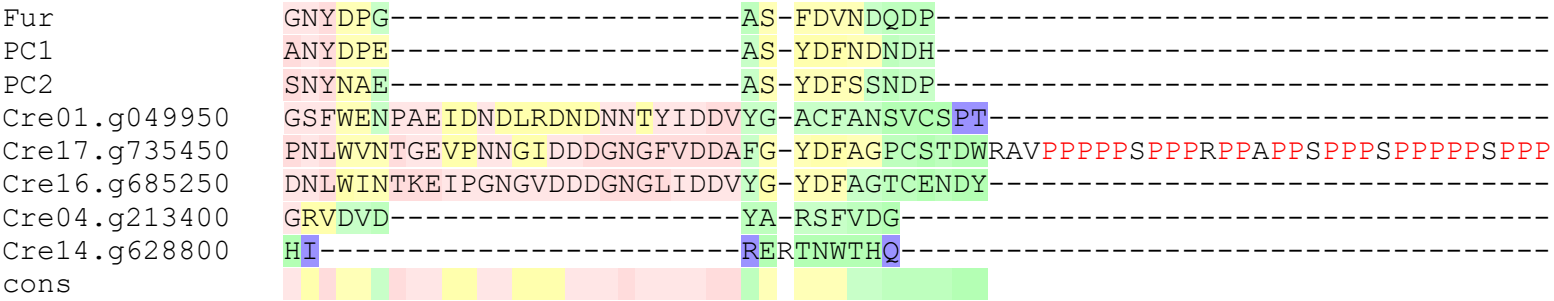
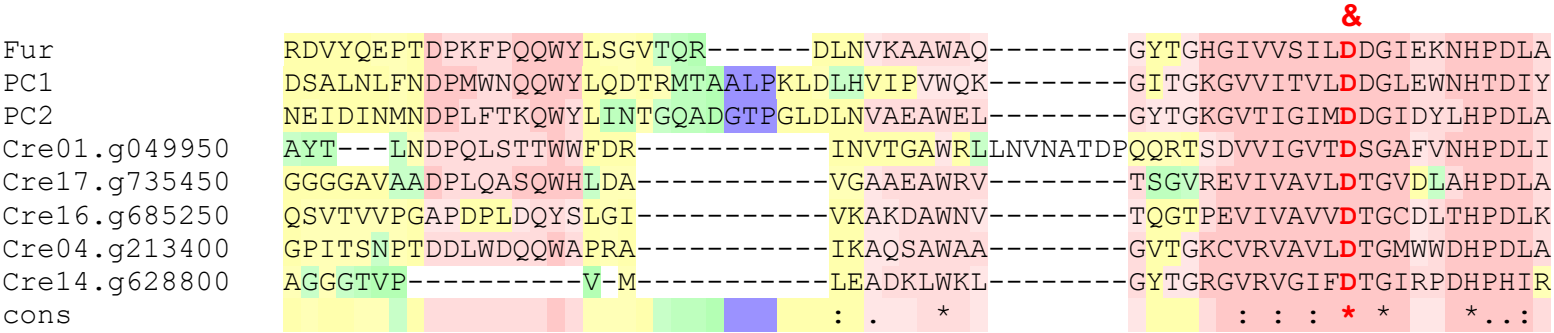


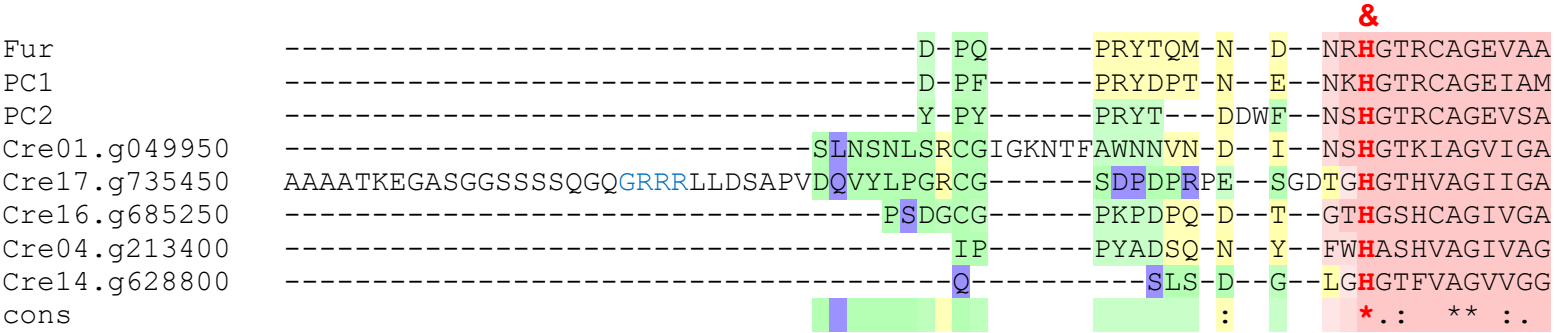
Supplemental Fig.1. Alignment of five S8 subtilisin-like *C. reinhardtii* and human proteins.

MSA *The multiple sequence alignment result as produced by T-coffee.*

T-COFFEE, Version_11.00.8cbe486 (2014-08-12 22:05:29 - Revision 8cbe486 - Build 477)
Cedric Notredame; SCORE=62. *BAD AVG GOOD* TCS (transitive consistency score) color code goes from 0 (blue) to 9 (red).



Cre17.g735450 RPRPPSPPPPPKRAPPPARKTSGRRLVSGAAATHGGRLLSALLGSVWGAGSSEALEATESLQHAQQAATA



Fur	VANNGVCGVGVAYNARIGGVRMLDGE-----VTDAVEARS	LGL-----NPNHIHIYSASWGPED-DGKTVDG
PC1	QANNHKCGVGVAYNSKVGGIRMLDGI-----VTDAIEASSIGF-----NPGHVDIYSASWGPND-DGKTVEG	
PC2	AANNNICGVGVAYNSKVAGIRMLDQPF---MTDIIIEASSISH-----MPQLIDIYSASWGPTD-NGKTVDG	
Cre01.g049950	RPNNGIGLAGVAPNLRQMVLKVVDITYFQYAYS	SDVVRAIDYGY-----AKGARIFSMSFGQDARTSATPTN
Cre17.g735450	VRNGNVGGSGIAPGVRVMPPLKVS	SDPQGNFW-ASNIVAAAYDYAH-----RMGAHVVSCSFGPYQ-PTLNPD
Cre16.g685250	VRNNGVGISGIAPRVKIMCLKVSTPSNNFY-TAHILKAYDYAY-----RMGAHVVSCSFGPAE-PNMKTIS	
Cre04.g213400	SANNG-GIVGIAPEATIIPIKVL	DNNAG-D-NSFLIQGLLYAASPLGPDGAGADIINLSLGGFGNPNDKN--
Cre14.g628800	QD--G-ACPGFAPDVEMLTFKVFTDDQVSY-TSWFLDAFN	YAI-----VVKVNVINLSIGGP-D--Y----
cons		

\$ (oxyanion residue)

Fur	-PARL---AEEAFFRGVSQGRGGLGSIFVWAS	NGGREHDSCNCDGYT--NSI-----YTLSISSATQF		
PC1	-PGRL---AQKAFEYGVKQGRQKGSI	FVWASNGGRQGDNCDCDGYT--DSI-----YTISISSASQQ		
PC2	-PREL---TLQAMADGVNKG	RGGKGSIVWASGGG-SYDDCNCDGYA--SSM-----WTISINSAIND		
Cre01.g049950	KPSLD---AAATAYRNLFN--KYSNALFVAAAG	NEWTSL	EGWRS	GNITYSPCMIA---TDNTLCVGGTN-V
Cre17.g735450	-I-DLAQAAAETRLYERAVRPLADKGVLLVAAAG	NEDTNLDRARAVGMPYLPCTLD---LPNVLCVAASN-T		
Cre16.g685250	-PDELAAYVAERNLYRAALQPMVDKNMLIIAAAG	NENTNLDQLDQYNTTYMPCFMAKEFPENMLCVTATN-Q		
Cre04.g213400	-----DTWALYNKVLNYVASQGV	LVVAAAGNDALNLGRTR--NTFVSPCEN---GGNTVCVSSTG-P		
Cre14.g628800	-----LDQPFVDKVLE-ITSNGVLLVSAIGN	DGPLYGTLN--NP---A-D---QNDVIGVGGID-N		
cons				

Fur	GNVPWY-----SEACS---STLA-TTYSSGNQ-NEKQIVTTDL-----		
PC1	GLSPWY-----AEKCS---STLA-TSYSSGDY-TDQRITSADV-----		
PC2	GRTALY-----DESCS---STLA-STFSNGRKR	NP	PEAGVATTDL-----
Cre01.g049950	NDTIFY-----VF	AFNQAGTNFGPTTVDM--GAP	AHSIYTDDI-----
Cre17.g735450	QGGIWR-----EIVANRWVGSNTGPTTVDL--AAPGS	RIVSTLP-----	
Cre16.g685250	DDVRWA-----EIAN	NKPVGSNFG	LKYADI--GAPGRLIMSSTP-----
Cre04.g213400	LDVCKTCQDD-WINTCLESPSTWDLKQQLALYSNYGPSAISV--AGPGGDYRSAFEWYAKWYDPATDSFLP		
Cre14.g628800	WDNI--ASFSSR---GMSTW-----ELPRGYGRVKPDV--MAYAKDVQG-----		
cons			

&

Fur	-----R--QK--CTESHTGT	SASAPLAAGIIALTLEANKN-----
PC1	-----D--ND--CTETHGT	SASAPLAAGIFALALEANPN-----
PC2	-----Y--GN--CTLRHSGT	SAAAPEAAGVFALALEANLG-----
Cre01.g049950	-----A--VNR--NYSAPSGT	SFATPMVAAGVLSALGG-----TGRA-
Cre17.g735450	-----G--AT--YGDKTGS	SMATPLVSGSAAVLSALGAVDGN-----
Cre16.g685250	---PV-----A--GNSMT--RWNLKTGS	SMATPMVSGIGALVLSVLGTGTGNY-----
Cre04.g213400	GTPPTCWFVDQVLSVGAPNRPDLP-AFIQYTGT	SQAAPAVSGVAALMIHDAAVRQGVDLCAASRAGLKGRAR
Cre14.g628800	-----SKITGGCRSLSGT	SVASPVVAGAVSLLASTVLP-----DKPGAGEDGRW-
cons		

Fur	--LTWRDMQHLLVVQTSK--P--A
PC1	--LTWRDMQHLLVVTSEYDP--L
PC2	--LTWRDMQHLLTVLTSK--RN-Q
Cre01.g049950	TPQTPLQIKNILMSS--GDL--L
Cre17.g735450	--FQAVRARSLLLEG--GDP--Q
Cre16.g685250	--FQGKEARKIIVES--A-----D
Cre04.g213400	GPLNVAQIKTALARG--AV-YPG
Cre14.g628800	ALLNPGSMKQALVEG--AVRLPN
cons	:

Supplemental Fig.1. Alignment of five S8 subtilisin-like *C. reinhardtii* proteins with human homologs. The *C. reinhardtii* proteins most homologous to human furin, PC1 and PC2 were identified using BLASTp. Their catalytic core sequences were then aligned using T-Coffee. The residues that form the catalytic triad (D, H and S) are marked in red. The oxyanion residue (N in all prohormone convertases except for PC2, where it is D) is also marked in red. In Cre17.g735450, an insert containing a Pro-rich region separates catalytic triad residues D and H.

Phylogram

Phylogenetic tree showing the relationships between various Cre recombinase proteins. The tree is rooted on the left and branches to the right. Proteins are grouped into three clusters:

- Cluster 1:** includes the 5 *C. reinhardtii* proteins most like human furin, PC1 or PC2 (green highlight)
 - Cre04.g213400 1.00037
 - Cre01.g049950 0.81436
 - Cre17.g713600 0.39923
 - Cre17.g708400 0.32139
 - Cre03.g145827 0.1853
 - Cre16.g685250 0.24055
 - Cre17.g735450 0.2168
 - Cre14.g628800 0.89151
 - Cre16.g675350 0.85988
- Cluster 2:**
 - Cre10.g459450 0.55592
 - Cre07.g329500 0.63461
 - Cre13.g585800 0.56373
 - Cre02.g076950 1.04512
 - Cre09.g406700 0.76895
- Cluster 3:** delete Cre05.g242700 from alignment – lacks catalytic triad D
 - Cre05.g242100 0.57412
 - Cre05.g242700 0.43396
 - Cre05.g242750 0.08948
 - Cre05.g242856 0.08695
 - Cre03.g190250 0.54155
 - Cre03.g170300 0.6062
 - Cre19.g750447 0.74282

*BAD AVG GOOD * TCS (transitive consistency score) color code goes from 0 (blue) to 9 (red).

Cluster 1: Cre01.g049950 (VLE1, sporangin); Cre03.g145827; Cre04.g213400; Cre14.g628800; Cre16.g685250; Cre17.g708400 (Pro-rich insert between H and oxyanion residue); Cre17.g713600 (Pro-rich insert between H and oxyanion residue and N → D for oxyanion residue); Cre17.g735450 (Pro-rich insert between H and oxyanion residue).

	&		&
Cre01.g049950	TSDVVGVTDSGAFVNHPDLIG---	SFWENPAEIDNLDLRDNDNTYIDDVYGACFANS--VCS-PTSLS-	NNVNDINSHGTKI
Cre03.g145827	VSDVVVAVIDTGC DLTHPDLQG---	QLWTNPGEIPGNGVDDNGYIDDVHGFDFAGS--CE-SDWRSTK-	PDPQDTHSHGTHC
Cre04.g213400	GKCVRVAVLDTGMWWDHPDLAG---	RVDVDYARSFVD-----GI-	ADSQNYFWHASHV
Cre14.g628800	GRGVRVGIFDTGIRPDHPHIRHIRER	TNWTHQQSL-----	SDGLGHGTFTV
Cre16.g685250	-PEVIVAVVDTGC DLTHPDLKD---	NLWINTKEIPGNGVDDDG NGLIDDVYGYDFAGT--CE-NDYPS-	PDPQDTGTHGSHC
Cre17.g708400	P-DVIVAVVDTGFDIDHPDLKP---	NLWVNPGVEVPGDGIDNDGN GFIDINGFDFAGSGDACQ-GDWRARA-(53 aa insert)--	PESSD-TGHGTHV
Cre17.g713600	--TVVVAVVDGGFDTRHPDLAG---	ALWVNPGIEIPGN GIDDDGN GFVDDVNGWDFGGN--CSSGAWRPAP-(43 aa insert)--	PEAGD-GGHGTHV
Cre17.g735450	VREVI AVLDTGVDLAHPDLAP----	NLWVNTGEVPPNN GIDDDGN GFVDDA FG YDFAGP--CS-TDWRAVP-(133 aa insert)-	PESGD-TGHGTHV
cons	* : : * *	** :	: *

Cre01.g049950	AGVIGARPNN	GIGLAGVAPNLRQ	MVLKVVDDTYFQYAY	SDVVRAIDYG	-----YAKGARIFSMSFGQD	ARTSA--	TPTNKPSLDAAA	TAYRNLFNK--	YS
Cre03.g145827	SGIVAALRN	-----IGIAPRVKVMCLKVATP	NGAFYT	SHILKAYDYA	-----YNMGAHIASCSFGPA	EPNLQP	TPQEVEEMWNETK	FYKTALDPMVKK	
Cre04.g213400	AGIVAGSANNG	-GIVGIAPEATIIP	IKVLDDNAG-DN	SFLIQGLLYAASPLGPD	GAGADIINLSLGGFG	GNPNDK-ND	-----TWALYNKVLNY	VASQ	
Cre14.g628800	AGVVGQD	--G-ACPGFAPDVEMLTFKVF	TDDQVSYT	SWFLDAFNYA	-----IVVKVNVINLSIGGP	DYLDQ---	P-----FVDKVL	EITSN	
Cre16.g685250	AGIVGAVRNNG	VGISGIAPRVKIMCLKVSTPSNNFYT		AHILKAYDYA	-----YRMGAHVVS	CSFGPA	EPNMKTISPDELA	AYVAERNLYRAALQPMVDK	
Cre17.g708400	AGIVA AVRNS	VAAVAGAAPRVKVMVLKVFDASGRVYA		SHVVAAYAYA	-----QRMGAHIV	VSCSFGPD	TPNLQ-PQPYEVA	EMVAQESLYSSAVTPLESK	
Cre17.g713600	AGVIAAVRK	DLAGTSGVAPHVRLMLLKVV	DGYGATYG	SRVAAALEYA	-----GRMGAHV	VVTSVGPS	APLPSAPRPYQ	LAANAAAAAVYAAAVGSLRDK	
Cre17.g735450	AGIIGAVRNG	VGSGIAPGVRVMPLKVSDPQGNFWA		SNIVAAYDYA	-----HRMGAHV	VVSCSFGPY	QPTLN-PDAIDLA	QAAAETRLYERAVRPLADK	
cons	:*::..		* ** : : **	: . . *		. : *. *		.	.

\$ (oxyanion residue)

Cre01.g049950	NALFVAAAAG	NEWTSLEGWRS	SGNYTYS	SPCMIA	-----TDNTLCVGGTNVNDTIFY	-----VF	AFN--	QQAGTNFGPTTV	DMGAPAHSIYT
Cre03.g145827	GMLVVAAAAG	NENTPLESLDAVNSTYNPCFMAR		--YFADSM	LCVMATNPNSRWF	-----EIAAN	---	KPIGSNYG	TKYADIGAPGRQIYS
Cre04.g213400	GVLVVAAAAG	NDALNLGRT	--RNTFV	SPCEN	-----GGNTVCVSSTGPLDVCKTC	QDDWINTCLES	PSTWDLKQQ	---	LALYSNYGPSAISVAGPGGDYRS
Cre14.g628800	GVLLVSAIGN	DGPLYGTILN	-----NPAD	-----QNDVIGVGGIDNWDN	IAS	-----FSSRGMST	WELPRGYGRV	KPDVMAYAKDVQG	
Cre16.g685250	NMLIIAAAAG	NENTNLQDLQYNTTYMPCFMAK		--EFPENML	CVTATNQDDV	RWAEIANN	---	KPVGSN	FGLKYADIGAPGRLIMS
Cre17.g708400	GVLVVAAAAG	NELTDLNLGLASVKSNYLPCTLP		-----NSNVMCITGSNPLDRAIT	-----GMAGN	---	MPVGVNYGTAV	VDMAAPGQDIYN	
Cre17.g713600	GVLVVAPAG	DDGIDLDAAKAAGLEYLPCTLAAPPYSLGNVLCVGAADASDRRIRI		-----AVPGLD	---	PALGSNYGSSSVSIAAPGV	DIYS		
Cre17.g735450	GVLVAAAAG	NEDTNLDRARAVGMPYLPCTLD		-----LPNVLCVAASNTQGGIWR	-----EIVAN	---	RWVGSN	TGPTTVDLAAPGSRIVS	
cons	. *.::.. *: :		*		. : : *	. : . .

&

Cre01.g049950	TDI	-----AVN	-----RNYSAPSGT	S	FATPMVA	AVAGLVLSAL	-----GGTG	--
Cre03.g145827	TVP	-----KLI	-----GN	-----GYATYAYKT	GSSMATPMTAGV	GALVLSVL	-----GAG	
Cre04.g213400	AFEW	-YAKWYDPATDSFLP	GTPTTCWFVDQVLSVGAPNR	PDLP	AFIQYTGTSQAAPAVSGVAALMIHDA	AVRQGV	DLCAASRAGL	KGR
Cre14.g628800	SKI	-----T	-----GGCRSLSGT	SVASPVVAGAVSLLAS	TVLPD	-----KPGAGEDG		
Cre16.g685250	STP	-----PVA	-----GN	-----SMTRWNLKTG	SSMATPMVSGIGALVLSVL	-----GTG		
Cre17.g708400	TFP	-----ISL	-----GSYGNKTG	SSMATPLVAGVAALVASV	-----GAVG	PL		
Cre17.g713600	TLPDRWAQV	VDPV	TG	-----GGYGNMSG	SSAAAPVVAGVAALVVG	-----GAV		
Cre17.g735450	TLP	-----G	-----ATYGDKTG	SSMATPLVSGSAALVLSAL	-----GAV			
cons	:				:*:* ** * . . . *			

Cre01.g049950	--RAT	-PQTPLQIKNILMSSGDL	LLPSLNN--	QFKSARR		
Cre03.g145827	--TGN	-YFQGQAARKIIVESADRLVDLPV	--R--	TGSRINAQRAVQ	QAR	
Cre04.g213400	---	ARGPLNVAQIKTALARGAV	-YPGGAV	DREAYGA		G
Cre14.g628800	--RWA	-LLNPGSMKQALVEGAVRLPNLNI	--YEQGS			G
Cre16.g685250	--TGN	-YFQKEARKI				
Cre17.g708400	VASPN	-YYQASSSVKSLLE	TADKLPGLPV--	I--	GGRRINAQRAVEA	A
Cre17.g713600						
Cre17.g735450	--DGN	-YFQAVRARSLLL				E
cons						

=====

MSA *The multiple sequence alignment result as produced by T-coffee.*

T-COFFEE, Version_11.00.d625267 (2016-01-11 15:25:41 - Revision d625267 - Build 507) Cedric Notredame SCORE=64

BAD AVG GOOD (transitive consistency score) color code goes from 0 (blue) to 9 (red).

Cluster 2: Cre02.g076950 (Insert between oxyanion residue and S); Cre07.g329500; Cre10.g459450; Cre13.g585800; Cre16.g675350 (Insert between D and H).

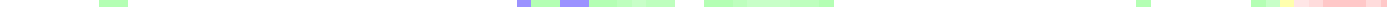
&
 Cre02.g076950 GWDVKVGIVDTGVDIRHPALG-----G-----GFGPG-----FKVAYGR-----DF-
 Cre07.g329500 GSGSHIYVLDTGIRSTHYAFS-----G-----RVGAGT-----SM-
 Cre10.g459450 -AGVTIYALDSGVFAQHDEFQ-----SW-----GTAPA-----AG-SGAATA-----GRASYGH-----DF-
 Cre13.g585800 -TGVTAIVVDSGVRGSHSTQFR-----PL-----LPPA-----QQ-PAAPTA-----PRAVSPSPGRSRVQHGY-----SVF-
 Cre16.g675350 GRGVKICIFDTGVDPGAAGLQITTDGKPK-(25 aa)-CIAGASGRKLRLNPDWKNPTGEWRVG-(56 aa)-KKERAELEGRVAALK-(58 aa)-GTF-
 cons . . * : * : : . : . : .

&

Cre02.g076950	AGDA---	FKPGGVPA----	PDADPMDCAHGHTHVAGIIAGSYSSADWSFSGVAPGVTLGAYKVFVGCTG--	GTTSELVVAALDAA----	AADGMQ--VINLSL
Cre07.g329500	-----	-----	LGSSWEDDHGHGTHVSGTAMGAT-----	YGVARAAIVHAVKVCDSQG--	SGTYSNFISGLGWVKNHVQNNGIKIQAVVSMSL
Cre10.g459450	VD-----	-----	GDAEAADCDGHGTHVASTAVGRS-----	VGVARGAELVAVRVLDCSG--	SGSIADTVAGLDWL---AKHKVRPAVAMLSL
Cre13.g585800	TATA---	GG---LGG----	SGGGSDSCYGYHGTHTVASLLGGVD-----	YGKANVTITVPVQVLDCNG--	NTSETLLKQGLEWV---GRNLVLPVVVHMSL
Cre16.g675350	SELDACNFV--	LNILDGGR	TLSVVVDDCGAHGTHVAGITAAHFDP--	DPSGNIGAPGAQIISKIGDTRLGSMETGTGVVRGLIAA----	RQH---GAHLINMSY
cons			. * . *	* : *	: . :

\$ (oxyanion residue)

Cre02.g076950	GD-EGAWG----	GPVAEAAARLA-G-	LGVLVVAAVG-NAGP-SG-	(231 aa insert)-	NALTPTVPVA--	TIPALSPQSPPPPS	GQGQGQTGGNGATQGS
Cre07.g329500	---SGPRS---	ASLNDAIQDVV-N-	NGITVVVVAAGNNGGDS	-----	CSYSPGSAPA--	AISV-----	-----
Cre10.g459450	GVPAGDW----	RVLGEAVSVLV-SQ	HGVPVVAASGNAAV-DS	-----	CGITPANLPE--	VITVAASN-----	LE-----G-
Cre13.g585800	---EGGYS---	AAVNAMVDSLT-A-	LGAVVVASAGNSNR-DT	-----	CAVSPASAQS--	AV-----	-----
Cre16.g675350	G---EPTTTPNAGR	FIELATELVRK-	HGVIFVASAGNAGP-AL	-----	TTVGAPGGTSSA	LFGVGAYV-----	-----
cons		:	* . * : *		.	.	

Cre02.g076950 QGRDSSPSSAVTASSPPPPSPPPSPAPPGSPKPGNSGNTPAKSPSPFFFPFPPSPSPPTGKPPSPPPPSPPPPAPPPPSPPPSPPPPSPPPAPPIKKW-
Cre07.g329500 -----
Cre10.g459450 -----KFNTSAR-----RPPPLQPPL-----A-----SPRPLPPLQPPPREPMYS-----
Cre13.g585800 -----
Cre16.g675350 --SP-----QLAAAGH-----
cons 

&

Cre02.g076950	(47 aa insert)-SSGSSGNSGSI-(84 aa insert)--LASPFSSWGPTDPLHIEPMLAAPGGAVLSTVPIVTRADGSRSTGYAYLSGTSQSAPYVAA
Cre07.g329500	-----GSSNRDD-----SLSSFNVGVCVF-----TFPGGSIVSAYNTADNA-----EATMSGTSMATHPVSG
Cre10.g459450	-----WSNTGKCVD-----LFAPGVEIFGACGGKDRCPAVTASAYTWASGTSMAVPAVAG
Cre13.g585800	-----AVAAVEADG-----TRWYRSNWGCVD-----LFAPGVSLLAAPVWTDSA-----SGLKTGTSM AAPFASG
Cre16.g675350	-----SVLEAPAEC-----LQYNWSSRGPTSDGHGTGVAFSAPGGAIAPVPQWTQR-----RQLMNGTSMSSPNACG
cons	* * ** : . *** *: ..

Cre03.g170300	GARLAFFDIMNKKK-----EYLL---PS---YDDL	LTHQLESGVLVQSDSWGNGKQT--FWYNG	LCHQLDEV	TWRHPELLVLTAAGNDGNM	QPP	PANGTVS
Cre03.g190250	RARLTFNDIGGTG-----GNLIN---IDN--VDE	LYLPFHKNVGAAITSDSWGPMWS---NYDGL	ARRFDSFLWRNPDM	TSFIAAGNSGDDNTR	-VG	GTVD
Cre05.g242100	GARLSVIDISTAG-----GGLTV---PSP--MDQ	TLLQVHASAGVSITSDSWGVLRS---DYTAA	ANDFDRYLWSHGDH	VALVAAGNYQSPGL	-RSG	QVVT
Cre05.g242750	RAKLSVVDISLPLV-RGDPEAGLQT---PEP--VDD	VYLRLHYEVGARVTSDSWGTDIN---IYEETA	QAYDAFLWRNPDS	IAFIAAGNNGSDALT	-PGGS	SIG
Cre05.g242856	RAKLSFVDISLTAAAAGGLQALGP---PEP--VDD	VYLRLHYDAGARITSDSWGAVAN---IYDDFA	QAYDAFLWRNPDS	IAFIAAGNNGSDALT	-PGGS	SIG
Cre09.g406700	MARLSFADLATSS-----ASLTGLSYPGDYEFEE	QTFGRHYGNGARLFSFSGVSDSSGLSYGS	WGAAADDYLWRHPET	LVVAAGNNGANGY---	MSTVQ	
Cre19.g750447	AARIAFMDLSDKG-----DGVWT---PGD--LGR	SYFNLTYRVGARVHSDSWGSDLT---AYDSMA	AALDRFTWNNPDFL	SIFAAGNYGTAS-GA-	TTTVT	
cons	*:::.*:			:	*.:***	*:*:*::.*.***

Cre03.g170300	NTLSVGSTLSLTASSTNGGALT	TLHVVTGGPEPPPEDPALVAASLPA	EDAAALMSGDT-----AALS	-----	APANAKALYG--
Cre03.g190250	NVVGVGVGKYPAGVSFA-T-LLKIRGYQPS	-----SAG-----D--VR	-----	-----	SPANAK
Cre05.g242100	NVIAVGSCLSPPFTFYYP-FYTVRGRRLG	-----SGQ-----AWSQVWPIECRDVG	ARDFLTGVGGPSGRNATLELV	HPAAAKAVSGWS	
Cre05.g242750	NIVAVGATYRVPAELGIGPQ-VLEVRGSTPG	-----TDG-----STPTE	-----	-----	TPATAK
Cre05.g242856	NPVAVGATWRVPAELGLGAR-LLQVRGYNPA	-----ADLD-----ASTSTE	-----	-----	APATAK
Cre09.g406700	NVLSAGASVNHPRDENPWANQWLPLFTYRDG	-----QGQ-----T-QQA	-----	-----	TPGSAK
Cre19.g750447	NCIAAGATPSMGTGDRP-GAYGVTVYNT	-----	-----	-----	SPAVAK-TA--
cons	*:.*.:				*.***

Cre03.g170300	SSVISALAA-QAAAAAQAASGGGSSGDS	SSSSDGSGAGFDGAVAAALVDISFGGR	KAPAAATSSPPAGAAV-KAADDE	(50 aa insert)-EGLTLE
Cre03.g190250	-----LFDLALQPA-----	-----ETGG-----	LPYLSSV-VSATSE-----	TL
Cre05.g242100	LQPRPPDFATGTLAQQQG-----	-----EDGC-----	GELPA-----PGPSF-----	AP
Cre05.g242750	-----VLRLSVYPM-----	-----ENPD-----	LPQLLNI-IPLGSQ-----	VP
Cre05.g242856	-----LFRLALYPE-----	-----ENPD-----	LPQLLRV-IPADSE-----	VP
Cre09.g406700	-----VWVT-----	-----ARGSS-----	SATWRNL-VEDQPP-----	FA
Cre19.g750447	T-----VNIGGGQGARSV-----	RVVGAAFGGD-----	LAAAVSAAAA-----	NGGVVA
cons			.	

Cre03.g170300	AVAASPLNACA---RLNDSSSASN	NNKTAA	DAASAGAAAPL	PKLLIT-----TRDG-----C--	MPYVKAAAAAAGAAALAIANDAP	GPA---VA-IL-
Cre03.g190250	VGVSTPIDACE---PLYPEA-----	LAHVP-----GRVVLVS-----	YSDR-----C--	DVDTRARNLQDAGAAAMMIH	SPNMEYT-ESPPLSN	
Cre05.g242100	AGPAAFSSSCSLPPGL-VAS-----	GALA-----GAVAVAP-----	ASPN-----C--	SLVEQVLALQRAGAAAVLLGYTE	PSNP---PTPVLVS	
Cre05.g242750	LVVADPPDACT---PLTNPA-----	AEVR-----GAVVLA-----	VRGE-----C--	WFYQKVLHVRAAGGGAVIFVNNVP	EMMDSPPAPVYPN	
Cre05.g242856	LVVADPPDACT---PLTNPA-----	AEVR-----GAVVLA-----	VRGE-----C--	WFYQKVLNVRAAGGGAIIFVNNVP	EMMDSPPYPIYPN	
Cre09.g406700	VVLASPLHACT---PLQVPVGSAS	-----PSQYQ-----	GAIVIIDPVADT	SPDGLTPPD	CQGVLDSDRASYAMEAGAVAVWFVWRY	YRTVL---SEAYVI
Cre19.g750447	LAAATPADACT---GLSNAGGV-----	R-----GRVLLV-----	QRGR-----C--	YFSDKARAAAAAGATAVIVYND	KQDGYFLMEAPDAS-	
cons	: :*	*		:	*	:. **.*:.


```

Cre03.g170300 ----VNPTGSTLPAAALTQDLGARIFAALAAG--QR-V-FARA-SWQ---PQGPYDHVPAFSGWGPTNDGRVKPDIVAPGWRVVSAYSD----FVSAGVS--
Cre03.g190250 ----SDGPAIRIPVSSIPQVMGGWLVGNVTAG--VE-L--RLT-WQD---YPQSVGDVVRWSGYGPTRDGRIKPDVIAPGYDIVSAKAVD-----GIT--
Cre05.g242100 SPSAALNSTGV TIPVGYITQSLCSWLLGNMSAS--SRPV--TFSVSL---ASQGINSVASFSSAGPALDGRVKPDLVAPGFQLTSAWKTGAGAGAGAGTT-Q
Cre05.g242750 ----SNYYAAKMATSVISQGLGQWLLGNASAG--VE-L--SLA-NTV---KNAGSDTVAFFSSYGPMADGRIKPDIVTPGVLITSAGAKG-----GIT--
Cre05.g242856 ----SNSYAANMATSVISQGLGQWLLGNASAG--VL-M--TLA-NIV---EAAGPDTVAYFSSYGPMADGRIKPDVVAPGLYITSAGARG-----GIT--
Cre09.g406700 ----QGVSVIPLPYSIITREQGDWLQQVMTDPSHSE-F--RFQ--NITYPLVNGGEESLAFFSSYGPTPDGRFKPDLVAPGQRLISASSGA-----GLRAV
Cre19.g750447 ----GGSGSGLVPMGSVPASTGRWLLAAAAAG--NTAVSIRRA-ADV---APLAFDDVAAYSSFSGPTTDGRIKPDITAPGHLLS-ANANPT--LAPDGAV--
cons      .      : . . : . :      .      .      .      : : * . * * * . * * : * : *

```

&

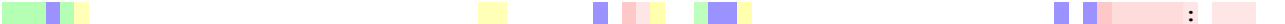
```


Cre03.g170300 D-GC-AV-RMTMEGTSMSAPLVAGAALLARQYFESGHYPAGRTDAPEAAPFSPSGVLLKAVLIGGAYDMSRSGGLSFSTLQPLLPAAPSAYQGFG
Cre03.g190250 GGQC-NDETQEMSGTSMATPLAAGSAALLRQYLRSGFYPGGWPDTQDSAPFKPSGILVKALVIAGAKS--LQGGWALQVARLLSTAPDYQGWG
Cre05.g242100 GDTG-GVSVAAARSGTSMSTPTAAGHTALVRQYLRTGFYPTGSPADAAAAPFTPSGMILLKAMVIAGAKS--MEGGEAMATRAPLGPPPDAYQGWG
Cre05.g242750 GGAC-SPAQANLSGTSMATPHAAGHTALVRQYLRTGFYPTGSPADAAAAPFTPSGMILLKAAIAGAKS--LQGGLAMALGIPMGPPPDAYQGWG
Cre05.g242856 GGAC-SPAQANLSGTSMATPHAAGHTALVRQYLRTGFYPTGSPADAAAAPFTPSGMILLKAAIAGAKS--LMGGLAMALGIPMGPPPDAYQGWG
Cre09.g406700 ETSC-SLEMRLAGTSMSTPVVAGNLAIARQYLTEGWYPAGWRGAPDSAPFVPSGPLLKAIAIAGAKS--LQGGLARAPGQLMGTPPDGYQGWG
Cre19.g750447 P-SCGAGATRILSGTSMATPVIAGSALL-----
cons      *      * * * : : * * * :


```

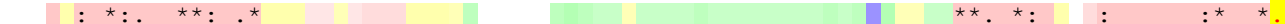
Supplemental Fig.2. Alignment of S8 domains of 21 *C. reinhardtii* family members. Phytozome identifies 21 proteins that contain the subtilisin-like S8 domain. A phylogenetic tree constructed using all 21 S8 domains, identified subclusters, which are marked. For each subcluster, catalytic core sequences were aligned using T-Coffee. Catalytic triad residues are marked, along with the oxyanion residue.


Supplemental Fig.3. Alignment of two M14 CPB-like *C. reinhardtii* and human proteins. Alignment of *C. reinhardtii* CPB-like enzymes CPZ2(Cre06.g309450) and CPZ3 (Cre07.g335900) with the three M14 domains of human CPD (CPD1, CPD2, CPD3) and the M14 domain of human CPE. M14 catalytic core sequences were aligned using T-COFFEE, Version_11.00. 8cbe486 (2014-08-12 22:05:29 - Revision 8cbe486 - Build 477). Cedric Notredame. SCORE=79. * **BAD AVG GOOD*** TCS (transitive consistency score) color code goes from 0 (blue, bad) to 9 (red, good).


CPD1	AHIKKA E -----ATTTTTSAGAEAA-EGQ F -----DRYYHEEE-L E S--
CPD2	-----Q-----PI-----Q-PK-D-F-----HHHHFPD-M E I--
CPD3	----TTK-----EF-----ETLIK D -LSAENGLESMLRSSSNLALALY R ---YHSYKD-L S E--
CPE	-----L-----QQ-----ED-GI-S-F-----EYHRY P E-L R E--
CPZ3	-WTSAAQGLRRPDADLPATSELDNGGTSAQQLGSAL-----RPL-SD T F-----VAT R NGQLTDGQ E FAS
CPZ2	-YPTFA-----Q-L S E--
cons	

CPD1	----AL---REAAA G LPGLA-----RLFSI-GRSVEGRPLWVLR L T A GLGSLI P EG D AGPD A AGPD A AG-PL L P-GR P
CPD2	----FL---RRFANE-YPNIT-----RLYSL-GKS V ESRELYVMEIS D -----NPGV H ---E-----PG-----EP
CPD3	----FL---RGLVMN-YPHIT-----NLTNL-GQ S TEYRHL I WSLEIS N -----KPN V S---E-----PE-----EP
CPE	----AL---VSVWLQ-CTAIS-----RIYTV-GR S FEGR L LVI E LS D -----NPGV H ---E-----PG-----EP
CPZ3	PPPGEFVGDEW A RY H N---HDEMVR L LQDVVN S NG A CT D EL D PPSV E GRVLYAVVAS P R-----NP-----AG G ---AV L PR P
CPZ2	----YL---A-----LRVAGSLGRCS L ASA-GKS V QGRDLWAVTIG D -----PAGVY Y PD-----PTNP D V P PFKA
cons	

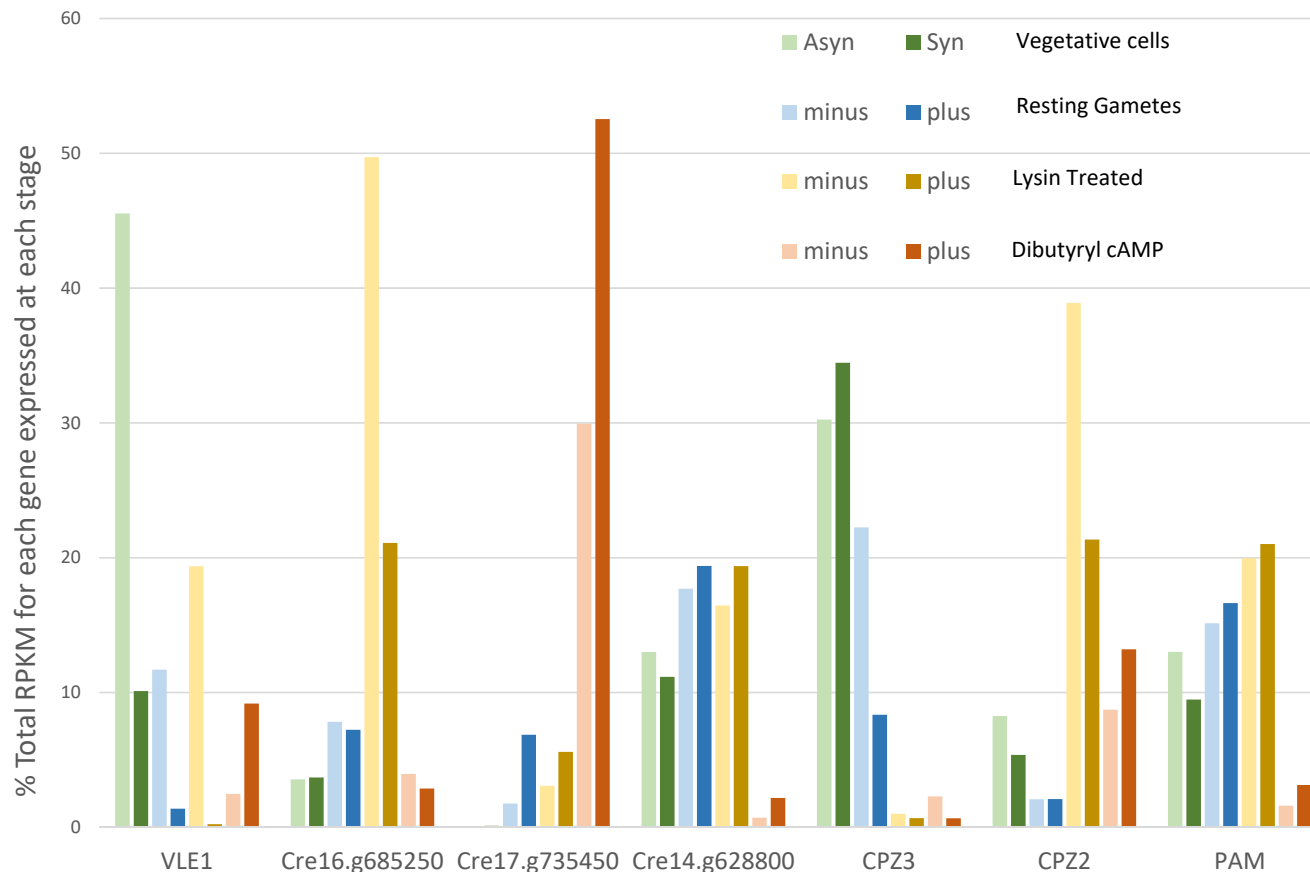
CPD1	QVKLVGNM H GD E TVSRQVLIYLARELAAGY R -----RGDPRLVR-----LLN-TTDVYLLPSLNPDGFER A REGDCGF D GGPSGA
CPD2	EFKYIGNM H GN E VVGRELLLN L IEYLCK N F-----GTDPEVTD-----LVH-N T RIHLMPSMNPDGYEK S QEGDSI-----SV
CPD3	KIRFVAGI H GNAPVGTELL L ALAEFLC L NY-----KKNPAV T Q-----LVD-R T RIVIVPSLNPDGRER A QEK D CT-----SK
CPE	EFKYIGNM H GN E AVGRELLI F LAQYLCNE Y Q-----KGNETIVN-----LIH-STRIHIMPSLNPDGF E KAA-SQ P G-----EL
CPZ3	SFTWIGNM H GD E TANRELL L RLAAG L CN-----G-----ELADSDARWKALQ S TTVRIIPTMNP D GYERR-----
CPZ2	RAAYIGVM H GD E KGHISAVLRLVGELCD P LSEPKFAPGGV L DSNVTD-----LL-GSTVLYV L PLMNP D GYT A T-----
cons	

CPD1	----SGRDNSRGRDLNR S FP-----DQ F -----STGE P PA--LD--EVPEVRALIE-WIRRN K FVLSGN L H
CPD2	----IGRNNSNNFDLNR N FP-----DQ F -----V-----QITDPTQ-----PETIAVMS-WMKS Y PFVLSAN L H
CPD3	----IGQTNARGKD L DT D FT-----N-----N-----AS-----Q P ETK A IIEN L IQKQ D FSLSVAL D
CPE	KDWFVGRSNAQ G IDLNR N FP-----DLDRIVYV N EKEGG P NNHLLKNMK I VDQ N TK--LAPETKAVI H -WIMDIPFVLSAN L H
CPZ3	----TRWNANKVDLNR N FW S VEY P YAK P TLD Q A-----RKRDRY G ATM P AA D W S KV G -N F TLQ P EA V AVSR-YLAA A PPDL S AN L H
CPZ2	----QRYNANGVDLNR N FY T SA F PFAM P TA D G-----YALQPG T SNALY N AA D W T DNGGG A HE P ETQ A VMS-WLASVRPHVSAD L H
cons	

CPD1	GGSVVASYPFDD S PE H KATGIY S KTS D DEVFKYLAKAYAS N HPIM K T-GE P -H-CP-GDEDET F KDGITNGAHWYDVEGG M QDYN Y VWAN
CPD2	GGSLV V NYPFDD D E--QGLATY S KSPDDAVFQ Q IALSY S KENSQ M FQ-GR P --CK N MY P NEYFPHGITNGASWY N VPGG M QDW N YLQ T N
CPD3	GG S MLV T YPYDK P V--Q--T V ENKE-----TLKHLASLYANN H PS M HM-GQ P -S-CP-NKSDENIPGGV M RGA E W H SLG S MKDYS V TY G H
CPE	GGDLV A NYPYDE T RS-G S AHE Y SS S PD D AI F QSLARAYSS F NP A MS D -P-NRPPCRK N DD S SFVDGTTNGGAWY S VPGG M QDF N Y L SS N
CPZ3	GGALV A NYP L DAC D -S V GALT N C P TA E E P LPRQLAA A Y V HN P N M SAQ G T-----A P FRAG T VQGA A WY P VLG S MQDWVY H HL D
CPZ2	GGALVGSYALD A CD-SRGALL D CP S PE A PLPGYLANV S M N HP S MR F -SW G -E-VQ-AS R QVQ F FN G T Q GATW P AI G TI A DWL H HT Y R
cons	

CPD1	C F EIT L EL S CKY P PA-SQLRQEWENN R ESLIT L IE-K V HIGVKG F VK D SITGS L ENATIS V -AG--IN H NIT T GR--FG D -FYRL L V P
CPD2	C F EV T IELGCVKY P LE-KELPN F WEQ N RRSLIQ F MK-Q V HQGV R GFVLDATDGRG I LNATIS V -AE--IN H PT V TY K --TG D -YWRL L V P
CPD3	C P EIT V YT S CCY F PSA-AR L PS L WAD N K R SL S ML V -EV H KGV H GFV K DK T -G K PI S KAVI-V L N E --GI-K V Q T -K--EG G Y F HV L L A P
CPE	C F EIT V EL S CEK F P P E-ET L KTYWED N K N SLIS Y LE-Q I HRGVKG F V R DLQ-G N PIANATIS V -EG--ID H D V TS A K--DG D -YWRL L IP
CPZ3	RLML T IEL H EVKD P PA-TALDD L WGQ N SA A FL R IME-LAG M GLR G R V LDQ Q T-RAP L AATIT V -T S PASSR P ADA-DAAR G GY F Y K PM A P
CPZ2	RHML T IEL H YKY A MF L ND K LNMYAT N R P S M LRLAG V GA H MGLR A LL R DAAS-GA A LAAT V TP-QQ P AGAW A PE V -E--SGGLV V K M AM P
cons	

Expression of subtilisin- and CPB-like and PAM transcripts: data from Ning et al., 2013



Supplemental Fig.4. Expression of *C. reinhardtii* subtilisin-like, CPB-like and PAM transcripts. Relative transcript levels (% total RPKM for each transcript in *plus* and *minus* mating type cells at each stage examined) are shown for the *C. reinhardtii* genes encoding proteins most like human furin, PC1, PC2, CPD, CPE and PAM [19]. Stages examined in this study were: Vegetative cells (Asy, asynchronous, and Syn, synchronous); Resting Gametes; Lysin-treated gametes; Dibutyryl-cAMP treated gametes.