

Using HHsearch to tackle proteins of unknown function – a pilot study with PH domains

David R. Fidler^{1*}, Sarah E. Murphy^{1*}, Katherine Courtis¹, Pantelis Antonoudiou¹, Rana El-Tohamy¹, Jonathan lent¹ and Timothy P. Levine¹

¹Department of Cell Biology, UCL Institute of Ophthalmology, 11-43 Bath Street, London EC1V 9EL, UK

* these authors contributed equally to this work

[¶] to whom correspondence should be addressed

Contact emails Timothy P. Levine: tim.levine@ucl.ac.uk

Supplemental Information

Supplemental Table 1.

Families of PH-like domains classified by different profile--sequence tools

	groupi	# entries	s in PDB	Sample			
Family Name	CDD	SCOP	Pfam family	groups full		structure	
classical PH	classical PH (cPH)	РН	PH, PH2-13, IQ_Sec7 Mcp5[Num1]	115	1=96 115 2=59 3=26 11 1=10		
FERM-C	FERM-C	FERM (3 rd domain)	FERM_C	14	3pvl_A 2j0j A		
РТВ	PTB, PID, IRS	РТВ	PTB, PID, PID2, ICAP1integrinBP	23	4wj7_A 2cy5 A		
RBD	Ran-BP1/BD WH1 YRB1	RanBD, Enabled/VASP hom.	WH1 Bon PD1	26	26 2=21 26 1=23		
	Dcp1	DCP1	DCP1		<u>2=22</u> 3	1ddw A 1q67 A	
AVO1 (Torc2 subunit)	PH_Avo1	x	SIN1		2	3voq_A	
Bact-PH	bPH	BPHL	bPH1-6		2	3hsa A	
CARM1-PH	CARM1	х	CARM1		2	2oqb A	
CARMIL	Х	Х	CARM1		1	4k17_A	
DOK-PTB	PH_DOK1,2,3	РТВ	PTB, IRS		7		
GLUE (Vps36)	Vps36	VPS36 N-term-like	Vps36_ESCRT-II	:	3	2cay_A	
GRAM	GRAM	GRAM	GRAM, BBL5(DM16)	4	4	1lw3_A	
hSac2-PH	Х	Х	hSac2		1		
ICIn	IcIn	х	Voldacs	:	3		
ISP3	Х	Х	Х	2		4chj_A	
Myosin1c-TH1	Myosin_TH1 (PH-like)	Х	Myosin_TH1	1		4r8g_A	
Necap	Necap1	Necap1 N-term	DUF1681	1		1tqz_A	
NF1-PH	PH-neurofibromin1 =NF1	х	х	2		3pg7_A	
OCRL1 & Inpp5b PH	PH_OCRL-like	х	х	2		2kig_A & 2kie A	
POB3N-PH	Х	х	SSrecog	2		4khb_D(N)	
Pre-BEACH	PH_BEACH	Pre-BEACH	PH_BEACH	:	2		
RGS-RhoGEF-PH	Х	Х	Х		1	4gou_A	
Rpn13	Rpn13	х	Proteasome Rpn13	;	3	2r2y_A	
RTT106C	Rtt106	х	Rtt106	;	3	3fss_A (C)	
SEC3-PH	PH-Sec3_like	х	Sec3- PIP2 bind	1	2	3hie_A	
Sharpin-PH	PH_Sharpin	Х	Sharpin_PH		1	4emo_A	
SNX17(FERM-C)	FERM-likeC_SNX17	Х	Х		1		
split (a-syntrophin)	PHsplit_syntrophin	a1syntrophin	PH	3		2adz_A	
SPT16D	FACT(SPT16/CDC68)	Х	SSrecog	1		4khb_A	
Spt16M	PH2_SSRP1-like rpt1	SSRP1-like	SSrecog	3		4kho_A (N)	
SSrecog	PH2_SSRP1-like rpt2	FACT Pob3M	SSrecog	1		3fss_A(C)	
TFIIH	TFIIH	TFIIH	PH_TFIIH	2		1pfj_A	
USP37-PH	PH_USP37_like	Х	UCH_N	1		3u12_A	
ZF21-PH	Х	Х	ZFYVE21_C	1		2rrf_A	
Subfamilies:	>350	15	39 + 1 wrong	34	39		
WRONG (<i>i.e.</i> not PH-like)			DUF1126 (EF hand)				

Supplemental Table 1. Classifications of PH-like domains.

Information on our classification of 39 PH-like families (see Figure 2), starting with the major groups that contain overlapping families in some classifications. Our classification is compared to those in Conserved Domain Database (CDD) at NCBI, the Structural Classification of Protein (SCOP) database at <u>scop.berkeley.edu/</u>, and Pfam 30.0. RED indicates PH-like families that were missing from the PH-like clan. "X" indicates that the family was missing completely. The penultimate column shows the number of PH-like domains for each family in PDB (using a version non-redundant at 70% identity), both for grouped families (excluding extensive overlap, n=34) or the full list (n=39). The final column shows the one structure we chose to represent the family for searches in PSI-BLAST and HHsearch. Where the structure includes more than one PH-like domain, the segment used is indicated. The bottom lines show the total number of PH-like families in each classification, and information on one domain wrongly assigned as PH-like in Pfam.

Supplemental Table 2: details of 91 yeast PH-like domains

A systematic proteins protein gene name residues 1 2 3 4 5 Known PH-like proteins gene name length residues sup- fissi family rebeat rebeat rebeat rebeat Askt0p Ygr097 1146 477-725 cPH cPH ePH 92.4 cPH[1/23] 99.9 h Atg26p-1 Ylr189c 1198 675-682 gram gram - 98.1 gram 97.6 h Avorp Yol078w 1176 1065-1172 - + - 100.0 PTB[Avo1] 100.0 c Bem3p Yp115c 1128 630-745 CPH CPH CPH 99.1 cPH[spiil 99.9 h Boitp Yb1085w 380 778-895 CPH CPH CPH 99.1 cPH[1/2/3] 99.9 h Cd124p Yp1092w 1448 1300-1420 CPH CPH 99.1 cPH[1/2/3]	!	1	1	!	1						
Systematic protein proteins residues plan PDB-to-yeast plan yeast plan not plan yeast plan not plan yeast plan not plan Ask10p YUr097 1146 477.725 CPH CPH CPH 2PH CPH 2PH <	Λ			method 1 2 3 4 5						5	
NAME Family -io-PDB Ask10p Yg097 1146 477.725 CPH CPH CPH 2PH 2PH[1/3] 99.9 h Alg26p-1 Ylr188c 1188 165.7482 gram gram 92. CPH[1/2/3] 99.9 h Avg26p-2 Ylr188c 1188 575-682 gram gram - 88.1 gram 97.6 h Bem2p Yer155c 2167 1780-1957 CPH CPH PH 92. CPH[1/2/3] 99.9 h Boitp Yer155c 2167 170-1467 BPH CPH CPH 99.5 CPH[1/2/3] 99.9 h Boitp Yer032 1448 130-1420 CPH CPH CPH 99.5 CPH[1/2/3] 99.9 h Cd24p Yin02x 1448 130-1467 BPH CPH CPH 98.1 CPH[1/2/3] 99.9 h Cd24p Yalotw 844 840.82	A			residues				PDB-to-veast			note
Aug26p-1 Yir189c 1188 165-340 CPH CPH GPH		gene name	length			SMART					
Aig26p-2 YIr189c 1198 575-682 gram gram - 98.1 gram 97.6 h Avo1p Y01078w 1176 1065-1172 - + - 100. PTE[Avo1] 100.0 c Bem2p Yer155c 1276 780-1957 CPH (2PH 6PH 6PH 6PH 6PH 6PH 6PH 99.5 CPH[1073] 99.9 h Boitp Y1015c 1128 630-745 CPH CPH 6PH 99.5 CPH[1723] 99.9 h Boitp Yr0124v 1448 1300-1420 CPH CPH 6PH 99.8 PF(1723) 99.9 h Cla4p Y1024v H448 1300-1420 CPH CPH 6PH 6PH 99.8 PH[1723] 99.9 h Cla4p Y1024v H44 810-6444 CPH CPH CPH 6PH CPH SPI CPH[1723] 99.9 Th Cla4p </td <td></td> <td>, v</td> <td></td> <td></td> <td></td> <td></td> <td>-</td> <td>-</td> <td></td> <td></td> <td></td>		, v					-	-			
Avo1p Yol078w 1176 1065-1172 - + - 100. PTB[Avo1] 100.0 c Bem2p Yer155c 2167 1780-1957 CPH CPH CPH 99.1 CPH[I2/3] 99.8 h Boi1p Yb0155c 2167 1370-1476 CPH CPH CPH 99.5 CPH[1/2/3] 99.9 h Boi2p Yer114c 1040 768-889 CPH CPH CPH OPH 99.5 CPH[1/2/3] 99.9 h Cal2p Yer114c 1040 768-889 CPH CPH CPH OPH 99.6 CPH[1/2/3] 99.9 h Cal2ap Yn128w 1060 74-214 CPH CPH CPH 98.1 CPH[1/2/3] 99.9 h Cal2ap Yn128w 842 60-183 CPH CPH CPH OPH DP1 Dop1 100.0 C Exo84p Ybr102c 753 352-443 <td< td=""><td><u> </u></td><td></td><td></td><td></td><td>-</td><td></td><td></td><td></td><td></td><td></td><td></td></td<>	<u> </u>				-						
Instruct Instruct Instruct Instruct Instruct Bern2p Yer155 216 1780-1957 CPH CPH CPH 92.7 CPH[I]/2/3 99.8 h Boilp Yer1145 1126 630-745 CPH CPH CPH 94.9 Sch[I]/2/3 99.9 h Boilp Yer114c 1040 769-889 CPH CPH CPH 99.5 CPH[I]/2/3 99.9 h Bud4p Yir052w 1448 1300-1467 BPH BPH BPA 98.0 CPH[I]/2/3 99.9 h Cat20p-2* Ynl29w 1060 74-214 CPH CPH CPH 98.1 CPH[1/2/3] 99.7 h Cat22p Yal041w 854 484-682 CPH CPH CPH 98.1 CPH[1/2/3] 99.7 h Cat22p Yal041w 854 484-680 C CPH CPH 98.1 CPH[1/2/3] 99.7 h 0					gram	gram	-				h
Bem3p YpI115c 1128 630-745 CPH CPH CPH OPL	Avo1p				-		-			100.0	С
Boitp Ybi085w 980 778-895 cPH cPH cPH ePH PH PH PH PH 99.5 cPH[1/2/3] 99.9 h Boltp Yer032 2167 1370-1467 PPH PPH PPH 98.1 cPH[1/2/3] 99.9 h Cat120p-2* Yn022w 1448 1300-1420 cPH cPH cPH 98.1 cPH[1/2/3] 99.9 h Cdc24p Yal041w 854 484-682 cPH cPH oPH 98.1 cPH[1/2/3] 99.9 h Cdc24p Yal041w 854 484-82 cPH cPH 98.1 cPH[1/2/3] 99.9 h Dc1p Yol149w 231 42-223 cPH cPH oPH 98.6 cPH[1/2] 99.8 cP Ira1p Ybr02c 753 352-434 cPH cPH 96.6 cPH[NF11 100.0 h Lam5p Yol081w 3072 2205-2301	1										
Boi2p Yer114c 1040 769-889 CPH cPH CPH PPH 99.8 PFe 93.4 c Bud4p Yip032w 1448 1300-1420 CPH CPH 98.8 PFe 93.4 c Caf120p-2° Yni278w 1060 74-214 CPH CPH 98.1 CPH[1/2/3] 99.9 h Caf24p Yni278w 1060 74-214 CPH CPH 98.1 CPH[1/2/3] 99.9 h Cla4p Yni28w 842 60-183 CPH CPH 98.1 CPH[1/2/3] 99.9 h Dcp1p Yol149w 231 42-223 + Dcp1 Dop1 Dop.1 100.0 c F Far1p Yb137c 830 444-540 - - 99.9 CPH[1/2] 79.1 h Iratp Yb140c 3079 2205-2301 - - 99.9 CPH[1/E] 100.0 h Lam4p Yn0301	Bem3p		-	630-745	-		-				
Bph1p Ycr032 2167 1370-1467 BPH BPH BPH PP 98.8 Pre- 93.4 c Buddap Yin/278w 1060 74-214 CPH CPH PCH 98.9 CPH[1/2/3] 99.9 n Cdc24p Yal041w 854 484-682 CPH CPH PCH 98.1 CPH[1/2/3] 99.9 n Dcq1p Yal041w 854 484-682 CPH CPH 98.9 CPH[1/2/3] 99.9 n Dcq1p Yol149w 231 422-23 + Dcp1 D0.0 Dcp1 100.0 n Ira1p Yb1702c 753 352-443 CPH CPH 98.6 CPH[1/2] 99.8 c Ira1p Yb1702c 630 244-540 - - - 99.9 CPH[1/2] 100.0 n Lam5p Yb1080c 1345 535-661 gram gram - 97.3 gram 96.6.4	Boi1p						-			99.9	
Bud4p Yjr092w 1448 1300-1420 CPH CPH CPH PH PH<	Boi2p			769-889	-	cPH		99.5		99.9	h
Caf120p-2* Yn1278w 1060 74-214 CPH CPH CPH 93.1 CPH[1/3] 99.9 h Cld24p Yn128w 842 60-183 CPH CPH PCH 98.1 CPH[1/2/3] 99.9 h Dcp1p Y0149w 231 42-223 + DCP1 DCp1 100. Dcp1 100.0 c Exo84p Ybr102c 753 352-443 CPH PCH PCH 98.6 CPH[1/2/3] 99.9 h Ira1p Yb1140c 3092 2204-2302 - - - 99.9 CPH[2] 79.1 h Lam6p Yh040c 1345 535-661 gram gram - 97.7 gram 96.6 h Lam6p Yh072w 693 170-270 gram gram - 97.7 gram 96.4 h Lam6p Yh070 665 462-600 - - 97.1 Gram 99.9	Bph1p	Ycr032	2167	1370-1467	BPH	BPH	BPH	99.8	pre-	93.4	С
Cdc24p Yal041w 854 484-682 CPH Dpp1	Bud4p	Yjr092w	1448	1300-1420	cPH	cPH	cPH	98.9	cPH[1/2/3]	99.9	h
Cla4p Ynl298w 842 60-183 CPH CPH CPH PD PH PH PB CPH[1/2] PD P	Caf120p-2 °	Ynl278w	1060	74-214	cPH	cPH	cPH	93.1	cPH[1/3]	99.9	h
Dcp1p Yol149w 231 42-223 + Dcp1 Dcp1 100. Dcp1 100.0 c Exo84p Ybr102c 753 352-443 cPH cPH cPH cPH cPH Sec cPH[12] 79.1 h Ira1p Ybr140c 3092 2204-2302 - - 99.9 cPH[NF1] 100.0 h Ira2p Yol081w 3079 2205-2301 - - 99.9 cPH[NF1] 100.0 h Lam5p Yl042c 674 223-305 gram gram - 97.3 gram 96.6 h Lam6p Yl072w 693 170-270 gram gram - 97.7 gram 96.4 h Lam6p Yl072w 693 170-270 gram gram - 97.5 gram 98.0 h Lmo1p Yl0072w 693 170-270 gram gram gram 97.5 gram </td <td>Cdc24p</td> <td>Yal041w</td> <td>854</td> <td>484-682</td> <td>cPH</td> <td>cPH</td> <td>cPH</td> <td>98.1</td> <td>cPH[1/2/3]</td> <td>99.7</td> <td>h</td>	Cdc24p	Yal041w	854	484-682	cPH	cPH	cPH	98.1	cPH[1/2/3]	99.7	h
Exo84p Ybr102c 753 352-443 CPH CPH CPH PH 98.6 CPH[1/2] 99.8 c Far1p Yjl157c 830 444-540 - - - 99.9 CPH[2] 79.1 h Ira1p Ybr140c 3092 2204-2302 - - - 99.9 CPH[2] 79.1 h Ira1p Ybr140c 3073 2205-2301 - - - 99.9 CPH[NF1] 100.0 h Lam4p Yhr080c 1345 535-661 gram gram - 97.3 gram 96.6 h Lam5p Ylf072w 693 170-270 gram gram 97.7 gram 96.4 h Last7p Vr0181 633 21-134 RBD	Cla4p	Ynl298w	842	60-183	cPH	cPH	cPH	98.9	cPH[1/2/3]	99.9	h
Far1pYjl157c830444-54095.2CPH[2]79.1hIra1pYbr140c30022204-230299.9CPH[NF1]100.0hIra2pYol081w30792205-230199.9CPH[NF1]100.0hLam4pYhr080c1345535-661gramgram-97.3gram96.6hLam5pYl1042c674223-305gramgram-97.7gram96.6hLam5pYl1072w693170-270gramgram-97.7gram96.4hLas17pYor18163321-134RBDRBDRBD99.5RBD[2]100.0cLot5pYkl183w17150-190ICInICIn99.7CPH[48]96.3hMyo3pYkl129c1272807-91199.7CPH[478]96.3hMyo3pYkl129c1278808-91199.7CPH[478]96.3hNum1pYdr15027482572-2687CPHCPHCPH98.8CPH[17/2]99.9hNu2p2Yir335w720602-718RBDRBDRBD99.9RBD[2]100.0hNu2p2Yir335w720602-718RBDRBDRBDP9.9RPH[17/2]99.9hOpy1p-2 <td< td=""><td>Dcp1p</td><td>Yol149w</td><td>231</td><td>42-223</td><td>+</td><td>Dcp1</td><td>Dcp1</td><td>100.</td><td>Dcp1</td><td>100.0</td><td>с</td></td<>	Dcp1p	Yol149w	231	42-223	+	Dcp1	Dcp1	100.	Dcp1	100.0	с
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	Exo84p	Ybr102c	753	352-443	cPH	cPH	cPH	98.6	cPH[1/2]	99.8	С
Ira2p Yol081w 3079 2205-2301 - - - 99.9 cPH[NF1] 100.0 h Lam4p Yhr080c 1345 535-661 gram gram - 97.3 gram 96.6 h Lam5p Yl1042c 674 223-305 gram gram - 97.7 gram 96.4 h Lam5p Yl072w 693 170-270 gram gram - 97.7 gram 96.4 h Las17p Yor181 633 21-134 RBD RBD PBD 99.5 RBD[2] 100.0 h, Lot5p Yk1183w 171 50-190 ICin ICin - 99.7 CPH[4r8g] 96.3 h Myo5p Yk1129c 1272 807-911 - - - 99.7 CPH[4r8g] 96.3 h Nug2p Yk129c 1272 807-911 - - - 99.7 CPH[4r8g]	Far1p	Yjl157c	830	444-540	-	-	-	95.2	cPH[2]	79.1	h
Lam4p Yhr080c 1345 535-661 gram gram - 97.3 gram 96.6 h Lam5p Ylf042c 674 223-305 gram gram - 97.9 gram 95.9 h Lam6p Ylf072w 693 170-270 gram gram - 97.7 gram 96.4 h Las17p Yorl81 633 21-134 RBD RBD RBD 99.5 RBD[2] 100.0 h, Lot5p Ykl183w 171 50-190 ICIn ICIn - - 99.7 Gram 98.0 h Mdr1p Ygr100 950 40-157 gram gram gram 97.5 gram 98.0 h Myo3p Ykl129c 1272 807-911 - - 99.7 CPH[4r8g] 96.3 h Num1p Ydr150 2748 2572-2687 CPH CPH CPH 98.8 CPH[1/2] <td>lra1p</td> <td>Ybr140c</td> <td>3092</td> <td>2204-2302</td> <td>-</td> <td>-</td> <td>-</td> <td>99.9</td> <td>cPH[NF1]</td> <td>100.0</td> <td>h</td>	lra1p	Ybr140c	3092	2204-2302	-	-	-	99.9	cPH[NF1]	100.0	h
Lam5p YII042c 674 223-305 gram gram - 97.9 gram 95.9 h Lam6p YI072w 693 170-270 gram gram - 97.7 gram 96.4 h Lam5p Y0r181 633 21-134 RBD RBD RBD 99.5 RBD[2] 100.0 c Lmo1p YI007c 665 462-600 - - 97.1 CPH[3] 100.0 h, Lot5p Yk112sw 171 50-190 ICIn ICIn - 99.9 pCIn 99.8 h Myo3p Yk12sc 1272 807-911 - - 99.7 CPH[4r8g] 96.3 h Num1p Ydr150 2748 2572-2687 CPH CPH CPH 98.1 CPH[4r8g] 96.3 h Nup2p Yr091c 770 125-268 CPH CPH PH 98.1 CPH[1/2] 99.9 h	Ira2p	Yol081w	3079	2205-2301	-	-	-	99.9	cPH[NF1]	100.0	h
Lam6p YIr072w 693 170-270 gram gram - 97.7 gram 96.4 h Las17p Yor181 633 21-134 RBD RBD RBD 99.5 RBD[2] 100.0 c Lmo1p YII007c 665 462-600 - - - 97.1 cPH[3] 100.0 h,d Lot5p Ykl183w 171 50-190 ICIn ICIn - 99.9 pICIn 99.8 h Mdr1p Ygr100 950 40-157 gram gram gram 97.2 gram 98.0 h Myo5p Ymr109 1219 808-911 - - - 99.8 cPH[478] 96.3 h Num1p Ydr150 Z748 2572-2687 CPH cPH GPH 98.1 CPH[1/2] 99.9 h Nup2p Ypr091c 770 125-268 CPH CPH PH 99.8 cPH[1/2] <td>Lam4p</td> <td>Yhr080c</td> <td>1345</td> <td>535-661</td> <td>gram</td> <td>gram</td> <td>-</td> <td>97.3</td> <td>gram</td> <td>96.6</td> <td>h</td>	Lam4p	Yhr080c	1345	535-661	gram	gram	-	97.3	gram	96.6	h
Las17p Yor181 633 21-134 RBD RBD RBD RBD RBD RBD[2] 100.0 c Lmo1p Yll007c 665 462-600 - - - 97.1 cPH[3] 100.0 h, Lot5p Ykl183w 171 50-190 ICIn ICIn - 99.7 cPH[3] 100.0 h, Mdr1p Ygr100 950 40-157 gram gram gram 97.5 gram 98.0 h Myo5p Ymr109 1219 808-911 - - - 99.8 cPH[4/8g] 96.3 h Num1p Ydr150 Zr48 2572-2687 CPH cPH PH 98.1 cPH[1/2] 99.9 h Nup2p Yir035w 720 602-718 RBD cPH[1/2/3] 99.9 h Opy1p-2	Lam5p	Yfl042c	674	223-305	gram	gram	-	97.9	gram	95.9	h
Lmo1p Yillo07c 665 Ab2_c Nbb <t< td=""><td>Lam6p</td><td>Ylr072w</td><td>693</td><td>170-270</td><td>gram</td><td>gram</td><td>-</td><td>97.7</td><td>gram</td><td>96.4</td><td>h</td></t<>	Lam6p	Ylr072w	693	170-270	gram	gram	-	97.7	gram	96.4	h
Lot5p Yki183w 171 50-190 ICin ICin ICin — 99.9 pICin 99.8 h Mdr1p Ygr100 950 40-157 gram gram gram 97.5 gram 98.0 h Myo3p YkI129c 1272 807-911 – – 99.7 CPH[4r8g] 96.3 h Myo5p Ymr109 1219 808-911 – – 99.7 CPH[4r8g] 97.2 h Num1p Ydr150 2748 2572-2687 CPH CPH PH 98.1 CPH[1/2] 99.9 h Nyp2p Ypr091c 770 125-268 CPH CPH CPH #N/A CPH[1/2] 99.9 h Opy1p-1 Ybr129c 328 213-320 CPH CPH CPH 99.1 CPH[1/2/3] 99.9 h Osh1p Yar042 1188 280-383 CPH CPH CPH 99.1 CPH[1/2/3]	Las17p	Yor181	633	21-134	RBD	RBD	RBD	99.5	RBD[2]	100.0	С
Mdr1p Ygr100 950 40-157 gram gram gram gram gram 97.5 gram 98.0 h Myo3p Ykl129c 1272 807-911 - - - 99.7 CPH[4r8g] 96.3 h Myo5p Ymr109 1219 808-911 - - - 99.8 CPH[4r8g] 97.2 h Num1p Ydr150 2748 2572-2687 CPH CPH CPH CPH CPH CPH CPH[1/2] 99.9 h Nup2p Yr335w 720 602-718 RBD RBD RBD P8.9 8.1 CPH[1/2] 99.9 h Opy1p-1 Ybr129c 328 65-155 CPH CPH CPH 99.8 CPH[1/2/3] 99.9 h Osh1p Yar042 1188 280-383 CPH CPH CPH 99.3 CPH[1/2/3] 99.9 h Osh2p Ydr073 996	Lmo1p	YII007c	665	462-600	-	-	-	97.1	cPH[3]	100.0	h,o
Myo3p Ykl129c 1272 807-911 - - - 99.7 CPH[4r8g] 96.3 h Myo5p Ymr109 1219 808-911 - - - 99.8 CPH[4r8g] 97.2 h Num1p Ydr150 2748 2572-2687 CPH CPH PH 99.9 RBD[2] 100.0 h Nup2p Ylr335w 720 602-718 RBD RBD RBD 99.9 RBD[2] 100.0 h Ny12p Ypr091c 770 125-268 CPH CPH CPH 99.9 RBD[2] 100.0 h Opy1p-1 Ybr129c 328 65-155 CPH CPH CPH 99.8 CPH[1/2/3] 99.9 h Osh1p Yar042 1188 280-383 CPH CPH CPH 99.1 CPH[1/2/3] 99.9 h Osh2p Ydl019c 1283 288-390 CPH CPH PH 99.1	Lot5p	Ykl183w	171	50-190	ICIn	ICIn	-	99.9	pICIn	99.8	h
Myo5p Ymr109 1219 808-911 - - - 99.8 CPH[4r8g] 97.2 h Num1p Ydr150 2748 2572-2687 CPH CPH CPH 98.1 CPH[1/3] 99.9 h Nup2p Yir335w 720 602-718 RBD RBD RBD 98.9 RBD[2] 100.0 h Nvj2p Ypr091c 770 125-268 CPH CPH CPH PH 99.8 CPH[1/2/3] 99.9 h Opy1p-1 Ybr129c 328 65-155 CPH CPH CPH 99.8 CPH[1/2/3] 99.9 h Osh1p Yar042 1188 280-383 CPH CPH CPH 99.1 CPH[1/2/3] 99.9 h Osh2p Ydl019c 1283 288-390 CPH CPH CPH 99.1 CPH[1/2/3] 99.9 h Osh2p Ydl019c 1283 282-318 CPH CPH CPH </td <td>Mdr1p</td> <td>Ygr100</td> <td>950</td> <td>40-157</td> <td>gram</td> <td>gram</td> <td>gram</td> <td>97.5</td> <td>gram</td> <td>98.0</td> <td>h</td>	Mdr1p	Ygr100	950	40-157	gram	gram	gram	97.5	gram	98.0	h
Num1p Ydr150 2748 2572-2687 CPH CPH CPH 98.1 CPH[1/3] 99.9 h Nup2p Yir335w 720 602-718 RBD RBD RBD PBD 99.9 RBD[2] 100.0 h Nyj2p Ypr091c 770 125-268 CPH CPH CPH PH #N/A CPH[1/2] 99.9 h Opy1p-1 Ybr129c 328 65-155 CPH CPH CPH PH 98.8 CPH[1/2/3] 99.9 h Opy1p-2 Ybr129c 328 213-320 CPH CPH CPH 99.8 CPH[1/2/3] 99.9 h Osh1p Yaf042 1188 280-383 CPH CPH CPH 99.3 CPH[1/2/3] 99.9 h Osh2p Ydl019c 1283 288-390 CPH CPH CPH 99.3 CPH[1/2/3] 99.9 h Osh2p Ydl019c 1283 288-390 <td< td=""><td>МуоЗр</td><td>Ykl129c</td><td>1272</td><td>807-911</td><td>-</td><td>-</td><td>-</td><td>99.7</td><td>cPH[4r8g]</td><td>96.3</td><td>h</td></td<>	МуоЗр	Ykl129c	1272	807-911	-	-	-	99.7	cPH[4r8g]	96.3	h
Nup2p YIr335w 720 602-718 RBD RBD RBD RBD P9.9 RBD[2] 100.0 h Nvj2p Ypr091c 770 125-268 cPH cPH cPH ePH dPH cPH cPH ePH dPH <td>Myo5p</td> <td>Ymr109</td> <td>1219</td> <td>808-911</td> <td>-</td> <td>-</td> <td>-</td> <td>99.8</td> <td>cPH[4r8g]</td> <td>97.2</td> <td>h</td>	Myo5p	Ymr109	1219	808-911	-	-	-	99.8	cPH[4r8g]	97.2	h
Nyi2p Ypr091c 770 125-268 CPH CPH CPH #N/A CPH[1/2] 99.9 h Opy1p-1 Ybr129c 328 65-155 CPH CPH CPH PB.8 CPH[1/2/3] 99.9 h Opy1p-1 Ybr129c 328 213-320 CPH CPH CPH 99.8 CPH[1/2/3] 99.9 h Osh1p Yar042 1188 280-383 CPH CPH CPH 99.1 CPH[1/2/3] 99.9 h Osh2p Ydl01pc 1283 288-390 CPH CPH CPH 99.3 CPH[1/2/3] 99.9 h Osh2p Ydl01pc 1283 288-390 CPH CPH 99.3 CPH[1/2/3] 99.9 h Osh3p Yhr073 996 20-318 CPH CPH CPH 99.4 CPH[1/2/3] 99.9 h Plo5p-1 Yml069 552 01-111 + - - 94.2 P	Num1p	Ydr150	2748	2572-2687	cPH	cPH	cPH	98.1	cPH[1/3]	99.9	h
Nyi2p Ypr091c 770 125-268 CPH CPH CPH #N/A CPH[1/2] 99.9 h Opy1p-1 Ybr129c 328 65-155 CPH CPH CPH 98.8 CPH[1/2/3] 99.9 h Opy1p-2 Ybr129c 328 213-320 CPH CPH CPH 99.8 CPH[1/2/3] 99.9 h Osh1p Yar042 1188 280-383 CPH CPH CPH 99.1 CPH[1/2/3] 99.9 h Osh2p Ydl01pc 1283 288-390 CPH CPH CPH 99.3 CPH[1/2/3] 99.9 h Osh3p Yhr073 996 220-318 CPH CPH CPH 99.3 CPH[1/2/3] 99.9 h Osh3p Yhr073 996 220-318 CPH CPH PH 99.5 CPH[1/2/3] 99.9 h Plotp Yh1069 552 102-232 + - - 90.4<	Nup2p	Ylr335w	720	602-718	RBD	RBD	RBD	99.9	RBD[2]	100.0	h
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$		Ypr091c	770	125-268	cPH	cPH	cPH	#N/A		99.9	h
Opy1p-2 Ybr129c 328 213-320 CPH CPH CPH 99.8 CPH[1/2/3] 99.9 h Osh1p Yar042 1188 280-383 CPH CPH CPH P9.1 CPH[1/2/3] 99.9 h Osh2p Ydl019c 1283 288-390 CPH CPH CPH 99.1 CPH[1/2/3] 99.9 h Osh2p Ydl019c 1283 288-390 CPH CPH CPH 99.5 CPH[1/2/3] 99.9 h Osh3p Yhr073 996 220-318 CPH CPH CPH 99.5 CPH[1/2/3] 99.9 h, Plc1p Ypl268w 869 122-217 + 812 CPH 99.4 cPH[1/2/3] 99.9 h, Pob3p-1 Yml069 552 01-111 + - - 90.3 N100.0 c Pob3p-3 Yml069 552 247-347 + + + 96.5 Strecog </td <td></td> <td>Ybr129c</td> <td>328</td> <td>65-155</td> <td>cPH</td> <td>cPH</td> <td>cPH</td> <td>99.8</td> <td></td> <td>99.9</td> <td>h</td>		Ybr129c	328	65-155	cPH	cPH	cPH	99.8		99.9	h
Osh1p Yar042 1188 280-383 CPH CPH CPH 99.1 CPH[1/2/3] 99.9 h Osh2p Ydl019c 1283 288-390 CPH CPH CPH CPH 99.3 CPH[1/2/3] 99.9 h Osh3p Yhr073 996 220-318 CPH CPH CPH 99.5 CPH[1/2/3] 99.9 h Plc1p Ypl268w 869 122-217 + 812 CPH 99.4 CPH[1/2/3] 99.9 h Pob3p-1 Yml069 552 01-111 + - - 94.2 Pob3N 100.0 c Pob3p-2 Yml069 552 01-232 + - - 100. Pob3N 100.0 c Pob3p-3 Yml069 552 247-347 + + + 96.9 SSrecog 100.0 c Pob3p-4 Yml069 552 365-470 Rtt106 Rtt106 99.5 <			328		cPH	cPH	cPH	99.8		99.9	h
Osh2p Ydl019c 1283 288-390 cPH cPH cPH 99.3 cPH[1/2/3] 99.9 h Osh3p Yhr073 996 220-318 cPH cPH cPH PH 99.5 cPH[1/2/3] 99.9 h Plc1p Ypl268w 869 122-217 + 812 cPH 99.4 cPH[1/2/3] 99.9 h, n Pob3p-1 Yml069 552 01-111 + - - 94.2 Pob3N 100.0 c Pob3p-2 Yml069 552 010-232 + - - 94.2 Pob3N 100.0 c Pob3p-3 Yml069 552 247-347 + + + 96.9 SSrecog 100.0 c Pob3p-4 Yml069 552 365-470 Rtt106 Rtt106 Rtt106 99.7 Rtt106C 100.0 c Psy2p Ynl201c 858 28-127 + PH RBD <t< td=""><td></td><td>Yar042</td><td>1188</td><td></td><td></td><td></td><td>cPH</td><td></td><td></td><td></td><td>h</td></t<>		Yar042	1188				cPH				h
Osh3p Yhr073 996 220-318 CPH CPH CPH 99.5 CPH[1/2/3] 99.9 h Plc1p Ypl268w 869 122-217 + 812 CPH 99.4 CPH[1/2/3] 99.9 h,t Pob3p-1 Yml069 552 01-111 + - - 94.2 Pob3N 100.0 c Pob3p-2 Yml069 552 102-232 + - - 100. Pob3N 100.0 c Pob3p-3 Yml069 552 247-347 + + + 96.9 Ssrecog 100.0 c Pob3p-4 Yml069 552 365-470 Rtt106 Rtt106 99.7 Rtt106C 100.0 c Psy2p Ynl201c 858 28-127 + PH RBD 99.5 RBD[2] 100.0 h Rgn19 Ygr070 1155 682-822 + PH cPH 99.1 cPH[1/2/3]		Ydl019c			cPH		cPH	99.3			h
Pic1p Ypi268w 869 122-217 + 812 CPH 99.4 CPH[1/2/3] 99.9 h,t Pob3p-1 Ymi069 552 01-111 + - - 94.2 Pob3N 100.0 c Pob3p-2 Ymi069 552 102-232 + - - 100. Pob3N 100.0 c Pob3p-3 Ymi069 552 247-347 + + + 96.9 SSrecog 100.0 c Pob3P-4 Ymi069 552 247-347 + + + 96.9 SSrecog 100.0 c Pob3P-4 Ymi069 552 247-347 + + + 96.9 SSrecog 100.0 c Psy2p Yni201c 858 28-127 + PH RBD 99.5 RBD[2] 100.0 h,d Rom1p Ygr070 1155 682-822 + PH CPH 99.1 CPH[1/2/3]											h
Pob3p-1 Yml069 552 01-111 + - - 94.2 Pob3N 100.0 c Pob3p-2 Yml069 552 102-232 + - - 100. Pob3N 100.0 c Pob3p-3 Yml069 552 102-232 + - - 100. Pob3N 100.0 c Pob3p-3 Yml069 552 247-347 + + + 96.9 Ssrecog 100.0 c Pob3p-4 Yml069 552 245-470 Rtt106 Rtt106 99.7 Rtt106C 100.0 c Psyzp Ynl201c 858 28-127 + PH RD 99.5 RBD[2] 100.0 h Rgc1p Ypr115 1083 493-587 cPH cPH ePH 99.1 cPH[1] 99.8 h Rom1p Ygr070 1155 682-822 + PH cPH 99.1 cPH[1/2/3] 99.5		Ypl268w									h,o
Pob3p-2 Yml069 552 102-232 + - - 100. Pob3N 100.0 c Pob3p-3 Yml069 552 247-347 + + + 96.9 SSrecog 100.0 c Pob3p-4 Yml069 552 365-470 Rtt106 Rtt106 99.7 Rtt06C 100.0 c Pob3p-4 Yml069 552 365-470 Rtt106 Rtt106 99.7 Rtt06C 100.0 c Pob3p-4 Yml021c 858 28-127 + PH RBD 99.5 RbD[2] 100.0 h Rgc1p Ypr115 1083 493-587 cPH cPH ePH 99.1 cPH[1]/2/3 99.6 h, o Rom1p Ygr070 1155 682-822 + PH cPH 99.1 cPH[1/2/3] 99.6 h, o Rtm10p Yir421c 156 20-132 Rpn13 Rpn13 - 100. Rpn13 10			552		+	_	-	94.2		100.0	
Pob3p-3 Yml069 552 247-347 + + + 96.9 SSrecog 100.0 c Pob3p-4 Yml069 552 365-470 Rtt106 Rtt106 99.7 Rtt106C 100.0 c Psy2p Ynl201c 858 28-127 + PH RBD 99.5 RBD[2] 100.0 h Rgc1p Ypr115 1083 493-587 CPH cPH RD 99.5 RBD[2] 100.0 h Rom1p Ygr070 1155 682-822 + PH cPH 99.1 cPH[1] 99.6 h,4 Rom2p Ylr371w 1356 875-1020 + PH cPH 99.1 cPH[1/2/3] 99.5 h,4 Rpn13p Ylr421c 156 20-132 Rpn13 Rpn13 - 100. Rpn13 100.0 c Rtt106p-1 Ynl206c 455 70-197 Rtt106 - - 99.8 SSrecog		Yml069			+	_	-				
Pob3p-4 Yml069 552 365-470 Rtt106 Rtt106 Rtt106 99.7 Rtt106C 100.0 c Psy2p Ynl201c 858 28-127 + PH RBD 99.5 RBD[2] 100.0 h Rgc1p Ypr115 1083 493-587 CPH cPH PH 89.3 cPH[1] 99.8 h Rom1p Ygr070 1155 682-822 + PH cPH 99.1 cPH[1/2/3] 99.6 h,d Rom2p Yir371w 1356 675-1020 + PH cPH 99.2 cPH[1/2/3] 99.5 h,d Rpn13p Yir421c 156 20-132 Rpn13 - 100. Rpn13 100.0 c Rtt106p-1 Ynl206c 455 70-197 Rtt106 - - 99.8 SSrecog 100.0 c Std3p-1 Ynl206c 455 215-305 Rtt106 Rtt106 - 100. <t< td=""><td></td><td>Yml069</td><td></td><td></td><td>+</td><td>+</td><td>+</td><td>96.9</td><td></td><td></td><td></td></t<>		Yml069			+	+	+	96.9			
Psy2p Ynl201c 858 28-127 + PH RBD 99.5 RBD[2] 100.0 h Rgc1p Ypr115 1083 493-587 cPH cPH cPH 89.3 cPH[1] 99.8 h Rom1p Ygr070 1155 682-822 + PH cPH 99.1 cPH[1]/2/3] 99.6 h,d Rom2p Yir371w 1356 875-1020 + PH cPH 99.2 cPH[1]/2/3] 99.6 h,d Rpn13p Yir421c 156 20-132 Rpn13 Rpn13 100.0 c Rtt106p-1 Ynl206c 455 70-197 Rtt106 - - 99.8 Srecog 100.0 c Rtt106p-2 Ynl206c 455 215-305 Rtt106 Rtt106 - 100.0 Rt106 100.0 c Seg3p Yer008c 1136 107-223 - PH - 100.0 cH[1/2/3] 99.9		Yml069									
Rgc1p Ypr115 1083 493-587 CPH CPH CPH 89.3 CPH[1] 99.8 h Rom1p Ygr070 1155 682-822 + PH CPH 99.1 CPH[1]/2/3 99.6 h, , Rom2p Yir371w 1356 875-1020 + PH CPH 99.2 CPH[1/2/3] 99.5 h, , Rpn13p Yir421c 156 20-132 Rpn13 Rpn13 - 100. Rpn13 100.0 c Rtt106p-1 Ynl206c 455 70-197 Rtt106 - - 99.8 Ssrecog 100.0 c Rtt106p-2 Ynl206c 455 215-305 Rtt106 - 100.0 Rtt106C 100.0 c Sec3p Yer008c 1136 107-223 - PH - 100.0 CPH[1/2/3] 99.9 h Skg3p-1 Ynl257c 1229 308-430 CPH cPH PH 93.3		Ynl201c									
Rom1p Ygr070 1155 682-822 + PH CPH 99.1 CPH[1/2/3] 99.6 h, d Rom2p Yir371w 1356 875-1020 + PH CPH 99.1 cPH[1/2/3] 99.6 h, d Rom2p Yir371w 1356 875-1020 + PH cPH 99.2 cPH[1/2/3] 99.6 h, d Rpn13p Yir421c 156 20-132 Rpn13 Rpn13 - 100. Rpn13 100.0 c Rtt106p-1 Ynl206c 455 70-197 Rtt106 - - 99.8 SSrecog 100.0 c Rtt106p-2 Ynl206c 455 215-305 Rtt106 Rtt106 - 100.0 Rt16C 100.0 c Sec3p Yer008c 1136 107-223 - PH - 100.0 cPH[sc3] 100.0 c Sip3p-1 Ynl257c 1229 308-430 cPH cPH cPH											
Rom2p Yir371w 1356 875-1020 + PH CPH 99.2 CPH[1/2/3] 99.5 h, f,											h,o
Rpn13p YIr421c 156 20-132 Rpn13 Rpn13 - 100. Rpn13 100.0 c Rtt106p-1 Ynl206c 455 70-197 Rtt106 - - 99.8 SSrecog 100.0 c Rtt106p-1 Ynl206c 455 70-197 Rtt106 - - 99.8 SSrecog 100.0 c Rtt106p-2 Ynl206c 455 215-305 Rtt106 Rtt106 - 100. Rtt106C 100.0 c Sec3p Yer008c 1136 107-223 - PH - 100. cPH[sc3] 100.0 c Sip3p-1 Ynl257c 1229 308-430 cPH cPH cPH 99.3 cPH[1/2/3] 99.9 h Skg3p-1 Ynl257c 1229 308-430 cPH cPH PH 97.3 cPH[1/z/3] 99.8 h Skg3p-2 Yir187w 1026 253-385 cPH cPH -		-									h,o
Rtt106p-1 Ynl206c 455 70-197 Rtt106 - - 99.8 SSrecog 100.0 c Rtt106p-2 Ynl206c 455 215-305 Rtt106 Rtt106 - 100. Rtt106C 100.0 c Sec3p Yer008c 1136 107-223 - PH - 100. cPH[sec3] 100.0 c Sip3p-1 Ynl257c 1229 308-430 cPH cPH cPH 99.3 cPH[1/2/3] 99.9 h Skg3p-1 Ynl27r 1229 308-430 cPH cPH cPH 97.3 cPH[1/2/3] 99.9 h Skg3p-1 Ynl47rw 1026 89-230 cPH cPH cPH 97.3 cPH[1/2/3] 99.8 h Skg3p-2 Yir187w 1026 253-385 cPH cPH - 94.6 cPH[1/3] 97.4 h											
Rtt106p-2 Ynl206c 455 215-305 Rtt106 Rtt106 - 100. Rtt106C 100.0 c Sec3p Yer008c 1136 107-223 - PH - 100. cPH[sec3] 100.0 c Sip3p-1 Ynl257c 1229 308-430 cPH cPH cPH 99.3 cPH[1/2/3] 99.9 h Skg3p-1 Ylr187w 1026 89-230 cPH cPH cPH 97.3 cPH[1/2/3] 99.8 h Skg3p-2 Ylr187w 1026 253-385 cPH cPH - 94.6 cPH[1/3] 97.4 h						-	_				
Sec3p Yer008c 1136 107-223 - PH - 100. cPH[sec3] 100.0 c Sig3p-1 Ynl257c 1229 308-430 cPH cPH ePH 99.3 cPH[1/2/3] 99.9 h Skg3p-1 Yirl87w 1026 89-230 cPH cPH cPH 97.3 cPH[1/2/3] 99.8 h Skg3p-2 Yirl87w 1026 253-385 cPH cPH - 94.6 cPH[1/3] 97.4 h		1				R#106	_				
Sip3p-1 Ynl257c 1229 308-430 CPH CPH CPH 99.3 CPH[1/2/3] 99.9 h Skg3p-1 Ylr187w 1026 89-230 CPH CPH CPH 97.3 CPH[1/2/3] 99.8 h Skg3p-2 Ylr187w 1026 253-385 CPH CPH – 94.6 CPH[1/3] 97.4 h					_		_				
Skg3p-1 YIr187w 1026 89-230 CPH CPH CPH 97.3 CPH[1/2/3] 99.8 h Skg3p-2 YIr187w 1026 253-385 CPH CPH – 94.6 CPH[1/3] 97.4 h											
Skg3p-2 Ylr187w 1026 253-385 cPH cPH – 94.6 cPH[1/3] 97.4 h											
	• ·						UPH				
SKITP TUTTSW 000 US-122 CET CET CET 99.3 [CET[1/2/3] 99.9 N							- -				
Slm1p Yil105c 686 467-583 cPH cPH cPH 98.2 cPH[1/2/3] 99.9 h											

continued overleaf

!	1									
A (cont'd)	systematic		residues	1	2	r 3	nethod	4	5	note
protein name	gene name	length	Tesidues	InterPro- Scan	SMART	SUP- FAM	PDE p[SS]	3-to-yeast family	yeast -to-PDB	note
Slm2p	Ynl047c	656	445-555	cPH	cPH	cPH	98.3	cPH[1/2/3]	99.9	h
Spo14p	Ykr031c	1683	487-668	cPH	cPH	cPH	97.4	cPH[1/2/3]	99.7	h
Spo71p-2 °	Ydr104c	1245	741-970	cPH	5	-	93.4	cPH[1/2/3]	97.1	h
Spo71p-3 °	Ydr104c	1245	1026-1240	cPH	cPH	cPH	97.4	cPH[1/2/3]	97.0	h
Spt16p-1	Ygl207w	1035	550-660	+	+	-	100.0	Spt16D	100.0	С
Spt16p-2	Ygl207w	1035	665-820	I	-	1	100.0	Spt16M	100.0	С
Spt16p-3	Ygl207w	1035	835-940	Rtt106	Rtt106	-	75.7	Rtt106C	100.0	С
Ste5p	Ydr103w	917	400-510	I	-	I	86.5	cPH[2]	98.0	h
Syt1p	Ypr095c	1226	855-1075	cPH	cPH	cPH	60.1	cPH[1/2/3]	99.7	h
Tfb1p	Ydr311w	642	01-111	TFIIH	TFIIH	TFIIH	99.9	Tfb1	100.0	С
Tus1p	Ylr425w	1307	710-888	cPH	cPH	cPH	99.2	cPH[1/2/3]	99.8	h
Vps36p	Ylr417w	566	15-160	+	gram	gram	99.9	glue	100.0	С
Yel1p	Ybl060w	687	395-551	-	cPH	-	95.8	cPH[1/3]	99.8	h
Yhr131cp	Yhr131c	850	130-306	+	cPH	cPH	92.7	cPH[split]	99.9	h
Ymr1p	Yjr110w	688	16-130	+	0.03	gram	99.7	gram	97.5	h
Ynl144cp	Ynl144c	740	135-306	+	+	cPH	95.1	cPH[split]	99.9	h
Yrb1p	Ydr002w	201	80-199	RBD	PH	RBD	99.9	RBD[2]	100.0	h
Yrb2p	Yil063c	327	205-327	RBD	PH	RBD	99.9	RBD[2]	100.0	h
Ysp1p-1	Yhr155w	1228	305-430	cPH	cPH	cPH	99.4	cPH[1/2/3]	99.9	h
Ysp2p	Ydr326c	1438	632-760	gram	gram	I	97.3	gram	97.2	h
Known positives identified				62	59	48	71		72	
False positives (non-PH domains identified)				3§	0	0	1§		1§	

В	svstematic	PSI F (0			PDB-to-yeast		Yeas						
D New PH-like proteins	gene name	protein length	residues	e-val	#	SMART	SFFAS	РНҮRЕ	p[SS]	family	p[SS]	PDB hit	HHalign
Age1p	Ydr524p	483	1-130						95	split	42.0	thiored.	0.1*
Bud2p-1	Ykl092c	1105	28-144 ¶	-6	3		-8.7	97	60.5	split	97.6	1upq_A	99.8
Bud2p-2	Ykl092c	1105	203-315	-24	3		-14	90	63.2	cPH[3]	99.8	1wi1_A	89.9
Caf120p-1	Ynl278w	1060	240-364	-27	1	+2	-11	96	97.4	cPH[1]	97.5	1dro_A	94.0
Gyp7p	Ydl234c	747	1-155				-12	25	66.1	Vps36	95.7	2cay_A	96.1
Lam1p-2	Yhr155w	1228	596-700	-14	2		-13	98	96.6	gram	96.5	2cay_A	97.9
Pkh1p §	Ydr490c	767	542-673						<5		97.9	1w1g_A	86.2
Pkh2p	Yol100w	1082	860-967	-2	2		-16	97	23.3	cPH[1]	98.1	1w1g_A	98.4
Rbh1p	Yjl181c	612	440-546				-11	64	<5		91.9	1rrp_B	88.4
Rbh2p	Yjr030c	746	566-680				-10	93	32.7	FERM1	95.8	1rrp_B	93.4
Rec114p	Ymr133w	429	1-130					68	18.3	RBD[1]	57.4	2oqb_A	71.0
Sip3p-2	Ynl257c	1229	599-699	-7	2		-14	98	96.0	gram	97.4	2cay_A	98.3
Spo71p-1	Ydr104c	1246	492-613	-18	1	-30	-11.9	97	<5		94.9	2dtc_A	94.5
Tph3p-1	Yjl016w	562	62-178				-6.1	97	<5		96.6	2rov_A	95.7
Tph3p-2	Yjl016w	562	182-407	-6	1	-9	-11	94	74.2	cPH[1]	88.4	2d9w_A	88.2
Vid27p-1	Ynl212w	783	16-165		1				<5[77]	indirect	51.5	UmVid27p-2	96.8
Vid27p-2	Ynl212w	783	228-339	-2 ^T	4		-29	98	96.2	RBD[2]	99.9	3m1i_B	99.8
Vps13p	YII040c	3144	3028-144				-10	95	74.5	bactPH	94.1	3hsa_A	95.0

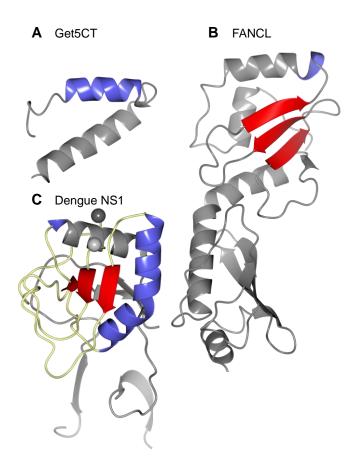
!

Supplemental Table 2. Description of PH-like domains in yeast A. 73 known PH-like domains in yeast identified by 5 different methods: InterProScan (available through yeastgenome.org), Smart, SuperFamily (methods 1–3) and HHsearch (PDB-to-yeast and yeast-to-PDB =

2

methods 4 and 5). For methods 1-3, any family specified is indicated (cPH = classical PH), while "+" indicates a generic PH-like identification. § False positives in InterProScan were produced by Gene3D in Cbk1, Fpk1p and Kin82p. HHsearch identified Age1p as a false positive (see Supplemental Figure 4). For method 4, the prob[SS] and family (among the 39 we defined – see Figure 2) of the strongest hit to the yeast domain is indicated. Other notes: c = crystal structure solved, h = homologous to a solved structure, o = omitted by Yu*et. al.*(2004) ⁽⁶⁴⁾; ° numbering of PH-like domains within individual proteins takes account of the new domains we identify (see Figure 4).

B. Newly identified PH-like domains in yeast. Details include discoverability by different tools: (1) PSI-BLAST, showing E-value (log₁₀) and number of iterations (#) when a known PH-domain (definitions 1 or 2 including yeast proteins in part A above) is found in the hit list; ^T indicates only temporarily in hit list, lost before convergence; (2) SMART domain prediction server, showing E-values (log₁₀) for any domains detected; (3) FFAS profile-profile tool, (score is more significant when more negative, threshold = –9) ⁽⁴⁷⁾; (4) PHYRE2 structural prediction tool (showing probability of PH-like fold, if top hit) ⁽⁴⁸⁾. HHsearch: method 5 = PDB-to-yeast (prob[SS]and query); method 6 = yeast-to-PDB (prob[SS] and target); method 7 = HHalign result with target in method 6. Prob[SS] values outlined in red were used to make identifications: 4 in PDBto-yeast, 11 in yeast-to-PDB; one further identification was made by indirect searches (outlined in black). Other notes: ¶ Bud2p-1 is indicated by SMART as overlapping a false-positive C2 domain; § new true positive in Pkh1p already in SGD (identified by Gene3D); "T" indicates that for Pkh2p and Vid27p-2 the PSI-BLAST hit was only temporary, and was missing at convergence. Grey shading = non-significant hit in Rec114p found using ssw=30%; yellow = false positive for Age1p-N, where the top hit in yeast-to-PDB searches was a thioredoxin, and pairwise alignment of Age1p-N and the PH domain from centaurin (human ARFGEF) had prob[SH]=0.1%. Fidler et al. Supplemental Figure 1.



Supplemental Figure 1: Structures of strongest hits that are not PH-like domains

Three of the strongest false hits to non-PH-like proteins among PDB-to-PDB HHsearches with 39 PH-like domains (indicated as black squares in Figure 3A): **A.** C-terminus of yeast Get5p ($3vej_A$) aligned at 14 of its 41 residues to the helix of 4chj_A, prob[SS]=77%; **B.** mid-section of human FANCL ($3zqs_A$) aligned at 40 of its 186 residues to strands 3–5 of 2kig_A, prob[SS]=83%. **C.** $\alpha\beta$ subdomain of NS1 from Dengue virus type 2 (within a multi-domain structure of 350 aa, not shown) aligned at 90 of its 114 residues to all 7 sheets in 4gou_A, prob[SS]=76%. Aligned regions shown in colors (blue = helix, red = sheet, yellow = loop). While some non-PH-like hits are very short and can be identified as non-PH-like by that criterion (A), others cover multiple structural elements and cannot be excluded on the basis of shortness (B & C).

Fidler et al., Supplemental Figure 2

Α	
Q ss pred CCCceeEEEEccCCCEEEEcccccccccCceEeccCCEEEec43ccCc	cc
Q Centaurin PH 62 LKTWKKKYVTLCSNGVLTYYSSLGDYMKNIHKKEIDLRTSTIKVP43FSP0	ΞI
Q Consensus ~k~WKrRwFVL~~~g~L~Yy~s~~d~~~~~g~I~L~~~tv~v~43~s~~	~~
.+. . .+ .++ + +	
T Consensus ~v~~~~v~~~v~~~y~~w~~~v~t~i~lk~~~c~~~	
T Agelp 12 VVPLFSKGTVAR-TASKAQYPSWCNNALKLTNILLKSLRCKFQ	
T ss_pred ccccccceeeee-ccchhcccchhccccccchHHHHHHHhhh	
Confidence 346999999999 688887765521 122234556655433211	

	*
Q ss_pred	cccccccC7cc8Cce EEEE eCCCc EEEE cCC HHHHHHHH-HHHHHH h
Q Centaurin PH	SSTTSPKLN7NK8TNNFMIVSATGQTWHFEATTYEERDAWV-QAIQSQIL 280
Q Consensus	~~~~p~~~7~r8~~~F~Iit~~~rty~f~A~S~eE~~~WI-~AI~~aI~
	+ .+ . . + ++++ ++++ .+ .+++-
T Consensus	n dr ~f~~~~v~~-~~a~~e~~~~i~~~~~i~~~~~
T Agelp	TNRCED DR GFEVYCVIL-KSIALLMAAKESLILLQI PP SL <mark>P</mark> SGF <mark>P 97</mark>
T ss pred	hccccc cc ceEEEEecc-chhhhhhhcchHHHHhhhhhhccccc
Confidence	011111 33 799999998 477899999999999999 66766653
	*

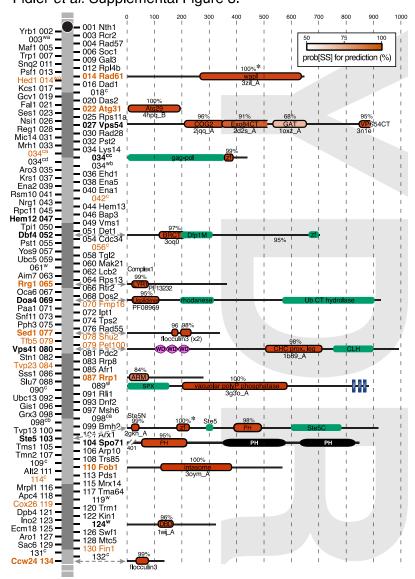
В

-
$+ + + + - + \pm - + - + - + - + - + - $
Age1p (S.cerevisiae)MDFY DINKNVVPLESKG VAK ASKAUYPSWCNN-ALKL NILLKSLKCKE
Age1p (S.cerevisiae) MDFY11D1NKNVVPLF5KG1VAH1A5KAUYPSWCNN-ALKL1N1LLK5LKCK Arf-GEF (Yarrowia) MDPSTKNINKKTVGWHKYWVVLAGGRLSEYSNWKQG1EANRTPIDLKMALCRE
Arf-GEF (Tetraodon)
+ + + +-+-++ +-++ ± ¥+ PP±+P+++P
Age1p (S.cerevisiae) QTNRCED-DRGFEVYCVILKSIALLMAAKESLILLQIPPSLPSGFPFRSPQ
Arf-GEF (Yarrowia) ARNS-DRRFCFEIVTPQTKRVYQATSEDDLATWVAAINGAISTSLEN
Age1p (S.cerevisiae) QTNRCED-DRGFEVYCVILKSIALLMAAKESLILLQIPPSLPSGFPFRSPQ Arf-GEF (Yarrowia) ARNS-DRRFCFEIVTPQTKRVYQATSEDDLATWVAAINGAISTSLEN Arf-GEF (Tetraodon)CEDnERRFCFEVVSPTKSCMLQAESEKLRQAWIqavQASIASAYKDITDN
*

Supplemental Figure 2: Dissection of a false positive in Age1p

A. Alignment showing a strong hit obtained in HHsearch between the PH-like domain of human ArfGEF Centaurin and the N-terminus of Age1p. For interpretation of the alignment, see Figure 5B legend. Three unstructured loops in centaurin of 43, 7 and 8 aa are omitted. The single-most conserved residue in PH-like domains is the tryptophan in the helix (asterisk), which is reported as making a very good hit (|) with a glutamine in Age1p. **B.** All sequence from the first 102 columns of the Age1p MSA in HHsuite. This part of the MSA includes only 3 of 164 sequences: Age1p N-terminus (1-102/482aa), *Y. lipolytica* ArfGEF (GI:50550125; 735-833/1099aa), and *T. nigroviridis* ArfGEF (GI:47212317; 425-474/991aa). Coloring by HHsuite AlignmentViewer. All residues where the HHsearch alignment in part A. reported a good or very good hit (+ or |) are scored for actual conservation between Age1p and the other sequence(s): +, ±, – for strength of hit. Residues 81-97 of Age1p predicted to align with alpha helix of centaurin in A. contains multiple prolines in its second half (highlighted in both A and B). The N-terminus of Age1 has homologs only in species very closely related to *S. cerevisiae*, so this part of the MSA should only have one sequence. However, the downstream GEF (Age1p residues 168-307) has 163 homologs, many of which have a domain structure BAR-PH-GEF. Two of these align to part of Age1p 1-130 (A). These give a PH-like appearance to the MSA at its N-terminus that is not present in the original seed (B).

Fidler et al. Supplemental Figure 3.



Supplemental Figure 3: New domains found in 132 consecutive yeast genes

132 consecutive verified open-reading frames on the long arm of chromosome IV ("YDR") starting at the centromere (YDR001C-134C) were analyzed by HHsearch. Names in bold indicate proteins (x16) where new domains were documented (x21). New domains are outlined in black and shaded by prob[SS] on a white–red scale; scores and main hit in PDB or Pfam are also given. All new hits were confirmed by pairwise alignments of domains alone. Asterisks indicate that the domain was detected by SMART even though absent from InterProScan data displayed at <u>yeastgenome.org</u>). Known domains in these proteins are shaded as in Figure 4 or in green. Naming: where a protein only has a systematic name, just the suffix is given (*i.e.* "XYZ W/c" from the full name YDRXYZw/c). Red indicates the proteins (19 = 14%) without prior structural/functional predictions (discounting assignment to a fungal protein family and prediction of transmembrane domains). Seven of these (37%) had newly predicted domains. By contrast, nine of 113 (8%) previously characterised proteins had newly identified domains.

Movies 1-3: PH-like domains share structure Rotating models of structures in Figure 1.