# Supplemental Materials: AYbRAH: a curated ortholog database for yeasts and fungi spanning 600 million years of evolution

October 8, 2018

## Contents

1	Comparison of AYbRAH orthology assignments to published ortholog databases	<b>2</b>

2 Sample webpages for homolog groups

6

# 1 Comparison of AYbRAH orthology assignments to published ortholog databases

Table S1: Comparison of manually curated acetyl-Coenzyme A synthetase orthologs in AYbRAH to highly cited ortholog databases. N/A indicates genomes that do not have orthology relationships in the public database but have been assigned orthology with AYbRAH. Omission indicates genomes that have annotations in the public ortholog database but do not have an annotation for the given gene. PANTHER is the only database that can distinguish between the three ACS ortholog groups; ACS3 is assigned to a different PANTHER family despite the shared ancestry of all the orthologs. KEGG can only differentiate between ACS1 and ACS3 ortholog groups, while all other database orthologly assignments are polyphyletic. EggNOG includes FOG07524 and FOG07525 in the ACS ortholog group, which both have predicted acetoacetate-CoA ligase activity.

AYbRA	H Annotation	Ortholog DB	spo	ang	ncr	yli	ppa	pic	kla	sce
ACS1	FOG00404		PTHR24095:SF14	N/A	PTHR24095:SF14	PTHR24095:SF14	N/A	N/A	N/A	PTHR24095:SF14
ACS2	FOG00405	PANTHER					N/A	N/A	N/A	PTHR24095:SF245
ACS3	FOG00406			N/A	PTHR43347:SF3					
ACS1	FOG00404		K01895	K01895	K01895	K01895	K01895	K01895	K01895	K01895
ACS2	FOG00405	KEGG					K01895	K01895	K01895	K01895
ACS3	FOG00406			K01908	K01908					
ACS1	FOG00404		OG5_126680	N/A	OG5_126680	OG5_126680	N/A	OG5_126680	OG5_126680	OG5_126680
ACS2	FOG00405	OrthoMCL					N/A	$OG5_{-}126680$	$OG5_{-}126680$	OG5_126680
ACS3	FOG00406			N/A	OG5_126680					
ACS1	FOG00404		EOG8HQC0H	EOG8HQC0H	EOG8HQC0H	EOG8HQC0H	EOG8HQC0H	EOG8HQC0H	EOG8HQC0H	EOG8HQC0H
ACS2	FOG00405	OrthoDB					EOG8HQC0H	EOG8HQC0H	EOG8HQC0H	EOG8HQC0H
ACS3	FOG00406			omission	EOG8HQC0H					
ACS1	FOG00404		KOG1175	KOG1175	KOG1175	KOG1175	KOG1175	KOG1175	KOG1175	KOG1175
ACS2	FOG00405						KOG1175	KOG1175	KOG1175	KOG1175
ACS3	FOG00406	EggNOG		KOG1175	KOG1175					
HYP	FOG07524			KOG1175	KOG1175					
HYP	FOG07525			KOG1175	KOG1175					

Table S2: Comparison of manually curated Type II NADH dehydrogenase (NDH2) orthologs in AYbRAH to highly cited ortholog databases. N/A indicates genomes that do not have orthology relationships in the public database but have been assigned orthology with AYbRAH. Omission indicates genomes that have annotations in the public ortholog database but do not have an annotation for the given gene. PANTHER is able to distinguish between most orthologs in the NDH2 family, with the exception of NDE1 and NDE2; NDE0 is in a different PANTHER family than the rest of the NDH2 genes. KEGG is the only other database that can differentiate between some NDH2 genes; the genes are split between the older NDI0/NDE0 ortholog group and more recent NDE1/NDI1 ortholog group. PANTHER and EggNOG contain additional genes not included in other ortholog databases, which may represent ancient paralogs having lower sequence similarities than other NDH2 paralogs. AIF1, which can localize to the mitochondria in S. cerevisiae, is in the same subfamily as NDE0 in PANTHER; the other inconsistency is an Aspergillus niger gene (FOG07265) paralogous to a characterized external NADH dehydrogenase in Neurospora crassa (FOG07264) in EggNOG (Carneiro et al., 2007).

	AH annotation	Ortholog Database	spo	ang	ncr	yli	ppa	pic	kla	sce
NDI0	FOG00837			N/A	PTHR43706:SF11					
NDE0	FOG00838			N/A	PTHR42913:SF2	PTHR42913:SF2	N/A		N/A	
AIF1	FOG11982	PANTHER					N/A			PTHR42913:SF2
NDE1	FOG00839		PTHR43706:SF1	N/A	PTHR43706:SF1	PTHR43706:SF1	N/A	N/A	N/A	PTHR43706:SF1
NDE2	FOG00845									PTHR43706:SF1
NDI1	FOG00846								N/A	PTHR43706:SF10
NDI0	FOG00837			omission	K03885					
NDE0	FOG00845			omission	K03885	K03885	K03885		K03885	
NDE1	FOG00845	KEGG	K17871	K17871	omission	K17871	K17871	K17871	K17871	K17871
NDE2	FOG00845									K17871
NDI1	FOG00845								K17871	K17871
NDI0	FOG00837			N/A	OG5_126960					
NDE0	FOG00838			N/A	OG5_126960	OG5_126960	N/A		OG5_126960	
NDE1	FOG00839	OrthoMCL	OG5_126960	N/A	OG5_126960	OG5_126960	N/A	OG5_126960	OG5_126960	$OG5_{-}126960$
NDE2	FOG00845									OG5_126960
NDI1	FOG00846								OG5_126960	OG5_126960
NDI0	FOG00837			EOG8P8D18	EOG8P8D18					
NDE0	FOG00838			EOG8P8D18	EOG8P8D18	EOG8P8D18	EOG8P8D18		EOG8P8D18	
NDE1	FOG00839	OrthoDB	EOG8P8D18	EOG8P8D18	EOG8P8D18	EOG8P8D18	EOG8P8D18	EOG8P8D18	EOG8P8D18	EOG8P8D18
NDE2	FOG00845									EOG8P8D18
NDI1	FOG00846								EOG8P8D18	EOG8P8D18
NDI0	FOG00837	EggNOG		KOG2495	KOG2495		<b>TTO CO.</b> 105		TTO 00 105	
NDE0	FOG00838			KOG2495	KOG2495	KOG2495	KOG2495		KOG2495	
NDE1	FOG00839		KOG2495	KOG2495	KOG2495	KOG2495	KOG2495	KOG2495	KOG2495	KOG2495
NDE2	FOG00845	-001.00								KOG2495
NDI1	FOG00846								KOG2495	KOG2495
HYP	FOG07265			KOG2495						

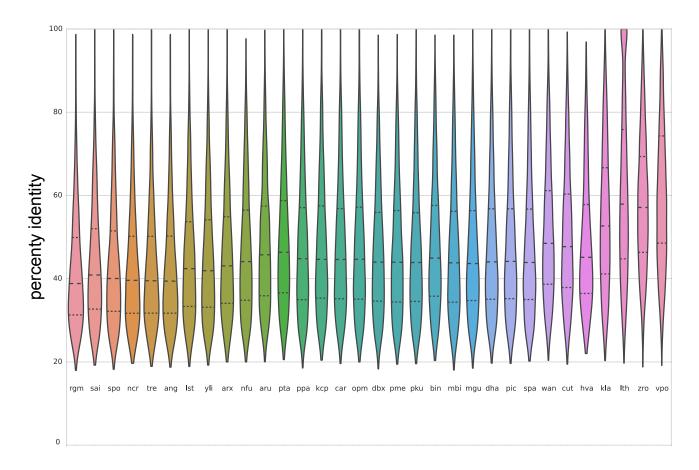


Figure S1: Distributions of BLASTP percent identities for proteins identified as orthologous to *Saccharomyces cerevisiae* in AYbRAH.

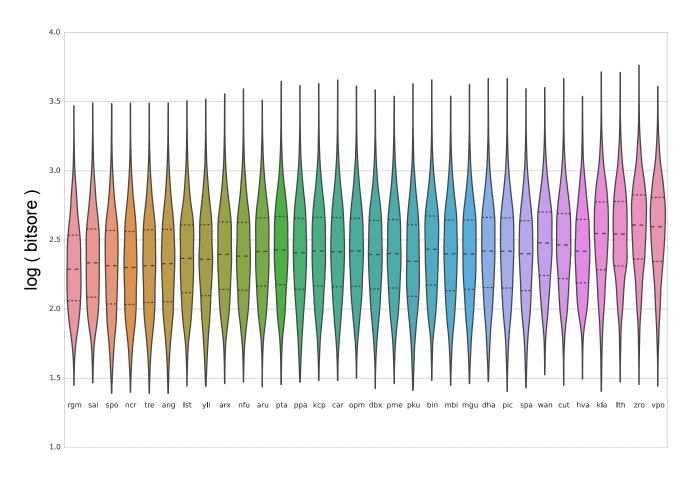


Figure S2: Distributions of logarithm BLASTP bitscores for proteins orthologous to *Saccharomyces cerevisiae* in AYbRAH.

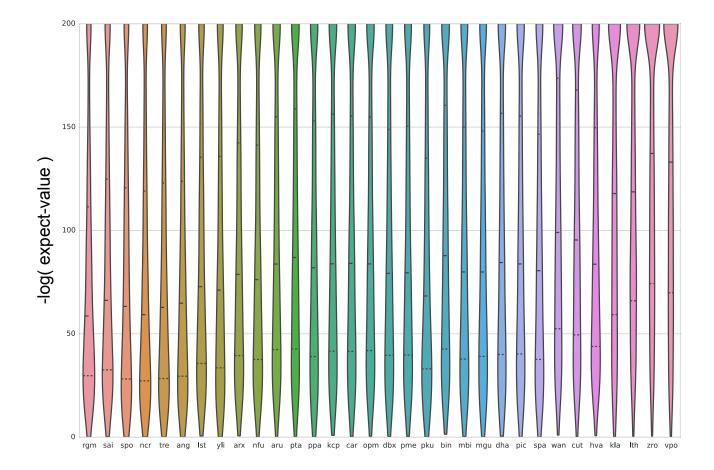


Figure S3: Distributions of negative logarithm of BLASTP expect-values for proteins orthologous to *Saccharomyces cerevisiae* in AYbRAH.

# 2 Sample webpages for homolog groups



# Analyzing Yeasts by Reconstructing Ancestry of Homologs



#### Home

HOG00227HAD-like hydrolase superfamily. DOG/GPP familyHOG00229acetyl-CoA hydrolase/transferase family

FASTA MAFFT sequence alignment Phyml trees Gblocks Phobius predictions

FOG00404 EOG8HQC0H ACS1 sce:ACS1

Genes: 34

#### **Protein description**

Acetyl-coA synthetase isoform expressed with non-fermentable carbon sources. Spo gene expressed with fermentable carbon sources.

#### **SGD Description**

Acetyl-coA synthetase isoform; along with Acs2p, acetyl-coA synthetase isoform is the nuclear source of acetyl-coA for histone acetylation; expressed during growth on nonfermentable carbon sources and under aerobic conditions

#### **PomBase Description** acetyl-CoA ligase (predicted)

AspGD Description Putative acetyl-CoA synthase

#### References

<u>Armitt S, et al. (1976 Feb)</u>. Analysis of acetate non-utilizing (acu) mutants in Aspergillus nidulans.

Payton M, et al. (1976 May). Agar as a carbon source and its effect on the utilization of other carbon sources by acetate non-utilizing (acu) mutants of Aspergillus nidulans.

<u>Frenkel EP, et al. (1977 Jan 25)</u>. Purification and properties of acetyl coenzyme A synthetase from bakers' yeast.

Hynes MJ, et al. (1977 Sep). Induction of the acetamidase of Aspergillus nidulans by acetate metabolism.

<u>Midelfort CF, et al. (1978 Oct 25)</u>. The stereochemical course of acetate activation by yeast acetyl-CoA synthetase.

Bal J, et al. (1979 Dec). Allele specific and locus non-specific suppressors in Aspergillus nidulans.

Kelly JM, et al. (1981 Apr). The regulation of phosphoenolpyruvate carboxykinase and the NADP-linked malic enzyme in Aspergillus nidulans.

Sandeman RA, *et al.* (1989 Jul). Isolation of the facA (acetyl-coenzyme A synthetase) and acuE (malate synthase) genes of Aspergillus nidulans.

<u>Connerton IF, et al. (1990 Mar)</u>. Comparison and cross-species expression of the acetyl-CoA synthetase genes of the Ascomycete fungi, Aspergillus nidulans and Neurospora crassa.

Lloyd AT, et al. (1991 Nov). Codon usage in Aspergillus nidulans.

<u>Sandeman RA, et al. (1991 Sep</u>). Molecular organisation of the malate synthase genes of Aspergillus nidulans and Neurospora crassa.

<u>Birch PR, et al. (1992)</u>. Nucleotide sequence of a gene from Phanerochaete chrysosporium that shows homology to the facA gene of Aspergillus nidulans.

Maconochie MK, *et al.* (1992 Aug). The acu-1 gene of Coprinus cinereus is a regulatory gene required for induction of acetate utilisation enzymes.

<u>De Virgilio C, *et al.* (1992 Dec)</u>. Cloning and disruption of a gene required for growth on acetate but not on ethanol: the acetyl-coenzyme A synthetase gene of Saccharomyces cerevisiae.

<u>Kujau M, et al. (1992 Mar</u>). Characterization of mutants of the yeast Yarrowia lipolytica defective in acetyl-coenzyme A synthetase.

<u>Saleeba JA, et al. (1992 Nov</u>). Characterization of the amdA-regulated aciA gene of Aspergillus nidulans.

<u>Steensma HY, et al. (1993 Apr)</u>. Genetic and physical localization of the acetyl-coenzyme A synthetase gene ACS1 on chromosome I of Saccharomyces cerevisiae.

Martínez-Blanco H, et al. (1993 Aug 25). Characterisation of the gene encoding acetyl-CoA synthetase in Penicillium chrysogenum: conservation of intron position in plectomycetes.

<u>Gouka RJ, et al. (1993 Jan)</u>. Development of a new transformant selection system for Penicillium chrysogenum: isolation and characterization of the P. chrysogenum acetylcoenzyme A synthetase gene (facA) and its use as a homologous selection marker.

<u>Sealy-Lewis HM, et al. (1994 Jan)</u>. A new selection method for isolating mutants defective in acetate utilisation in Aspergillus nidulans.

<u>Van den Berg MA, et al. (1995 Aug 1)</u>. ACS2, a Saccharomyces cerevisiae gene encoding acetyl-coenzyme A synthetase, essential for growth on glucose.

<u>Kratzer S, et al. (1995 Aug 8)</u>. Carbon source-dependent regulation of the acetyl-coenzyme A synthetase-encoding gene ACS1 from Saccharomyces cerevisiae.

van den Berg MA, *et al.* (1996 Nov 15). The two acetyl-coenzyme A synthetases of Saccharomyces cerevisiae differ with respect to kinetic properties and transcriptional regulation.

<u>de Jong-Gubbels P, et al. (1997 Aug 1)</u>. The Saccharomyces cerevisiae acetyl-coenzyme A synthetase encoded by the ACS1 gene, but not the ACS2-encoded enzyme, is subject to

glucose catabolite inactivation.

<u>Clutterbuck AJ, et al. (1997 Jun)</u>. The validity of the Aspergillus nidulans linkage map.

Todd RB, et al. (1998 Apr 1). FacB, the Aspergillus nidulans activator of acetate utilization genes, binds dissimilar DNA sequences.

<u>Stemple CJ, et al. (1998 Dec)</u>. The facC gene of Aspergillus nidulans encodes an acetateinducible carnitine acetyltransferase.

<u>Papadopoulou S, et al. (1999 Sep 1)</u>. The Aspergillus niger acuA and acuB genes correspond to the facA and facB genes in Aspergillus nidulans.

<u>Dessen P, et al. (2000 Feb 22)</u>. The PAUSE software for analysis of translational control over protein targeting: application to E. nidulans membrane proteins.

Brock M, et al. (2000 Mar). Methylcitrate synthase from Aspergillus nidulans: implications for propionate as an antifungal agent.

Jones IG, et al. (2001 Feb). ADHII in Aspergillus nidulans is induced by carbon starvation stress.

Lodi T, et al. (2001 Sep). Three target genes for the transcriptional activator Cat8p of Kluyveromyces lactis: acetyl coenzyme A synthetase genes KlACS1 and KlACS2 and lactate permease gene KlJEN1.

<u>Hynes MJ, et al. (2002 Jan)</u>. Regulation of the acuF gene, encoding phosphoenolpyruvate carboxykinase in the filamentous fungus Aspergillus nidulans.

Kumar A, et al. (2002 Mar 15). Subcellular localization of the yeast proteome.

<u>Flipphi M, et al. (2002 May 15)</u>. Characteristics of physiological inducers of the ethanol utilization (alc) pathway in Aspergillus nidulans.

<u>Flipphi M, et al. (2003 Apr 4)</u>. Onset of carbon catabolite repression in Aspergillus nidulans. Parallel involvement of hexokinase and glucokinase in sugar signaling.

Zeeman AM, et al. (2003 Jan 15). The acetyl co-enzyme A synthetase genes of Kluyveromyces lactis.

Sickmann A, et al. (2003 Nov 11). The proteome of Saccharomyces cerevisiae mitochondria.

<u>Flipphi M, et al. (2003 Sep)</u>. Relationships between the ethanol utilization (alc) pathway and unrelated catabolic pathways in Aspergillus nidulans.

Brock M, et al. (2004 Aug). On the mechanism of action of the antifungal agent propionate.

Sims AH, et al. (2004 Feb). Use of expressed sequence tag analysis and cDNA microarrays of the filamentous fungus Aspergillus nidulans.

Jogl G, et al. (2004 Feb 17). Crystal structure of yeast acetyl-coenzyme A synthetase in complex with AMP.

<u>Takasaki K, et al. (2004 Mar 26)</u>. Fungal ammonia fermentation, a novel metabolic mechanism that couples the dissimilatory and assimilatory pathways of both nitrate and ethanol. Role of acetyl CoA synthetase in anaerobic ATP synthesis.

Zhang YQ, et al. (2004 Oct). Connection of propionyl-CoA metabolism to polyketide biosynthesis in Aspergillus nidulans.

David H, et al. (2006). Metabolic network driven analysis of genome-wide transcription data from Aspergillus nidulans.

Mogensen J, et al. (2006 Aug). Transcription analysis using high-density micro-arrays of Aspergillus nidulans wild-type and creA mutant during growth on glucose or ethanol.

Takahashi H, et al. (2006 Jul 21). Nucleocytosolic acetyl-coenzyme a synthetase is required for histone acetylation and global transcription.

<u>Hynes MJ, et al. (2006 May)</u>. Regulatory genes controlling fatty acid catabolism and peroxisomal functions in the filamentous fungus Aspergillus nidulans.

<u>Salazar M, *et al.* (2009 Dec)</u>. Uncovering transcriptional regulation of glycerol metabolism in Aspergilli through genome-wide gene expression data analysis.

Shimizu M, et al. (2009 Jan). Proteomic analysis of Aspergillus nidulans cultured under hypoxic conditions.

<u>Flipphi M, et al. (2009 Mar)</u>. Biodiversity and evolution of primary carbon metabolism in Aspergillus nidulans and other Aspergillus spp.

<u>Hynes MJ, et al. (2010 Jul)</u>. ATP-citrate lyase is required for production of cytosolic acetyl coenzyme A and development in Aspergillus nidulans.

<u>Oh YT, et al. (2010 Mar)</u>. Proteomic analysis of early phase of conidia germination in Aspergillus nidulans.

<u>Wendland J, et al. (2011 Dec)</u>. Genome evolution in the eremothecium clade of the Saccharomyces complex revealed by comparative genomics.

<u>Saykhedkar S, et al. (2012 Jul 26)</u>. A time course analysis of the extracellular proteome of Aspergillus nidulans growing on sorghum stover.

<u>Georgakopoulos P, et al. (2012 Nov</u>). SAGA complex components and acetate repression in Aspergillus nidulans.

<u>Nakamura T, et al. (2012 Sep</u>). Impaired coenzyme A synthesis in fission yeast causes defective mitosis, quiescence-exit failure, histone hypoacetylation and fragile DNA.

<u>Carpy A, et al. (2014 Aug)</u>. Absolute proteome and phosphoproteome dynamics during the cell cycle of Schizosaccharomyces pombe (Fission Yeast).

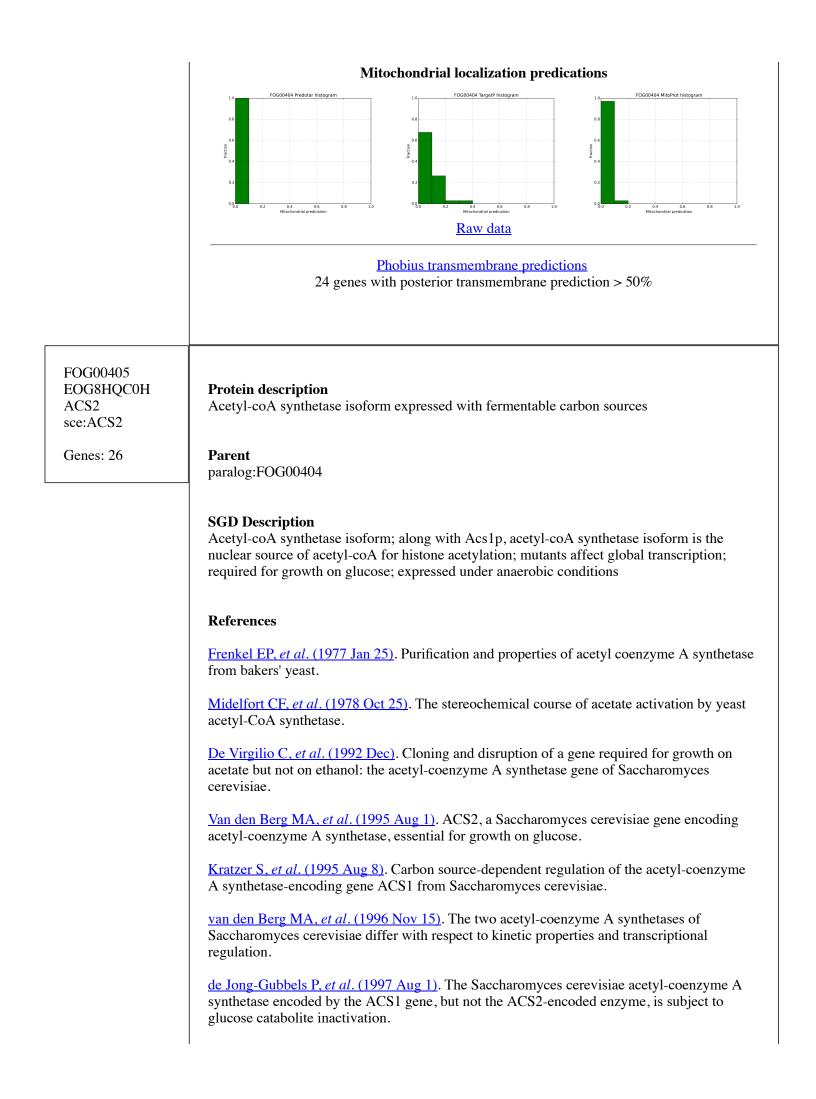
<u>Beckley JR, et al. (2015 Dec)</u>. A Degenerate Cohort of Yeast Membrane Trafficking DUBs Mediates Cell Polarity and Survival.

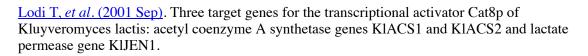
Malecki M, et al. (2016 Nov 25). Functional and regulatory profiling of energy metabolism in fission yeast.

Mitochondrial localization predications TargetP

MitoProt

Predotar

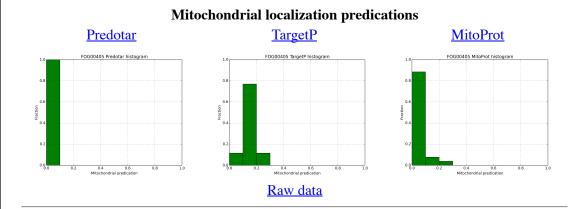


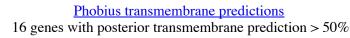


Peng J, et al. (2003 Aug). A proteomics approach to understanding protein ubiquitination.

Zeeman AM, et al. (2003 Jan 15). The acetyl co-enzyme A synthetase genes of Kluyveromyces lactis.

Takahashi H, et al. (2006 Jul 21). Nucleocytosolic acetyl-coenzyme a synthetase is required for histone acetylation and global transcription.





FOG00406 EOG8HQC0H ACS3 sce:absent

Genes: 3

**Protein description** Uncharacterized acetyl-coA synthetase paralog

**Parent** paralog:FOG00404

AspGD Description Putative acetyl-CoA synthase

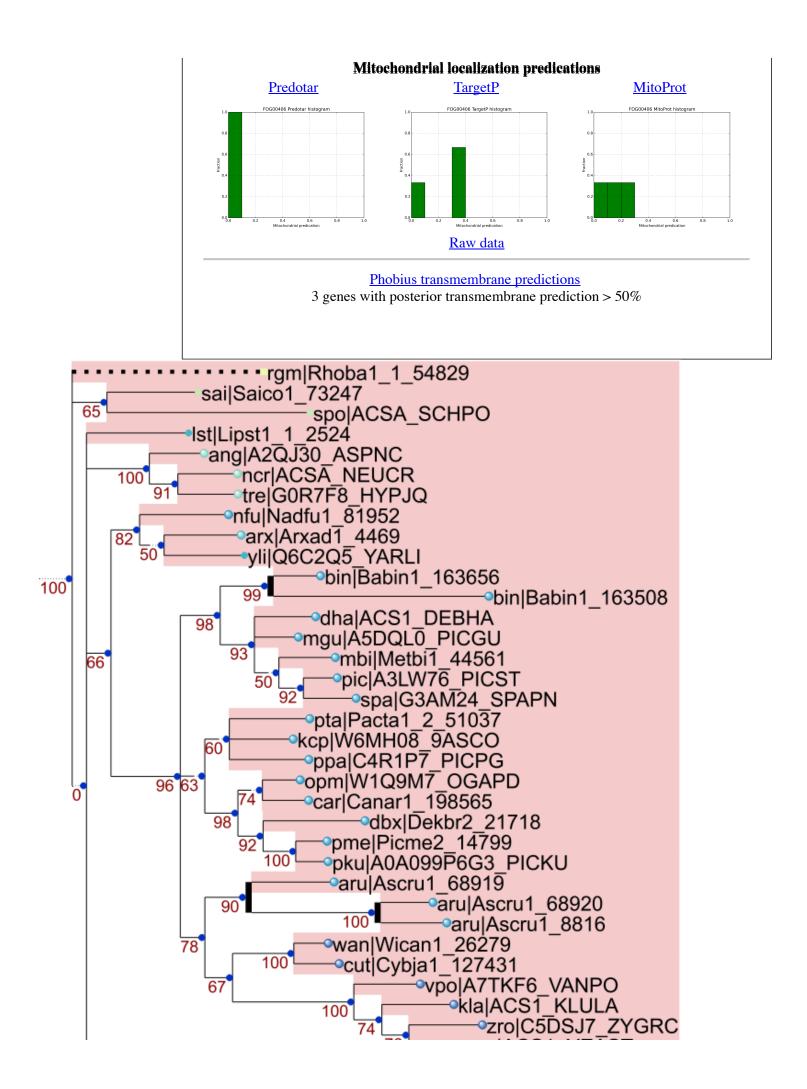
#### References

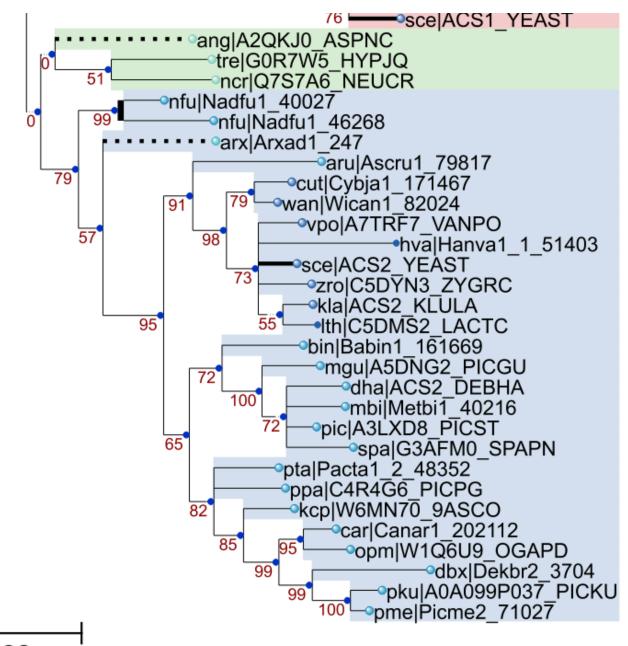
Zhang YQ, et al. (2004 Oct). Connection of propionyl-CoA metabolism to polyketide biosynthesis in Aspergillus nidulans.

<u>Salazar M, et al. (2009 Dec)</u>. Uncovering transcriptional regulation of glycerol metabolism in Aspergilli through genome-wide gene expression data analysis.

<u>Flipphi M, et al. (2009 Mar)</u>. Biodiversity and evolution of primary carbon metabolism in Aspergillus nidulans and other Aspergillus spp.

Lee MK, et al. (2014 May). NsdD is a key repressor of asexual development in Aspergillus nidulans.









Analyzing Yeasts by Reconstructing Ancestry of Homologs



#### <u>Home</u>

HOG00398FMC1 familyHOG00399HOG00403SUN family

FASTA MAFFT sequence alignment Phyml trees Gblocks Phobius predictions

FOG00837 EOG8P8D18 NDI0 sce:absent

Genes: 6

**Protein description** Internal NADH dehydrogenase, orthologous to NDH2 in E. coli

Features NADH

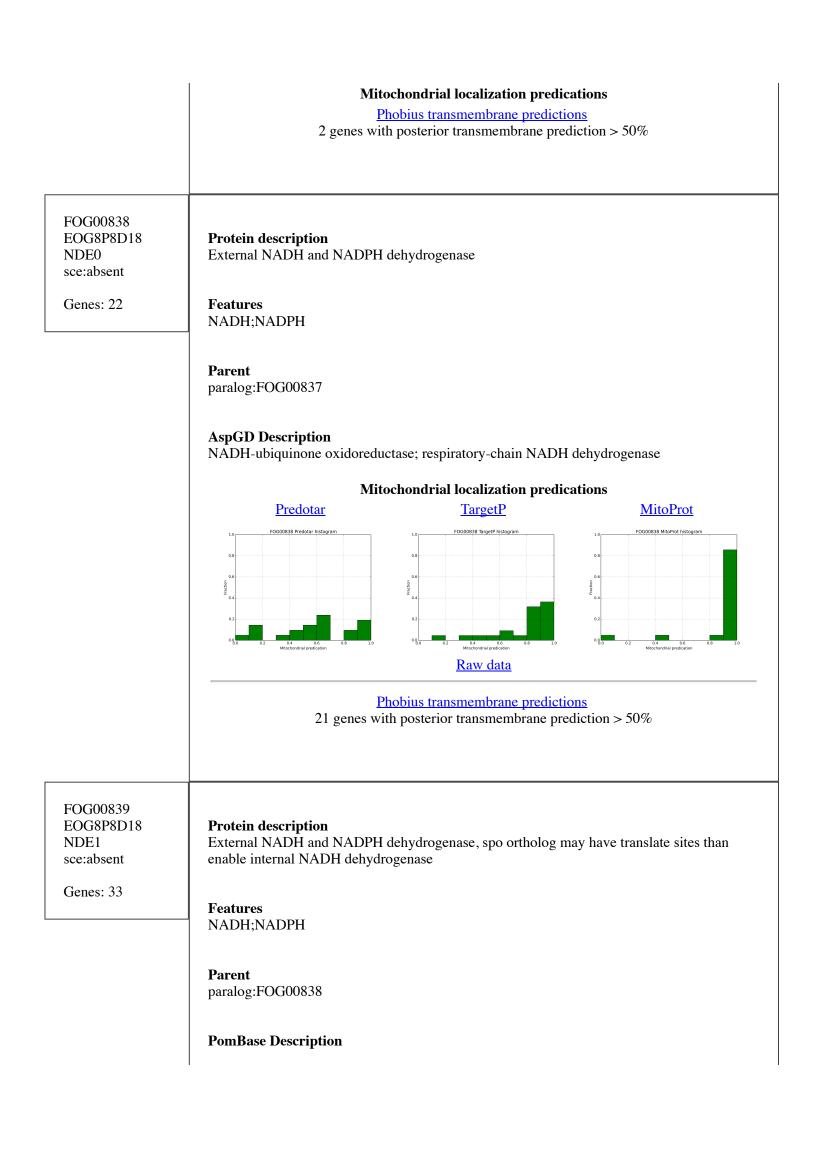
AspGD Description Ortholog(s) have NADPH dehydrogenase activity, role in spore germination and mitochondrial inner membrane localization

#### References

<u>Dinamarco TM, et al. (2010 Dec)</u>. The roles played by Aspergillus nidulans apoptosisinducing factor (AIF)-like mitochondrial oxidoreductase (AifA) and NADH-ubiquinone oxidoreductases (NdeA-B and NdiA) in farnesol resistance.

<u>Dinamarco TM, et al. (2011 Oct)</u>. Farnesol-induced cell death in the filamentous fungus Aspergillus nidulans.

<figure>Discontraction biscontraction biscontraction biscontractionOpenanTargetMicoProtOpenanOpenant



mitochondrial NADH dehydrogenase (ubiquinone) Nde2 (predicted)

AspGD Description

NADH-ubiquinone oxidoreductase; respiratory-chain NADH dehydrogenase

#### References

Kerscher SJ, et al. (1999 Jul). A single external enzyme confers alternative NADH:ubiquinone oxidoreductase activity in Yarrowia lipolytica.

Todd BL, et al. (2006 Apr). Sterol regulatory element binding protein is a principal regulator of anaerobic gene expression in fission yeast.

Harris SD, et al. (2009 Mar). Morphology and development in Aspergillus nidulans: a complex puzzle.

<u>Dinamarco TM, et al. (2010 Dec)</u>. The roles played by Aspergillus nidulans apoptosisinducing factor (AIF)-like mitochondrial oxidoreductase (AifA) and NADH-ubiquinone oxidoreductases (NdeA-B and NdiA) in farnesol resistance.

<u>Arita Y, et al. (2011 May)</u>. Microarray-based target identification using drug hypersensitive fission yeast expressing ORFeome.

<u>Dinamarco TM, et al. (2011 Oct)</u>. Farnesol-induced cell death in the filamentous fungus Aspergillus nidulans.

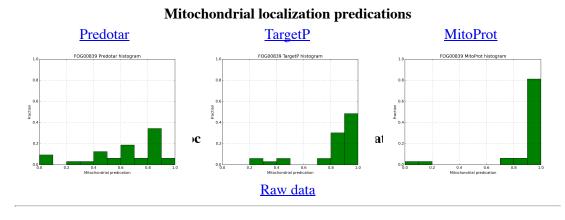
Lando D, et al. (2012). The S. pombe histone H2A dioxygenase Ofd2 regulates gene expression during hypoxia.

<u>Carpy A, et al. (2014 Aug)</u>. Absolute proteome and phosphoproteome dynamics during the cell cycle of Schizosaccharomyces pombe (Fission Yeast).

<u>Guo L, et al. (2016 Oct 13)</u>. Global Fitness Profiling Identifies Arsenic and Cadmium Tolerance Mechanisms in Fission Yeast.

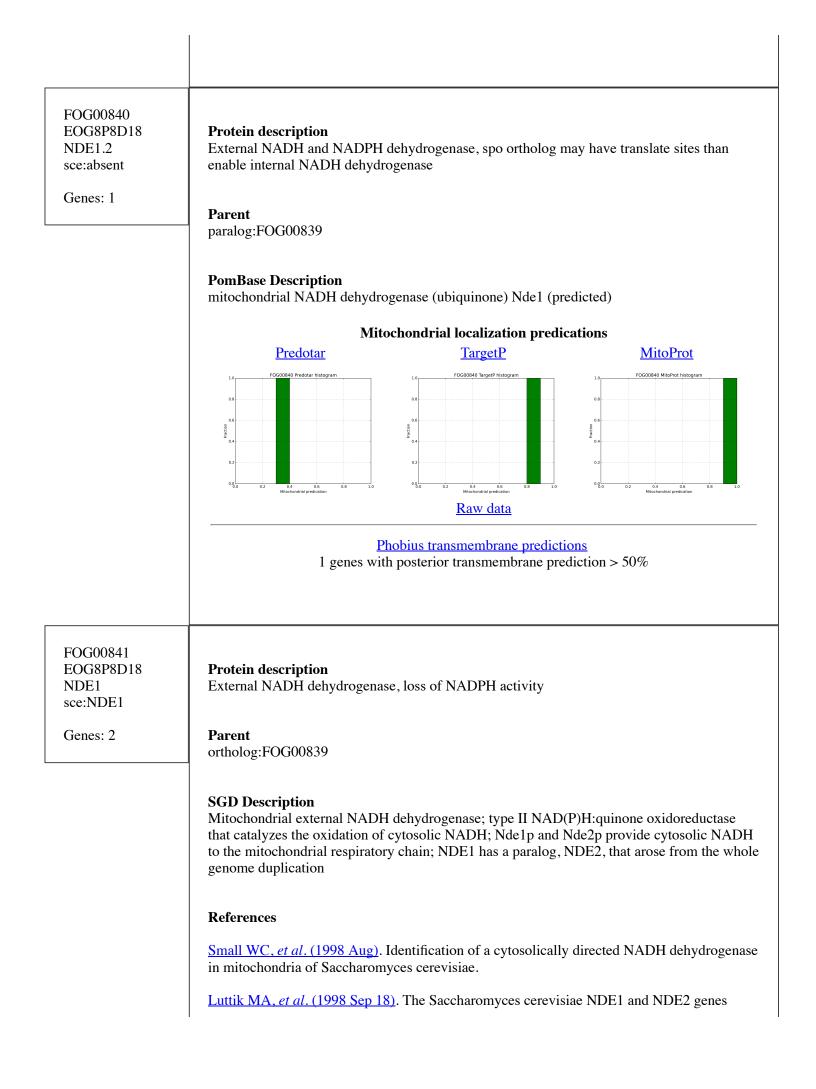
<u>Dudin O, et al. (2017 Apr</u>). A systematic screen for morphological abnormalities during fission yeast sexual reproduction identifies a mechanism of actin aster formation for cell fusion.

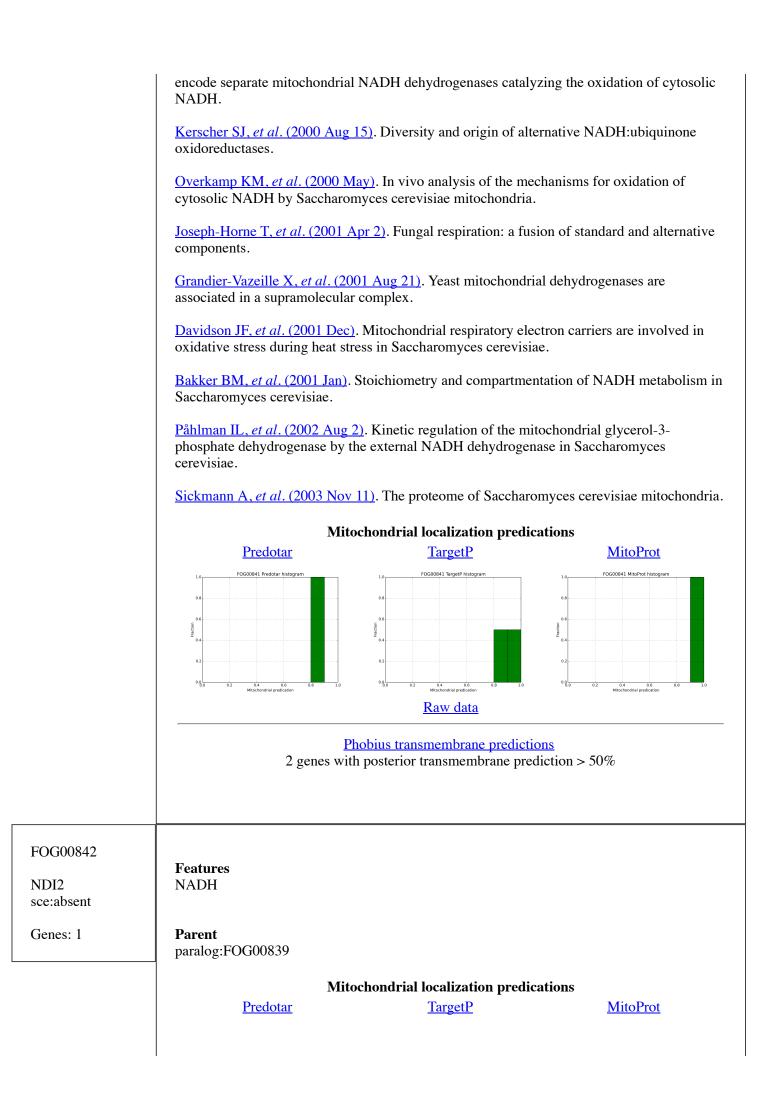
Lee J, *et al.* (2017 Feb 20). Chromatin remodeller Fun30<sup>Fft3</sup> induces nucleosome disassembly to facilitate RNA polymerase II elongation.

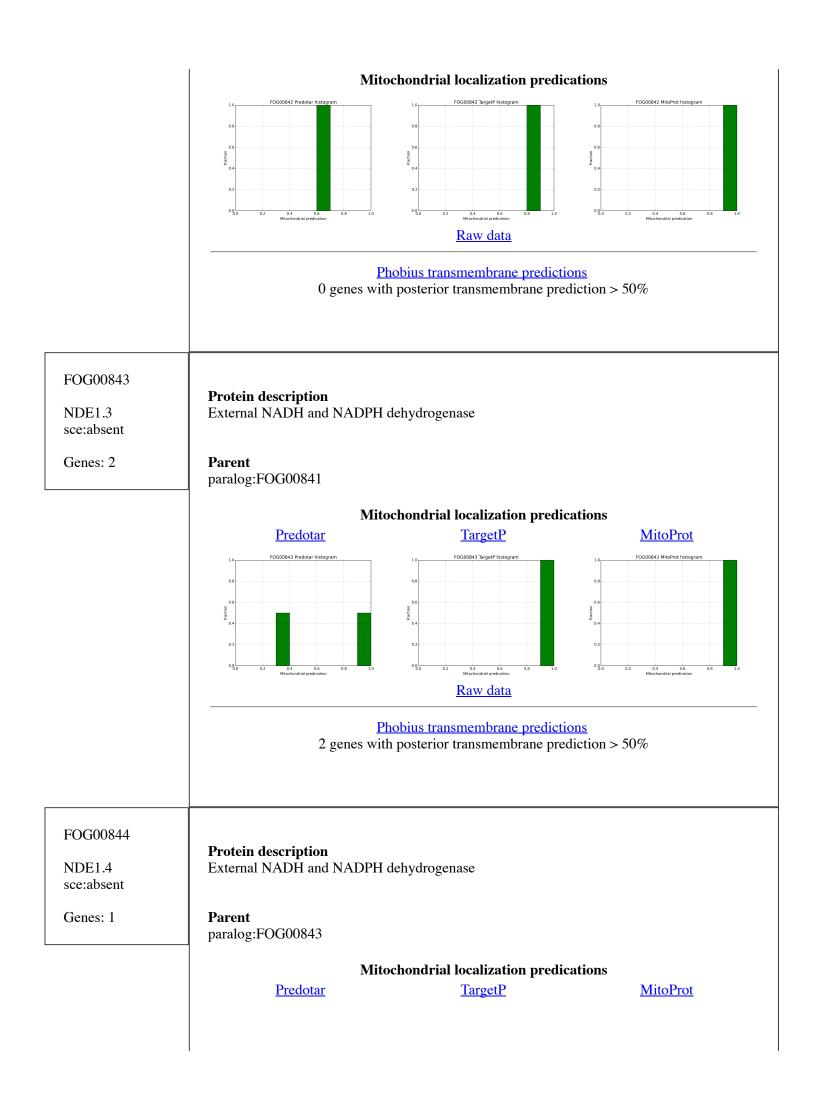


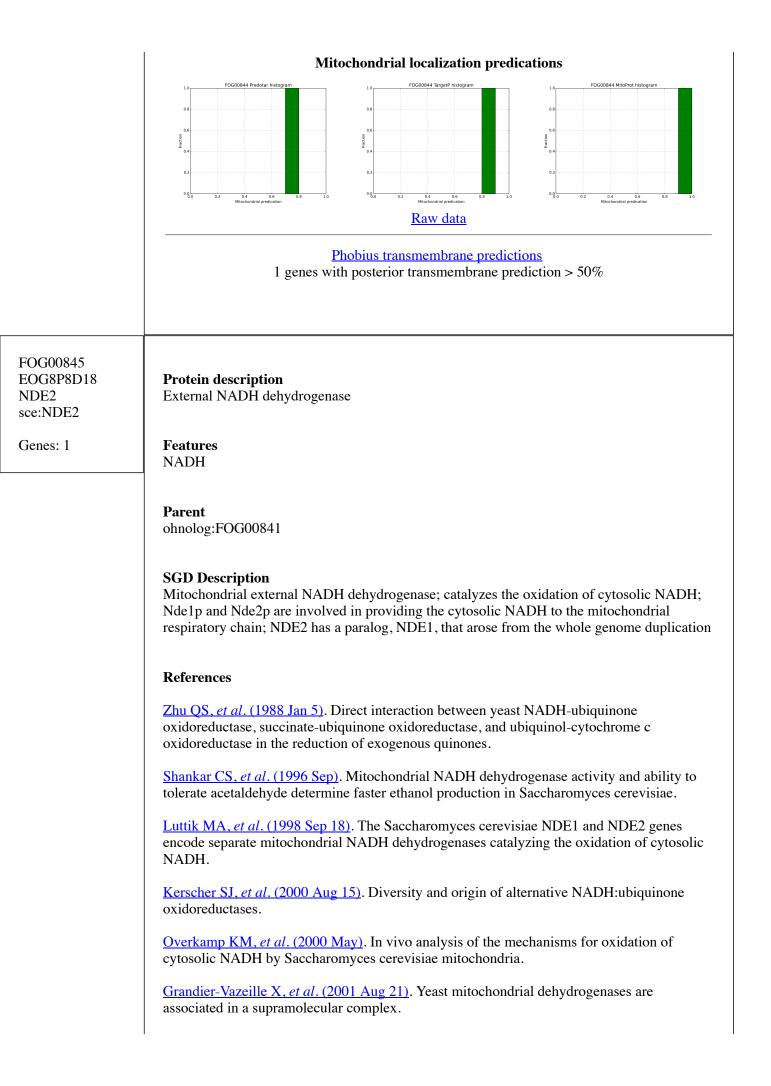
Phobius transmembrane predictions

31 genes with posterior transmembrane prediction > 50%

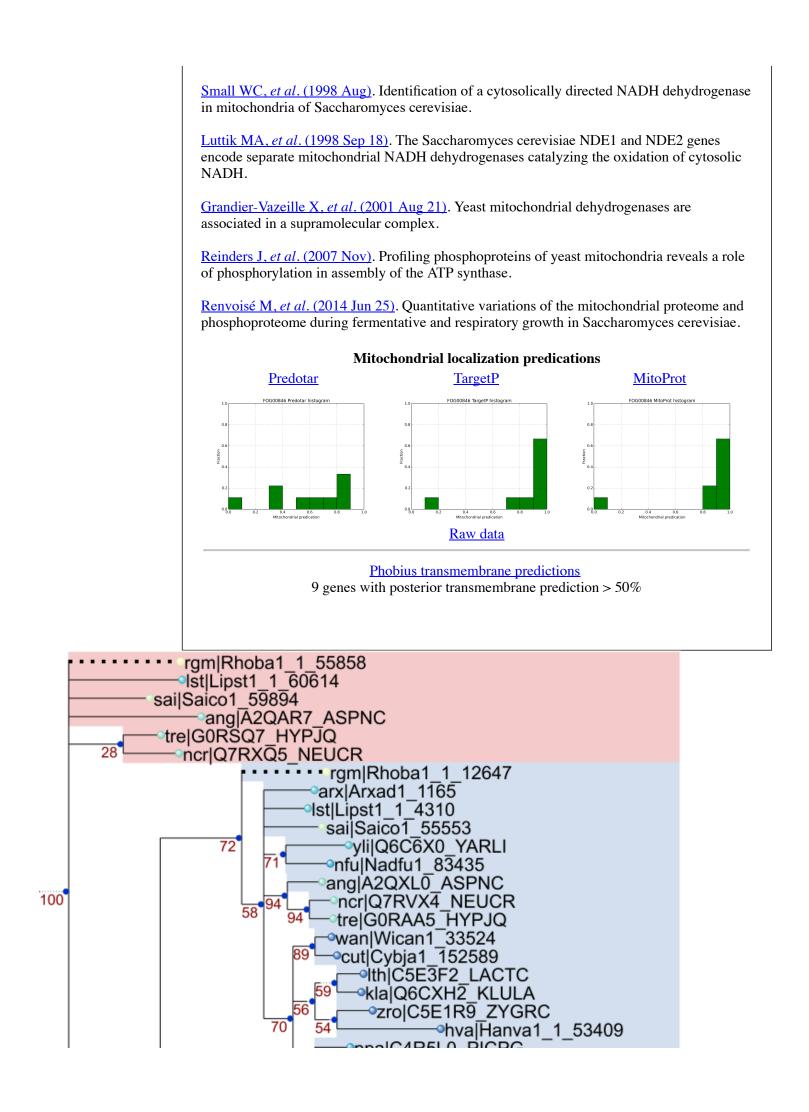


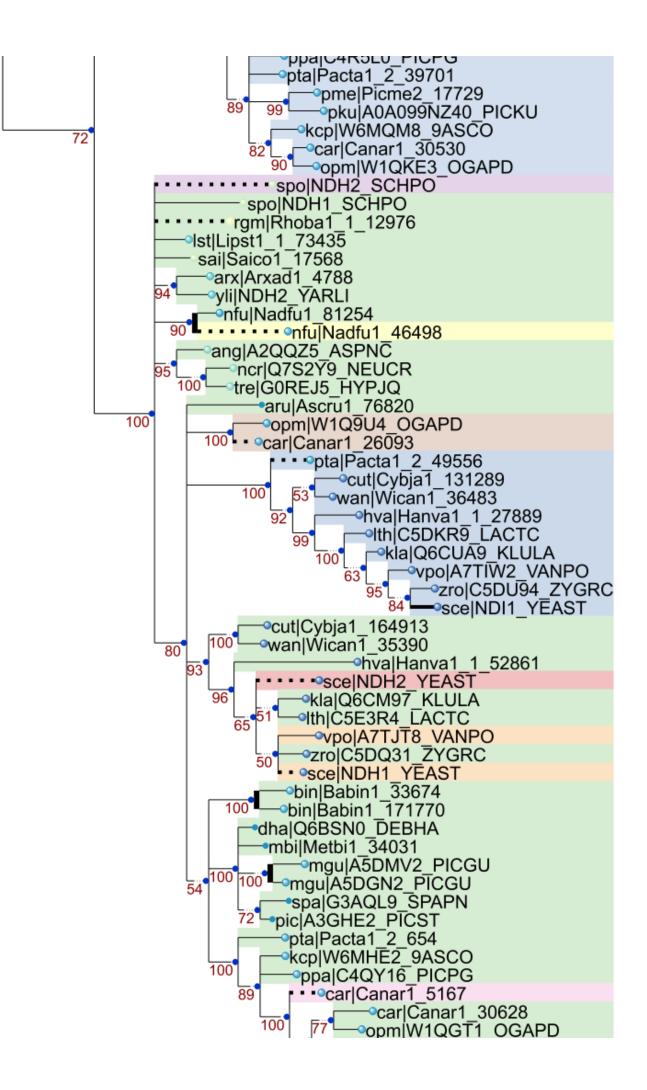












⊢\_\_\_\_| 1.11

### References

Patrícia Carneiro, Margarida Duarte, and Arnaldo Videira. The external alternative NAD(P)H dehydrogenase NDE3 is localized both in the mitochondria and in the cytoplasm of *Neurospora crassa*. *Journal of Molecular Biology*, 368(4):1114–1121, 2007.