Contribution of host plant resistance and geographic distance to the structure of Potato virus Y (PVY) populations in pepper in northern Tunisia

Mekki Ben Khalifa, Vincent Simon, M. Marrakchi, Hatem Fakhfakh, Benoît Moury

Version définitive du manuscrit publié dans / Final version of the manuscript published in : Plant Pathology, 2009, vol.58, no.4
DOI: 10.1111/j.1365-3059.2009.02064.x

Contribution of host plant resistance and geographic distance to the structure of *Potato virus Y* (PVY) populations in pepper in Northern Tunisia.

M. Ben Khalifa[a], V. Simon[b], M. Marrakchi[a], H. Fakhfakh[a] and B. Moury[b*]

[a] *Laboratoire de Génétique Moléculaire Immunologie et Biotechnologie, Faculté des Sciences de Tunis, Campus Universitaire, 2092, El-Manar, Tunisia.*

[b] *INRA, UR407 Pathologie Végétale, F-84143 Montfavet cedex, France.*

***Corresponding author:** Benoît Moury ; INRA, Unité de Pathologie Végétale, Domaine St Maurice, BP94, F-84143 Montfavet Cedex, France ; phone: 04 32 72 28 43 ; fax: 04 32 72 28 42 ; e-mail: moury@avignon.inra.fr

Abstract

The impact of plant resistances on the structure of targeted viral populations has rarely been studied, even though it is of importance to optimize the management of genetic control methods. We investigated the genetic structure of *Potato virus Y* (PVY) populations in naturally-infected pepper fields, collected at eight different localities of Northern Tunisia, where 23% of the sampled plants were homozygous for the *pvr2*^{1} recessive resistance allele, while the other plants carried the dominant susceptibility allele *pvr2*$^{+}$. Restriction fragment length polymorphism analysis at three PVY genome segments allowed detecting a high level of viral diversity, a majority of cases of co-infections of individual plants by several PVY haplotypes and a strong genetic differentiation of viral populations collected in the different localities. We detected a strong effect of geographic distances on the differentiation of PVY populations and isolation by distance among these populations was significant. On the opposite, the occurrence of the *pvr2*^{1} resistance allele did not contribute to the structure of viral populations, suggesting that the virulence properties of the virus did not affect significantly its fitness, that the larger deployment of the *pvr2*^{1} gene would probably not be a suitable strategy to control PVY and that other resistance genes should be preferred.

38 **Keywords:** eukaryotic initiation factor 4E (eIF4E), *Potyvirus*, restriction
39 fragment length polymorphism (RFLP), isolation by distance.

40

Introduction

Potato virus Y (PVY) is the type member of the genus *Potyvirus*, the largest among plant viruses, which belongs to the family *Potyviridae*, a group of single-stranded RNA viruses. PVY is transmitted in a nonpersistent manner by more than 70 aphid species (Powell, 1991; Lopez-Abella et al., 1988) and is the causal agent of major diseases in cultivated solanaceous crops including potato, pepper, tobacco and tomato. Based on their host range and the symptoms they induce in potato and tobacco as well as on serological characterization, PVY isolates have been divided into three main groups (O, N and C), which correspond to three monophyletic groups. Several recombinant isolates have also been characterized, mainly between the N and O groups (Revers et al., 1996; Glais et al., 2002; Ogawa et al., 2008). Almost all pepper isolates of PVY belong to the C group (Blanco-Urgoiti et al., 1998).

Genetic resistance in plants is an efficient way to control virus diseases but the more or less rapid selection of virus isolates adapted to resistance genes (*i.e.* virulent isolates) has been documented in several cases (*e.g.* Pelham et al., 1970; Fletcher 1992). The impact of plant resistance genes on the structure of the viral populations has however never been explored, though the analysis of that structure could provide essential information regarding the selective impact of plant resistance and the occurrence of fitness costs associated to virus virulence. In these regards, it is essential to disentangle

the effects of geographical location and host genotype on the structure of virus populations.

In this study, we analysed the genetic diversity of PVY infecting pepper in the regions of Bizerte and Cap Bon, the most important open-field pepper producing areas in Tunisia, with the aims to understand the structure of PVY populations and to estimate the respective impacts of pepper genotype, notably the presence of a recessive resistance allele, and geographic location on that structure.

Materials and methods

Field PVY isolates

PVY isolates were collected according to a hierarchical (nested) sampling design. Collections were made in 21 fields in six localities in the regions of Bizerte, northern Tunisia, and two localities in the region of Cap Bon, North-eastern Tunisia (Fig. 1) from October to December 2006. A total of 192 young pepper (*Capsicum annuum* L.) fruits (one fruit per plant) were randomly collected from symptomatic plants of the local cultivars 'Baklouti' and 'Bar Abid'. The variety Baklouti produces hot, large fruits which become red at maturity, whereas the variety Bar Abid produces long, thin, red-fleshed hot fruits. Samples were first analysed by double antibody sandwich-enzyme-linked immunosorbent assay (DAS-ELISA) with

antibodies raised against PVY and CMV, the two most prevalent pepper viruses in Tunisia (Mnari-Hattab et al., 1999).

cDNA synthesis and RFLP analysis of virus isolates

The viral population from PVY-infected samples was characterized by reverse transcription-polymerase chain reaction-restriction fragment length polymorphism (RT-PCR-RFLP) analysis of the cDNAs obtained from three genomic regions encoding respectively the P3 protein, the viral protein genome-linked (VPg) and the coat protein (CP). These regions were chosen because they affect PVY virulence properties (VPg; Moury et al., 2004) or because they are variable but flanked by relatively conserved regions, allowing the design of PVY-polyvalent primers (P3 and CP). Total RNAs were purified from a 0.5 g piece of flesh of PVY-infected pepper fruits with the Tri-Reagent kit (Molecular Research Center Inc., Cincinnati, USA), and were used as template for RT-PCR (Moury et al., 2004). Parts of the P3 and VPg cistrons, and the entire CP cistron were amplified using primers listed in Table 1. These primers were designed to be polyvalent for all PVY groups and amplified 1145 nucleotides of the P3 cistron, 547 nucleotides of the VPg cistron and a 1159 nucleotide fragment overlapping the CP cistron. Restriction endonucleases used for RFLP analyses of RT-PCR products were chosen based on genomic sequences available for PVY group C, which includes almost all pepper isolates, in order to reveal polymorphism

within this group. For RFLP analysis, 2 µl of RT-PCR products were digested with endonucleases *Rsa*I, *Cla*I and *Bgl*II for the P3 cistron, *Rsa*I, *Hind*III and *Hin*fI for the VPg cistron and *MseI* and *Hae*III for the CP cistron. A fourth endonuclease, *Sac*I, was used for the VPg cistron in order to reveal the presence of PVY isolates belonging to the N or O groups. Indeed, a *Sac*I site is present in 56 sequences of the VPg cistron of 58 available sequences from the N and O groups (or N×O recombinants) but absent among the 22 available sequences of the VPg cistron of the C group.

Genotyping of the *pvr2* locus of sampled pepper plants

The *pvr2* gene from eleven randomly-sampled pepper plants were amplified by RT-PCR using specific primers (Table 1) and the obtained PCR products were directly sequenced by Genome Express (Meylan, France). Occurrence of the two nucleotide substitutions that distinguish alleles $pvr2^{+}$ and $pvr2^{1}$ (Ruffel 2004; Ayme et al., 2007) was then checked for all sampled plants using dCAPS analysis (Neff et al., 2002) of RT-PCR products obtained from the RNA extracts previously used to analyse PVY diversity. The dCAPS analysis used mismatches in one of the two PCR primers flanking the mutations to create a specific *Nru*I recognition site in the $pvr2^{+}$ or $pvr2^{1}$ sequence (Table 1).

Statistical analyses

For each genomic region, haplotypes were defined as the different combinations of RFLP patterns observed with the different endonucleases

among the PVY isolates. A number of PVY samples showed mixtures of two different RFLP patterns for some of the genome segments and some of the endonucleases (Fig 2). These observations were considered to be the result of mixed infections of the plants by several PVY isolates belonging to different haplotypes. In these cases, when only one endonuclease revealed a mixture of two different RFLP patterns, we considered that the samples were composed of two different PVY isolates belonging to two haplotypes. When two or more endonucleases showed mixtures of two different RFLP patterns, neither the number of isolates composing these PVY samples nor the haplotypes of these isolates could be known. For instance, for a given genome segment and a given PVY sample let A and B be the two haplotypes observed for a first endonuclease and C and D be the two haplotypes observed for a second endonuclease. This PVY sample can be composed of two to four isolates whose haplotypes could be, non-exhaustively, A-C and B-D; A-D and B-D; A-C, A-D and B-C or A-C, A-D, B-C and B-D. Such PVY samples were excluded from further analyses. Further analyses were conducted separately for each genome region and/or for the combination of the three genome regions. In the latter case, haplotypes were defined in the same way as described above for the combination of results obtained with different endonucleases.

Completeness of sampling was evaluated by the use of “rarefaction curves”, where the cumulative number of observed haplotypes was plotted *versus*

sample size, and of the richness of the PVY populations (*i.e.* the actual number of haplotypes) estimated with the SPADE software (Chao and Shen, 2003).

The observed haplotype distribution and that expected in a neutral evolution model were compared using the Ewens-Watterson test (Watterson, 1978). Pairwise genetic differentiation between PVY populations in different localities (hereafter subpopulations) was estimated with the Wright's fixation index (F_{ST}; Wright, 1951). Two populations with similar distributions of haplotype frequencies will give an F_{ST} value not statistically different from zero. Following Rousset's recommendations (1997), the hypothesis of isolation by distance (IBD) was explored by examining the correlation between the matrices representing $F_{ST}/(1-F_{ST})$ and the natural logarithm of geographic distance (ln d) for pairs of subpopulations. Significance of the correlation was assessed by a Mantel test using 10,000 permutations using the R software (Ihaka and Gentleman, 1996).

The effects of the sampling locality, pepper variety or *pvr2* allele composition of sampled plants on the variation of haplotype distribution were estimated by hierarchically partitioning the sampled population among these factors, and determining the contribution of each separate factor to the observed genetic variation. For this purpose, comparison by analysis of molecular variance (AMOVA; Excoffier et al., 1992) was performed by

considering the type and frequency of haplotypes in each group of samples and by estimating the probability that their distribution was random.

The Ewens-Watterson test, *F* statistic and AMOVAs were performed using the software Arlequin version 3.11 (Excoffier et al., 2005). *F* statistics and AMOVAs were computed both using haplotypes distributions and distance matrices.

Results

High diversity level and frequency of mixed infections in PVY samples from Tunisian pepper plants

DAS-ELISAs revealed the infection of 114 pepper samples with PVY, 104 of which were simultaneously infected with CMV. Genetic diversity among these 114 samples was examined by RT-PCR-RFLP analysis of three genomic regions in the P3, VPg and CP cistrons. For each genomic region and each endonuclease, RT-PCR-RFLP analyses revealed from 2 to 4 DNA profiles (Fig. 2), except *Sac*I for the VPg cistron which was monomorphic and uncut. Seven different haplotypes were observed for the P3 and VPg cistrons and eight different haplotypes were observed for the CP cistron (Table 2). Evidence of mixed infections with isolates belonging to different haplotypes was found for 51.6% of the PVY samples, based on observation of two RFLP profiles for at least one endonuclease with one genome segment. In addition, for some of the mixed-infected samples, the haplotype

composition could not be known (see the Materials and Methods section). Consequently, only 98 of the 114 PVY positive samples were analysed for at least one genome segment and 75 were analysed for the three genome segments. Note that the number of analysed isolates can be larger than 98 (Table 2) due to the presence of two isolates belonging to different haplotypes in a single initial PVY sample.

The haplotype distribution and frequency for each genomic region was used to analyse the representativeness of sampling. Rarefaction curves presented quite similar shapes for each genome region and reached asymptotes as the number of samples increased, suggesting that PVY populations were exhaustively sampled (Fig. 3). This was supported by the nonparametric Chao1 richness estimates of the actual number of haplotypes (Table 3), which were close to the observed numbers of haplotypes and presented narrow confidence intervals. Thirty-six different haplotypes were observed when taking into account the RFLP analyses of the three genomic regions and an even larger number of haplotypes could have been expected, had the sampling effort been more intense (Table 3). This number of haplotypes is however somewhat smaller than that expected if the haplotypes defined separately in the three genome regions were randomly combined (95% confidence interval: 45.0-60.0; Table 3).

High genetic differentiation of PVY subpopulations

Overall, the highest diversity of haplotypes was found in Aousja and Touiba (from three to six haplotypes for each genomic region) and the lowest diversity was found in El-Marja (which contained only three samples) and Sounine (one or two haplotypes, depending on the genomic region). Whatever the genome segment examined, the distribution of haplotypes was quite dissimilar between subpopulations. For each segment, a predominant haplotype was observed which accounted for 45-63% of the total, but this major haplotype was minor or even absent in several subpopulations. There was no specific association between particular haplotypes and the sampled varieties of pepper. In contrast, several haplotypes were observed in one locality only (for Aousja, Utique, Touiba, Boucharraya and El-Marja), usually at low frequencies.

Comparing the observed and expected homozygosity values with the Ewens-Watterson test did not reveal any significant departure from neutrality for the P3 or CP cistrons (Table 4). This was the same for the VPg cistron, except for the subpopulation from Aousja, where the observed homozygosity was significantly lower than the expected one ($P<0.01$). In that locality and for that genome segment, six different haplotypes were observed, four of which at frequencies 15-27%. The Ewens-Watterson test was not significant when combining the three genome regions (Table 4). Consequently, further analyses of the structure of the PVY populations

based on these viral genomic segments will globally not be obscured by strong selection or demography effects.

Genetic differentiation between PVY subpopulations was shown by pairwise F_{ST}. From 16 to 21 of 28 pairwise F_{ST} comparisons were significant (P<0.05), depending on the genome segment and, remarkably, almost all pairwise F_{ST} were significant for at least one genome segment (data not shown). Similarly, for the combination of the three genome regions, 26 of 28 pairwise F_{ST} comparisons were significant (Table 5). Results were similar using distance matrices instead of raw haplotypes distributions to compute F_{ST} values. A hypothesis for this high genetic differentiation could be a lack of PVY migration between sites. Confirming that hypothesis, we noticed that five of seven haplotypes for the P3 cistron, two of seven haplotypes for the VPg cistron and six of eight haplotypes for the CP cistron were observed in two or less localities, suggesting lack of virus dissemination.

Combining the three genome regions and using the estimated F_{ST} values between subpopulations, a small but significant correlation was observed between geographic distances and PVY subpopulation differentiation (Pearson's coefficient of correlation r=0.36; P-value=0.013, Mantel test; Fig. 4). This correlation was higher and significance increased at smaller distances, *i.e*. using the samples from the region of Bizerte only, which were

collected in a 16×6 km area (Pearson's coefficient of correlation r=0.66; *P*-value=0.006, Mantel test; Fig. 4).

Relative contributions of geographical location and host genotype to the structure of the PVY population

Since the alleles present at the *pvr2* locus determine resistance or susceptibility of pepper plants to PVY isolates and could be important to understand the differentiation of PVY populations, each RNA sample was checked for occurrence of two nucleotide substitutions differentiating $pvr2^+$ from $pvr2^1$ (Ruffel 2004; Ayme et al., 2007), and both $pvr2^+$ and $pvr2^1$ from other *pvr2* alleles, by dCAPS analysis. Three kinds of plants were observed: 66% of $pvr2^+/pvr2^+$ and 11% of $pvr2^+/pvr2^1$ plants (both kinds of plants called "susceptible" in the following, given that resistance is recessive) and 23% of $pvr2^1/pvr2^1$ plants (hereafter called "resistant"). The $pvr2^2$ and $pvr2^3$ alleles were absent from the analysed plants since none of the samples showed uncut RT-PCR products simultaneously for both nucleotide substitutions. Resistant plants were present in all localities at frequencies between 12 and 28%, except in Sounine and El-Marja (0%) and Saaden (75%). The estimated haplotype richness (Chao1 estimators) was around four times higher in susceptible than in resistant plants (Table 3), although 95% confidence intervals were large and did slightly overlap. This could indicate a selective effect of the $pvr2^1/pvr2^1$ resistant plants and the lack of

fitness cost associated with virulence toward *pvr2*[1], since the richness in susceptible plants was similar to that in the whole population (Table 3).

AMOVAs were conducted to evaluate the contribution of various factors to the differentiation of PVY subpopulations. The total variation observed for the combination of the three genomic regions was shared between (i) variations among the different groups considered (localities, pepper varieties, susceptibility or resistance of sampled plant), (ii) between fields or plants within the same group and (iii) within the field or plant population. The locality contributed significantly to the total variation (10.2% of total variation) (Table 6). 8.1% of the variation was contributed by the field level within localities while the largest part of the variance (81.7%) was at the within-field level. On the contrary, there was no effect of the pepper variety or of the *pvr2* allele in the host plants on the differentiation of PVY populations (Table 6). Similar results were obtained using distance matrices to perform AMOVAs and by the separate analysis of the three genome regions (data not shown).

Discussion

Genetic variation in PVY populations was studied in pepper crops in Northern Tunisia by analysing RFLP haplotypes of three genomic regions. Analysis of the VPg cistron with the *Sac*I endonuclease suggested that all collected isolates belonged to the C group, since almost all available

sequences of PVY O and N groups possess at least one *Sac*I restriction site in the VPg cistron. This was confirmed later by sequencing the VPg cistron of 20 randomly chosen samples (data not shown). Potato crops are common in the vicinity of the sampled pepper fields around Bizerte and Cap Bon. These crops are mainly infected with N and O group PVY recombinants (Boukhris Bouhachem et al., 2007), which were not found in pepper crops, confirming that PVY does not spread from potato to pepper (Gebre Selassie et al., 1985).

Overall, a high PVY diversity was observed, since a limited number of restriction enzymes permitted to estimate that the actual richness in the PVY population examined varied between 45.5 and 131.0 haplotypes (Table 3). Also, a high percentage (51.6%) of plants was infected simultaneously by several PVY isolates belonging to different haplotypes. This percentage is certainly an underestimation because of the small number of endonucleases used and of the lack of power to detect PVY isolates present at low concentrations by RT-PCR-RFLP. Indeed, artificial mixtures of PVY isolates with different RFLP profiles showed that this method does not allow detecting isolates representing less than 10 to 20% of the entire virus population (data not shown). A potential consequence of this detection limit could have been to decrease the number of observed PVY isolates and, consequently, to decrease the ability to differentiate genetically the PVY

subpopulations. This was however not the case (Table 5). Co-infection of plants with different closely-related virus variants is usually rare (<10%; Bodaghi et al., 2004; Rubio et al., 2001; McNeil et al., 1996) although higher frequencies of mixed infections have occasionally been observed, especially in perennial plants (Vigne et al., 2004). However, detection of viral co-infections in plants depends largely on the level of polymorphism of the virus population and on the use of appropriate molecular tools, such as RFLP or sequencing of a number of DNA clones representing the viral population. Therefore, we still lack an exhaustive view of the occurrence of viral co-infections (García-Arenal et al., 2001).

There was no evidence of differences in the selective constraints exerted on the different genome regions analysed. Indeed, similar results were obtained for the three regions by analyses of homozygosity (Table 4), *F* statistics or AMOVAs (data not shown). Except one genome segment for one PVY subpopulation, the haplotype distributions did not deviate significantly from the neutral model, thus making our data suitable for analyses of migration patterns and population differentiation independently of demography and selection processes. The only exception concerns the VPg segment in the Aousja subpopulation, where there is an excess of haplotypes at intermediate frequencies compared to the neutral model. This could be due to repeated selection of VPg variants by the host plant, since the VPg was

shown to be involved in adaptation to resistance alleles at the *pvr2* locus (Moury et al., 2004; Ayme et al., 2006; 2007).

When analysing the haplotypes observed for the combination of the three genome regions, we noticed a near-random association of haplotypes between these regions (Table 3). This may be caused by high recombination and/or mutation frequencies in these populations. No estimations of recombination frequency are available for PVY. However, genome analyses have revealed the widespread occurrence of recombination at the intra-specific level (Revers et al., 1996; Glais et al., 2002; Ogawa et al., 2007) which, together with the high frequency of mixed-infected plants, makes recombination a plausible mechanism for the generation of the haplotype diversity observed.

A strong genetic differentiation was observed between the sampled subpopulations as assessed by the *F* statistic (Table 5). In theory, that structure could be attributed either to ecological substructuring of the virus population into ecotypes associated with distinct habitats, such as different plant genotypes, or to the geographical distance separating the sampled localities, which could be here considered as the null hypothesis for population differentiation. A significant IBD was indeed observed when comparing the geographic and genetic distances separating the subpopulations (Fig. 4). The correlation between geographic and genetic distances increased when analysing the samples from the region of Bizerte

only. It is known that sampling at large distances can reduce the capacity to detect IBD because the variance of F_{ST} estimators is probably higher at large distances (Rousset, 1997). Confirming this, significant correlations were observed between the logarithm of the distance (ln d) and the residuals of the linear regression between $F_{ST}/(1-F_{ST})$ and (ln d), either for the whole dataset or the region of Bizerte only (*P*-values=0.014 and 0.002, respectively; Mantel tests).

AMOVAs showed that the strong effect of the locality in PVY genetic differentiation was due to the location of the site itself but not to the pepper variety cultivated in these sites (Table 6). The varieties Bar Abid and Baklouti are local pepper populations which are propagated from seeds saved from one season to the next by the growers. Relatively high rates of allogamy have been observed in pepper (Tanksley, 1984) and can be responsible for maintaining polymorphism at the *pvr2* locus in these plant populations. In spite of the strong resistance level conferred by the $pvr2^1$ gene against avirulent PVY isolates (Moury et al., 2004), the polymorphism at the *pvr2* locus did not contribute significantly to the genetic differentiation of PVY, as assessed by AMOVA (Table 6). Eight of the ten haplotypes (defined from the combination of the three genome regions) observed in $pvr2^1/pvr2^1$ resistant plants were also observed frequently in susceptible plants carrying the $pvr2^+$ allele, suggesting limited, if any, fitness cost associated to the virulence toward $pvr2^1$. The remaining two

haplotypes were a singleton and a doubleton (haplotype observed twice). It should be noted that co-infections with CMV do not modify the level of PVY resistance conferred by the *pvr2*^{1} allele in pepper (M. Ben Khalifa and B. Moury, data not shown). These observations indicate that the host plant is only weakly involved in the genetic structure of PVY populations and the observed PVY diversity is mostly explained by the geographic distances separating the subpopulations. The strong genetic differentiation of PVY populations observed at relatively short distances could be due to the fact that nonpersistent viruses such as PVY bind only transiently to their aphid vectors (a few hours at maximum and frequently less when aphids feed on healthy plants), which is certainly responsible for a lack of long-distance virus dispersal and for a strong differentiation of PVY populations. In addition, genetic drift caused by narrow population bottlenecks both during plant infection (French and Stenger, 2003; Sacristán et al., 2003) and during PVY transmission by aphids between plants (Moury et al., 2007) may also contribute to strong population differentiation.

Measuring fitness costs imposed by virulence properties is important to analyse the emergence of pathogen variants and to optimize the use of plant resistance genes (Leach et al., 2001). In plant viruses, fitness comparisons between virulent and avirulent variants have been essentially performed in controlled conditions with a very limited number of virus genotypes and of environmental conditions (Ayme et *al*. 2006; Desbiez et *al*. 2003; Jenner et *al*.

2002). The representativeness of these experiments regarding the problems of plant resistance management and durability can therefore be questioned and measures of fitness costs in more realistic conditions would be highly desirable. This study provides a first analysis of the impact of a plant resistance gene on the genetic structure of targeted virus populations in epidemiological conditions. Our results suggest the lack of strong fitness costs associated to the virulence towards the *pvr2*^{1} resistance gene. This could be because the mutations responsible for the virulence towards *pvr2*^{1} are not costly for the virus or that an initial fitness cost conferred by these mutations was compensated for by additional mutations. Consequently, the larger deployment of pepper cultivars carrying the *pvr2*^{1} resistance would probably not help control PVY in pepper in Northern Tunisia. Instead, other resistance genes such as *pvr2*^{2} (Ayme et al. 2007) or *Pvr4*, which proved to be durable, would be more appropriate.

Acknowledgments

This study was supported by a Franco-Tunisian CORUS program from the French Ministère des Affaires Etrangères, the Tunisian Ministère de l'Education Supérieure, de la Recherche Scientifique et de la Technologie, a program “Ecologie pour la gestion des Ecosystèmes et de leurs Ressources” and a program “Agriculture et Développement Durable” from the Agence Nationale de la Recherche. We gratefully acknowledge Pierre Franck for

433 advice concerning the analyses and Carole Caranta and Carine Charron for
434 sharing oligonucleotides for the *pvr2* gene. We also acknowledge the
435 Tunisian Ministry of Agriculture for help with the surveys.

436

References

Ayme V, Souche S, Caranta C, Jacquemond M, Chadœuf J, Palloix A, Moury B, 2006. Different mutations in the VPg of Potato virus Y confer virulence on the *pvr2*^{3} resistance in pepper. *Molecular Plant-Microbe Interactions* **19,** 557-563.

Ayme V, Petit-Pierre J, Souche S, Palloix A, Moury B, 2007. Molecular dissection of the potato virus Y VPg virulence factor reveals complex adaptations to the *pvr2* resistance allelic series in pepper. *Journal of General Virology* **88,** 1594-1601.

Blanco-Urgoiti B, Sanchez F, Perez de san Roman C, Dopazo J, Ponz F, 1998. PVY^{C} isolates are a homogeneous pathotype but two different genetic strains. *Journal of General Virology* **79,** 2037–2042.

Bodaghi S, Mathews DM, Dodds JA, 2004. Natural incidence of mixed infections and experimental cross protection between two genotypes of *Tobacco mild green mosaic virus*. *Phytopathology* **94,** 1337-1341.

Boukhris-Bouhachem S, Khamassy N, Glais L, Kerlan C, 2007. Occurrence in Tunisia of potato tuber necrotic ringspot disease (PTNRD) caused by variant PVY^{NTN} of *Potato virus Y*. *New Disease Reports* **15**.

Chao A, Shen T-J, 2003. Program SPADE (Species Prediction And Diversity Estimation). Program and User's Guide published at http://chao.stat.nthu.edu.tw.

458 Desbiez C, Gal-On A, Girard M, Wipf-Scheibel C, Lecoq H, 2003. Increase
459 in Zucchini yellow mosaic virus symptom severity in tolerant zucchini
460 cultivars is related to a point mutation in P3 protein and is associated
461 with a loss of relative fitness on susceptible plants. *Phytopathology* **93**,
462 1478-1484.
463 Excoffier L, Laval G, Schneider S, 2005. Arlequin ver. 3.0: An integrated
464 software package for population genetics data analysis. *Evolutionary*
465 *Bioinformatics Online* **1,** 47-50.
466 Excoffier L, Smouse P, Quattro J, 1992. Analysis of molecular variance
467 inferred from metric distances among DNA haplotypes: Application to
468 human mitochondrial DNA restriction data. *Genetics* **131,** 479-491.
469 Fletcher JT, 1992. Disease resistance in protected crops and mushrooms.
470 *Euphytica* **63**, 33-49.
471 French R, Stenger DC, 2003. Evolution of *Wheat streak mosaic virus*:
472 Dynamics of population growth within plants may explain limited
473 variation. *Annual Review of Phytopathology* **41,** 199–214.
474 García-Arenal F, Fraile A, Malpica JM, 2001. Variability and genetic
475 structure of plant virus populations. *Annual Review of Phytopathology*
476 **39,**157-186.
477 Gebre Selassie K, Marchoux, G, Delecolle B, Pochard E, 1985. Variabilité
478 naturelle des souches du virus Y de la pomme de terre dans les cultures

de piment du sud-est de la France. Caractérisation et classification en pathotypes. *Agronomie* **5,** 621–630.

Glais L, Tribodet M, Kerlan C, 2002. Genetic variability in *Potato potyvirus Y* (PVY): Evidence that $PVY^{N}W$ and PVY^{NTN} variants are single to multiple recombinants between PVY^{O} and PVY^{N} isolates. *Archives of Virology* **147,** 363-378.

Ihaka R, Gentleman R, 1996. R: A language for data analysis and graphics. *Journal of Computational and Graphic Statistics* **58,** 299–314.

Jenner CE, Wang X, Ponz F, Walsh JA, 2002. A fitness cost for Turnip mosaic virus to overcome host resistance. *Virus Research* **86**, 1-6.

Leach JE, Vera Cruz CM, Bai J, Leung H, 2001. Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Annual Review of Phytopathology* **39**, 187-224.

Lopez-Abella D, Bradley RHE, Harris KF, 1988. Correlation between stylet paths made during superficial probing and the ability of aphids to transmit nonpersistent viruses. *Advances in Disease and Vector Research* **5,** 251-285.

McNeil JE, French R, Hein GL, Baezinger PS, Eskridge KM, 1996. Characterization of genetic variability among natural populations of wheat streak mosaic virus. *Phytopathology* **86,** 1222-1227.

Mnari-Hattab M, Ezzaier K, Gebre-Selassie K, Marchoux G, Gognalons P, 1999. Surveys of viruses affecting pepper (*Capsicum annuum* L.) in

Tunisia. In: *Proceedings of the 7th International plant virus epidemiology symposium on Plant virus epidemiology: Current status and future prospects*. *April 11-17, 1999*: Almeria, Spain. pp 143-144.

Moury B, Morel C, Johansen E, Guilbaud L, Souche S, Ayme V, Caranta C, Palloix A, Jacquemond M, 2004. Mutations in *Potato virus Y* genome-linked protein determine virulence toward recessive resistances in *Capsicum annum* and *Lycopersicon hirsutum*. *Molecular Plant-Microbe Interactions* **17,** 322-329.

Moury B, Fabre F, Senoussi R, 2007. Estimation of the number of virus particles transmitted by an insect vector. *Proceedings of the National Academy of Sciences of the USA* **104,** 17891-17896.

Neff MM, Turk E, Kalishman M, 2002. Web-based primer design for single nucleotide polymorphism analysis. *Trends in Genetics* **18,** 613-615.

Ogawa T, Tomitaka Y, Nakagawa A, Ohshima K, 2008. Genetic structure of a population of Potato virus Y inducing potato tuber necrotic ringspot disease in Japan; comparison with North American and European populations. *Virus Research* **131**, 199-212.

Pelham J, Fletcher JT, Hawkins JH, 1970. The establishment of a new strain of tobacco mosaic virus resulting from the use of resistant varieties of tomato. *Annals of Applied Biology* **65**, 293-297.

Powell G, 1991. Cell membrane punctures during epidermal penetrations by aphids: consequences for the transmission of two potyviruses. *Annals of Applied Biology* **119,** 313-321.

Revers F, Le Gall O, Candresse T, Le Romancer M, Dunez J, 1996. Frequent occurrence of recombinant potyvirus isolates. *Journal of General Virology* **77,** 1953-1965.

Rousset F, 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145,** 1219-1228.

Rubio L, Ayllón MA, Kong P, Fernández A, Polek M, Guerri J, Moreno P, Falk BW, 2001. Genetic variation of *Citrus tristeza virus* isolates from California and Spain: evidence for mixed infections and recombination. *Journal of Virology* **75,** 8054-8062.

Ruffel S, 2004. Résistances récessives aux potyvirus chez les solanacées et facteurs du complexe d'initiation de la traduction. PhD thesis, University of Aix Marseille II.

Sacristán S, Malpica JM, Fraile A, García-Arenal F, 2003. Estimation of population bottlenecks during systemic movement of *Tobacco mosaic virus* in tobacco plants. *Journal of Virology* **77,** 9906–9911.

Tanksley SD, 1984. High rates of cross-pollination in chile pepper. *Hortscience* **19,** 580-582.

Vigne E, Bergdoll M, Guyader S, Fuchs M, 2004. Population structure and genetic variability within isolates of *Grapevine fanleaf virus* from a

543 naturally infected vineyard in France: evidence for mixed infection and
544 recombination *Journal of General Virology* **85**, 2435–2445.
545 Watterson G, 1978. The homozygosity test of neutrality. *Genetics* **88,** 405-
546 417.
547 Wright S, 1951. The genetical structure of populations. *Annals of Eugenics*
548 **15,** 323-354.
549

Figure 1 Sampling sites in the regions of Bizerte (Northern Tunisia) and Cap-Bon (North-eastern Tunisia).

Figure 2 Restriction profiles obtained for the P3, VPg and CP cistrons of PVY by digestion of corresponding RT-PCR products with endonucleases *Rsa*I (lines: 1 to 3), *Cla*I (lines: 4 and 5) and *Bgl*II (lines 6 to 9) for P3, *Rsa*I (lines 1 and 2), *Hind*III (lines 3 to 5) and *Hinf*I (lines 6 and 7) for VPg, and *Mse*I (lines 1 to 4) and *Hae*III (lines 5 to 7) for CP. An example of mixed infection is shown for the CP cistron and *Mse*I endonuclease (lines 1, 2, 3, 5: profiles corresponding to single haplotypes; lines 4, 6, 7: profiles corresponding to a mixture of the same haplotypes). M: size marker. ND: non digested RT-PCR products.

Figure 3 Observed haplotype richness of *Potato virus Y* samples collected in Tunisian pepper fields versus sample size (rarefaction curves). The number of haplotypes observed for a given sample size was averaged over 100 simulations. Diversity was revealed from RFLP of three genomic regions.

Figure 4 Isolation by distance pattern between genetic differentiation, measured as F_{ST} / (1- F_{ST}), for haplotypes distributions combining the three genome regions, and geographic distances (natural logarithm of the distances in km) for pairwise PVY subpopulations. The correlation was significant as assessed by Mantel test, both for samples from the region of

571 Bizerte only (boxed), or for the whole dataset (P=0.006 and P=0.013,
572 respectively).
573

574 **Table 1** Primers used for reverse transcription or PCR amplifications
575

Name	Polarity	Primer sequence (5' to 3')[a]	Binding site
P3-F	+	TCACCNTTYAGAGARGGNGG	2445-2464 [b]
P3-R	-	CARTCRCTCCTTTCAGCATC	3570-3589 [b]
VPg-F	+	GAATYCAAGCHYTRAAGTTTCG	5734-5755 [b]
VPg-R	-	GCTTCATGYTCYACHTCCTG	6261-6280 [b]
CP-F	+	GCTGAACACAGGCTCGAAG	8289-8307 [b]
CP-R	-	TAAAAGTAGTACAGGAAAAGCCA	9425-9447 [b]
Pvr2-F	+	AAAAGCACACAGCACCAACA	9-28 [c]
Pvr2-R	-	TATTCCGACATTGCATCAAGAA	716-737 [c]
dC67[d]	-	GAGCTACCCCAAGCAGCTTGTTTCGATTTCGCG	229-261 [c]
dC79[d]	-	CTTCAACAGTGGAGAAAGTGTAGACGTTGCGT	265-296 [c]

576 [a] N: A, C, G or T; Y: C or T; R: A or G; H: A, C or T.
577 [b] referring to PVY strain SON41p (accession number AJ439544).
578 [c] referring to the cDNA of the *pvr2*$^{+}$ allele from *Capsicum annuum* cv. Yolo
579 Wonder (accession number AY122052).
580 [d] The primers were designed from sequences AY122052 (*pvr2*$^{+}$) and
581 AF521964 (*pvr2*^{1}) with the help of dCAPS Finder 2.0 (Neff et al., 2002).
582 The mismatches creating the specific *Nru*I restriction sites are underlined.

583 **Table 2** Distribution of RFLP haplotypes of three genome regions of *Potato virus Y* among pepper samples collected in different localities of
584 Northern Tunisia
585

Locality	Variety	No. fields	No. samples	P3 cistron							VPg cistron							CP cistron							
				1	2	3	4	5	6	7	a	b	c	d	e	f	g	α	β	χ	δ	ε	ϕ	γ	η
Aousja	Baklouti	2	16	6	10	1	2	1	-	-	5	6	8	5	2	1	-	17	11	-	1	-	-	-	-
	Bar Abid	1	7	1	1	5	-	-	-	-	1	1	1	-	-	3	-	6	-	-	5	-	-	-	-
Saaden	Baklouti	1	14	12	6	-	-	-	-	-	15	-	-	4	1	-	-	10	6	6	-	1	-	-	-
Sounine	Baklouti	1	7	7	-	-	-	-	-	-	-	6	1	-	-	-	-	7	-	-	-	-	-	-	-
Ghar el	Baklouti	2	6	5	-	-	-	-	-	-	9	4	2	4	-	-	-	6	-	-	-	-	-	-	-
Melh	Bar Abid	1	9	9	-	-	-	-	-	-	1	2	-	-	1	-	-	4	-	-	-	5	-	-	-
Utique	Baklouti	2	12	-	9	-	2	-	-	1	9	-	3	-	2	-	-	2	-	8	-	-	-	2	-
Touiba	Baklouti	5	6	1	2	-	-	3	-	-	2	1	1	-	-	-	-	2	1	-	1	-	2	-	1
Boucharraya	Bar Abid	5	18	4	9	-	-	-	-	-	13	-	-	2	1	-	1	6	8	-	-	-	4	-	-
El-Marja	Bar Abid	1	3	-	-	-	-	-	3	-	-	-	3	-	-	-	-	3	-	-	-	-	-	-	-
Total		21	98	45	37	6	4	4	3	1	55	20	19	15	7	4	1	63	26	14	7	6	6	2	1

586

Table 3 Estimations of plant virus population richness (actual number of haplotypes)

Genome region and dataset	N[a]	n[b]	S_{Chao1}[c]	$CI_{95\%}$
P3	100	7	7.0	7.0-7.0[c]
VPg	120	7	7.0	7.0-7.0[c]
CP	125	8	8.5	8.0-16.4[c]
Combination of the three regions	107	36	66.1	45.5-131.0[c]
Random sampling of the three regions	107	52.2	-	45.0-60.0[d]
Resistant (*pvr2*^{1}/*pvr2*^{1}) plants (3 regions)	25	10	14.5	10.5-50.9[c]
Susceptible (*pvr2*$^{+}$/-) plants (3 regions)	82	34	64.1	43.5-129.0[c]

[a] Number of virus samples.
[b] Observed number of haplotypes.
[c] nonparametric estimator of actual richness of the population and 95% confidence intervals ($CI_{95\%}$) obtained with the SPADE software (Chao and Shen, 2003).
[d] $CI_{95\%}$ obtained with 1,000 Monte Carlo simulations with the R software (Ihaka and Gentleman, 1996) for random sampling of the three genomic regions.

Table 4 Observed and expected homozygosity according to the neutral model with infinite alleles and probability of the Ewens-Watterson test of neutrality

	P3 cistron			VPg cistron			CP cistron			Three genome regions		
Localities	Obs.	Exp.	*P*	Obs.	Exp.	*P*	Obs.	Exp.	*P*	Obs.	Exp.	*P*
Aousja	0.29	0.39	0.20	0.19	0.34	0.007*	0.43	0.61	0.13	0.09	0.09	0.52
Saaden	0.56	0.72	0.24	0.60	0.56	0.69	0.33	0.46	0.14	0.20	0.20	0.62
Sounine	1.00	-	-	0.76	0.65	1.00	1.00	-	-	0.68	0.61	1.00
Ghar el Melh	1.00	-	-	0.30	0.37	0.28	0.56	0.71	0.25	0.18	0.25	0.10
Utique	0.60	0.51	0.84	0.48	0.53	0.47	0.50	0.51	0.53	0.34	0.23	1.00
Touiba	0.39	0.43	0.60	0.38	0.38	1.00	0.30	0.32	0.72	0.38	0.38	1.00
Boucharraya	0.57	0.70	0.32	0.60	0.42	0.95	0.36	0.55	0.06	0.21	0.26	0.23
El-Marja	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-

Obs.: observed haplotype distributions; Exp.: haplotype distributions expected under the neutral model. *Significant difference ($p<0.05$).

602 **Table 5** F_{ST} values between pairs of PVY subpopulations for the combination of three coding
603 regions using haplotypes distributions
604

	Aousja	Saaden	Sounine	Ghar el Melh	Utique	Touiba	Boucharraya
Saaden	0.103*						
Sounine	0.163*	0.312*					
Ghar el Melh	0.067*	0.074*	0.190*				
Utique	0.148*	0.080*	0.397*	0.201*			
Touiba	0.019	0.161*	0.403*	0.151*	0.233*		
Boucharraya	0.070*	0.098*	0.319*	0.118*	0.211*	0.094	
El-Marja	0.301*	0.375*	0.742*	0.363*	0.472*	0.520*	0.385*

605 *Significant difference ($p<0.05$).
606

607 **Table 6** Analysis of molecular variance for haplotypes distributions when the subpopulations are
608 grouped by locality, pepper variety or resistance status of sampled plants
609

Source of variation	d.f.	Sum of squares	Components of variance	Percent variation	*F*-statistics
Between localities	7	9.3	0.05	10.2*	F_{CT}= 0.10*
Between fields in each locality	10	5.8	0.04	8.1*	F_{SC}= 0.09*
Within fields	91	36.1	0.40	81.7*	F_{ST}= 0.18*
Total	108	51.1	0.49		
Between varieties	1	1.1	0.003	0.5	F_{CT}= 0.01
Between fields with the same variety	16	14.0	0.08	17.2*	F_{SC}= 0.17*
Within fields	91	36.1	0.40	82.3*	F_{ST}= 0.18*
Total	108	51.2	0.48		
Between resistant and susceptible plants	1	0.5	-0.01	-2.6	F_{CT}=-0.03
Between fields with the same category of plants	24	18.0	0.09	18.9*	F_{SC}=0.18*
Within fields	83	32.7	0.39	83.7*	F_{ST}=0.16*
Total	108	51.2	0.47		

610 * Significant at $P<0.01$.