





## Figure S1

### Multi-species alignment of H orthologues from Drosophilids

Hairless protein sequences were predicted from genomic sequences. SBD, GBD and CBD are underlined. H orthologues from Drosophilids contain each four introns at identical positions (arrowheads). Presumptive start methionines are shown in cyan. Green boxes: presumptive nuclear localization signals. Blue, identical amino acids; red, highly similar amino acids; yellow, similar amino acids; dots represent gaps. Similarities and identities with respect to the *D. melanogaster* sequence are given in percent. Sequencing gap in *D. ananassae* (D.an) is shown by question marks. D.me: *D. melanogaster*, D.se: *D. sechellia*, D.si: *D. simulans*, D.ya : *D. yakuba*, D.er : *D. erecta*, D.ps: *D. pseudoobscura*, D.pe: *D. persimilis*, D.wi: *D. willistoni*, D.mo: *D. mojavensis*, D.vi: *D. virilis*, D.hy: *D. hydei*, D.gr: *D. grimshawi*.

The BESTFIT program was used for calculation of similarity and identity scores (gap weight 1, length weight 1, maximum penalty length 30). The multi-alignment was performed with PRRN program. Parameters of multi-alignment were: gap extension penalty 1, gap open penalty 4.



1201 1400

D.me SSC.PSPGERSASPE.....RHHNQ.QQPH.....LQ.....R.....SSPLH.Y.MVPPPPVNGNGSAGSPTSAPTSNSSSAAVAAAAAAAYTSPSI...YWP.VI.STMA.FL..HNPL..MHVQT..GASP....LLSPHPQGGSAAAAAA....

A.ga PS....RERSNSNASSSS.....H.KNDL...HPEGSDVVGASLNHY.GTTPSSPAGVIRPNTVIASPTAHPVNAFVPLYQ.AACL.....GAPSSSSLMN...AAVAAATSSVPPHPA.LT.YYQMYTAATMA.AYR...TPI...MH.YGLPGVAPHGPPHMICPHQ.APPSSVAPNISSAD

C.pi PSTDPVKRERSNSNASSTSEVNRSGAHLKELDYAERSAAVDMVDG..VHHHPSSSQASP..VIRPSTVIASPTPHHPVNPFF.MY..PPOI.....AASLL.....AAQSSAVPELPAEPLTSYYQHMYNQAATAATYR...NQV...MH.YQG.....LGPAA.AAAAAAAAV....

A.ae SAVDAARRERSNSNASSTSEIART.AHFKELDYAERSAASDVVDG..LPPTLSTSQSP..VIRPSTVIASPTPHHPVNPFF.MX..PPOI.....AASLL.....AAQSSAVPELPAEPLT.YYQHMYNQAATAATYR...NPI...MH.YQG.....ISPE.AAAAAAAAV....

B.mo SA...AREEW.....FDL.....E...Y.VY.....GYGYVPP.....Y.....GPR.APPPI...MH.YA.....

T.ca .....GPR.APPPI...MH.YA.....

A.me .....QHP.....LY.....GVPMLPPG.....P.YR...APL...MH.YPS.....PVE...YPPMPLY....

N.vi AS...SI.....RQP.....LY.....GLPMPPT.....GPR..APEFC...MH.YSQ.....PVE...YAPMPLY..A...

P.hc FY.....RSP.....LY..PH.....FLQVPA...HSHV.F.GPSSFGASAL..GMP..HH.....HSH...V....

1401 1600

D.me .....EARLS...QSAYHAFAYNGVGAAVAAAAAAAFGQPAFSPHHPH.....L.AHPH...QHPHPAALT.....HHSPAHLATPKLTSSTDQMSATSSHR.....TASTSP.....SSSSASAS.....SSAATSGASSAMHTSSLRNEQ...SSD

A.ga RRLIDRSPTTEP...PPPOFTSGSS...I.TDPA..ESORLCHSILAT...PS..STAAL..NYSVTALSSSSRDIING..GSFTSP...SGSSFTDAVGYQHHPDHSARSTAAVSGASATAKDET...SNG

C.pi ...SPPVAAT.....APFPSSGGS...RDAQ.RDSSRIHSVSTATA...I.SSSSSPSP.ASAA...NYSAAALASSSHHP...PHLHP...GIASHWST...PSSGRRTPPSSAVTSSAAKSDI...GAD

A.ae ...SPPSAFH.....PPSSSSSTNSTGTGNSIHRDAQQRDSNRHSAAVAATGYNHGANSSSSATTSTTTSTHSSNNSSSHIHGSSSSPSP.GAMAL..NYSAAALAASSHHPHHQHQHPHPRPHHSLAAAGPNS.....GAQHDAPTTGTSGTAATKEDH...GAD

B.mo ...VPE.....SEWAPL...T.P...I...LTDH.....KDES...ASD

T.ca .....ESV...DRLSIELS...V.SY...KIDAC...NKT

A.me ...APP.PPH...PSSPEV...HRYK...DY...REQTLTPPSD

N.vi ...APPQPH...HT...PSSPEV...HRYK...DY...REQTLTPPSD

P.hc ...SSSLGGGS.YHGL...VSFYETS...CY...SY...GQPPINEK...REDS...TSD

1601 1650

D.me .....MPLNLSKH~~ sim/id

A.ga ECNEGLRSCKIKKLLTSLHTSLSIKSNASVEFCGVLFIVCSFYVCFMPLNLSKH~~ 66/60

C.pi .....MPLNLSKH~~ 67/61

A.ae .....MPLNLSKH~~ 66/60

B.mo .....MPLNLSKH~~ 68/61

T.ca .....Y...MPLNLSKHAG 73/65

A.me .....MPLNLSKHAG 70/63

N.vi .....MPLNLSKHAG 66/59

P.hc .....MPLNLSKHAG 67/57

CBD

## Figure S2

### Multi-species alignment of H orthologues from different insects

Comparison of *Drosophila melanogaster* (*D.me.*), *Anopheles gambiae* (*A.ga.*), *Culex pipiens* (*C.pi.*), *Aedes aegypti* (*A.ae.*), *Bombyx mori* (*B.mo.*), *Tribolium castaneum* (*T.ca.*), *Apis mellifera* (*A.me.*), *Nasonia vitripennis* (*N.vi.*) and *Pediculus humanus corporis* (*P.hc.*) is shown. Protein sequences were predicted from genomic DNA apart from *D.m.H* and *A.m.H.* that are based on sequenced cDNAs. SBD, GBD and CBD are underlined. Sequence similarities and identities are given in percent with respect to the *D.m.H.* Blue, identical amino acids; red, highly similar amino acids; yellow, similar amino acids; dots represent gaps. For programs and parameters, see Figure S1.