Supplementary information for

Mycovirus hunting revealed the presence of diverse viruses in a single isolate of the phytopathogenic fungus *Diplodia seriata* from Pakistan

Haris Ahmed Khan[1], Paul Telengech[2], Hideki Kondo[2], Muhammad Faraz Bhatti[1]*, Nobuhiro Suzuki[2]*

1. Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), H-12, 44000, Islamabad, Pakistan.
2. Institute of Plant Science and Resources, Okayama University, Kurashiki, 710-0046, Japan.

Running title: **Virome of a Pakistani isolate of *Diplodia seriata***

*Corresponding authors
Nobuhiro Suzuki
Institute of Plant Science and Resources, Okayama University, Kurashiki, 710-0046, Japan.
e-mail: nsuzuki@okayama.u-ac.jp
Muhammad Faraz Bhatti
Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), H-12, 44000, Islamabad, PAKISTAN.
e-mail: mfbhatti@asab.nust.edu.pk

Supplementary Figures: 8

Supplementary Table : 1

Supplementary Figures

Supplementary Figure 1. Phylogenetic analysis of isolate L3 (*Diplodia seriata*) utilizing its internal transcribed spacer sequence (ITS) and other isolates within the genera *Diplodia* and *Botryosphaeria*. A maximum-likelihood tree was constructed with a best fit model using 1000 bootstrap value replicates. Values less than 500 were masked and the tree was visualized using an online version of iTOLL (ver. 6.0) (see Materials and Methods). The position of L3 is indicated by the red arrow.

Supplementary Figure 2. Molecular features of Diplodia seriata chrysovirus 1 (DsCV1). (A) A sequence comparison of the terminal untranslated regions (UTRs) of all four genomic segments (dsRNA1 to dsRNA4). The 5'-terminal sequence 5'- UGAUAAAAA---3' was conserved among all three completely sequenced segments. The 5' terminal of dsRNA3 could not be determined *via* RACE. The 3'-terminal sequence 5'----CCUUUAGUAGU-3' was found strictly conserved. (B) Conserved motifs in the RdRP domain of DsCV1 and other chrysoviruses. A total of eight conserved motifs (I-VIII) were observed. Asterisks (*) indicate identical amino acid residues, ":" and "." indicate conserved and semi-conserved amino acid residues. The level of conservation is shown above the aligned sequences as rectangles (silver color). The RdRP conserved motifs are shown in the same manner with the same symbols in the subsequent figures. Abbreviations of viruses utilized in this study are as follows: Diplodia seriata chrysovirus 1 (DsCV1; OM837790) Fusarium oxysporum f. sp. dianthi mycovirus 1 (FodV1; YP 009158913.1), Fusarium graminearum mycovirus China 9 (FgV-ch9; ADU54123.1), Fusarium graminearum dsRNA mycovirus 2 (FgV2; ADW08802.1), Botryosphaeria dothidea chrysovirus 1 (BdCV1; ARE59228.1,Tolypocladium cylindrosporum virus 2 (ToCV2; CBY84993.1), Aspergillus mycovirus 1816 (AsV1816; ABX79996.1), Amasya cherry disease associated chrysovirus (ACDCV; CAG77602.1), Fusarium oxysporum chrysovirus 1 (FoCV1; YP 009665200.1), Cryphonectria nitschkei chrysovirus 1 (CnV1; BBJ21307.1), Penicillium chrysogenum virus (PcV; YP 392482.1), Cryphonectria nitschkei chrysovirus 1 (CnV1; BBJ21307.1).

Supplementary Figure 3. Molecular features of Diplodia seriata polymycovirus 1 (DsPmV1). (A) Terminal sequences of the DsPmV1 genomic segments (dsRNA1 to dsRNA4). The 5'-CGAUUAAAACUU---3' was conserved among all segments. The 3' tetra-nucleotide 5'----GGGG-3' sequence was found conserved. (B) Conserved motifs in RdRP domain of DsPmV1 and other polymycoviruses. The viruses included in the analyses are: Diplodia seriata polymycovirus 1 (DsPmV1; OM837794) Plasmopara viticola lesion associated polymycovirus 1 (PvaPmV1; QHG11067.1), Beauveria bassiana polymycovirus 2 (BbPmV2; CUS18599.1), Magnaporthe oryzae polymycovirus 1 (MoPmV1; QAU09249.1), Aspergillus fumigatus polymycovirus 1 (AfuPmV1; YP 009551547.1), Botryosphaeria dothidea virus 1 (BdRV1; YP 009342446.1), Penicillium janthinellum polymycovirus 1 (PjPmV1; BCJ03666.1), Colletotrichum camelliae filamentous virus 1 (CcFV-1; ASV63092.1).

Supplementary Figure 4. Genomic terminal sequences and RdRP motifs of Diplodia seriata partitivirus 1 (DsPV1). (A) Comparison of the terminal sequences of DsPV1 with Metarhizium brunneum partitivirus 1(MbPV1) (accession no. QHB49873.1). The 5' penta- and 3' deca-nucleotides, 5'-CCCAA---3' and 5'---CCCCUUGGGG-3', were conserved between the two partitiviruses. (B) Conserved motifs in the RdRP domain of DsPV1 and other partitiviruses. The viruses included in the analyses are: Diplodia seriata partitivirus 1(DsPV1; OM837798), Metarhizium brunneum partitivirus 1 (MbPV1; QHB49873.1), Beauveria bassiana partitivirus 3 (BbPV3; QFP40245.1), Colletotrichum eremochloae partitivirus 1 (CePV1; AZT88590.1), Rhizoctonia solani partitivirus virus 5 (RsPV5; AZQ25369.1), Beihai barnacle virus 12 (BbV12; YP 009333370.1).

Supplementary Figure 5. Molecular features of Diplodia seriata victorivirus 1 (DsVV1). (A) Alignment of the terminal sequences of DsVV1 with Magnaporthe oryzae virus 2 (MoV2; LC573906.1) and Sphaeropsis sapinea RNA virus 2 (SsRV2; AF039080.1). No conservation was observed at the 5' terminal region among the three closely related viruses, while the tetra-nucleotide, 5'-GAUA---3' is shared by DsVV1 and MoV2. The 3' penta-nucleotide 5'---AAUGC-3' was conserved among the three viruses. (B)

RdRP conserved motifs of DsPV1 and other victoriviruses. The viruses included in the analyses are: Diplodia seriata victorivirus 1 (DsVV1; OM837800) Neofusicoccum parvum victorivirus 1 (NpVV1; QTE76048.1), Rosellinia necatrix victorivirus 1 (RnVV1; YP 008130308.1), Ustilaginoidea virens RNA virus 1 (UvRV1; YP 007761589.1), Macrophomina phaseolina victorivirus 2 (MpVV2; QKO02072.1), Phomopsis longicolla totivirus 1 (PlTV1; ALD89108.1), Alternaria arborescens victorivirus 1 (AbVV1; YP 009553478.1), Coniothyrium minitans RNA virus (CmRV; YP 392467.1), Penicillium digitatum virus 1 (PdV1; AMY26886.1), Magnaporthe oryzae virus 2 (MoV2; BBG92298.1), Sphaeropsis sapinea RNA virus 2 (SsRV2; NP 047560.1), Macrophomina phaseolina victorivirus 2 (MpVV2; QKI37143.1), Neofusicoccum parvum victorivirus 2 (NpVV2; QTE76050.1).

Supplementary Figure 6. Genomic terminal sequences and RdRP motifs of Diplodia seriata splipalmiviruses 1 (DsSpV1). (A) A comparison of the untranslated regions (UTRs) of the DsSpV1. The genomic segments RNA1 and RNA2 have relatively well-conserved nucleotides at both termini. DsSpV1 RNA3 remained undetermined. (B) Conserved motifs in RdRP domain of DsSpV1, other splipalmiviruses and narna-like viruses. The viruses included in the analyses of motifs G, F, A and B are: Diplodia seriata splipalmivirus 1 (DsSpV1; OM837803) Erysiphe necator associated narnavirus 4 (EnNV4; QJT93743.1), Aspergillus fumigatus narnavirus 1 (AfuNV1; AXE72933.1), Neofusicoccum parvum narnavirus 2 (NpNV2; QDB74995.1), Aspergillus flavus narnavirus 1 (AfNV1; UAW09567.1), Cryphonectria naterciae splipalmivirus 1 (CnSpV1; BCX55509.1), Magnaporthe oryzae narnavirus 1 (MoNV1; QJV68034.1), Oidiodendron maius splipalmivirus 1 (OmSpV1; QNN89179.1), Plasmopara viticola lesion associated narnavirus 20 (PvlaNV20; QIR30299.1), Neofusicoccum parvum narnavirus 3 (NpNV3; QTE76053.1), Aspergillus fumigatus narnavirus 2 (AfuNV2; BCH36622.1). The viruses included in the analyses of motifs C and D are: Diplodia seriata splipalmivirus 1 (DsSpV1; OM837804) Cryphonectria naterciae splipalmivirus 1 (CnSpV1; BCX55510.1), Magnaporthe oryzae narnavirus 1 (MoNV1; BCH36655.1), Oidiodendron maius splipalmivirus 1 (OmSpV1; QNN89180.1), Botrytis cinerea binarnavirus 2 (BcBNV2; QLF49184.1), Aspergillus fumigatus narnavirus 2a (AfuNV2a; BCH36634.1), Aspergillus fumigatus narnavirus 2b (AfuNV2b; BCH36628.1), Aspergillus fumigatus narnavirus 2c (AfuNV2c; BCH36623.1), Aspergillus flavus narnavirus 1 (AfNV1; UAW09566.1), Sclerotinia sclerotiorum narnavirus 2 (SsNV2; QZE12027.1).

Supplementary Figure 7. Molecular features of Diplodia seriata botourmiavirus 1 and 2 (DsBOV1 and 2). A). A comparison of the terminal sequences of the DsBOV2 genomic RNA with that of Neofusicoccum parvum ourmia-like virus 1 (NpBOV1; MK584837.1). The viruses share high levels of nucleotide conservation. (B) Conserved RdRP motifs of DsBOV1 and 2 with other botourmiaviruses. The viruses included in the analyses are: Diplodia seriata botourmiavirus 1 (DsBOV1; OM837801), Diplodia seriata botourmiavirus 2 (DsBOV2; OM837802) Epicoccum nigrum ourmia-like virus 2 (EnOulV2; QDB75005.1), Aspergillus neoniger ourmia-like virus 1 (AnOLV1; AZT88620.1), Pyricularia oryzae ourmia-like virus 1 (PoOLV1; BBF90576.1), Phomopsis longicolla RNA virus 1 (PlRV1; YP 009345044.1), Sclerotinia sclerotiorum ourmia-like virus 1 (SsOLV1; ALD89138.1), Rhizoctonia solani ourmia-like virus 1 (RsOLV1; YP 010084710.1), Magnaporthe oryzae ourmia-like virus (MOLV; SBQ28480.1), Soybean leaf-associated ourmiavirus 2 (SLOV2; ALM62250.1).

Supplementary Figure 8. Molecular features of Diplodia seriata RNA virus 1 (DsRV1). (A) Four trans-membrane helices domains predicted at N-terminus of the ORF1 using the TMHMM - 2.0 web server. (B) Conserved motifs in RdRP domain of DsRV1 other members of the proposed family Ambiguiviridae. Four conserved motifs (IV-VIII) were observed. The viruses included in the analyses are: Diploda seriata RNA virus 1 (DsRV1; OM837806), Phoma matteucciicola RNA virus 1 (PmRV1; QNC69246.1), Periconia macrospinosa ambiguivirus 1 (PmAV1; AZT88666.1), Diaporthe ambigua RNA virus 1 (DaRV1; NP_037581.1), Magnaporthe oryzae RNA virus (MoRV; QVU39989.1), Erysiphe necator associated ambiguivirus 1 (EnAV1; QKN22641.1).

Fig. S1

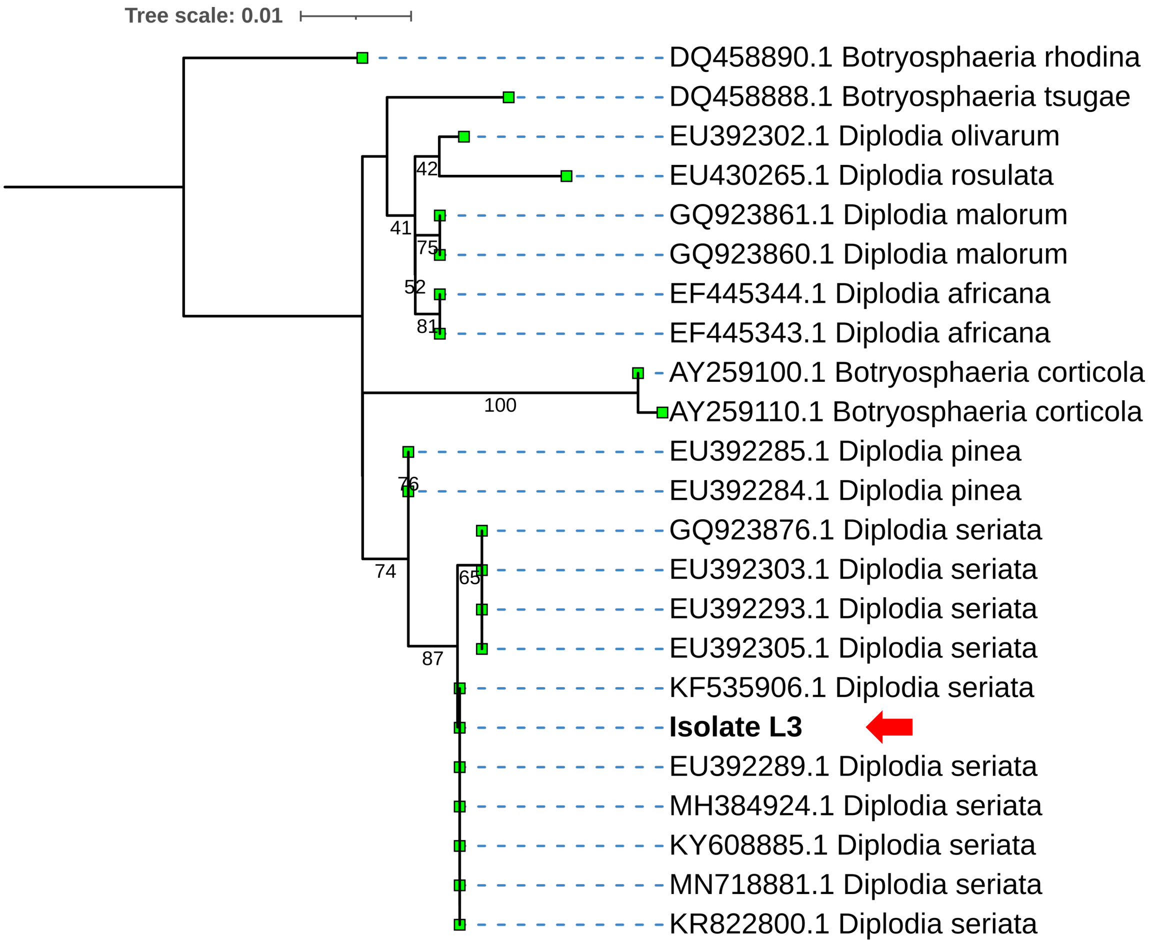
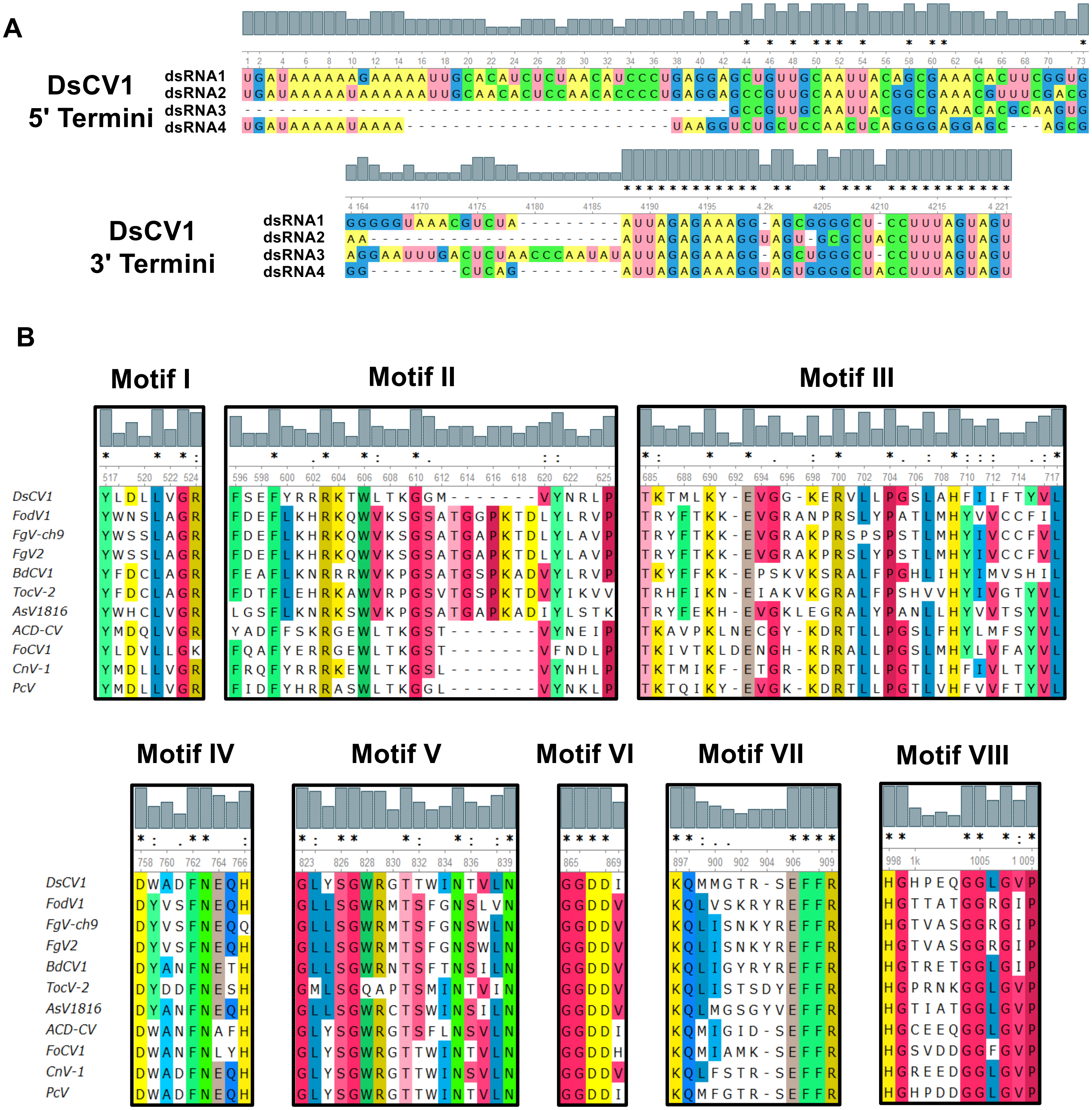

Tree scale: 0.01
DQ458890.1 Botryosphaeria rhodina
DQ458888.1 Botryosphaeria tsugae
EU392302.1 Diplodia olivarum
EU430265.1 Diplodia rosulata
GQ923861.1 Diplodia malorum
GQ923860.1 Diplodia malorum
EF445344.1 Diplodia africana
EF445343.1 Diplodia africana
AY259100.1 Botryosphaeria corticola
AY259110.1 Botryosphaeria corticola
EU392285.1 Diplodia pinea
EU392284.1 Diplodia pinea
GQ923876.1 Diplodia seriata
EU392303.1 Diplodia seriata
EU392293.1 Diplodia seriata
EU392305.1 Diplodia seriata
KF535906.1 Diplodia seriata
Isolate L3
EU392289.1 Diplodia seriata
MH384924.1 Diplodia seriata
KY608885.1 Diplodia seriata
MN718881.1 Diplodia seriata
KR822800.1 Diplodia seriata
42
41
75
52
81
100
76
74
65
87

Fig. S2



	DsCV1 5' Termini (1–73)	DsCV1 3' Termini (4 164–4 221)
dsRNA1	UGAUAAAAAAGAAAAAUUGCACAUCUCUAACAUCCCUGAGGAGCUGUUGCAAUUACAGCGAAACACUUCGGUG	GGGGGUAAACGUCUA---------AUUAGAGAAAGG-AGCGGGGCU-CCUUUAGUAGU
dsRNA2	UGAUAAAAAUAAAAAAUUGCAACACUCCAACACCCCUGAGGAGCCGUUGCAAUUACGGCGAAACGUUUCGACG	AA----------------------AUUAGAGAAAGGUAGU-GCGCUACCUUUAGUAGU
dsRNA3	--GCCGUUGCAAUUACGGCGAAACACGCAAGUG	AGGAAUUUGACUCUAACCCAAUAUAUUAGAGAAAGG-AGCUGGGCU-CCUUUAGUAGU
dsRNA4	UGAUAAAAAUAAAA-----------------------UAAGGUCUGCUCCAACUCAGGGGAGGAGC---AGCG	GG--------CUCAG---------AUUAGAGAAAGGUAGUGGGGCUACCUUUAGUAGU

	Motif I (517–524)	Motif II (596–625)	Motif III (685–717)
DsCV1	YLDLLVGR	FSEFYRRRKTWLTKGGM-------VYNRLP	TKTMLKY-EVGG-KERVLLPGSLAHFIIFTYVL
FodV1	YWNSLAGR	FDEFLKHRKQWVKSGSATGGPKTDLYLRVP	TRYFTKK-EVGRANPRSLYPATLMHYVVCCFIL
FgV-ch9	YWSSLAGR	FDEFLKHRKQWVKSGSATGAPKTDLYLAVP	TRYFTKK-EVGRAKPRSPSPSTLMHYIVCCFVL
FgV2	YWSSLAGR	FDEFLKHRKQWVKSGSATGAPKTDLYLAVP	TRYFTKK-EVGRAKPRSLYPSTLMHYIVCCFVL
BdCV1	YFDCLAGR	FEAFLKNRRRWVKPGSATGSPKADVYLRVP	TKYFFKK-EPSKVKSRALFPGHLIHYIMVSHIL
TocV-2	YFDCLAGR	FDTFLEHRKAWVRPGSVTGSPKTDVYIKVV	TRHFIKN-EIAKVKGRALFPSHVVHYIVGTYVL
AsV1816	YWHCLVGR	LGSFLKNRKSWVKPGSATGAPKADIYLSTK	TRYFEKH-EVGKLEGRALYPANLLHYVVTSYVL
ACD-CV	YMDQLVGR	YADFFSKRGEWLTKGST-------VYNEIP	TKAVPKLNECGY-KDRTLLPGSLFHYLMFSYVL
FoCV1	YLDVLLGK	FQAFYERRGEWLTKGST-------VFNDLP	TKIVTKLDENGH-KRRALLPGSLMHYLVFAYVL
CnV-1	YMDLLVGR	FRQFYRRRKEWLTKGSL-------VYNHLP	TKTMIKF-ETGR-KDRTLLPGTLIHFIVLTYVL
PcV	YMDLLVGR	FIDFYHRRASWLTKGGL-------VYNKLP	TKTQIKY-EVGK-KDRTLLPGTLVHFVVFTYVL

	Motif IV (758–766)	Motif V (823–839)	Motif VI (865–869)	Motif VII (897–909)	Motif VIII (998–1 009)
DsCV1	DWADFNEQH	GLYSGWRGTTWINTVLN	GGDDI	KQMMGTR-SEFFR	HGHPEQGGLGVP
FodV1	DYVSFNEQH	GLLSGWRMTSFGNSLVN	GGDDV	KQLVSKRYREFFR	HGTTATGGRGIP
FgV-ch9	DYVSFNEQQ	GLLSGWRMTSFGNSWLN	GGDDV	KQLISNKYREFFR	HGTVASGGRGIP
FgV2	DYVSFNEQH	GLLSGWRMTSFGNSWLN	GGDDV	KQLISNKYREFFR	HGTVASGGRGIP
BdCV1	DYANFNETH	GLLSGWRNTSFTNSILN	GGDDV	KQLIGYRYREFFR	HGTRETGGLGIP
TocV-2	DYDDFNESH	GMLSGQAPTSMINTVIN	GGDDV	KQLISTSDYEFFR	HGPRNKGGLGVP
AsV1816	DYANFNEQH	GLLSGWRCTSWINSILN	GGDDV	KQLMGSGYVEFFR	HGTIATGGLGVP
ACD-CV	DWANFNAFH	GLYSGWRGTSFLNSVLN	GGDDI	KQMIGID-SEFFR	HGCEEQGGLGVP
FoCV1	DWANFNLYH	GLYSGWRGTTWINTVLN	GGDDH	KQMIAMK-SEFFR	HGSVDDGGFGVP
CnV-1	DWADFNEQH	GLYSGWRGTTWINSVLN	GGDDV	KQLFSTR-SEFFR	HGREEDGGLGVP
PcV	DWADFNEQH	GLYSGWRGTTWINTVLN	GGDDI	KQMFGTR-SEFFR	HGHPDDGGLGVP

Fig. S3

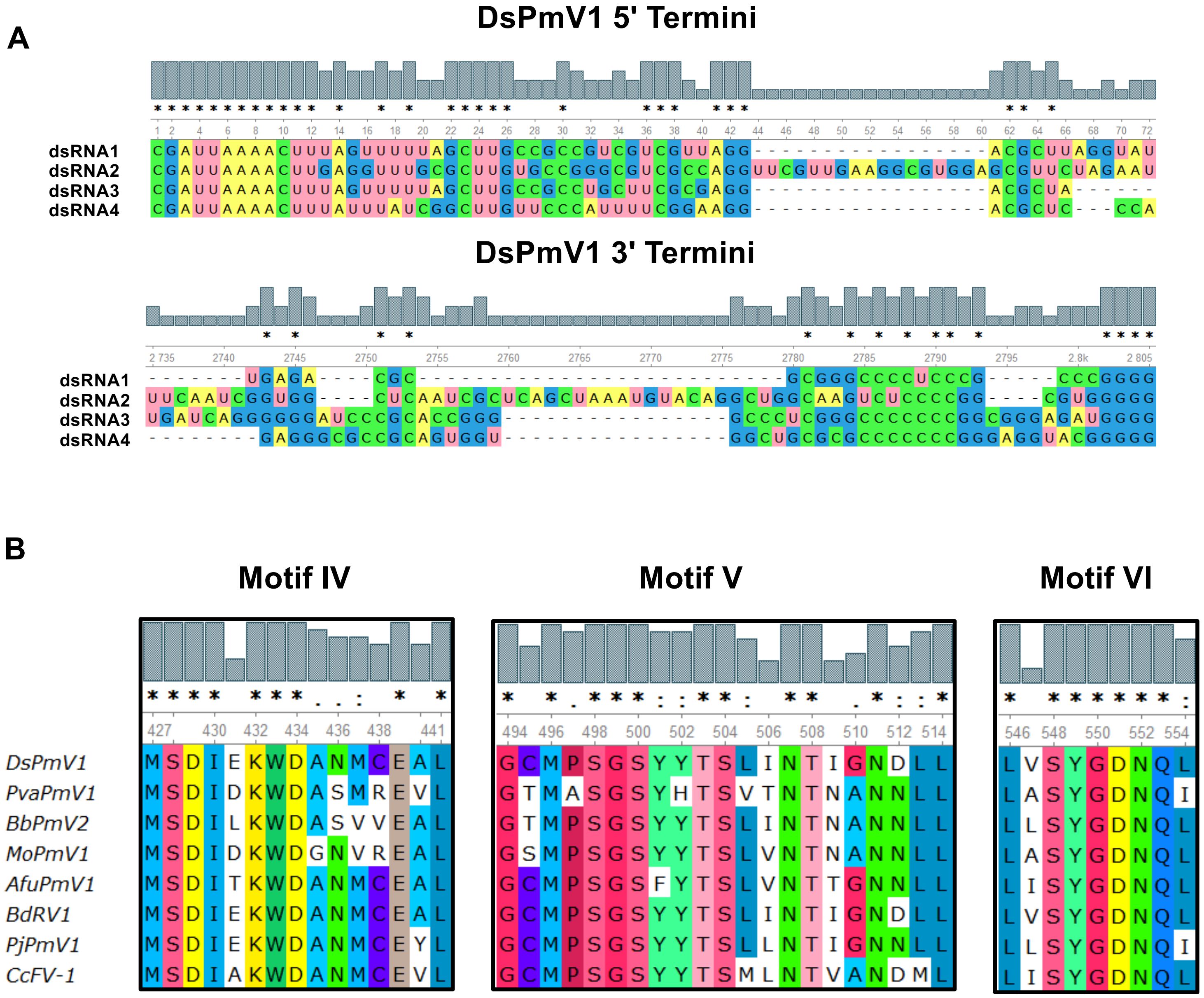
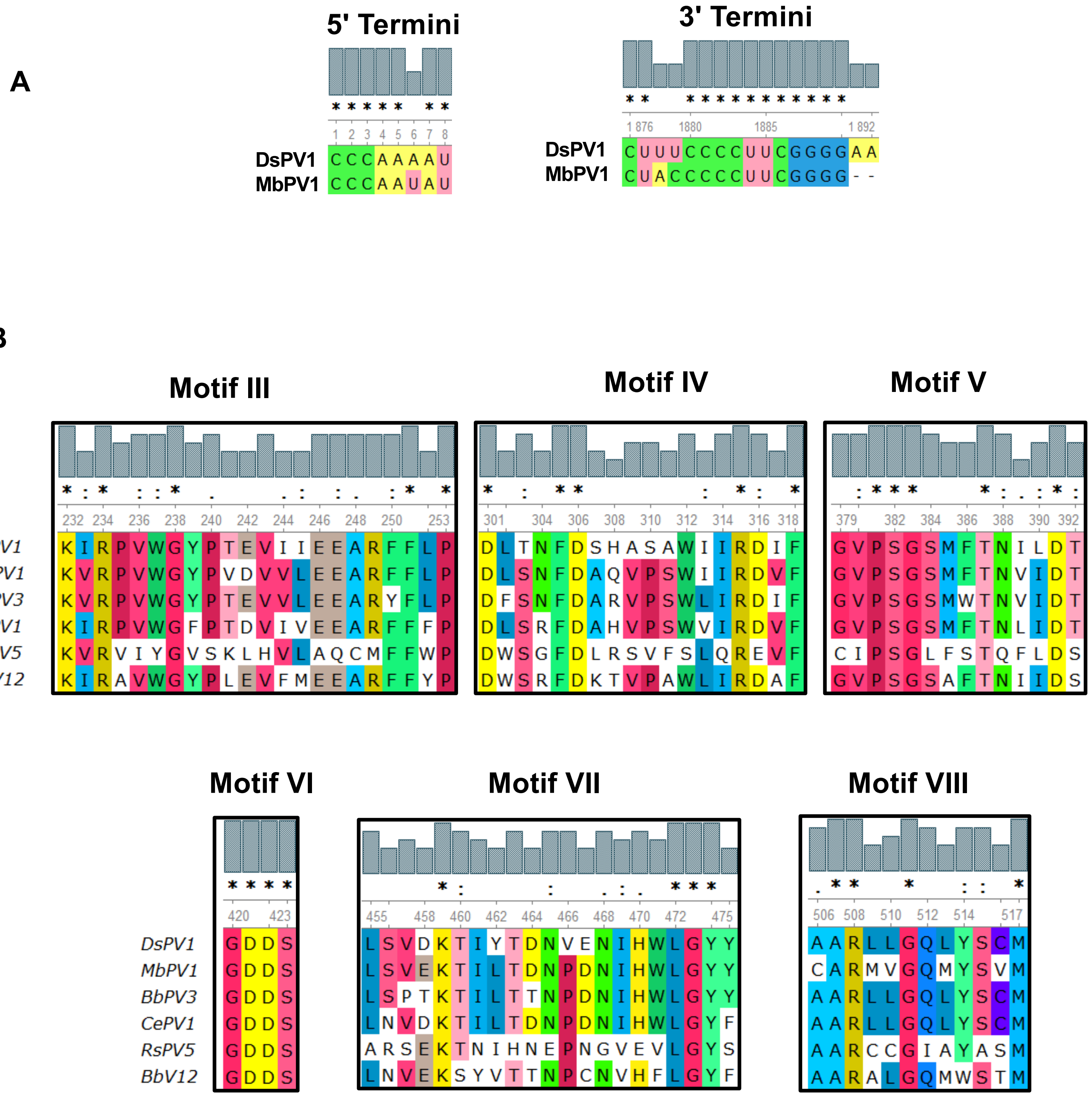

A
DsPmV1 5' Termini
dsRNA1 CGAUUAAAACUUUAGUUUUUAGCUUGCCGCCGUCGUCGUUAGG-----------------ACGCUUAGGUAU
dsRNA2 CGAUUAAAACUUGAGGUUUGCGCUUGUGCCGGGCGUCGCCAGGUUCGUUGAAGGCGUGGAGCGUUCUAGAAU
dsRNA3 CGAUUAAAACUUUAGUUUUUAGCUUGCCGCCUGCUUCGCGAGG-----------------ACGCUA------
dsRNA4 CGAUUAAAACUUUAUUUAUCGGCUUGUUCCCAUUUUCGGAAGG-----------------ACGCUC---CCA
DsPmV1 3' Termini
dsRNA1 -------UGAGA----CGC--------------------------GCGGGCCCCUCCCG-----CCCGGGG
dsRNA2 UUCAAUCGGUGG----CUCAAUCGCUCAGCUAAAUGUACAGGCUGGCAAGUCUCCCCGG----CGUGGGGG
dsRNA3 UGAUCAGGGGGGAUCCCGCACCGGG----------------GCCCUCGGGCCCCCCCGGCGGGAGAUGGGG
dsRNA4 --------GAGGGCGCCGCAGUGGU----------------GGCUGCGCGCCCCCCCGGGAGGUACGGGGG
B
Motif IV
Motif V
Motif VI
DsPmV1 MSDIEKWDANMCEAL GCMPSGSYYTSLINTIGNDLL LVSYGDNQL
PvaPmV1 MSDIDKWDASMREVL GTMASGSYHTSVTNTNANNLL LASYGDNQI
BbPmV2 MSDILKWDASVVEAL GTMPSGSYYTSLINTNANNLL LLSYGDNQL
MoPmV1 MSDIDKWDGNVREAL GSMPSGSYYTSLVNTNANNLL LASYGDNQL
AfuPmV1 MSDITKWDANMCEAL GCMPSGSFYTSLVNTTGNNLL LISYGDNQL
BdRV1 MSDIEKWDANMCEAL GCMPSGSYYTSLINTIGNDLL LVSYGDNQL
PjPmV1 MSDIEKWDANMCEYL GCMPSGSYYTSLLNTIGNNLL LLSYGDNQI
CcFV-1 MSDIAKWDANMCEVL GCMPSGSYYTSMLNTVANDML LISYGDNQL

Fig. S4


A
5' Termini
* * * * *   * *
1 2 3 4 5 6 7 8
DsPV1 CCCAAAAU
MbPV1 CCCAAUAU
3' Termini
* *   * * * * * * * * * * *
1 876 1880 1885 1 892
DsPV1 CUUUCCCCUUCGGGGAA
MbPV1 CUACCCCCUUCGGGG--
B
Motif III
232 234 236 238 240 242 244 246 248 250 253
DsPV1 KIRPVWGYPTEVIIEEARFFLP
MbPV1 KVRPVWGYPVDVVLEEARFFLP
BbPV3 KVRPVWGYPTEVVLEEARYFLP
CePV1 KIRPVWGFPTDVIVEEARFFFP
RsPV5 KVRVIYGVSKLHVLAQCMFFWP
BbV12 KIRAVWGYPLEVFMEEARFFYP
Motif IV
301 304 306 308 310 312 314 316 318
DsPV1 DLTNFDSHASAWIIRDIF
MbPV1 DLSNFDAQVPSWIIRDVF
BbPV3 DFSNFDARVPSWLIRDIF
CePV1 DLSRFDAHVPSWVIRDVF
RsPV5 DWSGFDLRSVFSLQREVF
BbV12 DWSRFDKTVPAWLIRDAF
Motif V
379 382 384 386 388 390 392
DsPV1 GVPSGSMFTNILDT
MbPV1 GVPSGSMFTNVIDT
BbPV3 GVPSGSMWTNVIDT
CePV1 GVPSGSMFTNLIDT
RsPV5 CIPSGLFSTQFLDS
BbV12 GVPSGSAFTNIIDS
Motif VI
420 423
DsPV1 GDDS
MbPV1 GDDS
BbPV3 GDDS
CePV1 GDDS
RsPV5 GDDS
BbV12 GDDS
Motif VII
455 458 460 462 464 466 468 470 472 475
DsPV1 LSVDKTIYTDNVENIHWLGYY
MbPV1 LSVEKTILTDNPDNIHWLGYY
BbPV3 LSPTKTILTTNPDNIHWLGYY
CePV1 LNVDKTILTDNPDNIHWLGYF
RsPV5 ARSEKTNIHNEPNGVEVLGYS
BbV12 LNVEKSYVTTNPCNVHFLGYF
Motif VIII
506 508 510 512 514 517
DsPV1 AARLLGQLYSCM
MbPV1 CARMVGQMYSVM
BbPV3 AARLLGQLYSCM
CePV1 AARLLGQLYSCM
RsPV5 AARCCGIAYASM
BbV12 AARALGQMWSTM

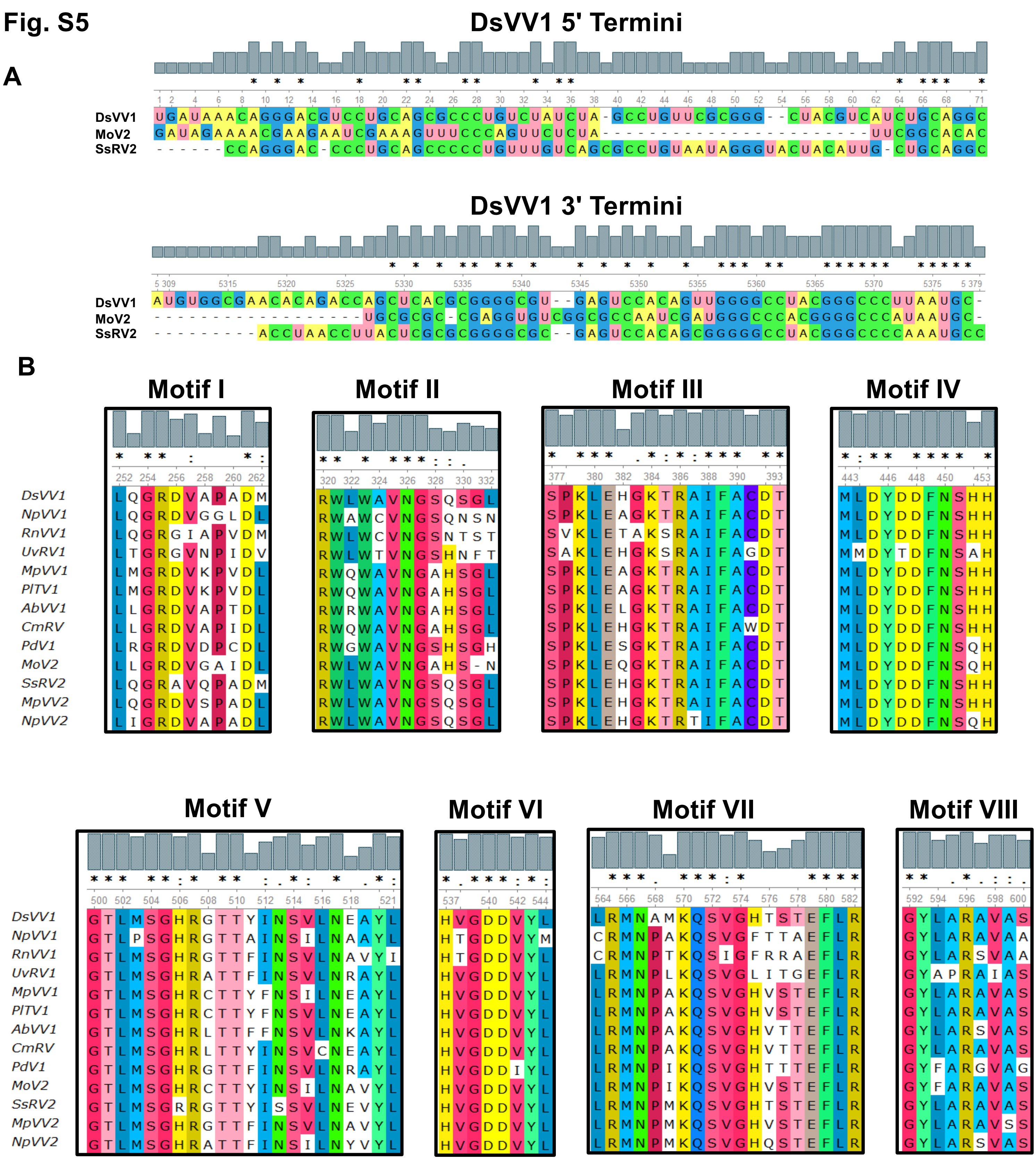


Fig. S5

A

DsVV1 5' Termini

```
           1 ... 71
           *  * *    *   **   **    * **                           * ***  *
DsVV1      UGAUAAACAGGGACGUCCUGCAGCGCCCUGUCUAUCUA-GCCUGUUCGCGGG--CUACGUCAUCUGCAGGC
MoV2       GAUAGAAAACGAAGAAUCGAAAGUUUCCCAGUUCUCUA-----------------------UUCGGCACAC
SsRV2      ------CCAGGGAC-CCCUGCAGCCCCCUGUUUGUCAGCGCCUGUAAUAGGGUACUACAUUG-CUGCAGGC
```

DsVV1 3' Termini

```
           5 309 ... 5 379
                               * * * ** ** *   * * * *  *  *** **   ****** *****
DsVV1      AUGUGGCGAACACAGACCAGCUCACGCGGGGCGU--GAGUCCACAGUUGGGGCCUACGGGCCCUUAAUGC-
MoV2       ------------------UGCGCGC-CGAGGUGUCGGCGCCAAUCGAUGGGCCCACGGGGCCCAUAAUGC-
SsRV2      ---------ACCUAACCUUACUCGCGCGGGGCGC--GAGUCCACAGCGGGGGCCUACGGGCCCCAAAUGCC
```

B

Motif I

```
           252 ... 262
           * ** :   *:
DsVV1      LQGRDVAPADM
NpVV1      LQGRDVGGLDL
RnVV1      LQGRGIAPVDM
UvRV1      LTGRGVNPIDV
MpVV1      LMGRDVKPVDL
PlTV1      LMGRDVKPVDL
AbVV1      LLGRDVAPTDL
CmRV       LLGRDVAPIDL
PdV1       LRGRDVDPCDL
MoV2       LLGRDVGAIDL
SsRV2      LQGRAVQPADM
MpVV2      LQGRDVSPADL
NpVV2      LIGRDVAPADL
```

Motif II

```
           320 ... 332
           ** * ***::.
DsVV1      RWLWAVNGSQSGL
NpVV1      RWAWCVNGSQNSN
RnVV1      RWLWCVNGSNTST
UvRV1      RWLWTVNGSHNFT
MpVV1      RWQWAVNGAHSGL
PlTV1      RWQWAVNGAHSGL
AbVV1      RWRWAVNGAHSGL
CmRV       RWQWAVNGAHSGL
PdV1       RWGWAVNGSHSGH
MoV2       RWLWAVNGAHS-N
SsRV2      RWLWAVNGSQSGL
MpVV2      RWLWAVNGSQSGL
NpVV2      RWLWAVNGSQSGL
```

Motif III

```
           377 ... 393
           * *** .*:*:*** **
DsVV1      SPKLEHGKTRAIFACDT
NpVV1      SPKLEAGKTRAIFACDT
RnVV1      SVKLETAKSRAIFACDT
UvRV1      SAKLEHGKSRAIFAGDT
MpVV1      SPKLEAGKTRAIFACDT
PlTV1      SPKLEAGKTRAIFACDT
AbVV1      SPKLELGKTRAIFACDT
CmRV       SPKLEHGKTRAIFAWDT
PdV1       SPKLESGKTRAIFACDT
MoV2       SPKLEQGKTRAIFACDT
SsRV2      SPKLEHGKTRAIFACDT
MpVV2      SPKLEHGKTRAIFACDT
NpVV2      SPKLEHGKTRTIFACDT
```

Motif IV

```
           443 ... 453
           *:** **** *
DsVV1      MLDYDDFNSHH
NpVV1      MLDYDDFNSHH
RnVV1      MLDYDDFNSHH
UvRV1      MMDYTDFNSAH
MpVV1      MLDYDDFNSHH
PlTV1      MLDYDDFNSHH
AbVV1      MLDYDDFNSHH
CmRV       MLDYDDFNSHH
PdV1       MLDYDDFNSQH
MoV2       MLDYDDFNSQH
SsRV2      MLDYDDFNSHH
MpVV2      MLDYDDFNSHH
NpVV2      MLDYDDFNSQH
```

Motif V

```
           500 ... 521
           *** **:* ** :.*: * .*:
DsVV1      GTLMSGHRGTTYINSVLNEAYL
NpVV1      GTLPSGHRGTTAINSILNAAYL
RnVV1      GTLMSGHRGTTFINSVLNAVYI
UvRV1      GTLMSGHRATTFINSVLNRAYL
MpVV1      GTLMSGHRCTTYFNSILNEAYL
PlTV1      GTLMSGHRCTTYFNSVLNEAYL
AbVV1      GTLMSGHRLTTFFNSVLNKAYL
CmRV       GTLMSGHRLTTYINSVCNEAYL
PdV1       GTLMSGHRGTTFINSVLNRAYL
MoV2       GTLMSGHRCTTYINSILNAVYL
SsRV2      GTLMSGRRGTTYISSVLNEVYL
MpVV2      GTLMSGHRGTTFINSVLNAVYL
NpVV2      GTLMSGHRATTFINSILNYVYL
```

Motif VI

```
           537 ... 544
           *.***:*:
DsVV1      HVGDDVYL
NpVV1      HTGDDVYM
RnVV1      HTGDDVYL
UvRV1      HVGDDVYL
MpVV1      HVGDDVYL
PlTV1      HVGDDVYL
AbVV1      HVGDDVYL
CmRV       HVGDDVYL
PdV1       HVGDDIYL
MoV2       HVGDDVYL
SsRV2      HVGDDVYL
MpVV2      HVGDDVYL
NpVV2      HVGDDVYL
```

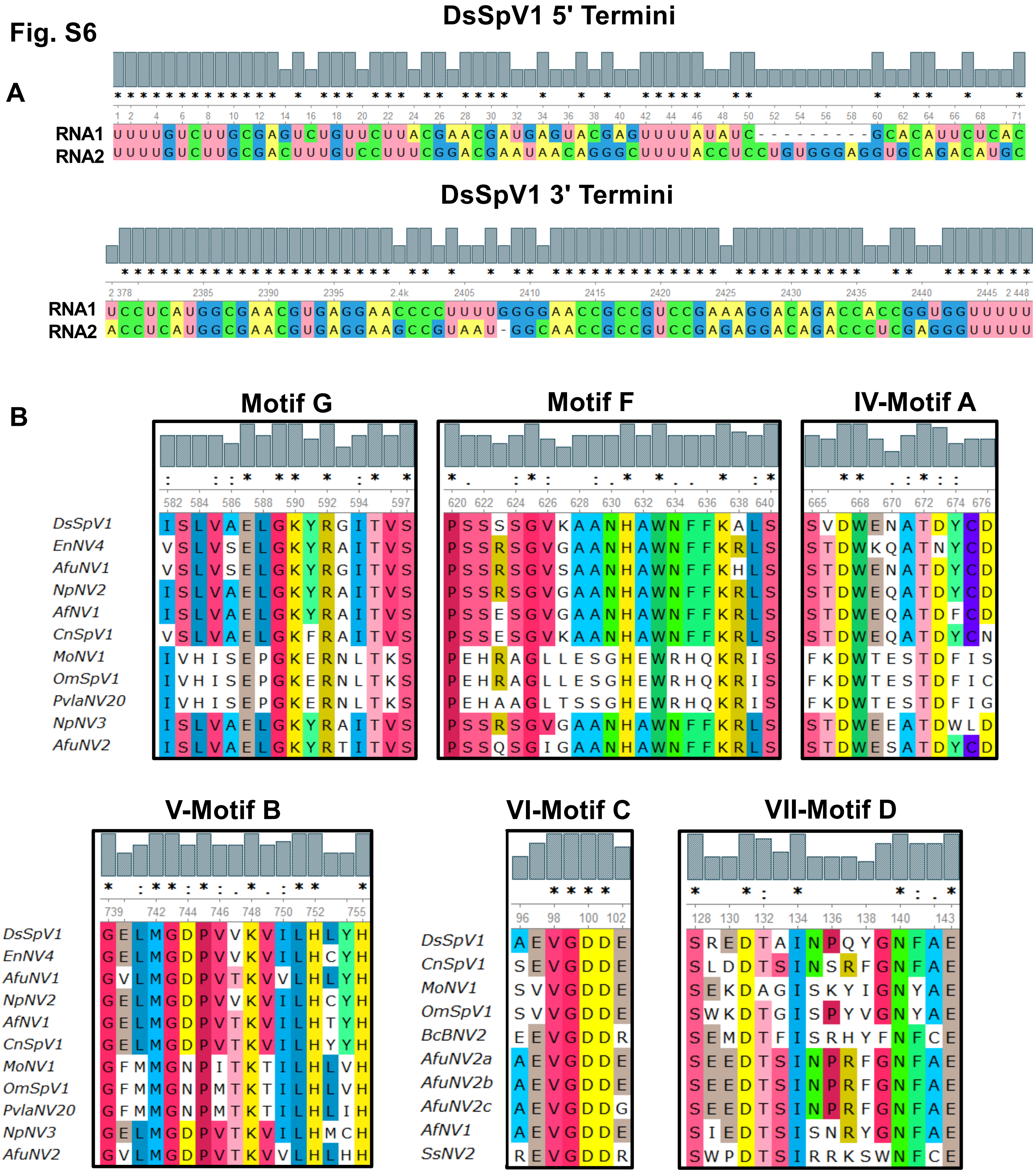
Motif VII

```
           564 ... 582
            ***. ***:*    ****
DsVV1      LRMNAMKQSVGHTSTEFLR
NpVV1      CRMNPAKQSVGFTTAEFLR
RnVV1      CRMNPTKQSIGFRRAEFLR
UvRV1      LRMNPLKQSVGLITGEFLR
MpVV1      LRMNPAKQSVGHVSTEFLR
PlTV1      LRMNPAKQSVGHVSTEFLR
AbVV1      LRMNPAKQSVGHVTTEFLR
CmRV       LRMNPAKQSVGHVTTEFLR
PdV1       LRMNPIKQSVGHTTTEFLR
MoV2       LRMNPIKQSVGHVSTEFLR
SsRV2      LRMNPMKQSVGHTSTEFLR
MpVV2      LRMNPMKQSVGHVSTEFLR
NpVV2      LRMNPMKQSVGHQSTEFLR
```

Motif VIII

```
           592 ... 600
           ** .*.::.
DsVV1      GYLARAVAS
NpVV1      GYLARAVAA
RnVV1      GYLARSVAA
UvRV1      GYAPRAIAS
MpVV1      GYLARAVAS
PlTV1      GYLARAVAS
AbVV1      GYLARSVAS
CmRV       GYLARAVAS
PdV1       GYFARGVAG
MoV2       GYFARAVAS
SsRV2      GYLARAVAS
MpVV2      GYLARAVSS
NpVV2      GYLARSVAS
```

Fig. S6



DsSpV1 5' Termini

```
RNA1 UUUUGUCUUGCGAGUCUGUUCUUACGAACGAUGAGUACGAGUUUUAUAUC---------GCACAUUCUCAC
RNA2 UUUUGUCUUGCGACUUUGUCCUUUCGGACGAAUAACAGGGCUUUUACCUCCUGUGGGAGGUGCAGACAUGC
```

DsSpV1 3' Termini

```
RNA1 UCCUCAUGGCGAACGUGAGGAACCCCUUUUGGGGAACCGCCGUCCGAAAGGACAGACCACCGGUGGUUUUU
RNA2 ACCUCAUGGCGAACGUGAGGAAGCCGUAAU-GGCAACCGCCGUCCGAGAGGACAGACCCUCGAGGGUUUUU
```

Motif G

```
DsSpV1    ISLVAELGKYRGITVS
EnNV4     VSLVSELGKYRAITVS
AfuNV1    VSLVSELGKYRGITVS
NpNV2     ISLVAELGKYRAITVS
AfNV1     ISLVAELGKYRAITVS
CnSpV1    VSLVAELGKFRAITVS
MoNV1     IVHISEPGKERNLTKS
OmSpV1    IVHISEPGKERNLTKS
PvlaNV20  IVHISEPGKERNLTKS
NpNV3     ISLVAELGKYRAITVS
AfuNV2    ISLVAELGKYRTITVS
```

Motif F

```
DsSpV1    PSSSSGVKAANHAWNFFKALS
EnNV4     PSSRSGVGAANHAWNFFKRLS
AfuNV1    PSSRSGVSAANHAWNFFKHLS
NpNV2     PSSRSGVGAANHAWNFFKRLS
AfNV1     PSSESGVGAANHAWNFFKRLS
CnSpV1    PSSESGVKAANHAWNFFKRLS
MoNV1     PEHRAGLLESGHEWRHQKRIS
OmSpV1    PEHRAGLLESGHEWRHQKRIS
PvlaNV20  PEHAAGLTSSGHEWRHQKRIS
NpNV3     PSSRSGVGAANHAWNFFKRLS
AfuNV2    PSSQSGIGAANHAWNFFKRLS
```

IV-Motif A

```
DsSpV1    SVDWENATDYCD
EnNV4     STDWKQATNYCD
AfuNV1    STDWENATDYCD
NpNV2     STDWEQATDYCD
AfNV1     STDWEQATDFCD
CnSpV1    STDWEQATDYCN
MoNV1     FKDWTESTDFIS
OmSpV1    FKDWTESTDFIC
PvlaNV20  FKDWTESTDFIG
NpNV3     STDWEEATDWLD
AfuNV2    STDWESATDYCD
```

V-Motif B

```
DsSpV1    GELMGDPVVKVILHLYH
EnNV4     GELMGDPVVKVILHCYH
AfuNV1    GVLMGDPVTKVVLHLYH
NpNV2     GELMGDPVVKVILHCYH
AfNV1     GELMGDPVTKVILHTYH
CnSpV1    GELMGDPVTKVILHYYH
MoNV1     GFMMGNPITKTILHLVH
OmSpV1    GFMMGNPMTKTILHLVH
PvlaNV20  GFMMGNPMTKTILHLIH
NpNV3     GELMGDPVTKVILHMCH
AfuNV2    GVLMGDPVTKVVLHLHH
```

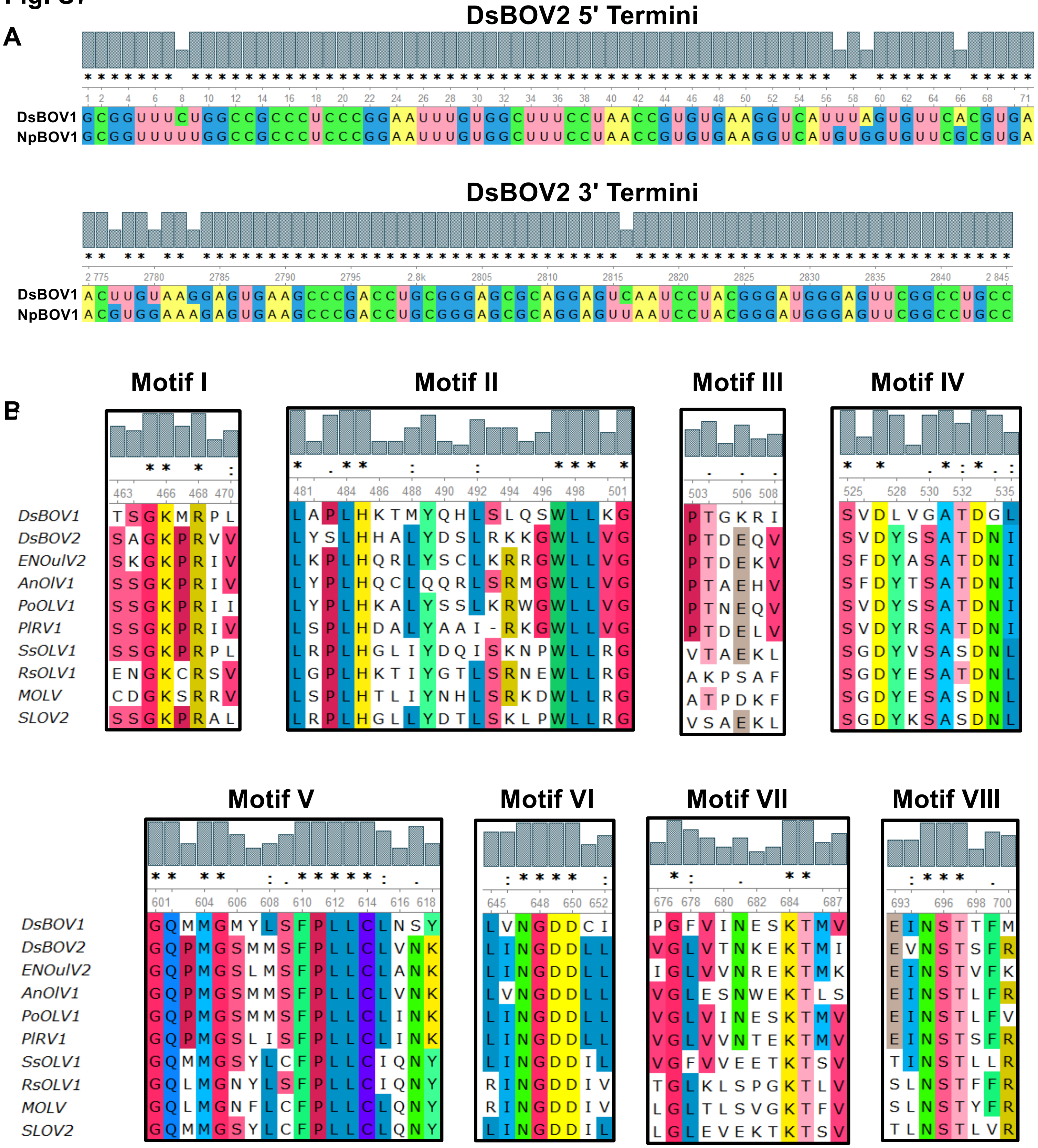
VI-Motif C

```
DsSpV1    AEVGDDE
CnSpV1    SEVGDDE
MoNV1     SVVGDDE
OmSpV1    SVVGDDE
BcBNV2    EEVGDDR
AfuNV2a   AEVGDDE
AfuNV2b   AEVGDDE
AfuNV2c   AEVGDDG
AfNV1     AEVGDDE
SsNV2     REVGDDR
```

VII-Motif D

```
DsSpV1    SREDTAINPQYGNFAE
CnSpV1    SLDDTSINSRFGNFAE
MoNV1     SEKDAGISKYIGNYAE
OmSpV1    SWKDTGISPYVGNYAE
BcBNV2    SEMDTFISRHYFNFCE
AfuNV2a   SEEDTSINPRFGNFAE
AfuNV2b   SEEDTSINPRFGNFAE
AfuNV2c   SEEDTSINPRFGNFAE
AfNV1     SIEDTSISNRYGNFAE
SsNV2     SWPDTSIRRKSWNFCE
```

Fig. S7



	DsBOV2 5' Termini
DsBOV1	GCGGUUUCUGGCCGCCCUCCCGGAAUUUGUGGCUUUCCUAACCGUGUGAAGGUCAUUUAGUGUUCACGUGA
NpBOV1	GCGGUUUUUGGCCGCCCUCCCGGAAUUUGUGGCUUUCCUAACCGUGUGAAGGUCAUGUGGUGUUCGCGUGA

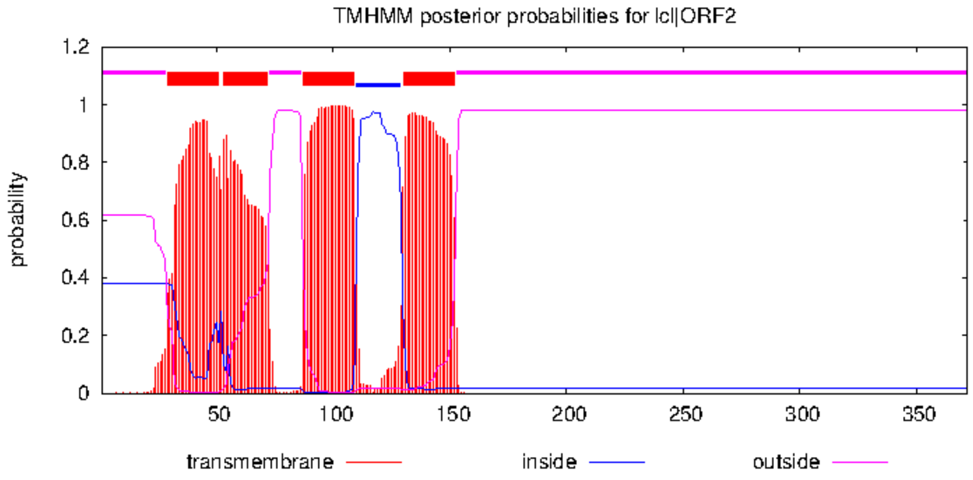
	DsBOV2 3' Termini
DsBOV1	ACUUGUAAGGAGUGAAGCCCGACCUGCGGGAGCGCAGGAGUCAAUCCUACGGGAUGGGAGUUCGGCCUGCC
NpBOV1	ACGUGGAAAGAGUGAAGCCCGACCUGCGGGAGCGCAGGAGUUAAUCCUACGGGAUGGGAGUUCGGCCUGCC

	Motif I	Motif II	Motif III	Motif IV
	** * :	* . ** : : *** *	. . .	* * . * : * . :
DsBOV1	TSGKMRPL	LAPLHKTMYQHLSLQSWLLKG	PTGKRI	SVDLVGATDGL
DsBOV2	SAGKPRVV	LYSLHHALYDSLRKKGWLLVG	PTDEQV	SVDYSSATDNI
ENOulV2	SKGKPRIV	LKPLHQRLYSCLKRRGWLLVG	PTDEKV	SFDYASATDNI
AnOlV1	SSGKPRIV	LYPLHQCLQQRLSRMGWLLVG	PTAEHV	SFDYTSATDNI
PoOLV1	SSGKPRII	LYPLHKALYSSLKRWGWLLVG	PTNEQV	SVDYSSATDNI
PlRV1	SSGKPRIV	LSPLHDALYAAI-RKGWLLVG	PTDELV	SVDYRSATDNI
SsOLV1	SSGKPRPL	LRPLHGLIYDQISKNPWLLRG	VTAEKL	SGDYVSASDNL
RsOLV1	ENGKCRSV	LGPLHKTIYGTLSRNEWLLRG	AKPSAF	SGDYESATDNL
MOLV	CDGKSRRV	LSPLHTLIYNHLSRKDWLLRG	ATPDKF	SGDYESASDNL
SLOV2	SSGKPRAL	LRPLHGLLYDTLSKLPWLLRG	VSAEKL	SGDYKSASDNL

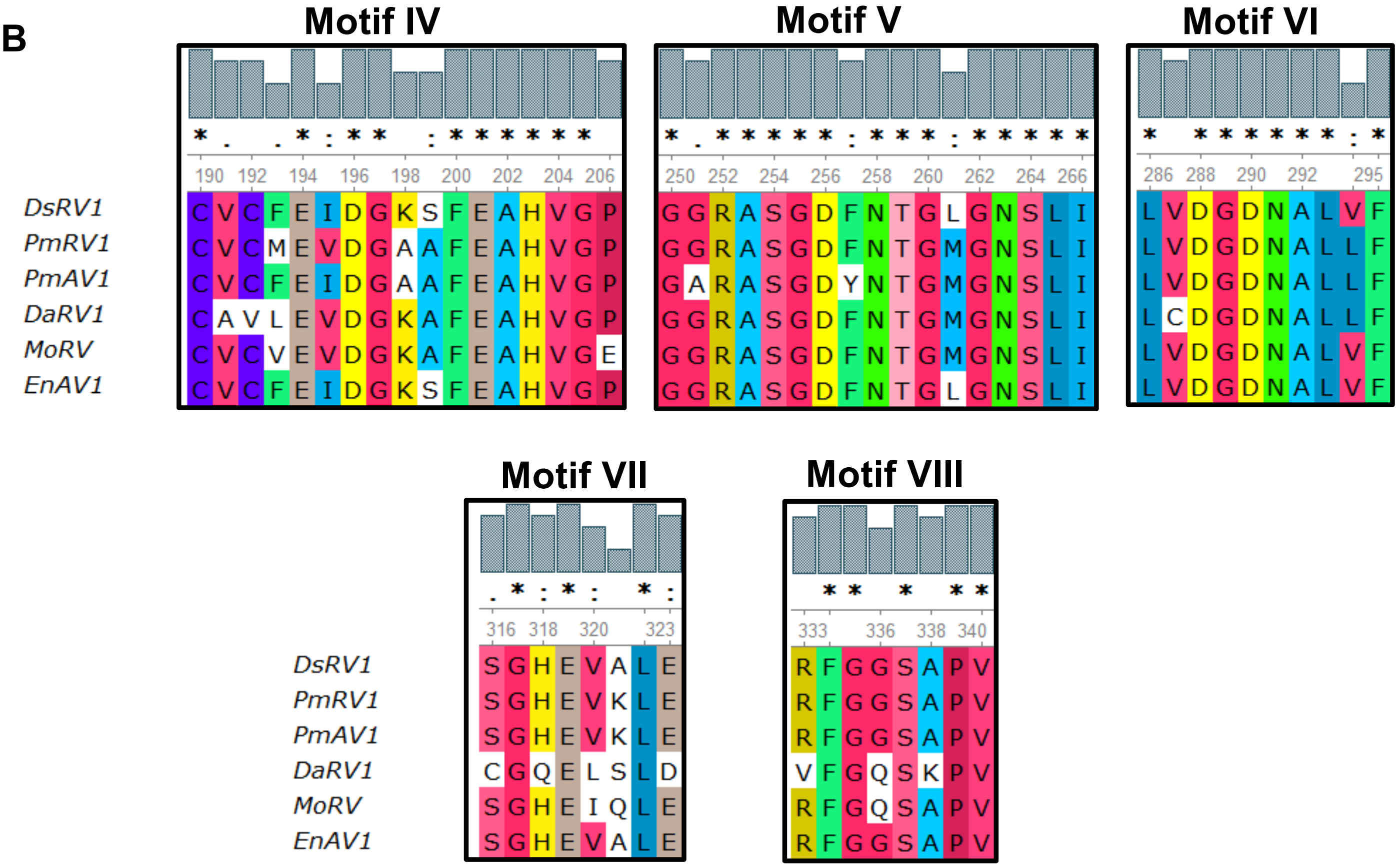
	Motif V	Motif VI	Motif VII	Motif VIII
	** ** : . ***** : .	: **** :	* : . **	: *** .
DsBOV1	GQMMGMYLSFPLLCLNSY	LVNGDDCI	PGFVINESKTMV	EINSTTFM
DsBOV2	GQPMGSMMSFPLLCLVNK	LINGDDLL	VGLVTNKEKTMI	EVNSTSFR
ENOulV2	GQPMGSLMSFPLLCLANK	LINGDDLL	IGLVVNREKTMK	EINSTVFK
AnOlV1	GQPMGSMMSFPLLCLVNK	LVNGDDLL	VGLESNWEKTLS	EINSTLFR
PoOLV1	GQPMGSMMSFPLLCLINK	LINGDDLL	VGLVINESKTMV	EINSTLFV
PlRV1	GQPMGSLISFPLLCLINK	LINGDDLL	VGLVVNTEKTMV	EINSTSFR
SsOLV1	GQMMGSYLCFPLLCIQNY	LINGDDIL	VGFVVEETKTSV	TINSTLLR
RsOLV1	GQLMGNYLSFPLLCIQNY	RINGDDIV	TGLKLSPGKTLV	SLNSTFFR
MOLV	GQLMGNFLCFPLLCLQNY	RINGDDIV	LGLTLSVGKTFV	SLNSTYFR
SLOV2	GQMMGSYLCFPLLCLQNY	LINGDDIL	LGLEVEKTKTSV	TLNSTLVR

Fig. S8

A



B



Motif IV

	190		192		194		196		198		200		202		204		206
	*	.		.	*	:	*	*		:	*	*	*	*	*	*	
DsRV1	C	V	C	F	E	I	D	G	K	S	F	E	A	H	V	G	P
PmRV1	C	V	C	M	E	V	D	G	A	A	F	E	A	H	V	G	P
PmAV1	C	V	C	F	E	I	D	G	A	A	F	E	A	H	V	G	P
DaRV1	C	A	V	L	E	V	D	G	K	A	F	E	A	H	V	G	P
MoRV	C	V	C	V	E	V	D	G	K	A	F	E	A	H	V	G	E
EnAV1	C	V	C	F	E	I	D	G	K	S	F	E	A	H	V	G	P

Motif V

	250		252		254		256		258		260		262		264		266
	*	.	*	*	*	*	*	:	*	*	*	:	*	*	*	*	*
DsRV1	G	G	R	A	S	G	D	F	N	T	G	L	G	N	S	L	I
PmRV1	G	G	R	A	S	G	D	F	N	T	G	M	G	N	S	L	I
PmAV1	G	A	R	A	S	G	D	Y	N	T	G	M	G	N	S	L	I
DaRV1	G	G	R	A	S	G	D	F	N	T	G	M	G	N	S	L	I
MoRV	G	G	R	A	S	G	D	F	N	T	G	M	G	N	S	L	I
EnAV1	G	G	R	A	S	G	D	F	N	T	G	L	G	N	S	L	I

Motif VI

	286		288		290		292			295
	*		*	*	*	*	*	*	:	*
DsRV1	L	V	D	G	D	N	A	L	V	F
PmRV1	L	V	D	G	D	N	A	L	L	F
PmAV1	L	V	D	G	D	N	A	L	L	F
DaRV1	L	C	D	G	D	N	A	L	L	F
MoRV	L	V	D	G	D	N	A	L	V	F
EnAV1	L	V	D	G	D	N	A	L	V	F

Motif VII

	316		318		320			323
	.	*	:	*	:		*	:
DsRV1	S	G	H	E	V	A	L	E
PmRV1	S	G	H	E	V	K	L	E
PmAV1	S	G	H	E	V	K	L	E
DaRV1	C	G	Q	E	L	S	L	D
MoRV	S	G	H	E	I	Q	L	E
EnAV1	S	G	H	E	V	A	L	E

Motif VIII

	333			336		338		340
		*	*		*		*	*
DsRV1	R	F	G	G	S	A	P	V
PmRV1	R	F	G	G	S	A	P	V
PmAV1	R	F	G	G	S	A	P	V
DaRV1	V	F	G	Q	S	K	P	V
MoRV	R	F	G	Q	S	A	P	V
EnAV1	R	F	G	G	S	A	P	V

Supplementary Table S1. List of primers used in this study.

Primer name	Primer sequence (5’-3’)	Amplicon size (bp)	Purpose
contig_73(F)	AAAGCGATGGTCTCCTACTTCA	552	RT-PCR
contig_73(R)	AACGTAGTCTTCAAGCTCCTGA		
contig_995 (F)	AATGCGCGGACTGTTCCGTTT	505	RT-PCR
contig_995 (R)	GATCAAGGAGGTTGGCTCGAT		
contig_218 (F)	AAGGGCACGTATTGGTTCGGAA	665	RT-PCR
contig_218 (R)	ACACCTTCGCAGATGTCGTAGT		
contig_310 (F)	GAAGTTGGCGACGATGAAGAAG	729	RT-PCR
contig_310 (R)	TTCTGCACGAACCTCGAACTC		
contig_38 (F)	AAGGTTAGGTCAGACACACACTT	493	RT-PCR
contig_38 (R)	TCCATACAAGTTGAGGCACACA		
contig_217 (F)	TCATGGTTGTTGAAGGGACCTC	770	RT-PCR
contig_217 (R)	TAGTCTGGTTCGGCCATAACAG		
contig_980 (F)	TTGCACCCGCGTTTCCTTGA	754	RT-PCR
contig_980 (R)	GCGCTACCTTGACCATGAGA		
contig_24 (F)	GCCAGTTAACTTGGTCTTCGCT	682	RT-PCR
contig_24 (R)	ACGAGGAAAGGCTCGTTGATC		
contig 33 (F)	TCATGGGGACGAACAAGTTGG	600	RT-PCR
contig 33 (R)	GGTGCTGCAACAGCCAAGTT		
Cont175 F	TTCTCTGCCATCAAATCATCGGA	519	RT-PCR
Cont175 R	TCGGAACTACTGCTTGATGCAATA		
Cont255 F	TTACATGGATCTCGTCAACACAGT	626	RT-PCR

Cont255 R	ACCAACCAAGCGACGAGTAAGT		
Cont 70 Chryso F	ACCAATGATACAGTCGCACGTCA	579	RT-PCR
Cont 70 Chryso R	AGCGCTGACAATTGGTACAGG		
Cont 348 Chryso F	GTTAGTTCGGCTCAGCGACTT	535	RT-PCR
Cont 348 Chryso R	CATCATCATGCCCATCCAGTTC		
Cont 151 Chryso F	ATCACCAGACCGTCTGTACCAT	569	RT-PCR
Cont 151 Chryso R	GTATCATCATCATCAGCTGAGGC		
Cont 965 polymy F	TGCTATCACCGTGGTGTCTCT	563	RT-PCR
Cont 965 polymy R	CTGTGTATTCCTCTCGATGCG		
Cont 245 polymy F	AGCGTAAGTACTCTGCCGTCAT	587	RT-PCR
Cont 245 polymy R	GACTCCGTGATCGACAGAATC		
Cont 706 polymy F	TCGAAATGGCCTCCGAACAAAC	533	RT-PCR
Cont 706 polymy R	CGTGAAACGGTAGGATCCATAG		
contig_980 (5P)	AACCGTTTCTTTGACATGTCCGA	397	RLM-RACE
contig_980 (3P)	TCCAGAAGCTCGGTTTCGCT	264	RLM-RACE
contig_965 (5P)	GTAGAGAGGATGTATATTGCGG	834	RLM-RACE
contig_965 (3P)	AGGATTGCCCCTCGTGTGAA	417	RLM-RACE
contig_245 (5P)	ATTGTCCTCGAGGGTTGTTAGT	561	RLM-RACE
contig_245 (3P)	GCTGCCGACGTTGAGGGTCTG	391	RLM-RACE
contig_706 (5P)	CGTACGTGGCGTGGTATTGA	357	RLM-RACE
contig_706 (3P)	GCATCACCCGCATCAAAGGAA	411	RLM-RACE
contig_24 (5P)	GGCGGTATAGTGTGTGTCCAT	714	RLM-RACE
contig_24 (3P)	TTCTACTGCGACAACGTTTTGG	654	RLM-RACE
contig 33 (5P)	CCAACTTGTTCGTCCCCATGA	300	RLM-RACE
contig 33 (3P)	TCGGCCCAAACAGCTGAAACT	500	RLM-RACE

contig_73(5P)	TAAGGACGTCCGCTTTGGTG	469	RLM-RACE
contig_73(3P)	GGAGCTTGAAGACTACGTTCG	379	RLM-RACE
contig_995 (5P)	CAACCTCGTAGAAGATCGACC	487	RLM-RACE
contig_995 (3P)	ATCCTGGGTGACGTTAAGAGAA	403	RLM-RACE
contig_218 (5P)	CCATATGTTTTCTGCATTCCAGC	448	RLM-RACE
contig_218 (3P)	AGGTATACCTGTTCTCCGTTGA	406	RLM-RACE
contig_310 (5P)	GGGTTTATCGCAGTATCTTCTC	462	RLM-RACE
contig_310 (3P)	GTTATGCCGTGAGGCAAACTT	464	RLM-RACE
contig_38 (5P)	ATCCGAAGACATACTTGCGTAAG	539	RLM-RACE
contig_38 (3P)	AAGTGAGTGGAATACGCAGTCAT	443	RLM-RACE
contig_217 (5P)	TTACACGTAAGGTTAGGGATCG	464	RLM-RACE
contig_217 (3P)	AAAGAAAAGAGAGTGGTTGGCG	512	RLM-RACE