

Supplemental Figure 1. Affinity purification and MS-based protein component identification of StrepII-tagged Clp assemblies. Native PAGE analysis of affinity-purified StrepII-tagged Clp assemblies (350 and 200kDa) from total leaf extractions showing the gel lanes at full-length. On-column salt incubation dissociated the individual rings from the core revealing the Strep-tagged Clp ring. Additional bands at 720 and 500 kDa were observed but were also found in mock purifications (with wild-type). MS/MS analyses of these bands indicate the co-purification of the Chaperone 60 complex (700 kDa) and the highly abundant Ribulose-1,5-bisphosphate carboxylase oxygenase (RUBISCO) complex (~550 kDa)



Supplemental Fig. 2 Legend on next page

Supplemental Figure 2. Representative extracted ion chromatograms (XICs) of the peptides derived from 100 fmol Clp-QconCAT protein. The numbers designate the position and identity of the peptide (see Supplemental Table 2). 32 out of the 37 Clp-QconCAT peptides can be detected and quantified. All the quantifiable Clp-QconCAT peptides eluted within the first 45 minutes of the run. The % acetonitrile concentration (red dashed line) was estimated based on the mobile phase gradient used with correction for dead volume during the chromatographic run.



Supplemental Data. Olinares et al. (2011). Plant Cell 10.1105/tpc.111.086454

Supplemental Figure 3. Phylogenetic analysis of the ClpP/R proteins from proteobacteria, photosynthetic bacteria, green algae and plants. The phylogenetic tree inferred for the 118 Clp sequences from various prokaryotes and photosynthetic eukaryotes employing the Maximum Likelihood (ML) approach (Felsenstein, 1981) with the General Time Reversible (GTR) model for nucleotide substitution (Tavar, 1986) and a gamma model of substitution rate heterogeneity (Yang, 1996). Support values are shown from 1,000 non-parametric bootstrap inferences. This includes the Escherichia coli ClpP; the ClpP from alphaproteobacteria presumed to be the ancestor of mitochondrial organelles: *Rickettsia prowazekii* (Rp) and *Wolbachia wmel* (Ww); the cyanobacteria which are the ancestor of plastids: *Anabaena sp. (strain PCC 7120)* (As), *Synechocystis sp. (strain PCC 6803)* (Ss), *Synechococcus elongatus* (strain PCC 7942) (Se), *Microcystis aeruginosa* (Ma) and *Prochlorococcus marinus* MED4 (Pm); the green alga *Chlamydomonas reinhardtii* (Cr) and representative plant species: the bryophyte *Physcomitrella patens* (Pp), the lycopod *Selaginella moellendorfii* (Sm), the dicots *Arabidopsis thaliana* (At) and *Populus trichocarpa* (Pt) as well as dicots *Zea mays* (Zm) and *Oryza sativa* (Os). The plastid ClpP1 homologues for the ferns *Psilotum nodum* (Pn) and *Adiantum capillus renesis* (Ac) and the gymnosperm *Pinus thunbergii* (Pnt) were also collected from recent chloroplast genome sequencing projects for these species. The Cr_P1* is the plastid-encoded ClpP1 in green alga with the large insertion (IS1) sequence removed. Ec_P was used as the outgroup.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol **17**: 368-376. **Tavar, S.** (1986). Some probabilistic and statistical problems on the analysis of DNA sequences. In Some mathematical questions in biology-DNA sequence analysis., R.M. Miura, ed (Providence, R.I.), pp. 57-86.

Yang, Z. (1996). Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol. Evol. 11: 367-372.

Cap	10	20 30	40	50		60 70	80	90 100	110
Ec_P Biero B		LVPMVIEQT	SRGER SFDIYS	- RLLKERVIFL	GQV	EDHMANL		- EKDIYLYINSPO	GVI
Wawm_P		LIPIVVEQTS	SRGERAYDIYS	-RLVKERIIFV	GPI	EDNMANL	I VAOLLFLESEN P	NKDICMYINSPO	G V V
CrP2		YV <mark>P</mark> IVIEQT	TRGERAYD I FS	- R L L <mark>K E R</mark> I V V I M	GG I	DDNTSNS	VVAQLLFL <mark>ES</mark> QAP	- DKPITMYINSPO	GV <mark>V</mark>
PhyP2 SmP2			SRGERAYDIFS ARGERAYDIFS	- R L L K ER I VC I I	GPI		VVAQLLYLESEHP-	- DKPIHLYINSPO	GVV
AIP2		LI <mark>P</mark> MVIEHS	SR <mark>G</mark> ERA <mark>YD</mark> I F <mark>S</mark>	RLLKERIICI	G P I	NDDTSHV	VVAQLLYLESENP-	- SKPIHMYLNSPO	GHV
PIP2-1 PIP2-2		LIPMVIEHS	SRGERAYDIFS SPGERAYDIES	- R L L K ER I VC IN	GPI ·		VVAOLLELESEN P	- SKPIHMYLNSPO	GHV
PIP2-3		LIPMVIEHS	SRGERAYDIFS	RLLKERIVCI	G P I	DDDTSHV	VVAQLLFLESENP	- SKPIHMYLNSPO	GQV
OsP2-1		LVPMVIEHT	SRGERAYDIFS	RLLKERIVCI	GPI		VVAQLLFLESENP-	- AKPVHLYINSPO	G V V
ZmP2m-1		LVPMVIEHTS	SRGERAYDIFS	RLLKERIVCI	GPI	TDDTASL	VVAQLLFLESENP	AKPVHLYINSPO	G V V
ZmP2m-2		LVPMVIEHTS	SRGERAYDIF S	- RLLKERIVCI	GPI ·		VVAQLLFLESENP-	AKPVHLYINSPO	• • • • • • • • • • • • • • • • • • •
ZmP2p-1 ZmP2p-2			SRGERAYDIFS	-RLLKERIVCI	GPI	ADDTASL	VVAQLLFLESEN P	LKPVHLYINSPO	GVV
SeP1		<mark>P</mark> I VVEES	GRGERAFDIYS	- R L L <mark>R E R</mark> I I F L C	EPV	TSD VANR	I VAQLLFLEAEDP	- EKDIYLYINSPO	••••• <mark>6</mark> 8 <mark>7</mark> ••••
SeP2 SsP1		PTVIETS	GRGDRAFDIYS	-RLLRERIVFLO	QEV	RD ENANL	VVAQLLFLEAEDP	- EKDIYLYINSPO	GSV
SsP2	• • • • • • • • • • • • • • • • • • • •	VV <mark>P</mark> MVVEQS	G M G E R A F D I Y S	- R L L <mark>R E R</mark> I I F L C	T P V	DDQVAD S	I VAQLLFLDA ED P	- EKDIQLYINSPO	<mark>G</mark> S <mark>V</mark>
AsP1 AsP2			GRGERAFDIYS GMGERAFDIYS	RLLRERIVFLO			I VAQLLFLDAED P	- EKDIQLYINSPO	GSV GSV
PmP1		<mark>P</mark> LVLEES	GGSERVFDIYS	-RLL <mark>RER</mark> IIFLO	EQV	T SD TANR	I VAQLLFLEAEDP	- EKDIYMYINSPO	••••••••••••••••••••••••••••••••••••••
PmP2 MaP1		AVPTVVEQS	GRGERAFDIYS		TGI	NDQVSDS			GSV
MaP2		VL PVVIEQS	GMGERAFDIYS	RLLRERIVFLO	T P V	DDQ VAD S	IVAQLLYLEAEDP	- EKDIQLYINSPO	G <mark>SV</mark>
C/P5			ER FQGVVS		GAV	DDDMANL	LVAQLLYLDSVDN -	- KRDITMYVNSPO	GSV
PhyP5-2			ER FQSVIT	-QLFQQR IIRC	GAV	DDDMANL	IVAGLLYLDAADP	- NRD I VMY VN SPO	GSV
SmP5			ERFQSVIT	- QLFQHRIIRCO	GAV	DDDMANL	IVAQLLYLDAADP-		G S V
PIP5-1			ERFQSVIS	-QLFQHRIIRC	GAV	DDDMANI	I VAQLLYLDAVDP	- NKD I VMYVN SPO	GSV
PIP5-2			ERFORMU		C	MAN I	I VAQLLYLDAVDP	- NKDIVMYVNSPO	GSV
ZmP5			ERFQSVVS	-QLFQHRIIRCO	GPV	EDDMANI	IVAQLLYLDAIDP	-NKDIIMYVNSPO	G S V
PhyP3-1			D LD PA T	- MLLQQR ILFL	SQV	D D F <mark>S</mark> AD V	II SQLLLLDAQDP	- TKDIKLFINSPO	GSV
PhyP3-2 SmP3				LLQQR IVFI	SQV	DDFSADL	II SQMLLLDAQDP		GSV GSV
AlP3			E LD T TN	- MLLRQR IVFL	SQV	DDMTADL	VI SQLLLLDAED S	- ERDITLFINSPO	GSI
PIP3 ZmP3			ELDTTN		SQV	DDVTADF	II SQLLFLDAEDP	-KKDIKLFINSPO	GSV
OsP3			E LD T T N	- MLL <mark>RQR</mark> I VFLC	S P V	DDMSADL	II SQLLLL <mark>D</mark> A EDK	- TKDIKLFINSPO	G S I
CrP4 PhyP4-1			QGD P F G		HQM	EDFGADA	IVSQLLLLDAVDP		G S V
PhyP4-2			EGD VMG	- LMLKOR IVFLO	NQM	DDFVADS	IISQLLLLDAIDP	-KODIRLFINC PO	G S I
SmP4			E SD VMC	- LLLROR TVFLO	HQL		II SQLLLLDAMDS		GSV
PIP4-1			E TDAMG	- LLLRERIVFLO	NNI	DDFVADA	II SQLLLLDAQDP	- TKD I R L F VNC PC	GSL
PIP4-2			ETDAMG	- LLLRERIVELO			II SQLLLLDAQDP		GSL ····
CrR6			EMDLYR	YLLSNRIIFIC	GYI	NDKMATQ	IVGSLMALEAVDE	- N ED I R I YINSPO	GQ P
PhyP6-1 PhyP6-2					SPV	NSEVAQS	VVSKLLTLAAIND -	- EEEIKMYINC PC	G S T
SmP6 SmP6			ADLTS	- IMLR SRIIFIC	Q P I	N SQ VAQR	VISQLVMLAAIDE.	- EKDIQLYINCPO	GST
AIP6	•••••		· · · · · · LDLSS	- VLFRNRIIFIC	Q P I	NAQ VAQR	VISQLVTLASIDD -	-KSDILMYLNCPO	••••••••••••••••••••••••••••••••••••••
PIP6-2				- VLFRNRIIFIC	QPI	N SQ VAQR	VI SQLVTLATIDE.	-NADILMYLNC PC	GST
OsP6	• • • • • • • • • • • • • • • • • • • •		LD LA T	- VLLGNRIIFIC	QYI	N SQ VAQR	VI <mark>SQLVT</mark> LAAVDE-	- EADILIYLNC PO	• • • • • • • • • • • • • • • • • • •
SeP3	MPIGVPSVPYRLPC	G S S - F	ERWIDIYN	- RLAMERIIFIC	QEV	TDGLANS	I VAQLLYLDSEDS-	- SKPIYLYINSPO	GSV
SsP3	MPIGVPSVPFR LPC	g sq - Y	ERWIDIYT	- R L <mark>SQ</mark> ER I I F L C	Q E V	NDSIANR	I VA FLL YL <mark>D SDD P</mark> -	- SKPIYLYINSPO	<mark>G</mark> S <mark>V</mark>
PmP3 MaP3	MPIGTPSVPYRLPC	GSQ - Y	ERWVDIYI	- RLGVERILFLO	QEV	TDGLANA	LVAQMLYLDSDDN -	- TKPIYLYINSPO	GSV GSV
AsP3	MPIGVPKVPYRMPC	GGQ - F	TD <mark>WISIY</mark> D	- R L Y <mark>R ER</mark> I I F L (R D V	D D E I A N Q	I I A VMLYLDSEDP-	GKDIYLYINSPO	G <mark>MV</mark>
CrP1 PhyP1	MPIGVPRIIYCWGEELP MPIGVPKVPFRLPC	GEEDA				DDELCNQ	LIGIMMYLNGEDE		G S V G S V
SmP1-1	TPIGVPKVSFR ISC	GEEDM	VR I DTHN	- R L YR ER L L F L C	QNM	DDEIANQ	IICIMMYLNGENR -	- SKDIYLYINSPO	GAV
PnP1 AcP1	MPIGVPKVPFRLPC	GEEDA				DDELANQ			GAV
PntP1	MPVGVPKVPFRAPC	GD EDA	TWVD L YN	- R L <mark>YR ER</mark> L L F L A	QDI	NHE ANQ	LMGLMVYL SA ED S -	- NKDIFSFINCPO	••••••••••••••••••••••••••••••••••••••
AIP1 pip1	MPIGVPKVPFRSPC	GEGDT			QEV		LISLMIYLSIEKD -	- TKDLYLFINSPO	GWV
OsP1-1	MPIGVPKVPYRIPC	GDEEA		- VMYRERTLFLC	QEI	RCEVTNH	ITGLMVYLSIEDG	- I SD I FLFINSPO	
OsP1-2 0sP1-2	MPIGVPKVPYRIPO				QEL	RCEVTNH	ITGLMVYLSIEDG	- I SDIFLFINSP	GWL
OsP1-4	MPIGVPKVPYRIPC	GDEEV		- VMYRERTLFLC	QEI	RCEVTNH	ITGLMVYLSIEDG	- I SD I FLFINSPO	GWL
ZmP1-1 ZmP1-2	MPIGVPKVPYRIPC	GDEEA			QEI ·	RCEITNH	ITGLMVYLSIEDG	- NSDIFLFINSPO	GWL
C/R2	MPVSVPKVLCR PPC	GQRQS	EWVD LWE	- A YTYQK VVFI	EAL	TED VANN	MIALTLYLDSLDQ	- KRIYYWLNVPC	GD V
PhyR2-1	APIGTPKVPYRTPC	GEGGW		- VLYRERIIFLO	Q Y I		VVATMLYLDSVDN -	- SKPMYLYINCPC	GD 1
SmR2-1	MPIGTPKVPYRTPC	GEGTW		- VLYRERIIFIC	QH		VLA TMLYLDSVDS	- SKDLHFYMSCPC	GDL
SmR2-2	MPLGVPKVPYRASS	SEGSW		- VLYQERIVFIC	QHI		VVATMLYLDSVDS-	- SKHLHFYISC PC	GSL
PIR2	MPIGTPRVPYRTPC	GEGTW		-ALYRERILFIC	QN I	DEEFSNQ	ILATMLYLDT IDD -	- SKRMYLYINGPO	GDL
OsR2 ZmR2	MPVGTPRVPYRTPC	GEGTW		ALYRERIIFIC	DSI	DEEFSNQ	VLASMLYLDSVDN -	- TKKILLYINGPO	GDL
SeR .	mr igirkvr tRTPC	a = a 11V	PPDLPS	LLLKERIIYLC	MPLFS	SDD VKRQ VG FD VTEL	I IAQLLYLEFDNP	- EKPIYFYINST	T SWY TGDA I G YE
SsR MaR			PPDLES	- LLLKERIVYLO	MPLFS	SD EVKQQ VG I D V TQL	I IAQLLYLQFDDP	- DKPIYFYINSTO	T SWY TGDA VG F E
maR PmR			PPDLES	LLLKERIVYLC	LPLFS	DDDAKRQLGMDVTEL	I IAQLLYLEFDDP.	DKPIYFYINST	TSWYTGDATGYE
AsR			PPDLP	LLLKERIVYLC	MPL	· · · · · · · · · V <mark>P</mark> A V T E L	I VAQLLYLQSDDP	-DKPIKIYINST	TSGYSGEPIGFE
PhyR1-1 CrR1			PPDLPS	- LLLQQRVIVI	MA	UPAVTEL	VVAQCYYLDFDDRN	IRQRPIYVYLNSS	CINDKGQAISAD
PhyR1-2			<mark>P PD L P</mark> S	- LLLDARICYL	MA I	V <mark>P</mark> A V T E L	I I A E L L WLD FDN P	- SKPIYFYINSS	TQN EKKESIGFE
SmR1 AIR1			PPDLPS		MAL	VPAVTEL		- SKPIYLYINC SC	TQNEKKESVGFE
PIR 1-1			· · · · · PPDLPS	- LLLDARICYLC	MPI	VPAVTEL	LVAQFMYLDYEDP	SKPIYLYINSS	TQNEKMETVGSE
PIR1-2 OsR1			APDL PS		MPI	VPAVTEL			TMDENNELVASE
ZmR1			A PD L PS	LLLDSRIIFLO	MPI	VPAVTEL	IAAQFLWLD YDDR	TKPIYLYINST	TMD ENNEL VASE
CrR3 PhyR2.1			PPDL PS		MPL	VPAVTEL	MVAELLYLEKQGA-	- TLPIEMLINSS	
PhyR3-2			PPDLPS	LLLNSRIVYIC	MPL	VPAVTEL	VIAELLYLQYADP	ROPIYVYINST	TARADGETVGLE
SmR3			PPDLPS		MPL		VIAELLYLQYSDP-	- TL PIYMYINSTO	TSRADGETVGME
PIR3-1			PPDLPS	- L L L HGR I VYIC	MPL	VAAVTEL	VVAELMYLQWMDP	-KAPIYIYINST	TTRDDGESVGME
OsR3 ZmR2			PPDLPS	- LLLHGRIVYIC	MPL	······································	VVAQLMYLEWMNS-	- KEPVYIYINSTO	TARDDGEPVGME
CrR4			PPDLAT	-RLFSERIMYLC	MPI		LTAQLEVLVQEAP	- DPIFFYINSTG	AK STTKFGNE
PhyR4-1			PPDLPS	- YLF <mark>K</mark> NRIVFL	MSL	· · · · · · · · · V <mark>P</mark> A V T E L	VMAELLYLQYDDP	- EKSIYMYINSTO	TTK - DGEKLGYE
PhyR4-2 PhyR4-3			PPDLPS		MSL	VPAVTEL	FMAELLYLQYDDP-	- EESIYMYINST	TTK-DGEKLGYE
SmR4			PPDLP	- YLFHNRIVFL	MSL	V <mark>P</mark> A V T E L	IMAEFLYLQFVDP	- EKPIYMYINSTO	TTK - DG EKLGYE
ABR4 PIR4			PPDLAS		MSL	VPSVTEL	I LAEFLYLQYEDE	- EKPIYLYINST	TTK - GGEKLGYD
OsR4			<mark>P PD</mark> LA <mark>S</mark>	- FLYKNRIVYLO	MCL	V <mark>P</mark> SVTELI	MLAEFLYLQYDDA	- EKPIYLYINSTO	TTK-NGEKLGYE
∠mR4			PPDLAS		MCL	VPSVTEL	MEAEFLYLQYDDA -	- ENPIYMYINST	EKLGYE

Supplemental Figure 4. Protein sequence alignment:

_	127	13	7 147	157	167	177	_	187 197	207	217	227	237
Ec_P Ripro P	TAGMSIYDTMQFIKF TAGLAIYDTMQFIKF	KVA	TICMGQAA SMGA FLLTAG TLCIGQAC SMG SLLLCGG	- A K G K I - EH G M	R F C L PN S R V M I H Q P L R Y S L PH S R I M I H Q P S	<mark>GGY</mark> - QC 6 - <mark>GGY</mark> - КС	8	- ATDIEIHAREILKV - ATDIEIHARETLKI	KGRMNELMALH Krllnelyskh	IGQSLEQIER KQELKYLER	SMERDN FMS	EEAKKFGLVDSIL
Wawm_P	TAGLSIYDTMQYINF	DVS	TLCIGQAA SMGSLLLAAG	- TKGK	YSLPHSRIMIHOPS	GGY-HO	<mark>0</mark> -	- ATDIEIHANEILRV	KKLNQIYEKH	GNSLKKIEG		EEARKIGLIDRVI
PhyP2	TAGLAIYDTMQYIRS	5 PVS	TLC VGQAA SMASLLLAAG	- Q PG E	R SLPN SKIMLHOPL	GGA - SC		- A SDIAIQAREILDM	RARLSRLYVKH	IGH TYEK I ES	SMERDHFMS	EEAKQFGLVDEVI
SmP2	TAGLAIYD TMQYIRS	PVS	T LC VGQAA SMGSLLLAAG	- EPGQ	R SL PNAR VMVHQ PS	GGA - SO	0	- A SDIAIQAREILDM	R SR LNKLYAKH	GSSIEKIDO	SMERDH FMS	E EAKEFGLIDEVI
PIP2-1	TAGLAIYDTMQYIRS	<mark>P V</mark> N	TICLGQAASMASLLLASG	AKGE	KALPNATIMINOPS	6 - <mark>GGY</mark> - SC		AKDLTIHTKQIVRV	WDALNQLYCKH	GKPIDVIQH		EEAKEFGIIDEVI
PIP2-2 PIP2-3	TAGLAIYD TMQYIR S	 PVT	TICLOGAA SMASILLAAG	- AKGE	KALPNATIMIHOPS	6 - GGY - SC - GGY - SC	8	- AKDMTIHTKOIVRV	WDALNOLYCKH			E EAKEFGIIDDVI
OsP2-1	TAGLAIYDTMQYIR	<mark>Р</mark> V Т	T L C I GQAA SMA S L L LAAG	ARGE	RALPNAR VMIHOPS	s - <mark>GGA</mark> - SC	0	- A SDIA IHAKEILKV	R DR LNK I YAKH	SQAIDRIEC	CMERDMEMD	EEAHDWGLIDEVI
OsP2-2 ZmP2m-1	TAGLAIYD TMQYIRC TAGLAIYD TMQYIRC	PV T	TLCIGQAA SMGSLLLAAG TLCIGQAA SMA SLLLAAG	- A R G E	RALPNARVMINOPS RALPNARVMINOPS	6 - GGA - QG 6 - GGA - SG	8	- A TO IA IQAKEILKL - A SO IA IHAKEILKV	R D R L N K I YQKH R D R L N K I YAKH	GQEIDKIEG	CMERDLFMDF CMERDMFMDF	EEARDWGLIDEVI
ZmP2m-2	TAGLAIYD TMQYIRS	PVT	TLCIGQAA SMA SLLLAAG	ASGE	RRAL PNAR VMIHOPS	GGA - SO	<mark>0</mark> -	- ASDIAIHAKEILKV	RDR LNK I YAKH	GQAIDRIEC	CMERDMEMD	EAHDWGLIDEVI
ZmP2p=1 ZmP2p=2	TAGLAIYDTMQYIRC	PVT	TLC I GQAA SMGSLLLAAG	- A PGE	RALPNAR VMIHOPS	GGA - QG		-ATDIAIQAKEILKL	R DR LNK I YQKH	ROPIDKIEG	CMERD LFMD	DEARDWGLIDEVI
SeP1	YDGLGIFDTMNHIRF		TVC VGLAA SMGA FLLAAG		TSLAHSRIMIHOPL	- GGA - QG	8	- AKDIEIQANEILYI	QN LN EVLA ER		D T D R D F F M SA	SEA VEYGLIDRVI
SsP1	SAGLGIFDTMNQIRF	D VC	TICI <mark>G</mark> LAA <mark>S</mark> M <mark>G</mark> AFLLSAG	AKGK	R M S L PN S R I M I H Q P L	- GGA - QG	0	- ATDIEIQAKEILYL	KALLNQHLANH	IGK SLEEI TA	DTERDFFMSA	EESKEYGLIDQVI
SsP2 AsP1	TAGMGIFDTMKHIRF		TIC FGLAA SMGA FLL SGG TIC TGLAA SMGA FLL SAG	- CKGK	RMALPSSRIMIHOPL RM <mark>S</mark> LPHSRIMIHOPL	- GGA - QG - GGA - QG		- AVEIEIQAREILYI	KORLNTMLVEH KRRLNDYLAEH	IGQ PMEKLQE IGQ PIERIAE	D T ERD F FM SA	EEAKEYGLIDQVI DEAKDYGLIDQVI
AsP2	YAGMAI YD TIQQIR F	DV V	TICFGLAA SMGAFLLTAG	- TKGK	RMSLPDSRIMIHOPL	- GGA - QG	<mark>9</mark> -	- AIDIEIQAREILYI	KAQLNQLLANH		DTDRDFFMSA	EEAKNYGLIDQVI
PmP2	TAGLAIYDTMQQISF	D VV	TICFGVAA SMGAFLLSGG	- AKGK	LAL PN SRIMINGPL	- GGA - QG	0	-AVEIEIQAKEILFL	KTLNSLLAKH	NOSLEKINE	DTERDYFLSF	EEAVEYGLIDKVI
MaP1 MaP2	SAGLGIFDTMNQIRF	DVS	TICIGLAA SMGAFLLSAG	- K PGK	RMSLPNSRIMIHOPL RMSLPSSRIMIHOPI	- GGA - QG	8	- A T D I E I QAKEILYL	COR LN EHLASH	GOPYDRIAN		E EALDYGLIDQVI
CrP5	TAGMAVEDTMRHIRE	D V S	TCC I <mark>g laa s</mark> m <mark>g</mark> afila <mark>s</mark> g	QAGK	R Y S L PN S R I M I H Q P L	- GGA - QG	0	- ATDIEIQANEILHH	KLTLNGYLAQF	IGQ SMET I TH	DTDRDFFMSF	QEAIEYGLVDAII
PhyP5-1 PhyP5-2	TAGMAIFDTMRHIRF TAGMAIFDTMRHIRF	DVS	TVC VG LAA SMGA FLLSGG TVC VG LAA SMGA FLLSGG	- TKGKI - TKGKI	RYSLPN SRIMINOPL RYSLPN SRIMINOPL	- GGA - QG - GGA - QG		- QTDIEIQANEILHH - QTDIEIQANEILHH	KAN LNGYLAYH Kan Lngylayh	IGQ PYEKIVO IGQ PYEKIVO	D T D R D F F M SA	KEAQEYGLIDAVI KEAKEYGLIDAVI
SmP5	TAGMAVED TMRHIRE	DVS	TVC V <mark>G</mark> LAA <mark>S</mark> M <mark>G</mark> A FLL SAG	- TKGK	RYSLPNSRIMIHOPL	- GGA - QG	<mark>0</mark> -	- QSDIEIQANEILHH	KANLNSYLSYH			K EAQ EYGLIDGVI
PIP5-1	TAGMAIFDTMRHIRF	D V S	TVC VG LAA SMGA FLL SAG	- TKGK	RYSLPNSRIMIHOPL	- GGA - QG	G	-QSDIDIQANEMLHH	KANLNGYLAYH	IGQ SLEKING	DTDRDYFMSA	KEAKDYGLIDGVI
PIP5-2 0xP5	TAGMAIFDTMRHIRF	DVS	TVC VGLAA SMGA FLL SAG TVC IGLAA SMGA FLL SAG	- TKGK	RYSLPNSRIMINOPL RYSLPNSRIMINOPL	- GGA - QG	G	- Q S D I D I QAN EMLHH	KAN LNGYLAYH			KEAKDYGLIDGVI
ZmP5	TAGMAIFDTMKHIRF	D V S	TVCIGLAA <mark>S</mark> MGAFLLSAG	- T <mark>KG</mark> K	R Y S L PN S R I M I H Q P L	- GGA - QG	<mark>0</mark>	- ETDLEIQANEMLHH	KANLNGYLAYH	IGQ P L D K I N V	DTDRDYFMS	K EAK EYGLIDGVI
PhyP3-1 PhyP3-2	TAGMGIYDAMKLCKA	DVS	TVC FGLAA SMGA FLLAAG TVC FGLAA SMGA FLL TAG	- TKGKI	RYCMPNAR VMIHOPL	- GGA - GG	T	- AIDGALOVREMMYH	KYKLQKIMSRL KYKLQKIMSRL			WEAKEYGLVDGVI
SmP3	SAGMG I YDAMK LC RA		TICFGLAA SMGAFLLASG		YCMPNAR VMIHOPL	- GGA - GG	V	- A I D VG I Q VR EMMYH	KAKLNKILSRL			WEAK EYGL UDAVI
PIP3	TAGLGIYDAMKLCKA	D V S	TICLGLAASMGAFLLSAG	- SKGK	FCMPNGRVMIHQPL	-GTA - GO	к	- A SEMSIRIREMSYH	KIKLNKILSRI	GKPLEQVEN		WEAKEYGLVDEVI
ZmP3 OsP3	TAGMGVYDAMK FCKA		TVC FGLAA SMGA FLLAAG TVC FGLAA SMGA FLLAAG	- TKGK	R YCMPNAR IMIHOPS R FCMPNAR IMIHOPS	GGA - GG	К к	- VTEMGLQIREMMYE	KIKINKIMSRI KIKINKILSRI	GK SEEQIDE		PSEAKDYGIVDNII
CrP4	TAGMGIYDAMMLCRA	D VN	TYCFGLAA SMGAFLLGAG	KRGK	R N SM PN SR IMIHQ PL	- GGA - SC	9	- A VD I E I QAKEIMYH	KAN LNR I MAD Y	QQPLSKIEE	DTDRDRYMS	LEAKEYGLIDHII
PhyP4-1 PhyP4-2	SAAMAIFDATOLCRA SAAMGIFDAMQLCK	DIS	TICFGLAASTAAIVLAGG TICFGFAA <mark>ST</mark> AALLLAAG	- TPGK	R LAMPN SRIMMHOPL R LAMPN SRIMMHOPN	- GGA - SC 1 - GGA - SC		- AIDVEIQAKEVMYH	K SN VIRILSEI	IGRAPEQVER	DIDRDRYMS	EAMEYGIIDGVI
SmP4	SAAMGIYDAIQLCRA		TIAFGLAASTASVILAGG	- TKGK		1 - GGA - SO	9	- AIDVEIQAKEIMFH	KNN FTRILAEI			VEAMEYGLIDGVI
P1P4-1	SA SMA I YD VVQ L VRA	DVS	TIALGISASTASIILGGG	TKGK	R FAMPN TRIMINQPL	- GGA - SC	0	AIDVEIQAREIMHN	KNNIIRIVSGF		DMDRDR YMS F	I EAVEYGIIDGVI
P1P4-2 OsP4	SA SMA IYD VVQLVRA SA TMA IYD VMQLVRA	DVS	TIALGISASTASIILGGG TIGLGIAGSTASIILGGG	- TKGK - TKGK	R FAMPN TRIMINOPL	GGA - SO / - GGA - SO	8	- AIDVEIGAREIMHN - ALDVEVGAKEILTN	K N N I I R I V SG F K R N V I R L I SG F	ISR IVEQ VQH I <mark>g</mark> r IPEQ Veh		LEAVEYGIIDGVI
CrR6	YSVLGVVDAMQSIK	D V Q	TVALGACYSYASLVVAAG	- TKGK	YAMKN TR LMMTOPN	1 - GG S - QG	D -	- I YQIKATVEELNAL	YQIFSRYYMKF		ATCRDHEMT	EOAKLEGLIDEII
PhyP6-2	YSIMAIFDAMQWIKF	D V S	TVCFGIAASQGALLLAGG	- KKGK	R FAMPNAR I MI HQ PQ	- GGC - GC	. т	- ADD VRRQVN EVMSS	RDKIDKMYSTF		FTERDYFFSA	AEALDFGLIDGLL
SmP6 AlP6	YSVLAIYDCMSWIKF YSVLAIYDCMSWIKF		TVC F G MAA S Q G A L I LA G G TVA F G VAA S Q G A L L LA G G	- TKGLI - EKGMI	R FAMPN SR VM I HQ PQ	9 - GGC - GC 9 - TGC - GC	T -	- MED VRROVNEVVAS	R D K I D K M YAA F Q K I D R M YAA F	GOPLEIVO1	YTERDHFFSA YTERDRFLSA	A EAMEFGLIDGLL
PIP6-1	YSVLAI YDCMSWIK	K VG	TVCFGVAA SQGALLLSGG	- EKGM	YAMPNAR IMIHQPQ	- GGC - GG	H -	· · VEDVRRQVNEAVQA	RHKIDQMYATF		YTERDRFLST	SEAMEFGLIDGIL
OsP6	YSILAIYDCMSWIK		TVC FG VVA SQGALLLAGG	EKGM	YAMPNAR VMIHQPQ	- G V S - E C	N		RDK VDKMFAAF		WTERDRFLS	SEAMDFGLVDALL
ZmP6 SeP3	YSILAIYDCMSWVS- TAGMAIYDTMOYIK	HLC	SLSF-LI [^]	SKGK	LALPH SRIMINOPL	GGTGRE	0		KLLNQIMADR			
SsP3	TAGMAIYD TMQYIKA	EVI	TICVGLAA SMGAFLLASG	- A PGK	R LA L PHAR I MI HQ PN	I-GGTGRF	i <mark>a</mark> -	- ATDIDIEAREILRI	RQQLNEIMAQR	I GQ T VEK I AK	DTDRDYFLSA	AEAKEYGLIDKVI
MaP3	TAGMA I YD TMQ Y I KA	EVV	TICVGLAASMGAFLLGAG	- SPGK	LAL PHARIMINOPL	I-GGTGRF	0	- ATDIDIEAKEILRI	RQQLNEIMANR	GQTIERIEM		E EA VAYGLIDKVV
AsP3 CrP1	T SGLAIYD TMQHIKS GNGITVHDALOFIKA		TICVGLAA SMGSFLLAAG TLALGVAA SAASLALAGG	- TKGK	RMALPHSRIMIHOPS YVTEGCHTMIHOPE	GGT-RG	8	- A T D I E I EAREILRI	RHQLNQI YANN RLDVAFI YSLS	SQ <mark>PL</mark> AKIEK VRPRHKILE		QEAMEYGLIDRVI
PhyP1	LAGISVYDAMQFVVF	D VH	TICMG LAA SMG SFILAGG	EITK	RIAL PHAR VMIHOPA	- SSYYD	0	-AGECIMEAEEVLKL	RDYITRVYVQR	G <mark>k plwv i</mark> se	DMERDVFMSA	QEAKTYGIVDLVA
SmP1-1 PhP1	L PG I SVYDAMQFVVF	DVH	TIC WG LAA SMG SFIL TGG TIC MG LAA SMG SFIL TGG	- EITK	RIAL PHAR VMIHOPA	- SSFYEC		- AGECIMEAEEILKL	RDCITSIYARK	EKPLWVVSE EKPLWVISE		KEARAYGI VDLIA
AcP1 DelP1	LAGISVYDAMQFVVF	P D VH	TICMGLAA SMGSSISAGG	- E I TR	RIAL PHAR VMIHOPA		<mark>0</mark>	- AGEC VMEAEEASKL	RDCITKVYAQR	GKPLWLISE		E EAQDYG VADLVA
AIP1	I SGMA I YD TMQ F VR F	D V Q	TICMGLAASIASFILVGG	AITK	RIAF PHAR VMI HQPA	- SSFYEA	9	TGEFILEAEELLKL	RETITRVYVQR	GKPIWVISE	DMERDVFMSA	TEAQAHGIVDLVA
PIP1 OsP1-1	I PGIAIYD TMOFVRF I SGMAICD TMOTVTF	D VQ	TVC MG LA A SMG SFIL VGG TIC LGIA A SMA SFIL LGG	- K I TK - E <mark>P</mark> TK	LAFPHARVMIHOPF IAFPHARIMLHOPA	- AAFYEA - SAYYRA	R -	- IGEFVLEAEELLKL - TPEFLLEVEELHKV	REILTRVYAQR REMITRVYALR	IGK PLWVVSE IGK PFWVVSE	DMERDVFMSA DMERDVFMSA	LA EAQ VHG I VD L VA
OsP1-2	I SGMA I FDTMQTVTF	DIY	TICLGIAA SMASFILLGG	- EPTK	RIAFPHARIMLHOPA	- SA YYRA	R	- TPEFLLEVEELHKV	REMITRVYALR	GK P FWV V S E		DEAKAYGL VDI VG
OsP1-4	I SGMA I FDTMQTVTF	D I Y	TICLGIAASMVSFILLGG	EPTK	RIAF PHAR I MLHQPA	- SA YYRA	R	TPEFLLEVEELHKV	HEMITRVYALR	I <mark>g</mark> k <mark>p</mark> Fwv v se	DMERDVFMSA	DKAKAYGLVDIVG
ZmP1-1 ZmP1-2	ISGMAIFDTMQTVTF ISGMAIFDTMQTVTF	DY	TICLGIAA SMA SFILLGG TICLGIAA SMA SFILLGG	- EPTK - EPTK	RIAFPHARIMLHOPA	- SAYYRA - SAYYRA	R -	- TPEFLLEVEELHKV	REMITRVYALR REMITRVYALR	IGKPFWVVSE IGKPFWVVSE	DMERDVFMSA DMERDVFMSA	DEAKAYGL VDIVG
CrR2	VPTLAL YDTMQ YVRS	K TA	TVC YGLCLGMGG FLL TAG	GEKGY	R FAMPHSILMMHHPS	GAS-RO	9	- A SEMHIESRELVRM	R D Y L S L L T S N A		ELSRNKWMD	KQAIEYGMIDKVL
PhyR2-7 PhyR2-2	TPCMAIYD TMESIKS	PVG	TLALGYAYNNAGLLLAAG	- AKGM	VAMPLTR VALMSPA	-GAA-RO	0	ADDIINEAKELNRV	RNYVYEQVSKK			NEALEYGIIDKVV
SmR2-1 SmR2-2	TPSMAIYDTMGSVKS	R MG	TMALGYAYN IAGFLLAAG	- EKGLI	R T S M P L T R I A M Q P P Q R M A M P L A R L A L Q P P F	- GAA - RO	K	- AADIQNEAKELSRV	R D Y L FKQ LA VK	IGO PEEKIHK	D F S R I K R F S A	
AIR2	TPSLAIYDTMKSLKS	PVG	THC VG LA YN LAG FLLAAG	EKGH	FAMPLSRIALQSPA	- GAA - RO	<mark>0</mark> -	ADDIQNEAKELSRI	RDYL FNELAKN			EEAIEYGLIDKIV
OsR2	TPCMALYDTMUSLKS	PVG	THC VG FAYN LAG FIL TAG THC LG FAFN LAG FILAAG	- EKGNI	R TGMPLCRISLQSPA	- GAA - RO - GAA - RO		- ADDICNEANELLRI	K NYLYSKL <mark>S</mark> EH	T <mark>gappekis</mark> k T <mark>ghpvd</mark> kihe		EGALEYGIIDRII
ZmR2 SeP	T PC MALYD TML SLKS	PIG	THCLGFAFNLAGFILAAG	- EKGS	TGMPLCRISLQSPA		<mark>8</mark>	- ADDIENEANELIRI	KNYLYGKLAGH KRTMLELEARN			EGALEYGIIDRIV
SsR	TEAFAICDTLNYIK	P P VH	TICIGQAMGTAAMILSSG	- TKGY	A SL PHATIVLNONR	- TGA - QG	0	- ATDIQIRAKEVISN	QTMLEILS LN	GQ TQEK LAN		AQAK EYGLIDRVL
MaR PmR	TEAFAICDTLN YIKF	 PIH	TICIGQAMGTAAMILSAG TICIGQAMGTAAVILSSG	- TKGFI - TKGHI	A SLPHATIVLNONR AALPHASIVLHOPI	- TGA - QC - SGA - RC	- 8	- ATDIQIRAKEVIAN	KQTMLEIFSKN KKSMLEILSRN	IGQTTEKLAN IGKSIEELSM	DMDR TFYLTF DSDRMSYLNF	QQAKDYGLIDRVL QEALDYGVIDRIL
AsR R	TEAFAIYDTMKYIK	PI H	TICIGSAMGMAAMLLSAG	- TKGC	ASLPSSSIILHOPK	- SYA - QO	<mark>0</mark> -	- ATDIQIRAREVLVN	KG SMVD I FART			YQAKDYGLIDRVF
CrR1	N E F YA I WAA LG F TRA	<mark>P</mark> L Y	TGVTWKAQNQAAVLLSAG	- QKGR	YSEPHAKISTAPPV	/MNRV-F0	- 0	- A VDAQLQANELDYA	TKY <mark>Y</mark> AA I LARS	GKDLETCQM	QYLSRKRYF	VKEAYEEGL VDKL
PhyR1-2 SmR1	TEAYAIADTIRYVKS	KVY	TINC GOAF GOAAMLLSIG TVNC GMAF GOAAMLLSVG	- EKGF	RALQPSSSTKLYSPK RAVOPSSSTKLYAPK	VNQS-SO	S	STDMWIKGKELDAN	TDYYIDLLASG	IGK SKEELAN		QEAIDYGLADKII
AIR1	TEA YA IADTI SYCKS	<mark>D</mark> V Y	TINCGMAFGQAAMLLSLG	KKGY	A VOPHSSTKLYLPK	VNRS - SC	A -	- A I DMWI KAKELDAN	TEYYIELLAKG	GK SK EQ I N E		QAAIDYGIADKIA
PIR1-1 PIR1-2	TEA TATAD IMSTCKS	Υ		ANGY			- A	AVUMINIKANELDAN	LIAKG			GEAID TGIADK LL
OsR1 ZmR1	TDA FAIAD FINR SKS	KVY	TINESMAYGQAAMLESEG		GVL PNSITKLYLPK	VHKS-GO	A -	- A I DMWI KAKELDTN	TDYYLELLSKG	VGK PKEELAE		QEALDYGLADTIL
C/R3	SEGVALTSTMGFIK	<mark>P</mark> I S	TVNMGLAVGWSCVVLSFG	RKGW	KSLPHSLAMIQQPR	- VPP - TO	QF	QAIEVHIKWREVLDY	KRELLRMF SLG	GLPVDKLDA	DMQRPLYMR	QDALEYGIIDEII
PhyR3-1 PhyR3-2	TEGFAIYDTLMN VRN		TVAVGAAIGQACLLLASG	- DKGM - DKGM	YMMPHATAMIQQPR	: - L P S - T C : - L <mark>P S</mark> - T C		ISAIDVHIRAKELINN	RDTLVGLLAKH	IGN PYEK VAR	VMCRPFYMNF VMERPFYMNF	KKAVEFGVADKIL
SmR3 AIR3	TEGFAIHDTMQLVKN		VGVGAAIGQACLLLAAG			- MPS - MO		A SDVI LPAKEVINN	DIL VELLA EN			KKAMD FG VVDK I L
PIR3-1	TEGFAIYDSMMQLK	I EIH	TVA VGAA I GQAC LLLAAG	- TKGH	YMMPHAKAMIQQPR	- V <mark>P S</mark> - SC	LMP	- A SD VL IRAKEA VIN	RDVLTELLAKH	IGN SVETVAN		RRAKEFGVIDKIL
OsR3 ZmR3	SEGFAIYDAMMRMKI SEGFAIYDAMMRMKI	EIH	TLCIGAAAGHACLVLAAG TLCIGAAAGHACLVLAAG	- KKGK	YMFPHAKAMIQQPR YMFPHAKAMIQQPR	t - I <mark>PS</mark> - YO t - I <mark>PS</mark> - YO	MMQ	- A SDVVI RAKEVVHN - A SDVVI RAKEVVHN	RNTLVRLLARH RNTLVKLLARH	IGN PPEKIDK Ign ppekidk	VMRGPFYMD VMRGPFYMD	LKAKEFGVIDKIL
CrR4 PhyP4.1	HEALAVY SMMKGVQK		LC VGNA FGEAALLL SAG	SPGK	AALRSSTIMLROPL	- ORL - GO	- MQ		TATMAK YLAAC			YEAVSYGLIDTVL
PhyR4-2	TEAFAIYDTMR YVK	P PIF	TLC VGNAWGEAALLLAAG	- SKGN	ACLPSATIMLKOPI	-AQF-RO	0	- A TD I D I ARKEVRNV	KEELVKLLSRH	IGQ SPEK I TO		DEAVDYGIIDKVL
PhyR4-3 SmR4	TEAFAIYDTMRYMK F	P PIF	TLC VGNAWG EAALLLAAG TLC VGNAWG EAALLLAAG	- SKGN - AKGN	ACLPSATIMLKOPI SALPSTTIMLKEPI	- AQF - RO	8	- A TO I DIARKEVRNV - A TO I DIARKEVRNU	KDLVELLERH	IGQ PPEKITC IGQ PADKIEE		D EA VD YGIIDKVL
AIR4 PIR4	TEAFAIYD VMGYVK	P P F	LC VGNAWG EAALLL TAG		SAL PSSTIMIKOPI	- ARF - QO	<mark>9</mark>	- A TO VE LARKEIKHI	TEMVKLYSKH			TEAVEYGIIDKVV
OsR4	TEAFAI YDAMR YVK	/PIF	TLC VGNAWG EAALLLAAG	AKGN	RAALPSSTIMIKOPI	- GR F - QG	0	ATDVDIARKEIRNV	KIEMIKLLSRH	GKSVEELAC	DIKRPKYFS	SEAVDYGIIDKVL
∠mR4	LALAVYOAMRYVK	• • • <mark>• 1</mark> F	LC VGNAWGEAALLLAAG	AKGN	GAALPSSIIMIKQPI	- GRE-Q	🖬		N I EM FIHF	1 📴 D 🚾 K N *		



cDNA sequence alignment:





	460		470 480	490	500	510	520	530	540	550	560	57) 580	590
Ec_P/1-564	TOCTOACCOCAGOG	· · · <mark>6</mark>	CAAAAGGTAAACGTTT	TT <mark>BCCTB</mark> C	CGAATTCGCGCGT	GATGATTCAC	CAACC <mark>GTTG</mark>	- GGCGGCTAC	CA <mark>GGGC</mark> -	· · · · · CA <mark>G</mark>		COACCEATA	TCGAAATTCATGC	CCGTGAAATTCTGAAA
Wowm P/1-564	TECTTECAECTEET	A	CAAAAGGTAAGCGCTA	CTCTCTAC	CTCATTCAAGAAT	TATGATACAT	CABCCATCT	GGTGGTTAT	····CATGGT	CAA		CAACTGATA	TAGAAATACATG	TAATGAAATTTTGCGG
CrP2/1-564	CTGCTGGCGGCGGGGG	<mark>G</mark>	CCAA <mark>BGGCCAG</mark> CGCCG	GAGTCTGC	CCAACTCGCGCAT	CATGCTGCAC	CA <mark>G</mark> CCTCTG	- <mark>GGCGCCG</mark> CC	··· <mark>gaggg</mark> ·	<mark>C</mark> A <mark>G</mark>	• • • • • • •	CCAC <mark>BB</mark> ACA	TCATGATCCGGGC	GCAGGAGATCATGCGC
PhyP2/1-555	TOCTAGCTGCAGGA	· · · · <mark>C</mark>	AACCTGGAGAGAGACG	CTCATTAC	CTAATGCCCGCGT	CATGATTCAT	CABCCTTCC	- GEGEGETECC	····AGTGGC·	CAA		BCGAGTGAT	TTECTATTCAAGC	TCGTGAGATACTTGAT
AtP2/1-564	TGTTGGCAGCGGGT		CCAAGGGGCAAAGGCG	GTCACTTC	CAAATGCAACTGT	TATGATTCAT	CAGCCTTCA	GGAGGATAT	· · · AGTGGA	CAG		CTAAAGATA	TAACGATTCATAC	AAAACAGATTGTTCGT
PtP2-1/1-564	CTTTT <mark>GGCTTCT</mark> GGG	· · · <mark>6</mark>	CTAAGGGTGAGAGGAA	GGCACTTC	CAAA <mark>TG</mark> CAACAAT	TAT <mark>G</mark> ATTCAT	CA <mark>G</mark> CCATCA	- <mark>ggt</mark> gggtac	···A <mark>GCGG</mark> A-	· · · · · CA <mark>G</mark>		CAAA <mark>gg</mark> a <mark>t</mark> i	TGACAATTCACAC	CAAGCAGATAGTTCGG
PtP2-2/1-564 0 PtP2-3/1-564 0	CTCTTGGCTGCCGGG		CTAAGGGCGAGAGGAA CTAAGGGCGAGAGGAA	GGCACTAC GGCACTAC	CAAATGCAACAAT CAAATGCAACAAT	CATGATTCAT	CAGCCATCA	- GGTGGGTAC	···AGCGGA·			CAAAGGACA Caaaggaca	TGACAATTCACAC	AAAGCAGATAGTTCGG AAAGCAGATAGTTCGG
OsP2-1/1-564	TCCTCGCCGCGGGGC	6	CGCGCGGCGAGCGGCG	GGCGCTCC	CCAACGCGCGCGT	CATGATCCAC	CAGCCGTCG	GGCGGCGCG	····	CAG		CGTCCGACA	TCGCCATCCACGC	CAAGGAGATCCTCAAG
OsP2-2/1-564	CTECTCECCECCEC	<mark>G</mark>	CECECEEEEAECECCE	GE <mark>CE</mark> CTEC	CCAAC <mark>GCGCGGGGT</mark>	CATGATTCAC	CA <mark>G</mark> CCA <mark>T</mark> CC	- <mark>GEGEGCGC</mark> G	• • • <mark>Ca<mark>gggc</mark> •</mark>	· · · · · <mark>CA</mark> G		BCCACC <mark>G</mark> ACA	TC <mark>B</mark> CCA <mark>T</mark> CCA <mark>BB</mark> C	CAA <mark>gg</mark> a <mark>gatteteaag</mark>
ZmP2m-1/1-564 C ZmP2m-2/1-564 C	CTCCTCGCCGCGGGGGGGGGGGGGGGGGGGGGGGGGGG		CEAGCEGEAGAGAGECE	GECECTEC	CGAACGCGCGGGG	CATGATCCAC	CABCCCTCC	- GGTGGCGCG				CCTCCCACACA	TCBCCATCCACCC	CAAGGAGATCCTCAAG
ZmP2p-1/1-564	TCCTCGCGGCTGGG	<mark>6</mark>	CGCCGGGTGAGAGGCG	CCCCCTTC	CCAAC <mark>GCCAGGGT</mark>	CATGATCCAC	CAGCCCTCG	- GGCGGCGCG	· · · CAGGGG	<mark>C</mark> AA		CCACC <mark>G</mark> ACA	TCGCCATCCAGG	CAA <mark>gg</mark> a <mark>gatcctc</mark> aaa
ZmP2p-2/1-564	CTCCTCGCGGCTGGG	· · · <mark>6</mark>	COCCOGGTGAGAGGCG	C <mark>BCBCT</mark> BC	CCAAC <mark>G</mark> CCA <mark>GGG</mark> T	CATGATCCAC	CA <mark>B</mark> CCC <mark>TC</mark> B	- GGCGGCGCG	· · · CA <mark>GGGG</mark> ·	· · · · · CA <mark>G</mark>		CCACTGACA	TCGCCATCCAGGC	CAA <mark>bg</mark> a <mark>gatcctcaa</mark> g
Se P1/1-558 Se P2/1-564	CIGCICGCAGCCGGA-		CCAAAGGTAAGCGCAC	CAGCIIGG a a cactto	c d a d t d c t c d d a t	cateatecac	caaccocte	- GGIGGIGCI		····CAG		CIAAGGACA	ttgaaatccaage	dadadaaateetetat
SsP1/1-558	CTCCTTAGTGCTGGA	<mark>6</mark>	CCAAA <mark>ggt</mark> aa <mark>g</mark> cgcat	GA <mark>gttt</mark> gc	CCAACTCCCGCAT	CATGATTCAC	CAACCCCTA	GGGGGAGCC	··· <mark>ca<mark>gggc</mark>·</mark>	<mark>C</mark> AA		CCAC <mark>GG</mark> ATA	TTGAAATTCAAGC	CAAAGAGATTCTTTAT
SsP2/1-564	tgctttccggtggc	· · · • •	gtaagggaaaacgtat	gg <mark>ctttg</mark> c Gabtctac	ctagttctcgcat	catgattcac	CAACCTOTA	- ggggggagcc	···· <mark>cagggt</mark> ·	····· <mark>c</mark> aa		CAACTOAT	ttgaaattcaggc	ca <mark>ggg</mark> aaatcot <mark>gt</mark> ac BCBABAAATCOTTTAC
AsP2/1-564	TGTTAACAGCAGGA	A	CTAAAGGTAAACGTAT	GTCCCTAC	CTGATTCCCGCAT	TATGATTCAC	CAACCCCTC	GGTGGCGCT	···CAAGGT	<mark>C</mark> AA		CCATTGATA	TAGAAATCCAAGC	AAGGGAAATTCTTTAC
PmP1/1-558	TACTTGCAGCTGGA	<mark>A</mark>	CTAAAGGTAAAAGAAG	TAGCCT <mark>G</mark> A	GACATTCAAGAAT	CATGATTCAT	CAACCTCTT	· GGAGG <mark>T</mark> GCA	· · · AGAGGT ·	<mark>с</mark> аа		CCA <mark>gt</mark> ga <mark>t</mark> a	TAA <mark>g</mark> aa <mark>tt</mark> caa <mark>g</mark> c	TGATGAGATTTTATAT
PmP2/1-564 MaP1/1-558	CTCCTCAGTGCCGGC	A	AACCGGGCAAAAGGTT	AGCITTAC Gageetac	CCAACTCGCGGAT	CATGATTCAC	CAACCACTT	- GGCGGCGCC	···CAAGGT	····CAA		CAGI IGAAA CCACCGATA	TCGAAATACAAGC	CAAAGAAATACTATTT CAAAGAAATTCTCTAT
Ma P2/1-564	CTECTTTETEETCE	<mark>G</mark>	C <mark>ggcaggt</mark> aaac <mark>ggat</mark>	GTCCCTAC	CCA <mark>GT</mark> A <mark>G</mark> CCGGAT	CATGATTCAC	CAACCCC <mark>T</mark> C	GGCGGCGCC	· · · CAA <mark>ggt</mark> ·	<mark>с</mark> аа		SCTAGTGATA	TCGCCATCCAAGC	CAAA <mark>g</mark> a <mark>g</mark> atcctctat
CrP5/1-528 F	ATCCTGGCCAGCGGC-	· · · · <mark>C</mark>	AGGCGGGGCAAGCGCTA	CAGCCTGC	CCAACAGCCGCAT	CATGATCCAC	CABCCCCTG	- GGTGGCGCC	CAGGGC-	· · · · · · CAG		BCCACT GACA	TTGAGATCCAGGC	CAACGAGATCCTGCAC
PhyP5-2/1-528	CTTCTTAGCGGGGGGT		CGAAAGGTAAAAGGTA	CAGTCTAC	CGAATTCACGAAT	CATGATTCAC	CAGCCTTTG	GGAGGTGCT	CAAGGG	CAA		CAGACTGACA	TTGAAATTCAGGC	TAACGAGATTCTACAC
SmP5/1-528	CTTCTTAGTGCAGGC	A	CAAAA <mark>gg</mark> aaa <mark>gcgtt</mark> a	TAGTTTAC	CAAA <mark>TTCT</mark> A <mark>gg</mark> at	CATGATTCAC	CAGCCGCTC	· GGTGGCGCT	· · · CAA <mark>GG</mark> A-	<mark>CA</mark> G		CA <mark>GT</mark> CC <mark>G</mark> ACA	TTGAGATCCAGG	CAA <mark>tg</mark> aaa <mark>tcttg</mark> cat
AtP5/1-528 0 PtP5-1/1-528 0	CTGCTTAGTGCTGGA	A	CCAAAGGAAAAAGATA	CAGECTAC CAGCTTGC	CAAACTCAAGGAT	AATGATCCAT	CAGCCGCTT	- GGTGGAGCT	···CAAGGT	· · · · · · · · · · · · · · · · · · ·		CAAACCUACA Caatctgaca	TAGATATCCAGGC	AAATGAAATGCTGCAT AAATGAAATGTTGCAT
PtP5-2/1-453	CTGCTTAGTGCTGGG	<mark>A</mark>	CCAAA <mark>gg</mark> aaaaaga <mark>g</mark> ata	CA <mark>GCTT</mark> GC	CGAATTCAAGGAT	AAT <mark>G</mark> ATCCAT	CABCCTCTT	- <mark>GGT</mark> GGAG <mark>CT</mark>	· · · CAA <mark>ggt</mark> ·	<mark>GGG</mark>		CAA <mark>TCTG</mark> ACA	TAGATATCCAGG	AAA <mark>cg</mark> aaa <mark>tgttg</mark> cat
OsP5/1-528	CTGCTTAGTGCTGGG	A	CAAAAGGGAAGCGATA	CAGCTTAC	CTAACTCAAGAAT	AATGATCCAT	CAACCTCTC	- GGAGGAGCC	CAAGGA-	CAA		BAGACTGAT (TTGAGATCCAGGC	TAATGAGATGCTGCAT
PhyP3-1/1-522	TECTEECCECAGE		CGAAAGGCAAACGGTA	CTGCATGC	CCAACGCAAGGGT	GATGATTCAC	CAGCCGTTG	GGTGGCGCC	· · · GGTGGC ·	ACA		CCATTGAC	GCGCTTTGCAGGT	GCGGGAGATGATGTAT
PhyP3-2/1-522	TTGCTGACAGCCGGG	A	C <mark>G</mark> AA <mark>GGGCAAGCGGT</mark> A	CTGCATGC	C <mark>B</mark> AAC <mark>BCCABBB</mark> T	GATGATTCAC	CAGCCGTTG	- <mark>GGT</mark> GG <mark>CG</mark> CC	· · · <mark>GG<mark>T</mark>GG<mark>C</mark> ·</mark>			CCA <mark>T</mark> AGAC	GCGCGTTGCAAGT	GAGGGAGA <mark>tgatgtt</mark> c
SmP3/1-519 AtP3/1-522	CTTCTTGCTTCTGGA	A	CAAAAGGGAAACGGTA	TTGTATEC	CAAAIGCTCBTGT CTAACTCTAAAGT	TATGATCCAT	CAACCACTE CACCACTE	- GGTACTOCT	····GGAGGC·			CAACCEACC	TGAGCATACAAGT	AAGAGAAAATGATGTAC
PtP3/1-522	TTCTTTCTCCTCCTCCTCCT	· · · <mark>1</mark>	CCAAA <mark>BGC</mark> AA <mark>B</mark> A <mark>GATT</mark>	TTGCATGC	CCAA <mark>T</mark> GGGAGGGT	GATGATTCAT	CA <mark>G</mark> CCACTT	- GGAACTGCT	· · · <mark>gg</mark> agg <mark>c</mark> ·	AAA		CATCA <mark>g</mark> aaa	TGAGCATCCGGAT	TAGAGAAATGAGTTAC
ZmP3/1-522	TACTTGCTGCTGGG-	A	CAAAGGGTAAGAGGTA	CTGCATGC	CAAATGCGAGGAT	TATGATCCAT	CAGCCATCA	- GGTGGCGCT	66 <mark>6666</mark> -	AAA		STCACAGAGA	TEGEACTACAGAT	AAGAGAGA <mark>TG</mark> ATGTAC
GrP4/1-522	CTGCTCGGCGCGGGGC		AGCGCGGCAAGCGCAA	CAGCATEC	CCAACTCCCGCAT	CATGATTCAC	CARCCCCTC	GGTGGCGCC	AGCGGC	CAG		CCGTGGAT	TTGAGATCCAGG	CAAGGAGATCATGTAC
PhyP4-1/1-522	GTACTTGCCGGAGGT	<mark>A</mark>	C <mark>GCCGGGGGAAGCGATT</mark>	<mark>gg</mark> caa <mark>t</mark> gc	CTAATTCTCGTAT	TAT <mark>GATGCA</mark> T	CA <mark>GCCCTTG</mark>	- <mark>ggaggtgct</mark>	· · · A <mark>GT</mark> GGA ·	· · · · · <mark>CA</mark> G	• • • • • • •	CAATT <mark>G</mark> AC	TCGAAATTCAAGC	CAAA <mark>gaggtt</mark> at <mark>gt</mark> ac
PhyP4-2/1-522 C SmP4/1-510 4	CTACTCGCAGCAGGT	A	CEAAGGGGGAAGCGACT	GGCGATGC TGCAATGC	CTAACTCTCGCAT	CATGATECAT	CABCCCATE	- GGAGGTGCT	···AGTGGG·	·····CAG		CAATTGAT	TTGAAATCCAAGC	TAAGGAGGTCATGTAC
AtP4/1-522 4	ATTCTTGGTGCGGGA	A	CTAAAGGCAAGCGCTT	TECTATEC	CCAACACGAGGAT	AATGATTCAT	CAACCTCTT	GGAGGTGCA	· · · AGCGGT ·	<mark>C</mark> AA		CTATAGAT	TTGAGATTCAAGC	TAAGGAAGTTATGCAT
PtP4-1/1-522 A	A <mark>TTCTTGG</mark> A <mark>GGTGG</mark> A-	A	C <mark>G</mark> AAA <mark>GG</mark> CAAAC <mark>GTTT</mark>	TECAATEC	CCAACACAC <mark>GG</mark> AT	AATGATTCAT	CAACCTCTT	- <mark>GG</mark> A <mark>GGAGCT</mark>	· · · AGTGGG ·	<mark>с</mark> аа	· · · · · · ·	CAATAGAT	TGGAAATTCAAGC	ACGAGAAATTATGCAT
Os P4/1-522	ATCCTTGGTGGTGGC	A	CAAAGGGCAAGCGATT	CECCATEC	CCAACACACGGGAT	AATGATCCAT	CAACCTCTT	GGAGGCGCA	···AGTGGG-	CAA		CAATAGAT CCTTGGAT	TCGAGGTCCAGG	CAAGGAGATATTGACT
CrR6/1-522	G <mark>TGGTGGCT</mark> GCGGGG	A	CCAA <mark>ggg</mark> caa <mark>gcg</mark> cta	C <mark>GC<mark>G</mark>A<mark>TG</mark>A</mark>	A <mark>g</mark> aacac <mark>gcggc</mark> t	CATGATGAC	CAGCCCATG	- <mark>GG</mark> C <mark>GG</mark> C <mark>T</mark> CC	· · · <mark>Caggg</mark> t -	<mark>G</mark> AC		A <mark>tct</mark> acca <mark>g</mark> a	TCAA <mark>gg</mark> ccacag <mark>t</mark>	GGAGGAGCTCAACGCA
PhyP6-1/1-519 PhyP6-2/1-519	TACTICCAGGAGGC	A	AGAAGGGCAAGCGTTT AAAAAGGCAAGCGGTT	CECAATEC CECAATEC	CAAATGCTCGTAT	CATGATICAT	CAACCGCAA	- GGAGGGTGT - GGAGGCTGC	··· GGGGGA·	AC		CAGACGATO CAGACGATO	TTAGAAGGCAAGT TTAGAAGACAAGT	GAACGAAGTAATGTCC Gaacgaagtcatgtcc
SmP6/1-519	ATT <mark>ctagcc<mark>ggt</mark>gg</mark> a	<mark>A</mark>	C <mark>G</mark> AAA <mark>ggcttg</mark> cgttt	TGCGATGC	CAAAC <mark>T</mark> CCA <mark>G</mark> AGT	CATGATCCAT	CAACCTCAA	- <mark>GG</mark> A <mark>GGCTG</mark> T	· · · <mark>GG<mark>T</mark>GG</mark> A-	AC <mark>G</mark>		A <mark>TGG</mark> AGG <mark>AT</mark> O	TGAGAA <mark>GGC</mark> AAGT	GAA <mark>tg</mark> aa <mark>gtggt</mark> agca
AtP6/1-519	CTTCTTGCTGGAGGT	· · · <mark>6</mark>	AAAAA <mark>gg</mark> aa <mark>tgcgct</mark> a	T <mark>GCAAT</mark> GC	CAAATACTCGTGT	CATGATACAT	CAACCACAA	- ACTGGATGC	····	· · · · · CAT		STAGAGGAC	TGAGGAGACAGG	CAA <mark>TG</mark> AAGCCATCGAA
PtP6-2/1-519	CTTCTTGCTGGTGGA		AGAAGGGAATGCGATA	TGCAATGC	CTAATGCACGTAT	TATCATACAT	CAACCGCAG	GGCGGCTGT	···· GGGGGT·	CAT		GTGGAGGAT	TGAGACGCCAAGT	AAATGAAGCAGTCCAA
Os P6/1-519 🖌	A <mark>tacttgctggcggt</mark> -	<mark>G</mark>	A <mark>gaagggaatgcgtt</mark> a	T <mark>g</mark> cca <mark>tg</mark> c	CAAA <mark>TGCT</mark> A <mark>g</mark> agt	AAT <mark>G</mark> ATTCAT	CAACCTCAA	- <mark>ggtgtatc</mark> a	··· <mark>GAGGGT</mark> ·	<mark>AA</mark> T		S <mark>T</mark> GGAGG <mark>AGC</mark>	TGAGGC <mark>G</mark> ACAGGT	T <mark>gggg</mark> aaaccatttat
ZMP6/1-234 -														
Se P3/1-582	CTGCTCTGCGCCGGC-	_A	GCAAAGGAAAACGGCT	C <mark>G</mark> CA <mark>TT</mark> GC	CTCATTCGCGGAT	TATGATTCAC	CAACCGCTG-	- GGGGGAACGI	GGTCGCCGT-	CAG		CTTCCGATA	TTGAAATCGAAGC	CAAA <mark>g</mark> aga <mark>ttctgcg</mark> a
Se P3/1-582 Ss P3/1-582	CTECTCTECECCEEC TTETTAECCTCCEEE	A	GCAAAGGAAAACGGCT Caccaggaaaacggtt	C <mark>GCATTG</mark> C Agccctgc	CTCATTCBCGGAT CCCATGCCCGGAT	TAT <mark>GATTCAC</mark> TAT <mark>GATCCAC</mark>	CAACCOCTO	- GGGGGGAACG - GG <mark>CGGT</mark> ACT	GG <mark>TCGCCGT</mark> - GGCCGCCGT	CAG		CTTCCGATA CAACGGACA	TTGAAATCGAAGC	CAAAGAGATTCTGCGA Cagggaaattctccgc
Se P3/1-582 Ss P3/1-582 Pm P3/1-582 Ma P3/1-582	CTECTCTECECCEGE ITETTACCCTCCEGE ITATTEGEAECTEGT ITACTEGECTCCEGT	A A	GCAAAGGAAAACGGCT Caccaggaaaacggtt, Ctaaaggtaaaagggt Gtccgggtaagcgtct	CCATTOC ACCCCTOC TCCTCCTC CCCCCTC	CTCATTCGCGGAT CCCATGCCCGGAT CTCATAGTCGAAT CCCATGCGCGGAT	TATGATTCAC TATGATCCAC TATGATTCAT TATGATTCAC	CAACCGCTG CAACCCATG CAACCACTG CAACCACTG	- GGGGGAACG - GGCGGTACT - GGAGGGACT - GGCGGAACC	GGTCGCCGT GGCCGCCGT TCACAACGT GGGCGCCGG	CAG		SCTTCCGATA SCAACGGACA SCTAGCGATA SCGACGGATA	TTGAAATCGAAGO TTGACATTGAAGO TTGAAATAGAAGO TCGACATTGAAGO	CAAAGAGATTCTGCGA CAGGGAAATTCTCCGC TAGAGAAATACTGAGA TAAGGAAATCCTGCGT
Se P3/1-582 Ss P3/1-582 PmP3/1-582 Ma P3/1-582 As P3/1-579	CTGCTCTGCGCCGGC TGTTAGCCTCCGGG TATTGGGAGCTGGT TACTGGCTCCGGT TGTTAGCTGCTGGT	A	GCAAAGGAAAACGGCT Caccaggaaaacggtt Ctaaaggtaaaaggt Gtccggtaaggtaa Ccaaaggtaagcgtct Ccaaaggtaagcgcat	CECATTE ASCCCTEC TECTCTTC CECCCTTC BECATTAC	CTCATTCGCGGAT CCCATGCCCGGAT CTCATAGTCGAAT CCCATGCGCGGAT CCCATGCGCGGAT	TATGATTCAC TATGATCCAC TATGATTCAT TATGATTCAC TATGATTCAC	CAACCOCTO CAACCCATO CAACCCATO CAACCCATO CAACCCATO CAACCTTCT	- GGGGGGAACG - GGCGGTACT - GGAGGGACC - GGCGGAACC - GGTGGAACC	GGTCGCCGT GGCCGCCGT TCACAACGT GGGCGCCGG ···CGTGGA	CAG CAA CAG CAA		SCTTCCGATA SCAACGGACA SCTABCGATA SCGACGGATA SCAACCGATA	TTGAAATCGAAGO TTGACATTGAAGO TTGAAATAGAAGO TCGACATTGAAGO TCGAAATCGAAGO	CAAAGAGATICIGCGA CAGGGAAATICICCGC TAGAGAAATACIGAGA TAAGGAAATICIGCGI CAGAGAAATICIACGG
Se P3/1-582 Ss P3/1-582 Pm P3/1-582 Ma P3/1-582 As P3/1-579 CrP1/1-579 Phy P1/1-585	CT GCT CT GCGCCGGC TGT TAGCCT CCGGG TAT TGGBAGCT GGT TT ACT GGCT TCCGGT TT TACT GCT GCT GGT CT TT AGCT GGT GGT CT TT AGCT GGT GGT TT TT AGCT GGG GGG	A A A A A	GCAAAGGAAAACGGCT CACCAGGAAAACGGTT CTAAAGGTAAAGGGT GTCCGGGTAAGCGCAT CCAAAGGTAAGCGCAT CAATTGGTGAACGTAT	CGCATTGC AGCCCTGC TGCTCTTC CGCCCTTC GGCATTAC TGTTACTG	CTCATTCECEGAT CCCATECCEGAT CTCATAETCEAAT CCCATECECEGAT CCCATECECEGAT AAGETTETCACEGAT AAGETTETCATAC	TATGATTCAC TATGATCCAC TATGATTCAT TATGATTCAC TATGATTCAC TATGATTCAC	CAACCOCTG CAACCCATG CAACCACTG CAACCACTG CAACCTTCT CCAACCTGAA CCAACCTGAA	- GGGGGGAACG - GGCGGGTACT - GGAGGGGACT - GGCGGAACC - GGTGGAACC - GGTGGTTTA - AGTTCTTAT	GGTCGCCGT GGCCGCCGT TCACAACGT GGGCGCCGG CGCGCGGA CGTGGA AATGGA	CAG CAG CAG CAA CAA	3	SCTTCCGATA SCAACGGACA SCTAGCGATA SCGACGGATA SCAACCGATA SCAACCGATA SCAGGAGAAA	TTGAAATCGAAGC TTGACATTGAAGC TTGAAATAGAAGC TCGACATTGAAGC TCGAAATCGAAGC TTTGGATTGATGAAGC	CAAAGAGATTCTGCGA CAGGGAAATTCTCCGC TABAGGAAATACTGAGA TAAGGAAATCCTGCGT CAGAGAAATTCTTACGG TCAAGGAAATTATGAAA
Se P3/1-582 Ss P3/1-582 Pm P3/1-582 Ma P3/1-582 As P3/1-579 CrP1/1-579 Phy P1/1-585 Sm P1-1/1-582	THE TRACESCOULD BE CONTRACTOR SCOULD BE CONTRACTOR	A A A A A A A	CCAAAGGAAAACGGCT CTCAAAGGTAAAACGGTT GTCAGGTAAAGCGTCT CCAAAGGTAAAGCGTCT CCAAAGGTAAGCGTAT CAATTGGTGAACGTAT AAATTACTAAACGTAT GGATCACCAAACGTAT	CGCATTGC AGCCCTGC TGCTCTTC CGCCCTTC GGCATTAC TGTTACTG AGCGCTAC	CTCATTCECEGAT CCCATECCEGAT CTCATAETCEAAT CCCATECECEGAT CCCATECECEGAT AAGETTETCACEGAT AAGETTETCATAECE CTCACECTAEGET	TATGATTCAC TATGATTCAT TATGATTCAT TATGATTCAC TATGATTCAC TATGATTCAC AATGATTCAT	CAACCOCTG	- GGGGGAACG - GGCGGTACT - GGAGGGACT - GGCGGAACC - GGTGGAACC - GGTGGTTA - AGTTCTTAT - AGCTCTTTC	GGTCGCCGT GGCCGCCGT TCACAACGT GGGCGCCGG - CGTGGA - CGTGGA TATGATGGA TATGATGGA TACGAAGGA	САС САС САС САС САС САС САС САС САС		SCTTCCGATA SCAACGGATA SCTAGCGATA SCGACGGATA SCAACCGATA SCACCGATA SCAGGAGAA SCGGGGGGAA	TTGAAATCGAAGC TTGAAATAGAAGC TCGAAATAGAAGC TCGAAATCGAAGC TTGGAATCGAAGC TTGGATTGAAGG GTATTAGGAGG GCCTGATGGAAGG	CAAAGAGAATTCTCCCC CAGGGAAATTCTCCCC TAAGGAAATACTGAGA CAGGAAATCCTGCG CAGAGAAATCCTACGG TCAAGAAATTCTACGG CAAGAAGTATTAAAA GAAGAGATATTGAAA
Se P3/1-582 Ss P3/1-582 Ma P3/1-582 As P3/1-579 CrP1/1-579 Phy P1/1-582 SmP1-1/1-582 As P1-1/1-582 As P11/1-582	TTGCTCTGCGCCGGC TTGTTAGCCTCCGGG TTATTGCGGCTCCGGT TTACTGCCTCCGGT TTGCTAGCTGCTGGT GCTTAGCTGCTGGGGGT ATTTAGCTGGGGGG ATTTCCCTGGAGGG ATTCCCCTGGAGGG	A 	ССАААББААААСБСТ Сассабсаааасвстт. Стааабстааасбст Втссбобтаабсста Ссааабстаабсбат сааттастаасстта Аааттастааасстат Абатсассааастат. Абатсассаасстат.	CGCATTGC AGCCCTGC TGCTCTTC CGCCCTTC GGCATTAC GGCATTACTG AGCGCTAC AGCACTAC AGCACTAC	CTCATTCGCGGAT CCCATGCCCGGAT CTCATAGTCGAA CCCATGCGCGGAT CCCATGCGCGGAT AAGGTTGTCATGC CTCACGCTAGGGT CCCACGCTAGGGT CCCACGCTAGGGT	TATGATTCAC TATGATTCAT TATGATTCAT TATGATTCAC TATGATTCAC TATGATTCAC AATGATTCAT TATGATTCAT TATGATTCAT	CAACCECTE CAACCATE CAACCATE CAACCATE CAACCTECT CAACCTEAA CAACCTECT CAACCTECT CAACCTECT	- GGGGGAACG - GGCGGTACT - GGAGGGAACT - GGTGGAACC - GGTGGAACC - GGTGGTTTA - AGTTCTTAT - AGTTCTTAT - AGTTCTTAT	GGTCGCCGT GGCCGCCGT TCACAACGT GGCCGCCGG 	CAG CAG CAG CAG CAA CAA CAA CAA		SCTTCCGATA SCAACGGACA SCTABCGATA SCGACCGATA SCACCGATA SCAGGAGAA SCAGGAGAA SCAGGAGAA	TTGAAATCGAAGC TTGACATTGAAGC TCGACATTGAAGC TCGACATTGAAGC TCGACATTGAAGC TTTGGATTGAAGC GCTTGATGAGGG GTATTATGGAGG GTGTTATGGAAGC GTGTCATGGAAGC	CAAAGAGATTCTGCGA CAGGGAAATTCTCCGC TAGGAAATTCTCGCG CAGAGAAATTCTGCGC CAGAGAAATTCTCACGG TCAAGAAATTCTACGG CAAGAAGATTTTGAAA AGAAGAGATATTGAAA AGAAGAGATATTGAAA
Se P3/1-582 Ss P3/1-582 Pm P3/1-582 Ma P3/1-582 As P3/1-579 CrP1/1-579 Phy P1/1-582 Sm P1-1/1-582 Ac P1/1-582 Prt P1/1-582	TOTTACCOGGC TTATTGGCAGCTCGG TTATTGGCAGCTCGGG TTACTGGCAGCTGCG TTACTGGCAGCTGCGGG TTACTGGCAGCTGCGGG TTTACCTGGCGGG TTTACCTGGCGGGA TATTCACTGGCGGGA TATCGGCAGGGGGA	A 	ВСАЛАВСВАЛАССВСТ САССАБСВАЛААССВСТ СТАЛАВСТААЛАССВСТ ССЛОВСТАЛВСССТ ССАЛАССТАЛВСССТ САЛАТСВСТАЛВСССТ АЛАТСВСТАЛВСССТ ПОЛТСАСТАЛАССТАЛ САЛТСАСТАЛАССТАЛ ВАТТАСТАЛАССТАЛ АДАТТАССЛАЛАССТАЛ АЛАТСССТАЛАССТАЛ	CGCATTGC AGCCCTGC TGCTCTTC CGCCCTTC GGCATTAC GGCATTAC GGCATTAC AGCGCTAC AGCACTGC AGCACTGC	CTCATTCGC66AT CCCATGCCC6AT CCCATGCCC66AT CCCATGCGC66AT CCCATGCGC66AT AGGCTGTCAL6A CTCACGCTA66GT CCCACGCCA66GT CCCACGCTA66GT CCCACGCTA66AT CTCACGCTA66AT	TATGATTCAC TATGATTCAT TATGATTCAC TATGATTCAC TATGATTCAC AATGATTCAC AATGATTCAT TATGATTCAT TATGATTCAT TATGATTCAT	CAACCECTE CAACCATE CAACCATE CAACCATE CAACCTECT CAACCTEAA CAACCTECT CAACCTECT CAACCTECC CAACCTECT CAACCTECT	- GGGGGAACG - GGCGGTACT - GGAGGGAACC - GGTGGAACC - GGTGGAACC - GGTGGTTTA - AGTTCTTAT - AGTTCTTAT - AGTTCTTAT - AGTTCATAT - AGTTCATAT	GGTCGCCGT GGCCGCCGT TCACAACGT GGCCGCCGG - CGTGGA - AATGGA TATGATGGA TATGATGGA TATGATGGA TATGACGGA TATGACGGA	CAG CAG CAG CAG CAA CAA CAA CAA CAA CAG CAG		CTTCCGATA CCACCGACA CTACCGATA CCACCGATA CCACCGATA CCACCGATA CCGCGGGGAA CCGCGGGGAA CCGCGAGAA CCGCGAGAA CCGCAGAA	TTGAAATCGAAG TTGACATTGAAG TTGACATTGAAG TCGACATTGAAG TCGACATTGAAG TTTGAAATCGAAG GTATTATGGAAG GCCTGATGGAGG GTATTATGGAAG GTGTCATGGAAG GTGTCATGGAAG	САААБАБАТТСТСССА САБСБАААТТСТСССС ТАВАБАААТАСТСАСС ТАВАБАААТССТВОВТ САБСАААТССТВОВТ САВСААТТСТАССВ ТСААБАААТТСТАССВ ТСААБАААТТАТСААА СБААБААБТАТТСААА АБААБАБАТАТТСААА БСАБСААССВТСКАА БСААБСАССВТСКАА
Se P3/1-582 Ss P3/1-582 Pm P3/1-582 Ma P3/1-582 As P3/1-582 CrP1/1-582 Sm P1-1/1-582 As P1/1-582 As P1/1-582 Art P1/1-582 At P1/1-582	TOTTATEGCACCEGC TTATEGCACCEGC TTATEGCACCEGC TTACEGCACCEGT TTACEGCACCEGC TTACEGCACCEGC TTACEGCACCEGC TTACEGCEGCG TACECCACEGCAGCA TTTCACEGCEGCAGCA TTCCACEGCEGCAGCA TTCCACEGCEGCAGCA	A 	ВСАААБСААААССВСТ САССАБСААААССВСТ СТАААССТАААССВСТ ССАССБСТААССВСТ ССААТСВСТААВСССТ ССААТСВСТААВСССТ СААТСВСТААСССТАТ АСАТТАСТАААССТАТ АСАТТАСТАААССТАТ АСАТТАССАААССТАТ СААТСАССАААССТАТ СААТСАССАААССТАТ СААТСАССАААССТАТ СААТСАССАААССТАТ	CECATIEC ACCCTEC TECTCTIC CECCTTC ECCATIC ECCATIC ACCETAC ACCECTAC ACCACTAC ACCACTAC ACCACTAC ACCACTAC	CTCATTCBCSGAT CCCATGCCCBAA CTCATABTCBAA CTCATGCCSGAT CCCATGCCSGAT CCCATGCCSGAT CCCACGCTASGGT CCCACGCTASGGT CCCACGCTASGGT CCCACGCTASGGT CTCACGCTASGAT	TATGATTCAC TATGATCAT TATGATCAT TATGATTCAC TATGATTCAC TATGATTCAC AATGATTCAT TATGATTCAT TATGATTCAT TATGATTCAT TATGATCAC AATGATCAC	CAACCGCTG CAACCATG CAACCATG CAACCCATG CAACCCATG CAACCTGCT CAACCTGCT CAACCTGCT CAACCTGCT CAACCTGCT CAACCTGCT CAACCTGCT CAACCGCT	- GGGGGAACG - GGCGGGACT - GGCGGAACC - GGTGGAACC - GGTGGAACC - GGTGGAACC - GGTGGATTA - AGTCCTTAT - AGTCCTTAT - AGTTCCTAT - AGTTCCTAT	GTCGCCGT GGCCGCCGT TCACAACGT GGCCCCCGG ACTGGA AATGGA TATGATGGA TATGATGGA TATGATGGA TATGACGGA TATGACGGA TATGACGGA	CAG CAG CAG CAA CAA CAA CAA CAA CAA CAA		CTTCCGATA CCAACGGACGACA CTABCGATA CGACCGATA CCACCGATA CCACCGATA CCTTCCACT CCGCGAGAA CCGCGGGAAA CTGCAGACA CTGCAGACA	ТТБАААТСБААБС ТТБАААТБААВС ТТБАААТБААВС ТСБАААТБААВС ТСБАААТБААВС ТСБАААТБААВС ТСБАААТСААВС ВТАТТАТБСАВС БССТВАТБСАВС БТАТТАТБСАВС БТЕТАТБСАВС БТЕТАТБСАВС ТСАТААСААТС ТСАТАССАЯТС ТСАТАСААТС	САААБАБАТТСТСССА САБСБАААТТСТСССС ТАБАБАААТССТСССС ТААБААААТССТСССС САБСАААТССТСССС САБСАААТССТСССС СААСАААТТАТСААА ССААСААТТАТСААА АБААБАБАТАТТСААА ВСААБААБТАТТСААА ВСААСАССТАССААС ВСААСААСТАСТААС
Se P3/1-582 Ss P3/1-582 PmP3/1-582 As P3/1-582 CrP1/1-579 CrP1/1-579 SmP1-1/1-582 As P1/1-582 PnP1/1-582 PnP1/1-582 PnP1/1-582 PrP1/1-582 PrP1/1-582	THE TO TREE CORE TTATTOR CONTINUES TTATTOR CONTINUES TTATTOR CONTINUES TTATTAC TO TO TO TTATTAC TO TO TO TTATTAC TO TO TO TTTAC TO TO TO TO TTTO TO TO TO TO TO TTTO T	A A A A - A A - A A - A A A A	ВСАААВВААААСВОСТ САССАВБААААСВОТ (ПАААВСТААААВОТ ВСОВБСТААВСОСТ ССАААВСТААВСОСТ САААТВОТБААСВОСТ СААТТВОТБААСВТАТ АВАТТАСТАААСВТАТ АВАТТАСТАААСВТАТ АААТВОСТАААСВТАТ АААТВАСТАААСВТАТ АААТТАССАААСВТАТ АААСТАССАВСТААВСТА	CECATICC AGCCCTEC CECCCTIC ESCATICC ESCATICC ASCCTICC ASCCCTAC ASCACTAC ASCACTEC ASCATICC ASCATICC	CTCATTCBCSGAT CCCATGCCCBAT CTCATGCCCGGAT CCCATGCCGGAT CCCATGCCGGAT CCCATCACBGAT CCCACGCTABGGT CCCACGCTASGGT CCCACGCTASGGT CCCACGCTASGGT CTCACGCTASGGT CTCACGCTASGGT CTCACGCTAGGGT CCCACGCTASGGT	TATGATTCAC TATGATCAC TATGATCAC TATGATTCAC TATGATTCAC TATGATTCAC TATGATTCAT TATGATTCAT TATGATTCAT TATGATTCAT TATGATCCA AATGATCCAT AATGATCCAT	CAACCGCTG CAACCATG CAACCATG CAACCATG CAACCATG CAACCTGA CAACCTGA CAACCTGA CAACCTGC CAACCTGCT CAACCTGCT CAACCGCT CAACCGCT CAACCGCT	- GGGGGAACG - GGCGGAACG - GGCGGAACC - GGCGGAACC - GGTGGAACC - GGTGGTTA - AGTCCTAT - AGTCCTAT - AGTCCTAT - AGTCCTAT - AGTCCTAT - GCTCCTAT	66 106 006 00 66 000 000 00 66 000 000 000 000 100 0000 100 000 100 000 100 000 100 000 100 000 100 0	CAG CAA CAA CAA CAA CAA CAA CAA CAA CAA		SCTTCCGAT SCAACGGACA SCAACGGAT SCGACGAT SCAACCGAT SCAGGAGAA SCTGCAGGAA SCGGGGGAA SCTGGAGAA SCTGGAGAA SCTGGAGAA ACGGGAGAA ACGCCAGAA	TTGAAATCGAAG TTGAAATGAAG TTGAAATGAAG TTGAAATGAAG TCGAAATGAAG TCGAAATGAAG GTATTATGAAG GTATTATGAAG GTATTATGAAG GTGTTATGAAG TTGATAAGAAAT TTATCTTGGAAG TTGTCCTGGAAG	CAAGGAGAATT CT CCCC CAGGGAAATT CT CCCC TAGGGAAATT CT CCC TAGGGAAATC TC CCC CAGGAAATT CT CCC CAGGAAGT TT CAAG CGAAGAAGT ATT CAAA AGAAGGAT ATT CAAA AGAAGAGAT TT CAAA AGAAGAGAT TT CAAA GAAGAACT CC CAAA GAAGAACT ACT TAAA GGAAGAATT CC CGAAA
Se P3'1-582 SP P3'1-582 Ma P3'1-582 Ma P3'1-582 As P3'1-573 PhyP1/1-585 Sm P1-1/1-582 PhP1'1-582 PhP1'1-582 PhP1'1-582 PhP1'1-582 PhP1'1-582 OsP1-1/1-582 OsP1-1/1-582	TICT CT C C C C C C C C C C C C C C C C	A 		CECATIC AGCCCTEC CECCCTIC CECCTIC CECCTIC CECATIC CECCTIC AGCCTIC AGCCTIC AGCACTEC AGCATICC AGCATICC AGCATICC	CTCATTCBC8GAT CCCATGCCC8GAT CCCATGC8C8GAT CCCATTCAC8BAT AABGTTGC8BAT CTCAC8CTASGAT CTCAC8CTASGAT CTCAC8CTASGAT CTCAC8CTASGAT CTCAC8CTASGAT CTCAC8CTASGAT CTCAC8CTASGAT CTCAC8CTASGAT CCCAC8CGASGAT	TATGATICAC TATGATCCAC TATGATCAC TATGATTCAC TATGATTCAC TATGATTCAC TATGATTCAC TATGATTCAT TATGATTCAT GATGATTCAT GATGATCCAT AATGATCCAT AATGATCCAT TATGCTTCAC	CAACCECTE CAACCATE CAACCATE CAACCATE CAACCATE CAACCTCT CAACCTEC CAACCTEC CAACCTEC CAACCTEC CAACCTEC CAACCTEC CAACCTEC CAACCTEC CAACCTEC CAACCTET CAACCTEC CAACCTECT	- GGGGGAACG GGCGGTACT GGCGGAACCI GGCGGAACCI GGTGGGAACCI GGTGGGAACCI GGTGGTTA AGTCCTTAT - AGTCCTTAT - AGTTCTAT - AGTTCCTAT - GCTGCTTAT - AGTGCTTAT	GG TCGCCGT GGCCGCCGT TCACAACGT TCACAACGT CGCGCCCCGG CGTGGA ATGATGGA TATGATGGA TATGATGGA TATGAGGA TATGAGGA TATGAGGA TATGAGGA TATGAGGA TATGAGGA TATGAGGA TATGAGGA TATGAGGA	CAG CAA CAA CAA CAA CAA CAA CAA CAA CAA		SCTTCCGAT SCAACGGACA SCAACGGAT SCGACCGATA SCACCGATA SCAGGAGAA SCTGCAGGAGAA SCAGGAGAA SCAGGAGAA SCTGCAGAGAT ACGCGAGAA ACACCAGAA	TTGAAATCGAAG TTGAATGAAG TGAAATGAAG TGGAAATGAAG TTGGAATGAAG GTATTATGAAG GTATTATGAAG GTATTATGAAG GTATTATGAAG TTGTCCTGGAAG TTTACTAGAAG TTTACTAGAAG	
S = P3/1-582 S = P3/1-582 M = P3/1-582 A = P3/1-582 PhyP1/1-585 PhyP1/1-585 S = P1/1-585 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 O = P1-1/1-582 O = P1-1/1-582 O = P1-1/1-582 O = P1-1/1-582 O = P1-1/1-582 O = P1-1/1-582 O = P1-1/1-582	TIGTACCICCCCCC TTALTGGGAGCICGT TTACTGGCTCCCGG TTACTGGCTCCGGT TTACTGCTCCGGT TTACCGCTGCGGGG TTTACCGGGGGG TTTTACCGGGGGGGA TTTACCGGCGGAGGA TTCCGCTGGAGGA TCCCGGTGGAGGA TTCCCGTGGAGGA TTCCCGGGGGAGGA	A A - A - A - A - A - A - A - A - A	CAAAAGGAAAACGGCT CACCAGGAAAACGGCT CACAAGGAAAGGCT CTAAAGTAAGCGT CCAAAGTACTAAGCGTA CAAATACTAACGAACGTA CAATACTAACGAACGTA AAATACTAACGAACGTA AAATACTAACGAACGTA AAATACCAAAACGTA CAATACCAAAACGTA CAATACCAACAACGTA AAACCACCAACAACGTA AAACCACCAACAACGTA AAACCACCAACGAGGTA AAACCACCAACGAGGTA	CECATTEC AGCCCTEC CECCCTIC GECCTTC GECATTAC TGTTACTE ASCETTAC AGCACTEC AGCACTEC AGCATTCC AGCATTCC AGCATTCC AGCATTCC	CTCATCGC66AT CCCATGCC66AT CCCATGCC6AT CCCATGC6C6AT CCCATGC6C6AT CCCATGC6C6AGAT CCCATGC6C6CATGC CCCACGCCAGGT CCCACGCCAGGT CCCACGCCAGGT CCCACGCCAGGT CCCACGCAGGAT CCCACGCAGGA CCCACGCAGGA CCCACGCAGGAGAT CCCACGCAGGAGAT	TATGATICAC TATGATCAT TATGATCAC TATGATTCAC TATGATTCAC TATGATTCAC AATGATTCAC AATGATTCAT GATGATTCAT GATGATCAC TATGCTCAC TATGCTTCAC TATGCTTCAC	CAACCECTE CAACCATE CAACCATE CAACCATE CAACCATE CAACCTATE CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT	- GGGGGAACG GGCGGAACCI GGCGGAACCI GGCGGAACCI GGTGGGAACCI GGTGGGAACCI GGTGGCTTTA AGTTCTTAT AGTTCATAT AGTTCATAT AGTTCCTTAT AGTGCTTAT AGTGCTTAT AGTGCTTAT	GG TCGCCGT GGCCGCCGGT TCACAACGT TCGCAACGT CGGCGCCCCGG CGGCGCCCGG TATGATGGA TATGATGGA TATGATGGA TATGAGCGA TATGAGCGA TATGAGCGA TATCGGCGA TATCGGCGA	CAG CAA CAA CAA CAA CAA CAA CAA CAA CAA		SCITCCGAT SCAACGAC SCAACGAT SCACCGAT SCACCGAT SCITCAGAT SCITCAGAT SCITCAGAT SCITCAGAT SCITCAGAT SCITCAGAT SCITCAGAT SCITCAGAT SCITCAGAT ACGCGAGAA ACGCCAGAA ACGCCAGAA	ITGAAATCGAAGG ITGAAATCGAAG ITGAAATGAAG ITGAAATGAAG ITGAAATGAAG ITGAAATGAAG GTATTAGATA GTATTAGAAG GTATTAGAAG GTATTAGAAG ITTATGAAG ITTATGAAG ITTTACTGAAAG ITTTACTGAAAG ITTTACTGAAAG ITTTACTGAAAG ITTTACTGAAAG ITTTACTGAAAG ITTTACTGAAAG ITTTACTGAAAG ITTTACTGAAAG	CAAGGAAATTC TG GG CAGGAAATTC TG CG TAAGGAAATTC TC CG TAAGGAAATC TG ACG TAAGGAAATC TG ACG TG AGGAAATTC TAACG GAGGAAATTA GAAA AGAAGGAATTT AAAA AGAAGGAATTT AAAA GGAGGAACTT CAAAA GGAGGAATTA CT TAAA GGAGGAATTA CT TAAA GGAGGAATTA CT TAAA GGAGGAATTA CAAAAA GGAGGATTA CAAAAA GGAGGATTA CAAAAA
Se P3/1-582 Ss P3/1-582 MaP3/1-582 MaP3/1-582 MaP3/1-579 CrP1/1-579 PhyP1/1-582 MaP1/1-582 AP1/1-582 AP1/1-582 AP1/1-582 OsP1-2/1-582 OsP1-2/1-582 OsP1-3/1-582 OsP1-3/1-582 OsP1-3/1-582	TOCOTORADA TOTALOCIONA TACTECTICE TACTECTICE TACTECTICE TOTALCECTORE TOTALCECTORE TOTALCECTORE TOTALCECTORE TOTALCECTORE TOTALCECTORE TOTECTO TOTECTORE TOTECTORE TOTECTORE TOTECTORE TOTECTORE	A A A A A A A A A A A A A A A A A A A	ВСАААВВААААСВОТ САССАВБААААСВОТ (ПАААВВТАААСВОТ СПАСАВСТААВСОСАТ ССААТВСТВАКССКАТ ССААТВСТВАКССКАТ АВАТТАСТАВСТВАКСТАТ АВАТТАСТАВАСТАТ АВАТТАСТАВАСТАТ АВАТТАСТАВАСТАТ АААТТАСТАВАСТАТ АААТТАСТАВАСТАТ АААТТАСТАВАСТАТ АААСТАССАСААВССТАТ ААСССАССАВССТАВССТ	CECATTEC ABCCCTEC TECTCTTC GECCTTC GECATTACTE ABCGCTAC ABCACTAC ABCACTAC ABCACTCC ABCATTCC ABCATTCC ABCATTCC ABCATTCC ABCATTCC	CTCATTOSCBGAT CCCATTOSCBGAT CCCATTOSCBAT CCCATTOSCCCATTOSCCCATTOSCCCCATTOSCCCCATTOSCCC	TATGATICAC TATGATCAT TATGATCAT TATGATCAT TATGATTCAC TATGATTCAC TATGATTCAC TATGATTCAC TATGATTCAT TATGATCAT TATGATCAT TATGCTCAC TATGCTTCAC	CAACCECTE CAACCATE CAACCATE CAACCATE CAACCATE CAACCTECTEC CAACCTECTEC CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCTECT	GGGGGAACG GGCGGACT GGCGGACC GGCGGACC GGCGGAACCI GGTGGAACCI GGTGGAACCI GGTGGACC AGTGCTTT AGTCCTAT AGTCCTAT AGTCCTAT AGTGCTTAT AGTGCTTAT	GGTCGCCGT GGCCGCCGT TCACAAGT GGCGCCCGG CGGCGCCGG GGGCGCCGG TATGATGGA TATGATGGA TATGATGGA TATGATGGA TATGAGGA TATGAGGA TATGAGCA TATGGGCA TATCGGCA	CAG CAA CAA CAA CAA CAA CAA CAA CAA CAA		SCITICCGAT SCAACGGAT SCAACGAT SCAACCGAT SCAACCGAT SCAGGAGAA SCAGGAGAA SCAGGAGAA SCAGGAGAA SCTGGAGAT SCTGGAGAT ACGGGAGAA ACGCCAGAA ACACCAGAA ACACCAGAA	IT GAAAT CGAAG IT GACAT CGAAG IT GACAT GGAAG IT GGAAT GGAAG IT GGAAT GGAAG GC TGAT GGAAG GC TGAT GGAAG GC TGAT GGAAG GC TGAT GGAAG GT GT GGAAG TG GT GAAG TG GT GAAG TG GT GAAG TG GT GAAG TT TT GC GAAG TT TT GC TG GAAG	CAABBABATT CT BOGA CABBAAATT CT COCC CABBAAATT CT COCC TABBAAATC CT COCC CABBAAATC CT BOCT TABBAAATC CT BOCT TABBAATC CT BOAT BOATCAATT CT BOAT ABABABAATT BOAT GAABAACT CT BOAT GAABAACT CT BOAT GAABAACT CACCAAA GAABABTT CACCAAA GAABABTT CACCAAA
S = P3/1-582 S = P3/1-582 PmP3/1-582 Map3/1-582 S = P1/1-582 PhyP1/1-585 S = P1-1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 OcP1-1/1-582 OcP1-1/1-582 CoP1-1/1-582	TGC C 10060000 TAT 1600AC 1001 TACT6C TC 000 TAT 1600AC 1001 TGT 1AC 100700 TGT 1AC 10070 TGT 1AC 100700 TGT 1AC 100700 TGT 1AC 100700 TGT 1007000 TGT 1007000 TGT 1007000 TGT 1007000 TGT 1007000 TGT 1007000 TGT 1007000 TGT 100700000 TGT 10070000000000000000000000000000000000	A A G G A A A A A A A A A A A A A A A A		CECATTEC ABCCCTCC CECCTCC CECCTCC CECCTCC CECCTCC CECCTCC ABCATTAC ABCBCTAC ABCACTCC ABCATTCC	CT CATT CGC GGAT CCATT CGC GGAT CCATT CGC TAT CCATT CGC TAT CCATT CGC TAG CCATT CAC GGAT CGC AGC TAG CGC T	TATGATICAC TATGATCAT TATGATCAT TATGATTCA TATGATTCAC TATGATTCAC TATGATTCAC TATGATTCAT TATGATTCAT GATGATTCAT TATGATCCAC AATGATCCAT TATGCTCAC TATGCTTCAC TATGCTTCAC TATGCTTCAC TATGCTTCAC	CAACCECTE CAACCATE CAACCATE CAACCATE CAACCATE CAACCATE CAACCTECT CAACCTECT CAACCTECT CAACCTECT CAACCECTECT CAACCECT CAACCECT CAACCECT CAACCECT CAACCECTECT CAACCECTECT CAACCECTECT CAACCECTECT CAACCECTECT	BGGGGAACG GGCGGACC GGCGGAACC GGTGGAACC GGTGGTTA AGCTCTTAT AGCTCTTAT AGCTCTTAT AGCTCCTAT GCTGCTTAT AGTCGTTAT AGTGCCTAT AGTGCCTAT	GGTCGCCGT GGCCGCCGT TCACAAGGT TCACAAGGT CACAAGGT CGTGA TATGATGGA TATGATGGA TATGATGGA TATGACGA TATGACGA TATGAGGA TATCGGCA TATCGGCA TATCGGCA TATCGGCA TATCGGCA	CAG CAA CAA CAA CAA CAA CAA CAA CAA CAA		CTTCCGAT GCACCGACC GCACCGAT GCACCGAT GCACCGAT GCACCGAT GCACCGAT GCACCGACA GCACGACAA GCACGACAA GCACCAGAA GCACCAGAA ACACCAGAA ACACCAGAA ACACCAGAA ACACCAGAA	TT GAAAT CGAAG TT GAAAT CGAAG TT GAAAT GAAG TT GAAAT GAAG TG GAAT GAAG TT GGAAT GAAG TT TG GAAG GC TG AT GAAG GC TG AT GAAG GC TG AT GAAG GC TG AT GAAG GT TG GAAG GT TG GAAG TT AT GAAG	
S = P3/1-582 S = P3/1-582 PaP3/1-582 A = P3/1-582 A = P3/1-572 C+P1/1-572 PhpP1/1-582 A = P1/1-582 A = P1/1-582 A = P1/1-582 A = P1/1-582 C = P1-1/1-582 C = P1-1/1-582 C = P1-1/1-582 Z = P1-1/1-582 Z = P1-1/1-582 C =	TICTC/100000000000000000000000000000000000	A G G G G G G G G G G G G G G G G G G G		CECATTEC ABCCCTCC CECCTCC CECCTCC CECCTCC CECCTCC CECATTAC GECATTAC ABCACTAC ABCACTAC ABCACTAC ABCATTC		TATGATICAA TATGATICAA TATGATICAA TATGATICAA TATGATICAA TATGATICAA TATGATICAA TATGATICAA TATGATICAA TATGATICAA TATGATICAA TATGCTICAA TATGCTICAA CATGATGCAA TATGCTICAA CATGATGCAA	CAACCB CTG CAACCATG CAACCATG CAACCATG CAACCATG CAACCATG CAACCTGAA CAACCTGAA CAACCTGAA CAACCTGCT CAACCGCT CAACCGCT CAACCGCT CAACCGCT CAACCGCT CAACCGCT CAACCGCT CAACCGCT CAACCGCT CAACCGCTGCT CAACCGCTGCT CAACCGCTGCT CAACCGCTGCT CAACCGCTGCT CAACCGCGCT	GGGGGAACG GCGGTAC GGCGGAACC GGGGGAACC GGTGGAACC GGTGGAACC GGTGGAACC GGTGGTTA AGTTCTTAT AGTTCTTAT AGTTCCTAT AGTGCTTAT AGTGCTTAT AGTGCTTAT AGTGCTTAT AGTGCTTAT AGTGCTTAT GGGCCGCG GGTGCAAC	GGTCGCCGT GGCCGCCGT TGACAACGT GGGCGCCGC AATGGA TATGATGGA TATGATGGA TATGATGGA TATGACGA TATGACGA TATGACGA TATGACGA TATGAGCA TATCGGCA TATCGGCA TATCGGCA TATCGGCA TATCGGCA TATCGGCA CGCGCC CGCGGC	CAG CAA CAA CAA CAA CAA CAA CAA CAA CAA		SCHTCCGAL SCACGGACA SCACGGACA SCACGGACA SCACGGACA SCACGGACA SCACGGACA SCACGAGAA SCAGGAGAA SCAGGAGAA SCAGGAGAA SCACGAGAA SCACCAGAA SCACCAGAA SCACCAGAA SCACCAGAA SCACCAGAA SCACCAGAA SCACCAGAA SCACCAGAA SCACCAGAA SCACCAGAA SCACCAGAA SCACCAGAA	TT GAAAT CGAAG TT GACAT GAAG TT GACAT GAAG CGACAT GAAG CGACAT GAAG CT GAAT GAAG CT TT GGAAT GAAG CT TT GAAG CT TT GAAG CT TT GAAG CT TT GAAG TT TAC GAAG CT TT GAAG TT TAC GAAG	
SP P3/1-582 SP P3/1-582 PaP3/1-582 AP P3/1-582 AP P3/1-573 CP P1/1-573 CP P1/1-573 CP P1/1-573 AP P1/1-582 AP P1/1-582 AP P1/1-582 AP P1/1-582 COF1-21/1-582 COF1-21/582 COF1-	TIC:CC:000000000000000000000000000000000	A A A A A A A A A A A A A A A A A A A		CECATTEC AGOCCTEC TECTOTE GECATTAC TETTACTE GECATTAC TETTACTE AGOCTTAC AGOCTTAC AGOCTTAC AGOCTTAC AGOCTTAC AGOCTTCC AGOCTTCC AGOCTTCC AGOCTTCC AGOCTTCC AGOCTTCC AGOCTTCC AGOCTTCC AGOCTTCC AGOCTTCC AGOCTTCC AGOCTTCC AGOCTTCC	CTOATT CGC BGAT CCAT CGC BGAT CCAT AT CGC BGAT CCAT CGC BGAT CCAT CGC BGAT CCAT CGC BGAT CGC BGC BGAT CGC BGC BGC BGAT CGC BGC BGC BGAT CGC BGC BGC BGAT CGC BGC BGC BGC BGAT CGC BGC BGC BGC BGC BGC BGC CGC BGC BGC BGC BGC CGC BGC BGC BGC CGC BGC BGC BGC CGC BGC BGC CGC BGC BGC CGC BGC BGC CGC BGC BGC CGC BGC BGC CGC BGC CGC CGC BGC CGC BGC CGC CGC BGC CGC BGC CGC BGC CGC BGC CGC CGC BGC CGC BGC	TATGATTCAC TATGATCAL TATGA	CAACCECTE CAACCATE CAACCATE CAACCATE CAACCATE CAACCATE CAACCTECT	BOGGGAACG BCCGGTACT BCACGGACT BCACGGACT BCCGGACC BCTGGAACC BCTGGTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT ACTCCTTAT	GGT C GC C GT GG C C GC C GT T C A C A A C G C C GC C GT C G C C G G C GT G G A C G A C G A T A C G A A G G A T A C G A A G G A T A C G A A G G A T A C G G A A G G A T A C G G C A T A C C G G C A T C C G C C A C C C C G C C A C C C C G C C A C C C C C C C C C C C C C C C C C	CAB CAA CAA CAA CAA CAA CAA CAA CAB CAB			TT GAAAG CGAAG C TT GAAAG CGAAG C TT GAAAG CAAT GAAG C TT GAAAG CAAT GAAG C TG GAAC TG GAAG C TG GAAC TG GAAG C TT GGAAG CGAAG C TT GAAG TG GAAG C TT GAC TG GAAG C TT TAC TG GAAG C TG TG TG TG GAAG C TG TG TG GAAG GAG C TG TG GAAG GAAG C TG TG GAAG GAG C TG TG GAAG GAAG C TT TG GAAG G TT TG GAAG GAAG C TT TG GAAG GAAG C TT TG GAAG GAAG C TT TG GAAG GAAG C TT TG GAAG G TT TG GAAG G TT TG GAAG C TT TG GAAG C TT TG GAAG C TG GAAG C TT TG GAAG C TT TG GAAG C TG GAAG C TT TG GAAG C TG TG T	
S = P3/1-582 S = P3/1-582 Pare P3/1-582 Are P3/1-582 Are P3/1-572 S = P3/1-579 Phy P1/1-582 Phy P1/1-582 Phy P1/1-582 Phy P1/1-582 Phy P1/1-582 Phy P1/1-582 C = P1/2-575 S = P1/2-1/575 S = P1/	IGC (C) CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC			CECATTEC AGCCCTEC TEGTTACTE GECATTAC TGTTACTE AGCGTTAC AGCCTTAC AGCCTTAC AGCATTAC AGCATTAC AGCATTAC AGCATTAC AGCATTAC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC CCCCATCC GCCATTCC CTCCATCC	CTCATTOGC6GAT COATTOGC6GAT COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC6GC COATTOGC7GC COATTOGC COATTOGC COATTOGC	TATGATTCAC TATGATCAT TATGATCAC TATGATTCAC TATGATTCAC TATGATTCAC TATGATTCAC TATGATTCAC TATGATCAC TATGATCAC TATGATCAC TATGATCAC TATGATCAC TATGATCAC TATGATCAC TATGCTCAC TATGCTCAC TATGCTCAC TATGCTCAC TATGCTCAC TATGCTCAC TATGCTCAC	CAACCECTG CAACCATG CAACCATG CAACCATG CAACCATG CAACCATG CAACCTCCC CAACCTCCC CAACCTCCC CAACCTCCC CAACCTCCCCCC CAACCTCCCCCCC CAACCTCCCCCCCC	00000000000000000000000000000000000000	60 T C C C C T 60 C C C C T 60 C C C C T 10 C C C C C T 10 C C C C C C C 10 T C A A A C T 10 T C A A C C C C C 10 T C A A C C C C C C 10 T C A C C C C C C C C C C C C C C C C C	CAB CAA CAA CAA CAA CAA CAA CAA CAA CAA			TT GAAAT CGAAGG TT GAAAT CGAAGG TT GAAAT GAAG TT GAAAT GAAG GAATT GAAAT GGAATT GAAG GC GGAAT GAAG GC GGAAT GAAG GC GGAAT GAAG GC GGAAG GC GGAAG GC GGAAG GC GGAAG TT GAAG GC GGAAG TT GAAG TT	
S = P3/1-582 S = P3/1-582 PaP3/1-582 Ma P3/1-582 A = P3/1-572 C+P1/1-572 PhyP1/1-582 PaP1/1-582 PaP1/1-582 QaP1-1/1-582 QaP1-1/1-582 QaP1-1/1-582 QaP1-1/1-582 QaP1-1/1-582 ZaP1-1/1-582 ZaP1-1/1-582 ZaP1-1/1-582 SaP2-2/1-573 SaR2-2/1-573 SaR2-2/1-573 SaR2-2/1-573	TGCTGCCGCCGCCGC TATTGGGAGCTGGT TACTGGCTCCGGT TGTTACCGCTGCGGT TGTTACCGCGCGGGGG TGTTACCGCGGGGGG TTTACCGCGGGGGG TACTCACTGGGGGG TACTCACTGGGGGGA TCCGCTGGAGGAGGA TCCGCTGGAGGAGGA TCCGCTGGAGGGGGG TCCGCTGGAGGCGGG TGCTGCTGGAGGCGGG TGCTGCGCGCGGGGGG TGCTGCGCGCGGGGGGGG	- A - A - A - A - A - A - A - A - A - A		CECATTEC AGECCTEC TEGETTEC GECCTTEC GECCTTEC GECATTAC GECATTAC GECATTAC AGECTAC AGECTAC AGECTAC AGECATTAC AGECATTAC AGECATTCC					00 1 00 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	САВ САВ САА САА САА САА САА САА САВ САВ				
S = P3/1-582 S = P3/1-582 PaP3/1-582 A&P3/1-582 A	THE CLE BEECEBEC THAT TROCK DUCODE THAT TROCK DUCODE THAT TROCK DUCODE THAT TROCK DUCODE THAT TRACTED TO COM THAT THAC TECHTED THAT THAC TECHTED THAT THAC TECHTED THAT THAC TECHTED THAT THAT THAT THAT THAT THAT THAT THAT			CECATTEC ABCCCTEC TOTACIE COCCTEC TOTACIE ABCSCTAC ABCSCTAC ABCSCTAC ABCACTAC ABCACTAC ABCATTCC ABCATTCC ABCATTCC ABCATTCC ABCATTCC ABCATTCC ABCATTCC CACATCC CCACATCC CCACATCC TCCATCC CCACATCC TCCATCC		TATGATIGAC TATGATCAL TATGA	CAACCECTG CAACCATG CAACCATG CAACCATG CAACCATG CAACCATG CAACCTGC CCACCGCG CCCCCCCGC CAACCTGC CAACCTGC CAACCTGC CACCTGC CA	00000000000000000000000000000000000000	001 C 00 C 01 00 C 00 C 01 00 C 00 C 01 00 C 00 C 01 00 C 00 C 00 00 C	САВ САВ САВ САВ САВ САВ САВ САВ САВ САВ			TT GAAAT CGAAG TT GAAAT CGAAG TT GAAAT GAAG CGAAAT GAAG CGAAAT GAAG CGAAAT GAAG TT GGAAT GAAG CT TA GAAG CT TA GAAG CT TA GAAG CT TA GAAG CT TA GAAG TT TA CA GAAG TT TA CT GAAG TT GAAG TO CT GAAG TT GAAG TO CT GAAG TT GAAG TO CT GAAG TT GAAG TO CT GAAG TT	CAABBABATT CT GOGA CABBAAATT CT COCO CABBAAATT CT COCO TABABAAATC CC COC TCABBAAATT CT COCO TCABBAAATT TAACA ABABAATT AT TAACA ABABABATT AT TAACA ABABABATT AT TBAAC ABABABATT AC TAACA ABABABATT AC TAACA GGABABATT AC CAACA GGABABATT AC CAACAA GGABABATT AC CAACAA GGABABATT AC CAACAA GGABABATT AC CAACAA GGABABATT AC CAACAA GGABABATT AC CAACAA GGABABATT AC CAACAACAACAA GGABABATT AC CAACAACAACAACAACAACAACAACAACAACAACAA
S + P3/1-582 S + P3/1-582 ParP3/1-582 Ar P3/1-57 Ar P3/1-57 Phy P1/1-582 Phy P1/1-582 Phy P1/1-582 Phy P1/1-582 Phy P1/1-582 Phy P1/1-582 Phy P1/1-582 CrP1-1/	1000000000000000000000000000000000000			CECATTEC ABCCCTEC TECTCTEC CECCCTEC TECTACTEC ABCSTACTEC ABCSTACTEC ABCSTACTEC ABCSTACTEC ABCSTACTEC ABCSTTCC ABCSTTCC ABCSTTCC ABCSTTCC ABCSTTCC ABCSTTCC ABCSTTCC CECATEC CECATEC TECATEC	CTCATTCGCGGAT CCATTGCCGGAT CCATGCCGATGCCGAT CCATGCGCGGCA CCATGCGCGGCA CCATGCCAGCTAGGC CCAGCTAGGCAGGC CCAGCTAGGCAGGC CCAGCTAGGAT CCAGCTAGGAT CCAGCTAGGA CCAGCTAGGA CCAGCTAGGAT CCAGCTAGGAT CCAGCTAGGAT CCAGCTAGGAT CCAGCGAGGAT CCAGCGAGGAGAT CCAGCGAGGAGAT CCAGCGAGGAGAT CCAGCGAGGAGAT CCAGCGAGGAGAT CCATGACCAGGAGAT CCTTGCCGGAAT CCCAGCGCGGAGAT		CAACCECTG CAACCATG CAACCATG CAACCATG CAACCATG CAACCATG CAACCTCC CAACCTCC CAACCTCC CAACCTCC CAACCTCC CAACCTCC CAACCTCC CAACCTCC CAACCTCC CAACCTCC CAACCTCC CAACCTCC CAACCTCC CAACCTCCC CAACCTCCC CAACCTCCCCCC		GGT CGC CGT GGC CGC CGT TCACAAC CGT CGT CGACAC CGT CGT CGACAC CGT CGT CGACAC CGT CGT CGACAC CGT CGT CGAC CGT	CAB CAA CAA CAA CAA CAA CAA CAA CAB CAA CAB CAA CAB CAB			TT GAAAT CGAAG TT GAAT CGAAG TT GAAT GAAG TT GAAT GAAG TT GGAAT GAAG CGAAT GAAG CGAAT GAAG CGGAAT GAAG CGGAAT GAAG CGGAAT GAAG CGGAAT GAAG CGGAAG TT GAAG TT G	CAAABABATT CI BOBA CAGBAAATT CI BOBA CAGBAAATT CI BOBA TAABAAATA CI BOBA TAABAAATA CI BOBA TAABAAATTA BAAA ABAABAATTA BAAA ABAABAATTA BAAA DI ABAAATTA BAAA DI ABAAATTA BAAA DI ABAAATTA CI BAAA DI ABAAATTA CI BAAA DI ABAAATTA CI BAAA DI ABAAATTA CI TAAA ABAABAATTA CI BAAA DI ABAAATTA CI TAAA DI ABAAATTA CI TAAA DI ABAAATTA CI TAAA DI ABAATTA CI TAAA
S = P3/1-582 S = P3/1-582 PaP3/1-582 As P3/1-573 As P3/1-573 PhyP1/1-582 As P1/1-573 PhyP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/1-582 CrP1/2-575 SaR2-2/1-575 PhR2/1-575 Ph2/	TICT CCC 100000000000000000000000000000000			CGCATTGC AGCCCTGC TGCTCTGC CGCCCTGC CGCCCTC CGCCCTC CGCCCTC AGCGTACAT AGCGTAC AGCGTAC AGCATCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC CGCCATGC CGCCCC CGCCCCCCCCCC					GB (1) C (2) C (3) GB (2) C (3) C (3) GB (3) C (3) C (3) <td>CA3 CA3 CA3 CA3 CA3 CA3 CA4 CA4 CA5 CA4 CA6 CA4 CA6 CA4 CA6 CA6 CA7 CA7 CA7</td> <td></td> <td></td> <td></td> <td></td>	CA3 CA3 CA3 CA3 CA3 CA3 CA4 CA4 CA5 CA4 CA6 CA4 CA6 CA4 CA6 CA6 CA7 CA7 CA7				
S = P3/1-582 S = P3/1-582 PaP3/1-582 As P3/1-575 As P3/1-575 CrP1/1-575 PhyP1/1-585 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 CrP1-	THE CLE DECESSE THAT TACCE CLE DECESSE THAT TAC TERE THAT TAGE TAC TO CONT THAT TAC THAT TAC TO CONT THAT TAC THAT TAC TO CONT THAT T	A A A A A A A A A A A A A A A A A A A	ВСАААВСААААССВОТ САССАБСААААССВОТ ПОСОБСАТААВСОТ ОТ САССБОТААВССТАТ СОСАААВСТАВСССАТ СОСАААВСТАВСССАТ СААТВОТБАВСССАТ ОДАТТАСТАВССАТАС ДАТТАСТАВССАТАССТАТ АДАТТАСТАВАССТАТ АДАТТАСТАВАССТАТ АДАТТАСТАВАССТАТ АДАТТАСТАВАССТАТ АДАТТАСТАВАССТАТ АДАТТАСТАВАССТАТ СААТВОСТАВССТАВССТАТ АДССАССААССТАВССТАТ АДССАСССАВССТАВССТАТ АДССАСССАВССТАВССТАТ АДССАСССАВССТАВССТАТ АДССАСССАВССТАВССТАТ АДССАСССАВССТАВССТАТ АДАССАСССАВССТАВССТАТ АДАССАСССАВССТАВССТАТ АДАССАСССАВССТАВССТАТ ССААВБСАТСССССАТ АДАВССАТССССАВСТАА АДААВССТАСССАТ АДАВССАТСАСССАТ АДАВССАТСАСССАТ АДАВССАТСАСССАТ АДАВССАТСАСССАТ АДАВССАТСССАСТАВСТАТС ССААВБСАТССССАТ АДАВССАТСССАТ	CECATTOC AGCCTTC GGCOTTC GGCOTTC GGCATTACTE AGCETACTE AGCETAC AGCETAC AGCATTACTE AGCATTACTE AGCATTC AGCATTC AGCATTC AGCATTC AGCATTC AGCATTC CACATCC GGCAA CATTCC CCCACATCC GGCAA CATTCC CCCACATCC CC				00000000000000000000000000000000000000	001 C 00 C 01 00 C 00 C 01 00 C 00 C 01 00 C 00 C 01 00 C 00 C 00 00 C 01 00 C	CAS CAS CAS CAS CAS CAS CAS CAS CAS CAS			TT GAAAT CGAAG TT GAAAT CGAAG TT GAAAT GAAG CGAAAT CGAAG CGAAAT CGAAG TT GGAAT GAAG CT TT GGAAT CGAAG CT TT GGAAT CGAAG CT TT ACT AGAAG CT TT ACT AGAAG CT TT ACT AGAAG CT TT ACT AGAAG TT AGAAAAT AGAAG TT AGAAAAG TT AGAAACT AGAG TT AGAAAT CGGG TC AGAAT CGGG TC AGAAT CGGG	
S + P3/1-582 S + P3/1-582 PaP3/1-582 An P3/1-582 An P3/1-582 An P1/1-573 PhyP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 SmR2-2/1-575 SmR2-1/1-575 SmR2	TICTIC DECESSE TATIONAL TACTECTICSE TACTECTICSE TACTECTICSE TACTECTICSE TOTTACCECTECTECTE TOTTACCECTECTECTECTECTECTECTECTECTECTECTECTEC	A A A A A A A A A A A A A A A A A A A		CGATTGC AGCCTGC TGCTCTC CGCCTGC TGCTATAC AGCGTAC ACGTAC AGCGTAC AGCGTAC AGCGTAC AGCATAC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATTCC AGCATCC CCCACCC TCCCACCC CCCCCCCCCC	CTCATTCGC6GAT CCATGCC6GAT CCATGCC6AT CCATGCC6AT CCATGCC6AT CCATGCC6AT CCATGCC6AT CCATGCC6AT CCATGCC7ATCC7AT CCATGCC7AT CCATGCC7AT CCATGCC7AT CCATGCC7AT CC			00000000000000000000000000000000000000	GGT CGC CGT GGC CGC CGT TCACAACGT CGT CGACACGT CGT CGACACGT CGT CGACACGT CGT CGACACGT CGT CGAC CGT CGAC CGT CGAC ACAAGGA ACGACGAC CGT CGAC CGT	CAB CAB CAA CAA CAA CAA CAA CAA CAA CAA			TT GAAAT CGAAGG TT GACAT CGAAG TT GACAT GAAG TT GACAT GAAG GC GAAT GAAG GC GAAT GAAG GC GAAT GAAG GC GAAT GAAG GC GAAT GAAG GC GAA GAAG GC GAAG TT GAAG GC GAAG GC GAAG TT GAAG GC GAAG TT GAAG TC GAAG TC GAAG TT GAAG TT GAAG TT GAAG TC GAAG TT GAAG TT GAAG TT GAAG TT GAAG TT GAAG TT GAAG TT GAAG TG GAG TG GAAG TG GAAG TG GAAG TG GAG TG GAAG TG GAG TG GAG T	CAAABABATTCC BOGA CABGAAATCC BOGA CABGAAATCC BOGA TAAGAAAATC BABA TAAGAAATTAC BOGA TCAABAAATTACAA ABAABABTAAATTACAA ABAABABTAAATTACAA ABAABABTAAATTACAA ABAABABTAAATTACAA ABAABABTAA CABAAATTACAAAA CABAABTAA
S = P3/1-582 S = P3/1-582 ParP3/1-582 A = P3/1-575 A = P3/1-575 C = P1/1-757 Phy P1/1-585 A = P1/1-7582 PhP1/1-582 PhP1/1-582 A = P1/1-582 C = P1-1/1-582 C = P1-1/1-582 C = P1-1/1-582 C = P1-1/1-582 C = P1-1/1-582 C = P1-2/1-582 C = P1-2	TGC TGC CGC CGC TAT TGC CC CGG TAT TGC CC CGG TAT TGC CC CGG TAT TGC CC CGG TTAT TGC CGC CGG TGT TAC TGC CGC TGT TAC TGC CGG TTT TCC CGC CGG TTT CGC CGG TGC CAC TGC CGC TC CGC CGC CGC CGC CGC CGC CGC TC CGC CGC CGC CGC CGC CGC CGC CGC CGC C	A A A A A A A A A A A A A A A A A A A							GGT CGCCGT GGCCGCCGT TCACACCGT TCACACCGT TCACACCGT GGCCGCCGG CGTCGAC CGTCGAC CGTCGAC ACGACGAC ACGACGAC ACGACGAC ACGACGAC CGTCG	САВ САВ САВ САВ САА САА САА САА САВ САВ				
S = P3/1-582 S = P3/1-582 PaP3/1-582 As P3/1-573 As P3/1-573 PhyP1/1-583 PhyP1/1-583 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 CaP1-1/1-582	TICTIC 1000000000000000000000000000000000000	A A A A A A A A A A A A A A A A A A A							GET CBCCBT GBCCBCCBT TACAACBT GBCCBCCBC CBTCBA ACTACAACBAACBA ACTACAACBAACBAACBAACBAACBAACBAACBAACBAAC					
S + P3/1-582 S + P3/1-582 PaP3/1-582 As P3/1-575 As P3/1-575 PhyP1/1-585 S mP1-1/1-572 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 S mP2-2/1-575 S mP2-2/1-575	Tig Trace (1000000000000000000000000000000000000	A A A A A A A A A A A A A A A A A A A		CECATTOC ACCATTOC				00000000000000000000000000000000000000	GGT CGC CGT GGC CGC CGT CGA CAC CGT CGT CGA CAC CGT CGT CGA CAC CGT CGT CGA CAC CGT CGT CGA CGT A CGA AGGA A CGA AGGA A CGA AGGA A CGA CGA A CGA				TT GAAAT CGAAGG TT GACAT GAAG TT GACAT GAAG TT GACAT GAAG TT GACAT GAAG TT GAAAT GAAG TT GAAAT GAAG GC GAAT GAAG GC GAAT GAAG GC GAAT GAAG GC GAA GAAG GC GAA GAAG GC GAA GAAG GC GAAG TT GACC GAAG TT GACC GAAG TT GACC GAAG TT GACC GAAG TT TAC GAAG TT TAC GAAG TT TAC GAAG TT GACC GAAG TC GAACAAG TC GAACAAG TC GAACAAG TC GAACAAG TG GAAG TG GAACAT GAAG TC GAACAAG TG GAAG TG GAG TG GAAG TG GAG TG GAG TG GAG TG GAAG TG GG TG GAG TG G	CAAABABATTCC 0000 CAABBABATTCC 0000 TABBAAAATC 0000 TABBAAATAC 0000 TABBAAATAC 0000 TABBAAATA 0000 CAABAAATA 0000 CAABAAATA 0000 CAABAAATA 0000 CAABAACTAC 0000 CAABAATAC 0000 CAABAATACAC 0000 CAABAATACAC 0000 CAABAATACACAAAA CAABAACTACACAAA
S + P3/1-522 S + P3/1-522 ParP3/1-522 Ar P3/1-527 Ar P3/1-527 Ar P3/1-577 PhyP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 CrP1-1/1-582 S = 27-1/1-575 S = 27-1/1-575 S = 27-1/575 S		A A A A A A A A A A A A A A G <td< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td>GGT CGC CGT GGC CGC CGT TCACAACGT GGC CGC CGT CGT CGACACGT AGT CGACACGT AGT CGACACGT AGT CGACGT AGT CGACGT AGT</td><td>CAB CAB CAA CAA CAA CAA CAA CAA CAA CAA</td><td></td><td></td><td>TT GAAAT CGAAG TT GAAAT CGAAG TT GAAAT GAAG TT GAAAT GAAG TT GAAAT GAAG GC GAAT GAAG GC GAAG TT GAAAC GAAG TT GC GAAG TC GAAAT GC GO TC GC GAAT GC GC TC GC GC GC GC GC TC GC GC GC GC GC TT GC GAAT GC GC TT GC GG GC GC GC GC TC GC GC GC GC GC TC GC GC GC GC GC TC GC GC GC GC GC TT GC GC GC GC GC GC GC TT GC GC</td><td>CAAABABATT CI BOBA CAABBABATT CI BOBA CABBAAATT CI BOBA TAABAAAATT CI BOBA TAABAAAATT ALBAA ABAABAAATT ALBAA ABAABABAATT ALBAA ABAABABAATT ALBAA ABAABABATT AAAA ABAABABAATT ALBAA DIAABAATT ACTTAAA ABAABABATT ACAABAT DIAABAATT ACTTAAA ABAABABATT ACAABAT DIAABAATT ACTTAAA ABAABABTT ACABAT DIAABAATT ACTTAAA ABAABABTT ACABAT DIAABAATT ACTTAAA ABAABABTT ACABAT DIAABAATT ACTTAAA AAABAABTT ACABAT DIAABAATTACTTAAA AAABAABTT ACABAT DIAABAATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACABAA DIAABAABTTACABAA DIAABAABTTACABAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA</td></td<>							GGT CGC CGT GGC CGC CGT TCACAACGT GGC CGC CGT CGT CGACACGT AGT CGACACGT AGT CGACACGT AGT CGACGT AGT	CAB CAB CAA CAA CAA CAA CAA CAA CAA CAA			TT GAAAT CGAAG TT GAAAT CGAAG TT GAAAT GAAG TT GAAAT GAAG TT GAAAT GAAG GC GAAT GAAG GC GAAG TT GAAAC GAAG TT GC GAAG TC GAAAT GC GO TC GC GAAT GC GC TC GC GC GC GC GC TC GC GC GC GC GC TT GC GAAT GC GC TT GC GG GC GC GC GC TC GC GC GC GC GC TC GC GC GC GC GC TC GC GC GC GC GC TT GC GC GC GC GC GC GC TT GC	CAAABABATT CI BOBA CAABBABATT CI BOBA CABBAAATT CI BOBA TAABAAAATT CI BOBA TAABAAAATT ALBAA ABAABAAATT ALBAA ABAABABAATT ALBAA ABAABABAATT ALBAA ABAABABATT AAAA ABAABABAATT ALBAA DIAABAATT ACTTAAA ABAABABATT ACAABAT DIAABAATT ACTTAAA ABAABABATT ACAABAT DIAABAATT ACTTAAA ABAABABTT ACABAT DIAABAATT ACTTAAA ABAABABTT ACABAT DIAABAATT ACTTAAA ABAABABTT ACABAT DIAABAATT ACTTAAA AAABAABTT ACABAT DIAABAATTACTTAAA AAABAABTT ACABAT DIAABAATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACABAA DIAABAABTTACABAA DIAABAABTTACABAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA AAABAABATTACTTAAA
S = P3/1-582 S = P3/1-582 PaP3/1-582 As P3/1-573 As P3/1-573 PhyP1/1-583 PhyP1/1-583 PhyP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 CaP1-1/1-582 CaP1-1/1-582 CaP1-1/1-582 CaP1-1/1-582 CaP1-1/1-582 CaP1-1/1-582 CaP1-1/1-582 CaP1-1/1-582 CaP2/1-579 PhyP2-2/1-579 PhyP2-2/1-579 PhyP2-2/1-579 PhyP2-2/1-579 PhyP2-2/1-579 PhyP2-2/1-579 PhyP2-2/1-579 PhyP2-1/1-582 CaP1-1/1-582 CAP1-1/1-58	TICTIC 1000000000000000000000000000000000000	A A A A A A A A A A A A A A A A A A A							GB (GC) GC)	CAS CAS CAS				
S + P3/1-582 S + P3/1-582 PaP3/1-582 A + P3/1-572 A + P3/1-573 C+P1/1-573 PhyP1/1-573 PhyP1/1-573 PhyP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-2/1-582 C+P1-2/1-582 C+P2-1/2-573 PhyR2-2/1-573 SimP2-1/1-573 SimP1-574 Sim	THE CLOCE OF CONCEPTION TO CLOCE OF CLO	A A A A A A A A A A A A A A A A A A A						Construction of the second sec	GGT CGC CGT GGC CGC CGT CGA CAC CGT CGA CAC CGT CGT CGA CGT C				TT GAAAT C GAAG TT GACAT C GAAG TT GACAT C GAAG TT GACAT GAAG TT TACT GAAG TC TACT GAAG TT TACT GAAG TT TACT GAAG TT TACT GAAG TT TACT GAAG TT TACT GAAG TT TACT GAAG TC TACT GAAG TT TACT TACT TACT TACT TACT TACT TACT T	
S + P3/1-582 S + P3/1-582 Pap3/1-582 An P3/1-582 An P3/1-582 An P1/1-573 PhyP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 OrF1-1/1-582 S = R/1-575 S = R/1-575		A A A A A A A A A A A A A A A A A A A		CECATEC ABCCCTEC GCCCTEC GCCCTEC GCCCTEC GCCCTEC GCCCTEC ABCBCACATACIE ABCBCACATACIE ABCCTACIE ABCACTACIE ABCATEC ABCA					GGT CGC CGT GGC CGC CGT TCACAACGT GGC CGC CGT CGT CGACACGT CGT CGACACGT CGT CGAC CGT CGAC CGT CGAC CGT CGAC CGT CGAC CGT CGAC CGT CGAC CGT CGAC CGT CGA CGT CGGA CGT CGA CGT CGA CGA CGT CGA CGT CGA CGA CGA CGA CGA CGA CGA CGA CGA CGA	CAB CAB CAA CAA CAA CAA CAA CAA CAA CAA				CAAABABATTC BOGA CABBAAATTC BOGA CABBAAATTC BOGA TAABAAATTA CBOGA TAABAAATTA CBOGA TCAABAAATTA BAAA ABAABAATTA BAAA ABAABAATTA BAAA CBAABAABTTA CAABA CAABAABTTA CABCAC CAABAATTO CTOR CAABAATTA CTOR CAABAACTA CACAAA CABBAACTA CACAAA CABBAACTA CACAAA CABBAACTA CACAAA CAABAACTTA CACAAA CAABAACTTA CACACAA
S = P3/1-582 S = P3/1-582 PaP3/1-582 As P3/1-573 As P3/1-573 PhyP1/1-583 PhyP1/1-583 PhyP1/1-583 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP2-1/1-582 S mP2/1-575 S mP2/1-575 S mP2/1-575 S mP2/1-575 S mP2/1-575 S mP2/1-575 S mP1/1-553 S	TICT CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	A A A A A A A A A A A A A A A A A A A							GGT CGC CGT GGC CGC CGT TCGC CGT GGC CGC CGT CGT CGA CCGT CGT CGA CGT AGA CGGC AGA CGGC AGA CGGC CGT CGA AGA CGGC CGT CGA CGT	CAB CAB CAB CAA CAA CAA CAA CAA CAA CAA				
S + P3/1-582 S + P3/1-582 PaP3/1-582 A + P3/1-573 A + P3/1-573 C + P1/1-573 PhyP1/1-583 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 C + P1/1-582 C + P1/1-582		A A A A A A A A A A A A A A A A A A A						B000000000000000000000000000000000000	GGT CGC CGT GGC CGC CGT CGA CAC CGT CGA CAC CGT CGA CAC CGT CGT CGA CGT CGA C	CAB CAB CAB			TT GAAAT C GAAG TT GACAT C GAAG TT GACAT C GAAG TT GACAT GAAG TT TACT GAAG TT GACAT GAAG TT GACAT TC GAAG TT GACAT TC GAAG TC GAAG TT GAAG	
S + P3/1-582 S + P3/1-582 PaP3/1-582 A + P3/1-575 A + P3/1-575 C+P1/1-575 PhyP1/1-585 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P1-1/1-582 C+P2-1/1-575 S = RP1-575 S = RP1-	Tig Trace (1000000000000000000000000000000000000	A A A A A A A A A A A A A A A A A A A		CECATEC AGCCTEC GGCCTTC GGCCTTC GGCCTTC GGCCTTC GGCCTTC GGCCTCC GGCCTCC GGCCTCC ACCTCC ACCTCC ACCTCC ACCTCC ACCTCC ACCTCC CCCCCCCC					GGT CGC CGT GGC CGC CGT TCACAAC CGT GGC CGC CGT CGT CGACAC CGT CGT CGACAC CGT CGT CGACAC CGT CGT CGACAC CGT CGACACAC CGT CGACACACACACACACACACACACACACACACACACACA	CAB CAB CAA CAA CAA CAA CAA CAA CAA CAA				
S = P3/1-582 S = P3/1-582 Pap3/1-582 As P3/1-573 As P3/1-573 PhyP1/1-583 PhyP1/1-583 PhyP1/1-583 PhyP1/1-582 PhyP1/1-582 Ph/1/-582 Ph/1/-582 Ph/1/-582 Ph/1/-582 Ph/1/-582 Ph/1/-582 Ph/1/-582 Ph/1/-582 Ph/1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-578 Ph/2-1/-578 Ph/2-1/-578 Ph/2-1/-578 Ph/2-1/-578 Ph/2-1/-578 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-588 Ph/2-1/-582 Ph/2-1/-582 Ph/2-1/-582 Ph/1-1/-5		A A A A A A A A A A A A A A A A A A A							GGT CGC CGT GGC CGC CGT TCACAACDT GGC CGC CGT CACAACDT ACT ACAACGA ACT ACAACGA AT AT ACAACGA AT AT ACAACGA AT AT ACAACGA AT AT ACGAACGA AT AT ACGAACGA CGT CGAA AT ACGAACGA CGT CGAACGA CGT CGAACGAACGA CGT CGAACGAACGAACGAACGAACGAACGAACGAACGAACGA	CAB CAB CAA CAA CAA CAA CAA CAA CAA CAA			TT GAAAT CGAAG TT GAAAT CGAAG TT GAAAT GAAG TT GAAAT GAAG CGAAT GAAG CGAAT GAAG CGAAT GAAG CGAAT GAAG CGT GAAT GAAG CGT GAAT GAAG CGT GAAG CGT GAAG CGT GAAG TT GAT GAAG CGT GAAG TT GAT TA GAG TT GAAT TA GAG TT GAAT TA GAG TT CAT T	
S + P3/1-582 S + P3/1-582 PaP3/1-582 A + P3/1-573 A + P3/1-573 C + P1/1-573 PhyP1/1-583 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 C + P1/1-582 C		A A A A A A A A A A A A A A A A A A A							GET CS C CS I BSC CS C CS I C CA CA CS I C C CA CA CS I C C C CS C					
S + P3/1-582 S + P3/1-582 PaP3/1-582 A + P3/1-573 C + P1/1-573 PhyP1/1-583 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 C + P1/1-582 C + P1/1-582 S = R2-17-575 S = R2		A A A A A A A A A A A A A A A A A A A		CECATEC ABCCCTCC GECCTTC GECCTTC GECCTTC GECCTTC GECCTTC GECCTTC GECCTTC GECCTCC ACCTCC ACCTCC ACCTCC ACCTCC ACCTCC ACCTCC ACCTCC ACCTCC CECCTCC GEGTATCC GEGTATCC GEGTATCC GEGTATCC GEGTATCC GEGTATCC GEGTATCC GEGTACCCC GEGTACCC GEGTACCC GEGTACCC GEGTACCC GEGTACCC GEGTACCC G					GGT CGC CGT GGC CGC CGT TCACAACGT GGC CGC CGT CGT CGA CGT CGA CGA CGA CGA CGA CGA CGA CGA CGA CGA	CAB CAB CAA CAA CAA CAA CAA CAA CAA CAA				
S = P3/1-582 S = P3/1-582 PaP3/1-582 As P3/1-573 As P3/1-573 PhyP1/1-583 PhyP1/1-583 PhyP1/1-583 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP2-1/1-582 CaP1-1/1-582		A A A A A A A A A A A A A A A A A A A							GGT CGC CGT GGC CGC CGT TCACAACDT GGC CGC CGT CGC CGG CGT CGA ACT CGA ACT CGA ACT CGA ACT CGA CGT CGA ACT CGA CGT CGA ACT CGG CGT CGA CGT CGA CGA CGT CGA CGT CGA CGA CGA CGA CGA CGA CGA CGA CGA CGA	CAB CAB CAA CAA CAA CAA CAA CAA CAA CAA			TT GAAAT CGAAGC TT GAAAT CGAAGC TT GAAAT CGAAGC TT GAAAT GAAGC TT GAAAT GAAGC TT GAAAT GAAGC TT GAAAT GAAGC TT GAAAT GAAGC TT GAAAT GAAGC TT TAC TAGAAGC TT TAGAAGC	
S + P3/1-582 S + P3/1-582 PaP3/1-582 A + P3/1-573 A + P3/1-573 S + P3/1-573 PhyP1/1-583 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 C + P1/1-582 C +		A A A A A A A A A A A A A A A A A A A		CECATEC ABCCCTCC CCCCCCC CCCCCCCC CCCCCCCCC CCCCCC				Bigging AACB Bigging AACB </td <td>GET CS C CS I BS C SC CS I C A A A CS I C A A A CS I C A A A CS I C A C A CS I C A C A CS I C A CS</td> <td></td> <td></td> <td></td> <td>TT GAAAT C GAAG TT GACAT C GAAG TT GACAT C GAAG TT GACAT GAAG TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT GACAT TT GACAT GACAT TC GAAG TT GACAT GACAT TT GACAT GACAT TT GACAT GACAT TT GACAT GACAT TT GACAT TT GACAT TC GACAT TT GACAT TC GACAT TC GACAT TC GACAT TC GACAT TC GACAT TC GACAT TT GACAT</td> <td></td>	GET CS C CS I BS C SC CS I C A A A CS I C A A A CS I C A A A CS I C A C A CS I C A C A CS I C A CS				TT GAAAT C GAAG TT GACAT C GAAG TT GACAT C GAAG TT GACAT GAAG TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT TT GACAT GACAT TT GACAT GACAT TC GAAG TT GACAT GACAT TT GACAT GACAT TT GACAT GACAT TT GACAT GACAT TT GACAT TT GACAT TC GACAT TT GACAT TC GACAT TC GACAT TC GACAT TC GACAT TC GACAT TC GACAT TT GACAT	
S + P3/1-582 S + P3/1-582 PaP3/1-582 A + P3/1-573 C + P1/1-573 PhyP1/1-583 PhyP1/1-583 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 PhyP1/1-582 C + P1/1-582 C + P1/1-582		A A A A A A A A A A A A A A A A A A A		CECATEC ABCCCTCC GCCCTCC GCCCTCC GCCCTCC GCCCTCC GCCCTCC ACCCCCCCC ACCCCCCCC ACCCCCCCCCC					GGT CGC CGT GGC CGC CGT CGC CGG CGT CGA CGT					
S + P3/1-582 S + P3/1-582 PaP3/1-582 A + P3/1-582 A + P3/1-573 C + P1/1-582 P + P1/1-582 P + P1/1-582 P + P1/1-582 P + P1/1-582 P + P1/1-582 P + P1/1-582 C +		A A A A A A A A A A A A A A A A A A A		CECATEC ABCCCTEC GCCCTC GCCCTC GCCCTC GCCCTC GCCCTC GCCCTC GCCCTC GCCCTC GCCCTC ACCCTC ACCCTC ACCCTC ACCCTC ACCCTCC GCCTCC CCCCCCCCC CCCCCCC CCCCCCCC					GGT CGC CGT GGC CGC CGT TCACAACGT AT CACAACGT AT CACAACGT AACGT AACGT AACGT AT CACAACGT AT CACAACGT AT CACAACGT AA	CAB CAB CAB CAA CAA CAA CAA CAA CAA CAA		CTTCCBAT CAACBACA CAACBAC	TT GAAAT CGAAG TT GACAT CGAAG TT GACAT GAAG TT GACAT GAAG TT GACAT GAAG GC GAAT GAAG TT GAAAT GAAG TT GAAAT GAAG TT GAAT GAA	
S + P3/1-582 S + P3/1-582 PaP3/1-582 A + P3/1-573 A + P3/1-573 C + P1/1-573 PhyP1/1-583 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 PhP1/1-582 C + P1/1-582 C +		A A A A A A A A A A A A A A A A A A A		CG ATT GC AG CC TT GC AG CC TT GC AG CC TT GC AG CC TT AC AG CC TT					GBT CBC CBT BBC CBC CBT BBC CBC CBT CAR CAC BT BBC CBC CBT BBC CBC CBT CBT CBA CAC BT CAT CBC CAC BT CBT CBA CAC BT CB					

C. D/4 604	610	620	630	640	650	660 670) 680	690	700	710	720	730
Ripro_P/1-564	ATCAAAAGGTTAC	TTAATGAATT/	ATATAGTAAAC	ATACTAAACAA	AATTAAAATATAT	CGAAAAAAGTAT	GGAACGTGATAATT	TTAT GTCAT	C <mark>GG</mark> AAGAAGCA	AAAAAATTT	GGATTAGTAG	ATAATATAATT
CrP2/1-564	ATGAAGAAAAAAC Atgaagagagacga	TAAATCAAAT T <mark>g</mark> acca <mark>gccti</mark>	TATGAAAAAC Stacatcaagc	ATACTGGAAATT Acacc <mark>ggctg</mark> ca	CACTGAAGAAAA CGCGCGAGA <mark>T</mark> GGO	CGGAGAAGAACACGCT	GGAAAGAGATAAAT GGACCGGGGAC <mark>T</mark> CGT	TCATGGATC TCATGAGCG	CTGAAGAAGCA C <mark>GC</mark> AGG <mark>AGG</mark> CC	AGGAAAATT AA <mark>gg</mark> ac <mark>t</mark> gg	GGCTTAATTG/ GGG <mark>CTG</mark> ATCG/	ACAGGGTGATA A <mark>cg</mark> aaatcctc
PhyP2/1-555 SmP2/1-564	A <mark>T</mark> GAGGGCGCGAC Atgagatcccggc	TCA <mark>G</mark> CC <mark>GT</mark> CT(TCAACAA G CT(CTA <mark>tgtg</mark> aagc Ctacgccaagc	ACACTGGTCACA ACACGGGGGAGTA	C <mark>gtatg</mark> aaaagat Gcatcgagaagat	ITGAATCAAGTAT ICGACCAGAGCAT	GGAGCGGGGA <mark>T</mark> CA <mark>TT</mark> GGAGAGGGGATCACT	TTAT <mark>GTCTC</mark> TCATGTCGC	CA <mark>G</mark> AGGAGGCC CCGAGGAGGCC	AA <mark>g</mark> caattt Aaggagttt	GGTCTCGTGG/ GGGCTCATCG/	A <mark>tgaagtgatc</mark> Atgaggtgatt
AtP2/1-564	GTGTGGGGATGCGT	TAAATGAATTO	TATETCAAAC	ACACGGGTCAAC	CTCTTGACGTGG	TCACAATAACAT	GGATAGGGATCATT	TCATGACTC	CTGAAGAGGCT	AA <mark>GGCGTTT</mark>	GGAATAATCG	ATGAAGTGATT
PtP2-2/1-564	TCTGGGATGCTC	TGAATCAATT	TACT GCAAGC	ATACAGGGCAGC	CACTTGATGTAA	TTCAGAAGAATAT	GGATAGAGATTATT	TTAT GACCC	CAGAAGAGGCA	AA <mark>GG</mark> AA <mark>TTT</mark>	GGAATTATCG	ATGATGTGATT
PtP2-3/1-564 OsP2-1/1-564	GTCTGGGATGCTC GTGCGCGACCGTC	TGAATCAATTO Tcaacaa <mark>g</mark> ato	BTACTGCAAGC Ctac <mark>g</mark> ccaagc	ATACAGGGCAGC Acac <mark>g</mark> agccagg	CACTTGATGTAA C <u>C</u> atc <mark>g</mark> acc <mark>gt</mark> at	I TCAGAAGAATAT I <mark>Cg</mark> agcagtgcat	GGATAGAGATTATT GGA <mark>gcggg</mark> aca <mark>tg</mark> t	TTATGACCC Tcatggatc	CAGAAGAGGCA Ca <mark>g</mark> aggaggcg	AA <mark>ggaattt</mark> Cac <mark>g</mark> actgg	GGAATTATCG/ GGG <mark>CT</mark> CATCG/	ATGATGTGATT A <mark>cg</mark> a <mark>ggt</mark> catc
OsP2-2/1-564 ZmP2m-1/1-564	CTGCGCGACCGCC GTGCGCGACCGGC	TCAACAA <mark>G</mark> AT(TCAATAAGAT(CTACCA <mark>g</mark> aago Ctatgccaago	ACACCGGCCAGG	A <mark>gategacaaga</mark> Ccategacegeat	ICGAGCAGTGCAT ICGAGCAGTGTAT	GGAGCGCGACCTCT GGAGAGGGGACATGT	TCATGGACC TCATGGACC	CC <mark>G</mark> AGGAGGCG CGGAGGAGGCA	CGCGATTGG .CATGAATGG	GGGCTCATCG GGGCTCATCG	AC <mark>G</mark> a <mark>ggt</mark> aatt Acgaggtcatc
ZmP2m-2/1-564	STOCOGOCACCOGC	TCAACAAAAT	CTATECCAAEC	ACACEGECCAE	CCATCGACCGCAT	TCCACCACTCTAT	GGAGAGGGGACATGT	TCATEGACC	C <mark>GG</mark> AGG <mark>AGGC</mark> A	CATGATTGG	GGGCTTATCG	ACCACCE ACCE
ZmP2p-2/1-564	TGCGCGACCGCC	TCAACAAGAT	TACCAGAAGC	ACACECECCAEC	CCATCGACAAGA	CGAGCAGTGCAT	GGAGAGGGACCTCT	TCATGGATC	CC <mark>GACGAGGCG</mark>	C G C G A C T G G	GGCCTCATCG	AC <mark>GAGGT</mark> CATC
Se P1/1-558 Se P2/1-564	ac <mark>aagagcacgc</mark>	tgaacgacttg	ctcgcccagc	acacaggccagc	ccctcgaaaaaa	ttgaagtcgacac	cgatcgcgacttct	tcatgtcgc	cggaagaagcc	aa <mark>ggott</mark> ac	ggattaattga	a <mark>tcaggt</mark> actt
SsP1/1-558 SsP2/1-564	CTTAAAGCACTAC	TTAACCAACA tcaataccat	TTTGGCCAACC	ACACTGGCAAAT	CCCTA <mark>G</mark> A <mark>GG</mark> AAA ccatggaaaaact	TACCGCCGATAC	GGAACGGGATTTCT ggaaagagatttct	TTAT <mark>GTCC</mark> G ttatgtctg	CTGAGGAATCG cggaggaagca	AA <mark>GG</mark> AA <mark>T</mark> AT aaggagtat	GGCTT <mark>G</mark> ATTG/ ggcctcattg	ACCA <mark>GGT</mark> AATT accaggtaatt
AsP1/1-558	CACAAGCGGCGGT	TAAATGATTA	CTTAGCCGAAC	ACACTGGTCAAC	CCATTGAGCGCAT	TECCEAAGATAC	TGAACGTGACTTTT	TCATGTCAC	CAGATGAAGCC	AAAGATTAC	GGCTTAATTG	ACCAA <mark>GT</mark> AATT
PmP1/1-558	TAAAAGAACGTT	TAAATACTGA	CTATCT GAAA	GAACAGGTAAAG	AACTT <mark>G</mark> AAACCA	TAAGGGAGATAC	TGATAGAGATTTTT	ATAT GTCTC	CACAGGAGGCA	GTAGAGTAT	GGTCTAATAG	ATTTGGTTTTA
PmP2/1-564 MaP1/1-558	CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	TAAATTCTCT TCAAT <mark>G</mark> AACA(CTA <mark>g</mark> caaaac CCTA <mark>g</mark> c <mark>t</mark> agtc	ATACTAATCAAT ATAC <mark>666</mark> TAAAC	CTTTAGAAAAAA Ctctc <mark>g</mark> acctcat	TCGCCGAGGATAC	GGAAAGGGGA <mark>tttc</mark> t	TTAT <mark>GTCCG</mark>	CTGAAGAAGCA CC <mark>G</mark> AA <mark>g</mark> aagcc	CTAGACTAC	GGACTAATIG GG <mark>TTT</mark> AATCG	ATAAAGTCATA Atca <mark>ggtg</mark> atc
Ma P2/1-564 CrP5/1-528	A <mark>TT</mark> AAACA <mark>g</mark> agac Cacaa <mark>g</mark> ctcacgc	TCAA <mark>T</mark> accati Tcaac <mark>gg</mark> ctai	SCTGGCTCACC CCTGGCCCAGT	ACACC <mark>GGT</mark> CAAC	CCTATGATCGCA CCATGGAGACCA	COCCAATGACAC Caccaaggacac	CGAACGGGGACTTTT CGACCGCGACTTCT	TCATGTCGG TCATGAGCC	CC <mark>GT</mark> CGAAGCC CGCAGGAGGCC	AAA <mark>g</mark> aatat Atcgagtac	GGGTTAATCG/ GGCCTGGTGG/	A <mark>tctggtg</mark> att Ac <mark>g</mark> ccatcatc
PhyP5-1/1-528		TGAACGGTTA	ICTA <mark>G</mark> CATACC	ACACTGGACAGC	CCTACGAGAAGA	TAGTCCAAGACAC	GGACCGAGATTTCT	TTATGAGTG	CCAAGGAAGCG	CAGGAGTAT	GGCCTCATCG	ACCCTCTTATT
SmP5/1-528	CACAAA <mark>g</mark> c <mark>g</mark> aac <mark>t</mark>	TGAACAGCTA	TCTA <mark>TCTT</mark> ATC	ACACC <mark>GGGG</mark> CA <mark>G</mark> C	CACTGGAGAGGA	TAA <mark>ttcaag</mark> acac	CGACCGGGGATTTCT	TCATGAGTC	CCAA <mark>gg</mark> aagc <mark>t</mark>	CAAGAGTAT	GGCCTCATTG	ACGGC <mark>GTCATT</mark>
PtP5-1/1-528	CACAA <mark>gg</mark> caaacc	TAAAC <mark>GGT</mark> TAC TAAAT <mark>GGG</mark> TA	ICTCBCATACC	ACACTGGTCAAA	G <mark>tcttggagaaga</mark> G <mark>tcttgagaaga</mark> 1	ICAACCAGGACAC	AGATCGTGATTACT	TCATGAGIG TCATGAGIG	CCAAAGAAGCA C <mark>g</mark> aaagaagco	AAAGAGTAT AA <mark>gg</mark> actac	GGG <mark>CTT</mark> ATTG	ACCECTETTATC A <mark>tcctctatc</mark>
PtP5-2/1-453 OsP5/1-528	CACAA <mark>gg</mark> caaacc Cacaa <mark>gg</mark> c <mark>t</mark> aacc	TAAAT <mark>GGG</mark> TA T <mark>G</mark> AATGGATA(TCTC <mark>G</mark> CCTACC CCTA <mark>G</mark> CATACC	ACACTGGTCAAA ACACTGGGCAGC	GTCTTGAGAAGA CCCTAGATAAGA	CAACCAA <mark>G</mark> ATAC	AGATCGTGATTACT	TCATGAGTG TCATGAGCG	CGAAAGAAGCC CGAAGGAGGCA	AA <mark>gg</mark> actat Aa <mark>gg</mark> a <mark>gt</mark> at	GGGCTCATTG GGTCTAATTG	AT <mark>ggtgttat</mark> c At <mark>ggagttat</mark> c
ZmP5/1-528	CACAA <mark>GGCG</mark> AAC <mark>T</mark>	TAAAT <mark>gg</mark> ata(COTTOCATACC	ACACTGGGCAGC	CCCTGGATAAGAT	TCAACGTAGATAC	TGACCGTGACTACT	TCATGAGCG	C <mark>G</mark> AAA <mark>G</mark> AAGCA	AA <mark>GG</mark> A <mark>GT</mark> AT	GGCCTTATTG	AT <mark>ggagtaatc</mark>
PhyP3-2/1-522	ACAA <mark>gt</mark> acaa <mark>g</mark> c	TTCA <mark>G</mark> AA <mark>G</mark> ATA	AATGTCGCGTC	TTACGGGCCAGC	C <mark>T</mark> GA <mark>GG</mark> C <mark>GAGGG</mark>	TC <mark>G</mark> AA <mark>GCTG</mark> ACAC	66A <mark>T</mark> A6A6ACAATT	TATGAATG	CG <mark>T</mark> GGGA <mark>GGC</mark> G	A <mark>tgg</mark> a <mark>gt</mark> at	GGGCTTGTGG	A <mark>tggcgtg</mark> att
AtP3/1-519	CACAAA <mark>B</mark> CAAABC	TTAACAAAAT	CTTCTCTAGAA	TAACTGGAAAG	CTGAAGAGCAGA CTGAA <mark>t</mark> cagagai	ICGAAAGTGACAC	AGACCAGAGACTACT	TCTTGAATC	CAT <mark>GGG</mark> AGGCA	AAA <mark>g</mark> aatat	GG <mark>TTT</mark> GATCG	AC <mark>GCT</mark> GTAATC
PtP3/1-522 ZmP3/1-522	CACAA <mark>G</mark> ATTAAGC Saaaagattaaga	T <mark>g</mark> aacaa <mark>g</mark> at <i>i</i> Tcaa <mark>t</mark> aaaati	ATTATCAA <mark>g</mark> aa Aat <mark>gtcg</mark> agaa	TAACAGGGAAGC TTACTGGAAAAT	CTCTAGAGCAGG Ctgaggagcagai	TTGAAGTTGACAC	GAAGTTTGACAATT	TCATGAATG TCATGAGTC	CTTCCCAACCT CTTCCCAACCC	AAA <mark>G</mark> AATAT AA <mark>GG</mark> ATTAT	GGG <mark>TT</mark> GGTAG/ GGGATAGTTG/	A <mark>tg</mark> aa <mark>gttatt</mark> Acaatatcata
OsP3/1-522	GAGAAGATAAAGA	TCAACAAAAT/	ACTOTCAA <mark>G</mark> AA	TCACCGGGGAAGC	CTGAAGAGCAGAT	TT <mark>GATGAGGAT</mark> AC	AAAGTTTGACTATT	TCATGAGTC	CTTGGGAAGCC CBCTGGABGCC	AA <mark>GG</mark> ATTAT	GGCATAGTTG	ACAGC <mark>GTT</mark> ATA
PhyP4-1/1-522	ACAA <mark>GAGCAA<mark>t</mark>g</mark>	TCAC <mark>GCGT</mark> AT	TTTATCTGAGA	TTACAGGGCGGG	CTCCTGAGCAGG	TAGAGAAGGACA <mark>T</mark>	TGATAGGGACCGGT	ACA <mark>TGTCT</mark> C	CTATCGAAGCA	A <mark>tgg</mark> aatat	GGCATTATTG	A <mark>TGGCGTCATT</mark>
SmP4/1-522 SmP4/1-510	CATAAGAGTAACG CACAA <mark>g</mark> aacaact	TTACCC <mark>GG</mark> AT	CCTC <mark>G</mark> CAGAGA	TAACTGGGAAAC	CATTT <mark>G</mark> AGCAGG	I C <mark>GAG</mark> AAGGACAT	AGACAGAGAGATCGTT	ACATGTCTC ACATGTC <mark>GG</mark>	CA <mark>gtgg</mark> aagc <mark>t</mark>	ATGGAGTAC Atggaatac	GGGTTGATCG	A <mark>tggtgttate</mark>
AtP4/1-522 PtP4-1/1-522	AACAA <mark>g</mark> aacaatg Aacaagaataata	TCACCA <mark>G</mark> CAT TAATTA <mark>G</mark> AATI	TATC <mark>GCGGGG</mark> AT T <mark>GTTTCAGGTT</mark>	GTACTAGTCGAT TCACTAGTCGCA	CGTTTGAGCAGG CGGTTGAGCAAGT	ITCTGAAAGATAT Igcaaaaagatat	TGATAGGGACCGGT GGATAGAGATCGTT	ACA <mark>TGTCT</mark> C Acat <mark>gtct</mark> C	CAA <mark>ttg</mark> aagca Caattgaagca	GTTGAGTAT GTTGAATAT	GGTTTAATTG GGAATAATTG	AT <mark>ggag</mark> ttatt Atggagttatt
PtP4-2/1-522	AACAA <mark>G</mark> AA <mark>T</mark> AA <mark>T</mark> A	TAATTA <mark>G</mark> AATT	T <mark>GTTTCAGG</mark> TT TATATCAGGTT	TCACTAGTCGCA	CEGETTGAGCAAGT	I <mark>G</mark> CAAAAA <mark>G</mark> A <mark>T</mark> AT	GGATAGAGATCGTT	ACATGTCTC ATATGGGTC	CAATTGAAGCA	GTTGAATAT	GGAATAATTG/	AT <mark>ggagttatt</mark> Acaststaatc
CrR6/1-522	TGTACCAGATCT	TCTCGCGCTA	TACATGAAG	TCACCGGCATGA	ACCAGGACCAGA	TT <mark>G</mark> AG <mark>CAGGCCAC</mark>	CTGCCGTGACCACT	TCATGACGC	CC <mark>G</mark> AGCAGGCC	AA <mark>GCTGG</mark> AG	GGCCTGATTG	ACGAGATCATC
PhyP6-1/1-519 PhyP6-2/1-519	CTCGTGATAAAA CTCGGGGATAAAA	TAGACCAGAT	TATTCAACAT	TCTGCAATCAAC	CTTTGGAAAGAG	IGCGCAAGTTTAC	AGAGCGAGATTACT	ТСТТСТСТ <mark>С</mark>	CTGCTGAGGCA	TT <mark>GG</mark> ATTTT	GGACTCATTG	ACGGTTT <mark>G</mark> TTA
SmP6/1-519 AtP6/1-519	CAC <mark>G</mark> A G ATAAAA SCCC G ACAAAAAA	TCGATAAGATO TTGACAGGATO	STAC <mark>GCGGCT</mark> T Statgcagctt	TCACTGGGCAGC TCACTGGACAAC	CACTGGAGA <mark>T</mark> agi Ctctggagaaagi	ICCAAACA <mark>T</mark> ACAC I <mark>g</mark> cagcaatacac	AGAGCGTGACCATT Tgaaagagatcgt	TTTTCTCGG TCTTATCAG	CACCCCACCC CATCTCACCCC	AT <mark>gg</mark> aattt Cttgagttc	GGATTGATCG/ GGGCTCATTG/	AC <mark>GGGTTG</mark> CTC A <mark>TGGTCTATT</mark> G
PtP6-1/1-519	CTCGTCATAAAA	TTGACCAAAT	STAT <mark>GCT</mark> ACAT	TCACTGGTCAAC	CCCTTGAGAAAGT	ACAGCAATACAC	TGAGAGAGAGATCGTT	TCTTATCCA	CTTCAGAGGCT	ATGGAGTTT ATGGAGTTT	GGGCTCATCG	AT <mark>GGCATCCTG</mark> Atggcatcctg
OsP6/1-519	CTCGTGATAAAG	TTGATAAGAT	ЗТТТ <mark>бстб</mark> стт	TTACTGGGCAAA	CCTTGGATATGG	ACAACAGTGGAC	AGAGAGGGATCGTT	TCATGTCTT	CATCTGAAGCC	ATGGACTTT	GGACTAGTTG	A <mark>TGCCCTGCTG</mark>
2mP6/1-234 Se P3/1-582	A <mark>T</mark> CAAAAAAC <mark>TG</mark> C	T <mark>G</mark> AACCAAA <mark>T</mark> (CA <mark>tgg</mark> cc <mark>g</mark> a <mark>t</mark> c	GCACCGGTCAGC	CTCTT <mark>G</mark> AAAAAA	CGAGAAGGACAC	CGATCGCGACTACT	TCATGTCTG	C <mark>TG</mark> AA <mark>G</mark> AAGCC	C G T G A G T A C	GGCTTGATCG	ACCAA <mark>gtg</mark> a <mark>tc</mark>
SsP3/1-582 PmP3/1-582	ATTAAA <mark>g</mark> atatgt	TGAACGAAAT Tgaa <mark>tcact</mark> c	TATGGCCCAAC Tatgt <mark>ctg</mark> ata	TGACAGGCCAAA TGACAGGACAAT	CTGTGGAAAAAAA CCTTT <mark>g</mark> aaaaaaa	TECCAAEGATAC TA <mark>g</mark> agaaaaga <mark>t</mark> ac	CGACCGGGACTATT TGATAGAGATTATT	TCCTATCAG TTCTAAGTG	CGGCAGAAGCA C <mark>T</mark> GAAG <mark>aagcT</mark>	AA <mark>gg</mark> agtac Aaaaactat	GGCTTGATCG/ GGG <mark>CT</mark> aatag/	ATAAAGTGATT Ata <mark>g</mark> agtaatt
MaP3/1-582 AsP3/1-579 A	ATTCCCCACCAAT Attccccaccac	TAAAC <mark>G</mark> AAAT TCAATCAAAT	TAT <mark>gg</mark> ctaatc Ttat <mark>g</mark> ccaaca	GCACCGGACAAA ACACCAGTCAAC	CCATCGAACGCA CCCTAGCCAAAA	ICGAAAAAGACAC ICGAAAAAGATAT	GGACCGCGATTATT GGATCGGGACTTTT	TCCTCTCCG TTAT <mark>G</mark> TCAG	CCGAGGAAGCT CTCAAGAGGCG	GTCGCCTAC Atggaatac	GGTTTAATTG GGTTT <mark>G</mark> ATTG	ATAAGGTGGTG Acc <mark>gtgtg</mark> att
CrP1/1-579	ATTC <mark>GTTTAG</mark> ATG	TAGCAGAAATT	TTATTCTTTAT	CAACGTACCGAC	CAA <mark>G</mark> ACATAAAA	TTTACGTGATTT	AGATCGTGATTTTT	ATTTAACTG	CAA <mark>TGG</mark> AAACA	ATTTATTAT AAAAACCTAT	GETTTAECTE	ATCAAATTGCA
SmP1-1/1-582	CTTCGTGACTGCA	TTACAA <mark>GT</mark> ATT	TTAT <mark>GCAC<mark>GG</mark>A</mark>	AAAC <mark>GGGT</mark> AAAC	CCCTAT GGGTAG	TTTCCGAGGATAT	GGAAAGAGAT <mark>GTTT</mark>	TCATGTCAG	C G A G A G A A G C C	AGGGTTTAT	GGCATCATCG	ATCTT <mark>G</mark> TA <mark>GT</mark> C
Ac P1/1-582	TACGTGATTGTA	TAACCAAAGT	TTAT <mark>G</mark> CCCAAA	GAACTGGCAAAC	CTTTGTGGTTAAT	TTTCCGAGGACAT	GGAAAGAGATGTAT	TTTTGTCAG	C <mark>GG</mark> AA <mark>G</mark> AAGCC	CA <mark>gg</mark> attac	GG <mark>TG</mark> TTG <mark>CGG</mark>	ATCTA <mark>g</mark> tagca
PntP1/1-582 (AtP1/1-582 (CTTCGCGATTACA C <mark>TT</mark> C <mark>G</mark> C <mark>G</mark> AAA <u>CC</u> A	TAACCAGATG Tcacaa <mark>ggg</mark> ti	TTATATA <mark>g</mark> aaa Ttat <mark>g</mark> tacaaa	GAACGGACCAAC Gaac <mark>ggg</mark> caaac	CC <mark>GGAGAGGT</mark> AAT C <mark>tat</mark> at <mark>gggtta</mark> t	TTCAACGGGACCT Ta <mark>tccg</mark> aagacat	GAACAGAGATGTTT <mark>GG</mark> AAC <mark>GGG</mark> A <mark>tg</mark> ttt	TTATGTCAG TTAT <mark>gtcag</mark>	CAACA g aagcc Caaca <mark>g</mark> aagcc	CAAGCTTAT CAA <mark>gctc</mark> at	GGCATTGTTG/ GGAATTGTTG/	ATGTCGTAGCG At <mark>cttg</mark> tagcg
PtP1/1-582 OsP1-1/1-582	CTGCGCGAAATCC GTTCGCGAAATGA	TCACAA <mark>GGGT</mark> TCACAA <mark>GGGT</mark>	TTAT <mark>g</mark> cacaaa Ttatgcactaa	GAACAGGCAAAC	CCTTAT <mark>GGG</mark> TTG1 CTTTTTGGGTTG1	TATCTGAAGACAT TATCCGAAGACAT	GGAAAGGGATGTTT GGAAAGGGATGTTT	TTAT <mark>GTCAG</mark> TTATGTCAG	CA <mark>G</mark> CA <mark>G</mark> AAGCC CAGACGAAGCC	CAA <mark>GTTCAT</mark> AAA <mark>GCTT</mark> AT	GGAATTGTTG GGACTTGTCG	A <mark>tctt</mark> gtagca Atattgtaggg
OsP1-2/1-582	TTCCCCCAAATCA	TCACAAGGGT	TTAT <mark>GCACT</mark> AA	GAACAGGCAAGC	CTTTTTT <mark>GGG</mark> TT <mark>G</mark>	TATCCGAAGACAT	GGAAAGGGATGTTT	TTATETCAE	CAGACGAAGCC	AAA <mark>GCTT</mark> AT	GGACTTGTCG	ATATT <mark>GTAGGG</mark>
OsP1-4/1-582	TTCATGAAATGA	TCACAAGGGT	TTAT <mark>GCACT</mark> AA	GAACAGGCAAGC	CTTTTT <mark>GGG</mark> TT <mark>G</mark> 1	TA <mark>TCCG</mark> AA <mark>G</mark> ACAT	GGAAAGGGAT <mark>GTTT</mark>	TTATGTCAG	CA <mark>G</mark> ACAAAGCC	AAA <mark>GCTT</mark> AT	GGACTTGTCG	A <mark>tatt</mark> gtaggg
ZmP1-1/1-582 ZmP1-2/1-582	STTCCCCGAAATGA Sttccccgaaatga	TCACAAGGGT TCACAA <mark>GGGT</mark>	TAT <mark>b</mark> cactaa Ttat <mark>b</mark> cactaa	GAACAGGCAAGC	CTTTTT <mark>GGG</mark> TTG CTTTTT <mark>GGG</mark> TTG	TATCCGAAGACAT	GGAAAGGGATGTTT GGAAAGGGATGTTT	TTATGTCAG	CAGACGAAGCC Cagacgaagcc	AAA <mark>gcttat</mark> Aaa <mark>gctt</mark> at	GGACTTGTCG/ GGACTTGTCG/	ATATTGTAGGG A <mark>tattgtaggg</mark>
CrR2/1-579 PhyR2-1/1-579	ATGCGCGACTACC Btgagaaactatg	TGTCGCTGCT(TATACGAACAA	CACCTCCAAC <mark>G</mark> A <mark>GTTTCG</mark> AAGA	CCACCGGGCAAC	CCTAC <mark>G</mark> acc <mark>ggg</mark> CT <mark>GTGG</mark> acaagg	T <mark>Gattcgcgag</mark> ct Taatgaaggactt	GTCACGCAACAAGT GAATCGTATCAAAC	GGATGGACC GTTTCAGTG	CCAA <mark>g</mark> ca <mark>gg</mark> cc Ctaaagaagct	ATCGAGTAC TTGGAGTAC	GGCATGATCG GGTATTATTG	ACAA <mark>ggtgctg</mark> Acaaggttgtg
PhyR2-2/1-579 SmR2-1/1-579	TCACCACTATC	T <mark>gt</mark> ac <mark>g</mark> aacaa Tcttcaagcaa	A <mark>gtttcaaag</mark> a Acta <mark>bctbtc</mark> a	AAACTGGTCAAC	CTTTT <mark>g</mark> ataaggi Cabaagaagaagai	TAGTGAAGGATTT	GAATCGTATTAAGC	GCTTTAGCG	CTAATGAAGCT CGCAAGAGGCT	TT <mark>GG</mark> AATAT	GGTATAATTG	ACAA <mark>ggttgt</mark> a Acaaaatcata
SmR2-2/1-579	TECEAGACTETC	TCTACEACCAL	BCTCBCCATCA	ACACTGGCCAGC	CA <mark>GTT</mark> GACAAGAT	CCACAAGGACTT	GGACAGGATGAAGA	TTTTCAGCG	CTCAAGAGGCT	CTAGACTAC	GGCCTTATTG	ACAAGATCGTG
PtR2/1-5/9	ATCAGGGACTACC	TTTACAATGA	GTTATCTAAGA	IAAACAGGACAGC	CTTTTGAAAAGA		GAGCAGAA TGAAGC	GCTTTGAAG	CCCAAGA <mark>T</mark> GCC	CTTGAGTAT	GGTCTCATTG	ATCECATAATC
OsR2/1-579 ZmR2/1-579	ATCAAGAACTATC Atcaa <mark>g</mark> aactatc	TTTACAGCAAU TCTA <mark>TGGC</mark> AAU	GCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	ACACA <mark>GGT</mark> CATC Acaca <mark>gg</mark> cca <mark>tt</mark>	CTGTTGATAAGA CTGTC <mark>G</mark> A <mark>G</mark> AA <mark>G</mark> A1	I TCATGAAGATCI I <mark>CC</mark> atgaggacc <mark>i</mark>	ATCTAGGGTGAAGC GTCTAGGGTGAAGC	GCTTCGATG GCTTCGACG	CTGAAGGAGCT CC <mark>G</mark> AA <mark>ggggg</mark> CT	CTCGAGTAC	GGGATTATTG/ GGG <mark>ATC</mark> ATCG/	ACCUTATIATI ACC <mark>ut</mark> ati <mark>c</mark> i
Se R/1-579 Ss R/1-579	AACAA <mark>b</mark> cbcacca Aacaaacabacta	TGCTGGAGAT(TGCTGGAAAT	CTTT <mark>G</mark> CTC <mark>G</mark> CA ICTTTCCCTCA	ACACA <mark>GGT</mark> CAA <mark>G</mark> ACAC TGG CCAAA	ATCCCGATCGCT CCCAGGAAAAGT	T <mark>gg</mark> cccgcgacac Tagctaaagacat	CGATCGCATGCTTT GGACCGAACCTTTT	A <mark>TATG</mark> ACTC ACCTCACCC	CGGCTCAGGCT CTGCCCAGGCC	GTGGAGTAT AAGGAGTAT	GGCTTGATCG/ GGACTGATTG/	A <mark>tcgcgtcct</mark> c Accgggtgtta
MaR/1-579	ACAAACA <mark>G</mark> AC <mark>G</mark> A	TOTTOCAAATT	TTTCTCCAAAAA	ATACEGETCAAA	CCACCGAAAAAC	I GGCGAAAGATAT	GGATCGCACCTTCT	ATCTCACCC	CACAACA <mark>GGCG</mark>	AAA <mark>G</mark> ACTAC	GGATTAATCG	ATCCCCCTTCTC
AsR/1-549	AACAAA <mark>GGTTC</mark> AA	TGGTTGATAT	TTTC <mark>G</mark> CCC <mark>G</mark> CA	CCACCGGACAGG	CACCAGAAAAAA	CACCAAAGACAT	GGATCGTCTCCTAT	ACATGACTC	CTTATCAAGCC	AA <mark>GG</mark> ATTAC	GG <mark>TTT</mark> GATTG	ATAGAGTTTTT
CrR1/1-558	CCACAGAGTACT	ACATABACCII Ac <mark>gCtg</mark> ccat(CCTGGCTCGCT	CCACCGGCAAG	ACCTGGAGACCT	CCAGAAGCAGTA	CCTGTCGCGCAAGC	GCTACTICA GCTACTICI	CACAGGAGGCA C <mark>GGT</mark> GAA <mark>GGAG</mark>	GCATAC <mark>G</mark> AG	GAGGGCCTGG	T <mark>gg</mark> acaaactg
PhyR1-2/1-552 SmR1/1-552	AACAC <mark>gg</mark> attatt Aacacgaagtact	ACATA <mark>G</mark> ACTTU Acatt <mark>g</mark> acttu	STTGGCTTCGG Sttggcgaagg	GAACGGGCAAAT	CCAAGGAAGAGAGA CGGAAGAAGAAGAAA	TT <mark>GCTAATGATAT</mark> TTGCTAAAGATAT	TCGTAGGCCTAGGT CCAACGTCCCAAGT	ACTTCA <mark>GG</mark> C ACTTCAA <mark>GG</mark>	CGCAAGAGGCT CCAAAGAAGCT	ATT <mark>G</mark> actat Gtcgagtat	GGTCTTGCTG/ GGTCTGGCCG/	ACAA <mark>g</mark> atcata Acaaa <mark>gtg</mark> att
AtR1/1-552	AATACCGAGTACT	ACATCGAGCT(STTAGCTAAGG	GAACAGGGAAA	CCAAAGAGCAGA	CAACGAGGACAT	CAAGCGACCTAAAT	ATCTCCAAG	CTCAAGCAGCC	ATTGACTAT	GGAATTGCAG	ACAAGATAGCT
PtR1-2/1-93												
ZmR1/1-552	ACACAGATTACT AACACAGATTACT	ACCTCGAACT	CTT <mark>GTCT</mark> AAAG	GAGTTGGAAAAC	CAAAGGAAGAGC Caaa <mark>gg</mark> aagagc	TICT CAGAGETTICT	TAGGGGCCCGAGGT	ACTTCCGAG	CACAAGAGGCC CacaagaggCT	CTTGATTAT	GGACTTGCAG	ATACAATCITG Atac <mark>gatatt</mark> a
CrR3/1-555 PhyR3-1/1-555	ACAAGCGCGAGC Aaca <mark>ggg</mark> atacac	TECTECECATO TTETECECATO	ACTAGCCTGG	GCACTGGCCTGC A <mark>t</mark> actggcaatc	CC <mark>GTGG</mark> ACAA <mark>GT</mark> CCTATGAGAAAGGI	IGGACGCGGACAT Itgctaaagtgat	GCAACGGCCGCTGT GCAGAGGCCTTTTT	ACATGC <mark>GG</mark> C Aca <mark>tg</mark> aatc	CCCA <mark>gg</mark> acgc <mark>g</mark> C <mark>t</mark> aagaaagca	CTGGAGTAC GTGGAGTTT	GGCATCATTG GG <mark>TGTTGCT</mark> G	AC <mark>GAGAT</mark> CATC A <mark>t</mark> aagatattg
PhyR3-2/1-555	AATAGGGATACAC	TTGTGGGTT/	ACTTOCAAAGC	ATACTGGCAATC	CTGTTGAGAAGG	TECAAAETEAT	GGAAAGGCCTTTCT	ACATGAATC	CTAAGAAAGCG	GTGGACTTT ATGGATTTT	GGCGTCGTCG	ATAAGATATTG
AtR3/1-555	AATAGGGATATAC	TTGTGGAACTA	ACTATCAAAGC	ATACTGGGAATT	CC <mark>GTGGAGACTG</mark>	AGCTAACGTAAT	GAGAAGGCCATATT	ACATGGATG	CACCAAAAGC	AAAGAATTT	GGAGTCATTG	ACA <mark>gg</mark> a <mark>ttctt</mark>
Os R3/1-555	ACAGGAACACCT	TAGTGAGAACT	TTTAGCAAGC	ACACCEGGAAATT	CACCA <mark>G</mark> AGACAC	AGACAAGGTGAAT	GAGAGGACCATTTT	ATATGGATT ATATGGATT	CC <mark>TTG</mark> AA <mark>ggc</mark> a	AAGGAATTC AA <mark>gg</mark> a <mark>gttt</mark>	GG <mark>T</mark> GT <mark>CATT</mark> G	ACAA GATCOTT
ZmR3/1-555 CrR4/1-552	ACA <mark>gg</mark> aacaca <mark>t</mark> Aa <mark>g</mark> acc <mark>g</mark> ccacca	TGGTCAAAC <mark>t</mark> Tggccaa <mark>gt</mark> a(TTTAGCAAGG CC <mark>T</mark> GGCCGCCT	ATACTGGAAATC GCACCAAGAAGA	CACCAGAGAAAA C <mark>TG</mark> aggagcagai	TAGACAAGGTAAT TCA <mark>tg</mark> accgactt	GAGAGGACCA <mark>TTTT</mark> Ca <mark>cccg</mark> cccccgct	ACA <mark>tgg</mark> ac <mark>t</mark> A <mark>ctt</mark> caacc	CCTTGAAGGCA CC <mark>T</mark> AC <mark>G</mark> AGGCG	AAGGAGTTT <mark>Gt</mark> cagctac	GGAGTCATTG GGG <mark>CT</mark> CATCG	ACAA <mark>gat</mark> cctt Acac <mark>ggt</mark> cctg
PhyR4-1/1-546	TGAAGGAAGAGC	TCGTCAAGCTC	ATCTCAAGAC	ATACAGGTCAGA	CT GT GG AAAAGA	CACCBACBATAT	CAGGCGACCCAAGT	ATTTCAAAC	CCGA <mark>T</mark> GAAGCT	GTGGACTAT	GGCCTTATTG	ACAA <mark>ggttct</mark> c
PhyR4-3/1-546	TGAAGGAAGAGT	TCGTGAAGTTC	CTCTCAAGAC	ACACAGGACAAC	CTCCTGAGAAGA	AACAGGTGATAT	CAGGCGGCCGAAAT	ACTTCAAGC	CGGA <mark>T</mark> GGAGCT	GTTGATTAC	GGCATCATTG	ACAA <mark>GGTTCTG</mark>
3mr:4/1-546 AtR4/1-546	TAAAGACAGAAA	TGGTCAAGCT	TAGAGAGGC	ATACAGGCCAAC	CCCC <mark>GG</mark> A <mark>gC</mark> AGA	T GAAGC T GACAT	GAAACGCCCGAAAT	ACTITICACCC ATTITACTC	CCAC <mark>TG</mark> AGGCT	GTTGAATAT	GGGATCATTG	ATAAGGTGGTT
PtR4/1-546 OsR4/1-546	STGAAAACAGAA <mark>T</mark> agaaa	TGGTCAATCTC TGATTAAGCTC	CTTGGCAAAG <mark>C</mark> CTGTCAAGGC	ATATT <mark>GG</mark> AAAAT A <mark>T</mark> A <mark>TCGGT</mark> AAAT	TACCTGAGGAGA Ca <mark>gt</mark> agaagaga	TTGAAGCAGACAT TTGCTCAGGACAT	TAGTCGTCCAAAG CAAACGGCCCAAAT	ATTTTA <mark>GT</mark> C A <mark>TTTC</mark> AGTC	CAAC <mark>TG</mark> AA <mark>g</mark> ca Caa <mark>gtg</mark> aagct	GTTGAATAT GTGGATTAT	GGCATCATAG GGGATCATTG	ACAA <mark>ggt</mark> aa <mark>t</mark> a Acaa <mark>ggtg</mark> ctg
ZmR4/1-459	<mark>STCAAG</mark> A <mark>T</mark> A <mark>g</mark> aaa	T <mark>G</mark>	TTCATTCACT	TCATTGGTGATT	GGAA <mark>G</mark> AAC <mark>T</mark> AG							

			R4-StrepII		,	P3-	StrepII	
		nc	salt	+ salt		no salt	-	+ salt
Accession	СІр	core (350 kDa)	ring(s) (200 kDa)	ring (200 kDa)	core (350 kDa)	ring(s) (200 kDa) ^b	ring(s) (180 kDa) [♭]	ring (200 kDa)
ATCG00670.1	P1	5	2	6	43	11	40	
AT1G49970.1	R1	8	1	19	19	10	28	
AT1G12410.1	R2	6	4	30	25	20	60	
AT1G09130.1	R3	11	2	33	35	34	39	
AT4G17040.1	R4	10	4	43	57	26	77	
AT1G66670.1	P3	3	1		35	50	70	18
AT5G45390.1	P4	9	6	3*		122	127	72
AT1G02560.1	P5	8	4	2*	66	86	109	41
AT1G11750.1	P6	10	6		43	86	76	40
AT4G25370.1	T1	2			38	36	26	
AT4G12060.1	T2	3			6	2		

Supplemental Table 1. MS/MS analyses of the affinity-purified CIp assemblies^a

^a MS/MS analyses was performed from peptides extracted from the in-gel trypsin-digested protein bands from native gel separations of the affinity-purified Clp complexes (see Fig.1 for the corresponding gel images). Shown here are the total adjusted spectral counts per protein.

^b Two bands were observed at slightly different masses and were excised separately for in-gel digestion and MS/MS analysis. Nevertheless, both bands contained all the ClpP/R subunits.

* These are spectral counts from minor contaminating hydrophobic peptides of ClpP4 and ClpP5 that could not be fully removed from the column ("sticky peptides") even after two blank LC-MS runs.

Accession	Peptide	Position ^a	Peptide	Residue order ^b	Charge states	Number of Residues	[M+2H] ²⁺	[M+3H] ³⁺	Q?°
Clp core subunits									
ATCG00670.1	P1-1	19	IAFPHAR	245-251	2,3	7	406.2323, 271.1573	271.1573	Ν
	P1-2	6	SPGEGDTSWVDIYNR	86-100	2	15	848.3841		Y
AT1C40070 1	R1-1	11	YLQAQAAIDYGIADK	143-157	2	15	820.4199		Y
ATTG49970.1	R1-2	24	TAPPDLPSLLLDAR	323-336	2	14	739.9143		Y
AT1C12/10 1	R2-1	25	IALQSPAGAAR	337-347	2	11	527.8038		Y
ATTO12410.1	R2-2	4	FNAEEAIEYGLIDK	56-69	2	14	806.3987		Y
AT1G09130 1	R3-1	12	EPIYIYINSTGTTR	158-171	2	14	814.4199		Ν
ATT603130.1	R3-2	26	DILVELLSK	348-356	2	9	515.3132		Y
AT4G17040 1	R4-1	27	GSAHEQPPPDLASYLFK	357-373	2, 3	17	928.9625	619.6441	Y
A14017040.1	R4-2	14	YFSPTEAVEYGIIDK	181-195	2	15	866.4274		Y
	P3-1	7	DNTNLPSER	101-109	2	9	523.2491		Ν
AT1G66670.1	P3-2	20	LPSFEELDTTNMLLR	252-266	2	15	889.9533		Y
	P3-2 (ox)	20	LPSFEELDTTNM(ox)LLR	252-266	2	15	897.9508		Y
AT5G45390 1	P4-1	8	SFEQVLK	110-116	2	7	425.7371		Y
/10040000.1	P4-2	21	ADVSTIALGIAASTASIILGAGTK	267-290	3	24	1101.579		Ν
AT1G02560 1	P5-1	22	ANLNGYLAYHTGQSLEK	291-307	2, 3	17	939.9709	626.983	Y
ATTO02500.1	P5-2	9	FQSIISQLFQYR	117-128	2	12	765.4092		Y
AT1G11750 1	P6-1	23	IIFIGQPINAQVAQR	308-322	2	15	834.4832		Y
////00///	P6-2	10	VISQLVTLASIDDK	129-142	2	14	751.4272		Y
Clp chaperones/ad	aptors								
AT4G25370 1	T1-1	13	DETLSLLGK	172-180	2	9	488.2715		Y
/14020070.1	T1-2	28	AIAWAIDEK	374-382	2	9	508.7742		Y
AT4G12060 1	T2-1	29	ALDSALDQNLK	383-393	2	11	594.317		Y
/11/012000.1	T2-2	15	ILATLGFTDEK	196-206	2	11	604.3321		Y
AT1G68660 1	S-1	30	VILHNDNFNK	394-403	2, 3	10	607.3198	405.2157	Ν
/11/00000011	S-2	16	GGGVLDKPIIEK	207-218	2, 3	12	613.3612	409.2432	Y
AT5G50920.1	C1-1	31	VPEPTVDETIQILK	404-417	2	14	791.4403		Y
AT5G50920.1,	C1/2-1	17	VLENLGADPSNIR	219-231	2	13	699.3728		Y
AT3G48870.1	C1/2-2	32	GSGFVAVEIPFTPR	418-431	2	14	738.8959		Y
AT5G51070.1	D1-1	18	VVGQDEAVAAISR	232-244	2	13	657.8542		Y
	D1-2	33	VFEAAVEYSR	432-441	2	10	585.7931		Y
Localization marke	rs ^d								
AT1G04410.1	MD-1	1	VLVVANPANTNALILK	8-23	2	16	825.5011		Y
	MD-2	34	LSVPVSDVK	442-450	2	9	472.2766		Y
AT5G20290.1	RPS8-1	3	VLDVVYNASNNELVR	41-55	2	15	852.9494		Y
	RPS8-2	36	SAIVQVDAAPFK	465-476	2	12	623.3455		Y
AT3G04790.1	PRI-2	35	SLGIPLVGLDIHPR	451-464	2, 3	14	/3/.9225	492.2841	Y
	PRI-1	2	LLSSGELYDIVGIPTSK	24-40	2	17	896.4906		Y
AT5G51820.1	PGM1-1	5	VAEIPDIDLSQVGVTK	/0-85	2	16	842.4618		Y
	PGM1-2	37	IYGNTLSISEIK	477-488	2	12	669.3692		Y

Supplemental Table 2. The peptides comprising Clp-QconCAT and their corresponding properties.

^a Order by which these proteotypic peptide appeared starting from the N-terminus of the Clp-QconCAT construct

 $^{\rm b}$ Residue numbers of the peptide sequence within the Clp-QconCAT construct.

^c Can be detected and quantified. Y=yes, N = no.

^dThese proteins serve as reference for organellar location. Malate dehydrogenase (MD) and the 40S ribosomal subunit S8 (RPS8) are both cytosolic whereas ribose 5-phosphate isomerase (PRI) and phosphoglucomutase (PGM) are both plastid-localized.

Peptid	e Position	slope	y-intercept	Regression coefficient
ClpP1/R ring	3			
P1-2	6	1.40	-0.05	0.995
R1-1	11	1.09	-0.04	0.971
R1-2	24	1.20	-0.04	0.990
R2-1	25	1.05	0.14	0.999
R2-2	4	1.12	0.04	0.987
R3-2	26	0.76	0.13	0.968
R4-1	27	1.13	-0.06	0.996
R4-2	14	1.32	-0.05	0.991
ClpP3-P6 rir	ng			
P3-2	20	0.89	0.26	0.987
P4-2	8	1.04	0.05	0.999
P5-1	22	1.827	-0.07	0.999
P5-2	9	0.57	0.42	0.956
P6-1	23	1.21	-0.05	0.994
P6-2	10	1.22	-0.33	0.975
Clp chapero	nes/adaptors			
T1-1	13	1.05	0.04	0.998
T1-2	28	1.10	0.07	0.998
T2-1	29	1.00	0.08	0.998
T2-2	15	0.72	0.19	0.987
S-2	16	0.86	0.12	0.998
C1/2-1	17	1.07	0.07	0.997
C1/2-2	32	0.70	0.30	0.971
C1-1	31	0.92	0.02	0.986
D1-1	18	1.08	0.09	0.999
D1-2	33	1.31	0.004	0.993
Localization	markers			
MD-1	1	1.11	-0.09	0.992
MD-2	34	1.03	0.002	0.998
PGM1-	1 5	1.09	-0.09	0.988
PGM1-	2 37	1.30	-0.06	0.996
PRI-1	2	0.87	0.13	0.981
PRI-2	35	1.14	0.03	0.996
RPS8-	1 3	1.26	-0.04	0.992
RPS8-2	2 36	0.85	0.15	0.989

Supplemental Table 3. Linear instrument response for individual Clp-QconCAT peptides.*

^{*} For 100 fmol of heavy Clp-Qconcat peptides mixed with varying amounts of their light versions spanning two orders of magnitude (10 fmol to 1 pmol) with L/H ratios of 0.1, 0.2, 0.5, 1, 2, 5 and 10.

Supplemental Table 4. Quantification of Clp subunits and determination of subunit stoichiometry in the Clp core complex and constituent Clp rings.^a

Clp core complex (350 kDa)

Clp	Рер			Quanti	ty (x 1	00 fmo	1)					Sto	oichior	netry		
		R4-1	%CV	R4-2	%CV	P3-1	%CV	P3-2	%CV	R4-1	R4-2	P3-1	P3-2	AVE	SD	%CV
P1	Q6	1.62	1.2	0.59	1.2	3.30	1.4	4.37	0.8	3.35	3.15	2.51	2.72	2.93	0.38	13.1
R1	Q11	0.59	0.9	0.28	0.9	1.31	0.9	1.67	0.1	1.22	1.50	1.00	1.04	1.19	0.23	19.0
D 2	Q25	0.74	4.9	0.29	4.9	1.45	1.0	1.82	0.8	1.53	1.54	1.10	1.13	1.33	0.24	18.2
112	Q4	0.33	0.5	0.15	0.5	0.90	1.2	1.30	1.1	0.69	0.82	0.69	0.81	0.75	0.07	9.5
R3	Q26	0.47	0.04	0.18	0.04	1.04	1.2	1.74	0.6	0.97	0.97	0.79	1.08	0.95	0.12	12.6
R/	Q27	0.51	1.5	0.17	1.5	0.92	3.2	1.22	0.4	1.05	0.92	0.70	0.76	0.86	0.16	18.4
114	Q14	0.46	1.9	0.20	1.9	1.11	0.1	1.45	0.6	0.95	1.08	0.85	0.90	0.94	0.10	10.5
TOTAL [♭]						7.85		10.68								

									Clp r	ing wi	th no s	salt ind	cubatio	ons (20	00 kDa)						
Clp	Рер					Qua	ntity (x	100 fm	ol)								Stoi	chiome	try			
		R4-3	%CV	R4-4	%CV	P3-3	%CV	P3-4	%CV	P3-5	%CV	P3-6	%CV	R4-3	R4-4	P3-3	P3-4	P3-5	P3-6	AVE	SD	%CV
P1	Q6	4.56	3.6	0.86	4.4	1.19	1.6	1.75	1.0	6.45	1.5	6.89	0.4	3.19	3.13	3.34	3.32	3.46	3.53	3.33	0.15	4.6
R1	Q11	1.54	0.8	0.38	0.2	0.56	0.9	0.67	0.4	2.17	0.4	2.28	0.5	1.08	1.40	1.57	1.28	1.16	1.17	1.28	0.18	14.3
D 2	Q25	1.84	2.4	0.49	3.6	0.67	0.6	0.81	1.9	2.30	1.9	3.04	1.1	1.29	1.80	1.89	1.54	1.24	1.55	1.55	0.26	16.9
112	Q4	0.93	1.2	0.24	3.3	0.27	0.7	0.47	0.9	1.74	15.5	1.79	0.3	0.65	0.89	0.75	0.90	0.93	0.92	0.84	0.11	13.4
R3	Q26	1.35	0.8	0.29	1.0	0.38	1.1	0.51	1.0	1.77	7.0	2.47	0.9	0.94	1.07	1.06	0.97	0.95	1.26	1.04	0.12	11.6
D4	Q27	1.43	0.5	0.23	2.5	0.37	1.3	0.48	3.1	1.69	2.5	1.82	0.7	1.00	0.84	1.03	0.91	0.91	0.93	0.94	0.07	7.3
1.4	Q14	1.43	0.2	0.32	1.9	0.34	1.3	0.57	0.8	2.04	1.4	2.09	1.6	1.00	1.16	0.97	1.09	1.09	1.07	1.06	0.07	6.4

		1	purifie	d Clp r	ing fro	om on-c	olumn	salt ind	cubatio	on (200) kDa)	
Clp		Qua	ntity (x	x 100 f	mol)			S	toichic	metry		
		%CV	R4-6	%CV	R4-7	%CV	R4-6	R4-7	R4-8	AVE	SD	%CV
P1	%CV R4-6 %CV R4-7 %CV R4-6 R4-7 R4-8 A Q6 1.5 2.74 1.0 1.15 1.2 2.79 3.38 2.68 2									2.95	0.38	12.7
R1	Q11	0.4	0.96	0.9	0.58	0.6	1.10	1.19	1.35	1.21	0.12	10.1
D 2	Q25	6.8	1.03	8.4	0.52	0.2	1.42	1.27	1.21	1.30	0.11	8.1
Π2	Q4	0.6	0.68	0.3	0.39	0.4	0.67	0.84	0.91	0.81	0.12	15.3
R3	Q26	0.9	0.81	0.6	0.58	1.2	0.62	1.00	1.34	0.99	0.36	36.9
D/	Q27	1.5	0.73	1.1	0.38	1.7	0.92	0.90	0.89	0.90	0.02	1.7
N4	Q14	0.8	0.89	1.7	0.48	0.6	1.08	1.10	1.11	1.10	0.02	1.4

	Clp core complex (350 kDa)															
Clp	Рер			Quanti	ity (x 1	00 fmo	I)					Sto	oichior	netry		
		R4-1	%CV	R4-2	%CV	P3-1	%CV	P3-2	%CV	R4-1	R4-2	P3-1	P3-2	AVE	SD	%CV
P3	Q20	0.47	1.8	0.26	0.5	1.60	2.8	1.84	0.4	0.98	1.38	1.22	1.15	1.18	0.16	13.8
P4	Q8					2.82	1.6	3.53	1.2			2.15	2.20	2.17	0.04	1.6
DE	Q22	1.60	1.0	0.48	1.5	2.51	1.3	3.74	3.4	3.31	2.56	1.91	2.33	2.53	0.59	23.2
гJ	Q9	1.08	0.9	0.50	0.4	4.86	0.9	6.25	0.7	2.24	2.68	3.70	3.90	3.13	0.80	25.5
De	Q23	0.72	1.8	0.32	1.1	1.50	1.2	1.86	0.8	1.48	1.68	1.14	1.16	1.37	0.26	19.3
FU	Q10	0.69	0.4	0.24	2.6	1.13	1.9	1.35	0.3	1.42	1.27	0.86	0.84	1.10	0.29	26.7
TOTAL)					9.43		11.97								
T1	Q28	0.42	0.4	0.16	2.2	0.96	0.4	1.29	0.4	0.86	0.83	0.73	0.80	0.81	0.06	7.1

	Clp ring with no salt incubations (200 kDa)																		
				Qua	ntity (x	100 fn	nol)							St	oichio	metry			
Clp	Рер	R4-3	%CV	P3-3	%CV	P3-4	%CV	P3-5	%CV	P3-6	%CV	R4-3	P3-3	P3-4	P3-5	P3-6	AVE	SD	%CV
P3	3 Q20 0.24 13.5 1.70 0.4 2.19 0.3 2.22 14.5 3.56 0.9 0.73 1															1.02	0.98	0.18	18.6
P4	P4 Q8 3.75 1.9 5.24 0.7 5.68 6.0 9.25 0.9 2.57 2.71 2.20 2.66															2.53	0.23	9.1	
P5	Q22	0.86	0.42	2.47	1.7	5.59	0.6	6.63	2.0	9.82	1.2	2.60	1.69	2.89	2.57	2.83	2.52	0.48	19.2
15	Q9	0.68	1.71	4.33	1.2	7.54	0.9	7.56	0.5	12.91	0.7	2.04	2.96	3.90	2.93	3.72	3.11	0.74	23.9
P6	Q23	0.35	0.87	1.69	0.3	2.27	0.7	3.01	1.3	3.78	1.2	1.05	1.16	1.17	1.16	1.09	1.13	0.05	4.8
10	Q10	0.31	2.33	1.23	0.2	1.60	1.4	2.16	0.5	3.16	2.0	0.95	0.84	0.83	0.84	0.91	0.87	0.05	6.2
TOTAL ^b																			

T1	Q28	0.85	0.72	1.31	0.74	0.72	2.35	1.31	0.23	0.58	0.68	0.28	0.38	0.48	0.18	38.1
----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

	purified Clp ring from on-column salt incubation (200 kDa)															
	Quantity (x 100 fmol)									Stoichiometry						
Clp	Рер	P3-7	%CV	P3-8	%CV	P3-9	%CV	P3-10	%CV	P3-7	P3-8	P3-9	P3-10	AVE	SD	%CV
P3	Q20	0.57	1.8	1.52	с	0.59	1.2	2.14	1.5	0.99	0.95	0.86	1.15	0.99	0.12	12.3
P4	Q8	1.37	1.2	3.23	3.4	1.48	1.6	4.92	2.1	2.40	2.02	2.14	2.64	2.30	0.28	12.0
D5	Q22	0.85	0.5	3.62	4.4	1.55	1.5	4.76	0.2	1.49	2.26	2.24	2.55	2.14	0.45	21.3
FJ	Q9	1.45	1.1	5.00	1.2	2.11	0.5	7.73	0.8	2.54	3.12	3.05	4.15	3.21	0.67	21.0
De	Q23	0.65	1.3	1.78	1.6	0.76	1.3	2.08	0.4	1.14	1.11	1.10	1.12	1.12	0.02	1.8
FU	Q10	0.49	1.7	1.43	1.8	0.63	0.3	1.65	1.8	0.86	0.89	0.90	0.88	0.88	0.02	2.2

^a Based on affinity purification and native gel separation of various Clp assemblies from R4-StrepII (n=8) or P3-StrepII (n=10) transgenic lines . Molar amounts are derived from peak area ratios between the endogenous Clp peptides and 100 fmol spiked Clp-QconCAT peptides). ClpP4 could not be quantified in runs R4-1, R4-2 and R4-5. Stoichiometry was calculated by normalization against the molar amounts of ClpR4 or ClpP6 proteins. Peak areas for peptides identified with multiple charge states (R4-1, Q27 and P5-1, Q22) or modifications (P3-2, Q20 abd P5-1, Q22) were combined prior to calculating the light-to-heavy ratios.

^b Sum of the molar amounts of the constituent Clp subunits per ring. The molar amounts for proteins quantified with two peptides were averaged. The average ratio between the total moles of ClpP1/R and ClpP3-6 within the Clp core is 1.16 for n=2 (Runs P3-1 and P3-2).

Protein Accessions (Prokaryotic Clps) Proteo-bacteria P Cyano P1 Cyano P2 Cyano P3 Cyano R POA6G7 E. coli (Ec) Rickettsia prowazekii (Rp) RP520 WD0319 Wolbachia wmel (Ww) Anabaena sp. (strain PCC Q8YXH5 Q8YQX8 Q8YP43 Q8YP42 7120) (As) Synechocystis sp. (strain P54416 Q59993 P74467 P74466 PCC 6803) (Ss) Synechococcus elongatus Q9L4P4 P54415 O34125 Q9L4P3 (strain PCC 7942) (Se) MAE 11870 MAE 62720 MAE 57180 MAE 57190 Microcystis aeruginosa (Ma) Prochlorococcus marinus PMM0742 PMM1656 PMM1314 PMM1313 MED4 (Pm) cDNA Accessions (prokaryotic Clps) Proteo-bacteria P Cyano P1 Cyano P2 Cyano P3 Cyano R ENA|AAA23588 E. coli (Ec) Rickettsia prowazekii (Rp) RP520 Wolbachia wmel (Ww) WD0319 gi|17227497: gi|17227497: Anabaena sp. (strain PCC gi|17227497: 5219511-5220104 gi|17227497: 5220184-5220846 7120) (As) 1470240 -1470854 4443275-4443973 Synechocystis sp. (strain gi|16329170: gi|16329170: gi|16329170: 2206551-2207159 gi|16329170: 2205774-2206451 PCC 6803) (Ss) 3212598-3213194 3245421-3246101 Synechococcus elongatus ENA|AAC67306 ENA|AAB68677 ENA|AAB68677 ENA|CAB81780 (strain PCC 7942) (Se) Microcystis aeruginosa (Ma) MAE_11870 MAE_62720 MAE_57180 MAE_57190 Prochlorococcus marinus PMM0742 PMM1656 PMM1314 PMM1313 MED4 (Pm) Protein Accessions (Plastid Clps) **P1 P**3 Ρ4 P5 P6/R6 Chlamydomonas reinhardtii P42380 A8IJ60 A8IL21 A8INX1 (Cr) Physcomitrella patens P3-1: A9RYV9, P3-2: P4-1: A9TZR6, P4-2 P6-1: A9TH43, P6-2: A9SWP4 Q6YXM7 P5-1: A9S5E8, P5-2: A9TE25 (PhvP) A9T6I1 A9RE44 Selaginella moellendorfii C7B2H1 XP_002960346.1 XP_002981945.1 XP_002963649.1, XP_002983780.1 XP_002968942.1 (Sm) ATCG00670.1 AT1G66670.1 AT5G45390.1 AT1G02560.1 AT1G11750.1 Arabidopsis thaliana (At) P4-1: B9GZW8, P4-Populus trichocarpa (Pt) A4GYT6 B9H362 P5-1: B9GST6, P5-2:B9I9I1 P6-1: A9PDP3, P6-2: B9HRQ3 2: A9PA38 Zea mays (Zm) ("GRM" ZM2G448161_P01, ZM2G001755 P01 ZM2G121456_P01 ZM2G092632 P01 ZM2G427444_P06 removed from Accessions) Os08g15270, Os10g21300, Osp1g005990, 0s01g32350.1 Os10g43050.1 Os03g19510.1 Os03g29810.1 Oryza sativa (Os) Os12q10590 Q8WHZ7 Psilotum nodum (Pn) Adiantum capillus renesis Q85FJ8 (Ac) Pinus thunbergii (Pnt) P41609 cDNA accessions (for plastid ClpPs) **P1 P**3 **P4** P5 P6/R6 gi|159465134: 119-Chlamydomonas reinhardtii ENA|ACJ50097 gi|159465360: 108-878 gi|159466487: 159-1010 (Cr) 1156 Physcomitrella patens gi|34501376:11632-11702, P3-1: gi|168013185, P4-1: gi|168064315 P5-1: gi|168018158: 395-1306, P5-2: P6-1: gi|168050942, P6-2: (PhyP) 12338-12632, 13074-13307 P3-2: gi|168043510 P4-2: gi|167999543 gi|168048811 gi|168036512 Selaginella moellendorfii gi|255961289: c>68286-XM 002960300.1 XM_002981899.1 XM_002963603.1 XM_002983734.1 XM_002968896.1 68216, c67451-66902 (Sm)

Supplemental Table 5. Accession numbers for protein and cDNA sequences of the Clp proteins from prokaryotes, green algae and plants.

Arabidopsis thaliana (At)	dopsis thaliana (At) ATCG00670.1		gi 145358888	gi 145334998	gi 145335409	
Populus trichocarpa (Pt)	ENA ABO36731	gi 224079412	P4-1: gi 224074499:77-739 , P4-2: EF145046	P5-1: gi 224068557: 123-1025, P5- 2: gi 224128301: 37-543	P6-1: gi 224080314: 72-869 , P6-2: gi 224103436:1-741	
Zea mays (Zm) ("GRMZM" removed from Accessions)	P1-1: 2G448161_T01,P1- 2: 2G427444_T06	2G001755_P01	2G001755_T01	2G121456_T03	2G121456_T03	
Oryza sativa (Os)	P1-1: 13108.m09245, P1-2: 13110.m07757, P1-3: gi 11466763: 67638- 68288, P1-4: 13112.m01122	13101.m03260	13110.m04073	13103.m02334	13103.m03417	
Psilotum nodum (Pn)	ENA BAB84241					
Adiantum capillus renesis (Ac)	ENA AAP29415					
Pinus thurnbergii (Pnt)	gi 7524593: 29611-30201					
	Prote	in accessions (p	lastid ClpR and r	nitochondrial ClpP2)		
	R1	R2	R3	R4	mito P2	
Chlamydomonas reinhardtii (Cr)	A8INX1	jgi Chlre4 183767	A8I547	A8IH07	A8IX06	
Physcomitrella patens (PhyP)	P6-1: A9TH43, P6-2: A9SWP4	R1-1:A9TVY1, R1-2: A9SI89	R2-1: A9SWI4, R2-2: A9T3W3	R3-1: A9SG20, R3-2: A9RZ41	R4-1: A9RQL9, R4-2: A9TF07, R4- 3: A9SSY7	
Selaginella moellendorfii (Sm)	XP_002968942.1	XP_002970082.1	XP_002968834.1	XP_002990376.1	XP_002982594.1	
Arabidopsis thaliana (At)	AT1G11750.1	AT1G49970.1	AT1G12410.1	AT1G09130.1	AT4G17040.1	
Populus trichocarpa (Pt)	P6-1: A9PDP3, P6-2: B9HRQ3	R1-1: B9HMY7, R1- 2: B9NHA6	B9N3L0	A9PFT0	A9PAS6	
Zea mays (Zm)	ZM2G092632_P01	ZM2G099529_P01	ZM2G148106_P01	ZM2G030072_P02	AC207652.3_FGP003	
Oryza sativa (Os)	Os03g29810.1	Os05g51450.1	Os06g04530.1	Os03g22430.1	Os01g16530.1	
	•DNA	, accessions (Di	otid ClpB and a	mitachandrial CInP2)		
	CDNA					
	R1	R2	R3	R4	mito P2	
Chlamydomonas reinhardtii (Cr)	jgi Chlre4 183767 estExt_fgenesh2 _kg.C_10275	gi 159467460: 144-992	gi 159464778: 1· 1248	gi 159470090: 140-970	gi 159477908	
Physcomitrella patens (PhyP)	R1-1: gi 168061570, R1-2: gi 168026988	R2-1: gi 168036369, R2-2: gi 168041626	R3-1: gi 168025571, R3-2: gi 168013354	R4-1: gi 168007319, R4-2: gi 168049468, R4-3: gi 168033938:1-714	gi 168031791	
Selaginella moellendorfii (Sm)	XM_002970036.1	XM_002968788.1	XM_002990330.1,	XM_002982548.1	XM_002983734.1	
Arabidopsis thaliana (At)	gi 145336568	gi 145335431	gi 145335309	gi 30683910	gi 186524962: 74-799	
Populus trichocarpa (Pt)	R1-1:gi 224103734:80- 1255, R1-2: gi 224097119	gi 224125071: 126-965	gi 224125177: 74-1099	gi 224126446: 25-957	P2-1: gi 224081362: 28-762 , P2-2: gi 224094112: 142-876 , P2-3: gi 224094112: 142-876	
Zea mays (Zm) ("GRMZM" removed from Accessions)	M" 2G099529_T01 2G148106_T01 2G030072_T02 AC207652.3_FGT003		AC207652.3_FGT003	P2m-1: 2G320135_T02, P2m-2: 2G474883_T01, P2p-1: 2G053236_T01, P2p-2: 2G111756_T01		
Oryza sativa (Os)	13105.m05548	13106.m00412	13103.m02699	13101.m01848	P2-1: 13104.m04426, P2-2: 13102.m04675	

Supplemental Table 6. Primers used in this study

Construct	Primers (5' to 3')
R4cDNAStrepII Step1, for PCR amplification	1: CCATGGAGGTAGCAGCAGCGA 2: AGGATGAGACCAAATGAGTTGTGCC
R4cDNAStrepII Step2, for PCR amplification	1: CACCATGGAGGTAGCAGCAGCGAC 2: TCACTTCTCGAATTGAGGATGAGACCA
R4StrepII in pEARLEYGATE100, for genotyping	1: GGGAATGTCTCTCGTACCTTCAGTT 2: GGCGCTCTATCATAGATGTCGCTATAAACC
R4StrepII in pEARLEYGATE100, for genotyping	1: CCATGGAGGTAGCAGCAGCGA 2: GGCGCTCTATCATAGATGTCGCTATAAACC
R4 gene specific, for genotyping	1; GGGAATGTCTCTCGTACCTTCAGTT 2: ACAACTGGACACTGTTGCATAATGA
P3 cDNA Strep II, for PCR amplification	1: ATGGAGATGAGTTTGCGTCTCGCTTC 2:CTACTTCTCGAATTGAGGATGAGACCATTCAA T GGCGGCATAACCATTCTGTGTC
P3StrepII in pEARLEYGATE100, for genotyping	1: ATGGAGATGAGTTTGCGTCTCGCTTC 2: GGCGCTCTATCATAGATGTCGCTATAAACC
TDNA insertion, for genotyping	1: GGCAATCAGCTGTTGCCCGTCTCACTGGTG 2: ACAACTGGACACTGTTGCATAATGA
ClpQconCAT in pET21, for confirmation	1: TTATGCTAGTTATTGCTCAGCGGTG 2: CCATTTCTCGTATCGCATTTCCAC
ClpQconCAT in pET21, for confirmation	1: TCCGGCGTAGAGGATCGAGATC 2: TGGAAATGCGATACGAGAAATGG

Supplementary Text. Selection of the Clp-QconCAT peptides. Sequence coverage of the subunits of the plastid-localized Clp protease complex from extensive MS analysis of *Arabidopsis* chloroplasts, total leaf extracts and purified Clp complexes. The identified tryptic peptides are show in red and are separated by bars. The predicted chloroplast transit peptide sequence is underlined. Chemically reactive residues (methionine and cysteine) and positions in the sequence that confer instability (e.g., N-terminal Asn, N-terminal Gln, Asn-Gly) or missed cleavages (e.g., Pro after Lys or Arg) within the identified peptides are shown in green. The selected Clp-QconCAT peptides are highlighted in yellow. Additional comments regarding the behavior of the selected Clp peptides during the Clp-QconCAT quantification experiments are also included.

ClpP1

MPIGVPKVPFR|<mark>SPGEGDTSWVDIYNR</mark>|LYRER|LFFLGQEVDTEISNQLISL<mark>M</mark>IYLSIEK|DTKD LYLFINSPGGWVISGMAIYDTMQFVR|PDVQTI<mark>CM</mark>GLAASIASFILVGGAITK|R|<mark>IAFPHAR</mark>|V <mark>MIHQPASSFYEAQTGEFILEAEELLK|LR|ETITR|VYVQR|TG<mark>KF</mark>IWVISED<mark>M</mark>ER|DVFMSATE AQAHGIVDLVAVQ</mark>

For ClpP1, only one peptide (SPGEGDTSWVDIYNR, P1-2) passed the requirements for a good quantifiable peptide. It harbors tryptophan which can be oxidized. However, peak area analyses of the unmodified peptide and its various oxidized forms (e.g., kynurenine, hydroxytryptophan and dihydroxytryptophan (see also (Perdivara et al., 2010)) showed that 98% of this peptide remained unoxidized in our Clp-QconCAT experiments. IAFPHAR (P1-1) was rarely detected in our extensive MS analysis of various *Arabidopsis* large-scale proteome experiments but we still included it in our Clp-QconCAT construct hoping that we would observe it in purified Clp complexes. However, we still did not detect this peptide in our Clp purifications most likely due to its short peptide length and high hydrophilicity. Most of the other peptides are long, methionine-bearing peptides that could be differentially modified and could be very hydrophobic.

ClpR2

MAVSFNTTLHQPSLSPSCSIKLYSGLKPQSASFLASGYQNLNKEFYGRVYKSLQ</u>SGTGKASR SRVKMMPIGTPRVPYR|NR|EEGTWQWVDIWNALYR|ER|VIFIGQNIDEEFSNQILATMLYL DTLDDSR|R|IYMYLNGPGGDLTPSLAIYDTMK|SLK|SPVGTHCVGLAYNLAGFLLAAGEK|G HR|FAMPLSR|IALQSPAGAAR|GQADDIQNEAK|ELSR|IR|DYLFNELAK|NTGQPAER|VFK|D LSR|VK|R|FNAEEAIEYGLIDK|IVRPPR|IKEDAPRQDESAGLG We were able to quantify the two ClpR2 peptide standards which exhibited good instrument responses. However, peptide IALQSPAGAAR (R2-1) yielded a slightly higher stoichiometry value than the other peptide standard (see Table 2) and this might have been due to a potential missed cleavage within the Clp-QconCAT construct (see Methods section for the Clp-QconCAT sequence) yielding slightly lower amounts of the labeled standard and higher sample-to-standard ratios.

ClpR1

MATALVSPLTSQLNHEAVCSKFVLPKSPFMSGSKLFSSNMPCSTVPRRTRRSHCFASAKDM SFDHIPKQFRGDNLKDGVMQNFKNVPQYFYGLNSAQMDMFMTEDSPVRRQAEKVTEESIS SRNNYLNNGGIWSMSGMNAADARRYSMSVQMYRGGGGGGGGSERPR|TAPPDLPSLLLDAR| ICYLGMPIVPAVTELLVAQFMWLDYDNPTKPIYLYINSPGTQNEK|METVGSETEAYAIADTI SYCK|SDVYTINCGMAFGQAAMLLSLGK|KGYRAVQPHSSTKLYLPKVNR|SSGAAIDMWIK| AK|ELDANTEYYIELLAK|GTGK|SK|EQINEDIK|R|PK|YLQAQAAIDYGIADK|IADSQDSSEK| R|DYDGTLAQR|AMRPGGGSPAAPAGLR|

The two peptides were both observed in our large-scale experiments. However, the peptide YLQAQAAIDYGIADK (R1-1) exhibited low signal-to-noise ratio even for the stable isotope-labeled version (see Figure 2) and is thereby not amenable for quantification. The rest of the tryptic peptides were not selected since they exhibited weak MS intensity response (e.g., peptides EQINEDIK and IADSQDSSEK) and were occasionally detected, was part of a missed cleavage (peptide DYDGTLAQR) or possessed methionine/cysteine residues.

ClpR3

MASCLQASMNSLLPRSSSFSPHPPLSSNSSGRRNLKTFRYAFRAKASAKIPMPPINPK|DPFLS TLASIAANSPEK|LLNRPVNADVPPYLDIFDSPQLMSSPAQVER|SVAYNEHRPR|TPPPDLPSM LLDGR|IVYIGMPLVPAVTELVVAELMYLQWLDPK|EPIYIYINSTGTTR|DDGETVGMESEG FAIYDSLMQLK|NEVHTVCVGAAIGQACLLLSAGTK|GKRFMMPHAK|AMIQQPRVSSGLMP ASDVLIR|AK|EVITNR|DILVELLSK|HTGNSVETVANVMR|R|PYYMDAPKAK|EFGVIDR|ILW RGQEK|IIADVVPSEEFDK|NAGIKSVV

Similar to ClpR1, the two selected peptides for ClpR3 were both observed in our large-scale experiments. However, the peptide EPIYIYINSTGTTR (R3-1) exhibited low signal-to-noise ratio even for the stable isotope-labeled version (see Figure 2) and is thereby not amenable for quantification. The

other tryptic peptides also exhibited weak MS intensity response (peptide EFGVIDR), was part of a missed cleavage (peptide EVITNR) or have reactive residues and are thus not good for quantification. Peptide IIADVVPSEEFDK could have been chosen but this was occasionally detected from our initial experiments and was thus not selected included then.

ClpR4

<u>MEVAAATATSFTTLRARTSAIIPSSTRNLRSKPRFSSSSSLRASLSNGFLSPYTGGSISS</u> <u>DLCGAKLR</u>AESLNPLNFSSSKPKRGVVTMVIPFSK|<mark>GSAHEQPPPDLASYLFK|</mark>NRIVYLGM SLVPSVTELILAEFLYLQYEDEEKPIYLYINSTGTTKNGEKLGYDTEAFAIYDVMGYVKP PIFTLCVGNAWGEAALLLTAGAKGNR|SALPSSTI<mark>M</mark>IK|QPIAR|FQGQATDVEIAR|K|EIKHI KTEMVKLYSKHIGK|SPEQIEAD<mark>M</mark>K|RPK|<mark>YFSPTEAVEYGIIDK</mark>|VVYNER|GSQDR|GVVSDL K|K|AQLI

Both ClpR4 peptides were reliably quantified.

ClpP3

MEMSLRLASSSTSNPICLLNPGKNLNFPIRNHRIPKTSKPFCVRSSMSLSKPPR|QTLSSNWDV SSFSIDSVAQSPSR|LPSFEELDTTNMLLR|QR|IVFLGSQVDDMTADLVISQLLLLDAEDSER|D ITLFINSPGGSITAGMGIYDAMK|QCK|ADVSTVCLGLAASMGAFLLASGSK|GKRYCMPNSK| VMIHQPLGTAGGK|ATEMSIRIREMMYHKIKLNKIFSR|ITGKPESEIESDTDR|DNFLNPWEA K|EYGLIDAVIDDGKPGLIAPIGDGTPPPK|TK|VWDLWK|VEGTKK|DNTNLPSER|SMTQNG YAAIE

For ClpP3, most identified peptides had methionine or had potential internal missed cleavage sites (KP in their sequences). We resorted to selecting a methionine-bearing peptide (LPFEELDTTNMLLR, P3-2) that exhibited good MS instrument response. We also included DNTNLPSER (P3-1) which was rarely observed. We did not observe peptide DNFLNPWEAK in our initial experiments and was thereby not selected but we had been detecting it in enriched Clp complex purifications.

ClpP4

MGTLSLSSSLKPSLVSSRLNSSSSASSSSFPKPNNLYLKPTKLISPPLRTTSPSPLR|FAN ASIEMSQTQESAIR|GAESDVMGLLLR|ER|IVFLGSSIDDFVADAIMSQLLLLDAK|DPK|KDI K|LFINSPGGSLSATMAIYDVVQLVR|ADVSTIALGIAASTASIILGAGTK|GK|R|FAMPNTR|I MIHQPLGGASGQAIDVEIQAKEVMHNK|NNVTSIIAGCTSR|<mark>SFEQVLK</mark>|DIDR|YMSPIEA

VEYGLIDGVIDGDSIIPLEPVPDR|VKPR|VNYEEISK|DPMK|FLTPEIPDDEIY

Most detected peptides in ClpP4 have methionine and peptide VNYEEISK was only detected as part of missed cleavages. Only SFEVLK (P4-1) worked. We included ADVSTIALGIAASTASIILGAGTK (P4-2) but it was a very hydrophobic sticky peptide that does not completely elute during an LCMS run (we still detected in subsequent blank runs).

ClpP5

MAHACVSTSASSLRFTAGFVSASPNGSSFDSPKLSLPFEPLRSRKTNKLVSDRKNWKNSTPK AVYSGNLWTPEIPSPQGVWSIRDDLQVPSSPYFPAYAQGQGPPPMVQER|FQSIISQLFQYR|I IR|GGGAVDDDMANIIVAQLLYLDAVDPTK|DIVMYVNSPGGSVTAGMAIFDTMR|HIR|PDVS TVCVGLAASMGAFLLSAGTK|GK|R|YSLPNSR|IMIHQPLGGAQGGQTDIDIQANEMLHHK|A NLNGYLAYHTGQSLEK|INQDTDR|DFFMSAKEAKEYGLIDGVIMNPLK|ALQPLAAA

Peptide YSLPNSR was only detected in one experiment and peptide INQDTDR was only observed as part of a missed cleavage. The selected peptide FQSIISQLFQYR (P5-2) elutes later in the LC gradient although this does not fully explain why it exhibits a higher stochiometry ratio than peptide P5-1 (ANLNGYLAYHTGQSLEK). Deamidation of Asn, most common in -Asn-Gly-sequences, yields aspartic acid (Asp) and isoaspartic acid (β Asp) and is usually observed in proteomic workflows involving overnight tryptic digestion (Krokhin et al., 2006). Two peaks were observed for the deamidated species (aspartic acid and isoaspartic acid are isomeric) of P5-1 and one species eluted closely with the unmodified P5-1. Previous studies revealed that reversed-phase LC using formic acid as ion-pairing modifier separated native and deamidated peptides in the order: -Asn- /- β Asp-/-Asp- (Krokhin et al., 2006).

ClpP6

MAGLAISPPLGLSFSSRTRNPKPTSFLSHNQRNPIRRIVSALQSPYGDSLK|AGLSSNVSGSPIK| IDNKAPRFGVIEAK|K|GNPPVMPSVMTPGGPLDLSSVLFR|NR|<mark>IIFIGQPINAQVAQR|VISQL</mark> VTLASIDDK|SDILMYLNCPGGSTYSVLAIYDCMSWIKPK|VGTVAFGVAASQGALLLAGGE K|GMR|YAMPNTR|VMIHQPQTGCGGHVEDVR|R|QVNEAIEAR|QKIDR|MYAAFTGQPLEK| VQQYTERDRFLSASEALEFGLIDGLLETEY

The two ClpP6 peptides worked well.