



HAL
open science

Constraints on evolution of virus avirulence factors predict the durability of corresponding plant resistances

Bérenger Janzac, Frederic Fabre, Alain Palloix, Benoît Moury

► To cite this version:

Bérenger Janzac, Frederic Fabre, Alain Palloix, Benoît Moury. Constraints on evolution of virus avirulence factors predict the durability of corresponding plant resistances. *Molecular Plant Pathology*, 2009, 10 (5), pp.599-610. 10.1111/j.1364-3703.2009.00554.x . hal-02666562

HAL Id: hal-02666562

<https://hal.inrae.fr/hal-02666562v1>

Submitted on 31 May 2020

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

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



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

A new predictor of plant resistance durability

1 **Constraints on evolution of virus avirulence factors predict the durability of** 2 **corresponding plant resistances.**

3 Bérenger JANZAC^{1,2}, Frédéric FABRE¹, Alain PALLOIX² and Benoît MOURY¹.

4 ¹*INRA, UR 407 Pathologie Végétale, F- 84140 Montfavet, France.*

5 ²*INRA, UR 1052 Génétique et Amélioration des Fruits et Légumes, F- 84140 Montfavet, France.*

6 7 **Summary**

8 Understanding the factors driving pathogen emergence and re-emergence is a major challenge,
9 particularly in agriculture where the use of resistant plant cultivars imposes strong selective
10 pressures on plant pathogen populations and leads frequently to “resistance breakdown”.
11 Presently, durable resistances are only identified after a long period of large-scale cultivation
12 of resistant cultivars. We propose a new predictor of the durability of plant resistance.
13 Because resistance breakdown involves modifications in the avirulence factors of pathogens,
14 we tested for correlations between the evolutionary constraint acting on avirulence factors or
15 their diversity, and the durability of the corresponding resistance genes in the case of plant-
16 virus interactions. An analysis performed on 20 virus species-resistance gene combinations
17 revealed that the selective constraint applied on amino acid substitutions in virus avirulence
18 factors correlates with the observed durability of the corresponding resistance genes. Based
19 on this result, a model predicting the potential durability of resistance genes as a function of
20 the selective constraint applied on the corresponding avirulence factors is proposed to help
21 breeders to select the most durable resistance genes.

22 23 **Corresponding authors:**

24 Bérenger Janzac and Benoît Moury, INRA, UR407 Pathologie Végétale, Domaine Saint
25 Maurice, BP94, F-84140 Montfavet cedex, France.

26 *E-mail:* berenger.janzac@avignon.inra.fr

27 *E-mail:* benoit.moury@avignon.inra.fr

Code de champ modifié

A new predictor of plant resistance durability

28 **Introduction**

29 Understanding the mechanisms linked to the emergence of pathogenic micro-organisms is a
30 major challenge considering their impact on human health and activities. Most attention has
31 been paid to the role of ecological factors (including migration, climate, agricultural practices
32 and spread of biological disease vectors) in the emergence of plant diseases (Anderson *et al.*,
33 2004). However, the role of the ecological factors is often difficult to disentangle from that of
34 the genetic changes operating within the pathogen populations (Fargette *et al.*, 2006; Desprez-
35 Loustau *et al.*, 2007; Holmes and Drummond, 2007). The emergence of pathogens requires
36 both their genetic adaptation to new hosts and their subsequent spread and maintenance
37 within host populations. Diverse strategies have been developed to control plant pathogens,
38 the use of resistant varieties by growers being the most efficient and simplest one. However,
39 the widespread use of genetic resistances imposes strong selective pressures on the targeted
40 pathogens which are able to adapt, resulting in the breakdown of the resistance. Resistance
41 breakdown is a case of pathogen emergence which involves the jump of a host barrier at the
42 intraspecific level and whose genetic bases have been best documented (Flor 1971; Gabriel
43 1999; Skamnioti and Ridout, 2005). Combining surveys of resistance breakdown on the long
44 term with genetic analyses of pathogen populations can therefore help understanding the
45 respective roles of intrinsic evolutionary processes and external ecological factors in pathogen
46 emergence.

47 As defined by Johnson (1979), a resistance is durable if it remains effective in a cultivar for a
48 long period of time during its widespread cultivation under environments favourable to the
49 disease development. According to this definition, the durability of resistances can be
50 measured only *a posteriori* after their large-scale deployment. Ability to predict the durability
51 of resistance represents a major economic issue in agriculture. However, the criteria currently
52 retained by breeders to evaluate resistance sources do not include their potential durability

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Molecular Plant Pathology, 2009, vol.10, no.5. DOI: 10.1111/j.1364-3703.2009.00554.x

A new predictor of plant resistance durability

53 because of the lack of predictors. According to several authors, the mechanism of action and
54 the genetic determinism of plant resistance to pathogens explain only a small part of their
55 durability (Fraser 1990; McDonald and Linde, 2002; García-Arenal and McDonald, 2003).
56 The first predictor of resistance durability to pathogens was proposed by McDonald and
57 Linde (2002) and further applied to virus resistances by García-Arenal and McDonald (2003).
58 This predictor, called ‘evolutionary potential’ of pathogens focused on processes that govern
59 pathogen population evolution (mainly effective population size, migration and reproduction
60 system). However, this approach ignores the molecular events involved in resistance
61 breakdown or the fitness cost associated to virulence, i.e. ability of a pathogen genotype to
62 overcome a resistance in the case of gene-for-gene interactions (Sacristán and García-Arenal,
63 2008), and therefore cannot explain why different resistance genes directed toward the same
64 pathogen can display different durabilities.

65 Flor (1955, 1971) was the first to show that resistance or susceptibility of plants to pathogens
66 result from an intimate molecular relationship governed by a gene-for-gene interaction model.
67 In this model, the interaction between two partners, the resistance gene of the plant (under at
68 least two allelic forms: “resistant” and “susceptible”) and the avirulence gene of the pathogen
69 (under at least two allelic forms: “avirulent” and “virulent”), determine the resistance or the
70 susceptibility of a plant to a pathogen (Flor 1971). Breakdown of resistance occurs when the
71 two following steps can be fulfilled by the pathogen: (i) Mutation and/or recombination
72 events should appear in the avirulence gene of a pathogen to generate virulent variants and (ii)
73 the fitness of these virulent variants should permit them to increase in frequency in the plant
74 population and spread in the agro-ecosystem (Leach *et al.*, 2001). Since resistance durability
75 is linked to the mechanisms and dynamics that control these two steps, we propose a new
76 predictor of plant resistance durability based on the properties of pathogen avirulence factors.
77 We focussed our study on plant viruses, for which a large number of avirulence factors

A new predictor of plant resistance durability

78 corresponding to diverse plant resistance genes have been determined (Kang *et al.*, 2005b)
79 and for which estimates of resistance durability are available from field data. We checked
80 whether the evolutionary constraint acting on amino acid substitutions in avirulence factors
81 determines the durability of the corresponding resistance. Indeed, it is likely that the higher
82 the constraint on the avirulence factor, the higher the fitness penalty conferred by mutations
83 and/or recombination in this avirulence factor and consequently, the higher the resistance
84 durability. In addition, we tested the hypotheses that the nucleotide or amino acid diversity of
85 avirulence factors determines resistance durability. Indeed, the higher the diversity in
86 avirulence factors, the higher the probability for the virus to evolve a virulent form of the
87 avirulence factor with a minimum number of nucleotide or amino acid changes and hence, the
88 lower the resistance durability. It should be noted that, in theory, there is no expected
89 correlation between the selective constraint applied on genes and their diversity (Kimura 1985;
90 Wyckoff *et al.*, 2005), justifying to consider these hypotheses separately.

92 Results

93 *Dataset construction and evaluation*

94 The first step to test the correlation between the resistance durability and its potential
95 predictors (evolutionary constraint and diversities of avirulence factors) was to build a dataset
96 allowing to estimate precisely all these variables. Altogether, 20 virus species-plant resistance
97 gene combinations satisfying this condition were retained (see Experimental procedures).
98 These combinations correspond to 17 different avirulence factors, 13 different virus species
99 and seven virus genera (Table 1). All viruses are RNA viruses, with a majority of potyviruses
00 and tobamoviruses, and they infect a large diversity of host species. Seven viruses are
01 transmitted by insects, two by soilborne fungi and four by contacts between plants. Six
02 recessive and 14 dominant resistance genes are included.

A new predictor of plant resistance durability

103 The 20 virus species-plant resistance gene combinations composing the dataset were
104 classified among four durability classes of increasing durability according to the time scale of
105 protection efficiency and the proportion of virulent variants described in the agroecosystem
106 (Table 1 and see Experimental procedures). The diversity of avirulence factors and the
107 evolutionary constraint acting on these factors were estimated from sequences available in
108 GenBank (see supplementary material). The p-distance was used to estimate the diversity of
109 avirulence factors both at the nucleotide and amino acid levels and the ω ratio, which is the
110 ratio between the non-synonymous and the synonymous substitution rates (Kimura 1983),
111 was used to estimate the evolutionary constraint acting on amino acid substitutions in the
112 avirulence factors (Yang *et al.*, 2000; Yang and Swanson, 2002) (see Experimental
113 procedures). We observed that the nucleotide diversity of the avirulence factors retained in the
114 dataset was usually low and varied from 0.005 (movement protein of ToMV) to 0.160 (VPg
115 of BYMV). Similarly, the amino acid diversity varied from 0.005 to 0.109 (Table 1). The ω
116 ratio was generally low (≤ 0.3) indicating that avirulence factors were globally submitted to
117 negative selection and that most amino acid changes were deleterious in these proteins
118 (Kimura 1983). However, for the P25 protein of BNYVV a higher value was estimated ($\omega =$
119 0.97). For all the avirulence factors analyzed, no amino acid sites were found under
120 significant positive selection ($\omega > 1$) with the SLAC method.

121 To get rid of possible biases in the estimation of the putative predictors of resistance
122 durability, we investigated the minimum number of sequences required for correct estimation
123 of the diversities and ω ratio. We noticed that eight to ten sequences were usually enough to
124 obtain satisfactory estimates (see Experimental procedures and Fig.1) and that the lower the
125 variable, the smaller the number of sequences required. Consequently, for the analyses we
126 have retained only datasets with a minimum of ten sequences of the avirulence factor except
127 in three datasets (P3-HcPro of SMV, 183 kDa protein of ToMV and 25 kDa protein of PVX)

A new predictor of plant resistance durability

153 0.003; ordinal multinomial model) than with three durability classes (P-value = 0.008) (Table
154 2). The relationship between the diversity of avirulence factors and resistance durability was,
155 on the whole, opposite to that expected: The higher the diversity, the higher the resistance
156 durability (Fig. 2 B and C). For the nucleotide diversity, a marginally significant relationship
157 was detected using the KW test (P-value = 0.052, Table 2). Note that the results obtained with
158 the ordinal multinomial model cannot be exploited because the proportional odds assumption
159 was not satisfied (score-test = 0.039). No significant relationship was found between the
160 resistance durability and either (i) the amino acid diversity, (ii) the evolutionary potential of
161 viruses defined by García-Arenal and McDonald (2003) or (iii) the type of resistance
162 (dominant or recessive) (Fig. 2 C and D, Table 2). Similarly, the separate analysis of the
163 relationships between resistance durability and the three components of the evolutionary
164 potential of viruses did not reveal any significance (data not shown). The smaller size of the
165 dataset used here, as well as differences in the analytical methods could explain that
166 difference with García-Arenal and McDonald's analyses. A likelihood ratio test indicates that
167 a model including the ω ratio estimated on avirulence factors and the evolutionary potential of
168 viruses or the type of resistance does not explain better the resistance durability than a model
169 involving only the ω ratio (Table 2).

71 Discussion

172 McDonald and Linde (2002) first proposed a predictor of the durability of resistances based
173 on the biological properties of the targeted pathogen. Later García-Arenal and McDonald
174 (2003) applied the same approach to virus resistances and defined a risk index for virus
175 species which correlates to resistance durability. With the increasing knowledge on viral
176 avirulence factors, it has been postulated that some relationships might exist between the
177 resistance durability and (i) the number of mutations involved in virulence (Harrison 2002,

A new predictor of plant resistance durability

178 Ayme *et al.*, 2007), or (ii) the fitness cost associated to these mutations (Leach *et al.*, 2001;
179 Jenner *et al.*, 2002). Fitness cost due to mutations involved in virulence property seemed
180 important to predict the resistance durability in several plant-virus pathosystems (Boyd 2006;
181 Lecoq *et al.*, 2004). We have consequently explored the possible relationships between the
182 observed resistance durability and properties of virus avirulence factors.

Causality relationships between properties of viral avirulence factors and resistance durability

185 We observed a marginal relationship between resistance durability and the nucleotide
186 diversity of the avirulence factor (Table 2). This relationship was opposite to that expected
187 since the higher the diversity, the higher the resistance durability (Fig. 2 B). This results
188 mainly from the lower diversity observed in durability class 1. A possible explanation is that
189 three (TSWV, BNYVV and PMMoV) out of four viruses in the durability class 1 emerged
190 relatively recently on a large geographical scale (independently of the resistance breakdown
191 phenomenon), while the emergence of viruses in other classes was older (at least 20 years ago
192 given our definition of durability classes 2, 3 and 4). As a consequence, viruses in the
193 durability class 1 may have had less time to get diversified than viruses corresponding to
194 classes 2, 3 and 4. The marginal correlation between resistance durability and nucleotide
195 diversity of avirulence factors would then be both indirect and due to a bias in our dataset.
196 More precise data about resistance breakdowns would be necessary to validate this
197 assumption. It is important to note that this putative bias has, in theory, no consequences on
198 our analyses involving the ω ratio (Kimura 1985; Wyckoff *et al.*, 2005).

199 A significant relationship was detected between resistance durability and the constraint
200 exerted on amino acid changes (ω ratio) in avirulence factors (Table 2). Predicting resistance
201 durability from the value of the ω ratio implies a causality relationship. The observed
202 association between resistance durability and the ω ratio could be due to the following

A new predictor of plant resistance durability

relationships: (i) The ω ratio determines the resistance durability; (ii) the resistance durability determines the ω ratio; (iii) both the resistance durability and the ω ratio are determined by at least one third factor. Since we could not find any evident third factor that would determine a negative correlation between resistance durability and ω ratio of avirulence factors, the first two scenarios only will be examined. Could resistance durability determine variations in the ω ratio of avirulence factors? Datasets corresponding to durability classes 1 and (to a lower extent) 2 contain a proportion of sequences of virulent isolates which have undergone amino acid substitutions in their avirulence factors. Since these particular amino acid positions could be subjected to positive selection (Moury 2004; Schirmer *et al.*, 2005), this could increase the ω ratio for durability classes 1 and 2 while affecting less (or not at all) the ω ratio for the durability classes 3 and 4. However, the facts that (i) in durability classes 1 and 2 a limited proportion of virulent isolates (0% to 22 %) are characterized (except for BNYVV-*Rz1*, BaYMV-*rym4*, BYMV-*wlv*, PMMoV-*L*³, SMV-*Rsv1* and PVY-*pvr2*¹) (Table 2 in supplementary materials), (ii) that the change in virulence properties generally depends only on a small number of amino acid substitutions, usually one or two (Harrison 2002), and (iii) that the presence of virulent isolates in datasets did not introduce significant biases into the ω ratio estimates for the BNYVV-*Rz1*, BaYMV-*rym4* and PMMoV-*L*³ combinations (the BYMV-*wlv*, SMV-*Rsv1* and PVY-*pvr2*¹ datasets did not include enough virulent and avirulent isolates for such analyses), suggest that the ω ratio in viral avirulence factors indeed determines the level of resistance durability. In addition, and most convincingly, when all virulent isolates were withdrawn from the sequence dataset, the negative correlation between the ω ratio and the resistance durability was still significant (P-value = 0.019; ordinal multinomial model) (Table 2). Note, however, that this relationship was not significant (P-value = 0.11) with the KW test, probably because four pathosystems were excluded from that analysis, due to insufficient number of sequences, and because the KW test has a low

A new predictor of plant resistance durability

228 statistical power (see Experimental procedures). The mechanism behind this causal link
229 resides in the fact that the ω ratio measures the propensity of viruses to change amino acids in
230 their avirulence factors without high fitness costs under the assumption that synonymous
231 nucleotide changes are neutral.

232 For viruses and other plant pathogens, gains of virulence toward resistant genotypes are
233 frequently associated to fitness costs in susceptible ones (Meshi *et al.*, 1989; Leach *et al.*,
234 2001; Jenner *et al.*, 2002; Desbiez *et al.*, 2003; Ayme *et al.*, 2007). Even though limited
235 numbers of amino acid positions have been shown to be involved in virulence properties
236 (Harrison 2002), measure of the evolutionary constraint on the whole avirulence factor
237 seemed adequate to our study for two reasons. First, it is plausible that much more virulence
238 mutations than presently known exist for a given virus species-resistance gene interaction
239 because reverse genetic analyses of avirulence mutations have frequently been performed
240 with a small number of virulent and avirulent isolates. Second, when virulence mutations in
241 the virus impose a fitness cost, compensatory mutations may appear somewhere else in the
242 avirulence factor, making worthwhile the measure of the ω ratio also in these parts of the
243 protein.

Predicting resistance durability as a function of the ω ratio

244 Since the evolutionary constraint acting on amino acid substitutions (ω ratio) in avirulence
245 factors is correlated to the durability of resistance, it is possible to assess, for a given
246 resistance gene, the probability of being in a given durability class as a function of a measured
247 ω ratio using Eq.1 (see Experimental procedures) (Fig. 3). According to this model, resistance
248 genes corresponding to avirulence factors with ω values higher than 0.50 are associated to the
249 class of lower durability (class 1) with a fairly high probability ($P > 0.80$). On the opposite,
250 resistance genes for which avirulence factors have ω values lower than 0.15 are assigned
251 either to the classes of medium or high durability (classes 2 or 3+4) with a probability $P \geq 0.80$.

A new predictor of plant resistance durability

253 On the whole, cross validation tests (see Experimental procedures) allowed assignments of
254 resistance genes to the correct durability class in 44% of the cases (33% of correct
255 assignments would be expected by chance). Even if our model still lacks power, 65% of
256 resistance genes were correctly assigned in the durability class 3+4 corresponding to the class
257 of genes of highest interest for resistance durability. It could also be noted that our model
258 cannot account for the fact that different resistance alleles or genes interacting with the same
259 viral avirulence gene show different durabilities as, for example, *pvr2*¹ and *pvr2*² for PVY
260 (Table 1). Additional virus species-resistance gene combinations or additional predictive
261 variables would be needed to improve the precision of those predictors. In the same way,
262 using ω ratio estimates on the specific domain of avirulence protein which interact with
263 resistance factors could improve the power of this model.

264 *A novel strategy for resistance screening: Targeted search for durable resistance*

265 Resistance durability is presently evaluated downstream of the breeding programs, when the
266 resistance genes have been deployed in commercial varieties at a large scale and during
267 several years. Incorporating resistance durability earlier in the breeding scheme would be
268 highly desirable (Boyd 2006). Beyond their predictive value for the durability of resistances
269 already incorporated in plant cultivars, ω ratios estimated from viral avirulence genes could
270 also be integrated upstream into resistance breeding programs in two ways. First, comparing
271 estimates of the ω ratios from avirulence factors corresponding to several candidate resistance
272 sources could allow breeders to evaluate and select the potentially more durable resistances.
273 Second, new resistance sources could be identified via the directed search for resistance
274 factors elicited by candidate avirulence factors. Thus, it was shown that transient expression
275 of avirulence proteins in plant tissues allowed to detect resistance genes in plants through the
276 elicitation of a necrotic response characteristic of dominant resistances (Laugé *et al.*, 1998;
277 Bendahmane *et al.*, 2000; Mestre *et al.*, 2000). Even if the efficiency of the detected

A new predictor of plant resistance durability

278 resistance should be further confirmed by inoculation tests since necrosis could also be
279 induced independently of the specific resistance phenomenon, using this method with the
280 most constrained viral proteins could allow the *a priori* identification of the most durable
281 resistances targeting them.

282 **Conclusion**

283 As proposed by McDonald and Linde (2002) and García-Arenal and McDonald (2003), we
284 analyzed the ability of some pathogen traits to predict the capacity of emergence of adapted
285 pathogens in crop plants carrying specific resistance genes. We focused on plant-virus
286 interactions, but our analytical approach could however have a broader interest. The validity
287 of the method depends on two functional assumptions: (i) The control mechanism (here plant
288 resistance genes) interacts either directly or indirectly with a specific protein target and (ii)
289 adaptation of the targeted organism to circumvent the control mechanism occurs via
290 mutations in that protein. The observed correlation between the ω ratio of viral avirulence
291 factors and the durability of corresponding plant resistance genes strongly suggests that these
292 assumptions are widely satisfied in the case of plant-virus interactions. Notably, the fact that
293 the molecular interactions between both partners operate at the protein but not the RNA level,
294 from the virus side, is not trivial since it has been shown that RNA segments can act as viral
295 avirulence factors independently of their possible coding capacity (Szittyá and Burgyan 2001;
296 Diaz *et al.*, 2004). Since the two above assumptions are commonplace among plant-pathogen
297 interactions, it would be worth evaluating the generalization of our approach to other kinds of
298 plant pathogens.

299 **Experimental procedures**

300 **Dataset**

A new predictor of plant resistance durability

303 To analyze the durability of plant resistance to viruses, we retained the virus species-
304 resistance gene (or allele) combinations which satisfied the three following conditions: (i) The
305 virus avirulence factor corresponding to the resistance gene was identified, (ii) nucleotide
306 sequences encoding the avirulence factor could be obtained from a sufficient number of virus
307 isolates to estimate their diversity and evolutionary constraint and (iii) plant cultivars carrying
308 the resistance gene have been largely exploited and field data about the durability of this
309 resistance are available. All resistance genes composing the dataset control qualitative
310 resistance (Table 1). Nucleotide sequences of avirulence factors were retrieved from Genbank,
311 aligned with the ClustalW program (Thompson *et al.*, 1994) and analyzed with the RDP
312 version 3 software (Martin *et al.*, 2005) to remove putative recombinant sequences which
313 could create biases in the estimation of the evolutionary constraints (Schierup and Hein, 2000).
314 For *Potato virus Y* (PVY) and *Tobacco etch virus* (TEV), five and twelve additional
315 sequences of the VPg-coding region, respectively, were obtained (accession numbers
316 EU334778 to EU334794). For the *pvr2¹* and *pvr2²* resistance alleles in pepper, only VPg
317 sequences from pepper isolates of PVY belonging to the C phylogenetic group were analyzed.
318 In case of overlapping open reading frames in the cistron coding for the avirulence factor (183
319 or 184 kDa proteins and movement protein of tobamoviruses and 25 kDa protein of potato
320 virus X), the overlapping region was removed to avoid artifacts in the estimation of their
321 diversity and evolutionary constraint. GenBank accession numbers of sequences retained are
322 available in supplementary materials (Table 2). To avoid artifactual polymorphism, we
323 discarded the sequences of virus isolates that have been maintained on the long term in
324 laboratory through repeated inoculations and sequences of virulent isolates selected
325 artificially.

326 ***Estimation of avirulence factors diversity, constraint exerted on amino acid substitutions***
327 ***and evolutionary potential of viruses***

A new predictor of plant resistance durability

328 For each virus species-resistance gene combination, four following putative predictors of
329 resistance durability were determined: (i) The diversity of the avirulence factors at the
330 nucleotide level, (ii) the diversity of the avirulence factors at the amino acid level, (iii) the
331 evolutionary constraint on amino acid substitutions in the avirulence factors and (iv) the
332 evolutionary potential of viruses as defined by García-Arenal and McDonald (2003).
333 Sequence alignments of avirulence factors were used to estimate the first three predictors. The
334 average pairwise p-distance was used to estimate the diversity of the avirulence factors, both
335 at the nucleotide and amino acid levels, using MEGA version 3.1 (Kumar *et al.*, 1994). The ω
336 ratio, which is the ratio between the non-synonymous (amino acid altering) and the
337 synonymous (silent at the amino acid level) substitutions, was used to estimate the
338 evolutionary constraint acting on amino acid substitutions in the avirulence factors (Kimura
339 1983; Yang *et al.*, 2000; Yang and Swanson, 2002). This ω ratio was estimated on whole
340 avirulence factor sequences with the SLAC (Single-Likelihood Ancestor Counting) method in
341 the HyPhy software (Kosakovsky Pond and Frost, 2005), a modification of the site-by-site
342 analytical method of Suzuki and Gojobori (1999). The probability threshold to detect
343 particular amino acid sites undergoing positive selection ($\omega > 1$) was set to $P = 0.10$.

344 The evolutionary potential of viruses (García-Arenal and McDonald, 2003) corresponds to a
345 compound risk index based on the biology of the viruses and takes into account: (i) The
346 population size of the virus, (ii) the amount of gene and genotype flow in the virus population,
347 and (iii) its frequency of recombination. Evolutionary potential of virus species were obtained
348 from García-Arenal and McDonald (2003), or estimated according to the criteria defined by
349 these authors (Table 1 and supplementary materials).

Precision of the estimates of diversity and evolutionary constraints of avirulence factors

351 The precision of estimates for the diversity and evolutionary constraint of avirulence genes
352 depends on representative samples of sequences of these genes. A method to evaluate the

A new predictor of plant resistance durability

353 representativeness of samples is to measure the variance of the studied variables with random
354 subsamples of increasing sizes (Hughes et al. 2001). Here, subsamples of sequences of
355 increasing size were randomly chosen among the five largest sequence sets (P25 of BNYVV,
356 CP of ToMV, VPg of BaYMV, CP of PVX and NSs of TSWV) presenting contrasted
357 diversity and evolutionary constraint (Table 1). In each subsample, ten random sets of
358 sequences were analyzed allowing to estimate the variance of the three variables (nucleotide
359 and amino acid diversities and ω ratio) and to perform mean comparisons. The minimal
360 number of sequences that allowed to obtain satisfactory estimates of the ω ratio or diversities
361 of avirulence factors (*i.e.* minimum sample delivering estimates that are not significantly
362 different from those obtained with the whole dataset) was comprised between eight (CP of
363 PVX and NSs of TSWV) and ten (CP of ToMV and VPg of BaYMV) (Fig. 1). The only
364 exception concerned the P25 protein of BNYVV that has a higher ω ratio and, consequently,
365 required a minimum of 20 sequences to be estimated properly (Fig. 1).

366 The same approach was used to determine the effect of increasing the number of geographical
367 localities (countries) on the precision of variable estimates. Subsamples of ten sequences with
368 increasing numbers of geographic origins chosen among the four sequence sets showing the
369 highest diversity of geographical origins (TSWV-NSs, TuMV-CI, PMMoV-CP, PVY-NiaPro)
370 were used to estimate the variance of the variables. No significant differences of predictor
371 estimates were revealed by rank tests performed on subsamples of sequences showing
372 increasing diversity of geographic origins. To evaluate the impact of virulent isolates on our
373 results, we also compared the estimates of the ω ratio obtained with sequence samples
374 including or not the virulent isolates for BNYVV-*Rz1*, BaYMV-*rym4* and PMMoV-*L*³ which
375 included sufficient numbers of virulent and avirulent isolates (See additional material Table 2).
376 We also conducted our analyses with a dataset excluding the sequences of virulent isolates to
377 get rid of any bias on the correlation between the ω ratio of avirulence factors and resistance

A new predictor of plant resistance durability

378 durability. This dataset was made up of 16 virus species-plant resistance gene combinations
379 satisfying the condition of a minimum of 8 sequences of the avirulence gene (BYMV-*w/v*,
380 SMV-*Rsv1*, PVY-*pvr2*¹ and PVX-*Nb* were excluded).

381 *Estimation of the durability of resistance*

382 The durability of resistance genes or alleles was assessed according to the timescale of the
383 protection efficiency before the appearance of crop damages due to resistance-breaking
384 isolates and according to the abundance of resistance-breaking isolates in the agro-ecosystem
385 (Table 1). A majority of these data was collected from García-Arenal and McDonald (2003)
386 and Harrison (2002) and additional references are available in supplementary Table 1. All the
387 resistances in the dataset have been used in the field for more than 20 years or were broken
388 down rapidly (Table 1). Resistances were considered as “broken-down” if the resistance was
389 defeated by resistance-breaking isolates and lead to important economic losses at least in
390 some growing regions. Resistance durability was categorized in four discrete durability
391 classes as follows.

392 Class 1: Rapid resistance breakdown (less than 12 years after initial deployment). Frequent
393 virulent isolates in the field prevent the efficiency of the resistance at least in some growing
394 regions.

395 Class 2: Slow resistance breakdown (more than 12 years after initial deployment). Frequent
396 virulent isolates in the field prevent the efficiency of the resistance at least in some growing
397 regions.

398 Class 3: Resistances deployed for more than 20 years with rare virulent isolates in the field.
399 Virulent isolates are restricted to small areas and did not economically compromise the use of
400 the resistant cultivars.

401 Class 4: Resistances deployed for more than 20 years and no virulent strains described in the
402 field.

A new predictor of plant resistance durability

403 In our dataset (Table 1), the distinction between the durability classes 1 and 2 was clear cut
404 since breakdown of resistance in class 1 occurred in one to eleven years while breakdown of
405 resistances in class 2 occurred mostly in more than 20 years. The SMV-*Rsv-1* gene
406 combination only had an intermediate durability (breakdown in about 15 years). However,
407 similar results were obtained when it was classified in class 1 or 2. Due to the small number
408 of resistances in class 4, classes 3 and 4 were grouped for further analyses.

409 *Statistical analysis*

410 Statistical analyses were used to evaluate the correlation between the response variable
411 “durability of resistance” and four potential predictor variables (nucleotide and amino acid
412 diversities of avirulence factors, ω ratios of avirulence factors and evolutionary potential of
413 viruses). We first used a non-parametric Kruskal-Wallis test for the null hypothesis: “The
414 average values of the considered explanatory variable are similar for all resistance durability
415 classes”. We also used a parametric approach based on ordinal multinomial models.
416 Multinomial models extend logistic regression to the case where the response variable has
417 three (or more) discrete outcomes (*i.e.* the response variable follows a multinomial
418 distribution) (McCullagh and Nelder, 1989). In our case, the durability of resistances is an
419 ordinal response variable because resistance durability increases from class 1 to class 3+4.
420 Ordinal multinomial models allow both (i) to test for a correlation between predictor and
421 response variables by taking into account the ordinal and multinomial nature of the response
422 variable and (ii) to model the probability to belong to a particular durability class as a function
423 of a set of predictor variables. More precisely, Y being the response variable “durability of
424 resistance”, let's denote $P_r = p(Y \leq r)$ the probability of Y being in the r or lower class of
425 durability ($r \in \{1, 2, 3+4\}$). In the proportional odds model, P_r is modelled with a logit link as

$$426 P_r = \frac{\exp(\alpha_r + \beta.X)}{1 + \exp(\alpha_r + \beta.X)} \text{ for } r=1,2 \text{ and } P_3=1 \text{ (Eq.1)}$$

A new predictor of plant resistance durability

427 where α_r are intercept parameters that depend only on the categories such that $\alpha_1 < \alpha_2$ and β
428 is a vector of the regression coefficients for the set of predictor variables X . Inherent in this
429 model is the proportional odds assumption, which states that the odds (*i.e.* the quantity
430 $p(Y \leq r) / p(Y > r)$) of being in a category less than or equal to r is $\exp[\beta \cdot (x_1 - x_2)]$ times higher at
431 $X = x_1$ than at $X = x_2$, x_1 and x_2 being two particular values of X . The PROC LOGISTIC of the
432 SAS software (version 8, SAS Inc., Cary, NC) was used to estimate α_r and β by maximum
433 likelihood for five models differing by the set X of predictor variables included (amino acid
434 diversity, nucleotide diversity, evolutionary potential, ω ratio and the combination of ω ratio
435 and evolutionary potential). Before fitting these models, a score test for the proportional odds
436 assumption was performed, a non significant test supporting the evidence that the assumption
437 is valid.

438 The predictive performance of the ordinal multinomial model was assessed with cross-
439 validation tests. Cross-validation tests were performed with the R software environment
440 (<http://cran.r-project.org/>) by (i) randomly leaving out 3 pathosystems (*i.e.* 15% of the dataset),
441 (ii) fitting the model to the remaining dataset and (iii) checking if the true class of durability
442 for each left-out resistance gene-avirulence factor combination corresponded to the highest
443 probability class predicted by the model (if so, the prediction is correct). The predictive
444 performance of the model was assessed as the percentage of correct assignments and
445 estimated by iterating 1,500 times the previous three steps.

447 Acknowledgements

448 We acknowledge Dr. D. Fargette (IRD, Montpellier, France) for sharing information about
449 the durability of the *Rymv-1* gene, C. Marchal (Clause-Tézier) for providing TEV isolates,
450 and K. Ezzaier (INRAT, Tunisia) and Dr. G. Marchoux (INRA, Avignon) for providing PVY
451 isolates. We also acknowledge the ANRT foundation and the companies Gautier Semences,

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Molecular Plant Pathology, 2009, vol.10, no.5. DOI: 10.1111/j.1364-3703.2009.00554.x

A new predictor of plant resistance durability

452 Clause Vegetable Seeds, Vilmorin SA, Sakata Seeds Europe and Rijk Zwaan for their
453 financial support.

A new predictor of plant resistance durability

References

- 454
455 **Acosta-Leal, R. and Rush, C. M.** (2007) Mutations associated with resistance-breaking
456 isolates of beet necrotic yellow vein virus and their allelic discrimination using TaqMan
457 technology. *Phytopathology* **97**, 325-330.
458
- 459 **Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R. and**
460 **Daszak, P.** (2004) Emerging infectious diseases of plants: pathogen pollution, climate
461 change and agrotechnology drivers. *Trends Ecol. Evol.* **19**, 535-544.
462
- 463 **Ayme, V., Petit-Pierre, J., Souche, S., Palloix, A. and Moury, B.** (2007) Molecular
464 dissection of the potato virus Y VPg virulence factor reveals complex adaptations to the
465 pvr2 resistance allelic series in pepper. *J. Gen. Virol.* **88**, 1594-1601.
466
- 467 **Bendahmane, A., Kohm, B. A., Dedi, C. and Baulcombe, D. C.** (1995) The coat protein of
468 potato virus X is a strain-specific elicitor of Rx1-mediated virus resistance in potato. *Plant J.*
469 **8**, 933-941.
470
- 471 **Bendahmane, A., Querci, M., Kanyuka, K. and Baulcombe, D. C.** (2000) Agrobacterium
472 transient expression system as a tool for the isolation of disease resistance genes:
473 application to the Rx2 locus in potato. *Plant J.* **21**, 73-81.
474
- 475 **Berzal-Herranz, A., De la Cruz, A., Tenllado, F., Díaz-Ruiz, J. R., López, L., Sanz, A. I.,**
476 **Vaquero, C., Serra, M. T. and García-Luque, I.** (1995) The *Capsicum* L3 gene-mediated
477 resistance against the tobamoviruses is elicited by the coat protein. *Virology* **209**, 498-505.
478
- 479 **Boyd, L. A.** (2006) Can the durability of resistance be predicted? *J. Sci. Food Agr.* **86**, 2523-
480 2526.
481
- 482 **Bruun-Rasmussen, M., Møller, I. S., Hensen, J. K. R., Lund, O. S. and Johansen, I. E.**
483 (2007) The same allele of translation initiation factor 4E mediates resistance against two
484 *Potyvirus* spp. in *Pisum sativum*. *Mol. Plant-Microbe Interact.* **20**, 1075-1082.
485
- 486 **Chiba, S., Miyanishi, M., Andika, I. B., Kondo, H. and Tamada, T.** (2008) Identification
487 of amino acids of the beet necrotic yellow vein virus p25 protein required for induction of
488 the resistance response in leaves of *Beta vulgaris* plants. *J. Gen. Virol.* **89**, 1314-1323.
489
- 490 **Dardick, C. D., Taraporewala, Z., Lu, B. and Culver, J. N.** (1999) Comparison of
491 tobamovirus coat protein structural features that affect elicitor activity in pepper, eggplant,
492 and tobacco. *Mol. Plant-Microbe Interact.* **12**, 247-251.
493
- 494 **Desbiez, C., Gal-On, A., Girard, M., Wipf-Scheibel, C. and Lecoq, H.** (2003) Increase in
495 *Zucchini yellow mosaic virus* symptom severity in tolerant zucchini cultivars is related to a
496 point mutation in P3 protein and is associated with a loss of relative fitness on susceptible
497 plants. *Phytopathology* **93**, 1478-1484.
498
- 499 **Desprez-Loustau, M. L., Robin, C., Buee, M., Courtecuisse, R., Garbaye, J., Suffert, F.,**
500 **Sache, I. and Rizz, D. M.** (2007) The fungal dimension of biological invasions. *Trends*
501 *Ecol. Evol.* **22**, 472-480.
502

A new predictor of plant resistance durability

- 503 **Diaz, J., Nieto, C., Moriones, E., Truniger, V. and Aranda, M.** (2004) Molecular
504 characterization of a Melon necrotic spot virus strain that overcomes the resistance in melon
505 and nonhost plants. *Mol. Plant-Microbe Interact.* **17**, 668-75.
506
- 507 **Eggenberger, A. L., M. R. Hajimorad and Hill, J. H.** (2008) Gain of virulence on Rsv1-
508 genotype soybean by an avirulent Soybean mosaic virus requires concurrent mutations in both
509 P3 and HC-Pro. *Mol. Plant-Microbe Interact.* **21**, 931-936.
510
- 511 **Eggenberger, A. L. and Hill, J. H.** (1997) Analysis of resistance-breaking determinants of
512 soybean mosaic virus. *Phytopathology* **87(suppl.)**, S27.
513
- 514 **Fargette, D., Konate, G., Fauquet, C., Muller, E., Peterschmitt, M. and Thresh, J. M.**
515 (2006) Molecular ecology and emergence of tropical plant viruses. *Annu. Rev. Phytopathol.*
516 **44**, 235-260.
517
- 518 **Flor, H. H.** (1955) Host-parasite interaction in flax rust - its genetics and other implications.
519 *Phytopathology* **45**, 680-685.
520
- 521 **Flor, H. H.** (1971) Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.* **9**,
522 275-296.
523
- 524 **Fraser, R. S. S.** (1990) The genetics of resistance to plant viruses. *Annu. Rev. Phytopathol.*
525 **28**, 179-200.
526
- 527 **Gabriel, D. W.** (1999) Why do pathogens carry avirulence genes? *Physiol. Mol. Plant P.* **55**,
528 205-214.
529
- 530 **García-Arenal, F. and McDonald, B. A.** (2003) An analysis of the durability of resistance to
531 plant viruses. *Phytopathology* **93**, 941-952.
532
- 533 **Hajimorad, M. R., Eggenberger, A. L. and Hill, J. H.** (2006) Strain-specific P3 of *Soybean*
534 *mosaic virus* elicits Rsv1-mediated extreme resistance, but absence of P3 elicitor function
535 alone is insufficient for virulence on Rsv1-genotype soybean. *Virology* **345**, 156-166.
536
- 537 **Hamada, H., Tomita, R., Iwadate, Y., Kobayashi, K., Munemura, I., Takeuchi, S.,**
538 **Hikichi, Y. and Suzuki, K.** (2007) Cooperative effect of two amino acid mutations in the
539 coat protein of *Pepper mild mottle virus* overcomes L3-mediated resistance in Capsicum
540 plants. *Virus Genes* **34**, 205-214.
541
- 542 **Harrison, B. D.** (2002) Virus variation in relation to resistance breaking in plants. *Euphytica*
543 **124**, 181-192.
544
- 545 **Hébrard, E., Pinel-Galzi, A., Bersoult, A., Sire, C. and Fargette, D.** (2006) Emergence of
546 a resistance-breaking isolate of *Rice yellow mottle virus* during serial inoculations is due to
547 a single substitution in the genome-linked viral protein VPg. *J. Gen. Virol.* **87**, 1369-1373.
548
- 549 **Holmes, E. C. and Drummond, A.** (2007) The evolutionary genetics of viral emergence.
550 *Curr. Top. Microbiol. Immunol.* **315**, 51-66.
551

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Molecular Plant Pathology, 2009, vol.10, no.5. DOI: 10.1111/j.1364-3703.2009.00554.x

A new predictor of plant resistance durability

- 552 **Hughes, J. B., Hellmann, J. J., Ricketts, T. H. and Bohannan, B. J. M.** (2001) Counting
553 the uncountable: Statistical approaches to estimating microbial diversity. *Appl. Envir.*
554 *Microbiol.* **67**, 4399-4406.
- 555
556 **Jenner, C. E., Sanchez, F., Nettleship, S. B., Foster, G. D., Ponz, F. and Walsh, J. A.**
557 (2000) The cylindrical inclusion gene of turnip mosaic virus encodes a pathogenic
558 determinant to the Brassica resistance gene TuRB01. *Mol. Plant-Microbe Interact.* **13**,
559 1102-1108.
- 560
561 **Jenner, C. E., Wang, X., Ponz, F. and Walsh, J. A.** (2002) A fitness cost for *Turnip mosaic*
562 *virus* to overcome host resistance. *Virus Res.* **86**, 1-6.
- 563
564 **Johnson, R.** (1979) The concept of durable resistance. *Phytopathology* **69**, 198-199.
- 565
566 **Kang, B.-C., Yeam, I., Frantz, J. D., Murphy, J. F. and Jahn, M. M.** (2005a) The pvr1
567 locus in *Capsicum* encodes a translation initiation factor eIF4E that interacts with Tobacco
568 etch virus VPg. *Plant J.* **42**, 392-405.
- 569
570 **Kang, B.-C., Yeam, I. and Jahn, M. M.** (2005b) Genetics of plant virus resistance. *Annu.*
571 *Rev. Phytopathol.* **43**, 581-621.
- 572
573 **Kimura, M.** (1983) *The neutral theory of molecular evolution*. Cambridge.
- 574
575 **Kimura, M.** (1985) The role of compensatory neutral mutations in molecular evolution. *J.*
576 *Genet.* **64**, 7-19.
- 577
578 **Kosakovsky Pond, S. L. and Frost, S. D. W.** (2005) Datamonkey: rapid detection of
579 selective pressure on individual sites of codon alignments. *Bioinformatics* **21**, 2531-2533.
- 580
581 **Kühne, T., Shi, N., Proeseler, G., Adams, M. J. and Kanyuka, K.** (2003) The ability of a
582 bymovirus to overcome the rym4-mediated resistance in barley correlates with a codon
583 change in the VPg coding region on RNA1. *J. Gen. Virol.* **84**, 2853-2859.
- 584
585 **Kumar, S., Tamura, K. and Nei, M.** (1994) MEGA: Molecular Evolutionary Genetics
586 Analysis software for microcomputers. *Comput. Appl. Biosci.* **10**, 189-191.
- 587
588 **Laugé, R., Joosten, M. H. A. J., Haanstra, J. P. W., Goodwin, P. H., Lindhout, P. and De**
589 **Wit, P. J. G. M.** (1998) Successful search for a resistance gene in tomato targeted against a
590 virulence factor of a fungal pathogen. *Proc. Natl. Acad. Sci. USA* **95**, 9014-9018.
- 591
592 **Leach, J. E., Vera Cruz, C. M., Bai, J. and Leung, H.** (2001) Pathogen fitness penalty as a
593 predictor of durability of disease resistance genes. *Annu. Rev. Phytopathol.* **39**, 187-224.
- 594
595 **Lecoq, H., Moury, B., Desbiez, C., Palloix, A. and Pitrat, M.** (2004) Durable virus
596 resistance in plants through conventional approaches: a challenge. *Virus Res.* **100**, 31-39.
- 597
598 **Malcuit, I., Marano, M. R., Kavanagh, T. A., De Jong, W., Forsyth, A. and Baulcombe,**
599 **D. C.** (1999) The 25-kDa movement protein of PVX elicits Nb-mediated hypersensitive cell
600 death in potato. *Mol. Plant-Microbe Interact.* **12**, 536-543.
- 601

A new predictor of plant resistance durability

- 602 **Margaria, P., Ciuffo, M., Pacifico, D. and Turina, M.** (2007) Evidence that the
603 nonstructural protein of *Tomato spotted wilt virus* is the avirulence determinant in the
604 interaction with resistant pepper carrying the *Tsw* gene. *Mol. Plant-Microbe Interact.* **20**,
605 547-558.
- 606
- 607 **Martin, D. P., Williamson, C. and Posada, D.** (2005) RDP2: recombination detection and
608 analysis from sequence alignments. *Bioinformatics* **21**, 260-262.
- 609
- 610 **McCullagh, P. and Nelder, J. A.** (1989) Generalized linear models. In *Mathematical*
611 *statistics of generalized linear model* (ed. C. a. Hall). London.
- 612
- 613 **McDonald, B. A. and Linde, C.** (2002) Pathogen population genetics, evolutionary potential,
614 and durable resistance. *Annu. Rev. Phytopathol.* **40**, 349-379.
- 615
- 616 **Meshi, T., Motoyoshi, F., Adachi, A., Watanabe, Y., Takamatsu, N. and Okada, Y.** (1988)
617 Two concomitant base substitutions in the putative replicase genes of *Tobacco mosaic virus*
618 confer the ability to overcome the effects of a tomato resistance gene, Tm-1. *EMBO J.* **7**,
619 1575-1581.
- 620
- 621 **Meshi, T., Motoyoshi, F., Maeda, T., Yoshiwoka, S., Watanabe, H. and Okada, Y.** (1989)
622 Mutations in the tobacco mosaic virus 30-kD protein gene overcome Tm-2 resistance in
623 tomato. *Plant Cell* **1**, 515-522.
- 624
- 625 **Mestre, P., Brigneti, G. and Baulcombe, D. C.** (2000) An Ry-mediated resistance response
626 in potato requires the intact active site of the NIa proteinase from *Potato virus Y*. *Plant J.*
627 **23**, 653-661.
- 628
- 629 **Moury, B.** (2004) Differential selection of genes of *Cucumber mosaic virus* subgroups. *Mol.*
630 *Biol. Evol.* **21**, 1602-1611.
- 631
- 632 **Moury, B., Morel, C., Johansen, E., Guilbaud, L., Souche, S., Ayme, V., Caranta, C.,**
633 **Palloix, A. and Jacquemond, M.** (2004) Mutations in *Potato virus Y* genome-linked
634 protein determine virulence toward recessive resistances in *Capsicum annuum* and
635 *Lycopersicon hirsutum*. *Mol. Plant-Microbe Interact.* **17**, 322-329.
- 636
- 637 **Padgett, H. S., Watanabe, Y. and Beachy, R. N.** (1997) Identification of the TMV replicase
638 sequence that activates the N gene-mediated hypersensitive response. *Mol. Plant-Microbe*
639 *Interact.* **10**, 709-715.
- 640
- 641 **Sacristán, S. and García-Arenal, F.** (2008) The evolution of virulence and pathogenicity in
642 plant pathogen populations. *Mol. Plant Pathol.* **9**, 369-384.
- 643
- 644 **Santa Cruz, S. and Baulcombe, D. C.** (1995) Analysis of potato virus X coat protein genes
645 in relation to resistance conferred by the genes *Nx*, *Nb* and *Rxl* of potato. *J. Gen. Virol.* **76**,
646 2057-2061.
- 647
- 648 **Schierup, M. H. and Hein, J.** (2000) Consequences of recombination on traditional
649 phylogenetic analysis. *Genetics* **156**, 879-891.
- 650

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Molecular Plant Pathology, 2009, vol.10, no.5. DOI: 10.1111/j.1364-3703.2009.00554.x

A new predictor of plant resistance durability

- 651 **Schirmer, A., Link, D., Cognat, V., Moury, B., Beuve, M., Meunier, A., Bragard, C.,**
652 **Gilmer, D. and Lemaire, O.** (2005) Phylogenetic analysis of isolates of beet necrotic
653 yellow vein virus collected worldwide. *J. Gen. Virol.* **86**, 2897-2911.
- 654
655 **Skamnioti, P. and Ridout, C. J.** (2005) Microbial avirulence determinants: guided missiles
656 or antigenic flak? *Mol. Plant Pathol.* **6**, 551-559.
- 657
658 **Strasser, M. and Pfitzner, A. J. P.** (2007) The double-resistance-breaking *Tomato mosaic*
659 *virus* strain ToMV1-2 contains two independent single resistance-breaking domains. *Arch.*
660 *Virol.* **152**, 903-914.
- 661
662 **Suzuki, Y. and Gojobori, T.** (1999) A method for detecting positive selection at single
663 amino acid sites. *Mol. Biol. Evol.* **16**, 1315-1328.
- 664
665 **Szittyá, G. and Burgyan, J.** (2001). Cymbidium ringspot tomosvirus coat protein coding
666 sequence acts as an avirulent RNA. *J. Virol.* **75**, 2411-2420.
- 667
668 **Thompson, J. D., Higgins, D. G. and Gibson, T. J.** (1994) CLUSTAL W: improving the
669 sensitivity of progressive multiple sequence alignment through sequence weighting,
670 position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673-4680.
- 671
672 **Weber, H., Schultze, S. and Pfitzner, A. J. P.** (1993) Two amino acid substitutions in the
673 tomato mosaic virus 30-kilodalton movement protein confer the ability to overcome the
674 Tm-22 resistance gene in the tomato. *J. Virol.* **67**, 6432-6438.
- 675
676 **Wyckoff, G. J., Malcom, C. M., Vallender, E. J. and Lahn, B. T.** (2005) A highly
677 unexpected strong correlation between fixation probability of nonsynonymous mutations
678 and mutation rate. *Trends Genet.* **21**, 381-385.
- 679
680 **Yang, Z., Nielsen, R., Goldman, N. and Pedersen, A. M. K.** (2000) Codon-substitution
681 models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**, 431-449.
- 682
683 **Yang, Z. and Swanson, W. J.** (2002) Codon-substitution models to detect adaptive evolution
684 that account for heterogeneous selective pressures among site classes. *Mol. Biol. Evol.* **19**,
685 49-57.
- 686

Tables

Table 1. Dataset used for analysis of resistance durability and diversity and evolutionary constraint of avirulence factors.

Plant		Virus				Resistance durability				Avirulence factor variables			Selected references ^(f)	
Crop	Resistance gene ^(a)	Species	Genus	Avirulence factor ^(b)	Number of sequence	EP ^(c)	Timescale of protection efficiency	Status of the resistance	Virulent isolates	Resistance durability class	Nucleotide diversity ^(d)	Amino acid diversity ^(d)		ø ratio ^(e)
Pepper	<i>Tsw</i>	TSWV	<i>Tospovirus</i>	NSs	28	7	1 year	Broken-down	Common	1	0,034	0,036	0,291	Margaria <i>et al.</i> , 2007
Tomato	<i>Tm-1</i>	ToMV	<i>Tobamovirus</i>	184 kDa protein	8	6	2 years	Broken-down	Common	1	0,007	0,005	0,169	Meshi <i>et al.</i> , 1988; Strasser and Pfitzner, 2007
Tomato	<i>Tm-2</i>	ToMV	<i>Tobamovirus</i>	MP	11	6	2-3 years	Broken-down	Common	1	0,005	0,006	0,314	Meshi <i>et al.</i> , 1989; Strasser and Pfitzner, 2007
Pepper	<i>L³</i>	PMMoV	<i>Tobamovirus</i>	CP	18	6	5 years	Broken-down	Common	1	0,029	0,020	0,204	Berzal-Herranz <i>et al.</i> , 1995; Hamada <i>et al.</i> , 2007
Sugar beet	<i>Rz-1</i>	BNYVV	<i>Benyvirus</i>	P25	33	4	6-11 years	Broken-down	Common	1	0,022	0,037	0,969	Acosta-Leal and Rush, 2007; Chiba <i>et al.</i> , 2008
Soybean	<i>Rsv-1</i>	SMV	<i>Potyvirus</i>	P3-HcPro	8	7	15 years	Broken-down	Common	2	0,034	0,022	0,132	Eggenberger and Hill, 1997; Eggenberger <i>et al.</i> , 2008 Hajimorad <i>et al.</i> , 2006
Barley	<i>rym-4</i>	BaYMV	<i>Bymovirus</i>	VPg	15	4	>20 years	Broken-down	Common	2	0,053	0,050	0,258	Kühne <i>et al.</i> , 2003
Pepper	<i>pvr2¹</i>	PVY	<i>Potyvirus</i>	VPg	11	6	>20 years	Broken-down	Common	2	0,090	0,065	0,137	Moury <i>et al.</i> , 2004
Potato	<i>Nb</i>	PVX	<i>Potyvirus</i>	25 kDa protein	9	4	>20 years	Broken-down	Common	2	0,035	0,021	0,096	Malcuit <i>et al.</i> , 1999
Rapeseed	<i>TuRB01</i>	TuMV	<i>Potyvirus</i>	CI	29	9	>20 years	Broken-down	Common	2	0,138	0,030	0,029	Jenner <i>et al.</i> , 2000
Pepper	<i>pvr2²</i>	TEV	<i>Potyvirus</i>	VPg	15	7	>20 years	Broken-down	Common	2	0,068	0,050	0,194	Kang <i>et al.</i> , 2005a
Rice	<i>rym-1</i>	RYMV	<i>Sobemovirus</i>	VPg	18	4	>20 years	Broken-down	Common	2	0,076	0,031	0,075	Hébrard <i>et al.</i> , 2006
Tomato	<i>Tm-2²</i>	ToMV	<i>Tobamovirus</i>	MP	11	6	>20 years	Not broken-down	Rare	3	0,005	0,006	0,314	Weber <i>et al.</i> , 1993; Strasser and Pfitzner, 2007
Potato	<i>Nx</i>	PVX	<i>Potexvirus</i>	CP	26	4	>20 years	Not broken-down	Rare	3	0,086	0,040	0,058	Santa Cruz and Baulcombe, 1995
Potato	<i>Rx</i>	PVX	<i>Potexvirus</i>	CP	26	4	>20 years	Not broken-down	Rare	3	0,086	0,040	0,058	Bendahmane <i>et al.</i> , 1995
Pepper	<i>pvr2²</i>	PVY	<i>Potyvirus</i>	VPg	11	6	>20 years	Not broken-down	Rare	3	0,090	0,065	0,137	Moury <i>et al.</i> , 2004
Bean	<i>wlv</i>	BYMV	<i>Potyvirus</i>	VPg	10	8	>20 years	Not broken-down	Rare	3	0,160	0,109	0,100	Bruun-Rasmussen <i>et al.</i> , 2007
Potato	<i>Ry</i>	PVY	<i>Potyvirus</i>	Nia-Pro	23	7	>20 years	Not broken-down	Not reported	4	0,128	0,053	0,068	Mestre <i>et al.</i> , 2000
Tobacco	<i>N</i>	TMV	<i>Tobamovirus</i>	183 kDa protein	10	6	>20 years	Not broken-down	Not reported	4	0,025	0,011	0,087	Padgett <i>et al.</i> , 1997
Pepper	<i>L¹</i>	ToMV	<i>Tobamovirus</i>	CP	17	6	>20 years	Not broken-down	Not reported	4	0,014	0,017	0,186	Dardick <i>et al.</i> , 1999; Hamada <i>et al.</i> , 2007

BNYVV (beet necrotic yellow vein virus); SMV (soybean mosaic virus); PMMoV (pepper mild mottle virus); ToMV (tomato mosaic virus); TSWV (tomato spotted wilt virus); BaYMV (barley yellow mosaic virus); PVY (potato virus Y); PVX (potato virus X); TuMV (turnip mosaic virus); TEV (tobacco etch virus); RYMV (rice yellow mottle virus); BYMV (bean yellow mosaic virus); TMV (tobacco mosaic virus).

^a : Upper- and lower-case initials indicate dominant and recessive resistances respectively.

^b : P25 : 25 kDa protein; P3-HcPro : protein 3 and helper component protease; CP : coat protein; MP: movement protein; NSs : non structural S protein ; VPg: genome-linked viral protein; CI: cylindrical inclusion protein; Nia-Pro: nuclear inclusion a protein.

^c : Evolutionary potential overall risk defined by García-Arenal and McDonald (2003) or from bibliographic data (see supplementary materials) .

^d : Average pairwise p-distance estimated with the MEGA 3.1 software.

^e : Nonsynonymous/synonymous substitution rate ratio estimated by the SLAC method implemented in the HyPhy software.

^f : Selected references for avirulence gene identification and amino acid substitutions involved in the resistance breakdown (when these are known).

A new predictor of plant resistance durability

Table 2. Statistical analysis of the relationships between the variable “durability of resistance” and the four potential predictor variables computed with the ordinal multinomial model and with the non-parametric Kruskal-Wallis test (the underrepresented durability classes 3 and 4 were grouped for the analyses).

Variable	Ordinal multinomial model				Kruskal-Wallis test
	Odds assumption	Likelihood ratio test			P-value
	Score test ^a	Chi-Square	DF ^b	P-value ^c	
ω ratio	0.120	7.008	1	0.008 **	0.031 *
ω ratio (Avirulent isolates only)	0.372	5.488	1	0.019 *	0.110 ^{ns}
ω ratio (Two durability classes: 1 and 2+3+4)	Not applicable	8.582	1	0.003 **	0.008 **
Nucleotide diversity	0.039	4.330	1	0.037 *	0.052 ^{ns}
Amino acid diversity	0.157	2.454	1	0.118 ^{ns}	0.216 ^{ns}
Evolutionary potential (EP)	0.976	0.006	1	0.935 ^{ns}	0.995 ^{ns}
Resistance type	0.051	0.264	1	0.607 ^{ns}	0.107 ^{ns}
ω ratio + EP	0.238	7.014	2	0.028 *	Not applicable
ω ratio + Resistance type	0.015	7.277	2	0.026 *	Not applicable

^a: A non significant test ($P > 0.05$) indicates that the proportional odds assumption is valid.

^b: DF: degree(s) of freedom.

^c: P-value of the global test ($\beta=0$): a significant test supports the evidence that at least one of the covariates is correlated with the resistance durability.

^{ns}: $P > 0.05$

*: $0.01 < P < 0.05$

** : $P < 0.01$

A new predictor of plant resistance durability

Figure legends

Figure 1. Estimates of the ω ratio of virus avirulence factors according to the size of the sequence set. Five large sequence sets were used to analyse random samples of sequences of increasing size (10 random samples for each size). Arrows indicate the smallest sequence set size for which the estimate of the ω ratio is not significantly different from the estimates obtained with larger datasets.

Figure 2. Distribution into the 3 resistance durability classes of the putative predictor variables related to virus avirulence factors or to biology of viruses (durability classes 3 and 4 have been grouped). (A) ω ratio of avirulence factors, (B) nucleotide diversity of avirulence factors, (C) amino acid diversity of avirulence factors, and (D) evolutionary potential of viruses (García-Arenal and McDonald 2003). Each virus species-resistance gene combination is represented by a diamond and grey bars indicate average values.

Figure 3. Probability (P_r) that a resistance gene belongs to the durability class $r \in \{1, 2, 3+4\}$ as a function of the ω ratio of the targeted avirulence factor based on Eq.1 (see the Experimental procedures) and our dataset (Table 1).