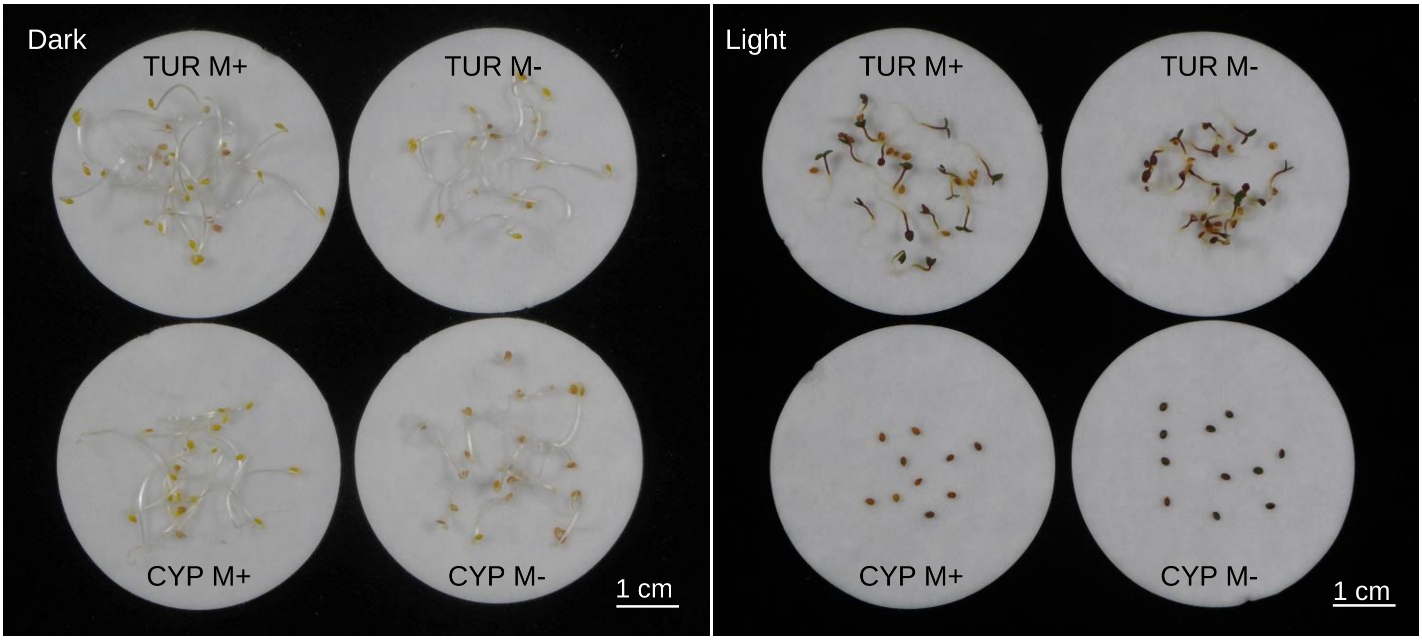
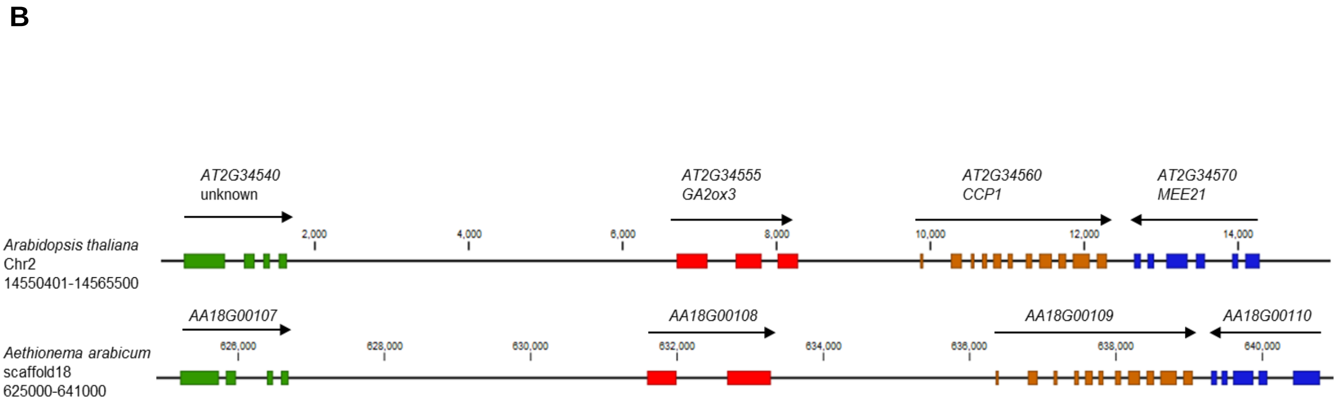


Supporting Figure 1. Germination of dimorphic seed types in response to light. Germination of mucilaginous (M+) and non-mucilaginous (M-) seeds were tested from TUR and CYP accessions in dark or under white light (100 µmol m^{-2} s^{-1}). Images were taken 7 days after imbibition.

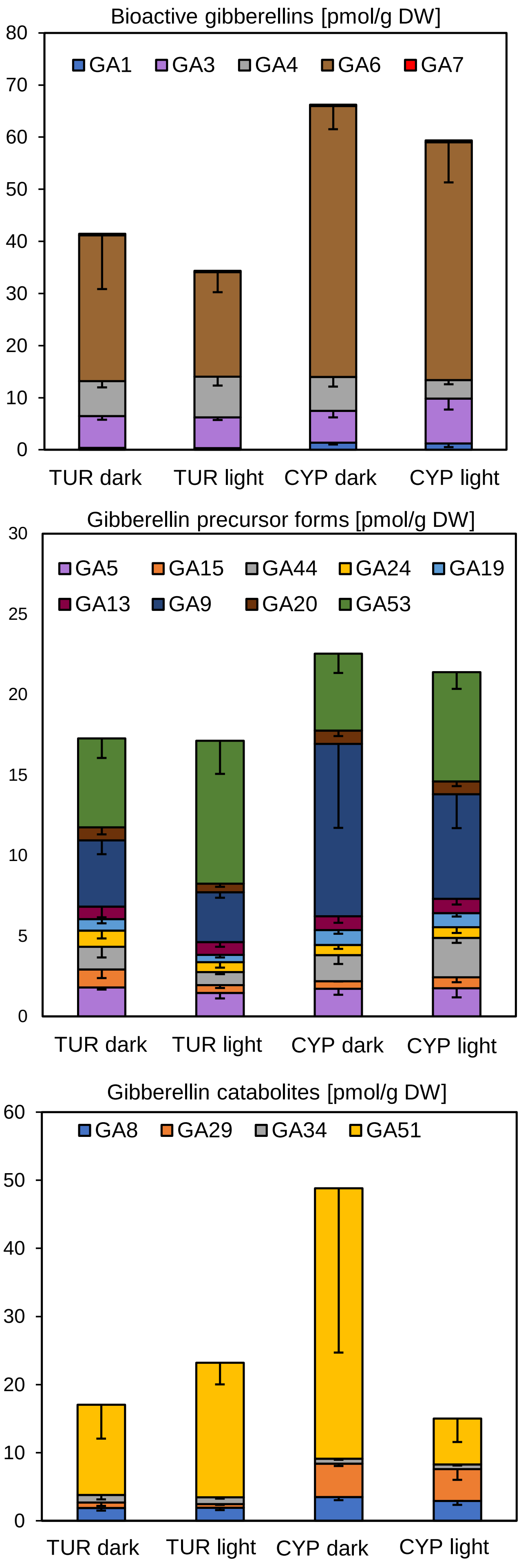
Aethionema accession	TAIR10DESCRIPTION
AA31G00304	ARATHwo_AT1G76600.1
AA18G00108	gibberellin 2-oxidase
AA18G00131	NA
AA6G00020	delay of germination 1
AA7G00013	alpha/beta-Hydrolases superfamily protein
AA21G00250	Matrixin family protein
AA96G00113	cytochrome P450, family 94, subfamily B, polypeptide 3
AA6G00035	Protein of unknown function, DUF584
AA31G00340	ARATHwo_AT1G76250.1
AA118G00057	ARATHwo_AT4G33980.2
AA9G00199	RING/U-box superfamily protein
AA53G01387	Integrase-type DNA-binding superfamily protein
AA35G00052	DRE-binding protein 2A
AA659G00001	RING/U-box superfamily protein
AA4G00240	Ethylene insensitive 3 family protein
AA13G00009	flavanone 3-hydroxylase
AA44G00071	Protein of unknown function (DUF1118)
AA60G00062	plasma membrane intrinsic protein 2
AA11G00062	UDP-Glycosyltransferase superfamily protein
AA65G00216	Protein of unknown function, DUF584
AA24G00057	ARATHwo_AT3G44450.1
AA4G00193	NA
AA62G00022	Histone superfamily protein
AA26G00398	sequence-specific DNA binding transcription factors
AA33G00153	basic leucine-zipper 2
AA57G00148	Leucine-rich repeat (LRR) family protein
AA26G00194	rhamnose biosynthesis 3
AA65G00256	extensin-like protein
AA39G00541	ADP-ribosylation factor A1F
AA40G00565	GAST1 protein homolog 4
AA297G00002	arabinogalactan protein 31
AA44G00506	Leucine-rich repeat (LRR) family protein
AA10G00399	UDP-glucose 6-dehydrogenase family protein
AA53G00911	Ferritin/ribonucleotide reductase-like family protein
AA31G00563	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AA466G00003	Eukaryotic aspartyl protease family protein
AA26G00651	cell elongation protein / DWARF1 / DIMINUTO (DIM)
AA44G00016	calnexin 1
AA53G01423	leucine-rich repeat/extensin 1
AA30G00236	NA
AA7G00012	beta-xylosidase 2
AA44G00068	RPA70-kDa subunit B
AA90G00040	UDP-glucosyl transferase 71C3
AA32G00601	homolog of yeast CDT1 A
AA21G00074	phy rapidly regulated 1
AA21G00008	uv-b-insensitive 4
AA428G00003	cytochrome P450, family 83, subfamily B, polypeptide 1
AA108G00006	Pectin lyase-like superfamily protein
AA17G00061	Protein of unknown function (DUF581)
AA805G00002	PYR1-like 4
AA398G00001	FASCICLIN-like arabinogalactan 2
AA32G00222	S-adenosylmethionine synthetase 2
AA44G00517	ARATHwo_AT5G13100.1
AA26G00318	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
AA53G01380	O-methyltransferase family protein
AA8G00033	SKU5 similar 17
AA26G00339	GDSL-like Lipase/Acylhydrolase superfamily protein
AA21G00251	GATA transcription factor 2
AA8G00303	FRA8 homolog
AA35G00022	expansin A2
AA8G00179	Plant invertase/pectin methylesterase inhibitor superfamily protein
AA8G00110	sulfurtransferase 18
AA26G00453	purple acid phosphatase 17
AA61G00301	phy rapidly regulated 2
AA26G00278	Ran BP2/NZF zinc finger-like superfamily protein
AA6G00288	RING/U-box superfamily protein
AA25G00031	ARATHwo_AT5G54970.1
AA32G00244	Heat shock protein DnaJ with tetratricopeptide repeat
AA26G00148	Subtilase family protein
AA87G00204	NA
AA134G00041	Pectate lyase family protein
AA18G00045	ovate family protein 6
AA56G00017	beta HLH protein 71
AA228G00013	ARATHwo_AT4G28310.1
AA30G00211	SKU5 similar 15
AA25G00163	expansin A14
AA21G00485	AtL5
AA79G00031	ARATHwo_AT4G35320.1
AA166G00005	Leucine-rich repeat (LRR) family protein
AA61G00462	FASCICLIN-like arabinogalactan-protein 10
AA8G00274	Stress responsive A/B Barrel Domain
AA19G00337	Glycosyl hydrolase superfamily protein
AA19G00338	Glycosyl hydrolase superfamily protein
AA53G01128	zinc finger protein 4
AA53G01281	early nodulin-like protein 8
AA21G00283	protein binding
AA20G00061	proline-rich protein 4

Supporting Figure 2. Heatmap of all 87 genes light-regulated in CYP seeds and differentially expressed in light-exposed TUR and CYP seeds based on RPKM values.

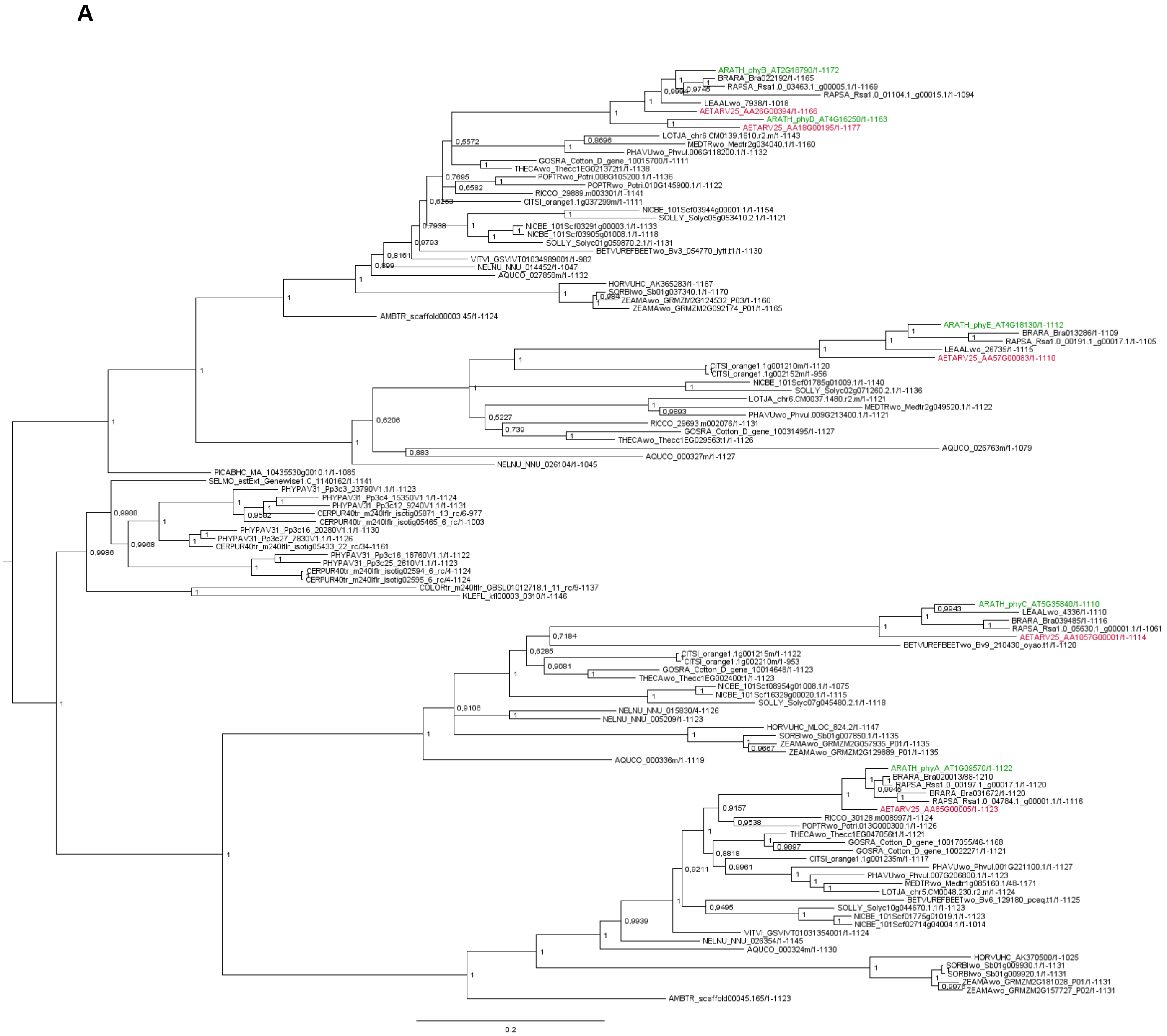
Supporting Figure 3



Supporting Figure 3. Identification of the *A. thaliana* orthologue of *Ae. arabicum* AA18G00108 as *GA2ox3.*
(a) Phylogenetic tree of gibberelin2-oxidases using Bayesian inference (2000000 generations, standard deviation of split frequencies 0.063371) and allows clear assignment of *Ae. arabicum* orthologues. Sequences of *A. thaliana* (ARATH) and *Ae. arabicum* (AETAR) are marked in green and red, respectively. For detailed assignment of five letter code see Supplemental Dataset 7. (b) Synteny of *GA2ox3* position in the genome of *A. thaliana* and *Ae. arabicum.*

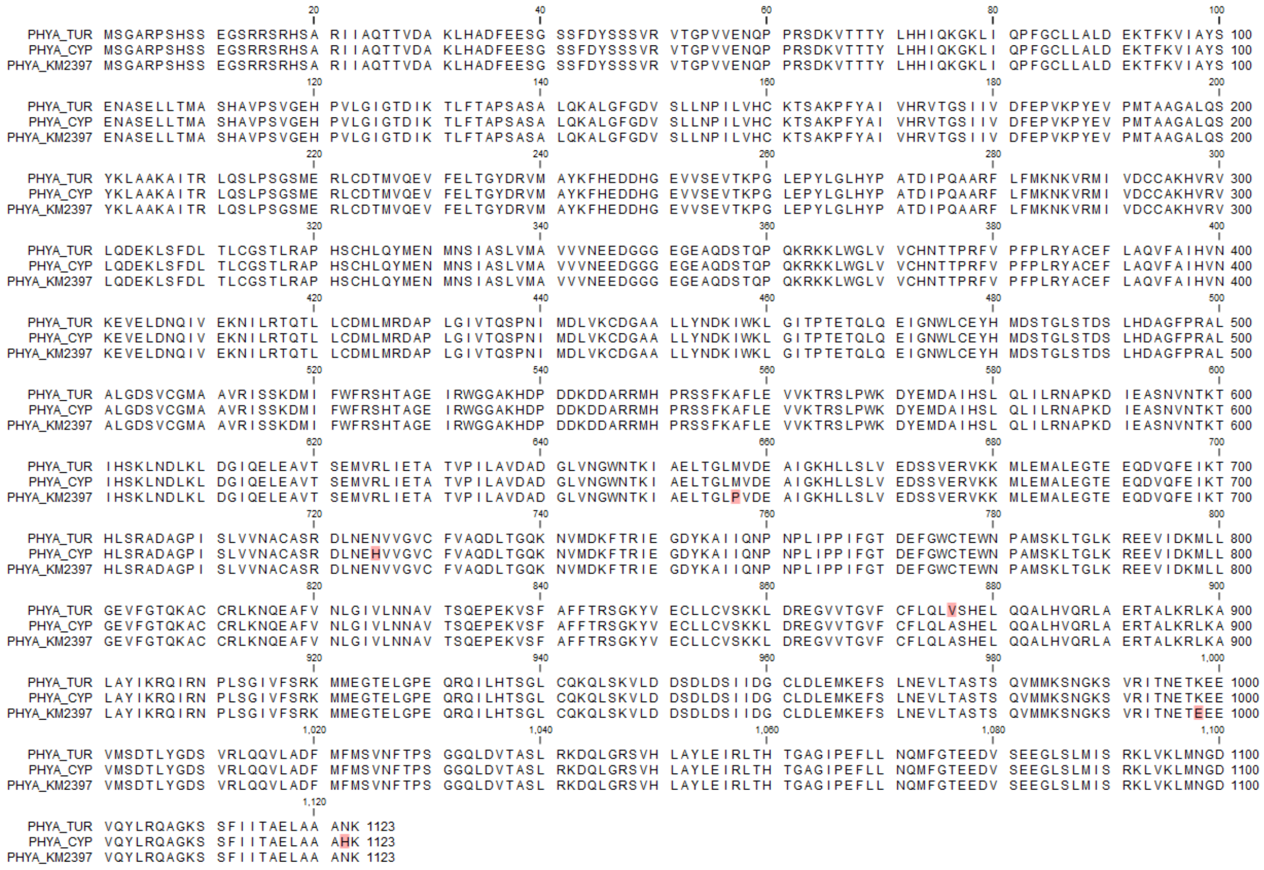


Supporting Figure 4. Accumulation of GA forms in *Ae. arabicum* TUR and CYP seeds under dark and light conditions.



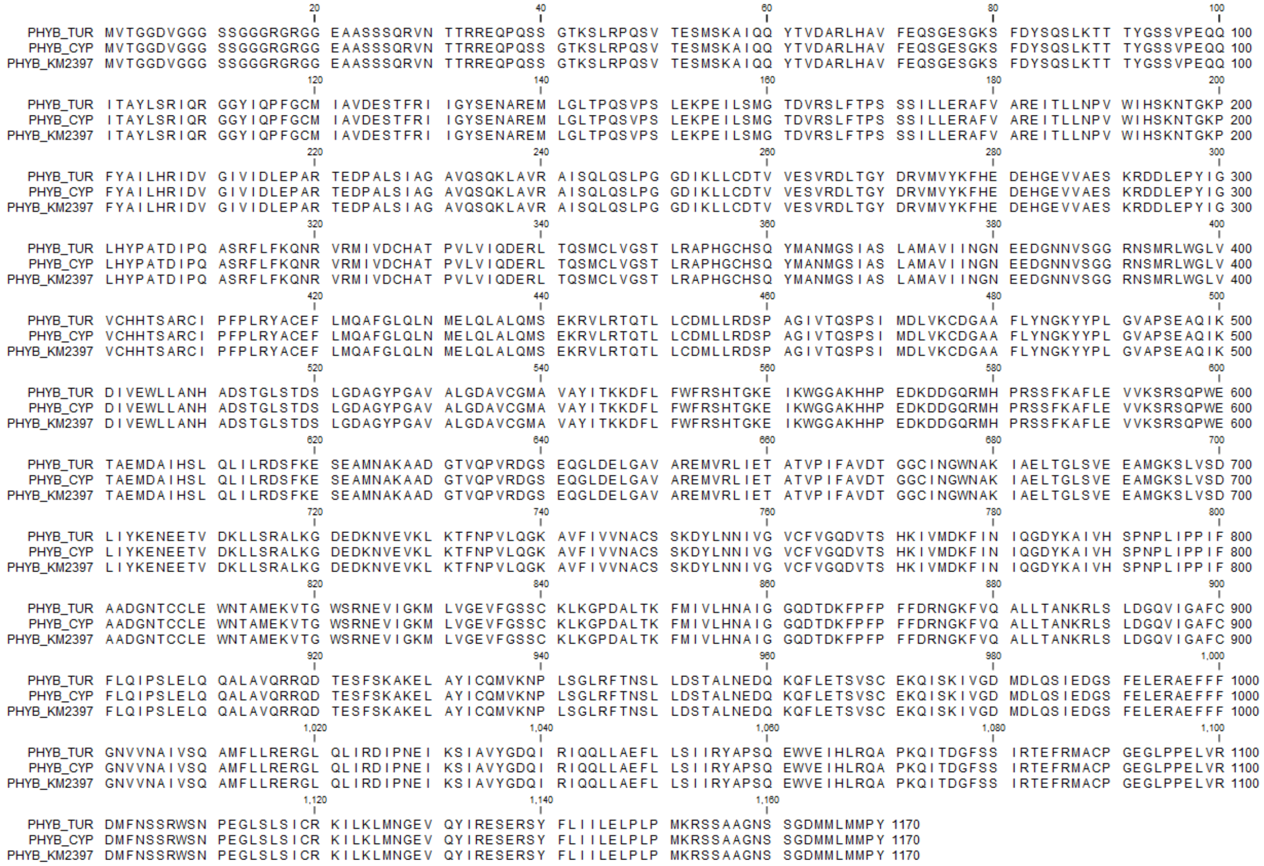
Supporting Figure 5

B



```
                               20                    40                    60                    80                   100
   PHYA_TUR MSGARPSHSS EGSRRSRHSA RIIAQTTVDA KLHADFEESG SSFDYSSSVR VTGPVVENQP PRSDKVTTTY LHHIQKGKLI QPFGCLLALD EKTFKVIAYS 100
   PHYA_CYP MSGARPSHSS EGSRRSRHSA RIIAQTTVDA KLHADFEESG SSFDYSSSVR VTGPVVENQP PRSDKVTTTY LHHIQKGKLI QPFGCLLALD EKTFKVIAYS 100
PHYA_KM2397 MSGARPSHSS EGSRRSRHSA RIIAQTTVDA KLHADFEESG SSFDYSSSVR VTGPVVENQP PRSDKVTTTY LHHIQKGKLI QPFGCLLALD EKTFKVIAYS 100
                              120                   140                   160                   180                   200
   PHYA_TUR ENASELLTMA SHAVPSVGEH PVLGIGTDIK TLFTAPSASA LQKALGFGDV SLLNPILVHC KTSAKPFYAI VHRVTGSIIV DFEPVKPYEV PMTAAGALQS 200
   PHYA_CYP ENASELLTMA SHAVPSVGEH PVLGIGTDIK TLFTAPSASA LQKALGFGDV SLLNPILVHC KTSAKPFYAI VHRVTGSIIV DFEPVKPYEV PMTAAGALQS 200
PHYA_KM2397 ENASELLTMA SHAVPSVGEH PVLGIGTDIK TLFTAPSASA LQKALGFGDV SLLNPILVHC KTSAKPFYAI VHRVTGSIIV DFEPVKPYEV PMTAAGALQS 200
                              220                   240                   260                   280                   300
   PHYA_TUR YKLAAKAITR LQSLPSGSME RLCDTMVQEV FELTGYDRVM AYKFHEDDHG EVVSEVTKPG LEPYLGLHYP ATDIPQAARF LFMKNKVRMI VDCCAKHVRV 300
   PHYA_CYP YKLAAKAITR LQSLPSGSME RLCDTMVQEV FELTGYDRVM AYKFHEDDHG EVVSEVTKPG LEPYLGLHYP ATDIPQAARF LFMKNKVRMI VDCCAKHVRV 300
PHYA_KM2397 YKLAAKAITR LQSLPSGSME RLCDTMVQEV FELTGYDRVM AYKFHEDDHG EVVSEVTKPG LEPYLGLHYP ATDIPQAARF LFMKNKVRMI VDCCAKHVRV 300
                              320                   340                   360                   380                   400
   PHYA_TUR LQDEKLSFDL TLCGSTLRAP HSCHLQYMEN MNSIASLVMA VVVNEEDGGG EGEAQDSTQP QKRKKLWGLV VCHNTTPRFV PFPLRYACEF LAQVFAIHVN 400
   PHYA_CYP LQDEKLSFDL TLCGSTLRAP HSCHLQYMEN MNSIASLVMA VVVNEEDGGG EGEAQDSTQP QKRKKLWGLV VCHNTTPRFV PFPLRYACEF LAQVFAIHVN 400
PHYA_KM2397 LQDEKLSFDL TLCGSTLRAP HSCHLQYMEN MNSIASLVMA VVVNEEDGGG EGEAQDSTQP QKRKKLWGLV VCHNTTPRFV PFPLRYACEF LAQVFAIHVN 400
                              420                   440                   460                   480                   500
   PHYA_TUR KEVELDNQIV EKNILRTQTL LCDMLMRDAP LGIVTQSPNI MDLVKCDGAA LLYNDKIWKL GITPTETQLQ EIGNWLCEYH MDSTGLSTDS LHDAGFPRAL 500
   PHYA_CYP KEVELDNQIV EKNILRTQTL LCDMLMRDAP LGIVTQSPNI MDLVKCDGAA LLYNDKIWKL GITPTETQLQ EIGNWLCEYH MDSTGLSTDS LHDAGFPRAL 500
PHYA_KM2397 KEVELDNQIV EKNILRTQTL LCDMLMRDAP LGIVTQSPNI MDLVKCDGAA LLYNDKIWKL GITPTETQLQ EIGNWLCEYH MDSTGLSTDS LHDAGFPRAL 500
                              520                   540                   560                   580                   600
   PHYA_TUR ALGDSVCGMA AVRISSKDMI FWFRSHTAGE IRWGGAKHDP DDKDDARRMH PRSSFKAFLE VVKTRSLPWK DYEMDAIHSL QLILRNAPKD IEASNVNTKT 600
   PHYA_CYP ALGDSVCGMA AVRISSKDMI FWFRSHTAGE IRWGGAKHDP DDKDDARRMH PRSSFKAFLE VVKTRSLPWK DYEMDAIHSL QLILRNAPKD IEASNVNTKT 600
PHYA_KM2397 ALGDSVCGMA AVRISSKDMI FWFRSHTAGE IRWGGAKHDP DDKDDARRMH PRSSFKAFLE VVKTRSLPWK DYEMDAIHSL QLILRNAPKD IEASNVNTKT 600
                              620                   640                   660                   680                   700
   PHYA_TUR IHSKLNDLKL DGIQELEAVT SEMVRLIETA TVPILAVDAD GLVNGWNTKI AELTGLMVDE AIGKHLLSLV EDSSVERVKK MLEMALEGTE EQDVQFEIKT 700
   PHYA_CYP IHSKLNDLKL DGIQELEAVT SEMVRLIETA TVPILAVDAD GLVNGWNTKI AELTGLMVDE AIGKHLLSLV EDSSVERVKK MLEMALEGTE EQDVQFEIKT 700
PHYA_KM2397 IHSKLNDLKL DGIQELEAVT SEMVRLIETA TVPILAVDAD GLVNGWNTKI AELTGLPVDE AIGKHLLSLV EDSSVERVKK MLEMALEGTE EQDVQFEIKT 700
                              720                   740                   760                   780                   800
   PHYA_TUR HLSRADAGPI SLVVNACASR DLNENVVGVC FVAQDLTGQK NVMDKFTRIE GDYKAIIQNP NPLIPPIFGT DEFGWCTEWN PAMSKLTGLK REEVIDKMLL 800
   PHYA_CYP HLSRADAGPI SLVVNACASR DLNEHVVGVC FVAQDLTGQK NVMDKFTRIE GDYKAIIQNP NPLIPPIFGT DEFGWCTEWN PAMSKLTGLK REEVIDKMLL 800
PHYA_KM2397 HLSRADAGPI SLVVNACASR DLNENVVGVC FVAQDLTGQK NVMDKFTRIE GDYKAIIQNP NPLIPPIFGT DEFGWCTEWN PAMSKLTGLK REEVIDKMLL 800
                              820                   840                   860                   880                   900
   PHYA_TUR GEVFGTQKAC CRLKNQEAFV NLGIVLNNAV TSQEPEKVSF AFFTRSGKYV ECLLCVSKKL DREGVVTGVF CFLQLVSHEL QQALHVQRLA ERTALKRLKA 900
   PHYA_CYP GEVFGTQKAC CRLKNQEAFV NLGIVLNNAV TSQEPEKVSF AFFTRSGKYV ECLLCVSKKL DREGVVTGVF CFLQLASHEL QQALHVQRLA ERTALKRLKA 900
PHYA_KM2397 GEVFGTQKAC CRLKNQEAFV NLGIVLNNAV TSQEPEKVSF AFFTRSGKYV ECLLCVSKKL DREGVVTGVF CFLQLASHEL QQALHVQRLA ERTALKRLKA 900
                              920                   940                   960                   980                 1,000
   PHYA_TUR LAYIKRQIRN PLSGIVFSRK MMEGTELGPE QRQILHTSGL CQKQLSKVLD DSDLDSIIDG CLDLEMKEFS LNEVLTASTS QVMMKSNGKS VRITNETKEE 1000
   PHYA_CYP LAYIKRQIRN PLSGIVFSRK MMEGTELGPE QRQILHTSGL CQKQLSKVLD DSDLDSIIDG CLDLEMKEFS LNEVLTASTS QVMMKSNGKS VRITNETKEE 1000
PHYA_KM2397 LAYIKRQIRN PLSGIVFSRK MMEGTELGPE QRQILHTSGL CQKQLSKVLD DSDLDSIIDG CLDLEMKEFS LNEVLTASTS QVMMKSNGKS VRITNETEEE 1000
                            1,020                 1,040                 1,060                 1,080                 1,100
   PHYA_TUR VMSDTLYGDS VRLQQVLADF MFMSVNFTPS GGQLDVTASL RKDQLGRSVH LAYLEIRLTH TGAGIPEFLL NQMFGTEEDV SEEGLSLMIS RKLVKLMNGD 1100
   PHYA_CYP VMSDTLYGDS VRLQQVLADF MFMSVNFTPS GGQLDVTASL RKDQLGRSVH LAYLEIRLTH TGAGIPEFLL NQMFGTEEDV SEEGLSLMIS RKLVKLMNGD 1100
PHYA_KM2397 VMSDTLYGDS VRLQQVLADF MFMSVNFTPS GGQLDVTASL RKDQLGRSVH LAYLEIRLTH TGAGIPEFLL NQMFGTEEDV SEEGLSLMIS RKLVKLMNGD 1100
                            1,120
   PHYA_TUR VQYLRQAGKS SFIITAELAA ANK 1123
   PHYA_CYP VQYLRQAGKS SFIITAELAA AHK 1123
PHYA_KM2397 VQYLRQAGKS SFIITAELAA ANK 1123
```

C



```
                               20                    40                    60                    80                   100
   PHYB_TUR MVTGGDVGGG SSGGGRGRGG EAASSSQRVN TTRREQPQSS GTKSLRPQSV TESMSKAIQQ YTVDARLHAV FEQSGESGKS FDYSQSLKTT TYGSSVPEQQ 100
   PHYB_CYP MVTGGDVGGG SSGGGRGRGG EAASSSQRVN TTRREQPQSS GTKSLRPQSV TESMSKAIQQ YTVDARLHAV FEQSGESGKS FDYSQSLKTT TYGSSVPEQQ 100
PHYB_KM2397 MVTGGDVGGG SSGGGRGRGG EAASSSQRVN TTRREQPQSS GTKSLRPQSV TESMSKAIQQ YTVDARLHAV FEQSGESGKS FDYSQSLKTT TYGSSVPEQQ 100
                              120                   140                   160                   180                   200
   PHYB_TUR ITAYLSRIQR GGYIQPFGCM IAVDESTFRI IGYSENAREM LGLTPQSVPS LEKPEILSMG TDVRSLFTPS SSILLERAFV AREITLLNPV WIHSKNTGKP 200
   PHYB_CYP ITAYLSRIQR GGYIQPFGCM IAVDESTFRI IGYSENAREM LGLTPQSVPS LEKPEILSMG TDVRSLFTPS SSILLERAFV AREITLLNPV WIHSKNTGKP 200
PHYB_KM2397 ITAYLSRIQR GGYIQPFGCM IAVDESTFRI IGYSENAREM LGLTPQSVPS LEKPEILSMG TDVRSLFTPS SSILLERAFV AREITLLNPV WIHSKNTGKP 200
                              220                   240                   260                   280                   300
   PHYB_TUR FYAILHRIDV GIVIDLEPAR TEDPALSIAG AVQSQKLAVR AISQLQSLPG GDIKLLCDTV VESVRDLTGY DRVMVYKFHE DEHGEVVAES KRDDLEPYIG 300
   PHYB_CYP FYAILHRIDV GIVIDLEPAR TEDPALSIAG AVQSQKLAVR AISQLQSLPG GDIKLLCDTV VESVRDLTGY DRVMVYKFHE DEHGEVVAES KRDDLEPYIG 300
PHYB_KM2397 FYAILHRIDV GIVIDLEPAR TEDPALSIAG AVQSQKLAVR AISQLQSLPG GDIKLLCDTV VESVRDLTGY DRVMVYKFHE DEHGEVVAES KRDDLEPYIG 300
                              320                   340                   360                   380                   400
   PHYB_TUR LHYPATDIPQ ASRFLFKQNR VRMIVDCHAT PVLVIQDERL TQSMCLVGST LRAPHGCHSQ YMANMGSIAS LAMAVIINGN EEDGNNVSGG RNSMRLWGLV 400
   PHYB_CYP LHYPATDIPQ ASRFLFKQNR VRMIVDCHAT PVLVIQDERL TQSMCLVGST LRAPHGCHSQ YMANMGSIAS LAMAVIINGN EEDGNNVSGG RNSMRLWGLV 400
PHYB_KM2397 LHYPATDIPQ ASRFLFKQNR VRMIVDCHAT PVLVIQDERL TQSMCLVGST LRAPHGCHSQ YMANMGSIAS LAMAVIINGN EEDGNNVSGG RNSMRLWGLV 400
                              420                   440                   460                   480                   500
   PHYB_TUR VCHHTSARCI PFPLRYACEF LMQAFGLQLN MELQLALQMS EKRVLRTQTL LCDMLLRDSP AGIVTQSPSI MDLVKCDGAA FLYNGKYYPL GVAPSEAQIK 500
   PHYB_CYP VCHHTSARCI PFPLRYACEF LMQAFGLQLN MELQLALQMS EKRVLRTQTL LCDMLLRDSP AGIVTQSPSI MDLVKCDGAA FLYNGKYYPL GVAPSEAQIK 500
PHYB_KM2397 VCHHTSARCI PFPLRYACEF LMQAFGLQLN MELQLALQMS EKRVLRTQTL LCDMLLRDSP AGIVTQSPSI MDLVKCDGAA FLYNGKYYPL GVAPSEAQIK 500
                              520                   540                   560                   580                   600
   PHYB_TUR DIVEWLLANH ADSTGLSTDS LGDAGYPGAV ALGDAVCGMA VAYITKKDFL FWFRSHTGKE IKWGGAKHHP EDKDDGQRMH PRSSFKAFLE VVKSRSQPWE 600
   PHYB_CYP DIVEWLLANH ADSTGLSTDS LGDAGYPGAV ALGDAVCGMA VAYITKKDFL FWFRSHTGKE IKWGGAKHHP EDKDDGQRMH PRSSFKAFLE VVKSRSQPWE 600
PHYB_KM2397 DIVEWLLANH ADSTGLSTDS LGDAGYPGAV ALGDAVCGMA VAYITKKDFL FWFRSHTGKE IKWGGAKHHP EDKDDGQRMH PRSSFKAFLE VVKSRSQPWE 600
                              620                   640                   660                   680                   700
   PHYB_TUR TAEMDAIHSL QLILRDSFKE SEAMNAKAAD GTVQPVRDGS EQGLDELGAV AREMVRLIET ATVPIFAVDT GGCINGWNAK IAELTGLSVE EAMGKSLVSD 700
   PHYB_CYP TAEMDAIHSL QLILRDSFKE SEAMNAKAAD GTVQPVRDGS EQGLDELGAV AREMVRLIET ATVPIFAVDT GGCINGWNAK IAELTGLSVE EAMGKSLVSD 700
PHYB_KM2397 TAEMDAIHSL QLILRDSFKE SEAMNAKAAD GTVQPVRDGS EQGLDELGAV AREMVRLIET ATVPIFAVDT GGCINGWNAK IAELTGLSVE EAMGKSLVSD 700
                              720                   740                   760                   780                   800
   PHYB_TUR LIYKENEETV DKLLSRALKG DEDKNVEVKL KTFNPVLQGK AVFIVVNACS SKDYLNNIVG VCFVGQDVTS HKIVMDKFIN IQGDYKAIVH SPNPLIPPIF 800
   PHYB_CYP LIYKENEETV DKLLSRALKG DEDKNVEVKL KTFNPVLQGK AVFIVVNACS SKDYLNNIVG VCFVGQDVTS HKIVMDKFIN IQGDYKAIVH SPNPLIPPIF 800
PHYB_KM2397 LIYKENEETV DKLLSRALKG DEDKNVEVKL KTFNPVLQGK AVFIVVNACS SKDYLNNIVG VCFVGQDVTS HKIVMDKFIN IQGDYKAIVH SPNPLIPPIF 800
                              820                   840                   860                   880                   900
   PHYB_TUR AADGNTCCLE WNTAMEKVTG WSRNEVIGKM LVGEVFGSSC KLKGPDALTK FMIVLHNAIG GQDTDKFPFP FFDRNGKFVQ ALLTANKRLS LDGQVIGAFC 900
   PHYB_CYP AADGNTCCLE WNTAMEKVTG WSRNEVIGKM LVGEVFGSSC KLKGPDALTK FMIVLHNAIG GQDTDKFPFP FFDRNGKFVQ ALLTANKRLS LDGQVIGAFC 900
PHYB_KM2397 AADGNTCCLE WNTAMEKVTG WSRNEVIGKM LVGEVFGSSC KLKGPDALTK FMIVLHNAIG GQDTDKFPFP FFDRNGKFVQ ALLTANKRLS LDGQVIGAFC 900
                              920                   940                   960                   980                 1,000
   PHYB_TUR FLQIPSLELQ QALAVQRRQD TESFSKAKEL AYICQMVKNP LSGLRFTNSL LDSTALNEDQ KQFLETSVSC EKQISKIVGD MDLQSIEDGS FELERAEFFF 1000
   PHYB_CYP FLQIPSLELQ QALAVQRRQD TESFSKAKEL AYICQMVKNP LSGLRFTNSL LDSTALNEDQ KQFLETSVSC EKQISKIVGD MDLQSIEDGS FELERAEFFF 1000
PHYB_KM2397 FLQIPSLELQ QALAVQRRQD TESFSKAKEL AYICQMVKNP LSGLRFTNSL LDSTALNEDQ KQFLETSVSC EKQISKIVGD MDLQSIEDGS FELERAEFFF 1000
                            1,020                 1,040                 1,060                 1,080                 1,100
   PHYB_TUR GNVVNAIVSQ AMFLLRERGL QLIRDIPNEI KSIAVYGDQI RIQQLLAEFL LSIIRYAPSQ EWVEIHLRQA PKQITDGFSS IRTEFRMACP GEGLPPELVR 1100
   PHYB_CYP GNVVNAIVSQ AMFLLRERGL QLIRDIPNEI KSIAVYGDQI RIQQLLAEFL LSIIRYAPSQ EWVEIHLRQA PKQITDGFSS IRTEFRMACP GEGLPPELVR 1100
PHYB_KM2397 GNVVNAIVSQ AMFLLRERGL QLIRDIPNEI KSIAVYGDQI RIQQLLAEFL LSIIRYAPSQ EWVEIHLRQA PKQITDGFSS IRTEFRMACP GEGLPPELVR 1100
                            1,120                 1,140                 1,160
   PHYB_TUR DMFNSSRWSN PEGLSLSICR KILKLMNGEV QYIRESERSY FLIILELPLP MKRSSAAGNS SGDMMLMMPY 1170
   PHYB_CYP DMFNSSRWSN PEGLSLSICR KILKLMNGEV QYIRESERSY FLIILELPLP MKRSSAAGNS SGDMMLMMPY 1170
PHYB_KM2397 DMFNSSRWSN PEGLSLSICR KILKLMNGEV QYIRESERSY FLIILELPLP MKRSSAAGNS SGDMMLMMPY 1170
```

Supporting Figure 5

D

```
                               20                    40                    60                    80
   PHYC_TUR MSSSSTERSN CSRRSSTRSI HNSRVTSQIL ADAKLHVNFE ESERLFDYSA SVNLNMSNST CNVSSSAVSA YLQMIQRGML IQPFGCLIVV 90
   PHYC_CYP MSSSSTERSN CSRRSSTRSI HNSRVTSQIL ADAKLHVNFE ESERLFDYSA SVNLNMSNST CNVSSSAVSA YLQMIQRGML IQPFGCLIVV 90
PHYC_KM2397 MSSSSTERSN CSRRSSTRSI HNSRVTSQIL ADAKLHVNFE ESERLFDYSA SVNLNMSNST CNVSSSAVSA YLQMIQRGML IQPFGCLIVV 90
                   100                   120                   140                   160                   180
   PHYC_TUR DEKSLKVIAY SENTQEMLGL TPHTVPNLEQ KETLTFGTNV QSLFLSPGSS ALQKAVEFDE ISILNPIMLH CRSSLKPFYA ILHRIEEGLV 180
   PHYC_CYP DEKSLKVIAY SENTQEMLGL TPHTVPNLEQ KETLTFGTNV QSLFLSPGSS ALQKAVEFDE ISILNPIMLH CRSSLKPFYA ILHRIEEGLV 180
PHYC_KM2397 DEKSLKVIAY SENTQEMLGL TPHTVPNLEQ KETLTFGTNV QSLFLSPGSS ALQKAVEFDE ISILNPIMLH CRSSLKPFYA ILHRIEEGLV 180
                              200                   220                   240                   260
   PHYC_TUR IDLEPVNPDE VPVTAAGALK SYKLAAKAIS RLQALTSGNL LLLCDVLVKE VSELTGYDRV MVYKFHEDGH GEVIAENCRL DVEPYIGLHY 270
   PHYC_CYP IDLEPVNPDE VPVTAAGALK SYKLAAKAIS RLQALTSGNL LLLCDVLVKE VSELTGYDRV MVYKFHEDGH GEVIAENCRL DVEPYIGLHY 270
PHYC_KM2397 IDLEPVNPDE VPVTAAGALK SYKLAAKAIS RLQALTSGNL LLLCDVLVKE VSELTGYDRV MVYKFHEDGH GEVIAENCRL DVEPYIGLHY 270
                   280                   300                   320                   340                   360
   PHYC_TUR SATDIPQASR FLFMRNKVRM ICDCSAVPIK VVQDKSLSQP ISLSGSTLRA PQACHAQYMS NMGSVASLVM SVTINGDDNY EMNRDPHTGR 360
   PHYC_CYP SATDIPQASR FLFMRNKVRM ICDCSAVPIK VVQDKSLSQP ISLSGSTLRA PQACHAQYMS NMGSVASLVM SVTINGDDNY EMNRDPHTGR 360
PHYC_KM2397 SATDIPQASR FLFMRNKVRM ICDCSAVPIK VVQDKSLSQP ISLSGSTLRA PQACHAQYMS NMGSVASLVM SVTINGDDNY EMNRDPHTGR 360
                              380                   400                   420                   440
   PHYC_TUR RLWGLVVCHH SNPRFVSFPL RYACELLMQV FTAQINKEVE SALLVKEKHI LQTQSVLCDM FFRNAPIGIV TESPNIMDLV KCDGAALCYR 450
   PHYC_CYP RLWGLVVCHH SNLRFVSFPL RYACELLMQV FTAQINKEVE SALLVKEKHI LQTQSVLCDM FFRNAPIGIV TESPNIMDLV KCDGAALCYR 450
PHYC_KM2397 RLWGLVVCHH SNPRFVSFPL RYACELLMQV FTAQINKEVE SALLVKEKHI LQTQSVLCDM FFRNAPIGIV TESPNIMDLV KCDGAALCYR 450
                   460                   480                   500                   520                   540
   PHYC_TUR DKFWLLGITP TEIQVRDIID WVLKNNEGKT GFTTESLMES GYPGASGLGE SVCGMAVVFI SEKDFLFWFR SGTAKQIKWG GARHDPNDRD 540
   PHYC_CYP DKFWLLGITP TEIQVRDIID WVLKNNEGKT GFTTESLMES GYPGASGLGE SVCGMAVVFI SEKDFLFWFR SGTAKQIKWG GARHDPNDRD 540
PHYC_KM2397 DKFWLLGITP TEIQVRDIID WVLKNNEGKT GFTTESLMES GYPGASGLGE SVCGMAVVFI SEKDFLFWFR SGTAKQIKWG GARHDPNDRD 540
                              560                   580                   600                   620
   PHYC_TUR GKRMHPRSSF KAFMEVIRWK SEAWVDMEMD AINSLQLIIK GSLQEHSKTD VNVPSMDNRV QKVDELCVIL NEMVRLIDTA AVPIFAVDAS 630
   PHYC_CYP GKRMHPRSSF KAFMEVIRWK SEAWVDMEMD AINSLQLIIK GSLQEHSKTD VNVPSMDNRV QKVDELCVIL NEMVRLIDTA AVPIFAVDAS 630
PHYC_KM2397 GKRMHPRSSF KAFMEVIRWK SEAWVDMEMD AINSLQLIIK GSLQEHSKTD VNVPSMDNRV QKVDELCVIL NEMVRLIDTA AVPIFAVDAS 630
                   640                   660                   680                   700                   720
   PHYC_TUR GVINGWNSKA AEVTGLAVEQ AIGKSLSDLV EEDSVTTVNN ILNLALQGSE ERGAEIKIKS FGPKRKSSPI DLVINTCCSR DTMRNVLGVC 720
   PHYC_CYP GVINGWNSKA AEVTGLAVEQ AIGKSLSDLV EEDSVTTVNN ILNLALQGSE ERGAEIKIKS FGPKRKSSPI DLVINTCCSR DTMRNVLGVC 720
PHYC_KM2397 GVINGWNSKA AEVTGLAVEQ AIGKSLSDLV EEDSVTTVNN ILNLALQGSE ERGAEIKIKS FGPKRKSSPI DLVINTCCSR DTMRNVLGVC 720
                              740                   760                   780                   800
   PHYC_TUR FIGQDVTGQK KLFESYSRVR EDYARIMWKP STLIPPIFMT NPNGVCTEWN DAMQKLSGIK REEAVEKLLL GEVFTLDDYG CRLKDHDTLV 810
   PHYC_CYP FIGQDVTGQK KLFESYSRVR EDYARIMWKP STLIPPIFMT NPNGVCTEWN DAMQKLSGIK REEAVEKLLL GEVFTLDDYG CRLKDHDTLV 810
PHYC_KM2397 FIGQDVTGQK KLFESYSRVR EDYARIMWKP STLIPPIFMT NPNGVCTEWN DAMQKLSGIK REEAVEKLLL GEVFTLDDYG CRLKDHDTLV 810
                   820                   840                   860                   880                   900
   PHYC_TUR KLRIALNAVS SGENVEKLFF GFYHREGNFV EALLSTNKKT DIEGKVTGVL CFLQVPSPEL QCALQVQQIS EKAMACALDK LAYLRQEVMN 900
   PHYC_CYP KLRIALNAVS SGENVEKLFF GFYHREGNFV EALLSTNKKT DIEGKVTGVL CFLQVPSPEL QCALQVQQIS EKAMACALDK LAYLRQEVMN 900
PHYC_KM2397 KLRIALNAVS SGENVEKLFF GFYHREGNFV EALLSTNKKT DIEGKVTGVL CFLQVPSPEL QCALQVQQIS EXAMACALDK LAYLRQEVMN 900
                              920                   940                   960                   980
   PHYC_TUR PENAIAFLQN LLHSSGLSDE QKWLLRTNIL CREQLAKVIS DSDLEGIEEG YGELSLNEFS LEESLEAVVK QVMELSDERK VQILCDYPEE 990
   PHYC_CYP PENAIAFLQN LLHSSGLSDE QKWLLRTNIL CREQLAKVIS DSDLEGIEEG YGELSLNEFS LEESLEAVVK QVMELSDERK VQILCDYPEE 990
PHYC_KM2397 PENAIAFLQN LLHSSGLSDE QKWLLRTNIL CREQLAKVIS DSDLEGIEEG YGELSLNEFS LEESLEAVVK QVMELSDERK VQILCDYPEE 990
                 1,000                 1,020                 1,040                 1,060                 1,080
   PHYC_TUR VSSIRLYGDY LRLQQILSET LSCSLSFTPA FKGLSVSFKV IARVEAIGRR TKRVHLEFRI IHPAPGLPGD LENEMFKPLR KGTSREGLGL 1080
   PHYC_CYP VSSIRLYGDY LRLQQILSET LSCSLSFTPA FKGLSVSFKV IARVEAIGRR TKRVHLEFRI IHLAPGLPGD LENEMFKPLR KGTSREGLGL 1080
PHYC_KM2397 VSSIRLYGDY LRLQQILSET LSCSLSFTPA FKGLSVSFKV IARVEAIGRR TKRVHLEFRI IHPAPGLPGD LENEMFKPLR KGTSREGLGL 1080
                            1,100
   PHYC_TUR HMTEKLVKLM EKGRMRYLRE SDISVFVILA EFPLI 1115
   PHYC_CYP HMTEKLVKLM EKGRMRYLRE SDISVFVILA EFPLI 1115
PHYC_KM2397 HMTEKLVKLM EKGRMRYLRE SDISVFVILA EFPLI 1115
```

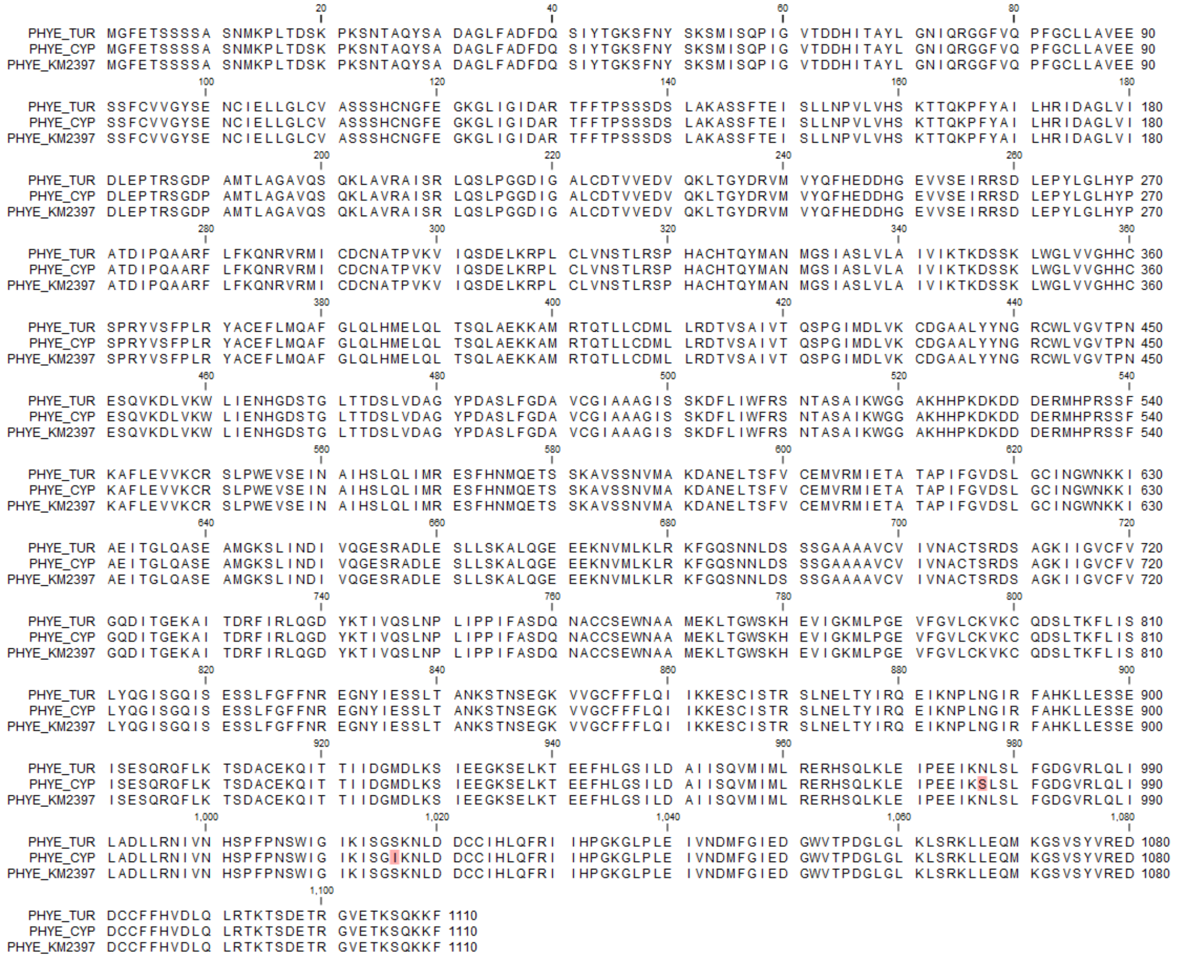
E

```
                               20                    40                    60                    80
   PHYD_TUR MVSGGGGEGS GSGGNRGGKA ASSGLRVAES NHTNSHRGEQ AQSSATKALR SQNQSQNHTD SKSKAIQQYT VDARLHAVFE QSGESGKSFD 90
   PHYD_CYP MVSGGGGEGS GSGGNRGGKA ASSGLRVAES NHTNSHRGEQ AQSSATKALR SQNQSQNHTD SKSKAIQQYT VDARLHAVFE QSGESGKSFD 90
PHYD_KM2397 MVSGGGGEGS GSGGNRGGKA ASSGLRVAES NHTNSHRGEQ AQSSATKALR SQNQSQNHTD SKSKAIQQYT VDARLHAVFE QSGESGKSFD 90
                   100                   120                   140                   160                   180
   PHYD_TUR YSNSLKTTTN WSSVPEQQIT AYLSRIQRGG YIQPFGCLIA VDESTFNIIG YSENAREMLG LMLQSVPSIE NSEVLSIGTD LRSLFKPSSI 180
   PHYD_CYP YSNSLKTTTN WSSVPEQQIT AYLSRIQRGG YIQPFGCLIA VDESTFNIIG YSENAREMLG LMLQSVPSIE NSEVLSIGTD LRSLFKPSSI 180
PHYD_KM2397 YSNSLKTTTN WSSVPEQQIT AYLSRIQRGG YIQPFGCLIA VDESTFNIIG YSENAREMLG LMLQSVPSIE NSEVLSIGTD LRSLFKPSSI 180
                              200                   220                   240                   260
   PHYD_TUR ILLERAFVAR EITLLNPVWI HSKNTGKPFY AILHRVDVGI VIDLEPARTE DPALSIAGAV QSQKLAVRAI SHLQSLPGGD IKLLCDTVVE 270
   PHYD_CYP ILLERAFVAR EITLLNPVWI HSKNTGKPFY AILHRVDVGI VIDLEPARTE DPALSIAGAV QSQKLAVRAI SHLQSLPGGD IKLLCDTVVE 270
PHYD_KM2397 ILLERAFVAR EITLLNPVWI HSKNTGKPFY AILHRVDVGI VIDLEPARTE DPALSIAGAV QSQKLAVRAI SHLQSLPGGD IKLLCDTVVE 270
                   280                   300                   320                   340                   360
   PHYD_TUR RVRDLTGYDR VMVYKFHEDE HGEVVAETKR DDLEPYFGLH YPATDIPQAS RFLFKQNRVR MIVDCHATPV RVVQDDRLTQ SICLVGSTLR 360
   PHYD_CYP RVRDLTGYDR VMVYKFHEDE HGEVVAETKR DDLEPYFGLH YPATDIPQAS RFLFKQNRVR MIVDCHATPV RVVQDDRLTQ SICLVGSTLR 360
PHYD_KM2397 RVRDLTGYDR VMVYKFHEDE HGEVVAETKR DDLEPYFGLH YPATDIPQAS RFLFKQNRVR MIVDCHATPV RVVQDDRLTQ SICLVGSTLR 360
                              380                   400                   420                   440
   PHYD_TUR APHGCHAQYM ANMGSIASLA MAVIINGNEE DGNGVTGGRN SMRLWGLVVC HHTSARCIPF PLRYACEFLM QAFGLQLNME LQLALQVSEK 450
   PHYD_CYP APHGCHAQYM ANMGSIASLA MAVIINGNEE DGNGVTGGRN SMRLWGLVVC HHTSARCIPF PLRYACEFLM QAFGLQLNME LQLALQVSEK 450
PHYD_KM2397 APHGCHAQYM ANMGSIASLA MAVIINGNEE DGNGVTGGRN SMRLWGLVVC HHTSARCIPF PLRYACEFLM QAFGLQLNME LQLALQVSEK 450
                   460                   480                   500                   520                   540
   PHYD_TUR RVLRMQTLLC DMLLRDSPAG IVTQSPSIMD LVKCNGAAFL YQGKYYSLGV APSEAQILDI VEWLLANHAD STGLSTDSLG DAGYPRASAL 540
   PHYD_CYP RVLRMQTLLC DMLLRDSPAG IVTQSPSIMD LVKCNGAAFL YQGKYYSLGV APSEAQILDI VEWLLANHAD STGLSTDSLG DAGYPRASAL 540
PHYD_KM2397 RVLRMQTLLC DMLLRDSPAG IVTQSPSIMD LVKCNGAAFL YQGKYYSLGV APSEAQILDI VEWLLANHAD STGLSTDSLG DAGYPRASAL 540
                              560                   580                   600                   620
   PHYD_TUR GDAVCGMAVA YITKKDFLFW FRSHTEKEIK WGGAKHHPED KDDGQRMHPR SSFKAFLEVV KSRCQPWEPA EMDAIHSLQL ILRDSFKESE 630
   PHYD_CYP GDAVCGMAVA YITKKDFLFW FRSHTEKEIK WGGAKHHPED KDDGQRMHPR SSFKAFLEVV KSRCQPWEPA EMDAIHSLQL ILRDSFKESE 630
PHYD_KM2397 GDAVCGMAVA YITKKDFLFW FRSHTEKEIK WGGAKHHPED KDDGQRMHPR SSFKAFLEVV KSRCQPWEPA EMDAIHSLQL ILRDSFKESE 630
                   640                   660                   680                   700                   720
   PHYD_TUR AINNSKAVAG AVQPLGDIAE QEIDELGSVA REMVRLIETA TVPIFAVDVD GCINGWNAKI AELTGLSVEE AMGRSLVHDL IYKEYEETVD 720
   PHYD_CYP AINNSKAVAG AVQPLGDIAE QEIDELGSVA REMVRLIETA TVPIFAVDVD GCINGWNAKI AELTGLSVEE AMGRSLVHDL IYKEYEETVD 720
PHYD_KM2397 AINNSKAVAG AVQPLGDIAE QEIDELGSVA REMVRLIETA TVPIFAVDVD GCINGWNAKI AELTGLSVEE AMGRSLVHDL IYKEYEETVD 720
                              740                   760                   780                   800
   PHYD_TUR KLLSCALKGD EGKNVELKLK TFGPELQGKA VFVVVNACSS KDYLDNIVGV CFIAQDVTGH KTVMDKFINI QGDYKAIIHS PNPLIPPIFA 810
   PHYD_CYP KLLSCALKGD EGKNVELKLK TFGPELQGKA VFVVVNACSS KDYLDNIVGV CFIAQDVTGH KTVMDKFINI QGDYKAIIHS PNPLIPPIFA 810
PHYD_KM2397 KLLSCALKGD EGKNVELKLK TFGPELQGKA VFVVVNACSS KDYLDNIVGV CFIAQDVTGH KTVMDKFINI QGDYKAIIHS PNPLIPPIFA 810
                   820                   840                   860                   880                   900
   PHYD_TUR ADENTCCMEW NTAMEKLTGW PRNEVIGKML VREVFGSCCR LKGPDALTKF MIVLHNVMGG QETDKFPFPF FDRRGKFIQT LLTLNKRVNL 900
   PHYD_CYP ADENTCCMEW NTAMEKLTGW PRNEVIGKML VREVFGSCCR LKGPDALTKF MIVLHNVMGG QETDKFPFPF FDRRGKFIQT LLTLNKRVNL 900
PHYD_KM2397 ADENTCCMEW NTAMEKLTGW PRNEVIGKML VREVFGSCCR LKGPDALTKF MIVLHNVMGG QETDKFPFPF FDRRGKFIQT LLTLNKRVNL 900
                              920                   940                   960                   980
   PHYD_TUR DGKVIGAFCF LQIPSPELQQ ALEIQRRQEA ECFSRGKELA YIFQVIKNPL SGLRFTHSML EATNLNEEQK QLLETGVSCE MQISKIVEYM 990
   PHYD_CYP DGKVIGAFCF LQIPSPELQQ ALEIQRRQEA ECFSRGKELA YIFQVIKNPL SGLRFTHSML EATNLNEEQK QLLETGVSCE MQISKIVEYM 990
PHYD_KM2397 DGKVIGAFCF LQIPSPELQQ ALEIQRRQEA ECFSRGKELA YIFQVIKNPL SGLRFTHSML EATNLNEEQK QLLETGVSCE MQISKIVEYM 990
                 1,000                 1,020                 1,040                 1,060                 1,080
   PHYD_TUR DVKSIEDGSF KLERTEFLLG NVINAVVSQV MLVLTERGLQ LIRDIPIEIK SMAVYGDQMR LQQVLAEFLS SIIRYAPVKS WVELHLLPVL 1080
   PHYD_CYP DVKSIEDGSF KLERTEFLLG NVINAVVSQV MLVLTERGLQ LIRDIPIEIK SMAVYGDQMR LQQVLAEFLS SIIRYAPVKS WVELHLLPVL 1080
PHYD_KM2397 DVKSIEDGSF KLERTEFLLG NVINAVVSQV MLVLTERGLQ LIRDIPIEIK SMAVYGDQMR LQQVLAEFLS SIIRYAPVKS WVELHLLPVL 1080
                            1,100                 1,120                 1,140                 1,160
   PHYD_TUR KQMADGFSAI HMEFRMACPG EGVPPERVQD MFHSSRWTSP EGLALSICRK ILKLMNGEVQ YIRECERSYF LIILELTIPP KRFSNTSLTA 1170
   PHYD_CYP KQMADGFSAI HMEFRMACPG EGVPPERVQD MFHSSRWTSP EGLALSICRK ILKLMNGEVQ YIRECERSYF LIILELTIPP KRFSNTSLTA 1170
PHYD_KM2397 KQMADGFSAI HMEFRMACPG EGVPPERVQD MFHSSRWTSP EGLALSICRK ILKLMNGEVQ YIRECERSYF LIILELTIPP KRFSNTSLTA 1170

   PHYD_TUR NIATMPA 1177
   PHYD_CYP NIATMPA 1177
PHYD_KM2397 NIATMPA 1177
```

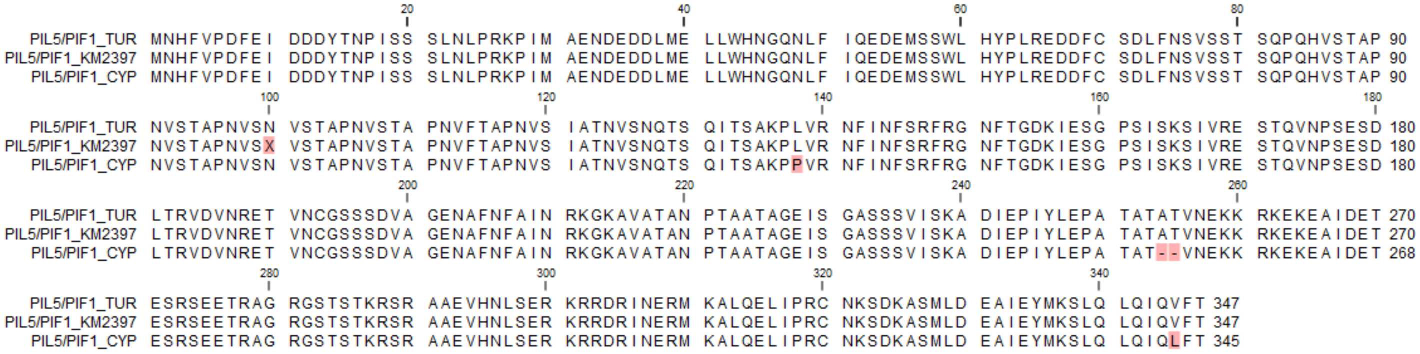
Supporting Figure 5

F



```
                               20                    40                    60                    80
   PHYE_TUR MGFETSSSSA SNMKPLTDSK PKSNTAQYSA DAGLFADFDQ SIYTGKSFNY SKSMISQPIG VTDDHITAYL GNIQRGGFVQ PFGCLLAVEE 90
   PHYE_CYP MGFETSSSSA SNMKPLTDSK PKSNTAQYSA DAGLFADFDQ SIYTGKSFNY SKSMISQPIG VTDDHITAYL GNIQRGGFVQ PFGCLLAVEE 90
PHYE_KM2397 MGFETSSSSA SNMKPLTDSK PKSNTAQYSA DAGLFADFDQ SIYTGKSFNY SKSMISQPIG VTDDHITAYL GNIQRGGFVQ PFGCLLAVEE 90
                   100                   120                   140                   160                   180
   PHYE_TUR SSFCVVGYSE NCIELLGLCV ASSSHCNGFE GKGLIGIDAR TFFTPSSSDS LAKASSFTEI SLLNPVLVHS KTTQKPFYAI LHRIDAGLVI 180
   PHYE_CYP SSFCVVGYSE NCIELLGLCV ASSSHCNGFE GKGLIGIDAR TFFTPSSSDS LAKASSFTEI SLLNPVLVHS KTTQKPFYAI LHRIDAGLVI 180
PHYE_KM2397 SSFCVVGYSE NCIELLGLCV ASSSHCNGFE GKGLIGIDAR TFFTPSSSDS LAKASSFTEI SLLNPVLVHS KTTQKPFYAI LHRIDAGLVI 180
                              200                   220                   240                   260
   PHYE_TUR DLEPTRSGDP AMTLAGAVQS QKLAVRAISR LQSLPGGDIG ALCDTVVEDV QKLTGYDRVM VYQFHEDDHG EVVSEIRRSD LEPYLGLHYP 270
   PHYE_CYP DLEPTRSGDP AMTLAGAVQS QKLAVRAISR LQSLPGGDIG ALCDTVVEDV QKLTGYDRVM VYQFHEDDHG EVVSEIRRSD LEPYLGLHYP 270
PHYE_KM2397 DLEPTRSGDP AMTLAGAVQS QKLAVRAISR LQSLPGGDIG ALCDTVVEDV QKLTGYDRVM VYQFHEDDHG EVVSEIRRSD LEPYLGLHYP 270
                   280                   300                   320                   340                   360
   PHYE_TUR ATDIPQAARF LFKQNRVRMI CDCNATPVKV IQSDELKRPL CLVNSTLRSP HACHTQYMAN MGSIASLVLA IVIKTKDSSK LWGLVVGHHC 360
   PHYE_CYP ATDIPQAARF LFKQNRVRMI CDCNATPVKV IQSDELKRPL CLVNSTLRSP HACHTQYMAN MGSIASLVLA IVIKTKDSSK LWGLVVGHHC 360
PHYE_KM2397 ATDIPQAARF LFKQNRVRMI CDCNATPVKV IQSDELKRPL CLVNSTLRSP HACHTQYMAN MGSIASLVLA IVIKTKDSSK LWGLVVGHHC 360
                              380                   400                   420                   440
   PHYE_TUR SPRYVSFPLR YACEFLMQAF GLQLHMELQL TSQLAEKKAM RTQTLLCDML LRDTVSAIVT QSPGIMDLVK CDGAALYYNG RCWLVGVTPN 450
   PHYE_CYP SPRYVSFPLR YACEFLMQAF GLQLHMELQL TSQLAEKKAM RTQTLLCDML LRDTVSAIVT QSPGIMDLVK CDGAALYYNG RCWLVGVTPN 450
PHYE_KM2397 SPRYVSFPLR YACEFLMQAF GLQLHMELQL TSQLAEKKAM RTQTLLCDML LRDTVSAIVT QSPGIMDLVK CDGAALYYNG RCWLVGVTPN 450
                   460                   480                   500                   520                   540
   PHYE_TUR ESQVKDLVKW LIENHGDSTG LTTDSLVDAG YPDASLFGDA VCGIAAAGIS SKDFLIWFRS NTASAIKWGG AKHHPKDKDD DERMHPRSSF 540
   PHYE_CYP ESQVKDLVKW LIENHGDSTG LTTDSLVDAG YPDASLFGDA VCGIAAAGIS SKDFLIWFRS NTASAIKWGG AKHHPKDKDD DERMHPRSSF 540
PHYE_KM2397 ESQVKDLVKW LIENHGDSTG LTTDSLVDAG YPDASLFGDA VCGIAAAGIS SKDFLIWFRS NTASAIKWGG AKHHPKDKDD DERMHPRSSF 540
                              560                   580                   600                   620
   PHYE_TUR KAFLEVVKCR SLPWEVSEIN AIHSLQLIMR ESFHNMQETS SKAVSSNVMA KDANELTSFV CEMVRMIETA TAPIFGVDSL GCINGWNKKI 630
   PHYE_CYP KAFLEVVKCR SLPWEVSEIN AIHSLQLIMR ESFHNMQETS SKAVSSNVMA KDANELTSFV CEMVRMIETA TAPIFGVDSL GCINGWNKKI 630
PHYE_KM2397 KAFLEVVKCR SLPWEVSEIN AIHSLQLIMR ESFHNMQETS SKAVSSNVMA KDANELTSFV CEMVRMIETA TAPIFGVDSL GCINGWNKKI 630
                   640                   660                   680                   700                   720
   PHYE_TUR AEITGLQASE AMGKSLINDI VQGESRADLE SLLSKALQGE EEKNVMLKLR KFGQSNNLDS SSGAAAAVCV IVNACTSRDS AGKIIGVCFV 720
   PHYE_CYP AEITGLQASE AMGKSLINDI VQGESRADLE SLLSKALQGE EEKNVMLKLR KFGQSNNLDS SSGAAAAVCV IVNACTSRDS AGKIIGVCFV 720
PHYE_KM2397 AEITGLQASE AMGKSLINDI VQGESRADLE SLLSKALQGE EEKNVMLKLR KFGQSNNLDS SSGAAAAVCV IVNACTSRDS AGKIIGVCFV 720
                              740                   760                   780                   800
   PHYE_TUR GQDITGEKAI TDRFIRLQGD YKTIVQSLNP LIPPIFASDQ NACCSEWNAA MEKLTGWSKH EVIGKMLPGE VFGVLCKVKC QDSLTKFLIS 810
   PHYE_CYP GQDITGEKAI TDRFIRLQGD YKTIVQSLNP LIPPIFASDQ NACCSEWNAA MEKLTGWSKH EVIGKMLPGE VFGVLCKVKC QDSLTKFLIS 810
PHYE_KM2397 GQDITGEKAI TDRFIRLQGD YKTIVQSLNP LIPPIFASDQ NACCSEWNAA MEKLTGWSKH EVIGKMLPGE VFGVLCKVKC QDSLTKFLIS 810
                   820                   840                   860                   880                   900
   PHYE_TUR LYQGISGQIS ESSLFGFFNR EGNYIESSLT ANKSTNSEGK VVGCFFFLQI IKKESCISTR SLNELTYIRQ EIKNPLNGIR FAHKLLESSE 900
   PHYE_CYP LYQGISGQIS ESSLFGFFNR EGNYIESSLT ANKSTNSEGK VVGCFFFLQI IKKESCISTR SLNELTYIRQ EIKNPLNGIR FAHKLLESSE 900
PHYE_KM2397 LYQGISGQIS ESSLFGFFNR EGNYIESSLT ANKSTNSEGK VVGCFFFLQI IKKESCISTR SLNELTYIRQ EIKNPLNGIR FAHKLLESSE 900
                              920                   940                   960                   980
   PHYE_TUR ISESQRQFLK TSDACEKQIT TIIDGMDLKS IEEGKSELKT EEFHLGSILD AIISQVMIML RERHSQLKLE IPEEIKNLSL FGDGVRLQLI 990
   PHYE_CYP ISESQRQFLK TSDACEKQIT TIIDGMDLKS IEEGKSELKT EEFHLGSILD AIISQVMIML RERHSQLKLE IPEEIKSLSL FGDGVRLQLI 990
PHYE_KM2397 ISESQRQFLK TSDACEKQIT TIIDGMDLKS IEEGKSELKT EEFHLGSILD AIISQVMIML RERHSQLKLE IPEEIKNLSL FGDGVRLQLI 990
                 1,000                 1,020                 1,040                 1,060                 1,080
   PHYE_TUR LADLLRNIVN HSPFPNSWIG IKISGSKNLD DCCIHLQFRI IHPGKGLPLE IVNDMFGIED GWVTPDGLGL KLSRKLLEQM KGSVSYVRED 1080
   PHYE_CYP LADLLRNIVN HSPFPNSWIG IKISGIKNLD DCCIHLQFRI IHPGKGLPLE IVNDMFGIED GWVTPDGLGL KLSRKLLEQM KGSVSYVRED 1080
PHYE_KM2397 LADLLRNIVN HSPFPNSWIG IKISGSKNLD DCCIHLQFRI IHPGKGLPLE IVNDMFGIED GWVTPDGLGL KLSRKLLEQM KGSVSYVRED 1080
                            1,100
   PHYE_TUR DCCFFHVDLQ LRTKTSDETR GVETKSQKKF 1110
   PHYE_CYP DCCFFHVDLQ LRTKTSDETR GVETKSQKKF 1110
PHYE_KM2397 DCCFFHVDLQ LRTKTSDETR GVETKSQKKF 1110
```

Supporting Figure 5. Identification and alignments of phytochromes in *Ae. arabicum*. (a) Phylogenetic tree of phytochromes using Bayesian inference (1688500 generations, standard deviation of split frequencies 0.009992) and allows clear assignment of *Ae. arabicum* orthologues. Sequences of *A. thaliana* (ARATH) and *Ae. arabicum* (AETAR) are marked in green and red, respectively. For detailed assignment of five letter code see Supplemental Dataset 7. (b-f) Phytochrome protein alignments of three *Ae. arabicum* accessions. Germination of TUR seeds is light-insensitive while CYP and KM2397 both have light inhibited germination.



```
                                            20                    40                    60                    80
                                             |                     |                     |                     |
    PIL5/PIF1_TUR MNHFVPDFEI DDDYTNPISS SLNLPRKPIM AENDEDDLME LLWHNGQNLF IQEDEMSSWL HYPLREDDFC SDLFNSVSST SQPQHVSTAP 90
 PIL5/PIF1_KM2397 MNHFVPDFEI DDDYTNPISS SLNLPRKPIM AENDEDDLME LLWHNGQNLF IQEDEMSSWL HYPLREDDFC SDLFNSVSST SQPQHVSTAP 90
    PIL5/PIF1_CYP MNHFVPDFEI DDDYTNPISS SLNLPRKPIM AENDEDDLME LLWHNGQNLF IQEDEMSSWL HYPLREDDFC SDLFNSVSST SQPQHVSTAP 90
                                 100                   120                   140                   160                   180
                                  |                     |                     |                     |                     |
    PIL5/PIF1_TUR NVSTAPNVSN VSTAPNVSTA PNVFTAPNVS IATNVSNQTS QITSAKPLVR NFINFSRFRG NFTGDKIESG PSISKSIVRE STQVNPSESD 180
 PIL5/PIF1_KM2397 NVSTAPNVSX VSTAPNVSTA PNVFTAPNVS IATNVSNQTS QITSAKPLVR NFINFSRFRG NFTGDKIESG PSISKSIVRE STQVNPSESD 180
    PIL5/PIF1_CYP NVSTAPNVSN VSTAPNVSTA PNVFTAPNVS IATNVSNQTS QITSAKPPVR NFINFSRFRG NFTGDKIESG PSISKSIVRE STQVNPSESD 180
                                            200                   220                   240                   260
                                             |                     |                     |                     |
    PIL5/PIF1_TUR LTRVDVNRET VNCGSSSDVA GENAFNFAIN RKGKAVATAN PTAATAGEIS GASSSVISKA DIEPIYLEPA TATATVNEKK RKEKEAIDET 270
 PIL5/PIF1_KM2397 LTRVDVNRET VNCGSSSDVA GENAFNFAIN RKGKAVATAN PTAATAGEIS GASSSVISKA DIEPIYLEPA TATATVNEKK RKEKEAIDET 270
    PIL5/PIF1_CYP LTRVDVNRET VNCGSSSDVA GENAFNFAIN RKGKAVATAN PTAATAGEIS GASSSVISKA DIEPIYLEPA TAT--VNEKK RKEKEAIDET 268
                                 280                   300                   320                   340
                                  |                     |                     |                     |
    PIL5/PIF1_TUR ESRSEETRAG RGSTSTKRSR AAEVHNLSER KRRDRINERM KALQELIPRC NKSDKASMLD EAIEYMKSLQ LQIQVFT 347
 PIL5/PIF1_KM2397 ESRSEETRAG RGSTSTKRSR AAEVHNLSER KRRDRINERM KALQELIPRC NKSDKASMLD EAIEYMKSLQ LQIQVFT 347
    PIL5/PIF1_CYP ESRSEETRAG RGSTSTKRSR AAEVHNLSER KRRDRINERM KALQELIPRC NKSDKASMLD EAIEYMKSLQ LQIQLFT 345
```

Supporting Figure 6. Alignment of PIL5/PIF1 protein sequence of three *Ae. arabicum* accessions. Germination of TUR seeds is light-insensitive while CYP and KM2397 both have light inhibited germination.