Emergence of new virulent populations of apple scab from nonagricultural disease reservoirs

Christophe Lemaire1*, Marie De Gracia2*, Thibault Leroy1,3, Monika Michalecka4, Hanne Lindhard-Pedersen5, Fabien Guerin2,6, Pierre Gladieux2,7,8 and Bruno Le Cam2

1Université d’Angers, UMR1345 Institut de Recherche en Horticulture et Semences (IRHS), SFR4207 QUASAV, PRES L’UNAM, Angers 49075, France; 2INRA, UMR1345 Institut de Recherche en Horticulture et Semences (IRHS), Beaucouzé 49071, France; 3INRA, UMR 1202 BIOEGEO, Cestas 33612, France; 4Department of Plant Pathology, Institute of Horticulture, Skierniewice 96-100, Poland; 5Department of Food Science, Aarhus University, Aarslev 5792, Denmark; 6Université de la Réunion, UMR Peuplements Végétaux et Bioagresseurs en Milieu Tropical (PVBMT), Saint Pierre 97715, La Réunion, France; 7INRA, UMR BGPI, Montpellier 34398, France; 8CIRAD, UMR BGPI, Montpellier 34398, France

Summary

- Plant pathogens adapt readily to new crop varieties in agrosystems, and it is crucial to understand the factors underlying the epidemic spread of new virulent strains if we are to develop more efficient strategies to control them.
- In this study we used multilocus microsatellite typing, molecular epidemiology tools and a large collection of isolates from cultivated, wild and ornamental apples to investigate the origin of new virulent populations of Venturia inaequalis, an ascomycete fungus causing apple scab on varieties carrying the Rvi6 resistance gene.
- We demonstrated a common origin at the European scale of populations infecting apples (Malus × domestica) carrying the Rvi6 resistance and Malus floribunda, the progenitor of the Rvi6 resistance. Demographic modeling indicated that the Rvi6-virulent lineage separated several thousands of years ago from populations infecting non-Rvi6 hosts, without detectable gene flow between the two lineages.
- These findings show that ‘breakdowns’ of plant resistance genes can be caused by the selection and migration of virulent genotypes from standing genetic variation maintained in environmental disease reservoirs, here ornamental crabapples. This work stresses the need to take better account of pathogen diversity in resistance screenings of breeding lines and in resistance deployment strategies, in order to enhance sustainable disease management.

Introduction

It has been admitted for almost a century that plant pathogens evolve rapidly in agrosystems and gene-for-gene resistances only control diseases for a few years (Johnson, 1961; Brown, 1994; Crute & Pink, 1996; McDonald, 2010). Understanding the mechanisms driving the appearance of new pathogen populations infecting new resistant crop varieties is crucial to achieve durable resistance in agrosystems. The emergence of new virulent plant pathogens is often conceptualized using the classical model of positive selection (Haldane, 1927; Fisher, 1930) under a generic arms race model (Dawkins & Krebs, 1979). Theoretical works predict that human management of resistance gene frequencies in crop varieties most often drives plant pathogen coevolutionary dynamics to instability in farming systems where the factors maintaining genetic diversity in natural systems have been excluded (Barrett et al., 2006; Brown & Tellier, 2011). This process is usually referred to as boom and bust cycle, because of the often dramatic rise and fall in the effectiveness of plant resistance. The standard model for describing the population genetics of pathogen adaptation to new resistant varieties assumes that positive selection of new virulence alleles primarily acts on single de novo mutations (Stukenbrock et al., 2007; Stukenbrock & Baillelon, 2012; but see also Karasov et al., 2010). Under this model, adaptation occurs in a mutation-limited regime, with the speed of adaptation governed by the waiting time until the arrival of a new adaptive mutation, which itself depends on the mutation rate toward adaptive alleles, and the population size (Charlesworth, 2009). Most fungal pathogens of domesticated plants have small long-term population size, as derived from levels of standing neutral variation, and adaptation should thus be substantially retarded. However, periods of huge census size and large mutational targets can increase the speed of adaptation (Garud et al., 2015; Gladieux et al., 2015).

Another explanation for the rapid adaptation of pathogens to a new host is the selection and migration of virulent genotypes from standing variation maintained in environmental disease reservoirs. Studies of pathogen of humans and bacterial pathogens of plants have recently shown the impact of pathogen
populations inhabiting environmental reservoirs on the evolution of enhanced pathogen fitness (Burdon & Thrall, 2008; Morris et al., 2009; Monteil et al., 2013 and references therein), but our understanding of the role of environmental populations in the emergence of new virulent populations of fungal plant pathogens remains limited (Leroy et al., 2014). Epidemiological and evolutionary studies on the emergence of new virulent fungal genotypes have focused overwhelmingly on populations sampled on cultivated crops, because of a lack of knowledge about the identity and distribution of wild hosts, or because of impediments in the characterization of symptoms on these hosts. Strains isolated from environmental reservoirs, and in particular from the wild relatives of domesticated crops, are also rarely used in screening of breeding lines for disease resistance (Zhan et al., 2005; Lê Van et al., 2012; Burdon et al., 2014). Taking better account of pathogen diversity and reaching a more complete understanding of the factors that drive the evolution of fungal plant pathogens are critical to enhance sustainable disease control strategies, and to better predict the emergence and spread of novel harmful genotypes (Morris et al., 2009).

Venturia inaequalis is the ascomycete responsible for scab disease on apples (Malus spp.). Recent sampling and characterization of fungal populations of V. inaequalis on wild Malus relatives, and even on other host species belonging to the Rosaceae family, have revealed substantial genetic and phenotypic diversity of the fungus outside agricultural habitats (Le Cam et al., 2002; Gladieux et al., 2010; Lê Van et al., 2012; Leroy et al., 2013). The cultivated apple, Malus domestica, was domesticated in Central Asia from the wild apple Malus sieversii (Velasco et al., 2010; Cornille et al., 2012, 2014). Domesticated apples were then spread westward to Europe and eastward to China, and they entered into contact and hybridized with wild apple relatives, such as Malus sylostris in Europe, Malus baccata in the Himalayas and Siberia, and Malus orientalis in the Caucasus (Janick et al., 1996; Forsline et al., 2010; Cornille et al., 2013, 2015). Collections of V. inaequalis in the center of origin of its host and in M. x domestica orchards from all continents revealed that the fungus has followed its host during its domestication and world-wide spread (Gladieux et al., 2008, 2010; Lê Van et al., 2012). Apple scab populations from isolated stands of the wild M. sieversii in Kazakhstan exhibit both genetic and phenotypic differences from those sampled within or near M. domestica orchards (Gladieux et al., 2010; Lê Van et al., 2012).

Many of the previously characterized interactions between V. inaequalis and Malus spp. follow a ‘gene-for-gene’ model (Flor, 1942, 1955, 1971) where a pathogen effector triggers immunity in host genotypes carrying the matching resistance gene (Jones & Dangl, 2006). Seventeen such interactions for disease resistance have been recorded by Bus et al. (2011) and two resistance genes have been cloned (Belfanti et al., 2004; Schouten et al., 2014). The resistance gene Rvi6 – formerly known as Vf (Bus et al., 2011) – initially identified in the wild relative Malus floribunda has been the most widely introgressed in commercial varieties released by plant breeders. Malus floribunda is a putative hybrid between the Siberian and Japanese wild Malus species M. baccata and Malus toringo (Juniper & Mabberley, 2006), M. baccata being the source of the Rvi6 gene (Parisi et al., 1993; Dunemann et al., 2012). It is widely planted as an ornamental in gardens, urban areas, and along roads. Despite the scarce use of the Rvi6 resistance in European orchards, resistance has been rapidly overcome. The first occurrence of virulence on resistant varieties was reported in Germany in 1988 (Parisi et al., 1993), and thereafter in various other European countries (Gessler & Pertot, 2012). Apple scab symptoms on M. floribunda were reported in a private English garden as early as 1989, even before plantation of Rvi6 apple varieties in English orchards (Roberts & Crute, 1994). At the scale of northern France, population genetics analyses of microsatellite data revealed that all pathogen genotypes able to infect Rvi6 apple varieties shared a common origin, with limited genetic diversity relative to neighboring populations sampled on other varieties (Guérin & Le Cam, 2004; Guérin et al., 2007). The existence of strong differentiation between V. inaequalis from Rvi6 and V. inaequalis from non-Rvi6 apple varieties suggested limited gene flow (Guérin et al., 2007), which was surprising given the frequent sympatry of the pathogen populations, which are often present in the very same orchards. Further multilocus microsatellite genotyping and cross-inoculations suggested host specificity as the strongest and probably most efficient barrier to gene flow between emerging virulent populations and populations infecting non-Rvi6 apples. As the fungus mates within its host after successful infection, such high host specificity necessarily and automatically strongly restricts gene flow (Giraud et al., 2010; Servedio et al., 2011). With such a life cycle, strict host specificity has a pleiotropic effect on reproductive isolation, and it can alone prevent gene flow between populations from divergent habitats without requiring any form of assortative mating (Giraud et al., 2006). This kind of ‘magic trait’ scenario is one of the most favorable for ecological speciation, as the strength of divergent selection acting on the genes affecting reproductive isolation is maximal, and it is often claimed to be particularly common in fungal pathogens, driving the emergence of new populations infecting novel hosts (Gladeieux et al., 2014). What remains unknown is the origin of the Rvi6-virulent populations of apple scab, and in particular whether they emerged de novo from the populations infecting non-Rvi6 varieties or whether they are derived from a population in which Rvi6-virulent alleles were already present.

Here, we investigate the origin of Rvi6-virulent populations, using a comprehensive collection of apple scab samples collected in the center of origin of the pathogen on Kazakh mountain populations of the wild M. sieversii, in Europe on Rvi6 and non-Rvi6 varieties, and on M. floribunda, the wild progenitor of the Rvi6 resistance. Indeed, as V. inaequalis is also present on the Rvi6-resistant M. floribunda, it is a good candidate for the ancestor of populations infecting domestic Rvi6 varieties. More specifically, we used multilocus microsatellite genotyping and a combination of approximate Bayesian computation (ABC) (Beaumont et al., 2002) and Bayesian clustering to investigate the relationships among Rvi6-virulent populations infecting M. domestica and M. floribunda, and populations infecting other agricultural or nonagricultural hosts.
Materials and Methods

Data collection

We state that no permit was necessary to sample apple tree leaves containing the fungus *V. inaequalis* in both wild and agricultural locations in Kazakhstan and other countries. The field study did not involve endangered or protected species. We used 821 strains of *V. inaequalis* collected at the European scale and in Kazakhstan (Table 1). Strains from *Rvi6* varieties were collected in France in 2001, 2003 and 2004 (*n*= 177), in Sweden in 2005 (*n*= 76), in Poland in 2010 (*n*= 156) and in Denmark in 2003 (*n*= 66). Strains from non-*Rvi6* varieties were collected in France in 2000, 2003 and 2004 (*n*= 186), in Sweden in 2010 (*n*= 25), in Poland in 2010 (*n*= 94) and in Denmark in 2003 (*n*= 30). Twenty-six strains were collected on *Malus floribunda* (Siebold ex Van Houtte, 1864) in different countries in Europe between 1989 and 1998 (France, the UK, the Netherlands, Denmark and Switzerland). We also collected 60 strains on *Malus sieversii* (Ledeb.) at two sampling sites in Kazakhstan.

Genomic DNA was directly extracted from scabbed apple leaf or derived from isolated conidia according to the procedure previously described (Le Cam *et al.*, 2002; Guérin *et al.*, 2007). All samples were genotyped at eight microsatellite loci: 1tc1a, 1tc1b, 1tc1g and 1aac3b from Tenzer *et al.* (1999), V1c1/2, Vtg1/17/20 and Vicag8/42 from Guérin *et al.* (2004) and M42 following a method previously described (Guérin *et al.*, 2004). As sampling occurred during the asexual phase, identical multilocus genotypes were treated as clones and removed from the analysis.

Population structure

Summary statistics of genetic variation, including expected heterozygosity (*H*<sub>e</sub>) (Nei, 1987) and pairwise fixation indices (*F*<sub>st</sub>) (Weir & Cockerham, 1984), were estimated using *GENETIX* 4.05 (Belkhir *et al.*, 2004). Significant departures from the null hypothesis (*F*<sub>st</sub> = 0) were tested using a permutation procedure. Allelic richness was estimated using the method implemented in *ADZE* (Szpiech *et al.*, 2008). Population subdivision was inferred using a Bayesian model-based clustering method implemented in *STRUCTURE* v2.3 (Pritchard *et al.*, 2000; Falush *et al.*, 2003). We used a model of admixture with correlated allele frequencies. We set the burn-in to 500 000 followed by 1500 000 iterations. For each number *K* of populations, ranging from 2 to 10, five independent runs were performed to check for convergence. Outputs were processed using *CLUMPP* (greedy algorithm) (Jakobsson & Rosenberg, 2007).

Finally, a discriminant analysis of principal components (DPAC) was performed using the R package adegenet (v.1.4.2; Jombart, 2008).

Comparison of demographic models

We used the ABC approach, as implemented in the software *DIYABC* (Cornuet *et al.*, 2008), to infer the origin of the *Rvi6*-virus vectors. We simulated microsatellite data sets under five different scenarios (Fig. 1) using a generalized stepwise model (Estoup *et al.*, 2002) with two parameters: *µ*, the mean mutation rate among loci, and *P*, the mean parameter of the geometric distribution of the mutation event in terms of the number of repeats. Mean mutation rates were drawn from a uniform distribution bounded by extreme values of 10<sup>-6</sup> and 10<sup>-3</sup>. Per locus mutation rates were drawn from a gamma distribution (mean = *µ*; shape = 2). The parameter *P* was drawn from a uniform distribution with extreme values 0.1–0.3. In order to compare observed with simulated data, we used the following summary statistics: mean number of alleles per locus, mean diversities, mean size variance, genetics of comparisons between

<table>
<thead>
<tr>
<th>Sample</th>
<th>Country</th>
<th>Location</th>
<th>Host</th>
<th>Year of collection</th>
<th>Collector/provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>France</td>
<td>Pont Audemer</td>
<td>Rvi6</td>
<td>2001</td>
<td>FG and BLC</td>
</tr>
<tr>
<td>F2</td>
<td>France</td>
<td>Le Sacq</td>
<td>Rvi6</td>
<td>2000–2003</td>
<td>FG and BLC</td>
</tr>
<tr>
<td>F4</td>
<td>France</td>
<td>Yvrandes</td>
<td>Rvi6</td>
<td>2002–2003</td>
<td>FG and BLC</td>
</tr>
<tr>
<td>F5</td>
<td>France</td>
<td>Pont Audemer</td>
<td>Non-Rvi6</td>
<td>2000</td>
<td>FG and BLC</td>
</tr>
<tr>
<td>F6</td>
<td>France</td>
<td>Le Sacq</td>
<td>Non-Rvi6</td>
<td>2005</td>
<td>FG and BLC</td>
</tr>
<tr>
<td>F7</td>
<td>France</td>
<td>La Folletière</td>
<td>Non-Rvi6</td>
<td>2004</td>
<td>FG and BLC</td>
</tr>
<tr>
<td>F8</td>
<td>France</td>
<td>Yvrandes</td>
<td>Non-Rvi6</td>
<td>2003</td>
<td>FG and BLC</td>
</tr>
<tr>
<td>SE1</td>
<td>Sweden</td>
<td>Alnarp</td>
<td>Rvi6</td>
<td>2010</td>
<td>HN</td>
</tr>
<tr>
<td>SE2</td>
<td>Sweden</td>
<td>Alnarp</td>
<td>Non-Rvi6</td>
<td>2005</td>
<td>HN</td>
</tr>
<tr>
<td>PL1</td>
<td>Poland</td>
<td>Pęczniew, Rokietnica, Lublin Wola Branicka</td>
<td>Rvi6</td>
<td>2010</td>
<td>MM</td>
</tr>
<tr>
<td>PL2</td>
<td>Poland</td>
<td>Pęczniew, NoweRowiska</td>
<td>Non-Rvi6</td>
<td>2010</td>
<td>MM</td>
</tr>
<tr>
<td>DK1</td>
<td>Denmark</td>
<td>Aarslev</td>
<td>Rvi6</td>
<td>2003</td>
<td>HLP</td>
</tr>
<tr>
<td>DK2</td>
<td>Denmark</td>
<td>Aarslev</td>
<td>Non-Rvi6</td>
<td>2003</td>
<td>HLP</td>
</tr>
<tr>
<td>FLO</td>
<td>Europe</td>
<td>a</td>
<td>Rvi6</td>
<td>1989–1998</td>
<td>CP and FL</td>
</tr>
<tr>
<td>CAM</td>
<td>Kazakhstan</td>
<td>Close to Almaty</td>
<td>Non-Rvi6</td>
<td>2006</td>
<td>CP and FL</td>
</tr>
</tbody>
</table>

FG, F. Guérin; BLC, B. Le Cam; PG, P. Gladieux; HN, H. Nybom; MM, M. Michalecka; HLP, H. Lindhard Perdersen; CP, C. Peix; FL, F. Laurens. CAM, Central Asia Mountains; FLO, *Malus floribunda*.

*Strains were collected in different European countries such as France, the UK, Spain, the Netherlands, Luxemburg, Germany and Switzerland.
populations \((F_D)\) and genetic distance \((\delta_{st})\). A polychotomous logistic regression was used to estimate the relative posterior probability of each model using 0.1% of the simulated data sets closest to the observed data. We used the limiting distribution of the maximum likelihood estimators to estimate confidence intervals for the posterior probabilities. We then used a local linear regression to estimate the posterior distributions of parameters under the most likely model. The 0.1% simulated data sets closest to the observed data were used for the regression, after the application of a logit transformation to parameter values.

**Results**

Patterns of genetic variation

We investigated the mode and tempo of emergence of *V. inaequalis* populations infecting apple varieties carrying the *Rvi6* resistance gene. We used multilocus microsatellite genotyping of 821 strains of *V. inaequalis* collected in the center of origin of the pathogen on mountain populations of the wild *M. sieversii* (the Central Asia Mountains (CAM) population), and in Europe on *Rvi6* cultivated apple varieties (the ‘*Rvi6*’ population), on non-*Rvi6* varieties (the ‘non-*Rvi6*’ population), and on the wild *M. floribunda*, the wild progenitor of the *Rvi6* resistance (the ‘FLO’ population).

Genetic variation mean was lowest within the *Rvi6* population (unbiased heterozygosity: \(H_u = 0.38 \pm 0.11\); allelic richness: \(A = 4.46 \pm 1.92\); Table 2), and it was only slightly higher in the FLO population sampled on *M. floribunda* (\(H_u = 0.42\); \(A = 3.62\)). This is consistent with histories of founder events, associated with epidemic spread across Northern Europe for *Rvi6*-virulent strains, and with the introduction of infected *M. floribunda* crabapples in Europe from an unknown source. Genetic variation was the highest in the non-*Rvi6* (\(H_u = 0.67 \pm 0.06\); \(A = 7.68 \pm 1.89\)) and CAM populations on wild *M. sieversii* (CAM) (\(H_u = 0.64\); \(A = 11\)). The very close levels of genetic variation in these two nonemerging populations of apple scab mirror patterns of genetic diversity in wild and domesticated apple trees, for which there is no evidence of a domestication bottleneck (Cornille et al., 2012).

**Population subdivision**

Model-based Bayesian clustering using STRUCTURE revealed a partition between strains sampled on *M. floribunda* and *Rvi6* varieties on the one hand and strains sampled on non-*Rvi6* and wild CAM *Malus* on the other hand (Fig. 1). The clustering of

**Table 2** Polymorphism and diversity among eight microsatellite loci at the 14 samples of *Venturia inaequalis*

<table>
<thead>
<tr>
<th>Samples</th>
<th>(n^a)</th>
<th>(Hnb^b)</th>
<th>(A^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1 <em>Rvi6</em></td>
<td>22</td>
<td>0.29</td>
<td>2.37</td>
</tr>
<tr>
<td>F2 <em>Rvi6</em></td>
<td>42</td>
<td>0.27</td>
<td>2.63</td>
</tr>
<tr>
<td>F3 <em>Rvi6</em></td>
<td>79</td>
<td>0.39</td>
<td>5.00</td>
</tr>
<tr>
<td>F4 <em>Rvi6</em></td>
<td>29</td>
<td>0.26</td>
<td>2.25</td>
</tr>
<tr>
<td>F5 non-<em>Rvi6</em></td>
<td>73</td>
<td>0.72</td>
<td>11.25</td>
</tr>
<tr>
<td>F6 non-<em>Rvi6</em></td>
<td>30</td>
<td>0.71</td>
<td>7.00</td>
</tr>
<tr>
<td>F7 non-<em>Rvi6</em></td>
<td>57</td>
<td>0.72</td>
<td>9.00</td>
</tr>
<tr>
<td>F8 non-<em>Rvi6</em></td>
<td>26</td>
<td>0.68</td>
<td>6.25</td>
</tr>
<tr>
<td>Sweden</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE1 <em>Rvi6</em></td>
<td>76</td>
<td>0.22</td>
<td>3.25</td>
</tr>
<tr>
<td>SE2 non-<em>Rvi6</em></td>
<td>25</td>
<td>0.64</td>
<td>5.25</td>
</tr>
<tr>
<td>Poland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL1 <em>Rvi6</em></td>
<td>156</td>
<td>0.55</td>
<td>8.00</td>
</tr>
<tr>
<td>PL2 non-<em>Rvi6</em></td>
<td>94</td>
<td>0.71</td>
<td>10.37</td>
</tr>
<tr>
<td>Denmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DK1 <em>Rvi6</em></td>
<td>66</td>
<td>0.42</td>
<td>3.66</td>
</tr>
<tr>
<td>DK2 non-<em>Rvi6</em></td>
<td>26</td>
<td>0.51</td>
<td>4.66</td>
</tr>
<tr>
<td>Europe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLO <em>Rvi6</em></td>
<td>30</td>
<td>0.42</td>
<td>3.62</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAM non-<em>Rvi6</em></td>
<td>60</td>
<td>0.64</td>
<td>11.00</td>
</tr>
</tbody>
</table>

\(\text{CAM}, \text{Central Asia Mountains}; \text{FLO}, \text{Malus floribunda}.\)

\(^a\)Sample size.

\(^b\)Nonbiased expected haplotype diversity.

\(^c\)Allelic richness.
Research

Values below the diagonal indicate $F_{st}$ estimates between populations; values in brackets represent the limits of the 95% confidence interval. Values in italic on the diagonal indicate within-population variation estimated by expected heterozygosity (Nei, 1987). All $F_{st}$ estimates were significantly different from zero ($P < 0.001$).

Table 3 Pairwise differentiation ($F_{st}$) between the four populations of Venturia inaequalis

<table>
<thead>
<tr>
<th></th>
<th>Agricultural Rvi6</th>
<th>Agricultural non-Rvi6</th>
<th>Wild Rvi6</th>
<th>Wild non-Rvi6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural Rvi6</td>
<td>0.509</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agricultural non-Rvi6</td>
<td>0.108 (0.054–0.164)</td>
<td>0.747</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild Rvi6</td>
<td>0.033 (0.002–0.069)</td>
<td>0.103 (0.073–0.132)</td>
<td>0.420</td>
<td></td>
</tr>
<tr>
<td>Wild non-Rvi6</td>
<td>0.264 (0.108–0.387)</td>
<td>0.092 (0.052–0.146)</td>
<td>0.233 (0.135–0.321)</td>
<td>0.646</td>
</tr>
</tbody>
</table>

Fig. 2 Multivariate clustering of the Venturia inaequalis strains using discriminant analysis of principal components. As indicated in the key at the top right of the panel, agricultural Rvi6 strains are in light blue, wild Rvi6 in purple, agricultural non-Rvi6 in orange and wild non-Rvi6 in green. Axis 1 (horizontal) differentiates Rvi6 from non-Rvi6 strains, and axis 2 (vertical) separates wild from agricultural non-Rvi6 strains.
and the two clusters corresponding to non-\textit{Rvi6} strains (orange and green clusters in Fig. 1) were summed. Genotypes with membership proportions between 0.30 and 0.70 for the pathogen population clusters found on their host of origin were considered as admixed; genotypes with membership proportions > 0.70 for the other population clusters were considered as migrants. Using this approach, 8.57% and 12.57% of genotypes were identified as admixed, and 4.03% and 4.84% of genotypes were identified as migrants on \textit{Rvi6} and non-\textit{Rvi6} hosts, respectively.

Comparison of demographic models

The different scenarios compared using ABC are presented in Fig. 3. The four scenarios can be summarized as follows.

Model 1: Parallel evolution We assume here that virulence against \textit{Rvi6} might have emerged independently in nonagricultural and agricultural environments. Thus, strains virulent against \textit{Rvi6} resistance emerge from non-\textit{Rvi6} populations in orchards whereas populations FLO and CAM on wild apple trees derived from a common ancestor somewhere in Asia.

The three remaining scenarios describe a single emergence from the wild of populations virulent against \textit{Rvi6} domestic apple trees.

Model 2: Emergence from the wild without gene flow \textit{Rvi6} populations share a recent common ancestor with the FLO population and remain isolated from other populations.

Model 3: Admixture 1 \textit{Rvi6} populations are formed by admixture between a ghost population that shared a recent common ancestor with the FLO population and agricultural non-\textit{Rvi6} populations.

Model 4: Admixture 2 \textit{Rvi6} populations were formed following an admixture event between the FLO population and the non-\textit{Rvi6} population.

The two last models (models 3 and 4) differ from each other by the introduction in model 3 of a ghost population that represents a nonsampled population related to the FLO population (e.g. an unknown environmental reservoir). Following model 3, this ghost population – and not the FLO population – admixed with agricultural non-\textit{Rvi6} populations. Conversely, in model 4, the FLO population is directly involved in admixture with agricultural non-\textit{Rvi6} populations.

Among the four scenarios implemented in \textsc{divABC}, scenario 2 was clearly the most supported (posterior probability, \(P = 0.9598\); 95% confidence interval (CI) 0.9554–0.9642; Table 4), with all other models having very low posterior probabilities (\(P < 0.02\)). This scenario assumes that \textit{Rvi6} and non-\textit{Rvi6} populations were independently derived from the two nonagricultural apple species \textit{M. floribunda} and \textit{M. sieversii}, respectively. Testing for confidence in model choice (Supporting Information Table S1) showed strong support for model 2. Parameter estimates under model 2 are presented in Table 5. Estimates of effective population sizes were on the order of \(10^4\) for all populations (Table 5). Nevertheless, effective sizes of \textit{Rvi6} populations on agricultural and wild apple trees (respectively 36 300, 95% CI 9200–63 000, and 25 300, 95% CI 11 000–67 400) are smaller than those of non-\textit{Rvi6} populations on agricultural and wild apple trees (respectively 68 500, 95% CI 24 900–121 000, and 86 000, 95% CI 27 200–170 000; Table 4). Assuming a generation time of 1 yr (MacHardy, 1996), the divergence between the CAM/non-\textit{Rvi6} and FLO/\textit{Rvi6} populations was estimated to be between \(c. 8000\) and 50 000 yr before present (bp). Divergence time estimates between non-\textit{Rvi6} populations and their central Asian source population (CAM) ranged from \(c. 1900\) to 4000 yr BP (95% CIs), matching previous estimates obtained using an

---

**Table 4** Relative posterior probabilities \((P)\) for the four historical models compared using approximate Bayesian computations

<table>
<thead>
<tr>
<th>Model</th>
<th>(P^a)</th>
<th>CI 2.5</th>
<th>CI 97.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.006</td>
<td>0.005</td>
<td>0.007</td>
</tr>
<tr>
<td>2</td>
<td>0.960</td>
<td>0.955</td>
<td>0.964</td>
</tr>
<tr>
<td>3</td>
<td>0.016</td>
<td>0.014</td>
<td>0.018</td>
</tr>
<tr>
<td>4</td>
<td>0.018</td>
<td>0.016</td>
<td>0.020</td>
</tr>
</tbody>
</table>

CI, confidence interval.

*aProbabilities are derived from a logistic approach, using the 40 000 closest values for each model.*
isolation with migration (IM) model (Gladiex et al., 2010). Divergence time between emerging Rvi6 populations and non-agricultural populations associated with *M. floribunda* (FLO) ranged from c. 70 to 160 yr BP, a timeframe that includes the date of introgression of the Rvi6 resistance gene from *M. floribunda* into domestic apple varieties.

While the common ancestry of domestic and wild Rvi6 populations is clear, our data set did not allow accurate testing of the source–sink relationships between these two populations. Indeed, the scenario of wild Rvi6 populations as a source for agricultural ones as well as its reciprocal scenario (agricultural as a source for wild) provided poor confidence in scenario choice (*P* < 0.2). In the same way, the data set did not permit inclusion of bottlenecks at divergence times in the different scenarios. Divergence times could therefore be overestimated as a result of strong drift leading to high *F_{IS}*. 

**Discussion**

Nonagricultural hosts as the source of Rvi6-virulent populations

The processes underlying the emergence of new virulent populations of fungal plant pathogens remain largely unknown. It has generally been assumed that new favorable mutations arise *de novo* from populations already present in the agrosystem, in response to the introduction of new resistant hosts. The alternative hypothesis is that new favorable mutations are selected from standing variation, allowing quick response to changing selective pressures exerted by resistant varieties. However, these two hypotheses have rarely been challenged, probably because of a lack of samples from potential source populations in agrosystems or nonagricultural habitats (Burdon & Thrall, 2008; Monteil et al., 2013). Here, we used a comprehensive sample of *V. inaequalis* populations from wild and cultivated hosts to investigate the processes underlying the emergence of new populations infecting varieties carrying the Rvi6 resistance gene. We found that the populations infecting hosts carrying or not the Rvi6 resistance gene form two distinct lineages that diverged several thousands of generations ago, without subsequent detectable gene flow. Although we could not test specifically for a source–sink model involving populations infecting wild and cultivated hosts because of insufficient statistical power to distinguish competing demographic scenarios, this finding is nonetheless consistent with a scenario in which populations infecting Rvi6 varieties emerged from standing variation present on the nonagricultural progenitor of the Rvi6 resistance, *M. floribunda*. We favor this hypothesis because the alternative – which assumes that populations infecting Rvi6 varieties are the source of populations infecting *M. floribunda* – would imply that Rvi6-virulent populations coexisted in sympatry for thousands of generations with populations associated with non-Rvi6 hosts despite the absence of the Rvi6 resistance gene in apple agrosystems. The ornamental Japanese crabapple *M. floribunda* was first used in 1914 as a donor of the Rvi6 gene for resistance against apple scab in cultivated varieties. Strains overcoming the Rvi6 resistance were first observed on *M. floribunda* in 1988 in a germplasm repository in Germany (Parisi et al., 1993) and in 1989 in the UK, even before the first introduction of Rvi6 varieties in British orchards (Roberts & Crute, 1994). A likely scenario is that a small number of Rvi6-virulent individuals have immigrated to Europe with *M. floribunda*, possibly from Japan (Juniper & Mabberley, 2006), and they remained cryptic before the deployment of apple varieties obtained by hybridization between *M. floribunda* and *M. domestica*.

Previous works revealed that all pathogen genotypes able to infect Rvi6 apple varieties share a common origin and are strongly differentiated from populations from non-Rvi6 apples (Guérin & Le Cam, 2004; Guérin et al., 2007). Temporal sampling confirmed the lack of gene flow between sympatric populations from Rvi6 and non-Rvi6 hosts, and cross-inoculation tests identified strong selection against immigrants (i.e. host
specificity) from different host varieties as an efficient barrier to gene flow between local and emerging populations (Gladieux et al., 2011; Lê Van et al., 2012). Host specificity in pathogens that mate within their hosts pleiotropically induces assortative mating, because only pathogens able to infect the same individual host can then mate and exchange genes (Giraud et al., 2010; Servedio et al., 2011). In the present study, we used comparisons of demographic models to demonstrate that the lack of gene flow between populations from Rv16 and non-Rv16 hosts also holds on a larger scale, between populations sampled at multiple sites, on multiple host species and varieties.

Estimates of divergence time and admixture levels between Rv16 and non-Rv16 populations were consistent with divergence in allopatry, and comparisons among models did not support a scenario in which Rv16 populations emerged from sympatric non-Rv16 populations. Allopatry, however, does not obviously preclude a role of divergent selection as a barrier to gene flow between Rv16 and non-Rv16 lineages. Adaptive divergence between lineages is even facilitated in allopatry, as it can proceed unimpeded by gene flow. Allopatric V. inaequalis populations coevolving independently with divergent apple species could readily evolve differentiated sets of pathogenicity factors preventing infection of their reciprocal hosts. The relative rarity of migrants and F1 hybrids in our data set, despite interfertility, confirms that the main barrier to gene flow occurs before host infection and therefore before the stage of mating, and that immigrant inviability (i.e. host specificity) is the main barrier to gene flow in this system. If nonecological prezygotic (e.g. assortative mating by mate choice) or early postzygotic barriers (e.g. hybrid inviability because of genetic incompatibilities) were the main barriers to gene flow, migrants would have been frequent, and F1 hybrids absent. If late postzygotic barriers (e.g. F1 sterility) were the main barriers to gene flow, substantial numbers of both migrants and F1 hybrids would have been observed, without nevertheless pervasive introgression between the two populations. Proportions of F1 hybrids and migrants, however, are higher at the European scale (this study) than the scale of the north of France (54), suggesting that the barrier of immigrant inviability may be eroding, as a consequence of possible mating events between Rv16 and non-Rv16 strains on host varieties susceptible to both kinds of pathogen.

Further studies are required to assess whether other intrinsic or extrinsic postzygotic barriers could complete reproductive isolation and explain why the observed hybrids do not lead to detectable gene flow across years. The two lineages were estimated to have diverged several thousands of years ago, providing sufficient time for the accumulation of multiple Dobzhansky–Muller incompatibilities, that is, hybrid dysfunctions caused by epistatic interactions between alleles at different loci (Barton & Hewitt, 1985; Orr & Turelli, 2001). The spread of Rv16 populations amid already established non-Rv16 populations can be viewed as a case of secondary contact. Endogenous (i.e. incompatibilities) and exogenous (i.e. genes under divergent selection) barriers to gene flow could be identified using genome scans of hybrid zones (Bierne et al., 2011), such as apple varieties that could be infected by both the Rv16 and non-Rv16 lineages, should such a situation occur.

**Implications for disease management**

A few empirical examples have shown the importance of interactions between plant pathogens of agrosystems and nonagricultural environments, as reviewed by Burdon & Thrall (2008). In such examples, nonagricultural reservoirs are often found to provide new strains that are pathogenic on crops. For instance, two clonal lineages of Fusarium oxysporum f. sp. vasinfectum, infecting cotton (Gossypium hirsutum L.) crops in Australia, were found to originate from local wild cotton areas (Wang et al., 2010). Here, our study shows that resistance traits introduced from wild relatives can act as a gateway for the agrosystem, facilitating the introduction of pathogen populations carrying new alleles with detrimental effects on agricultural production. This highlights the point that the current focus on coevolution of pathogens with their domesticated plants and current paradigms concerning the factors that drive pathogen evolution need to be expanded to develop a more holistic approach including biological interactions with multiple host species embedded in agricultural, natural and urban ecosystems (Morris et al., 2009). As noted by Monteil et al. (2013), the importance of environmental reservoirs is widely accepted for animal and human pathogens, and current concepts from the fields of medical and veterinary epidemiology could serve as a useful foundation for novel hypotheses about the rapid adaptation of fungal plant pathogens in agrosystems.

Given the limited genetic diversity of elite cultivated gene pools for many major crops, a high strategic priority for varietal improvement is to enrich the cultivated gene pools by incorporating favorable genomic features from wild relatives, including resistance genes (Feuillet et al., 2007). Our study suggests that exploiting wild relatives for new resistance alleles is doomed to failure if the diversity and distribution of pathogen populations in nonagricultural disease reservoirs is not taken into account in screenings for new resistance traits and subsequent strategies of resistance deployment. Venturia inaequalis is a common pathogen of nonagricultural hosts such as pyracantha (Pyracantha coccinea), loquat (Eriobotrya japonica), and various Sorbus species (Le Cam et al., 2002; Gladieux et al., 2010), suggesting important disease reservoirs that could also be the source of future disease emergence and should therefore be included in resistance screenings and risk-prediction frameworks. Leroy et al. (2014) underlined the necessity to protect newly introgressed resistance genes from pathogen populations infecting nonagricultural hosts by taking drastic sanitary measures targeting specifically wild relatives to prevent future introductions of virulent populations.

We showed that the breakdown of the Rv16 resistance gene was caused by the emergence of a population that is reproducively isolated from local populations infecting non-Rv16 trees. No matter what the cause, the existence of reproductive barriers between emerging and native populations is a double-edged sword. Such barriers impede the infusion of virulence alleles in local populations and thereby decrease the risk of transmission to areas that remained free from virulent strains, and in the meantime facilitate the rapid spread of emerging populations by limiting competition and maladaptive hybridization with populations.
adapted to different hosts. In 40 yr of monitoring in experimental orchards in different areas of the world, *Rvi6* apples remained free of apple scab symptoms, and the *Rvi6* resistance was seen as an example of sustainable resistance (Crosby et al., 1992). The *Rvi6* resistance actually remains efficient in commercial orchards outside Europe (Gessler & Pertot, 2012), suggesting that the mutation of the corresponding virulence gene may incur a high fitness cost for the pathogen, preventing emergence from *de novo* mutations selected from the populations infecting non-*Rvi6* hosts. It is therefore crucial to prevent the use of varieties that would be host to both *Rvi6* and non-*Rvi6* pathogen populations. This would increase the chance of recombination between the two populations, and thus potentially allow escape of the *Rvi6* virulence gene, facilitating introductions in areas where the resistance gene remains efficient.

**Acknowledgements**

We thank Dr H. Nyblom for providing samples from Sweden. We also thank Marie Noëlle Bellanger for help with strain curation and the ANAN platform (SFR QUASAV) for genotyping.

**Author contributions**

M.D.G., C.L. and B.L.C. designed the research. F.G., P.G. and M.M. performed the genotyping. C.L. and M.D.G. performed the statistical analyses. H.L-P. provided samples. C.L., M.D.G., T.L., F.G., P.G. and B.L.C. wrote the manuscript.

**References**


**Supporting Information**

Additional supporting information may be found in the online version of this article.

**Table S1** Confidence of the four different models

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.