

## Supplementary data 1 (CIR alignment)

## a) CIR amino acid sequence alignment:

	..... .....	..... .....	..... .....	..... .....	..... .....	..... .....	..... .....	..... .....	..... .....	..... .....	..... .....
	5	15	25	35	45	55	65	75	85	95	
PCHAS_000950	-----MAL	NLCNALKDI-	EEFLPNS--L	SSESKDVNRN	LYKTYCPIDQ	RGKQECRTDG	ERIGAMFKLL	LQQL-FVNNE	GDLELENKND	EYSEYAILWL	
PCHAS_001010	-----MTL	NLCSAFKGI-	GELLPDS--L	SPESDDDNHS	LYKIYCPDSK	TGNPECKTDG	QRISAMFRYL	LEYL-FVNND	DDLKSENQNN	EYSEYAILWL	
PCHAS_140090	-----MNN	-LCGLINVI-	DTITSVN-VN	NSILFIGNNE	TIKECCTNDT	GKKEECDSEYI	KITNSAIIAV	LNYP--KNIS	VADGDDLEDD	KIAEYAILWL	
PCHAS_000730	-----MSE	NLCKLINSI-	DKSITVN-VN	NSLAFIEFDN	ILTEYCTEEG	GGKKECFSEY	VLISSIFIKL	LKIF-----	-MNLDDLEND	KHAQYAILWF	
PCHAS_070160	-----MNN	-LCTIFKGI-	DDILVVK-SN	NPDYRHSFS	SLNSYCPSSD	GKNQECNNYE	EITGSAFISL	LIYL-----	---KNYVEND	KLSQYAILWL	
PCHAS_030020	-----MSK	KLCGLIIGI-	DNGIIVK-LN	GSNVQIKNDT	TFNDYCPEGM	KTNGECDSNI	QNVNFAIIKL	LTYF-----K	SVVDDILEDN	KLSEYAILWL	
PCHAS_000360	-----MSK	KLCGLINVI-	DHGVKLT-MN	GTNVQIKNDA	TFNDYCPEGM	KTNGECDSNI	QNVNSAIIKL	LTYF-----K	SVVDDILEDN	KLSEYAILWL	
PCHAS_110020	-----MDN	-LCEAIDQI-	DEKITLD-TN	GSKQLFHYNT	GSYPYINGE	GGTQECFSEY	VIVSSLFIKM	VNHF-----	-MKGDDLKNN	KLAEYAILWL	
PCHAS_050060	-----MDN	-LCKLVNAI-	DGKFSVD-VN	NLKAIIEFDN	IQNQYCT-GD	NGGGECFPYE	AIVSSYFIKM	LNYP-----	-MNDKDLNMK	KFGEYVILSL	
PCHAS_010030	-----MSK	EVCSIINTI-	DDKVSVN-AN	GLIAMVDFNE	IHSKYCD--S	EEGGFCYSYE	AIVSSVFTKI	LKHF----ES	CNDDDNLENA	KLAQYAILWL	
PCHAS_030030	-----MDN	-LCKVINRL-	DRKITVK-VD	NLIAMIEFDK	IQSKYCT-AK	DEEGTCYSYE	AIISSYFITL	LNHF-----K	SCNDNNLEND	KLAQYAVLWL	
PCHAS_000350	-----MSE	KLCKLFSAI-	DNGVTVT-VN	NSKAMIKFDN	IFTEYCYEED	GGKKECFSEY	VLVTSYFIGL	LNYP-----M	KVDDLK--HN	KNAEYAILWL	
PCHAS_011530	-----MDT	-LCQLFKGF-	DENLVVK-PN	NSEFSDSNFP	SFNNFCPTDG	GKKQKCNSEY	EMVISAFLTL	IINF----AS	INDGGNIEND	KIAQYAILWL	
PCHAS_073160	-----	--CEFINTI-	DVLLDVK-WD	GSKISFKYS	SLDNYCSMGN	EKQEDCNNE	EMISSAFLTL	LHHF---KSV	NAGDENLEDD	KLAQYALLWL	
PCHAS_070090	-----MLK	DLCGAIDKI-	DTKISLD-AN	DPNLSDDYM	ALNKYCPVNN	GNEQKCKNNE	EIIISAFMTF	LVSF----IS	SNGNEYLDRD	NFAEYITLWL	
PCHAS_000580	-MAESSYDIE	KVCEAINKI-	DHLSVSN-PD	DSLNSDFP	PLINACPVNN	DSEHNCKNNE	ELVSSAFITM	LVSF--ISSN	HGEYLD--RD	KFAEYAILWL	
PCHAS_140130	-----MTE	DMCRSFISVW	TDFPDQL--D	GGNYKFSDDG	YCEALFT---	--DNKCNDF	DKINAVCFWL	FKKI-MWDSS	SSSIKEKINT	DIHYIMTWL	
PCHAS_140040	-----MTE	NMCRGFKTIW	DEFPNSV-GK	DGDNQSNINN	YCETLFN---	---KKIYDDI	DIVNAVSWL	FEQN-FGGGS	SLTKNVNSNT	NIFHYIMIWL	
PCHAS_000740	-----MAD	FMCRMFYNVW	DDFPDQL--D	NGNYKFNDDE	YFDNLFT---	--DKNYGNDI	DKVNAISFWL	FEKN-LWDSS	SSSINAKSNT	DIVHYIVIWL	
PCHAS_100030	-----MAE	NMCRSFYNVW	TDFLDQL--N	EGDYKFNDDE	YSTNLFT---	---NKSDNDI	DKVNAVSWL	FEQN-LWDSS	SSSINAKSNT	DIVYYIVIWL	
PCHAS_001120	-----MAE	LMCQRFNNVW	TDFLDQL--D	EGKYKFTDDG	YSDNLFT---	---NKSDDDL	NKINAVSWL	FEQN-LWDSS	SSSINEKRNT	DIVYYIVIWL	
PCHAS_011520	-----MTE	NMCRRFKTVW	GDFPDQL--D	NGNYNFKDDT	YFKDFFT---	--NKSCDNDI	DKVNAISFWL	FIQN-FGDGS	SLTKDVNSNT	NIFHYIMIWL	
PCHAS_070030	-----MAE	LMCQRFNNVW	NDFPDQL--D	NGNYKFNDDE	YFKNFFT---	--NKCCDNDI	DKVNAISFWL	FKQN-LWDSS	SSSINAKSNT	DIVHYIVIWL	
PCHAS_073190	-----MAE	NMCRRFKTVW	GDFPDQL--D	NGNYNFKDDT	YFKDFFT---	--NKSCDNDI	DKVNAISFWL	FKQN-FGDSS	SSSINAKSNT	DIVYYIVIWL	
PCHAS_000570	-----MAE	LMCRKFNNVW	TDFPESV-GE	DGNCQSITSD	YWETLFN---	--NKKCGNYI	DQVNAVSWL	LVQN-FGDGS	KFTKNVISDT	NIFHYIMIWL	
PCHAS_114730	-----MAE	NMCQRFKTVW	DDFPEGV-GE	DGNSQIKIKE	YWDILFK---	--NKNYDNYI	DKVNAVSWL	FVQN-FGDNF	SFTKNVDSNT	NIFHYIMIWL	
PCHAS_010020	-----MSE	NMCKYFKLVW	DDFPDQTV--R	DGKYQFNDHN	IYSTYCD---	NNTDKCQTDI	DKINAGYLSF	FKTL-FGDSE	SIKNHEKNNM	NISQYIIIWL	
PCHAS_000280	-----MNE	DMCKIFEMMW	EDFPDTL-GN	DGNYQFKNED	IYSKYCD---	SNCDNCRDLD	DKINAACFVL	FKIF-FEDSE	SFMKNAKCNM	NIAQYIIIWL	
PCHAS_000490	-----MNE	DMCKIFEMMW	EDFPDTL-GN	DGNYQFKNED	IYSKYCD---	SNCDNCRDLD	DKINAACFVL	FKIF-FEDSE	SFMKNAKCNM	NIAQYIIIWL	
PCHAS_000030	-----MNE	DMCKFFEVDW	TDFPDTL-DD	DGNYQFKSDD	TYQTYCI---	--NETCNNDI	DKINAWCLCL	LNLI-FGNFE	SFTGYAKSNM	NNVQYILAWL	
PCHAS_000090	-----MNE	DMCKFFEVDW	TDFPDTL-DD	DGNYQFKSDD	TYQTYCI---	--NETCNNDI	DKINAWCLCL	LNLI-FGNFE	SFTGYAKSNM	NNVQYILAWL	
PCHAS_000270	-----MNE	DMCKFFEVDW	TDFPDTL-DD	DGNYQFKSDD	TYQTYCI---	--NETCNNDI	DKINAWCLCL	LNLI-FGNFE	SFTGYAKSNM	NNVQYILAWL	
PCHAS_000120	-----MNE	DMCKFFEVDW	TDFPDTL-DD	DGNYQFKSDD	TYQTYCI---	--NETCNNDI	DKINAWCLCL	LNLI-FGSFE	SFTDHAKSNM	NNVQYILAWL	
PCHAS_001090	-----MFM	NLCDAINGI-	DDLIAKAE-VK	GEEIYIENNR	ILNANCPYIT	EEGGQICIGYI	AAVISAFIGM	LKTF-----E	NSDIDVSENG	EFAQYAILWL	
PCHAS_011510	-MAQSNHTIK	DVYDDIFKI-	NDYFYET-EQ	GQLMLDKTHG	SIEDYCYGNP	PEKGNCDNYL	KMASSGVIHL	LKNY-----	--KKYGLGHD	KLAEYAILFL	
PCHAS_070050	-MTNTSDNLK	DVYRKFATI-	DDYFYVT-DD	GQVRVDTQHK	SIIQYCNKNN	SVTYDCNDYF	QLVSCGFIYL	LKTL-----	-KDNYKLEDD	KLAEYVILWL	
PCHAS_000720	-MTNPSYDIE	KVYRDIYAI-	EDYFYVT-ND	DQIRVDRKHK	SIIQYCNKNN	SVTYDCNDYF	QLVSCGFIYL	LKTL-----	-KDNYKLEDD	KLAEYVILWL	
PCHAS_070020	-MENS SYDIG	DVYRDIYAI-	EDYFYVM-DD	GEIRVDIAHE	SIIQYCNKNN	SVTYDCNDYF	QLTSCGFIYL	LKTL-----	-KDNYKLEDD	KLAEYVILWL	
PCHAS_104260	-MENS SYDIG	DVYRDIYAI-	EDYFYVM-DD	GQVRVDRQHQ	SIIQYCNKNN	SGPYNCNDYF	QLTSCGFIYL	LKTL-----	-KDNYKLEDD	KLAEYVILWL	
PCHAS_140020	-MENS SYDIG	DVYRDIYAI-	EDYFYVM-DD	GEFRVDRQHE	SILKYCNKNT	SGTYKCDNYI	EMTSCSVIYL	LKTL-----	-KDNYKLEDD	KLAEYVILWL	
PCHAS_030060	-MAESSYKIE	DVYSTIKTI-	SNYFYED-EK	DGTKTQVTNE	EIHNYCDEDR	SDSNKCLDYF	EMTSSGVIHL	LKDL-----	-KKKCNLDDD	KLAEYAILWL	
PCHAS_000750	-MTESSDNLK	DLYSAIKTI-	SNYFEGGNG	QLTVNKKYAD	LIDKYCDGDK	SEKGNCFNEHF	KIISGVIINL	LDNL-----	--KKYKLEDD	KLADYAILWL	

## Supplementary data 1 (CIR alignment)

PCHAS_000680	-MTKSSYDIE	K-YKEFVTI-	NSYFGEEKKQ	NGEISLKCKE	EIHNYCHGNT	SEKGNCFNEF	KMTSSGVIYL	LENL-----	--KKYGLDDD	KLADYAILWL
PCHAS_000420	-MRNPSYKIE	DVYKEFVTI-	SNYFDED-KN	DGTKTKVTNE	AIHNYCHKDR	SDSNKCFDYY	EMTSSGVIYL	LENL-----	-KKKCNLDDD	KLAEYAILWL
PCHAS_000100	-MRNPSYKIE	DVYNDIFKI-	SNYFDED-EN	DGTKTKVTNK	AIHNYCHDDR	QEKDKCNDYY	EMTSSGVIHL	INNLL-----	-KGKNVLDYD	KLAEYAILWL
PCHAS_000310	-MRNPSYKIE	DVYKEFATI-	DGYFYVD-ED	DGSITQVTNK	AIHNYCDDRQ	ENGDKCSGY	EMTSSGVIYL	LENL-----	-KKKCNLDDD	KLAEYAILWL
PCHAS_040040	-MPNPSYNIE	DVYKEFATI-	DGYFYVD-ED	DGSKTKVNNQ	AIHNYCDDRQ	EKDKCCSGY	EMTSSGVIYL	LENL-----	-KKKCNLDDD	KLAEYAILWL
PCHAS_120060	-MRNPSYKIE	DVYKEFVTI-	SNYFNED-EN	DGAKTKVINE	AIHNYCYDKQ	LEKDKCSGY	EMTSSGVIHL	INNLL-----	-KGKNVLDYD	KLAEYAILWL
PCHAS_104230	-MRNPSYKIE	DVYKEFATI-	DGYFYVD-ED	DGAKTKVINK	AIHNYCDDK	REKDKCSGY	EMTSSGVIHL	INNLL-----	-KDKNVLDYD	KLAEYAILWL
PCHAS_030040	-MRNPSYKIE	DVYKEFVTI-	NGYFDDD-ED	DGTQTKVNE	AIHNYCDDK	KNNDKCSGY	EMISSGVIYL	LENL-----	-KTKCNLDDD	KLAEYAILWL
PCHAS_000340	-MPNPSYNIE	DVYKEFATI-	DGYFYVD-QD	DGSITHVTNK	AIHNYCDDRS	RN-CKCCGY	EMTSSGVIYL	LENL-----K	KKCNLD--DD	KLAEYAILWL
PCHAS_001110	-----MSE	EVCEAFKKV-	DDCLQIG-MV	STGDICYIDE	VLNEYCPNIE	GQNGKCDTNN	EKISAGFIWL	LITF--EDLC	EGQCSDNENE	KYAEYAILWL
PCHAS_120040	-----MSE	EVYKEIKKV-	DDSLPIG-II	STGSSSSTR	VLIDYCP-KK	NGERQCETIC	DKISAGFIWL	LISF--ENLC	EGKCSDDENE	KYAEYGILWL
PCHAS_000130	-MEESSNLK	DVCNFFNKIV	DDCLPTG-II	STGVSCSND	VLNEYCPKRE	GGNRQCDTNN	DKISAGFIWL	LISF--ENLC	DSQCSDDENE	KYAEYGILWL
PCHAS_000400	MAESSYKIED	VVCEAFKKV-	DDCLQIG-MK	STGGSCSIDY	AFTDYCPKIE	GQNVNCEANN	EKMSAGFIWL	LITF--ENLC	VGQCSDNENE	KYAEYAILWL
PCHAS_040020	-MAESSYDIE	KVCEAFKKV-	DDCLQIG-MK	STGGSCSIDY	AFTDYCPKIE	GQNVNCEANN	EKMSAGFIWL	LITF--ENLC	VGQCSDNENE	KYAEYAILWL
PCHAS_000110	-----MSK	EVCEEINIM-	GKYIVVK-QK	DSGVDIEFDP	KLNDYCPKNN	GINEQCETND	EKISAGFIWL	LVMF--EHL	DEECPQNEK	RYVEYAILWL
PCHAS_120050	-MENSDDLK	YVCKDIKKI-	DDCLQID--T	IYTGVCSDT	LYSDYCPMKN	GKKGQCETNN	DKISAGFIWL	LVMF-EHICD	DDECSQNEK	QYAGYAILWL
PCHAS_104250	-----MFQ	EVCEEINNM-	GKYIVVK-QE	GSGVNIIEFDN	YLSDYCPKNN	GRNGQCETND	EKITAGFIWL	VVMF--EYLC	EEKYPQNEK	QYVEYAILWL
PCHAS_000320	-----MSK	GVCEEINNM-	GKYIVVK-QE	GSGVNIIEFDN	KLNDYCPKNN	GRNGQCETND	EKISAGFIWL	LVMF--EHW	DEECPQNEK	RYVEYAILWL
PCHAS_000410	-----MSK	GVCEEINNM-	GKYIVVK-QK	DSGVDIEFDN	ILNDYCPKNN	GRNGQCETND	EKISAGFIWL	LVMF--EHW	DEECPQNEK	RYVEYAILWL
PCHAS_140030	-----MNE	NMCRILEGVW	GDFPDTL-DG	AGNYQFKNDE	INATYCK---	DNIENCKTDL	DKINVGCFVL	FEIF-FKDS	SLMENAKNNI	NIAGYILAWL
PCHAS_114740	--MGLMYLCK	YLCKILSFIW	EEFPDTL-GD	DGNYQLKSDG	AYKTFCI---	--NETCSDL	DKINAWVLSL	FEIF-FKNS	SLMESAKSNI	NIAAYILAWL
PCHAS_040030	-----MIK	EVCGTISII-	DKFIWVE-QK	GVVEVIQYYN	LVNAYCPKST	PKKNECHSYD	EMISSAVLFF	LKWL---ETS	YDYKDDLKND	KLAEYAILWL
PCHAS_070130	-----MAS	AVCNAIKAI-	DNFIVVK-EG	NIGVNISFKE	ILNPYCTKSL	VNKEECQSYN	EMVSSAFILL	LKFL---NLV	DVYDGLTND	RLAEYAILWL
PCHAS_073180	-----MAF	GVCGAIKAI-	DKFFSVK-QK	NSVDVIEHYN	LINTYCPKNN	PRKNECHSYE	EMVSSAVIYL	LKYF---EFD	YDYKDDFKND	KIAEYAILWL
PCHAS_000300	-----MAS	GVCGAINGI-	ERLIDVK-EE	DSRVSFKCNV	VLDAYCPKST	SRKNECHSYA	EMVSSVFLFL	LKSL---ENN	YHYEDDLKNA	KLAEYAILWL
PCHAS_120070	-----MAS	GVCNAIKAI-	EKFIVVK-EE	DSGVYFTPNI	ALKAYCNAKE	GGTGVCFSYA	EMVSSVFLFL	LKSL---ETS	YDYGDYKND	KLAEYAILWL
PCHAS_001130	-----MAY	GVCNAIKAI-	EKLITVK-EE	DSRVYFTPNI	ALKAYCNVKE	DGIGVCFSYA	EMVSSAVLFF	LKWL---ETS	YDYGDYKND	KLAEYAILWL
PCHAS_100040	-----MAS	AVCGAINTI-	ERLIDVK-EK	DSRVYFTPNI	ALKAYCNVKE	RGTGVCFSYA	EIVSSAVLFL	LKLL---EID	YHYEDDLKNA	KLAEYAILWL
PCHAS_000560	-----MAY	GVCGAINVI-	EKLIVVN-EK	DLGVYFSPNQ	LLKAYCTRVN	DQTGICFSYA	EIVSSAVLFL	LKLL---EID	YHYEDDLKNA	KLAEYAILWL
PCHAS_050020	-----MNK	DLCDVIKGI-	DDLIEVK-RN	TEAIEIRST	LFDTYCPKRD	GQEGHCIGYS	ETVISAFIHL	QETL----KI	NYSPEELERD	KLAQYAILWL
PCHAS_000020	-----MNK	DLCDVIKGI-	DDLIEVE-VK	AEGIETIRDE	LFNTYCPROQ	GQDGHICIGYS	ETVISAFIYL	QETL----KS	NYSPEELES	KLAQYAILWL
PCHAS_000290	-----MNK	DLCDVIKGI-	DDLIEVE-VK	AEGIETIRDE	LFNTYCPROQ	GQDGYCIGYS	ETVISAFIHL	QETL----KN	NDSQKKLDRD	KLAQYAILWL
PCHAS_060020	-----MSK	EVCKAFKDI-	DKLIKVE-SS	GEDVFIENNK	ILNAYCPYVN	GEGGQCIGYI	EIVNSGFISL	LEKF----NN	FDNGRNLESE	RRFQYAILWL
PCHAS_001040	-----MSK	EGCTYINFV-	NKLIKSN-TT	GTEGKIENHK	SLNTNCP---	--NEKCDTDS	QKLSSAFILL	IKFF-----	-SIDDNLEND	KHAEYAILWL
PCHAS_114750	-----MSN	EVCMYINKV-	DKLIKSN-AA	GTEGKIENHS	RLNANCP---	--NKNCDTDA	HKLSSAFILL	LKYF-----	-EVADDLKDD	KCAEYAILWL
PCHAS_000140	-----MSN	EVCMYINKV-	DKLIKSN-AA	GTEGKIENHS	RLNANCP---	--NKNCDTDA	HKLSSAFILL	LKYF-----	-EVADDLKDD	KCAEYAILWL
PCHAS_000070	-----MSN	EVCMYINMV-	DKLIKSN-AA	GTEGKIENHN	RLNANCP---	--NKNCDTDA	HKLSSAFILL	LKYF-----	-EVADDLKDD	KCAEYAILWL
PCHAS_030080	-----MFK	EVCEHIIHKI-	DTLIKLE-VT	DV-SYVEYDD	MLKSYCPKKN	YNTIECFGYD	EIVSAAFIF	LENF----NS	SNDEEKFE	KLAEYAILWL
PCHAS_073200	-----MSK	GVCEAINSV-	NELINVK-ME	TEGTYIYYND	ILKANCPKAV	LKTNECHGYE	EIVGSTFIAL	LTLF---KSN	DDEVDAL	KLIEYAILWL
PCHAS_120020	-----MSK	KVCEAIKSA-	NELFNVE-KD	GSNTNINYDK	SLKAYCPPNH	LNNQECVHYE	QVVASAFIAL	LTLF--KNDD	DDEEDVLEDD	KLAEYAILWL
PCHAS_000060	-----MSK	KVCEAIKSA-	NELFNVE-KD	GSNTNINYNE	SLKAYCPPKN	LTKQECTNYE	QIVSSAFIAL	LTLF-RNFDD	DGDEDVLEDD	KLAEYAILWL
PCHAS_130020	-----MSK	KVCEAIKSA-	NDLFKVE-KN	DLGSFIKYDD	SLKAYCPPNH	LNKQECTYEE	QVVASAFIAL	LTLF--KNDD	DDEEDVLEDD	KLAEYAILWL
PCHAS_000150	-----MSK	KVCEAIKSA-	NDLFKVE-KK	GLETNIIYDD	SLKAYCPPKN	LNNKECVHYE	AIVSSAFIAL	LTLFKNGDDD	DEEEDVLEDD	KLAEYAILWL
PCHAS_000500	-----MIS	EVCGVIKEV-	DKCLSKT-IL	STGDECLDEL	GYTAYCPAKN	GKLGECRTNG	GKLSAWFIWL	LEMF--KTLT	NVDDFKDING	QYVEYAILWL
PCHAS_001100	-----MLE	EVCGTIKKI-	DDCLSRD-IL	STGDKCSDAE	LYTVYCSKKE	GVIGKCETNG	GRISAGFIWL	LVMF---DV	LCGGECSEK	QYDEYAILWL
PCHAS_042010	-----MHH	EVCKAFRKA-	DYAFKDK-KP	DLDVINSQNG	PYVQDCPDNK	SKSHRCKNDL	EGMYVLCCLHL	FHDL-FNLPO	ETLKRENDN	QYVEYIMMWL
PCHAS_000040	-----MLE	EVCGAINQI-	NKVLSDDILT	LGQHFPPDEL	GYTAYCPAKE	GEKGCVTNG	DRISAGFIWL	LEMF--KALD	DVENLKDIND	QYVEYAILWL

## Supplementary data 1 (CIR alignment)

PCHAS_000390	-----MIE	V-CGVINQI-	NKVLSDDILT	LGQYFPDDEL	GYTAYCPAKE	GEK GKCVTNG	DRISAGFIWL	LEMF--KALD	GVENLKDIND	QYVEYAILWL
PCHAS_030180	-----MSK	EVCEAIKFA-	DEIIVFD--K	KIKNYTFKND	IFKVYCPQRK	GPNRKCDSGG	KILGAVFVAL	LNFF---GSV	EDYEENLKND	SLSEYAILWL
PCHAS_130070	-----MSK	EVCEAIKFA-	DEIIVFD--K	KIKNYTFKDE	IFKVYCPQKR	GPNRKCDSGG	KILGAGFIAL	LKTL---DGV	DDYEENLKND	RLSEYAILWL
PCHAS_011500	-----MSK	QVCEAIKEI-	DENVIFN--S	VSEKYEFKDK	MYGAYCANRK	GGKGECDSDG	KILGSAFISL	LEIL---DGV	DDYEELKND	RLSEYAILWL
PCHAS_146850	-----MSK	EVCKVINDI-	EKNILFD--S	GSQKYEFKND	IFKAYCPRKR	GGKRQCDSNG	KILGAAFIAL	LKNL-DSVED	YEENLK--KD	ILSEYAILWL
PCHAS_070100	-----MSK	EVCEAIKGI-	DENVVFD--P	VSQKYAFTDE	IYGVYCSTRK	GEKGQCDSNG	KLLGSAFIAL	LNIF---DSF	GDEEEKLENN	KLGEYAILWL
PCHAS_060050	-----MSK	ELCDAIDFT-	DKNVVD--Q	KSKNYTFKDD	IFKVFCSSGK	GGKGC-NSA	QKLSSGFMGL	LEYF-----K	SIEDENLDGD	KLAQYAILWF
PCHAS_030110	-----MSK	GVCEQINFV-	DENIVFD--Q	KNKKYTYKDK	IFETYCSSRN	GEKGQCNDG	ELLSSAFIAL	LKNF-----E	NVDENKLESD	KLAQYAILWF
PCHAS_130050	-----MSK	ELCKAIDEI-	DKNVVFN--T	VSQKYTLEDK	IYGAYCPSEN	GEKGQCDSDE	ELLGSAFIAL	LKFF-----K	NIDNGHLEDD	KLAQYAILWF
PCHAS_030090	-----MSE	ELCGAIKFV-	DENVVFD--K	NTQNYTFKDE	IFKAFCSGK	DGK GKCDTNR	EFLNSGIMGL	LEYV-----K	SIDEEELDGD	KLAQYAILWF
PCHAS_146870	-----MSK	ELCGGIDGI-	EDFIVFD--S	GSQNYKFKDD	ILNAYCT---	--NNNCDSDG	KKLGSAFIAL	LKYF-----N	SIDNVNLEDD	KLAQYATLWF
PCHAS_130030	-----MSE	ELCGGIDDI-	EDFIVFD--S	ESESYRFNDD	ILKVYCSKEN	GVKGECDSDS	LKLSSAFIAL	LEYF-----K	SIEDKNLDDD	KRAQYAILWF
PCHAS_040110	-----MSK	ELCOGINFA-	DKNVVD--Q	KSESYKFNDD	IFKVYCP---	--NNNCDSDG	LKIGSAFTAL	LEYF-----K	NIDNENSEDD	KLTOYAILWF
PCHAS_042030	-----MSE	ELCEGIKFA-	DENVVFD--S	ESQNYTFKDE	IFKVFCSSGK	GEKGQCDSNG	LKVSSGFMGL	LEYF-----K	SIDEEGLDGD	KLAQYAILWF
PCHAS_060140	-----MSE	ELCDGIKFA-	DENVVFD--S	ESQKYTFKDK	IFDAYCPSEN	GGKGCDSG	LKVGSFTAL	LEYF-----K	KIDGENPEGD	KLAQYAILWF
PCHAS_030210	-----MSE	ELCKAIDEI-	EDFVVD--S	ESKKYTFKDK	ILEAYCSSGK	GENGCDSDE	LKLSSGFMGL	LEYF-----K	SIEDENLDDD	KRAQYAILWF
PCHAS_060060	-----MTE	ELCEAIKFA-	DENVVFN--S	ESQDYTFKDD	IFKAFCS--L	SGKGCASDE	LKVGSFTAL	LEYV-----K	SIDEEELDGD	KLAQYAILWF
PCHAS_050070	-----MSK	GVCEQIDQI-	EKYITFD---	SGSQNYKFND	ILKAYCP---	--NKNCDTDE	KILGSAFTAL	VENL-----K	SIDDEKPEDI	KLQYAILWL
PCHAS_040050	-----MSK	GVCDAINEI-	DKNVVD--S	ASQNYKFDEE	IYGAYCP---	--NKICDTNE	KLLGSAFMSL	IELF---KNV	DDYDDNFKND	KLSQYAILWF
PCHAS_001050	-----MSK	GVCDAINEI-	DKNVVD--S	ASQNYKFVDQ	IYGGHCH---	--NKNCDTDE	KILGSAFMSL	LELF---KDV	NDYEDNLKNY	KLAQYAILWF
PCHAS_083760	-----MSK	ELCEKINFA-	DENVVFD--P	ESQNYKFNDE	ILKAYCP---	--NKNCDSES	LKLSSSFVGL	LENF-----R	SIDNEKSEDD	KLSQYAILWF
PCHAS_140070	-----MSK	ELCGAIKFA-	DENVVFD--S	KSERYSFKDD	ILKTYCP---	--NQICDGDG	KLLGSAFTAL	LYYF-----K	SIDEEKLEDD	KLAQYAILWF
PCHAS_000180	-----MSE	ELCDAINAI-	DQNVVVD--S	KSQKYKFENQ	IYGAFCP---	--NKNCDTDE	KTLGSAFMSL	LKLF-----K	SIDENKLESD	KLSQYAILWF
PCHAS_001060	-----MSE	ELCKAIDEI-	DKNVVFN--T	VSQTYTLEDK	IYGAYCPSGN	GEKGQCSDS	LKLGSFMSL	LNNF-----K	SIDKLESD	KLSQYAILWF
PCHAS_011450	-----MSE	ELCEKINQI-	EYITFD--S	TSQKYMVKNE	ILNSYCPSEN	GGKGCDSNE	ELLGSFMSL	LVNF-----K	GIEDEKLEDD	KIYQYAVLWL
PCHAS_060030	-----MSE	ELCKQIDQI-	EYITFD--S	ESQNYTSLDG	ILKAYCT---	--NQKCDSDS	LKLGSFMSL	LENF-----N	SNDGEDPESD	KLYQYAILWL
PCHAS_120030	-----MSE	ELCGAIKFA-	DENVVFD--S	TSQNYNFNDG	ILKAYCPKDK	DGNQCSDS	LKLGSFMSL	LNNF-----K	SIDGENPEGD	KLYQYAILWL
PCHAS_030120	-----MSK	ELCKAINEA-	DEKIVLD--L	ESKDYMFKDK	ILEAFCPKKN	GGKRECDSDG	LKVSAFIAL	LKYF---KSI	DDYEDNLKND	KLAEYAILWL
PCHAS_114640	-----MSS	GVCEAINEV-	DKYIVLD--P	ESKDYTFKDN	ILEAYCPKKK	GGKRECDSDG	LKVSAFIAL	LKLF--KSD	DEKLE--SD	KLAEYAVLWL
PCHAS_130350	-----MSY	HVCEMFNSL-	YSQLPDD-DN	EKQRNGEKL	LYESFCPKIG	ETYQCSSEF	EKISAGFVYL	LGQL-FGNVN	SEGEHEDQK	HYVYGFLLWI
PCHAS_030150	YTIKFFYKNV	FQCMFLEI-	NRQLPDE-TN	DLGEYIYNDE	VYKAYCPNEL	TGDHTCNSNI	ERINAGCILL	LNML-FNIDD	DILESENQDT	SFAQYIILWL
PCHAS_060120	-----MSA	HVCEIFGNLY	MQLPDGD--D	EVNLGNGASL	LYESYCPNND	GQHGTCNTDY	DKIGAGFGYL	LTQL-FNNVE	SEEEHENQK	NYVHYGLMWI
PCHAS_041960	-----MSA	HVCRIFQSLY	NELPDDG--D	EPKRENGAVV	LYETYCPIND	SGNQCNNSL	DKIIAGFGYL	LPQL-FGNVD	NEEDHEDQED	NYVYALLWV
PCHAS_130110	-----MSA	HVCEMFDLS-	YKQLPDN-GD	EANQSDDTIL	LYKTYCPING	DEPGKCNNSL	DKISAGLGYL	LPQL-FGNVE	SEEDHEEQEG	NYVYGLLWT
PCHAS_041950	-----MSY	KACNIFRDV-	DNQFNGK-SI	NVEQFNTNSG	LYAQYCP-VK	QGSKRCVTDY	EKLNAISGHA	FMEL---TKN	NKINLYSGHD	PSIDFLVMGW
PCHAS_060130	-----MGM	K-CETFLDV-	DKLFINY-KA	NEEQFNTSSG	LYYQYCP-VK	NGLRTCCTDY	EKLSAIFGYA	LMEL---AKN	PKMDLYSEYD	PSIAFLVMGW
PCHAS_130120	-----MGM	KVCETFLDV-	DNLFINY-EA	NEEQNLGGYG	SYKKYCP-IK	RGVRKCTNY	EKLRAISEYG	FMEL---AKN	DKVNLGSEYD	PSADFVVMGW
PCHAS_070170	-----MEI	E-YNLFLV-	DELFDK-SV	DV-VQFNANR	KVAAYCP-EK	NGYRSCNDY	QRINAIGSHL	HMEL--DTKD	PNIYGGGNK	RFMEYLIMWL
PCHAS_100060	-----MEI	E-SNLFLGV-	EELFDGK-SV	DV-GKFNANY	QVPTYCP-EK	KGYRSCNDY	QRINAIGSHL	HMEL--DTKD	TNIYGGGNK	RFMEYLIMWL
PCHAS_146770	-----MAN	KACGYLLDA-	DEYFNNE-VV	DE-NKFNKND	SLKNRCPYEN	KVSRPCQNNY	ERINALAVYL	YQNL-NSTST	SLSGKGHNGD	RHIEFFMMWL
PCHAS_146790	-----MAS	RACKFLLEA-	DEYFNNG-IV	DE-DKFKKNT	ELHSYCPYDN	GKRRPCKNNY	ERINALGVHL	YQKL-NKTTK	NLNGTGHHEN	RHIEFFMIWL
PCHAS_060160	-----MAN	KACSYLIEA-	DEYFNNG-VV	DE-NKFNKNS	SLQDRCPYEN	RSPRPCKNNY	ERINAVGAYL	YQNL-SKTGK	NISGTGHNGD	RHIEFFMMWL
PCHAS_000260	-----MAI	RACKFLLEA-	DEYFNNG-VV	DE-TKFNKNS	SLENRCPYEN	KRARPCNNY	ERINALAVYL	YQNL-TKTTK	NLNGTGHHEN	RHIEFFMIWL
PCHAS_130280	-----MAN	KACSYFIEA-	DEYFNNG-IV	DE-NKFNKND	SLKYRCPYEN	RTLRCQNNY	ERINAVGAYL	YNNL-LNVGS	SFKGIGNNDN	RHVEFFMMWL
PCHAS_083720	-----MAS	RACKFLLEA-	DEYFNNG-VV	DE-IKFNKND	SLENRCPYEN	KKVRPCKNNY	ERINALAVYL	YQNL-TKTTK	NLNGTGHHEN	RHIEFFMIWL
PCHAS_090010	-----MSS	KLCKLLLDV-	DDLFTDG-TV	DV-NKFNKLD	SYKKYCPYEN	GKLRDCKNNY	ERINALGEYL	YQKL-PKNSD	EFKGEGINNN	LYIEFFMMWL
PCHAS_114600	-----MSS	KLCKLLLDV-	DEYFNNG-IV	DE-GKFDKKN	SLQSKCP-QK	GQENKCTNY	ERINALGEYL	YQKM-PKNSN	EFKGEGINNN	LDIEFFMMWL

## Supplementary data 1 (CIR alignment)

PCHAS_146860	-----MNY	SACDLFFEVE-	DELIDDK-SV	DVDSFNDLSS	YFNEYCP-EK	NGSKICSTDY	ERMIAVGAYL	FMEF---INE	NNIDLNSDKE	YHSEYFIMWI
PCHAS_130090	-----MSY	KTCDFREV-	DELFDGK-SI	DVDNFNAFSS	LYNEYCP-VK	NGSKSCDTEY	ERIAAVAGYI	FMNF---ILD	NSINLHSDDD	RHLEYFIWI
PCHAS_030170	-----MDT	ESYKLFREV-	DKLFDEK-SV	NV-GKFNTYQ	DVADYCP-EK	NGYKSCDNDY	ERINAIGGHL	IKQL--YFKN	SSINGGNNNN	RFIEYFIMWL
PCHAS_130060	-----MDT	ESYKLFLEV-	DKLFDDQ-SV	KV-ANFNAYH	NVSDYCP-EK	DGYKSCDNDY	QRINAIGAYL	ITQL--YLKN	SSINGGNNNDK	RFIEYFIMWL
PCHAS_030200	-----MSY	NACKLFSSEL-	DELFNNS-SV	DEDSFNGLSV	LYNEYCPQDK	NGDKKCITDY	DRISAVAGYL	FMYF---ITE	NGIDINNNND	YYIQYFVMWI
PCHAS_060080	-----MSY	NACNLFHGL-	DELFEEG-FG	DVKELDSL SI	LFNDYCP-ER	GESKRCDDY	ERISAVAGYL	FMNF---IAD	NNIDINNDNN	YYIQYFVMWI
PCHAS_060110	-----MDT	KLCNLFINV-	DKLFTKG-NV	NE-TLFNTSN	LYKKYCP---	--NGKCNTNY	DRIGALCEYL	LAEI-PKLN	KPNGSKDNVN	QYVEYVFMWL
PCHAS_030140	-----MET	KLCKLFDV-	DKLFTKG-NI	NV-NKFNSD	SYKKFCP---	--KGGCNTNY	DRIGALCGYL	LAEI-PKNYD	KQKGGNNNGN	LDYEYIMWL
PCHAS_130100	-----MFP	FDCKLFDV-	DKLFTKG-NI	NE--QIFNK-	SYKQFCP---	--KGCNNNY	DRIGALCEYL	LAEI-PKND	KLKGGNNNGN	RDYEYIMWL
PCHAS_041990	-----MDT	KLCKLFDV-	DKLFTKG-NV	NE-TLFNTSN	LYKKYCP---	--NGKCNTNY	DRIGALCEYL	LAEI-PKLN	KPNGSKDNVN	QYVEYVFMWL
PCHAS_130080	-----MDK	NVCKLFDV-	DEVFNHG-TV	NE-SIFNTSN	LYKKYCP---	--NGKCNTNY	DRISALCDYL	LTEL-QKYDK	KQNDSDNDVN	QNYEYVFMWL
PCHAS_030190	-----MDK	NVCKLFDV-	DKLFTKG-SV	NE--TLFN-D	PYKNFCP---	--KGGCNTNY	DRIGALCGYL	LAEI-PKLN	NPKGSEDNAH	QNYEYVFMWL
PCHAS_060090	-----MDK	NVCKLFDV-	DKLFTKG-SV	DE--TLFN-D	PYKNFCP---	--KGGCNTNY	DRIGALCGYL	LAEI-PKLN	NPKGSEDNAN	QNYEYVFMWL
PCHAS_041970	-----MNK	NVCKLFDV-	DKSFKKG-TG	NV---NTFDN	TYTKYCP---	--EEGCNTNY	DKISALCEYL	LEEL-PKIDD	KQKGEHNVY	RDYEYVFMWL
PCHAS_041980	-----MDK	NVCKLFDV-	DKLFTKG-SV	NE---TLFND	SYKNFCP---	--KGGCNTNY	DRIGALCEYL	LEEL-SKND	KQKGDNNNVN	QNYEYVFMWL
PCHAS_042020	-----MYR	DVCKIFQEV-	DEVFKDN-KP	DL-NKIFED	KYRRYCPWDI	TGVNQCQHDV	AGIDAVYSYL	FKEI-HQLPL	DQOKYENNDN	QYTEYMLIWL
PCHAS_060070	-----MSK	EVCKIFKDA-	DNFFKGG-NP	DLDKINKEIG	ENIRDCLGNK	SDTYTCKTDF	DGINALYCHV	VTEL-HKIPI	ETQEHENNDY	QYIEYAIWI
PCHAS_011330	-----MDN	EACKLFRMV-	NANFNKE-IV	DL-DKFKQFT	VCHRYCPGKK	QQRDCNNNY	ERINAIGGYL	RSQL-SGVKK	EIQTRENNND	QYIEYMLIWL
PCHAS_000220	-----MDN	EACKLFRMV-	NANFNKE-VV	DA-DKFKQFT	ECHRYCPGRK	KQRDCNNNY	ERINAIGGYL	RVKL-SSIKN	DIQTRENNND	QYIEYMLIWL
PCHAS_042070	-----MAI	NACKLLELV-	DGYFKNE-NV	DE-EKFNKNN	SLVYRCPYED	RKLRPCENNY	ERINALGVYL	YEKL-GKIAN	DLKQKGNNDN	RHIEIFIMWL
PCHAS_110030	-----MAD	KACTLLREV-	DAYFNNE-NV	NE-EKFNNSG	LFTYRCP-RK	GKEYICTTNN	ERINTLGVYL	YENL-NKISK	DFKGEANEAN	RHIEIFMMWL
PCHAS_140140	-----MAI	EACKILRDV-	DGYFKDE-IV	DE-SKFNNSG	LFTYKCPKQN	RRFRPCENNN	ERVNTLGVYL	YEKL-NGIAN	KLNGEENAN	RHVEIFMMWL
PCHAS_010040	-----MAG	KACKLLRDV-	DAYFNNG-IV	DE-NKFNSSG	SFIYKCPYEN	RKPRGCKNNN	ERINALGGDL	YDKL-ARIAN	DFKGEANGN	RHIEIFMMWL
PCHAS_073150	-----MAS	EACKILRDV-	DGYFKDE-IV	DE-SKFNNSG	LFTYKCP-RK	GKESKCTTNN	ERINTLGAYL	YTKL-FKISN	ILKGDGNNAN	RHIEIFMMWL
PCHAS_073130	-----MGY	QMCETFIVA-	DQVINGENAN	ITMDDISKNP	GFKQYCP---	--NQKCESTR	KRISALSTYL	FMQL-----	---RTMKSAG	QYDEYFLMWL
PCHAS_040060	-----MTT	YMCETFLEAD	KIINRENDAR	MKMEDIDKKS	PYYEFCP---	--NRKCVTDV	ORIGVMTTHV	FLKV-----	----KADKNN	EYGEYFLMWL
PCHAS_000430	-----MDP	NMCETFLEAD	KIINGENGAR	MKMEDIDKKS	SYEFCP---	--NKKCLTDV	ORIGVMTTHV	F-----	-LK GKADKNN	EYGEYFLMWL
PCHAS_104200	-----MDP	NMCETFLEAD	KIINGENGAR	MKMEDIRKSQ	SFNGFCP---	--NNKCVTDE	QCIGAMTYV	FSKV-----	----GADKNN	EYGEYFLMWL
PCHAS_114720	-----MDP	NMCETFLEAD	KIINGENGAR	MKMEDIRKSP	SFNGFCP---	--NNKCVTDE	QCIGAMTYV	FSKV-----	----GADKNN	EYGEYFLMWL
PCHAS_000770	-----MDK	HMCEFLIEA-	DGYFNGK-NV	NT-EKINEDP	TIKGYCR---	--NGGCKTNE	DSINALTAYI	IKEF-----	KKSIKSDEYN	KYDECFFMWI
PCHAS_130220	-----MDK	HMCEFLIEA-	DSYFNGK-NV	KM-NEINKHA	TIKGYCY---	--NDDCKTNE	DSINALTAYI	IKEF-----	KTSINQHDYN	KYDECFLMWL
PCHAS_070040	-----MDT	EMCKLFREA-	DSYFNGK-DV	DT-TKFNEHK	TIKSYCR---	--DDGGCKTNE	ERINALIAYI	IMDF-----	KRSIEKNEYN	DYDEYLLMWI
PCHAS_011490	-----MDP	QMCETFLIKA-	DKYFAGN-KV	DI-KEINKNK	SIKSYCH---	--NSDCKTNE	DSINALAAYI	IMIF-----	KRSIKTNEYS	HYDECLLMWL
PCHAS_070060	-----MDS	KMCEFLIKA-	DKYFAGN-KV	DI-KEINKNK	SIKSYCH---	--NSDCKTNE	DSINALAAYI	IMIF-----	KRSIKTNEYS	HYDECLLMWL
PCHAS_030070	-----MDH	KMCKYLNIA-	DSYFKGE-NV	NT-KIINKHS	TIKYCY---	--NGVCKTNE	AGINALAAYI	FKQF-----	KVSIEANEYN	KYDEYLLMWL
PCHAS_137110	-----MDH	KMCKYLNIA-	DSYFKGK-NA	DM-KIINKDP	AIKYCY---	--NGVCETNE	AGINALVAYI	FNQF----KR	SIEANE--YN	KYDEYFLMWL
PCHAS_114700	-----MTT	RLCEFLIEA-	DGYFNGK-DI	DM-KEINKNT	KIIGYCS---	--NGGCKKNE	EHINALTLYI	YMGF-----K	NSMKRQSEYN	KYDECLLMWL
PCHAS_137030	-----MTT	KLCEFLIEA-	DGYFNGK-DV	DM-KEINKNR	KIIGYCS---	--NGGCKKNE	EHINALALYI	FKEF-----K	NSIKRQTKYN	DYDECFLMWL
PCHAS_011480	-----MRK	HMCKLLLEG-	DSYFNGK-DL	DT-NEIIKKT	SIKGYCR---	--NDRCKKNE	ARINALTLYI	HMEF-----K	NLIKKKSEYS	NYDEWLLMWI
PCHAS_050040	-----MNH	ELCGLLIEA-	DGYFNGK-DV	DA-QRINDEP	TIKGYCS---	--NDGCKTNE	ARINALAAYI	LMLF-----	KKSIRKDEYN	DYDECFLMWL
PCHAS_000470	-----MNM	RRCKLLLEG-	DSYFNGK-DV	NT-NEFNKHS	TIKAYCR---	--NDDCKTNE	ERINALTAYI	FKEY-----K	GKTSRRTKYN	DYDECFLMWL
PCHAS_000170	-----MNH	ELCGLLIEA-	DGYFNGK-DV	DA-QRINDEP	TIKGYCS---	--NDGCKTNE	AYINALAAYI	FMLF-----	KRSIQKDEYS	EYDECFLMWL
PCHAS_130170	-----MSK	HLCMLLLEG-	DSYFNDE-NV	DT-EKINKDI	TIKGYCR---	--NGSCKTNE	ESIDALAAYI	FKKF-----K	DSIKVKQRYN	NYDECLLMWI
PCHAS_000660	-----MDS	KMCNLFREA-	DGYFNDE-NV	DT-KKINKNT	KIKGYCS---	--NNGCKTNE	DYINALSAYI	YKEF-----K	NLIKRNQKHN	DYDEYLLMWI
PCHAS_040080	-----MNT	KMCKLLLEG-	DGYFNDE-NV	DT-KKINKNT	IIGYCS---	--NNGCKTNE	DYINALSAYI	YKEF-----K	NLIKRNQKHN	DYDEYLLMWI
PCHAS_070070	-----MQN	HMCNLLLEG-	DSYFNDE-NV	DT-QKFNKHS	TIKAYCR---	--NGGCKTNE	ERINALNAYI	FKTF-----K	DSIVIKKKYN	HYDECLLMWV
PCHAS_137090	-----MSK	HMCKLLLEG-	DSYFNGK-NV	NT-EKINEDP	TIKGYCN---	--NGVCKTNE	ESISALTAYI	FMKF-----K	DSIVIKRKYN	DYDECILMWL

## Supplementary data 1 (CIR alignment)

	.... ....	.... ....	.... ....	.... ....	.... ....	.... ....	.... ....	.... ....	.... ....	.... ....
	105	115	125	135	145	155	165	175	185	195
PCHAS_000950	SNIIRHFNYK	HTPHVQDFYT	TFIKEG----	-----NW	YNDFRDKINK	KNDIMKLQIN	QIYNLYELLN	ILCKAITKYD	ENPSNSSECS	NFANKWEERS
PCHAS_001010	SNIIRHFNYK	HTPHVRDFYA	TFIKDS----	-----DL	YKDFHGKINE	KNEIMKIQIG	QMYNLYELLN	ILCNAITKYE	ENPSNCPECS	KFTNKWEEKT
PCHAS_140090	NHKLNQYPEY	GVSNLDDFHN	TYIKDI----	KETTD--IEG	YNAYMDFIEK	KQDMMNMDIK	DMSKFYASLK	SLCDMQTKID	ANKSNCASYL	EDSQKFVDEY
PCHAS_000730	CHKLNKNSQN	GINNLNDVYN	KYIKGNKEFF	EKIND--AGA	YNSYEEFINK	KQDMVNVNIK	DMSKFYEALQ	ILCKMYTKID	ESKSNCTKCL	DDAQKFADYEY
PCHAS_070160	CYELNKK--N	EISNLNEFYN	KYIKAIEEYV	KGSHA--AEA	YNSCMNIINK	KQYLMVSDIK	EMSKIYEALK	FLCKLYTGCN	DKEKNYSNYS	QDANDFVKEF
PCHAS_030020	CYKINLMSYD	GIRNLNDFHN	KYIKDKESDI	QKMVD--VEA	YNIYMDFINK	NQDLNMMDIN	GISNFYAPLK	SLCNMYNKF	AKNNSCAKCS	EDAKEFVEKY
PCHAS_000360	CYKLNLMSEY	GISNLNDFHN	KYIKDKESDI	QKIVD--VEA	YNIYKDLINK	KGDLVNMMDIK	IMSKFYETLK	FLCKLYTGCN	EKETNYPNCS	QDANAFVKEF
PCHAS_110020	CYKLNKKKEN	EISNLNDFHN	KYIKGNENHI	DKING--AGS	YNSYKNIINN	KKYLNTIDIK	EMSKFYVPLK	SLCKLYTEYS	KKKKNYTTCS	QDAQDFAKHF
PCHAS_050060	FYKLNQKKEY	EISNLNDFQK	KYIKDNENFI	DKING--DGS	YNSYLDIINK	NPYLMSIDIK	EMTKLYVPLK	SICKLYTECD	GKKKSYTNCL	QDAQDFAKHF
PCHAS_010030	CHKLNKNSQN	VVSNLKDIDY	VYIKDYEKDI	EKISG--ADA	YNSCKDIINK	KIYSMRIDIK	EMSKLYEALK	TLCKLYTECN	GKKQNTKCL	QHAEFVKSF
PCHAS_030030	CHKLNKNSQN	GVSNLKDIDY	VYIKDNKKDI	EKMSG---EA	YNSCKDIINK	KIYSMPIDIK	EMSRLYEALK	ALCKLYTECD	EKKEKYTSCS	RDAQDFANEF
PCHAS_000350	CHKLNKNSQN	KVSNIKYIYD	AYIKENEKDI	EKISG--AEA	YNSCKDIINK	KIYSMSFDIK	EMSRLYEALK	ALCKLYTECN	EKKKSYTNCS	KDAQDFAKHF
PCHAS_011530	CYKLNQRSEN	GTSDLNDFHN	KYINGIETHI	LKLPN--VNA	YNSYKDIINK	KQDMKVPIDIK	DISKLYEPLK	NLCKLYTGCD	EKKQNYTNCS	KDAHDFASNF
PCHAS_073160	CYEINKRNNN	EISNLNDFHN	KYIKNIEKYI	MKIPN--LGA	YNSYKGIIDK	QQNSMPIGVK	EMFKLYEALK	ILCDMYTECD	KKKQNYINCL	QDANAFVKEF
PCHAS_070090	CYQLNQKTKI	SIDNLNEFYN	NYIKGIEKYS	NK LDA---DN	YNSCMDIINK	KQYLMNMDIK	EMSKLYEALK	KLCKLYNECN	EKKNNYLNCS	QDAQDFANNF
PCHAS_000580	CYQLNKKTVN	TIDNLNDYYN	TYIKGIEKYS	DKLED--AEN	YNSCKDIINK	KQYLMNMDIN	KMSKLYEALK	ILCDMYTECN	GKEESYTDCS	KDAQDFVQKF
PCHAS_140130	IYMLRLKDDD	KTGNVTKFNN	KCINYDMNYI	NSIKD--TNA	YKIYKDLINS	KLYLMNNDIK	NISMFYDAFK	SLCNVYNEFN	EDDPVCSTYL	VKAKEFVKKY
PCHAS_140040	IYTLKLKSD	KTDNVMKFDY	TCINGGDNYI	NSIKNTNINA	YNIYKDPINS	KLLSMNKGIN	DISKFYDAFK	SLCNMYNGFN	PDSPNCTQCL	DDAKKFAEKY
PCHAS_000740	IYMLRLKGDG	QNGNVMKFDY	TCINPAQGYT	NSIKDTNTNA	YNIYKDLIDS	KLLLMNEGIK	DISKIYDAFK	SLCNMYNGFD	DDSDCTKNL	NDAKKFVEKY
PCHAS_100030	IHMLRLKNDG	QIDNVINFYN	TCIQPTQEYI	NSIKDTNSNA	YNIYKDLIDN	KLYLMNKDIDK	DISIFYDALK	SLCNMYNKF	DDSDCTQCL	DDAKNFAKKY
PCHAS_001120	IHMLRLKGDG	QIDNVIKFYN	TCIQPTREYI	NSIKDTNANA	YNIYKDLIDN	KLYLMNRDIDK	DISKFYDAFK	SLCNMYNGFN	PDDSDCTQCL	DDTKKFAKKY
PCHAS_011520	IYTLKLKSDG	NV---MEFYN	TCINGDDNYI	KSIKD--TNG	YNIYKDPIS	KLLSMNTGIE	DISKFYDALK	SLCKMYNEFD	RDKPDCTQCL	DDANKFAEEY
PCHAS_070030	IYMLRLKGDG	QNGNVMKFDY	KCINDDGNYI	KSIKDTNTNA	YKSYMDFISR	KLFLMNKDIDK	DISIFYDALK	SLCNMYNEFD	GDKPDCTQCL	DDAKKFAKKY
PCHAS_073190	IYMLRLKGD	NV---TEFYN	KCINPAQGYT	NSIKDTNTNA	YKSYMNLINT	KLFLMNGGIN	DISKFYDAFK	SLCNMYNEFD	GDSPYCTQCL	DDAKKFAKKY
PCHAS_000570	IYTLKLKSD	KIDKVMFYN	TCINPTSEYI	NYIKD--TNA	YNTYKDLINS	KLYFMNTDIDK	DISNFYDAFK	SLCNMYNEFA	EDDSRCTQNL	EDAKKFAEKY
PCHAS_114730	IYTLKLNGD	KV---MEFYS	KCINPAEGYT	NSIKDTNTNA	YNIYKDPINS	KLLSMNKGIN	DISIFYDALK	SLCKMYNDFA	DDSDCTKNL	GDAKKFVEKY
PCHAS_010020	SYMLNQINDK	TIKLNDFDY	QYVNRDDYK	TSIND--VSE	YNNYKQIIDK	RKDFFNMDMK	IISKFYNTLK	LLCNMYNEIN	DEVTNCTKY	GTAIKITDEY
PCHAS_000280	SYILSLKENN	NISNLKDFYD	LYINSNDKYN	DKITG--VKD	YTSYKDLIDK	NNYFLSMNMK	IVSKFYTSFK	SLCSMYNEIE	GNKSNSTKCL	EEAKNFVDEY
PCHAS_000490	SYILSLKENN	NISNLKDFYD	LYINSNDKYN	DKITG--VKD	YTSYKDLIDK	NNYFLSMNMK	IVSKFYTSFK	SLCSMYNEIE	GNKSNSTKCL	EEAKNFVDEY
PCHAS_000030	SYILSRKPNE	EISNLKCFYE	KYIESGEGYE	TSIDD--VSDY	YSSYKNLIDK	KNDFMNMDVN	YISTFYEAFK	ILCSMYNDIN	SDPPDCTKHF	TKTKEFVSKY
PCHAS_000090	SYILSRKPNE	EISNLKCFYE	KYIESGEGYE	TSIDD--VSDY	YSSYKNLIDK	KNDFMNMDVN	YISTFYEAFK	ILCSMYNDIN	SDPPDCTKHF	TKTKEFVSKY
PCHAS_000270	SYILSRKPNE	EISNLKCFYE	KYIESGEGYE	TSIDD--VSDY	YSSYKNLIDK	KNDFMNMDVN	YISTFYEAFK	ILCSMYNDIN	SDPPDCTKHF	TKTKEFVSKY
PCHAS_000120	SYILTRKPNE	EISNLKCFYE	KYIENGEEYE	TPIDD--VSDY	YSSYKNLIDK	KNDFMNMDVN	YISTFYELFK	ILCSMYNDIN	NDMPDCSKYF	TKAKEFSKY
PCHAS_001090	SYKLK--NKK	ITDEISDIYG	ELKKN-----	----NK-ISE	YNYKEQIN-	--NLMNMDFN	FISKFYEALQ	ILCDISTQFD	DDTTNCENCL	GDAKKFVAEY
PCHAS_011510	SYKLNQKSEY	SDMKLSDFYT	KYIVNN--NSY	NENING--DD	GLTYKEIINR	NNDLMN--IK	EISKFNYPPG	MLLSLYLVH	LKNLDCKNCS	KNANEFVQNF
PCHAS_070050	SYKLNHRSKC	SGTNLNDFYT	NHIEKN--EY	NKKIKD--DD	TTTTYKDIIDK	KKDLMD--IK	EISKFSSLS	ILFYLYNVFH	DKSLNCQONL	NLAKNFADIF
PCHAS_000720	SYKLSRIRQH	NFAKLNDFST	KYIVNN--ECY	NEKIKD--GD	TTTTYKDIIDK	KKDLMDMNI	EISKFNGLFS	ILFYLYLVFH	GECLDCEKNS	SLANNFADIF
PCHAS_070020	SYKLNTRVTKD	CATNLNEIYT	SYIETN--NYY	NKNIKD--GD	TTTTYKDIIDK	KKDLNMMDIK	EISKFNGLFS	ILFYLYLVFH	GERLDCQIQW	ELAKNFADIL
PCHAS_104260	SYKLNHRSKC	SATNLNEFYT	NHIKTN--KCY	NEKIKD--GD	TTTTYKDIIDK	KKDLMD--IK	EISKFSSLS	ILSYLYLVFH	GECLDCEKI	KLAKNFADIL
PCHAS_140020	SYKLNHRSKC	SATNLNEFYT	NHIKTN--KCY	NEKIKD--GD	TTTTYKDIIDK	KKDLMD--IK	EISKFSSLS	ILSYLYLVFH	GECLDCEKI	KLAKNFADIL
PCHAS_030060	IYKLIKIKNN	TVTNLNDFYN	NYIEKN--KCY	KDKIN---GD	GSTYKDIINK	KKDLMDMNI	EIFKLEAPFN	ILYLYYHEIS	DQKSFCKEYS	KYAANFADQF
PCHAS_000750	SYILNLYLKD	KLTLNLFYF	NYIEKN--KCY	NEKING---D	GSTYKEIIDK	KKDLMDMNI	EIFKLEAPFN	ILYLYYHQIY	DQTSFCSEYS	KYATKFSQF
PCHAS_000680	SYKLIKIKENN	IVNKLSDFYN	SYIETN--NDY	NKNING--ND	GPTYKEIIDK	KKDLMDMNI	EIFKLEAPFN	ILYLYYHEIS	DVAVDCKKNL	SIA--NNFAKE

## Supplementary data 1 (CIR alignment)

PCHAS_000420	SYKLIKIKENT	I IKKLSDFYD	SYIETN-EYY	KDKING--DD	GLTYKEIIDK	KKDLMDMNIS	EIFKLDAPFN	ILYFLYYAYN	YDYWDCAKNS	NYANSFVNQF
PCHAS_000100	SYKLIKIKENP	I IKKLSVFDYD	SYIERN-EYY	NKNING---D	NLTYKEIIDK	KKDLMDMDIN	EIFKLEAPFN	ILYYLYHQIY	DKNSFCSEYS	DYANQFVQKF
PCHAS_000310	SYKLIKIKENP	I IKKLSVFDYD	SYIERN-EYY	NKNING---D	NLTYKEIIDK	KKDLMDMDIN	EIFKLEAPFN	ILYYLYHQIY	DKNSFCSEYS	DYANSFVNQF
PCHAS_040040	SYKLIKIKENN	MIKKLSDFYN	SYIERN-EYY	NKNING--GD	GLTYKAIIDK	KKDLMDMNIS	EIFKLEAPFN	ILYYLYYKIS	DKHTDCEKNL	NDAKKFVDKI
PCHAS_120060	SYKLIKIKENN	RIKKLSDFYN	SYIETN-DCY	NKNING---G	NLTYKAIIDK	KKDLMDMDIN	EIFKLEAPFN	ILYYLYDQIS	DEKSFCTKNL	DYAKTFAEKY
PCHAS_104230	SYKLIKIKENE	KIKKLSVFDYD	SYIKTN-DCY	NKNING--GD	GLTYKAIIDK	KKDLMDIDIN	EIFKLEAPFN	ILYYLYNVIH	DEHPDCEKNL	DSAKNFAEKY
PCHAS_030040	SYKLIKIDNP	I IKKLSDFYK	SYIDTN-NYY	NENING--GD	GLTYKAIIDK	KKDLMDMNIN	DIFKLEAPFN	ILYYLYDQIS	DEHPDCEKNL	DYANDFAEKY
PCHAS_000340	SYKLIKIDNP	I IKKLSDFYK	SYIDTN-NYY	NENING--GD	GLTYKAIIDK	KKDLMDMNIN	DIFKLEAPFN	ILYYLYNVIH	DEHPDCEKNL	DYAKNFAEKY
PCHAS_001110	NYKLNQNSD	GINTLNDFYT	KYIEKN-THY	NQKISK-AND	NKTYKNIIDK	KKNLMNMDIN	DIFRFYDAFE	SLCSMYNELG	DENPNCGKCS	QKADEFVNKY
PCHAS_120040	GYKINQILNG	EFTTLKDFYT	KHIKNN----	-----KD	DVDYKDHLDN	KINSMDINIK	DTSNIYEAFE	ILCNMYTVYN	ENDKKCTNCS	QNAEFVQKI
PCHAS_000130	GYRLNQILDE	EITTLKDFYT	KYIKDN----	-----KD	DVDYKDRLLD	KINSMDINIK	DISNIYEAFG	ILCEMYTAHN	ENDKECTNCS	QYADEFVQKI
PCHAS_000400	NYKLNQISNE	ETTLKDFYT	KYIKDN----	-----KE	HVDYKDHLDN	KIYSMNIDSE	KIYNIYEAFG	ILCKIYTAHN	ENDKKCTNCS	QNAEEFVQKI
PCHAS_040020	NYKLNQISNE	ETTLKDFYT	KYIKDN----	-----KE	HVDYKDHLDN	KIYSMNIDSE	KIYNIYEAFG	ILCKIYTAHN	ENDKKCTNCS	QNAEEFVQKI
PCHAS_000110	SYMLNQISNE	GIPTLKDFYT	NNIEKN-TKY	TRHVSS--AN	DSNYKGVVDK	KINLMNINKN	IISKFYDVLK	SLCSMHTECI	GCKSNCTDCL	EKAKDFADKY
PCHAS_120050	SYILNQMPNE	GIHTLKNFYT	NHIETN-TNY	TSHVSS--AS	DSNYKGVVDK	KIDLNMNKS	IIPKFYDIFK	SLCNMYNELD	KNESNYTNCL	KDAQNFVDEY
PCHAS_104250	SYMLNQISNE	GTPTLKDFYA	NNIEKN-KKY	TNNVTS--AS	DSNYKEIIDK	KINLMNANKN	IIPKFYDIFK	SLCKMYNELD	QNEPNYTNCL	ADAQNFVNEY
PCHAS_000320	SYMLNQISNE	GISTLKNFYT	NHIETN-TNY	TNNVTR--AS	DSNYKEIIDK	KINLMNANKN	IISKFYDIFK	SLCNMYNELD	ENKPNYTKLC	ADAQKFVDEY
PCHAS_000410	SYMLNQISNE	GISTLKNFYT	NHIETN-TNY	TNNVTR--AS	DSNYKEIIDK	KINLMNANKN	IIPKFYDVFK	SLCSMYSELD	DDEPDYTNCL	KNAQNFVDEY
PCHAS_140030	SYILTRKTND	GINNLNDFYN	MYIKDK-EKY	TAKING-VTG	YTSYQNLIDK	YKELMTIDIQ	DMSKLYPPLK	SLCDLYNGCI	ICQSDYCNCL	EKANDFAKKY
PCHAS_114740	SHILSLKENE	GIKNLNGFYE	KYIKDK-EKY	NNHIPG-VHG	YTSYQNIIDK	YKELMTDDIE	YMLQFYGPLK	SLCEMYNACN	LNRSKCTNCL	EKAKDFADKY
PCHAS_040030	SYKINENTQT	I IKS LNDFYT	KHIEKN-ANY	SEKITK-SGD	NKTYKDIINK	KQNLNMNIN	DISQFYDALK	LLCSMYNELS	DDNPNCCKCS	KDAKEFVNKY
PCHAS_070130	SYKLNQNPQK	GINTLNDFYT	RNIGKN-THY	TSTTDD-IDV	YKSYKDFIDK	NNDLMTVNIK	EISQFYAPFK	SLCGMYIECN	GKKTNHTKCL	EKANEFVENF
PCHAS_073180	SYKLNQNPQN	RINRLNDFYT	KHIEKN-AQY	NTKITT-VSD	NKTYRNIIDK	KKDLMDMDIN	DIFKFYAPFK	SLCDMYNELD	KENQECENYF	RKANELVENF
PCHAS_000300	SYKLNKYPQN	KITTLNDFYT	EHIEKN-KYY	NVKIT---KS	SKTYKDIIDR	KHDLMNIGIK	EISQFYDAFK	SLCDMYNELD	KENQECANCS	KKANEFVANF
PCHAS_120070	SYKLNKYPQK	GINTLNDFYT	EHIEKN-EYY	NVKIT---KS	SKTYKDIIDR	KHDLMNIGIK	EISQFYDAFK	SLCDMYTELD	KKNQDCENYF	RKANKLVENF
PCHAS_001130	SYKLNKYPQN	KITTLNDFYT	KHIEKN-EYY	NVEIT---KS	SKTYKDIYR	KHDLMNIGIK	EISQFYDAFK	SLCDMYNELD	EENQDCENYF	RKANKLVANF
PCHAS_100040	SYKLNKYSQK	GITKLNDFYT	QHIEKN-KYY	NVEITK-SSD	KKTYKDIIDR	KHDLMNIGIK	EISQFYDAFK	SLCDMYTELD	KENQDCENYF	RKANELVENF
PCHAS_000560	SYELNKYSQK	GITKLNDFYT	QHIEKN-EYY	NVKIT---KS	SKTYKDIIDR	KHDLMNIGIK	EISQFYDAFK	SLCDMYNELD	KENQDCENYF	RKANEFVKNF
PCHAS_050020	SYKMNQHPSQ	KF-GTNDIYN	NFKQYG----	-----YW	NRKHNNYIEQ	IKKY--VDIK	DMTKLHEAFI	LLCNMYTEID	ENKSNCTKCS	QKASEFVKKF
PCHAS_000020	SYKINQHPNK	KF-GTNDIYN	NFKKYN----	-----YW	NLNHNNYIEQ	IKKY--VDIK	DMTKLHEAFI	LLCKMYTEFN	DETKNCTKCS	QKASEFVKSF
PCHAS_000290	SYKINQHPNQ	TF-GTNDIYN	NFKQYG----	-----YW	NRKHNNYIEQ	IKKY--VDIK	DMTKLHEAFI	LLCNMYTEID	ENKSNCTKCS	QKASEFVKKF
PCHAS_060020	GYKLNKYSNI	KLTLNDFYT	KIKINS--KY	IMHITG-NET	YKNYKDHDIDK	KPDLNMNINIK	TIFKFYEALQ	ILCNMHNEDK	EKNTDCTKCS	QDAEFVSEF
PCHAS_001040	CYKIQOYQNG	EIALKDFIT	EHIETN-TNY	NWKTCK-GIG	NKSYKEFIDK	KQSLINMDVK	IMSKFHEALQ	ILCNMYNENK	EKNINCTKCS	QNSEFLKIF
PCHAS_114750	SYKIQOYWNI	ETTLKKNFIT	EHIETN-TNY	NEKIKQ-GKG	NKNYKEFIDI	KQSLINMDVK	IISKFHEALQ	ILCNMYNEDN	AKNIDCAKCS	SKASEFVEKY
PCHAS_000140	SYKIQOYWNI	ETTLKKNFIT	EHIETN-TNY	NEKIKQ-GKG	NKNYKEFIDI	KQSLINMDVK	IISKFHEALQ	ILCNMYNEDN	AKNIDCAKCS	SKASEFVEKY
PCHAS_000070	SYKIQOYWNI	ETTLKKNFIT	EHIETN-TNY	NEKIKQ-GKG	NKNYKEFIDI	KQSLINMDVK	IISKFHEALQ	ILCNMYNEDN	AKNIDCAKCS	SKASEFVEKY
PCHAS_030080	SYKLNQNPQK	GITTVYDLYT	KHIEKN-TNY	NQKIID-DSD	DKISKKHINK	KLELSMEIN	VISKFYEALQ	ILCNMYDEDK	EKNPDCTKCS	QYAEFVKSF
PCHAS_073200	SYKLNQKQNI	KFKNLKDFYD	KYIKDN-GNI	YNNDITGVDA	YNSYKDIINE	KINMKTINIK	EMSFYDLLK	LLCDIYTEFD	SNNSDCTKCS	QKANEFVENY
PCHAS_120020	CYKINQ-ETH	KFSNLNKFYN	KYIKGI-EKH	FSEENGA-KA	YNSYKDLINN	KICSNTIDIS	NMSKFYEACE	LLCKMYEENK	GSSKDCTACL	EKAKEFSEHF
PCHAS_000060	CYKINQ-EKH	KLNNLSAFYK	KYIMDI-EKD	FMEENGA-NA	YKSYKDIINN	IICSKTIDIS	NMSKFYEGFE	LLCKMYEENK	GSSKDCTACL	EKAKEFSEHF
PCHAS_130020	CYKINQ-ETH	KFSNLNKFYN	KYIMNN-EKN	IKESNG-AEA	YKSYKDIINN	KICSNTINIK	ETSKLYEAFE	ILCKMYNELD	DDAKNCTKCL	EDANKFVSKY
PCHAS_000150	CYKIDQ-GTY	KYSNLNKFYK	KYIMDI-EKD	FMEENGA-NA	YKIYKDIINK	QICLKTINIK	AMSKLYEALE	ILCKMYSELG	DDAKNCTKCL	DDANKFVSKY
PCHAS_000500	NSKNNLINS	TYISTTTIYD	ILEINS----	-----SYW	YNEFRDKIEK	KRNAMNFGDY	YMHKLYNLLK	EICSTINKYN	EDKSYPTIYL	KHANKCVNTY
PCHAS_001100	SSKYNLISPE	YDVNITGIYD	ILERNN----	-----PIW	YNKYRDRIGK	KRNVMDFGDF	HMGNLYNLLK	EMCILITKYN	EDRSYQDVYL	KYANNCANIY
PCHAS_042010	GDRLFNIESY	SSSTLIDFYD	NYLIKL----	-----PK	FNSYKSLIEK	KKHLKDSHIV	YMRRFYLLFQ	EVCYISLKYS	KNNLNMDKMK	RDYTFLNKY
PCHAS_000040	SSKNILINPD	MLVSVTGIYD	ILERNN----	-----PIG	YKEFRNKIEK	KKQLMNFYDY	HMGKLYELLK	EMCTLITKYD	QDRSFPDSYL	NHANKCANLY
PCHAS_000390	CFKNKLINPD	MYVSVTGIYD	ILERNN----	-----PIW	YNEFRNKIEK	KKRLMNFYDY	HMGKLYELLK	EMCTLITKYN	VDSSHPDAYL	NYA-NKCANT

## Supplementary data 1 (CIR alignment)

PCHAS_030180	NSKIDRY-EE	RLIGVDAVYD	MLTRND----	-----W	FGEHYNSIKK	KSDMMKLFYT	YLTRLRYELLK	EICNTITKCN	-NSPNTTEECE	KHAKKCVQLY
PCHAS_130070	NSKIDQY-EE	RLIGVYAVYD	MLTRND----	-----W	FGEHYNSIRE	KTDMMKIFYT	YLNRLYGLLK	EICNTINTCK	-DSSNTNACL	NHAKKCVQLY
PCHAS_011500	CYKNNQPNR	LI-GVDDIYN	MLTRNP----	-----W	FGKYYSIKE	KRGMMKIYFS	YLNKLYALLK	GICDAINKCD	-YPSNADECI	KYANKCANLY
PCHAS_146850	GYKNDQPNR	LI-GVDNVYD	MLTRND----	-----W	FREHYNSIKG	KKNIMKIYFS	YLNPLYTLLK	KICDTINKCN	-DPSNNKECI	DYANKCVKLY
PCHAS_070100	SYKLYQYKER	SI-EVDDVYD	ILTGND----	-----W	FQEHYKSIEE	KKDMMKFHLI	YLNPLYTLLK	EICNTITNCS	-DSSNSSECC	ESAKKCAASY
PCHAS_060050	SYKISQNPKE	EI-AKNTMYD	VLTQNE----	-----W	FNEHSESIK	IKDMMNFHFI	HLKNLYEFLK	GICETINKCN	-GSSNTNECI	ESAKKCADSY
PCHAS_030110	SSKIKENTKA	GH-GINDIYN	TFIKNN----	-----SW	FSEHKSIEK	KKDMMNFHFI	YLTLNYKFLK	GICETINICN	-DPSKTSECI	KSANKCADSY
PCHAS_130050	SSKIKENTKA	GY-GINDIYN	TFIKNN----	-----SW	FSELSIESIEK	KKDMMNIHLI	CLKNLYALLK	GICETINKCN	-ESSNSSECI	KAGKECASLY
PCHAS_030090	NYKISQDQNI	EI-IRGTMYE	ILKQND----	-----W	FGEHSESIK	IKDMMGIHSE	YLNRLYRLLK	GICETINTCN	-GSSDTNKCI	ESAKKCVDSY
PCHAS_146870	SYKISQNPNI	EI-AKNTMHD	ILTQNE----	-----W	FNEHSESIK	IKDIMGIHFL	YLNKLYEFLK	GICETINKCN	-GSSDSNECM	ESAKNCKSLY
PCHAS_130030	SYKISQNTNI	EI-IRGTMYD	ILKKNND----	-----W	FGEHSDSIDN	RKDTMEIHYL	YLNKLYNFLK	GICETINKCK	-VSSNSSGCK	ESAEKCSYLY
PCHAS_040110	NSKLSQKPI	EI-ERYTMYN	ILKQYN----	-----W	FGEDGKSIEK	KKDIMGIHFL	YLNKLYDFLK	GICETITNCN	-GHSKTSECI	ESAKKCSGLY
PCHAS_042030	NYKISQDQNI	EI-IRGTYLN	ILTQNN----	-----W	FSEYSESIEN	RKDTMKIHYL	YLNKLYDFLK	GICETINKCK	-DSPTSNECK	KYGEKCSPLY
PCHAS_060140	NSKITQNSTI	EI-EKYTMYN	ILKQNN----	-----W	FGEHSEPIET	KKEIMGIHFL	YLNKLYDFLK	GICETINKCK	-DSSHSSDCQ	ESAEKCSGLY
PCHAS_030210	SYKIRQNPNI	EI-EKNTMYD	ILRQNG----	-----W	LSEHIESIEK	IKDIMGIHFE	YLNRLYRFLK	GICETITNCN	-DPSKTSDCQ	ESAKKCSGLY
PCHAS_060060	SYKISQNKNI	EL-IRGTMYG	ILKQNG----	-----W	FGEHSDSIEN	RKDTMEIHFE	YLNRLYRFLK	GICETITNCK	-DSSTSSKCC	DSAKKCSGLY
PCHAS_050070	SYKIRENTNI	KL-IRNTIYD	ILTQNE----	-----W	FSEYRQYTDK	NEDIMGFHLI	YFKRLYDLLK	GICETINKCS	-NSSNNEECI	NSAKKCEFEY
PCHAS_040050	NSKIRGTAEM	EA-EINNIYN	VLTENG----	-----W	ISEYRQYTNK	NEDIMKFHFL	YLKFFYKFLK	GICETITNCS	-GPSKTEECI	KSAKKCEFEY
PCHAS_001050	NSKIRGTAKM	EA-EIKNIYE	ILIEND----	-----W	ISEYRQYTDK	NENELKFHFL	YLNKLYALLK	GICETITNCS	-DSSKTEECI	KSAKKCEFEY
PCHAS_083760	SSKIKENTKI	DI-EKNTMYD	IFTQNE----	-----W	FSEHSESINN	KKDIMGLHFI	YLNPLYKFLK	GICDTINKCK	-SSSNPTECC	ESAEKCGELY
PCHAS_140070	NYKITQNTKI	E-DGINDIYN	ILIKNN----	-----SW	FSEHRQYTDK	NENVMKIYFS	YLNPLYALLK	EICNTIDKCS	-DPSKTSECC	ESGKKCASLY
PCHAS_000180	NSKVKVNEVF	E-PVTKNIYN	ILNEHN----	-----L	LSEYREYTDK	NEDVMKIYVY	FLKNLYEFLK	GICDTINKCK	-GSSTSNECC	ESGKKCVELY
PCHAS_001060	NSKVKVNEGF	E-PETDNIYN	ILNEHN----	-----L	LSEYREYTDK	NEDVMKIYYA	FLKNLYTLLK	EICNTINKCK	-GSSTSTECQ	NSGEKCVELY
PCHAS_011450	SYKIRENPNI	EF-IRSTIPG	ILEKNQ----	-----W	YKELSKSADD	KENTMGFHLI	YLNPLYKFLK	GICETINKCK	-GSPTSSECK	QYGEKCSGLY
PCHAS_060030	SYKIKQNPKE	EF-IRTTIDD	ILKKNNE----	-----W	YNELNIPADD	KKNMIRFHYI	YFTNLYNFLK	GICETINKCK	-SSSNDECI	ESAKKCGELY
PCHAS_120030	SYKIKQNPKE	DF-IRTTIDD	ILKQNE----	-----W	YNELGITVDD	KKSTIRFHYI	YMNLYNFLK	GICETINKCK	-VSSNSNDCQ	ESAKKCVPLY
PCHAS_030120	NYKIRQNKKI	EA-GINDIYD	MITTNE----	-----W	FTEHRQYTDK	KKDMMKFFYT	YLNPLYALLK	GICDTINKCS	-ESSSTNECI	KSAERCVSPLY
PCHAS_114640	SYKIKQ-NKK	ILFGVSNIEE	MITHNE----	-----W	YTDYREYIDK	KKNIMGLYDS	YLNPLYKLLK	KICDTINKCN	-DPSKTKECI	ENAKMCATLY
PCHAS_130350	SYKLOQSNRK	SSKPIGLYDF	FNNHVV-NG-	-----DWYE	VVEEHVQPKI	SLLNNNINID	LMSDIYYILK	EMCKKFSNKN	DNFDYLEDFM	NYKRCGENI
PCHAS_030150	NHKLOQLPKS	ELIDLNNFYD	KYIKDN----	-----EW	YASFISEIDT	QKELMNV DAT	MLSDLYYIFK	EMCNKFSNTE	NHEDFAYFYN	NNNNNFVKTY
PCHAS_060120	SYKLOQLNEK	NSQSIGLNDF	FSNYII-NG-	-----EWYD	DIEEYVEPKM	FLLNNDINIE	HMSDIYKIFK	QMCKIFSND	DDIDDLDFTG	YYEKSCGENI
PCHAS_041960	SYKLOQLNKN	NDKPIGLNEF	FNKHIA-KG-	-----DWYE	EFKEQIESKI	HLVDKEIDID	LMSDVYIFK	QMCGLSNQE	DDFNYLEDFS	TYKICA EKL
PCHAS_130110	SYKLOQLPPS	ELIDLNQFYD	NYIKNN----	-----VW	FESYLSIDIT	KKNLMSVDTK	TLSDLYHIFK	EMCKIFYNNG	DDYNDLDFDT	GYINCAKNY
PCHAS_041950	CHRLYKISKD	YKLPLKNSLE	EYLGKS----	-----IG	SFNYRGILFS	KKHLMDSNIA	IMNMLYLLFQ	KICETINIYE	AHNAKPHEYI	NKGLECHAIY
PCHAS_060130	CHRLYKISKG	HNLPLKHSLE	KYLGKS----	-----TG	NINYRGILNN	KTYLMNSSIE	IMNMFYLLFQ	QLCEIINTYE	TNKLQPYQYI	NSA-AQFYII
PCHAS_130120	CHRLYKISKD	HNLNQLYLG	NNLGKS----	-----RG	DFNYKGILNN	KIYLMNSNVA	IMNMLYLLFQ	KICETINTYE	THNAQPHEHI	NKGIEYHFMY
PCHAS_070170	SHILYKKLEN	NTITLKS AHE	KHLKEN----	-----FG	NYNYWNL LAK	KYILTDSNIA	VMNVLYLLFQ	QICDTIKEYR	KENVLGHEYA	NKA-FQSYII
PCHAS_100060	SHILYKKLEN	NTITLKS AHE	KHLKEN----	-----FG	NYNYWNL LKG	KFYLTDSNIA	VMNVLYFLFQ	QICDTIKEYN	KENVLGHEYA	NKA-VQSYIM
PCHAS_146770	SDKLFKIEND	YKATLEESYD	KHLKNS----	-----MG	NFNYWRV LNS	KQIYKATIR	KMSAMYTLT	YICKLITEYN	KHNKSKNVSE	NDRTQCRNYY
PCHAS_146790	GDKLFKLEKN	YKTMEESYK	KYLDNH----	-----IG	NYKYWNVINS	KRKYKATIR	KMNELYNLLS	YICKLITEYN	KNPKNRDTL	GNYAQCRNYY
PCHAS_060160	SDKLFKIEND	YKATLEESYN	NHLKNY----	-----MG	NFNYWRV LNN	KKVYKATIW	KMSAMYTLT	YICKLITEYN	KNVKNRDTL	VKSSQSVNYY
PCHAS_000260	GDKLFKLEKN	YKATMEESYK	KHLDKH----	-----TG	NYIYWNVINS	KKLYKATIR	KMNELYNLLS	YICKLITEYN	KNIKNDRDL	GNYAQCRNYY
PCHAS_130280	GDKLFKIDND	YKITMEESYE	KNLKNS----	-----MG	NFNYWNLRS	KNIYKNATIR	KMNAMYTLT	YICKLVTEYN	KNNKNKNSD	NDRTQCRNYY
PCHAS_083720	GDKLFKLEKN	YKATMEESYK	KHLDKH----	-----TG	NYIYWNVINS	KKLYKATIR	KMNELYNLLS	YICKLITEYN	KNIKNDRDL	GNYAQCRNYY
PCHAS_090010	GDKIFKVGKD	YKSTLEESYE	KNLKNI----	-----TG	KFEYWDALSS	KEVYKATIV	RMIFYNLLN	SICKTINELN	KNLNDAESL	KNYTKCLYLY
PCHAS_114600	GDKIYKVDKD	YKSTLEESYE	KHLKNI----	-----TG	KFEYWDALNS	KQVYKNATIV	RMIFYNLLN	SICKTINELD	KNLNDAAIL	KKYTHCLDLY
PCHAS_146860	SHLLYKMETN	NLISLENSYE	NNLGKS----	-----IG	DFNFLNLLHN	KKYLMNANIV	IMNVFYHLFK	EICRTIKIYQ	TENTLSHEYI	NSATQIYFIY

## Supplementary data 1 (CIR alignment)

PCHAS_130090	SNILYKIAPT	HLESLGESYE	KHLSKS----	-----VG	NFVFWNLLYN	KRHLLFDGNIS	ILSIFYILFK	QMCETFDIYN	KRGISGHEYA	INAAQSYIYY	
PCHAS_030170	SHILYRKLED	HTITLKSAYK	KHLKDN----	-----FG	EFRHWYLLQD	KKYL TNSNMA	IMNLIYLLFFQ	QICDTIKIYK	KPHVLGHEYA	SKAFECYFIY	
PCHAS_130060	GHILYRMTED	HAITLKSGYD	KYLKDN----	-----FG	EFRHWYLLQD	KKYL TNSNMA	IMNLLYLLFFQ	QFFEMIKKYR	TPHVLGHEYG	NMALECYFIY	
PCHAS_030200	SNKLYEIATN	NSKYLNQSYE	INLDKS----	-----IG	NFDFLNSRDD	TKELKDANIT	IMNILYLLFFK	EICKAVWMDQ	GKSTEMYKHT	TQITQCFFIY	
PCHAS_060080	SNKLHEIATN	NYKYLNQPYE	NNLDKS----	-----IG	NFDFLNSRDD	TKELKDANIT	IMNILYLLFFK	EICKTVWMDQ	GKNTKIYEHT	TQITQCFLIY	
PCHAS_060110	ADKFRKANNE	GFFTLDEYYE	EFLVNH----	-----ND	NFNYWYELDN	KMHLKDSNII	LMVEFYLLFT	NICNMLLESE	KSELDLNKIK	VFDNNCFRKY	
PCHAS_030140	ADKFLKVNND	YSFSLNDYYE	ELIVKP----	-----GD	NFSWWEKLLD	KMLWKDSNII	LMARFYLLM	DICNALLENE	KSEFDLKRIE	GIDKKCHITY	
PCHAS_130100	ADKFLKVNDD	YSFSLNDYYE	ELIVKH----	-----GG	NFGWWEKLLD	KMFWKDSNII	LMARFYLLM	DICNALLENE	KSELDLKNIE	ENDKKCYLNY	
PCHAS_041990	ADKFRKANNE	GFFTLDEYYE	EFLVNH----	-----ND	NFNYWYELDN	KMHLKDSNII	LMVEFYLLFT	NICNMLLESE	KSELDLNKIK	VFDHNCFRKY	
PCHAS_130080	ADKFLKISPN	RSFSLYDYYE	KVIVKR----	-----GE	NFNCWGKLDN	KKDLKDSNLS	IMTLFYQIFM	NICNAIMKNE	ISRLELNKFM	DIDHSYHQLY	
PCHAS_030190	ADKFLKISPD	RSFSLNDYYE	KVIVKQ----	-----GE	NFNCWEKLDN	KKDFKDSNLS	IMSLFYQLFM	NICSALMKNE	ISNFELKKFK	DIDYDYYQIF	
PCHAS_060090	ADKFLKISPD	RSFSLNDYYE	KVIVKQ----	-----GE	NFNCWEKLDN	KKDFKDSNLS	IMSLFYQLFM	NICSALMKNE	ISNFELKKFK	DIDYDYYQIF	
PCHAS_041970	ADKFLKITPD	HSYTLNEYD	EVIANN----	-----DG	NFNCWGKFDN	KKHFKYSNIS	IMSVFYQLFM	NICNAVEND	ISKLETKKFM	MIDHNYCQIY	
PCHAS_041980	AAKFLNITHD	VSFSLNDYYE	KFIVNQ----	-----GG	HFNCWGKLDN	KEYLKDSNLS	IMSVFYQLFM	NICKAVENE	ISKLETKKFM	MIDYNYQIY	
PCHAS_042020	GYRLFQTPSY	SSSTLVDDYD	NYILKS----	-----HL	PFKNDYLIDK	KKHLLHANFE	LINKLYKLLN	YLCNIITEPN	-IYTESTKIK	SNISMFQDEF	
PCHAS_060070	GYKLFQTSY	NSPNLIDFYN	NHLTKS----	-----IL	PSGLNYILEK	KKHLLYANLE	LINELYILFN	DMCNIITDPN	-IYTESDKIK	SNVSKFQDTF	
PCHAS_011330	SNILLQISKD	HNSILEASYK	TYLEKH----	-----MG	NYNYWNILGN	KRYLKDSNIS	YMSQFYKLF	QICNIINEYN	TNGPTSKNLG	KISIECSKTY	
PCHAS_000220	SNILLQISKD	HNSILEDYSK	AYLEKH----	-----MG	NYNYWNILGN	KRYLKDSNIS	YMSQFYKLF	QICNTINEYN	TNGAGSKNLR	NISMESKTY	
PCHAS_042070	SDKLYKLEEN	KIRTLDQFYE	KHLKDY----	-----TG	NFKYWNVIDS	KRVYKGANVW	YLSELYSLLN	DICSIVIEYN	KNKNKKNKIE	KASSQCHKKF	
PCHAS_110030	SGKLYKLENN	KSTTLEESYK	KYLENN----	-----MP	SFNYWNVLGS	KREYKIANVW	YMSRYSLLE	CICSIVIEYG	KSKRN-KQIE	QISQQCYQKF	
PCHAS_140140	GDKSYKLDKN	KTATLEDSYK	DNLEKY----	-----TG	NFKYWNVIDS	KRVYKRANVW	YMSELYKLLK	SICSIVIEYN	KNKSK-KKIE	QISSECYQKF	
PCHAS_010040	SDKLYKLEKN	KGVTLLEESYK	NYLDAH----	-----TG	NFKYWNIVDS	KRVYKANVW	YMSELYGLLK	YICDLVNEYN	-KKPNKEKIE	GISRKCQQKF	
PCHAS_073150	SDKLYKLEKN	NAATLEEYYN	KHLKEA----	-----MG	NYNYWNLLGG	KQYKRVANVW	YMSELYSLLN	DICSIVIEYG	-KNKRSKKIE	HISPOCYKFF	
PCHAS_073130	GDKLFKMHNK	SKGKGRNNNI	TLYSAY-EKY	LNNH----	RG	YLDYWALLNN	ISGLKEANLE	HMHKFYRLLN	DICKTIVYI	HKDSKNNNLI	MNSTECSNOY
PCHAS_040060	SDKLFKMYQK	DKRKGNNRI	TLDEAY-KKY	LDKD----	MG	DYKYWNRLDN	VKGLKDANLS	HMNEFYKLLS	HICKTIIYYK	IKDAKPTNLH	QNSTNSSNOY
PCHAS_000430	SDKLFNMYQK	DKRKGNNRI	TLDEAY-KKY	LDKD----	IG	DYKYWNLLGN	INGLKEANLS	HMNEFYKLLK	HICKTIMHHK	IKPTESASLI	QNSTNSSNOY
PCHAS_104200	SDKLFKMHEE	GKKKGQSNIT	TLDEAY-KSY	LDKN----	IG	NNKYWDALDN	VKGLKDANLS	HMNEFYKLLK	HICKTIMHHK	NKPTESANIL	QNSTNSYNQY
PCHAS_114720	SDKLFKMHED	GKKKSQSNIT	TLDEAY-KSY	LDKN----	IG	NNKYWDALDN	VKGLKDANLS	HMNEFYKLLK	HICKTIMHHK	IKPTEYTSHL	ENSTNSSNOY
PCHAS_000770	SDKLFKMYYE	SKGKDAKGF	VYNITL-NQA	YDKYLKHEV	KLNYWTLDDI	IPGLKEANLK	YMSEFYKLLS	NICKTIKEYN	DKGAESKLLS	KYSGSCLNOY	
PCHAS_130220	SDKLLKMYD	SKGKDVKKGF	VYLITL-NQA	YDKYLKKNV	KLDYWDLFGH	INGLKEANLK	YMSEFYKLLN	NICKTIADYK	DNGAGSKQLS	KYSKNCLNOY	
PCHAS_070040	SDKLFNMHKE	GKGKHIKGY	MDTFTL-KQA	YEEYLEKHG	IFGYWYLLDM	IKDLKEAYLK	YMSEFYKLLN	NICITTIKEYN	DKGAKSSQLS	KYSKNCLNOY	
PCHAS_011490	SDKLYKMHLK	SIGQKDTAEY	MDGTTL-NQA	YDNYLKNYKV	GLGYWDLDDM	IMGLKEANLK	YMAEFYKLLN	NICKIITDYN	DNGSESKKLS	KYSENCLNOY	
PCHAS_070060	SDKLYKMHLK	SIGQKDTAEY	MDGTTL-NQA	YDNYLKNYKV	GLGYWDLDDM	IMGLKEANLK	YMAEFYKLLN	NICKIITDYN	DNGSESTKLS	KYSENCLNOY	
PCHAS_030070	SDKLFKIHDK	SEEK----	D--KKITL-NQA	YEEYLEKHG	IFDYWIFLNM	QQGLKEANLR	YMSEFYKLLD	KICKTITDYE	KKRDEITNHI	TNSTECSNOY	
PCHAS_137110	SDKLFKI-HD	ESK---DKDN	--EITL-NQA	YDYLKHNHV	NFNYWNFFYN	IESLKEANLW	YMSEFYKLLD	KICKTITDYE	KNRDEITNLI	TNSTECSNOY	
PCHAS_114700	SDKLFKMDDE	GKEKNSKKPN	IHTITL-NQA	YEKYLKHNHV	KLDYWALLNI	MPGLKANLK	YMSEFYKLLN	HICKVITYYN	EKGAKSRKLS	KYFVDCRRQY	
PCHAS_137030	SDKLFKMDDE	GKEKNSKKPN	IHTITL-NQA	YEKYLKHNHV	KLDYWALLNI	MPGLKANLK	YMSEFYKLLN	HICKVITYYN	EKGAKSKKLS	KYFVDCRRQY	
PCHAS_011480	SDKLFKIHKE	AKNIGEGRM-	-DGTTL-NQA	YDYLKDHKV	GLDYWVLLDM	IKGLKEANLK	YMSEFYKLLN	IICKTIVDYK	DNGTESKLLS	KNSISCRROQY	
PCHAS_050040	SDKLFKIHSK	SENK-----	NKQITL-NQA	YDMYLKHKHA	KLDHWILFDN	IQGLKANLK	YMSEFYKLLN	KICITITDYS	ENGTESKNII	MNSTECSNOY	
PCHAS_000470	SDKLYKMHEE	SIGQKYENNY	MDGTTL-NEA	YEKYLKNYKV	KLDYWILFDM	IKDLKEANLK	YMSEFYKLLN	KICKIITDYN	-NSAQTQQLS	KYSVDSRRQY	
PCHAS_000170	SDKLFKIHKE	AKNIQRG--Y	MDATTTL-NQA	YKEYLEKHQ	RLDYWNILNM	QQGLKEANLK	YMSEFYKLLN	HICKIITYDE	TKDAKSREFS	INSIGCSRQY	
PCHAS_130170	SDKLFKMHLK	SIDKKDVNNY	MDGTTL-NEA	YKNYLEKYKG	IFDYWAFLDM	IKGLKEANLK	YMSEFYKLLN	LICKIITGY	-NGTQTKQFY	KYPADCSHOY	
PCHAS_000660	SDKLLKIHKK	GKGKIGIGR	MDDFTL-KRA	YEEYLKHNHRQ	RLDYWVLLDM	NPGLKEANLW	YMSGFYKLLN	LICKIITGY	-NGPKNKQFY	KYPADCSHOY	
PCHAS_040080	SDKLLKIHKK	GKGKIGIGR	MDDFTL-KRA	YEEYLKHNHRQ	RLDYWVLLDM	IKDLKEANLW	YMSEFYKLLN	LICKLITDYY	-NGVQTKQLY	KYPADCSHOY	
PCHAS_070070	SDKLFKMHLE	SKGIKDPDY	MDGITL-NKA	YKNYLEKHG	ILDYWTHLDM	IPSLKEANLK	YMSEFYKLLN	LICKIITGYN	-NGSQNKKLY	KYPSDCSFQY	
PCHAS_137090	SDKLFKMDDE	SK---DPNRN	IYTITL-NQA	YEKHLKKYKV	KLDYWTLFDI	IGDLKANLK	YISEFYKLLS	NICKAIADYE	KNGAKSKKLS	KNSADCSYOY	



## Supplementary data 1 (CIR alignment)

	.... ....  205	.... ....  215	.... ....  225	.... ....  235	.... ....  245	.... ....  255	.... ....  265	.... ....  275	.... ....  285	.... ....  295
PCHAS_000950	KELVNKVFED	EHCYNALLNL	KKAYEKFTNG	-----NN	T--NAFPKIN	EI-----	-----EKMN	NCKKLCEKAN	RSWKVIHVEV	KDVVDAKKVS
PCHAS_001010	KELVNKFFED	ENYCNVLLTL	KKAYENFRSG	-----SNN	T--STLPILN	EI-----	-----EGIS	NCNKLCEKAN	RSWKLIHVEV	KDVVEEKKVS
PCHAS_140090	EKLNENDTKG	VSykkILSTL	FNDYDRLRKK	-----	C--NNFQSF	SM-----	---KITQSYV	QD-----	-----	-----
PCHAS_000730	EKINSNITGN	DSYSQMLSNL	FVDYDNLKKN	-----	C--KNSSSLP	TI-----	---KKPQNF	QR-----	-----	-----
PCHAS_070160	QKLN-ITEN	DSYSQILYTL	FNDYNSFKND	--CAK-HCSK	C--IDIPTLP	DI-----	---NLPQNFV	QDKVES----	-----	-----
PCHAS_030020	NQFNE-VIGN	QSYRKILTIL	FNDYDNLKKN	-----	C--KDSSSLP	TI-----	---EQIQYSV	KGSEK-----	-----	-----
PCHAS_000360	QKLN-ITGN	NSYSQILYSL	STDYKNLKND	--CAK-KCTG	C--KDIPTLP	NI-----	---KTPQNSI	ECSAQIRVQD	KGESPGOG--	-----
PCHAS_110020	EDLNQ-ITEN	ISYREILSSL	SIDYDDFKNE	--CAK-NCSY	C--NDIPTLS	EI-----	-----	-----	-----	-----
PCHAS_050060	EGLNQ-ITGN	SSYREILYNL	FNDYDDFKNS	--RAK-NCSS	C--NDLPTLP	KI-----	-----	-----	-----	-----
PCHAS_010030	GTLNN-HKKG	SFYSQLSTL	SKDYNNLKEE	--CDN-GQSK	K----FPSLP	PI-----	---KPTKSYT	HTYGETDIOF	-----	-----
PCHAS_030030	NKLN-ITEN	NLYSQILYAL	FNDYNSFKNG	--CAK-NCSS	C--NDVPTLS	EI-----	-----	-----	-----	-----
PCHAS_000350	EGLNQ-ITGN	SSYREILYNL	FNDYDHLKTD	--CAE-KCTD	C--KDIPTLS	EV-----	-----	-----	-----	-----
PCHAS_011530	ENLNQ-IIGN	NSYREILLSL	STDYNNKFKNG	--CGK-KCID	C--KDMPTLS	EI-----	---KTPQSF-	-----	-----	-----
PCHAS_073160	EELNQ-ITGN	NSYREILSSL	LTNYNNFKSD	--CNK-KCSK	C--SDIPTLT	NI-----	---KAPQSF-	-----	-----	-----
PCHAS_070090	EELNQ-ITGN	NLYREILFSL	SNDYDDFKNY	--CAK-NCSF	C--NNLPTLS	EI-----	-----	-----	-----	-----
PCHAS_000580	ENLNN-ITGN	DSYREILFSL	STNYNNFKND	--CAK-KCSK	C--SEIPTLT	DI-----	---KVPKSYV	QDQVES----	-----	-----
PCHAS_140130	NELNNAITKD	SLYQVLSLTL	SNDYDNFIKK	-----RGAV	C--VDFPPLS	AY-----	-----	-----	-----	-----
PCHAS_140040	KILLNNDIEG	SLYQILSIL	STDYYNFISQ	--CYE-KKVG	C--DKFPPLP	DC-----	-----	-----	-----	-----
PCHAS_000740	KILLNNDIEG	SLYQILSIL	STDYYNFISQ	--CYE-KKVG	C--DKFPPLP	DC-----	-----	-----	-----	-----
PCHAS_100030	EILFNNDTEG	SLYSQILSIL	STDYYNFINL	--CYE-KKEG	C--GRFPCLP	PY-----	-----	-----	-----	-----
PCHAS_001120	EILFNDDTED	SLYSQILSIL	STDYHNFINQ	--CYE-KKKG	C--GRFPPLP	DC-----	-----	-----	-----	-----
PCHAS_011520	EILFNNDTEG	SLYSQILSIL	STDYYNFINL	--CYE-KKEG	C--GRFPPLP	VY-----	-----	-----	-----	-----
PCHAS_070030	EILFNKDTED	SLYSQILSIL	STDYYNFINL	--CYE-KKEG	C--YRFPLLP	VY-----	-----	-----	-----	-----
PCHAS_073190	ETLFNDDTEG	SLYSQILSIL	STDYYNFINQ	--CYE-IKEG	C--GRFPCLP	PY-----	-----	-----	-----	-----
PCHAS_000570	EILLNNGIED	SLYSQILSIL	STDYYNFINQ	--CYK-KKVG	C--GYFPPLP	VY-----	-----	-----	-----	-----
PCHAS_114730	EMLRNNGIEG	SLYSQILSIL	STDYYNFVKK	--CYE-TKEG	C--GRFPCIP	PY-----	-----	-----	-----	-----
PCHAS_010020	QNLLNDGADG	SSYRQMLSTL	STDYDNFKKY	--CDE-KCTG	C--SSIPLLP	TT-----	---KTTHFPA	HI-----	-----	-----
PCHAS_000280	EKHNE-INKD	SLYSQILSTL	LTDYNNFKSY	--FVE-KCSH	C--RDISAFP	VI-----	-----	-----	-----	-----
PCHAS_000490	EKHNE-INKD	SLYSQILSTL	LTDYNNFKSY	--FVE-KCSH	C--RDISAFP	VI-----	-----	-----	-----	-----
PCHAS_000030	QNLLNNNTKD	SPYSQILSTL	STDYTNFKSY	--CAE-KCNG	C--DNIPPLS	YI-----	---EPNPFV	QN-----	-----	-----
PCHAS_000090	QNLLNNNTKD	SPYSQILSTL	STDYTNFKSY	--CAE-KCNG	C--DNIPPLS	YI-----	---EPNPFV	QN-----	-----	-----
PCHAS_000270	QNLLNNNTKD	SPYSQILSTL	STDYTNFKSY	--CAE-KCNG	C--DNIPPLS	YI-----	---EANPFV	QN-----	-----	-----
PCHAS_000120	QNLLNN-NKD	SSYSQILSTL	STDYINFKSY	--CDE-KCGV	C--KEIPSNL	DA-----	---VTSQNSM	GSSGLSHVQD	REENPVVI--	-----
PCHAS_001090	EKLN-NTEN	SSYRKVLYSL	STEYNNLKNK	-----	C--SVFQPLS	DI-----	---KTTQSFV	EGSLES----	-----	-----
PCHAS_011510	EDLNK-DEEN	SSYKLLSTL	SNDYDNLKNI	-----FKNK	C--PNLQPLP	EL-----	---TPKKISV	ETSTPSHVQD	NTHSSLQG--	-----
PCHAS_070050	EDLIK-NKEK	SLYTQILSTL	SDDYKNFIKI	--YHN-KDKS	C--NFQTLP	EL-----	---TPQKISV	EN-----	-----	-----
PCHAS_000720	KDLIK-NKEN	SPFSQILSTL	LDDYNNLINK	-----YGNK	C--SIPQIS	QL-----	---TPQKISV	ESSLQ-----	-----	-----
PCHAS_070020	EELNE-NKGG	SLHTQILSTL	LDDYNNLINK	-----YGNK	C--SIFQSF	EL-----	---TPKNKPV	ESSLQ-----	-----	-----
PCHAS_104260	EELNE-NKEK	SLHTQILSTL	LDDYNNLINK	-----YGNK	C--SISQIS	QL-----	---TAQNKPV	ESSLQ-----	-----	-----
PCHAS_140020	EELNE-IKEK	SLHTQILSTL	LDDYNNLINK	-----YGNK	C--SIFQSF	EL-----	---TPKNKHV	ESSLQ-----	-----	-----
PCHAS_030060	KNLNE-HIID	SSFSQILSTL	LNDHNNLINK	-YCKG--SCS	----FSSIS	QL-----	---TPQKKT	DNSVKDSGET	LGDT-----	-----
PCHAS_000750	KELNN-NIED	SSFSQILSTL	LNDHNNLINK	-----YCKE	S--CSFSSIS	QL-----	---PPKKPV	EGSAKGGEQT	SLPI-----	-----
PCHAS_000680	FEKLNNKEN	SLYTQILSTL	SNDYKNLKKK	-----	C--POFQIS	EI-----	---KRSQSSA	ENSVKVSEQP	TVLI-----	-----

Supplementary data 1 (CIR alignment)

PCHAS_000420	NDLNN-NKEN	SSYNKLLSIL	SDDYNNLQKK	-----	C--TNFPFLP	VY-----	-----	-----	-----	-----
PCHAS_000100	EDLNN-NKEN	SSYNKLLSIL	SDDYNNLQKK	-----	C--TNFPSLP	VY-----	-----	-----	-----	-----
PCHAS_000310	NDLNN-NKEN	SSYNKLLSIL	SDDYNNLQKK	-----	C--TNFPFLP	VY-----	-----	-----	-----	-----
PCHAS_040040	KDLNN-NKEN	SPFSQILYTL	SDDYNNLQKK	-----	C--TNFPSLP	VY-----	-----	-----	-----	-----
PCHAS_120060	EILFNNDTED	SPFSQILYTL	SDDYNNLQKK	-----	C--TNFPSLP	VY-----	-----	-----	-----	-----
PCHAS_104230	EVLLNNGIDD	SLYSQILSIL	STDYNNLRKK	-----	C--TNFPSLP	VY-----	-----	-----	-----	-----
PCHAS_030040	EVLLNNGIDD	SLYSQILSIL	STDYNNLQKK	-----	C--TNFPSLP	VY-----	-----	-----	-----	-----
PCHAS_000340	EILLNNGIDD	SLYSQILSIL	STDYNNLQKK	-----	C--TNFPSLP	VY-----	-----	-----	-----	-----
PCHAS_001110	DKLNE-ITGN	NSYKKILTTL	FDDYGNLKKK	--CDD---DP	S--SDFPTLS	PI-----	---KTTQSF-	-----	-----	-----
PCHAS_120040	EEFNK-ITVN	NSYSKMLSIL	SNDYCNLKD	--YAN-NCSE	C--SNIPDFP	EI-----N	PPKGSPPHNSL	EKSGENRAQD	NVDGSVHSSL	QGHAGNSQHV
PCHAS_000130	EELNK-IARN	DSYSKILSTL	SNDYNYLKNH	--YAT-NCSE	C--SNIPNFP	EI-----	---KTPHFSV	ETSTPSHVQD	NPDSSLQG--	-----
PCHAS_000400	EKLNK-ITKN	ESYSKILSTL	SDDYCNLKD	--YDN-NCNG	C--TNIPNFP	EI-----	---KTSQISV	EGSTPSHVQD	NPDSSLQG--	-----
PCHAS_040020	EKLNK-ITKN	ESYSKILSTL	SDDYCNLKD	--YDN-NCNG	C--TNIPNFP	EI-----	---KTSQISV	EGSTPSHVQD	NPDSSLQG--	-----
PCHAS_000110	NKLNE-TNQS	DSYKNVLYSL	STDYNNLKYY	--LIE-NCTD	C--NDTSSFP	EI-----	---KAPQGSF	QI-----	-----	-----
PCHAS_120050	QKFLNDDTDD	SSYKQILPIL	SNGYDNLIKK	--CNN-GQHS	N---FPPLP	TT-----	---KTTQLSA	HI-----	-----	-----
PCHAS_104250	QKFLNDDTNG	SSYKQILPIL	SNGYDNFKKK	--CNN-TQSS	N---FPLLP	TT-----	---KTTQNVV	EI-----	-----	-----
PCHAS_000320	QKFLNDDTNG	SSYKQILPIL	SNGYDNFKKK	--C---NNTQ	S--SNFPLP	TT-----	---KTTQNVV	EI-----	-----	-----
PCHAS_000410	QTFLNNDIDD	SSYKQILPIL	SNCYDNFKKK	--CNN-TQSS	N---FPLLP	TT-----	---KTTQNVV	EI-----	-----	-----
PCHAS_140030	NELNE-NNQS	NSYKNVLYSL	STDYNNLKYY	--RSE-VCGS	C---YDASF	DI-----	---EAPQGSF	QI-----	-----	-----
PCHAS_114740	NKLNG-KNES	DPYKNVLYSL	STDYDNFKKY	--CAE-NCTD	C--NDTSSFP	EI-----	---KAPQGSF	QI-----	-----	-----
PCHAS_040030	NELNN-ITRN	ISYSQILSNL	FNDYNSFKSY	--RAE-KCSK	C--SNIPSLP	DI-----	---KTAQFSG	QDNAEILVDI	-----	-----
PCHAS_070130	EKLNG-ITGN	SSYRHILYTL	SADYNSLKSD	--CAE-KCTD	C--KDIPILS	EI-----	-----	-----	-----	-----
PCHAS_073180	EKLNG-ITGN	NSYRKILYTL	STDYDDFKNY	--FVE-KCSG	Y--TNLPTLS	KI-----	-----	-----	-----	-----
PCHAS_000300	EKLNG-ITGN	SSYRKILYTL	STDYDDLKND	--FAE-KCSG	Y--SDIPTLS	AI-----	-----	-----	-----	-----
PCHAS_120070	EKLNG-ITGN	NSYRKILYTL	STDYDDFKNY	--FVE-KCSG	Y--TNLPTLS	KI-----	-----	-----	-----	-----
PCHAS_001130	EKLNG-INEN	SSYRKILYTL	STDYDDFKNY	--FVE-KCSG	Y--TDLPTLS	KI-----	-----	-----	-----	-----
PCHAS_100040	EKLNG-INES	SSYRKILYTL	STDYDDFKNY	--FVK-KCSG	Y--TDLPLLS	KI-----	-----	-----	-----	-----
PCHAS_000560	EKLNG-INEN	SSYRKILYTL	STDYDDFKNY	--FVE-KCSD	Y--TYLPTLS	KI-----	-----	-----	-----	-----
PCHAS_050020	EILND-HIKD	SPYSQILSTL	SNDYENLKKK	--CND-AQSS	N---FPPLP	PI-----	---KSTKKA	QYSLESSGLT	SMQS-----	-----
PCHAS_000020	KELND-HIDG	SSYSQILLTL	SKDYDNLKKN	--CDN---GQ	S--SNFPSP	PI-----	---KPTKSST	QNNIEASVOL	-----	-----
PCHAS_000290	EILND-HIKD	SPYSQILLTL	SKDYDNFKNC	--CNK-KKGE	S--CDFPSP	QI-----	---SPKKSFA	QNSLESFGHT	SGHN-----	-----
PCHAS_060020	KELND-HIEG	SSYSQILSTL	SNDYDNFKKT	--CCN-GQSK	N---LPSLP	QI-----	---KPTKISS	QKSLESGGK	SAQKSLESFG	HTYGQN----
PCHAS_001040	KELNE-VTKD	SPYYQVWCTL	STDYDNLKNE	--CIK-NCKE	C--TELPTLP	EI-----	---KESKKYV	QDYAELSAQG	-----	-----
PCHAS_114750	KDLND-VTKD	SPYYQVWCTL	STDYDNLKNE	--CIK-NCKE	C--TELPTLP	EI-----	---KESKKYV	QDYAELSAQG	-----	-----
PCHAS_000140	KDLND-VTKD	SPYYQVWCTL	STDYDNLKNE	--CIK-NCKE	C--TELPTLP	EI-----	---KESKKYV	QDYAELSAQG	-----	-----
PCHAS_000070	KDLNN-VTKD	SPYYQVWSTL	STDYNNLKKK	--CIN----D	Q--SSFPTLP	EI-----	---KEAKKSI	QGYAESSGQF	-----	-----
PCHAS_030080	KKLNE-VAKD	SPYYQVWTL	SNDYDNLKKK	-----CGN	CQSINLPTLP	SI-----	---KLKKISP	PNSVESYGKT	PIDNSVHI--	-----
PCHAS_073200	KKLNG-IDKD	SPYYQVLYIL	SNDYNNLKD	-----	C--SEFQSLP	PI-----	---KIAHSSV	KGHPDNSVHI	-----	-----
PCHAS_120020	EKLNK-ITED	DPYKVLFTL	SDDYVNLKKN	--CDN---GQ	S--SNFPSP	KI-----	---TPKKRSV	QNFAQS----	-----	-----
PCHAS_000060	EKLNK-ITED	DPYKVLCTL	SNDYKLNKKN	--YDN---AE	S--SKFPSP	PI-----	---KPTKSST	RNTVEASGQL	-----	-----
PCHAS_130020	EDLNK-ITKN	NHYYQVWNIL	STGYDNFKKK	--CVN-DQSS	F--PSLPPLP	TI-----	-----EQI	QLPAKCEEL	SIQPTVLS--	-----
PCHAS_000150	EDLSK-ITKD	GYYKVLHTL	STGYDNFKKK	--CDN-----	YQSSFPPLS	TI-----	-----DQI	EFPKCSSEEQ	SVQSTVIS--	-----
PCHAS_000500	KDLVTNGKMC	TSYCDVLSTL	KNSYDKFREE	--IDD-DPEH	K-----LP	EF-----	----IEEGIQ	NCKNLCKSNE	QKLENEKARS	DILETVMDIP
PCHAS_001100	KNLVTNGKVC	DSYCEVLSTL	KNGYDKFREE	--KNE-STPN	F-----QLP	KF-----	----IEDGVE	TCKSLCKSND	QKLEAENLIA	ESEIFTHQ
PCHAS_042010	TSLFNHINTC	DSYLSLLNNL	KAAFEQYKKS	--LVL-SVSN	S---PRKRIN	E-----	---GSVLGFD	-CKE-CKQVH	IKADEKSPKL	GLKALQHEKS
PCHAS_000040	QNLVTEKMC	DSYCYVLSTL	KDAYDKFRRE	---NE-FDSD	Y-----QLP	EF-----	---TLPEGIK	SCEKLCCKKN	QELKVESSKI	DVSETVTPTE
PCHAS_000390	YKDLVSKVKN	CDSYCNVLST	LKNAYDKFRE	E--KIMHDPE	C-----KLP	EF-----	---NVEE-IE	SCESLCKKKS	QEPAIKNPIA	KESEIDTPPK

Supplementary data 1 (CIR alignment)

PCHAS_030180	QGFVKEYGYC	NPYCRILSNL	KKDYENFRKN	-----	---YTPNNLP	EL-----	---NLPNGIP	SCESLCKNKE	QETNTKKPIT	EVSEIATPQK
PCHAS_130070	QGFVKEHDYC	NPYCRILSNL	KKDYENFRKN	-----	---YTTNNLP	EL-----	---NLPNGIT	SCESLCKNKE	QEANAKKPIT	DVSEIDTLQK
PCHAS_011500	QEYAKTHEYC	NPYCNVLSNL	KSDYDKFRET	-----NNNK	N---DLPELK	LL-----	-----DGGE	SCESFCKSKR	QKLNAEKAKP	EDSKIIVTSP
PCHAS_146850	QEFAKTHEYC	HPYCSALSNL	KNDYEFKFRK	-----	---NNNKDLP	EM-----	---ALEDGAK	NCESLCSKE	PKLDIEELET	EDSEMDIFTE
PCHAS_070100	RSCMISTEIC	DPYCRILSNL	KNDYDKIRGN	-----	-----NSKLP	EL-----	---KPPEGRK	SCEIFCEEK	QKLKAEKAKV	EVSEMNI RTE
PCHAS_060050	RTCLNSGELC	NPYCSVLSNL	KKDYEFKFRKN	-----	-----HNNLP	KL-----	---EPPKGRE	NCENYCKSLT	QILNAEKSAI	EQA-----
PCHAS_030110	RTCMNSGELC	SPYCSVLSNL	KKDYENFREA	-----NINK	K-----DLP	EL-----	---KPPEGRE	NCENYCKSLT	QILNGKESAI	GGT-----
PCHAS_130050	RTCLTSRGLC	NPYCSVLSNL	KKDYEKIREN	-----	-----NDKLP	EL-----	---TPPSGRE	SCENYCEILT	QKLNAEGSAI	EGT-----
PCHAS_030090	HSCIIQKEIC	NPYCRVLSNL	KKDYDKFREA	-----NNNK	D-----QLP	TF-----	---TIPEGAK	SCEETCAKKW	QEPEAKEPAI	DVSEIGTP--
PCHAS_146870	RQCIIKIEIC	NPYCSVLSNL	KKDYEKIREN	-----	-----YKNLP	EL-----	---ELTEGLS	DCYGECSKQE	ERHNAGVATQ	NRLGGGSE--
PCHAS_130030	RACLLHREIC	NPYCSVLTNL	KNDYDKLKAK	-----	-----YNLP	EL-----	---TLPOGLS	DCNHECFKQE	ERYKARVAAE	NHLGYGSE--
PCHAS_040110	RACILHREIC	NPYCNVLTNL	KNDYDKIREN	-----	---NRDKKLP	EL-----	---TLPEGLY	DCKTECYKQE	EGHKARVAAL	NRSSDGS---
PCHAS_042030	RACILHREIC	NPYCSVLTNL	KNDYDKLKEK	-----	-----YKLP	EL-----	---KLPEGLY	DCKTECHNOE	ERHKASVAVQ	NRSSDGSEID
PCHAS_060140	RACILHREIC	NPYCSVLTNL	KNDYDKIKEK	-----	-----YNLP	EL-----	---KLPOGLS	DCNHECFKQE	EQYQARVGAQ	NHSSDGSEIG
PCHAS_030210	RACIVQREIC	NPYCSVLTNL	KNDYDKFRKN	-----	---NSDKDIP	EL-----	---ELTEGLS	ECNVECSKQE	KRHKGGGAAQ	NRSSDGSEID
PCHAS_060060	RACLMQREIC	NPYCSVLSNL	KKDYDKLREN	-----	-----NKNIP	EL-----	---ELTEGLS	DCYGECSKQE	KRYEEWLAAQ	KHLSGGSEIV
PCHAS_050070	RICILSKEIC	NPYCSVLSNL	KNDYDKLRKN	-----N	D-----QLP	EL-----	---TLPKQVE	SCENLCNGKQ	KWEFAKLKFE	VSKIDTP---
PCHAS_040050	RACILSEEIC	NPYCNLLSNL	KKDYDKIREK	-----YKT	K-----EIP	EL-----	---TLPVQVE	SCESLCNGKQ	QWKVEKLKIK	ELKTSTL---
PCHAS_001050	RACILSKEIC	NPYCNLLSNL	KNDYDKIKEK	-----	---YKTYELP	EL-----	---TIPDGVE	SCESLCNGKQ	QWEVEELKIK	ELKTSTL---
PCHAS_083760	RTCLITKEIC	NPYCSVLSNL	KKDYDKFRKN	-----	---HNDKNLP	EL-----	---KPPEGRE	SCESFCKSKQ	ESKVEEPTNE	VSEIVTP---
PCHAS_140070	RSCIISKEIC	NPYCSVLSNL	KKDYDKFRKN	-----	---HNDKNLP	EL-----	---KPPEGRE	SCESFCKSKQ	ESKVEEPTNE	VSEIVTP---
PCHAS_000180	NTCLIQKEIC	NPYCSVLSNL	KNDYDKLRKK	-----	-----YVKLP	EL-----	---KLKEGLL	ECNAECSKKN	IQYKDIVDAQ	NNSSGGSKIV
PCHAS_001060	NICHVTKEFC	NPYCRVLSNL	KNDYDNLRKN	-----	-----NDHLP	EL-----	---KPPPGVK	SCEETCAQKC	KEPEAKEPAT	EVSKIIVTP--
PCHAS_011450	RACL MHKEIC	NPYCSVLSNL	KNDYDKIREK	-----	-----YSLP	EL-----	---NLPQGLS	DCNAECSKKA	SARPNGLPG-	-----
PCHAS_060030	RSCLIAKEIC	NPYCNVLTNL	KNDYDKIRKK	-----	-----YSNLP	EL-----	---ELKEGLL	ECNEECSKKA	AT-----	-----
PCHAS_120030	RACIMQREVC	NPYCRVLSNL	KKDYEKIRGK	-----	-----YNNIP	EL-----	---KLTEGLF	ECNVECSKKN	IQYKDIVDAQ	NNSSGGSKIV
PCHAS_030120	QYCVNNKRNY	NLYCSVLSHL	KSDYDKFREN	-----H	N-----NLP	EM-----	---KLPNGIE	SCDSLCEKKK	QESSEKQTS	EQPEQVTTSE
PCHAS_114640	QHCADNKRDC	SPYCNVLSNL	KSDYDKFRET	-----NHNK	N-----DLP	EL-----	---KLPDQGE	SCESYCKSKG	QKLNAEDQVT	EEQKSVTPTE
PCHAS_130350	LKKKSNDNID	EIYSDLYDML	KNVYNDYRKY	--YYE-KNPQ	D--NAPPELP	KI-----	---EEIKKNF	TPDLGASDLQ	DSTEGSPEPQ	DGDGNSEQQQ
PCHAS_030150	EKLVTQSTMC	SSYTTLLNDF	KSSYEFKDG	--HKD----	---INLPELS	CL-----	---KDIKEIS	KPNLEISDSQ	DSTDGNPESQ	DGGSSENSGT
PCHAS_060120	LKKNSNDTRD	SHCRALYDML	EHVYNDFFKKD	-YVEV-MGPQ	T--NELPEIP	KI-----	---EEIKEII	ESKLDESQ	DDVPNSEHIN	STNDLEPOQN
PCHAS_041960	KKSSNNDISD	DQYLAIYDML	HHVYADFRKD	--YLK-NNPE	C--DKIPDLP	NF-----	---KEIEVTP	EPDLDEPDL	ELDPQENTNE	DPETQDGDVD
PCHAS_130110	SKQCNTSGN	-TNFVLYDML	KHVYDDFFKKG	--YK-NNPQ	A--NSFPEIP	EI-----	-----EEIK	EPLISDENL	DSL DSTDENI	ETQDGSSENS
PCHAS_041950	DNLSKSVNQC	GPYLKLEHL	KTIYNAFIKD	-AIKK-NSHD	KICDQLIEFP	SI-----	---DTKKFGSE	FKRTECKKLH	QMLEKNVPM	KDESQEEHDE
PCHAS_060130	YNKFSNVNQC	GPYLRLLNHL	KTIYDEFIN-	AVIKYNNHDQ	SLRSRLIGIS	SI-----	---DKTKFVSE	FNSTGCKQVN	QMLEKKISRL	KNEAQEEQDE
PCHAS_130120	DMLSKSVNQC	DPYIRLLNHL	KTIYNEYIK-	AVIKDN-DHD	ESLRNQLIEL	SSI-----	---NSKLVGSE	FNSEGCKKLH	KKLTQTPPNI	IKMGIKMLD
PCHAS_070170	YNKIYNIKQC	GPYLELLDNL	KEIYNSFIKT	AKSQNSHGDN	ILNQLVELPQ	I-----	---GNINDKYN	FKSKGCEKLH	KKLENNNSKL	IKMGIRMLKG
PCHAS_100060	YSQIYKINKC	DPYLELLDHL	KTPYDNFIKT	AKSKNLHGDD	ILNQLHELPP	I-----	---NKKKFESS	FESPGCKKVH	KRLENNNSKI	IKMGIKMLKD
PCHAS_146770	RTIHNSVNGC	KPYLHLLDSL	KKIYENFILT	KIINN-KNLE	KLSRRIKSLK	TF---KDEDK	YFISDNEELS	FNTEACGKVK	LEDEELGKQI	ASKGSPSKPK
PCHAS_146790	KTTHNAVKSC	KPYLHLLDSL	KRIYEDFRWE	--KII-NNND	SLSNRIESLK	TF---GNEDH	YFVSDSEELS	FNNEGCGKVK	LEDEELGKKI	ASKDSQKQK
PCHAS_060160	KTIHNSVNGC	KPYLHLLDSL	KKVYENFRST	RIINN-KSLD	KLSKRVEPFK	TF---GNEDK	YFVLDSEVLS	FDNEECGKVK	TQDEELGKQI	ASKGSPSKQQ
PCHAS_000260	KTTHNAVKDC	KPYLHLLDSL	KKVYEDFLWD	-KIIN-NSDS	SLSNRIKSLK	TF---GNKDE	YFVSDSEVLS	FDNEGCGKVK	SQDEEIGKKI	TLTNSQKQAQ
PCHAS_130280	RTTHSAVNGC	KPYLHLLDSL	KKMYENFIST	KLIDN-KSFE	KLTRRIEPLK	TF---KNENK	YFVSDNEALS	FDTEACGKVK	SQDEELGKQI	ALTKSQSISK
PCHAS_083720	KTTHKAVKDC	KPYLHLLDSL	KKVYEDFLWN	KIITN-NDSN	ILSNRIKPLK	TF---GNKDE	YFVSDSEVLS	FDNEGCGKVK	SQDEEIGKKI	ASTNSQSISK
PCHAS_090010	RNIYESVKEC	KSYTHLLDSL	KTIYIYIKSY	-KISN-NISI	EKVVLFDATP	SLTTSQYQNK	YFIPNYETLS	FSDEGCGRVK	SKDEKDGKKG	SLPDLQNTLQ
PCHAS_114600	RNIYESVKGK	KSYTHLLDSL	KTIYIYIKSY	-----KITG	NKVLLFVSTP	SLTTSQYQNK	YFIHNYETLN	FGDEGCCKLK	LKDEKDGKKG	SLPDLQSIQK
PCHAS_146860	DEISKFVKQC	GPYIELMHHI	KAIYEEFRQT	--IKE-NIED	PVFSQMLMELP	SI-----	---DKTKSGSE	FKSNKCNKVH	EKI IKKLPKL	IKKEKDKLNG

## Supplementary data 1 (CIR alignment)

PCHAS_130090	NELSKFVNQC	GPYRGLLDHL	KTIYNDFRQT	-AIRE-NAHN	KIFDQILEFP	SI-----	--DKTKYGSE	FESLECKQVH	DKLIKNLPRL	VKKENGELNY
PCHAS_030170	DKISKLIKRC	GPYIELLDHL	RTIYNEFIKR	-VKRE-HHHG	DILDRLMELP	EI-----	--DKTKLGCK	FKTGRCIRLH	QKLEKSNSKI	IKLGIKMLKD
PCHAS_130060	GQISNLIKQC	DPYLELLDHL	RTMYNEFIKS	-VKRE-NLHK	DIHAQLMELP	PI-----	--EKNMFGRN	FKTKRCIRLH	QKLENKNSKL	IKLGIKMLKD
PCHAS_030200	TALSKFVNPC	SPYLELLDHL	KTLYDDYRDA	-AIQN-KIHG	KTPDLLRRFP	EI-----	--DKTTQKFN	FTSPECKKVH	EQLIKNAQKL	IKEEKEKEEG
PCHAS_060080	TELSKFVNPC	SPYLELLDHL	KTLYDDYRED	-AIQN-KIHG	RTPELLRRFP	EI-----	--DNTTKKFN	LKSPECKKVH	EQLIKNAQKL	IKEEKEKEEG
PCHAS_060110	YFINSKAFKC	NPYIELLTNL	KKKYDEYKSL	--IIK-KFPK	NANNFFSDFP	PI-----N	NSNNNHPEQQ	FESKGCCELLH	VIYQRPGRKY	KPKKRQKMLK
PCHAS_030140	HFLKQRVSR	SPYAQLLDVL	KKTYDEYKIL	--VNK-KIEQ	GKFSDFPSIS	NY-----	--KYSPELE	FESSGCKRVH	LFFQKHGTTY	KPKSIMGVFK
PCHAS_130100	YFLNQIRSRC	SPYAQLLDVL	KKTYDEYKIL	--VNK-KIQQ	GDKIKFSELP	PL-----S	NNKNNQPELT	FDNIGCKLVH	LFLQKHGTTY	KPKKILKVPK
PCHAS_041990	YFINSKAFKC	NPYIELLTNL	KKKYDEYKSL	--IIK-KFPK	NANNFFSDFP	PI-----N	NSNNNHPEQQ	FESKGCCELLH	VIYQRPGRKY	KPKKRQKMLK
PCHAS_130080	ELINYQVSNC	DPYIQLFVNL	NKTYDEYRKL	--AIN-KFSK	DLKNILTFLP	II-----	-NNDDQKDLQ	FQSOGCKDLH	AMFEQSYKPR	PKRRQRMPKP
PCHAS_030190	DLINTQVSNC	DPYVQLLINF	KKTYDEYKGL	--VIN-KIPK	NPINNNDDQP	EL-----	-----L	FESNGCKELH	EFFQOIFOKH	KPKQLRRSK
PCHAS_060090	DLINTQVSNC	DPYVQLLINF	KKTYDEYKGL	--VIN-KIPK	NPINNNDDQP	EL-----	-----L	FESNGCKQLH	EFFQOIFOKH	KPKQLRRSK
PCHAS_041970	ELINSDFSDC	DPYVQLFINL	KNIYDKYRNL	--AIK-QIPK	DDSVNSLTRP	EI-----	NNKDNEPOLK	FQSNCKELH	LRFYKLRIKR	KPKKKPRILK
PCHAS_041980	DLINSQFSSC	DPYVQLLINL	KKLYDEYRNL	--AIK-QIPK	DDTVNSLTCP	EI-----	NNNDNQPNLQ	FQSNCKELH	DFFRQISRKR	KPKRPSKGSK
PCHAS_042020	TSLYDEVKEC	YPYFKLLKNF	KKTYDDYRNS	-IMYT-IGNN	SMSFLNLPLT	ELKIPNWSDE	YMDESLEIED	FMTSGCAKLH	SKYLOLEKKN	TSQDTQEQP
PCHAS_060070	TSIYNETKEC	YPYFKLLKNF	KKIYDDYRNN	IIYTI-VSNK	FLTYLNLPLT	ELKIPDWNDE	FMEESLEIDD	FMTSGCAKLH	SKYIEHQOQN	ISQDLPKEQP
PCHAS_011330	TKIYNELNKC	GPYIKLLSNL	KTTYDNFRTA	VINAD-KTKK	LNKVIQELKP	EIKINSATNE	FDDECKKAHS	NAEKNSLKDQ	PKDPPQGPPK	VPPKVPPHVP
PCHAS_000220	TKIYNEINRC	GPYVNLKLNL	KTTYDNFRTS	VISTD-KTKN	LNKVIQELKP	EINIDATKE	FDNECKKVHS	NAEKNNLKDQ	SKDPSKDPK	DPPQGTTPKVP
PCHAS_042070	KNIYNNVKDC	YSYVHLLKYL	KNIYDVIRND	-AIKE-IAAK	KLDASTISLM	DLTTSADWNQI	FLDESDQIID	FHTQKCEVELY	SKTVKQDEEH	ESKIIPKAEP
PCHAS_110030	INIYKDVKEC	YSYFHLLKFL	KNIYDVIRND	-AINE-TVAK	KLNALTISLK	DLTPKDWQOR	FLDESDQIID	LYTKNCNALH	SEISKQAKKD	ISKNPSTEQS
PCHAS_140140	MNINKDIKDC	YSYFHLLKYL	KSIYDGIKND	--AIK-DADI	KLNDLTISLK	DLTPPEWNQR	FSDSDQIID	FHTQKCEVELY	SELAEQVKKD	TPQNAPTEQS
PCHAS_010040	TSIYKDIKDC	YSYFHLLKFL	KNIYDGIKND	--AIKKTNT	KLNTPTISLI	DLTPSNWDKR	FSDSDKIVD	FHTQACINLY	SEFVQVKVKD	TPKNILEGQP
PCHAS_073150	MNIYKNIKDC	YSYFHLLKYL	KNIYDDIRND	--AIK-EAAN	KLDASTISLI	DLTPQGWQOR	FSDVSDQIID	LYTQNCVALH	SKIPQHAKKG	TSKDIPEAHP
PCHAS_073130	MFLYQNVFKC	NSYLHLLDNL	KKIYDSFRAA	--VKN-IDPK	S----AAYLQ	TLTTKENS DL	YFAKNFKTFE	FNGSGCQLQY	DDNIFETLEK	AKLHGNQTN
PCHAS_040060	MLLYQNFSKC	DSYLHLLDNL	KKTYEDFRTT	--TKN-GDPN	L-ASSLQTLT	TI---ENTDS	YLVKGFSTFD	FSNSKCQSEY	DDSILEKSKK	PEARTKQKDN
PCHAS_000430	MLLYQNVSEC	DSYLHLLDNL	KKTYEKFRST	--IKN-GDPN	L-ASSLQTLT	TI---ENTDS	YFVKGFSTFD	FSNSKCQSEY	DDEILEKWRK	DQARTKQKDN
PCHAS_104200	MLLYQNVSEC	DSYLHLLDNL	KKTYEKFRST	--IKN-GDPN	L-ASSLQTLT	TI---ENTGS	YLVEGFKNFD	FSDRKCQSEY	DDSILEKLEK	TKAQRKQKDN
PCHAS_114720	MLLYQNVSEC	DSYLHLLDNL	KKTYEKFRTT	--INN-GDSK	L-ASSLQTLT	TI---ENTDS	YFVEGFKNFD	FSDQKCQSKY	DDSILEKWKE	TQDRSEQKNN
PCHAS_000770	KTLYRNISEC	KSYLRLLNKL	KGIYDDFRNY	-VIOE-NSSK	N--NLETNLK	KLTPNGIEM	DAVRSFISYT	FSKKKCNSLD	KKNTNPPKAG	KSPLOPSNOQ
PCHAS_130220	KILYNNISEC	RSYLRLLNKL	KGIYDDFRVS	-AIEE-SDSK	N--NLETKLQ	KLTRDGVEM	NAVRSFISYK	FSKKKCYSQK	KKPNPGPAPT	ITIQUESPGS
PCHAS_070040	RILYMNIYEC	KSYLHLLNKL	KGIYDDFRVS	-AIKE-TSSN	N--NLETILQ	KLTLGNGTEL	EAVRSFKTYD	FSGSECKLPK	KKNPKPKKTG	SPPLOPTKLE
PCHAS_011490	RTLYINIYEC	KSYLHLLNKL	KGIYDDFRNS	-AINK-NNSN	N--NLATNLQ	KLTKPDGEM	NAVRSFISYK	FNKKICNSLH	KKATTSKPTN	PPGLPPEESH
PCHAS_070060	RTLYINIYEC	KSYLHLLNKL	KGIYDDFRNS	-AINK-NNSN	N--NLATNLQ	KLTKPDGEM	NAVRSFISYK	FNKKICNSLH	KKTTTQKPTN	PPGLSSSSKE
PCHAS_030070	ISIIYNDIHKC	QSYLNLLNKL	KGIYDDFRNY	-AIKE-TDSK	I--NLETNLK	KLTKPDGEM	NAVKGFKSYN	FNNSKCKSLH	NKITMSKPTD	PPGLPRS---
PCHAS_137110	ISIIYNDIPKC	QSYLDLLNKL	KGIYDDFRNY	-AIQD-TDSN	N--NLETDLK	KLTKPDGEM	DAVKGFNYSN	FSNSKCKVKK	KSASLKKEDS	PSLTPSPEQP
PCHAS_114700	RTLYMNVSEC	KPYLHLLSKL	KGLYDDFRSY	-AIKN-TDSN	N--NLVTNLE	KLTLENGEEI	GATKNFTSYN	FSNQPCAKK	KKKTDKPSLQ	SSNQLKDRQQ
PCHAS_137030	RTLYMNVSEC	KPYLHLLSKL	KGLYDDFRSY	-AIKN-TDSN	N--NLVTNLE	KLTLENGEEI	GATKNFTSYN	FSNQPCAKK	KKKTDKPSLQ	SSNQLKDRQQ
PCHAS_011480	RTLYNNIFEC	KSYLDLLNKL	KGLYDDFRSS	-AIKK-NGSK	N--NLETILQ	KLTLGNGTEM	KAVRGFKTYD	FSGSECKFPK	KKKKPEKSSL	QPSNQLKDSK
PCHAS_050040	MILYNLLSEC	KSHLHLLNKL	KYIYDQFREP	--TMK-EIRE	Q--NLSIPLQ	TLTTIAGVEM	RSSKVFKSYN	FSNTKCKRKV	KKKTDKPSLQ	SSSKKESLPP
PCHAS_000470	RTLHMNIYEC	KSYLNLLNKL	KGIYDDFSSA	--IKK-TGSN	N--ELATKLL	KFTLEDGTEM	AAVRGFKTYN	ISNTKCKSLH	KKMPKPKKAV	KSSLQSSSKK
PCHAS_000170	KTLYNNISKC	QSYLDLLNKL	KGIYDDFSSA	--IMR-NRSN	N--NLATKLL	KLTKPDGKEM	DAVRGFKTYN	FSNSKCKFPP	PKKKIVSSSK	AEPGPQTSS
PCHAS_130170	KNLYLNIYKC	KPYLDLLNKL	KGIYDDFSSF	--IKK-NSSN	S--KLAALKQ	TLTPKDGKEM	KAVRGFKTYD	ISNTKCKFPQ	KKITNPKKAD	KSPLOSSSKE
PCHAS_000660	KNLYLNISEC	KPYLDLLNKL	KGIYDDFSSV	--IKK-YNPN	S--ELATKLL	KLTPENGKEM	DAVRGYKTYD	FSTEQCKFPK	KKKTKPKKQD	KSPLOLSNQL
PCHAS_040080	KNLYLNISEC	KPYLDLLNKL	KGIYDDFSSS	--IKK-NDSN	N--ELATKLL	KLTPKDGKEM	DAVRGFKTYN	FSGSQCKFPK	KK-IPKPKPN	PPRLPPSSKE
PCHAS_070070	KTLYLNISKC	KPYLDLLNKL	KGIYDDFSSV	--IKK-NPN	S--ELATKLL	KLTPKDGKEM	EAVRGFKTYN	FSGSQCKFPK	KKNTKPKKAD	KSSLQPSPE
PCHAS_137090	KTLYMNVSEC	KPYLDLLNKL	KGIYDDFRNY	-GIKN-NSSN	N--ELETKLL	KLIPKDGEM	AAVRGFKLYN	FSNQPCGK	KKKTDKSKKA	DSPGLPSSSK

Supplementary data 1 (CIR alignment)

	.... ....  305	.... ....  315	.... ....  325	.... ....  335	.... ....  345	.... ....  355	.... ....  365	.... ....  375	.... ....  385	.... ....  395
PCHAS_000950	ENGEEGEEKT	KESNPVETPG	PVAIENGLIG	IGGVALLIPI	ISTVLYKYWF	ARWRNKSERK	KSVKKVINL-	-----	-----	-----
PCHAS_001010	GNGEGNVSQK	HTTEIQVNPP	KTEL--ELIG	IGGIALLIPV	ISAVLYKYWF	VGWRKKSTRK	KNTKKVINL-	-----	-----	-----
PCHAS_140090	-----	----SDVTSS	SS-IGRKLIP	VFLAF-AISV	FLGIAYKYSL	FGFDKRLQRK	QLREKIKK--	-----	-----	-----
PCHAS_000730	-----	----FEVTSS	NS-IATKLIP	TLIF-AMSF	FFGIAYKYSL	FGFGKRSQKQ	YLKRLK--	-----	-----	-----
PCHAS_070160	-----	----SGVTSP	SS-IAKTLIP	ALLIF-AIPV	FVGIAYKYSL	FGIDKRLQRI	HSRGKLRKI-	-----	-----	-----
PCHAS_030020	-----	----SEFASS	SS-ITIKLIP	VLLTF-TIPF	FLRIAYKYSL	FGFGKRSQKQ	YLGKRLK--	-----	-----	-----
PCHAS_000360	-----	----SEAILS	GS-VASKLIP	ALLVF-AMPV	FLGIAYKYSL	FGFGKRSQKR	YLRENIKK--	-----	-----	-----
PCHAS_110020	-----	----KTPP	SS-IASKLIP	VLLTC-SISF	FLGIAYKYSL	FGFDKRRQRO	YLRENLLK--	-----	-----	-----
PCHAS_050060	-----	----KTT	PT---KLIT	VLLIFAAIPI	SLGIAYKYSL	FGFDKRLQRH	DIRNRIKKI-	-----	-----	-----
PCHAS_010030	-----	----SEVETS	SS-IASKLIP	GLLVF-TIPI	FLGIAYKYSL	FGFGKRSQKQ	YLRKMLK--	-----	-----	-----
PCHAS_030030	-----	----KAPP	SS-IPRKLIP	VLLTF-SIPF	FLGIAYKNSL	FGFHKRVQRK	HLRERLK--	-----	-----	-----
PCHAS_000350	-----	----KAPP	SS-IAKTLIP	VLLTF-AIPV	FLGIAYKYSL	FGFDKRVQRQ	NLREKPKKI-	-----	-----	-----
PCHAS_011530	-----	----EHASS	SS-IASKLIP	GLLIF-AIPI	LLGVAYKYSL	FGFDKRLKRI	HSREKIKKI-	-----	-----	-----
PCHAS_073160	-----	----EHASS	SS-IASTLIP	VLLAF-TIPF	FLGVAYKYSL	FGFDKRLQRO	YIREKVKKI-	-----	-----	-----
PCHAS_070090	-----	----KAPQ	SS-IESKLIP	ALLIF-AIHF	SWELLI-YSL	FGFDKQRHRQ	YLREKLKKI-	-----	-----	-----
PCHAS_000580	-----	----SGVTSP	SS-IASTLIP	GLSTF-AIPA	FLGIAYKYSL	FGIDKLFQRO	YIRKLLKKI-	-----	-----	-----
PCHAS_140130	-----	----SQSSST	KS---AIIS	VTFIFVAIPI	FLGIAYKYSL	FGFDKRLHKQ	YLREKVKKI-	-----	-----	-----
PCHAS_140040	-----	----SRCSSST	KN---ALIS	ITFIFVAISI	FLGFAYKHSL	FGFGKRSQKQ	YLREKRKKI-	-----	-----	-----
PCHAS_000740	-----	----SRCSSST	KN---ALIS	ITFIFVAISI	FLGFAYKHSL	FGFGKRSQKQ	YLREKRKKI-	-----	-----	-----
PCHAS_100030	-----	----SRCSLT	KN-L-----	ITFIFVAIPI	FLGFAYKYSL	FGFGKRSQKQ	YLREKRKKI-	-----	-----	-----
PCHAS_001120	-----	----SR	CS-I--KNAL	ISITFIAIPI	FLGVAYKYSL	FGFGKRSQKQ	YLREKRKKI-	-----	-----	-----
PCHAS_011520	-----	----PR	SF-IKVTLIP	ITFIFVVIPI	FLEFAYKYSL	FGFGKRSQKQ	YLREKRKKA-	-----	-----	-----
PCHAS_070030	-----	----PR	SF-IKVTLIP	ITFIFVAIPI	FLEFAYKYSL	FGFGKRSQKQ	YLREKRKKA-	-----	-----	-----
PCHAS_073190	-----	----SRCSLT	KN-L-----	ITFIFVAISI	FLGFAYKHSL	FGFGKRSQKQ	YLREKRKKI-	-----	-----	-----
PCHAS_000570	-----	----PR	SF-IKMTLIP	ITFIFVAIPI	FLEFAYKYSL	FGFGKRSQKQ	HLREKRKKI-	-----	-----	-----
PCHAS_114730	-----	----SR	CP-IKNALIS	ITFIFVAIPI	FLEFAYKYSL	FGFGKRSQKQ	YLREKRKKA-	-----	-----	-----
PCHAS_010020	-----	----SEDAPS	SS-ITSKLIP	VLLAF-TIPV	FLGIAYKYSL	FGFDKRLQKH	DLRKLK--	-----	-----	-----
PCHAS_000280	-----	----KTTS	SL-IASKLIP	VLLMF-AIPI	FLGIAYKYSL	FGFDKRVQRQ	TLRKKYK--	-----	-----	-----
PCHAS_000490	-----	----KTTS	SL-IASKLIP	VLLMF-AIPI	FLGIAYKYSL	FGFDKRVQRQ	TLRKKYK--	-----	-----	-----
PCHAS_000030	-----	----SEDTSS	SS-AASKLIP	VISIFVAIPI	LLGVAYKYSL	FGFDKGLHRQ	YLREKVKKI-	-----	-----	-----
PCHAS_000090	-----	----SEDTSS	SS-AASKLIP	VISIFVAIPI	LLGVAYKYSL	FGFDKGLHRQ	YLREKVKKI-	-----	-----	-----
PCHAS_000270	-----	----SEDTSS	SS-AASKLIP	VISIFVAIPI	LLGVAYKYSL	FGFDKRLHRQ	YLREKVKKI-	-----	-----	-----
PCHAS_000120	-----	----SEAISS	GS-VSTKLIS	VLLIFSAPIPL	FLGIAYKYSL	FGIDKLFQRO	YLREKVKKI-	-----	-----	-----
PCHAS_001090	-----	----SEAISP	SS-TASKLIA	VLSIFVAMPI	FLGIAYKYSL	FGFGKRSQKQ	YLRDKLKKI-	-----	-----	-----
PCHAS_011510	-----	----SEVTSS	SS-ILNTVIP	VLSTF-AIPV	FLVVSYYKYSL	FGFDKLFQRO	YIRNKLKKV-	-----	-----	-----
PCHAS_070050	-----	----SEATTS	SS-ILNTVIP	VLSTFSVISL	FLGVSYKYSL	FGFGKRSQKR	YLRENIKK--	-----	-----	-----
PCHAS_000720	-----	----TSKGTS	SP-ILNTVIP	VLSTFSVISL	FLGVAYKYSL	FGFGKRSQKR	YLRENIKK--	-----	-----	-----
PCHAS_070020	-----	----SEGASS	SS-MLNTVIP	VLSTFSVISL	FLGVAYKYSL	FGFGKRSQKQ	YLREKRKKI-	-----	-----	-----
PCHAS_104260	-----	----SEGASS	SS-MLNTVIP	VLSTFYVISL	FLGVAYKYSL	FGFGKRSQKQ	YLREKRKKI-	-----	-----	-----
PCHAS_140020	-----	----SEGASS	SS-MLNTVIP	VLSTFSVISF	FLGVAYKYSL	FGFGKRSQKQ	YLREKRKKI-	-----	-----	-----
PCHAS_030060	-----	----PDATPS	SS-TLNTVIP	VLSTFSVISL	FLVVSYYKYSL	FGIDKLFQRO	YRRNKLKKT-	-----	-----	-----
PCHAS_000750	-----	----LEATPS	SS-TLNTVIP	ILSTF-AIPV	FLGVAYKYSL	FGFDKLFQRO	YIRNKLKKI-	-----	-----	-----
PCHAS_000680	-----	----PEATPS	SS-ILNTVIP	VLSTFSVISL	FLGVSYKYSL	FGFGKRSQKR	YLRENIKK--	-----	-----	-----

Supplementary data 1 (CIR alignment)

PCHAS_000420	-----	-----PR	SF-IKVTLP	ITFIFVAIP	FLELAYKYSL	FGFGKRSQKQ	YLREKRKKA-	-----	-----	-----
PCHAS_000100	-----	-----PR	SF-IKVTLP	ITFIFVAIP	FLEFAYKYSL	FGFGKRSQKQ	YLREKRKKA-	-----	-----	-----
PCHAS_000310	-----	-----PR	SF-IKVTLP	ITFIFVAIP	FLELAYKYSL	FGFGKRSQKQ	YLREKRKKA-	-----	-----	-----
PCHAS_040040	-----	-----PR	SF-IKVTLP	ITFIFVAIP	FLGFAYKHSL	FGFGKRSQKQ	YLREKRKKI-	-----	-----	-----
PCHAS_120060	-----	-----PR	SF-IKVTLP	ITFIFVAIP	FLGFAYKHSL	FGFGKRSQKQ	YLREKRKKI-	-----	-----	-----
PCHAS_104230	-----	-----PR	NF-IKVTLP	ITFIFVAIP	FLEFAYKYSL	FGFGKRSQKQ	YLREKRKKA-	-----	-----	-----
PCHAS_030040	-----	-----PR	GF-IKVTLP	ITFIFVAIP	FLEFAYKYSL	FGFGKRSQKQ	YLREKRKKA-	-----	-----	-----
PCHAS_000340	-----	-----PR	NF-IKVTLP	ITFIFVAIP	FLEFAYKYSL	FGFGKRSQKQ	YLREKRKKA-	-----	-----	-----
PCHAS_001110	-----	----EDASS	GS-VASKLIP	VLSIF-AISL	FLGIAYKYSL	FGFDKLFQRO	YIRKKLKKI-	-----	-----	-----
PCHAS_120040	-----	----SEFASS	SS-VASKLIP	VLLTF-AIPI	FLGIAYKYSL	FGIDKLFQRO	YIRKKLKKI-	-----	-----	-----
PCHAS_000130	-----	----SEVTSS	SS-VASKLIP	VLSIF-AISL	FLGIAYKYSL	FGFGKRVHRK	HLREKLKK--	-----	-----	-----
PCHAS_000400	-----	----SEVTSS	SS-VASKLIP	VLSIF-AISL	FVGIAYKYSL	FGIDKLFQRO	YRRNKLKKT-	-----	-----	-----
PCHAS_040020	-----	----SEVTSS	SS-VASKLIP	VLSIF-AISL	FVGIAYKYSL	FGIDKLFQRO	YIRKKLKKI-	-----	-----	-----
PCHAS_000110	-----	----FEATSP	IS-IASILIP	VLLTF-SITF	FLGIAYKYSL	FGFDKRLNRQ	YLREKIKKI-	-----	-----	-----
PCHAS_120050	-----	----SEDTSS	SS-IASKLIP	TLVIF-AIPV	FLGIAYKYSL	FGFDKRFRRK	YSREKLKNK-	-----	-----	-----
PCHAS_104250	-----	----SEATSS	SS-VASKLIP	VLSIF-AISL	FLGIAYKYSL	FGFDKQLQRO	YLREKLKKI-	-----	-----	-----
PCHAS_000320	-----	----SEATSS	SS-VASKLIP	VLSIF-AISL	FLGIAYKYSL	FGFDKQRHRQ	YLREKIKKI-	-----	-----	-----
PCHAS_000410	-----	----SEATSS	SS-VASKLIP	VLSIF-AISL	FLGIAYKYSL	FGFDKQRHRQ	YLREKLKKI-	-----	-----	-----
PCHAS_140030	-----	----FEATPS	SS-IASKLIP	ALLVF-AIPV	FLGVSYKYSL	FGFDKRFQRO	YLREKLKKI-	-----	-----	-----
PCHAS_114740	-----	----FEATSP	IS-IASILIP	VLLTF-SITF	FLGIAYKYSL	FGFDKRLQRO	YLREKLKKI-	-----	-----	-----
PCHAS_040030	-----	----SEDASS	SS-IASKLIP	ALLIC-SIPI	LLGIAYKYSL	FGFDKRLHRQ	YLRDKVKKI-	-----	-----	-----
PCHAS_070130	-----	----KTPP	FS-ITSKLIS	VLLIFVAIPI	SLGIAYKYSL	FGFDKRLHRQ	YLREKLKKI-	-----	-----	-----
PCHAS_073180	-----	----KTLQ	FS-ITSKLIP	VLLAF-AIPI	FLGIAYKYSL	LGFDKRLHIQ	YLREKRKKI-	-----	-----	-----
PCHAS_000300	-----	----KTPL	IL-IARKLIP	VLLAF-AIPI	FLGIAYKYSL	FGFDKRLHIQ	YLREKRKKI-	-----	-----	-----
PCHAS_120070	-----	----KTLQ	FS-ITSKLIP	VLLAF-AIPI	FLGIAYKYSL	LGFDKRLHIQ	YLREKRKKI-	-----	-----	-----
PCHAS_001130	-----	----KTLQ	FS-ITSKLIP	VLLAF-AIPI	FLGIAYKYSL	LGFDKRLHIQ	YLREKRKKI-	-----	-----	-----
PCHAS_100040	-----	----KTLQ	FS-ITSKLIP	VLLAF-AIPI	FLGIAYKYSL	LGFDKRLHIQ	YLREKRKKI-	-----	-----	-----
PCHAS_000560	-----	----KTLQ	FS-ITSKLIP	VLLAF-AIPI	FLGIAYKYSL	LGFDKRLHIQ	YLREKRKKI-	-----	-----	-----
PCHAS_050020	-----	----SEVTSQ	NS-IASKLIP	VLSIFVAIPI	FLGISYKYSL	FGFDKRRHRQ	YLREKVKKI-	-----	-----	-----
PCHAS_000020	-----	----SEYKPS	SS-VENALIP	VLSIFVAIPV	FLGIAYKYSL	FGFDKQRHRQ	YLREKIKKI-	-----	-----	-----
PCHAS_000290	-----	----SEDISS	KS-MANKLIP	GLLIFAAIPV	FLGIAYKYSL	FGFDKQRHRQ	YLREKIKKI-	-----	-----	-----
PCHAS_060020	-----	----SEDISL	KS-IANTLIP	VLSIF-AVPI	FLGIAYKYSL	FGFDKRLHRK	YLREKVQOI-	-----	-----	-----
PCHAS_001040	-----	----SEATSS	SS-IASKLVP	VLSIFVAIPI	FLGFAYKHSL	FGFDKRLNRQ	YLREKVKKI-	-----	-----	-----
PCHAS_114750	-----	----SEATSS	SS-IASKLVP	VLSIFVAIPI	FLGFAYKHSL	FGFDKRLNRQ	YLREKVKKI-	-----	-----	-----
PCHAS_000140	-----	----SEATSS	SS-IASKLVP	VLSIFVAIPI	FLGFAYKHSL	FGFDKRLNRQ	YLREKVKKI-	-----	-----	-----
PCHAS_000070	-----	----SEDASS	SL-IASKLIP	VLLVFGAIP	FLGVAYKYSL	FGLDKRLQRO	YLRDKVKKI-	-----	-----	-----
PCHAS_030080	-----	----SDVTTS	NS-IASKLIP	ILLIF-SIPI	FLGIAYKYSL	FGFDKRVQRO	NFREKLKKI-	-----	-----	-----
PCHAS_073200	-----	----SEDTSS	SS-IASTLIP	ILSIF-SIPI	FLGIAYKYSL	FGFDKRLQRO	YLREKLKKI-	-----	-----	-----
PCHAS_120020	-----	----SEVTTS	SS-VASKLIP	ALLVF-AIPI	FLGISYKYSL	FGFDKRRHRQ	YLREKVKKI-	-----	-----	-----
PCHAS_000060	-----	----SEYKPS	SS-VASKLIP	ALLIC-SIPI	FLGISYKYSL	FGFDKRLHRQ	YLREKLKKI-	-----	-----	-----
PCHAS_130020	-----	----PDAASS	SS-IASKLIP	ALLIF-AIPV	FLGISYKYSL	FGFDKRAHRQ	YLREKIKKI-	-----	-----	-----
PCHAS_000150	-----	----PESAPS	SP-IASKLIP	ALLVF-AIPI	FLGIAYKYSL	FGFDKRLHRQ	YLREKLKKI-	-----	-----	-----
PCHAS_000500	VSL-----	----SDEPPT	EV-LGNKVTY	IVVPFTLILI	ILGILYKYL	HGQRKLLKRR	KNANKIINL-	-----	-----	-----
PCHAS_001100	TSL-PDLSTT	ELRITQEEPQ	SN--GSKLPY	IAVPFILIP	IFAISYKYLT	LMWKKMKSK	KNARKIINL-	-----	-----	-----
PCHAS_042010	EPITKSPKSE	SAQSLSEKSG	SAGAPVGIIG	IGIIGISIFI	ILAIMYKYL	FGSRKNSKKK	KITKKVINLV	DGKKREKTFI	NSIDREKKEK	IIVNS----
PCHAS_000040	TILSYQTSTE	LSDSREESLP	KIEVGNKLPY	IAVPFILIPV	IFGISYKYLT	LMWKKMKSK	KNVRKIINL-	-----	-----	-----
PCHAS_000390	ISLPVRSSTE	LTDNRENTLS	VNEVGNKLPY	IAVPFILIPV	IFGISYKYLT	LMWKKMKSK	KNVRKIINL-	-----	-----	-----

Supplementary data 1 (CIR alignment)

PCHAS_030180	NSLPLLDLFLN	KLKGKEGES	SSGLSNKVIY	IVVAFISILI	ILGISYKYL	LGGVKKLKRK	KMKKIINM-	-----	-----	-----
PCHAS_130070	NSLQLIDFLN	KLKGKEGES	VISVSNKIIY	IVVAFISILI	ILGISYKYL	LVGVKKLKRK	KMKKIINM-	-----	-----	-----
PCHAS_011500	NSLSSLSRKD	GEKNEYDSTT	DKSVSGIIIIY	IVAGFISILI	IFGISYK---	--INNKKIIK	YKIEKIAK--	-----	-----	-----
PCHAS_146850	DS-----	----LSDQPS	TEVIANKLPY	IAVPFILVSI	ILGISYKYL	HGWRKRSNAK	KMKTIINL-	-----	-----	-----
PCHAS_070100	GSLSSQSSTK	LSDSLGNTLP	VKEVTNKLPI	IAIPFFLIPV	ILGISYKYL	HGQRKKSNGK	KKVKTIIINL-	-----	-----	-----
PCHAS_060050	-----	----KQITTH	NIGLGNKLLY	IAIPFILIP	ILGISYKYL	PVWRKMMKRK	TMKKIINSS-	-----	-----	-----
PCHAS_030110	-----	----KQVTTH	KIGLGNKLPY	IAIPFILIP	ILGISYKYIT	YRQRKRLNSK	KKYANDYKF-	-----	-----	-----
PCHAS_130050	-----	----EQVTTH	KISLGNKLPY	IAIPFILIP	ILGISYKYL	PLWRKMMKRK	-TMKKIINL-	-----	-----	-----
PCHAS_030090	-----	----PGISLP	GLPVGKLPY	IAPVLILIP	ILGISYKYL	PVWRKKSRRK	-AMKKIINL-	-----	-----	-----
PCHAS_146870	-----	----KVLTL	NSSLGNKLPY	IAPVLILIP	ILGISYKYL	PVWRKKTTRK	-AMKKIINL-	-----	-----	-----
PCHAS_130030	-----	----KVPTLP	NSSLGNKLPY	IAPVLILIP	ILGISYKYL	PVWRKMMKRK	-TMKKIINL-	-----	-----	-----
PCHAS_040110	-----	----EIDTTP	GISLGNKLPY	IAPVLILIP	ILGISYKYL	PVWRKKTTRK	-AMKKIINL-	-----	-----	-----
PCHAS_042030	TS-----	----TNASLP	GPSVGNKLPY	IAPVLILIP	ILGISYKYL	PVWRKMMKMK	-NMKKIINL-	-----	-----	-----
PCHAS_060140	IL-----	----TKANLP	GQPIGNKLPY	IAPVLILIP	ILGISYKYL	PVWRKKTTRK	-AMKKIINL-	-----	-----	-----
PCHAS_030210	TF-----	----TNASLP	GPSVGNKIPY	IAIPFILIP	ILGISYKYL	PVWRKKSRRK	-AMKKIINL-	-----	-----	-----
PCHAS_060060	TP-----	----TEDSLP	GSSVGNKIPY	IAPVLILIP	ILGISYKYL	PVWRKKTTRK	-AMKKIINL-	-----	-----	-----
PCHAS_050070	-----	----TKVSLP	VPSVANKLSY	IAPVLILIP	ILGFSYKYL	PVWRKKAKRK	-AMKKIINL-	-----	-----	-----
PCHAS_040050	-----	----TEVSLS	VPSVANKLPY	IAPVLILIP	ILGFSYKYL	HSRRKKSNAK	KMKKAIINL-	-----	-----	-----
PCHAS_001050	-----	----TEVSLS	VPSVANKLPY	IAPVLILIP	ILGFSYKYL	PVWRKKLNAK	KKVKAIINL-	-----	-----	-----
PCHAS_083760	-----	----TVVSLP	DSSVGNKLPY	IAPVLILIP	ILGISYKYL	PVWRKKAKRK	-AMKKIINL-	-----	-----	-----
PCHAS_140070	-----	----TVVSLA	DPSVGNKLPY	IAPVLILIP	ILGISYKYL	PVWRKMMKRK	-TMKKIINL-	-----	-----	-----
PCHAS_000180	TP-----	----AVVSLP	SPPVGNKLPY	IAPVLILIP	ILGISYKYL	PVWRKKTTRK	-AMKKIINL-	-----	-----	-----
PCHAS_001060	-----	----AVVSLP	DSSVGNKLPY	IAPVLILIP	ILGISYKYL	PVWRKMMKRK	-TMKKIINL-	-----	-----	-----
PCHAS_011450	-----	----LPVIPT	SINNGNKLPI	IAPVLILIP	ILGISYKYL	PVWRKMMKMK	-TMKKIINL-	-----	-----	-----
PCHAS_060030	-----	----RPNLSP	GLPVGKLPY	IAPVLILIP	ILGISYKYL	HGRRKRLNAK	KKVKAIINL-	-----	-----	-----
PCHAS_120030	TP-----	----AVVSLP	SPPVGNKLPY	IAIPFILIP	ILGISYKYL	PVWRKKAKRK	-AMKKIINL-	-----	-----	-----
PCHAS_030120	DSL-PVRSQI	KLADSREEPQ	SS--GNKLPY	IAPFILIP	ILVISYKFLA	PAWRKRMKMK	-NMKKIINL-	-----	-----	-----
PCHAS_114640	NSLSGSQEGS	VFGIEGDMHE	SNSIGNKLPY	IAPVFLIP	ILGISYKYL	HGRGKKVKKK	KNAKNIINL-	-----	-----	-----
PCHAS_130350	SSSEPPSEAP	PSEAPPTEPA	PSLSISKSV	FSVIVATPI	ILAIMYKYL	YGRKKPKRK	KNMKKVINS-	-----	-----	-----
PCHAS_030150	EN-----	----SDQITP	TNELINNKIK	IGVIALSIP	ILVFIYKYFP	WKRTKKPKKT	KMKRVVNL-	-----	-----	-----
PCHAS_060120	EETTILLDGF	AVPYEITTP	DYEIHQNRK	IGVIALSIP	VLVFIYKYFP	WKRTKKPKKT	KMKKVINLL	DRKTKKIDI	HSIDDKKTIO	-TIINS----
PCHAS_041960	NIPIEEVCI	ENWAPESP	ETELHQNRK	IGVIALSIP	ILVFIYKYFP	WKRTKKPKKT	KMKKVVNLL	DRKTKTIGI	NSIDDKKTMO	ITINSDDK--
PCHAS_130110	DFVE-----	----SDQVSP	ECELDQNRK	IGVIALSIP	VLVFIYKYFP	WKRTKKPKKT	KMKRVVNLL	DRKTKKIDI	NSIDDKKTIO	-T-----
PCHAS_041950	FSTLPEDLSK	QLKSLEEHP	SSSIGQIPNY	VIAIVISIP	ILLIMYKYL	YWRTKKSKGK	TKMKVINL-	-----	-----	-----
PCHAS_060130	LNALKEGDT	MKSIDNTQ	H-LTK-KLIA	FSVIGIAP	TLAIMYKYL	PWRTKKSKRK	TKMKIINL-	-----	-----	-----
PCHAS_130120	DKKRDNTSQ	HERDPINIS	QSSISDKLIA	FSVIGIAP	TLAIMYKYL	PWRTKKSKRK	TKMKIINL-	-----	-----	-----
PCHAS_070170	NAKRVKPGA	PEPAAPPAA	APTNDQDIV	IGIFISTM	VLAFLMYKYL	FSAKNSRKE	KNMKRVINL-	-----	-----	-----
PCHAS_100060	DAKKQIQHGQ	QGTSPTPPPG	P-ISN-DILG	IGIFISIS	VLAFLMYKYL	FSAKNSKKE	KNMKRVINL-	-----	-----	-----
PCHAS_146770	GTITTNLPPA	KPAPANP	RPPMSTDIV	IGIGTSLFV	FLAIMYKYL	FSAKNSKKE	KNMKRVINLV	GGKKRERTFI	NSGDGKKMK	IIINS-----
PCHAS_146790	DPTKSPQKPA	PAKPAPAKPA	TTPTQKDIV	IGIGISIS	FLAIMYKYL	FSAKNSKKE	KSMKRVINS-	-----	-----	-----
PCHAS_060160	GSGNDQSGKV	PTPGKAGVPS	PPPTKDDILG	IGIGAAIFI	FLAFTYKYL	FSAKNLKKE	KSMKSVINLV	SGKKREKRFI	NSGDGKKTEI	IINS-----
PCHAS_000260	DPGNKPAPAK	PAPTSPGMPP	KQPIQNDILG	IGIGAAIFI	FLAFTYKYL	FSAKNSKKE	KSMKSVINLV	SGKKREKRFI	NSGDGKKTEI	IINS-----
PCHAS_130280	GTGDAKPAPA	HTAAGKPTPP	PAPILQDILG	IGIGAAIFI	FLAFTYKYL	FSAKNSKKE	KNMKRVINLV	SGKKREKRFI	NSGDGKKTEI	IINS-----
PCHAS_083720	GTGDAHTAAG	KPTPPPAQPV	PTPIQNDILG	IGIGAAIFI	FLAFTYKYL	FSAKNSKKE	KNMKRVINLV	SGKKREKRFI	NSGNGKKTEI	IINS-----
PCHAS_090010	GPGKSQQGQQ	QSASSGASTP	SPPMSTDIVG	ISVIGVSIS	FLAIMYKYL	FSAKNSKKA	KNMKRVINS-	-----	-----	-----
PCHAS_114600	DKLPAKPAPA	KPVPPPQSP	PSPMASDLVG	ISVIGVSIS	FLAIMYKYL	FSAKNLKKE	KNMKRVINS-	-----	-----	-----
PCHAS_146860	YEDSVELQEE	NREQKVESKD	SHSITDDIVG	IGIGISIFI	VLAFLMYKYL	FGRNNSKKA	KITKVINLV	DGKKMEQTFI	KVDREKKPK	IIVNSDDNKK

## Supplementary data 1 (CIR alignment)

PCHAS_130090	YEEKSKDATP	SKAQQPKQQP	STSIGTNIIG	IGIIGISIFV	FLAFLYKYLP	FGSRKKSXXX	KITKKVINLV	DGRKMEKTFI	KSIDRGKKSX	IIINSGDNKK
PCHAS_030170	DAKRKFLSQL	ENSQQGTSPV	PAIKEKNIIG	IGIIGISIFA	FIAFMFKYLS	FGSRKKSXXX	KITKKVINLV	DGRKMEKTFI	KSIDREKKSX	IIINSGDNKK
PCHAS_130060	DAKRPAQQT	QTHAATKSV	TPGLIENIIG	IGIIGISIFV	FLAFMFYKYL	FGSRKKSXXX	KITKKVINLV	DGRKMEKTFI	KVDREKKSX	IIINSGDNKK
PCHAS_030200	NGKAHNTNAQ	PGKSTPATP	GKSTLNNIIG	IGIIGISIFV	FLAFMYKYLP	FGSRKKSXXX	KITKKVINLV	DGRKMEKTFI	KSIDREKKSX	IIINSGDNKK
PCHAS_060080	KGKTAQPGKK	STPATPGKPP	PGPISDNIIG	IGIIGISIFV	FLAFMFYKYL	FGSRKKSXXX	KITKKVINLV	DGRKMEKTFI	KSIDREKKSX	IIINSGDNKK
PCHAS_060110	PPDPQKDTK	QSAKEKDQPP	VKDIGENIIG	IGIIGISIFV	FLAFMFYKYL	FGSRKKSXXX	KITKKVINLV	DGRKMEKTFI	NSIDREKKSX	IIINSGDNKK
PCHAS_030140	SSSNAKEETI	QSIKKESQPS	EPSVHGDIIG	IGIIGISIFV	FLAFMFYKYL	FGSRKKSXXX	KITKKVINLV	DGRKMEKTFI	KSIDREKKSX	IIINSGDNKK
PCHAS_130100	SSSNTKEETK	QSIKKENQPS	EPSVAPDIIG	IGIIGISIFV	FLAFMFYKYL	FGSRKKSXXX	KITKKVINLV	DGRKMEKTFI	NSIDREKKSX	IIINSGDNKK
PCHAS_041990	PPDPQKDAK	QSTIKEKQSP	EPSLINKVAI	MALMAVSIPI	VLIIMYKYLY	YGCCKTSXXX	KMVKKIINSQ	DGKRRKKKVI	SPIDGKSNLK	TVINSIGGEN
PCHAS_130080	KSNDNSTEKH	NEPQKETQPP	ESSLIKVAI	MALMAVSIPI	ALVIMYKYLY	YGCCKTSXXX	KMVKKIINSH	NGKRTVKKII	NPIDGKRNK	TVINPIDGKN
PCHAS_030190	YSSNSKKDAQ	QSTKKENQPP	ESSLVDKVI	MALMVSIPI	VLAIMYKYLY	YGWGKTSXXX	KMVKKIINSH	NGKRVKKII	NPIDGKRTLK	TVINPIDGKN
PCHAS_060090	YSSNSKKDAQ	QSTKKENQPL	ESSLKVKVI	MALMAVSIPI	ALVIMYKYLY	YGWGKTSXXX	KMVKKIINSH	NGKRKIKKII	NPIDGKRTLK	TVINPIDGKN
PCHAS_041970	SPSNQTLKSK	PENIQEGSPN	NNPVPKAVV	ITLIVVSIPI	VLAIMYKYLY	YRCGKPKPKK	KVKKKIINS-	-----	--IDGKNTK	TVINLIGGKN
PCHAS_041980	SSSNSQKNTD	QSTIKENPPS	EPSLKIKVAI	MGLMAVSIPI	VLIIMYKYLY	YGCCKTSXXX	KIVKKIINSN	NEKRRTKKVI	SPIDGKRTLK	TVINPDRKN
PCHAS_042020	PSGSQKPKSK	PQSPKQEP	PKGLPTNIIG	IGIIGISIFV	FLVFMFKYLS	FGSRKKSXXX	KITKKVINLV	DGKMEKTFI	NSIDREKKEK	MIVNSGDNKK
PCHAS_060070	PSGSRPESE	QOSPQEP	KTGIITNIIG	IGIIGISIFV	FLAFMFYKYL	FGSRKKSXXX	KITKKVINLV	DGRKMEKTFI	KSIDRGKKSX	IIANSQDNKK
PCHAS_011330	PQVPDTSPGK	PPGKPPGKPP	GKQLQNDVVG	IGIIAISTLI	MISIMYKYLS	FGWGKTTTKK	RNMKVINLV	ESKTEKTVT	NSVTGKKTLO	IIINSSSQKK
PCHAS_000220	PQPGQKPKG	QKQPKGTPK	GTQLPSDVVG	IGIIVISTII	MISIMYKYLS	FGWGKTTTKK	RNMKVINLV	ESKTEKTVT	NSVTGKKTLO	IIINSSSQKK
PCHAS_042070	QAGSQNKDGD	NSHTQKQQA	STSLPTASTY	VKVIALLIPI	ILGIMYKCLS	SGWRKELKRR	KSMKVINLS-	-----	--IGGKTTQ	IIINSSTKXX
PCHAS_110030	QSGSTALPPP	PEPQKQDSPT	PPPKTNRPE	TVITVILIPI	ITLIIYKYLS	RERTKKSEKK	-NMKVINL-	-----	--AYGKRKTQ	IIIQCDRTK
PCHAS_140140	QSGGQEEPT	ALPQKQDPS	PSQPEKNRPE	IVITVILIPI	ISLIIYKYLS	RERTKKSEKK	-NMKVINL-	-----	--AYGKRKTQ	IIIKCDRTK
PCHAS_010040	QLGSETLPQ	PIPAQIQPQ	PPGMQKSLTG	VSVTVVLIPI	FLLVAYKFLS	REWTKKSEKK	-SMKVINL-	-----	--ADGNRKTQ	IIIKSYDRNK
PCHAS_073150	QLGSDPPSQ	AETPPQQPS	PQGTEKSLTG	ISVTVVLIPI	FLLAAYKCLS	REWTKKSEKK	-NMKRIINL-	-----	--SDGNRKTQ	IIIKSYDRNK
PCHAS_073130	KTKGELVPGS	QGATSDTTSS	KSGMDKKQVE	ISIIVLLIPI	ALAIMHKYK	FGWRKELKXX	KNMKHIINLF	DVNKAPKTVI	NPINGKRPMQ	IIINSPTQKK
PCHAS_040060	GADTGSQSKE	NAKGSEHTDP	GSDIGNKQVG	ILIIIVILIPI	TLAIMYKYLA	SGRKKELKRR	KNMKVINL-	-----	--MEGKRQMQ	IIIKSSSQKK
PCHAS_000430	GAN-----	----EDNNPQ	NPKVGSKQTR	ILIIIVILIPI	TLAIMYKYLS	FGRRELKRR	KNMKVINL-	-----	--MEGKRQMQ	IIIKSSSQKK
PCHAS_104200	GGNE-----	----KQNTQP	ESSVGSKQTR	ILIIIVILIPI	TLAIMYKYLS	FGRRELKRR	KNMKVINL-	-----	--MEGKRQMQ	IIIKSSSQKK
PCHAS_114720	EGNG-----	----KQSSQP	ESSVGSKQTR	ILIIIVILIPI	TLAIMYKYLS	FGRRELKRR	KNMKVINL-	-----	--MEGKRQMQ	IIIKSSSQKK
PCHAS_000770	KGGQEIKETK	QOSSTESSPQ	SAEVPNRPE	IAITILLIPI	ISLIIYKYFL	SGWRKELKXX	QKMKVINLF	GANKTAKTVI	NSSDGKKQVQ	IIIKSSSQKK
PCHAS_130220	HGV-SNAAGG	QLSNQENTSK	GSGTAQSLTG	ISVTVLVIPI	ILLIVYKYLS	SGWRKEMKRR	KNMKVINLF	GANKTTKTVI	NSTYGKNQVQ	IIIKSSSRKK
PCHAS_070040	STSSLPLTPS	PEMQQDLS	PSQLHNKKIG	IPIIVIIISI	TLAIMYKFLV	FERRKKLRR	K-MKKVPLNF	GVNKT-----	-----	-----
PCHAS_011490	KS--EKFSQS	EPNDSIGPE	GSGTGTKKIG	IPIIVIIISI	TLAIMYKSK	THIKTYKFR	WGKIPLNIY	KLMQALHHL	LINFCDFFC	DFCDFFCDF
PCHAS_070060	VPPPPEESHK	SEKISQSEPN	DSGTDQKIV	IPIIVIIISI	TLAIMYKYL	FDRKKLRR	K-MKKVDFL	GVNKT-----	-----	-----
PCHAS_030070	-----	----SEETPS	SSGVGTKKIG	IPIIVIIISI	TLAIMYKFLV	FDRKKLRR	K-MKKVPLNF	GVNKT-----	-----	-----
PCHAS_137110	KEPSSDSTG	GQDAINIVPA	ASSLGEKKIG	IPIIVIIISI	TLVIMYKSK	AHIKAYKFR	WGKSSLLNIY	KLMQDLHHL	LINFCDFFC	-----
PCHAS_114700	GTPPPETQHL	SSTTPPEEP	TKPIDSKKIV	ISIIIVILIPI	TLTILYKYLS	LGRRELKXX	NMKKVINMV	GVNKTTKTVI	NSSDGKKQIQ	IIIKSSSKXX
PCHAS_137030	GTPPPETQQL	SSTTPPEEP	TKGMDRKKIV	ILIIIVILIPI	TLTILYKYLS	SGWRKEMKXX	KNMTKVINMV	GVNKMVKMVI	NSSDGKKQIQ	IIIKSSSKXX
PCHAS_011480	RETPADIPSV	QDKAQETNPT	TSGMIDSITG	IPIILVLIPI	ILLIVCKYLS	SEWRKEMTRK	KNMTKVINVV	GVNKTVMVI	NSSDGKKQIQ	IIIKSSSQKK
PCHAS_050040	KEPLPETQOS	SSTTPSEEP	AKGVDTSITG	IPIILVLIPI	ILLIVCKYLS	SEWRKEMTRK	KNMTKVINVV	GVNKTVMVI	NSSDGKKQIQ	IIIKSSSQKK
PCHAS_000470	ESLPETQOS	SSTTPPEDPP	KKGIDTSITG	IPIILVLIPI	ILLIVCKYLS	SEWRKEMTRK	KNMTKVINVV	GVNKTVMVI	NSSDGKKQVQ	IIIKSSSKXX
PCHAS_000170	QASGTTQSD	QGLSGGSGS	SSGVSDKKIV	ISIIIVILIPI	TLTILYKYLS	FGRRELKXX	NMKKVINMV	VANKTTKTVI	NSNDGKKQIK	IIIKSSSQKK
PCHAS_130170	EPPPPPSGQ	LKNSQHETPP	SPGMDQKQFV	ILTIVFLIPI	ALAIMYQYLS	FGRKKLKGK	KNMKVINLF	GANKTTKTVI	NSSDGKKQIQ	IIIKSSGQKK
PCHAS_000660	KGRQPETQQL	SSTTPLEEP	AKGIDQKIV	ILIIIVILIPI	TLTIMYKYLS	FGRRELKXX	NMKKVINMV	GVNKTTKTVI	NSSDGKKQIQ	IIIKSYSQKK
PCHAS_040080	EPPPASNSE	GNTGSENRP	SSGMDRQFV	ILTIVLLIPI	ALAIYKYLS	FGRRELKXX	NMKKVINMV	GVNKTTKTVI	NSSDGKKQIQ	IIIKSYSQKK
PCHAS_070070	EPPPKEGASK	RTEHGGTSPP	GPSDQKIG	ISIIIVILIPI	TLAILYKYLS	SGWRKEMKRR	KKMTKVINVA	GVNKTIKTVI	NSSDGKKQIQ	IIIKSSSQKK
PCHAS_137090	KESLDPSDGK	GSQVNGDSE	NSGVPKIG	ISIIIVILIPI	TLAILYTYLS	SGWRKEMKRR	KKMTKVINMV	GVNKTVMVI	NSSDGKKQIQ	IIIKSSGQKK







Supplementary data 1 (CIR alignment)

PCHAS_030180	-----	-----	-CVKNKIKRN	LKYIIENNQR	GQL-----	-----	-----	-----	-----	-----
PCHAS_130070	-----	-----	-CVKNKIKRN	LKYIIENNQR	GQL-----	-----	-----	-----	-----	-----
PCHAS_011500	-----	-----	---DEKGYKF	V-----	-----	-----	-----	-----	-----	-----
PCHAS_146850	-----	-----	-CDVNKTKKE	VTNGSA----	-----	-----	-----	-----	-----	-----
PCHAS_070100	-----	-----	-CDANKIQKD	VTNGFVENI-	-----	-----	-----	-----	-----	-----
PCHAS_060050	-----	-----	-DQKKAQNGF	TNAFIEKNQS	E-----	-----	-----	-----	-----	-----
PCHAS_030110	-----	-----	V-----	-----	-----	-----	-----	-----	-----	-----
PCHAS_130050	-----	-----	-NDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_030090	-----	-----	-SDQKRA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_146870	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_130030	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_040110	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_042030	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_060140	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_030210	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_060060	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_050070	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_040050	-----	-----	-CDENKTKKC	VKNAFIEKNQ	SE-----	-----	-----	-----	-----	-----
PCHAS_001050	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_083760	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_140070	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_000180	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_001060	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_011450	-----	-----	-SDPKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_060030	-----	-----	-CDEKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_120030	-----	-----	-SDQKKA---	-----	-----	-----	-----	-----	-----	-----
PCHAS_030120	-----	-----	-CDKKKAKNG	VTNVFIEKNQ	LE-----	-----	-----	-----	-----	-----
PCHAS_114640	-----	-----	-CEEK-----	-----	-----	-----	-----	-----	-----	-----
PCHAS_130350	-----	-----	IGGKRSVKII	INPSTQKKQT	KKSKN----	-----SVR	GEKMSSLNMY	ELMKADPLPF	INLFFLLIFF	VYKRTRDSIE
PCHAS_030150	-----LDR	KKTKKIDINS	IDGKKTIQIT	INSNDKKKET	KKIIS----	-----SDN	GKTTLLFNII	KOMQLSPMPF	IHLFLLLIF	IFKRKKDSIE
PCHAS_060120	-----NDK	KKTPKRIINS	NDKEKTRKRI	INSNDKGKTR	KRIIN----	-----YNDK	EKTTLLFNII	KOMRLSPMPF	IHLFLLLIF	IFKRKKNSIE
PCHAS_041960	-----	KKTTKKIIDS	NDKEKTPKRI	INSNDKEKTA	KRIIS----	-----SNN	GKTTLLFNII	KOMQLSPMPF	IHLFLLLIF	IFKRKKDSIE
PCHAS_130110	-----	-----IINS	NDKKKPKKRI	ISSNDKEKTT	KRNIS----	-----SDN	GKASLLFNII	KKMQLSPMPF	IHLFLLLIF	IFKRKKDFIE
PCHAS_041950	-----	-----VG	VDKTKKTAIN	-----	-----	-----	-----	-----	-----	-----
PCHAS_060130	-----VEI	NKTKKTVINS	INGKRPMQII	INSSTKKKQT	KKFIT-----	-----SVY	GKNFPLLNIY	NLMQADPVPF	INLFFLLIFF	VYK-----
PCHAS_130120	-----VEI	NKTKKTVINS	INGKRPMQII	ISSSTKKKQT	KKIIT-----	-----SVY	GKNFPLLNIY	KLMEADPLPF	INLFFLLIFF	VYKRKDILLN
PCHAS_070170	-----VDK	KKSIKKGINS	HDGKRKTHIT	INSKHKKKST	KMVIN-LDDG	KKTKIIVNSV	NEKMPSLNIY	KLMQADPMPF	INLFLWLWIFF	IYKRKRDTIE
PCHAS_100060	-----VDK	KKSIEKGINS	RDGKQKTHII	INSKHKKKNT	KMIIN-LGDG	KKTKIIVNSV	NEKMPSLNIY	KLMQADPMPF	INLFLWLWIFF	IYKRKRDTIE
PCHAS_146770	-----DDK	NKSIKHEIYS	HDGKRKTHIT	INSGYKKKNT	KSVIN-SDDR	KTTKIIVNSV	NEKIPLLNIY	KLMHADPMPF	INLFFLLIFF	VYKRKRDTIE
PCHAS_146790	-----	-----	TSGKKQIQII	INSSTKKKQT	KKYIK-----	-----PVY	RGKPPLLNIY	KLMQADPMPF	INLFFLLIFF	VYKRKRDTIE
PCHAS_060160	-----VNE	KKSIKTVINS	RDGKRKTYIT	INSEYKKKYT	KSVIN-FGDG	KKTEIVINSV	NVKIPLLNIY	KLMQADPMPF	INLFFLLIFF	IYKRKRDTIE
PCHAS_000260	-----DDN	KKSIKTVINS	RNGKRKTYIT	INSEYKKKYT	KSVIN-SGDG	KKTEIINSV	NVKIPLLNIY	KLMQADPMPF	INLFFLLIFF	IYKRKRDTIE
PCHAS_130280	-----VNE	KKSIKTVINS	RDGKRKTYIT	INSEYKKKYT	KSVIN-FGDG	KKTEIVINSV	NVKIPLLNIY	KLMQADPMPF	INLFFLWIFF	IYKRKRDTIE
PCHAS_083720	-----VNE	KKSIKTVINS	RDGKRKTYIT	INSEYKKKYK	KSVIN-FGDG	KKTEIVINSV	NVKIPLLNIY	KLMQADPMPF	INLFFLWIFF	IYKRKRDTIE
PCHAS_090010	-----	-----	TSGKKQIQII	IKSSTKKKQT	KKSIK-----	-----PVY	REKSPLLNIY	KLMQADPMPF	INLFFLLIFF	VYKRKRDIIE
PCHAS_114600	-----	-----	TSGKRQIQII	INSSTKKKQT	KKSIK-----	-----PVY	RGKPPLLNIY	KLMQADPMPF	INLFFLLIFF	VYKRKRDTIE
PCHAS_146860	IAKIIIINDDM	DEPIKTVINP	WNEKRKTHAT	INSEYKKKYI	KSVIN-SGP-	-----	-----	-----	-----	-----

Supplementary data 1 (CIR alignment)

PCHAS_130090	IAKIIIINDDT	NKPIKTEINS	RDEKRKTHIT	INSEHAKKYT	KSVIN-SSDR	KKRKIIVNSV	NEKMPLLNIIY	KLMKADPIPF	INLFFLLIFF	VYKRKQDTI-
PCHAS_030170	LAKIIIINDDN	INPIKKVINP	WDEKQMTDIT	INSENTKKYT	KSVIN-SGDR	KTKKIIVNSV	NEKISLLNIY	KFMKADPMPF	INLFFLLIFF	VYKRKRNTIE
PCHAS_130060	IAKIIMNDDT	NKPIKTAINT	WDEKRMTHIT	INSEHTKKYT	KSVIN-SGDR	KTKKIIVNSV	NEKITLLNIY	KLMKADPMPF	INLFFLVIFF	VYKRKRNTIE
PCHAS_030200	IAKIIMNDDN	IKPIKEAINP	WDEKQKTDIT	INSEHTKKYT	KSVIN-SGDR	KTKKIIVNSV	NEKISLLNIY	KFMKADPMPF	INLFFLLIFY	VYKRKRNTIE
PCHAS_060080	IAKIIMNDDN	IKPIKKAINP	WDEKRMTHIT	INSEHTKKHT	KSVIN-SSDR	KTKKIIVNSV	NEKISLLNIY	KFMKADPMPF	INLFFLLIFF	VYKRKRSTIE
PCHAS_060110	IAKIIMNDDN	IKPIKKAIDP	WDEKQMTDIT	INSEHTKKYT	KSVIN-SSDR	KTKKIIVNSV	NEKISLLNIY	KFMKADPIPF	INLFFLLIFF	VYKRKYNFL-
PCHAS_030140	IAKIIMNDDN	IKPIKKAIDP	WDEKQMTDIT	INSEHTKKYT	KSVIN-SSDR	KTKKIIVNSV	NEKISLLNIY	KFMKADPMPF	INLFFLLIFF	VYKRKRNTIE
PCHAS_130100	IAKIIMNDDN	IKPIKKAIDP	WDEKQMTDIT	INSEHTKKYT	KSVIN-SSDR	KTKKIIVNSV	NEKISLLNIY	KFMKADPIPF	INLFFLLIFF	VYKRKYNFL-
PCHAS_041990	TSNTIISIDD	ENTTNTIISP	NYEETNVKTI	IDSDSGEKT	IVIINSYDEK	NITIHSIKSS	PPKITALNAY	KHIFTNPAPP	ISLFFLLIFF	FYK-----
PCHAS_130080	TSNIIISIDE	ENTENATISP	NCEETNVKTI	INDSCEKKT	IVIINSYDEK	DVTIQSVKSS	SPKTTSLKGY	KHIFANPAPP	INLFFLLIFF	VYKKKYNFL-
PCHAS_030190	TSNIIISIDE	ENTENATISP	NCEETNVKTT	INDSCEKKT	IVIINSYDEK	DVTIQSVKSS	SPKTTSLKGY	KHIFANPAPP	INLFFLLIFF	VYKRKYNFL-
PCHAS_060090	ISNTIISIDE	KNTENATISP	NSEETNVKTT	IDSDCEEKTT	ILIINSYDEK	DVTIQSVKSS	SPKTTSLKGY	KHIFANPAPP	INLFFLLIFF	VYKRKYNFL-
PCHAS_041970	ITIQNIK---	-----	-----	-----	-----	-----FS	SPKTTSLNAY	KHIFSNSAPP	INLFFLLIFF	VYKRKYNSL-
PCHAS_041980	TATTVIN---	-----	-----	-----	-----LIGGK	NIMIQSIKSS	FHQAISLNAY	KHIFSNSAPP	INLFFLLIFF	VYKGYISL-
PCHAS_042020	IAKIIIN---	-----	-----	-----	-----	-----PGD	GKKNPLLNIIY	KLMQADPMPF	INLFFLLIFF	VYKRKSTIE-
PCHAS_060070	IAKIIIN---	-----	-----	-----	-----SGP--	-----	-----	-----	-----	-----
PCHAS_011330	QTKKSIN---	-----	-----	-----	-----	-----SV	NEKIPLLNIY	KLMQADPVPF	INLFFLLIFF	VYKRKRDVIE
PCHAS_000220	QTKKSIN---	-----	-----	-----	-----	-----LV	NEKIPLLNIY	KLMQADPAPP	INLFFLLIFF	VYKRKDNSIE
PCHAS_042070	QSKKFIT---	-----	-----	-----	-----	-----SVY	RKNFPLLNIY	TLMEADPVPF	INLFFLLIFC	VYKRKLNIFE
PCHAS_110030	NLKPVIN---	-----	-----	-----	-----	-----SVD	RKKDSSLNIH	KLMQADSIPF	INLFFLLIFL	SIKENTIFWN
PCHAS_140140	DLKPVIN---	-----	-----	-----	-----	-----SVD	RKKDSSLNIP	KLMQANSIPF	INLFFLLIFL	SIKENTIFWN
PCHAS_010040	DLKPVIN---	-----	-----	-----	-----	-----SVS	RKKYPLLNIY	KLMQADPVPF	INLFFLLIFF	VYKENMIL--
PCHAS_073150	DLKPVIN---	-----	-----	-----	-----	-----PVG	RKKRPLLDIY	KLMQADPIPF	INLFFLLIFF	VYKRKYDFLE
PCHAS_073130	QIKKSIN---	-----	-----	-----	-----	-----PVY	RERFPLLNIY	KLMQADPVPF	INLFFLLTFL	FIKEKTILWN
PCHAS_040060	QTKKSIN---	-----	-----	-----	-----	-----PVY	GEKSPSINIY	KLMQADPVPF	INLIFLLIFL	FIKEKTILWK
PCHAS_000430	QTKKSIN---	-----	-----	-----	-----	-----PVY	GEKSPSINIY	KLIQADPVPF	INLIFLVIFL	FIKEKTILWK
PCHAS_104200	QTKKSIN---	-----	-----	-----	-----	-----PVY	GEKSPSINIY	KLMQADPVPF	INLIFLLIFL	FIKEKTILWK
PCHAS_114720	QTKKSIN---	-----	-----	-----	-----	-----PVY	VEKSPSINIY	KLMQADPVPF	INLIFLVIFL	FIKEKTIFWK
PCHAS_000770	QTKKSIN---	-----	-----	-----	-----	-----FVN	RKKPPFLNIY	KLMMADAIPF	INLFFLLIFF	VYKRKRNTIE
PCHAS_130220	KTKKSIN---	-----	-----	-----	-----	-----SIY	GEKFPLNVVY	KLMQADPAPP	INLFFLLIFF	VYKRKENTIE
PCHAS_070040	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
PCHAS_011490	-----	-----	-----	-----	-----	-----FVDFVVI	FLLIFFVDFV	VIFLLIFFVI	FFVNFFCQFF	LFIKENETL-
PCHAS_070060	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
PCHAS_030070	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
PCHAS_137110	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
PCHAS_114700	KTKKSIN---	-----	-----	-----	-----	-----SVY	GEKSLSLKIY	QLMQADPVPF	INLIFLLIFF	VYKRKRDFIE
PCHAS_137030	KTKKSIN---	-----	-----	-----	-----	-----SVY	GEKSPSLNIY	QLMQADPVPF	INLFFLLIFF	VYKRKRDFIE
PCHAS_011480	QTKKSIN---	-----	-----	-----	-----	-----SVN	RKKSPSLNMY	QLMQADPVPF	INLIFLLIFF	VYKRKRDFIE
PCHAS_050040	QTKKSIN---	-----	-----	-----	-----	-----FVY	GEKSPSLKIY	QLMQADPVPF	INLIFLLIFF	VYKRKRDFIE
PCHAS_000470	KTKKSIN---	-----	-----	-----	-----	-----SVY	GEKSPSLNIY	QLMQADPVPF	INLIFLLIFF	VYKRKRDFIE
PCHAS_000170	QTKKFIN---	-----	-----	-----	-----	-----FVY	GEKSPSLNIY	QLMQADPVPF	INLFFLLIFF	VYKRKRDFIE
PCHAS_130170	QTIKSIN---	-----	-----	-----	-----	-----SVY	GEKSPSLNIY	QLMQADPVPF	INLFFLLIFF	VYKRKENTIE
PCHAS_000660	KTKKFIN---	-----	-----	-----	-----	-----FVY	GEKSPSLNIY	QLMQADPVPF	INLIFLLIFF	VYKRKRDFIE
PCHAS_040080	PTKKSIN---	-----	-----	-----	-----	-----SVN	RKKSPSLNIY	QLMQADPVPF	INLFFLLIFF	VYKRKRDFIE
PCHAS_070070	QTKKSIN---	-----	-----	-----	-----	-----SVN	RKKSPSLNIY	QLMQADPVPF	INLIFLLIFF	VYKRKRDFIE
PCHAS_137090	QTKKSIN---	-----	-----	-----	-----	-----SVN	RKKSSSLNIY	QLMQADPVPF	INLFFLLIFF	VYKRKRDFIE

## b) CIRs excluded from the amino acid sequence alignment:

Reason for exclusion	<i>cir</i> genes
Partial <i>cir</i> gene	PCHAS_000080 PCHAS_000330 PCHAS_001070
Poor alignment	PCHAS_042000 PCHAS_010120 PCHAS_030160 PCHAS_030270 PCHAS_060100 PCHAS_100110 PCHAS_104110 PCHAS_093820 PCHAS_083690 PCHAS_104270
CIR-like genes	PCHAS_146840 PCHAS_090030 PCHAS_052480