CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element
Authors	Firth, Andrew E.;Wills, Norma M.;Gesteland, Raymond F.;Atkins, John F.
Publication date	2011
Original Citation	Firth, A. E., Wills, N. M., Gesteland, R. F. and Atkins, J. F. (2011) 'Stimulation of stop codon readthrough: frequent presence of an extended 3# RNA structural element', Nucleic Acids Research, 39(15), pp. 6679-6691. doi: 10.1093/nar/gkr224
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkr224 - 10.1093/nar/gkr224
Rights	© 2011, the Authors. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited. - http://creativecommons.org/licenses/by-nc/2.5
Download date	2024-03-29 15:50:46
Item downloaded from	https://hdl.handle.net/10468/5023

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Alphavirus

Venezuelan equine encephalitis virus (VEEV; RefSeq NC_001449 [derived from L04653])

Notes on the nucleotide sequences:

1. The readthrough stop codon is indicated in red. The predicted stem is indicated in yellow. Single substitutions that preserve the predicted base-pairings (e.g. G-C to G-U or G-U to A-U) are indicated in pale blue. Compensatory substitutions (paired substitutions that preserve the predicted base-pairings) are indicated with pink.
2. The readthrough codon is always UGA and the 3'-adjacent codon is CGG with just 2 exceptions (CGC in AF075257 and CGA in AF075258).
3. The non-degenerate positions of the 5' half of the predicted stem (here GCN GGN GCN UAY) are conserved also because they form the NSP3|NSP4 cleavage site: [VA]-G-[GA] | Y. This holds also for the other alphavirus species.
4. The predicted stem is supported by highly significant enhanced synonymous site conservation (main text Fig 1), some compensatory substitutions, and mutational experiments (main text).
5. The loop region also shows enhanced synonymous site conservation (albeit much lower than in the stem region; main text Fig 1) even though there are no completely conserved 3rd codon positions, and generally folds into an imperfect hairpin structure.
6. There is a completely conserved CUC codon (6-fold degenerate leucine) just 3' of the 3' component of the predicted stem and a nearly completely conserved CGC (6-fold degenerate arginine) just preceding this. The -1 and -3 codons (both CAA; 2-fold degenerate glutamine) are also completely conserved.

An extract around the readthrough stop codon is shown below for the VEEV sequences used in the synonymous site conservation analysis (main text Fig 1; see below for the WEEV and EEEV sequences used).

```
                                                  (  (((  ((   (((  (((
AF075257   GUU GCG CAA CAG CAA UGA CGC UAU GAC GCG GGU GCA UAC AUC UUC UCU UCG GAU ACC GGA
U55350     GUA GCA CAA CAA CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
U55347     GUA GCA CAA CAA CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
U55345     GUA GCA CAA CAA CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
AF375051   GUA GCA CAA CAA CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
NC_001449  GUA GCA CAA CAA CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
AY986475   GUA GCA CAA CAA CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
AY973944   GUA GCA CAA CAA CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
U55342     GUA GCA CAA CAA CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
L00930     GUA GCA CAA CAA CAA UGA CGG UUU GAU GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
DQ390224   GUA GCA CAA CAA CAA UGA CGG UUU GAU GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
AF100566   GUA GCA CAA CAG CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
U55360     GUA GCA CAA CAG CAA UGA CGG UUU GAC GCG GGU GCA UAC ACC UUU UCC UCC GAU ACC GGU
AF004459   GUA GCA CAA CAG CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
U55362     GUA GCA CAA CAA CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
AF004472   GUA GCA CAA CAG CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
AF004458   GUA GCA CAA CAG CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
L01442     GUA GCA CAA CAA CAA UGA CGG UUU GAU GCG GGU GCA UAC AUC UUU UCC UCC GAC ACC GGU
J04332     GUA GCA CAA CAA CAA UGA CGG UUU GAU GCG GGU GCA UAC AUC UUU UCC UCC GAC ACC GGU
L01443     GUA GCA CAA CAA CAA UGA CGG UUU GAU GCG GGU GCA UAC AUC UUU UCC UCC GAC ACC GGU
AY741139   GUA GCA CAA CAA CAA UGA CGG UUU GAU GCG GGU GCA UAC AUC UUU UCC UCC GAC ACC GGU
AF069903   GUA GCA CAA CAA CAA UGA CGG UUU GAU GCG GGU GCA UAC AUC UUU UCC UCC GAC ACC GGU
AF075251   GUA GCA CAA CAG CAA UGA CGG UUC GAC GCG GGU GCA UAC AUA UUU UCC UCC GAU ACC GGU
AF448539   GUU GCC CAA CAG CAA UGA CGG UUC GAC GCG GGC GCG UAC AUU UUC UCC UCG GAU ACU GGU
AF448538   GUU GCC CAA CAG CAA UGA CGG UUC GAC GCG GGC GCG UAC AUU UUC UCC UCG GAU ACU GGU
U34999     GUU GCC CAA CAG CAA UGA CGG UUC GAC GCG GGC GCG UAC AUU UUC UCC UCG GAU ACU GGU
AY823299   GUU GCC CAA CAG CAA UGA CGG UUC GAC GCG GGC GCG UAC AUU UUC UCC UCG GAU ACU GGU
AF448537   GUU GCC CAA CAG CAA UGA CGG UUC GAC GCG GGC GCG UAC AUU UUC UCC UCG GAU ACU GGU
AF448536   GUU GCC CAA CAG CAA UGA CGG UUC GAC GCG GGC GCG UAC AUU UUC UCC UCG GAU ACU GGU
AF448535   GUU GCC CAA CAG CAA UGA CGG UUC GAC GCG GGC GCG UAC AUU UUC UCC UCG GAU ACU GGU
AF075252   GUC GCC CAA CAG CAA UGA CGG UUC GAC GCG GGU GCA UAC AUU UUC UCC UCG GAU ACU GGC
AF075256   AUC GCU CAA CAG CAA UGA CGG UUU GAC GCG GGU GCA UAC AUC UUU UCC UCC GAC ACC GGU
AF075255   GUU GCU CAA CAG CAA UGA CGG UAU GAU GCG GGU GCA UAC AUC UUU UCC UCC GAU ACC GGU
AF075254   GUU GCG CAA CAG CAA UGA CGG UAC GAA GCG GGU GCG UAC AUU UUU UCC UCU GAU ACC GGU
```

```
AF075253   GUC GCG CAA CAA CAA UGA CGG UAC GAA GCG GGU GCG UAC AUC UUC UCC UCC GAC ACC GGU
AF075259   GUC GCA CAA CAG CAA UGA CGG UAU GAA GCG GGU GCG UAC AUC UUU UCC UCC GAU ACA GGC
AF075258   GUG GCG CAA CAA CAA UGA CGA UAU GAA GCG GGU GCA UAC AUA UUC UCC UCG GAC ACC GGC
            *  **  *** **  *** *** **  *   **  *** **  **  *** *   **  **  **  **  **  **

AF075257   CAA GGA CAU UUA CAA CAG AAA UCG GUG AGG CAA ACC GUC UUG UCC GAG GUG AUU CUA GAG
U55350     CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
U55347     CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
U55345     CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
AF375051   CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
NC_001449  CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
AY986475   CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
AY973944   CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
U55342     CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
L00930     CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG CUA UCU GAA GUG GUG UUG GAG
DQ390224   CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAG ACG GUG CUA UCC GAA GUG GUG UUG GAG
AF100566   CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
U55360     CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
AF004459   CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
U55362     CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
AF004472   CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
AF004458   CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG UUA UCC GAA GUG GUG UUG GAG
L01442     CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG CUA UCC GAA GUG GUG UUG GAG
J04332     CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG CUA UCC GAA GUG GUG UUG GAG
L01443     CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG CUA UCC GAA GUG GUG UUG GAG
AY741139   CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG CUA UCC GAA GUG GUG UUG GAG
AF069903   CAA GGG CAU UUA CAA CAA AAA UCA GUA AGG CAA ACG GUG CUA UCC GAA GUG GUG UUG GAG
AF075251   CAA GGG CAU UUA CAG CAA AAA UCA GUA AGG CAA ACG GUG CUA UCC GAA GUG GUA UUG GAG
AF448539   CAA GGA CAU UUG CAA CAA AAA UCA GUA AGG CAG ACA GUA UUG UCU GAA GUG GUG CUA GAG
AF448538   CAA GGA CAU UUG CAA CAA AAA UCA GUA AGG CAG ACA GUA UUG UCU GAA GUG GUG CUA GAG
U34999     CAA GGA CAU UUG CAA CAA AAA UCA GUA AGG CAG ACA GUA UUG UCU GAA GUG GUG CUA GAG
AY823299   CAA GGA CAU UUG CAA CAA AAA UCA GUA AGG CAG ACA GUA UUG UCU GAA GUG GUG CUA GAG
AF448537   CAA GGA CAU UUG CAA CAA AAA UCA GUA AGG CAG ACA GUA UUG UCU GAA GUG GUG CUA GAG
AF448536   CAA GGA CAU UUG CAA CAA AAA UCA GUA AGG CAG ACA GUA UUG UCU GAA GUG GUG CUA GAG
AF448535   CAA GGA CAU UUG CAA CAA AAA UCA GUA AGG CAG ACA GUA UUG UCU GAA GUG GUG CUA GAG
AF075252   CAA GGA CAU CUG CAA CAA AAA UCA GUA AGG CAG ACA GUG UUG UCU GAA GUG GUG CUA GAG
AF075256   CAG GGG CAC CUG CAA CAG AAA UCA GUC CGG CAA ACG GUU UUG UCG GAG GUG GUG CUG GAG
AF075255   CAA GGG CAU UUA CAA CAA AAG UCA GUA AGA CAG ACA GUG UUG UCU GAA AUU GUA CUG GAA
AF075254   CAG GGG CAU UUA CAA CAA AAG UCA GUG AGG CAG ACG ACG CUG UCC GAG GUA GUA CUU GAA
AF075253   CAA GGG CAU CUA CAA CAA AAA UCU GUG AGG CAA ACA GCG CUG UCC GAA GUG GUG CUU GAA
AF075259   CAA GGA CAU UUG CAA CAA AAA UCC GUA CGG CAG ACC GUU CUA UCU GAG GUC GUA UUG GAA
AF075258   CAA GGA CAU UUA CAG CAA AAA UCG GUG AGG CAG ACG GUG CUG UCC GAG GUC GUG CUA GAA
           **  **  **   *  **  **  **  **  **   *  **  **       *  **  **   *   *   *  **
                                      ) ))) ))) ))) ))
AF075257   CGC ACA GAA CUG GAA AUU GCG UAU GCC CCG CGC CUC GAC UUA GCU AAA GAA GAG CUG CUG
U55350     AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAG GAA AAA GAA GAA CUA CUA
U55347     AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAG GAA AAA GAA GAA CUA CUA
U55345     AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAG GAA AAA GAA GAA CUA CUA
AF375051   AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAG GAA AAA GAA GAA CUA CUA
NC_001449  AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAG GAA AAA GAA GAA CUA CUA
AY986475   AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAG GAA AAA GAA GAA CUA CUA
AY973944   AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAG GAA AAA GAA GAA CUA CUA
U55342     AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAG GAA AAA GAA GAA CUA CUA
L00930     AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAG UUA UUA
DQ390224   AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA UUA UUA
AF100566   AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA CUA CUA
U55360     AGG ACU GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA CUA CUA
AF004459   AGG ACU GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA CUA CUA
U55362     AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA CUA CUA
AF004472   AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA CUA CUA
AF004458   AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA CUA CUA
L01442     AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA UUA CUA
J04332     AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA UUA CUA
L01443     AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA UUA CUA
```

```
AY741139   AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA UUA CUA
AF069903   AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CAA GAA AAA GAA GAA UUA CUA
AF075251   AGG ACC GAA UUG GAG AUU UCG UAU GCC CCG CGC CUC GAC CUA GAG AAG GAA GAA UUA CUA
AF448539   AGG ACC GAG UUA GAG AUC UCG UAC GCC CCG CGC CUC GAC CUG AAC AAA GAA GAA AUA CUG
AF448538   AGG ACC GAG UUA GAG AUC UCG UAC GCC CCG CGC CUC GAC CUG AAC AAA GAA GAA AUA CUG
U34999     AGG ACU GAG UUA GAG AUC UCG UAC GCC CCG CGC CUC GAC CUG AAC AAA GAA GAA AUA CUG
AY823299   AGG ACC GAG UUA GAG AUC UCG UAC GCC CCG CGC CUC GAC CUG AAC AAA GAA GAA AUA CUG
AF448537   AGG ACC GAG UUA GAG AUC UCG UAC GCC CCG CGC CUC GAC CUG AAC AAA GAA GAA AUA CUG
AF448536   AGG ACC GAG UUA GAG AUC UCG UAC GCC CCG CGC CUC GAC CUG AAC AAA GAA GAA AUA CUG
AF448535   AGG ACC GAG UUA GAG AUC UCG UAC GCC CCG CGC CUC GAC CUG AAC AAA GAA GAA AUA CUG
AF075252   AGG ACC GAA UUA GAG AUC UCG UAU GCC CCG CGC CUC GAC CUG AAC AAA GAA GAG AUA CUG
AF075256   AGG ACC GAG UUG GAG AUA GCG UAU GCC CCG CGC CUC GAC CUG GAG AAG GAG GAA GCA CUA
AF075255   CGA ACA GAA UUG GAG AUA GCG UAU GCC CCG CGC CUC GAC CUG GAA AAA GAA GAA GCU UUA
AF075254   CGG ACA GAG UUA GAG CAU UCG UAC GCC CCG CGC CUC GAC CUA GAG AAA GAA GAG AUA CUA
AF075253   CGG ACG GAG UUG GAG AAU UCG UAC GCC CCG CGC CUC GAC UUA AUA AAA GAG GAA AGU UUA
AF075259   CGC ACG CUA UUG GAU GAC GUG UAC GCC CCG CGU CUC GAU CUU AAC AAG GAG GAG UUA CUG
AF075258   CGC ACG GAA UUG GAG AUA UCG UAU GCC CCG CGC CUC GAU UUA AAU AAA GAA GAA AUU UUA
             *  **       *  **        * **  *** *** **  *** **          **  **  **       *
```

Western and Eastern equine encephalitis/encephalomyelitis viruses (WEEV, EEEV)

RefSeqs:

NC_003899	Eastern equine encephalitis virus	(derived from X63135)
NC_003908	Western equine encephalomyelitis virus	(derived from AF214040)
NC_012561	Highlands J virus	(derived from FJ827631)
NC_013528	Fort Morgan virus	(derived from GQ281603)

Notes on the nucleotide sequences:

1. The readthrough codon is always UGA and the 3'-adjacent codon is CGG with just 1 exception (CGC in NC_013528).
2. Again, there is a completely conserved CUC leucine codon just 3' of the 3' component of the predicted stem and a completely conserved CGC arginine codon just preceding this.

An extract around the readthrough stop codon is shown below for the WEEV/EEEV sequences used in the synonymous site conservation analysis (main text Fig 1; see above for the VEEV sequences used), plus the additional partial sequences X74892 and U01065.

```
                                                  ((( ((  ((( ((    [ [[[ [[[ [
AY705240   CGG AGG CAC UCG AAU UGA CGG UAC GAA GCG GGC GCG UAC AUU UUC UCA UCC GAG ACG GGA
EF568607   CGG AGG CAC UCG AAU UGA CGG UAC GAA GCG GGC GCG UAC AUU UUC UCA UCC GAG ACG GGA
AY705241   CGG AGG CAC UCG AAU UGA CGG UAC GAA GCG GGC GCG UAC AUU UUC UCA UCC GAG ACG GGA
AY722102   CGG AGG CAC UCG AAU UGA CGG UAC GAA GCG GGC GCG UAC AUU UUC UCA UCC GAG ACG GGA
NC_003899  CGG AGG CAC UCG AAU UGA CGG UAC GAA GCG GGC GCG UAC AUU UUC UCA UCC GAG ACG GGA
EF151502   CGG AGG CAC UCG AAU UGA CGG UAC GAA GCG GGC GCG UAC AUU UUC UCA UCC GAG ACG GGA
U01034     CGG AGG CAC UCG AAU UGA CGG UAC GAA GCG GGC GCG UAC AUU UUC UCA UCC GAG ACG GGA
DQ241303   CGC CGA CAU UCA AAU UGA CGG UAU GAA GCG GGU GCG UAU AUU UUC UCG UCA GAG ACA GGG
DQ241304   CGU CGG CAC UCU AAC UGA CGG UAU GAA GCG GGU GCG UAU AUC UUC UCA UCA GAG ACA GGG
EF151503   CGC CAG CAC UCA AAU UGA CGG UAC GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGA
NC_013528  CAA CAG CAA CUG AAU UGA CGC UAC GAA GCG GGA GCA UAC AUA UUC UCA UCA GAG ACU GGU
GU167952   CGC CAG CAC ACG AAU UGA CGG UAC GAA GCG GGU GCG UAC AUC UUC UCC UCU GAG ACG GGC
NC_012561  CGC CAG CAC ACG AAU UGA CGG UAC GAA GCG GGU GCG UAC AUC UUC UCC UCU GAG ACG GGC
GQ227789   CGC CAG CAC ACG AAU UGA CGG UAC GAA GCG GGU GCG UAC AUC UUC UCC UCU GAG ACG GGC
GQ287646   CAC CAA CAC UCC AAU UGA CGG UAU GAA GCG GGA GCG UAU AUC UUC UCA UCG GAA ACA GGC
FJ786263   CGU CAA CAU UCC AAC UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
GQ287640   CGU CAA CAU UCC AAC UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
GQ287644   CGU CAA CAC UCC AAC UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
FJ786260   CGU CAA CAC UCC AAC UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
FJ786261   CGU CAA CAC UCC AAC UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
GQ287645   CGU CAA CAC UCC AAC UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
GQ287647   CGU CAA CAC UCC AAU UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
NC_003908  CGU CAA CAC UCC AAC UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
X74892     CGU CAA CAU UCC AAC UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
U01065     CGU CAA CAU UCC AAC UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
```

```
GQ287641   CGU CAA CAC UCC AAU UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
GQ287642   CGU CAA CAC UCC AAU UGA CGG UAU GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
GQ287643   CGU CAA CAC UCC AAC UGA CGG UAC GAA GCG GGA GCG UAU AUU UUC UCA UCG GAA ACA GGC
           *       **      **  *** **  **  *** *** **  **  **  **  *** **  **  **  **  **

AY705240   CAA GGG CAC CUG CAA CAA AAA UCU ACG CGG CAA UGC AAA CUC CAG UAU CCA AUC CUG GAG
EF568607   CAA GGG CAC CUG CAA CAA AAA UCU ACG CGG CAA UGC AAA CUC CAG UAU CCA AUC CUG GAG
AY705241   CAA GGG CAC CUG CAA CAA AAA UCU ACG CGG CAA UGC AAA CUC CAG UAU CCA AUC CUG GAG
AY722102   CAA GGG CAC CUG CAA CAA AAA UCU ACG CGG CAA UGC AAA CUC CAG UAU CCA AUC CUG GAG
NC_003899  CAA GGG CAC CUG CAA CAA AAA UCU ACG CGG CAA UGC AAA CUC CAG UAU CCA AUC CUG GAG
EF151502   CAA GGG CAC CUG CAA CAA AAA UCC ACG CGG CAA UGC AAA CUC CAG UAU CCA AUC CUG GAG
U01034     CAA GGG CAC CUG CAG CAA AAA UCU ACG CGG CAA UGC AAA CUC CAG UAU CCA AUC CUG GAG
DQ241303   CAA GGG CAU CUG CAA CAG AAA UCA ACU CGG CAG UGC AAA CUC CAA AAU CCU AUC UUG GAG
DQ241304   CAA GGG CAU CUG CAA CAG AAA UCC ACC AGG CAG UGC AAA CUC CAA CAC CCU AUU CUG GAG
EF151503   CAA GGG CAU CUC CAA CAG AAA UCC ACU AGA CAG UGC ACA CUU CAG CAU CCU AUA UUG GAG
NC_013528  CAG GGU CAC UUG CAA CAG AAA UCU AAC AGG CAA GGC CCU UAC GUG UAC CCG GUC CUG GAA
GU167952   CAA GGA CAC CUC CAG CAG AAA UCA GUA CGG CAA UGC AAA CUU CAA GAC GCA GUA CUA GAG
NC_012561  CAA GGA CAC CUC CAG CAG AAA UCA GUA CGG CAA UGC AAA CUU CAA GAC GCA GUA CUA GAG
GQ227789   CAA GGA CAC CUC CAG CAG AAA UCA GUA CGG CAA UGC AAA CUU CAA GAC GCA GUA CUA GAG
GQ287646   CAG AGU CAC CUU CAG CAG AAA UCA GUA CGU CAA UGU AAA CUU CAG GAU CCU AUA UUG GAA
FJ786263   CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAC CCU AUA UUU GAA
GQ287640   CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAC CCU AUA UUU GAA
GQ287644   CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAA CCU AUA UUG GAA
FJ786260   CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAA CCU AUA UUG GAU
FJ786261   CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAA CCU AUA UUG GAU
GQ287645   CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAA CCU AUA UUG GAU
GQ287647   CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAA CCU AUA UUG GAU
NC_003908  CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAA CCU AUA UUG GAU
X74892     CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAC CCU AUA UUU GAA
U01065     CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAC CCU AUA UUU GAA
GQ287641   CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAA CCU AUA UUG GAC
GQ287642   CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAA CCU AUA UUG GAC
GQ287643   CAA GGU CAC CUU CAA CAG AAA UCA GUA CGU CAA UGU AAA CUA CAA GAA CCU AUA UUG GAU
           **   *  **   *  **  **  *** **       *  **   *                 *   *   *   *  **
                        ]] ]]] ]]]     ))) ))) ))) )
AY705240   CGU UCC GUC CAU GAG AAA UUU UAC GCC CCG CGC CUC GAU CUC GAG CGU GAG AAG CUG UUG
EF568607   CGU UCC GUC CAU GAG AAA UUU UAC GCC CCG CGC CUC GAU CUC GAG CGC GAG AAG CUG UUG
AY705241   CGU UCC GUC CAU GAG AAA UUU UAC GCC CCG CGC CUC GAU CUC GAG CGU GAG AAG CUG UUG
AY722102   CGU UCC GUC CAU GAG AAA UUU UAC GCC CCG CGC CUC GAU CUC GAG CGU GAG AAG CUG UUG
NC_003899  CGU UCC GUC CAU GAG AAA UUU UAC GCC CCG CGC CUC GAU CUC GAG CGU GAG AAG CUG UUG
EF151502   CGU UCC GUC CAU GAG AAA UUU UAC GCC CCG CGC CUC GAU CUC GAG CGU GAG AAG CUG UUG
U01034     CGU UCC GUC CAU GAG AAA UUU UAC GCC CCG CGC CUC GAU CUC GAG CGU GAG AAG CUG UUG
DQ241303   CGC UCU GUC CAC GAG AAA UUC UAC GCC CCG CGC CUC GAC UUG GAA CGU GAG AAA CUG UUA
DQ241304   CGC UCU GUC CAU GAG AAG UUC UAC GCC CCG CGC CUC GAC CUG GAA CGU GAG AAA UUG CUG
EF151503   CGC UCU GUC CAU GAG AGA UUC UAC GCC CCG CGC CUC GAC CUU GAA AAA GAA AAA CUA CUA
NC_013528  AAA UCA GUA CAU GAG AAA UUU UAU GCC CCG CGC CUC GAU AUG GAG AAA GAG AAA AUU CUC
GU167952   AGG GUC ACU CAU GAG AAG UAU UAC GCC CCG CGC CUC GAC CUA GAA AAA GAG AAG UUG UUG
NC_012561  AGG GUC ACU CAU GAG AAG UAC UAC GCC CCG CGC CUC GAC CUA GAA AAA GAG AAG UUG UUG
GQ227789   AGG GUC ACU CAU GAG AAG UAC UAC GCC CCG CGC CUC GAC CUA GAA AAA GAG AAG UUG UUG
GQ287646   CGG GCC AUC CAU GAG AAG UAC UAC GCC CCG CGC CUC GAU CUC GAA AGG GAG AAG AUU CUA
FJ786263   CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAG AUG UUA
GQ287640   CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAG AUG UUA
GQ287644   CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAG AUG UUA
FJ786260   CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAA AUG UUA
FJ786261   CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAA AUG UUA
GQ287645   CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAA AUG UUA
GQ287647   CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAA AUG UUA
NC_003908  CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAA AUG UUA
X74892     CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAG AUG UUA
U01065     CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAG AUG UUA
GQ287641   CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAA AUG UUA
GQ287642   CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAA AUG UUA
GQ287643   CGG GCC GUC CAU GAG AAG UAU UAC GCC CCG CGC CUC GAU CUC GAA AGA GAG AAA AUG UUA
                     **  *** *   *   **  *** *** *** *** **   *  **      **  **   *   *
```

Sindbis and Aura viruses (SINV, AURAV; RefSeqs NC_001547 [derived from J02363], NC_003900 [derived from AF126284])

Notes:

1. Here the readthrough codon is UGA and the 3'-adjacent codon is CUA.
2. Longer loop region than in VEEV/WEEV/EEEV but, again, the positions of the two components of the predicted stem correspond to peaks in synonymous site conservation (main text Fig 1). Note that the position of the bulge in the 3' component of the stem is shifted by 1 nt in AURAV (NC_003900) relative to SINV (the other sequences).
3. Once again, there are conserved codons in the spacer region and just 3' of the 3' component of the predicted stem.
4. The SINV sequence U38305 lacks the stop codon.

An extract around the readthrough stop codon is shown below for the SINV/AURAV sequences used in the synonymous site conservation analysis (main text Fig 1), plus the additional partial sequence AY532326. Partial sequences AY532321, AY532322, AY532323, AY532324 and AY532325 (not shown) are locally identical to M69205.

```
                                                        ((( ((( ((( ((
NC_003900  GGA AAC AGA CAA UAU UGA CUA ACC GGG GUA GGU GGG UAC AUA UUC UCU UCU GAU ACA GGC
AF429428   AAG AAA GCG GAG UAC UGA CUA ACC GGG GUA GGU GGG UAC AUC UUC UCA ACG GAU ACU GGU
AF492770   AAG AAA ACU GAA UAU UGA CUA ACC GGG GUA GGU GGG UAU AUC UUC UCG ACU GAC ACG GGA
NC_001547  AGG AGG ACU GAA UAC UGA CUA ACC GGG GUA GGU GGG UAC AUA UUU UCG ACG GAC ACA GGC
M69205     AGG AGG ACC GAA UAC UGA CUA ACC GGG GUA GGU GGG UAC AUA UUU UCG ACG GAC ACA GGC
AY532326   AGG AGG ACC GAA UAC UGA CUA ACC GGG GUA GGU GGG UAC AUA UUU UCG ACG GAC ACA GGC
U38304     AGG AGG ACC GAA UAC UGA CUA ACC GGG GUA GGU GGG UAC AUA UUU UCG ACG GAC ACA GGC
AF103728   AAG CGG ACU GAA UAC UGA CUA ACC GGG GUA GGU GGG UAC AUA UUU UCG ACG GAU ACA GGG
AF103734   AGG AGG ACC GAA UAC UGA CUA ACC GGG GUA GGU GGG UAC AUA UUU UCG ACG GAC ACA GGC
U38305     AGG AGG ACC GAA UAC UGU CUA ACC GGG GUA GGU GGG UAC AUA UUU UCG ACG GAC ACA GGC
                     *  **  **  *** *** *** *** *** *** **  **  **  **   *  **  **  **

NC_003900  CCG GGG CAU CUG CAA CAA AAA UCA GUU AUU CAA AAC AGC ACC ACC GAG AUA CUA AUA GAA
AF429428   CCU GGA CAU UUG CAG AAG AAA UCA GUU CUG CAG AAC GUA CUU ACA GAA CCG UUG CUU GAG
AF492770   CCG GGU CAC CUC CAG AAG AAG UCG GUU CUA CAA AAC CAG CUU ACG GAA CCA ACU CUC GAG
NC_001547  CCU GGG CAC UUG CAA AAG AAG UCC GUU CUG CAG AAC CAG CUU ACA GAA CCG ACC UUG GAG
M69205     CCU GGG CAC UUG CAA AUG AAG UCC GUU CUG CAG AAC CAG CUU ACA GAA CCG ACC UUG GAG
AY532326   CCU GGG CAC UUG CAA AUG AAG UCC GUU CUG CAG AAC CAG CUU ACA GAA CCG ACC UUG GAG
U38304     CCU GGG CAC UUG CAA AUG GAG UCC GUU CUG CAG AAU CAG CUU ACA GAA CCG ACC UUG GAG
AF103728   CCA GGA CAU CUG CAA AAG AAG UCU GUC UUG CAG AAU CAA UUU UCC GAA CCG ACC UUG GAG
AF103734   CCU GGG CAC UUG CAA AAG AAG UCC GUU CUG CAG AAC CAG CUU ACA GAA CCG ACC UUG GAG
U38305     CCU GGG CAC UUG CAA AAG AAG UCC GUU CUG CAG AAC CAG CUU ACA GAA CCG ACC UUG GAG
           **  **  **   *  **       *  **  **   *  **  **           *  **           *  **

NC_003900  CGA AGU AGA CUA GAA AAA AUU CAU GCC CCU GUA CUG GAC CUA CAG AAA GAA GAG AUG CUG
AF429428   CGC AAU AAU CUA GAA CAA AUC UAU GCG CCG ACG UAC GAC CAG UCC AAA GAA GAG CUA CUA
AF492770   CGU AAU CAA UUA GAA CGA GUU UAU GCA CCC AGU CUU GAU GCC AAG AAA GAG GAA CUC UUG
NC_001547  CGC AAU GUC CUG GAA AGA AUU CAU GCC CCG GUG CUC GAC ACG UCG AAA GAG GAA CAA CUC
M69205     CGC AAU GUU CUG GAA AGA AUC UAC GCC CCG GUG CUC GAC ACG UCG AAA GAG GAA CAG CUC
AY532326   CGC AAU GUU CUG GAA AGA AUC UAC CCC CCG GUG CUC GAC ACG UCG AAA GAG GAA CAG CUC
U38304     CGC AAU GUU CUG GAA AGA AUC UAC GCC CCG GUG CUC GAC ACG UCG AAA GAG GAA CAG CUC
AF103728   CGU AAC GUG CUG GAA AAG AUA UAC GCU CCG ACG CUU GAU ACG UCG AAA GAA GAA CUA CUC
AF103734   CGC AAU GUU CUG GAA AGA AUC UAC GCC CCG GUG CUC GAC ACG UCG AAA GAG GAA CAG CUC
U38305     CGC AAU GUU CUG GAA AGA AUC UAC GCC CCG GUG CUC GAC ACG UCG AAA GAG GAA CAG CUC
           **  *        *  ***      *   *   *  **          **          *** **  **       *
                                  )))  )) ))) )))
NC_003900  AAG UGU AGG UAC CAG AUG UCA CCC ACC GUA GCC AAC AAG AGC AGG UAC CAG UCA CGG AAA
AF429428   AAA UUU AAG UAC CCA AUG AUA CCC ACC GAA GCC AAC AAG AGC AGG UAC CAA UCA CGG AAA
AF492770   AAA CUC AAG UAC CAA AUG AUG CCC ACC GAA GCC AAU AAA AGU AGG UAC CAG UCU AGA AAG
NC_001547  AAA CUC AGG UAC CAG AUG AUG CCC ACC GAA GCC AAC AAA AGU AGG UAC CAG UCU CGU AAA
M69205     AAA CUC AGG UAC CAG AUG AUG CCC ACC GAA GCC AAU AAA AGC AGG UAC CAG UCU CGA AAA
AY532326   AAA CUC AGG UAC CAG AUG AUG CCC ACC GAA GCC AAU AAA AGC AGG UAC CAG UCU CGA AAA
U38304     AAA CUC AGG UAC CAG AUG AUG CCC ACC GAA GCC AAC AAA AGC AGG UAC CAG UCU AGA AAA
AF103728   AAA UUU AGA UAC CAA AUG AUG CCC ACC GAA GCC AAU AAG AGC AGG UAC CAG UCC CGC AAA
AF103734   AAA CUC AGG UAC CAG AUG AUG CCC ACC GAA GCC AAC AAA AGC AGG UAC CAG UCU CGA AAA
```

```
U38305     AAA CUC AGG UAC CAG AUG AUG CCC ACC GAA GCC AAC AAA AGC AGG UAC CAG UCU CGA AAA
           **      *   *** *   ***     *** *** * * *** **  **  **  *** *** **  **   *  **
```

Other alphaviruses

The available alphavirus non-structural polyprotein coding sequences mostly fall into three major groups:

1. (VEEV+EEEV+WEEV) + (SINV+AURAV) = VEE/WEE/EEE complexes
2. (CHIKV+ONNV) + (SFV+MAYV+RRV+GETV) = SFV complex
3. SAV = salmonid alphaviruses

(Other alphaviruses - including Barmah Forest and Middelburg viruses - will not be discussed here due to lack of sequence data for informative comparative predictions.)

Group 3 members appear to lack stop codon readthrough (i.e. they have a sense codon at that position). Group 1 members nearly always have a readthrough site. Group 2 members are more variable with respect to readthrough - many species (CHIKV, SFV, GETV, ONNV) have some sequences with and some sequences without readthrough. For those species in which there appears to be a constant flux between having a stop codon and having a sense codon (it has been suggested that alternating replication cycles in insect and in vertebrate hosts provide opposing selective pressures; also passaging in cell culture can select for or against readthrough depending on the system; see main text for references) it is not unreasonable to suspect that the 3' structure, if any, will be present whether or not the stop codon is present in any particular sequence.

For viruses in group 2, however, we have found that the range of divergences between the available sequences is less useful for the synonymous site conservation analysis compared to the viruses in group 1. In general, therefore, we were unable to obtain independent support from conservation analysis for the following potential structures in SFV complex alphaviruses. This is in contrast to the readthrough associated structure predictions in the VEE/WEE/EEE complex alphaviruses and in tobra-, peclu, furo- and pomovirus RNA1 (see below), where the predicted structures are supported not only by many compensatory substitutions, but also by coinciding peaks in synonymous site conservation.

Ross River and Getah viruses (RRV, GETV; RefSeqs NC_001544 [derived from M20162], NC_006558 [derived from AY702913])

Notes on the nucleotide sequences:

1. Here the readthrough codon is UGA and the 3'-adjacent codon is CUA.
2. EF011023 and EU015062 lack the stop codon.
3. Although sequence inspection reveals a potential RNA stem, the synonymous site conservation analysis proved uninformative in this case.

```
                                                  ((( ((( (  (  ((( ((
AB032553   GAC UUU AAC CAG UCC UGA CUA GGC AGG GCU GGC GCG UAU AUC UUU UCG UCU GAC ACU GGC
NC_006558  GAC UUU AAC CAG UCC UGA CUA GGC AGG GCU GGC GCG UAU AUC UUU UCG UCU GAC ACU GGC
EF631998   GAC UUU AAC CAG UCC UGA CUA GGC AGG GCU GGC GCG UAU AUC UUU UCG UCU GAC ACU GGC
EF631999   GAC UUU AAC CAG UCC UGA CUA GGC AGG GCU GGC GCG UAU AUC UUU UCG UCU GAC ACU GGC
EU015063   GAC UUU AAC CAG UCC UGA CUA GGC AGG GCU GGC GCG UAU AUC UUU UCG UCU GAC ACU GGC
EU015062   GAC UUU AAC CAG UCC UGG CUA GGC AGG GCU GGC GCG UAU AUC UUU UCG UCU GAC ACU GGC
EF011023   GAC UUU AAC CAG UCC AGA CUA GGC AGG GCU GGC GCG UAU AUC UUU UCG UCU GAC ACU GGC
EU015061   GAC UUU AAC CAG UCC UGA CUA GGC AGG GCU GGC GCG UAU AUC UUU UCG UCU GAC ACU GGC
GQ433354   GAU UUU GAC CAA UUC UGA CUA GGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU ACC GGA
HM234643   GAU UUC GAC CAA UUC UGA CUA GGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU AUC GGA
NC_001544  GAU UUC GAC CAA UUC UGA CUA AGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU ACC GGA
M20539     GAU UUU GAC CAA UUC UGA CUA GGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU ACC GGA
DQ226993   GAU UUU GAC CAA UUC UGA CUA GGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU ACC GGA
GQ433355   GAU UUU GAC CAA UUC UGA CUA GGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU ACC GGA
GQ433356   GAU UUU GAC CAA UUC UGA CUA GGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU ACC GGA
GQ433357   GAU UUU GAC CAA UUC UGA CUA GGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU ACC GGA
GQ433358   GAU UUU GAC CAA UUC UGA CUA GGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU ACC GGA
```

```
GQ433359   GAU UUU GAC CAA UUC UGA CUA GGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU ACC GGA
GQ433360   GAU UUU GAC CAA UUC UGA CUA GGC AGA GCG GGG GCG UAC AUC UUC UCG UCU GAU ACC GGA
           **  **   ** **  * *  *  ***  ** **  **  **  *** **  *** **  *** *** **  *   **

AB032553   CCG GGU CAC CUA CAG CAG AGG UCC GUA AGG CAA CAU GAA UUG CCA UGC GAG ACU CUG UAU
NC_006558  CCG GGU CAC CUA CAA CAG AAG UCC GUA AGG CAA CAU GAA UUG CCA UGC GAG ACU CUG UAC
EF631998   CCG GGU CAU CUA CAG CAG AAG UCC GUA AGG CAA CAU GAA UUG CCA UGC GAG ACU CUG UAU
EF631999   CCG GGU CAC CUA CAG CAG AAG UCC GUA AGG CAA CAU GAA UUG CCA UGC GAG ACU CUG UAC
EU015063   CCG GGU CAC CUA CAA CAG AAG UCC GUA AGG CAA CAU GAA UUG CCA UGC GAG ACU CUG UAC
EU015062   CCG GGU CAC CUA CAA CAG AAG UCC GUA AGG CUA CAU GAA UUG CCA UGC GAG ACU CUG UAC
EF011023   CCG GGU CAC CUA CAG CAG AAG UCC GUA AGG CAA CAU GAA UUG CCA UGC GAG ACU CUG UAU
EU015061   CCG GGU CAC CUA CAG CAG AAG UCC GUA AGG CAA CAU GAA UUG CCA UGC GAG ACU CUG UAU
GQ433354   CCG GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCA UGC GAA AUG CUA UAC
HM234643   CCG GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCG UGC GAA AUG CUA UAC
NC_001544  CCG GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCG UGC GAA AUG CUA UAC
M20539     CCA GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCG UGC GAA AUG CUA UAC
DQ226993   CCA GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCG UGC GAA AUG CUA UAC
GQ433355   CCA GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCG UGC GAA AUG CUA UAC
GQ433356   CCA GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCG UGC GAA AUG CUA UAC
GQ433357   CCA GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCG UGC GAA AUG CUA UAC
GQ433358   CCA GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCU UGC GAA AUG CUA UAC
GQ433359   CCA GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCG UGC GAA AUG CUA UAC
GQ433360   CCA GGG CAC UUA CAA CAG AAG UCA GUA CGG CAA CAC GCA CUA CCG UGC GAA AUG CUA UAC
           **  **  **   ** **  *** * * **  ***  ** * * **  * *  *  **  *** **  *   **  **

                                  ))) )))  )) ))) ))
AB032553   GCC CAU GAA GAC GAA CGC AUA UAC CCG CCG GCA UUU GAC GGA GAG AAA GAG AAG AUA CUC
NC_006558  GCC CAU GAA GAC GAA CGC AUA UAC CCG CCG GCA UUU GAC GGA GAG AAA GAA AAG GUA CUC
EF631998   GCC CAU GAA GAC GAA CGC AUA UAC CCG CCG GCA UUC GAC GGA GAG AAA GAG AAG GUA CUC
EF631999   GCC CAU GAA GAC GAA CGC AUA UAC CCG CCG GCA UUC GAC GGA GAG AAA GAA AAG GUA CUC
EU015063   GCC CAU GAA GAC GAA CGC AUA UAC CCG CCG GCA UUU GAC GGA GAG AAA GAA AAG GUA CUC
EU015062   GCC CAU GAA GAC GAA CGC AUA UAC CCG CCG GCA UUU GAC GGA GAG AAA GAA AAA GUA CUC
EF011023   GCC CAU GAA GAC GAA CGC AUA UAC CCG CCG GCA UUU GAC GGA GAG AAA GAG AAG GUA CUC
EU015061   GCC CAU GAA GAC GAA CGC AUA UAC CCG CCG GCA UUU GAC GGA GAG AAA GAG AAG GUA CUC
GQ433354   GCC CAC GAG GAA GAA CGG ACG UAC CCC CCC GCA CUG GAU GAG GCC AGG GAG AAA CUG CUG
HM234643   GCC CAC GAG GAG GAA AGG ACG UAC CCC CCC GCA CUG GAU GAG GCU AGG GAG AAA CUG CUG
NC_001544  GCC CAC GAG GAA GAA CGG ACG UAC CCC CCC GCA CUG GAU GAG GCU AGG GAG AAA CUG CUG
M20539     GUC CAC GAG GAA GAA CGG ACG UAC CCC CCC GCA CUG GAU GAG GCC AGG GAG AAA CUG CUG
DQ226993   GUC CAC GAG GAA GAA CGG ACG UAC CCC CCC GCA CUG GAU GAG GCC AGG GAG AAA CUG CUG
GQ433355   GUC CAC GAG GAA GAA CGG ACG UAC CCC CCC GCA CUG GAU GAG GCC AGG GAG AAA CUG CUG
GQ433356   GUC CAC GAG GAA GAA CGG ACG UAC CCC CCC GCA CUG GAU GAG GCC AGG GAG AAA CUG CUG
GQ433357   GUC CAC GAG GAA GAA CGG ACG UAC CCC CCC GCA CUG GAU GAG GCC AGG GAG AAA CUG CUG
GQ433358   GUC CAC GAG GAA GAA CGG ACG UAC CCC CCC GCA CUG GAU GAG GCC AGG GAG AAA CUG CUG
GQ433359   GUC CAC GAG GAA GAA CGG ACG UAC CCC CCC GCA CUG GAU GAG GCC AGG GAG AAA CUG CUG
GQ433360   GUC CAC GAG GAA GAA CGG ACG UAC CCC CCC GCA CUG GAU GAG GCC AGG GAG AAA CUG CUG
           * * **  **  **  ***  *  *   *** **  **  ***  *  **  *   *   *   **  **   *  **
```

Semliki Forest and Mayaro viruses (SFV, MAYV; RefSeqs NC_003215 [derived from X04129], NC_003417 [derived from AF237947])

Notes on the nucleotide sequences:

1. Here the readthrough codon is UGA and the 3'-adjacent codon is CUA.
2. SFV NC_003215 and AJ251359 lack the stop codon.
3. Additional SFV sequences AY112987, DQ189079, DQ189080, DQ189081, DQ189082, DQ189083, DQ189084, DQ189085 and DQ189086 are locally identical to AJ251359.
4. Predicted structure not so stable for MAYV (underlined)

```
                                          ((( ((( ((( (    [[ [[[ [[[ [[[
Z48163     UUC GAC GAU GUC CUG UGA CUA GGC CGC GCG GGU GCA UAU AUU UUC UCC UCG GAC ACU GGC
EU350586   UUC GAC GAC GUC CUG UGA CUA GGC CGC GCG GGU GCA UAU AUU UUC UCC UCG GAC ACU GGC
AJ251359   UUC GAC GAC GUC CUG CGA CUA GGC CGC GCG GGU GCA UAU AUU UUC UCC UCG GAC ACU GGC
NC_003215  UUC GAC GAC GUC CUG CGA CUA GGC CGC GCG GGU GCA UAU AUU UUC UCC UCG GAC ACU GGC
DQ001069   AUC AGC AAC UCA UCC UGA CUA GGC CGA GCG GGG GCU UAU AUC UUC UCA UCA GAC GUC GGU
```

```
NC_003417  AUU AGU AAU UCA UCU UGA CUA GGC CGU GCG GGG GCC UAU AUU UUC UCA UCC GAC GUC GGU
             *       *           ** *** *** **  *** **  **  *** **  *** **  **  ***     **

Z48163     AGC GGA CAU UUA CAA CAA AAA UCC GUU AGG CAA CAC AAU CUC CAG UGU GCA CAA CUG GAU
EU350586   AGC GGA CAU UUA CAA CAA AAA UCC GUU AGG CAG CAC AAU CUC CAG UGC GCA CAA CUG GAU
AJ251359   AGC GGA CAU UUA CAA CAA AAA UCC GUU AGG CAG CAC AAU CUC CAG UGC GCA CAA CUG GAU
NC_003215  AGC GGA CAU UUA CAA CAA AAA UCC GUU AGG CAG CAC AAU CUC CAG UGC GCA CAA CUG GAU
DQ001069   CCA GGG CAU CUG CAA CAG AAA UCA GUG AGA CAG CAC GAU CUA GAG GUG CCG AUU AUG GAU
NC_003417  CCA GGG CAC CUG CAA CAG AAA UCA GUG AGG CAG CAU GAC UUA GAG GUG CCG AUU AUG GAU
               **  **   *  *** **  *** **  **  **  **  **   *   *   **      *       ** ***
                 ] ]]] ]]] ]]] ]   ))) ))) ))) )
Z48163     GCG GUC GAG GAG GAG AAA AUG UAC CCG CCG AAA UUG GAU ACU GAG AGG GAG AAG CUG UUG
EU350586   GCG GUC GAG GAG GAG AAA AUG UAC CCG CCA AAA UUG GAU ACU GAG AGG GAG AAG CUG UUG
AJ251359   GCG GUC GAG GAG GAG AAA AUG UAC CCG CCA AAA UUG GAU ACU GAG AGG GAG AAG CUG UUG
NC_003215  GCG GUC CAG GAG GAG AAA AUG UAC CCG CCA AAA UUG GAU ACU GAG AGG GAG AAG CUG UUG
DQ001069   CGC GUA GUC GAG GAG AAA GUC UAC CCG CCG AAA CUA GAU GAG GCA AAA GAG AAA CAG CUG
NC_003417  CGU GUG AUU GAG GAA AAG GUC UAC CCG CCU AAA UUA GAU GAG GCA AAA GAG AAA CAG CUA
               **      *** **  **   *  *** *** **  ***  *  ***     *   *   *** **  * *  *
```

Chikungunya and O'nyong-nyong viruses (CHIKV, ONNV; RefSeqs NC_004162 [derived from AF369024], NC_001512 [derived from M20303])

Notes:

1. This alignment shows a selection of sequences with the stop codon.
2. Here the readthrough codon is UGA and the 3'-adjacent codon is CUA. 5'-adjacent conserved UUA.
3. The potential structure is not so convincing as in VEEV/EEEV/WEEV - longer spacer, shorter stem, no supporting evidence from compensatory substitutions, and disrupted in ONNV (underlined).

```
                                                    ( ((( ((( ((
AY726732   ACG GAC GAC GAA UUA UGA CUA GAU AGG GCA GGU GGG UAC AUA UUC UCA UCU GAC ACC GGC
EU703761   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACU GGU
EU703760   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACU GGU
EU703759   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACU GGU
EF027141   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCA GAC ACU GGU
EF452494   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACU GGU
EF452493   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACU GGU
EF027140   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACU GGU
L37661     ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACU GGU
EU564335   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
EU244823   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
EF210157   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
EF027138   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
EF027137   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
EF027136   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
EF027135   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
EF027134   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
EF012359   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
DQ443544   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
AM258994   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
AM258993   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
AM258991   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
AM258990   ACG GAC GAC GAG UUA UGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
AF079456   ACA GAU GAA GAG UUA UGA CUA GAC AGA GCA GGG GGU UAC AUA UUC UCC UCU GAC ACU GGU
           **  **  **  **  *** *** *** **  **  *** **  **  **  *** *** **  **  *** **  **

AY726732   CCC GGC CAC CUG CAA CAG AGG UCU GUC CGU CAG ACA GUA CUG CCG GUA AAU ACC UUG GAG
EU703761   CCA GGC CAU UUA CAA CAG AAG UCG GUA CGC CAG UCA GUG CUG CCG GUA AAC ACC CUG GAG
EU703760   CCA GGC CAU UUA CAA CAG AAG UCG GUA CGC CAG UCA GUG CUG CCG GUA AAC ACC CUG GAG
EU703759   CCA GGC CAU UUA CAA CAG AAG UCG GUA CGC CAG UCA GUG CUG CCG GUA AAC ACC CUG GAG
EF027141   CCA GGU CAU UUA CAA CAG AAG UCG GUA CGC CAG UCA GUG CUG CCG GUA AAC ACC CUG GAG
EF452494   CCA GGC CAU UUA CAA CAG AAG UCG GUA CGC CAG UCA GUG CUG CCG GUA AAC ACC CUG GAG
```

```
EF452493   CCA GGC CAU UUA CAA CAG AAG UCG GUA CGC CAG UCA GUG CUG CCG GUA AAC ACC CUG GAG
EF027140   CCA GGU CAU UUA CAA CAG AAG UCG GUA CGC CAG UCA GUG CUG CCG GUA AAC ACC CUG GAG
L37661     CCA GGC CAU UUA CAA CAG AAG UCG GUA CGC CAG UCA GUG CUG CCG GUA AAC ACC CUG GAG
EU564335   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
EU244823   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC UUG GAG
EF210157   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
EF027138   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
EF027137   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
EF027136   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
EF027135   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
EF027134   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
EF012359   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
DQ443544   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
AM258994   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
AM258993   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
AM258991   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
AM258990   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
AF079456   CAA GGU CAU CUA CAG CAA AAA UCA GUA CGU CAA ACG ACG CUA CCG GUA AAC AUU GUU GAA
           *   **  **   *  **  **  *   **  **  **  **   *       **  *** **  **  *    *  **
                                          ))) ))) )))
AY726732   GAA GUU CAG GAG GAG AAA UGU UAC CCA CCU AAG UUG GAU GAA GUG AAA GAG CAG UUG UUA
EU703761   GAA GUU CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA UUA AAG GAG CAA CUA CUA
EU703760   GAA GUU CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA UUA AAG GAG CAA CUA CUA
EU703759   GAA GUU CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA UUA AAG GAG CAA CUA CUA
EF027141   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA UUA AAG GAG CAA CUA CUA
EF452494   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA UUA AAG GAG CAA CUA CUA
EF452493   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA UUA AAG GAG CAA CUA CUA
EF027140   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA UUA AAG GAG CAA CUA CUA
L37661     GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA UUA AAG GAG CAA CUA CUA
EU564335   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
EU244823   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
EF210157   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
EF027138   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
EF027137   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
EF027136   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
EF027135   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
EF027134   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
EF012359   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
DQ443544   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
AM258994   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
AM258993   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
AM258991   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
AM258990   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
AF079456   GAG GUC CAC GAA GAG AAA UGC UAC CCA CCU AAA UUG GAU GAG AUC AAA GAG CAA CUC UUA
           **  **  **  **  *** **  **  *** *** *** **   ** *** **      **  *** **   *   **
```

This alignment shows a selection of CHIKV and ONNV (underlined) sequences without a stop codon. Again the potential structure is disrupted in ONNV.

```
                                                 ( (((  ((( ((
AF490259   ACG GAC GAC GAG UUA CGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACU GGU
EF027139   ACG GAC GAC GAG UUA CGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
EU037962   ACG GAC GAC GAG UUA CGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
EU564334   ACA GAC --- --- --- AGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACC GGU
AF369024   ACG GAC GAC GAG UUA CGA CUA GAC AGG GCA GGU GGG UAU AUA UUC UCG UCG GAC ACU GGU
NC_001512  ACA GAC GAA GAG UUA CGA CUA GAC AGA GCA GGG GGU UAC AUA UUC UCC UCU GAC ACU GGU
AF079457   ACA GAU GAA GAG UUA CGA CUA GAC AGA GCA GGG GGU UAC AUA UUC UCC UCU GAC ACU GGC
           **  **               ** *** *** **  *** **  **  **  *** *** **  **  *** **  **

AF490259   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
EF027139   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
EU037962   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
EU564334   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
```

```
AF369024   CCA GGU CAU UUA CAA CAG AAG UCA GUA CGC CAG UCA GUG CUG CCG GUG AAC ACC CUG GAG
NC_001512  CAA GGU CAU CUA CAG CAA AAA UCA GUA CGU CAA ACG ACG CUA CCG GUA AAC AUU GUU GAA
AF079457   CAA GGU CAU CUA CAG CAA AAA UCA GUA CGU CAA ACG ACG CUA CCA GUA AAC AUU GUU GAA
           * * *** ***  ** **  **  **  *** *** **  **   *    * **  **  **  *** *    *  **
                                          ))) ))) )))
AF490259   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA CUA
EF027139   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
EU037962   GAA GUC CAC GAG GAG NAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
EU564334   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA UUA
AF369024   GAA GUC CAC GAG GAG AAG UGU UAC CCA CCU AAG CUG GAU GAA GCA AAG GAG CAA CUA CUA
NC_001512  GAG GUC CAC GAA GAG AAA UGC UAU CCA CCU AAA UUG GAU GAG AUC AAA GAG CAA CUA CUA
AF079457   GAG GUC CAC GAA GAG AAA UGC UAU CCA CCU AAA UUG GAU GAG ACC AAA GAG CAA CUC CUA
           **  *** *** **  ***  *  **  **  *** *** **   ** *** **      **  *** *** **   **
```

Tobra/Peclu/Furo/Pomovirus RNA1 (replicase)

Notes on the nucleotide sequences:

1. The readthrough stop codon is indicated in red or pink. Predicted stems are indicated in yellow (bulgeless base stem) and green or pale orange (extension stem). The loop region can also fold but the fold is less highly conserved. Single substitutions that preserve the predicted base-pairings (e.g. G-C to G-U or G-U to A-U) are indicated with pale blue. Compensatory substitutions (paired substitutions that preserve the predicted base-pairings) are indicated with pink or orange. Within the yellow stem, cross-family variations are indicated but, within the green stem (whose position varies between different genera), only cross-genus variations are indicated.
2. Completely conserved 'A' at -1 even though the -1 amino acid is not conserved.
3. High conservation in the 'spacer' sequence (i.e. between the stop codon and the predicted stem-loop) and also just 3' of the predicted stem-loop, e.g. the 3'-adjacent nearly conserved GGG codon.
4. The predicted stems are supported by large numbers of compensatory substitutions and, despite these variations, there is also very significant enhanced synonymous site conservation within the stems relative to the mean for the entire coding sequence (main text Fig 4).

RefSeqs:

NC_002036	Pea early-browning virus	Tobravirus	(derived from X14006)
NC_002041	Soil-borne wheat mosaic virus	Furovirus	(derived from L07937)
NC_002351	Soil-borne cereal mosaic virus	Furovirus	(derived from AJ132576)
NC_002358	Oat golden stripe virus	Furovirus	(derived from AJ132578)
NC_002359	Chinese wheat mosaic virus	Furovirus	(derived from AJ012005)
NC_003669	Pepper ringspot virus	Tobravirus	(derived from L23972)
NC_003672	Peanut clump virus	Pecluvirus	(derived from X78602)
NC_003723	Potato mop-top virus	Pomovirus	(derived from AJ238607)
NC_003805	Tobacco rattle virus	Tobravirus	(derived from AF166084)
NC_004014	Sorghum chlorotic spot virus	Furovirus	(derived from AB033691)
NC_004423	Broad bean necrosis virus	Pomovirus	(derived from D86636)
NC_004729	Indian peanut clump virus	Pecluvirus	(derived from X99149)
NC_003510	Beet virus Q	Pomovirus	(derived from AJ223596)
NC_003520	Beet soil-borne virus	Pomovirus	(derived from Z97873)

An extract around the readthrough stop codon is shown below for the sequences used in the synonymous site conservation analysis (main text Fig 4), plus the additional partial sequence Z97358. The alignment is gap-free in this region.

```
                                                  ( ((( ((( ((( (([ [[[ [[[ [[[ [
AF034622   ACU GAG ACC GUC UUA UGA CGG UUU CGG UCU AGG UUU GAU GUU UUU AGA CAU CAU GAA GGG
D00155     ACC GAG ACC GUC UUA UGA CGG UUU CGG UCU AGG UUU GAU GUU UUU AGA CAU CAU GAA GGG
X06172     ACC GAG ACC GUC UUA UGA CGG UUU CGG UCU AGG UUU GAU GUU UUU AGA CAU CAU GAA GGG
AF314165   ACU GAG ACC GUC UUA UGA CGG UUU CGG UCU AGG UUU GAU GUC UUU AGA CAU CAU GAA GGG
AF406990   ACU GAG ACC GUC UUA UGA CGG UUU CGG UCU AGG UUU GAU GUC UUU AGA CAU CAU GAA GGG
NC_003805  ACU GAG ACC GUC UUA UGA CGG UUU CGG UCU AGG UUU GAU GUU UUU AGA CAU CAU GAA GGG
Z97358     ACU GAG ACC GUC UUA UGA CGG UUU CGG UCU AGG UUU GAU GUC UUU AGA CAU CAU GAA GGG
AJ586803   ACU GAG ACC GUC UUA UGA CGG UUU CGG UCU AGG UUU GAU GUC UUC AGA CAU CAC GAA GGG
GQ903771   ACA GAA ACC GUC UUA UGA CGG UUU CGG UCU AGG UUU GAU GUC UUC AGA CAU CAU GAA GGG
NC_003669  ACU GAU GCU GCC UUA UGA CGG UGU CGG UCG AGA UUC GAU UCC UUU UCA CAC CAU GAU GGC
NC_002036  GAA GAU GCU AUG AAA UGA CGG UGU CGG UCA AGG UUC GAU GUC UUU AGA CAC CAU GAU GGU
NC_003672  UCA GAA CAG ACC AAA UGA CGG UUU GGG UCA AAG UUU GAG UCC AUU GUU GUA GCA GAU AGG
NC_004729  UCU GAG CAG ACC AAA UGA CGG UUU GGG UCA AAG UUU GAG UCC AUC GUG GUA GCA GAU AGA
AB299271   CCC GAA UUC GAC AAA UGA CGG UUU GGG UCG AAA UAU GAG UCC AUC UUA AUA UGC GAU AGA
AJ271838   CCC GAA UUC GAC AAA UGA CGG UUU GGG UCG AAA UAU GAG UCC AUC UUA AUA UGC GAU AGA
NC_002359  CCU GAA UUC GAC AAA UGA CGG UUU GGG UCG AAA UAU GAG UCC AUC UUA AUA UGU GAU AGA
NC_002041  UCU GAG CUU ACU AAA UGA CGG UUU GGG UCG AAG UAU GAG UCU AUU UUG AUA UGU GAU AGA
AF146278   UCU GAG UUG GUG AAA UGA CGG UUU GGG UCG AAG UAU GAG UCC AUA UUG AUU UGU GAU AGG
AF146279   UCU GAG UUG GUG AAA UGA CGG UUU GGG UCG AAG UAU GAA UCC AUA UUG AUU UGU GAU AGG
AF146280   UCU GAG UUG GUG AAA UGA CGG UUU GGG UCG AAG UAU GAA UCC AUA UUG AUU UGU GAU AGG
NC_002351  UCU GAG UUG GCG AAA UGA CGG UUU GGG UCG AAG UAU GAA UCC AUA UUG AUU UGC GAU AGG
AJ252151   UCU GAG UUG GCG AAA UGA CGG UUU GGG UCG AAG UAU GAA UCC AUA UUG AUU UGC GAU AGG
AJ298068   UCU GAG UUG GCG AAA UGA CGG UUU GGG UCG AAG UAU GAA UCC AUA UUG AUU UGC GAU AGG
AB033689   GUG GAG AAC GGG AAA UGA CGG UUU GGG UCG AAG UAU GAA UCA AUU UUG ACC UGU GAC CGU
NC_002358  CUC GAG AAU CAG AAA UGA CGG UUU GGG UCG AAA UAC GAA ACU AUC UUA AUA UGU GAU AGA
NC_004014  GGU GAG CAU ACC AAA UGA CGG UUU GGG UCG AAG UUU GAC UCU GUA GAU GUG UGU GAU AGG
NC_003723  GAG GAA GGU GUG AAA UGA CGC CAG GCG UCA AAG UAU GAG GAG AUA AUG GUG UCC GAU AAU
AY196959   GAG GAA GGU GUG AAA UGA CGC CAG GCG UCA AAG UAU GAG GAG AUA AUG GUG UCC GAU AAU
NC_004423  GAG GAA GGU CCU AAA UGA CGG UGU GGG UCG AAA UUU GAA UCU AUA CAA GUG AUG GAU AGG
NC_003510  UCG GAA UCU GUU CAA UAA CGG UUU GGG UCA AAG UAC GAA GAA UUC GUA UUU UCG GAU AUU
EF545138   ACG GAG AGU ACU CAA UAA CGG UGU GGG UCG AAG UAC GAG UCG AUA ACC GUU UCC GAU CGU
EF545139   ACG GAG AGU ACU CAA UAA CGG UGU GGG UCG AAG UAC GAG UCG AUA ACC GUU UCC GAU CGU
FN386612   ACG GAA AGU ACU CAA UAA CGG UGU GGG UCG AAG UAC GAG UCG AUA ACC GUU UCC GAU CGU
NC_003520  ACG GAG AGU ACU CAA UAA CGG UGU GGG UCG AAG UAC GAG UCG AUA ACC GUU UCC GAU CGU
FJ971717   ACG GAG AAU ACU CAA UAA CGG UGU GGG UCG AAG UAC GAG UCG AUA ACC GUU UCC GAU CGU
               **             * * * **       * **  *   *   **      *              **


AF034622   CCU UGC GCC GUU CCA GAU UCA GGU ACG AUU ACG GAC UUG GAG AUG UGG UAC GAC GCU UUG
D00155     CCU UGC GCC GUU CCA GAU UCA GGU ACG AUU ACG GAC UUG GAG AUG UGG UAC GAC GCU UUG
X06172     CCU UGC GCC GUU CCA GAU UCA GGU ACG AUU ACG GAC UUG GAG AUG UGG UAC GAC GCU UUG
AF314165   CCU UGC GCC GUU CCA GAU UCA GGU ACG AUU ACG GAC UUG GAG AUG UGG UAC GAC GCU UUG
AF406990   CCU UGC GCC GUU CCA GAU UCA GGU ACG AUU ACG GAC UUG GAG AUG UGG UAC GAC GCU UUG
NC_003805  CCU UGC GCC GUU CCA GAU UCA GGU ACG AUU ACG GAC UUG GAG AUG UGG UAC GAC GCU UUG
Z97358     CCU UGC GCC GUU CCA GAU UCA GGU ACG AUU ACG GAC UUG GAG AUG UGG UAC GAC GCU UUG
AJ586803   CCU UGC GCC GUU CCA GAU UCA GGU ACG AUU ACG GAC UUG GAG AUG UGG UAC GAC GCU UUG
GQ903771   CCU UGC GCU GUC CCA GAU UCA GGU ACG AUU ACG GAC UUG GAA AUG UGG UAC GAC GCU UUG
NC_003669  CCG UGU UCG GUU CCA GAC ACU GGA ACU ACU GUC GAU AUA CAA ACA UGG UAU GAU GAU AUU
NC_002036  CCU UGU UCA GUA CCU GAU UCU GGA ACU CCU GUG GAU AUC CAG AGU UGG UAU GAU GAA UUA
NC_003672  AAA GUU AUG GCU CCU GAU GUA GGA GAC CUA GUA ACU AUC CAA GAU CUU UAC GAC AGG ACC
NC_004729  AAG GUC AUG GCC CCU GAU GUU GGG GAU CUU GUG ACC AUU CAA GAC UUG UAC GAU CGG ACU
AB299271   GAG GUA AGG GUA CCU GAU GUA GGA ACA CCU GCA AUU AUG CAG GAC UUC UAC GAC AAG GUG
AJ271838   GAG GUA AGG GUA CCU GAU GUA GGA ACA CCU GCA AUU UUG CAG GAC UUC UAC GAC AAG GUG
NC_002359  GAG GUA AGG GUA CCC GAC GUA GGA ACA CCU GCA AUU UUG CAG GAC UUC UAC GAC AGG GUG
NC_002041  GAG GUA AGG GUA CCU GAC GUA GGA ACC CCU GUG AUU AUA CAG GAU UUC UAU GAC AGG GUA
AF146278   GAA GUG AGG GUA CCC GAU GUG GGA AAU UCU GUU AUU UUG CAG GAU UUU UAC GAC AGG GUU
AF146279   GAA GUG AGG GUA CCC GAU GUG GGA AAU UCU GUU AUU UUG CAG GAU UUU UAC GAC AGG GUU
AF146280   GAA GUG AGG GUA CCC GAU GUG GGA AAU UCU GUU AUU UUG CAG GAU UUU UAC GAC AGG GUU
NC_002351  GAA GUG AGG GUU CCC GAC GUG GGA AAU UCU GUU AUU UUG CAG GAU UUU UAC GAU AGG GUU
```

```
AJ252151   GAA GUG AGG GUU CCC GAC GUG GGA AAU UCU GUU AUU UUG CAG GAU UUU UAC GAU AGG GUU
AJ298068   GAA GUG AGG GUU CCC GAC GUG GGA AAU UCU GUU AUU UUG CAG GAU UUU UAC GAU AGG GUU
AB033689   GAA GUA AAA GUU CCU GAU ACU GGU GAU UUA GUC ACC AUU CAG GAU UUU UAU GAU CGG GUU
NC_002358  GAA GUG AAG GUU CCG GAC ACA GGA GAU CUU GUA AUG AUU CAA GAU UUU UAC GAU CGA GUG
NC_004014  GAA GUU AAA GUG CCA GAC GUA GGA GAA UUG CAA ACA AUU CAA GAC UUC UAU GAC CGC ACU
NC_003723  GCU GUG ACU GUC CCU GAU GUG GGU AAU UUG GUU GAU UUG CAA GAG AUG UAU GAC AUA GCA
AY196959   GCU GUG ACU GUC CCU GAU GUG GGU AAU UUG GUU GAU UUG CAA GAG AUG UAU GAC AUA GCA
NC_004423  AAA GUC UGU GUA CCU GAU GUC GGU AAC AUU ACG AUU AUG CAG GAU UUU UUC GAC GUU GUA
NC_003510  GCU ACA AAA GUA CCU GAU GUA GGU GAA GCU CAA GCU UUA CAA GAA AUG UAU GAU AUA GCG
EF545138   CCG GUA AGU GUU CCU GAC GUU GGA GCU GUU GAA GAU UUG CAG GUU AUG UAU GAC AUU GCU
EF545139   CCG GUA AGU GUU CCU GAC GUU GGA GCU GUU GAA GAU UUG CAG GUU AUG UAU GAC AUU GCU
FN386612   CCG GUA AGU GUU CCC GAC GUU GGA GCU GUU GAA GAU UUG CAG GUU AUG UAU GAC AUU GCU
NC_003520  CCG GUA AGU GUU CCU GAC GUU GGA GCC GUU GAA GAU UUG CAG GUU AUG UAU GAC AUU GCU
FJ971717   CCG GUA AGU GUU CCA GAC GUU GGA GCU GUU GAA GAU UUG CAG GUU AUG UAU GAC AUU GCU
                     *   **  **      **                      *   *          *   **

                              ]] ]]] ]]] ]]] ))) ))) ))) )))
AF034622   UUU CCA GGA AAU UCG UUA AGA GAC UCA AGC CUA GAC GGG UAU UUG GUG GCA ACG ACU GAU
D00155     UUU CCG GGA AAU UCG UUA AGA GAC UCA AGC CUA GAC GGG UAU UUG GUG GCA ACG ACU GAU
X06172     UUU CCG GGA AAU UCG UUA AGA GAC UCA AGC CUA GAC GGG UAU UUG GUG GCA ACG ACU GAU
AF314165   UUU CCG GGA AAU UCG UUA AGA GAC UCA AGC CUA GAC GGG UAU UUG GUG GCA ACG ACU GAU
AF406990   UUU CCG GGA AAU UCG UUA AGA GAC UCA AGC CUA GAC GGG UAU UUG GUG GCA ACG ACU GAU
NC_003805  UUU CCG GGA AAU UCG UUA AGA GAC UCA AGC CUA GAC GGG UAU UUG GUG GCA ACG ACU GAU
Z97358     UUU CCG GGA AAU UCG UUA AGA GAC UCA AGC CUA GAC GGG UAU UUG GUG GCA ACG ACU GAU
AJ586803   UUU CCG GGA AAU UCG AUA AGA GAC UCA AGC CUA GAC GGG UAC UUG GUG GCA ACG ACU GAU
GQ903771   UUU CCU GGA AAU UCA UUG AGA GAC UCA AGC CUA GAC GGG UAU CUG GUG GCA ACG ACU GAC
NC_003669  UUU CCG GGA AAU UCU GUG AGG GAU UCG AGU CUC GAC GGG UAC AUG GUC AGU ACC ACU GAC
NC_002036  UUU CCA UUG AAU UCG GUC AGA GAC ACG AGC CUU GAC GGG UAU AUG GUG AAU ACA GAG GAC
NC_003672  UUU CCA GGA AAU UCG ACU UUG GAC UCA AGC UUU GAC GGG UAC ACC GUG GCU UCU UCU GAC
NC_004729  UUC CCU GGU AAU UCC ACU UUG GAU UCA AGC UUU GAC GGG UAU ACU GUU GCU UCU UCU AAC
AB299271   UUA CCU GGA AAU UCA ACG AUG GAU UCA CAU UUC GAC GGG UAU GAA GUU UCU ACA UCU GAC
AJ271838   UUA CCU GGA AAU UCA ACG AUG GAU UCA CAU UUC GAC GGG UAU GAA GUU UCU ACU UCU GAC
NC_002359  UUU CCU GGA AAU UCA ACG AUG GAU UCA CAU UUC GAC GGG UAU GAA GUU UCU ACU UCU GAC
NC_002041  UUA CCU GGA AAU UCC ACA AUG GAU UCA CAC UUC GAC GGG UAC GAA GUU UCU ACU UCC GAC
AF146278   CUA CCU GGA AAU UCU ACA AUG GAU UCA CAC UUC GAC GGG UAU GAA GUG GCA ACG ACA GAC
AF146279   CUA CCU GGA AAU UCU ACA AUG GAU UCA CAC UUC GAC GGG UAU GAA GUG GCG ACG ACA GAC
AF146280   CUA CCU GGA AAU UCU ACA GUG GAU UCA CAC UUC GAC GGG UAU GAA GUG GCG ACG ACA GAC
NC_002351  CUA CCU GCA AAU UCU ACA AUG GAU UCA CAC UUC GAC GGG UAU GAA GUG GCA ACG ACA GAC
AJ252151   CUA CCU GGA AAU UCU ACA AUG GAU UCA CAC UUC GAC GGG UAU GAA GUG GCA ACG ACA GAC
AJ298068   CUA CCU GGA AAU UCU ACA AUG GAU UCA CAC UUC GAC GGG UAU GAA GUG GCA ACG ACA GAC
AB033689   UUU CCA GGU AAU UCA ACG GUG GAU UCU UAC UUC GAC GGG UAU GAG GUU GCG ACA ACU GAU
NC_002358  UUU CCG GGA AAU UCU ACG AUG GAA UCG UAU UUC GAC GGG UAU GAG GUA GCA ACU GGU GAC
NC_004014  UUU CCC GGC AAC UCC ACU AUA GAG ACA UUC UUC GAC GGG UAC GAA GUC GCA ACU GGU GGU
NC_003723  UUU CCC GGG AAU UCC AUU GUU GAC ACU UAC UUU GAC GGU UAU GAA GUG GCC ACA GGA GGU
AY196959   UUU CCC GGG AAU UCC AUU AUU GAC ACU UAC UUU GAC GGU UAU GAA GUG GCC ACA GGA GGU
NC_004423  UUU CCU GGA AAU UCU UGU AUU GAU ACA UUU UUC GAC GGG UAU GAA GUU GCG ACU GGU GGU
NC_003510  UUU CCG GGA AAU UCU ACG AUU GAC ACG UGC UUU GAC GGG UAU GAA GUU GCU ACA GGU GGU
EF545138   UUU CCU GGC AAU UCG GUU AUU GAG ACG UAU UUC GAC GGG UAC GAU GUU GCG ACG GGA GGU
EF545139   UUU CCU GGC AAU UCG GUU AUU GAG ACG UAU UUC GAC GGG UAC GAU GUU GCG ACG GGA GGU
FN386612   UUU CCU GGC AAU UCG GUU AUU GAG ACG UAU UUC GAC GGG UAC GAU GUU GCG ACG GGA GGU
NC_003520  UUU CCU GGC AAU UCG GUU AUU GAG ACG UAU UUC GAC GGG UAC GAU GUU GCU ACG GGA GGU
FJ971717   UUU CCU GGC AAU UCG GUU AUU GAG ACG UAU UUC GAC GGG UAC GAU GUU GCU ACG GGA GGU
            *  **      **  **          **   *       *  *** **  **      **       *
```

Furovirus RNA2 (CP extension)

Notes on the nucleotide sequences:

1. The readthrough codon is UGA and is followed by CGG, except for the highly divergent NC_004015 (mean 45% nt identity to other furoviruses for the readthrough CDS) which has UGA followed by UGG.
2. See notes on Tobra/Peclu/Furo/Pomovirus RNA1 (above) for a description of the colour coding used here.
3. Again, the predicted stem corresponds to tandem synonymous site conservation peaks (main text Fig 4).

RefSeqs:

NC_002042	Soil-borne wheat mosaic virus	Furovirus	(derived from L07938)
NC_002330	Soil-borne cereal mosaic virus	Furovirus	(derived from AJ132577)
NC_002356	Chinese wheat mosaic virus	Furovirus	(derived from AJ012006)
NC_002357	Oat golden stripe virus	Furovirus	(derived from AJ132579)
NC_004015	Sorghum chlorotic spot virus	Furovirus	(derived from AB033692)

Extracts around the readthrough stop codon are shown below for the furovirus RNA2 sequences used in the synonymous site conservation analysis (main text Fig 4). Additional relevant sequences that were omitted for various reasons (e.g. incomplete coverage of the full coding sequence) are shown below this alignment.

```
                                                  (( ((( ((( (([ [[[ [
AB033690   UGG GAA GGU ACA AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGA GCC CAG AUU GUC GAA
AJ749657   UGG GAA GGU ACC AGU UGA CGG GAC GGC GUU UCG GGG AAA CUG AGA ACC CAG AUU GUC GAA
AF146281   UGG GAA GGU ACG AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CGC AUU GUC GAA
NC_002330  UGG GAA GGU ACC AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CGC AUU GUC GAA
AJ252152   UGG GAA GGU ACG AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CGC AUU GUC GAA
AJ298069   UGG GAA GGU ACG AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CGC AUU GUC GAA
AF146283   UGG GAA GGU ACG AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CGC AUU GUC GAA
AF146282   UGG GAA GGU ACU AGU UGA CGG GAC AGC GUC UCG GGA AAA UUG AGA ACU CAG GUC GUC GAA
NC_002042  UGG GAA GGU UCG AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CAA AUC GUU GAA
X81639     UGG GAA GGU UCG AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CAA AUC GUU GAA
AB299272   UGG GAA GGU UCG AGU UGA CGG GAU GGC GUC UCG GAC AAG AGU AGG GUC CAA GUG GUG GAG
NC_002356  UGG GAA GGU UCG AGU UGA CGG GAU GGC GUC UCG GAU AAG AGU AGG GUC CAA GUG GUG GAG
AJ271839   UGG GAA GGU UCG AGC UGA CGG GAC AGC GUC UCG GAC AAG AGU AGG GUC CAA GUG GUG GAG
NC_002357  UGG GAA GGU AGU GCC UGA CGG GGC GGC GUC UCG GAG AAG ACU CGU GUG CAA GUU GUC GAA
NC_004015  AAA UGG AGU GGC AAA UGA UGG GGU GGC GUC UUG UCG GCG AGC CGC GAG UCU GUA GUG GAA
                 **         ***  ** *    ** **  * *             *          *  **  **

                                                                  ]]] ]]] ))) ))) )))
AB033690   GAC UUC CAG AAU AGG CUG GUG UUG GCU GAC GAU UUG GGC GUU UUU CCC AGA CGC AAU CUU
AJ749657   GAU UUC CAG AAC AGG CUG GUG UUG GCU GAC GAU UUG GGU GUU UUU CCG AGA CGC AAU CUC
AF146281   GAU UUC CAG AAU AGG CUG GUU UUG GCC GAC GAU UUG GGU GUU UUU CCG AGA CGC AAU CUU
NC_002330  GAU UUC CAG AAU AGG CUG GUU UUG GCC GAC GAU UUG GGU GUU UUU CCG AGA CGC AAU CUU
AJ252152   GAU UUC CAG AAU AGG CUG GUU UUG GCC GAC GAU UUG GGU GUU UUU CCG AGA CGC AAU CUU
AJ298069   GAU UUC CAG AAU AGG CUG GUU UUG GCC GAC GAU UUG GGU GUU UUU CCG AGA CGC AAU CUU
AF146283   GAU UUC CAG AAU AGG CUG GUU UUG GCC GAC GAU UUG GGU GUU UUU CCG AGA CGC AAU CUU
AF146282   GAC UUC CAG AAU AGG CUG GUG UUG GCA GAC GAU UUG GGU AUU UUC CCG AGA CGC AAU AUC
NC_002042  GAA UUU CAG AAU CGA CUC AUA AUU GCU GAC GAU UUG GGC AUC UUU CCG AGA CGC GAU GUU
X81639     GAA UUU CAG AAU CGA CUC AUA AUU GCU GAC GAU UUG GGC AUC UUU CCG AGA CGC GAU GUU
AB299272   GAU CUG UCU AAC AGA CUC ACA CUG GCU GAC GAC CUG GGU AUC UUA CCG AGA CGC GAC AAC
NC_002356  GAU CUG UCU AAC AGA CUC ACA CUG GCU GAC GAC CUG GGU AUC UUA CCG AGA CGC GAC AAU
AJ271839   GAU CUG UCU AAC AGA CUC ACA CUU GCU GAC GAC CUG GGU AUC UUA CCG AGA CGC GAC AAC
NC_002357  GAG CUA GCU GAU AGA AUA GUA UUA GCC GAC GAU UUA GGC AUU UUU CCG AGA CGC GGU GAA
NC_004015  AAU GCU GUU CUG CGG CGA UUA ACC GCU AAA GAU UUG GGG UUG CUG CCA AGG CGC GAU ACA
            *               *              **   *  **   *  **   *   *  **  **  ***
```

Additional furovirus RNA2 partial sequences:

```
                                                  (( ((( ((( (([ [[[ [
X89078     UGG GAA GGU UCG AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CAA AUC GUU GAA
HM133583   UGG GAG GGC UCG AGC UGA CGG GAC GGC GUC UCG GGG AAA UUG AGA GCA CAA AUC GUU GAA
X90903     UGG GAA GGU UCG AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CAA AUC GUU GAA
```

```
AJ298070   UGG AAA GGU UCG AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CGC AUU GUC GAA
AF519799   UGG GAA GGU UCG AGU UGA CGG GAC GGC GUC UCG GGA AAG UUG AGG GCA CAA AUC GUU GAA
AY016008   UGG GAG GGC UCG AGC UGA CGG GAC GGC GUC UCG GGG AAA UUG AGA GCA CAA AUC GUU GAA

                                                                ]]] ]]] ))) ))) )))
X89078     GAA UUU CAG AAU CGA CUC AUA AUU GCU GAC GAU UUG GGC AUU UUU CCG AGA CGC GAU GUU
HM133583   GAA UUU CAG AAU CGA CUC AUA AUU GCU GAC GAU UUG GGC AUC UUU CCG AGA CGC GAC GUU
X90903     GAA UUU CAG AAU CGA CUC AUA AUU GCU GAC GAU UUG GGC AUU UUU CCG AGA CGC GAU GUU
AJ298070   GAU UUC CAG AAU AGG CUG GUU UUG GCC GAC GAU UUG GGU GUU UUU CCG AGA CGC AAU CUU
AF519799   GAA UUU CAG AAU CGA CUC AUA AUU GCU GAC GAU UUG GGC AUC UUU CCG AGA CGC GAU GUU
AY016008   GAA UUU CAG AAU CGA CUC AUA AUU GCU GAC GAU UUG GGC AUC UUU CCG AGA CGC GAC GUU
```

Pomovirus RNA2 (CP extension)

Notes on the nucleotide sequences:

1. Unlike the furovirus CP extension readthrough site, in the pomovirus CP extension readthrough site the readthrough codon is UAG (or UAA, in the case of Broad bean necrosis virus, D86637) and the stop codon is preceded by an 'A' and followed by CAA UYA – i.e. similar to the tobamovirus readthrough site. Although a potential stem is annotated below, it is weaker than the stem predicted for the furoviruses and has little support from the synonymous site conservation analysis (main text Fig 4).
2. D86637 is highly divergent from the other pomovirus sequences (overall, < 30% amino acid identity).
3. See notes on Tobra/Peclu/Furo/Pomovirus RNA1 (above) for a description of the colour coding used here.

RefSeqs:

NC_003511	Beet virus Q	Pomovirus	(derived from AJ223597)
NC_003518	Beet soil-borne virus	Pomovirus	(derived from U64512)
NC_003724	Potato mop-top virus	Pomovirus	(derived from AJ243719)
NC_004424	Broad bean necrosis virus	Pomovirus	(derived from D86637)

Extracts around the readthrough stop codon are shown below for the pomovirus RNA2 sequences used in the synonymous site conservation analysis (main text Fig 4). The highly divergent sequence NC_004424 = D86637 is also shown below but was not used in the synonymous site conservation analysis. BVQ sequence NC_003511 has been shown to be defective with respect to the CP extension coding sequence (multiple indels due to 'lab' adaption; Crutzen et al 2009 J Gen Virol 90:754-758). Thus, although the NC_003511 sequence is shown below, for the synonymous site conservation analysis the CP CDS from NC_003511 was fused with the RT CDS from the BVQ partial sequence FM244652 to make a synthetic representative for BVQ, as there are no other full-length BVQ sequences currently available.

Additional relevant sequences that were omitted for various reasons (e.g. incomplete coverage of the full coding sequence) are shown below this alignment.

```
                                          (( ((( ((( ((
NC_004424  UAU AAA CCG ACA GCA UAA CAA UUA ACG UAU CCG AUA AAA GAU CCU AAU GAG AUG UUA CGU
NC_003511  AAG UGG ACC GGC UCA UAG CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
EF545140   UUG CGC UGG GUU GAA UAG CAA UCA ACU GCU CCG CCG GUU AAA GUG GAU GUA CGU GCU GAA
NC_003518  UUG CGC UGG GUU GAA UAG CAA UCA ACU GCU CCG CCG GUU AAA GUG GAC GUA CGU GCU GAA
EF545141   UUG CGC UGG GUU GAA UAG CAA UCA ACU GCU CCG CCG GUU AAA GUG GAC GUA CGU GCU GAA
FJ971718   UUG CGC UGG GUU GAA UAG CAA UCA ACU GCU CCG CCG GUU AAA GUG GAC GUA CGU GCU GAA
FN386613   UUG CGC UGG GUU GAA UAG CAA UCA ACU GCU CCG CCG GUU AAA GUG GGC GUA CGU GCU GAA
AM503632   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503630   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
DQ102381   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503631   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AY196094   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503618   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503619   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503620   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503621   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503622   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
```

```
AM503623   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503624   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503625   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503626   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503627   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503628   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503629   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503611   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503612   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503613   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503614   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503615   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503616   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503633   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
NC_003724  CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
                         * **  *** * * *      * *                   *

                                                                  ) ))) ))) )))
NC_004424  GAG UUU GCG CUG AGG CUA AAG CUU GCU AAC AAU UUA GGU AUU UUA CCG GGU CAA CCA AAG
NC_003511  AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUG CUA CCG GGA CAA
EF545140   AUU GAU UGU AUU AAC AAU CGA AUG AAA GCA GCG GCG CGU UUA GGU CUA CUA CCG GGG CAG
NC_003518  AUU GAU UAU AUU AAC AAU CGA AUG AAA GCA GCG GCG CGU UUA GGU CUA CUA CCG GGG CAG
EF545141   AUU GAU UGC AUU AAC AAU CGA AUG AAA GCA GCG GCG CGU UUA GGU CUA CUA CCG GGG CAG
FJ971718   AUU GAU UGU AUU AAU AAU CGA AUG AAA GCA GCG GCG CGU UUA GGU CUA CUA CCG GGG CAG
FN386613   AUU GAU UGU AUU AGC AAU CGA AUG AAG GCA GCG GCG CGU UUA GGU CUA CUA CCG GGG CAG
AM503632   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503630   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GUU GCG UUA AGA CUA UUG CCU GGG CAG
DQ102381   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503631   AUC GAC GAA AUU GAU AAU CGA CUG CUG GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AY196094   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503618   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503619   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503620   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503621   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503622   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503623   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503624   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503625   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503626   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503627   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503628   AUC GAC GAA AUU GAU AAU CGA CUG CUG GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503629   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503611   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503612   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503613   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503614   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503615   AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503616   AUC GAC GAA AUU GAU AAU CGA CUG CUG GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503633   AUC GAC GAA AUU GAU AAU CGA CUG CUG GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
NC_003724  AUC GAC GAA AUU GAU AAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
                     *                 *                         *              *       *

Additional pomovirus RNA2 partial sequences:
                                                [[ [[[ [[[ [[
AJ005497   UUG CGC UGG GUU GAA UAG CAA UCA ACU GCU CCG CCG GUU AAA GUG GAC GUA CGU GCU GAA
FM244652   --- --- --- --- --- --- CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
FM244648   --- --- --- --- --- --- CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
FM244649   --- --- --- --- --- --- CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
FM244650   --- --- --- --- --- --- CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
FM244644   --- --- --- --- --- --- CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
FM244645   --- --- --- --- --- --- CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
FM244643   --- --- --- --- --- --- CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
FM244651   --- --- --- --- --- --- CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
FM244646   --- --- --- --- --- --- CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
```

```
FM244647   --- --- --- --- --- --- CAA UCA AUU GUC CUG GGA UGG GGU GAG AGC GUG CAU GAU GCU
AF487407   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AM503617   CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
D16193     CGC UGG GCU GGU GCA UAG CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU
AJ224991   --- --- --- --- --- --- CAA UUA ACC GCU CAG GCU UUU UGG UUU GAC GUC AAC GCC GCU

                                                                                ]]] ]]] ]]]
AJ005497   AUU GAU UGU AUU AAU AAU CGA AUG A-- --- --- --- --- --- --- --- --- --- --- ---
FM244652   AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUG CUG CCG GGA CAA
FM244648   AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUU CUG CCG GGA CAA
FM244649   AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUG CUG CCG GGA CAA
FM244650   AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUG CUG CCG GGA CAA
FM244644   AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUG CUG CCG GGA CAA
FM244645   AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUG CUG CCG GGA CAA
FM244643   AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUG CUG CCG GGA CAA
FM244651   AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUG CUG CCG GGA CAA
FM244646   AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUG CUG CCG GGA CAA
FM244647   AUC AAU GCG AUU GAU AAU CGU AUU AUU GCA GCU GCA CGG UUG CUC AUG CUG CCG GGA CAA
AF487407   AUC GAC GAA AUU GAU GAU CGA CUG CUA GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AM503617   AUC GAC GAA AUU GAU AAU CGA CUG CUG GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
D16193     AUC GAC GAA AUU GAU AAU CGA CUG CUG GCU GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
AJ224991   AUC GAC GAA AUU GAU AAU CGA CUG CUG GCG GCG GCU GCG UUA AGA CUA UUG CCU GGG CAG
```

Coltivirus segment 9

Sequences EU789390, AF000720 and AF282475 were used for the synonymous site conservation analysis (main text Fig 4). There is enhanced synonymous site conservation over ~40 codons 3'-adjacent to the readthrough UGA codon. AF007183, AF007182 and AF007184 are partial (incomplete) sequences. EU789390 is also a partial sequence but has sufficient coverage of the full coding sequence for it to be useful to include in the synonymous site conservation analysis.

```
                                                ( (((  (((  (((  (
EU789390   ACU AAG GGU UGU UGU UGA CGG UGU UGG CCU GGU GAG CUC UAC CUC AUC AAC AUC GAG AGU
AF000720   ACC AAG GGC UGC UGU UGA CGG UGU UGG CCU GGU GAG CUG UAC CUC AUU AAC AUC GAG AGU
AF282475   ACU AAG GGU UGU UGU UGA CGG UGU UGG CCU GGU GAG CUC UAC CUC AUC AAC AUC GAG AGU
AF007183   ACC AAG GGU UGC UGU UGA CGG UGU UGG CCU GGU GAG CUG UAC CUC AUU AAC AUC GAG AGU
AF007182   ACC AAG GGC UGC UGU UGA CGG UGU UGG CCU GGU GAG CUG UAU CUC AUU AAC AUC GAG AGU
AF007184   ACC AAG GGC UGC UGU UGA CGG UGU UGG CCU GGU GAG CUG UAC CUC AUU AAC AUC GAG AGU
           **  *** **  **  *** *** *** *** *** *** *** *** **  **  *** **  *** *** *** ***

                                                        ))) ))) ))) ))
EU789390   GUA GGC UUU AGA GAA GGA GCU UCA GCA GCA GCG GCU UAC CAG GCG GGU UUC UUG CGG GCG
AF000720   GCA GGC UUU CGU GAA GGA GCU UCA GCG GCU GCG GCU UAC CAG GCG GGC UUU UUG CGU GCU
AF282475   GUG GGC UUU AGA GAG GGA GCU UCA GCA GCA GCG GCU UAC CAG GCG GGU UUC UUG CGG GCG
AF007183   GCA GGC UUU CGU GAA GGA GCU UCA GCG GCU GCG GCU UAC CAG GCG GGC UUU UUG CGU GCU
AF007182   GCA GGC UUU CGU GAA GGA GCU UCA GCG GCU GCG GCU UAC CAG GCG GGC UUU UUG CGU GCU
AF007184   GCA GGC UUU CGU GAA GGA GCU UCA GCG GCU GCG GCU UAC CAG GCG GGC UUU UUG CGU GCU
           *   *** ***  *  **  *** *** *** **  **  *** *** *** *** *** **  **  *** **  **
```

```
RefSeqs:
NC_004180  Colorado tick fever virus       Coltivirus  (derived from AF000720)
NC_003704  Eyach virus                     Coltivirus  (derived from AF282475)
```

Tobamovirus

Notes on the nucleotide sequences:

1. The readthrough context 'xxA UAG CAA UUA xxG' is completely conserved despite lack of amino acid conservation at the -1 and +3 codons.

An extract around the readthrough stop codon is shown below for the sequences used in the synonymous site conservation analysis (main text Fig 4). Note that the sequences as shown below are **NOT** aligned but simply lined up at the stop codon. An additional 10 sequences, with coverage of the readthrough site, but not used for the synonymous site conservation analysis for various reasons (e.g. incomplete coverage of the full coding sequence, high number of ambiguous nucleotide codes, etc.), also have conserved 'xxA UAG CAA UUA xxG'.

```
AB000709   GAU GUG UCG ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
AB069853   GAU GUG UCG ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
AB550911   GAU GUG UCG ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
AY859497   GAU GUG UCG ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
NC_003630  GAU GUG UCG ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
AB113116   GAU GUG UCG ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
AB113117   GAU GUG UCG ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
AB126003   GAU GUG UCG ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
AB276030   GAU GUG UCG ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
AB254821   GAU GUG UCG ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
AJ308228   GAU GUG UCA ACU CAA UAG CAA UUA CAG AUA GAA UCG GUG UAC AAA GGU GUU AAC CUU UUC
AB015145   GUG GCG UCC CCC AAA UAG CAA UUA CAG CAG AUG UCG UUC CAA AUC ACU GAG AAU AUA GUU
AB162006   GUA GCG UCU ACC AAA UAG CAA UUA CAG CAA AUG UCG UUC CAA AUA ACC GAA AAU GUA UUC
NC_003610  GUA GCG UCC ACC AAA UAG CAA UUA CAG CAA AUG UCG UUC CAC AUC ACU GAA AAC GUA GUU
AJ252189   GUG GCG UGC ACC AAA UAG CAA UUA CAG CAA AUG UCG UUC CAC AUU UCA GAA AAU GUA GUA
NC_003878  GUG GCG UGC ACC AAA UAG CAA UUA CAG CAA AUG UCG UUC CAC AUU UCA GAA AAU GUA GUA
NC_002633  GUG GCG GGG ACC AAA UAG CAA UUA CAG CAA CUA UCC GUC UAU GUG CAU UUG CCC GUG UCG
AB015146   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAC UCA CUG UAU GUC CAU CGU AAU AUU UUC
AB369274   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAC UCA CUG UAU GUC CAU CGU AAU AUU UUC
GQ277655   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAC UCA CUG UAU GUC CAU CGU AAU AUU UUC
FJ654657   ACU GUG UCU ACC AAA UAG CAA UUA AUG CAG AAC UCG CUG UAU GUC CAU CGU AAU AUU UUC
AF417242   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAC UCG CUG UAU GUC CAU CGU AAU AUU UUC
EU352259   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAC UCG CUG UAU GUC CAU CGU AAU AUU UUC
FJ654658   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAC UCG CUG UAU GUC CAU CGU AAU AUU UUC
FJ654659   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAC UCG CUG UAU GUC CAU CGU AAU AUU UUC
HM008919   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAC UCG CUG UAU GUC CAU CGU AAU AUU UUC
NC_001801  ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAU UCG CUG UAU GUC CAU CGU AAU AUU UUC
EF611826   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAU UCG CUG UAU GUC CAU CGU AAU AUU UUC
DQ767631   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAC UCG CUG UAU GUC CAU CGU AAU AUU UUC
AF417243   ACU GUG CCU ACC AAA UAG CAA UUA AUG CAG AAU UCG CUG UAU GUC CAU CGU AAU AUU UUC
FJ848666   ACU GUG CCU ACC AAA UAG CAA UUA ACG CAG AAC UCU CUG UAU GUU CAC CGC AAU AUU UUC
GQ411361   ACU GUG CCU ACC AAA UAG CAA UUA ACG CAG AAC UCU CUG UAU GUU CAC CGC AAU AUU UUC
GQ495274   ACU GUG CCU ACC AAA UAG CAA UUA ACG CAG AAC UCU CUG UAU GUU CAC CGC AAU AUU UUC
GQ495275   ACU GUG CCU ACC AAA UAG CAA UUA ACG CAG AAC UCU CUG UAU GUU CAC CGC AAU AUU UUC
NC_008614  GUG GCU GGG AAG GCA UAG CAA UUA AAG CAA AGU CAG UUA UAC UUC CAA AAG AAC UUG CAG
NC_008310  GAG AAU CAG AAA GCA UAG CAA UUA CAG CAG CAU CGC CUC UAC AUC CAC AAA AAC AUC UUU
FJ196834   GAA GGC CAA AAG GCA UAG CAA UUA CAG CAA AAC CGU CUC UAC AUC CAC AAA AAC AUU UUC
NC_014546  GAU UCU UCG ACC AAA UAG CAA UUA AUG GAC CUU CCC GUG UAU GAG UUU CGC AAC UUA CAC
NC_011803  CUA GGG UCG ACA GCA UAG CAA UUA ACG ACU CAG CCC AUC ACU AGA GGU GUC AAU AUG UAC
NC_004106  GAU AGU GUU AAC GCA UAG CAA UUA CAG GUA GGU AGU GUG UAU AUG UCC GAA AGU CUU UUU
L11665     GAU AGU GUG AGU GCA UAG CAA UUA CAG GUA AGC GGU GUG UAC CUA GCG GAG AAU CUC UUC
NC_003852  GAU AGU GUG AGU GCA UAG CAA UUA CAG GUA AGC GGU GUG UAC CUA GCG GAG AAU CUC UUC
NC_003355  GAG UCG GGA ACC CAA UAG CAA UUA CAG AUC GAU GCG GUG UUC AAA GGG GUC AAC UUA AAU
AB017504   GAG UCG GGA ACC CAA UAG CAA UUA CAG AUC GAU GCG GUG UUC AAA GGG GUC AAC UUA AAU
AB261175   GAG UCG GGG ACC CAA UAG CAA UUA CAG AUC GAC ACA GUG UUC AAA GGG ACA AAC CUG UUC
D38444     GAG UCG GGG ACC CAA UAG CAA UUA CAG AUC GAC ACA GUG UUC AAA GGG ACA AAC CUG UUC
DQ223770   GAG UCG GGA ACC CAA UAG CAA UUA CAG AUC GAC ACA GUG UUC AAA GGG ACG AAU UUG UUU
EU571218   GAG UCG GGA ACC CAA UAG CAA UUA CAG AUC GAC ACA GUG UUC AAG GGG ACA AAU UUG UUC
AY318866   GAG UCG GGG ACC CAA UAG CAA UUA CAG AUC GAC ACA GUG UUC AAA GGG ACG AAU CUG UUC
NC_002792  GAG UCG GGG ACC CAA UAG CAA UUA CAG AUC GAC ACA GUG UUC AAA GGG ACG AAU CUG UUU
NC_004422  GAG UCG GGU ACC CAA UAG CAA UUA CAG AUC GAC ACA GUG UUC AAA GGG ACA AAC UUG UUC
```

```
GQ401365   GAA GCU GGC AUC CAA UAG CAA UUA CAG AUA GAU GCA GUA UUC AAG GGA ACG AAU CUG UUC
GQ401366   GAA GCU GGC AUC CAA UAG CAA UUA CAG AUA GAU GCA GUA UUC AAG GGA ACG AAU CUG UUC
Z29370     GAA GCG GGG AUC CAA UAG CAA UUA CAG AUA GAU GCA GUA UUC AAG GGA ACG AAC UUA UUU
NC_001873  GAA GCG GGG GUC CAA UAG CAA UUA CAG AUC GAU GCA GUA UUC AGG GAC AGC AAC UUG UUU
NC_008365  GAA AAA CCU GUC CAA UAG CAA UUA CAG AUA GAG AGU GUC UUC AAG GGU UCG AAU UUA UUC
AY571290   GAU GCC GGG AUC UUA UAG CAA UUA CAG GUA GGG AGU AUC UUC AAA GGC GAA AAU CUA UUC
DQ139262   GAU GCC GGG AUC UUA UAG CAA UUA CAG GUA GGG AGU AUC UUC AAA GGC GAA AAU CUA UUC
U89894     GAU GCC GGG AUC UUA UAG CAA UUA CAG GUA GGG AGU AUC UUC AAA GGC GAA AAU CUA UUC
D13941     GAU GCC GGG AUC UUA UAG CAA UUA CAG GUA GGG AGU AUC UUC AAA GGC GAA AAU CUA UUC
S83257     GAU GCC GGG AUC UUA UAG CAA UUA CAG GUA GGG AGU AUC UUC AAA GGC GAA AAU CUA UUC
NC_001728  GAU GCC GGG AUC UUA UAG CAA UUA CAG GUA GGG AGU AUC UUC AAA GGC GAA AAU CUA UUC
U34586     GAU GCC GGG AUC UUA UAG CAA UUA CAG GUA GGG AGU AUC UUC AAA GGC GAA AAU CUA UUC
AB078435   GAA GCA GGU GGU AGA UAG CAA UUA CAG AUG GAU GCA GUG UUC AAA GGU CAU AAU CUC UUU
DQ821941   GAA GCA GGU GGU AGA UAG CAA UUA CAG AUG GAU GCA GUG UUC AAA GGU CAU AAU CUC UUU
EF469769   GAA GCA GGU AGU AGA UAG CAA UUA CAG AUG GAU GCA GUG UUC AAA GGU CAU AAU CUC UUU
NC_001556  GAA GCA GGU AGU AGA UAG CAA UUA CAG AUG GAU GCA GUG UUC AAA GGU CAU AAU CUC UUU
NC_010944  UCA GCA GGU CAG CAA UAG CAA UUA CAG GUC GGG UUC GAA UUU AUU CGU UCC AAU CUG UUU
NC_009642  GAU GCU GGU GUG CAA UAG CAA UUA CAG GUC GAC UCG GUA UUU UCA GGG ACA AAU CUC UUC
AB083196   GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA AAU UUC AAU CUU UUU
AJ417701   GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA AAU UUC AAU CUU UUU
NC_002692  GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA AAU UUC AAU CUU UUU
X02144     GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA AAU UUC AAU CUU UUU
Z92909     GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA AAU UUC AAU CUU UUU
AJ243571   GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA AAU UUC AAU CUU UUU
DQ873692   GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA AAU UUC AAU CUU UUU
GQ280794   GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA AAU UUC AAU CUU UUU
AB355139   GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA AAU UUC AAU CUU UUU
AJ132845   GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA AAU UUC AAU CUU UUU
FN985165   GAU GCA GGU ACU CAA UAG CAA UUA CAG GUC GAC UCU GUG UUU AAA GAU UUC AAU CUU UUU
NC_009041  GAU GCA GGU ACG CAA UAG CAA UUA CAG AUA AAC UCG GUG UUU AAA GGU UCU AAU CUU UUU
EF392659   GAU GCA GGU ACC CAA UAG CAA UUA CAG GUU GAC UCG GUG UUU AGA GGU UCG AAU CUU UUU
D78608     GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
AF165190   GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
AB369276   GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
D63809     GAU GCA GGC AUU CAA UAG CAA UUA CAG AUU GAC UCG GUG UUU AAA GGU UCA AAU CUU UUU
X68110     GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
AJ011933   GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
AB369275   GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
AF395129   GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
AF273221   GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
NC_001367  GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
AF395127   GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
AF395128   GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
V01409     GAU GCA GGA ACA CAA UAG CAA UUA CAG AUU GAC UCG GUG UUC AAA GGU UCC AAU CUU UUU
                         * *** *** ***   *                                       *
```

RefSeqs:

NC_001367	Tobacco mosaic virus	Tobamovirus	(derived from V01408)
NC_001556	Tobacco mild green mosaic virus	Tobamovirus	(derived from M34077)
NC_001728	Odontoglossum ringspot virus	Tobamovirus	(derived from X82130)
NC_001801	Cucumber green mottle mosaic virus	Tobamovirus	(derived from D12505)
NC_001873	Turnip vein-clearing virus	Tobamovirus	(derived from U03387)
NC_002633	Cucumber fruit mottle mosaic virus	Tobamovirus	(derived from AF321057)
NC_002692	Tomato mosaic virus	Tobamovirus	(derived from AF332868)
NC_002792	Ribgrass mosaic virus	Tobamovirus	(derived from AF254924)
NC_003355	Crucifer tobamovirus	Tobamovirus	(derived from AB017503)
NC_003610	Kyuri green mottle mosaic virus	Tobamovirus	(derived from AJ295948)
NC_003630	Pepper mild mottle virus	Tobamovirus	(derived from M81413)
NC_003852	Obuda pepper virus	Tobamovirus	(derived from D13438)
NC_003878	Zucchini green mottle mosaic virus	Tobamovirus	(derived from AJ295949)
NC_004106	Paprika mild mottle virus	Tobamovirus	(derived from AB089381)
NC_004422	Youcai mosaic virus	Tobamovirus	(derived from U30944)
NC_008310	Hibiscus latent Singapore virus	Tobamovirus	(derived from AF395898)

NC_008365	Streptocarpus flower break virus	Tobamovirus	(derived from AM040955)
NC_008614	Cucumber mottle virus	Tobamovirus	(derived from AB261167)
NC_009041	Rehmannia mosaic virus	Tobamovirus	(derived from EF375551)
NC_009642	Bell pepper mottle tobamovirus	Tobamovirus	(derived from DQ355023)
NC_010944	Brugmansia mild mottle virus	Tobamovirus	(derived from AM398436)
NC_011803	Cactus mild mottle virus	Tobamovirus	(derived from EU043335)
NC_014546	Frangipani mosaic virus	Tobamovirus	(derived from HM026454)

Luteo/Polerovirus

Notes on the nucleotide sequences:

1. The readthrough context 'AAA UAG GUA' is completely conserved (for amino acid conservation either an 'A' or a 'G' at -1, and any nucleotide at +3 would suffice).
2. 'C' at +6 is highly conserved, despite the corresponding amino acid not being conserved: 1 AAC, 2 AGC, 1 GAA, 164 GAC, 2 GAU, 9 GGC, 1 RGC (ambiguous nt 'R').
3. Enhanced synonymous site conservation extends over several further codons, especially codons -1 to +5.

An extract around the readthrough stop codon is shown below for the sequences used in the synonymous site conservation analysis (main text Fig 4). Note that the sequences as shown below are **NOT** aligned but simply lined up at the stop codon. An additional 66 sequences, with coverage of the readthrough site, but not used for the synonymous site conservation analysis for various reasons (e.g. incomplete coverage of the full coding sequence, high number of ambiguous nucleotide codes, etc.) also had conserved 'AAA UAG GUA' except for one sequence, FJ687407 (BYDV-PAV 562 nt fragment), with 'AGA UAG GUA' (perhaps a sequencing error or defective virus sequence?). A further sequence, NC_010806 = EU024678 is quite different and is discussed separately below.

```
NC_003056  ACG CAG AAU GCU AAA UAG GUA GAC GGA GAA CCG GGU CCG AAA CCA GGU CCA GAC CCA GCA
AB076038   ACU CAG AAU CCU AAA UAG GUA GAC GGA GAA CCG GGU CCG AAA CCA GGU CCA GAC CCA GCA
L24049     ACG CAG AAU GCU AAA UAG GUA GAC GGA GAA CCG GGU CCG AAA CCA GGU CCA GAC CCA GCA
AB038149   ACU CAG AAU CCU AAA UAG GUA GAC GGA GAA CCG GGU CCG AAA CCA GGU CCA GAC CCA GCA
AB038148   ACG CAG AAU GCU AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG AGC CCC GCA
FJ390404   ACU CAG AAU CCU AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
AB038150   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
DQ145545   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419573   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU095846   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU095847   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCG GGU CCG GAC CCC ACA
EU419574   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419575   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419576   ACU CAG AAY CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419577   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419578   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419579   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419580   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419581   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419582   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419570   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAG CCA GGU CCG GAC CCC ACA
EU419571   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
FJ390401   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419584   ACU CAG AAU CCC AAA UAG GUA GAC GGA GGA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
FJ390405   ACU CAG AAU CCU AAA UAG GUA GAC GGA GGA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
FJ390403   ACU CAG AAU CCC AAA UAG GUA GAC GGA GGA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
FJ390406   ACU CAG AAU CCC AAA UAG GUA GAC GGA GGA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
EU419572   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAG CCA GGU CCG GAC CCC ACA
EU419583   ACU CAG AAU CCA AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
FJ390402   ACU CAG AAU CCC AAA UAG GUA GAC GGA GAA CCG GGG CCG AAA CCA GGU CCG GAC CCC ACA
NC_002198  CUA CAA AAC CCG AAA UAG GUA GAC GCG GAA CCC GGC CCU AGU CCA GGG CCA UCA CCC GAC
D10206     CUA CAA AAC CCA AAA UAG GUA GAC GCG GAA CCC GGU CCC AGU CCA GGA CCA UCA CCC GAC
NC_004751  CUA CAA AAC CCA AAA UAG GUA GAC GCG GAA CCC GGU CCC AGU CCA GGA CCA UCA CCC GAC
EF521827   CUA CAA AAC CCA AAA UAG GUA GAC GCG GAA CCC GGU CCC AGU CCA GGA CCA UCA CCC GAC
EF521848   CUA CAA AAC CCA AAA UAG GUA GAC GCG GAA CCC GGU CCC AGU CCA GGA CCA UCA CCC GAC
EF521830   CUA CAA AAC CCA AAA UAG GUA GAC GCG GAA CCC GGU CCC AGU CCA GGA CCA UCA CCC GAC
EF521839   CUA CAA AAC CCA AAA UAG GUA GAC GCG GAA CCC GGU CCC AGU CCA GGA CCA UCA CCC GAC
NC_012931  UUU CAA AAC CCG AAA UAG GUA GAC GCG GAA CCC GGU CCU AGU CCA GGA CCA UCA CCC ACU
NC_002766  UUC CAU AAC CCC AAA UAG GUA GAC GAG GAA CCC GGC CCU GGC CCA GGG CCU UCU CCU UCU
EU022509   UUC CAU AAC CCC AAA UAG GUA GAC GAG GAA CCC GGC CCU GGC CCA GGG CCU UCU CCU UCU
EU022511   UUC CAU AAC CCC AAA UAG GUA GAC GAG GAA CCC GGC CCU AGC CCA GGG CCU UCU CCU UCU
EU022510   UUC CAC AAC CCC AAA UAG GUA GAC GAG GAA CCC GGC CCU GGC CCA GGG CCU UCU CCC UCU
AF352025   UUC CAC AAC CCC AAA UAG GUA GAC GAG GAA CCC GGC CCU AGC CCA GGG CCU UCU CCC UCU
NC_004756  UUC CAC AAC CCC AAA UAG GUA GAC GAG GAA CCC GGC CCU AGC CCA GGG CCU UCU CCC UCU
```

```
DQ132996   UUC CAC AAU CCC AAA UAG GUA GAC AAG GAA CCC GGC CCU AGC CCA GGG CCU UCU CCC UCU
EU148509   UUC CAC AAU CCC AAA UAG GUA GAC AAG GAA CCC GGC CCU AGC CCA GGG CCU UCU CCU UCU
NC_003491  UUC CAC AAU CCG AAA UAG GUA GAC AAG GAA CCC GGC CCU AGC CCA GGG CCU UCU CCC UCU
EU148510   UUC CAC AAU CCC AAA UAG GUA GAC AAG GAA CCC GGC CCU AGC CCA GGG CCU UCU CCC UCU
EU148508   UUC CAC AAU CCC AAA UAG GUA GAC AAG GAA CCC GGC CCU AGC CCA GGG CCU UCC CCC UCU
NC_003743  UUC CAC AAC CCC AAA UAG GUA GAC GAG GAA CCC GGC CCU AGC CCA GGG CCU UCU CCC UCU
X13062     UUC CAC AAU CCC AAA UAG GUA GAC GAG GAA CCC GGC CCU AGC CCA GGG CCU UCU CCC UCU
EU636990   UUU CAC AAC CCC AAA UAG GUA GAC GCG GAA CCA AGC CCU AGC CCA GGG CUU UCU CCC UCU
EU636991   UUU CAC AAC CCC AAA UAG GUA GAC GCG GAA CCG AGC CCU AGC CCA GGG CCU UCU CCC UCU
NC_003369  ACU CAG AAU CCA AAA UAG GUA GAC GGA GAA CCG GAU CCU CCG GGC CCU GGC CCA GAU CCA
AF271214   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
AF453394   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
AF453390   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
AF271215   UUG CAA AAC CCC AAA UAG GUA GAC UCC AGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
AF453393   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
EU717545   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
AF453388   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
EU717546   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
X74789     UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
X77321     UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
X77322     UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
X77323     UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
AF453391   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
X77325     UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
X77326     UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
AY138970   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
EU313202   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
X77324     UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
D13953     UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
D13954     UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
AF453392   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCA AGC CCA CAA CCG ACA
Y07496     UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA CCA GAG CCU GGU CCA AGC CCA CAA CCA ACA
NC_001747  UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU AGU CCA AGC CCA CAA CCA ACA
AF453389   UUG CAA AAC CCC AAA UAG GUA GAC UCC GGA UCA GAG CCU GGU CCG AGC CCA CAA CCA ACA
NC_006265  UUC CAA AAU CCU AAA UAG GUA GAC GCG GAA CCG GGG CCU AGU CCA GGC CCA AGC CCC CCC
AY236971   UUC CAG AAU CCC AAA UAG GUA GGC GAC GCU CCC CCA ACA CCC ACC CCA ACA CCA ACC CCA
AM072750   UUC CAG AAU CCC AAA UAG GUA GGC GAC GCU CCC CCA ACA CCC ACC CCA ACA CCA ACC CCA
AM072751   UUC CAG AAU CCC AAA UAG GUA GGC GAC GCU CCC CCA ACA CCC ACC CCA ACA CCA ACC CCA
NC_000874  UUC CAG AAU CCC AAA UAG GUA GGC GAC GCU CCC CCA ACA CCC ACC CCA ACA CCA ACC CCA
AM072754   UUC CAG AAU CCC AAA UAG GUA GGC GAC GCU CCC CCA ACA CCC ACA CCA ACA CCA ACC CCA
AM072755   UUC CAG AAU CCC AAA UAG GUA GGC GAC GCU CCC CCA ACA CCC ACA CCA ACA CCA ACC CCA
AM072756   UUC CAG AAU CCC AAA UAG GUA GGC GAC GCU CCC CCA ACA CCC ACA CCA ACA CCA ACC CCA
GU190159   UUC CAG AAU CCC AAA UAG GUA GAC GAC GCU CCC CCA ACA CCC ACC CCA ACA CCA ACC CCA
AM072752   UUC CAG AAU CCC AAA UAG GUA GGC GAC GCU CCC CCA ACA CCC ACC CCA ACA CCA ACC CCA
AM072753   UUC CAG AAU CCC AAA UAG GUA GGC GAC GCU CCC CCA ACA CCC ACC CCA ACA CCA ACC CCA
GU570006   UUC CAG AAU CCC AAA UAG GUA AGC GAC GCU CCC CCA ACA CCC ACC CCA ACA CCA ACC CCA
GU570007   UUC CAG AAU CCC AAA UAG GUA AGC GAC GCU CCC CCA ACA CCC ACC CCA ACA CCA ACC CCA
GU570008   UUC CAG AAU CCC AAA UAG GUA RGC GAC GCU CCC CCA ACA CCC ACC CCA AUA CCA ACC CCA
GU327735   UUC CAG AAU CCC AAA UAG GUA GAC GAC GCU CCC CCA ACA CCC ACA CCA CAA CCA ACC CCA
NC_008249  ACU CAG AAU CCG AAA UAG GUA GAU GAC AGU UCC CCC CCC GGA CCA AGU CCA ACA CCA CCU
NC_010809  GUC CAG AAC CCG AAA UAG GUA GAA GGC ACC UCC CCC UCA GCA CCA ACA CCA ACC CCC CCC
EU000535   GUC CAA AAC CCG AAA UAG GUA GAC GGC AGU UCC CCC CCC CCU CCC AGU CCU AGC CCG ACU
NC_003688  GUC CAG AAC CCG AAA UAG GUA GAC GGC AGU UCC CCC CCC CCC CCA AGU CCU AGU CCA ACC
EU636992   GUC CAA AAC CCG AAA UAG GUA GAC GGC AGC UCC CCC CCC CCC CCC AGU CCU AGC CCG ACU
GQ379224   UUC CAC AAU CCA AAA UAG GUA GAU GAU GGU CCC CCC CCA CCG GGU CCG UCU CCC CCU CCU
NC_014545  UUC CAC AAU CCA AAA UAG GUA GAC GAU GGU CCC CCC CCG CCG GGU CCG UCU CCC CCU CCU
U06865     AUG GCC AAC CCU AAA UAG GUA GAC CCC UCG CCG UCU CCC GCC CCC GCU CCC GCA CCC ACA
U06866     AUG GCC AAC CCC AAA UAG GUA GAC CCC UCG CCG UCC CCC GCU CCU ACU CCA ACA CCA ACC
AF338909   AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAG CCU ACA CCC CAA CCC CAA CCA
NC_004666  AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAG CCU ACA CCC CAA CCC CAA CCA
AY610953   AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAG CCU ACA CCC CAA CCC CAA CCA
AY610954   AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAG CCU ACA CCC CAA CCC CAA CCA
EU402390   AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAG CCU ACA CCC CAA CCC CAA CCA
EU402386   AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAG CCU ACA CCC CAA CCC CAA CCA
```

```
EU402387   AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAG CCU UCA CCC CAA CCC CAA CCA
EU402391   AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAG CCU UCA CCC CAA CCC CAA CCA
EU402388   AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAG CCU ACA CCC CAA CCC CAA CCA
EU402389   AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAG CCU ACA CCC CAA CCC CAA CCA
NC_003680  AAC AUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCA GAA CCU UCG CCA CAA CCC CAG CCA
EU332315   CUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU ACA CCU GCA CCA GCC CCA
EU332317   CUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU ACA CCU GCA CCA GCC CCA
EU332326   CUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU ACA CCU GCA CCA GCC CCA
EU332327   CUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU ACA CCU GCA CCA GCC CCA
AY855920   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU AAA CCC ACA CCA GCU CCA
EU332316   UUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU AAA CCC ACA CCA GCU CCA
EU332325   UUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU AAA CCC ACA CCA GCU CCA
EU332318   UUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU AAA CCC ACA CCA GCU CCA
EU332319   UUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU AAA CCC ACA CCA GCU CCA
EU332310   UUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU AAA CCC ACA CCA GCU CCA
EU332311   UUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCU AAA CCC ACA CCA GCU CCA
EU332308   CUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCC ACA CCA GCU CCA
EU332320   CUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCC ACA CCA GCU CCA
EU332330   CUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCA AAA CCC ACA CCA GCU CCA
EU332336   UUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAG CCC ACA CCA GCC CCA
EU332331   CUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCC ACA CCA GCC CCA
EU332321   UUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCC ACA CCA GCU CCA
EU332322   CUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCA AAA CCC ACA CCA GCU CCA
EU332323   CUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCC ACA CCA GCU CCA
EU332324   CUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCC ACA CCA GCU CCA
EU332312   CUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCA AAG CCC ACA CCA GCC CCA
EU332314   CUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCA AAG CCC ACA CCA GCC CCA
EU332334   CUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCA AAG CCC ACA CCA GCC CCA
EU332335   CUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCA AAG CCC ACA CCA GCC CCA
EU332307   UUA AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCC AAA CCC ACA CCA ACU CCA
EU332313   UUG AUG ACU GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCC AAA CCC ACA CCA ACU CCA
EF521833   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
NC_002160  AUG UUA ACU CCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCA GCA CCA GAG CCA
EF521850   AUG UUA ACU CCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCA GCA CCA GAA CCA
AJ810418   AUG CUG ACU CCC AAA UAG GUA GAC UCC UCA CCA CCG GAA CCG AAA CCA GCA CCA GAA CCG
EF521828   AUG CUG ACU CCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCA GCA CCA GAA CCA
AY040344   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCC
EF521832   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
EF521840   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCA AAA CCU GCA CCG GAA CCA
NC_004750  UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCA AAA CCU GCA CCG GAA CCA
EF521834   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
EF521845   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
EF521847   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
AF235167   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCA
EF521835   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCC GCA CCA GAA CCU
D11032     UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA UCG GAA CCG CAA CCU GCA CCA GAA CCU
EF521844   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA UCG GAA CCG AAA CCA GCA CCA GAA CCU
EF521829   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA UCG GAA CCG AAA CCU GCA CCA GAA CCU
EF521837   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA UUG GAA CCG AAA CCU GCA CCA GAA CCU
EF521842   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
EF521846   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
EF521849   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
EF521843   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
EU332333   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
AY040343   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCC
EF521841   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCA
EU332329   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCC
EF521838   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GGA CCC
EU332309   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCA
EF521831   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CAG AAA CCU GCA CCA GAA CCU
EF521836   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CAG AAA CCU GCA CCA GAA CCU
AJ007492   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAG CCA
EU332328   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
EU332332   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCA
```

```
AF391101   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCC
AJ007491   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCA
EF043235   UUG AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCU
DQ079609   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCC
DQ079610   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCC
DQ079611   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCC
D85783     UUG AUG ACG GCC AAA UAG GUA AAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCC
DQ079612   UUA AUG ACG GCC AAA UAG GUA GAC UCC UCA ACA CCG GAA CCG AAA CCU GCA CCA GAA CCC
                 *    *  *** *** ***             *
```

RefSeqs:

NC_000874	Sugarcane yellow leaf virus	Polerovirus	(derived from AF157029)
NC_001747	Potato leafroll virus	Polerovirus	(derived from D00530)
NC_002160	Barley yellow dwarf virus-PAS	Luteovirus	(derived from AF218798)
NC_002198	Cereal yellow dwarf virus-RPS	Polerovirus	(derived from AF235168)
NC_002766	Beet chlorosis virus	Polerovirus	(derived from AF352024)
NC_003056	Soybean dwarf virus	Luteovirus	(derived from AB038147)
NC_003369	Bean leafroll virus	Luteovirus	(derived from AF441393)
NC_003491	Beet mild yellowing virus	Polerovirus	(derived from X83110)
NC_003680	Barley yellow dwarf virus-MAV	Luteovirus	(derived from D11028)
NC_003688	Cucurbit aphid-borne yellows virus	Polerovirus	(derived from X76931)
NC_003743	Turnip yellows virus	Polerovirus	(derived from X13063)
NC_004666	Barley yellow dwarf virus-GAV	Luteovirus	(derived from AY220739)
NC_004750	Barley yellow dwarf virus-PAV	Luteovirus	(derived from X07653)
NC_004751	Cereal yellow dwarf virus-RPV	Polerovirus	(derived from L25299)
NC_004756	Beet western yellows virus	Polerovirus	(derived from AF473561)
NC_006265	Carrot red leaf virus	Polerovirus	(derived from AY695933)
NC_008249	Chickpea chlorotic stunt virus	Polerovirus	(derived from AY956384)
NC_010809	Melon aphid-borne yellows virus	Polerovirus	(derived from EU000534)
NC_012931	Wheat yellow dwarf virus-RPV	Polerovirus	(derived from FM865413)
NC_014545	Cotton leafroll dwarf virus	Polerovirus	(derived from GU167940)

NC_010806 = EU024678 (Luteovirus; Rose spring dwarf-associated virus) has UGA instead of UAG, and the 3'-adjacent CGG (instead of the GUA conserved in Luteo/Polerovirus UAG readthrough sites) is the same as in the Tobra/Peclu/Furo/Pomovirus RNA1 and several other UGA readthrough sites.

```
NC_010806  CAG AGU CCC CGU GAA UGA CGG GAC GCU GCU CCU GCC CCA UCC CCA UCA CCU CCU UCU CCG
           ACA CCU GCU CCU ACU CCU GCA CCA ACA CCU CAG CCA GAA AGA UUC UUU GUC UAU GCU GGC
           GUC CCU GGC GUU GAC AUU CAA ACU CGU GAG ACC GAU GAC AGC AUU AUU GUU GGU CGG UUG
```

Enamovirus

Also in the Luteoviridae family, but with significant differences from the luteo- and poleroviruses. Here, the readthrough codon is UGA instead of UAG. Little phylogenetic divergence among available sequences for comparative predictions of RNA structure or conserved regions.

```
AF082833   AUG AAU GCC UCC CUC UGA GGG GAC GAC GCU CCC CCA UCA CCA GGG CCU GAU CCC GGG CCC
Z48507     AUG AAU GCC UCC CUC UGA GGG GAC GAC GCU CCC CCG UCA CCA GGG CCC GAU CCC GGG CCC
NC_003629  AUG AAU GCC UCC CUC UGA GGG GAC GAC GCU CCC CCG UCA CCA GGG CCU GAU CCC GGG CCC
Y09099     AUG AAU GCC UCC CUC UGA GGG GAC GAC GCU CCC CCG UCA CCA GGG CCU GAU CCC GGG CCC
Y09100     AUG AAU GCC UCC CUC UGA GGG GAC GAC GCU CCC CCG UCA CCA GGG CCU GAU CCC GGG CCC
HM439775   AUG AAU GCC UCC CUC UGA GGG GAC GAC GCU CCC CCG UCA CCA GGG CCU GAU CCC GGG CCC
           *** *** *** *** *** *** *** *** *** *** ***  *** *** *** **  *** *** *** ***
```

NC_003629	Pea enation mosaic virus-1	Enamovirus	(derived from L04573)

Benyvirus RNA2

Notes on the nucleotide sequences:

1. Most of these sequences are similar to NC_003515.
2. Tobamovirus-like 'UAG CAA UUA' conserved in all sequences except EU099845.
3. Some species also have tobamovirus-like CAA as the -1 codon.
4. Possible RNA stem-loop. Some support from compensatory substitutions, but available sequence data is too limited to provide strong independent support from synonymous site conservation analysis (main text Fig 4).

An extract around the readthrough stop codon is shown below for the sequences used in the synonymous site conservation analysis (main text Fig 4), plus the additional sequence AF522462 which has only partial coverage of the readthrough CDS.

```
                                                  ((( ((( ((( (((           ) ))) ))) ))) ))
EU099845   UCG CCA GCG GAA GCA UAG GGG UAC GGG GCA GCG GCA CUC AGU GAU CUG GUU GCA GCU GCC
NC_003503  GUU CCA CGC ACC AAU UAG CAA UUA AAU GCU GCU CGA ACU AUG GCA CAU AUU CGA GCA GCA
EU330451   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
EU330452   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AJ634732   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AF197547   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AF197556   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AJ634733   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AJ634734   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AJ634735   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AJ634736   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AJ634737   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AJ634738   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AJ634739   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AJ634740   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AJ634741   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AM779752   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
EU864118   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
FM210683   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
HM117903   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCA GCG
NC_003515  AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
X04197     AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG ACG GCA CAC AUU CGA GCG GCG
AF522462   AGU CCA CCC GGA CAA UAG CAA UUA GCU GCU GCU CGG GUG GCG GCA CAC AUU CGA GCG GCG
               ***         ***     *       **  **             *   *    **   * **  **

EU099845   CAA UCA CAG GAC AUG UAC CCA GGG UAC GGU CCC GAA UGG AAU CUU CCU CCC UUU AAG CAU
NC_003503  UUA AAG GCG CGU CUG UAU CCU GGU GAC AGU AUU GAG UGG GUU GGA UGG GUG CAU AGU UAC
EU330451   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
EU330452   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AJ634732   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AF197547   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AF197556   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AJ634733   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AJ634734   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AJ634735   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AJ634736   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AJ634737   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AJ634738   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AJ634739   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AJ634740   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AJ634741   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AM779752   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
EU864118   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
FM210683   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
HM117903   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
NC_003515  AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
X04197     AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
AF522462   AAG CGG GCA CUA UUA UAU CCU GGU GAU AGU CCC GAG UGG GUU GGU UGG AAA CAU UUC UAU
                       *  **  **  **   *   **     **  ***   *             *      *
```

RefSeqs:

NC_003503	Beet soil-borne mosaic virus	Benyvirus	(derived from AF061869)
NC_003515	Beet necrotic yellow vein virus	Benyvirus	(derived from D84411)

Sequences used for Tombusviridae synonymous site conservation plots

Tombusviridae

AB044291 AB044292 AB086951 AB189943 AB189944 AB232925 AB232926 AB250684 AB250685 AB250686
AB250687 AF192772 AF266518 AF452884 AJ243370 AJ249740 AJ304989 AJ514833 AJ557804 AJ607402
AJ811998 AM711119 AY038069 AY100482 AY122286 AY163842 AY312063 AY330700 AY546104 AY579432
AY613852 AY616760 AY626780 AY830918 D00942 D12536 D86123 DQ011234 DQ083996 DQ219415 DQ227315
DQ256073 DQ339157 DQ392986 DQ922807 EF153268 EF207438 EF589670 EU081018 EU127904 EU358605
EU545828 EU835946 EU849616 FJ457015 FJ666328 FJ707484 FJ712214 FJ768020 FJ843021 FN543421
FN543466 FN543467 FN543468 FN543469 FN543470 FN543471 FN565520 GQ221829 GQ244431 GQ259476
GQ259477 GQ259478 GQ259479 GQ259480 GU138674 GU326337 GU480022 L16015 M21958 M22445 M25270
M29671 M33002 NC_003500 U20976 U55002 U62546 U72332 U80935 X02986 X14736 X15511 X58455
X62493 X83964 X85215 X85989 X86448 X87115 X94560 Y13463.

Tombusvirus

AF266518 AJ249740 AJ607402 AM711119 AY100482 AY163842 AY579432 AY830918 DQ011234 GQ259476
GQ259477 GQ259478 GQ259479 GQ259480 M21958 M25270 NC_003500 U80935 X15511 X62493 X85215

Necrovirus

AF452884 AY546104 AY616760 AY626780 D00942 DQ083996 EF153268 EU545828 FJ666328 FN543421
FN543466 FN543467 FN543468 FN543469 FN543470 FN543471 FN565520 GQ221829 GU326337 M33002
U62546 X58455 X85989 X94560

Carmovirus

AB044291 AB044292 AB086951 AB189943 AB189944 AB232925 AB232926 AB250684 AB250685 AB250686
AB250687 AF192772 AJ304989 AJ514833 AJ811998 AY122286 AY312063 AY330700 D12536 D86123
DQ219415 DQ256073 DQ339157 DQ392986 DQ922807 FJ457015 FJ707484 FJ843021 GU480022 L16015
M22445 M29671 U20976 U72332 X02986 X86448 Y13463

Drosophila* gene *kelch

```
                                                                ( (((  ((   (((  (((  ((
D. ananassae        AUC GAC AAG CCC AUG UGA AUG GAA GAG CAG GGU GCA UUG GCA CGA CAA GCG
D. erecta           AUC GAU AAG CCC AUG UGA AUG GAA GAG CAG GGU GCA UUG GCC CGC CAA GCU
D. grimshawi        AUA GAC AAG CCC UUG UGA AUG GAA GAG CAG GGC GCA UUG GCA AGG CAA GCG
D. melanogaster     AUC GAC AAG CCC AUG UGA AUG GAA GAG CAG GGU GCA UUG GCC CGC CAA GCU
D. mojavensis       AUA GAU AAG CCA UUG UGA AUG GAA GAG CAG GGC GCG UUG GCA AGG CAA ACG
D. persimilis       AUC GAU AAG CCC AUG UGA AUG GAA GAG CAG GGU GCA UUG GCA CGA CAA GCG
D. pseudoobscura    AUC GAU AAG CCC AUG UGA AUG GAA GAG CAG GGU GCA UUG GCA CGA CAA GCG
D. sechellia        AUC GAC AAG CCC AUA UGA AUG GAA GAG CAG GGU GCA UUG GCC CGC CAA GCU
D. simulans         AUC GAC AAG CCC AUG UGA AUG GAA GAG CAG GGU GCA UUG GCC CGA CAA GCU
D. virilis          AUC GAU AAG CCC UUG UGA AUG GAA GAG CAG GGC GCA UUG GCA CGC CAA GCG
D. willistoni       AUU GAU AAG CCC AUG UGA AUG GAA GAG CAG GGU GCA UUG GCA AGA CAA GCU
D. yakuba           AUC GAC AAG CCC AUG UGA AUG GAA GAG CAG GGU GCA UUG GCC CGC CAA GCU
                    **  **  *** **   *  *** *** *** *** *** **  **  *** **   *  ***  *

                       240 nt in
                    D. melanogaster                 ))) ))) ))) ))) ))
D. ananassae        GCU UCG ....... GCA AAU GCC AAU GCC AAU GCC CCU GCA AAU GCC GAG GAA
D. erecta           GCU UCG ....... GCC AAU GCC AAU GCC AAU GCC CCU GCA AAU GCC GAG GAA
D. grimshawi        GCU UCG ....... GUC AAU GCC AAU GCC AAU GCC CCU GCG AAU GUU GAA GAA
D. melanogaster     GCU UCG ....... GCC AUC GCC AAU GCC AAU GCC CCU GCA AAU GCC GAG GAA
D. mojavensis       GCC UCG ....... GCC AAU GUC AAU GCC AAU GCC CCU GCA AAU GUU GAG GAA
D. persimilis       GCU UCG ....... GCC GCC GCC AAC GUC AAU GCC CCU GCA AAU GUU GAG GAU
D. pseudoobscura    GCU UCG ....... GCC GCC GCC AAC GUC AAU GCC CCU GCA AAU GUU GAG GAU
D. sechellia        GCU UCG ....... GCC AUC GCC AAU GCC AAU GCC CCU GCA AAU GCC GAG GAA
D. simulans         GCU UCG ....... GCC AUC GCC AAU GCC AAU GCC CCU GCA AAU GCC GAG GAA
D. virilis          GCC UCG ....... GCC AAU GCC AAU GCC AAU GCC CCU GCA AAU GUU GAG GAA
D. willistoni       GCU UCG ....... --- --- --- AAU GCC AAU GCC CCU GCC AAU GUU GAG GAA
D. yakuba           GCU UCG ....... GCC AUC GCC AAU GUC AAU GCC CCU GCA AAU GCC GAG GAA
                    **  ***                     **  * * *** *** *** **  *** *   **  **
```

Drosophila* gene *headcase

```
                                                  [[[ [{{ ((( ((( ((( (
D. ananassae        CUG CCG ACC CUC GCA UAA CAU GAG GCC CCG UUG GAG UCG CCC GCU AAC GGU
D. erecta           CUG CCG ACG CUC GCA UAA CAU GAG GCC CCG UUG GAG UCG CCA GUG GGA ---
D. grimshawi        AUG CAG GCG AUC GCA UAA CAU GAG GCC CCG UUG GAG UCC CCA GCG GUC GCC
D. melanogaster     CUG CCG ACG CUC GCA UAA CAU GAG GCC CCG UUG GAG UCG CCC GUG GGA ---
D. mojavensis       CUG CAG GCG AUC GCA UAA CAU GAG GCC CCG UUG GAG ACA CCA ACG GUU GCU
D. persimilis       CUG CCG ACG CUC GCA UAA CAU GAG GCC CCG UUG GAG UCU CCG GCA GGC ---
D. pseudoobscura    CUG CCG ACG CUC GCA UAA CAU GAG GCC CCG UUG GAG UCU CCG GCA GGC ---
D. sechellia        CUG CCG ACU CUC GCA UAA CAU GAG GCC CCG UUG GAG UCG CCC GUG GGA ---
D. simulans         CUG CCG ACG CUC GCA UAA CAU GAG GCC CCG UUG GAG UCG CCC GUG GGA ---
D. virilis          CUG CAG GCG AUC GCA UAA CAU GAG GCC CCG UUG GAG CCA CCA ACG GUU GCU
D. willistoni       CUG CAG ACG CUC GCA UAA CAU GAG GCC CCG UUG GGG UCA CCA GCC --- ---
D. yakuba           CUG CCG ACG CUC GCA UAA CAU GAG GCC CCG UUG GAG UCG CCC GUG GGA ---
                     ** * *  *   ** *** *** *** *** *** *** * *  *  **

                                                                                ))) 
D. ananassae        ACU GGU ACC GCC AAU CCC ACC CAG GGA CAG GUG CCA GGC GCC CAG GGC ACU
D. erecta           --- --- --- ACA GGU GCC ACC ACC ACC CAG GUG CCA AAU GCC CAG GGC GCU
D. grimshawi        --- GGC AGU GCA ACA GCA ACA --- --- --- --- GCC AGU GCC AAU GGU GCU
D. melanogaster     --- --- --- ACA GGU GCC ACC ACC ACU CAG GUG CCA AAU GCC CAG GGC UCU
D. mojavensis       --- GGC GCA AAU GUA AAU GCG AAC GCA GUC ACA GCC AGU GGC AAU GGU GCU
D. persimilis       --- --- ACG GCA GCC GCA GCA GCA GCA GCA ACA GCC AGU GCC CCA GGU GCU
D. pseudoobscura    --- --- ACG GCA GCC GCA GCA GCA GCA GCA ACA GCC AGU GCC CCA GGU GCU
D. sechellia        --- --- --- ACA GGU GCC ACA ACC ACU CAG GUG CCA AAC GCC CAG GGC GCU
D. simulans         --- --- --- ACA GGU GCC ACC ACC ACU CAG GUG CCA AAC GCC CAG GGC GCU
D. virilis          --- GGC GCA AAU CCA AAU UCA --- --- --- --- GCC AGU GCC AAU GGU GCU
D. willistoni       --- --- --- --- --- --- --- --- --- --- --- --- --- --- UUA GGU GCU
D. yakuba           --- --- --- ACA GGU GCC ACC GCC ACU CAG GUG CCA AAC GCC CAG GGC GCU
                          } }                                                 **   **
                    ))) ))) )]] ]]
D. ananassae        CCA ACG GCC UCA AAU GGU UCC AAU AGU UGC AAC AGC AAC GGC AGC UCG UCA
D. erecta           CCA ACG GCC UCC GGC UGC AGC AGC AAC --- --- --- --- --- --- ACC AUC
D. grimshawi        UCA ACG GCC UCC AAG CAG CCA AUC UAC AAU --- --- --- --- --- GCU GCC
D. melanogaster     CCA ACG GCC UCC GGC UGC AGC AGC AAC --- --- --- --- --- --- ACC AUC
D. mojavensis       UCA ACG GCC UCC AAG CAG CCG CUA UAU AAU --- --- --- --- --- GCG GCC
D. persimilis       CCA ACG GCC UCC UCG ACG GCG AAG CAG CCG --- --- AUC UUC AAC GCG GCC
D. pseudoobscura    CCA ACG GCC UCC UCG ACG GCG AAG CAG CCG --- --- AUC UUC CAC GCG GCC
D. sechellia        CCA ACG GCC UCC GGC UGC AGC AGC AAC --- --- --- --- --- --- ACC AUC
D. simulans         CCA ACG GCC UCC GGC UGC AGC AGC AAC --- --- --- --- --- --- ACC AUC
D. virilis          UCA ACG GCC UCC AUG CAG CCG CUG UAU AAU --- --- --- --- --- GCA GCC
D. willistoni       CCA ACG GCC UCG AAA CAG CCG CUC UUU AAU --- --- --- --- --- CAA GCC
D. yakuba           CCA ACG GCC UCC GGC UGC AGC AGC AAC ACC --- --- --- --- AUC GCC AUC
                     ** *** *** **
```