Mitochondrial and nuclear gene sequences to infer the phylogeny of Pezizomycotina (Ascomycota)

Mario E. E. Franco[1]*, Silvina López[2], Mario C. N. Saparrat[2,3] & Pedro Balatti[1].

Centro de Investigaciones de Fitopatología, Facultad de Ciencias Agrarias y Forestales, Universidad Nacional de La Plata. La Plata, Buenos Aires, Argentina[1]; Instituto de Fisiología Vegetal, UNLP-CONICET. La Plata, Buenos Aires, Argentina[2]; Instituto Carlos Spegazzini, Facultad de Ciencias Naturales y Museo. La Plata, Buenos Aires, Argentina[3].
*e-mail: ernesto.franco@agro.unlp.edu.ar

INTRODUCTION

Pezizomycotina is the largest subphylum of Ascomycota. It includes filamentous species widespread in nature where they live as saprotrophs, mutualists and/or parasites of animal, plants and fungi (1). In some instances, the lack of distinctive morphological characters makes their morphology-based identification a real challenge. In order to overcome this limitation, several nuclear and mitochondrial genes have been used with diagnostic and taxonomic purposes (2). But, is the information contained in these genes subjected to different mutational forces phylogenetically congruent?

OBJETIVE

The aim of this work was to compare the phylogeny of pezizomycotina species inferred from the complete set of mitochondrial-encoded genes and from five nuclear-encoded genes widely used in fungal taxonomy.

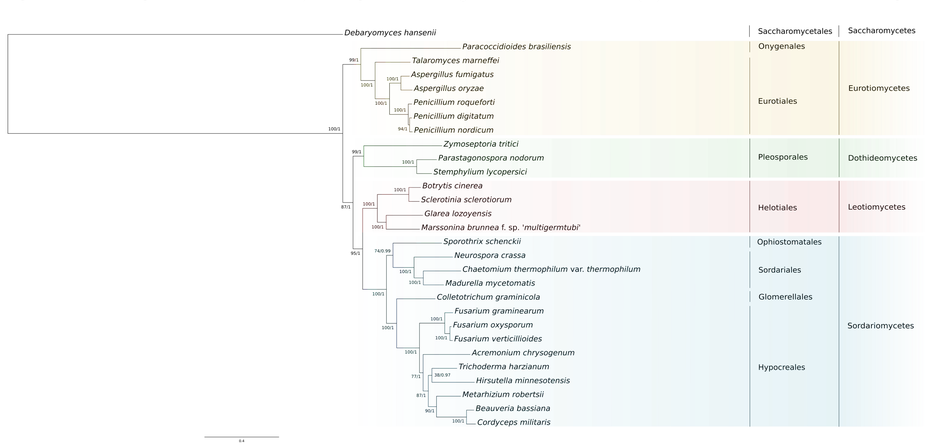
MATERIAL AND METHODS

The phylogenetic reconstitution was carried out by the concatenated nucleotide sequences of 12 mitochondrial genes and the individual and concatenated nucleotide sequence of 5 nuclear genes. The ingroup consisted of 28 pezizomycotina species, a sample size that was limited by the number of species for which all this nucleotide information was publicly available. The saccharomycotina *Debaryomyces hansenii* was used as outgroup (see table). Alignments were generated with ClustalW (3) and automatically curated with Gblocks 0.91b (4). Best-fit model of evolution was selected with jModelTest 2.1.7 (5) and data matrices were analyzed under maximum likelihood and bayesian criteria in PhyML 3.0 (6) and MrBayes 3.2 (7), respectively. Phylogenetic trees were edited in FigTree v1.4.2 (8) and Inkscape 0.91 (9).

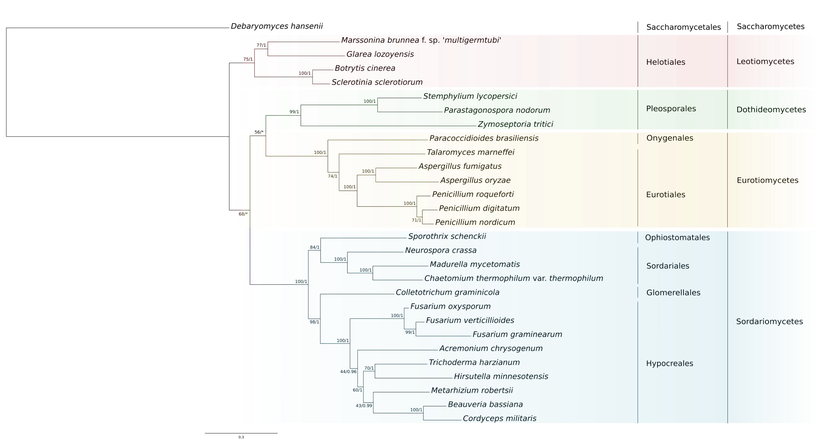
Organism	Beta-tubulin		Elongation factor 1-alpha		Glyceraldehyde 3-phosphate dehydrogenase		DNA-directed RNA polymerase II subunit RPB1		DNA-directed RNA polymerase II subunit RPB2		Mitogenome
	Locus tag	Accession number	Locus tag	Accession number	Locus tag	Accession number	Locus tag	Accession number	Locus tag	Accession number	Accession number
Acremonium chrysogenum ATCC 11550	ACRE_032630	JPKY01000025	ACRE_065030	JPKY01000086	ACRE_019650	JPKY01000012	ACRE_036870	JPKY01000031	ACRE_077950	JPKY01000126	NC_023268
Aspergillus fumigatus Af293	AFUA_1G10910	NC_007194	AFUA_1G06390	AAHF01000007	AFUA_5G01970	NC_007198	AFUA_1G14680	NC_007194	AFUA_7G01920	NC_007200	JQ346808
Aspergillus oryzae 3.042	Ao3042_04421	AKHY01000128	Ao3042_03515	AKHY01000119	Ao3042_05102	AKHY01000002	Ao3042_07218	AKHY01000156	Ao3042_06477	AKHY01000150	NC_018100
Beauveria bassiana D1-5	BBAD15_g10308	ANFO01001065	BBAD15_g4641	ANFO01000374	BBAD15_g6090	ANFO01000567	BBAD15_g2249	ANFO01000152	BBAD15_g12296	ANFO01001505	ANFO01001745
Botrytis cinerea B05.10	BC1G_00122	NW_001814571	BC1G_09492	NW_001814509	BC1G_11968	NW_001814480	BC1G_12445	NW_001814472	BC1G_09004	NW_001814520	KC832409
Chaetomium thermophilum var. thermophilum DSM 1495	CTHT_0041050	NW_006383034	CTHT_0012210	NW_006383030	CTHT_0004880	NW_006383023	CTHT_0058480	NW_006383036	CTHT_0045510	NW_006383034	NC_015893
Colletotrichum graminicola M1.001	GLRG_08786	NW_007361045	GLRG_02486	NW_007361009	GLRG_05186	NW_007361021	GLRG_00584	NW_007361006	GLRG_03562	NW_007361014	NW_007361658
Cordyceps militaris CM01	CCM_04055	NW_006271971	CCM_00809	NW_006271969	CCM_04549	NW_006271971	CCM_05485	NW_006271972	CCM_06694	NW_006271973	KP719096
Debaryomyces hansenii CBS767	DEHA2D08800g	NC_006046	DEHA2D02376g	NC_006046	DEHA2F04796g	NC_006048	DEHA2A13618g	NC_006043	DEHA2G07502g	NC_006049	NC_010166
Fusarium graminearum NRRL 31084	FGSG_06611	NC_026477	FGSG_08811	NC_026475	FGSG_06257	NC_026476	FGSG_00916	NC_026474	FGSG_02659	NC_026474	NC_009493
Fusarium oxysporum UASWS AC1	-	JNNQ01001150	-	JNNQ01001079	-	JNNQ01001163	-	JNNQ01001152	-	JNNQ01001085	KR952337
Fusarium verticillioides 7600	FVEG_05512	DS022247	FVEG_02381	DS022243	FVEG_04927	DS022246	FVEG_00683	DS022242	FVEG_09286	DS022253	NC_016687
Glarea lozoyensis 74030	M7I_0791	AGUE01000013	M7I_3044	AGUE01000067	M7I_3455	AGUE01000078	M7I_7592	AGUE01000229	M7I_4004	AGUE01000100	KF169905
Hirsutella minnesotensis 3608	HIM_10565	KQ030654	HIM_04902	KQ030515	HIM_00238	KQ030498	HIM_03184	KQ030507	HIM_03128	KQ030507	NC_027660
Madurella mycetomatis mm55	MMYC01_G07311	LCTW01000394	MMYC01_G0542	LCTW01000246	MMYC01_G03227	LCTW01000197	MMYC01_G06051	LCTW01000519	MMYC01_G00277	LCTW01000001	NC_018359
Marssonina brunnea f. sp. *'multigermtubi'* MB_m1	MBM_05801	NW_006763051	MBM_07351	NW_006763056	MBM_05497	NW_006763050	MBM_03515	NW_006763044	MBM_04681	NW_006763048	NC_015991
Metarhizium robertsii ARSEF 23	MAA_02081	NW_011942149	MAA_03797	NW_011942171	MAA_07675	NW_011942149	MAA_06658	NW_011942215	MAA_00336	NW_011942171	JELW01000367
Neurospora crassa OR74A	NCU04054	NC_026506	CU02003	NC_026501	MAA_07675	NCU01528	NCU02103	NC_026501	NCU05120	NC_026506	NC_026614
Paracoccidioides brasiliensis Pb18	PADG_02900	NW_011371360	PADG_00692	NW_011371358	PADG_02411	NW_011371360	PADG_06747	NW_011371366	PADG_04340	NW_011371361	NC_007935
Parastagonospora nodorum SN15	SNOG_06791	NW_001884558	SNOG_11663	NW_001884568	SNOG_01278	NW_001884550	SNOG_11556	NW_001884567	SNOG_11456	NW_001884567	NC_009746
Penicillium digitatum Pd1	PDIP_27420	NW_014574610	PDIP_69010	NW_014574583	PDIP_47870	NW_014574585	PDIP_66280	NW_014574583	PDIP_55940	NW_014574584	AKCU01000562
Penicillium nordicum DAOMC 185683	ACN38_g2403	LHQQ01000026	ACN38_g7728	LHQQ01000133	CN38_g7290	LHQQ01000121	ACN38_g3180	LHQQ01000011	ACN38_g2016	LHQQ01000021	KR952336
Penicillium roqueforti UASWS P1	-	JNNS01000301	-	JNNS01000412	-	JNNS01000385	-	JNNS01000373	-	JNNS01000134	NC_027416
Sclerotinia sclerotiorum 1980 UF-70	SS1G_04652	NW_001820831	SS1G_05520	NW_001820830	SS1G_07798	NW_001820826	SS1G_05928	NW_001820829	SS1G_00885	NW_001820835	KT283062
Sporothrix schenckii ATCC 58251	HMPREF1624_07151	KI440850	HMPREF1624_04957	KI440846	HMPREF1624_08681	KI440857	HMPREF1624_00281	KI440842	HMPREF1624_01203	KI440842	NC_015923
Stemphylium lycopersici CIDEFI 216	TW65_01392	LGLR01000178	TW65_04709	LGLR01000223	TW65_04473	KNG48731	TW65_04414	LGLR01000215	TW65_06600	LGLR01000278	KX453765
Talaromyces marneffei PM1	GQ26_0140870	JPOX01000014	GQ26_0060610	JPOX01000006	GQ26_0120900	JPOX01000012	GQ26_0110470	JPOX01000011	GQ26_0021640	JPOX01000002	NC_025256
Trichoderma harzianum T6776	-	JNNP01001389	-	JNNP01001073	-	JNNP01001364	THAR02_09524	JOKZ01000428	THAR02_08409	JOKZ01000322	KR952346
Zymoseptoria tritici IPO323	MYCGRDRAFT_102950	CM001196	MYCGRDRAFT_92705	NC_018215	MYCGRDRAFT_99044	NC_018217	MYCGRDRAFT_99873	NC_018215	MYCGRDRAFT_99853	NC_018215	NC_010222

RESULTS

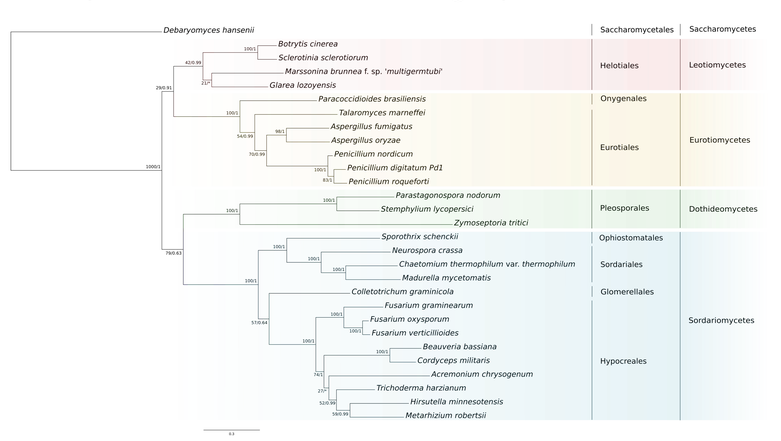
Pezizomycotina phylograms. Maximum likelihood trees are shown. Numbers at the nodes represents maximum likelihood bootstrap support values (as a percentage of 1000 replicates) and bayesian posterior probabilities, respectively. An asterisk (*) indicates when a particular node was not recovered in the bayesian inference. The scale bar represents the number of changes per sites.



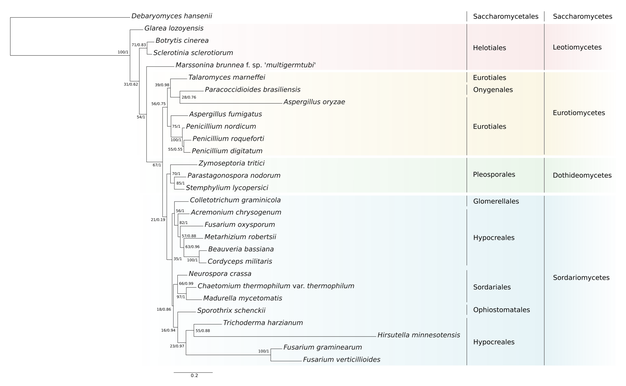
Phylogram inferred from the concatenated mitochondrial genes. Gblocks selected 9197 conservative positions from the original alignment of 17249. jModelTest (AIC) identified GTR +I+G as the best-fit nucleotide substitution model for the dataset.



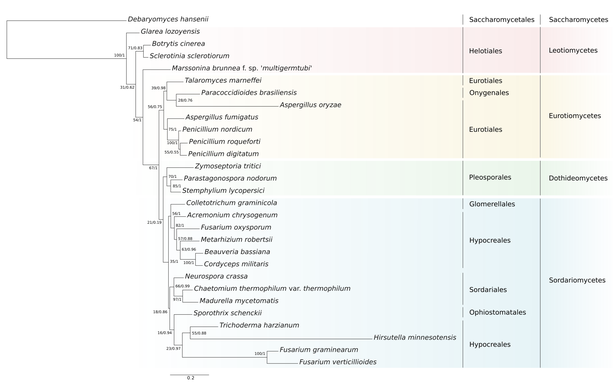
Phylogram inferred from the concatenated nuclear genes. Gblocks selected 10575 conservative positions from the original alignment of 26679. jModelTest (AIC) identified GTR +I+G as the best-fit nucleotide substitution model for the dataset.



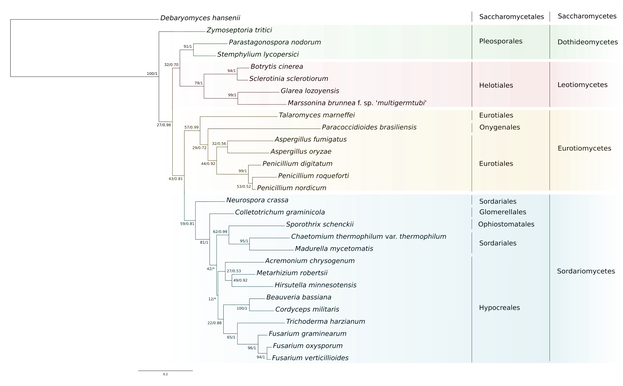
Phylogram inferred from the DNA-directed RNA polymerase II subunit RPB2 gene. Gblocks selected 3229 conservative positions from the original alignment of 5511. jModelTest (AIC) identified GTR+I+G as the best-fit nucleotide substitution model for the dataset.



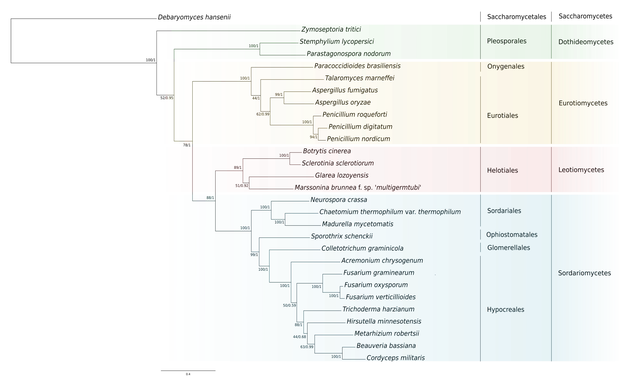
Phylogram inferred from the beta-tubulin gene. Gblocks selected 1100 conservative positions from the original alignment of 5272. jModelTest (AIC) identified TIM1+I+G as the best-fit nucleotide substitution model for the dataset.



Phylogram inferred from the elongation factor 1-alpha gene. Gblocks selected 1054 conservative positions from the original alignment of 5049. jModelTest (AIC) identified TrN+I+G as the best-fit nucleotide substitution model for the dataset.



Phylogram inferred from the glyceraldehyde 3-phosphate dehydrogenase gene. Gblocks selected 831 conservative positions from the original alignment of 3518. jModelTest (AIC) identified TrN+I+G as the best-fit nucleotide substitution model for the dataset.



Phylogram inferred from the DNA-directed RNA polymerase II subunit RPB1 gene. Gblocks selected 4361 conservative positions from the original alignment of 7329. jModelTest (AIC) identified GTR +I+G as the best-fit nucleotide substitution model for the dataset.

CONCLUSION

Rpb2 and the concatenated nuclear and mitochondrial datasets clustered species in concordance with the current systematic of this subphylum. Since different genes may have different histories, special care should be taken in order to obtain an accurate phylogenetic estimation.

BIBLIOGRAPHY: 1. Carris, L. M., C. R. Little, and C. M. Stiles. Introduction to Fungi. The Plant Health Instructor. 2012. DOI: 10.1094/PHI-I-2012-0426-01; 2. Shenoy, B. D. Multigene phylogeny of selected anamorphic ascomycetes. HKU Theses Online (HKUTO), 2007; 3. Chenna,R., Sugawara ,H., Koike,T., Lopez,R., Gibson,T.J., Higgins,D.G. and Thompson,J.D. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. 2003; 31: 3497-3500; 4. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 2007; 56: 564-577; 5. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods. 2012; 9: 772; 6. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 2003; 52: 696-704; 7. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 2001; 17: 754-755; 8. FigTree website: http://tree.bio.ed.ac.uk/software/figtree/; 9. Inkscape website: https://inkscape.org/en/