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

The genome of RYMV consists of a single-stranded, positive-sense RNA molecule with a viral genome-linked protein (VPg) that is covalently linked to its 5’ end. The VPg encoded by the central domain of ORF2a interacts with rice eIF(iso)4G1 (Hébrard et al., 2010). A single amino acid substitution in the middle domain of eIF(iso)4G1 results in the highly resistant Oryza sativa indica cultivars Gigante and Bekarosaka (Albar et al., 2006; Rakotomalala et al., 2008). The phenotype of this recessive allele is characterized by an absence of symptom expression and a lack of viral detection by ELISA. However, adaptation to the rymv1-2 allele has been reported, and the genetic determinism of this adaptation has been elucidated (Pinel-Galzi et al., 2007). The rymv1-2 resistance-breaking (RB) phenotype is caused by point mutations in the VPg, most often located at codon 48, but sometimes at codon 52. The associated major and minor mutational pathways have been described previously.

The rymv1-2 RB ability is related to an E/T polymorphism at the adjacent codon 49 of the VPg (Poulicard et al., 2012). Threonine at codon 49 (T49) confers a strong selective advantage over viral populations harbouring E49 in susceptible and resistant O. glaberrima cultivars, whereas T49 is a major constraint to overcome the rymv1-2 allele found in O. sativa indica cultivars Gigante and Bekarosaka. Antagonistic epistasis between T49 and RB mutations was established through mutagenesis of the infectious clone CIa. Phenotypic and genetic barriers prevented the major
and minor mutational pathways from being followed. The direct influence of the E/T polymorphism at codon 49 on the rymv1-2 RB ability of the WT genotype CIa (with T49) and the mutated genotype CIa*K531E was validated experimentally, and the T49E substitution was found to increase the ability to overcome rymv1-2 resistance from 5% to 40% (Poulicard et al., 2012). The RB pattern of the mutant CIa*K531E was therefore similar to that of other WT viral populations containing E49 (Pinel-Galzi et al., 2007) and mostly involved fixation of the R48G RB mutation (i.e. the first step in the major mutational pathway).

Although the adaptability of the CIa genotype has been assessed previously (Pinel-Galzi et al., 2007), rymv1-2 resistance breakdown has never been observed. In the present study, three of 53 plants (i.e. 5%) inoculated with the CIa genotype showed characteristic RYMV symptoms (Poulicard et al., 2012). The objective of this study was to identify and characterize the mutational pathways involved in rymv1-2 resistance breakdown in the unfit CIa genotype. The VPg of the RB populations of each plant was amplified and directly sequenced according to a method described previously (Fargette et al., 2004). Surprisingly, these RB populations did not show mutations in the VPg. To identify candidate mutations involved in the RB phenotype, the full-length viral genomes of these populations were amplified via reverse transcription-polymerase chain reaction (RT-PCR) and sequenced. Two RB genotypes were characterized by a single mutation (A2229G or G2278A), while no mutation was detected in the third RB genotype. This is the first report of an RB-associated mutation outside the VPg. To further investigate the frequency of these mutations, 20 infected plants of 50 plants inoculated with CIa49E were analysed following the same procedure. Three of the 20 plants infected with the CIa49E genotype also displayed single mutations outside the VPg (A2199G, G2275A and A2301G). These five mutations occurred within a 102 nt stretch in the ORF2a/VPg overlapping region, which was located 376 nt downstream of the VPg (Fig. 1a). These mutations always caused non-synonymous changes in the P2a polyprotein (Fig. 1b), and they generally involved the substitution of a positively charged amino acid, such as lysine or arginine, with the negatively charged amino acid glutamic acid. In contrast, these mutations did not always change the physicochemical properties of the residue in ORF2b. The genotypes harbouring the mutations A2199G, A2229G, G2275A, G2278A and A2301G were subsequently designated CIa49E*K531E, CIa*K541E, CIa*K541E*G556E, CIa*R557Q and CIa49E*K565E, respectively, in reference to the nature and position of the mutations in the P2a polyprotein.

Back-inoculations of five resistant O. sativa indica cultivar Gigante plants with each viral population confirmed the rymv1-2 RB ability of three of them, CIa49E*K531E, CIa*K541E and CIa*R557Q (100% infection rate). To establish the causal role of these mutations, directed mutagenesis of the infectious clone CIa was performed using the QuikChange Site-Directed Mutagenesis kit (Stratagene). Notably, the K531E mutation was introduced without the T49E mutation in the VPg to assess the independence of the two mutations. Transcription of the mutated clones and inoculation of the viral RNAs in planta were performed as previously described (Poulicard et al., 2010). Each mutated clone was inoculated in five susceptible and five resistant individuals of O. sativa indica cultivars IR64 and Bekarosaka, respectively. All mutants were infectious in the susceptible plants (100% infection rate). In all of the resistant plants, characteristic symptoms, high ELISA values, successful RT-PCR amplification and sequencing confirmed that the point mutations K531E, K541E and R557Q were directly involved in the rymv1-2 RB phenotype. Additional mutations did not emerge within the P2a or VPg coding regions. Interestingly, the role of the K531E mutation in resistance breakdown was validated in the absence of the T49E VPg mutation. Therefore, the emergence of K531E is sufficient to overcome rymv1-2 resistance.

To evaluate the accumulation of the three RB genotypes CIa49E*K531E, CIa*K541E and CIa*R557Q, three resistant and three susceptible individuals of O. sativa indica cultivars Bekarosaka and IR64, respectively, were back-inoculated and analysed via qRT-PCR following a protocol previously described (Poulicard et al., 2010). The accumulation of each mutant genotype was compared to that of the WT CIa genotype and two rymv1-2 RB genotypes with mutations in the VPg, i.e. CIa*R48E and CIa*H52Y (Fig. 2a). In this experiment, ~10^{12} RNA copies of each genotype were inoculated per plant. The flanking region of the VPg and the C-terminal region of P2a (nucleotides 1480–2900) were sequenced from the total RNA extracts used for qRT-PCR; no new mutations were detected. At 28 days post-inoculation (p.i.), ~10^{10} copies of the CIa*K531E, CIa*K541E and CIa*R557Q genotypes per milligram of leaf tissue had accumulated in the resistant plants; i.e. their levels were 10^5 times higher than those of the WT CIa genotype (Fig. 2a). No significant differences in RNA accumulation were detected between the CIa*K531E, CIa*K541E and CIa*R557Q genotypes (ANOVA, P>0.05). Interestingly, the accumulation of these RB genotypes was not significantly different from that of the suboptimal CIa*H52Y genotype (P>0.05). CIa*R48E showed maximal accumulation of ~10^{12} RNA copies per milligram of leaf tissue in the resistant cultivar at 28 days p.i., as observed previously (Poulicard et al., 2010), which was approximately 50 times higher than the accumulation of the other RB genotypes (P<0.0001). In the susceptible plants, the RNA accumulation of the CIa*K531E, CIa*K541E, CIa*R557Q and CIa*H52Y genotypes at 14 days p.i. was not significantly lower than that of the WT genotype (P>0.05; Fig. 2b). In contrast, the CIa*R48E genotype was strongly impaired in the susceptible hosts, and its RNA accumulation was ~10^5 times lower than that of the other genotypes. Surprisingly, sequencing of the RB viral population in the susceptible plants revealed the emergence of reverse mutations 28 days after inoculation. The artificially inserted substitution from arginine to glutamic acid (from AGG to GAG) at codon 48.
in VPg of the Cla*R48E genotype was displaced by a mutation that restored the positively charged residue, i.e. lysine (AAG). In addition, reversion of the K531E and K541E mutations was always observed following back-inoculation to *O. sativa indica* IR64. These reverse mutations suggested a fitness cost in susceptible hosts of the mutations

<table>
<thead>
<tr>
<th>Inoculum</th>
<th>Genotypes</th>
<th>Positions</th>
<th>ORF2a</th>
<th>ORF2b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cla</td>
<td>Cla*K541E</td>
<td>2229</td>
<td>AAG (K) → GAG (E)</td>
<td>AAA (K) → AGA (R)</td>
</tr>
<tr>
<td>Cla</td>
<td>Cla*R557Q</td>
<td>2278</td>
<td>CGG (R) → CAG (Q)</td>
<td>ACG (V) → ACA (A)</td>
</tr>
<tr>
<td>Cla49E</td>
<td>Cla49E*K531E</td>
<td>2199</td>
<td>AAG (K) → GAG (E)</td>
<td>GAA (E) → GGA (G)</td>
</tr>
<tr>
<td>Cla49E</td>
<td>Cla49E*G556E</td>
<td>2275</td>
<td>GGA (G) → GAA (E)</td>
<td>AGG (R) → AGA (R)</td>
</tr>
<tr>
<td>Cla49E</td>
<td>Cla49E*K565E</td>
<td>2301</td>
<td>AAG (K) → GAG (E)</td>
<td>CAA (Q) → CGA (R)</td>
</tr>
</tbody>
</table>

**Fig. 1.** Location of the alternative resistance-breaking mutations. (a) Genomic organization of RYMV. Alternative RB mutations emerged within the carboxy-terminal domain of polyprotein P2a (black rectangle), while the major mutational pathways involved the VPg (hatched square). Pro, protease; VPg, viral genome-linked protein; Pol, polymerase; CP, coat protein; fs, −1 frameshift signal. (b) Nature of the alternative *rymv1-2* RB mutations in the five viral populations from the infectious clone Cla (accession reference AJ608219) and the mutated clone Cla49E. Mutations are indicated in the two overlapping ORFs (ORF2a and ORF2b after the −1 frameshift).

**Fig. 2.** Fitness of the alternative resistance-breaking mutations. (a) Viral accumulation of the WT Cla and RB genotypes with mutations within and outside the VPg (RB VPg and RB CterP2a, respectively) in resistant plants. The number of viral RNA copies per milligram of fresh leaf tissue was estimated via quantitative RT-PCR at 28 days p.i. in the resistant cultivar *O. sativa indica* cultivar Bekarosaka. The vertical bars show the SE of the means. Groups i, ii and iii were significantly different after multiple mean comparison (ANOVA, *P* =0.05). (b) Viral accumulation at 14 days p.i. in the susceptible plants *O. sativa indica* cultivar IR64.
that emerged in the C-terminal domain of P2a. No other mutation emerged in the P2a coding region during this fitness experiment, and compensatory mutations were never observed in the susceptible plants.

The natural diversity of the positions involved in the alternative RB pathway and their genetic context in the ORF2a/ORF2b overlapping region were analysed (Table S1, available in JGV Online). Sequence alignment of 33 full-length sequences of viral populations that were representative of the genetic and geographical diversity of RYMV (Rakotomalala et al., 2013) revealed that these positions were strictly conserved. The nucleotide diversity ($\pi$) was estimated for each ORF using DNASP software (Librado & Rozas, 2009). As previously reported (Fargette et al., 2004), the nucleotide variability was lower in ORF2a and ORF2b ($\pi=0.054$ and 0.057) than in ORF1 and ORF4 ($\pi=0.109$ and 0.087). However, these values were still higher than that found for the recently described ORFx which overlaps the 5' end of sobemovirus ORF2a (Ling et al., 2013) ($\pi=0.029$). Interestingly, the RB mutations always emerged in the ORF2a domain, which is characterized by low nucleotide diversity (Fig. 3a). Overlapping viral regions (OVRs) have been reported to be highly conserved, showing strong constraints against synonymous changes (Simon-Loriere et al., 2013). Accordingly, the P2a OVR exhibited a lower diversity of synonymous sites than the VPg, while the proportion of non-synonymous sites was similar (Table S2).

The organization and the biological function of the C-terminal domain of P2a are unknown. However, the processing of the P2a polyprotein of another sobemovirus, sesbania mosaic virus (SeMV), has recently been elucidated. In addition to a serine protease and VPg, two new proteins, designated P10 and P8, were identified (Nair & Savithri, 2010b; Satheshkumar et al., 2004). Similar to the VPg of RYMV, the VPg and P8 of SeMV were demonstrated to be natively unfolded proteins (Hébrard et al., 2009; Nair & Savithri, 2010a; Satheshkumar et al., 2005). Detection of the disordered arrangement of the C-terminal domain of P2a of the RYMV was then performed using the software FOLDINDEX (Obradovic et al., 2005; Prilusky et al., 2005). Prediction analyses indicated an alternating arrangement of folded and unfolded domains in RYMV P2a that was similar to the profile obtained for SeMV, although the sequence identity between sobemoviruses is low (Fig. 3b).

Interestingly, all of the RB mutations described in this study were located within the predicted unfolded segment of the P2a OVR, which strongly suggested that all of these RB mutations also occurred in the RYMV homologue of P8. OVRs have been reported to exhibit more structural disorder than non-overlapping regions (NOVRs) (Rancurel

![Fig. 3. Distribution of RB mutations in the P2a polyprotein. Arrows, RB mutations; fs, −1 frameshift signal. (a) Diversity index (total substitutions/site) calculated using DNASP. (b) Prediction of the folded/unfolded arrangement using FOLDINDEX.](image-url)
et al., 2009). Moreover, the termini of proteins are, on average, more prone to be disordered than internal regions (Uversky, 2013). Because intrinsically disordered protein tails are engaged in a wide range of functions, the C-terminal domain of P2a may be directly or indirectly involved in the interaction between the VPg of RYMV and the eIF(iso)4G1 of rice. Thus, the RB mutations described in this study may favour the capture of eIF(iso)4G1 in resistant plants, which would compensate for the relatively low affinity of the WT VPg with the rymv1-2 eIF(iso)4G1 (Hébrard et al., 2010).

Despite the crucial roles of eFs/VPg interactions in successful infections and in plant resistance breakdown, described in numerous studies [for reviews see Truniger & Aranda (2009) and Wang & Krishnaswamy (2012)], exceptions are occasionally reported. The RB phenotypes induced by lettuce mosaic virus and clover yellow vein virus from the genus Potyvirus are sometimes related to the emergence of mutations in the cylindrical inclusion (CI) and P1 proteins, respectively (Abdul-Razzak et al., 2009; Nakahara et al., 2010). In this study, the detected RB mutations emerged in the C-terminal domain of the P2a polyprotein, which is downstream of the VPg, within the ORF2a/ORF2b overlapping region. This alternative RB mutational pathway was revealed under strong selective constraints. The threonine at position 49 of the VPg was previously reported to be the major genetic constraint blocking the emergence of RB mutations in the VPg. This strong constraint could lead this genotype to adopt an alternative strategy to overcome the rymv1-2 allele. However, the frequency of this alternative mutational pathway in less-constrained viral populations might be underestimated, as suggested by the detection of P2a C-terminal mutations in the artificial genotype Cla49E. Comparison of the identified RB mutational pathways revealed similarities. Although the RB mutations emerged in two different domains, they occurred in two unfolded regions of the same highly conserved P2a polyprotein. The fitness of RB genotypes with mutations in the C-terminal domain of P2a was suboptimal in rymv1-2-resistant plants and intermediate in susceptible plants, as were those of the artificially mutated genotypes Cla*H52Y and Cla*R48I (Poulicard et al., 2010). Similar to genotypes Cla*48E49E and Cla*48G49E (Poulicard et al., 2010, 2012), the reversions observed here suggested a fitness cost in susceptible hosts. Taken together, the results of this study show that, in spite of tight restrictions due to a wide spectrum of constraints, a priori unfit genotypes could adopt a wide array of solutions to overcome strong selective pressures efficiently.

References


