Supporting Information

C

Blaby-Haas et al. 10.1073/pnas.1421545111



Fig. S1. Schematics of alternative-splicing events and gene duplications. (*A*) Gene models presented in Fig. 1*A* (*Upper*) and predicted structural models of PCH1 and PAA1 (*Lower*). The N-terminal HMBD of PAA1 is colored purple to highlight identity with PCH1, also colored purple. (*B*) Gene models not presented in the main text, where alternative-splicing events producing both PCH1 and PAA1 are likely (because EST support for PCH1 is available). (*C*) Gene models of *PAA1* and *PCH1* presented in Fig. 1*B* (*Upper*) and predicted structural models of PCH1 and PAA1 (*Lower*). (*D*) Gene models from *G. max* and *P. trichocarpa*, where the presence of two genes with alternative splicing is unclear. chr, chromosome; sc, scaffold.



Fig. 52. PAA1/PAA2 protein cluster uniquely contains soluble proteins. The protein similarity network of the heavy metal associated (HMA) domains (CDD cl00207) in Cu⁺-ATPases and Cu⁺ chaperones from plants, cyanobacteria, *Saccharomyces cerevisiae* (ScATX1), and *Homo sapiens* (HsHAH1) is shown. Nodes that represent the HMA domain from the Cu⁺-ATPases are colored blue, whereas the nodes that represent the HMA domain from identified and predicted chaperones are colored red. Proteins that are more similar to one another than to other proteins in the network will cluster, allowing the delineation of protein subgroups based on amino acid similarity. Each distinct cluster is given a number designation, and representative *A. thaliana* proteins from each cluster are shown. AtHMA5 contains two HMA domains, one in cluster 2 and the other in cluster 3. The same network, but colored by taxonomic group, is provided to the right. Atx1, antioxidant 1; CCH, Copper chaperone; HIPP, heavy metal associated isoprenylated plant proteins; HPP, heavy metal associated plant proteins; RAN1, responsive-to-antagonist 1.



Fig. S3. Phylogenetic reconstruction (neighbor-joining tree) of the HMA domain from selected Cu⁺-ATPases (PAA1 in orange and PAA2 in yellow), PCH1 (lilac), and no-glycine stretch protein (pink) from the PAA1/PAA2 cluster identified in Fig. S2. The clade that represents grass-specific sequences is highlighted with a dashed line.

	-						
EhCopZ_1CPZ	A						
CrPCH1 AtPCH1 MgPCH1 PrPCH1 HsHAH1 ScATX1 AtCCH AtATX1	MQSVAQRGVVARP MESTLSAFSTVKATAMARSSGGPSLP MESTLSVGASTMS MEAVVAT	ATTHFVF LLTISKALNRHFTCARHLHP IAALSRSLNSH.STPIIAHL TTTRVSFP LFILSKALNRHFTTNTTRSL	2CPRRARPALVTLAV LLARCSPSVRRLGG IRRFSTTLPSQIALL LIPTAGRPLLP .RISSTQLVTRRLSS	SCPGNNTSLC FHGSRFTS 2LSRRVELRR IYSRNSDFLE	SQFQHASWAQK SNSA. LRSLG FGPGHLVRG.H SSPS. LKGLR	FVRATVVATTLQ AAVLPVIRHRLE DSV.DLRRNRLQ PVLARPRR VGV.PLVRLRLE LKDLFQ	FFAAAS CLSSSS FSVRSA GFASVS CVSSSA
	polyalycine	stretch	β1	ي 10	α1 200000000 20	TT <u>β2</u> 30	40
EhCopZ_1CPZ			QE <mark>F</mark> S	S <mark>VKGMSC</mark> NHC	<mark>v</mark> ar <mark>ieeav</mark> gri	. SGVKKVKVQLK	. K E K <mark>A</mark> V <mark>V</mark>
CrPCH1 AtPCH1 MgPCH1 PtPCH1 HsHAH1 ScATX1 AtCCH AtATX1	HAANASVGGGIGAGGGHRGGGGGGG PSFRSISSGGGSGFGGYNGGSGG SSGGGFDGIDGGAGGGGG ATSTSGAGGGGGRSSAGGCGDG ASFGTTSGGGGGGGEFGGG TASLSLFQAL.	GGDGHSSMSHGGSPTVLG.D GGGGGSESGDSKSKLGAN GGDGAAEGGDAKASAVTS GGDGAAEGGDAKASAVTALG GGGGSNGGDAESNSVAEAV	AEASSDLVEEVVLL SDGVSVPSSDIIL SAGASPLSEVIIL SEQQEDGDSDAIVL SEEAPVLSPDVIIL AAEAPVLSPDVIIL AEIKHYOPI SVVESKAMSQTVVL	MCCCC VGGMKCGGC VGGMTCGGC VGGMTCGGC VGGMTCGGC VGGMTCGGC VGMTCGGC V.DMTCGGC V.VMTCSGC V.AMTCEGC	VGHVKKILEEQ SASVKKILESQ AASVKRILESQ ADKVKRILESQ ADKVKRILESQ ASVKRILESQ SGAVNKVLKL SGAVNKVLTKL VGAVNKVLGKM VGAVKRVLGKM	. PGVTSASVNLA . PQVASASVNLT . PQVSSASVNLT . PEVAAATVDVE . SQVFSASVNLA . GGV. KYDIDLP EPDVSKIDISLE . EGVESFDIDLK . EGVESFDIDLK	TETALV TETALV TETALV KATAIV NKKVCI KQLVDV EQKVTV EQKVTV
	83	~?	84				
	=►	eeeeeee	u 🗕				
5k07 4007	50 K EDE ANUO		60 NELOVONENT				
EnCop2_1CP2	KFDEANVQ	A	NELGIQAEVI				
CrPCH1 AtPCH1	RVLVPRGSRQQPSGSSG	NGGSGGGGGALAALG <mark>E</mark> KLAQV	SKEGFASRVRVGGA	ASVAK			
MgPCH1	WPSSEA	NWREDIG <mark>E</mark> ALAKH	TNC <mark>G</mark> FKSNLRDQEA	LEGDIQL			
ZmPCH1 PtPCH1	WTTPEA	DWQKQLAEKLANH. NWQKQLG <mark>E</mark> ALAKH	TSCGFQSHLQDEGEA SCGFISNLRGQEA	AEQTDA FEGDISS			
HsHAH1	ESEHSM	DTLLAT	KKTCKEVPSCKOL				
AtCCH	KGNVEPEAVFQTVSKTGKKTS	YWPVEAEAEPKAEA <mark>D</mark> P <mark>KV</mark> ET	TET <mark>KTEAETKTEAK</mark>	DAKADVEPK	AAEAETKPSQV		
AtATX1	KGNVQP	D A <mark>VL</mark> QT	<mark>/T</mark> KT <mark>G</mark> KKTAFWEAEGI	ETAKA			
Fig. S4. Ami (Enterococcus	no acid alignment of plant PCH1 sequ hirae Protein Data Bank ID code 1C with a grav-blue box. At A thaliana: (ences with the sequences of PZ) secondary structure is sh Fr. C. reinbardtii: Hs. H. sanie	known Atx1-like cha nown above. The pol	perones. A ca yglycine stret	artoon represen tch found near	tation of the alig the N terminus	ined Cop of PCH1

pA 1 is Cr, C. reinhardtii; Hs, H. sapiens; Mg, M. g *uttatus*; Pt, cerevisiae; Zm, Z. mays.

AtPAA1 AtPCH1	1 MESTLSAI MESTLSAI	10 STVKAT STVKAT	2 AMARS AMARS	9 SGGPSL SGGPSL	30 PLLTIS PLLTIS	4 KALNRHFT KALNRHFT	GARHLH GARHLH	50 PLLLARC PLLLARC	SPSVRRI SPSVRRI	Q GGFHG GGFHG	70 SRFTSS SRFTSS	NSALR NSALR	8 Q Slgaav Slgaav	LPVIRH LPVIRH	90 RLECLSS RLECLSS	100 SSPSFF SSPSFF	SISSGGG SISSGGG	110 SGFGGYN SGFGGYN	120 GGSGGG GGSGGG	GGGGS GGGGS
AtPAA1 AtPCH1	130 ESGDSKSP ESGDSKSP	14 KLGANAS KLGANAS	DGVSV DGVSV	150 PSSDII PSSDII	ILDVGG ILDVGG	160 MTCGGCSA MTCGGCSA	170 SVKKIL SVKKIL	ESQPQVA ESQPQVA	180 SASVNLI SASVNLI	19 TETAI TETAI	0 VWPVPI VWPVPI	200 AKSVP AKSVP	DWQKSI DWQKSI	210 GETLANE GETLANE	22 ILTNCGF ILTNCGF	Q QSTPR QSTPR	EVPEDIA	DL GEFAP	230 VTENFF 	KVFET
AtPAA1 AtPCH1	240 KTKDKQAI	250 RLKESGF	ELAVS	260 WALCAV	2 CLVGHL	70 THFLGVNA	280 PWIHAI	29 HSTGFHV 	0 SLCLITI	300 LGPGR	KLVLDO	310 IKSLL	3 KGSPNM 	120 INTLVGLO	330 GALSSFS	VSSLAA	340 MIPKLGW	350 KTFFEEP 	VMLIAF	360 VLLGR
AtPAA1 AtPCH1	370 NLEQRAKI	CKATSDM	380 ITGLLS 	VLPSKA	390 RLLLDG 	400 DLQNSTVE	VPCNSL	410 SVGDLVV 	420 ILPGDRV	PADGV	430 VKSGR	TIDES	440 Sftgee 	450 LPVTKES	GSQVAA	460 GSINLN	47 GTLTVEV 	0 HRSGGET	480 AVGDII:	RLVEE
AtPAA1 AtPCH1	490 AQSREAP	500 VQQLVDK	VAGRF	510 IYGVMA	5: LSAATF 	20 IFWNLFGA	530 HVLPSA	54 LHNGSPM 	0 SLALQLS	550 CSVLV 	VACPCI	560 LGLAT	5 PTAMLV	GTSLGAP	580 RGLLLR	GGDILE	590 KFSLVDT 	600 VVFDKTG 	TLTKGH	610 PVVTE
AtPAA1 AtPCH1	620 VIIPENPE	C RHNLNDI	630 WSEVE	VLMLAA	640 AVESNT'	650 THPVGKAI	VKAARA	660 RNCQTMK 	670 AEDGTF1	EEPGS	680 GAVAIV	'NNKRV 	690 TVGTLE 	700 WVKRHG) ATGNSLL	710 ALEEHE	72 INNQSVV	0 YIGVDNT	730 LAAVIR:	FEDKV
AtPAA1 AtPCH1	740 REDAAQVV	750 /ENLTRÇ	GIDVYI	760 MLSGDK	7 RNAANY 	70 VASVVGIN	780 IHERVIA	79 GVKPAEK 	0 KNFINEI	800 800 800	IVAMV(810 DGIND	E AAALAS	SNVGVAN	830 Igggaga	ASEVSI	840 VVLMGNR 	850 LTQLLDA	MELSRQ	860 TMKTV
AtPAA1 AtPCH1	870 KQNLWWAI) FGYNIVG	880 IPIAA	GVLLPL	890 TGTMLT: 	900 PSMAGALM	IGVSSLG	910 VMTNSLL	920 LRYRFFS	NRNDK	930 NVKPE1	•KEGTK	940 QPHENI 	RWKQSS						

Fig. S5. Amino acid alignment of full-length AtPAA1 and AtPCH1. Identical amino acid residues are highlighted with a red background.

PNAS PNAS



Fig. S6. Predicted structural model of the pea PAA1 ortholog. Each cysteine is shown as a red space-filled model.





Fig. 57. Coomassie brilliant blue (CBB) staining and immunoblot of recombinantly expressed and purified Δ N-AtPAA1 and Δ N-AtPAA2 (A) and AtPCH1, AtPAA1-HMBD, AtPAA2-HMBD, and AtCCS (B).

Table S1. Primers

Primer name	R.E. site	Sequence $5' \rightarrow 3'$						
PAA1_His_For	Ndel	aaaaaacatatgCCTCGAGATTTGGTGACAGAGAA						
PAA1_His_Rev	Sall	aaaaaagtcgacAGAGCTTTGCTTCCATCTTG						
PAA2_His_For	Ndel	aaaaaacatatgAAATGGAAAGAAATGGTTAGC						
PAA2_His_Rev	Sall	aaaaaagtcgacCAGAGAATTTTTGCTGGTT						
AtPCH1_strep_For	Ndel	aaaaacatatgTCAAGTGGCGGCGGTTCT						
AtPCH1_strep_Rev	Sall	aaaaagtcgacAGGTGCAAACTCTCCTGCAA						
CCS_strep_For	Asel	<i>atgcattaat</i> ATGGTGGATATGACATGTGAGG						
CCS_strep_Rev	Sall	atgcgtcgacAACCTTACTGGCCACGAAATC						
PAA2HMBD_strep_For	Ndel	aaaaaacatatgATCGAAAGCGTGAAAAGTATTACG						
PAA2HMBD_strep_Rev	Sall	aaaaaagtcgacGGATTTGACCAGCAGGTCTTC						
PAA1HMBD_strep_For	Ndel	aaaaacatatgTCAAGTGGCGGCGGTTCT						
PAA1HMBD_strep_Rev	Sall	aaaaagtcgacGCGAGCCTGCTTGTCTTTTG						
PhoALacZ_For	BamHI	gatcggatccCCTGTTCTGGAAAACCGG						
PhoLacZ _Rev	Xmal	gatccccgggCCATTCGCCATTCAGGCT						
PAA1_PhoLac_For	Ncol	atgcccatggAGTCTACACTCTCAGCTTTC						
PAA1_PhoLac_Rev	Nhel	atgcgctagcCTAAGAGCTTTGCTTCCATCTTG						
PAA1_PhoLac_1Rev	BamHI	atgcggatccAACAGAGACACCATCACTTGCA						
PAA1_PhoLac_1For	Agel	atgcaccggtCCGTCGTCAGATATCATTATTCTC						
PAA1_PhoLac_2Rev	BamHI	atgcggatccGGGAGCATTAACCCCTAGAAA						
PAA1_PhoLac_2For	Agel	<i>atgcaccggt</i> TGGATTCATGCGATCCATT						
PAA1_PhoLac_3Rev	BamHI	atgcggatccACCTTTTAAAAGACTCTTGATACCATC						
PAA1_PhoLac_3For	Agel	atgcaccggtTCCCCAAACATGAACACG						
PAA1_PhoLac_4Rev	BamHI	atgcggatccTGTCTTCCAGCCCAATTTTG						
PAA1_PhoLac_4For	Agel	atgcaccggtTTTTTTGAGGAACCAGTTATGTTAA						
PAA1_PhoLac_5Rev	BamHI	atgcggatccACCATCAAGCAGAAGGCG						
PAA1_PhoLac_5For	Agel	<i>atgcaccggt</i> GATCTGCAGAATTCAACTGTTGA						
PAA1_PhoLac_6Rev	BamHI	atgcggatccAACATGTGCACCAAATAGATTCC						
PAA1_PhoLac_6For	Agel	<i>atgcaccggt</i> CTTCCTTCTGCCTTGCATAA						
PAA1_PhoLac_7Rev	BamHI	atgcggatccCCCATTATGCAAGGCAGAAG						
PAA1_PhoLac_7For	Agel	<i>atgcaccggtA</i> GCCCAATGTCTTTGGCC						
PAA1_PhoLac_8Rev	BamHI	atgcggatccATTGTGTCTTGGATTTTCAGGAA						
PAA1_PhoLac_8For	Agel	atgcaccggtTTGAATGATACTTGGTCAGAAGTAGA						
PAA1_PhoLac_9Rev	BamHI	atgcggatccTGTTTGACAATTACGAGCCC						
PAA1_PhoLac_9For	Agel	atgcaccggtATGAAGGCAGAGGATGGAAC						

R.E., restriction enzyme.

Other Supporting Information Files

Dataset S1 (XLSX)

PNAS PNAS