

CLUSTAL 2.1 multiple sequence alignment

```

Dmel_trn-PB      IRETIKGLWGNLSALGRKEREYQKTFCEDEYMSRQHHPHPCSLGIHSTFPNTYAPHH--P 58
Dpse_trn        IREYLKGLWGSALGRKEREYQKTFCEDEYMARHQQHPCSLGIHSTFPNTYAPHH--P 58
Dvir_trn        IREYLKGGVWGNLSALGRKEREYQKTFCEDEYMSRLQHHPCSLGIHSTFPNTYAPHQATA 60
Dgri_trn        IREYLKGLWGNLSALGRKEREYQKTFCEDEYMSRLQHHPCSLGIHSTFPNTYAPHQTTA 60
Dmel_caps-PD    IREMLKG---HSALGRKEREYQKTFSEDEYMSRPP-PGGGG-VHPAAGG-----YP 46
Dpse_caps       IREMLKG---HSALGRKEREYQKTFSEDEYMTTRPP-PGCGG-VHPAAA-----YP 45
Dgri_caps       IRELKFG---HSALGRKEREYQKTFSEDEYMTTRPPAGCVGGVHPGSG-----YP 47
Dvir_caps       IREMLKG---HSALGRKEREYQKTFSEDEYMTTRPP-PGCGG-VHPASG-----YP 45
* * : * *      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
* * * * *      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Dmel_trn-PB      GAT---HHYGMCP-MPVNDLGAII-DPQKFKQLVVP-----TATMIS-EKK 98
Dpse_trn        SAPG---HHYGMCTGMPINDLNAAGDPQKFKQLVVP-----TGLSMN-EKK 102
Dvir_trn        SAP---HHHYGMCP-MPINDLNAV--DGQHKFKQLVVP-----ITATLMH-EKK 102
Dgri_trn        AAAAAHHHYGMCP-MPINDLNSG--DGQHKLQQLVPA-----MSATLLHTEQK 107
Dmel_caps-PD    YIAG-----NSRMIPVTELXLEAPPPQLRGRGG-----GGGASTAS--GA 85
Dpse_caps       CSQSQ---YMGSRPIPVTLELXLEAPPPQMRGRGGIASTTTTTTSSGSGSSGSAAP 101
Dgri_caps       CPSSYNTTYQLGSRPIPVTLELXLEVPPPPQLRGRGG-----APTSH-----GAS 91
Dvir_caps       CPSGINSSQYLGSRSIPVTELXLEAPPPQLRGRGG-----APSSHNGPMASSTSGTP 98
* * * * *      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Dmel_trn-PB      LNNNKALVSQGAIDDSASFVLMKATMGRDQVH-----QNPQ-----136
Dpse_trn        LNNNKALASQAAIDDSASFVLMKATLARDHLQ-----QHPHPHQHPQHPQ--QHQS 155
Dvir_trn        LNNNKSLA--GSVDSASFVLMKATMGRERQQQ-----AQLQQHQQLQQQQL--QQQHP 153
Dgri_trn        LNNNKALA--GSIDDSANFVLMKATMNRDQVH-----QQQQQLQQQQQLPHQQHP 160
Dmel_caps-PD    VQQLQVPSAVDQAS--NSFAQLSHIHYMTNNGQ-----QQAQQQSTSKMHHSQ 133
Dpse_caps       LQQLQVPSVDHAAAASSFAQLSHIHYMTN-----QTIASPMTNPQMHHSQ 148
Dgri_caps       LQQLQVPSAIDA--HAPFAQLSHIHYMTNPLTASSAASAAAATTTATTPRSHHSQ 149
Dvir_caps       LQQLQ-VPSAVDAS-HAPFAQLSHIHYMTNPS-----AATTTATTTTPRSHHSQ 146
* * * * *      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Dmel_trn-PB      ---LNHYTKPQFLSATAVGDSCYS---YADVPMVHGAPLGGP--NQQLRLTQEHFK 186
Dpse_trn        VQSKLNHYTKPQFLSATAAVGDSCYS---YADVPMVHAAPLGAAPLQQLRLTHEHFK 211
Dvir_trn        HQSKLNHYTKPQFLAATAVADSCYS---YADVPMVHAP-----QQLRVTHEHFK 200
Dgri_trn        QQSKLNHYTKPQFLAATAVADSCYS---YADVPLVHAP-----QQLRITHEHFK 207
Dmel_caps-PD    QDMRLLACNGGKTLNA-TSLPRHRP---PVVQESTLSHYSQP-----LANGIRLTQDHFN 184
Dpse_caps       QDMRLLANGGGKTLNA-ASLPRHR---MQESTLSHYSQPLA---LANGIRLTQDHFN 198
Dgri_caps       QDMRLLANG--KALGLNASLPRHMAGRQCGVQESTLSHYSQVPG-----IRLTQDHFN 200
Dvir_caps       QDMRLLANGGGKVLGLNASLPRHMAGRQCGVQESTLSHYSQPLAN--GHAGIRLTQDHFN 204
* * * * *      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Dmel_trn-PB      QR-----ELYDQEMGS-EILDHNYIYSNTHYSMPLE 216
Dpse_trn        HRVAGT-----GEHYDNEVNS-EILDPNYIYSNAHYSMPLE 246
Dvir_trn        QREQR-----PRDFEADINGEMDPNYIYSNAHYSMPLE 236
Dgri_trn        QRE-----RDFDENPLGEMDPNYIYSNAHYSMPLE 239
Dmel_caps-PD    HNQQ-----SHNQHYGG-VYAKPCDAMSEPGYIHNNSHYSPLD 223
Dpse_caps       HNSHGHHGHHGHHGHPHSLGHPHSHTTHGGGVYAKACDAMTEPGYIHNNSHYSPLD 258
Dgri_caps       HN-----GGVYAKPCDAMTEPGYIHNNSHYSPLD 230
Dvir_caps       HNGIGG-----GAGVYAKPCDAMAEAGYIHNNSHYSPLD 239
* * * * *      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Dmel_trn-PB      QLGRSKTPTPPPMPALPLRNLG-----CATTGRRSFQKSAQKQQQNNNTLRQFTHX 270
Dpse_trn        QMGRSKTPTPPPMPALPLRNLG-----CATTGRRSFQKSA-----NNNTLRQFTHX 294
Dvir_trn        QMGRNKTPTPPVPALPLRNLG-----CATTGRRSFQK-TPAHNN-NTSTMRFTHX 288
Dgri_trn        QMGRSKTPTPPPLPALPLRNLG-----CATTGRRLQHRPTPATN---TLRQFTHX 288
Dmel_caps-PD    HDLPP-SPTPTPPPMPALPLRNGVMALIHGNTTGRSFSN-----NNNVSTLSNNNH 275
Dpse_caps       HDLPP-SPTPTPPPMPALPLRNGVMALIHGNTTGRSFSNSNSNS-NNNVATLSNNNH 315
Dgri_caps       HDMPPTPTPPPMPALPLRNLGMALVHGNTTGRSFSNSNSNSNNNNVATLSNNNH 289
Dvir_caps       HDMPPTPTPPPMPALPLRNGMALVHGNTTGRSFSN-----NNNVATLSNNNH 292
* * * * *      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Dmel_trn-PB      -----SSTYRRRQLSIYA-- 283
Dpse_trn        -----STVDYRRRQLSIYA-- 308
Dvir_trn        -----SSEHYRRRQLSIYA-- 302
Dgri_trn        -----SSEHYRRRQLSIFA-- 302
Dmel_caps-PD    ---GIGGGV-VAVGGTVGNNGSLRRYH-- 300
Dpse_caps       ---GGGGGGGGGAILASSNNGSLRRYH-- 341
Dgri_caps       ---GLVLANG--CSG-INNNNGSLRRYH-- 311
Dvir_caps       GVGGGAVLVNGNYNTNGSNNNNGSLRRYH-- 322
* * * * *      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

Figure S2 Conservation of protein sequence beyond stop-codon readthroughs. Clustal alignment (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) of the carboxy end of *tartan* gene proteins from several *Drosophila* species (if annotated with a short read-through extension) and the predicted carboxy end of *caps* gene proteins (if annotated with a long read-through extension) from several *Drosophila* species. The *caps* readthrough (denoted by 'X') is at alignment position 82; the *trn* readthrough is at alignment position 300. Note regions of low complexity, variable conservation and variable length interspersed with regions of protein sequence conservation.