New Phytologist Supporting Information

Article title: **Early evolution of the land plant circadian clock**

Authors: Anna-Malin Linde, D. Magnus Eklund, Akane Kubota, Eric R. A. Pederson, Karl Holm, Niclas Gyllenstrand, Ryuichi Nishihama, Nils Cronberg, Tomoaki Muranaka, Tokitaka Oyama, Takayuki Kohchi and Ulf Lagercrantz

Article acceptance date: 18 January 2017

The following Supporting Information is available for this article:

**Fig. S1** Alignments used for phylogenetic construction

**Fig. S2** Inferred phylogeny of homologs to the ELF3 family

**Fig. S3** Inferred phylogeny of homologs to the ELF4 family

**Fig. S4** Inferred phylogeny of homologs to the LUX family

**Fig. S5** Inferred phylogeny of homologs to the GI gene family

**Fig. S6** Inferred phylogeny of homologs to the ZTL gene family

**Fig. S7** Temporal expression patterns of putative circadian clock genes in *M. polymorpha* (Mp) under ND, LL and DD conditions

**Fig. S8** Generation of MpRVE knockout mutant

**Fig. S9** Generation of MpPRR knockout mutant

**Fig. S10** Generation of MpTOC1 knockout mutants

**Fig. S11** Temporal expression pattern of MpPRR and 35S<sub>pro:LUC</sub> under ND and LL conditions

**Fig. S12** Temporal expression pattern of MpPRR in WT, Mprve<sup>ko</sup>, Mptoc<sub>1</sub><sup>ko</sup> and restored lines of Mprve<sup>ko</sup>
**Fig. S13** pro:LUC bioluminescence for MpELF3, MpGI, MpLUX and MpRVE

**Fig. S14** Luciferase imaging in transgenic *M. polymorpha* plants expressing luciferase under the control of *M. polymorpha* promoters

**Fig. S15** MpPRR<sub>pro</sub>:GUS expression in mature thallus

**Fig. S16** Temporal expression patterns of putative circadian clock genes in *A. agrestis* (Aa) under ND, LL and DD conditions

**Table S1** Gene names, family/sub-clade, gene ID or accession number

**Table S2** Oligonucleotides used in this study

**Methods S1** Supplemental materials and methods describing sequence retrieval, sequence analysis and phylogenetic reconstruction
Fig. S1 Alignments used for phylogenetic construction. Amino acid sequences were aligned using the M-Coffee algorithm in T-Coffee (Notredame et al., 2000; Wallace et al., 2006). Alignments were filtered using Transitive Consistency Score (TCS) in the T-Coffee distribution (Chang et al., 2014). (a) CCA1/LHY/RVE, (b) PRR/TOC1, (c) ELF3, (d) ELF4, (e) LUX, (f) GI, (g) ZTL.

(a)

#NEXUS
BEGIN DATA;
dimensions ntax=31 nchar=215;
format missing=?
symbols="ABCDEFGHIJKLMNOPQRSTUVWXYZ"
interleave datatype=PROTEIN gap=- match=.
matrix
OsRVE6      MDRSSSEMPGKKARKPYTITKPRERWSEEEHERFLDALIMYGRDWKKIEE
PaCCA1-3    MLPFSIDEFFSKVRKPYTTKQERWTEDEHKKFLDALKLYGRSWRIE
PaCCA1      MMSSGDELAAKVRKPYTITKQERWSEEEHLKFLAELKMYGRAWRKIEE
KfRVE       MLGGKGDMMHKKVRKPYTITKQRENWTEQEHNKFLAELFDRDWKKIEE
MpRVE       MA-QGSEEKSKKVRKPYTITKSRSWTEQEHDKFLEALQLFDWKKIEE
AtRVE5      MVSSFSEDPTTKIRKPYTITKSRWTDQEHDKFLEALHLFDWKKIEA
AtRVE3      MVTSSFEDPTTKVRKPYTITKSRWTEQEHDKFLEALHFWDWKKIK
AtRVE4      MTESEAGEAPEKKVRKAYTITKSRWTEQEHDFLEALQLFDWKKIEA
AtRVE8      MSSAVAEGGSSKKVRKPYTITKSRWTEQEHDKFLEALQLFDWKKIEA
OsRVE5      MSSEVDDGGRVRKPYTITKSRWTDQEHDKFLEALQLFDWKKIEA
AaRVE       MF-GTVEEGSKKIRKPYTITKSRWTEQEHDFLEALQLFDWKKIEA
SmRVE       ---PSDASVSKKVRKPYTITKSRWTEQEHDFLEALQLFDWKKIEA
OsRVE4      MVSSAGEDASKKVRKPYTITKSRWTEQEHDKFLEALQLFDWKKIEA
AtRVE6      MVSSSEEDLSKIRKPYTITKSRWTEQEHDKFLEALQLFDWKKIEA
PpRVE1      MNASVSEEGSSKKIRKPYTITKSRWTEQEHDFKLDLALQFDWKKIEA
PpRVE2      MNPVSEEGSSKKIRKPYTITKSRWTEQEHDFKLEDALQFDWKKIEA
PpRVE3      MNASVSEEGSSKKIRKPYTITKSRWTEQEHDFKLEDALQFDWKKIEA
KfCCA1      MDDTSTESLSTKIRKPYTITKQTERWTEAEHDFLEALKLHGRAWRKIEE
PaCCA1-2    MASSGDEFTTKVRKPYTITKQTERWEHEHDFLEALKLHGRAWRKIEE
OsCCA1      ---SSGEEAVVKVRKPYTITKQTERWTEHNRFLERALKLYGRRAWQIEE
AtLHY       ---TSGEELAKKVRKPYTITKQTERWTEHERFLEALKLYGRRAWQIEE
AaCCA1      MNSSSGDEALAKVRKPYTITKQTERWTEHEHTKLEALKLYGRRAWQIEE
AtCCA1      ---SSGEDLVIKTRKPYTITKQTERWTEEHFRFLEALKLYGRRAWQIEE
OsRVE2      MARMMDEARIKPARKPYTITKQREKWTEDEHLKFLAELQLHGRAWRRIQE
OsRVE1      MAACLSEGEEHPARKPYTITKQREKWTEDEHRLKFLAELQLHGRAWRRIQE
AtRVE7      MLCISHNENVKVRKPYTITKQREKWSEEEEHRDFLEALKLYGRQRIQE
AtRVE7Like  MVMNHENVENKVRKPYTITKQREWSEEEEHRDFLEALKLYGRQRIQE
PpCCA1a     -----------VRKPYTITKQREWTEHEHQRFLERALKLYGRRAWQIEE
PpCCA1b     ---SGDEGASTKVRKPYTITKQTERWTEHEHQRFLERALKLYGRRAWQIEE
OsRVE3      MASEGMEEFPVKVRKPYTITKQREKWTEEHDFLEALKLYGRRAWQIEE
AtRVE2      M--SSDAFYLKTRKPYTITQREKWTEAEHEKFVEALKLYGRAWRRIEE
AtRVE1      MASFGNDYAPKVRKYITQERERWTDEEHKKEALKLYGRAWRRIEE
OsRVE6      HVTGTKTTIQIRSHAQKFKFLKVKQKGLMLAA--GLPPQYPRRRLVMQQQQQS
OsCCA1-3    HIKGSKAQVRSHAQKFKFTKLEKAGYELIEPPPFRKKPGHPYPKKTG
PaCCA1      HIGTKTAQIRQSHAQKFKSFLVGRSGSDDIPRPKRPKHPYPRKAV
KfRVE       YHVTGTKTQIRQSHAQKFKFLKVKQLDQA--HPVPPRPRKRKSSQPYPKAS
MpRVE       FVGSKSVIRQSHAQKFKFLKVKQKNTGE--HPVPPRPRKRKSAGQPYPKAP
AtRVE5      FVGSKTVQIRQSHAQKFKFLKVKQKSGANEM--HLPPPRPRKKAHPYPRKAPH
AtRVE3      FVGSKTVQIRQSHAQKFKFLKVKQKSGTNKE--HLPPPRPRKKAHPYPRKAPH
AtRVE4      FVGSKTVQIRQSHAQKFKFLKVKQKNGTGE--HLPPPRPRKKAHPYPRKAPH
AtRVE8      FVGSKTVQIRQSHAQKFKFLKVKQKSGTGE--HLPPPRPRKKAHPYPRKAPH
OsRVE5      YVGTKTVIQIRSHAQKFPYVQKNGTGE--HLPPPRPRKKAHPYPRKAPH
AaRVE       FVGSKTVQIRQSHAQKFKFLKVKQKNGTGE--HLPPPRPRKKAHPYPRKAPH
SmRVE       FVGSKTVQIRQSHAQKFKFLKVKQKNGTGE--HPVPPRPRKRKSAGQPYPKAP
OsRVE4      FVGSKTVQIRQSHAQKFKFLKVKQKNGTSE--HPVPPRPRKKAHPYPRKAPH
AtRVE6      FIGSKTVIQIRSHAQKFKFLKVKQKNTGE--HLPPPRPRKKAHPYPRKAPH
PpRVE1      FVGSKTVQIRQSHAQKFKFLKVKQKNGTGE--HPVPPRPRKRKSAGQPYPKAP
PpRVE2      FVGSKTVQIRQSHAQKFKFLKVKQKNTGE--HPVPPRPRKRKSAGQPYPKAP
PpRVE3      FVGSKTVQIRQSHAQKFKFLKVKQKNTGE--HPVPPRPRKRKSAGQPYPKAP
KfCCA1      HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
OsCCA1-2    HIGSKTAVQIRSHAQKFKSFLERASAGDISIPRPKRPKSHYPKAG
OsCCA1      HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
AtLHY       HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
AaCCA1      HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
AtCCA1      HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
OsRVE2      HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
OsRVE1      HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
AtRVE7      HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
AtRVE7Like   HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
PpCCA1a     HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
PpCCA1b     HIGSKTAVQIRSHAQKFKSFLERAGTSAEIDIPRPKRPKSHYPKAG
OsRVE3      HIGSKTAVQIRSHAQKFSKSVREPGS--AIEIPPRPPKRPKPLHPYPRKCA
AtRVE2      HIGSKTAVQIRSHAQKFKFTKVARDFGV--SEIPPRPPKRPKMPHPYPRKLV
AtRVE1      HIGSKTAVQIRSHAQKFSKSVREPGS--AIEIPPRPPKRPKPLHPYPRKCA
OsRVE6      SPA--MQNELEWASTGTAWGGNHGLPFVSFE--------TT-----
OsCCA1-3    VID-----KSTIHSSVSFTSNWSPA--LSYTHQ--QDEKQRNK----L
PaCCA1      GTA------EAVAVAVATVAASAWLSYGASLHPR--AENISSV--KAREGIQ
KfRVE       KVV------TPSKANPDFAF--V-YRFLGDVFEIGVAGNQLKCMPPIDR
MpRVE       KAV------TDLRATPDFAE--V-YKFIIGNVFDPGSINNHKLKEMAPIDR
AtRVE5      KNV------KPHRVMNFAE--V-YSFIGSVFPNTSGHQLRKLQMDPINM
AtRVE3      KFT------KPHRVAFLNAE--V-YNFIGSVFPKTGGVNRKLEMADNNL
AtRVE4      KNA------PSMHGLPDFAE--V-YNFIGSVFDPDKGMRKJMELMDNPINF
AtRVE8      KNA------PVLHGVPDFAE--V-YNFIGSVFDPETRHEKJEMADNNL
OsRVE5      KRA------PTLRAVDPDFAQ--V-YNFLGSFDPETSGLHRLEMDNPIDV
AaRVE       KSV------ASLRAVDPDFAQ--V-YKFIGSVFPDTAGHLKJLEMDADNNL
SmRVE       KPG------VTLSAAPAFSE--V-YKFIGSIFDPGTAGHLKREMIDR
OsRVE4      KNE------PSLRLMPDFAQ--V-YSFLGSVFDPSTSGLHKLEMNIDV
AtRVE6      KNV------HSLRVLPDFAQ--V-YGFIGSVFDPYASNLHQLKMDPIDV
PpRVE1      KCV------SCIRAAPDFAE--V-YKFIGSIFDPGTAGHLKREMIDR
PpRVE2      KTA------SCIRAAPDFAE--V-YKFIGSVFDPGVSHLRKLEMISAI
PpRVE3      KCG------SCIRAAPDFTE--V-YKFIGSVFDPGVSHLRKLEMISPIDR
KfCCA1      GNM------AAVAAATTMAASAWWTLQNNAQFP-QAKESEEQK-G--SMG
OsCCA1-2    TAI------IMAAVTAATVAAASAWWTLHGAPLPG-GNSSPL-KTEDGNE
OsCCA1      LSSGIH--MASIVTATVAAASAWWATQGLL--ADVQRAKEKDNDE
AtLHY       NNG------ITAIAAATVAAATAWWASHGLV--API--VAVPTAMTDQDS
AaCCA1      GSFS---Y-AAVAAATAIAASAWWALQGSPVHP--APTRTQKQE--DEGV
AtCCA1      SGT------AAVAAATVAAASAWWAANGLL--APL--HPSFPGSCQEHDS
OsRVE2      STA------PNGQVHSAIPCFTYHNEF-NEQ-------------
OsRVE1      GAA----------GFAAQPVMPWLSYNGSCF-----------AAA
AtRVE7      VPY----------GFAAQPVMPWLSYNGSCF-----------AAA
AtRVE7Like  VPY----------GFAAQPVMPWLSYNGSCF-----------AAA
PpCCA1a     RSF------AVSAMAATIAAASTWWAMHGGPMHPAV---------QCGENIP
PpCCA1b     RSF------AVSAMAATIAASAWWAMQGGPMHPAA---------QCGQEVL
OsRVE3      NSG------GESAADARIPPLHVWWPYYGFI--HPR-----------SDESS
AtRVE2      IP----------GFAAQPVMPWLSYNGSCF-----------AAA
AtRVE1      NE--------LS--SSKTL----------PI-RNSQEELLCW---PL
OsRVE6      -----LP--LSP----------------------------------
OsCCA1-3    EAIVL--------------------------------------QS--
PaCCA1      ESNDARC---HAEKQ------------------GERSQCASNTRRGAF-
KfrVE       ETIIILMRNLAINLASPVFAEQQHILALAD--QPRM-IPEQE---EQ-
MpRVE       ETVLLLMRLSINLSSPEFEEAKLIFMPAFE--QLA--MLSAQ--PHRT-----
AtRVE5      ETVLLLMQNLNLINLSPEFEEQRLILSSYS----------------
AtRVE3      ETVLLLMKNSVLNTSPEFDEQRLIYNS--SDESS
AtRVE4      ETVLLLMRLNTVLSNPDQPTSETYVDAAE----SDESS
AtRVE8      ETVLLLMRLNTVLSNPDLESTRKVLSSYD----------SDESS
OsRVE5      ETVLLLMKNSLNTNPFEAHRKLASHG--HE------SDESS
AaRVE       ETVLLLMRLSINLSSPDFEEHKLFLSVYD--QING-PSTAPASKHRY-
SmRVE       ETVLLLMRLNAINLSSPDFQKFRAK--------CF------SDESS
OsRVE4      ETALLLMRLSINLSDQKKLLSSYS--EL-----SDESS
AtRVE6      ETVLLLMRLSINLSDQKDHRLLSSYD--TDH-----SDESS
PpRVE1      ETVLLLMRLSINLSSPDFEHKLFLMVPID--DST--IGPAQPTPSY-
PpRVE2      ETVLLLMHLNLSINLDPFEQ-------TARSH
PpRVE3      ETVLLLMRLSINLSDPDFFKLYLSFYV--DLTS--APSAAQ-TPSY-
KfCCA1      APGGVPR--EAAIRRAERANKLREL-----KGLRTHQASNSFPRGAE-
OsCCA1-2    ESDVEM--QIEQ------------------IHRSSSGLTNPAEGAF
OsCCA1      S--EVLK--HTEL------------------KQDRSSCSSGTNPAEGAF
AtLHY       D--VTKL--NADS------------------LVDRSSCSSGTNPAEGAF
AaCCA1      DEAPSRK--CSAARERANSRMKEQ------ARVRSVSGNTPRGADG
AtCCA1      E------E---ENKS------------------QVDRSSCSSGTNPAEGAF
OsRVE2      ---------------------QREGSLTGSNTA------
OsRVE1      ---------------------QREGSLTGSNTA------
AtRVE7 ------------------ETFCTGSNAFW-----
AtRVE7Like -----------------------------------FCTGS---S------
PpCCA1a INGN-SR--SSAMRRAERAQLVKELE-------SSNSSACGSSDGRRGAFL
PpCCA1b SNNNSHR--SSAMRRRAERAQRVKELE-------SSNSSACGSSDGRRGAFL
OsRVE3 S--------------------------------------------------S-L--
AtRVE2 -----------------------------------------------SETQSQCSSSTS------
AtRVE1 KQEDVEN--------------------------------------------QNEGSSSTGSNTG------

OsRVE6 ------FEPIRRLLS
OsCCA1-3 --SGAGFVPYKGCQK
PaCCA1 -KSGVGFPYERCS--
KfRVE ---------P----
MpRVE -GFGMPLPTLT----
AtRVE5 --------------------------------------------
AtRVE3 -----------------------------------------------
AtRVE4 -----------------------------S---
AtRVE8 -----------------------------A---
OsRVE5 -----------------------------K---
AaRVE -VYGLGMPPLA-----
SmRVE ---------KVRK
OsRVE4 ---------P----
AtRVE6 --------------------------------------------
PpRVE1 -VYLPLPLSPLP----
PpRVE2 -VYS--LPSLP----
PpRVE3 -VYSAFLPSLP----
KfCCA1 HGGSSGFVPYKGYQG
OsCCA1-2 -NFFQGSVTWETLCF
OsCCA1 -SRRTGFKPYKRCSV
AtLHY -TRQTGFKPYKRCSM
AaCCA1 -FSGVGFPYQRNST
AtCCA1 -RGRTGFKPYKRCSM
OsRVE2 -NCRRGFVPYKRCVA
OsRVE1 -LLQRGFMPYKRCAA
AtRVE7 -VNASGRFPYKRCLS
AtRVE7Like -VNASGRFPYKRCLS
PpCCA1a -YSGVGFPYQRVSI
PpCCA1b -YSGVGFPYQRAST
OsRVE3 -EVIRGFVPYKRCKF
AtRVE2 -NNRGFMPYKRRVK
AtRVE1 -SNSRGFGPYKRRKM

; end;
(b)
#NEXUS
BEGIN DATA;
dimensions ntax=25 nchar=361;
format missing=?
symbols="ABCDEFGHIJKLMNOPQRSTUVWXYZ"
interleave datatype=PROTEIN gap=- match=.;
matrix
KfTOC1  M----------HDVTTFHRSSVHILLADRLQTRKEVFELLRRCSYKVT
PaPRR1  M----------AGSSVLDRCVRILLCDKDPTNSQQLLELLRKCMYQVT
MgTOC1  M----------GRVVFDRSRRLKICLDPNTSP--EVVDLLQRCYSQVV
SmTOC1  M----------EGKIGDRSRVILLCDKDSATAQVEKELLRCQYSQVS
AtTOC1  M----------DGFIARDSRVLICNDTSSTGEVTLLSECSYQVT
OsPRR1  MV---------EGQQFVDRSKRVILLCDSDPSSREVLRLNCQYSQVT
OsPRR59 -----------------------------
OsPRR95  MG---------RETRMLPRMPVRLAEGDSRTHIALLRKCGRYRA
AtPRR7  MN---------CWERFLHVRTIRVLLVENDDCTRYTVTALLRNCSYEVV
OsPRR37 MM----------CWERFIIKTTIKVLLVSDDDSTRQVSAALLRHCMEVII
OsPRR73 -------------------------------
PaPRR3  MQ---------VKDGHQPPQRCLKVLLVEDDDSTRKVVRALLCHRQYEV
AtPRR5  MW---------KWERFLPKIALRVLLEVADDDSTRQIIAALLRKCQYRA
KfPRR  MT---------SQWATLPSKHRLVLLVEDDDCTRHVAAALLRNCGYEV
AtPRR9  MG---------QWEKYPKTVLRLVVESDYSTRQIIALLRRCYKV
SmPRR7b  M---------QAEAFQHHHLKVLLVEDDDSTRHVAVALLRNCGQEV
AtPRR3  MC---------KWERYLPVRLKVLLVEDDDSTRHVALTALKNCYEV
SmPRR7a  MT---------GWENLFQPRQLVLVLLVEDDDSTRHVAAALLRNCGQVV
PaPRR7  MW---------RWSLLVHLQPLVD----------VV-------SVV
AaPRR  MT---------GWESFLPQRLRLVLLVEDDDSTRHVAGALLRNCSYEV
MpPRR  MS---------AEQGWESFLPRKHIRVLLAEDDDSTRHVAGALLRNCGYEV
PpPRR4  MDVARAGE------GWESFLKRNKLVLLVEDDDATHVAGALLRNCNYEV
PpPRR3  M----------MLSL-L-N-------------VE
PpPRR1  ---------GWESFLKRSRLILLVEYDADTHVAGALLRNCDYEVT
PpPRR2  ---------SWESFLKRKRLVLLVEDDDATHVAGALLRNCNYEV

KfTOC1  AVESAKHVLLEVLRAAGPAVDVLISEVEFPHEGKGLHKLLKHMREELKQIP
PaPRR1  AVSTARGVVSILNAEGQIEIDLILADVLPLSKGKMLKGYIRSTRCLQRIP
MgTOC1  AVQXARQVMVLRNMGSDVLILAEVELPKKGGKMLKHITKEARKLQIP
SmTOC1  VXKTARQVEVNLHITDSDKVLVSLVSEVNLPNGGKFLKHLIKVSNFKHPI
AtTOC1  AVKSARQVIDALNAEGPDIDIILAEIDLPMAKGMLRKEYTRDDELRI
OsPRR1  CASKPRQVINLRLCEAGIDIIIAEVLPSKCFKMLKYIRNRLERH
OsPRR59 -------MKAWGVMRERAYAFDLVLTEVTMLPSGIELLSRIVASDECKNIP
OsPRR95  ASGDKLYADKLFLEGNDLVLVTEVLPMSGLLLSTIMEHADCKNI
AtPRR7  EASNQIQAQKWDELNHHIDILVETEIVMYLSGILLCIKILNHSRRNIP
OsPRR37 PAENGQQTWYLEDLMNQISDLVLETEVMPGVSGILSLSRIMNHICKN
OsPRR73 PAENGTLHAWCLEDIQLNQHIDLVLETEVMPLSLGIGILSWHSHKICDIN
PaPRR3  AVALQAWNLDLETRQFDVLTEAMGPCLSIGIDLLSKIMSHRTKYHIP
| Protein | Sequence                          |
|---------|----------------------------------|
| AtPRR5  | AVPDGLKAWEMLKGKPEVSDLILTEVDLPSISGYALLTLIMEHDICKNIP |
| KfPRR   | PAADGFQALKLSDMKQFDLVLTEVMPGLSGIQLLSRIOQRDOHKRMP |
| AtPRR9  | AVSDGLAWEVKHELKNIDILTELDDLPSISGFALLALVMHEACKNIP |
| SmPRR7b | PAASGLQAEILETRSSVDLVLTDMMPSRLMGSSLHKKRSPKRP |
| AtPRR3  | AVPDVLAEWRLEDKESCIDLVLTEVMDPVHSTGLLLSISHMTKLKNIP |
| SmPRR7a | PAANGLQAEILDNNREDLVLTDVLTPGSLGGLLSSIMIKNNHQKVP |
| PaPRR7  | AVANGQQAWKLEDPSNHLVLTEVMPCLSGLLCLLKERKCHTCKNIP |
| AaPRR   | SANGQLAEILQLLERLLQLGLLGLLLKMKDNISPKRP |
| MpPRR   | SAANMQAEILEDNNFDLVLTDVMPCLSGVLKIMHRKAGKRIP |
| PpPRR4  | PVANGSLGLLEEANSFDLVLTDVMPYLSGGLSMMKREACKRVP |
| PpPRR3  | P-------------PELEDANNSFDLVLTDVMPCLSGGLSMMKREACKRVP |
| PpPRR1  | AVANGSLGLLEEANSFDLVLTDVMPRLSGLMSKKRPKCRVP |
| PpPRR2  | SVANGSLGLLEEANSFDLVLTDVMPCLSGGLSMMKREACKRVP |
| KfTOC1  | VVMMSQRDEMALVVLRLGAADLYKPLRKNELNLWTHWRRRMGLM |
| PaPRR1  | IVMMSRARDETVVMKLPLGADLYKLPRRNELNLWWMWRRRMGLL |
| MpTOC1  | IVMMSRDEMAMVVLKLMGAADLYKPLRKNELNLWTHWRRRMGLM |
| SmTOC1  | IVMMSRDEMAMVVLKLMGAADLYKPLRKNELNLWTHWRRRMGLM |
| AtTOC1  | VIMMSRQDEAMVVLKLMGAADLYKPLRKNELNLWTHWRRRMGLM |
| OsPRR1  | IIMMSNREDEVSVVVLGCLGAELKPLRNELNLWTHWRRRMGLM |
| OsPRR59 | VIMMSQDSIGTVLQCMQKAVDFLKVPRKELNLWLQHVWRHAMSQ |
| OsPRR95 | VIMMSNDSMVFLKMLGAADLYKPLRKNELNLWTHWRRRMGLM |
| AtPRR7  | VIMMSHMDMLVFKLKALGAGDFVLFKPRKELNLWLQHVWRKCRQSSSG |
| OsPRR37 | VIMMSNNDSMVFLKMLGAADLYKPLRKNELNLWTHWRRRMGLM |
| OsPRR73 | VIMMSNDSMVFLKMLGAADLYKPLRKNELNLWTHWRRRMGLM |
| PaPRR3  | VIMMSHMDMLVFKLKALGAGDFVLFKPRKELNLWLQHVWRKCRQSSSG |
| AtPRR5  | VIMMSQTSDEVYVKLMGAADLYKPLRRNLWLQHVWRQTSLAP |
| KfPRR   | VIMMSKDATDVVFKCFQRGAADLYKPLRKNELNLWTHVWRKCCSSSG |
| AtPRR9  | VIMMSQDSIKMVLKMLGAADLYKPRKELNLWLQHVWRRLTLD |
| SmPRR7b | VVMMSCLDMSDVVLKMLGAADLYKPLRKNELNLWTHVRMLLSSSG |
| AtPRR3  | VIMMSHMDMLVFKLKALGAGDFVLFKPRKELNLWLQHVWRKCSVS |
| SmPRR7a | VMMMSHDTNVVFKLKALGAGDFVLFKPRKELNLWLQHVWRKCRASSSG |
| PaPRR7  | VIMMSHMDMLVFKLKALGAGDFVLFKPRKELNLWLQHVWRKCSVS |
| AaPRR   | VMMMSHMDMLVFKLKALGAGDFVLFKPRKELNLWLQHVWRKCSVS |
| MpPRR   | VIMMSHMDMLVFKLKALGAGDFVLFKPRKELNLWLQHVWRKCSVS |
| PpPRR4  | IVIMMSYDSLIGVRLFKLCAGCLYKVLKPRKELNLWLQHVWRKCSVS |
| PpPRR3  | IVIMMSYDSLIGVRLFKLCAGCLYKVLKPRKELNLWLQHVWRKCSVS |
| PpPRR1  | IVIMMSYDSLIGVRLFKLCAGCLYKVLKPRKELNLWLQHVWRKCHSL-- |
| PpPRR2  | IVIMMSYDSLIGVRLFKLCAGCLYKVLKPRKELNLWLQHVWRKCSVS |
| KfTOC1  | TRLEGQVNEPDER-EDSRPQMS-YGSPAKFRTPPAP--------F--VPS |
| PaPRR1  | RNHAG-------PEIS-------ASH----------EL---------- |
| MpTOC1  | RSVMDVANGLSQ----------YLLTTSV----------KP----------VLTVCNHI-- |
| SmTOC1  | KKLKNPSEE-------------EI---------- |
| AtTOC1  | RSTNPQGRNLSNH-EWSVATA-DHLAGSHHEPMKR-- |
| OsPRR1  | ENINQETSTSNQ-EHNSDA--------EGGQ----------|
| OsPRR59 | SKTGEHSDEESDAQSSGSKREV-HALKGNDAPSNG--------LRSSDP |
OsPRR95     HKNMECSEQESDAQSSCSTRSEL-FSKPDHKNTEKNG--------CEKASR
AtPRR7      LNASDGSSDGSQ ASSWTTKAV-DVTMGRLDLEISR--------ESPS SK
OsPRR37     LNARDGSNGSTGQSSWTTKRAV-DFDKGKDEIGSP---------NEPTTQ
OsPRR73     HNARDGSNGSTGQSSWTTKRAV-DSMKGKDEIGAP---------NEPTTQ
PaPrR3      SKI VDGSSGDGKGTQAG--------DCA MGKLE IAPI--------VEPLQE
AtPRR5      DHVVS G-NDGAQQSCSTRPEM-------------------QSKQAE
KfP R R     VNVGGGS DNGSG TR N E SSVPVD-QRE-----ETPER------SR SSEK
AtPRR9      -D SRYHSDQGSQGAQAI----------------------KETF DV
SmPrR7b     LNVGGGS DNGSGTQSSWTTKPV-DRDCG- EVAMA EK--------GGGE PS
AtPRR3      LQDGS DNGSGTQSSWT KRAS-------------------QIGTG--------SKKAEE
SmPrR7a     IRISDQSDTES----SCTKKA-PK------------CPPS R------TDRAME
PaPrR7      LNRRDGSSDNGSGTQSSWTKRAV-DFAMQDSEIAVP--------SEVATK
AaPrR       LNLRRGSSDNGSGTQSSWTKRAT- EIESAHVDSVKK----YEQCE
M PPr R      LNRVGGSDNGSGTQTVAEVQ-DEQM QDLEMATR--------SGGS AK
PpPrR4      VNAQGGSDNGSGNACMQPVQ-DEMGHDLMATR--------EESSPK
PpPrR3      LNAQGGSDNGSGNACMQPVQ-DEEMGDLEMATR--------EESSPK
PpPrR1 ------LAA--------SS-----------SS-------------ES
PpPrR2      LNAQGGSDNGSGNACVPQPVQ-LDSKMGDLEMATR--------EESPK
KfTOC1      SGWHPN A QHAPPP PEYH RSQA QGDARR GTEG E----GSWQNSHRPPYP-...
PaPrR1      --------QK--------------------------- NVSVSPEIPI G
MpTOC1      -----------PA--------------------------CMGPPAA----
SmTOC1      ------------------P---------------------- HHQHSIPWS S
AtTOC1      SLHLHRGLAEKF----------GTEQ------YH-------SGQTLQNGA YPH
OsPrR1      RLDS------S---------------TPPV------YH-------FP------ FYYPG
OsPrR59     GKL DL LV V AQ TDRSHFVEINLKEKQHRNGYTNKL NEKDIFNHNS SSA FSRY
OsPrR95     DMELVHIIDNQQAHLQLESLVRS RDSYS RLEQ KNERRTLNHST TSPS LY
AtPrR7      QMHEDGGSSFK--IMHVEHSSKRHRGH TDKG DTRDNLRVLSRSEGSAFS RY
OsPrR37     AADLIGM AKNMMPSELSLKR SRSTGDGAN QEEQRNVLRRSDLSAFT RY
OsPrR73     TVDLISSIARSTM TSEEGLKRLKT GSGATSEIHDERNILKRS DLSAFT RY
PaPrR3      AIDLIGSIARKTPLLELTRRHSPSDGD--VHEPHVLRS HSGASAFS RY
AtPrR5      AIDFMGASFRRTRIE DLSSLRPNAS-ENQSGDRPSLP HSASAF SRY
KfP R R     AADLIGNFVEEKM PGLELSLKR GRPAGDSSE AQQMRTLRO QS GSAF SRY
AtPrR9      TMDL I GIDKRPGPEGLGLS LKRSCVSVSFENQDESHQKQL LSLDA SAS AF SRY
SmPrR7b     CKDLASIQQADSALELSLKTRVKGPDPDGDE GQKRLLLHSGS AF SRY
AtPrR3      PGDLEKNAKYSVLQSLQLEQTLKKTRED-----RDYKVGRSVLRHNSLNASFSKY
SmPrR7a     AVDLIGNIAE PSA TPLELKASPTP------DHAYEPSYVLNQ SGPSA SF RY
PaPrR7      AADLIGAIANKPLPFLELTKPRQNGKEDGEPEDHRVLRQGYS AFS RY
AaPrR       PMDLTATISRQALPTLELSLRKRPRAAGDHEGETDDRRVLRQGYS AFS RY
MPPR        AIDLIGIACQNMPTLELSLK RSRVPGDNSSEE AQQMRTLRO QS GSAF SRY
PpPrR4      IIDLINVIACQPSTLSELKRPRSAVNGGECLERQPLRHS GSAF SRY
PpPrR3      AVDSLHRRRVQSLVDCEILGR----------------------WY
PpPrR1      AVDLINIAQICPGSMLELSLKRPRSAVDNDGTELQPLRHS GSAF SRY
PpPrR2      AIDLINVVACQPMPLELSLKRPRSAVDNDGTELQPLRHS GSAF SRY
KfTOC1      -PQFGAMPQRESQLPSDL M------PHLAMFNLP----- PQSKPPQVTG--
PaPrR1      -AAAG--------SMHLC-H-GVHDV-----------------------
| Protein   | Sequence                                      |
|-----------|----------------------------------------------|
| MpTOC1    | PSEIQPM--MMTHKETL                            |
| SmTOC1    | TGAA--I-MVSHI                              |
| AtTOC1    | ERSR--DGSGF-S-APNAYP-YYMH                   |
| OsPRR1    | GKRRELLSCVE-GAAIPYHYGAIMQPMYQPA----RKG     |
| OsPRR59   | NCRTAS--LTFDGQPFNPVAPASLYFPQSA---KTLS     |
| AtPRR7    | NPASNA--IKVKh--SSFQPL--LPPQCGSSNVNYSVN     |
| OsPRR37   | HTPVASAMKTDVKANGHTSAFHPAQH--GAPQCGSSNVNYGVN |
| OsPRR73   | HHTVASALKTDVKATQHTSAFHPVQR----SDACQCGSSNVNYSVN |
| PaPRR3    | NTCKNHFPMPEILYNGVPQYGAAPAMFPHTCTVPNCGLSSM-ESV |
| AtPRR5    | VHRPLQ--SNQLNSWPQGSSYPTPT---GGL          |
| KfPRR     | SGGAMPQAFMPGPFPFQPPFHFQPYQPQESGLAGSSRSQ---ANN |
| AtPRR9    | EESKSA----NQENI---G------------------------------- |
| SmPRR7b   | SNNPLL----VAFNDAAYAQSLHHFFHYGSSSTQCGSSNVNGSVN |
| AtPRR3    | NNQAA----LSYDQLSPARGATY--YQTQPAPTCGGSNTNV|GN |
| SmPRR7a   | NDQAA----LSYDQLSPARGATY--YQTQPAPTCGGSNTNV|GN |
| PaPRR7    | NTSGGQAVAIEMPYDTTPRAYGSAIHPIYSHGTTPCLGSSNMGN|AN |
| AaPRR     | STSVVNGLGKDMTDYDGLSAAFGSMHPGFYPRPAGHRCGSSNVNGGN |
| MpPRR     | STSGVISTAFDIIYDVGVCASYGPMHPSFYSHPA-----NGSVN |
| PpPRR4    | GSGTIGMDLADYGIGTYGAPMHPMYAHPSAPRCGSTVGSMN |
| PpPRR3    | GSGTIGMDLADYGIGTYGAPMHPMYAHPSAPRCGSTVGSMN |
| PpPRR2    | SDGGTIGMDLADYGIGSYGAPMHPMYAHPSAPRCGSTVGSMN |
| KfTOC1    | GPVIGSVS-MDQAEREEALSKFRKQKRDKRCFZKIRYASRRKLAECARP |
| PaPRR1    | ISQVERREAAANKFRQKRKRDFDKIKYRSRKLAECARP |
| MpTOC1    | STRTTANTYFEQAEREEALNKFQKRKRKDFEKIRYSRKLAECARP |
| SmTOC1    | STRTTANTYFEQAEREEALNKFQKRKRKDFEKIRYSRKLAECARP |
| AtTOC1    | QHSQMSVNKLDRREEALLKFRKRNQRCFDKKIRYVRNKRALKAEPPR |
| OsPRR1    | QSNQLPCRSRERRAALAKFRLKRKFCFDKVKYVRNKRKLAEERP |
| AtPRR5    | QSNQLPCRSRERRAALAKFRLKRKFCFDKVKYVRNKRKLAEERP |
| OsPRR37   | GDGGTQMGMDLADYGIGTYGAPMHPMYAHPSAPRCGSTVGSMN |
| OsPRR95   | GSGTIGMDLADYGIGTYGAPMHPMYAHPSAPRCGSTVGSMN |
| AtPRR7    | GSGTIGMDLADYGIGTYGAPMHPMYAHPSAPRCGSTVGSMN |
| OsPRR73   | GSGTIGMDLADYGIGTYGAPMHPMYAHPSAPRCGSTVGSMN |
| PaPRR3    | ASEHGSNNRPNQSAQREAAALKFRKLCRCEFVKRYQSRKLAECARP |
| AtPRR5    | QDRDCSMVIGQQLQEREEALLKFRKRNQRCFDKIKYVRNKRKLAEERP |
| KfPRR     | TSGGGSNFGLVRDAREREEALNKFQKRKERNFEKIVKRYQSRKLAECARP |
| AtPRR9    | KAASKERWRSRSQREAAALKFRKLCRCEFVKRYQSRKLAECARP |
| SmPRR7b   | GSVGGSGNLSAblrTRREVALYKFQKRKERCFCFEVKRYQSRKLAECARP |
| AtPRR3    | GSSSSDLNDPSDWAQREAAALKFRKLCRCEFVKRYQSRKLAECARP |
| SmPRR7a   | GSVGGSGNLSAblrTRREVALYKFQKRKERCFCFEVKRYQSRKLAECARP |
| PaPRR7    | SAUTGSNNRPNQSAQREAAALKFRKLCRCEFVKRYQSRKLAECARP |
| AaPRR     | GSGSNNATTKFRSREELAlanKFRQKRKERCFCFEVKRYQSRKLAECARP |
| MpPRR     | GSGSNNLQNFREREEALNKFQKRKERCFCFEVKRYQSRKLAECARP |
| PpPRR4    | GSGSNTQMRFREEREEALNKFQKRKERCFCFEK----PKQETAP |
| PpPRR3    | GSGSNTQMRFREEREEALNKFQKRKERCFCFEK----PKQETAP |
(c)

#NEXUS
BEGIN DATA;
dimensions ntax=11 nchar=204;
format missing=?
symbols="ABCDEFGHIJKLMNOPQRSTUVWXYZ"
interleave datatype=PROTEIN gap=- match=.

matrix
PpELF3-2 AGPALFPRLHVAETKKAGPRGPPRNKMALYEQTLPSHRYKMSPLPPSKK
PpELF3-1 PEIALFPRLHVAETKKAGPRGPPRNKMALYEQTLPSHKFKPSPOMPPTNK
PpELF3-3 VGAALFPRLHVAETKKAGPRGPPRNKMALSEQTLIPHSKFKPSPAPIPSKK
KfELF3 EAAALFPRLHVSHAKRNGPFPRAKMNALYEQVTLPASHPKAAGASSRPR
SmELF3 PGGALFPRLHVKTNPAGPRPPRNKMALYEQTLPSHRFQQQQSDKTRA

PpPRR1 GSAGSNTGQIRFARREAALNKFRQKRERCFEKK-RYQSRKLAEQRPR
PpPRR2 GSAGSNTGQMRFARREAALNKFRQKRERCFEKK-RYQSRKLAQRPR
KfTOC1 IKGQFVRQKEG
PaPRR1 VRGQFVRQTNND
MTOC1 NRGQFVRRTSD
SmTOC1 VKGQFVRRAEE
AtTOC1 VKGQFVRKMNG
OsPRR1 VRGQFVRQANY
OsPRR59 VKGQFVSQKLL
OsPRR95 VKGQFVRQDHG
AtPRR7 VRGQFVRKTAAG
OsPRR37 VRGQFVRQAVQ
OsPRR73 IRGQFVRQSQG
PaPRR3 VRGQFVCQAVC
AtPRR5 IKGQFVRQVQS
KfPRR IRGQFVRQATF
AtPRR9 VKGQFVRTVNS
SmPRR7b IKTJRFDNNEA
AtPRR3 VRGQFIRKRD
SmPRR7a VRGQFVRQAVF
PaPRR7 VRGQFVRQTV
AaPRR VRGQFVRQAVY
MPPR VRGQFVRQTVY
PpP RR4 C-------------
PpP RR3 VGAKCCIAALR
PpP RR1 VRGLFVRQAAH
PpP RR2 VRGLFVRQAVY

; end;

AtEEC       TIPPLFPVRHVNDTGRGGL------------------------SQQFDGKTMS-----
AtELF3      ILEPMFPRLHVNDADKGGPRAPPNKMLEYEQLISPSQRFGDHGMQRKM
OsELF3-1    VMGPLFPRLHVNDAAKGGPRAPPNKMLEYEQTFTPSHRFSGGSGGRK
OsELF3-2    VMGPLFPRLHVNDAAKGGPRAPPNKMLEYEQTFTPSHRFSGGGSAGGRK
AaELF3      ETGLALFPRLHVKETKNVPRAPPNKLMEYEQTFTPSHRFMSSLPAAKK
MpELF3      ETGLALFPRLHVKETKNVPRAPPNKLMEYEQTFTPSHRFMSSLPAAKK

PpELF3-2    GKEEDDFAVPTYSSAETEESILEDPLGSSCQFQELLDHFQGQGELWKAQK
PpELF3-1    ESLGDDFAVPTYSSAMEETGSGKLMLENVVMGSIERFDVVDFAGQQGFFWKVQM
PpELF3-3    GRDDEDFAVPTAHGSGARTGSGLMLENVPTDFQVLRDVNFQGGFQWTQK
KfELF3      QGEEDGCVIPASYGPDVESNGASAAAAAIGLPKDVGSVITQHGFQWKSRR
SmELF3      PRVDDVAVPFTPSEKDSATKVVNAAPSAITSAEVMSAVGDEEFWKR
AtEEC       --EKNQFS--PIYNTKVASNCSAESLSSGASSYDARVIGEKRFWKMRT
AtELF3      VREEEDFAVVPYINSDDSVSMDSSISSIDVSPPDVVGILGQKFWRAR
OsELF3-1    LADDDEFMPSVFN3DSDDSVSSVCITGWEISPDKIVGAIGKTFWKARR
OsELF3-2    LADDDEFIVSPVSDADDLSDDSVSECTAWEISPDEIVGAIGAKHTFWKARR
AaELF3      HRNEDDFPTVSPDDSSSMVESITIHRISPWHVGAGQEFEWTRRK
MpELF3      GRNDDCVPITPYTAPEESDATMEDPSHKITPKDINAVGQGFEWTARK

PpELF3-2    AIIRQKQVFQRSQVFELHRLVIEQVQLLAKPLSLFRL-SNPYARSHMYVP
PpELF3-1    MIIRQKLFALQVFEHRAEIVQQLLAKLCSFLHV-ANPCATRMYVPP
PpELF3-3    VIIRQDRSKOLFDLHRIMEVQHLLAKEKCFSTV--TNPWAPMYVYP
KfELF3      AMQQQKKFSAQLFEMHRLVKQKMFAENPDFTETEQVPNPWLQNAAY-A
SmELF3      AMQRQQTFOQQLFELHRLIKVQKLIAASSSHKLTVQHMANSKID---YNYQWYAP---P
AtEEC       YMINQQKIFAGQFQVELHRLLIMVQKMVASKPNLFLES-GNQWLVPVLVYK
AtELF3      AIANQQRVFAQVFEHLRHLIKQVQLLIAASPDIDEQG<OPWILFVLYK
OsELF3-1    AIMNRQRVFAQVFEHLHKLKVQKLIAASPHLIES-QNQWLVPVLVYK
OsELF3-2    AIMNRQRVFAQVFEHLHKLKVQKLIAASPHLIES-QNQWLVPVLVYK
AaELF3      AIRQQRISFSQVFELRHLKQVLQMLAMPSVEYTV-NPWLGAMYPCH
MpELF3      ILQRQQMFASQVFELHCLIQKVQQLVGEDEEDWE-ANGWFGALMY-P

PpELF3-2    YPGPQSFGPWFPIIQCGVINVKKVVPRAVATQESAASILLSIQKERQR--
PpELF3-1    YSGPCPPWRISRFPTLYGGVKKVVPRAVATQESTASILLSIQTERR--
PpELF3-3    YPPTCSPWARFPQILQQGIVPLIKVAPAHTSASANILLSIQTKEWR--
KfELF3      FSNPAAQWELQFPVKAAPKIVPKPRASETPDGLVSILSIQKRER-----
SmELF3      FQFPFPPV--------SHKAIKVTPTAARPVTAESAASIALHISQKERPS-
AtEEC       FPGCPCTY-------MLRAIKAVPHNSTSAEASAIRFRIQΕΕΕΕΕEE
AtELF3      HPGMAHTY-------VTTRIKVVPNKLASEANARIFQSIQERKR
OsELF3-1    YSGPCPPY-------QTNIKVVPHNSTASEAARIFRISIQMERQRD
OsELF3-2    YSGPCPPY-------QTNIRVPHNSQTSEASEARIFRISIQMERQD
AaELF3      SAGPVGWLPIFLPVPVVQGQVKKVAPRAVATSEASAILLISIQQERRR--
MpELF3      YPGFQAHWALPLFPLAFFGKIVKVPVRAMVTAESAALILLSIQQKRQ--

PpELF3-2    ----
PpELF3-1    ----
PpELF3-3    ----
KfELF3      ----
SmELF3 ----
AtEEC    HMIS
AtELF3    SSKP
OsELF3 1----
OsELF3 2  S----
AaELF3 ----
MpELF3 ----

;
end;

(d)
#NEXUS
BEGIN DATA;
dimensions ntax=21 nchar=87;
format missing=?
symbols="ABCDEFGHIJKLMNOPQRSTUVWXYZ"
interleave datatype=PROTEIN gap=-- match=.;

matrix
AtELF4    MK--GGEDPAMWENLDRNFRQVQSVLDRNRSLIQVQVNDNHQSRMADNMSK
AtEFL1    MEASGVDEVEWVDTESNFGKRAQLYLDQNRDLIQVRVENHMSRIPDNVSR
KfELF4-1   MDLSTLGISKEWPG-AANFDKVMILHDQNRLLIKEINQHNCARRTEGLSR
SmELF4-2   MESSLGAEGKAWGKFQRTFHFQVQFLLDHNRLLIKEINVNQESQIPESLSR
KfELF4-2   MDLTVFDPHEWPGFEK-FIIVQAILIDHNKLLNEINLHERRRPEGLTR
AaELF4-1   MDE----KFNAEETFHRFFKEVCMILDHNNLLISEINHNHETRIPERRORLTR
OsELF4 3    MEEDSGGGGKQVQVLQRFGEVQILEQNRVLIEISQHARDADGLTR
OsELF4 2    MGDSTAEDGKVLHAFQTSFVQVQSLIDQRNEINEQHESKVPDGLSR
OsELF4 1    MGDSQVQNLKLIQTFHKSFQVQVQIDQRMLLEINEQHESRAPDNLTR
AtEFL2    MEQDVMQGLQINFQKFQVQDILDQNRLLINEQHESQADHLGR
AtEFL4    MEQDVQMDGKLLQSFQSFVDQVQDILDQNRLLINEQHESQPDNLGR
AtEFL3    MEQDVGQMDGKLLQSFQSFVDQVQDILDQNRLLINEQHESQPDNLGR
PaELF4-1   MGDAIQDKVMQTFQKSFGQVQNILQHNRLLINEQHESKQRSRTPNLTR
PaELF4-2   MGDAIQDKVMQTFQKSFGQVQNILQHNRLLINEQHESKQRSRTPNLTR
PaELF4-3   MGDAIQDKVMQTFQKSFGQVQNILQHNRLLINEQHESKQRSRTPNLTR
PaELF4-4   MGDAIQDKVMQTFQKSFGQVQNILQHNRLLINEQHESKQRSRTPNLTR
PaELF4-5   MGDAIQDKVMQTFQKSFGQVQNILQHNRLLINEQHESKQRSRTPNLTR
AaELF4-2   ------MDRAKLSKHFHSLSQDFVQDIDHNRVIDEIINQNEEKIPESLDR
SmELF4-1   MEADAVDRKMWAAFERGFSQVFQVQFLLDHRNLLEINEQHESKQPSL
MpELF4    MDRDQMDKTVAGFHQQNSQVQLDHRNLLEINEQHESKQPSL
PpELF4    MDRQPMQVGAHQQNSQVQLDHRNLLEINEQHESKQPSL

AtELF4    NVALIQELNGISKVVMYDSLNTSFSSGFHGTR---
AtEFL1    NVLIGNEINISQVMEISLDLSNFAKFDQ-----
KfELF4-1   NVLLIRELHSNVAKVELEYREISESVVVFVGN--VQ--
SmELF4-2   NVMLIENKYNKREYVAGLPTSIFRSNKLARS

13
\( \text{KfELF4-2} \) NVQ LI R E L N E N V T K V M K L Y E E L S Q A F V Q S F G K --GQK
\( \text{AaELF4-1} \) NV TL I R E L N H N I L K V V G M Y S D L S A V F E K S F E N A K M R I
\( \text{OsELF4-3} \) N VAL I R E L N T N I AR V V D L Y A N L S G S F S R S V T A S K S R A
\( \text{OsELF4-2} \) NV TL I R E L N N N I R R V V D L Y A D L S L F A A S P A H K V R S
\( \text{OsELF4-1} \) NV TL I R E L N N N I R R V V G L Y A D L S A F S A R T M D A Q K V R P
\( \text{AtELF2} \) NV GL I R E L N N R I T V A S L Y G D L S H S F A R S V D A Q K F R S
\( \text{AtELF4} \) NV GL I K E L N N N I R R V A S L Y G D L S H S F A R S V D A Q K F R S
\( \text{AtELF3} \) NV GL I R E L N N N I R R V A S L Y G D L S H S F A R S V D A Q K F R S
\( \text{PaELF4-1} \) NV TL I R E L N N N I R R V V D L Y A D L S S F A S P A H K V R S
\( \text{PaELF4-2} \) NV TL I R E L N N N I R R V V D L Y A D L S S F A S P A H K V R S
\( \text{PaELF4-3} \) NV GL I R E L N N N I R R V V D L Y A D L S S F A S P A H K V R S
\( \text{PaELF4-4} \) NV GL I R E L N N N I R R V V D L Y A D L S S F A S P A H K V R S
\( \text{PaELF4-5} \) NV GL I R E L N N N I R R V V D L Y A D L S S F A S P A H K V R S
\( \text{AaELF4-2} \) NV SL I T E L N Q N I K V V D L Y S G I S S F A S T F Q S ----- ---
\( \text{SmELF4-1} \) NV VL I K E L N K N I G Q V V S L Y S T I S S F V K S F E N H K V R S
\( \text{MpELF4} \) NV LL I R E L N N T K V V D L Y A S L S V S F K L F D N M K A R S
\( \text{PpELF4} \) NV ML I R Q L N S N I G K V V D L Y A N L S N F S ----- ---

\( \text{e} \)
#NEXUS
BEGIN DATA;
dimensions ntax=13 nchar=176;
format missing=?
symbols="ABCDEFGHIJKLMNOPQRSTUVWXYZ"
interleave datatype=PROTEIN gap=- match=.;
matrix
KfLUX-1 MSRESISKVTGLGSEDTSLEGELTLTLGLSAGLGSVEQAIAAAEA
KfLUX-2 --KEITRRWGVGISPEELAPLSTRLITPVLACAFNIEGDSGASEEEVDE
AaLUX -----MQLGVS-----------------------------------------
AtBOA MGEYRIPEIQLPNGDLLTPLSRLYVPSILALAFSMIPERSRTIHVPN
OsPCL1 MGDRVMEWETGLPGADELTPLSQQPLPGCLAAAFRIPPEPGRSTL
PpLUX2 --DDQVSDWTKGLPTSEELTPLSHSTLISRIASAFRIKHEEPMD
PpLUX1 -------------------------------------------------
PpLUX3 MHGDHVVEWNPGLPTGEELNPTSKSLSLISLVLASLSMKPEPLKTAADVSG
MpLUX MANNRVREWEAGLSADLTPLHLHSLITVRLARAFCIPTNATSAPADVLQ
PaLUX --EDWLGKWEDELPSPPELMPPLTQNLTPDAAAElfKHPSSASAP---
AtLUX MGGRDVEWEMGLPSDEDSLSPLSILPNLAMAFSPRTSRIQDVNR
PaLUX2 --DDRVEWEAGLPGPDELTPSQLPVLFPELACAFSISPECRSQLDVSR
PaLUX2 --NDRVEWEAGLPGPDELTPSQLPVLFPELACAFSISPECRSQLDVSR
KfLUX-1 PGEQAGAQAGGTEGEEEQREGRGQGNKKQRLVWTAELHRSRFFMNAVNLG
KfLUX-2 AALNTVTPSEQMEHTSDDPPGRALKRPRLWTPQLHHRFCDAVNHLGVE
AaLUX     -----------GGPENSNEEPNARTLKRPRLVWTPQLHKRFVDAVGHLGIK
AtBOA     ASQITSSLR-----GPEDASGKTSKRPRLVWTPQLHKRFVDVVAHLGIK
OsPCL1    ASAATVSLRGGADGGNTNNSSRASKARLWVTPLHKRFVEVVAHLGKM
PpLUX2    ESQATTHMLFGGLMNSNDE-----PKRARLWVTPLHKRFVEVVAHLGIK
PpLUX1    -----------GGPVSNEEPNARTLKRPRLVWTPQLHKRFVDAVGHLGIK
PpLUX3    ESASFSLQGGPVSNRENTARTLKRPRLVWTPQLHKRFVDAVGHLGIK
MpLUX     ASKTVVQQLDCP-----DGNTARTLKRPRLWTPQLHKRFVDAVGHLGIK
PaLUX     -----------GVTTENGG-PARTLKRPRLVWTPQLHKRFVDAHLGIK
AtLUX     ASETTLSLRR-----GTDLSGKTLKRPLWTPQLHKRFVEVVAHLGIK
PaLUX2    ASKNTMSALRSFENGSSDEHSARTLKRPRLWTPQLHRFVDVSAHLGIK
PaLUX2    ASIAMTVTALRGFPSEFEEPAARLKRPRLWTPQLHRFVDAHLGIK

KfLUX-1    NAVPKTILQLMNVEGVTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
KfLUX-2    NAVPKTIMQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
AaLUX     NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
AtBOA     NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
OsPCL1    NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
PpLUX2    NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
PpLUX1    NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
PpLUX3    NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
MpLUX     NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
PaLUX     NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
AtLUX     NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
PaLUX2    NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG
PaLUX2    NAVPKTQLMNVEGLTENVDASHLQPSAPIARPPHHACKYRLYKLKLAG

KfLUX-1    LPPNARMMAPPLHQGPYPPPPPHRMQ
KfLUX-2    DAPAVPSSSDAVFANGQPPPGRGGAH
AaLUX     VPTDGPANDQLFASTPLPPNGLFP
AtBOA     LTTEEDSSSDQLFSSTPVPPQSFVND
OsPCL1    LSNGGSPSFHDIFASTPVPHA-----
PpLUX2    LSNQPASDHLAFASMLPPGMLFP
PpLUX1    LPSDGMPANDQLFASTSLPSNLGLFP
PpLUX3    LSSDGPPANDQLFSSSTPLPPNLGLFP
MpLUX     LSNGGSAADPLFASAPLNNLQFA
PaLUX     LSSEGPSADDQLFASTPPVEDF--LFP
AtLUX     LTNEGPSASDKLFSSTPVPPQSFANE
PaLUX2    LSNGGPSLSDHLFASTPVPHASHA-----
PaLUX2    LSSEGPSADDQLFASTPVPHASHA-----

;
end;
#NEXUS
BEGIN DATA;
dimensions ntax=9 nchar=1001;
format missing=?
symbols="ABCDEFGHIKLMNOPQRSTUVWXYZ"
interleave datatype=PROTEIN gap=- match=.;

matrix
CcGI ------------------------------------------MAFVDLLLLQ
CiGI PGDDQDRSEDDLPKLRWLQGLVSSSSLRRPPSNNQQRQVEVLAYTELFQG
AaGI ---------MSVARPEWLLGLQSSSLFRPPPRNHQEMEILGYLEILMVQ
PaGI ---------MSISEQKWHLGLQSSSSLFRPQPQDAQRQRQAELAYVELFAQ
AtGI ---------SSSSSERWIDQLQFSSSLWPPRPDPQHQKDQVAYVEYFGQ
OsGI ---------MSASNKEWQDLQFSSSSLFPWPPQDQKQAQLAYVEYFGQ
MpGI ---------MSFPQGWGWLQGSSLFQPPESAIQRKQAELASVELFGH
SmGI ---------MSQPPQWHTLGYLSTSSLFPRPPLDHQERTKTVAYVEYFGQ
TlGI ---------RQQWLRGLQSSSSYRPPPANHRQRQTEVLEYVELFGH

CcGI LSQDLFAHVIALRILQYPREHPILLDVLASIVLHSPHEHAILHVLLS
CiGI FASELLSDDISQLVRNYHQPQNSSTLLDDVLATFVHLHPEHDATVHLLS
AaGI FASESFVDENMLEVRSHYPKEEACLVDLLAASFVHLHPEHDATVHLLS
PaGI FTS-EFPDDIAELVHSHYPNGEASLLDDVLAIVFVHLHPEHAILHLLS
AtGI FTS-QFPDDIAELVRHQYQPSTKRRLLDDVLAMIFVHLHPEHAVLPIIS
OsGI FTAQFQFEDIQLIQSCYPSEKELRQLVDLEFVHLHPEHAVBVPIIS
MpGI FASESFVDIGELVRAHETYNECHLLDDVLATFVHLHPEHAVHLLLS
SmGI FASDSFPEDALVDRHPKHEPCCLUDLLDDVLATFVHLHPEHAGTLHLLS
TlGI FASDAFVDDIGELVRHYPSEEEVCLDDDLDDVLATFVHLHPEHAILHLLS

CcGI SLVHGNLRYSRGISPSFSAFVAIFSPSSQRCAPCEQYELACAELQLLTHYN
CiGI CIIDGSLYESRKSPFPFSFVALFTNSTERMGLRQWALACTEILRVLTHYN
AaGI CVIDGTLYSEKGSFPRCFIALSFSSANEKHFSEQWAVACETLRLTHYN
PaGI CIIDGTLYDJKMSPSNFSNLFSPPSENDYSEQWALCEILRVLTHYN
AtGI CLIDGSLYVESKHFPFASFISTLVCPSSENDYSEQWALEILRILTHYN
OsGI RIIDGTLYSDRNGFPPFSFISLFSHTSEKEYSEQWALCEILRVLTHYN
MpGI CVIDGTLYNKTPPSFVSLFSPSTDFSEQWALCEITLRLTHYN
SmGI CVIDGTLAYSFTPPGSFVSFGVSSERDLTEQWALCEILRLLLLTHYN
TlGI CVIDGTLYDKTPPSFVSLFSPSTEFSEQWALCEITLRLTHYN

CcGI RPVQRSTSS---------GRVAPERVLTQWTIDCLLASPLGFRDYFRWSCGGV
CiGI KPVHKNRQEEDTNRNERRDARFRRLTMITIDCLLLAAPS1MTNYFRWCGGV
AaGI RPYVKSDP---------EKRVPSRLTPWITDSLAAAPGLVRSDFYFRWCGGV
PaGI RPYVYKIDR---------QDKKPLRLQTPWITDSLAAAPGLGIRDYFRCGGV
AtGI RPIYKTEQ---------HERKPLRPLSPWISDLLAAGPLGISDYFRCGGV
OsGI RPIIFKVDD---------PDRLPLRPLSPWITDSLAAAPGLGIRDYFRCGGV
MpGI RPTFRSDT---------DRRSPTRLLTPWITDSLAAAPGSRDYFQWCGGV
SmGI RPIYKSES---------GRRAPKRLLLTPWITDSLAAAPGLGSKDYFRCGGV
| Protein | Sequence |
|---------|----------|
| TlGI    | RPTYKAET-\-\-\-EKRVPTRLTPWITESLIAAPLGRSDYFRWCGGV |
| CcGI    | VGQYATSGDLHPLPSP--SGPGKHQPQLMPS\-TPRWAVANGAAAVIVSVDDE |
| CiGI    | TGYAAVGAL\-RPTTTL\-GNGQGKAPQLQPSPTPMWAMANGAAAVILWVDDE |
| AaGI    | MGKYAAGVDLKPSTV奶奶GKQPLLPSPTPMWAPANGAAVILSVCDDE |
| PaGI    | KGKTAAGGKLKPPTTAGGRGPGKHLMPSTPMWAPANGAVIGLSVSVDDE |
| AtGI    | MGKYAAG-ELKPTTAS-RGSKGPQLMPSPTPMWAPANGAGVILSVCDDE |
| OsGI    | MGKYAAGKPTTTAYRSRSGKHQPQLMPSPTPMWAPANGAVIGL SVSVDDE |
| MpGI    | MGKYTAGEELRPTTASSRQQGKQPQLPSPTPMWAPANGAVIGLSVSVDDE |
| SmGI    | LGKYAGGGLRPPTTGDCKGKHGPQLLLPSPTPMWAPANGAVIGL SVSVDDE |
| TlGI    | MGKYAAGVLRPTTASSGQRQGKHQLLLPSPTPMWAPANGAVIGL SVSVDDE |
| CcGI    | VARWGSVELTAA\-APALLLP\-EVEDEHLI-GLP\-PLEF\-ARLFHRYYAI |
| CiGI    | VTRFGTADLTAAAVPALLLP\-PDHDDDARAVALP\-PLEYTLRHFHVYTM |
| AaGI    | VERYRSAHLTAAAV\-APALLLP\-PEPA\-ALNILV\-APLEYQLFHRYYAI |
| PaGI    | VTRYETANLTAAPALLLP\-PTTTLDEHLA\-PLPLEYACLHFHVYAI |
| AtGI    | VARYETALTLAADV\-PALLLP\-PTTTLDEHLA\-PLPLEYARLFHRYYAI |
| OsGI    | VARYETAL\-PALLLP\-PPPTTLDPEHLA\-PLPLEYARLHFHVYAI |
| MpGI    | VRYETADLAT\-PALLLP\-PPPTTLPDEHLA\-PLPLEYP\-ARLFHRYYAI |
| SmGI    | VRLYETADLAT\-PALLLP\-PPPTTLPDEHLA\-PLPLEYARLHFHVYAI |
| TlGI    | VMRYETAL\-TAA\-PALLLLPP\-PATALDQL\-VLA\-PLPLEYARLHFHVYAI |
| CcGI    | ATSGATQRLLGLLEAL\-RPDALDA\-AVQLVELLRL\-SAEDGYSGACQLPNN |
| CiGI    | GAPVATQRLLLGLLTDVG\-P\-L\-ADAMEA\-AVLH\-RS-EDFGPPK\-LPP |
| AaGI    | ATPSATQR\-LLL\-LPALLES\-SWAPDALDA\-AVQLVELLA\-EDYAST\-LMR LPN |
| PaGI    | A\-TPSATQR\-LGP\-LPALLES\-SWAPDALDA\-AVQLVELLRA\-EDYGVR\-LPRN |
| AtGI    | A\-TPSATQR\-LGP\-LPALLES\-SWAPDALDA\-AVQLVELLRA\-EDYDSG\-LMR LPN |
| OsGI    | A\-TPSATQR\-LGP\-LPALLES\-SWAPDALDA\-AVQLVELLRA\-EDYDSG\-LMR LPN |
| MpGI    | VPSRATQR\-L\-LPALLES\-SWAPDALDA\-AVQLVELLRA\-EDYDSG\-LMR LPN |
| SmGI    | VLSRATQR\-L\-LPALLES\-SWAPDALDA\-AVQLVELLRA\-EDYDSG\-LMR LPN |
| TlGI    | VPSRATQ\-L\-LPALLES\-SWAPDALDA\-AVQLVELLRA\-EDYDSG\-LMR LPN |
| CcGI    | WLHLHFLRP\-G\-V\-AMAK\-Q\-S\-G\-A\-S\-ADAAA\-LLFR\-SFQP\-V\-LL\-F\-TP\-P\-HP\-GA\-AQGYPD |
| CiGI    | WLLHFLRP\-G\-V\-AMSSRAAV\-S\-A\-S\-ALLFR\-SFQH\-T\-L\-F\-PPP\-A\-LPD |
| AaGI    | WLLHFLRP\-G\-SAMAR\-Q\-SVI\-A\-D\-AA\-LLFR\-SFQP\-V\-LL\-F\-PR\-GA\-Q VR |
| PaGI    | WLLHFLRA\-T\-G\-AMRMV\-G\-IA\-AD\-T\-AA\-LLFR\-SFQ\-P\-LL\-F\-PL\-LAQ\-GDV |
| AtGI    | WMHLHFL\-R\-G\-AMSM\-RG\-AVA\-DA\-A\-ALLFR\-SFQ\-P\-PLL\-P\-QLV\-E |
| OsGI    | WMHLHFL\-R\-G\-AMSM\-RG\-AVA\-DA\-A\-ALLFR\-SFQ\-P\-PLL\-P\-QLV\-E |
| MpGI    | WLT\-H\-L\-R\-P\-G\-AMQ\-A\-R\-G\-ST\-TA\-D\-AA\-ALLH\-F\-R\-P\-L\-P\-QQG \-L |
| SmGI    | WFR\-L\-H\-L\-R\-P\-G\-AM\-T\-K\-Q\-G\-I\-A\-S\-D\-A\-A\-ALLH\-F\-R\-P\-P\-P\-R\-G\-H\-A \-G |
| TlGI    | WLHLHFL\-R\-P\-G\-AM\-T\-V\-R\-G\-I\-A\-S\-D\-A\-A\-ALLH\-F\-R\-P\-P\-P\-G\-H\-A \-G |
| CcGI    | AGRFV\-V\-KE\-Q\-A\-A\-A\-A\-A\-A\-D\-A\-T\-A\-L\-G\-A\-V\-E\-E\-W\-R\-C\-L\-W\-E\-A\-A\-F\-G\-L |
| CiGI    | TSS\-LR\-M\-M\-Q\-V\-L\-H\-A\-E\-E\-T\-A\-R\-G\-A\-L\-M\-C\-H\-G\-P\-D\-W\-E\-R\-I\-C\-S\-V\-L\-E\-A\-A\-Y\-G\-L \-T |
| AaGI    | VIPFR\-D\-E\-I\-E\-M\-A\-A\-V\-\-A\-\-Q\-\-N\-\-G\-A\-T\-C\-A\-A\-A\-L\-M\-E\-H\-G\-P\-D\-W\-E\-R\-I\-C\-S\-V\-L\-E\-A\-A\-Y\-G\-L |
| PaGI    | VQ\-C\-\-S\-G\-E\-E\-I\-E\-M\-Y\-Q\-\-S\-I\-E\-T\-A\-Q\-G\-V\-A\-S\-L\-M\-C\-L\-H\-G\-P\-D\-W\-E\-R\-I\-C\-S\-V\-L\-E\-A\-A\-Y\-G\-L |
| AtGI    | IQHYR\-K\-Q\-I\-E\-V\-A\-E\-E\-I\-T\-A\-Q\-G\-I\-A\-S\-M\-L\-C\-H\-G\-P\-D\-W\-E\-R\-I\-C\-S\-V\-L\-E\-A\-A\-Y\-G\-L |
OsGI        LHHYKRQLEVPASEATIDATAQGIASMLCAHGPDVEWRICTIWEAAYGLL
MpGI        STQSREAAKAAAQESEVATASGGLALTGHGIDVECHICAIWEAAYGLR
SmGI        VVQFHAQMEALATQVNEEATAKGVASLMRDHGRDVEWRICALWEAAYGLL
TlGI        INIQEVAADAAAQAANEETATAAQAVAALMCTHGLDVEWRICALWEAAYGLL
CcGI        PPQSSSSQLLPLLVANLAIYPFLPVLSWNVLRPLRRFHPHPNPSQSSLGR
CiGI        PVDSTVIALSHVGGISEVRHVLVLSWNLFRPLLLHFEHLPQPGPSHACLRR
AaGI        PVDLSVNVQQLDMTQLPFPVLVSWNLHFPLRRVMEHLRTGTRSQACLRR
PaGI        PLSSSTVFLPEIVATQLPPPVLSWNLQPLLRLVLEYLPGRSPSEASLMR
AtGI        PLNSSSAVDLPEIVATPLQLPPILSWSWNLPLLKLVEYLPGSPAELMKG
OsGI        PLSSSAVDLPEIVTAPQPLTFKLPLLLRALSILFYEYLPGRSPSEACLMK
MpGI        SLTSSVLPLDPVLSTPLQPPVLSWNLRLAlFRLISLYLPSPSQAACLKR
SmGI        PLDKSVDLMPVEMTQLPLSSLPFLTVPPLVEQGCQSTQLCRR
TlGI        PLNSSLKPDVLPEIVATPLQPPFLSWNLVLPRLVYMELPREPSLSLVMRR
CcGI        IFSATVQLGLERTQFSDESAGARLVGMTEIRALVHHLTEPCPSAPALASK
CiGI        IFSATLSSVLFSFPLTGHPPQTPQTLQPOQNDQQALHHHHHHHDLSSP
AaGI        VFSATVEAIVQRTFPCEENAGAKNVGMGALRMAVGLHFTEACPSDVAA
PaGI        IFTATVQLIALRRTAPQPSSTKSNRAELRTMVHLSFTESCVPIDLASR
AtGI        IFVATVETIRLSFPSEPSSATKNLASEMRLAVHALFLESCAGVFLASR
OsGI        IFVATVEAIRLRTSFQSETSSQSNRAELRTMHLFVECSADLAR
MpGI        IFVATVEAILRTFPLEDEVAGAKSVGMELRAMLHCLFTESFLSPFALAH
SmGI        IFSATVAIIRSFPLDVSAGMAELRALVHLFTETAFPSALASQ
TlGI        IFSATVEAILRTFQPSVEGAASQGVMELRATLHCLFTETSCPSDLGSQ
CcGI        LFTRLLTSLAYDAISARSVIGVATGRGRRREVCAGGGVGSRMNGG
CiGI        LQFPLPLHSQSSSSSSFP--K---FPATQ--IPIAwapla------WPPPS
AaGI        LLADALLCLESQKRGQADDQGRAVATFDSYLMAMCALAECVMAFSR
PaGI        LLFMSITVCLSHDAVQKAKKERGAVATFGSIALAAVCAQACEVQLFSSTS
AtGI        LLLFVVLTVQSVHESQSSNVGQGPPVAAPDSYAVLAAVCAACVMAQYPMIS
OsGI        LLLFVVLTVQSVHESQSSNVGQGPPVAAPDSYAVLAAAACLSELQFPFS
MpGI        LLLSALSLCSSDASQQGKQGRHAVATFDSYAVLAAVCAACVQPFPSP
SmGI        LLSDALTVCLSHDTRLQSNKDRGAVASFDSYIAAVCAACVQQLTFS
TlGI        LLSDVLTVCLSHDQSLVSQKRAGAVATFDSYIAVCACVQAMCQSLPLR
CcGI        VRSQEGQGGQEGGQIEEPSREMAGVDQAFFPFRETEAEGVGTGDEMSG
CiGI        PTHT-EHGPFPYHHHHPLHT-----HTTPSVPAAAENORSATQKMLATP
AaGI        STSS-FPALSNGTPVGIQQAARKVLELLERLLAEVPASLAEQSSSSIDVGD
PaGI        PMPV-VEGPPNGFQAVNTRRLLGIAMELSKPAAGCSSESISIIAAA
AtGI        GGGN-SKEYGAGIDSAISHTTRILAILEALFSKLPSVSQYTSSSEIIVA
OsGI        KNGN-SNEIHSISISAILHTTRILGILEALFSLKPSVSQYTSSSEIIVA
MpGI        AYFT-SNRFPGYFANVLKETKLRLGGELLGVPSFAGTTSTNEILQG
SmGI        ADGT-------GVTNSAQARLMSVEGLLVPEFSPVGNPNDLVSEA
TlGI        SMFT-PNVYVPEVAADAHTVRVLLGLESSLALVSHPSSHGTSSGESIVA
CcGI        SGGQIETKSQGTPMANGAQMTEKTTGGSFQKRKRRSEKSVRGARGV
CiGI        STLSLPLLPPGFLLLLLLLQLQTRLPWPPSLLLSFYPYHHHHHPSSL
AaGI  ISAAHISGVLGKSACLQLAGAVRCKWDSEVCTRASHALELVQAVCQTM
PaGI  MAVAAHVDLLGSKCAMHAI5SVMMRCKWDSELCARASRTVLALIDVNRKAV
AtGI  MAVAAHISEFRRSKALTHSLGMRCKWDKEIKHRASSLYNLIIVDSKVV
OsGI  MAVAHISFRRSRPCLNLASALKQCKWDAEISTRASSLYHIDLHFKT
MpGI  VGAAMHDSLQSLHSLTGLRMIRCKWPGILKASALSASIESGDLV
SmGI  IVAAMHISLRSQTHAILRCKWPGVSSKALSILAVDGNDKAV
TlGI  VVAAMHSSLGRSNACISHLTAIRCKWPGIQEKAASVSLAVEGNDVV

CcGI  KGAARARRGGLTSC-AAFSMSRSRSGSPLQLGADADASCQD-KVGAIMLL
CiGI  LLLLLLAAAGGGGGA-VETSSAPSSTEP-HPSGLGKGVPSSRLALRGQSSTMQWSGPIRVI
PaGI  ASIJKSESSSVSHV-E-DVSGKVIALNSMDGSVSLTVDNDVGNVFHAV
AtGI  ASIVDKAEPLAYLTDGSEGKLKIDFILLAASDNFLTDATGQLLRSV
OsGI  TISVNAKPLEAHL-ELNGSGKAILSLQVEASDLNFLMDSQTLRSLV
MpGI  AIATHGRNGFGGD-GENVGEARSFLPLEASDVLCSKTLTSDLKSV
SmGI  EAVFYNAKDLSIGE-ASGDKVDALSLSVDSTNLCC-TVSDLKAV
TlGI  ATAVHASNYESKID-RINGMEKIANGILLPLAADASSCNVKSICNLKLV

CcGI  LMSQQDVTAVPVLWQRLVLPHEVSVSQEGTASSRSRQVVAEALCNLVV
CiGI  LMEKOQLLAAAPLWQRLMCAPVLSTEGTSAEKWRQVLEGCLNIVV
AaGI  LTKEKDELVSSSVPVIWVRLAIPDPNSTIMNGASPRQGWRQVEAICLKVE
PaGI  LEEKQDLVESPLWQRLWPLITPEMSVETSAKQGWRQVVDALCNIVL
AtGI  LAEKPELSFSVLSSWHKILAAPIQDTEASTSAQQGWRQVVDALCNVVS
OsGI  LSEKQELCSVSVLWKLASPEMQMSAESTSAHQGWRQVVDALCDIVS
MpGI  LVGKQHILAIGATALLCQRLILAPDNPNAEGTSAERGWQVVEANCEFSA
SmGI  LKQKRLDVAIPLWQRLMAEELTPESTKEAQSAGWRQVVDAVCNVV
TlGI  LTKWQLAVAAVPPLLWQRLIFAPAEPVSSEGTSAKGRGWQVVEACINNVV

CcGI  AFSPGKASSAILAQAERDLKPVMPNVEVATAIAPDWSVNTRIMSLLAEILQQR
CiGI  APLQAOAVAILAIAEERLEPAWAEEDVGGDVCGRNGNIRVAISLAEMLLIR
AaGI  TFPGBKTSAILLQVQRDSLWSWGDQSDMRVENRKVRNRTIILVEHRLKL
PaGI  ASPAKAATAIVQVAAERDLQPWVARHDSQGQQWLRQIRVSLAEILLYR
AtGI  ATPAKAAAVAILQVQARLQPWIAKDEEQKMKWNQRIKVLVELRNH
OsGI  ASPTKAADIAVQQLAQEKLQPIARDDEGQQKMKVHRQVRKLAEILRNH
MpGI  AFSQVKVARFVQSAEAELRPWTVKDDGPSQSLWIRTRVCCULLSDILRT
SmGI  TEOKATSVVLQAIRGQIPWIGIDG--GEKWRMNTRIVFSLLELRNL
TlGI  GLPAQAAAIAIVQVQERDLHPWTRDDLQVQTVWVRNTRLSSLTELLRL

CcGI  RSREAWTAVACSESSLRLRATDGLIALAGVETPLLQELLEAVSVARSAQA
CiGI  QCPAAAVGVAHAAHALLHRATADIGIEGEACTPLLQELLEAATAADVARH
AaGI  HCEPAAILQALAKSSDLLRATDGLIEEADAYLPLLQELLEAAANIAVQRSLS
PaGI  NAPWALMNASDLMRATDGLMVDGCACTTPLLQELLEAMAVTAQLSLG
AtGI  DRPESLVIALASDLDLLRATDGLMVDGCACTPLLQELLEATARAPVLA
OsGI  DSPEALVILASDLDLLRATDGLMVDGCACTPLLQELLEVTARAVHLVE
MpGI  HLPKIVVQVHDGALLRRATDGMGLDGEAACLPLLQELLEAAALAAQBAALK
SmGI  D-PQVLGILANAGTLVQATDGSVDGPECTLPLLQELLEAAAALK
TlGI  TSEAVKVVQVSSDLHRATDGMVTGEACTPLLQELLEAMAVQVLLR
**CcGI**

```
WPIGGCQVAGKLLDLLKVRIPSLKVCKVCHENAHVRALATSLLQTILDTPD
```

**CiGI**

```
WGHSVTELLFMLKERRPVTVRCLSSTSTHVRALAMALLRHLAYEAV
```

**AaGI**

```
WDNSGVGQAQKLVAVLKDRIPOAVCRCLHSVCDARVLSQTFIQLDRVADS
```

**PaGI**

```
WVPFGKAMAEGLWNLLKYRLPATVQCLHSSAHHVRLSTSVLRDLHAES
```

**AtGI**

```
WGPSGLAVVDGLSNLKLKCRRLPATIRCLHSHPSAHHVRLSTSVLRDINQQSS
```

**OsGI**

```
WGDSGVSVADGLSNLLLKCRRLSTTIRCLHSHPSAHHVRLSMVSVLRLNSGQ
```

**MpGI**

```
WEEPNEVADQFLSRLRELPATVRCLHSNHSTHVROMSVARLRDMLYMES
```

**SmGI**

```
WKVSS----RGLLILLKITERALPAICLIHDSPIRASSASLLRREIVSTDV
```

**TlGI**

```
SEVSGKNALDKLILLKDLPLPAICLIHDSPIRASSASLLRREIVSTDV
```

**CcGI**

```
SPMAASWQAVERAVSLVSRDRQRHGLEVRILEEVSFALGNCNT-C------
```

**CiGI**

```
LAARCEWRRAVEQCLSEWAHFRSGLSTLLVDAASTLHCQIAV------
```

**AaGI**

```
VRLDSDKRAVEEALMFMEAQHRRAARLSTLSLLAAATVGLTIPN------
```

**PaGI**

```
LNFVQDWNKAVEQCAWAEHNRQARGMISIALALLAANLGSAN------
```

**AtGI**

```
IPISIDWKADIQNCEAHLSSLLSTMPTQFILTAARELGCSTIL------
```

**OsGI**

```
INSIINWQADVRCWEQAEHSSRAATGLTLAFLTAALKGCLPL------
```

**MpGI**

```
LRSTGNWRETYVQCVVWEKGSSASS-TIQIVEEYLSRGTSVWIWKL
```

**SmGI**

```
LRAAGAWLEDVQSIAWETHYRRAEGLSESFLASAIALGCLKLF------
```

**TlGI**

```
LQGEDWRRAVEKCEWEAYYRATGMSLSSASSLGCSP------
```

**CcGI**

```
-
```

**CiGI**

```
-
```

**AaGI**

```
-
```

**PaGI**

```
-
```

**AtGI**

```
-
```

**OsGI**

```
-
```

**MpGI**

```
I
```

**SmGI**

```
-
```

**TlGI**

```
-
```

; end;

(g)

```nexus
#NEXUS
BEGIN DATA;
dimensions ntax=12 nchar=559;
format missing=?
symbols="ABCDEFGHIKLMNOPQRSTUVWXYZ"
interleave datatype=PROTEIN gap=- match=.
```

```
matrix
KfZTL       MEDRIEVLHSSPHGLVTDALAEHPITYVNTIFQYTYGDAESILGRNC
AtFKF1      HAIVGMYMPMTPPSIVSDALEPDPFIYVNRVEVFYTGYRASEVLGRNC
OsFKF1      GDRDEAAAEGERAAIIVSVAEVDFPVIYVNAFEAATGYRASEVLGRNC
```
AaZTL       MDGRVKMLQTSCPALVTDALELDQPVIYVNSIFEHATGYKAEEILGRNC
AtLKP2      MIEWPGSLTAPCGFVVSDELPDNPIIYVNTVEIYTGRIAEEVI-----
SmZTL       -------MLGPCSVVTDALDVDPFPIIYVNNIFEFITGKYEAVLGRNC
OsZTL1      MEWAIEGV--GACGLVVSDELPDFPIIYVNRGFDATGRIAEEVLGRNC
OsZTL2      MEWAIEGMASPCGMLVTDALDPCPIIYVNCGFFEEATGRIAEEVLGRNC
PaZTL       MEWAMDNLSTPCGFAVTDELPDPIIYVNAFFEYVTGKAEEILGRNC
AtZTL       MEWPVGNLHTAPCGFVVTDAVEPDPIIYVNTVFEMTVGRIAEEVLGNC
MpFKF1      --MRELVHSSPCGMLVTDALPDPIIYVNTVFEGITYKAEEILGKNC
TlZTL       MDWRLEMLHSSPCGMLVTDALPDPIIYVNTVFEGITYKAEEILGRNC

KfZTL       RFLQMRGEFADKRPAVDLKTVRKMEIAAIGQEFKGEILLNFKSKDGPTLI
AtFKF1      RFLQYRDPAQRHRLVDVPVSVSREIRCLEGHIEFQGELLNFKSKDGPTLI
OsFKF1      RFLQFRPFAQRHRLVDPMVPVESIRECLNEIGEFQGELLNFKSKDGPTLI
AaZTL       RFLQNRGPFAQRHRLVDEATTISIEIRRCISGVEFEGILLNFKSKDGPTLI
AtLKP2      -------GPTKRHPMDVSTIARAQMRCLEIGEFQGELLNFKSKDGPTLI
SmZTL       RFLQCRGPFAQRHRLVDASAVTEIRRCREIGEFGQGELLNFKSKDGPTLI
OsZTL1      RFLQCRGPFAQRHRLVDSATTVEIRRCREIGEFGQGELLNFKSKDGPTLI
OsZTL2      RFLQGRGPFAQRHRLVDAMVCTEIRRCRevIGEFGQGELLNFKSKDGPTLI
PaZTL       RFLQRYRGPFAQRHRLVDPMVESIRECLDEEGEFQGELLNFKSKDGPTLI
AtZTL       RFLQCRGPFAQRHRLVDPMVESIRECLDEEGEFQGELLNFKSKDGPTLI
MpFKF1      RFLQCRGPFAQRHRLVDVTAVTIRKCTEGFEGILLNFKSKDGPTLI
TlZTL       RFLQYRGPFAQRHRLVDVSTVEIRRCRGEFEGILLNFKSKDGPTLI

KfZTL       NNLLMTPIHGEDIVTHFIGIQSFPIKMDLGPLPERPWHDSRHNAPVFA
AtFKF1      NTRLAPIRDDDGTIHTVIFGQFSETTIIDLDVRVSPVFKKHDPCGIQ
OsFKF1      NTRLAPIPMHDDDFQVTHTIGIQLFSEANIDLSNVYVPYKQSEYCCILQ
AaZTL       NKLFLTPILDEGAVTHVIGQSFQAEKIGMGPLSNAILKERSSTCGILQ
AtLKP2      NKLRLVPIDDGAVTHVIGQSFQAEKIGMGPLSNAILKERSSTCGILQ
SmZTL       NKLCLPIRQEDRGITIIGIQSFSEVKDLGILPLPPLWRNSKRDGCILR
OsZTL1      AKLQLTPIYGDEDTITMYGQFFNDNVDDLSPLVSTKEDSEHSDLFL
OsZTL2      NKLHLPITGDITDITMYGIQFFTANVDLGPLPLPSLTKEDSEHSLFL
PaZTL       NKLRLTPITGHGIITIIGIQSFTEALGDSPLSAIKEKSDGCILQ
AtZTL       NKLRLTPIGDITDIITIIGIQSFTEALGDSPLSAIKEKSDGCILQ
MpFKF1      NKLCPITIADGTVITHVIGIQSFTEAKLGLPLPGW---WKDSDKACGLLQ
TlZTL       NKMCLTPIHAGDGVTHIIGIQSFMEALKDLGPLPCWSKDRSDDCGLLH

KfZTL       LSDEVLVSRILGRLAPKDVAICSVMCRCVRFFRRLQDDYIWKRCRNWSNGH
AtFKF1      LSDEVLHNLISRSLTRPDVAGISGACRRRLQILTKESVRKMCQNAWGE
OsFKF1      LSDEVLHNLISRSLTRPDVAGISGACRRRLQILTKESVRKMCQNAWGFD
AaZTL       LSNEVLSYRILAVSPRDVAGISGLVCRRHLQHTKNDLLKLCVQNAWGD
AtLKP2      LSDEVIAIKIQLSGDGIAVGCVRRLNLKNDLWVVRMCQNTWGTE
SmZTL       LSDEVLVQKLQALQRTPRDVSVALCRRFNETNKTLWLVRCNAWGE
OsZTL1      LSDEVLQKLQALQRTPRDVSVALCRRFNETNKTLWLVRCNAWGE
OsZTL2      LTDEVLCQISLRSRPRDASVVSSCVCLRLLLRTNECLWVRMCQNAWGE
PaZTL       LSDEVLAQKILSLRTPRDASVVSSVCRLQLTKNDLWVRMCQNAWGE
AtZTL       LSDEVVSMKLISLRTPRDASVVSSCRLVLTKNDLWVRMCQNAWGE
MpFKF1      LSDEVLTHKIIAYVAPRDAALGLVCRRLHEITKNDLWRSVQNSWGE
TlZTL       LSDEVLAQKILALVAPRDVAIGSVCKRLHELTKNEDLWRMVCQNAWGSE
KfZTL       TAAAAIQALDTPSLGWARIARELTLEAAAWRKFTVGGVESPRCNFSEC
AtFKF1      ITGTLEIM--TKKLWGRGRELTLTEAVCRCWFTVGGVQFESPRCNFSEC
OsFKF1      VTVRLEMS--TKMLWGRGRELTLTEAAAWRKFTVGGREVESPRCNFSEC
AaZTL       ATMSLQALPGTSLWGRGRELTLTEAVWRKFTVGGAVEPRCNFSEC
AtLKP2      ATRVLESVGAKIGRWLRELATEFTTATEAWKFsVGGTVEPESRCNFSEC
SmZTL       TTAVERVHNPRLSDWGLARELTLEAAAWRKLVGGAVEPRCNFSEC
OsZTL1      ATQVLETAGTSLWGRGRELTLTEAVTRKFTVGGAVEPRCNFSEC
OsZTL2      TTRALETVPAAKRLWGRGRELTLTEAVWRKFTVGGAVEPRCNFSEC
PaZTL       TTRVLETVGAKRGWLRELATELLEAAAWRKLVGGAVEPRCNFSEC
AtZTL       TTRVLETVGAKRGWLRELATELLEAAAWRKLVGGAVEPRCNFSEC
MpFKF1      ATKALESVPAGDNLWGRGRELTLTEAAAWRKLVGGAVEPRCNFSEC
TlZTL       ATAMLEAVPPAGRIWGLARLELTLEAAAWKLTLVGGGVEPESRNCNFSEC

KfZTL       AVGNKVLFLFGGEVNMQPMDTFVLDSLLEHPAWRHVDVSAAPPGGWGHT
AtFKF1      AVGNRLFLFGGEVNMQPMDTFVLDSLLEHPAWRHVDVSAAPPGGWGHT
OsFKF1      AVGNRLFLFGGEVNMQPMDTFVLDSLLEHPAWRHVDVSAAPPGGWGHT
AaZTL       AVGNRLFLFGGEVNMQPMDTFVLDSLLEHPAWRHVDVSAAPPGGWGHT
AtLKP2      AVGNRLFLFGGEVNMQPMDTFVLDSLLEHPAWRHVDVSAAPPGGWGHT
SmZTL       AVGNKVLFLFGGEVNMQPMDTFVLDSLLEHPAWRHVDVSAAPPGGWGHT
OsZTL1      AAGNRVVLFLFGGEVNMQPMDTFVLDNSKPEWRHINRASAPPGGWGHT
OsZTL2      AVGNRVLFLFGGEVNMQPMDTFVLDNSKPEWRHINRASAPPGGWGHT
PaZTL       AVGNRVLFLFGGEVNMQPMDTFVLDNSKPEWRHINRASAPPGGWGHT
AtZTL       AVGNRVLFLFGGEVNMQPMDTFVLDNSKPEWRHINRASAPPGGWGHT
MpFKF1      AVGNKVLFLFGGEVNMQPMDTFVLDSLKVPAWQHVNKSAPPGGWGHT
TlZTL       AVGNKVLFLFGGEVNMQPMDTFVLDSLKVPAWQHVNKSAPPGGWGHT

KfZTL       LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
AtFKF1      LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
OsFKF1      LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
AaZTL       LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
AtLKP2      LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
SmZTL       LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
OsZTL1      LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
OsZTL2      LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
PaZTL       LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
AtZTL       LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
MpFKF1      LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
TlZTL       LCCLNGSWLVLVFGGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW

KfZTL       HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
AtFKF1      HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
OsFKF1      HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
AaZTL       HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
AtLKP2      HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
SmZTL       HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
OsZTL1      HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
OsZTL2      HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
PaZTL       HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
AtZTL       HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
MpFKF1      HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
TlZTL       HSSCTVGDYLVFVGCGTDGPLNDVFVLDDLDAEHPAWREVAGAPPLPRSW
AtFKF1      HELCLASRQ
OsFKF1      HELCLASRP
AaZTL       HELSLASKM
AtLKP2      HELLLATST
SmZTL       HELSLSSTS
OsZTL1      HELSLMFPT
OsZTL2      HELSLASST
PaZTL       HELSLASKH
AtZTL       HELSLASYL
MpFKF1      HELSITHKS
TlZTL       HELSLASKV

;
end;
**Fig. S2** Inferred phylogeny of homologs to the ELF3 gene family. The tree was constructed using MrBayes and PhyML on an amino acid alignment of proteins retrieved from *Arabidopsis thaliana* (At), *Oryza sativa* (Os), *Picea abies* (Pa), *Selaginella moellendorffii* (Sm), *Physcomitrella patens* (Pp), *Marchantia polymorpha* (Mp), *Anthoceros agrestis* (Aa), and *Klebsormidium flaccidum* (Kf). The Bayesian tree is shown with posterior probabilities (above) and bootstrap proportions from PhyML analysis (below) for each node. Nodes with conflicting support from the two methods were collapsed. Branch length is relative to the thickness of individual branches: the shortest branches have a straight line and the longest are increasingly triangular.
**Fig. S3** Inferred phylogeny of homologs to the ELF4 gene family. The tree was constructed using MrBayes and PhyML on an amino acid alignment of proteins retrieved from *Arabidopsis thaliana* (At), *Oryza sativa* (Os), *Picea abies* (Pa), *Selaginella moellendorffii* (Sm), *Physcomitrella patens* (Pp), *Marchantia polymorpha* (Mp), *Anthoceros agrestis* (Aa), and *Klebsormidium flaccidum* (Kf). The Bayesian tree is shown with posterior probabilities (above) and bootstrap proportions from PhyML analysis (below) for each node. Nodes with conflicting support from the two methods were collapsed. Branch length is relative to the thickness of individual branches: the shortest branches have a straight line and the longest are increasingly triangular.
**Fig. S4** Inferred phylogeny of homologs to the LUX gene family. The tree was constructed using MrBayes and PhyML on an amino acid alignment of proteins retrieved from *Arabidopsis thaliana* (At), *Oryza sativa* (Os), *Picea abies* (Pa), *Selaginella moellendorffii* (Sm), *Physcomitrella patens* (Pp), *Marchantia polymorpha* (Mp), *Anthoceros agrestis* (Aa), and *Klebsormidium flaccidum* (Kf). The Bayesian tree is shown with posterior probabilities (above) and bootstrap proportions from PhyML analysis (below) for each node. Nodes with conflicting support from the two methods were collapsed. Branch length is relative to the thickness of individual branches: the shortest branches have a straight line and the longest are increasingly triangular. The SHAQKYF-like motif in LUX homologs is SHLQKYR.
**Fig. S5** Inferred phylogeny of homologs to the GI gene family. The tree was constructed using MrBayes and PhyML on an amino acid alignment of proteins retrieved from *Arabidopsis thaliana* (At), *Oryza sativa* (Os), *Picea abies* (Pa), *Selaginella moellendorfii* (Sm), *Marchantia polymorpha* (Mp), *Anthoceros agrestis* (Aa), *Cylindrocystis cushleckae* (Cc), *Coleochaete irregularis* (Ci) and *Takakia lepidodzioides* (Tl). The Bayesian tree is shown with posterior probabilities (above) and bootstrap proportions from PhyML analysis (below) for each node. Nodes with conflicting support from the two methods were collapsed. Branch length is relative to the thickness of individual branches: the shortest branches have a straight line and the longest are increasingly triangular. † indicates gene loss in Pp (as well as all other mosses except Takakia).
**Fig. S6** Inferred phylogeny of homologs to the ZTL gene family. The tree was constructed using MrBayes on an amino acid alignment of proteins retrieved from *Arabidopsis thaliana* (At), *Oryza sativa* (Os), *Picea abies* (Pa), *Selaginella moellendorffii* (Sm), *Takakia lepidiozioides* (Ti), *Marchantia polymorpha* (Mp), *Anthoceros agrestis* (Aa), and *Klebsormidium flaccidum* (Kf). The Bayesian tree is shown with posterior probabilities (above) and bootstrap proportions from PhyML analysis (below) for each node. Nodes with conflicting support from the two methods were collapsed. Branch length is relative to the thickness of individual branches: the shortest branches have a straight line and the longest are increasingly triangular. † indicates gene loss in *Physcomitrella patens* (Pp; as well as all other mosses except Takakia).
**Fig. S7** Temporal expression patterns of putative circadian clock genes in *M. polymorpha* (Mp) under ND, LL, and DD conditions. Plants were sampled every fourth hour during 2 d. Quantitative RT-PCR expression values with standard errors are based on two biological replicates and were normalized using three reference genes (see the Materials and Methods section for details).
**Fig. S8** Generation of MpRVE knockout mutant. The exon-intron structure of MpRVE and the construct used for MpRVE transformation are shown in (a). Primer sets used in (b) are illustrated as arrows with numbers. The result of genomic PCR of MpRVE knock-out line and WT are shown in (b). Numbers refer to primer sets as shown in (a).
**Fig. S9** Generation of MpPRR knockout mutant. The exon-intron structure of MpPRR and the construct used for MpPRR transformation are shown in (a). Primer sets used in (b) are illustrated as arrows with numbers. The result of genomic PCR of MpPRR knock-out line and WT are shown in (b). Numbers refer to primer sets as shown in (a).
Fig. S10  Generation of MpTOC1 knockout mutants. The exon-intron structure of MpTOC1 and the construct used for MpTOC1 transformation are shown in (a). Primer sets used in (b) are illustrated as arrows with numbers. The result of genomic PCR of three MpTOC1 knock-out lines and WT are shown in (b). Numbers refer to primer sets as shown in (a).
**Fig. S11** Temporal expression pattern of *MpPRR* (a) and *35S:LUC* (b) under ND and LL conditions. Quantitative RT-PCR expression values with standard errors are based on 3 biological replicates and were normalized using *MpEF1*.
**Fig. S12** Temporal expression pattern of MpPRR in WT, Mprve<sup>ko</sup>, Mptoc1<sup>ko</sup> and restored lines of Mprve<sup>ko</sup>. MpPRR expression in WT, Mprve<sup>ko</sup> and a restored line are shown in (a). MpPRR expression under ND conditions (b) and LL conditions (c) in WT and two additional Mptoc1<sup>ko</sup> lines. Quantitative RT-PCR expression values with standard errors are based on 3 biological replicates and were normalized using MpEF1.
**Fig. S13** *pro:LUC* bioluminescence for MpELF3, MpGI, MpLUX and MpRVE. Plants were entrained in 12 h light, 12 h dark photoperiod and transferred to T=12 photocycles. Light intensity was set to 5 μmol m$^{-2}$ s$^{-1}$. Averages from three replicates of one transformant per gene are shown. Expression patterns were readily adjusted to a T=12 photocycle without frequency demultiplication.
Fig. S14 Luciferase imaging in transgenic *M. polymorpha* plants expressing luciferase under the control of *M. polymorpha* promoters. (a–c) MpRVE<sub>pro</sub>:LUC. (d–f) MpELF3<sub>pro</sub>:LUC. (g–i) MpGI<sub>pro</sub>:LUC. (j–l) MpLUX<sub>pro</sub>:LUC. (a, d, g, j) 4-d-old gemmaling showing expression at the apical notches. (b, e, h, k) 7-d-old gemmaling with expression at the recently split apical notches. (c, f, j, l) 4-wk-old thallus showing strong expression in apical regions and young gemmae cups. Bioluminescence is pseudocolored in green.
**Fig. S15** $MpPRR_{pro:GUS}$ expression in mature thallus. $MpPRR_{pro:GUS#7}$ stained overnight (a, c) or for 4 h (b). Dorsal side (a), and ventral side (b) of mature thallus. Close-up of dorsal side showing air chambers with expression in chlorenchyma cells (c). Arrowheads points at strongly stained meristematic regions.
Fig. S16 Temporal expression patterns of putative circadian clock genes in *A. agrestis* (Aa) under ND, LL, and DD conditions. Plants were sampled every fourth hour during 2 d. Quantitative RT-PCR expression values with standard errors are based on two biological replicates and were normalized using two reference genes (see the Materials and Methods section for details).
| Species                        | Gene name | Gene family/Clade | ID/Accession     |
|-------------------------------|-----------|-------------------|-----------------|
| *Chlamydomonas reinhardtii*   | Data from Phytozome v11.0 (https://phytozome.jgi.doe.gov) |                   |                 |
| CrLHY (ROC40)                 | RVE       |                   | Cre06.g275350   |
| CrPRR-1                       | PRR       |                   | Cre02.g094150   |
| CrPRR-2                       | PRR       |                   | Cre16.g676421   |
| CrLUX-1 (ROC75)               | LUX<sup>+</sup> |               | Cre02.g083750   |
| CrLUX-2 (ROC15)               | LUX<sup>+</sup> |               | Cre09.g410450   |
| CrEFL                         | ELF4<sup>+</sup> |               |                |
| *Ostreococcus tauri*          | Data from JGI (http://genome.jgi.doe.gov) |                   |                 |
| OtCCA1                        | RVE       |                   | 4802            |
| OtTOC1                        | PRR       |                   | 24394           |
| *Klebsormium flaccidum*       | Data from Klebsormidium flaccidum genome V1.0 (http://www.plantmorphogenesis.bio.titech.ac.jp/~algae_genome_project/klebsormidium/) |       |                 |
| KfCCA1                        | RVE/CCA1-LHY-clade |               | kfl00255_0190   |
| KFRVE                         | RVE/LCL clade |               | kfl00237_0160   |
| KFTOC1                        | PRR/TOC1 clade |               | kfl00059_0230   |
| KFPRR                         | PRR/PRR clade |               | kfl00165_0030   |
| KfZTL                         | ZTL       |                   | kfl00059_0220   |
| KfEFL3                        | ELF3      |                   | kfl00240_0090   |
| KfEFL-1                       | ELF4      |                   | kfl00065_0150   |
| KfEFL-2                       | ELF4      |                   | kfl00184_0080   |
| KfLUX-1                       | LUX       |                   | kfl00118_0040   |
| KfLUX-2                       | LUX       |                   | kfl00537_0090   |
| *Anthoceros agrestis*         | Data from the oneKP database (www.onekp.com) |                   |                 |
| AaCCA1                        | RVE/CCA1-LHY-clade |               |                 |
| AaRVE                         | RVE/LCL-clade |               |                 |
| AaPRR                         | PRR/PRR clade |               |                 |
| AaEFL3                        | ELF3      |                   |                 |
| AaEFL-1                       | ELF4      |                   |                 |
| AaEFL-2                       | ELF4      |                   |                 |
| AaLUX                         | LUX       |                   |                 |
| AaGI                          | GI        |                   |                 |
| AaZTL                         | ZTL       |                   |                 |
| *Marchantia polymorpha*       | Data from Phytozome v11.0 (https://phytozome.jgi.doe.gov) |                   |                 |
| MpRVE<sup>1</sup>             | RVE/CCA1-LHY-clade |               | Mapoly0042s0058.1 |
| MpTOC1<sup>2</sup>            | PRR/TOC1 clade |               | Mapoly0085s0081.1 |
|     |     |     |
|-----|-----|-----|
|MpPRR$^2$ | PRR/PRR clade | Mapoly0122s0007.1 |
|MpELF3   | ELF3     | Mapoly0014s0139.1 |
|MpEFL    | ELF4     | Mapoly0033s0009.1 |
|MpLUX$^3$ | LUX      | Mapoly0101s0068.1 |
|MpGI     | GI       | Mapoly0019s0145.1 |
|MpFKF    | ZTL      | Mapoly0004s0235.1 |

**Physcomitrella patens**  
Data from Phytozome v11.0 (https://phytozome.jgi.doe.gov)  
PpCCA1a RVE/CCA1-LHY-clade BAI39991.1  
PpCCA1b RVE/CCA1-LHY-clade BAI39992.1  
PpRVE1 RVE/LCL clade Pp1s198_135  
PpRVE2 RVE/LCL clade Pp1s160_6  
PpRVE3 RVE/LCL clade Pp1s46_272  
PpPRR1 PRR/PRR clade XP_001784613  
PpPRR2 PRR/PRR clade XP_001766102  
PpPRR3 PRR/PRR clade XP_001784616  
PpPRR4 PRR/PRR clade XP_001766105  
PpELF3 ELF3 Pp3c1_12790  
PpELF3-3 ELF3 Pp3c11_14750  
PpELF3-2 ELF3 Pp3c7_10610  
PpEFL ELF4 Pp1s180_31  
PpLUX1 LUX Pp3c9_14170  
PpLUX2 LUX Pp3c21_6770  
PpLUX3 LUX Pp3c15_13960  

**Selaginella moellendorffii**  
Data from Phytozome v11.0 (https://phytozome.jgi.doe.gov)  
SmRVE RVE/LCL clade 78045  
SmTOC1 PRR/TOC1 clade 438647  
SmPRR7a PRR/PRR clade XP_002972852†  
SmPRR7b PRR/PRR clade XP_002979868†  
SmELF3 ELF3 415241  
SmELF-1 ELF4 118674  
SmELF-2 ELF4 91083  
SmZTL ZTL 174189  
SmLUX LUX 36646  
SmGI GI 140066  

**Picea abies**  
Data from congenie.org  
PaCCA1 RVE/CCA1-LHY-clade MA_115536g0010  
PaCCA1-2 RVE/CCA1-LHY-clade MA_11267g0020  
PaCCA1-3 RVE/CCA1-LHY-clade MA_102199g0010  
PaPRR1 PRR/TOC1 clade
| Plant          | Data Source                                      | Gene Name          | Description       | Accession   |
|---------------|-------------------------------------------------|--------------------|-------------------|-------------|
| **Oryza sativa** | Data from IRGSP-1.0 (http://rapdb.dna.affrc.go.jp) | PaPRR3             | PRR/PRR clade     | MA_3352g0010 |
|                |                                                  | PaPRR7             | PRR/PRR clade     | MA_303578g0010 |
|                |                                                  | PaLUX              | LUX               | MA_70291g0010 |
|                |                                                  | PaLUX2             | LUX               | MA_303578g0010 |
|                |                                                  | PaZTL              | ZTL               | MA_70291g0010 |
|                |                                                  | PaGI               | GI                |             |
|                |                                                  | PaEFL-1            | ELF4              | MA_57007g0010 |
|                |                                                  | PaEFL-2            | ELF4              | MA_99665g0010 |
|                |                                                  | PaEFL-3            | ELF4              | MA_866497g0010 |
|                |                                                  | PaEFL-4            | ELF4              | MA_4172g0010  |
|                |                                                  | PaEFL-5            | ELF4              | MA_8565294g0010 |
| **Arabidopsis thaliana** | Data from TAIR 10 (www.arabidopsis.org) | OsCCA1             | RVE/CCA1-LHY-clade| Os08t0157600 |
|                |                                                  | OsRVE1             | RVE/CCA1-LHY-clade| Os02t0685200 |
|                |                                                  | OsRVE2             | RVE/CCA1-LHY-clade| Os04g0583900 |
|                |                                                  | OsRVE3             | RVE/CCA1-LHY-clade| Os06t0728700 |
|                |                                                  | OsRVE4             | RVE/LCL clade     | Os02t0680700 |
|                |                                                  | OsRVE5             | RVE/LCL clade     | Os06t0105800 |
|                |                                                  | OsRVE6             | RVE/LCL clade     | Os01t0156000 |
|                |                                                  | OsPRR1             | PRR/TOC1 clade    | Os02g0618200 |
|                |                                                  | OsPRR37            | PRR/PRR clade     | Os07g0695100 |
|                |                                                  | OsPRR73            | PRR/PRR clade     | Os03g0284100 |
|                |                                                  | OsPRR95            | PRR/PRR clade     | Os09g0532400 |
|                |                                                  | OsZTL1             | ZTL               | Os11t0547000 |
|                |                                                  | OsZTL2             | ZTL               | Os06t0694000 |
|                |                                                  | OsFKF1             | ZTL               | Os11t0547000 |
|                |                                                  | OsPCL1             | LUX               | Os01g0971800 |
|                |                                                  | OsELF3-1           | ELF3              | Os06t0142600 |
|                |                                                  | OsELF3-2           | ELF3              | Os01t0566100 |
|                |                                                  | OsEFL-1            | ELF4              | OS11G0621500 |
|                |                                                  | OsEFL-2            | ELF4              | OS03G0410300 |
|                |                                                  | OsEFL-3            | ELF4              | OS08G0366200 |
|                |                                                  | OsGI               | GI                | Os01g0182600 |

---

**Note:**

- **Pa** stands for *Populus trichocarpa*.
- **Os** stands for *Oryza sativa*.
- **At** stands for *Arabidopsis thaliana*.

---

45
| Genes          | Clades                  | Genes          |
|---------------|-------------------------|---------------|
| AtRVE3        | RVE/LCL clade           | At1g01520     |
| AtRVE4        | RVE/LCL clade           | At5g02840     |
| AtRVE5        | RVE/LCL clade           | At4g01280     |
| AtRVE6        | RVE/LCL clade           | At5g52660     |
| AtRVE7        | RVE/CCA1-LHY-clade      | At1g18330     |
| AtRVE7-like   | RVE/CCA1-LHY-clade      | At3g10113     |
| AtRVE8        | RVE/LCL clade           | At3g09600     |
| AtTOC1        | PRR/TOC1 clade          | At5g61380     |
| AtPRR3        | PRR/PRR clade           | Ag5g60100     |
| AtPRR5        | PRR/PRR clade           | Ag5g24470     |
| AtPRR7        | PRR/PRR clade           | Ag5g02810     |
| AtPRR9        | PRR/PRR clade           | Ag2g46790     |
| AtZTL         | ZTL                     | At5g57360     |
| AtFKF1        | ZTL                     | At1g68050     |
| AtLKP2        | ZTL                     | At2g18915     |
| AtLUX         | LUX                     | At3g46640     |
| AtBOA         | LUX                     | At5g59570     |
| AtELF3        | ELF3                    | At2g25930     |
| AtEEC         | ELF3                    | At3g21320     |
| AtELF4        | ELF4                    | At2g40080     |
| AtEFL-1       | ELF4                    | At2g29950     |
| AtEFL-2       | ELF4                    | At1g72630     |
| AtEFL-3       | ELF4                    | At2g06255     |
| AtEFL-4       | ELF4                    | At1g17455     |
| AtGI          | GI                      | At1g22770     |

1 Uncertain homology.
2 The second most similar *M. polymorpha* gene from the BLAST searches, Mapoly0026s0070, was excluded from further analyses as it resembled At5g08520 and related genes with high similarity to *DIVARICATA* in reciprocal BLAST searches (Galego & Almeida, 2002).
3 The third best *M. polymorpha* BLAST hit was Mapoly0101s0006. This protein contains a MYB domain not present in PRR family members and it also lack the C-terminal CCT domain. Reciprocal BLAST searches against the Arabidopsis genome gave RESPONSE REGULATOR1 (RR1) as the best hit, indicating Mapoly0101s0006 belongs to the RR family.
4 As with MpTOC1/PRR the most similar protein to MpLUX in the *M. polymorpha* genome is the RR-family member Mapoly0101s0006.1.
| Primer name     | Sequence - 5' to 3'                  | Comments       |
|-----------------|--------------------------------------|----------------|
| AaCCA1 F        | CGAAGCCATGCCGCGAGAAATT               | qRT-PCR        |
| AaCCA1 R        | CGGTGTTGAGGAATCTATGT                | qRT-PCR        |
| AaELF3 F        | AAGGACGCTATTACTGACACA               | qRT-PCR        |
| AaELF3 R        | GTGAGTACGGTGCTTTTCA                 | qRT-PCR        |
| AaELF4_2011664 F | CCCAGTACCGTGATGATGA                | qRT-PCR        |
| AaELF4_2011664 R | TCGTGGACGGTCTAAAC                | qRT-PCR        |
| AaELF4_2045507 F | TCACTGATACGGGACGTGAA              | qRT-PCR        |
| AaELF4_2045507 R | CCAACACTGAGACATGCTGTT            | qRT-PCR        |
| AaELF4_2048939 F | TGTCCGGTTTGGCATGAGAG            | qRT-PCR        |
| AaELF4_2048939 R | GCCACACATTTGCTTCTTT              | qRT-PCR        |
| AaGI F          | GGAGAAGCTGGTCTGATTG                 | qRT-PCR        |
| AaGI R          | TCACTGTTCTCTCTCTGGT                | qRT-PCR        |
| AaLUX F         | ATGGCTGTCTACATGGGATT               | qRT-PCR        |
| AaLUX R         | CCATAGGGTCCCATGCTCA                 | qRT-PCR        |
| AaPRR F         | ACCGATATCCGAAATCTGGA                | qRT-PCR        |
| AaPRR R         | CGCGCAAAGAACATGAGGTT               | qRT-PCR        |
| AaRVE F         | CGGCCAACAGGTGAAATAC                 | qRT-PCR        |
| AaRVE R         | CGGCCCAGCTGAAATGTA                 | qRT-PCR        |
| AaTUB F         | AGCCTGACATGCTCTGCTGGA              | qRT-PCR        |
| AaTUB R         | CTGGCAGAACAAGAAGCTGGG               | qRT-PCR        |
| AaTTL F         | GGGGGAGAGGGATTAAAGA                 | qRT-PCR        |
| AaTTL R         | GAGTGAAGAGGGAGCCAA                 | qRT-PCR        |
| CPEP23_2        | AGTACCTGAGCAGCCAAGACTGCAACC       | Rev MpRVEpro 1.9kb |
| CPEP24          | cacccAGGTACCAGAGGAGTTGTA^2       | Fwd MpRVEpro 1.9kb |
| CPEP29          | cacccTGTAGCTCTCTGCTGCT^2         | Fwd MpGIpro 5.4kb |
| CPEP30          | AGGCCAATTCTCCCAAAGG                 | Rev MpGIpro 5.4kb |
| CPEP32          | cacccATGTAGCTGGCAGGAGGC^2         | Fwd MpLUXpro 5.5kb |
| CPEP33          | CGCTGCCCATTTCACTTGAGGGTGT          | Rev MpLUXpro 5.5kb |
| CPEP49          | cacccGGGGTTTCTCTTCTGCTG^2         | Fwd MpELF4pro 5.8kb |
| CPEP50          | CGTCCCTGACATGCTGAAATCTCCTACT     | Rev MpELF4pro 5.8kb |
| EFpro_R         | CAACCTTTCTGCAGGGCACATC           | Genotyping^19 |
| gMpPRR_F        | CACAGCTCGACATCTCTTTGCT           | Fwd Compl. test |
| gMpPRR_R        | TGAAGGTAAAACAGTGGCAAAACA          | Rev Compl. test |
| gMpRVE_F        | CACCATTTTTGGGACTGCCCCCTTTTT     | Fwd Compl. test |
| gMpRVE_R        | CGTGGTGGAGGAGGGAGAATT            | Rev Compl. Test |
| HPT_F           | GTTAAGATCATAGCTTACAGGAATCTATGT    | Genotyping^20 |
| ME292           | cacccTGTTTGGATCTCGACGTTGGGAAA^2   | Fwd MpPRRpro 4.0kb |
| ME303           | TAAAGGCATCTATCTGGCGAACGACAA      | Rev MpPRRpro 4.0kb |
| ME367           | CGAAGACGCGAACGAGCTACC             | Fwd MpAPT qRT-PCR |
| ME368           | GTACCCCGGGTTGGCAATAAG             | Rev MpAPT qRT-PCR |
| ME369           | AGGCACTTGATCTACGAGG              | Fwd MpACT qRT-PCR |
| ME370           | ACATGGTCTTTCTCCACGAC               | Rev MpACT qRT-PCR |
| ME428           | cacccGGAAAACCATGGGAGTGA^2        | Fwd MpELF3pro 2.1kb |
| ME429           | GCCACACCTCTACGGCAAA             | Rev MpELF3pro 2.1kb |
| MpELF3 F        | ATCTTTCTCCTCCATCTCTTGCG          | qRT-PCR        |
| MpELF3 R        | ACAGCAGATGATATTTCCGAT            | qRT-PCR        |
| MpELF4 F        | GGGCAGCGAATTCTGCAAGAAA           | qRT-PCR        |
| MpELF4 R        | AGGTCAATATTCCGCCAGAGT            | qRT-PCR        |
| MpGI F          | TTGATGCTGGACCTCTCTAGT            | qRT-PCR        |
| MpGI R          | CATGACCTTTGGAGGAGGACT            | qRT-PCR        |
| MpLUX F         | TGGGAGACGATAGAGAATCTCTG          | qRT-PCR        |
| MpLUX R         | ATCTTCTACAGCTTCTCCTCTT           | qRT-PCR        |
| MpPRR F         | CACAGCTCTCTTGGCAAAACCA           | qRT-PCR        |
| MpPRR R         | GCCGTGAGGAGGAAGGAAAT             | qRT-PCR        |
| Primer          | Sequence                                | Method   | Notes                                      |
|-----------------|-----------------------------------------|----------|--------------------------------------------|
| MpPRR_GT3_F     | AACACTAGTGCGCCGCTGGGAAAGCTCTACCT       | Ko-plasm.|
| MpPRR_GT3_R     | TTATCCCCCTAGCGCCGCTGGGAAGGCAACTGAC     | Ko-plasm.|
| MpPRR_GT5_F     | CTAAAGGTAGCGATAATCAAAACTGACCATCTCAG   | Ko-plasm.|
| MpPRR_GT5_R     | CGGGGCAAGCTTTTAATAAGGCGGACCCAAAGGATT  | Ko-plasm.|
| MpRVE F         | AAACCTCGGCAAAATCAGGAGT                 | qRT-PCR  |
| MpRVE R         | GGCGAGGCAATTTTCAAAGCTG                 | qRT-PCR  |
| MpRVE_GT3_F     | AACACTAGTGCGCCGCTGGGAAAGCTCTACCT     | Ko-plasm.|
| MpRVE_GT3_R     | TTATCCCCCTAGCGCCGCTGGGAAGGCAACTGAC   | Ko-plasm.|
| MpRVE_GT5_F     | CTAAAGGTAGCGATAATCAAAACTGACCATCTCAG | Ko-plasm.|
| MpRVE_GT5_R     | CGGGGCAAGCTTTTAATAAGGCGGACCCAAAGGATT | Ko-plasm.|
| MpTOC1 F        | CGAAGGAAGAACGACTGAAGCA                 | qRT-PCR  |
| MpTOC1 R        | TCTGAGACATTTGACGACGACA                 | qRT-PCR  |
| MpTOC1_GT3_F    | AACACTAGTGCGCCGCTGGGAAAGCTCTACCT     | Ko-plasm.|
| MpTOC1_GT3_R    | TTATCCCCCTAGCGCCGCTGGGAAGGCAACTGAC   | Ko-plasm.|
| MpTOC1_GT5_F    | CTAAAGGTAGCGATAATCAAAACTGACCATCTCAG | Ko-plasm.|
| MpTOC1_GT5_R    | CGGGGCAAGCTTTTAATAAGGCGGACCCAAAGGATT | Ko-plasm.|
| MpTUB3 F        | AGGGGCTGAACTCATAGACTCT                 | qRT-PCR  |
| MpTUB3 R        | TATGAGGAGGTACCCTATGCGCA                | qRT-PCR  |
| MpZTL F         | TGTTCGAGTTCATAACAGGCTACA               | qRT-PCR  |
| MpZTL R         | CTCTGTTAATCTGGGCAACAT                 | qRT-PCR  |
| prrko_F1        | GTTTCCGAAAAGAGCCTCCCA                 | Genotyping |
| prrko_R1        | GATCCTGGAACCCGATTTTT                  | Genotyping |
| prrko_F2        | AAAATGAAGAGCCGCAATTTT                 | Genotyping |
| prrko_R3        | CTCAGCATGCAACCTGAT                    | Genotyping |
| RP_LUC_F1       | AATCCCACTTCTGCTCCACAC                 | qRT-PCR   |
| RP_LUC_R1       | CGGTGCTCCAAAACAACAC                    | qRT-PCR   |
| RP_PRR_F1       | CAGAAGACAGTATGCTCCACAGG               | qRT-PCR   |
| RP_PRR_R1       | CATCTCCCGGCTGCTACATT                  | qRT-PCR   |
| rveko_F1        | TGAAGATATCTTTAGATGCC                   | Genotyping |
| rveko_R1        | TACATCTCAGGGGGTCCG                    | Genotyping |
| rveko_F2        | TGATCAAATGCAGCAAATG                   | Genotyping |
| rveko_R3        | AAGATTGCGCTCTGTTTT                    | Genotyping |
| tocko_F1        | GCTAGAAGCTAGCAAAACAGG                 | Genotyping |
| tocko_R1        | TGTCTAAGCAGAAATGGAGGT                 | Genotyping |
| tocko_F2        | TTTACCCAGTTGAATTTGACTG                | Genotyping |
| tocko_R3        | CCATGAGAAAGCTGGGAGG                   | Genotyping |

**Notes**

Primer in **bold** were used in the Kyoto lab. All other primers were used in the Uppsala lab.

1. Saint-Marcoux et al., 2015
2. “cacc” is added to the 5’ end of the F primer for directional TOPO cloning in pENTR/D-TOPO (Thermofisher)
3. The gene ID for MpTUB2 is Mapoly0158s0010.1
4. MpPRR 3’ homologous arm
5. MpPRR 5’ homologous arm
6. MpRVE 3’ homologous arm
7. MpRVE 5’ homologous arm
8. MpTOC1 3’ homologous arm
9. MpTOC1 5’ homologous arm
10. Illustrated as primer set 1 in S8
11. Illustrated as primer set 2 in S8
12. Illustrated as primer set 3 in S8
13) Illustrated as primer set 1 in S9
14) Illustrated as primer set 2 in S9
15) Illustrated as primer set 3 in S9
16) Illustrated as primer set 1 in S10
17) Illustrated as primer set 2 in S10
18) Illustrated as primer set 3 in S10
19) Illustrated as primer set 2 in S8-10
20) Illustrated as primer set 3 in S8-10
Methods S1 Supplemental materials and methods describing sequence retrieval, sequence analysis and phylogenetic reconstruction.

Inventory of putative bryophyte and charophyte circadian clock genes

Homologs to *Arabidopsis thaliana* circadian clock genes were first identified in the liverwort *Marchantia polymorpha*, in the hornwort *Anthoceros agrestis* and in the charophyte *Klebsormidium flaccidum*. Initial gene family classification was supported by BLAST e-values, reciprocal BLAST searches to the Arabidopsis genome, characterization of conserved protein domains and by BLAST searches against the oneKP database (www.onekp.com).

Sequence retrieval

*M. polymorpha* genes were identified using Arabidopsis amino acid sequences as queries in tBLASTn searches in the publically available *M. polymorpha* genome v. 3.1 (Phytozome 11; https://phytozome.jgi.doe.gov). *A. agrestis* genes were likewise obtained searching the oneKP database (www.onekp.com). Algal gene sequences were retrieved from the publically available *K. flaccidum* genome V1.0 database (http://www.plantmorphogenesis.bio.titech.ac.jp/) and from transcriptome NGS data for the streptophyte algae *Cylindrocystis cushleckae* (Cc; Acc. No. ERR364373) and *Coleochaete irregularis*. (Ci; Acc. No. ERR364367), which were obtained from the NCBI short read archive (SRA). Sequences were assembled using Trinity r20131110 (Haas et al., 2013), tBLASTn searches were performed and gene prediction was done using the web-based FGENESH+ program. Homologs were also retrieved from the following species: *A. thaliana, Oryza sativa, Selaginella moellendorffii, Picea abies* and *P. patens*. Arabidopsis sequences were downloaded from TAIR (www.arabidopsis.org). Genes of other species were identified by BLASTp searches with Arabidopsis queries against databases at JGI (http://www.jgi.doe.gov/), NCBI (http://ncbi.nlm.nih.gov) and plantGDB (http://www.plantgdb.org). To investigate the distribution of identified clock homologs among charophyte algae, hornworts, liverworts and mosses, BLAST searches were performed against the oneKP database (www.onekp.com). All previously identified clock genes as well as genes
identified in this study are listed with accession numbers in Supporting Information Table S1.

**Sequence analysis and phylogenetic reconstruction**

Conserved protein motifs and domains were identified using MEME Suite MAST (http://meme-suite.org; Bailey & Gribskov, 1998; Bailey *et al*., 2009), SMART (Schultz *et al*., 1998; Letunic *et al*., 2015), and NCBI CD-search (CDD v.3.14; Marchler-Bauer & Bryant, 2004; Marchler-Bauer *et al*., 2015). Amino acid sequences were aligned using the M-Coffee algorithm in T-Coffee (Notredame *et al*., 2000; Wallace *et al*., 2006). Alignments were filtered using Transitive Consistency Score (TCS) in the T-Coffee distribution (Chang *et al*., 2014) and are available in Supporting Information Fig. S1. Phylogenetic reconstructions were done using PhyML 3.0 (Guindon *et al*., 2010) and MrBayes 3.2.6 (Huelsenbeck & Ronquist, 2001; Ronquist *et al*., 2012). For MrBayes the substitution model used was decided individually for each alignment using Modelgenerator v.85 (Keane *et al*., 2006). The final phylogenograms were visualized and edited in TreeGraph2 (Stöver & Müller, 2010), where the Bayesian tree was used as reference tree on which both posterior probabilities and bootstrap proportion from PhyML were mapped. Branches resulting in conflicting topologies from the two methods were collapsed. Trees were rooted with charophyte sequences, except for the CCA1/LHY/RVE and PRR families that were midpoint rooted.
References for Supporting information

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* 37: W202–W208.

Bailey TL, Gribskov M. 1998. Methods and statistics for combining motif match scores. *Journal of Computational Biology* 5: 211–221.

Chang J-M, Di Tommaso P, Notredame C. 2014. TCS: a new multiple sequence alignment reliability measure to estimate alignment accuracy and improve phylogenetic tree reconstruction. *Molecular Biology and Evolution* 31: 1625–1637.

Galego L, Almeida J. 2002. Role of *DIVARICATA* in the control of dorsoventral asymmetry in *Antirrhinum* flowers. *Genes & Development* 16: 880–891.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59: 307–321.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8: 1494–1512.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics (Oxford, England)* 17: 754–755.

Keane TM, Creevey CJ, Pentony MM, Naughton TJ, Mclnerney JO. 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evolutionary Biology* 6: 29.

Letunic I, Doerks T, Bork P. 2015. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Research* 43: D257–D260.

Marchler-Bauer A, Bryant SH. 2004. CD-Search: protein domain annotations on the fly. *Nucleic Acids Research* 32: W327–W331.
Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DL et al. 2015. CDD: NCBI’s conserved domain database. *Nucleic Acids Research* 43: D222–D226.

Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* 302: 205–217.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.

Saint-Marcoux D, Proust H, Dolan L, Langdale JA. 2015. Identification of reference genes for real-time quantitative PCR experiments in the liverwort *Marchantia polymorpha*. *PLoS ONE* 10: e0118678.

Schultz J, Milpetz F, Bork P, Ponting CP. 1998. SMART, a simple modular architecture research tool: identification of signaling domains. *Proceedings of the National Academy of Sciences, USA* 95: 5857–5864.

Stöver BC, Müller KF. 2010. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics* 11: 7.

Wallace IM, O’Sullivan O, Higgins DG, Notredame C. 2006. M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Research* 34: 1692–1699.