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

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

Running title: Alternative mutational pathways outside the RYMV VPg

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

Corresponding author: eugenie.hebrard@ird.fr, +33(0)4 67 416 289
Word count: summary 134, main text 2496

Summary

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

Main text

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, Sanjuan & 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, 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 protein genome-linked (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 *rymv1-2* allele found 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

characterised 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 punctual 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 (cv.) 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 wild-type genotype CIa (with T49) and the mutated genotype CIa49E 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 CIa49E was therefore similar to that of other wild-type 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 characterise 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 RT-PCR and sequenced. Two RB genotypes were characterised 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, twenty infected plants of fifty plants inoculated with CIa49E were analysed following the same procedure. Three of the twenty plants infected with the CIa49E genotype also displayed single mutations outside the VPg (A2199G, G2275A and A2301G). These five mutations occurred within a 102-nucleotide-long stretch in the ORF2a/ORF2b overlapping region, which was located 376 nucleotides downstream of the VPg (Figure 1a). These mutations always caused non-synonymous changes in the P2a polyprotein (Figure 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 physico-chemical properties of the residue in ORF2b. The genotypes harbouring the mutations A2199G, A2229G, G2275A, G2278A and A2301G were subsequently designated CIa49E*K531E, CIa*K541E, CIa49E*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.s. indica* cv. 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.s. indica* cv. IR64 and cv. 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 punctual 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.s. indica* cv. Bekarosaka and cv. 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 wild-type CIa genotype and two *rymv1-2* RB genotypes with mutations in the VPg, i.e., CIa*R48E and CIa*H52Y (Figure 2a). In this experiment, $c.10^{12}$ RNA copies of each genotype were inoculated per plant. The flanking region of the VPg and the C-terminal region of P2a (nucleotides 1,480-2,900) were sequenced from the total RNA extracts used for qRT-PCR, no new mutations were detected. At 28 dpi, $c.10^{10}$ copies of the CIa*K531E, CIa*K541E and CIa*R557Q genotypes had accumulated per milligram of leaf tissue in the resistant plants; i.e., their levels were 10^5 times higher than those of the wild-type CIa genotype (Figure 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 $c.10^{12}$ RNA copies per milligram of leaf tissue in the resistant cultivar at 28 dpi, as observed previously (Poulicard et al. 2010), which was approximately fifty 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 dpi was not significantly lower than that of the wild-type genotype ($P>0.05$; Figure 2b). In contrast, the CIa*R48E genotype was strongly impaired in the susceptible hosts, and its RNA accumulation was $c.10^5$ 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 CIa*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.s. indica* cv. IR64. These reverse mutations suggested a fitness cost in susceptible hosts of the mutations 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 (Suppl. Table 1). Sequence alignment of 33 full-length sequences of viral populations that were representative

of the genetic and geographic diversity of RYMV (Rakotomalala et al., 2013) revealed that these positions were strictly conserved. The nucleotide diversity (π) 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 (π=0.054 and 0.057) than in ORF1 and ORF4 (π=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) (π=0.029). Interestingly, the RB mutations always emerged in the ORF2a domain, which is characterised by low nucleotide diversity (Figure 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 (Suppl. Table 2).

The organisation 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 (Figure 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 et al., 2009). Moreover, the termini of proteins are, on average, more prone to be disordered than internal regions (Uversky). 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 wild-type VPg with the *rymv1-2* eIF(iso)4G1 (Hébrard et al., 2010).

Despite the crucial roles of eIFs/VPg interactions in successful infections and in plant resistance breakdown, described in numerous studies (for review (Truniger & Aranda, 2009, 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 CIa49E.

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 CIa*H52Y and CIa*R48I (Poulicard et al., 2010). Similar to genotypes CIa*48E*49E and CIa*48G*49E (Poulicard et al., 2010, Poulicard et al., 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 efficiently overcome strong selective pressures.

Figure legends:

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

a. Genomic organisation of RYMV. Alternative RB mutations emerged within the carboxy-terminal domain of polyprotein P2a (black square), while the major mutational pathways involved the VPg (hatched square). ORF, open reading frame; 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 CIa (accession reference AJ608219) and the mutated clone CIa49E. Mutations are indicated in the two overlapping ORFs (ORF2a and ORF2b after the -1 frameshift).

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

a. Viral accumulation of the wild-type CIa 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 reverse transcriptase-polymerase chain reaction (qRT-PCR) at 28 days post-inoculation (dpi) in the resistant cultivar *O.s. indica* cv. Bekarosaka. a, b and c, groups that were significantly different after multiple mean comparison (ANOVA, P=0.05).

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

Figure 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.

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 (Cl) 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., Ndjiondjop, 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., Dorsselaer, 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., Sérémé, D., Poulicard, N., Rabenantaondro, Y., Séré, Y., Konaté, G., Ghesquière, 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.

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

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

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

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

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

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

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

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

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

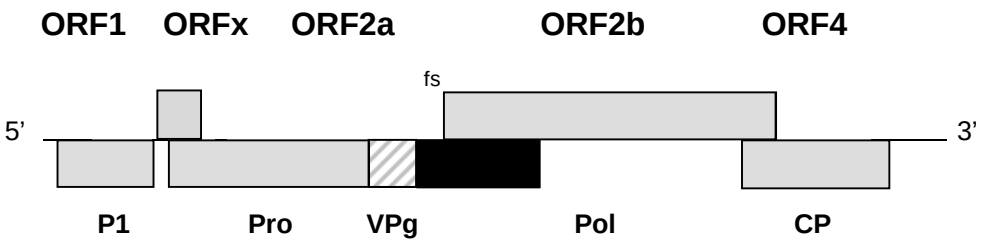
Truniger, V. & Aranda, M. A. (2009). Recessive resistance to plant viruses. *Adv Virus Res* 75, 119-59.

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

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

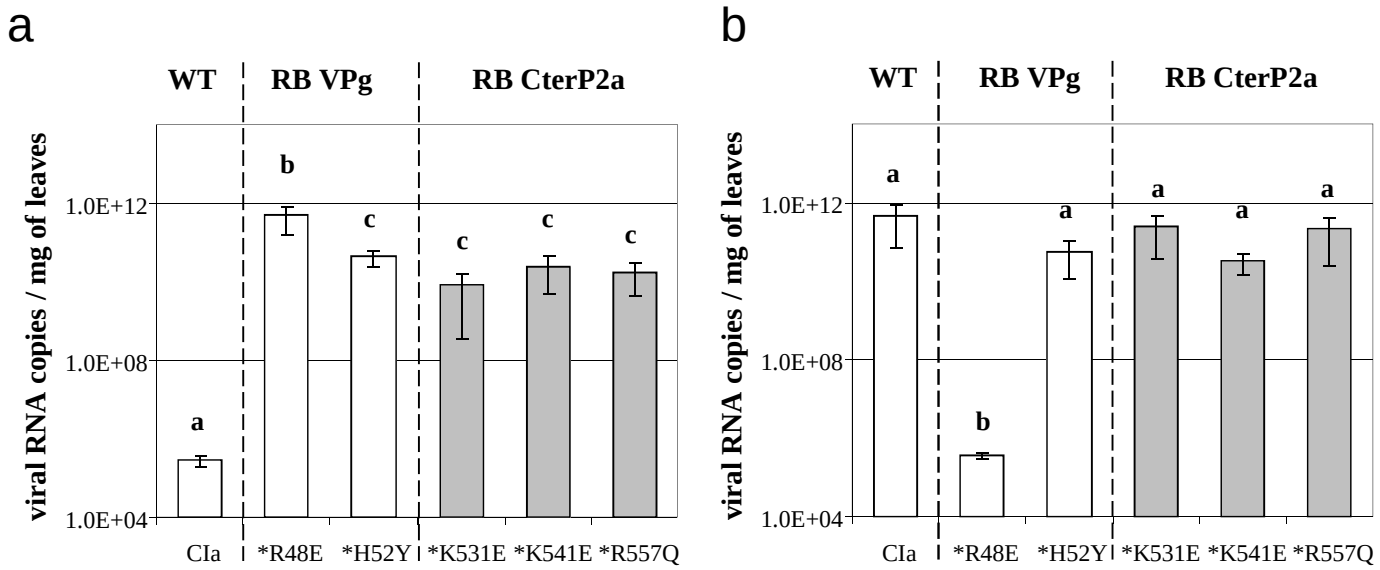
Weinreich, D. M., Watson, R. A. & Chao, L. (2005). Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59, 1165-74.

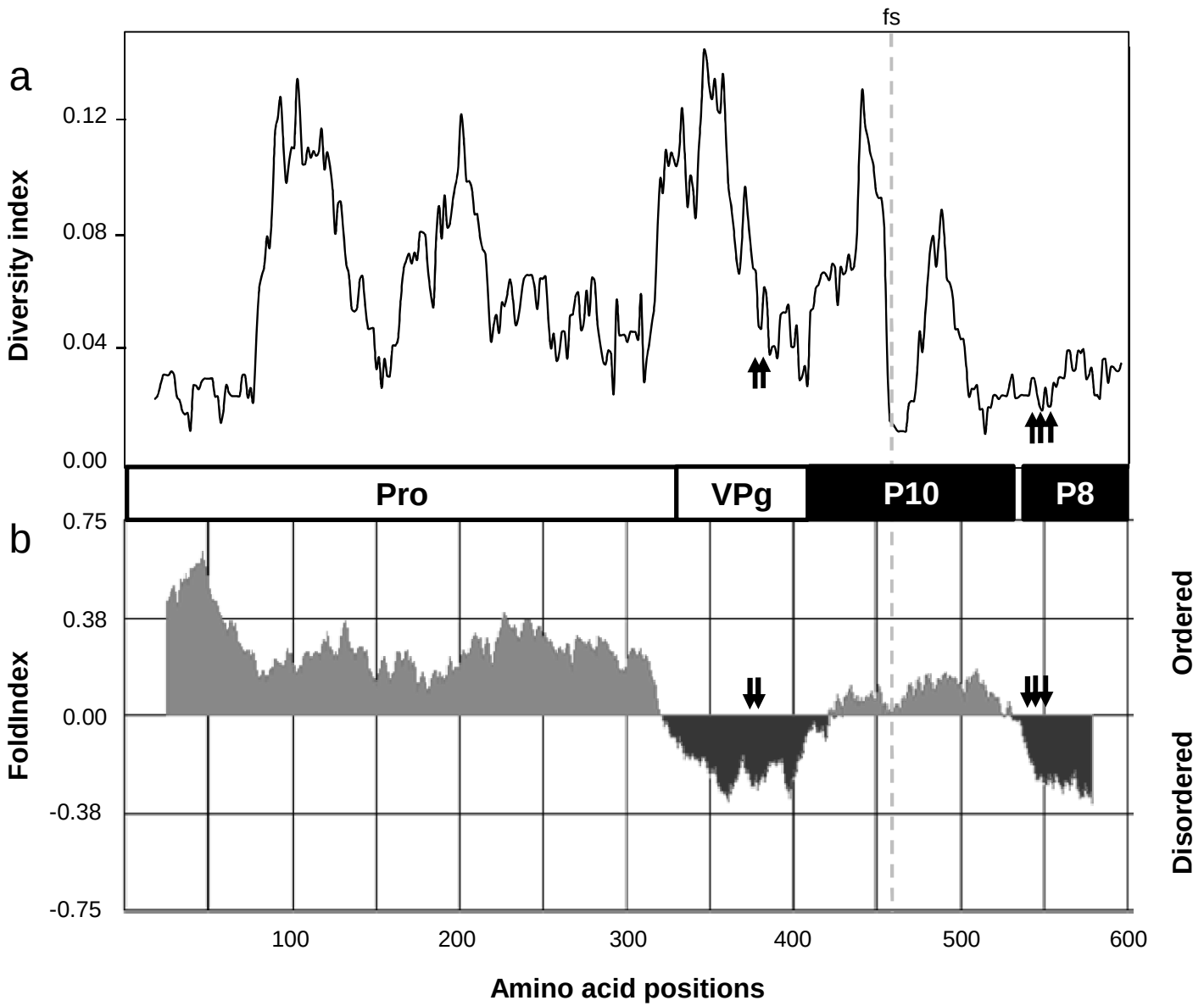
a



b

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


a
WT
RB VPg
RB CterP2a
viral RNA copies / mg of leaves
1.0E+12
1.0E+08
1.0E+04
a
b
c
c
c
c
CIa
*R48E
*H52Y
*K531E
*K541E
*R557Q
b
WT
RB VPg
RB CterP2a
viral RNA copies / mg of leaves
1.0E+12
1.0E+08
1.0E+04
a
b
a
a
a
a
CIa
*R48E
*H52Y
*K531E
*K541E
*R557Q


a
fs
Diversity index
0.12
0.08
0.04
0.00
Pro
VPg
P10
P8
b
0.75
0.38
0.00
-0.38
-0.75
FoldIndex
Ordered
Disordered
100
200
300
400
500
600
Amino acid positions