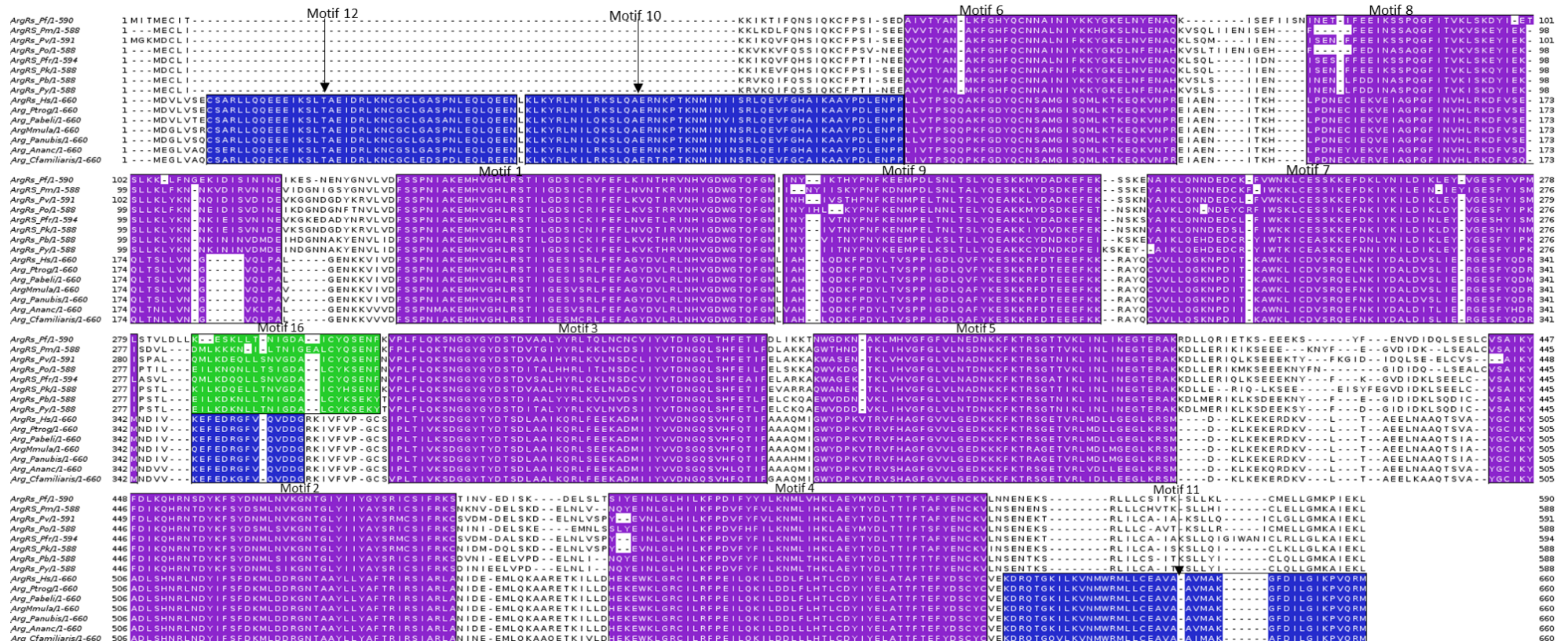
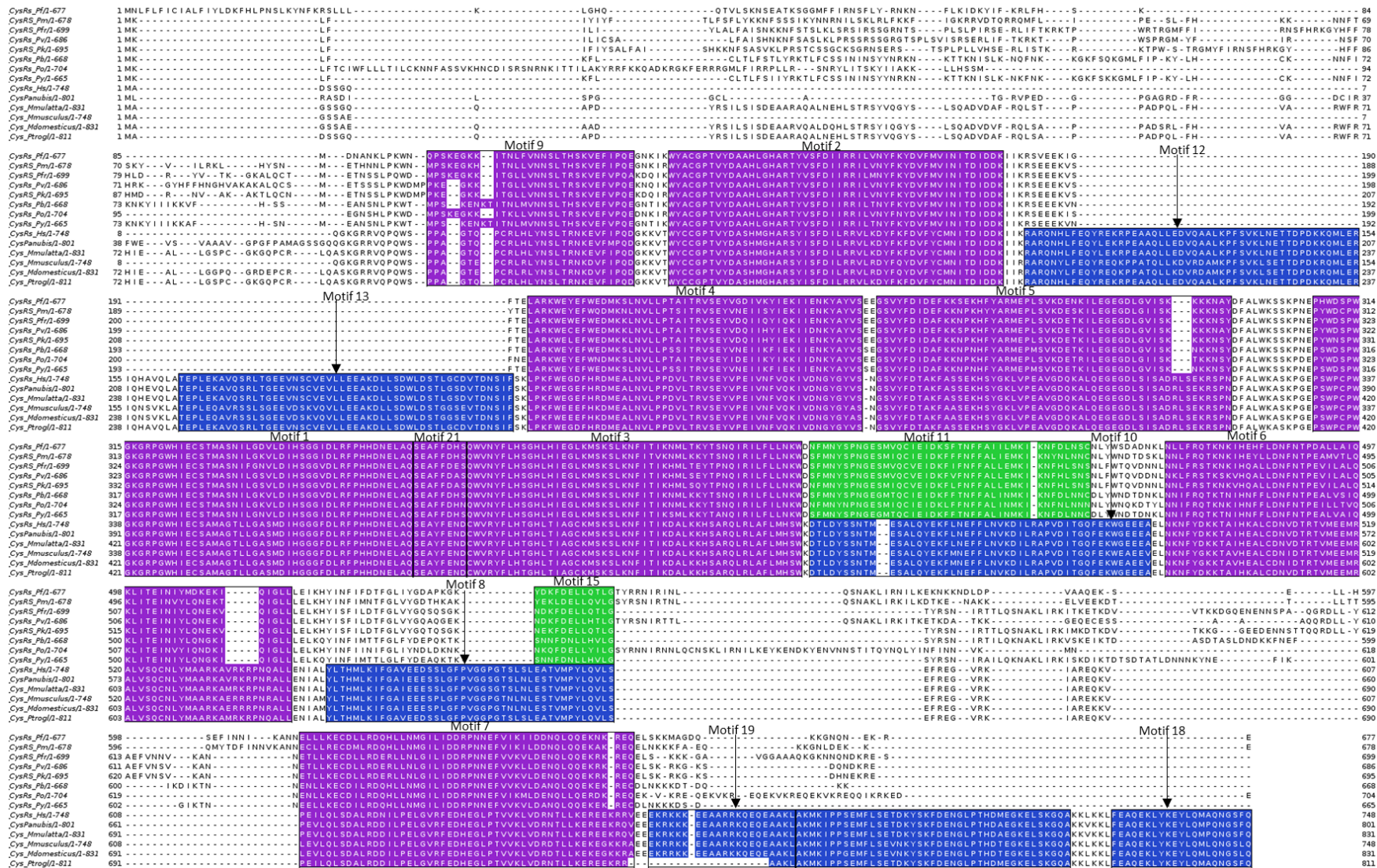


Additional file 4: Results on mapping of discovered motifs on alignments for the 20 aaRS families. Multiple sequence alignment was performed using TCOFFEE software with default parameters. A purple colour shows motifs conserved in all sequences, blue colour shows motifs only conserved in mammalian species and green shows motifs that are conserved in only the plasmodium species. Motif numbering is based on MEME results.



Additional file 4A: Mapping of motifs discovered in ArgRS family to the multiple sequence alignment



Additional file 4B: Mapping of motifs identified in CysRS family to the multiple sequence alignment

Additional file 4C: Mapping of motifs discovered in GlnRS family to the multiple sequence alignment

Glur5_cPfl-863	1MI	-----E-NTKV-NI	-----FYGNNKYPFICK	-----SILNVYKN	-----KNKKILEKT	-----ETTV-NVD	-----EN	-----IKFIK	-----DAN	-----VE	-----E	58
Glur5_Pfl-802	1MMSDO	-IIIIYGKKYFF-V-CR-TIL	-----NVHK-KL	-SOKENTAGSDHSA-DD	-----KVFKVDENVEEIKVEFV	-SKEKEGKNCI-NTT	-DK	-----IFAKN	-----LDL	-----IL	-----Q	59
Glur5_Pfl-789	1MV	-----D-CMKIYF	-----YGKKYFICRTVLNIIHRSK	-----TSENLED	-SITFIKDEKVEDIKVEFSEDSK	ILSSDKIFAKN	-----DK	-----IFAKT	-----LDVLLKTLN	-----	-----Q	92
Glur5_Pfl-802	1MIGDOK	-IIIIYGKKYFF-V-CR-TI	-----LNVHKILSO	-KENVEGRSNHFPNDK	KVFKVDENVEEIKVEFVNAKEEENG	-DKCI-NTT	-DK	-----IFAKT	-----LDE	-----IL	-----Q	91
Glur5_Pfl-787	1MVESDG	-IIIIYGKKYFF-V-CR-TIL	-----NAHK-SL	-ROKENPAGSD-HSAEE	-----KVFKVDENVEELKVEFV	-SKEEAO-GEKCI-NTT	-DK	-----IFAKS	-----LDO	-----IL	-----Q	91
Glur5_Pfl-804	1MIKD	-SKINIIYGKNYFF-L-CR-TVFNIYQNNIKKKTANNLNDK	-KIEECIVN-FINDKTIVE-DI	-----KIEFVKN	-----VA--DESR-KNVN-T-N	-----NS	-----VSSDKIFAINLDF	-----LL	-----V	-----KTNLY	103	
Glur5_Pfl-803	1MIKD	-SKINIIYGKNYFF-L-CR-TVFNIYQNNIKKKTANNLNDK	-KIEECIVN-FINDKTIVE-DI	-----KVEFVRNT	-----VSDQTRKNTNTNSV-TSS	-----DK	-----IFAIN	-----LDF	-----LL	-----K	99	
Glur5_Pfl-802	1MA	-----DEL-KI-NI	-----YGGKYPFL-CR	-----TVRNIYNN	-----IV-KKENL	-NSEC-LVG	-----EN	-----IKFVK	-----NEK	IQIEFVNDVSDVKF	IQSSDKIFAKSLD	83
Glur5_Hs2-1512	1MA	-----T-L-SL-TVN	-----SGDPLGA-LL	-----AVEHVKD	-----VSI	SVEEGK-ENIL-HVS	-----EN	-----VIFTD	-----VNS	-----IL	-----R	56
Glur5_Panabis2-1512	1MA	-----T-L-CL-TVN	-----SGDPLGA-LL	-----AVEHVKD	-----VSI	SVEEGK-ENIL-HVS	-----EN	-----VIFTD	-----VNS	-----IL	-----R	56
Glur5_Mmurinus2-1514	1MA	-----T-L-CL-TVN	-----SGDPLGA-LL	-----AVEHVKD	-----VSI	SVEEGK-ENIL-RIS	-----EN	-----VFTD	-----VNS	-----IL	-----R	56
Glur5_Cfamiliaris2-1509	1MA	-----T-L-CL-TVN	-----AGDPLGA-LL	-----AVEHVKD	-----VSI	SVEEGK-ENIL-HVS	-----EN	-----VFTD	-----VNS	-----IL	-----R	56
Glur5_Cayricha2-1512	1MA	-----T-L-CL-TVN	-----SGDPLGA-LL	-----AVEHVKD	-----VSI	SVEEGK-ENIL-RVS	-----EN	-----VFTD	-----VNS	-----IL	-----R	56
Glur5_Btaurus2-1511	1MA	-----T-L-CL-TVN	-----SGDPLGA-LL	-----AVEHVKD	-----VSI	SVEEGK-ENIL-HVS	-----EN	-----VFTD	-----VNS	-----IL	-----R	56
Glur5_Mcaroli2-1512	1MA	-----A-L-CL-TVN	-----AGNPPLEA-LL	-----AVEHVKD	-----VSV	SVEEGK-ENIL-RVS	-----EN	-----VFTD	-----VNS	-----IL	-----R	56
Glur5_Mmusculus2-1512	1MA	-----A-L-CL-TVN	-----AGNPPLEA-LL	-----AVEHVKD	-----VSI	SVEEGK-ENIL-RVS	-----ET	-----VFTD	-----VNS	-----IL	-----R	56
Motif 16												
Glur5_cPfl-863	59	LKIEFLTNAP	-----NNNNNNNNNNNNNNNNNNNNNN	INGKMMHCKGNKEEEKNLKSNNI	IEKVPYICSSDKIFAKILDHLKLTDLNI	SEKIDPNKNVNIYVQ	TOYEELWDFFRSRNVND	IV-FICEHNLKHLNLT	FVSEVYLTSD	IFIYFQMYKFHVS	STOYAKYSKMYKNINR	237
Glur5_Pfl-802	91	TK	-----LYHSF	-----CEN	-----NKKEDKQTDVYVQ	QOYDEWVDYFRNKQVQKEIO	-IICDHLNKLHLNLT	ICVDYLTSLD	VYVYEMKYNFVNA	NCYVVKYSKOFNRNVR	188	
Glur5_Pfl-789	93	YF	-----F	-----EGK	-----DQNKNI	INIVYVQ	SOYEEWVYFHHKDDTSSN	FLTRCDHINKHLNLT	ISSNFLTSLD	IYIYEMKYNFNL	SCYVVKYSKOFNRNINR	187
Glur5_Pfl-802	92	TK	-----LYHSF	-----CEN	-----NKKEDKQTDVYVQ	QOYDEWVDYFRNKHMQEIL	-IICDHLNKLHLNLT	ICDYLTLSD	VYVYEMKYNFVNA	NCYVVKYSKOFNRNINR	187	
Glur5_Pfl-787	90	TK	-----LYHSF	-----REN	-----NKKEDKQTDVYVQ	QOYDEWVDYFRNKQLEVL	-IICDHLNKLHLNLT	ICDYLTLSD	IYIYEMKYNFVNA	NCYVVKYSKOFNRNINR	187	
Glur5_Pfl-803	104	YFT	-----S	-----YR	-----ENIKNI	ITNVFQ	QOYNEWIDFLRNKDIEKNI	-IPICEHINKHLNLT	FLSHYLTSLD	IYIYEMKYNFSSN	ITNLKYPQVKNINR	198
Glur5_Pfl-803	100	TNL	-----Y	-----YF	-----ENIKNI	ITNVFQ	QOYNEWIDFLRNKDIEKNI	-IPICEHINKHLNLT	FLSHYLTSLD	IYIYEMKYNFSSN	ITNLKYPQVKNINR	198
Glur5_Pfl-802	84	YLL	-----KTNLYHFYDEN	-----K	-----KKKKILNVYI	QOYDWDIFFRYKDIQEN	IO-HVCEINHLNLT	ISSNFLTSLD	IYIYEMKYNFNL	TCYNIKYTKQYKNLR	184	
Glur5_Hs2-1512	57	Y	-----LARVA	-----T	-----TAAGLYGSLNMEHT	IDHWLEFSATKLSSCD	STSTINELNHCLSLR	TYLVGNLSL	ADLCVWATLK	-GNAA	MOEOLKOKKAPVHVKR	149
Glur5_Panabis2-1512	57	Y	-----LARVA	-----T	-----TAAGLYGSLNMEHT	IDHWLEFSATKLSSCD	STSTINELNHCLSLR	TYLVGNLSL	ADLCVWATLK	-GNAA	MOEOLKOKKAPVHVKR	149
Glur5_Mmurinus2-1514	57	Y	-----LARVA	-----T	-----TAAGLYGSLNMEHT	IDHWLEFSATKLSSCD	STSTINELNHCLSLR	TYLVGNLSL	ADLCVWATLK	-GNAA	MOEOLKOKKAPVHVKR	149
Glur5_Cfamiliaris2-1509	57	Y	-----LARVA	-----T	-----TAAGLYGSLNMEHT	IDHWLEFSATKLSSCD	STSTINELNHCLSLR	TYLVGNLSL	ADLCVWATLK	-GNAA	MOEOLKOKKAPVHVKR	149
Glur5_Cayricha2-1512	57	Y	-----LARVA	-----T	-----TAAGLYGSLNMEHT	IDHWLEFSATKLSSCD	STSTINELNHCLSLR	TYLVGNLSL	ADLCVWATLK	-GNAA	MOEOLKOKKAPVHVKR	149
Glur5_Btaurus2-1511	57	Y	-----LARVA	-----T	-----TAAGLYGSLNMEHT	IDHWLEFSATKLSSCD	STSTINELNHCLSLR	TYLVGNLSL	ADLCVWATLK	-GNAA	MOEOLKOKKAPVHVKR	149
Glur5_Mcaroli2-1512	57	Y	-----LARVA	-----T	-----TAAGLYGSLNMEHT	IDHWLEFSATKLSSCD	STSTINELNHCLSLR	TYLVGNLSL	ADLCVWATLK	-GNAA	MOEOLKOKKAPVHVKR	149
Glur5_Mmusculus2-1512	57	Y	-----LARVA	-----T	-----TAAGLYGSLNMEHT	IDHWLEFSATKLSSCD	STSTINELNHCLSLR	TYLVGNLSL	ADLCVWATLK	-GNAA	MOEOLKOKKAPVHVKR	149
Motif 3												
Glur5_cPfl-863	238	WHKLID	-----SLI	-----YF	-----EDAE	-----WKKN	-LKCLIDCNKMF	-----D	DRK-DNLKNNKSNMNTD	DDVCKKKTVT	TSY	533
Glur5_Pfl-802	189	WFKLIN	-----CLV	-----NY	-----PDAQLTLMKKE	-IEI	IQ-V	DIKDANTNMQLKQIV	S	-----TSY	-----GGKLN	292
Glur5_Pfl-789	188	WFKLIN	-----HLLCEDNELNMYLKKEA	-----E	-----NOVD	-----ITKD	-----TNT	-----MOLK	-----QKIV	S	-----TSY	291
Glur5_Pfl-802	190	WFKLIN	-----CLV	-----NYPD	-----EELS	-----TMVKE	-----LEKNO	-----VNI	TKDANSMLLKQKTV	C	-----TSY	293
Glur5_Pfl-787	188	WFKLIN	-----CLV	-----NYPD	-----EELS	-----TMVKE	-----LEKNO	-----VNI	TKDANTMOLLKQIV	S	-----TSY	291
Glur5_Pfl-804	199	WFKLIN	-----HLLDHI	THDAELVONLKEIKKIFNT	-----TNYGSKSS	-----NNE	-----K	-----SSTP	-----NKSHTK	-----V	-----GGTSL	318
Glur5_Pfl-803	199	WFKLIN	-----ALL	-----H	DHVATDAELI	QNLVKEKIFNTAG	-----SGKRLNGNNGK	-----VSAHS	-----DHTN	T	-----KGGATSL	317
Glur5_Pfl-802	185	WFKLIT	-----SL	-----YKND	DELNMYLQ	-----KEALMNKVN	-----ITKD	-DAN-S	-----VHMLK	-----QKAT	N	289
Glur5_Hs2-1512	150	WFGFLE	-----AQ	-----AF	-----QSVG	-----TKWD	-----VSTTK	-----ARVAP	-----EKKQ	D	-----V	240
Glur5_Panabis2-1512	150	WFGFLE	-----AQ	-----AF	-----QSVG	-----TKWD	-----VSTTK	-----ARVAP	-----EKKQ	D	-----V	240
Glur5_Mmurinus2-1514	150	WFGFLE	-----AQ	-----AF	-----QSVG	-----TKWD	-----VSTTK	-----ARVAP	-----EKKQ	D	-----V	240
Glur5_Cfamiliaris2-1509	150	WFGFLE	-----AQ	-----AF	-----QSVG	-----TKWD	-----VSTTK	-----ARVAP	-----EKKQ	D	-----V	240
Glur5_Cayricha2-1512	150	WFGFLE	-----AQ	-----AF	-----QSVG	-----TKWD	-----VSTTK	-----ARVAP	-----EKKQ	D	-----V	240
Glur5_Btaurus2-1511	150	WFGFLE	-----AQ	-----AF	-----QSVG	-----TKWD	-----VSTTK	-----ARVAP	-----EKKQ	D	-----V	240
Glur5_Mcaroli2-1512	150	WFGFLE	-----AQ	-----AF	-----QSVG	-----TKWD	-----VSGNR	-----ATVAP	-----DKQD	D	-----V	240
Glur5_Mmusculus2-1512	150	WFGFLE	-----AQ	-----AF	-----QSVG	-----TKWD	-----VSGNR	-----ATVAP	-----DKQD	D	-----V	240
Motif 5												
Glur5_cPfl-863	354	VLEDIKYKES	IIEDLENL	SLKYEKISYSDSDHFDLEKYCIDMIKMKHAYADD	TGVEDMNRORGEIES	VRNRNS	VEONLQLEFEMRKGTE	IGOKNCIRAKIDMS	SNKNCMRDPVLYRC	VDTPHHRHGF	KFKCYPTDYFACPI	536
Glur5_Pfl-802	293	VALEDIKYETS	IMEDLENL	SLKYEKISYSDSDHFDLEKYCIDMIKMKHAYADD	TGVEDMNRORGEIES	VRNRNS	VEONLQLEFEMRKGTE	IGOKNCIRAKIDMS	SNKNCMRDPVLYRC	VDTPHHRHGF	KFKCYPTDYFACPI	475
Glur5_Pfl-789	292	VALEDIKYETS	IMEDLENL	SLKYEKISYSDSDHFDLEKYCIDMIKMKHAYADD	TGVEDMNRORGEIES	VRNRNS	VEONLQLEFEMRKGTE	IGOKNCIRAKIDMS	SNKNCMRDPVLYRC	VDTPHHRHGF	KFKCYPTDYFACPI	474
Glur5_Pfl-802	292	VALEDIKYETS	IMEDLENL	SLKYEKISYSDSDHFDLEKYCIDMIKMKHAYADD	TGVEDMNRORGEIES	VRNRNS	VEONLQLEFEMRKGTE	IGOKNCIRAKIDMS	SNKNCMRDPVLYRC	VDTPHHRHGF	KFKCYPTDYFACPI	476
Glur5_Pfl-787	294	VALEDIKYETS	IMEDLENL	SLKYEKISYSDSDHFDLEKYCIDMIKMKHAYADD	TGVEDMNRORGEIES	VRNRNS	VEONLQLEFEMRKGTE	IGOKNCIRAKIDMS	SNKNCMRDPVLYRC	VDTPHHRHGF	KFKCYPTDYFACPI	474
Glur5_Pfl-804	319	VLEDVKEFNS	IIEDLETL	SLKYEKISYSDSDHFDLEKYCIDMIKMKHAYADD	TGVEDMNRORGEIES	VRNRNS	VEONLQLEFEMRKGTE	IGOKNCIRAKIDMS	SNKNCMRDPVLYRC	VDTPHHRHGF	KFKCYPTDYFACPI	501
Glur5_Pfl-803	318	VLEDVKEFNS	IIEDLETL	SLKYEKISYSDSDHFDLEKYCIDMIKMKHAYADD	TGVEDMNRORGEIES	VRNRNS	VEONLQLEFEMRKGTE	IGOKNCIRAKIDMS	SNKNCMRDPVLYRC	VDTPHHRHGF	KFKCYPTDYFACPI	500
Glur5_Pfl-802	290	VALEDI	IYETS	IIIDLEKL	SLKYEKISYSDSDHFDLEKYCIDMIKMKHAYADD	TGVEDMNRORGEIES	VRNRNS	VEONLQLEFEMRKGTE	IGOKNCIRAKIDMS	SNKNCMRDPVLYRC	VDTPHHRHGF	472
Glur5_Hs2-1512	241	EKEKEDFEKVL	IEDVAML	IKKPDFTYSDHFEIMKYAEKLIQEGKAYVDD	TPAEOMKAEERQIES	KHRKRN	IEKNLQWMEEMKKGSG	FGOSCCLR	AKIDMS	SNNGCMRDP	TYLYRC	422
Glur5_Panabis2-1512	241	EKEKEDFEKVL	IEDVAML	IKKPDFTYSDHFEIMKYAEKLIQEGKAYVDD	TPAEOMKAEERQIES	KHRKRN	IEKNLQWMEEMKKGSG	FGOSCCLR	AKIDMS	SNNGCMRDP	TYLYRC	422
Glur5_Mmurinus2-1514	241	EKEKEDFEKVL	IEDVAML	IKKPDFTYSDHFEIMKYAEKLIQEGKAYVDD	TPAEOMKAEERQIES	KHRKRN	IEKNLQWMEEMKKGSG	FGOSCCLR	AKIDMS	SNNGCMRDP	TYLYRC	422
Glur5_Cfamiliaris2-1509	241	EKEKEDFEKVL	IEDVAML	IKKPDFTYSDHFEIMKYAEKLIQEGKAYVDD	TPAEOMKAEERQIES	KHRKRN	IEKNLQWMEEMKKGSG	FGOSCCLR	AKIDMS	SNNGCMRDP	TYLYRC	422
Glur5_Cayricha2-1512	241	EKEKEDFEKVL	IEDVAML	IKKPDFTYSDHFEIMKYAEKLIQEGKAYVDD	TPAEOMKAEERQIES	KHRKRN	IEKNLQWMEEMKKGSG	FGOSCCLR	AKIDMS	SNNGCMRDP	TYLYRC	422
Glur5_Btaurus2-1511	241	EKEKEDFEKVL	IEDVAML	IKKPDFTYSDHFEIMKYAEKLIQEGKAYVDD	TPAEOMKAEERQIES	KHRKRN	IEKNLQWMEEMKKGSG	FGOSCCLR	AKIDMS	SNNGCMRDP	TYLYRC	422
Glur5_Mcaroli2-1512	241	EKEKEDFEKVL	IEDVAML	IKKPDFTYSDHFEIMKYAEKLIQEGKAYVDD	TPAEOMKAEERQIES	KHRKRN	IEKNLQWMEEMKKGSG	FGOSCCLR	AKIDMS	SNNGCMRDP	TYLYRC	422
Glur5_Mmusculus2-1512	241	EKEKEDFEKVL	IEDVAML	IKKPDFTYSDHFEIMKYAEKLIQEGKAYVDD	TPAEOMKAEERQIES	KHRKRN	IEKNLQWMEEMKKGSG	FGOSCCLR	AKIDMS	SNNGCMRDP	TYLYRC	422
Motif 10												
Glur5_cPfl-863	537	SRLAFVKTVM	SKRKLKWF	VEINIVD	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	705
Glur5_Pfl-802	476	SRLAFVKTVM	SKRKLKWF	VEINIVD	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	647
Glur5_Pfl-789	475	SRLAFVKTVM	SKRKLKWF	VEINIVD	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	647
Glur5_Pfl-802	477	SRLAFVKTVM	SKRKLKWF	VEINIVD	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	649
Glur5_Pfl-787	475	SRLAFVKTVM	SKRKLKWF	VEINIVD	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	646
Glur5_Pfl-804	502	SRLAFVKTVM	SKRKLKWF	VEINIVD	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	674
Glur5_Pfl-803	501	SRLAFVKTVM	SKRKLKWF	VEINIVD	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	673
Glur5_Pfl-802	473	YSRLNLNNTVL	SKRKLKWF	VNIGSL	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	640
Glur5_Hs2-1512	423	YSRLNLNNTVL	SKRKLKWF	VNIGSL	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	590
Glur5_Panabis2-1512	423	YSRLNLNNTVL	SKRKLKWF	VNIGSL	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	590
Glur5_Mmurinus2-1514	423	YSRLNLNNTVL	SKRKLKWF	VNIGSL	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	590
Glur5_Cfamiliaris2-1509	423	YSRLNLNNTVL	SKRKLKWF	VNIGSL	GWDDPRPMTIKGILRRGLTKEALFQ	ILEOGPSKAGNL	MOWDKLWSIN	KQID	IPNIPRFAA	VDKKGAVILK	LDLT	590
Glur5_Cayricha2-15-												

Additional file 4D: Mapping of motifs discovered in GluRS family to the multiple sequence alignment

[illegible]

Additional file 4F: Mapping of motifs discovered in LeuRS family to the multiple sequence alignment

Additional file 4F: Mapping of motifs discovered in LeuRS family to the multiple sequence alignment

	Motif9	Motif7	Motif4	Motif6	Motif14	Motif2	Motif12											
MetRs_Pf1-889	408 KRNEILQRLKEPLADLSCSRTKFTWG	QVFN	NDP	--KHVMYVWMDALINYYSSCF	ID--EGKKKY	WP--A--NV	HIIGKDIWFHFTVIFPTILLSA	--NPLSKSVFSGHGFVLAADGKKMSKSLGN	VINPLEIIDKYGDAFRFHYV	KETKRGCDTRFDVNDL	VD	MCNSDLADTIGNLVQRTLSL	581					
MetRs_Pm1-880	402 KRNEILQRLKEPLEDLSCSRTKFTWG	IAPV	SDS	--KHVMYVWMDALLNYYSSCF	MN--BEKKVY	WP--A--DV	HIIGKDIWAFHFTVIFPTILLMSA	--NIELSKSVFSGHGFVLAADGKKMSKSLGN	VINPVEIIDTYSGDAFRFHYV	KESKRGFDLRFIDNLD	ID	MCNSDLADTIGNLVQRTLSL	575					
MetRs_Po1-914	436 KKNEILQRLKEPLADLSCSRTKFTSWG	SVP	SDS	--KHVMYVWMDALLNYYSSCF	LD--NDKMKY	WP--A--DV	HIIGKDIWVHFTVIFPTILLMSV	--NLEIPKCVFSGHGFVLAADGKKMSKSLGN	VINPQEIIDSYGDAFRFHYV	KETKRGDFMRFDIDNM	VD	MCNSDLADTIGNLVQRTLSL	609					
MetRs_Pv1-885	405 KRNEILQRLKEPLADLSCSRTKFTSWG	PVP	SDN	--KHVMYVWMDALINYYSSCF	IV--EGKEDY	WP--A--DV	HLIGKDIWVHFTVIFPTILLMSV	--NVDLPKSVFCHGFVLAADGKKMSKSLGN	VINPFEIIEYSGDAFRFHYV	KETKRGDFMRFDIDNL	VD	MCNSDLADTIGNLVQRTLSL	578					
MetRs_Pv1-886	405 KRNEILQRLKEPLADLSCSRTKFTSWG	PVP	SDH	--KHVMYVWMDALINYYSSCF	IV--EGKEDY	WP--A--DV	HLIGKDIWVHFTVIFPTILLMSV	--NVALPKSVFCHGFVLAADGKKMSKSLGN	VINPFEIIEYSGDAFRFHYV	KETKRGDFMRFDIDNL	VD	MCNSDLADTIGNLVQRTLSL	578					
MetRs_Pv1-885	405 KRNEILQRLKEPLADLSCSRTKFTSWG	PVP	SDN	--KHVMYVWMDALINYYSSCF	IV--EGKEDY	WP--A--DV	HLIGKDIWVHFTVIFPTILLMSV	--NVDLPKSVFCHGFVLAADGKKMSKSLGN	VINPFEIIEYSGDAFRFHYV	KETKRGDFMRFDIDNL	VD	MCNSDLADTIGNLVQRTLSL	578					
MetRs_Pv1-898	419 NRNEILQRLKEPLEDLSCSRTKFTWG	IPV	NDP	--KHVMYVWMDALINYYSSCF	INNDEKIKY	WP--A--DI	HIIGKDIWVHFAVIFPTILLMSI	--NIELPKCVFSGHGFVLAADGKKMSKSLGN	VIDPKEIIEYSGDAFRFHYV	KESKRGVDMRFDVNDL	VD	MCNSDLADTIGNLVQRTLSL	593					
MetRs_Pv1-898	419 NRNEILQRLKEPLEDLSCSRTKFTWG	IPV	NDP	--KHVMYVWMDALINYYSSCF	INNDEKIKY	WP--A--DI	HIIGKDIWVHFAVIFPTILLMSI	--NIELPKCVFSGHGFVLAADGKKMSKSLGN	VIDPKEIIEYSGDAFRFHYV	KESKRGVDMRFDVNDL	VD	MCNSDLADTIGNLVQRTLSL	593					
MetRs_Hs1-900	478 AQFIIRSWLRDGLKPCRCITRDCLKWGTVPVLEGE	FED	KDVFYVWF	DATIGYLSITANYTDQWERWKNPEQVDLYQFMAKDNVFPFHG	IVFPC	SALGAEDNYTLVSHL	IATEYLNIEDGK	F--SKSRGV	VFGDMAQDTGIPAD	IWRFFYLLYIRPEGQDSAF	SWTDLL	LLKNNSELLNNLGNFINRAGMF	660					
Met_Mmulatta1-900	478 ARFIIRSWLRDGLKPCRCITRDCLKWGTVPVLEGE	FED	KDVFYVWF	DATIGYLSITANYTDQWERWKNPEQVDLYQFMAKDNVFPFHG	IVFPC	SALGAEDNYTLVSHL	IATEYLNIEDGK	F--SKSRGV	VFGDMAQDTGIPAD	IWRFFYLLYIRPEGQDSAF	SWTDLL	LLKNNSELLNNLGNFINRAGMF	660					
Met_Panubis1-908	478 ARFIIRSWLRDGLKPCRCITRDCLKWGTVPVLEGE	FED	KDVFYVWF	DATIGYLSITANYTDQWERWKNPEQVDLYQFMAKDNVFPFHG	IVFPC	SALGAEDNYTLVSHL	IATEYLNIEDGK	F--SKSRGV	VFGDMAQDTGIPAD	IWRFFYLLYIRPEGQDSAF	SWTDLL	LLKNNSELLNNLGNFINRAGMF	660					
Met_Mmurinus1-900	478 ARFIIRSWLRDGLKPCRCITRDCLKWGTVPVLEGE	FED	KDVFYVWF	DATIGYLSITANYTDQWERWKNPEQVDLYQFMAKDNVFPFHG	IVFPC	SALGAEDNYTLVSHL	IATEYLNIEDGK	F--SKSRGV	VFGDMAQDTGIPAD	IWRFFYLLYIRPEGQDSAF	SWADML	LLKNNSELLNNLGNFINRAGMF	660					
Met_Mfascic1-869	447 ARFIIRSWLRDGLKPCRCITRDCLKWGTVPVLEGE	FED	KDVFYVWF	DATIGYLSITANYTDQWERWKNPEQVDLYQFMAKDNVFPFHG	IVFPC	SALGAEDNYTLVSHL	IATEYLNIEDGK	F--SKSRGV	VFGDMAQDTGIPAD	IWRFFYLLYIRPEGQDSAF	SWTDLL	LLKNNSELLNNLGNFINRAGMF	629					
Met_Btaurus1-898	478 ARFIIRSWLRDGLKPCRCITRDCLKWGTVPVLEGE	FED	KDVFYVWF	DATIGYLSITANYTDQWERWKNPEQVDLYQFMAKDNVFPFHG	IVFPC	SALGAEDNYTLVSHL	IATEYLNIEDGK	F--SKSRGV	VFGDMAQDTGIPAD	IWRFFYLLYIRPEGQDSAF	SWTDML	LLKNNSELLNNLGNFINRAGMF	660					
Met_Caynichta1-900	478 ARFIIRSWLRDGLKPCRCITRDCLKWGTVPVLEGE	FED	KDVFYVWF	DATIGYLSITANYTDQWERWKNPEQVDLYQFMAKDNVFPFHG	IVFPC	SALGAEDNYTLVSHL	IATEYLNIEDGK	F--SKSRGV	VFGDMAQDTGIPAD	IWRFFYLLYIRPEGQDSAF	SWMDML	LLKNNSELLNNLGNFINRAGMF	660					
Met_Mmusculus1-902	480 ARFIIRSWLRDGLKPCRCITRDCLKWGTVPVLEGE	FED	KDVFYVWF	DATIGYYSITANYTDQWERWKNPEQVDLYQFMAKDNVFPFHGLVFP	PCSVLGAEDNYTLVKHI	IATEYLNIEDGK	F--SKSRGI	VFGDMAQDTGIPAD	IWRFFYLLYIRPEGQDSAF	SWTDLL	LLKNNSELLNNLGNFINRAGMF	662						
MetRs_Pf1-889	582 COLCNESKIPAFYPEYIHLPIIDLHF	INRVEYH	--MKTFCVQKCEKSVQVCKDLNKF	LDTMAPWKYKDD	LL--NKKL	HIIRIMLEAIYFIAHYLDI	IVPSIASQIFKKLNAPKKC	IVDLNPWLNNDQ	EG	VINNDNIILFKKFEVESAK	-----	-----	728					
MetRs_Pm1-880	576 COLSNDKIPPLFKEYNIDLPPSLLFY	INRVEYH	--MKSFCVQKCEKTVNVCKDLNKF	LDTAPWKYKDE	DK--NKKL	HIIRIMLEAIYFIAHYLDI	IVPSIASQIFKKLNAPKKC	IVDLNPWLNNDQ	EG	VINNDNIILFKKFEVESAK	-----	-----	722					
MetRs_Po1-914	610 COLSNEAKIPPLFEEYIDLPPCLLY	INRVEYH	--MOYFCVQKCEKTVNVCKDLNKF	LDTAPWKYKEEN	NK--NKKL	HIIRIMLEAIYFIAHYLDI	IVPSIASQIFKKLNAPKKC	IVDLNPWLNNDQ	EG	VINNDNIILFKKFEVESAK	-----	-----	756					
MetRs_Pv1-885	579 COLSNEKIPPLFEEYIDLPPSLLHF	INRVEFH	--MKTFCVQVCCERTVNVCKDLNKF	LDTAPWKYKNE	EG--NKKL	HIIRIMLEAIYFIAHYLDI	IVPSIASQIFKKLNAPKKC	IVDLNPWLNNDQ	EG	VINNDNIILFKKFEVESAK	-----	-----	725					
MetRs_Pv1-886	579 COLSNEKIPPLFEEYIDLPPSLLHF	INRVEFH	--MKTFCVQVCCERTVNVCKDLNKF	LDTAPWKYKNE	EG--NKKL	HIIRIMLEAIYFIAHYLDI	IVPSIASQIFKKLNAPKKC	IVDLNPWLNNDQ	EG	VINNDNIILFKKFEVESAK	-----	-----	725					
MetRs_Pv1-885	579 COLSNEKIPPLFEEYIDLPPSLLHF	INRVEFH	--MKTFCVQVCCERTVNVCKDLNKF	LDTAPWKYKNE	EG--NKKL	HIIRIMLEAIYFIAHYLDI	IVPSIASQIFKKLNAPKKC	IVDLNPWLNNDQ	EG	VINNDNIILFKKFEVESAK	-----	-----	725					
MetRs_Pv1-885	579 COLSNEKIPPLFEEYIDLPPSLLHF	INRVEFH	--MKTFCVQVCCERTVNVCKDLNKF	LDTAPWKYKNE	EG--NKKL	HIIRIMLEAIYFIAHYLDI	IVPSIASQIFKKLNAPKKC	IVDLNPWLNNDQ	EG	VINNDNIILFKKFEVESAK	-----	-----	725					
MetRs_Pv1-898	594 CMLSNNAKIPPLYDDIDLPEFLCMHF	INRIEYH	--MKNFHHKCEKTVNVCKDLNKF	LDTAPWKYKEESEK	NK--NKKL	HIIRIMLEAIYFIAHYLDI	IVPMIASQIFNQLNTPKKK	IVDLNPWLNNDQ	EG	VINNDNIILFKKFEVESAK	-----	-----	741					
MetRs_Pv1-898	594 CMLSNNAKIPPLYDDIDLPEFLCMHF	INRIEYH	--MKNFHHKCEKTVNVCKDLNKF	LDTAPWKYKEESEK	NK--NKKL	HIIRIMLEAIYFIAHYLDI	IVPMIASQIFNQLNTPKKK	IVDLNPWLNNDQ	EG	VINNDNIILFKKFEVESAK	-----	-----	741					
MetRs_Hs1-900	661 VSKFFGGYVPEMVLTPDDO	--RLLAHVTELEQHYHOLLEKVR	RDALRS	ILTI	SRHGNQYIQVNEPWKRIKGEADRRORAGT	VTGLAVN	IAALLSVMLQPYMPTVSAT	IOAQQLPPACSI	ILLTNFLCTLPAGH	IGTVSPLFQKLENDQ	IESLRORFGGGQ	-----	AKTSPKPAVVE	835				
Met_Mmulatta1-900	661 VSKFFGGYVPEMVLTPDDO	--RLLAHVTELEQHYHOLLEKVR	RDALRS	ILTI	SRHGNQYIQVNEPWKRIKGEADRRORAGT	VTGLAVN	IAALLSVMLQPYMPTVSAT	IOAQQLPPACSI	ILLTNFLCTLPAGH	IGTVSPLFQKLENDQ	IESLRORFGGGQ	-----	AKTSPKPTVIE	835				
Met_Panubis1-908	661 VSKFFGGYVPEMVLTPDDO	--RLLAHVTELEQHYHOLLEKVR	RDALRS	ILTI	SRHGNQYIQVNEPWKRIKGEADRRORAGT	VTGLAVN	IAALLSVMLQPYMPTVSAT	IOAQQLPPACSI	ILLTNFLCTLPAGH	IGTVSPLFQKLENDQ	IESLRORFGGGQ	LEETFDL	AKTSPKPTVIE	843				
Met_Mmurinus1-900	661 VSKFFGGYVPEMVLTPDDO	--RLLAHVTELEQHYHOLLEKVR	RDALRS	ILTI	SRHGNQYIQVNEPWKRIKGEADRRORAGT	VTGLAVN	IAALLSVMLQPYMPTVSAT	IOAQQLPPACSI	ILLTNFLCTLPAGH	IGTVSPLFQKLENDQ	IESLRORFGGGQ	-----	AKVSPKPEVVE	835				
Met_Mfascic1-869	630 VSKFFGGYVPEMVLTPDDO	--RLLAHVTELEQHYHOLLEKVR	RDALRS	ILTI	SRHGNQYIQVNEPWKRIKGEADRRORAGT	VTGLAVN	IAALLSVMLQPYMPTVSAT	IOAQQLPPACSI	ILLTNFLCTLPAGH	IGTVSPLFQKLENDQ	IESLRORFGGGQ	-----	AKTSPKPTVIE	804				
Met_Btaurus1-898	661 VSKFFGGYVPEMVLTPDDO	--RLLAHVTELEQHYHOLLEKVR	RDALRS	ILTI	SRHGNQYIQVNEPWKRIKGEADRRORAGT	VTGLAVN	IAALLSVMLQPYMPTVSAT	IOAQQLPPACSI	ILLTNFLCTLPAGH	IGTVSPLFQKLENDQ	IESLRORFGGGQ	-----	AKASPKT	833				
Met_Caynichta1-900	661 VSKFFGGYVPEMVLTPDDO	--RLLAHVTELEQHYHOLLEKVR	RDALRS	ILTI	SRHGNQYIQVNEPWKRIKGEADRRORAGT	VTGLAVN	IAALLSVMLQPYMPTVSAT	IOAQQLPPACSI	ILLTNFLCTLPAGH	IGTVSPLFQKLENDQ	ENHLRORFGGGQ	-----	AKAAPKPAE	835				
Met_Mmusculus1-902	663 VSKFFGGYVPEMVLTPDDO	--RLLAHVTELEQHYHOLLEKVR	RDALRS	ILTI	SRHGNQYIQVNEPWKRIKGEADRRORAGT	VTGLAVN	IAALLSVMLQPYMPTVSAT	IOAQQLPPACSI	ILLTNFLCTLPAGH	IGTVSPLFQKLENDQ	ENHLRORFGGGQ	-----	AKGSPKPAVVE	837				
MetRs_Pf1-889	729 ----KLQKVIIRVCKIVNIINTES	-SQNT	TICE	IEVD	----NDKITA	VLVY	PNYNNNNKLVNLLTALINIKP	ITINNITVNAI	IPHVNKEIFTFPIEER	PTGTL	IOGKNYKTLVKQORDNLT	TKKEVNSLDS	IINNCKFEKNI	IPLVFAS	HED	KQVYHSTODNGSLTFF	889	
MetRs_Pm1-880	723 ----KIQKVIIRVCKITNIINKNEE	-LLKA	TVCE	IEVD	----HEKYPA	VLVY	PNYNNNNKLVNLLTALINIKP	ITINNITVNAI	IPHVNKEIFTFPIEER	PTGTL	IOGKNYKTLVKQORDNLT	TKKEVNSLDS	IINNCKFEKNI	IPLVFAS	HED	KQVYHSTODNGSLTFF	880	
MetRs_Po1-914	757 ----KIQKVIIRVCKIVNIINKNEE	-LQKV	TICE	IDVD	----NEKVS	AVLVY	PNYNNNNKLVNLLTALINIKP	ITINNITVNAI	IPHVNKEIFTFPIEER	PTGTL	IOGKNYKTLVKQORDNLT	TKKEVNSLDS	IINNCKFEKNI	IPLVFAS	HED	KQVYHSTODNGSLTFF	914	
MetRs_Pv1-885	726 ----KIQKVIIRVCKIVNIINKNEE	-NOKA	TIFE	IEVD	----DNEKH	LA	VLVY	PNYNNNNKLVNLLTALINIKP	ITINNITVNAI	IPHVNKEIFTFPIEER	PTGTL	IOGKNYKTLVKQORDNLT	TKKEVNSLDS	IINNCKFEKNI	IPLVFAS	HED	KQVYHSTODNGSLTFF	885
MetRs_Pv1-886	726 ----KIQKVIIRVCKIVNIINKNEE	-NOKA	TIFE	IEVD	----DNEKH	LA	VLVY	PNYNNNNKLVNLLTALINIKP	ITINNITVNAI	IPHVNKEIFTFPIEER	PTGTL	IOGKNYKTLVKQORDNLT	TKKEVNSLDS	IINNCKFEKNI	IPLVFAS	HED	KQVYHSTODNGSLTFF	886
MetRs_Pv1-885	726 ----KIQKVIIRVCKIVNIINKNEE	-NOKA	TIFE	IEVD	----DNEKH	LA	VLVY	PNYNNNNKLVNLLTALINIKP	ITINNITVNAI	IPHVNKEIFTFPIEER	PTGTL	IOGKNYKTLVKQORDNLT	TKKEVNSLDS	IINNCKFEKNI	IPLVFAS	HED	KQVYHSTODNGSLTFF	886
MetRs_Pv1-885	726 ----KIQKVIIRVCKIVNIINKNEE	-NOKA	TIFE	IEVD	----DNEKH	LA	VLVY	PNYNNNNKLVNLLTALINIKP	ITINNITVNAI	IPHVNKEIFTFPIEER	PTGTL	IOGKNYKTLVKQORDNLT	TKKEVNSLDS	IINNCKFEKNI	IPLVFAS	HED	KQVYHSTODNGSLTFF	886
MetRs_Pv1-885	742 ----KIQKVIIRVCKIVNIINKNEE	-LQKV	TICE	IEVD	----DNEKH	LA	VLVY	PNYNNNNKLVNLLTALINIKP	ITINNITVNAI	IPHVNKEIFTFPIEER	PTGTL	IOGKNYKTLVKQORDNLT	TKKEVNSLDS	IINNCKFEKNI	IPLVFAS	HED	KQVYHSTODNGSLTFF	889
MetRs_Pv1-885	742 ----KIQKVIIRVCKIVNIINKNEE	-LQKV	TICE	IEVD	----DNEKH	LA	VLVY	PNYNNNNKLVNLLTALINIKP	ITINNITVNAI	IPHVNKEIFTFPIEER	PTGTL	IOGKNYKTLVKQORDNLT	TKKEVNSLDS	IINNCKFEKNI	IPLVFAS	HED	KQVYHSTODNGSLTFF	889
MetRs_Pv1-885	742 ----KIQKVIIRVCKIVNIINKNEE	-LQKV	TICE	IEVD	----DNEKH	LA	VLVY	PNYNNNNKLVNLLTALINIKP	ITINNITVNAI	IPHVNKEIFTFPIEER	PTGTL	IOGKNYKTLVKQORDNLT	TKKEVNSLDS	IINNCKFEKNI	IPLVFAS	HED	KQVYHSTODNGSLTFF	889
MetRs_Hs1-900	836 TAKPQIQALMDEVTKQGNIVRELK	-AOKA	-DKNEVAAE	VAKL	LDLKKQLAY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	900	
Met_Mmulatta1-900	836 TAGPQIQALMDEVTKQGNIVRELK	-AOKA	-DKNEVAAE	VAKL	LDLKKQLAY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	900	
Met_Panubis1-908	844 TAGPQIQALMDEVTKQGNIVRELK	-AOKA	-DKNEVAAE	VAKL	LDLKKQLAY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	908	
Met_Mmurinus1-900	836 TAGPQIQALMDEVTKQGNIVRELK	-AOKA	-DKNEVAAE	VAKL	LDLKKQLAY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	900	
Met_Mfascic1-869	805 TAGPQIQALMDEVTKQGNIVRELK	-AOKA	-DKNEVAAE	VAKL	LDLKKQLAY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	869	
Met_Btaurus1-898	834 TAGPQIQALMDEVTKQGNIVRELK	-AOKA	-DKNEVAAE	VAKL	LDLKKQLAY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	898	
Met_Caynichta1-900	836 TARPPQIQALMDEVTKQGNIVRELK	-AOKA	-DKNEVAAE	VAKL	LDLKKQLAY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	900	
Met_Mmusculus1-902	838 AAGQHIQTLTDEVTKQGNIVRELK	-AOKA	-DKNEVAAE	VAKL	LDLKKQLAY	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	902	

Additional file 4G: Mapping of motifs discovered in MetRS family to the multiple sequence alignment

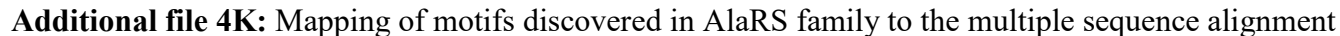
Additional file 4J: Mapping of motifs discovered in ValRS family to the multiple sequence alignment

Additional file 4J: Mapping of motifs discovered in ValRS family to the multiple sequence alignment

	Motif 4	Motif 21	Motif 31	Motif 3	Motif 40	
ValRS_Pf1-1090	1459	1459	1459	1459	1459	618
ValRS_Pm2-1089	447	447	447	447	447	620
ValRS_Py1-1077	463	463	463	463	463	613
ValRS_Pf2-1046	442	442	442	442	442	581
ValRS_Pf2-1071	447	447	447	447	447	507
ValRS_Pf2-1070	447	447	447	447	447	606
ValRS_Pf2-1077	461	461	461	461	461	613
ValRS_Pf2-1138	523	523	523	523	523	671
ValRS_Hs1-1263	707	707	707	707	707	851
Val_Csyrichita1-1265	708	708	708	708	708	852
Val_Cfamiliaris1-1264	707	707	707	707	707	851
Val_Mcaroli1-1299	742	742	742	742	742	885
Val_Mmusculus1-1263	706	706	706	706	706	850
Val_Mpahani1-1263	706	706	706	706	706	850
Val_Panabis1-1261	704	704	704	704	704	848
Val_Bmutus1-1275	712	712	712	712	712	862
ValRS_Pf1-1090	619	619	619	619	619	801
ValRS_Pm2-1089	621	621	621	621	621	803
ValRS_Py1-1077	614	614	614	614	614	796
ValRS_Pf2-1046	592	592	592	592	592	764
ValRS_Pf2-1071	608	608	608	608	608	789
ValRS_Pf2-1070	607	607	607	607	607	789
ValRS_Pf2-1077	614	614	614	614	614	796
ValRS_Pf2-1138	672	672	672	672	672	854
ValRS_Hs1-1263	852	852	852	852	852	1030
Val_Csyrichita1-1265	853	853	853	853	853	1031
Val_Cfamiliaris1-1264	852	852	852	852	852	1030
Val_Mcaroli1-1299	887	887	887	887	887	1065
Val_Mmusculus1-1263	853	853	853	853	853	1029
Val_Mpahani1-1263	849	849	849	849	849	1029
Val_Panabis1-1261	851	851	851	851	851	1027
Val_Bmutus1-1275	863	863	863	863	863	1041
ValRS_Pf1-1090	802	802	802	802	802	976
ValRS_Pm2-1089	804	804	804	804	804	976
ValRS_Py1-1077	797	797	797	797	797	963
ValRS_Pf2-1046	769	769	769	769	769	932
ValRS_Pf2-1071	791	791	791	791	791	956
ValRS_Pf2-1070	790	790	790	790	790	956
ValRS_Pf2-1077	797	797	797	797	797	963
ValRS_Pf2-1138	855	855	855	855	855	1024
ValRS_Hs1-1263	1031	1031	1031	1031	1031	1160
Val_Csyrichita1-1265	1032	1032	1032	1032	1032	1161
Val_Cfamiliaris1-1264	1031	1031	1031	1031	1031	1160
Val_Mcaroli1-1299	1066	1066	1066	1066	1066	1195
Val_Mmusculus1-1263	1030	1030	1030	1030	1030	1159
Val_Mpahani1-1263	1029	1029	1029	1029	1029	1159
Val_Panabis1-1261	1028	1028	1028	1028	1028	1157
Val_Bmutus1-1275	1042	1042	1042	1042	1042	1171
ValRS_Pf1-1090	977	977	977	977	977	1090
ValRS_Pm2-1089	977	977	977	977	977	1089
ValRS_Py1-1077	964	964	964	964	964	1077
ValRS_Pf2-1046	939	939	939	939	939	1046
ValRS_Pf2-1071	958	958	958	958	958	1071
ValRS_Pf2-1070	964	964	964	964	964	1070
ValRS_Pf2-1077	964	964	964	964	964	1077
ValRS_Pf2-1138	1025	1025	1025	1025	1025	1138
ValRS_Hs1-1263	1161	1161	1161	1161	1161	1263
Val_Csyrichita1-1265	1162	1162	1162	1162	1162	1265
Val_Cfamiliaris1-1264	1161	1161	1161	1161	1161	1264
Val_Mcaroli1-1299	1196	1196	1196	1196	1196	1299
Val_Mmusculus1-1263	1160	1160	1160	1160	1160	1263
Val_Mpahani1-1263	1160	1160	1160	1160	1160	1263
Val_Panabis1-1261	1160	1160	1160	1160	1160	1261
Val_Bmutus1-1275	1172	1172	1172	1172	1172	1275

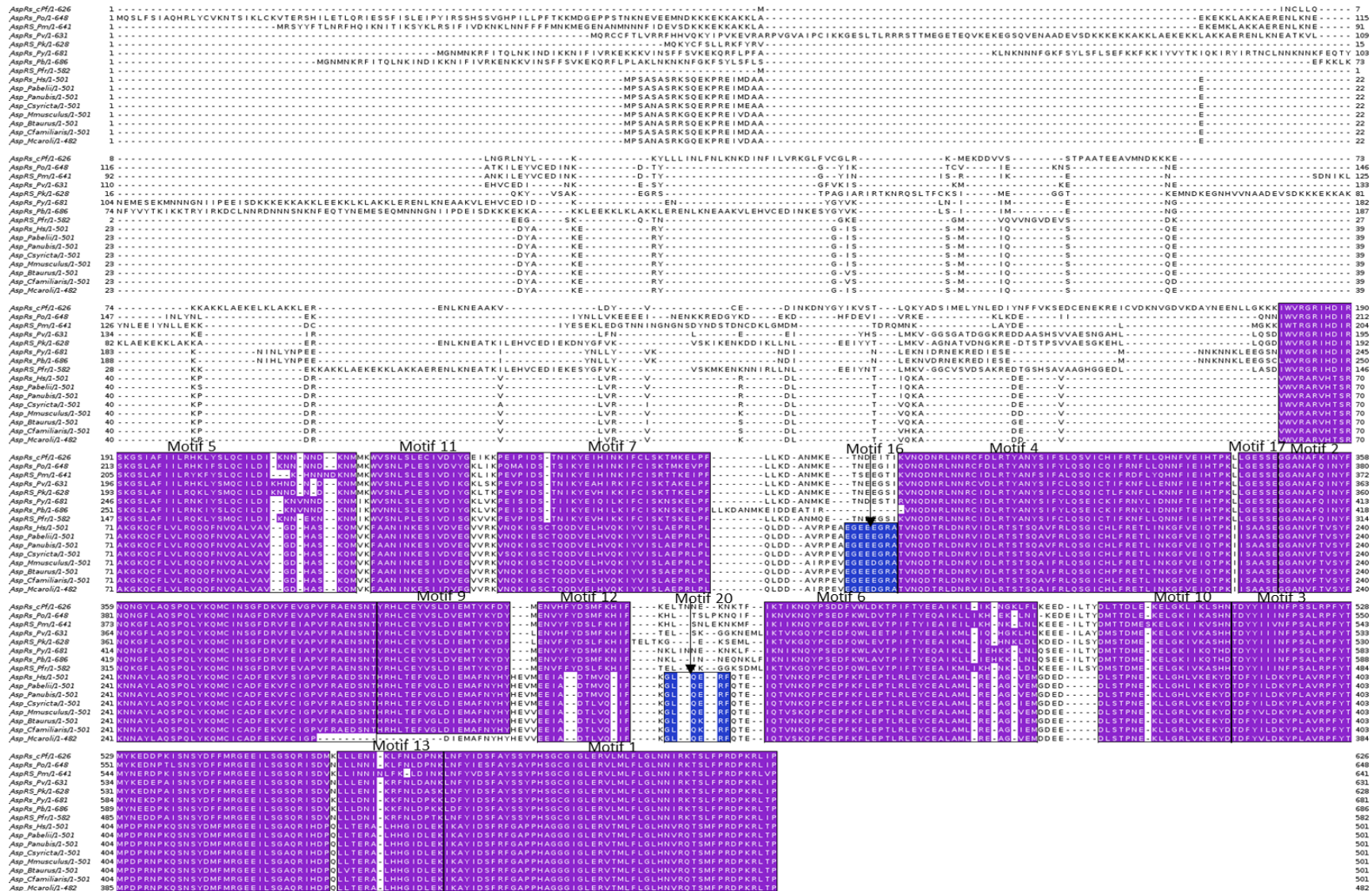
Additional file 4J: Mapping of motifs discovered in ValRS family to the multiple sequence alignment

Additional file 4K: Mapping of motifs discovered in AlaRS family to the multiple sequence alignment



[illegible]

Additional file 4L: Mapping of motifs discovered in AsnRS family to the multiple sequence alignment



Additional file 4M: Mapping of motifs discovered in AspRS family to the multiple sequence alignment

<i>GlyRs_Pf1-889</i>	1		Motif 17		Motif 19		Motif 14		Motif 28		5
<i>GlyRS_Pm1-884</i>	1	MRFTGLVLFSMFFQGRYTSFISSSLSNKRIFKTYNTIGTYIRIKFAAYVVPVSNFHKKKKKKKKNNYALFCKNINSSINAHNSMGSMNMNVSNANSSVNRN							SSSVNSNVNSSSIDETRVQQIEKYSQEISLIERNESIKVL		145
<i>GlyRS_Pk1-830</i>	1										11
<i>GlyRs_Pb1-812</i>	1										48
<i>GlyRs_Py1-806</i>	1										48
<i>GlyRS_Pf1-857</i>	1										48
<i>GlyRs_Po1-806</i>	1										49
<i>GlyRs_Pv1-842</i>	1										21
<i>GlyRs_Hu1-739</i>	1										11
<i>Gly_Ptrog1-739</i>	1										119
<i>Gly_Panab1-738</i>	1										119
<i>Gly_Csyricta1-739</i>	1										118
<i>Gly_Mmurina1-740</i>	1										119
<i>Gly_Btaurus1-739</i>	1										120
<i>Gly_Bmuta1-739</i>	1										119
<i>Gly_Mmusculus1-729</i>	1										119
<i>Gly_Mcarol1-729</i>	1										109
<i>GlyRs_Pf1-889</i>	6										82
<i>GlyRS_Pm1-884</i>	146										109
<i>GlyRS_Pk1-830</i>	12										89
<i>GlyRs_Pb1-812</i>	49	FRKNIILKPKKKVYVFFSNKPTIKMINDSEIEKKIEKISQEISLIERNEAIKIIDNDPLKKEYNNKLLKOKKTLNEVCNIFSKPLNLL									151
<i>GlyRs_Py1-806</i>	49	FRKNIILKPKKKVYVFFSNKPTIKMINDSEIEKKIEKISQEISLIERNEAIKIIDNDPLKKEYNNKLLKOKKTLNEVCNIFSKPLNLL									151
<i>GlyRS_Pf1-857</i>	50										90
<i>GlyRs_Po1-806</i>	22										96
<i>GlyRs_Pv1-842</i>	12										124
<i>GlyRs_Hu1-739</i>	120										124
<i>Gly_Ptrog1-739</i>	120										123
<i>Gly_Panab1-738</i>	119										124
<i>Gly_Csyricta1-739</i>	120										125
<i>Gly_Mmurina1-740</i>	121										124
<i>Gly_Btaurus1-739</i>	120										124
<i>Gly_Bmuta1-739</i>	120										124
<i>Gly_Mmusculus1-729</i>	110										114
<i>Gly_Mcarol1-729</i>	110										114
<i>GlyRs_Pf1-889</i>	83										171
<i>GlyRS_Pm1-884</i>	190										193
<i>GlyRS_Pk1-830</i>	84	SHVEASGSGAAEVTPIEE									169
<i>GlyRs_Pb1-812</i>	152										155
<i>GlyRs_Py1-806</i>	152										155
<i>GlyRS_Pf1-857</i>	91										129
<i>GlyRs_Po1-806</i>	93	ISSLIERNESIKCMDED									142
<i>GlyRs_Pv1-842</i>	87										129
<i>GlyRs_Hu1-739</i>	125										128
<i>Gly_Ptrog1-739</i>	125										128
<i>Gly_Panab1-738</i>	124										127
<i>Gly_Csyricta1-739</i>	125										128
<i>Gly_Mmurina1-740</i>	126										129
<i>Gly_Btaurus1-739</i>	125										128
<i>Gly_Bmuta1-739</i>	125										128
<i>Gly_Mmusculus1-729</i>	115										118
<i>Gly_Mcarol1-729</i>	115										118
<i>GlyRs_Pf1-889</i>	172										250
<i>GlyRS_Pm1-884</i>	194										296
<i>GlyRS_Pk1-830</i>	170										272
<i>GlyRs_Pb1-812</i>	156										262
<i>GlyRs_Py1-806</i>	156										272
<i>GlyRS_Pf1-857</i>	130	IDDELLKKEHDANEKSLKELKRLKLYIPSEPLHLATNRIKVDNLAK									289
<i>GlyRs_Po1-806</i>	149										252
<i>GlyRs_Pv1-842</i>	130	GLNDEELKKEHDANEKSLKELKRLKLYIPSEPLHLAKNRTKVDNLAK									302
<i>GlyRs_Hu1-739</i>	129										231
<i>Gly_Ptrog1-739</i>	129										231
<i>Gly_Panab1-738</i>	128										230
<i>Gly_Csyricta1-739</i>	129										231
<i>Gly_Mmurina1-740</i>	130										232
<i>Gly_Btaurus1-739</i>	129										231
<i>Gly_Bmuta1-739</i>	129										231
<i>Gly_Mmusculus1-729</i>	119										221
<i>Gly_Mcarol1-729</i>	119										221
<i>GlyRs_Pf1-889</i>	251										320
<i>GlyRS_Pm1-884</i>	297										356
<i>GlyRS_Pk1-830</i>	286	EG-SMDKR									325
<i>GlyRs_Pb1-812</i>	273	NCENINTN									311
<i>GlyRs_Py1-806</i>	262	-DGNKR									309
<i>GlyRS_Pf1-857</i>	290	-EKEKODEVNRTEDEPVGKQDQ									350
<i>GlyRs_Po1-806</i>	253	LVEMKKD									303
<i>GlyRs_Pv1-842</i>	303	AA-SOPHSA									396
<i>GlyRs_Hu1-739</i>	232	-SVEKKK									262
<i>Gly_Ptrog1-739</i>	232	-SVEKKK									262
<i>Gly_Panab1-738</i>	231	-SVEKKK									262
<i>Gly_Csyricta1-739</i>	232	-SAEKKK									262
<i>Gly_Mmurina1-740</i>	233	-SAEKKK									263
<i>Gly_Btaurus1-739</i>	232	-SAEKKK									262
<i>Gly_Bmuta1-739</i>	232	-SAEKKK									262
<i>Gly_Mmusculus1-729</i>	222	-SAEKKK									250
<i>Gly_Mcarol1-729</i>	222	-SAEKKK									252

Additional file 4N: Mapping of motifs discovered in GlyRS family to the multiple sequence alignment

[illegible]

Additional file 40: Mapping of motifs discovered in HisRS family to the multiple sequence alignment

<i>LysRs_cPfl-583</i>	1	-----MTSKSFLLSFLKYKHVNTYIFKESKFS	Motif 11	-----KSCVFTMNEKKEHVLGEKK	Motif 13	-----KRVVNASKD	Motif 8	-----KLENGEHELEDTINLTGRIMRV	142
<i>LysRS_Pm2-583</i>	1	-----MLKVLPLFLKSNMCTNIFNRSYSSLTLK		-----KXVHKVFTTMNEKKEHVLGEKK		-----KXVHKVFTTMNEKKEHVLGEKK		-----KLENGEHELEDTINLTGRIMRV	142
<i>LysRS_Pk2-585</i>	1	-----MFYRLPLILRYHKNFSGSHFYKQHSFPLTILNKNKNIICPVNK		-----KQTFKTMSEKKEHVMGEKKVPSKQOVKD		-----KQTFKTMSEKKEHVMGEKKVPSKQOVKD		-----KLENGEHELEDTINLTGRIMRV	144
<i>LysRS_Pfr2-585</i>	1	-----MFDRDPLILRYHKNFSGSYFYKQHSYSSAILNKNKNIIFPVNK		-----KQTFKTMSEKKEHVMGEKKVAKNHOAKD		-----KQTFKTMSEKKEHVMGEKKVAKNHOAKD		-----KLENGEHELEDTINLTGRIMRV	144
<i>LysRS_Pv2-585</i>	1	-----MFSDLLPLILRYHKNIGSHFYKQHSYSSAILNKNKNIICPVNK		-----KQTFKTMSEKKEHVMGEKKVAKNHOAKD		-----KQTFKTMSEKKEHVMGEKKVAKNHOAKD		-----KLENGEHELEDTINLTGRIMRV	144
<i>LysRS_Pb2-578</i>	1	-----MLTFLSLILKNKNFNROFFFNKNFLTLEFHNFKANII		-----IKVHFTMTTEKREHVTSNQKKNTPPPVD		-----IKVHFTMTTEKREHVTSNQKKNTPPPVD		-----KLENGEHELEDTINLTGRIMRV	147
<i>LysRS_Py2-581</i>	1	-----MHMLTLFLPLILKNKNLRROFFFNKNFLTLEFHNFKTNI		-----IIKAQFTTMTTEKREHVTSNQKKNTPPPVD		-----IIKAQFTTMTTEKREHVTSNQKKNTPPPVD		-----KLENGEHELEDTINLTGRIMRV	130
<i>LysRS_Po2-601</i>	1	-----MLSVLLPLFLKRRKKRPTEFLYHNSFSLVHLVHLSVIFGKNI		-----IPRIISKHPFTMNEKKEHINTTEDKISTITENKGNTPTE		-----IPRIISKHPFTMNEKKEHINTTEDKISTITENKGNTPTE		-----KLENGEHELEDTINLTGRIMRV	153
<i>LysRS_Hs2-597</i>	1	-----MAAVOAAAEVKVDGSE		-----KLSNKLKRRKRAEKKAKEAKOKELSEKOLSOATAAATHTTDN		-----KLSNKLKRRKRAEKKAKEAKOKELSEKOLSOATAAATHTTDN		-----KLENGEHELEDTINLTGRIMRV	153
<i>Lys_Panubis2-647</i>	1	-----MRSAAALSLRNSRRSAMKAPASPPEEVGPTLTWLKTYTPPYI		-----IWGSPSGKMAVQAAAEVKVDGSE		-----IWGSPSGKMAVQAAAEVKVDGSE		-----KLENGEHELEDTINLTGRIMRV	185
<i>Lys_Pabelli2-625</i>	1	-----MLTQAAVRLVRGSLRKTSWAEWGHRELRLGQLAPFTAPHDKDSF		-----SDQRSELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----SDQRSELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----KLENGEHELEDTINLTGRIMRV	162
<i>Lys_Csyricha2-594</i>	1	-----AAVPEVEIKVDG		-----GEPKLSKNELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----GEPKLSKNELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----KLENGEHELEDTINLTGRIMRV	137
<i>Lys_Mmulatta2-625</i>	1	-----MLTQAAVRLVRGSLRKTSWAEWGHRELRLGQLAPFTAPHDKDSF		-----SDQRSELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----SDQRSELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----KLENGEHELEDTINLTGRIMRV	163
<i>Lys_Mmusculus2-595</i>	1	-----MATLQSEVVKDGE		-----DKLSKNELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----DKLSKNELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----KLENGEHELEDTINLTGRIMRV	133
<i>Lys_Mmusculus2-624</i>	1	-----LMQAAVRLVRGSLRKTSWAEWGHRELRLGQLAPFTALHKDTP		-----LSDQRSELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----LSDQRSELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----KLENGEHELEDTINLTGRIMRV	162
<i>Lys_Btaurus2-623</i>	1	-----LVQAAARLVRGSLRKTSWAEWGHRELRLGQLAPFTTHHKDTP		-----LALSDQRSELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----LALSDQRSELKRRLKAEKKVAEKAOKELSEKOLSOATAAATHTTDN		-----KLENGEHELEDTINLTGRIMRV	161
<i>LysRS_cPfl-583</i>	143	ASGGOKLRFDFDLVGGGKIQVLANYSFHNEKGN	Motif 6	AECDYKIRRGDIVGIVGFGPKS	Motif 26	GLKGLSIFPKETILL	Motif 10	TEIRYRORYLDLINESST	319
<i>LysRS_Pm2-583</i>	143	ASGGOKLRFDFDLVGGGKIQVLANYAFHDTKTNP		AECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	319
<i>LysRS_Pk2-585</i>	143	ASGGOKLRFDFDLVGGGKIQVLANYAFHDTKTNP		AECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	321
<i>LysRS_Pfr2-585</i>	143	ASGGOKLRFDFDLVGGGKIQVLANYAFHDTKTNP		AECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	321
<i>LysRS_Pv2-585</i>	143	ASGGOKLRFDFDLVGGGKIQVLANYAFHDTKTNP		AECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	321
<i>LysRS_Pb2-578</i>	138	SSGOKLRFDFDLVGGGKIQVLANYFHDKEKSNF		VECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	314
<i>LysRS_Py2-581</i>	143	SSGOKLRFDFDLVGGGKIQVLANYFHDKEKSNF		VECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	314
<i>LysRS_Po2-601</i>	134	ASGGOKLRFDFDLVGGGKIQVLANYAFHDTKTNP		AECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	314
<i>LysRS_Hs2-597</i>	136	ASGGOKLRFDFDLVGGGKIQVLANYFHDKEKSNF		VECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	312
<i>Lys_Panubis2-647</i>	164	ASGGOKLRFDFDLVGGGKIQVLANYFHDKEKSNF		VECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	362
<i>Lys_Pabelli2-625</i>	138	SSGOKLRFDFDLVGGGKIQVLANYFHDKEKSNF		VECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	314
<i>Lys_Csyricha2-594</i>	134	ASGGOKLRFDFDLVGGGKIQVLANYFHDKEKSNF		VECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	314
<i>Lys_Mmulatta2-625</i>	164	ASGGOKLRFDFDLVGGGKIQVLANYFHDKEKSNF		VECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	362
<i>Lys_Mmusculus2-595</i>	134	ASGGOKLRFDFDLVGGGKIQVLANYFHDKEKSNF		VECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	314
<i>Lys_Mmusculus2-624</i>	164	ASGGOKLRFDFDLVGGGKIQVLANYFHDKEKSNF		VECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	362
<i>Lys_Btaurus2-623</i>	162	ASGGOKLRFDFDLVGGGKIQVLANYFHDKEKSNF		VECDYKIRRGDIVGIVGFGPKS		GLKGLSIFPKETILL		TEIRYRORYLDLINESST	362
<i>LysRS_cPfl-583</i>	320	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK	Motif 1	WSEDFSSLVHMLFGTYKILYNNKDGPNNDP	Motif 7	IEIDTPPYKVSIVEELE	Motif 9	LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	502
<i>LysRS_Pm2-583</i>	320	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	502
<i>LysRS_Pk2-585</i>	320	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	502
<i>LysRS_Pfr2-585</i>	322	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	504
<i>LysRS_Pv2-585</i>	322	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	504
<i>LysRS_Pb2-578</i>	315	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	497
<i>LysRS_Py2-581</i>	319	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	500
<i>LysRS_Po2-601</i>	331	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	513
<i>LysRS_Hs2-597</i>	313	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	504
<i>Lys_Panubis2-647</i>	363	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	546
<i>Lys_Pabelli2-625</i>	341	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	524
<i>Lys_Csyricha2-594</i>	310	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	494
<i>Lys_Mmulatta2-625</i>	341	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	524
<i>Lys_Mmusculus2-595</i>	313	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	494
<i>Lys_Mmusculus2-624</i>	339	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	524
<i>Lys_Btaurus2-623</i>	339	DVYIEIGKVFRLNIDINTHNPEFTSCEFYWAYADYNDLIK		WSEDFSSLVHMLFGTYKILYNNKDGPNNDP		IEIDTPPYKVSIVEELE		LE--OPFDSNETIEKMINIIEKHKIELPNPPTAAKLLDOLASHF	522
<i>LysRS_cPfl-583</i>	503	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ	Motif 3	LDFAFCTSL	Motif 14	AN--	583
<i>LysRS_Pm2-583</i>	503	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	583
<i>LysRS_Pk2-585</i>	503	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	585
<i>LysRS_Pfr2-585</i>	503	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	585
<i>LysRS_Pv2-585</i>	503	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	585
<i>LysRS_Pb2-578</i>	498	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	585
<i>LysRS_Py2-601</i>	501	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	578
<i>LysRS_Po2-601</i>	514	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	581
<i>LysRS_Hs2-597</i>	497	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	601
<i>Lys_Panubis2-647</i>	547	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	597
<i>Lys_Pabelli2-625</i>	525	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	647
<i>Lys_Csyricha2-594</i>	494	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	625
<i>Lys_Mmulatta2-625</i>	525	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	625
<i>Lys_Mmusculus2-595</i>	494	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	594
<i>Lys_Mmusculus2-624</i>	524	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	624
<i>Lys_Btaurus2-623</i>	523	NAYTELNDPFFKO		KECFKLOQDKREKGDTEAFQ		LDFAFCTSL		AN--	623

Additional file 4P: Mapping of motifs discovered in LysRS family to the multiple sequence alignment

Additional file 4Q: Mapping of motifs discovered in PheRS family to the multiple sequence alignment

ProRs_cPf1-746	190	ED	Motif 23	EMAKNEELQNNH-N	HNKNNSSNNNNNNNNNNH	I-L-KDTI-L-KGKL-L	SN	NEVEDNKS	248
ProRs_Pm1-747	175	ED		QIEEDKDKKKKE-DTKAE	EETGK-QNKNSNSN	R-NDKR-NDNNSN-SKSSNNNNNANNKNIYHNSGT		GVYNE	249
ProRs_Pm2-776	216			EGMNKGDTSKKEDI	QEEEEK-ICKQCKENSGGGNGDKTRSNRDTITSSIG-T			GGGKDDK	278
ProRs_Pkl1-715	174			HIEEGQNKMKKE-D	HMHEEO-KMEKQK	GTNAGG-S	GG	NAGKDDNAK	217
ProRs_Pyl1-728	169			TPEFINLSDSK-DTENQIKRDDSNINEESSKDTAKLAHGEN-NNNKHGKNDK	I-NDRD			KGKDG	230
ProRs_Pv1-773	208			HIEEGKDKLKE-D	QLLEED-KLEKGSGSGI	G-SGVS	SGVS	SGVS	275
ProRs_Pfr1-748	208			HIEEGRDIKKE-D	HMOMEIE-KLEKQK		CA	NAGKDDHAK	251
ProRs_Pb1-755	196			TPEFINLSDSK-DTENQIKRDDTNINEEISKGTTKLPNGEN-NNNKHGENDEK	I-NEED			KGKDE	257
ProRs_Hs1-1512	895	GLETPKAVL	FDKVASQGEVVRKLTAEKAPKDDQIVAVQELLQLKAAQYKSLTGVEYKPVYSAT	GAEDKDKKKKE-K	ENKSEK-QNKPKQNDQ	R-KDPSKNQSGSL-S	SS	GAGEGQPKK	1010
Pro_Panubis1-1512	895	GLETPKAVL	FDKVASQGEVVRKLTAEKAPKDDQIVAVQELLQLKAAQYKSLTGVEYKPVYSAT	GAEDKDKKKKE-K	ENKSEK-QNKPKQNDQ	R-KDPSKNQSGSL-S	SS	GAGEGQPKK	1010
Pro_Mmurinus1-1514	895	DDPAPEAKVL	FDKVASQGEVVRKLTAEKASQDQVETAVQELLRLKAAQYKSLTGVEYKPVYSAT	GAEDKDKKKKE-K	ENKSEK-QSKPKQNDQ	K-KDSSKNQSGSL-P	SSG	GPAGEGQPKK	1012
Pro_Cfamiliaris1-1519	892	GPEAPEAKVL	FDKVASQGEVVRKLTAEKASQDQVDVAVQELLQLKAAQYKSLTGVEYKPVYSAT	GAEDKDKKKKE-K	ENKSEK-QNKPLKQSDQ	K-KDSSKNQSGSL-S	SS	GAGEGQPKK	1017
Pro_Csyricha1-1519	902	DPETPKAVL	FDKVASQGEVVRKLTAEKASQDQVDVAVQELLQLKAAQYKSLTGVEYKPVYSAT	GAEDKDKKKKE-K	ENKSEK-QNKPKQNDQ	R-KDSSKNQSGSL-S	SS	GAGEGQPKK	1017
Pro_Btaurus1-1511	894	GPETPEAKLL	FDKVASQGEVVRKLTAEKASQDQVDVAVQELLQLKAAQYKSLTGVEYKPVYSAT	GAEDKDKKKKE-K	ENKSEK-QNKPKQNDQ	K-KDSSKNQSGSL-S	SS	GAGEGQPKK	1017
Pro_Mmusculus1-1512	895	GAEPKAVL	FDKVASQGEVVRKLTAEKASQDQVDVAVQELLQLKAAQYKSLTGVEYKPVYSAT	GAEDKDKKKKE-K	ENKSEK-QNKPKQNDQ	R-KDSSKNQSGSL-S	SS	GAGEGQPKK	1017
Pro_Mpahari1-1512	895	GSEKPEAKVL	FDKVASQGEVVRKLTAEKASQDQVDVAVQELLQLKAAQYKSLTGVEYKPVYSAT	GAEDKDKKKKE-K	ENKSEK-QNKPKQNDQ	G-KDSSKNQSGSL-S	SS	GAGEGQPKK	1010
ProRs_cPf1-746	249	SNILGI	Motif 2	QVIVKSELEYDIDSGCYILRPASYYIWEICQVAFNKEIKKLVNVSFYFL	LVTKNKLKEKKNHIEGFSPEVAWVTKYGDLSHLPPEIAIRPT	SETIMYSVFSKWI	Motif 1	TAKHNE	431
ProRs_Pm2-747	250	SNILGI		QVIVKSELEYDIDSGCYILRPASYYIWEICVOTFFNKEIKKLVNVSFYFL	LVTKNKLKEKKNHIEGFSPEVAWVTKYGSDTLPEEIAIRPT	SETIMYSVFSKWI		TAKHNE	432
ProRs_Pm2-776	279	SNLLGI		QVIVKSELEYDIDSGCYILRPASYYIWEICVQNFNKEIKKLVNVSFYFL	LVTKNKLKEKKNHIEGFSPEVAWVTKYGSDTLPEEIAIRPT	SETIMYSVFSKWI		TAKHNE	461
ProRs_Pkl1-715	218	ANILGI		QVIVKSELEYDIDSGCYILRPASYYIWEICLQTFNKEIKKLVNVSFYFL	LVTKNKLKEKKNHIEGFSPEVAWVTKYGSDTLPEEIAIRPT	SETIMYSVFSKWI		TAKHNE	400
ProRs_Pyl1-728	231	SNILGI		QVIVKSELEYDIDSGCYILRPASYYIWEICQVAFNKEIKKLVNVSFYFL	LVTKNKLKEKKNHIEGFSPEVAWVTKYGSDTLPEEIAIRPT	SETIMYSVFSKWI		TAKHNE	413
ProRs_Pfr1-773	276	ANILGI		QVIVKSELEYDIDSGCYILRPASYYIWEICQVAFNKEIKKLVNVSFYFL	LVTKNKLKEKKNHIEGFSPEVAWVTKYGSDTLPEEIAIRPT	SETIMYSVFSKWI		TAKHNE	413
ProRs_Pfr1-748	252	ANILGI		QVIVKSELEYDIDSGCYILRPASYYIWEICQVAFNKEIKKLVNVSFYFL	LVTKNKLKEKKNHIEGFSPEVAWVTKYGSDTLPEEIAIRPT	SETIMYSVFSKWI		TAKHNE	434
ProRs_Pb1-755	258	GNILGI		QVIVKSELEYDIDSGCYILRPASYYIWEICQVAFNKEIKKLVNVSFYFL	LVTKNKLKEKKNHIEGFSPEVAWVTKYGSDTLPEEIAIRPT	SETIMYSVFSKWI		TAKHNE	440
ProRs_Hs1-1512	1011	QTRLGLEAKKEENLADWYS	QVITKSEMIEYDIDSGCYILRPWAYIAWEIKDFFDAEIKKLVNVSFYFL	FMVSGALEKEKTHVADFPEVAWVTYSGKTELAEPIAIRPT	SETVMYPAYAKWVSHRDLPIRLNQWCVNVRWEFKHPQPLRTREFLWQEGH	SFAFATF		EEAAEVLQILEY	1193
Pro_Panubis1-1512	1011	QTRLGLEAKKEENLADWYS	QVITKSEMIEYDIDSGCYILRPWAYIAWEIKDFFDAEIKKLVNVSFYFL	FMVSGALEKEKTHVADFPEVAWVTYSGKTELAEPIAIRPT	SETVMYPAYAKWVSHRDLPIRLNQWCVNVRWEFKHPQPLRTREFLWQEGH	SFAFATF		EEAAEVLQILEY	1193
Pro_Mmurinus1-1514	1013	QTRLGLEAKKEENLADWYS	QVITKSEMIEYDIDSGCYILRPWAYIAWEIKDFFDAEIKKLVNVSFYFL	FMVSGALEKEKTHVADFPEVAWVTYSGKTELAEPIAIRPT	SETVMYPAYAKWVSHRDLPIRLNQWCVNVRWEFKHPQPLRTREFLWQEGH	SFAFATF		EEAAEVLQILEY	1193
Pro_Cfamiliaris1-1519	1008	QTRLGLEAKKEENLADWYS	QVITKSEMIEYDIDSGCYILRPWAYIAWEIKDFFDAEIKKLVNVSFYFL	FMVSGALEKEKTHVADFPEVAWVTYSGKTELAEPIAIRPT	SETVMYPAYAKWVSHRDLPIRLNQWCVNVRWEFKHPQPLRTREFLWQEGH	SFAFATF		EEAAEVLQILEY	1193
Pro_Btaurus1-1511	1010	QTRLGLEAKKEENLADWYS	QVITKSEMIEYDIDSGCYILRPWAYIAWEIKDFFDAEIKKLVNVSFYFL	FMVSGALEKEKTHVADFPEVAWVTYSGKTELAEPIAIRPT	SETVMYPAYAKWVSHRDLPIRLNQWCVNVRWEFKHPQPLRTREFLWQEGH	SFAFATF		EEAAEVLQILEY	1193
Pro_Mmusculus1-1512	1010	QTRLGLEAKKEENLADWYS	QVITKSEMIEYDIDSGCYILRPWAYIAWEIKDFFDAEIKKLVNVSFYFL	FMVSGALEKEKTHVADFPEVAWVTYSGKTELAEPIAIRPT	SETVMYPAYAKWVSHRDLPIRLNQWCVNVRWEFKHPQPLRTREFLWQEGH	SFAFATF		EEAAEVLQILEY	1193
Pro_Mpahari1-1512	1011	QTRLGLEAKKEENLADWYS	QVITKSEMIEYDIDSGCYILRPWAYIAWEIKDFFDAEIKKLVNVSFYFL	FMVSGALEKEKTHVADFPEVAWVTYSGKTELAEPIAIRPT	SETVMYPAYAKWVSHRDLPIRLNQWCVNVRWEFKHPQPLRTREFLWQEGH	SFAFATF		EEAAEVLQILEY	1193
ProRs_cPf1-746	432	RRWYEELAVP	Motif 6	IKSGEKFGGANFTSTAEAFISENGRAIQAAATSHYLGNTFAKMKKIEFEDQ	N-EVKQVYHOTS	Motif 4		TAKHNE	608
ProRs_Pm2-747	432	RRWYEELAVP		IKSGEKFGGANFTSTAEAFISENGRAIQAAATSHYLGNTFAKMKKIEFEDQ	N-EVKQVYHOTS			TAKHNE	609
ProRs_Pm2-776	401	RRWYEELAVP		IKSGEKFGGANFTSTAEAFISENGRAIQAAATSHYLGNTFAKMKKIEFEDQ	N-EVKQVYHOTS			TAKHNE	577
ProRs_Pkl1-715	411	RRWYEELAVP		IKSGEKFGGANFTSTAEAFISENGRAIQAAATSHYLGNTFAKMKKIEFEDQ	N-EVKQVYHOTS			TAKHNE	590
ProRs_Pyl1-728	401	RRWYEELAVP		IKSGEKFGGANFTSTAEAFISENGRAIQAAATSHYLGNTFAKMKKIEFEDQ	N-EVKQVYHOTS			TAKHNE	577
ProRs_Pv1-773	435	RRWYEELAVP		IKSGEKFGGANFTSTAEAFISENGRAIQAAATSHYLGNTFAKMKKIEFEDQ	N-EVKQVYHOTS			TAKHNE	635
ProRs_Pfr1-748	435	RRWYEELAVP		IKSGEKFGGANFTSTAEAFISENGRAIQAAATSHYLGNTFAKMKKIEFEDQ	N-EVKQVYHOTS			TAKHNE	610
ProRs_Pb1-755	441	RRWYEELAVP		IKSGEKFGGANFTSTAEAFISENGRAIQAAATSHYLGNTFAKMKKIEFEDQ	N-EVKQVYHOTS			TAKHNE	617
ProRs_Hs1-1512	1194	QVYEELLA		IPVVGKRTKEKFAAGDQYTTTLEAFISASGRAIQGATSHHLLGONFSKMFIVFEDQ	KTPGKEQFAYONS	Motif 5		TAKHNE	1376
Pro_Panubis1-1512	1194	QVYEELLA		IPVVGKRTKEKFAAGDQYTTTLEAFISASGRAIQGATSHHLLGONFSKMFIVFEDQ	KTPGKEQFAYONS			TAKHNE	1376
Pro_Mmurinus1-1514	1196	QVYEELLA		IPVVGKRTKEKFAAGDQYTTTLEAFISASGRAIQGATSHHLLGONFSKMFIVFEDQ	KTPGKEQFAYONS			TAKHNE	1378
Pro_Cfamiliaris1-1519	1191	QVYEELLA		IPVVGKRTKEKFAAGDQYTTTLEAFISASGRAIQGATSHHLLGONFSKMFIVFEDQ	KTPGKEQFAYONS			TAKHNE	1373
Pro_Csyricha1-1519	1201	QVYEELLA		IPVVGKRTKEKFAAGDQYTTTLEAFISASGRAIQGATSHHLLGONFSKMFIVFEDQ	KTPGKEQFAYONS			TAKHNE	1383
Pro_Btaurus1-1511	1191	QVYEELLA		IPVVGKRTKEKFAAGDQYTTTLEAFISASGRAIQGATSHHLLGONFSKMFIVFEDQ	KTPGKEQFAYONS			TAKHNE	1376
Pro_Mmusculus1-1512	1194	QVYEELLA		IPVVGKRTKEKFAAGDQYTTTLEAFISASGRAIQGATSHHLLGONFSKMFIVFEDQ	KTPGKEQFAYONS			TAKHNE	1376
Pro_Mpahari1-1512	1194	QVYEELLA		IPVVGKRTKEKFAAGDQYTTTLEAFISASGRAIQGATSHHLLGONFSKMFIVFEDQ	KTPGKEQFAYONS			TAKHNE	1376
ProRs_cPf1-746	609	NSCVIV	Motif 16	RRDNNKCNVKKESVLLLETOQMLVDIHKNLFLKAKKKLDDSI	QVQVTSFSEVMNALNKKMM	LAPWCEDI	Motif 11	PPMPNNTKCF	746
ProRs_Pm1-747	610	NSCVFV		RRDNNKCNVKKESVLLLETOQMLVDIHKNLFLKAKKKLDDSI	QVQVTSFSEVMNALNKKMM	LAPWCEDI		PPMPNNTKCF	747
ProRs_Pm2-776	639	NSCIEF		RRDNNKCNVKKESVLLLETOQMLVDIHKNLFLKAKKKLDDSI	QVQVTSFSEVMNALNKKMM	LAPWCEDI		PPMPNNTKCF	776
ProRs_Pkl1-715	578	NSCLFV		RRDNNKCNVKKESVLLLETOQMLVDIHKNLFLKAKKKLDDSI	QVQVTSFSEVMNALNKKMM	LAPWCEDI		PPMPNNTKCF	715
ProRs_Pyl1-728	591	NSCVFV		RRDNNKCNVKKESVLLLETOQMLVDIHKNLFLKAKKKLDDSI	QVQVTSFSEVMNALNKKMM	LAPWCEDI		PPMPNNTKCF	728
ProRs_Pv1-773	636	NSCLCV		RRDNNKCNVKKESVLLLETOQMLVDIHKNLFLKAKKKLDDSI	QVQVTSFSEVMNALNKKMM	LAPWCEDI		PPMPNNTKCF	773
ProRs_Pfr1-748	611	NSCVFV		RRDNNKCNVKKESVLLLETOQMLVDIHKNLFLKAKKKLDDSI	QVQVTSFSEVMNALNKKMM	LAPWCEDI		PPMPNNTKCF	748
ProRs_Pb1-755	618	NSCVFV		RRDNNKCNVKKESVLLLETOQMLVDIHKNLFLKAKKKLDDSI	QVQVTSFSEVMNALNKKMM	LAPWCEDI		PPMPNNTKCF	755
ProRs_Hs1-1512	1377	QCFVAV		RRDTEGKLTVAENAEATKLQALILEDIHVTLFTRASEDLKTHM	VANTEMDFQKILDSGKI	QVQVTSFSEVMNALNKKMM		PPMPNNTKCF	1512
Pro_Panubis1-1512	1377	QCFVAV		RRDTEGKLTVAENAEATKLQALILEDIHVTLFTRASEDLKTHM	VANTEMDFQKILDSGKI	QVQVTSFSEVMNALNKKMM		PPMPNNTKCF	1512
Pro_Mmurinus1-1514	1379	QCFVAV		RRDTEGKLTVAENAEATKLQALILEDIHVTLFTRASEDLKTHM	VANTEMDFQKILDSGKI	QVQVTSFSEVMNALNKKMM		PPMPNNTKCF	1514
Pro_Cfamiliaris1-1519	1374	QCFVAV		RRDTEGKLTVAENAEATKLQALILEDIHVTLFTRASEDLKTHM	VANTEMDFQKILDSGKI	QVQVTSFSEVMNALNKKMM		PPMPNNTKCF	1509
Pro_Btaurus1-1511	1384	QCFVAV		RRDTEGKLTVAENAEATKLQALILEDIHVTLFTRASEDLKTHM	VANTEMDFQKILDSGKI	QVQVTSFSEVMNALNKKMM		PPMPNNTKCF	1519
Pro_Mmusculus1-1512	1376	QCFVAV		RRDTEGKLTVAENAEATKLQALILEDIHVTLFTRASEDLKTHM	VANTEMDFQKILDSGKI	QVQVTSFSEVMNALNKKMM		PPMPNNTKCF	1511
Pro_Mpahari1-1512	1377	QCFVAV		RRDTEGKLTVAENAEATKLQALILEDIHVTLFTRASEDLKTHM	VANTEMDFQKILDSGKI	QVQVTSFSEVMNALNKKMM		PPMPNNTKCF	1512

Additional file 4R: Mapping of motifs discovered in ProRS family to the multiple sequence alignment

Additional file 4S: Mapping of motifs discovered in SerRS family to the multiple sequence alignment

	Motif 10	Motif 5	Motif 28	Motif 11	Motif 4
ThrRs_Pf1-1013	412	412	412	412	412
ThrRs_Po1-903	314	314	314	314	314
ThrRs_Pm1-890	301	301	301	301	301
ThrRs_Pk1-888	312	312	312	312	312
ThrRs_Pf1-875	307	307	307	307	307
ThrRs_Pb1-890	318	318	318	318	318
ThrRs_Py1-935	363	363	363	363	363
ThrRs_Pv1-924	392	392	392	392	392
ThrRs_Hs1-723	205	205	205	205	205
Thr_Panubis1-723	205	205	205	205	205
Thr_Csyrichia1-723	205	205	205	205	205
Thr_Btaurus1-723	205	205	205	205	205
Thr_Cfamilaris1-723	205	205	205	205	205
Thr_Mpahari1-722	204	204	204	204	204
Thr_Mmusculus1-722	204	204	204	204	204
Thr_Mcaroli1-722	204	204	204	204	204
	Motif 13	Motif 1	Motif 7	Motif 16	Motif 3
ThrRs_Pf1-1013	594	594	594	594	594
ThrRs_Po1-903	496	496	496	496	496
ThrRs_Pm1-890	483	483	483	483	483
ThrRs_Pk1-888	494	494	494	494	494
ThrRs_Pf1-875	489	489	489	489	489
ThrRs_Pb1-890	501	501	501	501	501
ThrRs_Py1-935	546	546	546	546	546
ThrRs_Pv1-924	514	514	514	514	514
ThrRs_Hs1-723	385	385	385	385	385
Thr_Csyrichia1-723	385	385	385	385	385
Thr_Btaurus1-723	385	385	385	385	385
Thr_Cfamilaris1-723	385	385	385	385	385
Thr_Mpahari1-722	384	384	384	384	384
Thr_Mmusculus1-722	384	384	384	384	384
Thr_Mcaroli1-722	384	384	384	384	384
	Motif 14	Motif 8	Motif 17	Motif 2	Motif 19
ThrRs_Pf1-1013	926	926	926	926	926
ThrRs_Po1-903	821	821	821	821	821
ThrRs_Pm1-890	808	808	808	808	808
ThrRs_Pk1-888	800	800	800	800	800
ThrRs_Pf1-875	787	787	787	787	787
ThrRs_Pb1-890	806	806	806	806	806
ThrRs_Py1-935	851	851	851	851	851
ThrRs_Pv1-924	836	836	836	836	836
ThrRs_Hs1-723	644	644	644	644	644
Thr_Panubis1-723	644	644	644	644	644
Thr_Csyrichia1-723	644	644	644	644	644
Thr_Btaurus1-723	644	644	644	644	644
Thr_Cfamilaris1-723	644	644	644	644	644
Thr_Mpahari1-722	644	644	644	644	644
Thr_Mmusculus1-722	644	644	644	644	644
Thr_Mcaroli1-722	643	643	643	643	643

Additional file 47: Mapping of motifs discovered in ThrRS family to the multiple sequence alignment