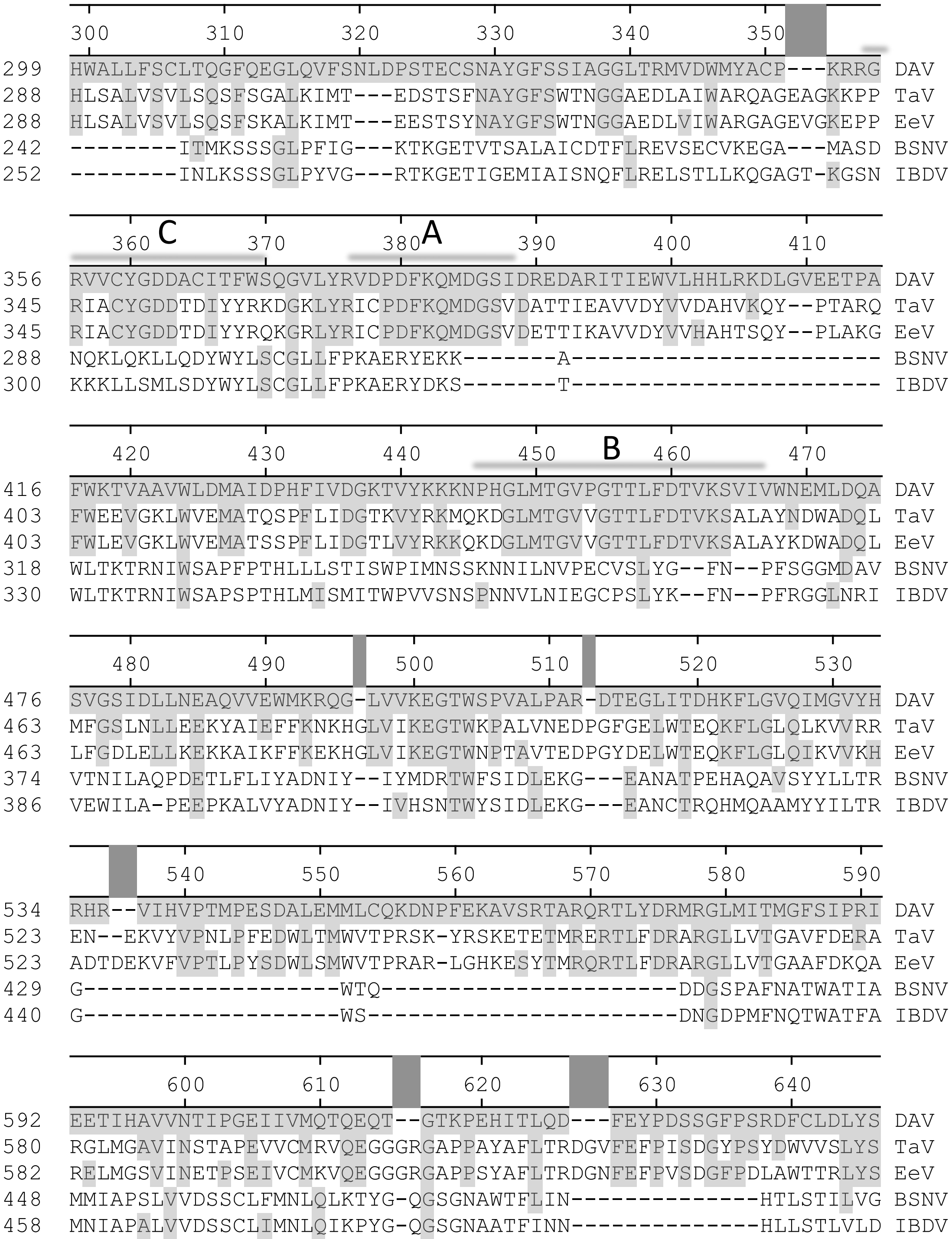

```
         300       310       320       330       340       350
299 HWALLFSCLTQGFQEGLQVFSNLDPSTECSNAYGFSSIAGGLTRMVDWMYACP---KRRG DAV
288 HLSALVSVLSQSFSGALKIMT---EDSTSFNAYGFSWTNGGAEDLAIWARQAGEAGKKPP TaV
288 HLSALVSVLSQSFSKALKIMT---EESTSYNAYGFSWTNGGAEDLVIWARGAGEVGKEPP EeV
242 --------ITMKSSSGLPFIG---KTKGETVTSALAICDTFLREVSECVKEGA---MASD BSNV
252 --------INLKSSSGLPYVG---RTKGETIGEMIAISNQFLRELSTLLKQGAGT-KGSN IBDV

          360 C     370       380A      390       400       410
356 RVVCYGDDACITFWSQGVLYRVDPDFKQMDGSIDREDARITIEWVLHHLRKDLGVEETPA DAV
345 RIACYGDDTDIYYRKDGKLYRICPDFKQMDGSVDATTIEAVVDYVVDAHVKQY--PTARQ TaV
345 RIACYGDDTDIYYRQKGRLYRICPDFKQMDGSVDETTIKAVVDYVVHAHTSQY--PLAKG EeV
288 NQKLQKLLQDYWYLSCGLLFPKAERYEKK-------A---------------------- BSNV
300 KKKLLSMLSDYWYLSCGLLFPKAERYDKS-------T---------------------- IBDV

          420       430       440       450   B   460       470
416 FWKTVAAVWLDMAIDPHFIVDGKTVYKKKNPHGLMTGVPGTTLFDTVKSVIVWNEMLDQA DAV
403 FWEEVGKLWVEMATQSPFLIDGTKVYRKMQKDGLMTGVVGTTLFDTVKSALAYNDWADQL TaV
403 FWLEVGKLWVEMATSSPFLIDGTLVYRKKQKDGLMTGVVGTTLFDTVKSALAYKDWADQL EeV
318 WLTKTRNIWSAPFPTHLLLSTISWPIMNSSKNNILNVPECVSLYG--FN--PFSGGMDAV BSNV
330 WLTKTRNIWSAPSPTHLMISMITWPVVSNSPNNVLNIEGCPSLYK--FN--PFRGGLNRI IBDV

          480       490        500       510        520       530
476 SVGSIDLLNEAQVVEWMKRQG-LVVKEGTWSPVALPAR-DTEGLITDHKFLGVQIMGVYH DAV
463 MFGSLNLLEEKYAIEFFKNKHGLVIKEGTWKPALVNEDPGFGELWTEQKFLGLQLKVVRR TaV
463 LFGDLELLKEKKAIKFFKEKHGLVIKEGTWNPTAVTEDPGYDELWTEQKFLGLQIKVVKH EeV
374 VTNILAQPDETLFLIYADNIY--IYMDRTWFSIDLEKG---EANATPEHAQAVSYYLLTR BSNV
386 VEWILA-PEEPKALVYADNIY--IVHSNTWYSIDLEKG---EANCTRQHMQAAMYYILTR IBDV

                540       550       560       570       580       590
534 RHR--VIHVPTMPESDALEMMLCQKDNPFEKAVSRTARQRTLYDRMRGLMITMGFSIPRI DAV
523 EN--EKVYVPNLPFEDWLTMWVTPRSK-YRSKETETMRERTLFDRARGLLVTGAVFDERA TaV
523 ADTDEKVFVPTLPYSDWLSMWVTPRAR-LGHKESYTMRQRTLFDRARGLLVTGAAFDKQA EeV
429 G-------------------WTQ----------------------DDGSPAFNATWATIA BSNV
440 G-------------------WS-----------------------DNGDPMFNQTWATFA IBDV

                600       610         620          630       640
592 EETIHAVVNTIPGEIIVMQTQEQT--GTKPEHITLQD---FEYPDSSGFPSRDFCLDLYS DAV
580 RGLMGAVINSTAPEVVCMRVQEGGGRGAPPAYAFLTRDGVFEFPISDGYPSYDWVVSLYS TaV
582 RELMGSVINETPSEIVCMKVQEGGGRGAPPSYAFLTRDGNFEFPVSDGFPDLAWTTRLYS EeV
448 MMIAPSLVVDSSCLFMNLQLKTYG-QGSGNAWTFLIN--------------HTLSTILVG BSNV
458 MNIAPALVVDSSCLIMNLQIKPYG-QGSGNAATFINN--------------HLLSTLVLD IBDV
```

Supplementary Fig. S1. Alignment of RdRP showing C–A–B arrangement of RdRP motifs. Virus RdRP regions as shown for Drosophila A virus (DAV), Thosea asigna virus (TaV), Euprosterna elaeasa virus (EeV), blotched snakehead virus (BSNV) and infectious bursal disease virus (IBDV). The regions identified as the C, A and B motifs are indicated. Amino acids matching the DAV sequence are boxed in grey and dashes represent gaps introduced to improve alignment. Alignment was performed using standard parameters for CLUSTAL W in the DNASTAR software package.

Ambrose, R. L., Lander, G. C., Maaty, W. S., Bothner, B., Johnson, J. E. & Johnson, K. N. (2009). Drosophila A virus is an unusual RNA virus with a *T*=3 icosahedral core and permuted RNA-dependent RNA polymerase. *J Gen Virol* **90**, 2191–2200.

```
       1        10        20        30        40        50        60        70        80        90        100       110       120
DaV    MATIVTTTKRISNNSGNNFSTRNRRTPKVTVQSVSSRQIVGKNCPMVQGPRRRGRNRPRKSGSQVAYSAMPKTTLSGVDIIGVIEVQPQRAVGASLFAQAICPTAMSDSRLQQQSKLYSR
RYMV   ............TNSNQGQQGKRKSRRPRGRSAEPQLQRA----PVAQASRISGTVGPLSSNT---------WPLHSVEFLADFKRS---STSADATTYDCVPFNLP--RVWSLARCYSM
CoMV   ......................................................................DWFDTGMITYLGGFQRTAGTTD...SQVFIVSPAALD..RVGTIAKAYAL
CarMV  ......AVKLVTRGWASLSTNQKRRAEMLAGYTPAILAFTPRRPRMTNPPPRTSRNSPGQAGKSM........TMSKTELLS..TVKGTTGVIPSFEDWVVSPRNVA.PQLSLLATNFNK
TBSV   ........GLGKKALNKVRNRRKQGNQQIITHVGGVGGSI--MAPVAVSRQLVGS-KPKFTGR----TSGGVTVTSHREYLTQVGFVVNGGIVGNSL-QLNPSNGTLFSWLPALASNFDQ
SBMV   ........AQAIQNTLPNPPRRKRRAKRRAAQVPKPTQAGVSMAPIAQGTMVKLR.PPMLR......SSMDVTILSHCELSTELAV....TVTIVVTSELVMPFTVG.TWLRGVAQNWSK
TNV    EANGIFVEQRWNGKRWVTEASWRAGQAPNKERAATVTKSIARKVRGGVSRAGGFVTAPVIGAMVTRPTVPRFGMRGNSTVVSNSELILNLTPIALAYTVQSLPLIAT.AWLGTIADNYSK

                130       140       150       160       170       180       190       200       210       220       230       240
DaV    WRPKKLRVLVVGSGSANTFGSLAIGWSPDPANNISGNDTQNLNRVMACRPSSMNRMNQTTVLNIPVATTRKWYLTKGNPDDANHGFINVVVASTTGGYTGSTTFTVTLEWVIEFEGAELG
RYMV   WKPTRWDVVYLPEVSATVAGSIEMCFLYDYADTIPRYTGKMSRGAEGCHLLSGGSARNAVVASMDCSRVGWKRVTSSIPSSVDPNVVNLAVRSSIKPTVSDTPGKLYVIASMVL......
CoMV   WRPKHWEIVYLPRCSTQTDGSIEMGFLLDYADSVPTNTRTMASDGSSLLHTSMKSMGNAVTSALPDEFSNKWFKLSWSTPEESENPARFVVRSDFPVVTADQPGHLWLRSRILLPSTNL.
CarMV  YRITALTVKYSPACSFETNGRVALGFNDDASDTPPTTKVG....FYDLGKHVETAAQTAKDLVIPVDGKTRFIRDSASDDAKLVDF GRIVLSTYGFDKADTVVGELFIQYTIV......
TBSV   YSFNSVVLDYVPLCGTTEVGRVALYFDKDSQDPEPADRVE----LANFGVLKETAPWAEAMLRIPTDKV-KRYCNDSATVDQKLIDLGQLGIATYGGAGADAVGELFLARSVTLYFPQPT
SBMV   YAWVAIRYTYLPSCPTTTSGAIHMGFQYDMADTLPVSVNQLSNLKGYVTGPVWEGQSLCFVNNTKCPDTSRAITIALDTNEVSEKRYPLVTAMEGGSSKTAVNTRLYASYTIRLIEPIAA
TNV    WRWVSLRIIYSPKCPTTTSGTVAMCLSYDRNDVAPGSRVQLSQTYKAINFPPYAGYDGAAILNTDVTPTSAIYVDVDVTRFDKAWY.TVHIGSDGGPAVAVPPGDIFFKYVIELIEPINP

       250       260       270       280       290       300       310       320       330       340       350     360       370       380
DAV    AVSIDEDITPDTGFQDLFTTSDGSWNSEYLTMKMHHGGQMVPFSAAHQAYVYTNAPGTNVTYYLEDKTTLKTANFFAVIQGYSIPGFVMFASYTDAKDYIRTGDTGKCLKYYTQGPVISPPIPKFRVADEAVQTVETNL
RYMV   ...........................................................................................................................................
CoMV   ...........................................................................................................................................
CarMV  ................LSDPTKTAKISQASNDKVSDGPTYVVPSVNGNELQLRVAAGK...WCIIVRGTVEGGFTKPTLIGPGISGDVDYES...............................................
TBSV   NTLLSSKRLDLTG.................SLADATGPGYLVLTRTPTVLTHTRATGTFNLSGGLRCLTSLT...................................................................
SBMV   ALNL.......................................................................................................................................
TNV    TMNV.......................................................................................................................................
```

Supplementary Fig. S2. Alignment of deduced capsid protein sequence of DAV with those of other RNA viruses; rice yellow mottle virus (RYMV), cocksfoot mottle virus (CoMV), carnation mottle virus (CarMV), tomato bushy stunt virus (TBSV), southern bean mosaic virus (SBMV) and tobacco necrosis virus (TNV). The alignment is based on structural motifs that are conserved between the viruses. The sequence of the DAV capsid is shown in the top line. For all other sequences, amino acids shown in red represent α-helices, blue represent β-strands and black represents an insertion. Each grey dash represents a gap inserted into the sequence for better alignment, and dots represents residues that are ignored in the alignment due to low similarity.

Ambrose, R. L., Lander, G. C., Maaty, W. S., Bothner, B., Johnson, J. E. & Johnson, K. N. (2009). Drosophila A virus is an unusual RNA virus with a *T*=3 icosahedral core and permuted RNA-dependent RNA polymerase. *J Gen Virol* **90**, 2191–2200.