

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20

**Luria-Delbrück Estimation of *Turnip mosaic virus***

**Mutation Rate *in vivo***

Francisca de la Iglesia,<sup>1</sup> Fernando Martínez,<sup>1</sup> Julia Hillung,<sup>1</sup> José M. Cuevas,<sup>1</sup> Philip J. Gerrish,<sup>2</sup> José-Antonio Daròs,<sup>1</sup> and Santiago F. Elena<sup>1,3,\*</sup>

*Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-UPV, 46022 València, Spain<sup>1</sup>; Center for Theoretical and Evolutionary Immunology, University of New Mexico, Albuquerque NM 87131, USA<sup>2</sup>; The Santa Fe Institute, Santa Fe NM 87501, USA<sup>3</sup>*

*Running title:* Mutation rate in TuMV

*Abstract word count:* 57

*Text word count:* 1964

\*Corresponding author. Mailing address: Instituto de Biología molecular y Celular de Plantas, CSIC-UPV, Campus UPV CPI 8E, Ingeniero Fausto Elio s/n, 46022 València, Spain. Phone: +34 963 877 895. Fax: +34 963 877 859. E-mail: santiago.elena@csic.es.

21           A potential drawback of recent antiviral therapies based on the transgenic  
22 expression of artificial microRNAs is the ease with which viruses may generate escape  
23 mutations. Using a variation of the classic Luria-Delbrück fluctuation assay, we estimated  
24 that the spontaneous mutation rate in the artificial microRNA (amiR) target of a plant  
25 virus was ca.  $6 \times 10^{-5}$  per replication event.

26

27           The rate of spontaneous mutation is a key parameter to understand the genetic structure of  
28 populations over time. Mutation represents the primary source of genetic variation on which  
29 natural selection and genetic drift operate. Although the exact value of mutation rate is  
30 important for several evolutionary theories, accurate estimates are available only for a reduced  
31 number of organisms (15). In the case of RNA viruses, mutation rates are orders of magnitude  
32 higher than those of their DNA-based hosts (7). This high mutation rates have important  
33 practical implications. For instance, for the long-term durability of vaccination strategies (6)  
34 and antiviral drugs (2), for the stability of live attenuated vaccines (26), for the eventual success  
35 of antiviral therapies based on the concept of lethal mutagenesis (1), or to determine the risk of  
36 new emerging viruses (14). The spontaneous mutation rate of a virus can be evaluated *in vivo*  
37 using a variety of experimental approaches. Among the most commonly used are: (i) estimating  
38 the frequency of mutants contained in a population generated from a single clone (7, 23), (ii)  
39 counting the number of mutant alleles accumulated in a locus which was protected against the  
40 action of purifying selection (20, 25), (iii) counting the number of lethal alleles present in a  
41 population (10), (iv) estimating the mean and variance in fitness declines among independent  
42 lineages during a mutation-accumulation experiment and then applying the Mukai-Bateman  
43 method (8), or (v) using a fluctuation assay (19). Among all, the latter is considered as the most  
44 flexible, robust and reliable method (9). The fluctuation test, originally developed by Luria and  
45 Delbrück (19), allows estimating the rate at which mutations arise in a genetic locus associated  
46 to an easy-to-score phenotype. The estimates obtained are independent of generation time and  
47 replication mode, factors that are not available for most RNA viruses. Advanced mathematical  
48 tools for the analysis of the distribution of the number of mutants across replicated cultures (the  
49 so-called Luria-Delbrück distribution) are readily available and easy to adapt to each  
50 experimental design (9).

51           The transgenic expression of 21-nt long artificial microRNAs (amiR) complementary to  
52 viral genomes has been proposed as a new antiviral strategy. Niu *et al.* (22) used the pre-  
53 miRNA159a precursor to engineer an amiR containing a sequence complementary to the RNA

54 genome of Turnip mosaic potyvirus (TuMV). Transgenic expression of this amiR in  
55 *Arabidopsis thaliana* conferred high levels of specific resistance. Similarly, a gene-silencing  
56 mechanism (RNAi) has been used in *in vitro* assays as antiviral therapeutics to inhibit the  
57 replication of several human viruses (5, 11, 16). However, a major issue of these amiR-based  
58 antiviral therapies has been the emergence of escape mutant viruses (3, 13, 17). These escape  
59 variants differ from the wild-type virus by at least one point mutation in the 21-nt target, leading  
60 to imperfect matching with the amiR. To evaluate the durability of amiR-mediated resistance in  
61 plants, Lafforgue *et al.* (17) performed an evolution experiment in which multiple independent  
62 lineages of TuMV were founded with an ancestral virus clone and allowed to evolve and  
63 diversify by serial passages in two different hosts. The first host was a wild-type *A. thaliana*  
64 and the second one the partially resistant 10-4 transgenic *A. thaliana* line that expressed amiR at  
65 subinhibitory concentrations. Periodically, the evolving populations were used to challenge the  
66 resistance of the 12-4 transgenic *A. thaliana* line, which was fully resistant to the ancestral  
67 virus. It was found that all lineages evolved in wild-type plants accumulated mutations in the  
68 amiR target and acquired the capacity to successfully infect 12-4 plants (17). The median time  
69 for lineages evolved in wild-type plants to break resistance was 14 passages, while lineages  
70 evolved in partially resistant plants only took 2 passages. The easiness to break this resistance  
71 correlated to the existence of natural variation for the 21-nt target sequence (Lafforgue *et al.*,  
72 unpublished results), suggesting that this genomic region shall not be under strong purifying  
73 selection.

74         The frequency at which mutations may be produced in the amiR target locus in evolving  
75 TuMV populations is fundamental to understand the observed dynamics of resistance breaking.  
76 Here, we report the results of a fluctuation assay experiment designed to evaluate the  
77 spontaneous mutation rate at the amiR target locus of TuMV. In this case, the phenotype  
78 associated to the mutants was the ability to replicate in the 12-4 transgenic plants expressing the  
79 antiviral amiR. We used a modification of the analytical method proposed in (12) that provides

80 improved accuracy and is especially well suited to large populations and/or high mutation rates.  
81 This method is a generalization of the statistical modeling developed by Lea and Coulson (18).

82 Fig. 1 shows a scheme of the experimental design for this fluctuation assay. A large  
83 stock of infectious sap was obtained from *Nicotiana benthamiana* plants inoculated with a  
84 plasmid containing TuMV cDNA (4, 17). This amplification step was necessary to overcome  
85 the low efficiency infecting *A. thaliana* plants with the TuMV cDNA. Sap was obtained by  
86 grinding infected tissues in a mortar with liquid N<sub>2</sub> and 20 volumes of extraction buffer (50 mM  
87 potassium phosphate pH 7.0, 3% polyethylene glycol 6000). One hundred wild-type *A.*  
88 *thaliana* plants were inoculated with 5 µL sap containing 10% Carborundum applied on three  
89 different leaves and gentle rubbing with a cotton swab (Fig. 1). After inoculation, plants were  
90 maintained in a growth chamber (16 h light 25 °C/8 h darkness 24 °C). TuMV replicated and  
91 systemically colonized the plants until reaching a population size of  $N_i$ , where the subscript  
92 denotes the  $i$ th plant. From each of these plants, virus was extracted from symptomatic tissue  
93 14 days post-inoculation (dpi), as described above. Heterogeneity in virus accumulation on  
94 leaflets of different age was minimized by pooling them into a single extraction. However, as only  
95 a fraction of the virus-infected host tissues was extracted, only a fraction  $d$  of total virus  
96 produced was obtained and successfully transmitted. This extract was divided into 10 parts,  
97 each of which was used to inoculate a 12-4 resistant plant as described above; i.e., there were 10  
98 resistant plants per each susceptible one (Fig. 1). Fourteen dpi, the number of resistant plants  
99 on which infection was successfully established was recorded. From this vector of counts,  $\mathbf{R}$ , a  
100 mutation rate was estimated using the following procedure. The number of mutants after  
101 growth and extraction from wild-type plants has a distribution whose probability generating  
102 function is (12):

103 
$$h(z) = (p - pz)^{\frac{p-pz}{1-p+pz}\mu N_i},$$

104 where  $p = d/10$  is the total dilution factor (dilution due to extraction,  $d$ , and dilution due to  
105 partitioning the extract into 10 parts),  $\mu$  is mutation rate per amiR target locus, and  $z$  denotes the  
106 argument of the generating function  $h(\cdot)$ .

107 Infection is not a deterministic process, and a single virion has probability  $q$  of infecting a  
108 plant and  $1 - q$  of not doing so. If the diluted inoculum contains  $m$  mutants, then the probability  
109 of not establishing an infection on the 12-4 resistant plants is  $(1 - q)^m$ . The number of mutants  
110 is unknown and so is treated as a random variable, and the total probability of not establishing  
111 an infection is therefore  $\sum_{m=0}^{\infty} (1 - q)^m \phi(m)$ , where  $\phi(m)$  is the  $m$ th coefficient in the  
112 expansion of  $h(z)$ . The total probability of not establishing infection is therefore  $h(1 - q)$ . Since  
113 the probability of establishing an infection is 50% when  $1 - (1 - q)^{\lambda_{0.5}} = \frac{1}{2}$ , where  $\lambda_{0.5}$  is the  
114 median infectious dose, the parameter  $q$  may thus be calculated as  $q = 1 - 2^{-1/\lambda_{0.5}}$ .

115 The log-likelihood function for  $\mu$  is thus  $l(\mu|\mathbf{R}, N_i, d, \lambda_{0.5}) = \sum_{i=1}^{1000} \log \xi(i)$ , where  $100 \times$   
116  $10 = 1000$  is the total number of resistant plants used in the fluctuation assay and

$$\xi(i) = \begin{cases} h(1 - q) & \text{if the } i\text{th plant is not infected} \\ 1 - h(1 - q) & \text{if the } i\text{th plant is infected.} \end{cases}$$

117 This function is maximized at  $\mu = \hat{\mu}$ , the maximum likelihood (ML) estimate of the mutation  
118 rate. Therefore, in addition to the vector  $\mathbf{R}$  with the counts of infected 12-4 plants for each  
119 wild-type plant, the other relevant parameters to be experimentally determined are  $N_i$  ( $i = 1, \dots,$   
120  $100$ ),  $d$  and  $\lambda_{0.5}$ .

121 First, the concentration of TuMV genomic (+) RNA strand in the original stock as well as  
122 resulting from each of the 100 wild-type *A. thaliana* (e.g.,  $N_i$ ) was measured by absolute RT-  
123 qPCR using an external standard as described in (21). In short, the standard curve was  
124 constructed using 1/5-fold dilution intervals of TuMV (+)-RNA in the range from  $1.28 \times 10^8$  to  
125  $4 \times 10^4$  molecules. Aliquots of 100 ng of total RNA were reverse transcribed in triplicate in the  
126 presence of 250 nM of primer PI (5'-TAACCCCTTAACGCCAAGTAAG-3', sequence

127 complementary to TuMV GenBank accession AF530055.2 positions 9599-9620) with M-MuLV  
128 reverse transcriptase (Fermentas) in 20  $\mu$ L reactions for 10 min at 25 °C, 45 min at 42 °C and 5  
129 min at 50 °C. Reactions were stopped by heating at 72 °C for 15 min. Sequence specific qPCRs  
130 were performed with 2  $\mu$ L of the reverse transcription products in 20  $\mu$ L final volume using the  
131 Maxima SYBR Green Master Mix reagent (Fermentas) and primers PI and PII (5'-  
132 CAATACGTGCGAGAGAAGCACAC-3', sequence homologous to TuMV positions 9448-  
133 9470) at 95 °C for 10 min followed by 40 cycles of 15 s at 95 °C and 1 min at 60 °C.  
134 Considering the total aerial plant masses (see below), measured  $N_i$  values ranged from  
135  $3.845 \times 10^{10}$  to  $3.429 \times 10^{11}$ , with an average value of  $1.226 \times 10^{11}$  TuMV (+) RNA molecules per  
136 plant (95% CI around the mean:  $1.110 \times 10^{11} - 1.341 \times 10^{11}$ ).

137 Second, the dilution factor  $d$  was approximated as the fraction of wild-type plant tissue  
138 used to generate the sap that was later used to inoculate the corresponding set of 12-4 resistant  
139 plants. On average, the aerial part of the infected wild-type plants weighted  $0.922 \pm 0.090$  g ( $\pm 1$   
140 SEM) and the average weight of the tissue ground to produce the 100 inocula was  $0.122 \pm 0.007$   
141 g, which corresponds to a dilution factor of  $d = 0.132 \pm 0.021$ .

142 Third, a dose-infectivity assay was used to evaluate  $\lambda_{0.5}$ . To do so, the original TuMV  
143 stock was serially diluted with 1/5-fold intervals in the range 1/1 – 1/500 and each dilution was  
144 used to inoculate sets of 10 plants. Twenty dpi the number of symptomatic plants was recorded.  
145 Infectivity data were subjected to a probit analysis that rendered an estimate of the median  
146 infectious dose of  $\lambda_{0.5} = 8.826 \times 10^6$  TuMV genomes per 12-4 plant (95% CI:  $4.543 \times 10^6 -$   
147  $1.779 \times 10^7$ ; goodness of fit test:  $\chi^2 = 2.694$ , 5 d.f.,  $P = 0.747$ ).

148 Finally, the fluctuation assay rendered the following results. From a total of 100 wild-  
149 type *A. thaliana* plants used as source a of TuMV inocula, only 11 contained escape mutants  
150 that produced at least one 12-4 plant infected (five cases of 1/10 and two cases of 2/10, 3/10,  
151 and 5/10). Feeding all data to the ML algorithm, the estimate of the mutation rate for the amiR  
152 target locus was  $\hat{\mu} = 5.545 \times 10^{-5}$  mutations per replication event (95% CI:  $2.886 \times 10^{-5} -$

153  $9.507 \times 10^{-5}$ ). Since the amiR target is 21-nt long, this estimate can be expressed in a more  
154 common per nucleotide scale as  $2.640 \times 10^{-6}$  s/n/r (95% CI:  $1.374 \times 10^{-6} - 4.527 \times 10^{-6}$ ). This  
155 empirical estimate is between 17 to 30 times lower than the value suggested by the simulations  
156 performed by Lafforgue *et al.* (17), which given the many assumptions behind the simulations,  
157 can be considered a reasonable discrepancy.

158 Direct estimates of mutation rates for plant RNA viruses are scarcer than for their animal  
159 and bacterial counterparts. The first estimate ever reported for a plant virus was for *Tobacco*  
160 *mosaic virus* and it was ca.  $1.8 \times 10^{-5}$  s/n/r (20). Later on, the mutation rate for *Tobacco etch*  
161 *virus* (TEV) was estimated to range between  $2.960 \times 10^{-5}$  (23) to  $4.754 \times 10^{-6}$  s/n/r (25). Our data  
162 for TuMV are in good agreement with those reported for TEV, another potyvirus. Furthermore,  
163 all these estimates are well within the range  $10^{-6}$  to  $10^{-4}$  recently reported for several animal  
164 RNA viruses and bacteriophages (24). All together, the recent estimates obtained for plant  
165 RNA viruses and the reanalyzes made of previous data (24) suggest that the mutation rate of  
166 RNA viruses may be lower than previously proposed (7).

167

168 This work was supported by grants BFU2009-06993 from the Spanish Ministerio de  
169 Ciencia e Innovación, RGP12/2008 from Human Frontier Science Program Organization and  
170 PROMETEO2010/019 from Generalitat Valenciana to S.F.E., by CSIC grant 2010TW0015 to  
171 J.A.D., and by USA National Institutes of Health grants R01GM079843-01 and ARRA  
172 PDS#35063 and by EC grant FP7231807 to P.J.G. F.M. was supported by a fellowship from  
173 Universidad Politécnica de Valencia, J.H. was supported by a fellowship from the Spanish  
174 Ministerio de Ciencia e Innovación, and J.M.C. was contracted under the CSIC JAE-Doc  
175 program.

176

177

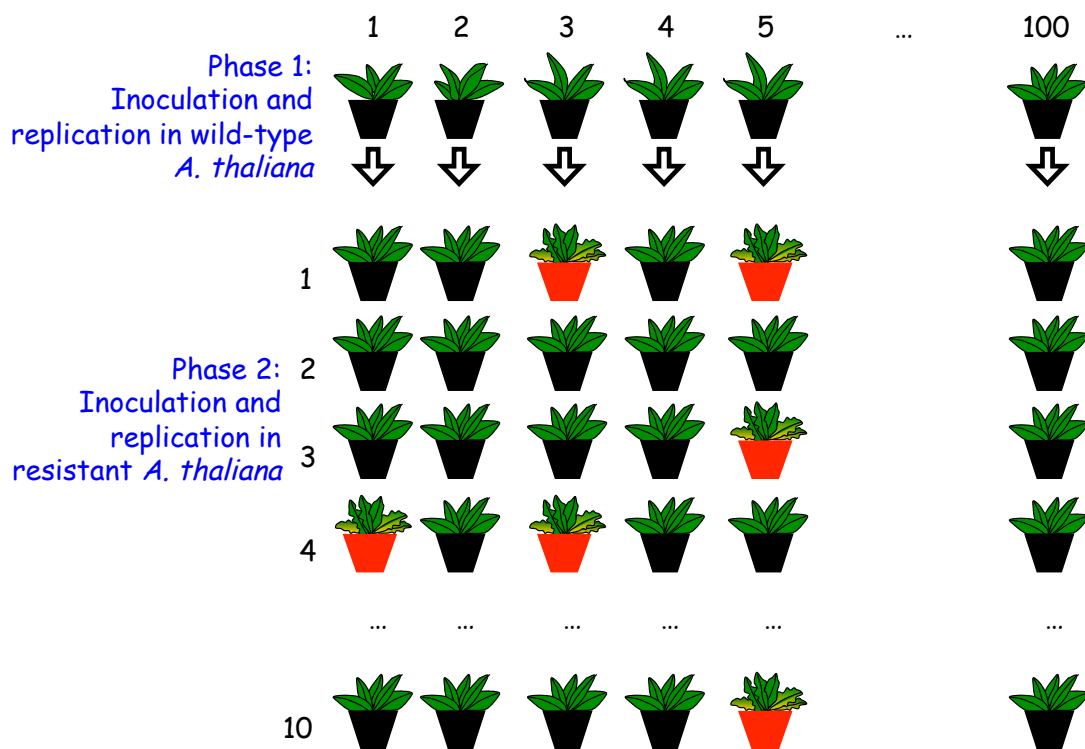
## REFERENCES



- 178 1. **Anderson, J. P., R. Daifuku, and L. A. Loeb.** 2004. Viral error catastrophe by mutagenic  
179 nucleosides. *Annu. Rev. Microbiol.* **58**:183-205.
- 180 2. **Arribas, M., L. Cabanilles, and E. Lázaro.** 2011. Identification of mutations conferring 5-  
181 azacytidine resistance in bacteriophage Q $\beta$ . *Virology* **417**:343-352
- 182 3. **Boden, D., O. Pusch, F. Lee, L. Tucker, and B. Ramratnam.** 2003. *Human*  
183 *immunodeficiency virus* type 1 escape from RNA interference. *J. Virol.* **77**:11531-11535.
- 184 4. **Chen, C. C., et al.** 2003. Identification of Turnip mosaic virus isolates causing yellow stripe  
185 and spot on calla lily. *Plant Dis.* **87**:901-905.
- 186 5. **Coburn, G. A., and B. R. Cullen.** 2002. Potent and specific inhibition of *Human*  
187 *immunodeficiency virus* type 1 replication by RNA interference. *J. Virol.* **76**:9225-9231.
- 188 6. **Davenport, M. P., L. Loh, J. Petravic, and S. J. Kent.** 2008. Rates of HIV immune escape  
189 and reversion: implications for vaccination. *Trends Microbiol.* **16**:561-566.
- 190 7. **Drake, J. W., and J. J. Holland.** 1999. Mutation rates among RNA viruses. *Proc. Natl.*  
191 *Acad. Sci. USA* **96**:13910-13913.
- 192 8. **Elena, S. F., and A. Moya.** 1999. Rate of deleterious mutation and the distribution of its  
193 effects on fitness in Vesicular stomatitis virus. *J. Evol. Biol.* **12**:1078-1088.
- 194 9. **Foster, P. L.** 2006. Methods for determining spontaneous mutation rates. *Meth. Enzymol.*  
195 **409**:195-213.
- 196 10. **Gago, S., S. F. Elena, R. Flores, and R. Sanjuán.** 2009. Extremely high mutation rate of a  
197 hammerhead viroid. *Science* **323**:1308.
- 198 11. **Ge, Q., M. T., et al.** 2003. RNA interference of *Influenza virus* production by directly  
199 targeting mRNA for degradation and indirectly inhibiting all viral RNA transcription. *Proc.*  
200 *Natl. Acad. Sci. USA* **100**:2718-2723.
- 201 12. **Gerrish, P. J.** 2008. A simple formula for obtaining markedly improved mutation rate  
202 estimates. *Genetics* **180**:1773-1778.
- 203 13. **Gitlin, L., and R. Andino.** 2005. *Poliovirus* escape from RNA interference: short interfering  
204 RNA-target recognition and implications for therapeutic approaches. *J. Virol.* **79**:1027-1035.

- 205 14. **Holmes, E. C.** 2009. The evolutionary genetics of emerging viruses. *Annu. Rev. Ecol. Evol.*  
206 *Syst.* **40**:353-372.
- 207 15. **Kondrashov, F. A., and A. S. Kondrashov.** 2010. Measurements of spontaneous rates of  
208 mutation in recent past and the near future. *Phil. Trans. R. Soc. B* **365**:1169-1176
- 209 16. **Krönke, J., et al.** 2004. Alternative approaches for efficient inhibition of *Hepatitis C virus*  
210 RNA replication by small interfering RNAs. *J. Virol.* **78**:3436-3446.
- 211 17. **Lafforgue, G., et al.** 2011. Tempo and mode of plant RNA virus escape from RNA  
212 interference-mediated resistance. *J. Virol.* **85**:9686-9695.
- 213 18. **Lea, D. E., and C. A. Coulson.** 1949. The distribution of the numbers of mutants in  
214 bacterial populations. *J. Genetics* **49**:264-285.
- 215 19. **Luria, S. E., and M. Delbrück.** 1943. Mutations of bacteria from virus sensitivity to virus  
216 resistance. *Genetics* **28**:491-511.
- 217 20. **Malpica, J. M., et al.** 2002. The rate and character of spontaneous mutation in an RNA  
218 virus. *Genetics* **162**:1505-1511.
- 219 21. **Martínez, F., J. Sardanyés, S. F. Elena, and J. A. Daròs.** 2011. Dynamics of a plant RNA  
220 virus intracellular accumulation: stamping machine vs. geometric replication. *Genetics*  
221 **188**:637-646.
- 222 22. **Niu, Q., et al.** 2006. Expression of artificial microRNAs in transgenic *Arabidopsis thaliana*  
223 confers virus resistance. *Nat. Biotech.* **24**:1420-1428.
- 224 23. **Sanjuán, R., P. Agudelo-Romero, and S. F. Elena.** 2009. Upper limit mutation rate  
225 estimation for a plant RNA virus. *Biol. Lett.* **5**:394-396.
- 226 24. **Sanjuán, R., M. R. Nebot, N. Chirico, L. M. Mansky, and R. Belshaw.** 2010. Viral  
227 mutation rates. *J. Virol.* **84**:9733-9748.
- 228 25. **Tomas, N., and S. F. Elena.** 2010. The rate and spectrum of spontaneous mutations in a  
229 plant RNA virus. *Genetics* **185**:983-989.
- 230 26. **Vignuzzi, M., E. Wendt, and R. Andino.** 2008. Engineering attenuated virus vaccines by  
231 controlling replication fidelity. *Nat. Med.* **14**:154-161.
- 232

233 **Figure 1. Schematic representation of the fluctuation assay.** A TuMV stock was produced  
 234 from infected *N. benthamiana* plants that were previously inoculated with a TuMV clone. This  
 235 stock was used to mechanically inoculate 100 wild-type *A. thaliana* plants (phase 1). During  
 236 this phase, erroneous viral replication produces spontaneous mutants in the amiR 21-nt target  
 237 that accumulate in the population. Fourteen dpi, virus was purified from each of these plants  
 238 and used to inoculate batches of 10 *A. thaliana* 12-4 plants expressing the antiviral amiR (phase  
 239 2). During this phase, only those genomes carrying a mutation in the 21-nt target would  
 240 eventually escape from the RNA silencing. Plants inoculated with these mutants will develop  
 241 symptoms whereas plants inoculated with the wild-type TuMV will not. The number of  
 242 infected 12-4 plants showing symptoms of infection was recorded 14 dpi (red pots).



243