Adaptation of turnip mosaic virus to *Arabidopsis thaliana* involves rewiring of VPg - host proteome interactions

José L. Carrasco[1], Silvia Ambrós[1], Pablo A. Gutiérrez[2], Santiago F. Elena[1,3,*]

[1]Instituto de Biología Integrativa de Sistemas (CSIC - Universitat de València), Paterna, 46182 València, Spain.

[2]Laboratorio de Microbiología Industrial, Facultad de Ciencias, Universidad Nacional de Colombia, Medellín, Colombia.

[3]The Santa Fe Institute, Santa Fe NM87501, USA.

[*]Correspondence, e-mail: santiago.elena@csic.es.

ORCIDs:

J.L.C. https://orcid.org/0000-0002-5905-8042

S.A. https//orcid.org/0000-0003-3390-3102

P.A.G. https://orcid.org/0000-0002-7003-6485

S.F.E. https://orcid.org/0000-0001-8249-5593

Keywords: experimental evolution; protein-protein interactions; targets of adaptation; virus evolution

Abstract

The outcome of a viral infection depends on a complex interplay between the host physiology and the virus, mediated through numerous protein-protein interactions. In a previous study we used high-throughput yeast two-hybrid (HT-Y2H) to identify proteins in *Arabidopsis thaliana* that bind to the proteins encoded by the turnip mosaic virus (TuMV) genome. Furthermore, after experimental evolution of TuMV lineages in plants with mutations in defense-related or proviral genes, most mutations observed in the evolved viruses affected the VPg cistron. Among these mutations, D113G was a convergent mutation selected in many lineages across different plant genotypes. In contrast, mutation R118H specifically emerged in the *jin1* mutant with affected jasmonate signaling. Using the HT-Y2H system, we analyzed the impact of these two mutations on VPg's interaction with plant proteins. Interestingly, both mutations severely compromised the interaction of VPg with the translation initiation factor eIF(iso)4E, a crucial interactor for potyvirus infection. Moreover, mutation D113G, but not R118H, adversely affected the interaction with RHD1, a zinc-finger homeodomain transcription factor involved in regulating DNA demethylation. Our results suggest that RHD1 enhances plant tolerance to TuMV infection.

1. Introduction

Viruses establish multiple contact points with the host proteome, primarily through a limited number of proteins, particularly reduced in the case of RNA viruses (Belshaw, Pybus and Rambaut 2007; Mahmoudabadi et al. 2018). This interaction involves virus-host protein-protein interactions (PPIs) that result in the manipulation of various cellular pathways for creating favorable replication conditions. This manipulation can occur by sequestering cell resources for the virus's benefit or by interfering with immune responses. Simultaneously, host cell receptors play a role in sensing foreign elements and responding accordingly. Recent years have seen significant progress in creating detailed, high-quality maps of virus-host PPIs (Uetz et al. 2006; de Chassey et al. 2008). Integrative approaches have identified both general and specific molecular mechanisms employed by different viruses (Dyer et al. 2008; Mukhtar et al. 2011; Pichlmair et al. 2012; Ahmed et al. 2018; Aguirre and Guantes 2023). These findings, coupled with additional omics data, have been used to predict phenotypic outcomes of infection (Tisoncik-Go et al. 2016; Cervera et al. 2018; Tarazona, Forment and Elena 2020). These advancements build upon the continuous development of a comprehensive large-scale map of host PPIs, which, while not fully complete, proves valuable in recognizing disease-associated modules (Mukhtar et al. 2011; Menche et al. 2015).

In a previous study Martínez et al. (2023) systematically identified direct PPIs between turnip mosaic virus strain YC5 (TuMV; species *Turnip mosaic virus*, genus *Potyvirus*, family *Potyviridae*) and one of its natural hosts, *Arabidopsis thaliana* (L.) HEYNH, using high-throughput yeast two-hybrid (HT-Y2H) screenings. This study uncovered 378 new PPIs between TuMV and plant proteomes. Among the viral proteins, the RNA-dependent RNA polymerase NIb emerged with the highest number of contacts, including crucial salicylic acid (SA)-dependent transcription regulators. The authors constructed and analyzed a network consisting of 399 TuMV-*A. thaliana* interactions, incorporating intravirus and intrahost connections. Notably, their findings revealed that TuMV-targeted host proteins (*i*) were enriched in various aspects of plant responses to infections, (*ii*) exhibit greater connectivity, (*iii*) had enhanced capacity to disseminate information throughout the cell proteome, (*iv*) showed higher expression levels, and (*v*) were under stronger purifying selection than expected by chance.

In previous evolution experiments with potyviruses, the VPg cistron has been identified as a significant target of selection. For instance, Agudelo-Romero *et al*. (2008) observed that a single amino acid replacement in VPg significantly increased tobacco

81 etch virus (TEV) infectivity, severity of symptoms, and viral load in *A. thaliana*. Later
82 on, Hillung et al. (2014) further described additional VPg mutants in several *A. thaliana*
83 ecotypes that differed in their susceptibility to TEV infection. Likewise, Navarro et al.
84 (2022) and Ambrós et al. (2024) evolved TuMV lineages in different defense-deficient
85 genotypes of *A. thaliana*, observing pervasive mutations affecting VPg, with remarkable
86 examples of parallel evolution with the same mutations fixing in lineages evolved in the
87 same host genotype. Gallois et al. (2010) found that *A. thaliana* plants with knock-out
88 mutations in the *EUKARYOTIC TRANSLATION INITIATION FACTORS (ISO)4E*,
89 *(ISO)4G1* and *(ISO)4G2* (*eIF(iso)4E*, *eIF(iso)4G1* and *eIF(iso)4G2*, respectively) genes
90 were resistant to TuMV infection, demonstrating the proviral role of these eIFs.
91 However, two mutations in VPg were sufficient to overcome this resistance and revert to
92 the original infection phenotype. Notably, Y2H assays showed that these mutations did
93 not preclude the binding of VPg to eIF(iso)4E (Gallois et al. 2010). As another example,
94 the *pvr2* locus, one of the most extensively used resistance genes against potato virus Y
95 (PVY) in pepper cultivars, encodes the eIF4E factor, which physically interacts with VPg.
96 Interestingly, all resistance-breaking viral isolates identified so far contain mutations in
97 the VPg cistron (Duprat et al. 2002; Moury et al. 2004; Ayme et al. 2006).
98 VPg has been described as a scaffolding protein that interacts with other potyviral
99 proteins (Bosque et al. 2014; Hafrén, Löhmus and Mäkinen 2015) and with many host
100 proteins (Martínez et al. 2016, 2023), most notably factors involved in genome
101 transcription, being linked to the 5'-end of the viral genome and providing the hydroxyl
102 group that primes the synthesis of complementary strands by the viral NIb replicase
103 (Eskelin et al. 2011). Additionally, it plays a role in protein synthesis by directly
104 interacting with canonical translation factors eIF(iso)4E and eIF(iso)4G (Gallois et al.
105 2010; Hafrén, Löhmus and Mäkinen 2015) and with the antiviral factor eIF4E1 that
106 protects the translational machinery during TuMV infection (Zafirov et al. 2023). Indeed,
107 Martínez et al. (2023) HT-Y2H analyses and literature curation resulted in the
108 identification of 43 direct interactors, involved in diverse cellular processes. Amongst
109 the most highly expressed and connected genes, *CALNEXIN 1* (*CNX1*) and *OBERON 1*
110 (*OBE1*) were shown to directly interact with VPg. Both CNX1 and OBE1 proteins were
111 proviral factors whose knocked-out expression resulted in late and less severe symptoms.
112 Other genes such as *NONEXPRESSER OF PATHOGENESIS-RELATED GENES 1*
113 (*NPR1*), *SUMO 3* (*SUM3*), *SUMO CONJUGATING ENZYME 1* (*SCE1*), and specially

114 *TGACG SEQUENCE-SPECIFIC BINDING FACTOR 1* (*TGA1*) also resulted in milder
115 disease symptoms, although their interaction with VPg was mediated by NIb.
116 In various evolution experiments of TuMV isolate YC5 in different *A. thaliana*
117 genotypes, it has consistently been observed that VPg fixed nonsynonymous mutations;
118 14 identified as being subject to positive selection (González, Butković, and Elena 2019;
119 González et al. 2021; Navarro et al. 2022; Melero, González, and Elena 2023; Ambrós et
120 al. 2024). The majority of these positively selected mutations are concentrated within a
121 helix-coil-helix structural domain, spanning amino acids 110 to 120 [Fig. 4B in Ambrós
122 et al. (2024)]. Certain mutations were consistently present across lineages that evolved
123 in different host genotypes and environmental conditions. For instance, the VPg^{D113G}
124 mutation was observed in lineages evolved in nine distinct plant genotypes, including
125 wild-type Col-0 plants, and two environmental conditions (González, Butković, and
126 Elena 2019; González et al. 2021; Navarro et al. 2022; Ambrós et al. 2024). Conversely,
127 some mutations were specific to particular host genotypes, such as the VPg^{R118H} mutation,
128 found exclusively in lineages evolved in mutants of *JASMONATE INSENSITIVE 1* (*JIN1*)
129 and *DECREASED DNA METHYLATION 1* (*DDM1*) genes, both being particularly
130 permissive to TuMV infection (Navarro et al. 2022; Ambrós et al. 2024).
131 Given the prominent role played by VPg during potyvirus infections, its apparent
132 prominence as target of natural selection and its promiscuous interaction with other
133 potyviral and host proteins, a question that immediately arises from all previously
134 mentioned studies is the extent to which these mutations affect the ability of VPg to
135 engage into PPIs resulting in fitness increases. To tackle this question, we have selected
136 mutants VPg^{D113G} and VPg^{R118H} as examples of what can be considered generalist and
137 specialist versions of VPg, respectively. The two VPg mutants were cloned and used as
138 bait proteins in HT-Y2H experiments to identify host interactors that differ from those
139 obtained for VPg^{wt} by Martínez et al. (2023). Then, the relative affinity of the differential
140 interactors was evaluated by a quantitative assay and compared with the VPg^{wt}. To
141 further characterize the effect of these mutations in TuMV performance, we inoculated
142 *eif(iso)4e* and *ros1-associated homeodomain protein 1* (*rhd1*) mutant plants and
143 characterized disease progression and viral accumulation. Finally, we performed *in silico*
144 evaluations of the effect of the two mutations of interest in VPg folding and used docking
145 and molecular dynamics techniques to evaluate their effect on the interaction with
146 eIF(iso)4E and RHD1.
147

148

149 ## 2. Materials and Methods

150 ### 2.1. Plasmid construction

151 The mutated variants of the VPg cistron were amplified by PCR using Phusion High-
152 Fidelity DNA polymerase (Thermo Scientific, Waltham MA, USA) according to
153 manufacturer's recommendations and using primers specifically designed for the In-
154 Fusion cloning system (Takara Bio, Mountain View CA, USA). The amplified PCR
155 fragments were recombined into the yeast bait two-hybrid vector pGBKT7 (Takara Bio,
156 Kusatsu, Japan), previously digested with *Eco*RI and *Bam*HI, using the In-Fusion
157 enzyme. Thus, translational fusions of all three VPg protein variants (bait protein) with
158 the GAL4 DNA-binding domain were generated.

159

160 ### 2.2. High-throughput yeast two-hybrid screening

161 To search for differential interactors among the original VPg protein and the two selected
162 mutated variants, an *A. thaliana* Col-0 cDNA library (Takara Bio) transformed into the
163 *Saccharomyces cerevisiae* Y187 strain (prey strain) was screened using the Matchmaker®
164 Gold Yeast Two-Hybrid System (Takara Bio). For that purpose, the three bait VPg
165 proteins cloned in pGBKT7 were transformed in the yeast strain Y2HGold. Then, the
166 Y187 prey strain was mated to each of the Y2HGold haploid bait strains expressing the
167 wild-type and the mutated variants of VPg protein and plated on a double dropout medium
168 (SD/-Leu/-Trp) containing 40 μg/mL X-α-Gal and 200 ng/mL aureobasidin A. Co-
169 transformants displaying α-galactosidase activity were subjected to a second round of
170 selection on a quadruple dropout medium (SD/-Leu/-Trp/-Ade/-His) also containing X-
171 α-Gal. Plasmids pGBKT7-T-antigen, pGADT7-laminin C, and pGADT7-murine p53
172 (Takara Bio) were used as negative controls. cDNA inserts from positive yeast clones
173 were amplified by colony PCR, digested with *Alu*I and analyzed by agarose gel
174 electrophoresis to eliminate duplicate clones. Selected prey plasmids were rescued using
175 a glass-bead extraction protocol and transformed into *Escherichia coli* DH5α for
176 amplification and sequencing. DNA and protein sequence analyses were performed with
177 the WU-BLAST algorithm. For each positive interaction, the prey plasmid was re-
178 transformed into the yeast strain Y2HGold and mated to all three bait strains and also to
179 a control Y187 yeast strain transformed with the empty pGBKT7 vector to remove false
180 positives.

181

182 **2.3. Plant genotypes, virus clones and inoculation procedures**

183 Wild-type and two *A. thaliana* mutants were used in this study, all in the same Col-0
184 background. Seeds of the T-DNA insertion line SALK_092897C (*rhd1-1* mutant) were
185 obtained from the Nottingham Arabidopsis Stock Center (NASC), whereas *eif(iso)4e*
186 seeds were a kind gift from Dr. Jean-Luc Gallois (INRA, Avignon, France) (Duprat et al.
187 2002).

188 TuMV infectious sap was obtained from TuMV-infected *Nicotiana benthamiana*
189 DOMIN plants inoculated with the infectious plasmid p35STunos that contains a cDNA of
190 the TuMV genome isolate YC5 from calla lily (*Zantesdeschia* sp) (GeneBank
191 AF530055.2) under the control of the cauliflower mosaic virus 35S promoter and the *nos*
192 terminator (Chen et al. 2003) as described elsewhere (González, Butković and Elena
193 2019; Corrêa et al. 2020). After plants showed symptoms of infection, they were pooled
194 and frozen with liquid N_2. This frozen plant tissue was homogenized into a fine powder
195 using a Mixer Mill MM400 (Retsch GmbH, Haan, Germany). For inoculations, 0.1 g of
196 powder was diluted in 1 mL inoculation buffer [50 mM phosphate buffer (pH 7.0), 3%
197 PEG6000, 10% Carborundum] and 5 µL of the inoculum was gently rubbed onto two
198 leaves per plant. As a control, another set of plants were mock-inoculated only with
199 inoculation buffer. Plants were all inoculated when they reached growth stage 3.5 in the
200 Boyes et al. (2001) scale. This synchronization ensures that all hosts were at the same
201 phenological stage when inoculated.

202 VPg mutations D113G and R118H were introduced by site-directed mutagenesis in
203 the p35STunos plasmid using the Quikchange II XL kit (Agilent Technologies, Santa
204 Clara CA, USA) and following the indications of the manufacturer, resulting in plasmids
205 p35STunos-VPg^{D113G} and p35STunos-VPg^{R118H}. Plasmids were purified using the
206 ZymoPURE II Plasmid Miniprep kit (Zymo Research, Irving CA, USA) and resuspended
207 in sterile water.

208 For evaluation of the mutated VPg alleles, wild-type, *eif(iso)4e* and *rhd1-1* plants
209 were inoculated with p35STunos, p35STunos-VPg^{D113G} and p35STunos-VPg^{R118H}. One
210 leaf per plant was inoculated by gentle abrasion after applying 3 µg of plasmid DNA
211 diluted in 6 µL of water supplemented with 10% Carborundum.

212

213 **2.4. Infectivity, symptoms progression and quantification of viral load**

214 The number of infected symptomatic plants, out of 20, as well as the severity of symptoms
215 in a semi-quantitative scale in which zero means no symptoms and 5 systemic necrosis
216 [Fig. 1 in Butković et al. (2021)] were monitored every day until 14 days post-inoculation
217 (dpi). As a control, another set of mock-inoculated plants was treated in the same way.
218 Viral loads were measured 12 dpi in four different pools of five infected plants each,
219 by absolute real-time quantitative RT-PCR (RT-qPCR) using a TuMV standard curve and
220 the primers TuMV F117-fw ($^{5'}$CAATACGTGCGAGAGAAGCACAC$_{3'}$) and F118-rv
221 ($^{5'}$TAACCCCTTAACGCCAAGTAAG$_{3'}$) that amplify a 173 nucleotides fragment from
222 the CP cistron of TuMV genome, as previously described (Corrêa et al. 2020). Briefly,
223 standard curve consisted of nine serial dilutions of the *in vitro* synthesized TuMV genome
224 prepared in total plant RNA purified from healthy *A. thaliana* plants. Amplification
225 reactions were run in a 20 µL volume using the qPCRBIO SyGreen 1-step Go Hi-ROX
226 System (PCRBiosystems, London, UK) and the recommended manufacturer's
227 instructions in an ABI StepOne Plus Real-time PCR System (Applied Biosystems, Foster
228 City CA, USA). An initial RT phase of 10 min at 45 ºC and 2 min at 95 ºC was followed
229 by a PCR stage with the following cycling conditions: 40 cycles of 5 s at 95 ºC and 25 s
230 at 60 ºC, and a final melting curve profile analysis that consisted of 15 s at 95 ºC, 1 min
231 at 60 ºC and 15 s at 95 ºC. As negative controls, total plant RNA (noninfected control)
232 and water were included in the analysis. Quantitative reactions were run as three
233 technical replicates per sample and results were analyzed using the StepOne software
234 2.2.2 (Applied Biosystems).
235
236 **2.5. Determination of TuMV consensus sequence**
237 The consensus sequences of the viral genomes accumulated at the end of the infection
238 processes were generated as described elsewhere (*e.g.*, Navarro et al. 2022). In short,
239 total RNAs were obtained and amplified by high-fidelity RT-PCR using the AccuScript
240 Hi-Fi (Agilent Technologies) reverse transcriptase and Phusion DNA polymerase
241 (Thermo Scientific) following the manufacturer's instructions. Each complete TuMV
242 genome was amplified as three overlapping amplicons using three specific sets of primers
243 and the amplification conditions described in Navarro et al. (2022). PCR products were
244 purified with the MSB Spin PCRapace kit (Stratec Molecular, Coronado CA, USA) and
245 then Sanger-sequenced by Novogene Europe Co. Ltd. (Cambridge, UK). Full-length
246 consensus viral sequences were obtained assembling the sequences of the three amplified

247 products by using the Genious R9.0.2 program (Dotmatics, Bishop's Stortford, UK). No
248 additional mutations were found in any of the VPg clones.
249
250 **2.6. Host gene expression analysis by relative RT-qPCR**
251 For the analysis of *RHD1* and *TGA1* expression, RNA was purified with the Quick-RNA
252 Plant Miniprep kit (Zymo Research) following the protocol recommended by the
253 manufacturer. RT-qPCR was performed in an ABI OneStep Plus Real-time PCR System
254 (Applied Biosystems) from 50 ng RNA using the qPCRBIO SyGreen 1-step Go Hi-ROX
255 System from (PCRBiosystems). The *AT1G13320* corresponding to the *PROTEIN*
256 *PHOSPHATASE 2A SUBUNIT A3* (*PP2AA3*) gene was used as endogenous reference
257 (Czechowski et al. 2005). The primers used for the amplification were pp2aa3-fw
258 ($^{5'}$TTGGTGCTCAGATGAGGGAGAG$_{3'}$), pp2aa3-rv
259 ($^{5'}$TTCACCAGCTGAAAGTCGCTTAG$_{3'}$), rhd1-fw
260 ($^{5'}$TATCGGAACCACCGCTTACG$_{3'}$), rhd1-rv ($^{5'}$CGGTGGAAGTTACGGTGACA$_{3'}$),
261 tga1-fw ($^{5'}$ACGAACCTGTCCATCAATTCGG$_{3'}$), and tga1-rv
262 ($^{5'}$CCATGGGAAGTATCCTCTGACACG$_{3'}$). For expression in leaves, plants were
263 grown for about 4 weeks [growth stage ~3.70 in the Boyes et al. (2001) scale].
264
265 **2.7. Statistical analyses**
266 Quantitative specificity α-galactosidase assays were conducted per interactor and VPg
267 allele, with repetitions ranging from 1 to 5 times (median 2) in independent blocks. To
268 mitigate potential block effects, individual measures were normalized by the mean of the
269 VPg^{wt} allele estimated in the corresponding block. Normalized data underwent analysis
270 using Welch's robust one-way ANOVA, and effect magnitudes were assessed using the
271 η^2 statistic. Conventionally, $\eta^2 < 0.05$ are considered small, $0.05 \leq \eta^2 < 0.15$ as medium,
272 and $\eta^2 \geq 0.15$ as large effects.
273 Relative expression data for *RHD1* and *TGA1* were fitted to a generalized linear
274 model (GLM) with infection status (mock-inoculated *vs*. TuMV-VPg^{wt} infected) and days
275 post-infection (dpi) as orthogonal factors, utilizing a Gamma distribution and log-link
276 function. In this context, the magnitude of effects was assessed using the η_P^2, with similar
277 criteria than for η^2 above.
278 The number of infected plants, out of 20, observed for wild-type and *rhd1* plants
279 were independently analyzed using Kaplan-Meier regression. Pairwise comparisons of
280 the median time to the appearance of first symptoms for each VPg allele were performed

281 using log-rank tests. Symptom severity progression data were independently fitted to
282 repeated-measures ANOVAs, with dpi as an intra-individual factor and VPg allele as an
283 inter-individual factor. The magnitude of effects was evaluated using the η_P^2 statistic.
284 Viral load data were fitted to a GLM with plant genotype and VPg allele as
285 orthogonal factors, and experimental block nested within the interaction of the two
286 orthogonal factors, employing a Gaussian distribution and identity-link function. The
287 magnitude of effects was assessed using the η_P^2 statistic.
288 In all reported pairwise *post hoc* tests, the Bonferroni sequential method was
289 applied.
290
291 **2.8. Homology modelling**
292 Models were generated with MODELLER version 10.4 (Webb and Sali, 2016). For the
293 *A. thaliana* eIF(iso)4E, and its paralogs eIF4E and nCBP, the X-ray structures of eIF4E
294 from *Cucumis melo* (Miras et al. 2017), *Pisum sativum* (Ashby et al. 2011), and *Mus*
295 *musculus* (Rydzik et al. 2017; Wan et al. 2020; Wojcik et al. 2021) were used as templates
296 (Supplementary Fig. S1). The model of the VPg protein from TuMV was inferred using
297 the NMR structure of the PVY homolog (Coutinho de Oliveira et al. 2019). Input
298 alignments were generated with MUSCLE (Edgar, 2004). In each case, the model lowest-
299 energy model out of 100 structures was selected and submitted to a final round of
300 refinement using ROSETTA's relax protocol (Conway et al. 2013).
301
302 **2.9. Docking**
303 Docking was performed with the Monte Carlo based multi-scale docking algorithm
304 implemented in RosettaDock (Chaudhury et al. 2011). In each case, high-energy
305 rotamers were removed prior to docking using the Docking Prepack application from
306 ROSETTA (Wang et al. 2005). For the eIF(iso)4E and VPg docking model, structures
307 were prepositioned near each other with the binding pockets facing each other using
308 HADDOCK (Dominguez et al. 2003; van Zundert et al. 2016; Honorato et al. 2021). In
309 this case, ambiguous interaction restraints (AIRs) were generated using a combination of
310 NMR chemical shift perturbations (Coutinho de Oliveira et al. 2019), mutagenesis data
311 on the eIF4E/eIF(iso)4E - VPg system (Monzingo et al. 2007; Roudet-Tavert et al. 2007;
312 German-Retana et al. 2008; Ashby et al. 2011; Pérez et al. 2012; Svanella-Dumas et al.
313 2014), and experimental data from Martínez et al. (2023). Unfolded and mobile regions

314 were removed from both proteins and docking site constraints for residues known to be
315 in contact in the eIF(iso)4E/VPg interaction were included in local docking protocol
316 (Supplementary Fig. S2). A total of 1000 decoys were generated during the docking
317 perturbation runs. Residues in the interface of the selected model were refined using a
318 high-resolution full atom minimization protocol (Conway et al. 2013). After this step,
319 the mobile regions of VPg and eIF(iso)4E were included using MODELLER, and an
320 optimal conformation was recalculated using the kinematic closure method in ROSETTA
321 (Mandell et al. 2009) (Supplementary Fig S2). The final complex was further refined
322 using a molecular dynamic simulation in explicit solvent with NAMD (Phillips et al.
323 2020). This step was performed using periodic boundary conditions in a water box with
324 a 10 Å padding, 0.15 M NaCl, and a temperature of 298 K (Spivak et al. 2023). Docking
325 of RHD1 and the VPg protein was performed using a global docking protocol using the
326 AlphaFold RHD1 model available at Uniprot (Jumper et al. 2021). Docking was
327 performed separately on N- and C- terminal Zn-finger and homeobox domains,
328 respectively. After minimization, structures were fitted using a global docking protocol
329 using 10,000 decoys for each RHD1 domain. Other RHD1 PPIs were pulled out from
330 STRING database version 12.0 (Szklarczyk et al. 2023).
331
332

333 ## 3. Results and Discussion

334 ### 3.1. HT-Y2H screens and evaluation of differences in affinity between VPg variants
335 and host interactors

336 Three screens were conducted in parallel using wild-type and two mutated VPg variants,
337 D113G and R118H, as the respective bait proteins. The number of clones screened
338 ranged from $2 \cdot 10^7$ to $2.5 \cdot 10^7$ per variant, with a mating efficiency of 30% to 35%.
339 Initially, approximately 400 clones were selected, from which six for VPg^{wt}, seven for
340 VPg^{D113G}, and eight for VPg^{R118H} were further validated through a second round of more
341 rigorous selection and semiquantitative evaluation of binding specificity after plasmid
342 rescue and yeast retransformation. Subsequently, a quantitative α-galactosidase activity
343 assay was performed with the selected clones. Only clones displaying significant
344 statistical differences among alleles were subjected to gene identification through
345 sequencing. Following sequencing, two clones selected for VPg^{D113G} corresponded to
346 *AT5G42780*, while two clones selected for VPg^{wt} matched with *AT5G35620*.

347 Consequently, we identified 19 host proteins that interact differentially among the three
348 VPg alleles.

349 Fig. 1 depicts a heatmap constructed based on the binding affinities similarity
350 matrix among the 19 host proteins and the three VPg alleles. Among the interactors, nine
351 exhibited a significant overall difference among the three VPg alleles (Supplementary
352 Fig. S3). The smallest significant effect was observed with *AT2G19270*, which encodes
353 the *PROLINE RICH MITOTIC CHECKPOINT CONTROL FACTOR* (*PRCC*). PRCC is
354 involved in regulating protein modifications and signal transductions (Welch's robust
355 ANOVA: $F_{2,25.451} = 4.326$, $P = 0.024$, $\eta^2 = 0.171$). On the other hand, the most substantial
356 effect was observed with *AT5G35620*, the locus encoding the well-characterized VPg
357 interactor eIF(iso)4E protein ($F_{2,67.604} = 675.852$, $P < 0.001$, $\eta^2 = 0.958$).

358 Applying Bonferroni *post hoc* pairwise difference tests, we compared the affinity
359 values between VPg^{wt} and each of the 19 host interactors with those observed for the two
360 VPg mutant alleles, revealing a remarkably diverse pattern of effects. As illustrated in
361 Fig. 1, these 19 proteins can be categorized into three distinct groups. In the following
362 discussion, we will delve into the host proteins exhibiting differential binding strength
363 with the VPg alleles. The first group comprises two proteins whose interactions are
364 significantly more robust with VPg^{wt} than with either of the two mutant alleles. These
365 proteins are: (*i*) the aforementioned eIF(iso)4E ($P < 0.001$; 90.50% weaker interaction)
366 and (*ii*) the FRUCTOKINASE-LIKE 1 (FLN1) protein, a member of the pfkB-
367 carbohydrate kinase family encoded by locus *AT3G54090* ($P < 0.001$; 36.54% weaker
368 interaction). FLN1 serves as a potential plastidial target of thioredoxin z and is crucial
369 for proper chloroplast development. Its role extends to the regulation of plastid-encoded
370 polymerase (PEP)-dependent chloroplast transcription. Consequently, *fln1* mutants
371 exhibit aberrant chloroplast development, general developmental defects, and impaired
372 PEP-dependent transcription (Gilkerson et al. 2012).

373 The second group in Fig. 1 is the most heterogeneous, encompassing three host
374 proteins that exhibit either no differences with VPg^{wt} or variable effects with VPg^{D113G}
375 and VPg^{R118H}. (*i*) The *AT5G62580* locus encodes ARM-REPEAT SUPERFAMILY
376 PROTEIN and demonstrates a 28.76% reduction in binding strength with VPg^{D113G}
377 compared to VPg^{wt} ($P = 0.025$), while it shows no significant effect on VPg^{R118H} ($P =$
378 1.000). This family of proteins binds to microtubules, impacting cellular organization
379 and organ growth (Buschmann et al. 2004). Intriguingly, Martínez et al. (2023) described
380 this protein as an interactor of NIb replicase. (*ii*) The second protein in this group is

381 RHD1 (encoded by *AT5G42780*). RHD1, a zinc finger and homedomain (ZF-HD)
382 protein, interacts with ROS1-ASSOCIATED METHYL-DNA BINDING PROTEIN 1
383 (RMB1), ROS1-ASSOCIATED WD40 DOMAIN-CONTAINING PROTEIN (RWD40)
384 and REPRESSOR OF SILENCING 1 (ROS1) in a multiprotein complex involved in the
385 base excision repair pathway through DNA methylation. This interactor was not
386 described by Martínez et al. (2023). Interestingly, RHD1 interacts with the transcriptional
387 regulator TGA1, known to act as a proviral factor via its interaction with NIb (Martinez
388 et al. 2023). RHD1 exhibits a 22.62% decrease in binding affinity with VPg^{D113G} ($P <$
389 0.001) but a 17.35% increase with VPg^{R118H} ($P = 0.003$) compared to VPg^{wt}. (*iii*) The
390 third member of this group is the product of the *PRCC* gene, involved in the regulation
391 of protein modifications and signal transductions. While VPg^{D113G} shows no differences
392 in binding compared to VPg^{wt} ($P = 1.000$), VPg^{R118H} exhibits 22.5% stronger binding (P
393 $= 0.029$).
394 The third group in Fig. 1 comprises four host proteins that exhibit no differences in
395 affinity between VPg^{wt} and VPg^{D113G} but consistently show increased binding strength
396 with VPg^{R118H}. (*i*) The product of the gene *PYRUVATE ORTHOPHOSPHATE*
397 *DIKINASE (PPDK) REGULATORY PROTEIN 1* (*RP1*) (*AT4G21210*) demonstrates
398 47.15% stronger binding with VPg^{R118H}. RP1, a bifunctional serine/threonine kinase and
399 phosphorylase, is involved in the dark/light-mediated regulation of PPDK by catalyzing
400 its phosphorylation/dephosphorylation (Chastain et al. 2007). Notably, RP1 was not
401 described by Martínez et al. (2023). (*ii*) The second host protein in this group is
402 OBERON 1 (OBE1), encoded by locus *AT3G07780*; this interaction was previously
403 described by Martínez et al. (2023). OBE1, a nuclear PHD finger protein, interacts with
404 WRKY transcription factors, forming complexes that bind to histones (via the PHD
405 finger) and repress the transcription of many stress-responsive genes (Du et al. 2023).
406 (*iii*) The MITOCHONDRIAL GLYCOPROTEIN FAMILY PROTEIN encoded by locus
407 *AT1G15870* binds 36.19% more strongly with VPg^{R118H} than with VPg^{wt}. Unfortunately,
408 no information about the function of this glycoprotein in the context of infection is
409 currently available. (*iv*) The MYOSIN 3 (MYOS3) protein encoded by locus
410 *AT3G58160*, a class XI myosin gene, is involved in the trafficking of Golgi stacks,
411 peroxisomes, and mitochondria in root hairs and leaf epidermal cells (Peremyslov et al.
412 2008). Class XI myosin genes have been shown to play important and diverse roles in
413 virus replication. For example, Amari et al. (2014) demonstrated that the inactivation of
414 myosin XI genes in *Nicotiana benthamiana* affected the functioning of the endoplasmic

415 reticulum, resulting in the aggregation of the movement protein of tobacco mosaic virus
416 and the altered intracellular distribution of the viral replicase. MYOS3 binds to VPg^{R118H}
417 27.51% more strongly than to the VPg^{wt} allele ($P = 0.007$).
418
419 **3.2. TuMV infection affects the expression of *RHD1* and *TGA1***
420 As outlined in the Introduction, VPg interacts with both the viral replicase NIb and the
421 host factor RHD1. Additionally, both NIb and RHD1 interact with the transcriptional
422 regulator TGA1, identified as a proviral factor by Martínez et al. (2023) for the first time.
423 To assess the impact of TuMV infection on the expression of these crucial genes, a time-
424 course analysis of *RHD1* and *TGA1* expression was conducted via relative RT-qPCR in
425 three pools of mock- and TuMV-VPg^{wt}-inoculated wild-type plants. Fig. 2 presents the
426 results of these experiments.
427 In the case of *RHD1*, overall effects were observed between infected and non-
428 infected plants (GLM: $\chi^2 = 28.568$, 1 d.f., $P < 0.001$, $\eta_P^2 = 0.814$) and across the
429 experimental time ($\chi^2 = 87.164$, 3 d.f., $P < 0.001$, $\eta_P^2 = 0.942$). More notably, a highly
430 significant interaction between these two factors was detected ($\chi^2 = 102.188$, 3 d.f., $P <$
431 0.001, $\eta_P^2 = 0.953$). This interaction is explained by gene expression in mock-inoculated
432 plants being consistently low and stable over time. In sharp contrast, expression in
433 infected plants was significantly lower (1.5- to 2.4-fold range; $P < 0.001$) at all time points
434 except at 6 dpi when *RHD1* expression surged 20-fold higher in infected plants ($P <$
435 0.001). Such pulse-like gene expression, is well suited for coordinating innate immune
436 antiviral responses, fine-tuning cellular processes without committing to prolonged
437 changes in gene activity (*e.g.*, Czerkies et al. 2018).
438 Concerning the effect of infection on *TGA1* expression, both infection status ($\chi^2 =$
439 100.553, 1 d.f., $P < 0.001$, $\eta_P^2 = 0.980$) and dpi ($\chi^2 = 39.247$, 3 d.f., $P < 0.001$, $\eta_P^2 = 0.953$)
440 had significant effects on *TGA1* expression levels. These effects are not independent (χ^2
441 $= 89.968$, 3 d.f., $P < 0.001$, $\eta_P^2 = 0.975$), as infected plants exhibit enhanced expression
442 over the time course. Maximum differences with non-infected plants were observed at
443 the two earlier time points (7.2- and 11.2-fold higher, respectively; $P < 0.001$), followed
444 by a reduction to non-significant differences at 9 dpi ($P = 0.570$) or a 1.9-fold higher
445 expression at the latest sampled time point ($P < 0.001$). This biphasic step-down
446 expression profile reflects a dynamic regulation of *TGA1* gene activity. The initial high
447 expression phase would be associated with an acute response to the initial infection, while

448 the reduction may represent a shift to a more sustained level of expression compatible
449 with TuMV overcoming defenses and stablishing a successful infection.
450 TGA1 regulates plant defense in a manner independent of NPR1 (Shearer et al.
451 2012). It accomplishes this by inducing the expression of key transcription factors, such
452 as *SYSTEMIC ACQUIRED RESISTANCE DEFICIENT 1* and *CALMODULIN-BINDING*
453 *PROTEIN 60g*. These transcription factors, in turn, target various essential regulators of
454 plant defense, including enzymes crucial for the synthesis of SA (Sun et al. 2018). Upon
455 a pathogen attack, SA levels surge, triggering two distinct responses: (*i*) the activation of
456 *NPR1*, leading to the expression of pathogenesis-related genes critical for reinforcing
457 systemic acquired resistance (SAR) against subsequent pathogen assaults (Zavaliev and
458 Dong 2023) and (*ii*) the enhancement of RNA-silencing antiviral defense (Alamillo,
459 Saénz and García 2006). Notably, despite the unaffected SAR in *tga1* plants, they exhibit
460 increased susceptibility to *Pseudomonas syringae* (Shearer et al. 2012). However, these
461 plants paradoxically display heightened resistance to TuMV infection (Martínez et al.
462 2023). This discrepancy raises intriguing questions about whether the activation of *TGA1*
463 during infection is a specific plant response or a consequence of a virus-induced
464 overexpression.
465
466 **3.3. *eif(iso)4e* resistance to TuMV infection is not affected by the three VPg alleles**
467 Sets of twenty plants of genotypes *eif(iso)4e* and wild-type were individually inoculated
468 with the three different VPg alleles. Daily records were kept for both the number of
469 infected plants and the severity of symptoms. A notable initial observation was that none
470 of the 20 *eif(iso)4e* plants exhibited any infection symptoms, irrespective of the VPg allele
471 carried by the inoculated TuMV. The possibility of this negative result being attributed
472 to failed inoculation trials was considered. To eliminate this potential explanation, an
473 additional 68 wild-type plants were inoculated with TuMV-VPg^{wt}. With a sample size of
474 88 wild-type plants, the probability of infecting a fully susceptible wild-type plant with
475 TuMV-VPg^{wt} was estimated based on the number of positive events (86) observed after
476 inoculating all the wild-type plants. Using the LaPlace point estimator and the adjusted
477 Wald 95% confidence interval method, the success rate was calculated to be 0.967 (0.916,
478 0.999). Therefore, even in the less favorable situation (*i.e.*, the lower limit of the
479 confidence interval), the likelihood of failing 20 independent inoculation events of fully
480 susceptible plants with a wild-type virus would be given by the Bernoulli process (1 –
481 $0.916)^{20} = 3.06 \cdot 10^{-22}$.

482 The same rationale was applied to the two mutant VPg alleles, where 20 out of 20
483 inoculated wild-type plants resulted in symptomatic infections. In this case, the LaPlace
484 estimation of success rate was 0.955 (0.859, 1.000), and the Bernoulli measure for the
485 less favorable situation was $1.007 \cdot 10^{-17}$. Consequently, we can confidently conclude that
486 *eif(iso)4e* plants are fully resistant to TuMV YC5 infection, regardless of the VPg allele
487 carried by the virus.
488 Many recessive resistance genes against potyviruses described in wild and crop
489 species encode members of the eIF(iso)4E and eIF(iso)4G families (Truninger and
490 Aranda 2009). Here, we have observed that TuMV YC5 replication strictly depends on
491 the interaction between VPg and a functional eIF(iso)4E. The complete resistance
492 observed for *eif(iso)4e* plants suggests its function cannot be complemented by other
493 translation initiation factors. By contrast, Navarro et al. (2022) were able of
494 experimentally evolving lineages derived from the YC5 isolate in *A. thaliana* mutants
495 *i4g1* (deficient for eIF(iso)4G1), observing different amino acid substitutions in VPg
496 associated with significant fitness increases. These results confirm that eIF(iso)4G is not
497 as essential as eIF(iso)4E for TuMV infection or that its function can be complemented
498 by other translation initiation factors.
499

500 **3.4. Disease progression in *rhd1* mutant plants depends on the VPg allele carried by**
501 **TuMV**

502 In stark contrast to *eif(iso)4e* plants, *rhd1* plants exhibited variable responses to infection
503 with the three TuMV variants. Firstly, concerning changes in infection rate (Fig. 3A),
504 the mean time to produce visible symptoms was the same for the three viral variants in
505 wild-type plants (log-rank test: $\chi^2 = 5.806$, 2 d.f., $P = 0.055$), averaging 6.817 ±0.151 dpi.
506 However, significant differences among VPg alleles were observed in *rhd1* plants ($\chi^2 =$
507 5.806, 2 d.f., $P = 0.002$): TuMV-VPg^{wt} produced visible symptoms at 7.000 ±0.218 dpi
508 (±1 SE), significantly later than TuMV-VPg^{D113G} (6.700 ±0.242; $\chi^2 = 3.881$, 1 d.f., $P =$
509 0.049) and TuMV-VPg^{R118H} (5.950 ±0.170; $\chi^2 = 16.028$, 1 d.f., $P < 0.001$), although the
510 difference between both mutants was not statistically significant ($\chi^2 = 3.839$, 1 d.f., $P =$
511 0.050). Therefore, we conclude that VPg^{D113G} and VPg^{R118H} accelerate the appearance of
512 disease symptoms in plants deficient for RHD1.
513 Secondly, Fig. 3B illustrates the disease progression curves for the three virus
514 variants in both wild-type and *rhd1* plants. In wild-type plants, large and significant inter-
515 individual differences were observed between the three VPg alleles ($F_{2,57} = 45.487$, $P <$

516 0.001, η_P^2 = 0.615), as well as substantial intra-individual differences between dpi
517 ($F_{3.933,224.203}$ = 502.504, P < 0.001, η_P^2 = 0.898) and the interaction between intra- and
518 inter-individual factors ($F_{7.867,224.203}$ = 13.586, P < 0.001, η_P^2 = 0.323). Overall, TuMV-
519 VPg^{wt} induced milder symptoms (0.879 ±0.043) compared to viruses expressing the two
520 mutant VPg variants (1.417 ±0.043 for VPg^{D113G} and 1.338 ±0.043 for VPg^{R118H}; pairwise
521 comparisons: P < 0.001). The progression of symptoms was similar for the three variants
522 during the first 7 dpi but diverged between VPg^{wt} and the two mutants thereafter. The
523 two mutant VPg variants were indistinguishable throughout the entire time course of the
524 experiment (P = 0.595). In the case of *rhd1* plants, inter-individual effects were large and
525 significant ($F_{2,57}$ = 8.322, P < 0.001, η_P^2 = 0.226), as were intra-individual effects of dpi
526 ($F_{2.790,159.046}$ = 293.746, P < 0.001, η_P^2 = 0.837), but not the interaction of the two factors
527 ($F_{5.581,159.046}$ = 2.090, P = 0.062, η_P^2 = 0.068). This suggests that differences among virus
528 variants remained constant over time. Overall, indistinguishable symptoms were induced
529 by TuMV-VPg^{wt} and TuMV-VPg^{D113G} (1.092 ±0.059 and 1.058 ±0.059, respectively; P
530 = 1.000), but both were significantly milder than those induced by TuMV-VPg^{R118H}
531 (1.371 ±0.059; $P \leq 0.005$).
532 Thirdly, we assessed the efficiency of TuMV carrying different VPg alleles in terms
533 of viral accumulation at 12 dpi. Fig. 3C illustrates the comparison between the two plant
534 genotypes and the three VPg alleles. The two orthogonal main factors had a significant
535 effect on viral load, although the magnitude of the effect was small among VPg alleles
536 (plant genotype: χ^2 = 226.733, 1 d.f., P < 0.001, η_P^2 = 0.459; VPg allele: χ^2 = 75.113, 2
537 d.f., P < 0.001, η_P^2 = 0.065). Nonetheless, the interaction between both factors was highly
538 significant and of large magnitude (χ^2 = 196.689, 2 d.f., P < 0.001, η_P^2 = 0.353), suggesting
539 that the effect on viral load of the VPg alleles was strongly dependent on the expression
540 of *RHD1*. As illustrated in Fig. 3C, viral load was reduced, on average, by 47.93% in
541 *rhd1* plants ($P \leq 0.018$ in the three pairwise comparisons). Focusing on wild-type plants,
542 TuMV-VPg^{R118H} showed the highest log-viral load (7.839 ±6.200; ±1 SE), followed by
543 TuMV-VPg^{wt} (7.754 ±6.122) and TuMV-VPg^{D113G} (7.563 ±5.992). However, in *rhd1*
544 plants, the situation reversed, and the highest log-viral load was observed for TuMV-
545 VPg^{D113G} (7.597 ±6.515), followed by TuMV-VPg^{wt} (7.329 ±5.824) and TuMV-VPg^{R118H}
546 (7.375 ±5.997), showing very similar accumulations ($P \leq 0.031$ in the three pairwise
547 comparisons).

548 At first glance, the observation of earlier and more severe symptoms in *rhd1* plants
549 suggests that the RHD1 protein plays an antiviral role in *A. thaliana*. However, higher
550 TuMV accumulation was noted in wild-type plants. To reconcile these two observations,
551 one could argue that symptoms may not necessarily correlate with viral accumulation if
552 *RHD1* enhances the plant's tolerance to viral infection (Pagán and García-Arenal 2020).
553 In such a scenario, wild-type plants expressing *RHD1* might exhibit higher viral
554 accumulation while maintaining mild symptoms. Conversely, if *RHD1* expression is
555 knocked out, even low viral accumulation could rapidly result in more severe symptoms.
556 Interestingly, the level of tolerance appears to be influenced by VPg alleles: viruses
557 carrying the wild-type and R118H alleles demonstrate larger reductions in virus
558 accumulation, whereas viruses carrying the generalist D113G allele are less affected (Fig.
559 3C).
560 Together with RMB1 and ROS1-ASSOCIATED WD40 DOMAIN-
561 CONTAINING PROTEIN (RWD40), RHD1 is part of the RWD40 complex, which
562 regulates active DNA demethylation by specifically directing ROS1 DNA demethylase
563 to specific target sequences in the *A. thaliana* genome (Liu et al. 2021). The absence of
564 expression of any of the components of the RWD40 complex results in ROS1
565 mislocalization, hypermethylation of targeted genomic DNA regions and
566 hypersusceptibility to the hemibiotrophic bacterial pathogen *Pseudomonas syringae*,
567 suggesting a role of the RWD40 DNA demethylation complex in pathogen defense (Liu
568 et al. 2021). One of these regions is localized in the *ROS1* promoter, leading to slightly
569 increased *ROS1* expression. Therefore, in *rhd1* plants, *ROS1* expression is upregulated
570 but its methylated target sequences are not recognized and become hypermethylated in
571 compared to wild-type plants. In addition, TuMV infection also downregulates the
572 expression of *ROS1* (Corrêa et al. 2020), in agreement with the results shown for wild-
573 type plants in Fig. 2 for most dpi. The sharp *RHD1* induction observed during TuMV
574 infection might be involved in temporally downregulating *ROS1* expression and the
575 subsequent temporal increase in the methylation rate of specific genome sequences,
576 thereby altering gene expression. Moreover, the pulse-like expression of *RHD1* and the
577 drop in expression in *TGA1* observed after 6 dpi (Fig. 2) coincides with the slow-down
578 in symptoms progression observed for TuMV-VPg^{wt}, which stay subsequently milder
579 (roughly speaking in the range 1 to 2). The two VPg mutants escape from this control,
580 resulting in worsened symptoms.
581

582 **3.5. Integrating interaction results with structural analyses**

583 To better understand these differences in interactions, we modeled the interactions of VPg
584 with eIF(iso)4E (Fig. 4) and RHD1 (Fig. 5). First, we modeled the TuMV-
585 VPg/eIF(iso)4E complex using a docking procedure guided by experimental data
586 (Supplementary Fig. S2). Several lines of evidence suggest that the proposed model is
587 close to the real structure: (*i*) the contact surface involves the central domain of VPg and
588 the cap-binding region of eIF(iso)4E, in agreement with experimental data; (*ii*) conserved
589 residue R103 in the central domain of VPg binds to eIF(iso)4E in a position equivalent to
590 the guanine base in the m^7G cap, (*iii*) acidic residues in the segment comprising the α1-
591 α2 loop bind to a region close to the phosphate binding region in eIF(iso)4E, (*iv*)
592 experimental data on residues involved int eIF4E - VPg interactions are mostly clustered
593 in the docking interaction surface (Grzela et al. 2006; Roudent-Tavert et al. 2007; Ashby
594 et al. 2012; Svanella-Dumas et al. 2014; Coutinho de Oliveira et al. 2019). Our model
595 also provides a plausible explanation of the effect of the R118H and D113G substitutions
596 on the eIF(iso)4E - VPg interaction (Fig 4). R118 and D113 are located in a small helical
597 segment predicted to bind to a variable region of complementary charge in eIF(iso)4E in
598 the m^7G cap binding pocket (Fig. 4). Both mutants are predicted to have a significant
599 effect on the net charge with respect to the VPg^{wt} (pI ≈ 7.23), with D113G making the
600 protein more positive (pI ≈ 8.02) and R118H more negative (pI ≈ 6.74). Therefore, we
601 postulate that these pI changes in the R118H and D113G mutants remove favorable ionic
602 interactions, resulting in the observed reduction in the stability of the eIF(iso)4E - VPg
603 complex.

604 Our model also explains the resistance of *eif(iso)4e* plants against the three TuMV
605 mutants. Potyviruses selectively require different eIF4E paralogs to establish infection
606 (Duprat et al. 2002; Sato et al. 2005; Nicaise et al. 2007; Gómez et al. 2019). In *A.*
607 *thaliana*, lettuce mosaic virus (LMV), TEV and TuMV use eIF(iso)4E, while other
608 potyviruses, such as clover yellow vein virus and PVY, use eIF4E (Sato et al. 2005;
609 Zafirov et al. 2023). Structurally, one of the most salient differences between eIF(iso)4E
610 and its paralogs eIF4E and nCBP is the presence of a bulkier segment in the VPg binding
611 region. We propose that due to this steric effect, neither eIF4E nor nCBP can substitute
612 for eIF(iso)4E in the *eif(iso)4e* plants (Fig. 4).

613 We also used docking to identify a potential VPg binding pocket in the RHD1
614 protein from *A. thaliana* (Fig. 5). A prediction of structural domains revealed that this

615 protein contains an N-proximal Zn finger domain, a C-terminal homeobox domain and
616 two unstructured segments comprising the N-terminal and central regions (Fig. 5A).
617 Assuming that binding occurs in the homeobox regions, we performed a global docking
618 of VPg against the homeobox domains using a refined AlphaFold model of RHD1 (Fig.
619 5B). The lowest energy complex fitted the VPg core domain within a crevice separating
620 the helical and beta-stranded regions of the ZF-HD homeobox. This region comprises
621 domains of positive and negative charge that might be affected by the ionization state of
622 the VPg protein.
623 Fig. 5C shows the RHD1 PPIs network generated with STRING. Interestingly,
624 RHD1 interactors mostly belong to the ZF-HD-containing family protein of transcription
625 factors. The direct interactor PATHOGENESIS RELATED HOMEODOMAIN
626 PROTEIN A (PRHA) is a hub that bridges ZF-HDs involved in the regulation of different
627 aspects of plant development and of circadian clock, with MVA3.30, a disease resistance
628 protein of the TIR-NBS-LRR class family, and with MORPH RELATED GENE 1
629 (MRG1) and 2 (MRG2), both readers of H3K4m3/H3K36m3 methylation marks. Taken
630 together, these interactions suggest that VPg affects epigenetic regulations by directly
631 interacting with RHD1 and indirectly affecting its interaction with other transcription
632 factors and with components of the histone methylation machinery. The role of such
633 epigenetic pathways in TuMV infection has been already stablished (Corrêa et al. 2020;
634 Navarro et al. 2022) and might contribute to the enhanced tolerance to infection of plants
635 accumulating RHD1.
636
637

638 **4. Concluding remarks**

639 The results shown above shed light on the complex interplay between TuMV VPg
640 variants, host proteins, and plant defense response, providing valuable insights into the
641 molecular mechanisms underlying virus-host interactions and disease progression in *A.*
642 *thaliana*. The three different VPg alleles used in our study show differential interactions
643 with plant proteins. Infected plants show altered expression of key genes *eIF(iso)4E*,
644 *RHD1* and *TGA1*. Mutant plants *eif(iso)4e* show complete resistance to TuMV infection
645 regardless of the VPg allele, confirming its well described essential role in virus
646 replication. Finally, different VPg alleles affect disease progression in *rhd1* mutant
647 plants. VPg^{D113G} and VPg^{R118H} accelerate symptom appearance, while VPg^{wt} induces

648 milder symptoms. This observation suggests that RHD1 modulates the tolerance to viral
649 infection, likely via epigenetic modifications.
650
651
652 **Data availability**
653 All raw data are available at https://git.csic.es/SFElena/tumv-evolution-in-a.-thaliana-
654 genotypes-with-deficient-immune-responses.
655
656
657 **Supplementary data**
658 Supplementary data is available at *Virus Evolution* online.
659
660
661 **Acknowledgments**
662 We thank Francisca de la Iglesia and Paula Agudo for excellent technical assistance and
663 the rest of the EvolSysVir lab members for fruitful discussions.
664
665
666 **Funding**
667 This work was supported by grants PID2022-136912NB-I00 funded by
668 MCIU/AEI/10.13039/501100011033 and by "ERDF a way of making Europe", and
669 CIPROM/2022/59 funded by Generalitat Valenciana. to S.F.E.
670
671
672 **References**
673 Agudelo-Romero, P. et al. (2008) 'Virus adaptation by manipulation of host's gene
674 expression', *PLoS ONE*, 3: e2397.
675 Aguirre, J., and Guantes R. (2023) 'Virus-host protein co-expression networks reveal
676 temporal organization and strategies of viral infection', *iScience*, 26: 108475.
677 Ahmed, H. et al. (2018) 'Network biology discovers pathogen contact points in host
678 protein-protein interactomes', *Nature Communications*, 9: 2312.

679 Alamillo, J.M., Saénz, P., and García, J.A. (2006) 'Salicylic acid-mediated and RNA-
680 silencing defense mechanisms cooperate in the restriction of systemic spread of plum
681 pox virus in tobacco', *Plant Journal*, 48: 217-27.
682 Amari, K. et al. (2014) 'Myosins VIII and XI play distinct roles in reproduction and
683 transport of *Tobacco mosaic virus*', *PLoS Pathogens*, 10: e1004448.
684 Ambrós, S. et al. (2024) 'Phenotypic and genomic changes during turnip mosaic virus
685 adaptation to *Arabidopsis thaliana* mutants lacking epigenetic regulatory factors',
686 *Evolution*, 28: 69-85.
687 Ashby, J. A. et al. (2011) 'Structure-based mutational analysis of eIF4E in relation to
688 *sbm1* resistance to pea seed-borne mosaic virus in pea', *PLoS ONE*, 6: e15873.
689 Ayme, V. et al. (2006) 'Different mutations in the genome-linked protein VPg of potato
690 virus Y confer virulence on the $pvr2^3$ resistance in pepper', *Molecular Plant-Microbe*
691 *Interactions*, 19: 557-63.
692 Belshaw, R., Pybus, O. G., and Rambaut, A. (2007) 'The evolution of genome
693 compression and genomic novelty in RNA viruses', *Genome Research*, 17: 1496-
694 504.
695 Bosque, G. et al. (2014) 'Topology analysis and visualization of *Potyvirus* protein-protein
696 interaction network', *BMC Systems Biology*, 8: 129.
697 Boyes, D. C. et al. (2001) 'Growth stage-based phenotypic analysis of Arabidopsis: a
698 model for high throughput functional genomics in plants', *Plant Cell*, 13: 1499-510.
699 Buschmann, H. et al. (2004) 'Helical growth of the *Arabidopsis* mutant *tortifolia1* reveals
700 a plant-specific microtubule-associated protein', *Current Biology*, 14: 1515-21.
701 Butković, A. et al. (2021) 'A genome-wide association study identifies *Arabidopsis*
702 *thaliana* genes that contribute to differences in the outcome of infection with two
703 turnip mosaic potyvirus strains that differ in their evolutionary history and degree of
704 host specialization', *Virus Evolution*, 7: veab063.
705 Cervera, H. et al. (2018) 'Viral fitness correlates with the magnitude and direction of the
706 perturbation induced in the host's transcriptome: the tobacco etch potyvirus-tobacco
707 case study', *Molecular Biology and Evolution*, 35: 1599-615.
708 Chastain, C. J. et al. (2007) 'The pyruvate, orthophosphate dikinase regulatory proteins
709 of Arabidopsis possess a novel, unprecedented Ser/Thr protein kinase primary
710 structure', *Plant Journal*, 53: 854-63.
711 Chaudhury, S. et al. (2011) 'Benchmarking and analysis of protein docking performance
712 in Rosetta v3.2'. *PLoS ONE*, 6: e22477.

713 Chen, C. C. et al. (2003) 'Identification of turnip mosaic virus isolates causing yellow
714 stripe and spot on calla lily', *Plant Disease*, 87: 901-5.
715 Conway, P. et al. (2014) 'Relaxation of backbone bond geometry improves protein energy
716 landscape modeling', *Protein Science*, 23: 47-55.
717 Corrêa, R. L. et al. (2020) 'Viral fitness determines the magnitude of transcriptomic and
718 epigenomic reprogramming of defense responses in plants', *Molecular Biology and*
719 *Evolution*, 37: 1866-81.
720 Coutinho de Oliveira, L. et al. (2019) 'Structural studies of the eIF4E-VPg complex reveal
721 a direct competition for capped RNA: implications for translation', *Proceedings of*
722 *the National Academy of Sciences of the USA*, 116: 24056-65.
723 Czechowski, T. et al. (2005) 'Genome-wide identification and testing of superior
724 reference genes for transcript normalization in Arabidopsis', *Plant Physiology*, 139:
725 5-17.
726 Czerkies, M. et al. (2018) 'Cell fate in antiviral response arises in the crosstalk of IRF,
727 NF-κB and JAK/STAT pathways', *Nature Communications*, 9: 493.
728 De Chassey, B. et al. (2008) 'Hepatitis C virus infection protein network', *Molecular*
729 *Systems Biology*, 4: 230.
730 Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) 'HADDOCK: a protein-protein
731 docking approach based on biochemical or biophysical information', *Journal of the*
732 *American Chemical Society*, 125: 1731-7.
733 Du, P. et al. (2023) 'WRKY transcription factors and OBERON histone-binding proteins
734 form complexes to balance plant growth and stress tolerance', *EMBO Journal*, 42:
735 e113639.
736 Duprat, A. et al. (2002) 'The Arabidopsis eukaryotic initiation factor (iso)4e is
737 dispensable for plant growth but required for susceptibility to potyviruses', *Plant*
738 *Journal*, 32: 927-34.
739 Dyer, M. D., Murali, T. M., and Sobral, B. W. (2008) 'The landscape of human proteins
740 interacting with viruses and other pathogens', *PLoS Pathogens*, 4: e32.
741 Edgar, R. C. (2004) 'MUSCLE: multiple sequence alignment with high accuracy and
742 high throughput', *Nucleic Acids Research*, 32: 1792-7.
743 Eskelin, K. et al. (2011) 'Potyviral VPg enhances viral RNA translation and inhibits
744 reporter mRNA translation *in planta*', *Journal of Virology*, 85: 9210-21.
745 Gallois, J. L. et al. (2010) 'Single amino acid changes in the turnip mosaic virus viral
746 genome-linked protein (VPg) confer virulence towards *Arabidopsis thaliana* mutants

747 knocked out for eukaryotic initiation factors eIF(iso)4E and eIF(iso)4G', *Journal of*
748 *General Virology*, 91: 288-93.
749 German-Retana, S. et al. (2008) 'Mutational analysis of plant cap-binding protein eIF4E
750 reveals key amino acids involved in biochemical functions and potyvirus infection',
751 *Journal of Virology*, 82: 7601-12.
752 Gilkerson, J. et al. (2012) 'The plastid-localized pfkB-type carbohydrate kinases
753 FRUCTOKINASE-LIKE 1 and 2 are essential for growth and development of
754 *Arabidopsis thaliana*', *BMC Plant Biology*, 12: 102.
755 Gómez, M. A. et al. (2019) 'Simultaneous CRISPR/Cas9-mediated editing of cassava
756 eIR4E isoforms nCBP-1 and nCBP-2 reduce cassava brown streak disease symptom
757 severity and incidence', *Plant Biotechnology Journal*, 17: 421-34.
758 González, R., Butković, A., and Elena, S. F. (2019) 'Role of host genetic diversity for
759 susceptibility-to-infection in the evolution of virulence of a plant virus', *Virus*
760 *Evolution*, 5: vez024.
761 González, R. et al. (2021) 'Plant virus evolution under strong drought conditions results
762 in a transition from parasitism to mutualism', *Proceedings of the National Academy*
763 *of Sciences of the USA*, 118: e2020990118.
764 Grzela, R. et al. (2006) 'Potyvirus terminal protein VPg, effector of host eukaryotic
765 initiation factor eIF4E', *Biochimie*, 88: 887-96.
766 Hafrén, A., Löhmus, A., and Mäkinen, K. (2015) 'Formation of potato virus A-induced
767 RNA granules and viral translation are interrelated processes required for optimal
768 virus accumulation', *PLoS Pathogens*, 11: e1005314.
769 Hillung, J. et al. (2014) 'Experimental evolution of an emerging plant virus in host
770 genotypes that differ in their susceptibility to infection', *Evolution*, 68: 2467-80.
771 Honorato, R. V. et al. (2021) 'Structural biology in the clouds: the WeNMR-EOSC
772 ecosystem', *Frontiers in Molecular Biosciences*, 8: 729513.
773 Jumper, J. et al. (2021) 'Highly accurate protein structure prediction with AlphaFold',
774 *Nature*, 596: 583-9.
775 Liu, P. et al. (2021) 'A novel protein complex that regulates active DNA demethylation
776 in Arabidopsis', *Journal of Integrative Plant Physiology*, 63, 772-86.
777 Mahmoudabadi, G, and Phillips, R. (2018) 'A comprehensive and quantitative
778 exploration of thousands of viral genomes', *eLife*, 7: e31955.
779 Mandell, D. J. et al. (2009) 'Sub-angstrom accuracy in protein loop reconstruction by
780 robotics-inspired conformational sampling', *Nature Methods*, 6: 551-2.

781 Martínez, F. et al. (2016) 'Interaction network of tobacco etch potyvirus NIa protein with
782 the host proteome during infection', *BMC Genomics*, 17: 87.
783 Martínez, F. et al. (2023) 'A binary interaction map between turnip mosaic virus and
784 *Arabidoposis thaliana* proteomes', *Communications Biology*, 6: 28.
785 Melero, I., González, R. and Elena, S.F. (2023) 'Host developmental stages shape the
786 evolution of plant RNA virus', *Philosophical Transactions of the Royal Society B:*
787 *Biological Sciences*, 378: 20220005.
788 Menche, J. et al. (2015) 'Uncovering disease-disease relationships through the incomplete
789 interactome', *Science*, 347: 841.
790 Miras, M. et al. (2017) 'Structure of eIF4E in complex with an eIF4G peptide supports a
791 universal bipartite binding mode for protein translation', *Plant Physiology*, 174:
792 1476-91.
793 Monzingo, A. F. et al. (2007) 'The structure of eukaryotic translation initiation factor-4E
794 from wheat reveals a novel disulfide bond', *Plant Physiology*, 143: 1504-18.
795 Moury, B. et al. (2004) 'Mutations in potato virus Y genome-linked protein determine
796 virulence towards recessive resistances in *Capsicum annuum* and *Lycopersicon*
797 *hirsutum*', *Molecular Plant-Microbe Interactions*, 17: 322-9.
798 Mukhtar, M. S. et al. (2011) 'Independently evolved virulence effectors converge onto
799 hubs in a plant immune system network', *Science*, 333: 596-601.
800 Navarro, R. et al. (2022) 'Defects in plant immunity modulate the rates and patterns of
801 RNA virus evolution', *Virus Evolution*, 8: veac059.
802 Nicaise, V. et al. (2007) 'Coordinated and selective recruitment of eIF4E and eIF4G
803 factors for potyvirus infection in *Arabidopsis thaliana*', *FEBS Letters*, 581: 1041-6.
804 Pagán, I., and García-Arenal, F. (2020) 'Tolerance of plants to pathogens: an unifying
805 view', *Annual Review of Phytopathology*, **58**: 77-96.
806 Peremyslov, V. V. et al. (2008) 'Two class XI myosins function in organelle trafficking
807 and root hair development in Arabidopsis', *Plant Physiology*, 146: 1109-16.
808 Pérez, K. et al. (2012) 'Tobacco etch virus infectivity in *Capsicum* spp. is determined by
809 a maximum of three amino acids in the viral virulence determinant VPg', *Molecular*
810 *Plant Microbe Interactions*, 25: 1562-73.
811 Phillips, J. C. et al. (2020). 'Scalable molecular dynamics on CPU and GPU architectures
812 with NAMD', *Journal of Chemical Physics,* 153: 044130.
813 Pichlmair, A. et al. (2012) 'Viral immune modulators perturb the human molecular
814 network by common and unique strategies', *Nature*, 487: 486-90.

815 Roudet-Tavert, G. et al. (2007) 'Central domain of a potyvirus VPg is involved in the
816 interaction with the host translation initiation factor eIF4E and the viral protein
817 HCPro', *Journal of General Virology*, 88: 1029-33.
818 Rydzik, A. M. et al. (2017) 'mRNA cap analogues substituted in the tetraphosphate chain
819 with CX2: identification of O-to-CCl2 as the first bridging modification that confers
820 resistance to decapping without impairing translation'. *Nucleic Acids Research*, 45:
821 8661-75.
822 Sato, M. et al. (2005) 'Selective involvement of members of the eukaryotic initiation
823 factor 4E family in the infection of *Arabidopsis thaliana* by potyviruses', *FEBS*
824 *Letters*, 579: 1167-71.
825 Shearer, H.L. et al. (2012) 'Arabidopsis clade I TGA transcription factors regulate plant
826 defenses in an NPR1-independent fashion', *Molecular Plant Microbe Interactions*,
827 25: 1459-68.
828 Spivak, M. et al. (2023) 'VMD as a platform for interactive small molecule preparation
829 and visualization in quantum and classical simulations', *Journal of Chemical*
830 *Information and Modeling*, 63: 4664-78.
831 Sun, T. et al. (2018) 'TGACG-BINDING FACTOR 1 (TGA1) and TGA4 regulate
832 salicylic acid and pipecolic acid biosynthesis by modulating the expression of
833 *SYSTEMIC ACQUIRED RESISTANCE DEFICIENT 1* (*SARD1*) and
834 *CALMODULIN-BINDING PROTEIN 60g* (*CBP50g*)', *New Phytologist*, 217: 344-
835 54.
836 Svanella-Dumas, L. et al. (2014) 'Adaptation of lettuce mosaic virus to *Catharanthus*
837 *roseus* involves mutations in the central domain of the VPg', *Molecular Plant*
838 *Microbe Interactions*, 27: 491-7.
839 Szklarczyk, D. et al. (2023) 'The STRING database in 2023: protein-protein association
840 networks and functional enrichment analyses for any sequenced genome of interest',
841 *Nucleic Acids Research*, 50: gkac1000.
842 Tarazona, A., Forment, J., and Elena, S. F. (2020) 'Identifying early warning signals for
843 the sudden transition from mild to severe tobacco etch disease by dynamical network
844 biomarkers', *Viruses*, 12: 16.
845 Tisoncik-Go, J. et al. (2016) 'Integrated omics analysis of pathogenic host responses
846 during pandemic H1N1 influenza virus infection: the crucial role of lipid
847 metabolism', *Cell Host Microbe*, 19: 254-66.

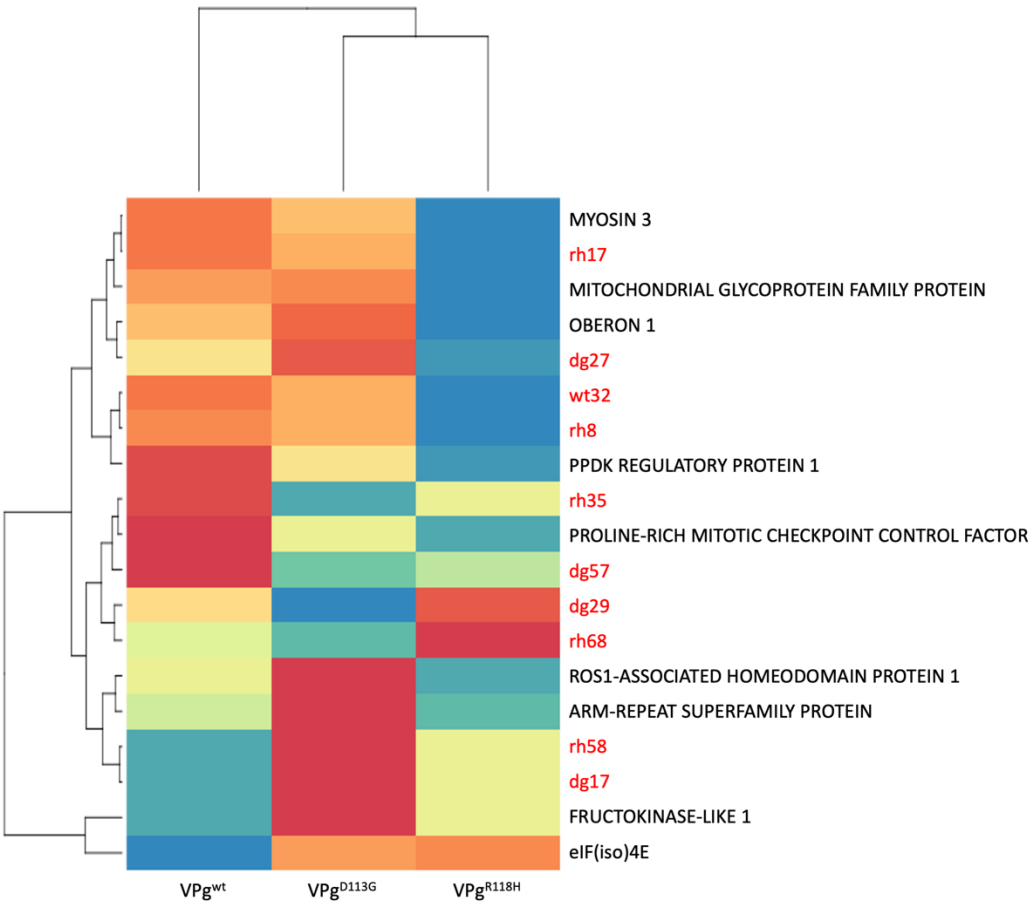
848 Truninger, V., and Aranda M.A. (2009) 'Recessive resistance to plant viruses', *Advances*
849 *in Virus Research*, 75: 119-59.
850 Uetz, P. et al. (2006) 'Herpesviral protein networks and their interaction with the human
851 proteome', *Science*, 311: 239-42.
852 Van Zundert, G. C. P. et al. (2016) 'The HADDOCK2.2 web server: user-friendly
853 integrative modeling of biomolecular complexes', *Journal of Molecular Biology*,
854 428: 720-5.
855 Wan, X. et al. (2020) 'Discovery of lysine-targeted eIF4E inhibitors through covalent
856 docking'. *Journal of the American Chemical Society*, 142: 4960-4.
857 Wang, C., Bradley, P., and Baker, D. (2007). 'Protein-protein docking with backbone
858 flexibility', *Journal of Molecular Biology*, 373: 503-19.
859 Webb, B., and Sali, A. (2016) 'Comparative protein structure modeling using
860 MODELLER'. *Current Protocols in Bioinformatics*, 54: 5.6.1-37.
861 Wojcik, R. et al. (2021) 'Novel N7-arylmethyl substituted dinucleotide mrna 5' cap
862 analogs: synthesis and evaluation as modulators of translation', *Pharmaceutics*, 13:
863 1941.
864 Zafirov, D. et al. (2023) 'Arabidopsis *eIF4E1* protects the translational machinery during
865 TuMV infection and restricts virus accumulation', *PLoS Pathogens*, 19: e1011417.
866 Zavaliev, R., and Dong, X. (2023) 'NPR1, a key immune regulator of plant survival under
867 biotic and abiotic stresses', *Molecular Cell*, 84: 131-41.
868

869 **Figures**

870

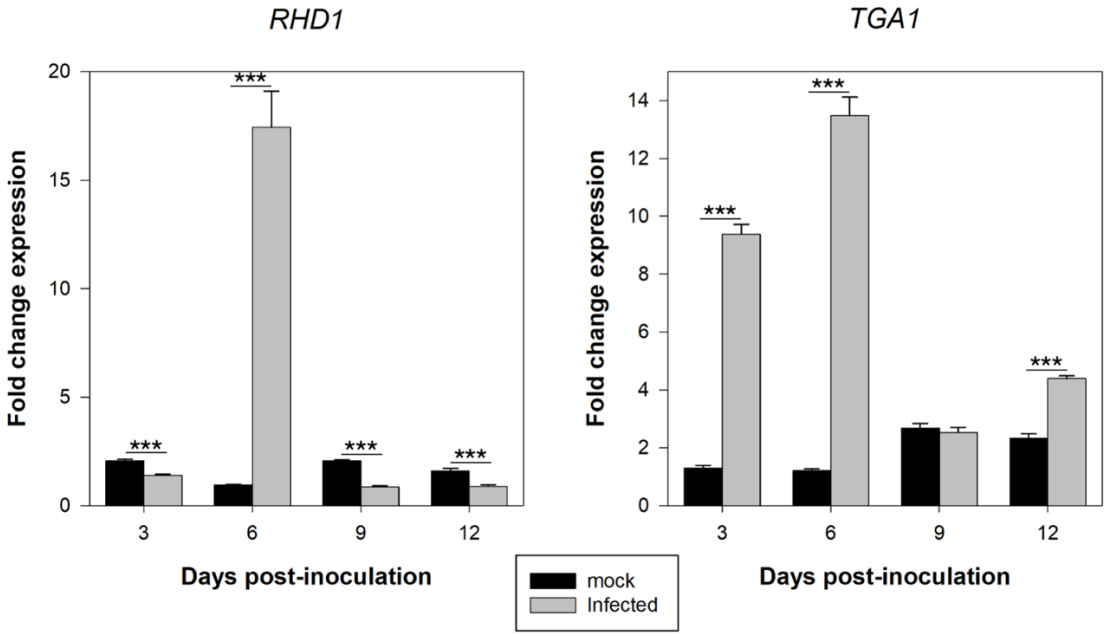
871 **Figure 1.** Heatmap constructed from the estimated binding affinities between the three
872 different VPg alleles (wild-type, D113G and R118H) and the 19 host proteins identified
873 in the HT-Y2H screening. Gene names are only provided for interactors for which
874 statistically significant differences with VPg^{wt} were found and further identified by
875 sequencing. Interactors for which no significant differences in binding affinity were
876 observed are indicated in red. Labels starting with "dg" correspond to those found in the
877 screening against VPg^{D113G}, "rh" to those found in the screening against VPg^{R118H} and
878 "wt" against the VPg^{wt} allele. Color scale: the more intense blue corresponds to stronger
879 affinities than for VPg^{wt}, whereas stronger red corresponds to weaker affinities than for
880 VPg^{wt}.

881



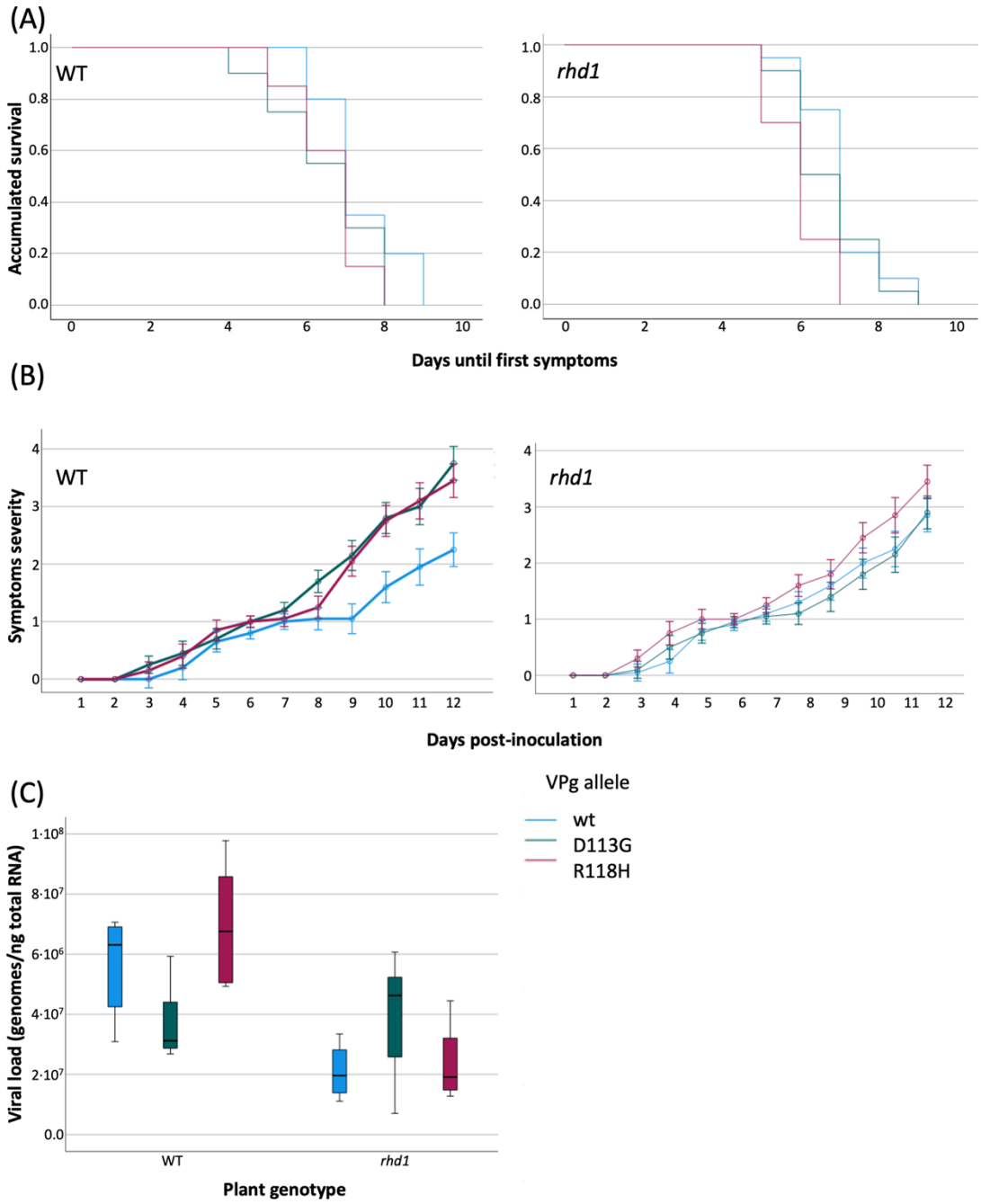
882

883 **Figure 2.** Effect of TuMV VPg^{wt} infection upon expression of VPg direct (*RHD1*) and
884 indirect (*TGA1*) interactors. Significance levels (sequential Bonferroni *post hoc* tests: ***
885 $P < 0.001$.
886



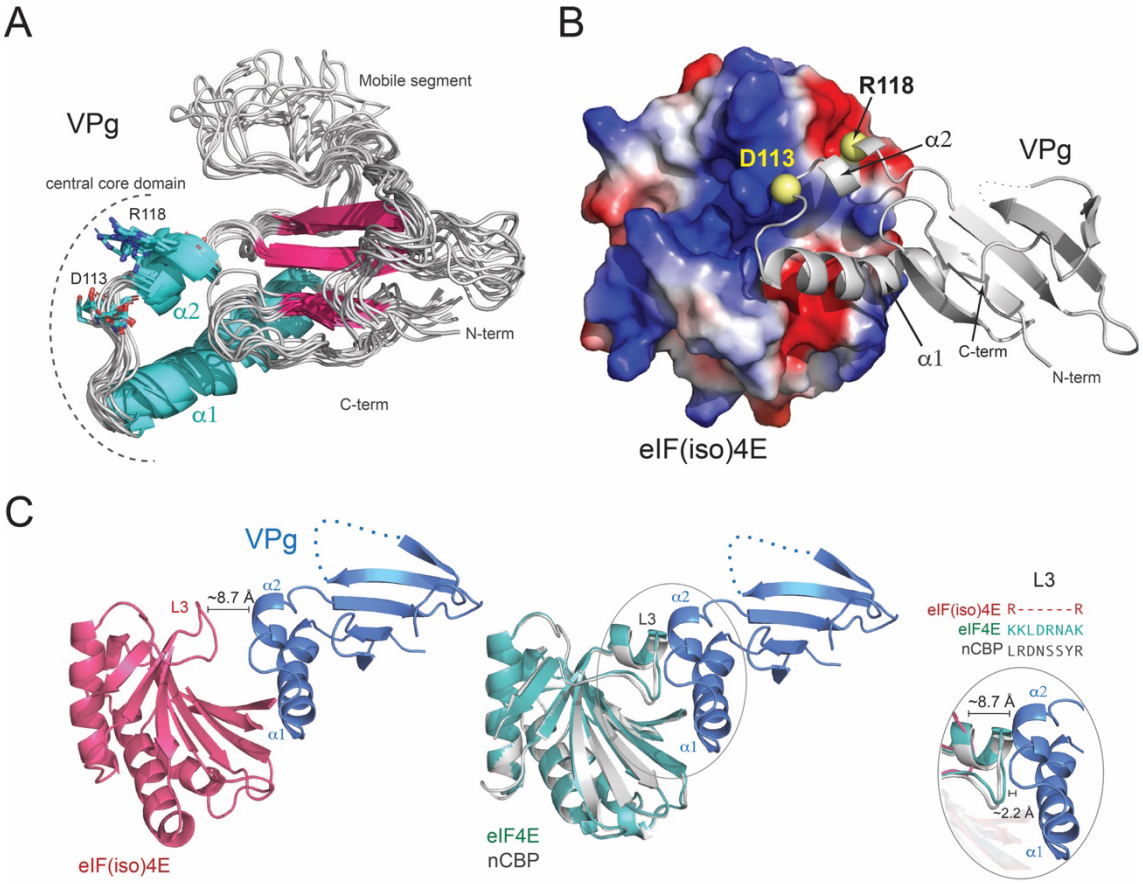
887
888

889 **Figure 3.** (A) Survival curves representing the number of non-symptomatic plants along
890 time. (B) Progression of disease severity in wild-type and *rhd1* plants infected with
891 TuMV expressing the three different VPg alleles. Error bars represent 95% CI. (C)
892 Differences in viral load. Boxes represent interquartile ranks, horizontal lines the median
893 and error bars represent 95% CI. WT: wild-type plants.
894



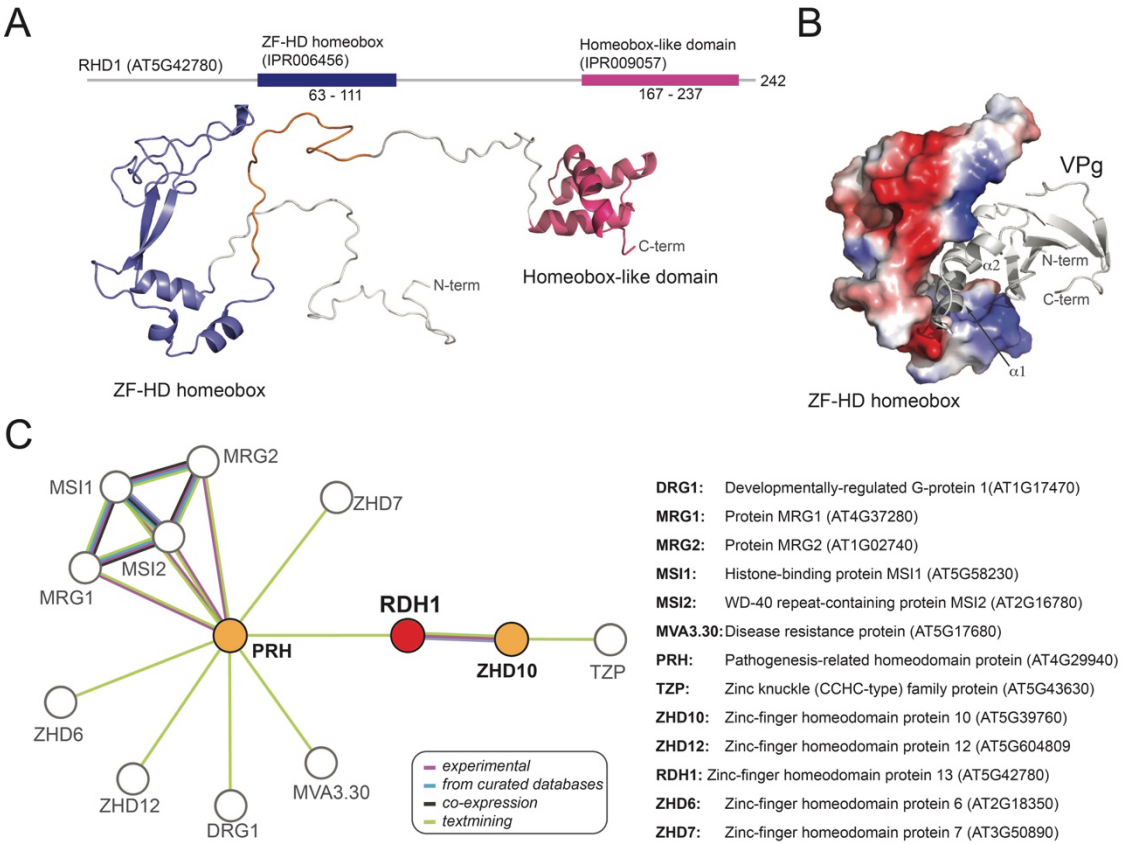
895
896

897 **Figure 4.** Structural analysis of the *A. thaliana* eIF(iso)4E - TuMV VPg complex. (A)
898 The model of TuMV VPg reveals that D113 and R118 project outwards in a segment
899 comprising the α1-α2 loop and helix α2. The structural ensemble illustrates the
900 dynamical properties of the VPg proteins as determined by molecular dynamics
901 simulations. (B) Analysis of the eIF(iso)4E - VPg docking model suggests that residues
902 D113 and R118 bind to pockets of complementary charges in eIF(iso)4E. The reduced
903 affinity of the D113G and R118H variants probably results from the removal of these
904 ionic interactions. (C) The eIF(iso)4E - VPg (left) docking model explains the resistance
905 in *eif(iso)4e* plants as the paralogous protein eIF4E and nCBP have a bulkier VPg binding
906 region (L3) that is likely to impede binding to TuMV VPg (center). The sequence of the
907 L3 region of the eIF4E homologs in *A. thaliana* and their relative distance to VPg is
908 illustrated in the right panel.
909

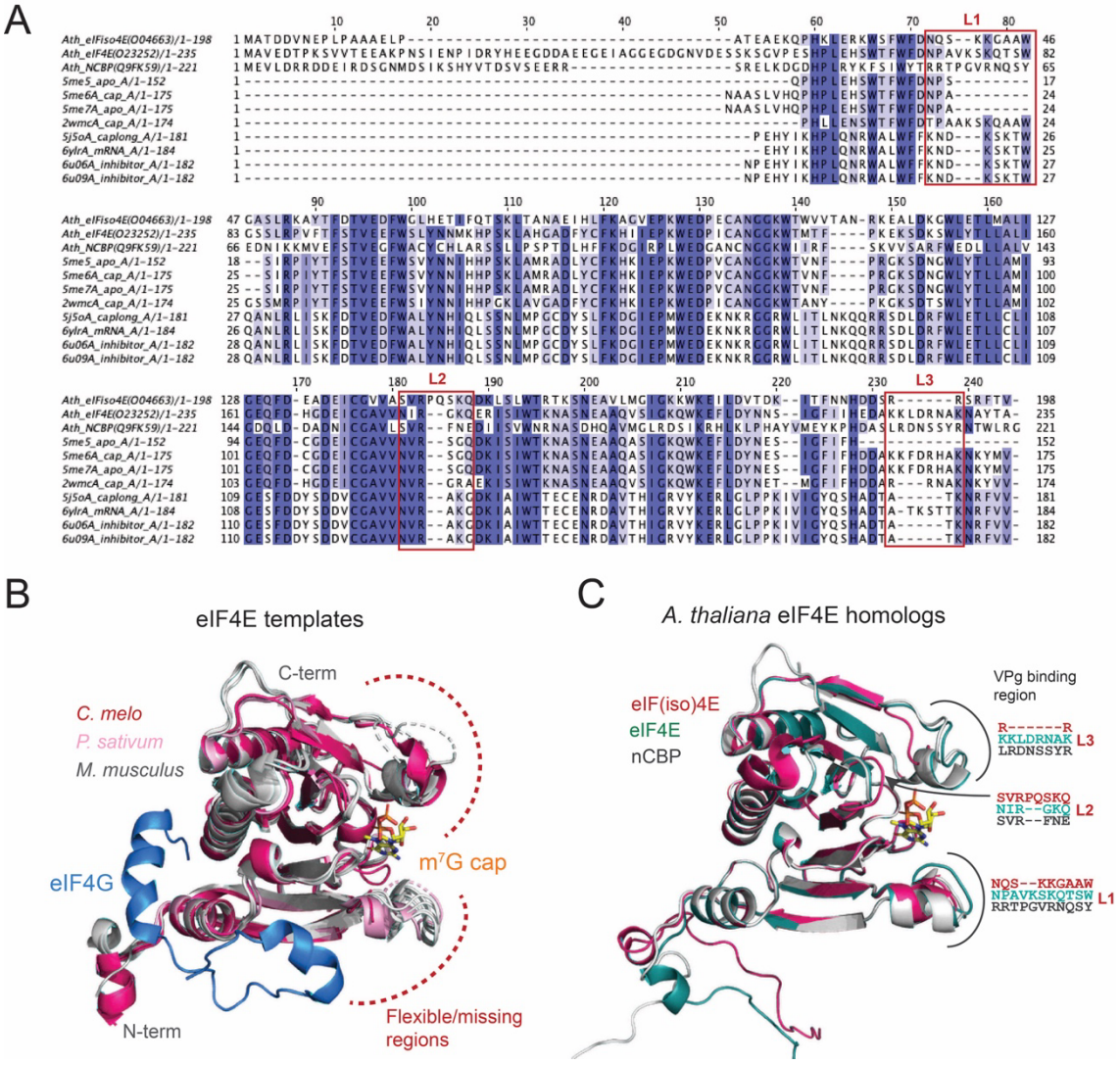


910
911

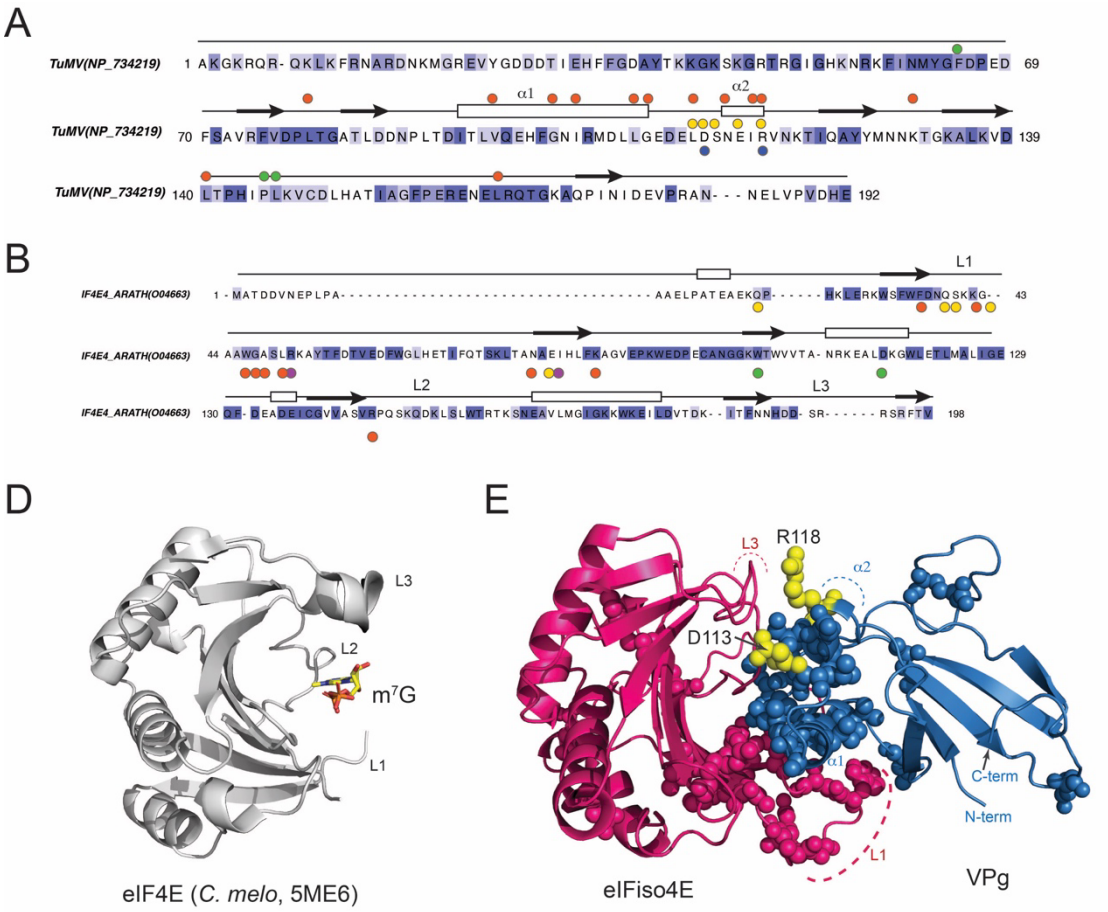
Figure 5. Structural properties of the RHD1 protein from *A. thaliana*. (A) The AlphaFold modeling and sequences analysis reveal that this protein contains Zn-finger (ZF-HD) and a C-terminal homeobox domain, and unstructured N-terminal and central regions. (B) Global docking of VPg against the homeobox domains of RHD1using ROSETTA revealed a potential binding site comprising the central domain of VPg and a crevice in the ZF-HD homeobox region. (C) PPI network of the RHD1 with other *A. thaliana* proteins according to STRING database.



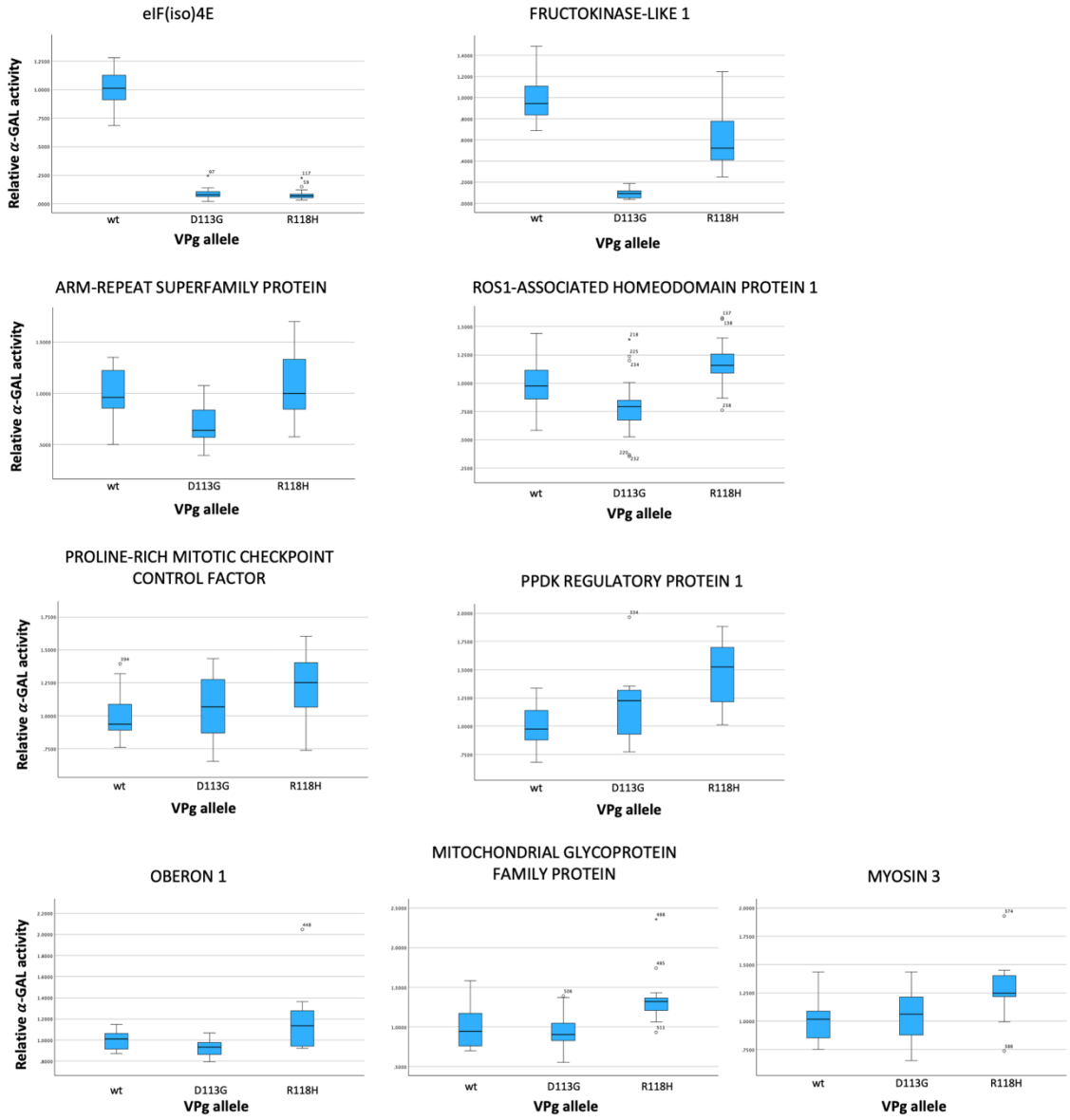
Supplementary Figure S1. Sequence and structural analysis of the eIF(iso)4E structure from *A. thaliana*. (A) Sequence alignment of *A. thaliana* eIF(iso)4E homologs and the eIF4E sequences used as structural templates. Variable loop regions predicted to be involved in the interaction with VPg are enclosed in red boxes. (B) Superposition of the structural templates used in this work showing position of the m^{7}G and eIF4G binding regions. (C) Superposition of the eIF(iso)4E, eIF4E and nCBP model from *A. thaliana* highlighting sequence and structural differences in the VPg-binding loops.



Supplementary Figure S2. Constraints used in the construction of the eIF(iso)4E - TuMV VPg model. (A) Identification of potential eIF(iso)4E interacting residues in the VPg based on a multiple sequence alignment of VPg protein from LMV, PVY, TEV, and TuMV. Filled circles indicate regions shown to be involved in the interaction with eIF4E/eIF(iso)4E: Coutinho de Olivera et al. (2019) (orange), Roudet-Tavert et al. (2007) (purple), Pérez et al. (2012) (yellow), Svanella-Dumas (2014) (green), and this study (blue). (B) dentification of potential VPg interacting residues in the eIF(iso)4E from *A. thaliana*. In this case, filled circles correspond to interaction data from Coutinho de Olivera *et al*. (2019) (orange), German-Retana et al. (2008) (purple), Pérez et al. (2012) (yellow), Ashby et al. (2011) (green), and Monzingo et al. (2007) (blue). (C) Structure of the eIF4E from *C. melo* in complex with m^7G, illustrating the position of the putative VPg-binging loops. (D) Cartoon representation of the predicted eIF(iso)4E - VPg complex derived from the interaction data presented in panels A and B. Interacting residues are shown as spheres. Residues D113 and R118 are shown in yellow.



949 **Supplementary Figure S3**. Distribution of affinity binding values estimated for the nine
950 host interactors for which a significant difference between the mutant VPg alleles and the
951 VPg^{wt} one has been observed.
952



953