Received: 8 May 2020 | Accepted: 22 July 2020

DOI: 10.1111/ppa.13258

ORIGINAL ARTICLE

Plant Pathology | WILEY

Molecular evolution of tomato black ring virus and de novo generation of a new type of defective RNAs during long-term passaging in different hosts

Daria Budzyńska[1] | Julia Minicka[1] | Beata Hasiów-Jaroszewska[1] | Santiago F. Elena[2,3]

[1]Department of Virology and Bacteriology, Institute of Plant Protection – National Research Institute, Poznań, Poland

[2]Instituto de Biología Integrativa de Sistemas, CSIC-Universitat de València, València, Spain

[3]The Santa Fe Institute, Santa Fe, NM, USA

Correspondence
Beata Hasiów-Jaroszewska, Department of Virology and Bacteriology, Institute of Plant Protection – National Research Institute, Wł. Węgorka 20, 60-318 Poznań, Poland.
Email: b.hasiow@iorpib.poznan.pl

Funding information
Spain Ministerio de Ciencia e Innovación - FEDER, Grant/Award Number: SPID201900X03998IV1; Narodowe Centrum Nauki, Grant/Award Number: 2015/17/B/NZ8/02407

Abstract

Tomato black ring virus (TBRV) is a worldwide-distributed RNA virus infecting a wide range of different host plants, including crop species, trees, shrubs, and weeds. Here, we investigated the molecular evolution of TBRV and its adaptability to different plant species. The TBRV-Pi isolate was used to generate five independent evolution lineages serially passaged in either quinoa, tobacco, or tomato plants. After 15 passages, the genetic variability present in all the lineages was characterized for the movement (*MP*) and coat (*CP*) coding cistrons. We addressed two main questions: to what extent does the amount of genetic variability in the TBRV genome depend on the host species, and are there host species-specific adaptive mutations? Overall, 201 different nucleotide substitutions emerged during the evolution experiment, with some of them appearing multiple times in different lineages; two of them (one in *CP* and one in *MP*) were unique for a particular host plant. We have shown that the degree of genetic differentiation depends on the host species in which the virus evolved, and that positive selection is operating upon certain residues, particularly in *CP*. Moreover, we have characterized new types of defective RNAs that arose during the TBRV-Pi evolution in tobacco. Furthermore, this is the first report of a defective RNA from the RNA2 of TBRV.

KEYWORDS

defective particles, experimental evolution, host adaptation, TBRV, virus evolution

1 | INTRODUCTION

RNA viruses have a great ability to generate genetic diversity. There are three major factors affecting this property: (a) a much shorter generation time relative to their hosts; (b) very large population sizes; and (c) a high amount of nucleotide misincorporations in RNA virus genomes as a consequence of the low fidelity of RNA-dependent RNA polymerase, which lacks a proofreading mechanism (Sanjuán *et al*., 2010). The analysis of genetic polymorphism can provide some interesting information concerning the evolution both of a single gene and of entire populations. In many cases, the virus within a single host is essentially a mixture of genetic variants that can be characterized by a consensus sequence and evolves as a single unit, the viral quasispecies. The distribution and frequency of genetic variants in populations is regulated by the joint action of genetic drift, natural selection, and migratory events, and may change over time. The ability of viruses to colonize new ecological niches ultimately depends on distribution of these genetic variants (Elena *et al*., 2014).

Tomato black ring virus is a representative of subgroup B in the *Nepovirus* genus (family *Secoviridae*) and is widely distributed in regions with moderate climate conditions. Tomato black ring virus (TBRV) was discovered for the first time in tomato (Smith, 1946) and since then has been found in a wide range of economically important crops (i.e., potatoes, cucurbitaceous species), ornamental plants (phloxes), as well as trees and shrubs (Lister, 1960; Pospieszny and Borodynko, 1999; Šneideris and Staniulis, 2014). Recently, TBRV has been detected in numerous black locust trees in Poland (Budzyńska *et al*., 2020). TBRV causes severe symptoms on infected plants, such as necrosis or mosaic on leaves, with the most characteristic symptoms being necrotic rings on leaves and fruits. The virus can be transmitted by longidorid nematodes (Harrison *et al*., 1961). Moreover, the spread of the virus occurs by seed or pollen, which promotes the dissemination of the virus over long distances (Sanfaçon *et al*., 2012; Pospieszny *et al*., 2020).

TBRV, as well as other nepoviruses, has a bipartite positive-sense RNA genome consisting of two single-stranded RNA segments: RNA1 and RNA2 (c. 7,400 and 4,600 nucleotides [nt] long, respectively). RNA1-encoded polyprotein p1 (254 kDa) is processed into peptides responsible for virus replication and genome expression: the protease cofactor Pro-cof, the nucleotide triphosphate-binding protein NTB, the viral genome-linked protein VPg, the cysteine-like protease Pro, and the RNA-dependent RNA polymerase RdRp. RNA2-encoded polyprotein p2 (149 kDa) is involved in virion formation and virus movement and is processed into the homing protein HP, the movement protein MP, and the coat protein CP. Furthermore, previous research has indicated that TBRV isolates may be accompanied by additional, subgenomic elements such as satellite RNAs (satRNAs) and defective RNAs (D RNAs). SatRNAs are around 1,000–1,300 nt long and do not share any significant identity with the genome of the virus with which they are associated (Oncino *et al*., 1995). D RNAs are shorter than satRNAs, do not encode proteins, and result from single or multiple deletions in the parental virus genome (Pathak and Nagy, 2009). Until now, two different types of TBRV-derived D RNAs have been described, both of which arise from RNA1. The first type contains sequences from the 5′ and 3′ untranslated regions (UTR) and from the 5′ region of the single open reading frame, whereas the second type consists of a portion of the coding region for the RdRp flanked by a short fragment of the 5′ UTR and almost the entire 3′ UTR. Previous studies have shown that these small molecules reduce the accumulation of the virus and affect symptoms on infected plants; due to this fact, they were described as defective interfering RNAs (DI RNAs; Hasiów-Jaroszewska *et al*., 2018).

Despite numerous studies on TBRV biology and epidemiology, the molecular evolution of TBRV has not been analysed yet. Herein, we have investigated the impact of the long-term passaging of the TBRV-Pi tomato isolate through different hosts on the genetic variability and population structure of the virus. The point mutations generated during prolonged passaging were identified and their possible adaptive value to the corresponding local host species was analysed. Moreover, we conducted molecular characterization of the new types of D RNAs generated de novo during the prolonged passaging of TBRV-Pi through tobacco.

2 | MATERIALS AND METHODS

2.1 | Virus origin, particle purification, and viral RNA isolation

In the present study, the TBRV-Pi isolate originally collected from greenhouse tomato in Poland was used. Fragments of tomato leaves with visible symptoms of infection were ground with a 0.05 M phosphate buffer (8.7 g/L of KH_2PO_4; pH 7.2) at a ratio of 1:5 (wt/vol) and then carborundum-dusted quinoa (*Chenopodium quinoa*) leaves were inoculated with the obtained tomato sap by rubbing. Mock plants were inoculated only with buffer. Plants were grown in greenhouse conditions at a temperature of 22–23°C and with a photoperiod of 16 hr light:8 hr dark. After the appearance of systemic infection symptoms, the plant material was harvested and a purified virus preparation was obtained in a sucrose gradient as described previously (Rymelska *et al*., 2013). Viral RNA was extracted using the phenol-chloroform procedure (Green and Sambrook, 2012) and the RNA profile was analysed by electrophoresis in 1% agarose gel. The concentration of RNA was measured in a ND-1000 spectrophotometer (Thermo Fisher Scientific) and adjusted to 1 μg/μl.

2.2 | Experimental evolution

The long-term passaging of the TBRV-Pi isolate was performed in quinoa, tomato (*Solanum lycopersicum* 'Moneymaker'), and tobacco (*Nicotiana tabacum* 'Xanthi'). Five individual plants of each host were inoculated with the TBRV-Pi isolate and used as founders of the 15 independent evolutionary lineages. Hereafter, lineages will be referred to as C1–C5, S1–S5, and N1–N5, according to the corresponding local host quinoa, tomato, and tobacco, respectively. Carborundum-dusted leaves were inoculated with the previously isolated and diluted viral RNA (c.3.3 μl RNA per leaf) and infected plants were maintained in greenhouse conditions as above. Seven days postinoculation (dpi), 500 mg leaves with symptoms were collected and used for mechanical inoculation of a new set of plants as described above. This serial passage protocol was repeated up to 15 times, over a total period of 105 days. The plants were observed for development of symptoms of viral infection throughout the entire experiment.

2.3 | Total RNA isolation, *CP* and *MP* coding regions amplification and sequencing

After the 15th passage, total RNAs from all evolutionary lineages were isolated using the RNeasy Plant Mini Kit (Qiagen) according

to the manufacturer's protocol. cDNAs were synthesized using 1 μl of these total RNAs and from the viral RNA used as inoculum for the first passage at a final concentration of 1 μg/μl, oligo (dT) primer (200 nM) using the Transcriptor High Fidelity cDNA Synthesis Kit (Roche) according to the manufacturer's protocol. The *CP* and *MP* cistrons were amplified using the Expand Long Range dNTP Pack (Roche). Two pairs of specific primers were used: those previously described for *CP* amplification (Šneiderіs and Staniulis, 2014), and those designed in this study MPF 5′-ACCTCCAGTTCTCAAATAG-3′ and MPR 5′-CCACAGGCAAAATCCCC-3′ primers for the amplification of the *MP* sequence. The obtained PCR products were separated in 1% agarose gel and analysed in terms of the appropriate size. Subsequently, the products were ligated into pCR4-TOPO vector (Invitrogen) and the recombinant plasmids were introduced into One Shot TOP10 chemically competent *Escherichia coli* cells (Invitrogen). For each lineage, 10 independent plasmid DNAs were sequenced in both directions using M13F and M13R primers by Genomed (Warsaw, Poland). The obtained *CP* and *MP* sequences were compiled and analysed in BioEdit (Hall, 1999).

2.4 | Molecular population genetics and analysis of the selection

Basic parameters describing the type and frequency of mutations were estimated using MEGA X (Kumar *et al*., 2018), separately for each cistron. For identification of mutations, each individual sequence was treated as a unique haplotype and compared with the ancestral TBRV-Pi *CP* or *MP* sequences. Multiple sequence codon alignments were generated using the MUSCLE program as implemented in MEGA X (Files S1 and S2). The best-fitting models of nucleotide substitutions, based in the smallest Bayesian information criterion, were T92 (Tamura, 1992) for *CP* and K2P (Kimura, 1980) for *MP*. The transition to transversion rates ratio (κ) was obtained using the maximum composite likelihood model (separately for each lineage).

Different nucleotide diversity (π) indices were calculated with Arlequin v. 3.5.2.2 (Excoffier and Lischer, 2010) for *MP* and *CP* separately. Genetic differentiation between lineages evolving in the same or different hosts, as well as among host species, was evaluated using the analysis of molecular variance (AMOVA) approach implemented in Arlequin. Wright's *F* fixation indices (Wright, 1950) were used to assess the fraction of the total observed genetic variability attributable to differences between host species (F_{CT}), between lineages evolved in the same host species (F_{SC}), and among clones within lineages (F_{ST}).

To identify whether and when cistrons underwent natural selection during our evolution experiments, Tajima's *D* tests of neutrality (Tajima, 1989) were performed as implemented in Arlequin. Furthermore, to detect codon sites in *MP* and *CP* that underwent positive selection, codon-by-codon analyses were conducted using the HyPhy package (Kosakovsky Pond *et al*., 2005), as implemented in MEGA 7 (Kumar *et al*., 2016). Synonymous (d_S) and nonsynonymous (d_N) substitution rates per synonymous and nonsynonymous sites were estimated using the modified Nei and Gojobori (1986) method. The normalized difference between d_N and d_S was used to evaluate the intensity and sign of the selective pressure operating on particular codons. Finally, residue covariation analysis for each gene was performed using the algorithms implemented in the CAPS2 server (caps.tcd.ie; Fares and McNally, 2006).

2.5 | Amplification, cloning, and sequencing of D RNAs

After the 15th passage, purified virus preparations were obtained (Rymelska *et al*., 2013). Viral RNAs were extracted using the phenol-chloroform procedure (Green and Sambrook, 2012) and the RNA profiles were analysed in 1% agarose gel. Molecules shorter than TBRV-genomic RNAs were extracted from the gel (Zhang and Papakonstantinou, 2009).

The cDNA was obtained using 1 μl of extracted RNA, a specific reverse primer hybridizing to the 3′ UTR of both RNA segments (Hasiów-Jaroszewska *et al*., 2012; Rymelska *et al*., 2013) and the Transcriptor High Fidelity cDNA Synthesis Kit (Roche), according to the manufacturer's protocol. Small RNAs were amplified with the Expand Long Template PCR System (Roche) and primers complementary to the conserved 5′ and 3′ UTRs of genomic RNA segments (Hasiów-Jaroszewska *et al*., 2012; Rymelska *et al*., 2013). The PCR products were cloned and the plasmid DNAs were purified and sequenced as described above. The obtained sequences of D RNAs were compared with 5′ and 3′ fragments of TBRV-Pi genomic RNAs (GenBank accessions MG458221.1, MG458220.1) and sequences of other DI RNAs (including those already available in GenBank) using BioEdit (Hall, 1999) and MEGA X. The sequences of new D RNAs obtained during the experiment were deposited in GenBank under the accession numbers MT439848, MT439849, and MT439850.

In order to investigate whether the obtained particles were not artefacts and replicated together with genomic particles, quinoa plants were reinoculated with viral RNAs from lineages N1 and N2 as described in Section 2.2. At 10dpi, purified virus preparations were obtained and viral RNAs were extracted. D RNAs were amplified, cloned, and then sequenced as described above.

3 | RESULTS

3.1 | Effect of experimental evolution on symptom development

During the passaging on infected plants, symptoms common for the TBRV infection were observed. They occurred in the form of necrotic rings on tobacco, chlorotic spots, mosaic, and leaf malformation on quinoa, or necrosis on tomato (Figure 1); in most of the evolutionary lineages, no significant changes were observed in the symptoms during the passaging. Lineages N2 and N5, which evolved

FIGURE 1 Symptoms of infection caused by TBRV-Pi isolate on the three plant species used in this study: (a) tomato, (b) quinoa, and (c) tobacco [Colour figure can be viewed at wileyonlinelibrary.com]

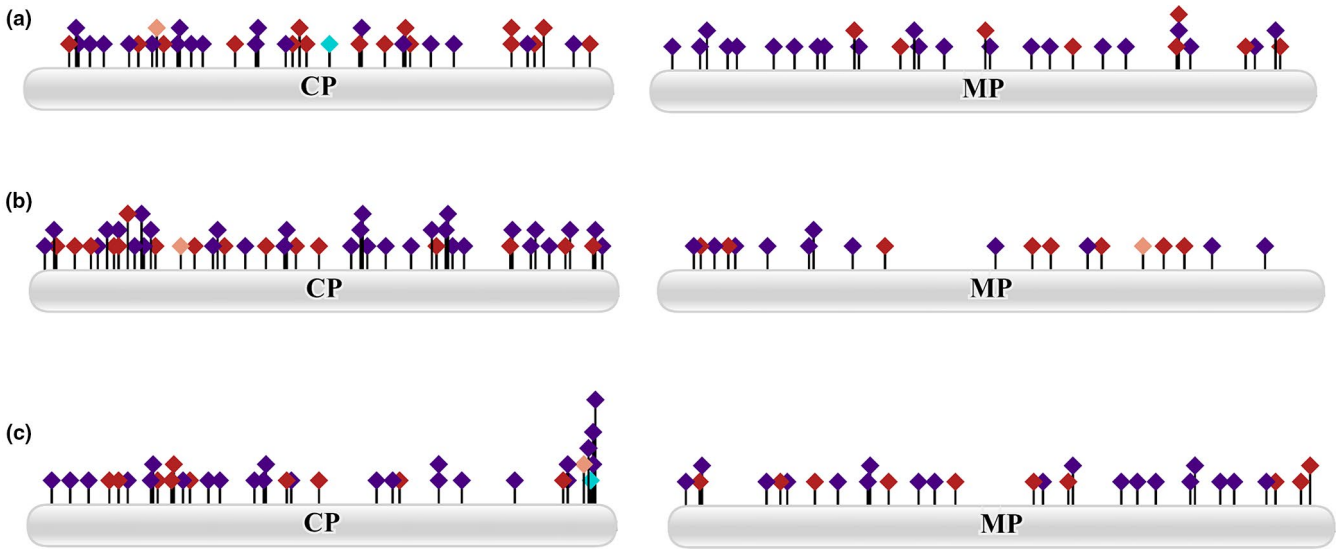

FIGURE 2 Location of particular mutations found in lineages evolved in (a) quinoa, (b) tomato, and (c) tobacco. The violet diamonds represent nonsynonymous mutation; red, synonymous; blue, two different nonsynonymous mutations occurring in the same codon; pink, occurrence of both synonymous and nonsynonymous mutation in the same codon. This figure was prepared using Pfam website (http://pfam.xfam.org/generate_graphic) [Colour figure can be viewed at wileyonlinelibrary.com]

in tobacco, were the only exception. For these two lineages, after the 10th passage, the necrotic mosaic on infected plants was stronger, covered most of the plant, and appeared earlier than in the other three evolutionary lineages.

3.2 | Number and nature of mutations arising during evolution in different host species

The resulting sequences obtained after the evolutionary experiment were 1,533 and 1,734 nt long for the *CP* and *MP*, respectively. The comparison of the 150 derived sequences per cistron with their corresponding ancestral sequences revealed the presence of 201 different mutations in both genomic regions, with 132 being nonsynonymous (65.67%; Figure 2). No insertion/deletion was found. For the *CP*, 80 out of the 123 different mutations were nonsynonymous (65.0%). Similarly for the *MP*, 52 out of the 78 different mutations were nonsynonymous (66.7%; Figure 2). Considering the number of mutations per individual lineage, it was found that for the *MP*, the lowest number of polymorphic sites occurred in lineages C4, S2, and S4, with three mutations each. In contrast, the largest number of mutations was found for lineage C1, with 10 mutations. For the *CP*, the lowest number of mutations was found for lineage C3, with six mutations, whereas the largest number of mutations was in lineage S4, with 15 mutations (Tables S1–S3).

Interestingly, many of these mutations were observed in multiple lineages as convergent mutational events. To further analyse all the mutational events observed, the counts shown in Tables S2–S4 were fitted to a generalized linear model (GLM) with a Poisson distribution of counts and a log-link function. In this model, the host species (quinoa, tobacco, and tomato), genomic region (*MP* and *CP*), and the type of mutation (synonymous and nonsynonymous) were treated as orthogonal factors, whereas lineage was nested within host species. The statistical significance of each factor, as well as of their interactions, was assessed by means of likelihood ratio tests (LRT)

that follow a χ^2 probability function. The results are summarized in Table S5.

Several interesting observations can be drawn based on the nucleotide sequences of the two analysed cistrons (Files S1 and S2). First, though no overall difference exists between host species, lineages within each host showed a significant differentiation in the number of fixed mutations. Secondly, overall, the two cistrons significantly differed in the number of mutations that fixed across lineages. On average, *CP* fixed 6.368 ± 0.555 (± *SEM*) mutations per lineage, while *MP* fixed 69.5% fewer mutations (1.940 ± 0.215) per lineage. Thirdly, this difference significantly depends on the host species in which TBRV-Pi lineages were evolved. The lowest number of mutations observed both for *CP* and *MP* were for the lineages evolved in tomato (average number of mutations 4.919 ± 0.818 and 0.449 ± 0.840, respectively). The largest number of mutations observed for *CP* was for the lineages evolved in tobacco (10.505 ± 1.237), while the largest observed for *MP* was for the lineages evolved in quinoa (4.743 ± 0.854). Fourthly, the type of mutations fixed across lineages also depended on the host species in which they were evolved. Synonymous mutations were more abundant in lineages evolved in tomato (6.836 ± 1.070) while nonsynonymous mutations were more common in lineages evolved in tobacco (7.969 ± 1.163). Fifthly, the frequency of synonymous and nonsynonymous mutations also significantly differed between the two cistrons. Synonymous mutations were less common in *CP* (4.381 ± 0.649) than in *MP* (6.162 ± 0.860), while the observation was reversed for nonsynonymous mutations (4.381 ± 0.649 for *CP* and 9.257 ± 0.849 for *MP*). Finally, this latter effect was significantly dependent on the host species in which lineages were evolved: the lowest number of synonymous mutations was observed for the *CP* of lineages evolved in tomato (3.277 ± 0.947) while the largest amount of synonymous mutations was observed for the *MP* of lineages evolved in tomato (14.260 ± 1.719). Likewise, the largest amount of nonsynonymous mutations was observed for the *CP* of lineages evolved in tobacco (18.323 ± 1.936), while the lowest number of nonsynonymous mutations was observed for the *MP* of lineages evolved in tomato (0.014 ± 0.005).

Most of the resulting mutations were singletons—sites at which only one sequence has a different nucleotide while all other sequences are identical; 88 (72.73%) in *CP* and 62 (80.52%) in *MP* (Tables S2–S4). However, some mutations were pervasively observed among different lineages. The first case of such convergent mutation was the synonymous transition U1371C in the *MP* gene, observed in all 50 clones generated from the five lineages evolved in tomato and in the 20 clones from lineages C3 and C4 (Tables S2 and S4). The second case of convergent fixed mutation was identified for lineages evolved in tobacco: nonsynonymous transition U1070C that results in amino acid substitution I357T of the CP, which involves replacing a hydrophobic side chain by a polar uncharged one. This mutation was exclusive to all the lineages evolved in tobacco. Two more mutations got fixed in the *CP* gene of specific lineages evolved in tobacco: synonymous transition U1395C was present in all 10 clones sequenced from lineage N1, and nonsynonymous transition G1404A (giving rise to amino acid replacement M468I, which maintains a hydrophobic side chain) was found in all 10 clones from lineage N2. In addition, two more synonymous transitions were observed in 9 out of the 10 clones sequenced from lineages N3 (U1671C) and N4 (U744C; Tables S3 and S4).

The above analyses did not show a clear association between the enhancement of infection symptoms observed in lineages N2 and N5 and the number and nature of mutations fixed in these lineages compared to the other three lineages. The two fixed mutations shared by all five lineages may explain some common adaptation to the host, but not differences in symptomatology among lineages. So far, only the amino acid replacement M468I in the CP of lineage N2 may be an interesting candidate for a host-specific adaptive mutation.

Furthermore, for each lineage, the transition to transversion rates ratio (κ) was computed (Table S1). The estimated standard deviations are larger than the actual κ values for many lineages and both genes. Indeed, a significant negative correlation exists between both statistics (partial correlation coefficient controlling for gene: $r_p = -0.598$, $df = 27$, $p = .001$), thus suggesting that the smaller estimates of κ are inherently noisier. Keeping this source of uncertainty in mind, it is still possible to infer some interesting conclusions from these computations. In all cases, $\kappa > 1$, though with a wide range of estimates. For the *CP* cistron, eight lineages had $\kappa \leq 9.567$, while the other seven had $\kappa \geq 360.085$. For the *MP*, only three lineages showed $\kappa \leq 14.009$, while all other showed values $\kappa \geq 235.524$. Indeed, the difference between κ values estimated for *MP* and *CP* was significant (Wilcoxon signed ranks test, $p = .005$), indicating that the ratio between transition and transversion rates was more strongly biased towards the first type of mutation in *MP* than in *CP*.

3.3 | Genetic differentiation between hosts and lineages

The amount of genetic differentiation generated among the five independent lineages evolving in the same host species was assessed, as well as among those evolving in alternative host species. Table S6 shows the estimated values of π for each lineage and both cistrons. In the case of *CP*, π values ranged between 0.784×10^{-3} (lineage C3) and 2.165×10^{-3} (lineage S4), with a mean of $1.284 \times 10^{-3} \pm 0.103 \times 10^{-3}$, while in the case of *MP* values ranged between 0.231×10^{-3} (lineages C4, S2, and S4) and 1.734×10^{-3} (lineage C5), with a mean of $0.721 \times 10^{-3} \pm 0.107 \times 10^{-3}$. The 1.781-fold larger π for *CP* compared to *MP* was significantly different (Wilcoxon signed ranks test, $p = .003$).

Next, we sought to test for genetic differences in the three hierarchical levels of our experimental design. To this end, we used the Wright's fixation *F* indices estimated using the AMOVA approach. Table 1 shows the summary of the AMOVA run for both genes. In this case, results are highly consistent for the two cistrons. Approximately 56% of all observed genetic variability was explained by nucleotide diversity within the lineages (i.e., differences among the 10 clones sequenced for each lineage; measured by F_{ST}), and

TABLE 1 Results of the population structure analyses

Gene	Source of variation	*df*	Sum of squares	Variance component	Explained variation (%)	Wright's *F*	*p*
CP	Among hosts	2	42.833	0.336	19.98	$F_{CT} = 0.415$	<.0001
	Among lineages	12	55.273	0.362	21.53	$F_{SC} = 0.269$	<.0001
	Within lineages	135	132.847	0.984	58.49	$F_{ST} = 0.415$	<.0001
	Total	149	230.953	1.682			
MP	Among hosts	2	31.523	0.243	20.77	$F_{CT} = 0.208$	<.0001
	Among lineages	12	43.555	0.300	25.71	$F_{SC} = 0.325$	<.0001
	Within lineages	135	84.428	0.625	53.52	$F_{ST} = 0.465$	<.0001
	Total	149					

was the most significant effect. These genetic differences reflect the highly diverse nature of the quasispecies inside each host plant. Net differences among lineages evolved in the same host species explained approximately 24% of the total observed variability (quantified by F_{SC}), which suggests that during experimental evolution, despite convergent evolution at particular sites, different alleles arose and shifted to fixation on different lineages. Finally, approximately 20% of the total observed π was explained by differences among the three different host species (quantified by F_{CT}), indicating that some alleles were exclusive of each particular host species. The latter finding may be a fingerprint of host-specific local adaption.

3.4 | Detection of natural selection

To investigate whether natural selection could affect the population genetic structure described in the previous section, we took three different approaches. First, a population-genetics approach based in the Tajima's *D* test of neutrality was conducted for each lineage (Table S6). In the case of *CP*, lineages C3, C4, N4, N5, S1, S2, and S4 had significant $D < 0$ values. In the case of *MP*, lineages C2, N1, N5, S1, S2, S4, and S5 also had significant $D < 0$ values. $D < 0$ values indicate that nucleotide diversity is lower than expected given the number of segregating sites. This situation is compatible with the action of purifying selection, the presence of slightly deleterious mutations segregating in the populations, or fast population expansions. However, after applying a more stringent false discovery rate correction for multiple tests of the same null hypothesis (Benjamini and Hochberg, 1995) for each cistron independently, none of them remained significant. Consequently, being conservative, we should reject purifying selection as the dominant force driving the maintenance of genetic diversity observed within the evolved lineages. However, in an expanding population, many new mutations may be segregating and will be observed in the data as singletons. As we already mentioned above, over 70% of all observed mutations were singletons, thus giving support to this latter possibility.

The observation of convergent mutations associated to a particular host species described above strongly suggests the action of positive selection. This prompted us to take a second approach based on a codon-by-codon analysis of positive selection using the normalized $d_N - d_S$. In the case of the *CP*, the null hypothesis of neutral evolution could be rejected in seven codons with $d_N > d_S$ values. These sites under positive selection were: M45V ($p = .003$) and M91V ($p = .021$) in lineage C4, W110R ($p = .003$) in lineage C1, M316I ($p < .001$) in lineage N4, M349T ($p = .004$) in lineage S1, M360I ($p < .001$) in lineage S4, and the substitution M468I ($p = .002$) found in lineage N2 already mentioned in Section 3.2. However, convergent mutation I357T described in Section 3.2 failed to reject the null hypothesis of neutral evolution ($p = .543$). Interestingly, six out of these seven amino acid replacements involve substituting methionine by valine (two cases in linage C4), isoleucine (lineages N2, N4, and S4), or threonine (lineage S1). In principle, none of these replacements may induce strong conformational changes, though a possible mechanism for this abundance of methionine replacements is presented in Section 4. However, replacing tryptophan by arginine (lineage C1) will necessarily induce a strong perturbation in the CP folding, by the fact of replacing a hydrophobic side chain by a long and positively charged one.

In the case of *MP*, only amino acid replacement W28R, observed in lineage N1, had a large enough $d_N - d_S > 0$ value as to reject the neutrality hypothesis ($p = .016$).

Finally, the third approach to detect the effect of positive selection (based on identifying groups of coevolving amino acids) allowed us to identify one group of three coevolving sites in the CP of lineage S3. These sites were L13P, C130R, and D278G (pairwise correlations between sites $r \geq .925$ and bootstrap support $p \geq .954$). Likewise, a group of three coevolving amino acids (L66P, I150V, and K228R) were observed in the MP of lineage C5 ($r \geq .185$, $p \geq .809$). These triplets of amino acids should be considered as targets of selection as a whole.

3.5 | Structure and origin of D RNAs

Due to the changes in the symptoms of infection observed in lineages N2 and N5 during the course of evolution in tobacco, viral RNAs were purified from the five lineages at the end of the experiment. Subsequently, RNA profiles were compared with the RNA profile of the ancestral TBRV-Pi isolate (not containing defective particles). In each lineage, the analysis revealed the presence of the

two genomic segments RNA1 and RNA2, and additional small RNAs of approximately 500 nt. The structure analysis of additional RNAs revealed that the approximately 500 nt long RNAs obtained for N3, N4, and N5 lineages contained an initial fragment of the RNA2 molecule, and probably arose as degradation products of RNA2. Thus, we discarded them from further studies. In the case of lineages N1 and N2, three types of RNA molecules of 570, 480, and 335 nt in length were identified. The first two types were isolated from lineage N1, whereas the third one was from lineage N2. The obtained molecules were designated as TBRV-Pi D RNA N1.1, TBRV-Pi D RNA N1.2, and TBRV-Pi D RNA N2, respectively.

The comparison of obtained D RNAs with the ancestral TBRV-Pi isolate revealed that the origin of the three D RNAs is compatible with a single deletion event; however, the junction site and the parental genomic RNA are different. The analyses revealed that D RNAs N1.1 and N1.2 have similar sequence structure and are derived from RNA1 as a result of a single deletion. Despite the fact that both are composed of a 5′ UTR fragment, a portion of the C-terminal end of RdRp, and almost the entire 3′ UTR (Figure 3), the site of recombination shift for each of these D RNAs is different. The deletion occurred between sites 185 and 6,830 for D RNA N1.1 and between sites 206 and 6,941 for D RNA N1.2. Consequently, D RNA N1.1 is 570 nt long and D RNA N1.2 is substantially shorter, at 480 nt long. The sequence structure of D RNA N2, 335 nt in length, unlike the other TBRV-associated defective molecules described to date, arose as a result of a single deletion between sites 318 and 4,488 of RNA2, being composed of a 5′ UTR fragment followed by the first nucleotide of HP (191 nt) and a part of the 3′ UTR (144 nt; Figure 3).

Based on their structure, D RNAs N1.1 and N1.2 were classified into the second type of TBRV D RNAs (sharing a similar structure with different junction sites; Hasiów-Jaroszewska *et al*., 2012, 2018; Rymelska *et al*., 2013), whereas the structure of D RNA N2 has not been described in the literature previously.

An analysis of the nucleotide sequences near the junction sites of D RNAs N1.1 and N1.2 revealed the presence of two sequence motifs, 6820-CUUCAGUA and 6808-GAUGGGAAGGUGCU, within the RdRp-coding region, which seem to be a complementary inverted repeat of the 5′ UTR sequences 176-UACUGAAG and 218-AGCGCCUACCAAUC, respectively, suggesting the formation of a hairpin structure contributing to the generation of these D RNAs. In the case of D RNA N2 the specific motifs near junction sites were also determined; the sequence motif 4432-AAAUAACUC in the 3′ UTR of RNA2 seems to be a complementary inverted repeat of the 330-GAGUUGUUU sequence within the 5′ end of the HP-coding sequence.

The comparison of the obtained D RNAs with others available in GenBank (accessions MG458222.1, JN133850.1, JN133847.1, AY772163.1, KC839482.1, KC839480.1, KC839481.1, and MG458223.1) revealed that, despite the differences in the structure caused by a different location of junction sites, there are fragments common for all D RNAs associated with TBRV. The comparison of small defective RNAs with the corresponding regions of the TBRV-Pi genomic sequences revealed 99.2% sequence identity for D RNA N1.1, 99.5% for D RNA N1.2, and 98.5% for D RNA N2.

Finally, additional analyses confirmed that the three described D RNAs replicate together with genomic particles. The sequences of D RNA obtained from original viral preparations (after passage 15) and viral preparations after the reinoculation of quinoa with previously obtained viral RNA from lineages N1 and N2, were identical.

4 | DISCUSSION

RNA viruses are described as the organisms with the highest evolutionary potential, due to their large population size, fast replication, and high mutation rates (Domingo *et al*., 2012). The error-prone viral RdRp has an error frequency of approximately one error per genome per replication cycle (Sanjuán *et al*., 2010). It has been suggested that the evolution of large diverse populations, like viral ones, is strongly affected by selection of linked mutations in the same haplotype, as well as by the interactions between different haplotypes in the population (Domingo *et al*., 2012). In viral populations, different forms of selection may be operating: positive, increasing the frequency of variants that are the fittest in a certain environment, and negative or purifying, removing less-fit variants (Domingo *et al*., 2012).

The research on molecular evolution of the TBRV-Pi isolate described in this paper allowed us to answer two main questions: (a) how does the host species affect the accumulation of mutations in the studied regions of TBRV-Pi genome?; and (b) are there some beneficial mutations responsible for host-specific adaptations?

Our results suggest that the evolved populations of TBRV-Pi were heterogeneous and that their genetic diversity may have depended

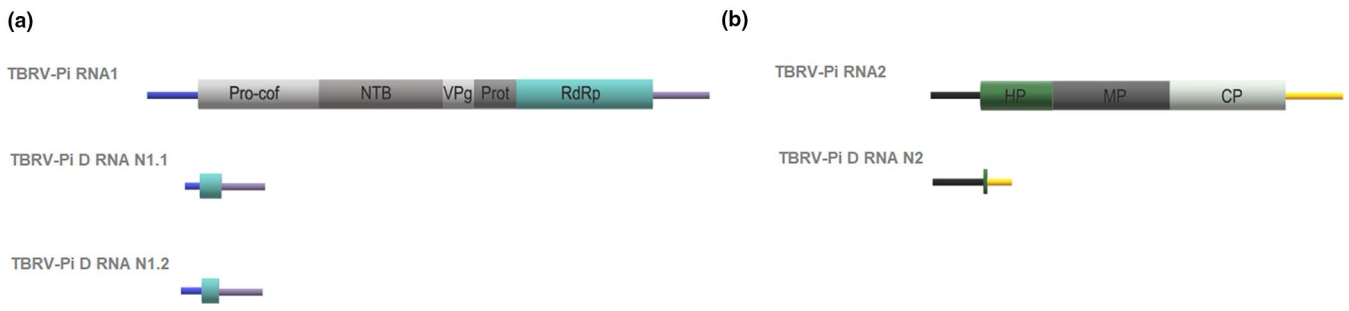

FIGURE 3 Genomic structure of D RNAs obtained for (a) lineage N1 and (b) lineage N2 evolved in *Nicotiana tabacum*. The coloured fragments in the D RNAs correspond to the same coloured regions in the parental RNAs [Colour figure can be viewed at wileyonlinelibrary.com]

on the local host plant within which the lineages evolved. The number of mutations that arose along the experimental evolution varied both in terms of the lineage and the particular local host species. Our AMOVA study concluded that significant genetic differences had accumulated at the three levels of our study after 15 passages of experimental evolution. First, multiple mutations coexisted within lineages, most as singletons, but several ones being fixed. Secondly, there were differences between lineages evolved in the same host (despite one mutation being fixed in all tobacco-evolved lineages), indicating that different adaptive walks were taken by each lineage. Thirdly, there were net differences between host species, which suggests the existence of possible host-specific adaptations taking place during our evolution experiment. An alternative nonadaptive hypothesis to explain the observed differences in TBRV-Pi genetic variability among different host species would rely on the assumption that the RdRp error rate may be host-dependent. This has been previously shown to be the case for cucumber mosaic virus (Pita *et al*., 2007). Unfortunately, the results from our study did not allow this very interesting hypothesis to be tested.

We have taken three different approaches to test whether the observed patterns of genetic variability were compatible with the action of some sort of selection. The first population genetics approach was based upon the Tajima's *D* test of neutrality. In many lineages, $D < 0$, which, combined with the existence of many singletons, points towards fast population expansion as the most parsimonious explanation for the observed levels of genetic variability within lineages (Slatkin and Hudson, 1991). This possibility is consistent with the exponential growth of the viral population after each inoculation. However, after adjusting the significance levels of the multiple tests, all failed to reject the null hypothesis of neutral evolution. Prompted by this failure and the pervasive observation of convergent mutations in different lineages, which can only be explained by strong selective constraints and limited number of accessible adaptive pathways, we undertook a second approach to detect specific codons under positive selection. In this case, we found seven codons in the *CP* and one in the *MP* that may be candidates for beneficial mutations. Interestingly, six of the amino acid replacements in the *CP* involved changing a methionine for another hydrophobic but shorter side chain. Unfortunately, in the absence of a crystal structure of this protein, it is not easy to propose a grounded mechanistic explanation for this situation. However, it is tempting to speculate that the increase in reactive oxygen species, especially H_2O_2, produced by plants upon viral infections may result in oxidation of methionine into methionine sulphoxide (Jacques *et al*., 2015; Veredas *et al*., 2017), which is a polar hydrophilic radical that may strongly affect the proper folding of CP. Therefore, replacements of methionine at structurally important sites by other nonpolar amino acids may be positively selected because they may contribute to keeping the right CP fold under oxidative stress conditions. Our third test of the action of positive selection was based in the identification of coevolving groups of amino acids. A protein's function is the result of the functional and structural communication between amino acid residues. Site constraints are hence dependent on the interactions with other residues of the molecule. Mutations at either nearby sites or functionally related distant sites in the structure will change the selective constraints. Therefore, testing coevolution between sites is an essential step to complement molecular selection analyses, providing more biologically realistic results (Fares and Travers, 2006). We found two triplets of amino acids that were significantly changing in a coordinated manner in two different lineages, one in CP and another in MP, giving extra support to the role of positive evolution operating upon the TBRV-Pi experimental lineages.

Serial passage experiments of viruses in novel hosts followed by characterization of the newly generated genetic variability are a classic approach in evolutionary virology to address questions related to the molecular basis and local adaptation and evolution of virulence (Elena and Sanjuán, 2007). After 20 passages of pepino mosaic virus in different host species, Minicka *et al*. (2015) indicated the presence of 143 nucleotide substitutions arising during the prolonged passaging of the virus. Interestingly, among them, the most abundant type were the nonsynonymous mutations (56.6%). This latter result is in good agreement with our research, which revealed that up to 65.7% of the described mutations were nonsynonymous. Previous research suggests that nonsynonymous point mutations have a greater fitness impact than synonymous or noncoding mutations (Cuevas *et al*., 2012).

The evaluation of the transition to transversion rates ratio, κ, for particular lineages has shown that the nucleotide substitution pattern in the studied regions was non-random, as transitions were far more common than transversions. For all the lineages, κ was greater than 1, with strong differences between hosts and genomic regions. Two main hypotheses have been brought forward to explain the transitions to transversions substitution bias. The first one, the mutational hypothesis, is based on the statement that the transition mutation rates of polymerases are higher than the transversion rates. The second one, the selective hypothesis, assumes that natural selection disfavours transversions (Lyons and Lauring, 2017). Lyons and Lauring (2017) tried to find whether the observed bias was supported by the mutational or selective hypothesis. Analyses conducted based on the nucleotide sequences of two RNA viruses, influenza A virus and human immunodeficiency virus type 1, have revealed that missense transversions are more detrimental to fitness than transitions and, therefore, the analyses directly support the selective hypothesis.

In addition to single nucleotide substitutions, intramolecular recombination events resulting in the formation of D RNAs also play an important role in many plant viruses (both negative- and positive-stranded) from different genera, that is, *Tospovirus*, *Bromovirus*, *Cucumovirus*, *Potexvirus*, and *Nepovirus* (Inoue-Nagata *et al*., 1998; Pathak and Nagy, 2009; Hasiów-Jaroszewska *et al*., 2018). D RNAs are a deleted/rearranged version of the parental virus genome, can be formed as a result of single or multiple deletions, and may arise spontaneously during the replication of the viral genome. All D RNAs described for TBRV so far have arisen during the prolonged passages of the virus in quinoa, exclusively as a result of single deletions in RNA1, varying in length from 480

to 540 nt depending on the parental isolate (Hasiów-Jaroszewska *et al*., 2012, 2018; Rymelska *et al*., 2013). Here, we described new types of D RNAs, which were formed de novo during the passaging of the TBRV-Pi isolate in tobacco. Two D RNAs were found in lineage N1 (originating from RNA1) and one in lineage N2 (originating from RNA2). This is the first report of a D RNA generated from the RNA2 of TBRV.

It has been previously shown that variability of D RNAs occurs not only between different isolates of the same virus, but even between D RNAs of the same isolate. Heterogeneous populations of D RNAs were also described for brome mosaic virus (BMV). Two types of D RNAs accompanying BMV isolate R25 (referred as D1 and D2) were found. Moreover, the sequencing of D2 clones revealed the presence of variants that differed in the location of junction sites (Damayanti *et al*., 1999). Consistent with this observation, in our case D RNAs N1.1 and N1.2 (showing a similar structure to each other and those described previously, except for the different location of the junction sites), could be considered as two different variants of the same type of D RNA rather than as different D RNA species.

Despite the fact that the mechanisms of D RNA formation are still poorly known, it has been postulated that the presence of secondary RNA structures may promote the template-switching of polymerase during the viral replication and formation of D RNAs (White and Morris, 1999; Pathak and Nagy, 2009). The identification of short inverted repeats near the potential "junction sites" in RNA1 and RNA2 of TBRV-Pi suggests that hairpin heteroduplexes may play a key role in the formation of D RNAs.

Previous research by Hasiów-Jaroszewska *et al*. (2018) showed that D RNAs interfere with parental virus replication and affect symptom development. In our study, the D RNAs in one tobacco lineage (N2) resulted in more severe symptoms, but for another lineage (N1) there were no apparent changes in the symptoms of infection. Moreover, symptom enhancement in lineage N2 was noticed, rather than the more commonly observed attenuation associated with the presence of DI RNAs. However, D RNAs may also be effectors of defence responses in the host plant, and thus may exacerbate the symptoms on infection (Perrault, 1981; Simon and Nagy, 1996).

Finally, we tried to find a correlation between the symptoms caused by the virus in the infected plants and the molecular changes in the viral genome. Despite the fact that, during the serial passaging of the virus in tobacco, we noticed evident changes in symptomology in two out of five lineages (N2 and N5), it was not easy to establish an unequivocal connection between mutations and the necrosis enhancement. However, we established that for these lineages, the number of occurring mutations both for *CP* and *MP* was the lowest compared to the other three lineages evolved in tobacco. All tobacco-evolved lineages shared CP mutation I357T and, in addition, lineage N2 also had mutation M468I. Mutation I357T per se did not justify the necrotic symptoms, but perhaps the combination of both mutations may do so. This is a hypothesis worth testing in a future study.

ACKNOWLEDGEMENTS

Work in Poznań was supported by project 2015/17/B/NZ8/02407 from the National Science Centre in Poland to B.H.J. Work in València was supported by grant SPID201900X03998IV1 from Spain Ministerio de Ciencia e Innovación – FEDER to S.F.E.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author.

ORCID

Daria Budzyńska https://orcid.org/0000-0002-3923-0829
Julia Minicka https://orcid.org/0000-0002-7773-9079
Beata Hasiów-Jaroszewska https://orcid.org/0000-0002-8267-023X
Santiago F. Elena https://orcid.org/0000-0001-8249-5593

REFERENCES

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B (Methodological)*, 57, 289–300.

Budzyńska, D., Hasiów-Jaroszewska, B. and Elena, S.F. (2020) Dynamika zmienności i zróżnicowanie genetyczne wirusa czarnej pierścieniowej plamistości pomidora (Tomato black ring virus, TBRV). In: *60th Scientific Session of The Institute of Plant Protection - National Research Institute "International Year of Plant Health", 11–13 February 2020, Poznań, Poland*. Abstracts: p. 22.

Cuevas, J.M., Delaunay, A., Visser, J.C., Bellstedt, D.U., Jacquot, E. and Elena, S.F. (2012) Phylogeography and molecular evolution of Potato virus Y. *PLoS ONE*, 7, e37853.

Damayanti, T.A., Nagano, H., Mise, K., Furusawa, I. and Okuno, T. (1999) Brome mosaic virus defective RNAs generated during infection of barley plants. *Journal of General Virology*, 80, 2511–2518.

Domingo, E., Sheldon, J. and Perales, C. (2012) Viral quasispecies evolution. *Microbiology and Molecular Biology Reviews*, 76, 159–216.

Elena, S.F. and Sanjuán, R. (2007) Virus evolution: insights from an experimental approach. *Annual Review of Ecology, Evolution, and Systematics*, 38, 27–52.

Elena, S.F., Fraile, A. and García-Arenal, F. (2014) Evolution and emergence of plant viruses. *Advances in Virus Research*, 88, 161–191.

Excoffier, L. and Lischer, H.E.L. (2010) Arlequin suite ver 3.5; a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567.

Fares, M.A. and McNally, D. (2006) CAPS: coevolution analysis using protein sequences. *Bioinformatics*, 22, 2821–2822.

Fares, M.A. and Travers, S. (2006) A novel method for detecting intramolecular coevolution: adding a further dimension to selective constraints analyses. *Genetics*, 173, 9–23.

Green, M.R. and Sambrook, J. (2012) *Molecular Cloning: A Laboratory Manual*, 4th edition. New York: Cold Spring Harbor Laboratory Press.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95–98.

Harrison, B.D., Mowat, W.P. and Taylor, C.E. (1961) Transmission of a strain of *Tomato black ring virus* by *Longidorus elongatus* (Nematoda). *Virology*, 22, 480–485.

Hasiów-Jaroszewska, B., Borodynko, N., Figlerowicz, M. and Pospieszny, H. (2012) Two types of defective RNAs arising from the *Tomato black ring virus* genome. *Archives of Virology*, 157, 569–572.

Hasiów-Jaroszewska, B., Minicka, J., Zarzyńska-Nowak, A., Budzyńska, D. and Elena, S.F. (2018) Defective RNA particles derived from

Tomato black ring virus genome interfere with replication of parental virus. *Virus Research*, 250, 87–94.
Inoue-Nagata, A.K., Kormelink, R., Sgro, J.Y., Nagata, T., Kitajima, E.W., Goldbach, R. *et al*. (1998) Molecular characterization of tomato spotted wilt virus defective interfering RNAs and detection of truncated L proteins. *Virology*, 248, 342–356.
Jacques, S., Bhesquière, B., De Bock, P.J., Demol, H., Wahni, K., Willems, P. *et al*. (2015) Protein methionine sulfoxide dynamics in *Arabidopsis thaliana* under oxidative stress. *Molecular and Cellular Proteomics*, 14, 1217–1229.
Kimura, M. (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16, 111–120.
Kosakovsky Pond, S.L., Frost, S.D.W. and Muse, S.V. (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics*, 21, 676–679.
Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33, 1870–1874.
Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution*, 35, 1547–1549.
Lister, R.M. (1960) Occurrence of soil-borne virus diseases of strawberry in Britain. *Plant Pathology*, 9, 102–105.
Lyons, D.M. and Lauring, A.S. (2017) Evidence of the selective basis of transition to transversion substitution bias in two RNA viruses. *Molecular Biology and Evolution*, 34, 3205–3215.
Minicka, J., Rymelska, N., Elena, S.F., Czerwoniec, A. and Hasiów-Jaroszewska, B. (2015) Molecular evolution of *Pepino mosaic virus* during long-term passaging in different hosts and its impact on virus virulence. *Annals of Applied Biology*, 166, 389–401.
Nei, M. and Gojobori, T. (1986) Simple methods for estimating the number of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, 3, 418–426.
Oncino, C., Hemmer, O. and Fritsch, C. (1995) Specificity in the association of Tomato black ring virus satellite RNA with helper virus. *Virology*, 213, 87–96.
Pathak, K.B. and Nagy, P.D. (2009) Defective interfering RNAs: foes of viruses and friends of virologists. *Viruses*, 1, 895–919.
Perrault, J. (1981) Origin and replication of defective interfering particles. *Current Topics in Microbiology and Immunology*, 93, 151–207.
Pita, J., de Miranda, J.R., Schneider, W.L. and Roossinck, M.J. (2007) Environment determines fidelity for an RNA virus replicase. *Journal of Virology*, 81, 9072–9077.
Pospieszny, H. and Borodynko, N. (1999) Wirus czarnej pierścieniowej plamistości pomidora (Tomato black ring nepovirus, TBRV) – metody wykrywania i identyfikacji. *Postępy w Ochronie Roślin*, 39, 327–331.
Pospieszny, H., Borodynko-Filas, N., Hasiów-Jaroszewska, B., Czerwonka, B. and Elena, S.F. (2020) An assessment of the transmission rate of *Tomato black ring virus* through tomato seeds. *Plant Protection Science*, 56, 9–12.
Rymelska, N., Borodynko, N., Pospieszny, H. and Hasiów-Jaroszewska, B. (2013) Analysis of the biological and molecular variability of the Polish isolates of *Tomato black ring virus* (TBRV). *Virus Genes*, 47, 338–346.
Sanfaçon, H., Iwanami, T., Karasev, A.V., van der Vlugt, R., Wellink, J., Wetzel, T. *et al*. (2012) Family *Secoviridae*. In: King, A., Adams, M., Carstens, E. and Lefkowitz, E. (Eds.) *Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses*. London, UK: Elsevier Inc, pp. 890–893.
Sanjuán, R., Nebot, M., Chirico, N., Mansky, L.M. and Belshaw, R. (2010) Viral mutation rates. *Journal of Virology*, 84, 9733–9748.
Simon, A.E. and Nagy, P.D. (1996) RNA recombination in *Turnip crinkle virus*: its role in formation of chimeric RNAs, multimers and in 39-end repair. *Seminars in Virology*, 7, 373–379.
Slatkin, M. and Hudson, R.R. (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, 129, 555–562.
Smith, K.M. (1946) Tomato black-ring: a new virus disease. *Parasitology*, 37, 126–130.
Šneideris, D. and Staniulis, J. (2014) Phylogenetic analysis of Lithuanian tomato black ring virus isolates. *Zemdirbyste-Agriculture*, 101, 193–198.
Tajima, F. (1989) Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics*, 123, 585–595.
Tamura, K. (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Molecular Biology and Evolution*, 9, 678–687.
Veredas, F.J., Cantón, F.R. and Aledo, C. (2017) Methionine residues around phosphorylation sites are preferentially oxidized in vivo under stress conditions. *Scientific Reports*, 7, 40403.
White, K.A. and Morris, T.J. (1999) Defective and defective interfering RNAs of monopartite plus-strand RNA plant viruses. In: Vogt, P.K. and Jackson, A.O. (Eds.) *Satellite and Defective Viral RNAs. Current Topics in Microbiology and Immunology*. Berlin, Heidelberg: Springer, pp. 1–17.
Wright, S. (1950) Genetical structure of population. *Nature*, 166, 247–249.
Zhang, C. and Papakonstantinou, T. (2009) Purification of nucleic acids from gels. In: Liu, D. (Ed.) *Handbook of Nucleic Acid Purification*. Boca Raton, USA: Taylor & Francis CRC Press, p. 552.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Budzyńska D, Minicka J, Hasiów-Jaroszewska B, Elena SF. Molecular evolution of tomato black ring virus and de novo generation of a new type of defective RNAs during long-term passaging in different hosts. *Plant Pathol.* 2020;69:1767–1776. https://doi.org/10.1111/ppa.13258