



HAL
open science

Alternative mutational pathways, outside the VPg, of Rice yellow mottle virus to overcome eIF(iso)4G-mediated rice resistance under strong genetic constraints

Nils Poulicard, Agnès Pinel-Galzi, Denis Fargette, Eugénie Hébrard

► To cite this version:

Nils Poulicard, Agnès Pinel-Galzi, Denis Fargette, Eugénie Hébrard. Alternative mutational pathways, outside the VPg, of Rice yellow mottle virus to overcome eIF(iso)4G-mediated rice resistance under strong genetic constraints. *Journal of General Virology*, 2014, 95 (1), pp.219-224. ird-01521875

HAL Id: ird-01521875

<https://ird.hal.science/ird-01521875v1>

Submitted on 12 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2 **Title: Alternative mutational pathways, outside the VPg, of *Rice yellow mottle virus* to**
3 **overcome eIF(iso)4G-mediated rice resistance under strong genetic constraints**

4
5 Running title: Alternative mutational pathways outside the RYMV VPg

6
7 Authors: Nils Poulicard*, Agnès Pinel-Galzi, Denis Fargette and Eugénie Hébrard
8 Institut de Recherche pour le Développement (IRD), UMR RPB, Montpellier, France
9 * present address: Centre for Plant Biotechnology and Genomics U.P.M. – I.N.I.A. Parque
10 Científico y Tecnológico de la U.P.M. Campus de Montegancedo, Madrid, Spain

11
12 Corresponding author: eugenie.hebrard@ird.fr, +33(0)4 67 416 289

13 Word count: summary 134, main text 2496

14
15 **Summary**

16 The adaptation of *Rice yellow mottle virus* (RYMV) to *RYMV1*-mediated resistance
17 has been reported to involve mutations in the viral protein genome-linked (VPg). In this
18 study, we analysed several cases of *rymv1-2* resistance breakdown by an isolate with low
19 adaptability. Surprisingly, in these rarely occurring resistance-breaking (RB) genotypes
20 mutations were detected outside the VPg, in the ORF2a/ORF2b overlapping region. The
21 causal role of three mutations associated with *rymv1-2* resistance breakdown was validated
22 via directed mutagenesis of an infectious clone. In resistant plants, these mutations increased
23 viral accumulation as efficiently as suboptimal RB mutations in the VPg. Interestingly, these
24 mutations are located in a highly conserved, but unfolded domain. Altogether, our results
25 indicate that under strong genetic constraints, *a priori* unfit genotypes can follow alternative
26 mutational pathways, i.e., outside the VPg, to overcome *rymv1-2* resistance.

27
28 **Main text**

29 High mutation rates, high levels of recombination or reassortment, short replication
30 cycles and high accumulation rates should allow plant viruses to adapt rapidly to new host
31 species and resistant hosts. However, adaptation dynamics depend on the number and nature
32 of mutations (genetic barrier) and on their fitness cost (phenotypic barrier) (Domingo et al.,
33 2012, Harrison, 2002). In particular, structural constraints and antagonistic epistasis
34 dramatically reduce the emergence of adaptive mutations (Camps et al., 2007, Sanjuan &
35 Nebot, 2008, Weinreich et al., 2005). This is exemplified by *Rice yellow mottle virus*
36 (RYMV), belonging to the genus *Sobemovirus*. RYMV shows a high virus content in plants
37 (Poulicard et al., 2010), evolves rapidly (Fargette et al., 2008) and is able to adapt to highly
38 resistant rice cultivars (Pinel-Galzi et al., 2007, Traoré et al., 2010). However, strong
39 demographic constraints (bottlenecks and random genetic drift), genetic constraints (codon
40 usage and mutational bias) and phenotypic constraints (epistasis antagonism and fitness costs)
41 have been identified, they modulate the ability to overcome the high resistance mediated by
42 the *RYMV1* gene which encodes the translation initiation factor eIF(iso)4G1 (Poulicard et al.,
43 2010, Traoré et al., 2010).

44
45 The genome of RYMV consists of a single-stranded, positive-sense RNA molecule with
46 a viral protein genome-linked (VPg) that is covalently linked to its 5' end. The VPg encoded
47 by the central domain of ORF2a interacts with rice eIF(iso)4G1 (Hébrard et al., 2010). A
48 single amino acid substitution in the middle domain of eIF(iso)4G1 results in the *rymv1-2*
49 allele found in the highly resistant *Oryza sativa indica* cultivars Gigante and Bekarosaka
50 (Albar et al., 2006, Rakotomalala et al., 2008). The phenotype of this recessive allele is

51 characterised by an absence of symptom expression and a lack of viral detection by ELISA.
52 However, adaptation to the *rymv1-2* allele has been reported, and the genetic determinism of
53 this adaptation has been elucidated (Pinel-Galzi et al., 2007). The *rymv1-2* resistance-breaking
54 (RB) phenotype is caused by punctual mutations in the VPg, most often located at codon 48,
55 but sometimes at codon 52. The associated major and minor mutational pathways have been
56 described previously.

57
58 The *rymv1-2* RB ability is related to an E/T polymorphism at the adjacent codon 49 of
59 the VPg (Poulicard et al., 2012). Threonine at codon 49 (T49) confers a strong selective
60 advantage over viral populations harbouring E49 in susceptible and resistant *O. glaberrima*
61 cultivars, whereas T49 is a major constraint to overcome the *rymv1-2* allele found in *O. sativa*
62 *indica* cultivars (cv.) Gigante and Bekarosaka. Antagonistic epistasis between T49 and RB
63 mutations was established through mutagenesis of the infectious clone CIa. Phenotypic and
64 genetic barriers prevented the major and minor mutational pathways from being followed.
65 The direct influence of the E/T polymorphism at codon 49 on the *rymv1-2* RB ability of the
66 wild-type genotype CIa (with T49) and the mutated genotype CIa49E was validated
67 experimentally, and the T49E substitution was found to increase the ability to overcome
68 *rymv1-2* resistance from 5% to 40% (Poulicard et al., 2012). The RB pattern of the mutant
69 CIa49E was therefore similar to that of other wild-type viral populations containing E49
70 (Pinel-Galzi et al., 2007) and mostly involved fixation of the R48G RB mutation (i.e., the first
71 step in the major mutational pathway).

72
73 Although the adaptability of the CIa genotype has been assessed previously (Pinel-Galzi
74 et al., 2007), *rymv1-2* resistance breakdown has never been observed. In the present study,
75 three of 53 plants (i.e., 5%) inoculated with the CIa genotype showed characteristic RYMV
76 symptoms (Poulicard et al., 2012). The objective of this study was to identify and characterise
77 the mutational pathways involved in *rymv1-2* resistance breakdown in the unfit CIa genotype.
78 The VPg of the RB populations of each plant was amplified and directly sequenced according
79 to a method described previously (Fargette et al., 2004). Surprisingly, these RB populations
80 did not show mutations in the VPg. To identify candidate mutations involved in the RB
81 phenotype, the full-length viral genomes of these populations were amplified via RT-PCR and
82 sequenced. Two RB genotypes were characterised by a single mutation (A2229G or
83 G2278A), while no mutation was detected in the third RB genotype. This is the first report of
84 an RB-associated mutation outside the VPg. To further investigate the frequency of these
85 mutations, twenty infected plants of fifty plants inoculated with CIa49E were analysed
86 following the same procedure. Three of the twenty plants infected with the CIa49E genotype
87 also displayed single mutations outside the VPg (A2199G, G2275A and A2301G). These five
88 mutations occurred within a 102-nucleotide-long stretch in the ORF2a/ORF2b overlapping
89 region, which was located 376 nucleotides downstream of the VPg (Figure 1a). These
90 mutations always caused non-synonymous changes in the P2a polyprotein (Figure 1b), and
91 they generally involved the substitution of a positively charged amino acid, such as lysine or
92 arginine, with the negatively charged amino acid glutamic acid. In contrast, these mutations
93 did not always change the physico-chemical properties of the residue in ORF2b. The
94 genotypes harbouring the mutations A2199G, A2229G, G2275A, G2278A and A2301G were
95 subsequently designated CIa49E*K531E, CIa*K541E, CIa49E*G556E, CIa*R557Q and
96 CIa49E*K565E, respectively, in reference to the nature and position of the mutations in the
97 P2a polyprotein.

98
99 Back-inoculations of five resistant *O.s. indica* cv. Gigante plants with each viral
100 population confirmed the *rymv1-2* RB ability of three of them, CIa49E*K531E, CIa*K541E

101 and CIa*R557Q (100% infection rate). To establish the causal role of these mutations,
102 directed mutagenesis of the infectious clone CIa was performed using the QuikChange Site-
103 Directed Mutagenesis Kit (Stratagene). Notably, the K531E mutation was introduced without
104 the T49E mutation in the VPg to assess the independence of the two mutations. Transcription
105 of the mutated clones and inoculation of the viral RNAs *in planta* were performed as
106 previously described (Poulicard et al., 2010). Each mutated clone was inoculated in five
107 susceptible and five resistant individuals of *O.s. indica* cv. IR64 and cv. Bekarosaka,
108 respectively. All mutants were infectious in the susceptible plants (100% infection rate). In all
109 of the resistant plants, characteristic symptoms, high ELISA values, successful RT-PCR
110 amplification and sequencing confirmed that the punctual mutations K531E, K541E and
111 R557Q were directly involved in the *rymv1-2* RB phenotype. Additional mutations did not
112 emerge within the P2a or VPg coding regions. Interestingly, the role of the K531E mutation in
113 resistance breakdown was validated in the absence of the T49E VPg mutation. Therefore, the
114 emergence of K531E is sufficient to overcome *rymv1-2* resistance.

115

116 To evaluate the accumulation of the three RB genotypes CIa49E*K531E, CIa*K541E
117 and CIa*R557Q, three resistant and three susceptible individuals of *O.s. indica* cv.
118 Bekarosaka and cv. IR64, respectively, were back-inoculated and analysed via qRT-PCR
119 following a protocol previously described (Poulicard et al., 2010). The accumulation of each
120 mutant genotype was compared to that of the wild-type CIa genotype and two *rymv1-2* RB
121 genotypes with mutations in the VPg, i.e., CIa*R48E and CIa*H52Y (Figure 2a). In this
122 experiment, $c.10^{12}$ RNA copies of each genotype were inoculated per plant. The flanking
123 region of the VPg and the C-terminal region of P2a (nucleotides 1,480-2,900) were sequenced
124 from the total RNA extracts used for qRT-PCR, no new mutations were detected. At 28 dpi,
125 $c.10^{10}$ copies of the CIa*K531E, CIa*K541E and CIa*R557Q genotypes had accumulated per
126 milligram of leaf tissue in the resistant plants; i.e., their levels were 10^5 times higher than
127 those of the wild-type CIa genotype (Figure 2a). No significant differences in RNA
128 accumulation were detected between the CIa*K531E, CIa*K541E and CIa*R557Q genotypes
129 (ANOVA, $P>0.05$). Interestingly, the accumulation of these RB genotypes was not
130 significantly different from that of the suboptimal CIa*H52Y genotype ($P>0.05$). CIa*R48E
131 showed maximal accumulation of $c.10^{12}$ RNA copies per milligram of leaf tissue in the
132 resistant cultivar at 28 dpi, as observed previously (Poulicard et al. 2010), which was
133 approximately fifty times higher than the accumulation of the other RB genotypes
134 ($P<0.0001$). In the susceptible plants, the RNA accumulation of the CIa*K531E, CIa*K541E,
135 CIa*R557Q and CIa*H52Y genotypes at 14 dpi was not significantly lower than that of the
136 wild-type genotype ($P>0.05$; Figure 2b). In contrast, the CIa*R48E genotype was strongly
137 impaired in the susceptible hosts, and its RNA accumulation was $c.10^5$ times lower than that
138 of the other genotypes. Surprisingly, sequencing of the RB viral population in the susceptible
139 plants revealed the emergence of reverse mutations 28 days after inoculation. The artificially
140 inserted substitution from arginine to glutamic acid (from AGG to GAG) at codon 48 in VPg
141 of the CIa*R48E genotype was displaced by a mutation that restored the positively charged
142 residue, i.e., lysine (AAG). In addition, reversion of the K531E and K541E mutations was
143 always observed following back-inoculation to *O.s. indica* cv. IR64. These reverse mutations
144 suggested a fitness cost in susceptible hosts of the mutations that emerged in the C-terminal
145 domain of P2a. No other mutation emerged in the P2a coding region during this fitness
146 experiment, and compensatory mutations were never observed in the susceptible plants.

147

148 The natural diversity of the positions involved in the alternative RB pathway and their
149 genetic context in the ORF2a/ORF2b overlapping region were analysed (Suppl. Table 1).
150 Sequence alignment of 33 full-length sequences of viral populations that were representative

151 of the genetic and geographic diversity of RYMV (Rakotomalala et al., 2013) revealed that
152 these positions were strictly conserved. The nucleotide diversity (π) was estimated for each
153 ORF using Dnasp software (Librado & Rozas, 2009). As previously reported (Fargette et al.
154 2004), the nucleotide variability was lower in ORF2a and ORF2b ($\pi=0.054$ and 0.057) than in
155 ORF1 and ORF4 ($\pi=0.109$ and 0.087). However, these values were still higher than that
156 found for the recently described ORFx which overlaps the 5' end of sobemovirus ORF2a
157 (Ling et al., 2013) ($\pi=0.029$). Interestingly, the RB mutations always emerged in the ORF2a
158 domain, which is characterised by low nucleotide diversity (Figure 3a). Overlapping viral
159 regions (OVRs) have been reported to be highly conserved, showing strong constraints
160 against synonymous changes (Simon-Loriere et al., 2013). Accordingly, the P2a OVR
161 exhibited a lower diversity of synonymous sites than the VPg, while the proportion of non-
162 synonymous sites was similar (Suppl. Table 2).

163
164 The organisation and the biological function of the C-terminal domain of P2a are
165 unknown. However, the processing of the P2a polyprotein of another sobemovirus, *Sesbania*
166 *mosaic virus* (SeMV), has recently been elucidated. In addition to a serine protease and VPg,
167 two new proteins, designated P10 and P8, were identified (Nair & Savithri, 2010b,
168 Satheshkumar et al., 2004). Similar to the VPg of RYMV, the VPg and P8 of SeMV were
169 demonstrated to be natively unfolded proteins (Hébrard et al., 2009, Nair & Savithri, 2010a,
170 Satheshkumar et al., 2005). Detection of the disordered arrangement of the C-terminal domain
171 of P2a of the RYMV was then performed using the software FoldIndex© (Obradovic et al.,
172 2005, Prilusky et al., 2005). Prediction analyses indicated an alternating arrangement of
173 folded and unfolded domains in RYMV P2a that was similar to the profile obtained for
174 SeMV, although the sequence identity between sobemoviruses is low (Figure 3b).
175 Interestingly, all of the RB mutations described in this study were located within the predicted
176 unfolded segment of the P2a OVR, which strongly suggested that all of these RB mutations
177 also occurred in the RYMV homologue of P8. OVRs have been reported to exhibit more
178 structural disorder than non-overlapping regions (NOVRs) (Rancurel et al., 2009). Moreover,
179 the termini of proteins are, on average, more prone to be disordered than internal regions
180 (Uversky). Because intrinsically disordered protein tails are engaged in a wide range of
181 functions, the C-terminal domain of P2a may be directly or indirectly involved in the
182 interaction between the VPg of RYMV and the eIF(iso)4G1 of rice. Thus, the RB mutations
183 described in this study may favour the capture of eIF(iso)4G1 in resistant plants, which would
184 compensate for the relatively low affinity of the wild-type VPg with the *rymv1-2* eIF(iso)4G1
185 (Hébrard et al., 2010).

186
187 Despite the crucial roles of eIFs/VPg interactions in successful infections and in plant
188 resistance breakdown, described in numerous studies (for review (Truniger & Aranda, 2009,
189 Wang & Krishnaswamy, 2012), exceptions are occasionally reported. The RB phenotypes
190 induced by *Lettuce mosaic virus* and *Clover yellow vein virus* from the genus *Potyvirus* are
191 sometimes related to the emergence of mutations in the cylindrical inclusion (CI) and P1
192 proteins, respectively (Abdul-Razzak et al., 2009, Nakahara et al., 2010). In this study, the
193 detected RB mutations emerged in the C-terminal domain of the P2a polyprotein, which is
194 downstream of the VPg, within the ORF2a/ORF2b overlapping region. This alternative RB
195 mutational pathway was revealed under strong selective constraints. The threonine at position
196 49 of the VPg was previously reported to be the major genetic constraint blocking the
197 emergence of RB mutations in the VPg. This strong constraint could lead this genotype to
198 adopt an alternative strategy to overcome the *rymv1-2* allele. However, the frequency of this
199 alternative mutational pathway in less-constrained viral populations might be underestimated,
200 as suggested by the detection of P2a C-terminal mutations in the artificial genotype CIa49E.

201 Comparison of the identified RB mutational pathways revealed similarities. Although the RB
202 mutations emerged in two different domains, they occurred in two unfolded regions of the
203 same highly conserved P2a polyprotein. The fitness of RB genotypes with mutations in the C-
204 terminal domain of P2a was suboptimal in *rymv1-2*-resistant plants and intermediate in
205 susceptible plants, as were those of the artificially mutated genotypes CIa*H52Y and
206 CIa*R48I (Poulicard et al., 2010). Similar to genotypes CIa*48E*49E and CIa*48G*49E
207 (Poulicard et al., 2010, Poulicard et al., 2012), the reversions observed here suggested a
208 fitness cost in susceptible hosts. Taken together, the results of this study show that, in spite of
209 tight restrictions due to a wide spectrum of constraints, *a priori* unfit genotypes could adopt a
210 wide array of solutions to efficiently overcome strong selective pressures.

211

212 **Figure legends:**

213 **Figure 1:** Location of the alternative resistance-breaking mutations.

214 a. Genomic organisation of RYMV. Alternative RB mutations emerged within the
215 carboxy-terminal domain of polyprotein P2a (black square), while the major mutational
216 pathways involved the VPg (hatched square). ORF, open reading frame; Pro, protease; VPg,
217 viral genome-linked protein; Pol, polymerase; CP, coat protein; fs, -1 frameshift signal.

218 b. Nature of the alternative *rymv1-2* RB mutations in the five viral populations from the
219 infectious clone CIa (accession reference AJ608219) and the mutated clone CIa49E.
220 Mutations are indicated in the two overlapping ORFs (ORF2a and ORF2b after the -1
221 frameshift).

222

223 **Figure 2:** Fitness of the alternative resistance-breaking mutations.

224 a. Viral accumulation of the wild-type CIa and RB genotypes with mutations within and
225 outside the VPg (RB VPg and RB CterP2a, respectively) in resistant plants. The number of
226 viral RNA copies per milligram of fresh leaf tissue was estimated via quantitative reverse
227 transcriptase-polymerase chain reaction (qRT-PCR) at 28 days post-inoculation (dpi) in the
228 resistant cultivar *O.s. indica* cv. Bekarosaka. a, b and c, groups that were significantly
229 different after multiple mean comparison (ANOVA, $P=0.05$).

230 b. Viral accumulation at 14 dpi in the susceptible plants *O.s. indica* cv. IR64.

231

232 **Figure 3:** Distribution of RB mutations in the P2a polyprotein.

233 Arrows: RB mutations; fs: -1 frameshift signal.

234 a. Diversity index (total substitutions/site) calculated using Dnasp.

235 b. Prediction of the folded/unfolded arrangement using FoldIndex.

236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286

References

- Abdul-Razzak, A., Guiraud, T., Peypelut, M., Walter, J., Houvenaghel, M. C., Candresse, T., O, L. E. G. & German-Retana, S. (2009). Involvement of the cylindrical inclusion (CI) protein in the overcoming of an eIF4E-mediated resistance against Lettuce mosaic potyvirus. *Mol Plant Pathol* 10, 109-13.
- Albar, L., Bangratz-Reyser, M., Hébrard, E., Ndjioudjop, M., Jones, M. & Ghesquiere, A. (2006). Mutations in the eIF(iso)4G translation initiation factor confer high resistance of rice to Rice yellow mottle virus. *Plant Journal* 47, 417-426.
- Camps, M., Herman, A., Loh, E. & Loeb, L. A. (2007). Genetic constraints on protein evolution. *Crit Rev Biochem Mol Biol* 42, 313-26.
- Domingo, E., Sheldon, J. & Perales, C. (2012). Viral quasispecies evolution. *Microbiol Mol Biol Rev* 76, 159-216.
- Fargette, D., Pinel, A., Abubakar, Z., Traoré, O., Brugidou, C., Fatogoma, S., Hébrard, E., Choisy, M., Séré, Y., Fauquet, C. & Konaté, G. (2004). Inferring the evolutionary history of rice yellow mottle virus from genomic, phylogenetic, and phylogeographic studies. *Journal of Virology* 78, 3252-3261.
- Fargette, D., Pinel, A., Rakotomalala, M., Sangu, E., Traore, O., Sereme, D., Sorho, F., Issaka, S., Hebrard, E., Sere, Y., Kanyeka, Z. & Konate, G. (2008). Rice yellow mottle virus, an RNA plant virus, evolves as rapidly as most RNA animal viruses. *J Virol* 82, 3584-9.
- Harrison, B. D. (2002). Virus variation in relation to resistance-breaking in plants. *Euphytica*.
- Hébrard, E., Bessin, Y., Michon, T., Longhi, S., Uversky, V. N., Delalande, F., Dorselaer, A. V., Romero, P., Walter, J., Declerk, N. & Fargette, D. (2009). Intrinsic disorder in Viral Proteins Genome-Linked: experimental and predictive analyses. *Virology Journal* 6, e23.
- Hébrard, E., Poulicard, N., Gérard, C., Traoré, O., Wu, H. C., Albar, L., Fargette, D., Bessin, Y. & Vignols, F. (2010). Direct interaction between the Rice yellow mottle virus VPg and the central domain of the rice eIF(iso)4G1 factor correlates with rice susceptibility and RYMV virulence. *Mol Plant Microbe Interact* 23, 1506-1513.
- Librado, P. & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451-1452.
- Ling, R., Pate, A. E., Carr, J. P. & Firth, A. E. (2013). An essential fifth coding ORF in the sobemoviruses. *Virology*.
- Nair, S. & Savithri, H. S. (2010a). Natively unfolded nucleic acid binding P8 domain of SeMV polyprotein 2a affects the novel ATPase activity of the preceding P10 domain. *FEBS Lett* 584, 571-6.
- Nair, S. & Savithri, H. S. (2010b). Processing of SeMV polyproteins revisited. *Virology* 396, 106-17.
- Nakahara, K. S., Shimada, R., Choi, S. H., Yamamoto, H., Shao, J. & Uyeda, I. (2010). Involvement of the P1 cistron in overcoming eIF4E-mediated recessive resistance against Clover yellow vein virus in pea. *Mol Plant Microbe Interact* 23, 1460-9.
- Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P. & Dunker, K. A. (2005). Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins: Structure, Function, and Bioinformatics* 61, 176-182.
- Pinel-Galzi, A., Rakotomalala, M., Sangu, E., Sorho, F., Kanyeka, Z., Traoré, O., Séré, Y., Poulicard, N., Rabenantaondro, Y., Séré, Y., Konaté, G., Ghesquiere, A., Hébrard, E. & Fargette, D. (2007). Theme and variations in the evolutionary pathways to virulence of an RNA plant virus species. *PLoS Pathogens* 3, e180.
- Poulicard, N., Pinel-Galzi, A., Hébrard, E. & Fargette, D. (2010). Why Rice yellow mottle virus, a rapidly evolving RNA plant virus, is not efficient at breaking rymv1-2 resistance. *Molecular Plant Pathology* 11, 145-154.
- Poulicard, N., Pinel-Galzi, A., Traore, O., Vignols, F., Ghesquiere, A., Konate, G., Hebrard, E. & Fargette, D. (2012). Historical contingencies modulate the adaptability of Rice yellow mottle virus. *PLoS Pathog* 8, e1002482.

287 Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., Silman, I.
288 & Sussman, J. L. (2005). FoldIndex: a simple tool to predict whether a given protein
289 sequence is intrinsically unfolded. *Bioinformatics* 21, 3435-3438.

290 Rakotomalala, M., Pinel-Galzi, A., Albar, L., Ghesquière, A., Ramavovololona, P., Rabenantaandro,
291 Y. & Fargette, D. (2008). Resistance to Rice yellow mottle virus in the rice germplasm in
292 Madagascar. *European Journal of Plant Pathology* 122, 277-286.

293 Rakotomalala, M., Pinel-Galzi, A., Mpunami, A., Randrianasolo, A., Ramavovololona, P.,
294 Rabenantoandro, Y. & Fargette, D. (2013). Rice yellow mottle virus in Madagascar and in
295 the Zanzibar Archipelago; island systems and evolutionary time scale to study virus
296 emergence. *Virus Research* 171, 71-79.

297 Rancurel, C., Khosravi, M., Dunker, A. K., Romero, P. R. & Karlin, D. (2009). Overlapping genes
298 produce proteins with unusual sequence properties and offer insight into de novo protein
299 creation. *J Virol* 83, 10719-36.

300 Sanjuan, R. & Nebot, M. R. (2008). A network model for the correlation between epistasis and
301 genomic complexity. *PLoS One* 3, e2663.

302 Satheskumar, P. S., Gayathri, P., Prasad, K. & Savithri, H. S. (2005). "Natively Unfolded" VPg is
303 essential for sesbania mosaic virus serine protease activity. *Journal of Biological Chemistry*
304 280, 30291-300.

305 Satheskumar, P. S., Lokesh, G. L. & Savithri, H. S. (2004). Polyprotein processing: cis and trans
306 proteolytic activities of *Sesbania mosaic virus* serine protease. *Virology* 318, 429-438.

307 Simon-Loriere, E., Holmes, E. C. & Pagan, I. (2013). The effect of gene overlapping on the rate of
308 RNA virus evolution. *Mol Biol Evol* 30, 1916-28.

309 Traoré, O., Pinel-Galzi, A., Issaka, S., Poulicard, N., Aribi, J., Aké, S., Ghesquière, A., Séré, Y., Konaté,
310 G., Hébrard, E. & Fargette, D. (2010). The adaptation of Rice yellow mottle virus to the
311 eIF(iso)4G-mediated rice resistance. *Virology* 408, 103-108.

312 Truniger, V. & Aranda, M. A. (2009). Recessive resistance to plant viruses. *Adv Virus Res* 75, 119-59.

313 Uversky, V. N. (2013). The most important thing is the tail: multitudinous functionalities of
314 intrinsically disordered protein termini. *FEBS Lett* 587, 1891-901.

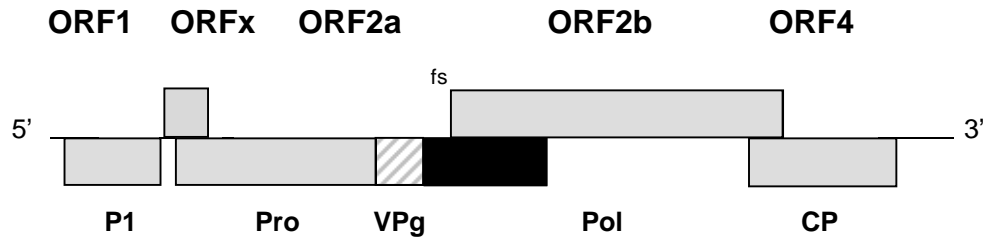
315 Wang, A. & Krishnaswamy, S. (2012). Eukaryotic translation initiation factor 4E-mediated recessive
316 resistance to plant viruses and its utility in crop improvement. *Mol Plant Pathol* 13, 795-
317 803.

318 Weinreich, D. M., Watson, R. A. & Chao, L. (2005). Perspective: Sign epistasis and genetic
319 constraint on evolutionary trajectories. *Evolution* 59, 1165-74.

320

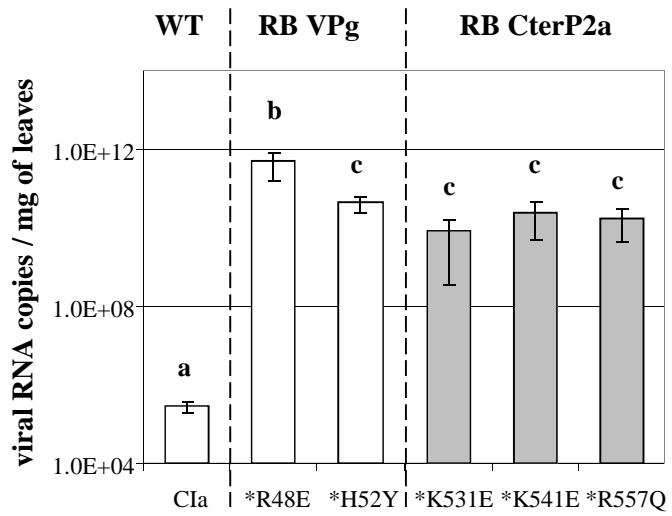
321

a



b

Inoculum	Genotypes	Positions	ORF2a	ORF2b
Cla	Cla*K541E	2229	<u>A</u> AG (K) → <u>G</u> AG (E)	AAA (K) → AGA (R)
	Cla*R557Q	2278	C <u>G</u> G (R) → C <u>A</u> G (Q)	AC <u>G</u> (V) → ACA (A)
Cla49E	Cla49E*K531E	2199	<u>A</u> AG (K) → <u>G</u> AG (E)	G <u>A</u> A (E) → G <u>G</u> A (G)
	Cla49E*G556E	2275	G <u>G</u> A (G) → G <u>A</u> A (E)	AG <u>G</u> (R) → AG <u>A</u> (R)
	Cla49E*K565E	2301	<u>A</u> AG (K) → <u>G</u> AG (E)	CA <u>A</u> (Q) → CG <u>A</u> (R)

a**b**