239	99% <sup>1</sup>	0.0
SA2546	571	571	<i>P. norvegensis</i>	FJ972223	99% <sup>1</sup>	0.0
SA1460	0-12	0	NA	NA	NA	NA
Strains not identified using MALDI-TOF						
CBS11169	604	604	<i>P. insulana</i>	EU747343	99% <sup>1</sup>	0.0
CBS6929	575	575	<i>P. pseudocactophila</i>	EF550242	99% <sup>1</sup>	0.0

<sup>1</sup> percentage could be 100% when considering heterozygous point mutations <sup>2</sup> fully uninterpretable

Table S4: Blastn results of EF-1 $\alpha$  sequences obtained from the strains tested in this study.

Strains	fragment size in bp	Blastn best hit of consensus sequence species	GenBank best hit (accession number)	Maximum % identity	E value
Strains identified as <i>C. inconspicua</i> or <i>P. cactophila</i> using MALDI-TOF					
SA952, SA1457, SA1458, SA1459, SA1461, SA1462, SA1463, SA1635, SA1729, SA1730, SA2185 CBS180, CBS7103, CBS2155, CBS7059, CBS6928	780	<i>C. inconspicua</i> <i>P. cactophila</i>	EF552464 EF552465	99% <sup>1</sup> 98% <sup>1</sup>	0.0
CBS6926	780	<i>P. cactophila</i> <i>C. inconspicua</i>	EF552465 EF552464	98% <sup>1</sup> 98% <sup>1</sup>	0.0
Strains identified as <i>P. norvegensis</i> using MALDI-TOF					
SA1460, SA2546, CBS6564	780	<i>P. norvegensis</i>	EF552463	100%	0.0
Strains not identified using MALDI-TOF					
CBS11169	809	<i>C. inconspicua</i> , <i>P. cactophila</i> , <i>P. pseudocactophila</i>	EF552464.1 EF552465.1 EF552466.1	98% 98% 98%	0.0
CBS6929	780	<i>P. pseudocactophila</i>	EF552466	100%	0.0

<sup>1</sup>percentage could be 100% when considering heterozygous point mutations

Table S5: Nucleotide divergence within and between the species of the *P. cactophila* clade. ITS region, D1D2 domain and EF-1 $\alpha$  gene fragment, were analysed, using complete sequences obtained from our study, and from GenBank. The number of base substitutions per site from averaging over all sequence pairs within each group and between groups are shown. Analyses were conducted using the Maximum Composite Likelihood model. PCCI: *P. cactophila* plus *C. inconspicua*, PC: *P. cactophila*, CI: *C. inconspicua*, PN: *P. norvegensis*, PP: *P. pseudocactophila*, PI: *P. insulana*, CK: *C. krusei*

A- ITS1/4 sequences

PCCI							
PCCI	0.010	PC					
PC	NA	0.007	CI				
CI	NA	0.011	0.011	PN			
PN	0.093	0.090	0.097	0.005	PP		
PP	NA	NA	NA	NA	NA	PI	
PI	0.048	0.046	0.050	0.115	NA	NA	CK
CK	0.208	0.204	0.213	0.166	NA	0.209	NA

B- D1/D2 sequences

PCCI							
PCCI	0.001	PC					
PC	NA	0.000	CI				
CI	NA	0.001	0.003	PN			
PN	0.054	0.054	0.055	0.003	PP		
PP	0.033	0.032	0.033	0.040	0.000	PI	
PI	0.015	0.015	0.016	0.051	0.030	NA	CK
CK	0.098	0.100	0.102	0.115	0.098	0.095	NA

C- EF-1 $\alpha$  gene fragment

PCCI		PC		CI		PN		PP		PI		CK	
PCCI	0.000												
PC	NA	0.004											
CI	NA	0.003	0.001										
PN	0.019	0.020	0.0019	0									
PP	0.026	0.028	0.027	0.017	NA								
PI	0.029	0.029	0.034	0.025	0.033	NA							
CK	0.061	0.062	0.062	0.052	0.057	0.063	NA						

NA: not applicable