Self-cleaving ribozymes conserved in RNA viruses unveil a new role in protein translation

María José López-Galiano[1], Sotaro Chiba[2], Marco Forgia[3], Beatriz Navarro[4], Amelia Cervera[1], Artem Babaian[5,6], Francesco Di Serio[4], Massimo Turina[7], Marcos de la Peña[1*]

[1] Instituto de Biología Molecular y Celular de Plantas, Universidad Politécnica de Valencia-CSIC, Valencia, Spain

[2] Graduate School of Bioagricultural Sciences, Nagoya University, Chikusa, Nagoya, Japan

[3] Institute for Sustainable Plant Protection, National Research Council of Italy, Torino, Italy

[4] Institute for Sustainable Plant Protection, National Research Council of Italy, Bari, Italy

[5] Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada

[6] Terrence Donnelly Centre for Cellular & Biomolecular Research, University of Toronto, Toronto, ON, Canada

[7] Institute for Sustainable Plant Protection, National Research Council of Italy, Brescia, Italy

*corresponding author: rivero@ibmcp.upv.es

Abstract

Small self-cleaving ribozymes are catalytic RNAs originally discovered in viroid-like agents, which are infectious circular RNAs (circRNAs) postulated as relics of a prebiotic RNA world. In the last decade, however, small ribozymes have also been detected across the tree of life, from bacterial to human genomes, and more recently, in viral agents with circRNA genomes. Here we report the conserved occurrence of small ribozymes within the linear genomes of typical ds and ssRNA viruses from fungi and plants. In most 5'-UTR regions of chrysovirids and fusarivirids, we find conserved type I hammerhead ribozymes (hhrbzs) showing efficient self-cleaving activity *in vitro* and *in vivo*. Similar hhrbzs, as well as hepatitis delta and twister ribozymes, were also detected in megabirna-, hypo-, fusagra- and toti-like viruses. These ribozymes occur as isolated motifs but also as close tandem pairs, suggesting that they are involved in the formation of ~300 nt circRNAs. *In vivo* characterization of a chrysovirid hhrbz revealed its unexpected role in protein translation as an internal ribosome entry site (IRES). RNA structural comparison between the hammerhead three-way junction and the core domain of picornavirus IRES elements allow us to suggest that these simple ribozymes may follow a similar strategy to achieve cap-independent translation. We conclude that self-cleaving ribozymes, historically involved in the rolling circle replication of viroid-like agents, have been exapted towards translational functions in linear RNA viruses.

Introduction

RNA's dual nature as carrier of genetic information and biochemical catalyst has established it as a candidate molecule for the origin of life. RNA predates DNA, proteins and the genetic code in the primordial "RNA world" hypothesis[1–3]. Genetic remnants of this hypothetical RNA world would be the extant catalytic RNAs, such as the ribosomal RNAs, the RNAse P, or the small self-cleaving ribozymes. Therefore, the simplest RNA-based entities, such as viruses, viroids and other RNA mobile genetic elements, can be regarded as the living fossils of Earth's first life[4,5]. The simplest known ribozymes belong to the family of small (50-200 nt) self-cleaving RNAs, with nine well characterized members: the hammerhead (hhrbz)[6,7], hairpin (hprbz)[8], human Hepatitis Delta (dvrbz)[9], Varkud-Satellite (vsrbz)[10], glmS[11], twister (twrbz)[12], and twister sister, hatchet and pistol[13] ribozymes. During the '80s, the first small ribozymes were described in infectious circRNAs of plants (hhrbz and hprbz in viroid-like agents), animals (dvrbz in Hepatitis Delta Virus) and fungi (vsrbz in a *Neurospora crassa* plasmid), but also in some DNA genomes[14,15]. In the last decade, these motifs have been discovered to be widespread in the DNA genomes of phages, bacteria, or eukaryotes[16–18], including the human genome[19,20]. The precise biological functions of many of these genomic ribozymes is not well understood, but numerous hhrbz, dvrbz and twrbz motifs strongly associate with autonomous and non-autonomous DNA retrotransposons across plant and metazoan genomes[21–26].

Likely due to its simplicity[27], the hhrbz is one of the most frequent ribozymes detected in nucleic acids. It is composed of three double helixes (I to III) that surround a core of 15 conserved nucleotides, and folds into a γ-shaped helical junction where the loops of helix I and II interact[28,29]. Depending on the open-ended helix, three circularly permuted topologies are possible for the hhrbz (type I, II, or III) (Fig. 1A). The dvrbz, another small ribozyme widespread among DNA genomes, shows a characteristic nested double pseudoknot structure with five helical regions (Fig. 1B).

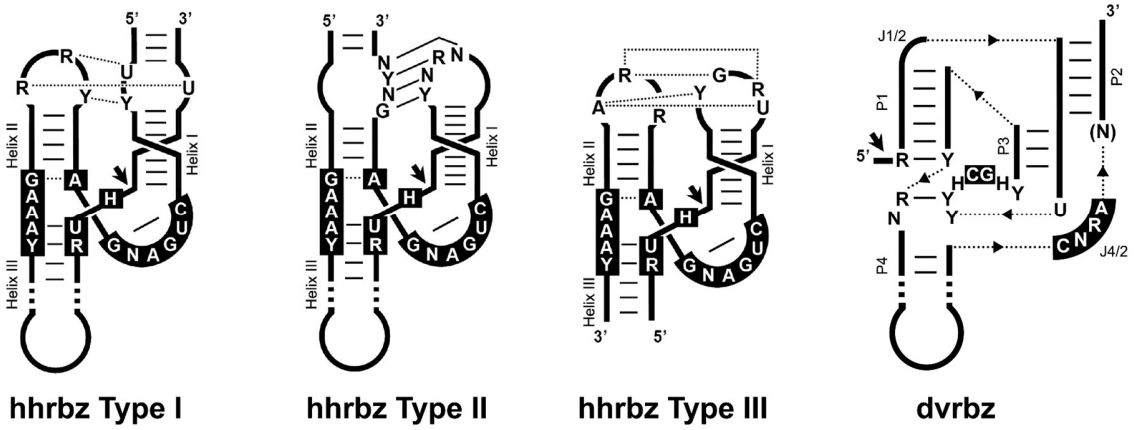


Fig. 1. Schematic representation of the three hammerhead topologies (hhrbz Type I, II, and III) and the hepatitis delta virus (dvrbz) ribozymes. The conserved nucleotides are shown in black boxes. Typical tertiary loop-loop interactions of the hhrbz are indicated (dotted and continuous lines refer to non-canonical and Watson–Crick base pairs, respectively). Dotted lines with triangles represent connections between the helixes. Helical stems and single-stranded junction strands are indicated. Black arrows indicate the self-cleavage site.

Sequencing and computing have driven an exponential growth in the characterized biodiversity of microbial life forms. This expansion has relied on the use of protein hallmark genes or ribosomal RNA. However, we are starting to appreciate the existence of an extant and highly diverse RNA world of minimal agents. Previous analyses have unveiled hundreds of novel circRNA genomes with self-cleaving ribozymes in either one or both polarities[30,31]. More recently, these minimal circRNA agents were expanded to more than 20,000 viroid-like taxonomic units, including examples of novel RNA virus-viroid hybrids such as ambiviruses and some mitoviruses with circular genomes and paired ribozymes[32,33]. This discovery blurs the distinction between viroidal agents, ribozyme-bearing satellite viruses, and RNA viruses, but confirms the existence of a diverse "modern RNA world" of mobile genetic elements with small self-cleaving ribozymes[33].

Here, we extend the conserved occurrence of small self-cleaving ribozymes, notably specific variants of the type I hhrbz, to diverse families of fungal and plant RNA viruses with linear genomes. To our surprise, these motifs do not appear to be involved in the RNA processing of intermediates during rolling-circle replication, but instead have been exapted to perform novel roles in the life cycle of linear RNA viruses.

Results

Type I hammerhead ribozymes in dsRNA viruses

In silico screening for the nine known families of small ribozyme structures in viral sequences deposited in public databases revealed the conserved occurrence of significant motifs across evolutionarily diverse RNA viruses (see Methods for full details). Notably, we identified putative ribozyme motifs in the *Chrysoviridae* family of multipartite dsRNA viruses, where 42% (96 out of 226 contigs >2 kb) of the alphachrysovirus sequences contain a distinct variant of the type I hhrbz class (Fig. 1). These motifs are found in the 5'-UTRs (positive strand) of each of the four RNA segments making up the viral genome (Fig. 2A and Supplementary Table 1). Exceptionally, a few alphachrysoviral segments exhibit a second hhrbz at the 3'-UTR region of the RNA (ie. Raphanus sativus chrysovirus 1, RNA segments 1 and 3. Supplementary Table 1). No examples of hhrbzs or any other ribozymes were detected in any of the betachrysovirus sequences analyzed (134 contigs >2 kb).

Previously described classic type I hhrbzs from eukaryotic retroelements[14,15,22,25,34] usually have a very short/absent helix III, which prevents efficient self-cleavage of monomeric but not of dimeric hhrbz molecules[6,35]. However, in the chrysoviral type I hhrbzs, the helix III is predicted to be longer and more similar to type I hhrbzs encoded in the DNA genomes of bacteria/phages[16], fungi[22] or mammals[19]. The newly uncovered hhrbzs from chrysoviruses can be categorized into two conserved architectures with distinct properties (Fig. 2B); one third of the hhrbzs have a long helix II of 6 bp capped by a tetraloop and usually harboring a bulged adenosine at the middle of the stem, whereas two thirds of the chrysoviral hhrbzs show a very short (1-2 bp stem) or even non-existent helix II (Fig. 2B and Supplementary Fig. 1).

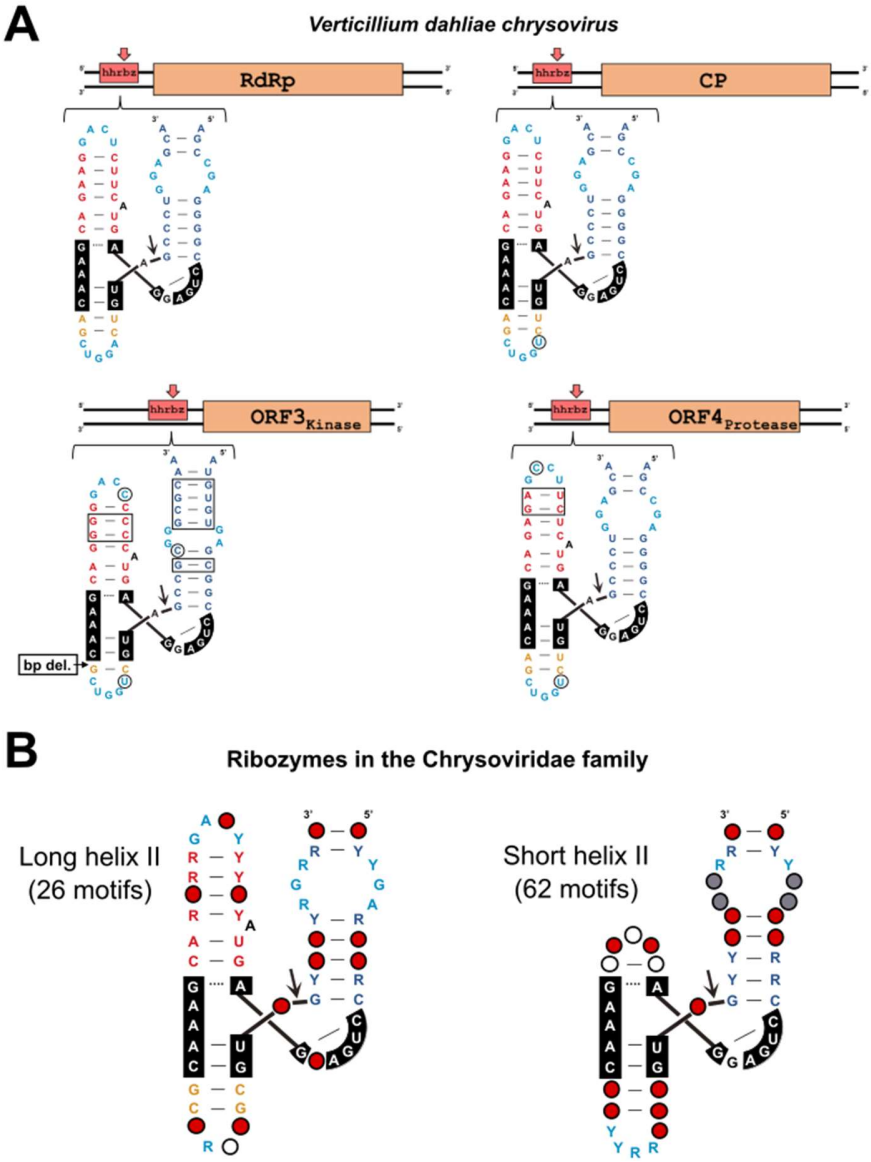


Fig. 2. Fungal and plant chrysovirus hammerhead ribozymes (hhrbz). (A) The multipartite Verticillium dahliae chrysovirus[36]contains type I hhrbzs in the 5'-UTRs of the 4 RNA segments that make up the viral genome. Nucleotide changes in the ribozyme motifs with respect to the hhrbz at the RNA 1 encoding the RNA-dependent RNA polymerase (RdRp) are indicated by circles (single changes) and rectangles (base-pair covariations) (B) A summary of the covariance models for the two hhrbz architectures detected among most chrysoviruses. Red, grey and white dots refer to nucleotides present in more than 97%, 75% or 50% of the sequences, respectively. Y: C or U. R: G or A (See Supplementary Fig. 1A for details). CP, coat protein. ORF, open reading frame.

In vitro analyses confirmed that chrysovirus hhrbzs with long helix II self-cleave efficiently during transcription (k_{obs}~2-3 min^{-1}) (Fig. 3). However, the observed rates of self-cleavage for the hhrbz with short (1 bp) or even absent helix II were low (k_{obs} ~10^{-2} min^{-1}), or very low (k_{obs} ~10^{-3} min^{-1}) respectively, either under co-tanscriptional

or post-transcriptional conditions, even at high pH (8.5) and Mg^{+2} concentrations (10 mM) (Fig. 3, Supplementary Fig. 2).

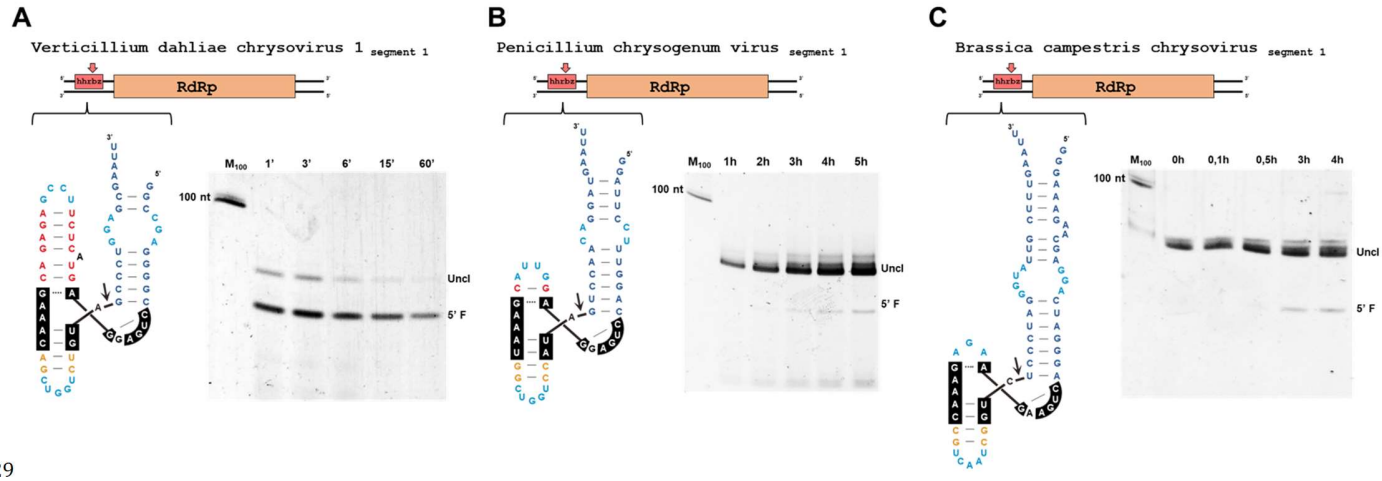


Fig. 3. Self-cleavage kinetics of hammerhead ribozymes from fungal and plant chrysoviruses. (A) long helix II, (B) short helix II, or (C) totally absent helix II hammerhead ribozymes are catalytically competent *in vitro* but show distinct self-cleaving efficiencies. RdRp, RNA-dependent RNA polymerase. (A) and (B) correspond to co-transcriptional cleavage analysis under standard conditions (pH 7.5, 1 mM Mg^{2+}), whereas (C) shows a post-transcriptional assay performed at higher pH (8.5) and Mg^{2+} concentration (10 mM) to increase its self-cleaving capabilities.

Other ds and ssRNA viruses encode hammerhead, but also deltavirus and twister ribozyme motifs

Fusariviruses are a family of mycoviruses with monopartite ssRNA genomes that encode one to four ORFs[37,38]. Our bioinformatic analyses found that up to 114 out of 156 analyzed fusariviral sequences encode one or several type I hhrbzs each (160 hhrbz hits in total) (Supplementary Table 1). The sequences and helix sizes of the motifs are quite variable depending on the viral genome, but as described for chrysoviruses, they either follow a canonical type I hhrbz architecture, with medium size helixes I, II and III (4-8 bp stems), or a variant with a very short helix II (1-2 bp stem) (Supplementary Fig. 1). These observations suggest a high and low self-cleavage efficiency, respectively, for each of the two fusarivirus hhrbz architectures. Again, fusariviral ribozymes occur in the untranslated regions preceding most of the viral ORFs (Fig. 4A and Supplementary Fig. 3A). We made similar observations in the family

150 of bipartite dsRNA megabirnaviruses[39], which show the presence of slightly different
151 variants of the type I hhrbz with longer helix II (Supplementary Fig. 3B). The motifs
152 occur again in the large 5'-UTR of diverse megabirnavirus RNA segments (17 out of 32
153 analyzed contigs) (Supplementary Table 1). Interestingly, not only hhrbz motifs, but
154 also an example of a *bonafide* dvrbz was detected in a genomic segment of the
155 Rhizoctonia cerealis megabirnavirus-like RNA (Fig. 4B, Supplementary Fig. 2B).
156 Similarly, instances of dvrbz, as well as type I hhrbz and even examples of twrbz motifs,
157 were detected in the dsRNA genomes of diverse totiviruses. In these examples,
158 however, the motifs were found as pairs of ribozymes (either from the same or
159 different classes) placed in tandem very close to each other (~200-400 nt) (Fig. 4C).
160 This arrangement resembles the one observed in plant and animal retrozymes, where
161 close tandem ribozyme pairs result in the formation of small circular RNAs[24,25]. In the
162 case of totiviruses, the predicted RNA circles can adopt highly stable secondary
163 structures, have nucleotide sizes multiple of 3, and may encode a never-ending
164 polypeptide in their positive polarity (Fig. 4C, Supplementary Fig. 4). The features of
165 these sub-genomic circRNAs are strikingly similar to those described for the intriguing
166 group of viroid-like Zeta viruses detected in diverse environmental
167 metatranscriptomic studies[30,33] (Supplementary Fig. 4). However, most of the
168 predicted circRNAs in totiviruses are not keeping in their sequence any of the two
169 flanking ribozymes (usually a type I hhrbz and a dvrbz at the 5' and 3' ends,
170 respectively), suggesting that they would not be capable of replication through a classic
171 rolling circle mechanism.

172 Finally, our bioinformatic analyses also revealed the presence of analogous type I
173 hhrbzs in the UTRs of some dsRNA genomes from additional fungal viral families,
174 including fusagraviruses and hypoviruses (Supplementary Table 1). Altogether, these
175 results confirm the prevalence of small self-cleaving ribozymes encoded within the
176 linear ss and dsRNA genomes of diverse fungal and plant viruses.

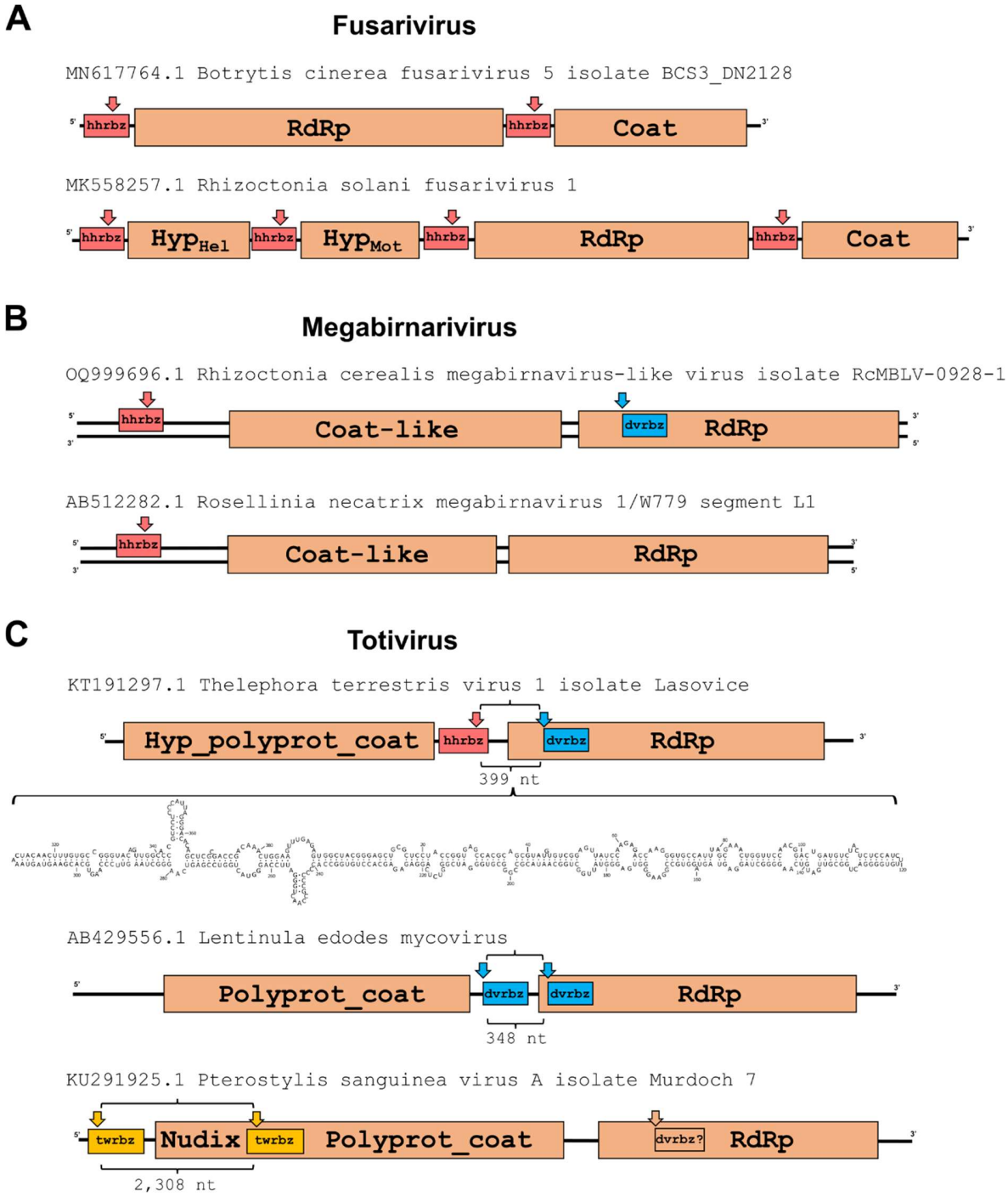


Fig. 4. Small ribozymes are widespread in fungal ss and dsRNA viruses with linear genomes. (A) fusariviruses, (B) megabirnaviruses and (C) totiviruses encode conserved self-cleaving ribozymes of the hhrbz, dvrbz and twrbz classes. In the case of diverse totivirus genomes, pairs of hammerhead and delta self-cleaving ribozymes occur very close to each other (~300 nt). Similar arrangements of close ribozymes in tandem involved in circRNA formation have been previously found encoded in plant (~600 nt[24]) and metazoan (~300 nt[25]) genomes, which suggests that ribozyme pairs in totivirus could allow the formation of similar circular RNAs, as depicted for Telephora terrestris virus 1 (See Supplementary Fig. 4 for more details).

186

187 ## Hammerhead ribozymes in RNA viruses show self-cleaving activity *in vivo*

188 Given the widespread occurrence of type I hhrbzs among chrysovirids and fusarivirids,
189 we chose to look for *in vivo* evidence of ribozyme self-cleaving activity during viral
190 infection. We selected two chrysoviruses (either infecting a plant or a fungus) and one
191 fungal fusarivirus as examples of ds and ssRNA viruses with hhrbzs in their 5'-UTRs,
192 respectively.

193 Brassica campestris chrysovirus 1 (BcCV1) is a tripartite dsRNA chrysovirus infecting
194 brassica plants[40], which shows conserved hhrbzs (short helix II variants with a minimal
195 self-cleaving activity *in vitro*, see above) in the 5'-UTRs of each RNA segment. The
196 expected cleavage sites of the hhrbzs map at positions 58, 60 and 60 of RNA segments
197 1, 2 and 3, respectively (Supplementary Fig. 5A). Rapid amplification of RNA ends
198 (RACE) experiments using total RNA from a BcCV1-infected *B. oleracea* plant generated
199 two amplification products for each of the chrysovirus RNAs, which are more evident
200 in the case of RNA2 and RNA3 (Supplementary Fig. 5B). Cloning and sequencing of the
201 amplicons confirmed the existence, for each genomic RNA segment, of molecules with
202 two different 5' ends, one corresponding to the expected viral full-length RNA segment,
203 and the other one with the self-cleavage site predicted for the respective embedded
204 ribozyme (Supplementary Fig. 5C).

205 The recently characterized Gnomognopsis castanea chrysovirus 1 (GcCV1) has 4
206 genomic segments encoding an ORF each, with the exception of the bicistronic RNA3[41].
207 We detected the presence of 5 different hhrbz motifs (long helix II variants with
208 efficient self-cleaving activity *in vitro*, see above), one in each 5'-UTR preceding an ORF
209 (Supplementary Fig. 6). Here, RACE experiments carried out on total RNA extracted
210 from an infected fungus isolate showed that the population of viral RNAs is also
211 composed of RNA segments with two different 5' termini, as supported by clones either
212 mapping to the expected 5' end of the genomic RNA, or to the predicted 5' site of self-
213 cleavage of each hhrbz (Supplementary Fig. 6). Such a mix of RNA molecules with two
214 different ends is particularly evident for RNA2, which gives two abundant and distinct

215 bands in PCR RACE amplifications (Supplementary Fig. 6E). It is of note that, in this
216 virus, the two copies of hhrbz present in each one of the two UTRs of the bicistronic
217 RNA3 were found to self-cleave *in vivo* at their predicted sites.

218 Finally, Pleospora tiphycola fusarivirus 1 (PtFV1) is a ssRNA virus present in an
219 ascomycetes fungus isolated from the seagrass *Posidonia oceanica*[42]. Members of the
220 recently established *Fusariviridae* family are a group of non-segmented viruses with
221 heterogeneous genome organization according to the genus they belong to[38]. The RNA
222 genome of PtFV1 can potentially express up to three proteins from three distinct
223 ORFs[42]. It also contains a predicted hhrbz with a short helix II, quite similar to the
224 Penicilium chrysogenum virus ribozyme (Fig. 3B). The catalytic RNA essentially spans
225 the entire small intergenic UTR (66 nt), with the self-cleavage site (position 4,735) just
226 7 nucleotides upstream of the proposed ORF2 start codon (position 4,742). RACE
227 analysis using reverse transcription products primed around 200 nt downstream the
228 cutting site confirmed the existence *in vivo* of an RNA species with the predicted 5' end
229 after self-cleavage (Supplementary Fig. 7A). Moreover, northern blot analysis showed
230 the accumulation of a positive-sense RNA fragment of a size compatible with that of a
231 subgenomic RNA originated by self-cleavage of the ribozyme (Supplementary Fig. 7B).
232 Overall, experiments with total RNA purified from either infected plants or fungi
233 followed by RACE analyses confirm that the population of viral RNAs *in vivo* is
234 heterogeneous, being composed of uncleaved and self-cleaved RNA molecules.

235

236 Hammerhead ribozymes show a role in cap-independent translation 237 initiation

238 Translation initiation activity, presumably via an internal ribosome entry site (IRES)
239 element, has been reported *in vivo* for the 5'-UTR and subsequent 72 nt of coding region
240 of fungal dsRNA viruses such as chrysoviruses and hypoviruses[43]. Among these, the 5'-
241 UTR of the alphachrysovirus Helminthosporium victoriae virus 145S (HvV145S) RNA2
242 was found to carry an IRES element. We found that, as most alphachrysoviruses,

243 HvV145S RNA segments also contain a conserved type I hhrbz motif at their 5'-UTRs
244 (Supplementary Table 1, Fig. 5A). These ribozymes correspond to the hhrbz variants
245 with a long helix II (Fig. 2B), which are expected to reach significant self-cleaving
246 activity *in vitro* and *in vivo*. To understand whether these type I hhrbzs are required to
247 promote internal initiation of protein translation, on- and off-target mutations were
248 introduced into the self-cleaving motif and analyzed *in vivo*. The transgenic expression
249 of bicistronic RNAs in *C. parasitica* mycelia was carried out as previously described[43],
250 in which the upstream ORluc and downstream OFluc were translated in cap- and IRES-
251 dependent manner, respectively (Fig. 5B). The original HvV145S-2 5'-UTR sequence
252 and its truncated variants in the 5' proximal region preceding the ribozyme (Δ60 and
253 Δ120) showed similar ratios of OFluc to ORluc chemiluminescence intensity (RLU),
254 suggesting that this off-target 5' region is dispensable for the IRES-mediated
255 translation initiation (Fig. 5D). In contrast, the deletion of the whole region
256 corresponding to the hhrbz motif (Δ124-183) abolished the IRES-mediated translation
257 initiation of OFluc, resulting in a significantly lower OFluc/ORluc RLU ratio. In addition,
258 either a point mutation in G159, the general base catalyst in the cleavage reaction[29], or
259 deletion of 5 essential nucleotides, including the conserved U-turn motif[44] of the
260 conserved core of the ribozyme (Δ134-138), abolished the translation initiation ability
261 observed for the wt 5'-UTR sequence with a hhrbz (Fig. 5C and D).

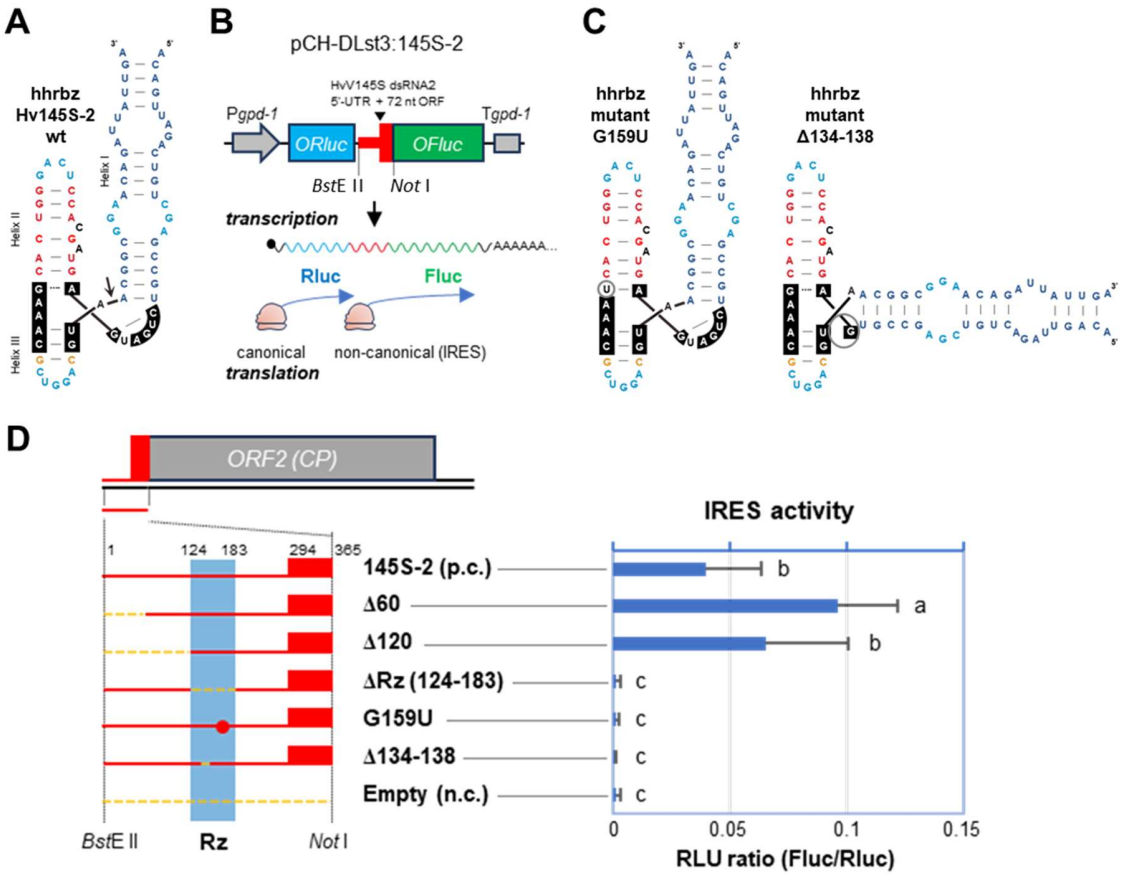


Fig. 5. Mutational analyses of the hammerhead ribozyme present in the 5'-UTR of the chrysovirus HvV145S RNA2. (A) Predicted structure of the HvV145S hammerhead ribozyme (hhrbz) present in the 5'-UTR of the RNA2. (B) A bicistronic dual-luciferase reporter system in fungal mycelia for investigation of the IRES function located at the 5'-UTR and coding region (72 nt) of HvV145S dsRNA2. Codon-optimized Renilla luciferase (ORluc, blue box) and firefly luciferase (OFluc, green box) genes are translated either in a canonical (cap-dependent) or a non-canonical (IRES-dependent) manner, respectively. Red line and box indicate HvV145S-2 sequence. (C) Predicted secondary structures of the 145S-2 hhrbz mutants containing either a substitution (left, G159U) or a deletion (right, Δ134-138), as indicated with grey circles. (D) Left, schematic representation of the 5'UTR and first 72 nt of coding region (in red) of HvV145S-2 (in grey) used for the bicistronic dual-luciferase reporter system. The details of mutations are represented at the bottom as the sequence variant names. Right, dual luciferase

274 reporter assay results for each construct showing the ratio of luminescent intensities (Fluc RLU/Rluc
275 RLU=IRES activity). p.c., positive control, n.c., negative control.

276 Discussion

277 Almost four decades after their discovery, the origin and evolutionary history of small
278 self-cleaving ribozymes remains a complex puzzle. It is assumed that RNAs capable of
279 sequence-specific cleavage and ligation could be among the first catalytic motifs to
280 emerge during the RNA world, with a role in processing primordial RNA and/or
281 circRNA genomes during replication[45]. Our knowledge about small ribozymes has been
282 mostly restricted to those from a tiny family of viroid-like circRNAs (~10-20 species),
283 regarded as reminiscences of prebiotic genomes. In the last years, however, small
284 catalytic RNAs have been found as ubiquitous motifs encoded in DNA genomes of
285 organisms throughout the tree of life[16,46] and, more recently, as part of thousands of
286 novel viroid-like circRNAs with unknown eukaryotic and, likely, prokaryotic hosts[30–
287 33]. Thus, this plethora of catalytic motifs are expected to have more biological roles
288 beyond multimer self-processing.

289 In this work, we have expanded our knowledge about the occurrence of small
290 ribozymes to the linear genomes of typical ss- and dsRNA viruses of fungal origin.
291 Notably, novel variants of the type I hhrbz are found to be conserved in many 5'-UTRs
292 of fungal and plant RNA viruses of the *Chrysoviridae* and *Fusariviridae* families. It
293 should be highlighted that nearly all hhrbz previously detected in infectious circRNAs,
294 such as plant viroids and viral satellites[47], obelisks[48], deltaviruses[49], ambiviruses or
295 mitoviruses[32,33], correspond to the type III architecture, whereas type I hhrbzs were so
296 far DNA-encoded motifs found in either bacteriophages and bacteria[16,17,50] or
297 eukaryotes (mostly in metazoan retroelements)[22,25,34]. On the other hand, natural
298 hhrbzs adopt a similar three-way junction fold with loop-loop interactions between
299 helixes I and II (Fig. 1), which are crucial to achieve high activity under *in vivo* low
300 magnesium concentrations [28,29,51]. The viral type I hhrbzs described in this work,
301 however, do not seem to have kept any of the typical loop-loop interactions conserved
302 in known hhrbzs (Fig. 1) or in other three-helical RNAs with remote tertiary contacts[52].

303 Moreover, stem lengths of helix II in viral hhrbzs are either slightly longer (6-8 bp) or
304 shorter (0-2 bp) (Fig. 2B) than the ones present in most type I hhrbzs (4-5 bp). These
305 observations, together with a medium to low *in vitro* self-cleaving activity for viral
306 hhrbzs, point to a lack of extensive loop-loop interactions in these motifs. These viral
307 hhrbzs are however functional *in vivo*, as revealed by the co-existence of uncleaved and
308 cleaved RNA genomes in plant and fungal infected cells, indicating that viral sub
309 genomic RNAs can arise through a novel mechanism catalyzed by small self-cleaving
310 ribozymes.

311 Regarding their biological role, most of the hhrbzs reported here are small motifs (50-
312 70 nt), which nevertheless span a significant portion of the viral UTR sequences (~100-
313 300 nt) where they are embedded. It is known that, among other functions, viral 5'-
314 UTRs such as the one from HvV145S RNA2 (293 nt) can work as IRES elements[43].
315 Deletion mutants of this UTR show that the hhrbz motif (nucleotides 124-182) is
316 required for IRES function. More intriguingly, a single mutation in the catalytic G base
317 within the hhrbz core, which is known to inhibit RNA catalysis without dramatically
318 affecting the whole ribozyme structure[53], was enough to drastically reduce the levels
319 of protein expression obtained with the wt ribozyme-containing UTR. This result
320 suggests that not only the ribozyme structure, but also its self-cleaving activity, would
321 be crucial for the role of the UTR as an IRES. Consequently, we hypothesize that viral
322 type I hhrbzs act as an IRES due to a combination of both ribozyme structure and self-
323 cleaving activity, which may allow protein translation through an unknown
324 mechanism. As a second alternative, we can also envisage that any mutation preventing
325 self-cleavage is indeed affecting the global fold of the hhrbz, which could disrupt its
326 role as an IRES. This possibility suggests that RNA self-cleavage could not be required
327 for IRES function but, more likely, as negative regulator of translation. In either case,
328 the central core of the hhrbz shows a clear secondary and tertiary structural similarity
329 with the core domain of diverse IRES elements from picornaviruses[54] (Fig. 6).
330 Regarding the 5'-UTR of the encephalomyocarditis virus (EMCV), it contains a type II
331 IRES known to engage the translation initiation factor eIF4G through a three-helical
332 motif known as the J-K region[55]. The structure of this motif shows a Y-shaped fold

without loop-loop interactions[56] but with a central core comparable to the one predicted for the hhrbz[29] (Fig. 6). The 5'-UTR of the Hepatitis A virus (HAV) has also a type III IRES composed of domains II to VI[57], with domain V (dV) as a core region structurally very similar[58] to the J-K region from EMCV or the hhrbz, and also capable of recruiting translation initiation factors such as eIF4G and eIF4F[59] (Fig. 6). As a last connection between protein translation and RNA self-cleaving motifs, it should be noted that the dvrbz present in the 5'-UTR of diverse eukaryotic retrotransposons has been found to promote protein translation initiation *in vitro* and *in vivo*[21]. Despite the structural differences between hhrbz and dvrbz (Supplementary Fig. 8), we cannot rule out a common and more general mechanism in protein translation for these and other self-cleaving RNAs.

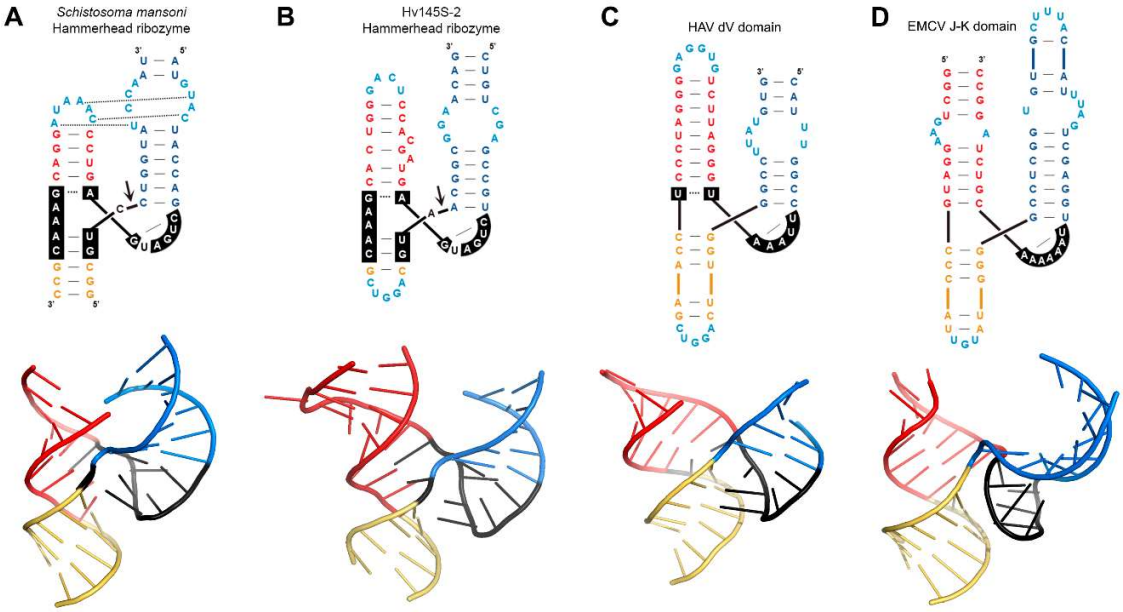


Fig. 6. (A) On the top, schematic representation of a type I hammerhead ribozyme (hhrbz) from *Schistosoma mansoni*. Tertiary loop-loop interactions are shown with dotted lines. On the bottom, 3D-structure (pdb 3ZD5) of the central core and the three stems, but not the loops, of this hhrbz (conserved core in black and stem I, II and III in blue, red and yellow, respectively). As above, similar schematic (top) and 3D structures (bottom) of (B) the Hv145S-RNA2 hhrbz (predicted through the Alphafold 3 server[60]), (C) Hepatitis A virus (HAV) IRES domain V (pdb 6mwn)[58] and (D) the J-K domain of the encephalomyocarditis virus (EMCV) IRES element (pdb 2nbx)[56].

Our discovery unveils an unexpected link between two unrelated kinds of ribozymes, such as simple self-cleaving RNAs and the complex ribosome. It could be argued that

355 this link may represent a primordial "interaction" dating back to the development of
356 the genetic code, where the common presence of small self-cleaving ribozymes would
357 be a selection force for their recognition by proto-ribosomes. Nevertheless, our
358 observations are restricted for the moment to "modern" eukaryotic viruses, which are
359 expected to have emerged after the origin of the eukaryotic cell itself (~2-3 billion
360 years ago[61]). Moreover, the structural similarities of viral hhrbzs with some
361 picornaviral IRES point to a recent event of domestication of the hhrbz to hijack
362 translation initiation factors in eukaryotes instead of direct recruitment of the
363 ribosome, as described for Hepatitis C virus (HCV) or cricket paralysis virus (CrPV)
364 IRES elements[62,63]. While these groundbreaking findings offer valuable insights into
365 RNA viral biology and potential biotechnological applications, further research will be
366 needed to elucidate this intriguing viral mechanism of cap-independent translation
367 mediated by small self-cleaving ribozymes.

368

369 # Methods

370 ## Bioinformatic analysis of viral sequences

371 Computational searches through INFERNAL 1.1 software[64] were carried out using
372 covariance models of well-known small ribozymes[33,65]. A first screening of a collection
373 of 852 thousand contigs (between 2,000 and 50,000 bp) of DNA or RNA viral origin
374 retrieved from the Genbank NR database resulted in hundreds of ribozyme hits that
375 were manually inspected. Hits present in either viral vectors, patented sequences, DNA
376 bacteriophages[16] or in RNA ambiviruses and mitoviruses[33] were discarded, whereas
377 those hits with low E-values such as type I hammerhead (E-val < 10^{-3}), deltavirus (E-
378 val < 0.1) or twister (E-val < 10^{-3}) ribozymes detected in RNA viruses from diverse
379 families were used to build new covariance models and re-screenings
380 (https://github.com/delapenya/RNA_virus_rbz). Analysis of specific RNA viral
381 families showing the presence of small ribozymes and refinement of covariance models
382 was performed iteratively.

383 The consensus features of the ribozyme motifs were calculated based on the obtained
384 ribozyme sequences for each family. Nucleotide frequencies and covariation, as well as
385 the corresponding drawing, were the output of the R2R software[66]. Multiple sequence
386 alignments were performed using MUSCLE[67].

387

388 ## Cloning of ribozyme sequences and transcriptions

389 Sequences corresponding to the hhrbz motifs of the RNA 1 segments of Verticillium
390 dahliae chrysovirus 1, Penicilium chrysogenum virus and Brassica campestris
391 chrysovirus 1 (Fig. 3) preceded by the T7 RNA polymerase promoter were purchased
392 as gBlock Gene Fragments (Integrated DNA Technologies). DNA fragments were cloned
393 into a linearized pUC18 vector by BamHI and EcoRI cohesive-end ligation. RNAs of the
394 cloned sequences were synthesized by *in vitro* run-off transcription of the linearized
395 plasmids containing the hhrbzs. The transcription reactions contained 40 mM Tris–
396 HCl, pH 8, 6 mM $MgCl_2$, 2 mM spermidine, 0.5 mg/ml RNase-free bovine serum albumin,
397 0.1% Triton X-100, 10 mM dithiothreitol, 1 mM each of ATP, CTP, GTP, and UTP, 2 U/µl
398 of Ribonuclease Inhibitor (Takara Inc), 20 ng/µl of plasmid DNA, and 4 U/µl of T7 RNA
399 polymerase. After incubation at 37°C during the indicated time, the products were
400 fractionated by 10% polyacrylamide gel electrophoresis (PAGE) in the presence of 8 M
401 urea.

402

403 ## Kinetics of self-cleavage

404 Analyses of hhrbz self-cleavage activity under co-transcriptional and post-
405 transcriptional conditions were performed as previously described[68]. Appropriate
406 aliquots of the transcription reactions (smaller volumes were taken at longer
407 incubation times in co-transcriptional assays) were removed at different time
408 intervals, quenched with a fivefold excess of stop solution (8 M urea, 50% formamide,
409 50 mM EDTA, 0.1% xylene cyanol and bromophenol blue) at 0°C, and analyzed as
410 previously described[22,49]. For co-transcriptional cleavage analysis, the uncleaved and

411 cleaved transcripts obtained at different times were separated by PAGE in 10%
412 denaturing gels and detected by Sybr Gold staining (Thermo Fisher Scientific). For
413 post-transcriptional cleavage analysis, uncleaved primary transcripts were eluted by
414 crushing the gel pieces and extracting them with phenol saturated with buffer (Tris–
415 HCl 10 mM, pH 7.5, ethylenediaminetetraacetic acid [EDTA] 1 mM, sodium dodecyl
416 sulfate 0.1%), recovered by ethanol precipitation, and resuspended in deionized sterile
417 water. To determine the cleaving rate constants, uncleaved primary transcripts (from
418 1 nM to 1 μM) were incubated in 20 μl of 50 mM Tris–HCl at the appropriate pH (either
419 7,5 or 8,5) for 1 min at 95 °C and slowly cooled to 25 °C for 15 min. After taking a zero-
420 time aliquot, self-cleavage reactions were triggered by adding $MgCl_2$ at a final
421 concentration of either 1 or 10 mM. Aliquots were removed at the appropriate time
422 intervals and quenched with a fivefold excess of stop solution at 0 °C. Substrate and
423 self-cleaved products were separated by PAGE in 10% denaturing gels and detected by
424 Sybr Gold staining (Thermo Fisher Scientific). In both cases, the product fraction at
425 different times, Ft, was determined by quantitative scanning of the corresponding gel
426 bands and fitted to the equation $F_t=F_\infty(1 - e^{-kt})$, where F_∞ is the product fraction at the
427 reaction endpoint, t is the time in minutes, and k is the first-order rate constant of
428 cleavage (k_{obs}).

429

430 **RACE and northern blot analyses**

431 Analysis of fusarivirus-infected fungi was carried out using the isolate MUT4379 from
432 the Mycotheca Universitatis Taurinensis. Gnomognopsis castanea chrysovirus 1
433 infected isolate 1 was recently described[41]. A symptomless plant of cabbage "Torzella"
434 (*Brassica oleracea* L. var. acephala) infected with BcCV1 (isolate TC-3) was used for
435 plant chysovirus 5' RACE analysis[69].

436 Plant and fungal total RNA was extracted from leaves from cabbage "Torzella" plants
437 and 4-day old mycelial liquid cultures grown in PDB following protocols previously
438 described[41,69]. Hirtzman's protocol for RACE was performed as previously described in
439 detail[70]. For each RACE two primers were designed downstream the predicted

440 ribozyme cleavage site. Regarding GcCV1, the primers used to demonstrate the *in vivo*
441 activity of the ribozymes found in the 5' end of each genomic segment are reported in
442 a previous work[41]. For the RACE demonstrating the *in vivo* activity of the ribozyme
443 found in the intergenic region of the RNA3 of GcCV1, two RACE primers were designed
444 (Supplementary Table 2); the oligonucleotides used to prime RACE for PtFV1 and
445 BcCV1 are also displayed in Supplementary Table 2. For each RACE a minimum of 5
446 clones were sequenced.

447 Northern blot of fungal extracts for MUT4379 was performed as previously described[71]
448 using a riboprobe (detecting the positive sense viral RNA) obtained by *in vitro*
449 transcription of a cDNA fragment cloned into pGEM-T-Easy (PROMEGA, Madison, WI)
450 corresponding to the genomic region between nt 4,800 and nt 4,980 of the deposited
451 fusarivirus sequence (NCBI NC_028470).

452

453 **Dual-luciferase (DL) reporter assay**

454 The DL assay using fungal mycelia was performed as previously described[43]. Briefly,
455 the pCH-DLst3 vector and PCR-amplified fragments of 5'-UTR and adjacent 72 nt of
456 coding sequence (5'-UTR+) from HvV145S dsRNA2 were digested with BstE II and Not
457 I, and ligated. The clone obtained was used for transformation of *Chryphonectria*
458 *parasitica* EP155, and mycelia of transformants were subjected to the DL assay as
459 described[43].

460 The vector pCH-DLst3 carries a codon-optimized Renilla luciferase (ORluc) and a
461 Firefly luciferase (OFluc) gene in a bicistronic manner, under the control of the gpd-1
462 promoter/terminator system. BstE II and Not I recognition sites in between ORluc and
463 OFluc were used for inserting the 5'-UTR and adjacent 72 nt coding cDNA sequence (5'-
464 UTR+; original and mutant variants), expecting that ORluc and OFluc are translated
465 cap- and IRES-dependently, respectively. *C. parasitica* EP155 was transformed with the
466 different constructs and mycelia of transformants were subjected to DL assay using a
467 dual-luciferase reporter assay system (Promega). A small patch of mycelia was
468 homogenized in a microtube with 200 µl of Cell-Lysate solution, and 10 µl of solution

469 were mixed with 100 µl of the firefly luciferase substrate solution, then the first
470 luminescence intensity was measured using GloMax 20/20 (Promega) with a detection
471 time of 1 s. This was followed by the addition of 100 µl of the Renilla luciferase
472 substrate solution, and the second luminescence measurement. IRES activities were
473 calculated as a ratio of Fluc RLU (relative luminescent unit)/ Rluc RL.

474

475 ## Supplementary Figures and Tables

476 **Supplementary Table 1.** Accession numbers of all the RNA virus genomes/segments
477 containing self-cleaving ribozymes identified in this study.

478 EXCEL FILE

479

480 **Supplementary Table 2.** Oligonucleotide sequences and their specific use in this
481 work.

482

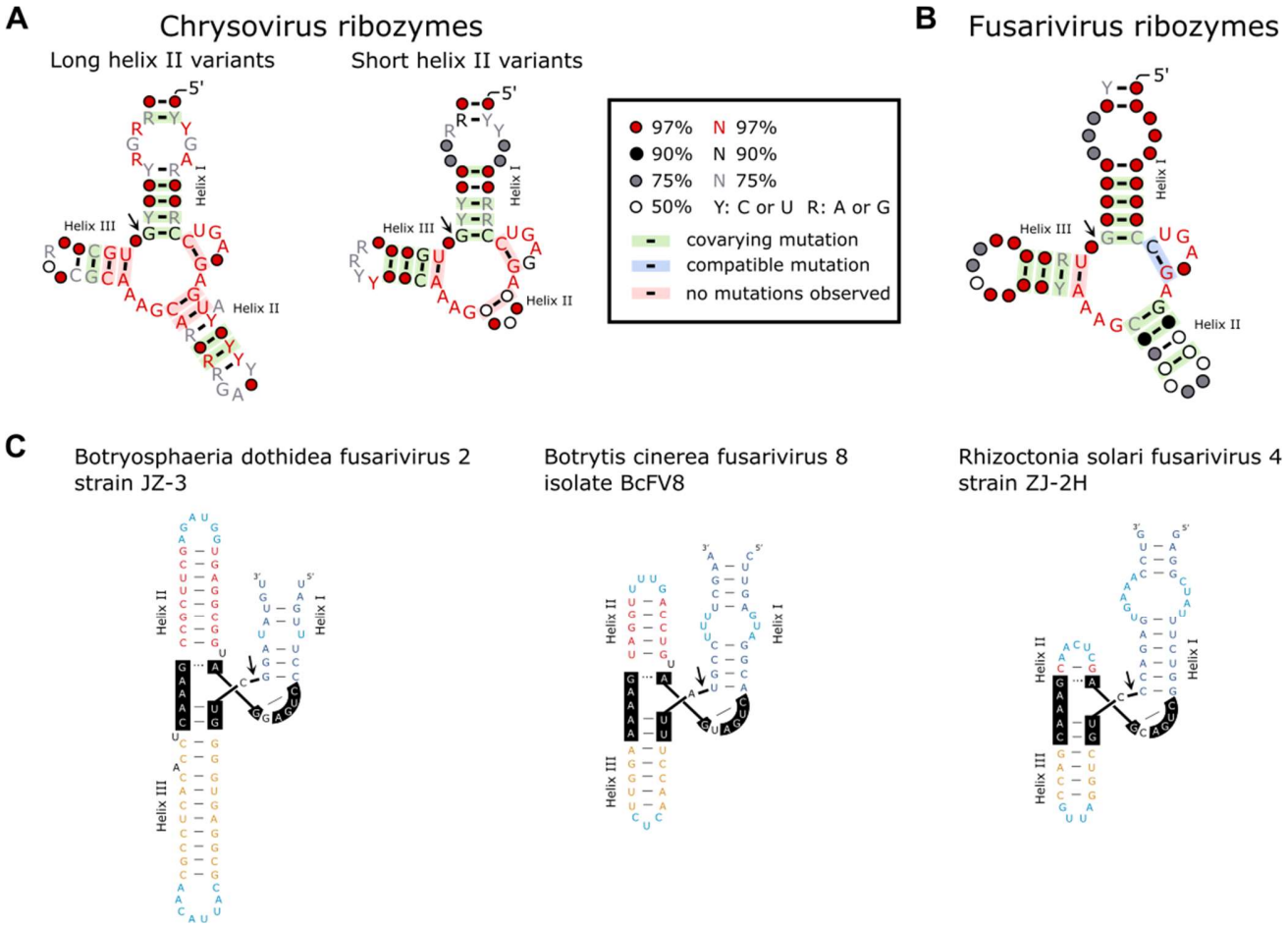
Oligonucleotide Name	Sequence (5' to 3')	Use
BcCV1-RNA1-77Rev	CAAAGATGCGTTTTCCCATGC	RACE
BcCV1-RNA1-87Rev	TGCTTTGCAGCACCGTTTA	RACE
BcCV1-RNA1-88NRev	TCGTCTGTGTAGCAACACTAAC	RACE
BcCV1-RNA2-81Rev	GCGCAAGCAACCAGGTTCA	RACE
BcCV1-RNA2-82Rev	CCTTCCATTGTGCTATTAAGGTAT	RACE
BcCV1-RNA3-85Rev	CTCTTCATGTTGCATGAATTCG	RACE
BcCV1-RNA3-86Rev	TGCAGACGTGTTTTCCATTGT	RACE
Fusari-5000Rev	TGTTTCCTTGACGCTCGTGT	RACE
Fusari-4800For	ACCCGGAGAGCTCAAATGTT	RACE
Fusari-4980Rev	GTAACTTAAGAAGTTTGGCT	RACE
GcCV1_RNA3_1163Rev	AATCTCTCCGCCGCTGTCGT	RACE
GcCV1_RNA3_1104Rev	TGTTCTCCCGTTCCTCGGTG	RACE

483

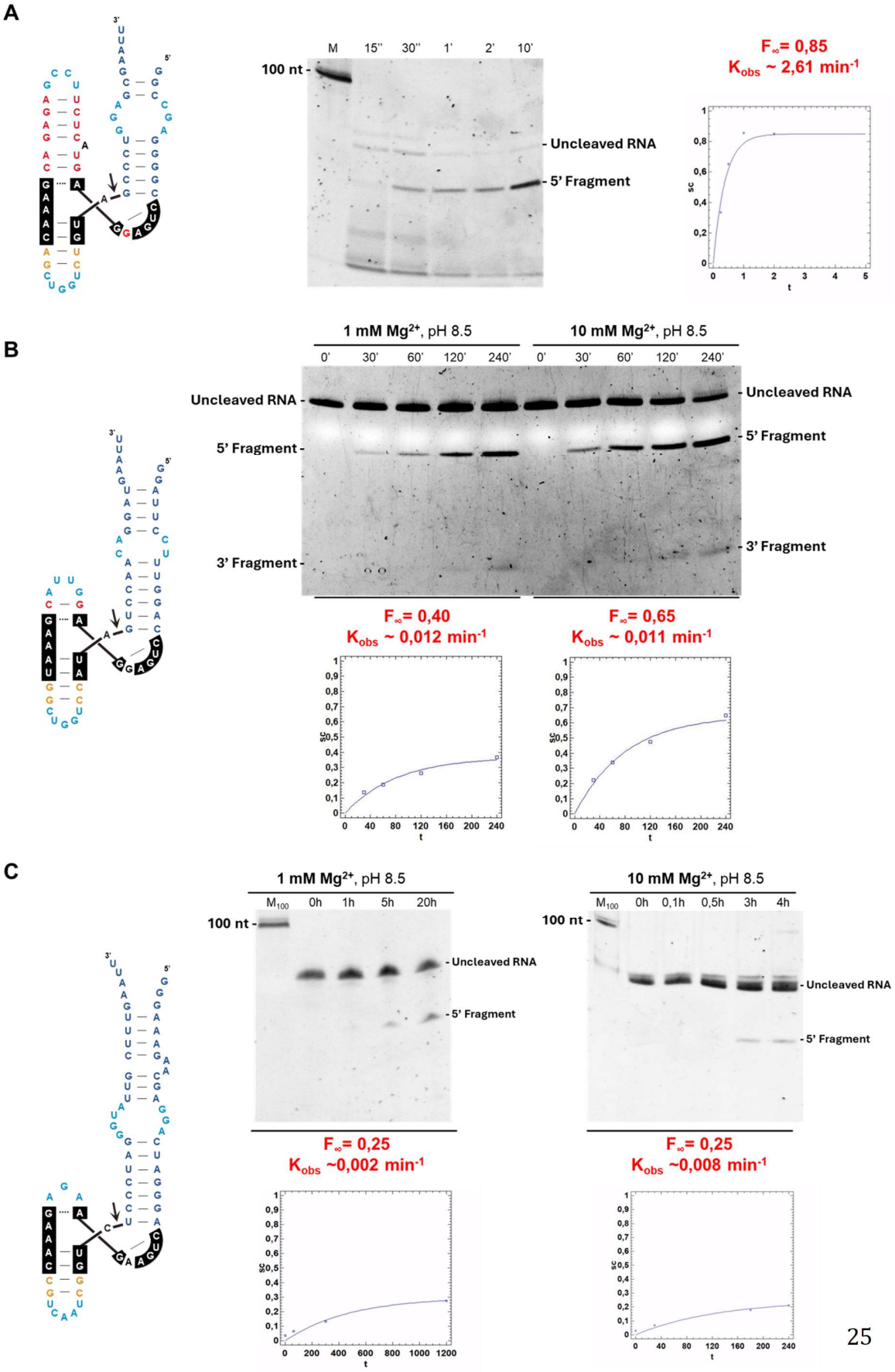
484

Supplementary Figures

Supplementary Fig. 1. (A) Detailed consensus models and secondary structures of the hammerhead ribozymes either with long helix II (left, based on 26 ribozyme sequences) or short helix II (right, based on 62 ribozyme sequences) from chrysovirus genomes. (B) Consensus model of the hammerhead ribozyme present in most fusarivirus genomes (based on 123 ribozyme sequences). The legend inset applies to the three models shown. (C) Three representative examples of hammerhead ribozymes from fusarivirus genomes with variable stem lengths for helix I, II and III. The site of ribozyme-mediated RNA cleavage is identified by an arrowhead. Consensus models were built as previously described[66].

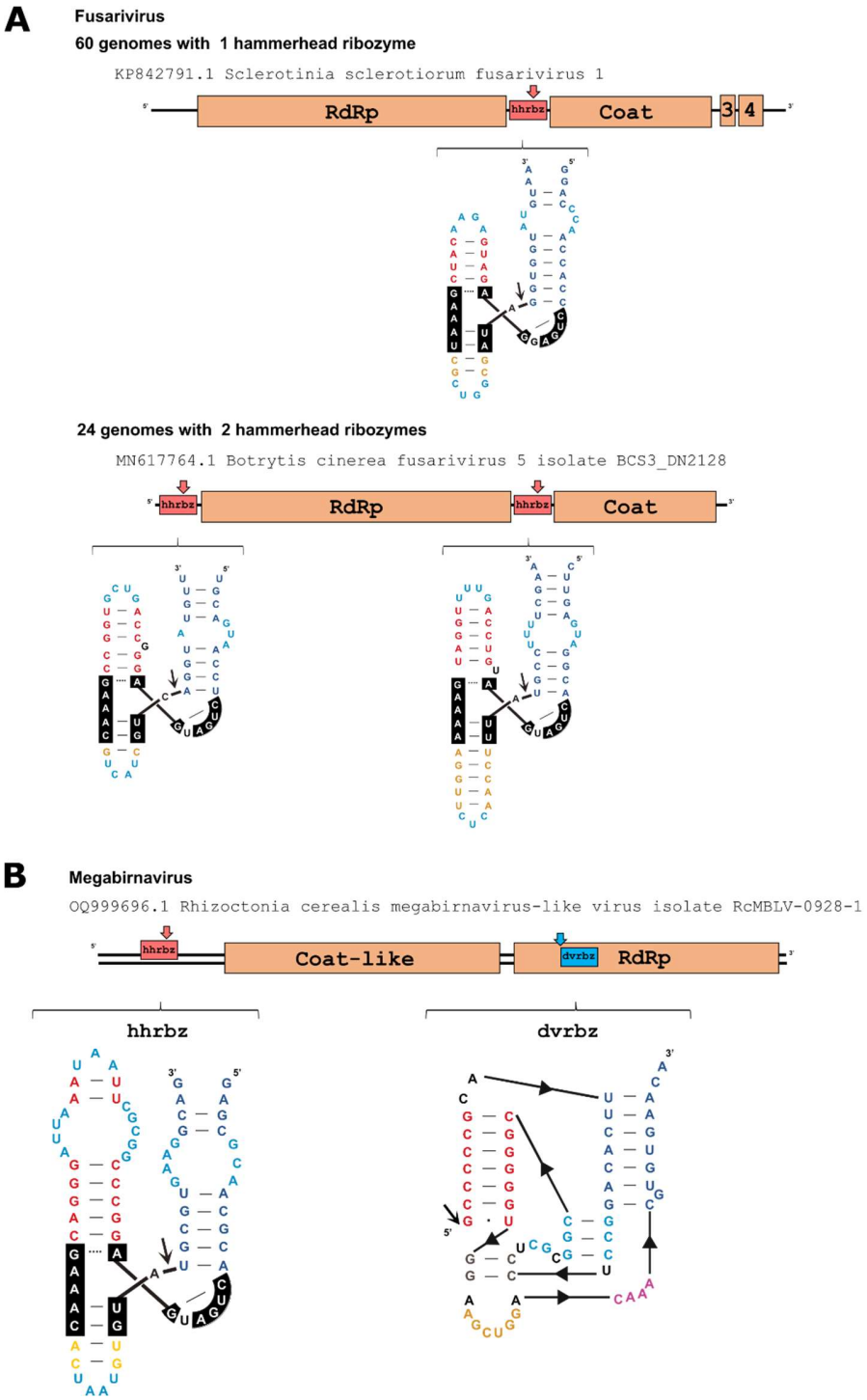


Supplementary Fig. 2. Representative assays of self-cleavage kinetics for three hammerhead ribozymes encoded in fungal and plant chrysoviruses. (A) Structure of the hammerhead ribozyme from Verticillium dahliae chrysovirus 1, RNA 1 (left), a 10% PAGE showing the resulting RNAs after different transcription times (standard conditions, pH 7.5, 1 mM Mg^{2+}) (center), and the kinetic analysis and quantification graph of the cleavage rates under these co-transcriptional conditions (right). (B) Structure of the hammerhead ribozyme from Penicillium chrysogenum virus, RNA 1 (left), a 10% PAGE showing a post-transcriptional assay of self-cleavage with purified RNA of this ribozyme incubated at pH 8.5 and 1 or 10 mM Mg^{2+}(Methods) (right), and their kinetic analysis and quantification graph of the observed cleavage rates (bottom). (C) Structure of the hammerhead ribozyme from Brassica campestris chrysovirus (left), two 10% PAGEs showing post-transcriptional assays of self-cleavage with purified RNA of this ribozyme incubated at pH 8.5 and 1 or 10 mM Mg^{2+}(Methods) (right), and their kinetic analysis and quantification graph of the cleavage rates (bottom).

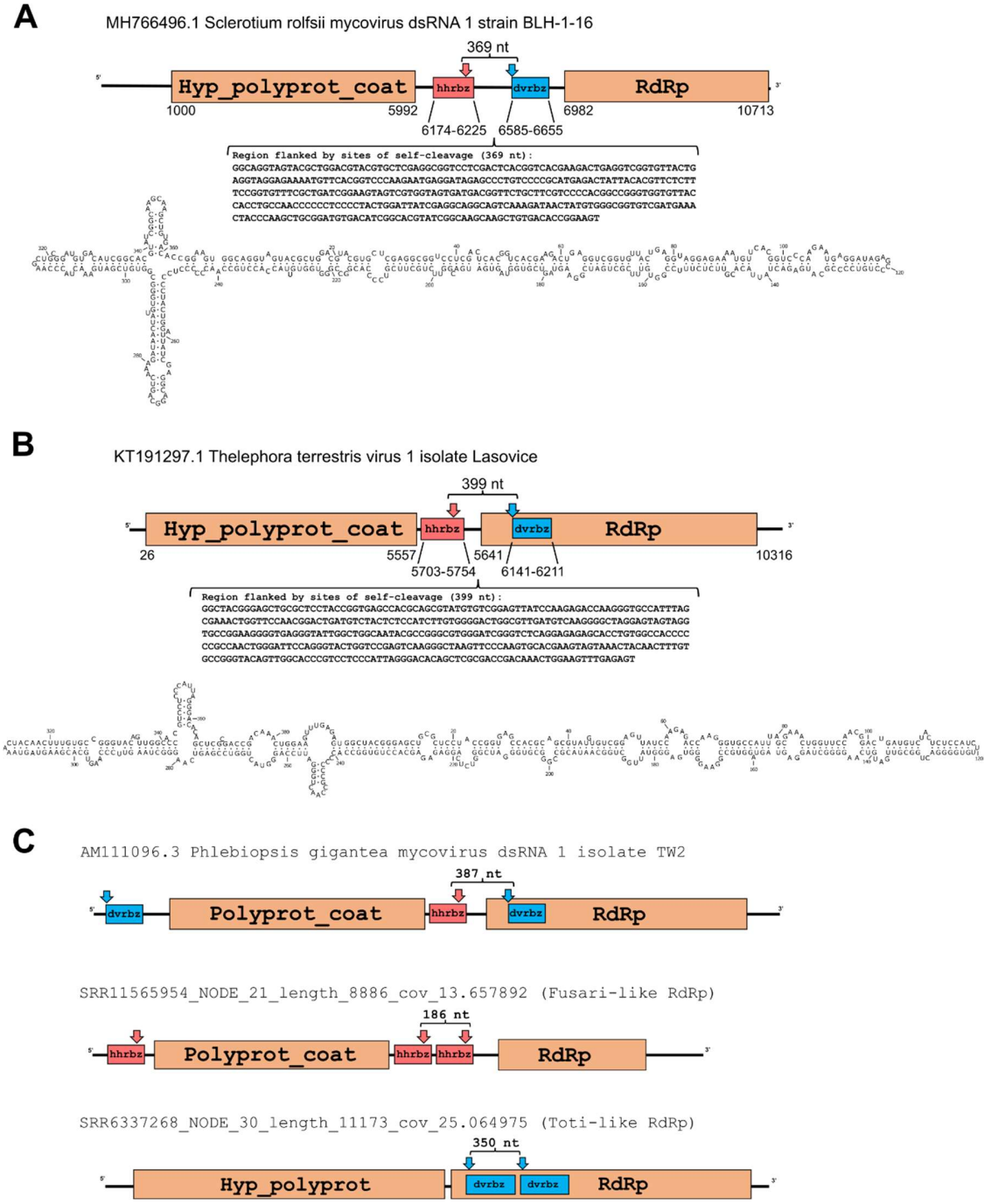

A
M
15''
30''
1'
2'
10'
100 nt
Uncleaved RNA
5' Fragment
F∞= 0,85
Kobs ~ 2,61 min-1
B
1 mM Mg2+, pH 8.5
10 mM Mg2+, pH 8.5
0'
30'
60'
120'
240'
Uncleaved RNA
5' Fragment
3' Fragment
F∞= 0,40
Kobs ~ 0,012 min-1
F∞= 0,65
Kobs ~ 0,011 min-1
C
1 mM Mg2+, pH 8.5
10 mM Mg2+, pH 8.5
M100
0h
1h
5h
20h
0,1h
0,5h
3h
4h
100 nt
Uncleaved RNA
5' Fragment
F∞= 0,25
Kobs ~0,002 min-1
F∞= 0,25
Kobs ~0,008 min-1

512

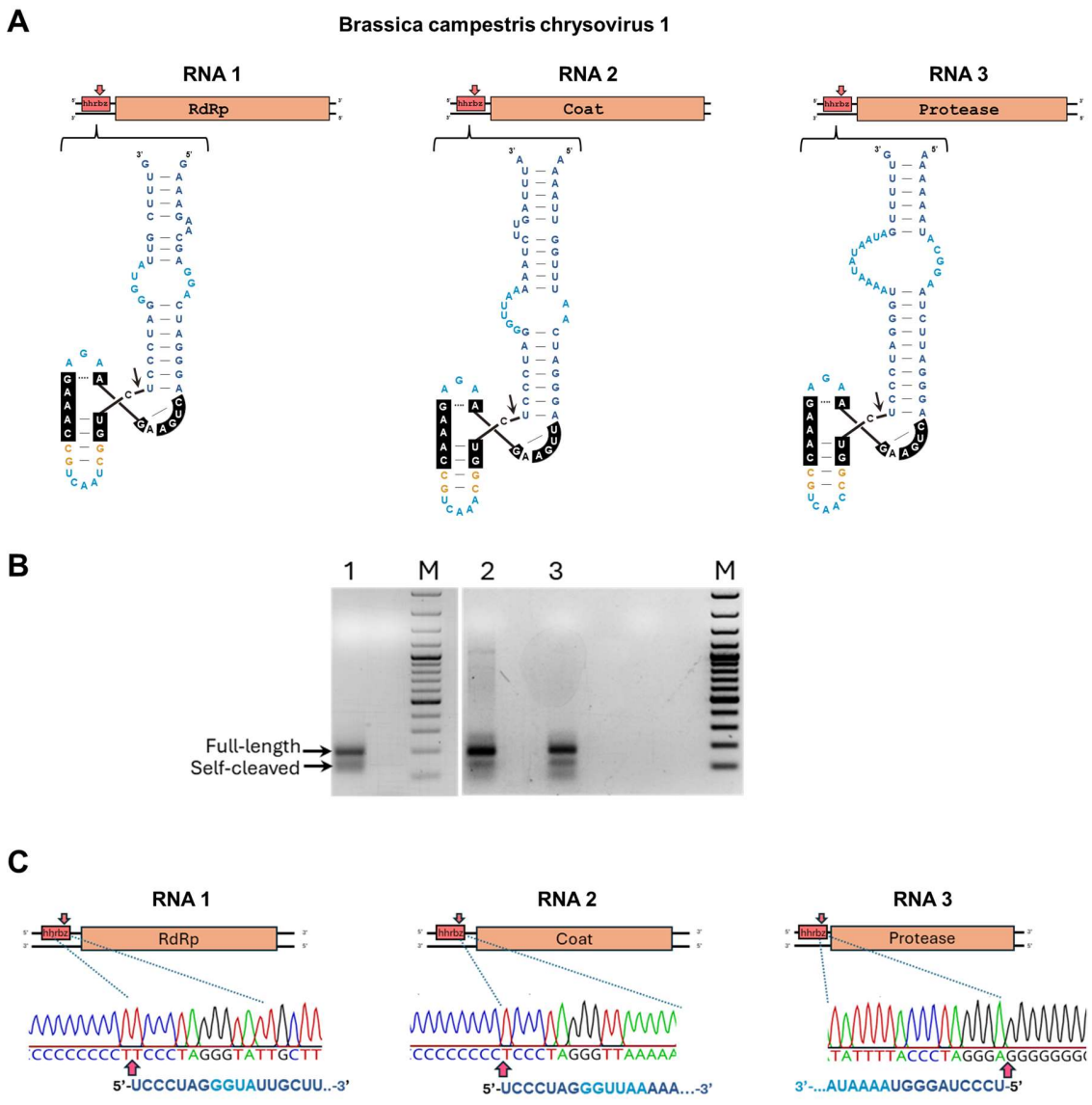
Supplementary Fig. 3. (A) Representative examples of hammerhead ribozymes (hhrbzs) detected in the UTRs of two ssRNA fusariviruses. (B) An example of a dsRNA megabirnavirus-like genome showing the presence of a hhrbz in its 5'-UTR and a deltavirus ribozyme encoded within the predicted RNA dependent RNA polymerase (RdRp) ORF. Sites of ribozyme self-cleavage are indicated by arrows.



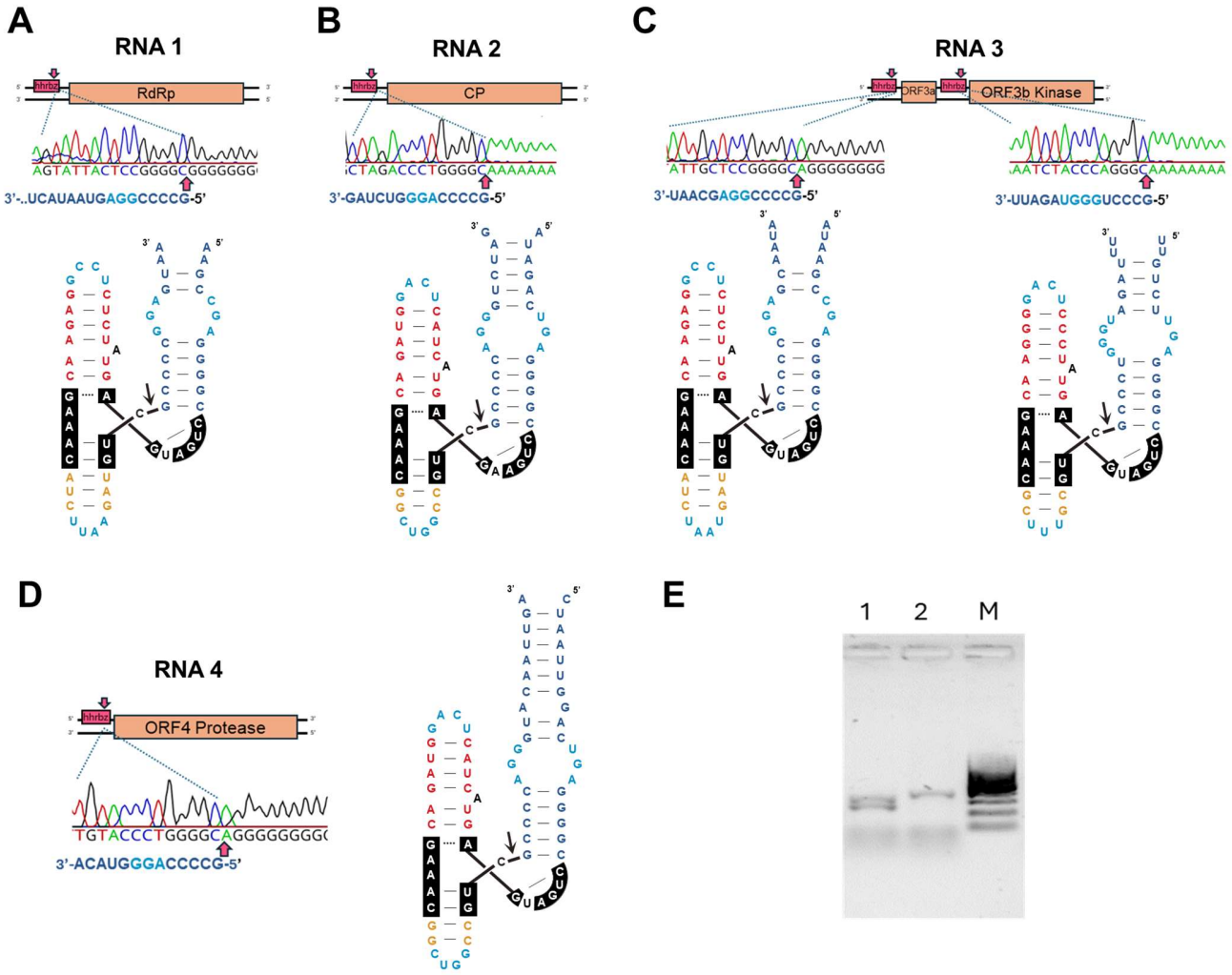
Supplementary Fig. 4. Subgenomic circular RNAs (circRNAs) produced by closely located tandem ribozyme pairs are likely encoded in some totivirids and similar RNA virus genomes. The small RNA fragments resulting after ribozyme self-cleavage usually show a nucleotide size multiple of 3, potentially encode an endless ORF (absence of stop codons) in the viral polarity, and are predicted to be highly base-paired circRNAs similar to viroid-like agents, zeta-viruses[30] or those encoded by eukaryotic retrozymes[24,25]. Hyp; hypothetical. polyprot; polyprotein. hhrbz; hammerhead ribozyme. dvrbz; deltavirus ribozyme. RdRp; RNA dependent RNA polymerase.



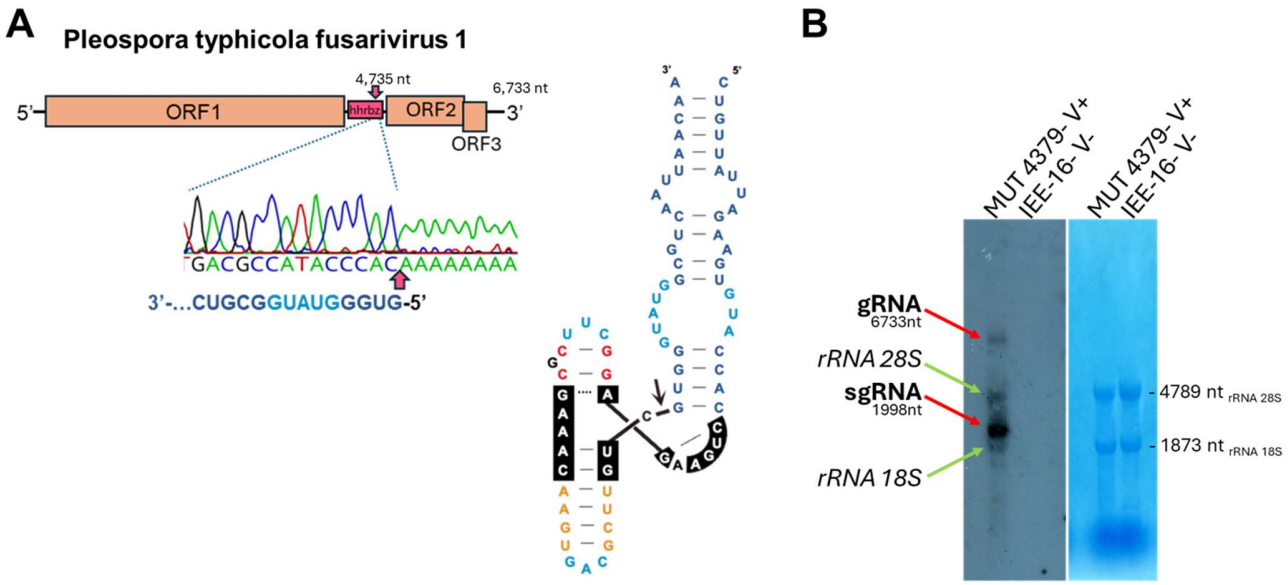
Supplementary Fig. 5. 5' RACE of Brassica campestris chrysovirus 1 (BcCV-1) from infected cabbage Torzella plants. (A) Schematic representation of the BcCV-1 hhrbzs in the 5'-UTRs of each of the three genomic RNA segments. (B) Agarose gel analysis of 5' RACE PCR of BcCV-1 RNA1, RNA2 and RNA3 (lanes 1, 2 and 3, respectively) shows two bands corresponding to the amplification of the full-length (upper and more abundant band) and the self-cleaved (lower band) RNAs. For the RNA1 (lane 1), a nested PCR, using a reverse BcCV-1 RNA1 (BcCV1-RNA1-88NRev, Supplementary Table S2), nested to the primer used in the first RACE amplification (BcCV1-RNA1-87Rev, Supplementary Table S2), is shown. M, DNA molecular weight marker (100 bp DNA ladder, Thermo Fisher). (C) Sequencing electropherograms of the cloned RACE product of smaller size for each RNA genomic segment. The 5' end coincides with that predicted to be generated by the hhrbz self-cleavage (indicated by arrows).



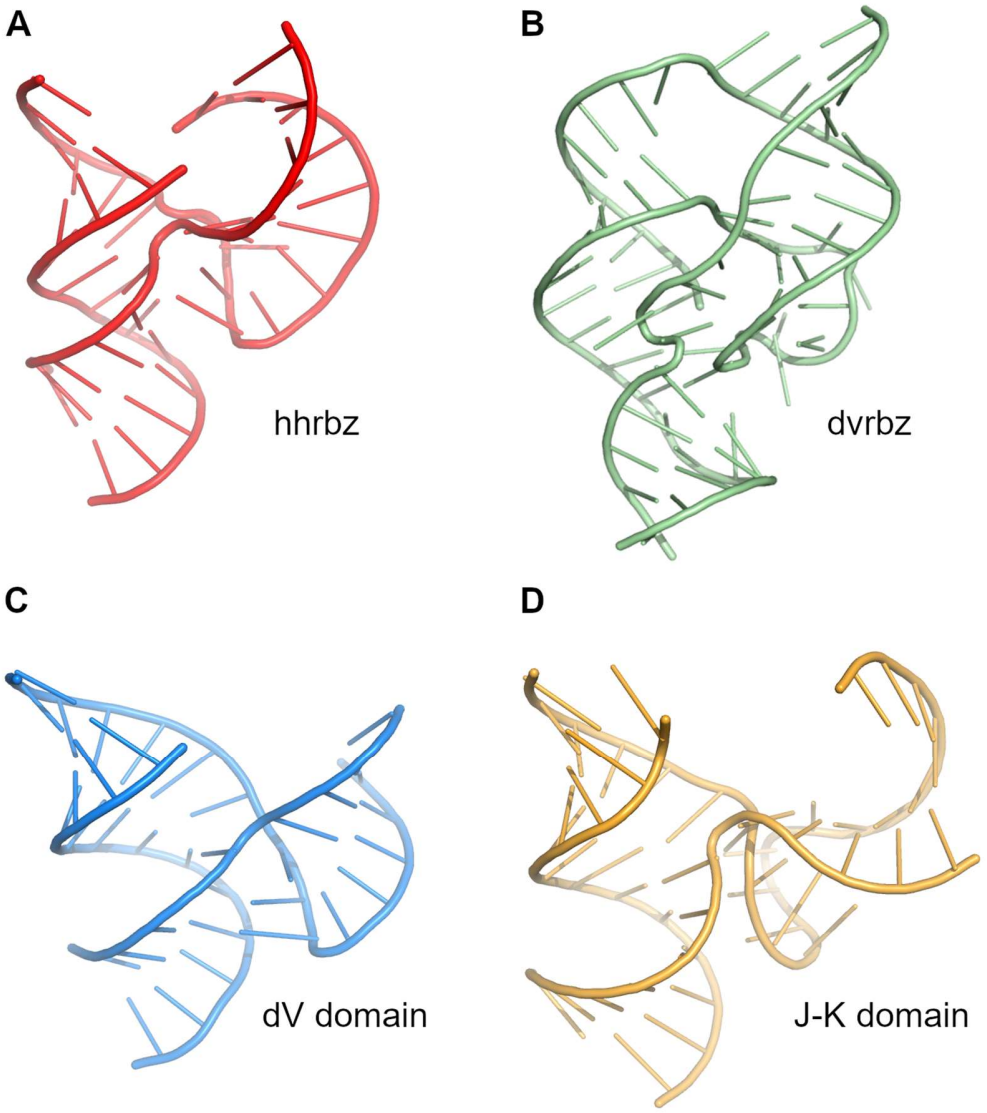
Supplementary Fig. 6. 5' RACE of Gnomoniopsis castaneae chrysovirus 1 (GcCV-1) from infected *G. castaneae* isolate 1[41]. Schematic representation of GcCV-1 RNAs, hammerhead ribozymes (hhrbz) secondary structure and sequencing electropherograms of the cloned RACE product of smaller size for each (A) RNA 1, (B) RNA 2 (C) RNA 3 and (D) RNA 4 genomic segments. Hhrbzs are present in all the 5'-UTR of RNAs, whereas RNA3 shows a second intergenic hhrbz motif between the two conserved ORFs (3a and 3b). The 5' ends of the sequencing electropherograms coincide with those predicted to be generated by the hhrbz self-cleavage (indicated by arrows). (E) Agarose gel analysis of 5' RACE PCR (lane 1) and 3'RACE PCR (lane 2) of GcCV-1 RNA2. Lane 1 shows two bands corresponding to the amplification of the full-length (upper band) and the self-cleaved (lower band) RNAs. M; DNA molecular weight marker of 100 bp DNA (Thermo Fisher). RdRp; RNA dependent RNA polymerase. CP; coat protein. ORF; open reading frame.



Supplementary Fig. 7. 5' RACE of Pleospora typhicola fusarivirus 1 (PtFV1) from infected MUT 4379 [42]. (A) Genome organization, with the position of the hammerhead ribozyme (hhrbz) placed in the intergenic region. The secondary structure of the ribozyme is also displayed (right), as well as the sequence electropherogram showing the 5' end of the cleavage site (red arrow). (B) Northern blot analysis of RNA extracts from virus-infected isolate MUT4379 and virus-free *Rhizoctonia solani* isolate IEE-16[72]. On the left panel, red arrows indicate the PtFV1 genomic RNA (gRNA) and PtFV1 sgRNA positions, while green arrows indicate some possible cross-hybridization with the two fungal ribosomal RNAs. The right panel shows a methylene blue stain of the membrane displaying the two ribosomal RNA bands.



Supplementary Fig. 8. Three dimensional models for (A) the hammerhead ribozyme from *Schistosoma mansoni* retrozymes (pdb 3ZD5, only central core and stems, loops 1 and 2 not shown), (B) the human Hepatitis Delta ribozyme (pdb 1DRZ), (C) the Hepatitis A virus IRES dV domain (pdb 6mwn) and (D) the encephalomyocarditis virus IRES J-K domain (pdb 2nbx).



Acknowledgments

This work was funded by: University of Valencia Margarita Salas Fellowship MS21-067 (MJLG); Generalitat Valenciana Grant PROMETEO CIPROM/2022/21 (MdlP); Ministry of Economics and Competitivity of Spain-FEDER grant PID2020-116008GB-I00 (MdlP). We are grateful to Daniela Alio and Maria Minutolo for providing plants infected by Brassica campestris chrysovirus 1.

References

1. Woese, C. R. The fundamental nature of the genetic code: prebiotic interactions between polynucleotides and polyamino acids or their derivatives. *Proc Natl Acad Sci U S A* **59**, 110–117 (1968).

2. Crick, F. H. The origin of the genetic code. *J Mol Biol* **38**, 367–379 (1968).

3. Gilbert, W. The RNA World. *Nature* **319**, 618 (1986).

4. Flores, R., Gago-Zachert, S., Serra, P., Sanjuan, R. & Elena, S. F. Viroids: survivors from the RNA world? *Annu Rev Microbiol* **68**, 395–414 (2014).

5. Forterre, P. The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res* **117**, 5–16 (2006).

6. Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R. & Bruening, G. Autolytic processing of dimeric plant virus satellite RNA. *Science (1979)* **231**, 1577–1580 (1986).

7. Hutchins, C. J., Rathjen, P. D., Forster, A. C. & Symons, R. H. Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. *Nucleic Acids Res* **14**, 3627–3640 (1986).

8. Buzayan, J. M., Gerlach, W. L. & Bruening, G. Satellite tobacco ringspot virus RNA: A subset of the RNA sequence is sufficient for autolytic processing. *Proc Natl Acad Sci U S A* **83**, 8859–8862 (1986).

9. Kuo, M. Y., Sharmeen, L., Dinter-Gottlieb, G. & Taylor, J. Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis delta virus. *J Virol* **62**, 4439–4444 (1988).

10. Saville, B. J. & Collins, R. A. A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria. *Cell* **61**, 685–696 (1990).

607 11. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. & Breaker, R. R. Control of gene
608 expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281–286
609 (2004).

610 12. Roth, A. *et al.* A widespread self-cleaving ribozyme class is revealed by
611 bioinformatics. *Nat Chem Biol* **10**, 56–60 (2014).

612 13. Weinberg, Z. *et al.* New classes of self-cleaving ribozymes revealed by
613 comparative genomics analysis. *Nat Chem Biol* **11**, 606–610 (2015).

614 14. Ferbeyre, G., Smith, J. M. & Cedergren, R. Schistosome satellite DNA encodes
615 active hammerhead ribozymes. *Mol Cell Biol* **18**, 3880–3888 (1998).

616 15. Epstein, L. M. & Gall, J. G. Self-cleaving transcripts of satellite DNA from the
617 newt. *Cell* **48**, 535–543 (1987).

618 16. Hammann, C., Luptak, A., Perreault, J. & de la Peña, M. The ubiquitous
619 hammerhead ribozyme. *RNA* **18**, 871–885 (2012).

620 17. De la Peña, M. & García-Robles, I. Ubiquitous presence of the hammerhead
621 ribozyme motif along the tree of life. *RNA* **16**, 1943–1950 (2010).

622 18. Webb, C. H. T., Riccitelli, N. J., Ruminski, D. J. & Lupták, A. Widespread
623 occurrence of self-cleaving ribozymes. *Science (1979)* **326**, 953 (2009).

624 19. De la Peña, M. & Garcia-Robles, I. Intronic hammerhead ribozymes are
625 ultraconserved in the human genome. *EMBO Rep* **11**, 711–716 (2010).

626 20. Salehi-Ashtiani, K., Luptak, A., Litovchick, A. & Szostak, J. W. A genomewide
627 search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene.
628 *Science (1979)* **313**, 1788–1792 (2006).

629 21. Ruminski, D. J., Webb, C. H. T., Riccitelli, N. J., Lupta, A. & Luptak, A. Processing
630 and translation initiation of non-long terminal repeat retrotransposons by
631 hepatitis delta virus (HDV)-like self-cleaving ribozymes. *J Biol Chem* **286**,
632 41286–41295 (2011).

633 22. Cervera, A. & de la Peña, M. Eukaryotic penelope-like retroelements encode
634 hammerhead ribozyme motifs. *Mol Biol Evol* **31**, 2941–2947 (2014).

635 23. Eickbush, D. G. & Eickbush, T. H. R2 retrotransposons encode a self-cleaving
636 ribozyme for processing from an rRNA cotranscript. *Mol Cell Biol* **30**, 3142–
637 3150 (2010).

638 24. Cervera, A., Urbina, D. & de la Peña, M. Retrozymes are a unique family of non-
639 autonomous retrotransposons with hammerhead ribozymes that propagate in
640 plants through circular RNAs. *Genome Biol* **17**, 135 (2016).

641 25. Cervera, A. & de la Peña, M. Small circRNAs with self-cleaving ribozymes are
642 highly expressed in diverse metazoan transcriptomes. *Nucleic Acids Res* **48**,
643 5054–5064 (2020).

644 26. Liu, G. *et al.* The function of twister ribozyme variants in non-LTR
645 retrotransposition in Schistosoma mansoni. *Nucleic Acids Res* **49**, 10573–10588
646 (2021).

647 27. Salehi-Ashtiani, K. & Szostak, J. W. In vitro evolution suggests multiple origins
648 for the hammerhead ribozyme. *Nature* **414**, 82–84 (2001).

649 28. De la Peña, M., Gago, S. & Flores, R. Peripheral regions of natural hammerhead
650 ribozymes greatly increase their self-cleavage activity. *EMBO J* **22**, 5561–5570
651 (2003).

652 29. Martick, M. & Scott, W. G. Tertiary contacts distant from the active site prime a
653 ribozyme for catalysis. *Cell* **126**, 309–320 (2006).

654 30. Edgar, R. C. *et al.* Petabase-scale sequence alignment catalyses viral discovery.
655 *Nature 2022* **602**, 1–6 (2022).

656 31. Weinberg, C. E., Olzog, V. J., Eckert, I. & Weinberg, Z. Identification of over 200-
657 fold more hairpin ribozymes than previously known in diverse circular RNAs.
658 *Nucleic Acids Res* **49**, 6375–6388 (2021).

659 32. Lee, B. D. *et al.* Mining metatranscriptomes reveals a vast world of viroid-like
660 circular RNAs. *Cell* **186**, 646-661.e4 (2023).

661 33. Forgia, M. *et al.* Hybrids of RNA viruses and viroid-like elements replicate in
662 fungi. *Nature Communications 2023 14:1* **14**, 1–11 (2023).

663 34. Arkhipova, I. R., Yushenova, I. A. & Rodriguez, F. Giant Reverse Transcriptase-
664 Encoding Transposable Elements at Telomeres. *Mol Biol Evol* **34**, 2245–2257
665 (2017).

666 35. Epstein, L. M. & Pabon-Pena, L. M. Alternative modes of self-cleavage by newt
667 satellite 2 transcripts. *Nucleic Acids Res* **19**, 1699–1705 (1991).

668 36. Cao, Y. fen *et al.* Genomic characterization of a novel dsRNA virus detected in
669 the phytopathogenic fungus Verticillium dahliae Kleb. *Virus Res* **159**, 73–78
670 (2011).

671 37. Zhang, R. *et al.* A novel single-stranded RNA virus isolated from a
672 phytopathogenic filamentous fungus, Rosellinia necatrix, with similarity to
673 hypo-like viruses. *Front Microbiol* **5**, 92402 (2014).

674 38. Chiba, S. *et al.* ICTV Virus Taxonomy Profile: Fusariviridae 2024: This article is
675 part of the ICTV Virus Taxonomy Profiles collection. *Journal of General Virology*
676 **105**, 001973 (2024).

677 39. Chiba, S. *et al.* A novel bipartite double-stranded RNA mycovirus from the white
678 root rot fungus Rosellinia necatrix: Molecular and biological characterization,
679 taxonomic considerations, and potential for biological control. *J Virol* **83**,
680 12801–12812 (2009).

681 40. Zhang, J. *et al.* The genome sequence of Brassica campestris chrysovirus 1, a
682 novel putative plant-infecting tripartite chrysovirus. *Arch Virol* **162**, 1107–
683 1111 (2017).

684 41. Ferilli, F., Lione, G., Gonthier, P., Turina, M. & Forgia, M. First detection of
685 mycoviruses in Gnomoniopsis castaneae suggests a putative horizontal gene
686 transfer event between negative-sense and double-strand RNA viruses.
687 *Virology* 110057 (2024) doi:10.1016/J.VIROL.2024.110057.

688 42. Nerva, L. *et al.* Multiple approaches for the detection and characterization of
689 viral and plasmid symbionts from a collection of marine fungi. *Virus Res* **219**,
690 22–38 (2016).

691 43. Chiba, S., Jamal, A. & Suzuki, N. First evidence for internal ribosomal entry sites
692 in diverse fungal virus genomes. *mBio* **9**, (2018).

693 44. Doudna, J. A. Hammerhead ribozyme structure: U-turn for RNA structural
694 biology. *Structure* **3**, 747–750 (1995).

695 45. Kristoffersen, E. L., Burman, M., Noy, A. & Holliger, P. Rolling Circle RNA
696 Synthesis Catalysed by RNA. *Elife* **11**, (2022).

697 46. Jimenez, R. M., Polanco, J. A. & Luptak, A. Chemistry and Biology of Self-Cleaving
698 Ribozymes. *Trends Biochem Sci* **40**, 648–661 (2015).

699 47. De la Peña, M., Garcia-Robles, I. & Cervera, A. The Hammerhead Ribozyme: A
700 Long History for a Short RNA. *Molecules* **22**, 1–11 (2017).

701 48. Zheludev, I. N. *et al.* Viroid-like colonists of human microbiomes. *bioRxiv*
702 2024.01.20.576352 (2024) doi:10.1101/2024.01.20.576352.

703 49. De La Peña, M., Ceprian, R., Casey, J. L. & Cervera, A. Hepatitis delta virus-like
704 circular RNAs from diverse metazoans encode conserved hammerhead
705 ribozymes. *Virus Evol* **7**, (2021).

706 50. Perreault, J. *et al.* Identification of hammerhead ribozymes in all domains of life
707 reveals novel structural variations. *PLoS Comput Biol* **7**, e1002031 (2011).

708 51. Khvorova, A., Lescoute, A., Westhof, E. & Jayasena, S. D. Sequence elements
709 outside the hammerhead ribozyme catalytic core enable intracellular activity.
710 *Nat Struct Biol* **10**, 708–712 (2003).

711 52. De la Peña, M., Dufour, D. & Gallego, J. Three-way RNA junctions with remote
712 tertiary contacts: a recurrent and highly versatile fold. *RNA* **15**, 1949–1964
713 (2009).

714 53. Schultz, E. P., Vasquez, E. E. & Scott, W. G. Structural and catalytic effects of an
715 invariant purine substitution in the hammerhead ribozyme: Implications for
716 the mechanism of acid-base catalysis. *Acta Crystallogr D Biol Crystallogr* **70**,
717 2256–2263 (2014).

718 54. Martínez-Salas, E., Francisco-Velilla, R., Fernandez-Chamorro, J., Lozano, G. &
719 Diaz-Toledano, R. Picornavirus IRES elements: RNA structure and host protein
720 interactions. *Virus Res* **206**, 62–73 (2015).

721 55. Kolupaeva, V. G., Pestova, T. V., Hellen, C. U. T. & Shatsky, I. N. Translation
722 Eukaryotic Initiation Factor 4G Recognizes a Specific Structural Element within
723 the Internal Ribosome Entry Site of Encephalomyocarditis Virus RNA. *Journal*
724 *of Biological Chemistry* **273**, 18599–18604 (1998).

725 56. Imai, S., Kumar, P., Hellen, C. U. T., D'Souza, V. M. & Wagner, G. An accurately
726 preorganized IRES RNA structure enables eIF4G capture for initiation of viral
727 translation. *Nature Structural & Molecular Biology 2016 23:9* **23**, 859–864
728 (2016).

729 57. Brown, E. A., Day, S. P., Jansen, R. W. & Lemon, S. M. The 5' nontranslated region
730 of hepatitis A virus RNA: secondary structure and elements required for
731 translation in vitro. *J Virol* **65**, 5828–5838 (1991).

732 58. Koirala, D. *et al.* A conserved RNA structural motif for organizing topology
733 within picornaviral internal ribosome entry sites. *Nature Communications 2019*
734 *10:1* **10**, 1–13 (2019).

735 59. Avanzino, B. C., Fuchs, G. & Fraser, C. S. Cellular cap-binding protein, eIF4E,
736 promotes picornavirus genome restructuring and translation. *Proc Natl Acad*
737 *Sci U S A* **114**, 9611–9616 (2017).

738 60. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions
739 with AlphaFold 3. *Nature 2024* 1–3 (2024) doi:10.1038/s41586-024-07487-w.

740 61. Strassert, J. F. H., Irisarri, I., Williams, T. A. & Burki, F. A molecular timescale for
741 eukaryote evolution with implications for the origin of red algal-derived
742 plastids. *Nature Communications 2021 12:1* **12**, 1–13 (2021).

743 62. Fernández, I. S., Bai, X. C., Murshudov, G., Scheres, S. H. W. & Ramakrishnan, V.
744 Initiation of Translation by Cricket Paralysis Virus IRES Requires Its
745 Translocation in the Ribosome. *Cell* **157**, 823–831 (2014).

746 63. Pfingsten, J. S., Costantino, D. A. & Kieft, J. S. Structural basic for ribosome
747 recruitment and manipulation by a viral IRES RNA. *Science (1979)* **314**, 1450–
748 1454 (2006).

749 64. Nawrocki, E. P. & Eddy, S. R. Infernal 1.1: 100-fold faster RNA homology
750 searches. *Bioinformatics* **29**, 2933–2935 (2013).

751 65. Kalvari, I. *et al.* Rfam 14: expanded coverage of metagenomic, viral and
752 microRNA families. *Nucleic Acids Res* **49**, D192–D200 (2021).

753 66. Weinberg, Z. & Breaker, R. R. R2R - software to speed the depiction of aesthetic
754 consensus RNA secondary structures. *BMC Bioinformatics* **12**, 1–9 (2011).

755 67. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high
756 throughput. *Nucleic Acids Res* **32**, 1792–1797 (2004).

757 68. Long, D. M. & Uhlenbeck, O. C. Kinetic characterization of intramolecular and
758 intermolecular hammerhead RNAs with stem II deletions. *Proc Natl Acad Sci U*
759 *S A* **91**, 6977–6981 (1994).

760 69. Minutolo, M. *et al.* A polyvalent tool for detecting coguviruses in multiple hosts
761 allowed the identification of a novel seed-transmitted coguvirus infecting
762 Brassicaceae. *Phytopathology* (2023) doi:10.1094/PHYTO-10-23-0362-
763 R/ASSET/IMAGES/LARGE/PHYTO-10-23-0362-RT2.JPEG.

764 70. Rastgou, M. *et al.* Molecular characterization of the plant virus genus
765 Ourmiavirus and evidence of inter-kingdom reassortment of viral genome
766 segments as its possible route of origin. *J Gen Virol* **90**, 2525 (2009).

767 71. Sutela, S. *et al.* The virome from a collection of endomycorrhizal fungi reveals
768 new viral taxa with unprecedented genome organization. *Virus Evol* **6**, (2020).

769 72. Picarelli, M. A. S. C. *et al.* Extreme diversity of mycoviruses present in isolates of
770 rhizoctonia solani AG2-2 LP from Zoysia japonica from Brazil. *Front Cell Infect*
771 *Microbiol* **9**, 457597 (2019).

772