

Additional File 12. Mating type locus

Figure S12 Matalpha protein alignments.

Proteins were aligned using Multalin¹⁷. The Alpha box (DNA binding domain) is boxed in the *S. cerevisiae* sequence. UniProtKB references for the aligned proteins are: F2Z6B8 *Yarrowia lipolytica*, Q707Y7.1 *Ogataea angusta*, Q6BNT6 *Debaryomyces hansenii*, G8XZ51 *Millerozyma farinosa*, F1D944 *Candida orthopsisilosis*, F2QXX3 *Komagataella pastoris*, C4Y9N3 *Clavispora lusitaniae*, G8ZQV3 *Torulaspora delbrueckii*, B2G3Y8 *Zygosaccharomyces rouxii*, P0CY06 *Saccharomyces cerevisiae*. The tree was built using ClustalW and visualized with NJ Plot; bootstrap values are indicated at nodes.

Figure S12 MTAL1 homologues

	1	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	168	
<i>A_adeninivorans</i>	MAIKDTTKSNTIKSQRRLKPTLRFQPNCEIKQRFLOPQYAL	TYFADGQSSPGTKTLPPIPSVEELKEAISKL										KIISDHNEAKKRDRPSRANTGSPLKFSSFLAFKTYGG-LHAN-QTEISSYLSAAH							
<i>Y_lipolytica</i>	MHQDEYHIKQRIFSISKAITHIKIKTEHKVYKARTPKKKHPGMARFRQAVPIINQNNNRYLE	SEISFEKLEPVPMPHYPELQKVLDYI										HSRKVPFDGSSGPYP--SSRRQVVNGYVAFKLYHHVPTDLS-AIDITHLLQGPW							
<i>O_angusta</i>	MHFHVLCFFGNTRSYANLCLQKFPPHNDSPKMHALDELNL	SFLTHLPQIPPTTSPPLKRLLETEYGAIDN-										GKDSLNVYAPKKSSRHYQEKKRLNGMFAFRSFYSRNIAYNQKHLSSQLADAH							
<i>D_hansenii</i>	MTSEVA--KVIPKRFITSRTFTSKTGKRSNSNIKEAIDIG-RQRKVSKHQVVFTELPSPSLPTPSDOLQEILIRFNQATRNCAGDS											SMGSQLFDMKVP-T-RKSIRCKTRCLNGMFAFRSFYSRSIYMEHORELSTLLGKLH							
<i>M_farinosa</i>	MTSKRT--KTVPKNLTSRFAVSGSNYCHODRRNRRAAPRSSFASDKLSSSILFLTELPSLSPSPSLELKNIILNFSSANNEFDNDWEOFF-											GLGSDNMSPKKTTTPRRYKPKARRTRNLNGMFAFRSFYSRSRISMVHQRLSSLLGSLH							
<i>C_orthopsilosis</i>	MSAKKGKNNKLIPKDFVTLFRVCRSLQKVTNN--LDRKPAPFKNPKLPIPEPSEELQSLLLDYSLLN--											KRTGENQYSEKRTKSSRSYMSKRIMAFIAFRSFYSRSQVNDYNKQISISRSLSYAH							
<i>K_pastoris</i>	MLPPQPQKSFKGQKPKCSQYAHSTLDNGNHLQHTKHP--YTYQLETPTRTSLQDIESHLIQTETGI											KLRQLQEYKVSQKQRKRRKGQSQRRTSINGFIAFRFTYYSKHLRNATLQQLSCQLGKVH							
<i>C_lusitaniae</i>	MVSSSKSFRGRFRVNRPLFKYQINKDSRTSQQCAKTFKKAALKCOPENSGRLEERWKTFVYILRPPIKTPLPPMILLLEKEIEQEKRKVYSETSEPIGSLWDPSIAHDNEDLLFSI	SRSQWGNIYKKPARETAWSKPLNSMFAFRRAYNSOFGYGLK-QNILSSLASAH										YFSWSSSKPLTGKIKKSKSQPKINSFIAFRSFYSRSRISPPYEQRELSTKLAKLH							
<i>T_delbrueckii</i>	MNTKSFRGFKSNQPLFKYVKEKKGRS--KRPTLQDTQTTSKCIAGKC--IEDQGVNVFMTVSKLIIKIPSPPEVYLRLKIEEKKHLKDDDFNMNILLWDPPVYHGEGYFY	SIADSSLNNLSTGKRGCGSEKPLNSMFAFRRAYNSOFGNGLK-QNILSSLASAH										SIADSSLNNLSTGKRGCGSEKPLNSMFAFRRAYNSOFGNGLK-QNILSSLASAH							
<i>Z_rouxii</i>	MFTSKPRAFIK-NKASKS--YRNTAVSKKLKEKRLAEH--VRPSCFNIIIRPLKDDIQIPVPSSSRFLNKQIHR-											IASGSQNTQFRQFNKTSIKS-EKYYNSCERFRHYSQEGSGGVK-DNYVSSLTREEM							
<i>S_cerevisiae</i>																			
Consensus	
	169	178	188	198	208	218	228	238	248	258	268	278	288	298	308	318	328	335	
<i>A_adeninivorans</i>	KSYPHKH- IADLYCDHFQI --LERKEPFYEWLLKTRTRSECPEENAA-QSVTNSDOCDSMHSHSSNLSNDMHLIPENIPLENME-											GQSISVEELLFSDTFDTAVNSD-SLHTHDTAYDACINYDQSDSELYSSGLWQEFSSDLSSYYHEFL							
<i>Y_lipolytica</i>	RACPDRH- IATKYADHYRS --RGROVAFKTWLLSVT-SPAPEKGLIIOPQTOSEQGYNFS-SSSDOFSSSPHESETLPVSSGPSTPVDPFLIEFDQGIAKEASFTSPCNQNYDLOFSQLQPLSIRVPYFNQYASSNSGLDYDYFQDGSTLDTPSYIDTFTPLH											NREKHHQ- IKSLYAIQYNH --SKSKEPFSRWALEMKLSNLGKFNSDDPPTTRREINVW---SHIVYEDOVFQECSSQY							
<i>O_angusta</i>													KNEPKNA- IANRYAIIEYNS --QHINQDFYDVLFKRGLGLEVEITFPVPIPYTSTVNRNEATFSSRNNEENRVED IYYYIGTV						
<i>D_hansenii</i>													QTEPNKN- IANRYAIIEYNT --RRTNQDFYEWLLCKALSLPLDQVSIPSPSTSTVKNKKHLFTSMNNNNYNAVED YYYRV						
<i>M_farinosa</i>													KEESHDRD- IADQYTFLYNNYLLKSNMNLTFYDWL QMLNLNESONNTPIOKSHVEDSLSKL---KSGTIQDQYLT						
<i>C_orthopsilosis</i>													KNE-NKE-LKRSYALLYNS--SNHNPFPYDWLQTGPKNSKRTSTVTKSFISHWELL---STRNKIEDOVFLDLQ						
<i>K_pastoris</i>													ENEPSRE- IACCYHYOYNE --TGGDESFDIWL YIKLGLLTHEKKSTRSQACKSYIF---KNVEDOVFYD						
<i>C_lusitaniae</i>													HSHPEQQGTHDTFQAQFNF--VKPKCGFVEWVQDQRYERES						
<i>T_delbrueckii</i>													HSHPEQQGTHDTFQAQFNF--VKPKCGFVEWVQDQRYERES						
<i>Z_rouxii</i>																			
<i>S_cerevisiae</i>																			
Consensus	
	..ep.....	IH..%a..%#.	.	Fv..Wl.....	d.....	

