

SUPPLEMENTARY INFORMATION

The characterization of the circadian clock in the olive fly *Bactrocera oleae* (Diptera: Tephritidae) reveals a *Drosophila*-like organization

Bertolini Enrico^{1,+},
Kistenpfennig Christa^{2,+},
Menegazzi Pamela¹,
Keller Alexander³,
Koukidou Martha²,
Helfrich-Förster Charlotte^{1,*}

¹ Neurobiology and Genetics, Theodor Boveri Institute, Biocentre, University of Würzburg, 97074 Würzburg, Germany

² Oxitec Ltd, 71 Milton Park, Oxford, OX14 4RQ, UK

³ Center for Computation and Theoretical Biology and Department of Bioinformatics, Biocentre, University of Würzburg, 97074 Würzburg, Germany

* charlotte.foerster@biozentrum.uni-wuerzburg.de

+ these authors contributed equally to this work

Supplementary Figure 1

CLUSTAL 2.1 multiple sequence alignment - PER

```

B_ol_per      MEG-ESTESTHNTKVSDSAYSNSCSNSQSQRSGSSKSRLSGSHSSGSSGYGGKPSTQASS 59
B_neohum_per  MEG-ESTESTHNTKVSDSAYSNSCSNSQSQRSGSSKSRLSGSHSSGSSGYGGKPSTQASS 59
D_mel_perA    MEGGESTESTHNTKVSDSAYSNSCSNSQSQRSGSSKSRLSGSHSSGSSGYGGKPSTQASS 60
                ***

B_ol_per      SDMHKIRVKDKTRKKKKLK-----CSAQTNILENQDDVPNTSEP-KSEEQITL 106
B_neohum_per  SDMHKIRVKDKTRKKKKLK-----CSAQTNILENQDDVPNTSEP-QNEEQIAL 106
D_mel_perA    SDMIKRNKDKSRKKKKNKGAGQGAGQAQTLISASTSLEGRDEEKPRPSGTGCVEQQICR 120
                ***   ***   ***   ***   *               **.*.:   .:::  *.*   .   *.:**

B_ol_per      GCD-----KAIAQCSRQKENKETSLLLEKETTE-QTKSEISLQFPTSPPLSTAT 154
B_neohum_per  GCD-----KAIAQCSRQKENKETSLLLEKETTE-QTKSEISLQFPTSPPLSTTT 154
D_mel_perA    ELQDQQHGEDHSEPPQAIEQLQQEEEEEDQSGSESEADRVEGVAKSEAAQSFPIPSPLSVTI 180
                :               : ** *   .:::*.::..   * : . *   : *** : . *   *****.:

B_ol_per      QQG-----GKSERTCESAPGKLESSG-----KAEKLKEESF 185
B_neohum_per  QQG-----GKSERTCESAPGKLESSG-----KAEKLKEESF 185
D_mel_perA    VPPSMGGCGGVGHAAGLDGLAKFDKTWEAGPGKLESMTGVGAAAAGTGQGERVKEDSF 240
                . *   :.*  *.:*****   :.*::*:**

B_ol_per      CCVISMHDGIVLFTTPSITDVLGFPRDMWLGRSFIDFVHPKDRATFASQITTGIPIAESR 245
B_neohum_per  CCVISMHDGIVLFTTPSITDVLGFPRDMWLGRSFIDFVHPKDRATFASQITTGIPIAESR 245
D_mel_perA    CCVISMHDGIVLYTTPSITDVLGYPRDMWLGRSFIDFVHLKDRATFASQITTGIPIAESR 300
                *****.:*****.:*****.:*****.:*****.:*****.:*****
                ‡Val243 Dmel perL

B_ol_per      NSAPKDARSTFCVMLRRYRGLKSGGYGVIGRPVXYEPFRLGLTFREAPPEARPDNNLVS 305
B_neohum_per  NSAPKDARSTFCVMLRRYRGLKSGGYGVIGRPVSYEPFRLGLTFREAPPEARPDNNLVS 305
D_mel_perA    GSVPKDAKSTFCVMLRRYRGLKSGGFVIGRPVSYEPFRLGLTFREAPPEARPDNYMVS 360
                .*.*****.:*****.:*****.:*****.:*****.:*****.:*****.:***

B_ol_per      GTNMLLVICATPIKSSYIVPDEILSHKSPKFSIRHTATGLISHVDSAAVSTLGYPQLDI 365
B_neohum_per  GTNMLLVICATPIKSSYIVPDEILSHKSPKFSIRHTATGLISHVDSAAVSTLGYPQLDI 365
D_mel_perA    GTNMLLVICATPIKSSYKVPDEILSQKSPKFAIRHTATGIISHVDSAAVSALGYLPQLDI 420
                *****.:*****.:*****.:*****.:*****.:*****.:*****

B_ol_per      GRSIFDFYHPEDLMVLKEIYETVMKKGQTAGASFCSKPYRFLIQNGCYILLETEWTSFVN 425
B_neohum_per  GRSIFDFYHPEDLMVLKEIYETVMKKGQTAGASFCSKPYRFLIQNGCYILLETEWTSFVN 425
D_mel_perA    GRSIMDFYHHEDLSVMKETYETVMKKGQTAGASFCSKPYRFLIQNGCYVLLLETEWTSFVN 480
                ****.:****   **   *: **   *****.:*****.:*****.:*****.:*****
                ‡Glu464 per01

B_ol_per      PWSRKLEFVIGHHRVFQGPKHCVNFDSAPANKPKLSDEVSRNMRKDEILKLEESISR 485
B_neohum_per  PWSRKLEFVIGHHRVFQGPKHCVNFDSAPANKPKLSDEVSRNMRKDEILKLEESISR 485
D_mel_perA    PWSRKLEFVVGHRVFQGPQCNVFEAAPTCKLKISEEAQSRNTRIKEDIVKRLAETVSR 540
                *****.:*****.:*****.:***: *   *.*:..:***   *****.:*: *   *:***

B_ol_per      PSDTVKQEVSRRCQALASFMETLMDEVTRTDLKLELPHENELTVSERDSVLMGEISPHHD 545
B_neohum_per  PSDTVKQEVSRRCQALASFMETLMDEVTRTDLKLELPHENELTVSERDSVLMGEISPHHD 545
D_mel_perA    PSDTVKQEVSRRCQALASFMETLMDEVSRADLKLELPHENELTVSERDSVLMGEISPHHD 600
                *****.:*:*****.:*****.:*****.:*****.:*****
                ‡Ser589 perS

B_ol_per      YYDSKSSSTETPPSYNQLNYNENLQRRFFNSKPVTAPEVDPMKNEQSYISADARNTLSPV 605
B_neohum_per  YYDSKSSSTETPPSYNQLNYNENLQRRFFNSKPVTAPEVDPMKNEQSYISADARNTLSPV 605
D_mel_perA    YYDSKSSSTETPPSYNQLNYNENLLRFFNSKPVTAPELDPPKTEPPEPRGTCVSGASGPM 660
                *****.:*****.:**.*   .   .   .   .   .   .:

B_ol_per      Q-CFEGSGSGSGSGNFTSGSHIHMSITNTSNAGTG----- 640
B_neohum_per  Q-CFEGSGSGSGSGNFTSGSHIHMSITNTSNAGTG----- 640
D_mel_perA    SPVHEGSGSGSGSGNFTTASNIIHMSVNTNTSIAGTGGTGTGTGTGTGTGTGTGTGTGTGT 720
                .   .*****.:*.:*****.:***   ****

```

```

B_ol_per      -----TSSGS-----AQLVTLTESLLNKHNDEMEKFM 667
B_neohum_per  -----TSSGS-----AQLVTLTESLLNKHNDEMEKFM 667
D_mel_perA    CTGTGTGTGTGTGTGTGNGTNSGTGTGTASSSKGGSAAIPPVTLTESLLNKHNDEMEKFM 780
               *.*.*:*****

B_ol_per      LKKHREARGRCGEKSKK---ATEKVM EYSGPGHGLKRGGSHSWEGDANKPKHSH TNVMDA 724
B_neohum_per  LKKHREARGRCGEKSKK---ATEKVM EYSGPGHGLKRGGSHSWEGDANKPKHSH TNVMDA 724
D_mel_perA    LKKHRESRGRTEKSKKSANDTLKMLEYSGPGHGIKRGGSHSWEGEANKPKQQL TLGTDA 840
               *****:*** ***** * *:*****:*****:*****:. * **

B_ol_per      QRDYVDHHN-----MATNSTSKAYGTLQTTLGETS-----FA 756
B_neohum_per  QRDYVDHHN-----MATNSTSKAYGTLQTTLGETS-----FA 756
D_mel_perA    IKGAAGSAGGAVGTGGVGGSGGAGVAGGGSGTGVAGTPEGRATTSTGTGTPGGAGGGGGA 900
               :. . . .*. :... * ** * : *

B_ol_per      GSYAAAGLS-----CTRNVLWPPFSVGIN--TSHTSQMVAQSGFTPQHG 799
B_neohum_per  GSYAAAGLS-----CTRNVLWPPFSVGIN--TSHTSQMVAQSGFTPQHS 799
D_mel_perA    GAAAAAGASSSVGSSTPGPSSYPTCTQNINLWPPFSVGITPPVHSTHTAMAQSSFS-SAG 959
               *: *** * **:*****. * :***:*. . .

B_ol_per      IFPTFYIYPATAAAAAVAATQTQVKP---NLADMPSTS-----AQALPLQYMTGVMYP 849
B_neohum_per  IFPTFYIYPATAAAAAAATQAQVKP---NLADMPSTS-----AQALPLQYMTGVMYP 849
D_mel_perA    LFPTFYIYPASLTPTSPTRSPRMHKHPHKG GTDMPTTSQQAAAAAAQAMPLQYMGVMYP 1019
               :*****: :.: : * . :***:** ***:***:*****

B_ol_per      HPSLFYTHP---ATAMMYQPMSFSNMANS LALSEQ--GNSASAFKTNQPVMLAPTPTKT 903
B_neohum_per  HPSLFYTHP---ATVMMYQPMSFSNMANS LALSEQ--GNSASAFKTNQPCMLAPTPTKT 903
D_mel_perA    HPSLFYTHPAAAAATAMMYQPMFPFGMANALQIPERPLGSQSAYNKSVYTTTTPASMTKKV 1079
               ***** **.*****.*.***:* :.*: *...: *: . *. ...

B_ol_per      QGAFHSIT-PAQFQRPSSQATSVKAEPGSNMAPSDSSKKGIADS-PIPSVIGDYVSD--- 958
B_neohum_per  QGAFHSIT-PAQLQRPSSQATSVKAEPGSNMAPSDSSKKGIADS-PIPSVIGDYVSD--- 958
D_mel_perA    PGAFHSVTTTPAQVQRPSSQSASVKTEPGSSAAVSDPCKKEVPDSSPIPSVMGDYNSDPPC 1139
               *****:* **.******:***:*****. * **..** :.* *****:*** **

B_ol_per      -QLNPNDLK-PHTDSNANSDDMDGSSSFSSFYSSFIKTTDGSDSPQENDKDGKHKRKYKVQT 1016
B_neohum_per  -QLNPNDLKQPNTDSNANSDDMDGSSSFSSFYSSFIKTTDGSDSPQENDKDGKHKRKYKVQT 1017
D_mel_perA    SSSNPANNK-KYTDSNGNSDDMDGSSSFSSFYSSFIKTTDGSESPDTEKDPKHKRLKSMS 1198
               . ** : * ****.******:*** :.:** ***** * :

B_ol_per      DSKS--MDNAEEQTLHG DG 1034
B_neohum_per  DSKS--MDNAEEDQTLHG DG 1035
D_mel_perA    TSESKIMEHPEEDQTQHG DG 1218
               *: * :.:* ** ** **

```

NLS: nuclear localization signal
 PAS-A and PAS-B domain: PER/ARNT/SIM
 CLD: cytoplasmic localization domain
 CCID: CKL:CYC inhibition domain - not always specified
 TG repeats

CLUSTAL 2.1 Multiple Sequence Alignments

```

Sequence type explicitly set to Protein
Sequence format is Pearson
Sequence 1: B_ol_per      1034 aa
Sequence 2: B_neohum_per 1035 aa
Sequence 3: D_mel_perA    1218 aa

```

```

Sequences (1:2) Aligned. Score: 98.55
Sequences (1:3) Aligned. Score: 68.96
Sequences (2:3) Aligned. Score: 68.89

```

Percent Identity Matrix - created by Clustal2.1

```

1: B_ol_per      100.00   98.55   71.54
2: B_neohum_per  98.55   100.00   71.64
3: D_mel_perA    71.54   71.64   100.00

```

Supplementary Figure 2

CLUSTAL 2.1 multiple sequence alignment - CYC

```
B_ol_CYC      -----MDEVDDETFDDAKSARTSDENRKQNHSEIEKRRRDKMNTYINELSSMIPMC 51
B_curc_CYC    -----MDEVDDETFDDAKSARTSDENRKQNHSEIEKRRRDKMNTYINELSSMIPMC 51
D_mel_CYC     MEVQEFCEENMEEIEDENYDEEKSAARTSDENRKQNHSEIEKRRRDKMNTYINELSSMIPMC 60
               *: *: *: *: *: *: *****

B_ol_CYC      YVVPKRLDKLTVLKYTVQHLSIRG--SVHPYSGGDYKPSFLSDQELKMLILQASEGFLF 109
B_curc_CYC    YVVPKRLDKLTVLKYTVQHLSIRG--SVHPYSGGDYKPSFLSDQELKMLILQASEGFLF 109
D_mel_CYC     FAMQKRLDKLTVLRMAVQHLSIRGSGSLHPFNGSDYRPSFLSDQELKMIILQASEGFLF 120
               :.: *****: :*****.* ** *: *: :. *: *****: *****

B_ol_CYC      VVDCDRGRILYVSESVSQVLNCSQMDLLGQSWFDILHFKDVAKVKEQLSSLDPCPRDRLI 169
B_curc_CYC    VVDCDRGRILYVSESVSQVLNCSQMDLLGQSWFDILHFKDVAKVKEQLSSLDPSPRDRLI 169
D_mel_CYC     VVGCDRGRILYVSDSVSSVLNSTQADLLGQSWFDVLHFKDIGKVKEQLSSLEQCPRERLI 180
               **.* *****:***.***.:* *****:*****:*****:.***.***

B_ol_CYC      DAKXMLPVKTDIPQSLCRLCPGARRSFFCRMKLKSN-NNQIKEESDTSSSSRSSTKRKSK 228
B_curc_CYC    DAKTMLPVKTDIPQSLCRLCPGARRSFFCRMKLKSN-NNQIKEESDTSSSSRSSTKRKSK 228
D_mel_CYC     DAKTMLPVKTDVPQSLCRLCPGARRSFFCRMKLRTASNQIKEESDTSSSSRSSTKRKSR 240
               *** *****:*****:*****:.: *****:*****:*****:

B_ol_CYC      LSVDHKYRVIQCTGYLKSWTPIKNEEQDSESEDNLTNHSSLVAIGRIPPNVLESNVPPSL 288
B_curc_CYC    LSGDHKYRVIQCTGYLKSWTPIKNEEQDSESEDNLTNHSSLVAIGRIPPNVLESNVPPSL 288
D_mel_CYC     LTTGHKYRVIQCTGYLKSWTPIKDEQDADSDEQTTNLSCLVAIGRIPPNVRNSTVPASL 300
               *: .*****:*****:*****:*****: * *.******:.***.***

B_ol_CYC      DNHPNIRHVLFI SRHSVDGKFLFIDQRATLVIGFLPQEMLGTSFYDYFHHEDIPALAES 348
B_curc_CYC    DNHPNIRHVLFI SRHSVDGKFLFIDQRATLVIGFLPQEMLGTSFYDYFHHDDVPALAES 348
D_mel_CYC     DNHPNIRHVLFI SRHSGEGKFLFIDQRATLVIGFLPQEI LGTSFYEYFHNEDIAALMESH 360
               *****:*****:*****:*****:*****:*****:*****:

B_ol_CYC      K M V I Q V P E K V T T Q V Y R F R C K D N T F I Q L Q S E W R A F K N P W T T D I E Y I I A K N T V F L 401
B_curc_CYC    K M V I Q V P E K V T T Q V Y R F R C K D N T F I Q L Q S E W K A F K N P W T T D I E Y I I A K N T V F L 401
D_mel_CYC     K M V M Q V P E K V T T Q V Y R F R C K D N S Y I Q L Q S E W R A F K N P W T S E I D Y I I A K N S V F L 413
               ***.* *****:*****:*****:*****:*****:*****:***
```

bHLH domain

PAS-A and PAS-B domain

Dmel cycO mutation is a nonsense mutation at K159

CLUSTAL 2.1 Multiple Sequence Alignments

Sequence type explicitly set to Protein

Sequence format is Pearson

Sequence 1: B_ol_CYC 401 aa

Sequence 2: B_curc_CYC 401 aa

Sequence 3: D_mel_CYC 413 aa

Sequences (1:2) Aligned. Score: 98.50

Sequences (1:3) Aligned. Score: 82.54

Sequences (2:3) Aligned. Score: 81.80

Percent Identity Matrix - created by Clustal2.1

1: B_ol_CYC	100.00	98.50	82.54
2: B_curc_CYC	98.50	100.00	81.80
3: D_mel_CYC	82.54	81.80	100.00

Supplementary Figure 3

CLUSTAL 2.1 multiple sequence alignment - CRY

```

B_ol_CRY      MAKR-ANVIWFRHGLRFHDNRALLEAISDKTEGIALIPLFIFDGESAGTKTVGYNRMSFL 59
B_tryi_CRY    MAKR-ANVMWFRHGLRLHDNPALLEAISDKTEGIALIPLFIFDGESAGTKTVGYNRMSFL 59
D_mel_CRY     MATRGANVIWFRHGLRLHDNPALLAALADKDQGIALIPVFIFDGESAGTKNVGYNRMRF 60
               **.* ***:*****:*** ** *.:** :*****:*****:*****.* **

B_ol_CRY      LNSLADIDKQLKAIRGANDIAGKLYLFQGNPTTVFRRNLNEYFRLNKICFEQDCEPIWNR 119
B_tryi_CRY    LNSLAEIDKQLKAIRXASDIAGKLYLFQGNPTTVFRRNLNEYFRLNKICFEQDCEPIWNR 119
D_mel_CRY     LDSLQDIDDQLQAAT---DGRGRLLVFEQEPAYIFRRLHEQVRLHRCIEQDCEPIWNER 117
               *.* :*.**:* * *.* :*:** :*****: ***:**:******.*

B_ol_CRY      DDSVRALCNDLGIEAVEKVSHTLWDPRTVISTNGGIPPLTYQMFLHTVEIIGAPPRVQD 179
B_tryi_CRY    DDSVRALCNDLDIEAVEKVSHTLWDPRTVISTNGRIPPLTXQMFLHTVEIIGAPPRVED 179
D_mel_CRY     DESIRSLCRELNIDFVEKVSHTLWDPQLVIETNGGIPPLTYQMFLHTVQIIGLPPRPAD 177
               *.:**:*.*.*: *****: **.*** ***** *****:*** ***. *

B_ol_CRY      PEWDGVKFLKLTDNMLMELNAFWQFPTPEDFNI FPDNINIVAKVKWLGGEQQALLHLAER 239
B_tryi_CRY    PEWDGVEFLKLTDNMLMELNAFWQFPTPEDFNI FPDNISYVAKVKWRGGEQQALXHLAER 239
D_mel_CRY     ARLEDATFVELDPEFCRSLKLFELPTPEHFNVIYGDNMGFLAKINWRGGETQALLLLDER 237
               .. :.. *.:* :. *.* *.:*****: ***:**:***:* ** * **

B_ol_CRY      LKVEERAFKNGYLPNQANPNILES PKSMSAHLRFGCLSVRRFYWSVHDLFKHVQIEALY 299
B_tryi_CRY    LKVEERAFKNGYLPNQANPNILES PKSMSAHLRFGCLSVRRFYWSVHDLFKHVQIEAFY 299
D_mel_CRY     LKVEQHAFAFERGYLPNQALPNIHDS PKSMSAHLRFGCLSVRRFYWSVHDLFKNVQLRACV 297
               ***:.*:.*.*:***** ***:*****:*****:*****:*****:***:..*

B_ol_CRY      QRIHMAGGEHITGQLIWRXYFYTMSVNNPYDRMEGNAICLNIPWAPPNQELLQSWRSQG 359
B_tryi_CRY    HRIHMAGGEHITGQLIWRXYFYTMSVNNPYDRMEGNEICLNIPWAPPNQELQSWRSQG 359
D_mel_CRY     RGVQMTGGAHITGQLIWRXYFYTMSVNNPNYDRMEGNDICLSIPWAKPNENLLQSWRLGQ 357
               :. :*:** ***** ***** ***** ***** ***.***** *: : ***** **

B_ol_CRY      TGFPLIDAAMRQLLAEGWHLHTLRNTVATFLTRGALWQSWEHGLRHFLKYLLDADNSVCA 419
B_tryi_CRY    TGFPLIDAAMRQLLAEGWHLHTLRNTVATFLTRGALWQSWEHGLRHFLKYLLDADNSVCA 419
D_mel_CRY     TGFPLIDGAMRQLLAEGWHLHTLRNTVATFLTRGGLWQSWEHGLQHFLKYLLDADNSVCA 417
               ***** .*****:*****:*****:*****:*****:*****:*****

B_ol_CRY      GNWMWVSSSAFERLLDSSLVSCPIAFSKRLDPKGEYIRQYVPELAKVPQEYIYEPWRMPQ 479
B_tryi_CRY    GNWMWVSSSAFERLLDSSLVSCPIAFSKRLDPKGEYIRQYVPELAKIPQEYIHEPWRMPQ 479
D_mel_CRY     GNWMWVSSSAFERLLDSSLVTCPVAKRLDPDGTIYIKQYVPELMNVKPEFVHEPWRMSA 477
               *****:*****:***.* ***:***** ***:***** :*:**:******.

B_ol_CRY      DMQENYECVIGVQYPERIVDLAKVSKRNTLAMQALRQSLIAGGAPDEGPPHCRPSNEEEV 539
B_tryi_CRY    EMQENYECVIGVQYPERIVDLAKVSKRNTLAXQTLRQSLIAGGAPDEGPPHCRPSNEEEV 539
D_mel_CRY     EQQEQQYECLIGVHYPERIIDLSMAVKRNMLAMKSLRNSLITP-----PPHCRPSNEEEV 531
               : ***:***:***:*****:***: . *** ** *:**:***:* *****

B_ol_CRY      HQFFWLVD--- 547
B_tryi_CRY    HQFFWLVD--- 547
D_mel_CRY     RQFFWLADV 542
               :*****.*

```

N-terminal DNA photolyase domain classI (4-195)
 C-terminal FAD binding domain (213-516)
 classI conserved C-terminal site (360-372)

aa putatively involved in binding of

MTHF (7)

FAD (13)

CPD (13)

FAD+CPD (1)

Dmel cry^b missense mutation at D410 (FAD binding domain)

CLUSTAL 2.1 Multiple Sequence Alignments

Sequence type explicitly set to Protein
Sequence format is Pearson

Sequence 1: B_ol_CRY 547 aa
Sequence 2: B_try_CRY 547 aa
Sequence 3: D_mel_CRY 542 aa

Sequences (1:2) Aligned. Score: 95.43
Sequences (1:3) Aligned. Score: 72.14
Sequences (2:3) Aligned. Score: 71.40

1: B_ol_CRY	100.00	95.43	73.05
2: B_try_CRY	95.43	100.00	72.30
3: D_mel_CRY	73.05	72.30	100.00

Supplementary Figure 4

CLUSTAL 2.1 multiple sequence alignment - CLK

```
B_ol_CLK      MEDESDDKDDTKRKSRNLSEKKRRDQFNTLVNDLSALISTSNRKMDKSTVLKSTIAFLKH 60
B_curc_CLK    -----KRRRDQFNTLVNDLSALISTSNRKMDKSTVLKSTIAFLKH 40
D_mel_CLK      MDESDDKDDTKRKSRNLSEKKRRDQFNSLVNDLSALISTSSRKMDKSTVLKSTIAFLKN 60
               *****:*****.*****:*****:

B_ol_CLK      HNEATDRSKVFEIQQDWKPTFLTNDFTHLMLESLDGFIIVFSGIGSICYASESITPLLG 120
B_curc_CLK    HNEATDRSKVFEIQQDWKPTFLTNDFTHLMLESLDGFIIVFSGIGSICYASESITPLLG 100
D_mel_CLK      HNEATDRSKVFEIQQDWKPAFLSNDEYTHLMLESLDGFMMVFSSMGSIFYASESITSQLG 120
               *****:*.**:*****:***.*** *****. **

B_ol_CLK      YLPSDLVNMTIFDLTYEMDHESLLNIFLNPKPVIEPLQTDINSSNQITFYHLRGGIDK 180
B_curc_CLK    YLPSDLVNMTIFDLTYEMDHESLLNIFLNPKPVIEPLQTDINSSNQITFYHLRGGIDK 160
D_mel_CLK      YLPQDLNMTIYDLAYEMDHEALLNIFMNPTVIEPRQTDISSNQITFYTHLRGGMEK 180
               ***.*** ***:**:*:*****:*****:*.***** ***.***** *****:.*

B_ol_CLK      VDANAYELVKFVG YFRNDVNLD SIQSQ-----NTLALPRIFQMNPSAEVDKKLIFVG 232
B_curc_CLK    VDANAYELVKFVG YFRNDVNLD SIQSQ-----NTLALPRIFQMNPSAEVDKKLIFVG 212
D_mel_CLK      VDANAYELVKFVG YFRNDTNTSTGSSSEVSNGSNGQPAVLPRIFQONPNAEVDKKLVFVG 240
               *****:*****.* :.*. :.***** **.******:***

B_ol_CLK      TGRIQTPQLIREMSAIDPTCNEFTSKHSMWKFLFLDHRAPPIIGYMPFEVLGTSGYDYY 292
B_curc_CLK    TGRIQTPQLIREMSAIDPTCNEFTSKHSMWKFLFLDHRVPPIIGYMPFEVLGTSGYDYY 272
D_mel_CLK      TGRVQNPQLIREMSIIDPTCNEFTSKHSMWKFLFLDHRAPPIIGYMPFEVLGTSGYDYY 300
               ***:*.***** ***.*****:*****:*****:*****:*****

B_ol_CLK      HFDDLESIVACHEELMQKGECKSCYYRFLTKGQQWIWH----- 330
B_curc_CLK    HFDDLESIVACHEELMQKGECKSCYYRFLTKGQQWIWLQT----- 312
D_mel_CLK      HFDDLESIVACHEELRQTGEGKSCYYRFLTKGQQWIWLQTDYYVSYHQFNSKPDYVCTH 360
               *****:***** *.** *****:*****

B_ol_CLK      -----
B_curc_CLK    -----
D_mel_CLK      KVSYSAEVLKDSRKEGQKSGNSNSITNNGSSKVIASGTSSKSASATTTLRDFELSSQNL 420

B_ol_CLK      -----
B_curc_CLK    -----
D_mel_CLK      DSTLLGNSLASLGTETAATSPAVDSSPMWSASAVQPSGSCQINPLKTSRPASSYGNISST 480

B_ol_CLK      -----
B_curc_CLK    -----
D_mel_CLK      GISPKAKRKCIFYNNRGNDSDSTSMSTDVTSRQSMTHVSSQSQRQRSHHREHHRENHH 540

B_ol_CLK      -----
B_curc_CLK    -----
D_mel_CLK      NQSHHMHMQQQQHQSQQQQHQHQHQQQLQQQLQHTVGTGPKMVPLLPIASTQIMAGNACQFPQ 600

B_ol_CLK      -----
B_curc_CLK    -----
D_mel_CLK      PAYPIASQPLVAPTFLPEPPQYLTAIPMQPVIAPFPVAPVLSPLPVQSQTDMPLPDTVMVTP 660

B_ol_CLK      -----
B_curc_CLK    -----
D_mel_CLK      TQSQLQDQLQRKHDELQKLILQQQNELRIVSEQLLLSRYTYLQPMMSMGFAPGNMTAAAV 720

B_ol_CLK      -----
B_curc_CLK    -----
```

D_mel_CLK GNLGASGQRGLNFTGSNAVQPQFNQYGFALNSEQMLNQDQDQMMMQQQQNLHTQHQHNLQ 780

B_ol_CLK -----
B_curc_CLK -----
D_mel_CLK QQHQSHSQLOQHTQQQHQQNDIL 840

B_ol_CLK -----
B_curc_CLK -----
D_mel_CLK LREDIDDIDAFLNLSPLHSLGSQSTINPFNSSNNNNNQSYNGGSNLNNGNQNNNNNRSSNP 900

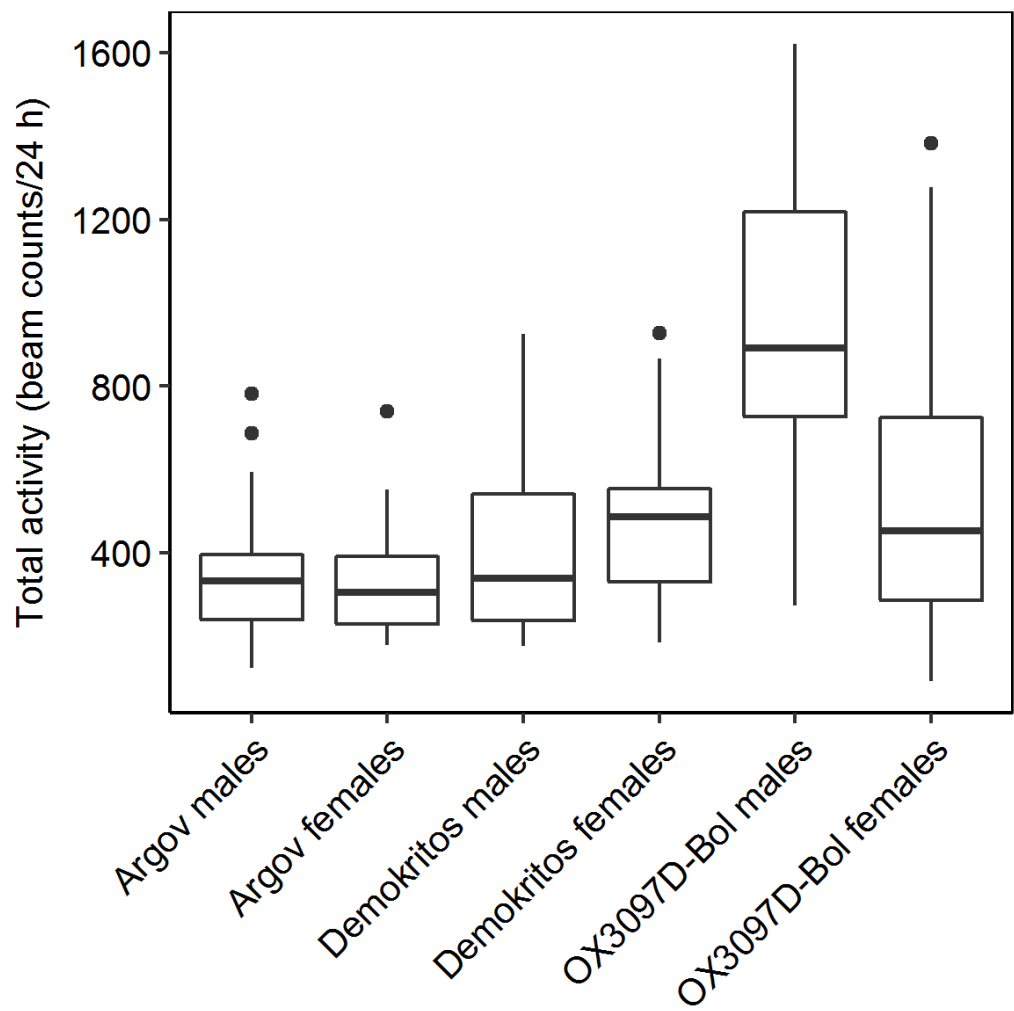
B_ol_CLK -----
B_curc_CLK -----
D_mel_CLK PQNNNEDSLLSYMQMATESSPSINFHMGISDDGSETQSEDNKMMHTSGSNLVQQQQQQQQ 960

B_ol_CLK -----
B_curc_CLK -----
D_mel_CLK QQQILQQHQQQSNSFFSSNPFLNSQNQNQNQLPNDLEILPYQMSQEQSQNLFNSPHTAPG 1020

B_ol_CLK ---
B_curc_CLK ---
D_mel_CLK SSQ 1023

bHLH domain
PAS-A and PAS-B domain
PAS-B C-terminal conserved region (incomplete)
Q-rich stretches and poly-Q (associated with transcriptional activation)

Supplementary Figure 5

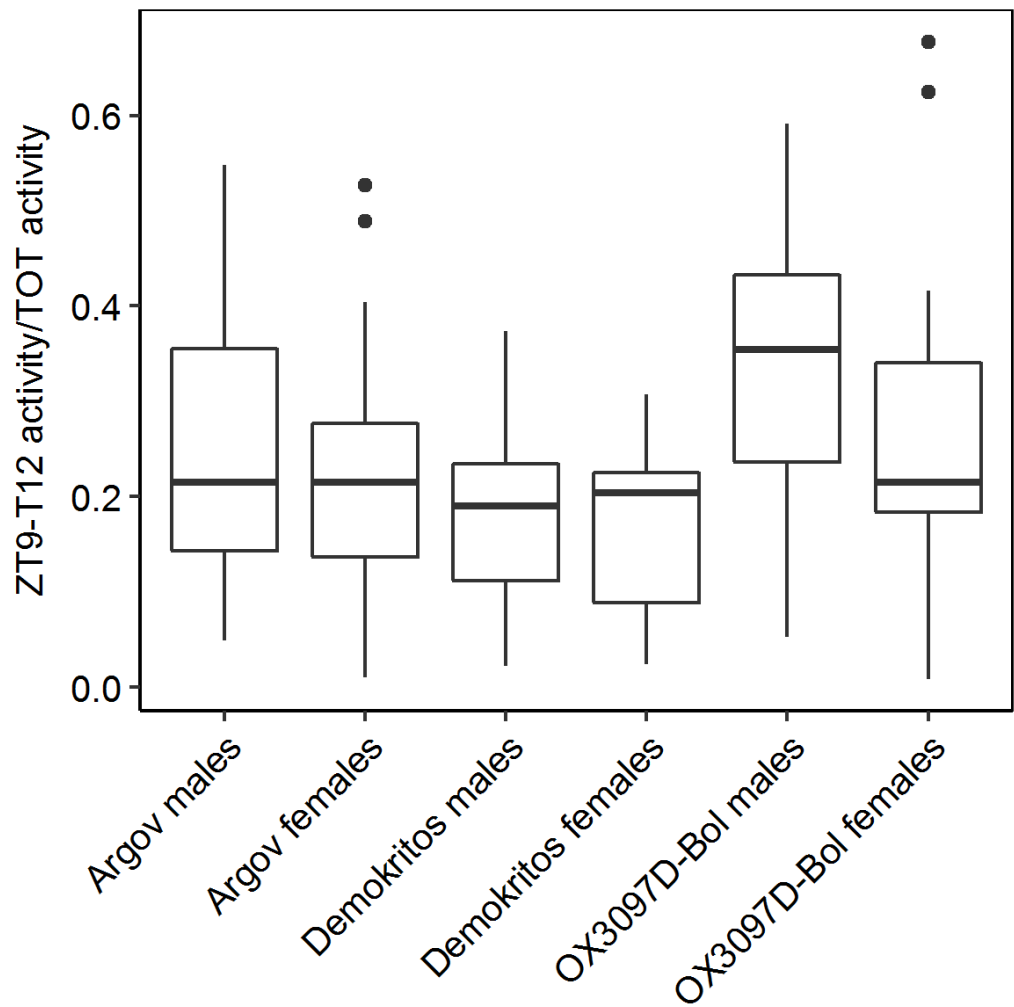


Kruskal-Wallis chi-squared = 56.169, df = 5, p-value = 7.501e-11

	Argov males	Argov females	Demokritos males	Demokritos females	OX3097D-Bol males
Argov females	1.00				
Demokritos males	1.00	1.00			
Demokritos females	0.92	0.38	1.00		
OX3097D-Bol males	0.00	0.00	0.00	0.00	
OX3097D-Bol females	0.07	0.02	0.44	1.00	0.00

t tests with pooled SD (Bonferroni correction)

Supplementary Figure 6

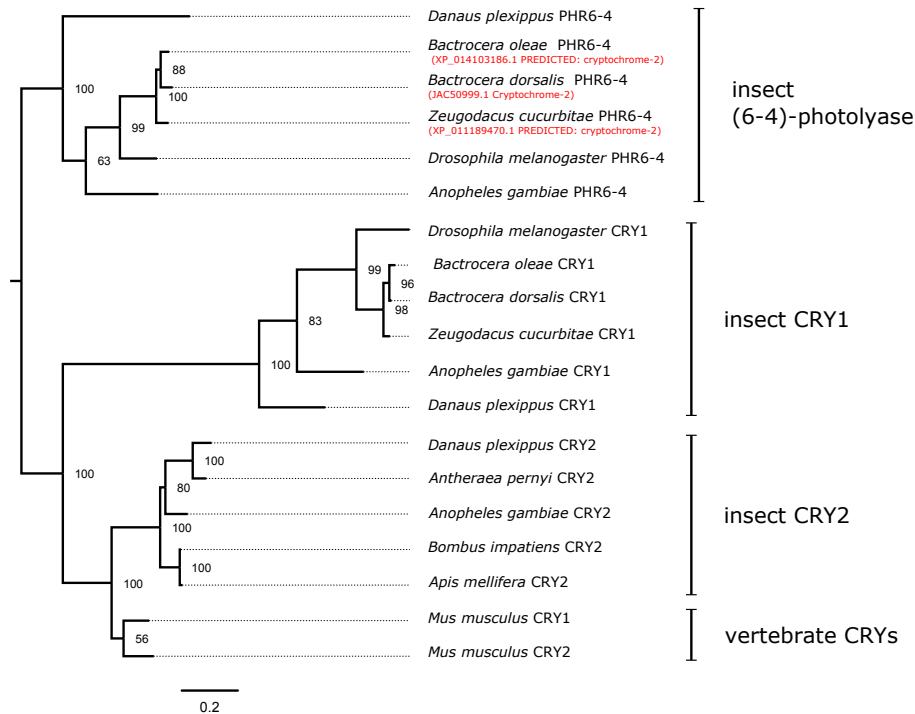


	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	5	0.55	0.11	6.81	0.0000
Residuals	158	2.55	0.02		

	Argov males	Argov females	Demokritos males	Demokritos females	OX3097D-Bol males
Argov females	1.00				
Demokritos males	0.89	1.00			
Demokritos females	0.50	1.00	1.00		
OX3097D-Bol males	0.10	0.01	0.00	0.00	
OX3097D-Bol females	1.00	1.00	0.57	0.31	0.13

t tests with pooled SD (Bonferroni correction)

Supplementary Figure 7



Gene tree of CRY1, CRY2, and insects (6-4)-photolyases reconstructed from amino acids with Maximum Likelihood using RaxML.

The tree shows that sequences annotated in Genbank as "cryptochrome 2" of *B. oleae* (XP_014103186.1), *B. dorsalis* (JAC50999.1), and *Z. cucurbitae* (XP_011189470.1) (highlighted in red) are closely related to insects (6-4)-photolyases. Values at the nodes represent bootstrap values, determined with 100 replicates. GenbankID: *A. gambiae* CRY1 (ABB29886.1), *A. gambiae* CRY2 (ABB29887.1), *A. gambiae* PHR6-4 (EAA10141.3), *A. mellifera* CRY2 (NP_001077099.1), *A. pernyi* CRY2 (ABO38435.1), *B. dorsalis* CRY1 (XP_011206676.1), *B. impatiens* CRY2 (ABO31112.1), *B. oleae* CRY1 (XP_014096643.1), *D. melanogaster* CRY1 (NP_732407.1), *D. melanogaster* PHR6-4 (NP_732407.1), *D. plexippus* CRY1 (AAX58599.1), *D. plexippus* CRY2 (ABA62409.1), *D. plexippus* PHR6-4 (ABO38436.1), *M. musculus* CRY1 (NP_031797.1), *M. musculus* CRY2 (AAD46561.1), *Z. cucurbitae* CRY1 (NP_001291679.1).

Supplementary Table 1

gene	gender	interaction	Df	Sum Sq	Mean Sq	F value	Pr(>F)
<i>per</i>	male	timepoint:strain	12	0.19	0.02	1.19	0.3206
<i>Clk</i>	male	timepoint:strain	10	0.13	0.01	1.47	0.1885
<i>cyc</i>	male	timepoint:strain	11	0.26	0.02	1.33	0.2427
<i>cry</i>	male	timepoint:strain	11	0.20	0.02	1.05	0.4200
<i>per</i>	female	timepoint:strain	9	0.09	0.01	0.65	0.7434
<i>Clk</i>	female	timepoint:strain	8	0.11	0.01	0.80	0.6101
<i>cyc</i>	female	timepoint:strain	13	0.35	0.03	1.24	0.2847
<i>cry</i>	female	timepoint:strain	9	0.20	0.02	0.69	0.7165

2-way ANOVA (interaction timepoint:strain)

Supplementary Table 2

Gene/experiment	Oligonucleotide	Sequence (5' → 3')
<i>period</i>		
identification	per F1	GGARGGCGAATCAACRGAATCTAC
	per R1	TTGTGCGAGAGTATCTCATCTGGCAC
	per F2	GTGARTTATGAACCCTTCCGTTTGG
	per R2	CTGTAAGACSGTTCRTTCTTCATTGGG
	per F3	GCAAAAGTTCGACGGAAACRCC
	per R3	CMCCAGTCATATATTGMAGCGG
	per F4	GCSGACATGCCGAGCACWTC
	per R4	GGGTTCATAACCAYARCGTCGTAC
	per 5'RACE	CTCGCCTGCGTTGATGGTTTGCCG
	per 3'RACE	GGCAACGTGACGATCAGTAAGCGGCA
	per 5'RACE nested F	GCCGCTGCTGCCAGATGAGTGACTION
	per 5'RACE nested R	TTCCCATACAAGTCCACTTTGGCAC
ORF Argov	per F5	GTAAAGCGCCAGCAGTTAGTG
	per R5	ATATGCATGTCGCTGCTGCTC
expression profile	per qPCR F	GCTGCCGTTGCCGCCACACAGACG
	per qPCR R	CCGAAGGCGCCATATTAGAGCCGGG
<i>cryptochrome</i>		
identification	cry F1	TTCAYYGSTGTGCTGTAATATTTGTG
	cry R1	GTKYTTGAAGAGATCATGTACGCTCC
	cry F2	GGCAAATCCMAACATASTYGAATCTCC
	cry R2	CGYTCGGGATACTGCACKCC
	cry 5'RACE	CCAGCAACGCCCCGATTGTCGTGG
	cry 3'RACE	CGAGGGCTGGTTGCACCACACGC
expression profile	cry qPCR F	CATGGCGAAAGTGCTGGTACC
ORF Argov	cry F3	CAGTATGTACCGGAATTGGC
	cry R3	TTTCGGACGATTTGATCACGT
	cry qPCR R	TAGGTCAGTGCGGGATTCCG
<i>Clock</i>		
identification	clk F1	CCNCARYTSATMMGNARATG
	clk R1	YTGSTRCCADATCCAYTGYTG
	clk 5'RACE	CGTTGCATGTCGGATCTATGGCGG
expression profile	clk qPCR F	GGAGGACGAGAGTGACGACAAGG
	clk qPCR R	TCGCTCGGAAGGTAACCCAGCA
<i>cycle</i>		
identification	cyc F1	AATCACAGCGARATCGAGAAGCG
	cyc R1	TTGACGGGYAACATGGTCTTYGC
	cyc 5'RACE2	GCACATCGGTATCATGGACGAGAGC
	cyc 3'RACE	GCTTTCTTCGTTAGATCCGTGCCCCG
expression profile	cyc qPCR F	CTGCAATGGACGAGGTGGACG
	cyc qPCR R	TGCGTCCACGATCACAAATCCA
<i>17S rRNA</i>		
identification	17S rRNA 5'	MKKYGARGARRTBGCIATYATTCCHAC
	17S rRNA 3'	KGAARTCVAGMWDYTTCAACATVTC
	17S rRNA 3'RACE	CGACGTTCTCGCTCTTCTTCTTGACGT
expression profile	17S rRNA qPCR F	CGTCACTCTCAAGTTAGAGGCATC
	17S rRNA qPCR R	TGTCCTGTTCCAAAGCAGATACAG

Supplementary Table 3

	species	GeneBank ID/reference
Cyc	<i>Antheraea pernyi</i>	AAR14937.1
	<i>Bombyx mori</i>	NP_001036982.1
	<i>Danaus plexippus</i>	EHJ64590.1
	<i>Apis mellifera</i>	XP_016770595.1
	<i>Bombus impatiens</i>	XM_012382960.1
	<i>Trachymyrmex zeteki</i>	XM_018454278.1
	<i>Aedes aegypti</i>	JN573265.1
	<i>Anopheles gambiae</i>	Meireles-Filho et al. (2006)
	<i>Anastrepha fraterculus</i>	AQV08515.1
	<i>Zeugodacus cucurbitae</i>	NP_001291680.1
	<i>Drosophila melanogaster</i>	AAC39124.1
	<i>Musca domestica</i>	XM_005180588.3
	<i>Neobellieria bullata</i>	ACJ08742.1
	<i>Mus musculus</i>	AAH25973.1
	<i>Homo sapiens</i>	D89722.1
Per	<i>Apis cerana</i>	NP_001315410.1
	<i>Apis mellifera</i>	ARB43935.1
	<i>Camponotus floridanus</i>	XP_011264661.1
	<i>Mythimna separata</i>	AQY60264.1
	<i>Bombyx mori</i>	NP_001036975.1
	<i>Danaus plexippus</i>	AAO48719.1
	<i>Anopheles sinensis</i>	KFB40661.1
	<i>Aedes albopictus</i>	AEX14535.1
	<i>Culex quinquefasciatus</i>	XP_001849299.1
	<i>Ceratitis capitata</i>	ABB20914.1
	<i>Zeugodacus cucurbitae</i>	NP_001291681.1
	<i>Neobellieria bullata</i>	ACJ08740.1
	<i>Musca domestica</i>	AAD39163.1
	<i>Drosophila melanogaster</i>	NP_001259194.1
	<i>Mus musculus</i>	AAC53592.1
	<i>Homo sapiens</i>	NP_073728.1