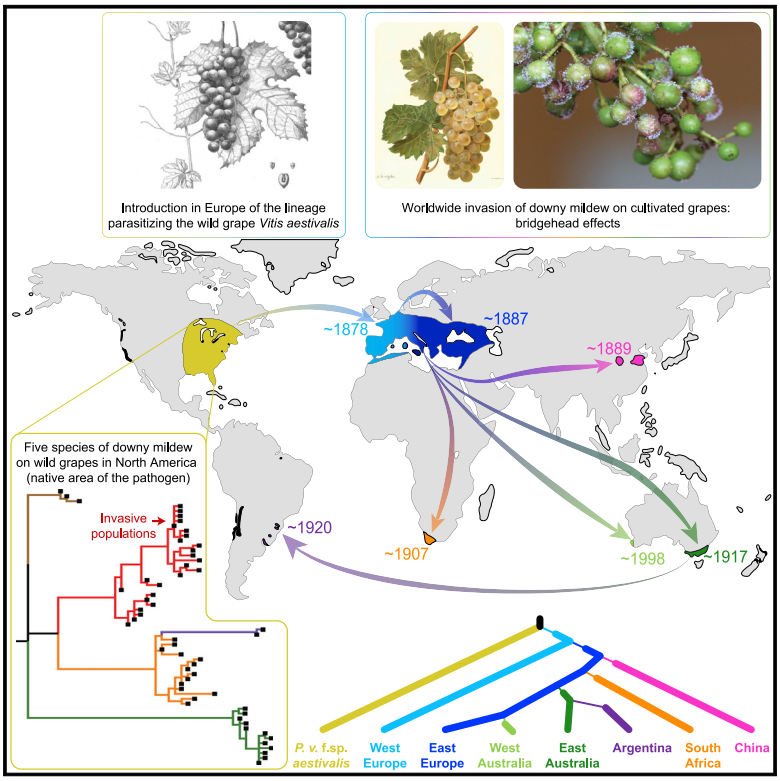
Article

Current Biology

Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*

Graphical abstract



Highlights

- Population genetics approaches infer the origin and routes of *P. viticola* invasion
- All invasive populations belong to only one of the five North American lineages
- A first introduction occurred into Europe in the 1870s with wild grape import
- European populations served as source for introductions into vineyards worldwide

Authors

Michael C. Fontaine, Frédéric Labbé, Yann Dussert, Laurent Delière, Sylvie Richart-Cervera, Tatiana Giraud, François Delmotte

Correspondence

michael.fontaine@cnrs.fr (M.C.F.), francois.delmotte@inrae.fr (F.D.)

In brief

Downy mildew is a devastating grapevine disease. Fontaine et al. show that invasive populations worldwide originate from the North American lineage from wild *Vitis aestivalis* grapes. A first introduction of the pathogen into Europe occurred in the 1870s, from where further invasions occurred into other wine-producing areas worldwide.

Fontaine et al., 2021, Current Biology *31*, 1–12
May 24, 2021
https://doi.org/10.1016/j.cub.2021.03.009

CellPress

Please cite this article in press as: Fontaine et al., Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.03.009

Current Biology

CellPress

Article

Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*

Michael C. Fontaine,[1,2,3,5,8,10,11,*] Frédéric Labbé,[1,2,5,6] Yann Dussert,[4,7] Laurent Delière,[4] Sylvie Richart-Cervera,[4] Tatiana Giraud,[2] and François Delmotte[4,9,*]

[1]Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, P.O. Box 11103 CC, Groningen, the Netherlands
[2]Ecologie Systématique et Evolution, UMR 8079, Université Paris-Saclay, CNRS, AgroParisTech, Orsay 91400, France
[3]Laboratoire MIVEGEC (Université de Montpellier, CNRS 5290, IRD 229) et Centre de Recherche en Écologie et Évolution de la Santé (CREES), Institut de Recherche pour le Développement (IRD), 34394 Montpellier, France
[4]SAVE, INRAE, Bordeaux Sciences Agro, Université de Bordeaux, 33140 Villenave d'Ornon, France
[5]These authors contributed equally
[6]Present address: Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA
[7]Present address: School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
[8]Twitter: @MikaFontaine1
[9]Twitter: @PacoDelmotte
[10]Twitter: @GenEcoEvo
[11]Lead contact
*Correspondence: michael.fontaine@cnrs.fr (M.C.F.), francois.delmotte@inrae.fr (F.D.)
https://doi.org/10.1016/j.cub.2021.03.009

SUMMARY

Europe is the historical cradle of viticulture, but grapevines (*Vitis vinifera*) have been increasingly threatened by pathogens of American origin. The invasive oomycete *Plasmopara viticola* causes downy mildew, one of the most devastating grapevine diseases worldwide. Despite major economic consequences, its invasion history remains poorly understood. We analyzed a comprehensive dataset of ~2,000 samples, collected from the most important wine-producing countries, using nuclear and mitochondrial gene sequences and microsatellite markers. Population genetic analyses revealed very low genetic diversity in invasive downy mildew populations worldwide and little evidence of admixture. All the invasive populations originated from only one of the five native North American lineages, the one parasitizing wild summer grape (*V. aestivalis*). An approximate Bayesian computation-random forest approach allowed inferring the worldwide invasion scenario of *P. viticola*. After an initial introduction into Europe, invasive European populations served as a secondary source of introduction into vineyards worldwide, including China, South Africa, and twice independently, Australia. Only the invasion of Argentina probably represents a tertiary introduction, from Australia. Our findings provide a striking example of a global pathogen invasion resulting from secondary dispersal of a successful invasive population. Our study will also help designing quarantine regulations and efficient breeding for resistance against grapevine downy mildew.

INTRODUCTION

Global changes (e.g., climate warming and international exchanges) are favoring increases in the numbers of emerging diseases caused by invasive pathogens on crops worldwide, incurring substantial economic, social, and environmental costs.[1–4] Infamous recent examples include the emergence of new races of the stem rust fungus in Eastern Africa[5] and Europe[6,7] and of the fungus causing wheat blast disease in Bangladesh,[8] both threatening wheat production and becoming invasive. Most emerging diseases result from biological invasions bringing the native host-parasite association back together after crop introduction into new areas, or host shifts following pathogen introductions.[9,10] An understanding of the evolutionary processes responsible for emerging crop diseases is important for preventing further devastating biological invasions and for controlling introduced populations. This requires elucidation of the invasion mechanisms, pathways, and demographic processes occurring during pathogen invasions (e.g., bottlenecks and hybridization). Important questions include whether pathogen invasions result from host shifts, whether limited genetic variation has been introduced, whether multiple introductions and admixture are required for successful invasions,[10] and whether the invaded areas are colonized directly from native populations or whether an initial successful invasive population serves as the source for secondary introductions.

Grapevine (*Vitis vinifera* subsp. *sativa* L.) is one striking case of crop threatened by invasive pests. This emblematic crop has a prominent place in the history of European civilization. Grapevine domestication started 8,000 years ago from

Please cite this article in press as: Fontaine et al., Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.03.009

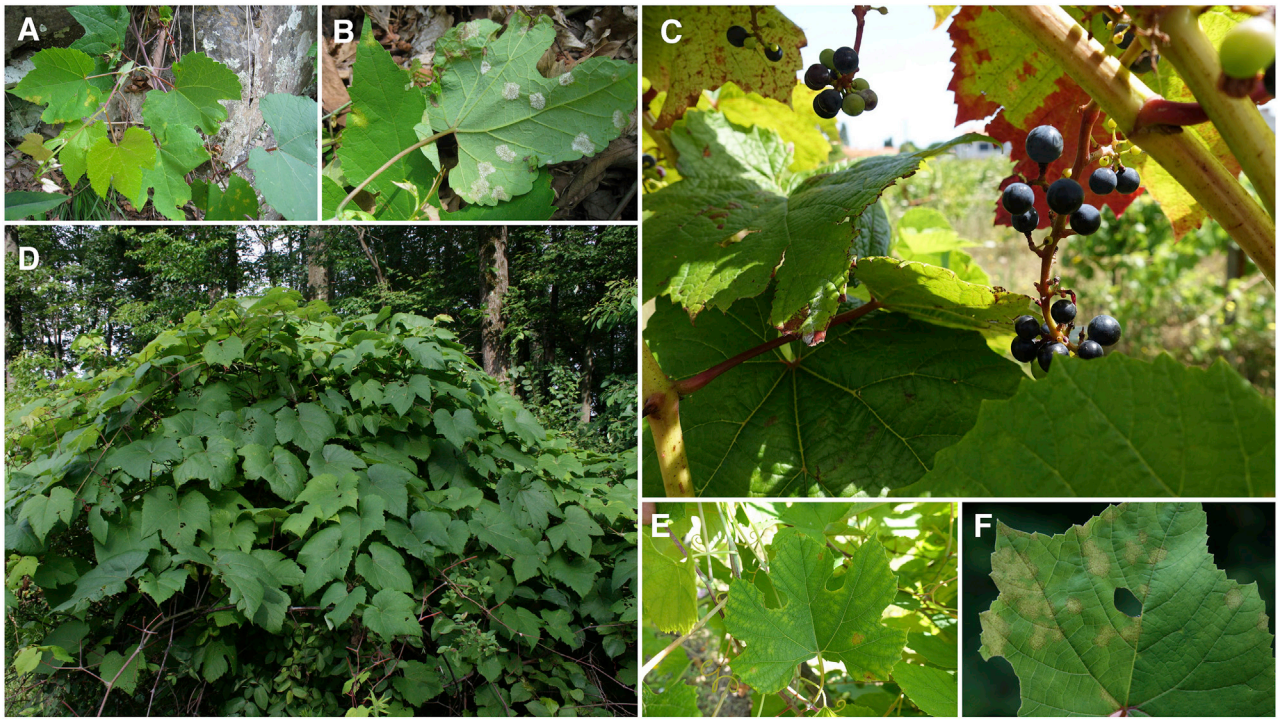


Figure 1. Grapevine downy mildew pathogen, *Plasmopara viticola*, on wild *Vitis aestivalis*

(A, B, E, and F) Typical pale-yellow lesions (oil spots) on a diseased *V. aestivalis* leaf at spring.

(C) Cluster of the wild plant.

(D) Infected plant on the edge of the forest.

See also Figure S1.

populations of *V. vinifera* subsp. *sylvestris* around the Caspian sea[11,12] and cultivated varieties were later introduced to the Mediterranean region.[13–16] While Europe is the historical cradle of viticulture, grapevine has been increasingly cultivated outside its traditional area since the 17th century (e.g., in South and North America, Australia, New Zealand, South Africa, and more recently Northeastern China). Grapevine has more recently been threatened by several devastating diseases caused by invasive pathogens native from North America, such as grapevine powdery mildew caused by *Erysiphe necator* and the soil-borne aphid phylloxera;[17] these diseases were likely introduced by scientists who established collections of American wild *Vitis* species in botanical gardens at the beginning of the 19th century. In the 1860s, the ''phylloxera plague'' destroyed most vineyards in Europe as no resistance had been identified in the cultivated *V. vinifera* germplasm.[18,19] Viticulture was rescued by grafting European cultivars onto resistant rootstocks bred at the end of the 19th century or at the beginning of the 20th century by hybridizing various wild North American *Vitis* species.

Grapevine downy mildew (Figure 1), one of the most damaging diseases of grapevine, is also caused by an invasive pathogen of North American origin, *Plasmopara viticola*.[18,20] This oomycete was introduced into Europe in the 1870s,[21] probably with the wild American *Vitis* species that were imported into Europe in attempts to identify species or generate hybrids that would be resistant to phylloxera. After its first description in France (1878 in Coutras), the grapevine downy mildew disease rapidly reached Southern and Central Europe and was soon reported in nearly all wine-producing countries worldwide.[17,22]

Advances in *P. viticola*-omics[23–28] and population genetic studies[29–44] have improved our knowledge of the grapevine downy mildew pathosystem. In its native range, five cryptic species (also called *formae speciales*) have recently been identified in the *P. viticola* species complex, with genetic differentiation and contrasting host ranges on various *Vitis* and related species.[42,43] The five *P. viticola formae speciales* (f. sp.) are found on wild *Vitis* species across North America,[42,43] and it remains unknown which lineages were responsible for grapevine downy mildew invasions in vineyards across the world. In most temperate regions, *P. viticola* populations present widespread footprints of recombination, indicating the occurrence of frequent sexual reproduction.[30,33,35,36,38,40,45] European invasive populations display little genetic diversity and have a weak but significant population structure at the continental scale.[31,34] Despite some notable advances, however, these studies have been restricted to a small number of countries and were performed with different markers, making it difficult to develop a comprehensive understanding of the pathways of *P. viticola* invasion worldwide.

In this study, we used phylogenetic and population genetic approaches, together with scenario testing by approximate Bayesian computation (ABC), to infer the origin and routes of

Please cite this article in press as: Fontaine et al., Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.03.009

P. viticola invasion worldwide. We analyzed almost 2,000 *P. viticola* samples, collected from wild and cultivated grapes in Northeast America, and from the main grape-growing regions in which grapevine downy mildew occurs. Using nuclear and mitochondrial gene sequences, we found that all invasive grapevine downy mildew populations worldwide belonged to a small clade of the species *P. viticola* f. sp. *aestivalis*, which parasitizes the *V. aestivalis* summer grape in Northeast America (Figure 1). Using microsatellite data from *P. viticola* f. sp. *aestivalis* samples, we found genetically differentiated populations in the different wine-producing continents, with little admixture between them and a low genetic diversity in invaded areas. The European populations had the highest level of diversity of all invasive populations and harbored all but one of the haplotypes present in other invaded areas. Using ABC scenario testing, we confirmed that *P. viticola* f. sp. *aestivalis* was first introduced into Western Europe, whence it spread to Central and Eastern Europe. The successful invasive populations in Europe then served as the source for secondary introductions into other grape-growing regions of the world, such as Northeast China, South Africa, and Australia. A third bridgehead effect occurred later, with the spread of the disease from Southeastern Australia to Argentina. All grapevine downy mildew invasions therefore stem from an initial single introduction event in Europe, followed by secondary and tertiary introductions via bridgehead effects. These introductions were probably mediated by the transfer of grapevine material by humans, as settlers imported European cultivars for the establishment of "New World" vineyards during the 19th century.[46] Our findings of a strong bottleneck following the introduction into Europe and the common origin of all introduced populations worldwide provide essential knowledge for guiding breeding for resistance to grapevine downy mildew. The identification of these historical pathways also improves our understanding of biological pest and pathogen invasions.

RESULTS

The cryptic species *Plasmopara viticola* f. sp. *aestivalis* is the origin of all invasive downy mildew populations worldwide

We reconstructed the genetic relationships between an extensive set of *P. viticola* strains collected from wild or cultivated grape species in native areas and areas of introduction, by sequencing DNA fragments from the mitochondrial cytochrome *b* (*cytb*, n = 1,299 samples), β-tubulin (*tub*, n = 424), and ribosomal 28S (*r28S*, n = 536) genes (Figures 2 and S1; Table S1). The β-tubulin tree provided the highest resolution (Figure 2A; with 56 distinct haplotypes versus 32 haplotypes for *cytb* and 9 haplotypes for *r28S*; Figure S2). Nevertheless, all the trees presented highly differentiated lineages with contrasting host ranges (Figures 2A and S2), consistent with the previously demonstrated existence of five cryptic species.[42,43] Three *formae speciales* were found on cultivated grapes in North America, but only *P. viticola* f. sp. *aestivalis* was found in introduced populations worldwide (Figures 2 and S2). In North America, this *forma specialis* was found on *V. aestivalis* and on *V. labrusca sensu lato* (i.e., including *V. labrusca* and its main artificial hybrids), two *Vitis* species that have recently diverged.[47]

A host shift from *V. aestivalis* at the origin of the invasive downy mildew populations and population subdivision in invasive populations

We used highly polymorphic microsatellite markers to obtain a finer resolution of the population structure within *P. viticola* f. sp. *aestivalis*, to elucidate its worldwide population structure and invasion routes. We genotyped 1,974 diploid strains with eight microsatellite markers (Table S2), which, together with the three sequenced fragments (Table S1), revealed 1,383 distinct genotypes with this 11-marker dataset. We found that the native and introduced *P. viticola* f. sp. *aestivalis* populations were strongly differentiated, along the first axis of a discriminant analysis on principal components (DAPCs)[48] (Figure S3). The second axis revealed a strong differentiation in North America between *P. viticola* f. sp. *aestivalis* populations collected on different *Vitis* species (Figure S3A). Indeed, we identified specific genetic clusters on *V. aestivalis* (in yellow), *V. labrusca* (in black), and two differentiated clusters (brown and red) on the cultivated grape *V. vinifera* (Figure S3). The cluster on cultivated grapes (in brown) was genetically close to the cluster on the wild species *V. labrusca* (in black), probably indicating a host shift from *V. labrusca* in the native range. We inferred that the invasive populations originated from a host shift from *V. aestivalis*, as the population from this wild summer grape (yellow cluster) appeared genetically closest, and even overlapping in the DAPCs, with the invasive populations (Figure S3A). However, the eight microsatellite markers had a low power to resolve the genetic structure between invasive populations (Figure S3C).

We therefore genotyped a subset of strains (n = 181) with 32 microsatellite markers, focusing on invasive populations and on the native *P. viticola* f. sp. *aestivalis* clusters that may have served as origin of the invasive populations (Table S3, 174 unique genotypes). This 32-marker dataset confirmed the genetic patterns observed in the native range based on the 11-marker dataset, and, in particular, confirmed that the cluster on *V. aestivalis* (yellow) was the likely origin of all invasive populations worldwide (Figures 3 and 4). This cluster was, again, the closest to all the invasive populations on the DAPCs (Figure 3A). Furthermore, using the 32-marker dataset, we showed that the cluster on *V. aestivalis* was the only cluster having common genetic ancestry with the invasive populations in the STRUCTURE Bayesian clustering analysis[49,50] (see the light blue, green, or pink ancestry from $K = 2$ to $K = 6$ in genotypes collected from *V. aestivalis* on Figure 4A). With the 32-marker dataset, we also showed that this was the closest cluster to all the invasive populations in the distance-based neighbor-joining tree (Figure 3E). The 32-marker dataset further identified two genetic clusters on *V. aestivalis* in the native range in the STRUCTURE analyses (Figure 4A): the yellow cluster identified above, genetically close to the invasive populations found on cultivated grapes worldwide (Figures 3A and 3B), and the red genetic cluster, corresponding to strains collected on cultivated grapes in North America (Figures 3 and 4). This suggests the occurrence of two distinct host shifts from the wild summer grape *V. aestivalis* to cultivated grapes, one in the native range, from the red genetic cluster, and the other, from the yellow cluster, giving rise to all invasive populations worldwide.

The 32-marker dataset also revealed genetic differentiation between invasive populations in the main wine-producing

Please cite this article in press as: Fontaine et al., Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.03.009

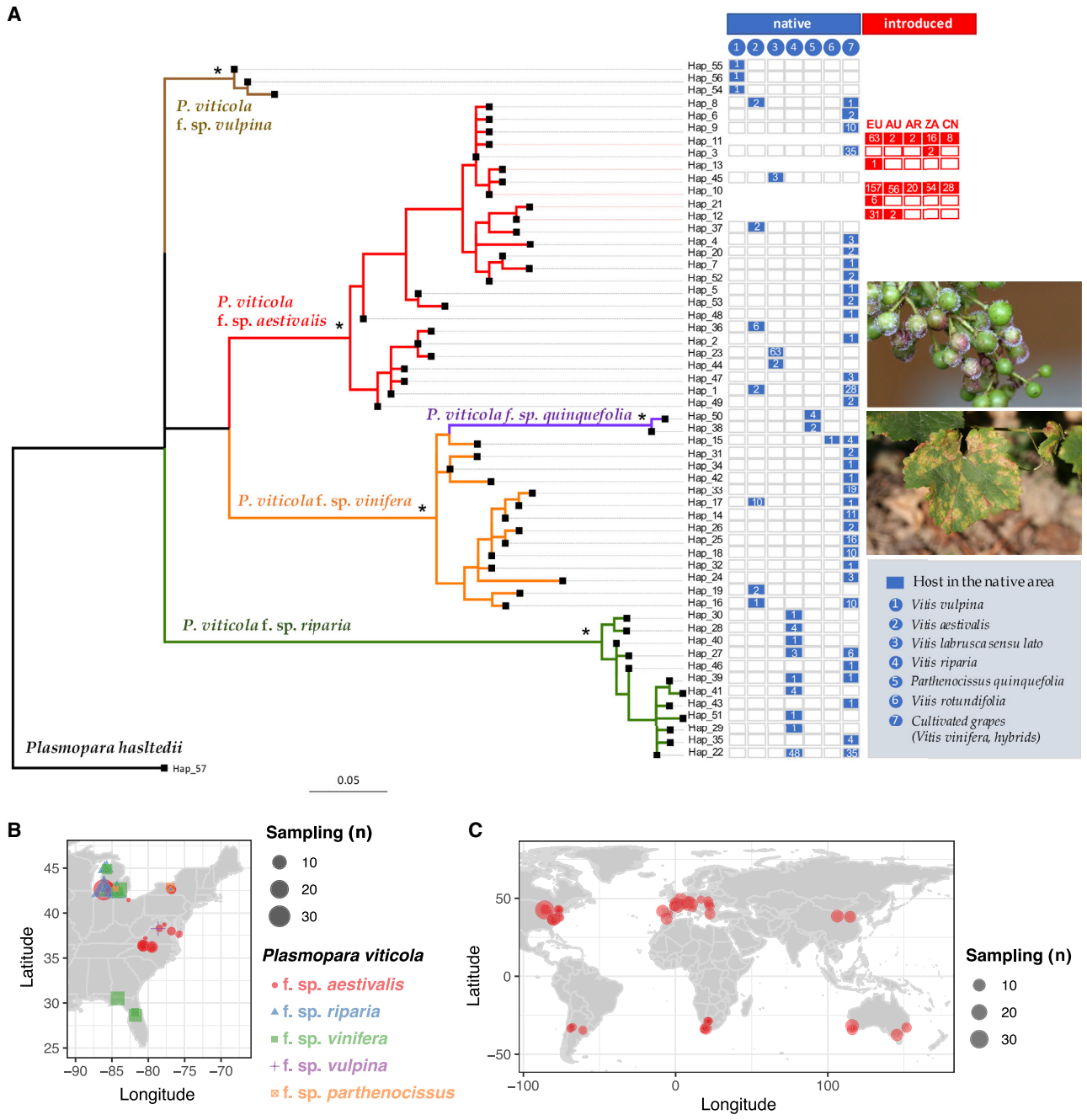


Figure 2. Phylogenetic relationships, sampling, and geographic distribution of *Plasmopara viticola* haplotypes of the β-tubulin (*tub*) gene

(A) The maximum likelihood phylogenetic tree is rooted with the *Plasmopara hasltedii tub* sequence. Nodes with an asterisk (*) are supported with a bootstrap value greater than 90%. The branches of the tree are color coded according to the five *P. viticola formae speciales* (ff. spp.). Colored and empty boxes on the right side of the tree show the host plant on which the haplotype was found in the native area or its geographic location on the introduced areas. Geographic codes include Europe (EU), Australia (AU), Argentina (AR), South Africa (ZA), and China (CN). The number of isolates carrying a given haplotype is indicated by the number within the boxes. Photos on the right show the grapevine downy mildew pathogen, *P. viticola*, infecting young grape berries (top) and typical lesions on a leaf (bottom) of *Vitis vinifera* in Europe.

(B) Map showing the sample size (n) and geographic location of origin of the strains from the five *formae speciales* occurring in Northeast America.

(C) Distribution of the sampling sites of the *P. v.* f. sp. *aestivalis* strains across the main wine-producing regions worldwide.

See also Figures S1 and S2 and Table S1.

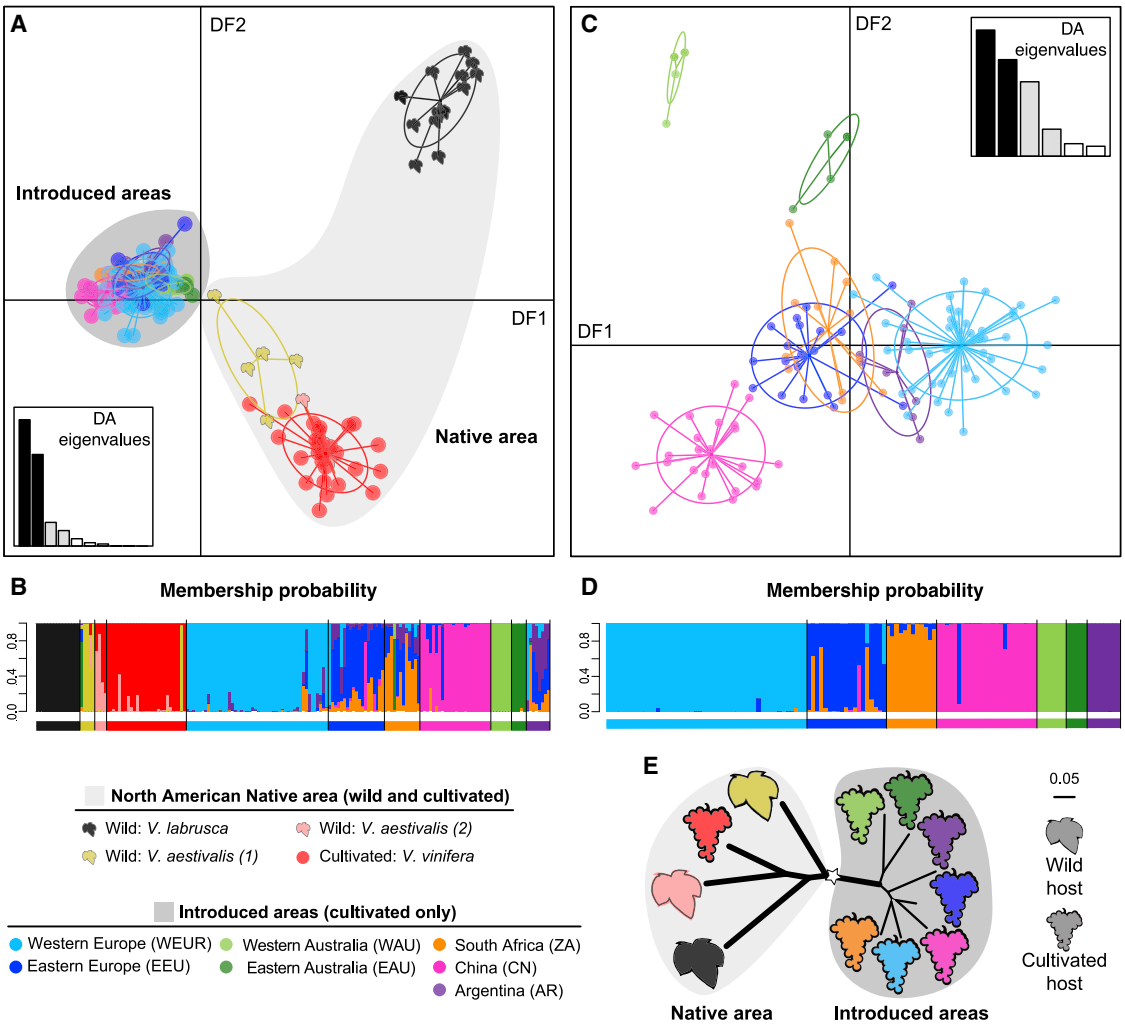


Figure 3. Worldwide population genetic structure of *Plasmopara viticola* inferred from the 32-microsatellite marker dataset using a discriminant analysis on principal components (DAPCs) and a distance-based neighbor-joining (NJ) tree

(A and C) Scatterplot of the first two discriminant functions (DFs) from the DAPCs, showing *P. viticola* individual genotypes for the full dataset (A) and for the subset excluding populations from the native North American range (C). The histogram insets of each DAPC scatterplot show the proportion of variance explained by each DF, according to their respective eigenvalues.

(B and D) The bar plots below each scatterplot represent the probability of the strain belonging to each group, based on all DFs of the DAPC in (A) and (C), respectively.

(E) NJ tree based on the Cavalli-Sforza chord distance. Population branching with high bootstrap support (i.e., >75%), indicated by large widths. Wild and cultivated grapes are represented by different symbols. North America and other continents are separated into different gray circles. The populations studied are as follows: North American *P. viticola* strains collected on wild *V. labrusca*, *V. aestivalis* group 1 (yellow) and 2 (red); North American strains collected on cultivated *V. vinifera*; strains from the rest of the world collected on cultivated *V. vinifera* including strains from Western and Eastern Europe (WEUR and EEUR), Western and Eastern Australia (WAUS and EAUS), China (CN), South Africa (ZA), and Argentina (AR).

See also Figure S3 and Data S1.

regions worldwide (Figures 3 and 4). The two distinct genetic groups of *P. viticola* previously identified in Western and Eastern European vineyards,[31] respectively, were also detected here in STRUCTURE clustering analyses (Figure 4) and with the DAPCs (Figures 3C and 3D). Furthermore, the populations in the different invaded areas also displayed significant differentiation, as shown by the significant F_{ST} values (Data S1A), STRUCTURE analyses (Figure 4), and the DAPC analyses focusing on invasive populations (Figures 3C and 3D). We also found two differentiated populations in Australia (Figures 3 and 4).

Lower diversity in invasive *P. v.* f. sp. *aestivalis* populations

Genetic diversity in invasive *P. v.* f. sp. *aestivalis* populations worldwide was much lower than that in native populations. Only a few closely related haplotypes of the three sequenced DNA fragments were present in the invaded areas, whereas considerable diversity was detected in the native *P. v.* f. sp. *aestivalis* populations (Figures 2 and S2; Table S1). The resulting low nucleotide (π) and haplotype (H) diversities in the invasive populations (Figure 5A; Table S1) indicated the occurrence of severe bottlenecks

Please cite this article in press as: Fontaine et al., Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.03.009

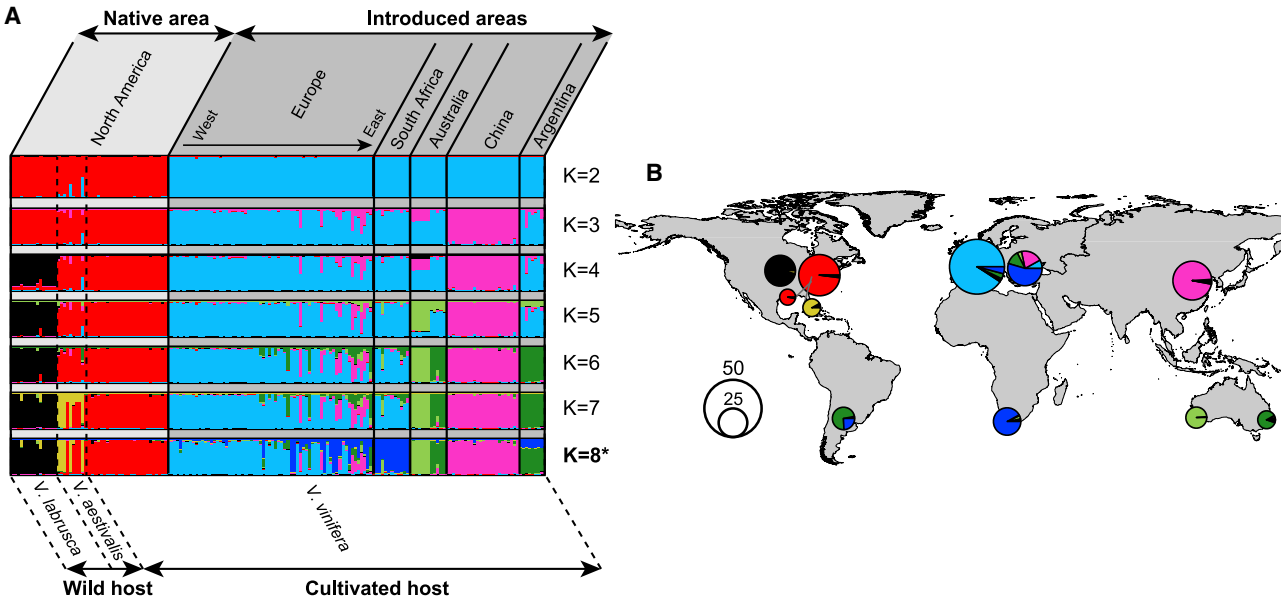


Figure 4. Worldwide population genetic structure of *Plasmopara viticola* inferred from the 32-microsatellite marker dataset using the Bayesian clustering approach of STRUCTURE

(A) Estimated individual ancestry proportions to each cluster (*K*) testing two to eight distinct genetic clusters. Each individual is represented by a thin vertical line, partitioned into *K*-colored segments representing the estimated genome ancestry fractions for each cluster. Individuals from different continents (labeled at the top of the figure) are separated by black continuous lines; strains from different host species (labeled at the bottom of the figure) are separated by black dotted lines; within these categories, individuals are sorted by increasing longitude. The figure shown for each *K* corresponds to the highest probability run from 10 replicates, with the best-fitting *K* value indicated by a star.

(B) STRUCTURE admixture proportions for samples averaged across populations for $K = 8$. Pie chart size is proportional to sample size.

See also Figure S3 and Data S1.

during the invasion process. European populations displayed a smaller decrease in genetic diversity than the populations of other invaded areas (Figures 2, 5A, and S2; Table S1), suggesting that the bottlenecks occurring in Europe were milder. Furthermore, all but one of the invasive haplotypes worldwide were present in Europe, for the three DNA fragments analyzed (Figures 2 and S2). Together, these findings suggest that Europe may have acted as a bridgehead for secondary invasions in the rest of the world.

Both microsatellite datasets confirmed the much lower levels of genetic diversity in invasive populations than in native populations, as shown by the compact clustering of invasive genotypes in the DAPCs (Figures 3A and S3A), and by diversity indices (Figure 5B; Tables S2 and S3; Data S1C and S1D). For example, private allelic richness was one order of magnitude lower in the invasive populations worldwide than in the native range. Microsatellite genetic diversity was also much lower in introduced than in native populations, with European populations having intermediate values (Figure 5B; Tables S2 and S3; Data S1C and S1D). Some F_{IS} values were significantly different from zero and positive, suggesting a Wahlund effect due to further population subdivision (Table S3).

This pattern of diversity is, thus, consistent with the hypothesis that Europe acted as a bridgehead, with a first wave of invasion to Europe originating from the yellow *P. v.* f. sp. *aestivalis* cluster (Figures 3, 4, and S3) occurring on the wild summer grape in the native Northern American range, and a second wave of invasions subsequently occurring from Europe to other vineyards throughout the world.

Worldwide invasion history of *P. viticola* reconstructed by ABC-RF scenario testing

We formally compared the likelihoods of alternative invasion history scenarios involving the most likely population of origin in the native range (i.e., the yellow genetic cluster on *V. aestivalis*), as identified based on previous analyses, and all the invasive introduced populations, using the 32-microsatellite marker dataset and an approximate Bayesian computation random forest (ABC-RF) statistical framework.[51] In ABC-RF, genetic data are simulated under different demographic scenarios, and summary statistics from the resulting simulated data are statistically compared with those obtained from the observed data.[52–54] We used an iterative process to infer the various invasion events while keeping a tractable number of scenarios to be compared with the ABC-RF.[55] We first identified the most likely demographic scenario(s) composed of bifurcation and admixture events considering the native populations and the invasive populations that corresponded to the most ancient introduction events (i.e., with dates of first disease records outside the native range between 1878 and 1889; Figure 6B). We then considered the other invasive populations successively following increasing dates of first disease records and tested using ABC-RF what was their population of origin.

As a first step, we considered 18 scenarios of introduction from the most likely population of origin in the native range (the yellow cluster on *V. aestivalis*) and the first reported invasive populations from historical records (i.e., in Western and Eastern Europe, in ~1878 and ~1887, respectively, and in China, in

Please cite this article in press as: Fontaine et al., Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.03.009

CellPress

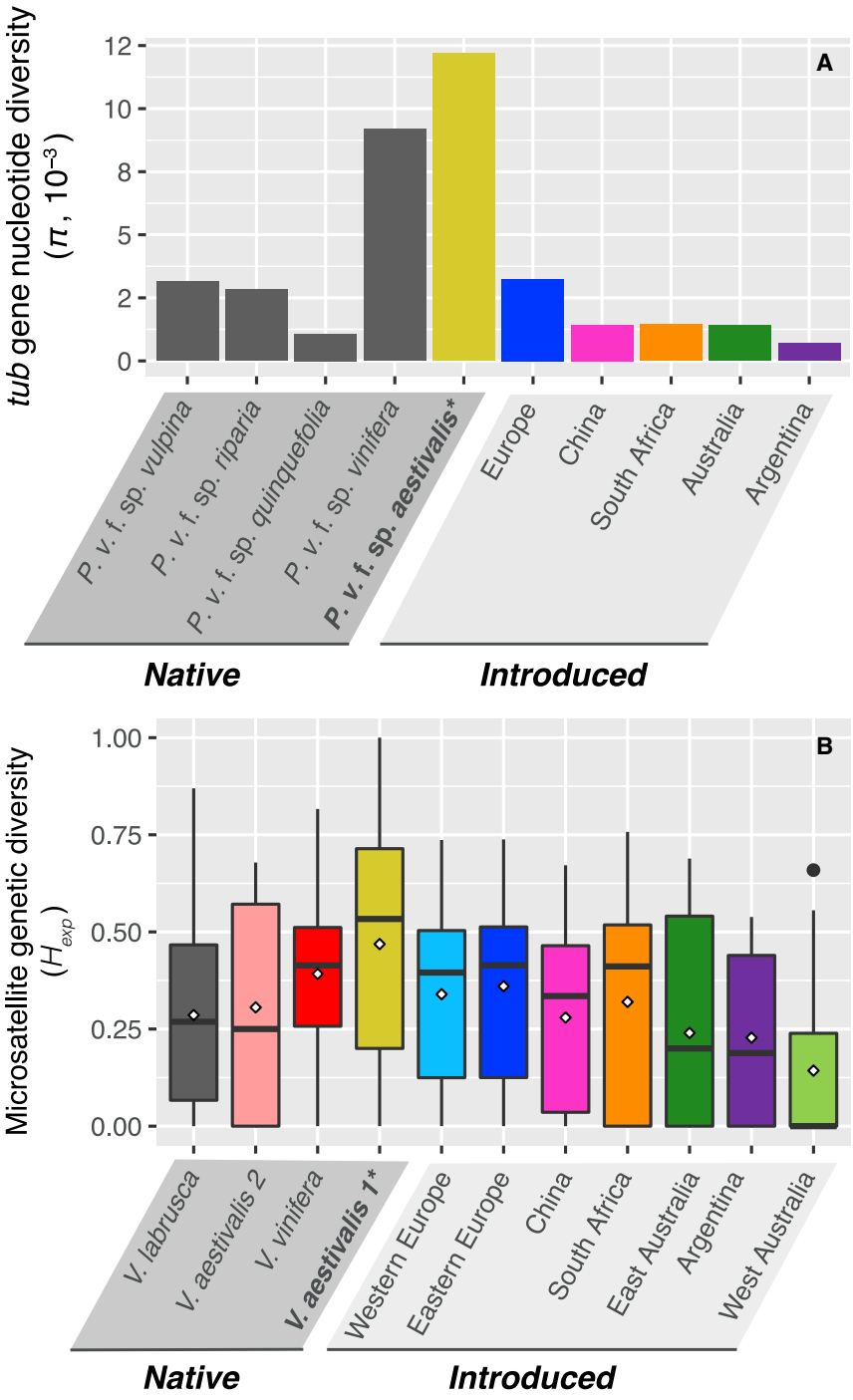


Figure 5. Lower diversity in invasive populations of *Plasmopara viticola* f. sp. *aestivalis*

(A) Bar plot showing the nucleotide diversity (π, per site) of the β-tubulin (*tub*) gene sequence.

(B) Boxplot showing the distribution of expected heterozygosity (H_{exp}) for the 32 microsatellite markers and the white diamonds show the average values.

See also Tables S1–S3 and Data S1.

~1889; Table S4; Figure S4). The most likely scenario identified (scenario 11 in Table S4 and Figure S4) involved an initial introduction into Western European vineyards, followed by a spread to Eastern Europe (as previously shown[31]), and then an expansion to Eastern China from Eastern Europe. This scenario was the most strongly supported, with a posterior probability of 0.42 ± 0.01 (Table S4).

The disease was later reported in South Africa in ~1907 and in eastern Australia in ~1917. For these two introduction events (steps 2 and 3 in Table S4; Figure S4), the ABC-RF again identified Eastern Europe as the most likely population of origin, with posterior probabilities of 0.78 ± 0.03 and 0.71 ± 0.02, respectively. The disease was then reported in Argentina in ~1920, and the ABC-RF analysis suggested that this introduction was probably from Eastern Australia, with a posterior probability of 0.70 ± 0.02 (step 4 in Table S4; Figure S4). The most recent first report of the disease was in Western Australia in ~1917, for which an Eastern Australian origin has been suggested.[37] However, the ABC-RF statistical framework, based on a larger set of microsatellite markers and more extensive sampling of *P. viticola* populations worldwide, again identified Eastern Europe as the most likely population of origin for the pathogen in Western Australia, with a posterior probability of 0.73 ± 0.02 (step 5 in Table S4 and Figure S4).

The most likely global invasion history is summarized in Figure 6 (see Data S2B for the posterior parameter estimates). The classification error rate (i.e., prior error rate) for the first step was relatively high (38%), reflecting a difficulty in distinguishing between a first introduction in Western Europe or China. By contrast, all subsequent steps in the ABC-RF analysis were well resolved, with low prior error rates, ranging from 7% to 12% (Table S4). The final global scenario (Figure 6) was strongly supported by the genetic data, as shown by the low prior error rate of 7% (Table S4). Posterior model checking and goodness-of-fit assessment of the model-posterior distributions showed that the inferred global invasion scenario generated genetic summary statistics consistent with the observed data, providing high confidence in the inferred scenario. The 10,000 simulations from posterior parameter distributions under the best model produced summary statistics very close to those obtained from the observed dataset (Figure S5), with only 24 of 256 statistics falling in the tail of the predictive probability distribution of statistics calculated from the posterior simulations (i.e., $p < 0.05$ or $p < 0.95$). Moreover, none of the p values remained significant after correction for multiple

Please cite this article in press as: Fontaine et al., Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.03.009

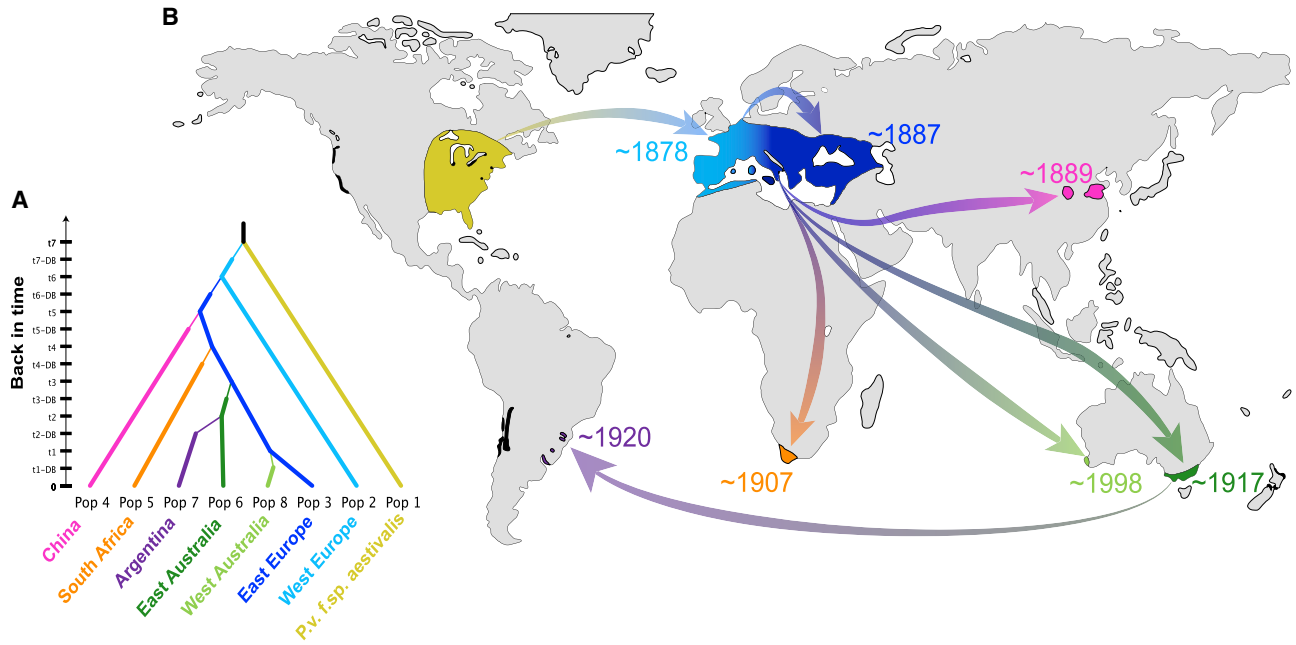


Figure 6. Worldwide invasion history of *Plasmopara viticola* f. sp. *aestivalis* inferred by the approximate Bayesian computation random forest (ABC-RF) approach

(A) The best population divergence scenario inferred by ABC-RF (Table S4).

(B) Geographic representation of the invasion scenario with the highest likelihood based on the ABC-RF analysis; areas are shown in color, based on their population assignment, as identified in Figures 3 and 4. Dates represent the first report of the disease in each area of introduction. Arrows indicate the most likely invasion pathways. Vineyards in other regions of the world not included in this study are indicated in black.

See also Figures S4 and S5, Tables S4, and Data S2.

comparisons.[56] The inferred invasion scenario (Figure 6; Data S2B) thus fitted the observed genetic data well.

DISCUSSION

Using extensive sampling and a powerful statistical framework, we have elucidated the worldwide invasion history of grapevine downy mildew. We show here that a specific genetic cluster of *P. v.* f. sp. *aestivalis*, which parasitizes the wild summer grape *V. aestivalis* in the native range, was the source of the invasion responsible for a devastating pandemic on cultivated grapevines in Europe at the end of the 19th century.[18,22] Severe bottlenecks occurred, with very few haplotypes from the native lineage introduced and established into Europe, leading to much lower genetic diversity in invaded continents than in the native range. In the inferred invasion scenario (Figure 6; Table S4; Data S2B), Europe then served as a bridgehead for a second wave of invasions worldwide, spreading the disease further afield, to Eastern China, South Africa, and Eastern Australia. A third wave of expansion probably occurred from Eastern Australia to Argentina. The much more recent introduction in Western Australia (1998) also seems to have originated from Europe, raising questions about the efficacy of quarantine regulations. The contribution of European populations to grapevine downy mildew invasions worldwide reflects the key role of Europe in the trading of plant material during the development of "New World" viticulture. The phylloxera crisis intensified the importation of plant material from North America to Europe in the second half of the 19th century and then from Europe to "New World" vineyards.[57] "New World" wine regions have indeed massively imported French or Italian *V. vinifera* varieties to establish their vineyards.[58] A role for Europe as a hub for the invasion of "New World" vineyards by grapevine pathogens has also been suggested for phylloxera and for the soil-born nematode vector of the fanleaf virus.[57,59]

Controlling the introduction of pests and diseases is one of the greatest challenges in viticulture, with important economic and environmental consequences. The genetic relationships and diversity inferred here could be used to guide quarantine regulations for grape-growing countries, to prevent further invasions. Introducing new strains or lineages of *P. viticola* would significantly increase the diversity of invasive populations and create opportunities for genetic admixture. Indeed, we found that, in North America, the cultivated grape (*V. vinifera*) was parasitized by a *P. viticola* lineage (f. sp. *vinifera*) other than the one identified here as invasive (f. sp. *aestivalis*). The *P. viticola* f. sp. *vinifera* lineage has not yet been found on other continents but could also become invasive should it be introduced. An increase in genetic variation through migration and/or admixture would probably facilitate the adaptation of *P. viticola* to pesticides and resistant cultivars.[60] New invasions by additional *P. viticola* strains or lineages would have a major effect on the wine industry, by destabilizing grapevine protection and canceling out the sustained efforts of breeders to obtain good-quality grapevine varieties resistant to downy mildew. The rich historical records available for grape diseases make these diseases excellent case studies

for obtaining fundamental insight into the processes underlying biological invasions, showing that worldwide biological invasions can result from the secondary dispersal of a particularly successful invasive population, as for grape phylloxera.[57,61] Our findings show that even pathogens subject to very strong bottlenecks and without admixture as a means of generating diversity can achieve successful invasions worldwide, by retaining an ability to evolve rapidly, and thus to develop new virulence against plant resistance[62–64] and resistance to fungicides.[30,65,66]

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Sample collection and DNA extraction
- METHOD DETAILS
  - DNA sequencing
  - Microsatellite genotyping
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - DNA fragment phylogenetic analyses and diversity
  - Inference of population genetic structure and diversity from microsatellite data
  - ABC-RF-based inferences of global invasion history

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2021.03.009.

ACKNOWLEDGMENTS

We would like to thank all the people involved in *P. viticola* sampling: Anton Baudoin (Virginia Tech, USA), Odile Carisse (Agriculture and Agri-Food, Canada), David Gadoury (Cornell University, USA), Mizuho Nita (Virginia Tech, USA), Annemiek Schilder (University of California, Davis, USA), Andrew S. Taylor (DPI, Australia), Davide Gobbin (ETH, Switzerland), Marie-Laure Panon (Comité Champagne, France), Patrice Retaud (SRPV, France), Marie-Pascale Latorse (Bayer Crop-Science, France), Jorge Silva (Bayer CropScience, Portugal), Hanns Kassemeyer (Staatliches Weinbauinstitut Freiburg, Germany), Pal Kozma (University of Pécs, Hungary), Mauro Jermini (Agroscope, Switzerland), Sara Legler (UCSC, Italy), Atanas Atanassov (JGC, Bulgaria), Pere Mestre (INRAE, France), and Jiang Lu (Shanghai JiaoTong University). This work was funded by French National Research Agency grant ANR-07-BDIV-003 (*EMERFUNDIS* project). Y.D. received a grant from the French National Research Agency (*GANDALF* project, ANR-12-ADAP-0009), and M.C.F. received a postdoctoral grant from the Région Ile-de-France. We thank the Center for Information Technology of the University of Groningen and, in particular, Bob Dröge and Fokke Dijkstra for their continuous support and for providing access to the *Peregrine* high-performance computing cluster.

AUTHOR CONTRIBUTIONS

F.D., T.G., and M.C.F. designed the study; F.D. and L.D. collected the biological materials; S.R.-C. conducted the molecular work in the laboratory and collected the data under the supervision of F.D.; M.C.F. and F.L. analyzed the data with help from Y.D. and F.D.; M.C.F., F.L., T.G., and F.D. interpreted the results; and M.C.F., T.G., F.L., and F.D. wrote the manuscript, with feedback and final approval from all the co-authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 4, 2020
Revised: February 4, 2021
Accepted: March 2, 2021
Published: March 25, 2021

REFERENCES

1. Ristaino, J.B., and Records, A. (2020). Emerging Plant Diseases and Global Food Security (American Phytopathological Society).
2. Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R., and Daszak, P. (2004). Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends Ecol. Evol. *19*, 535–544.
3. Desprez-Loustau, M.L., Robin, C., Buée, M., Courtecuisse, R., Garbaye, J., Suffert, F., Sache, I., and Rizzo, D.M. (2007). The fungal dimension of biological invasions. Trends Ecol. Evol. *22*, 472–480.
4. Fisher, M.C., Henk, D.A., Briggs, C.J., Brownstein, J.S., Madoff, L.C., McCraw, S.L., and Gurr, S.J. (2012). Emerging fungal threats to animal, plant and ecosystem health. Nature *484*, 186–194.
5. Singh, R.P., Hodson, D.P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., Herrera-Foessel, S., Singh, P.K., Singh, S., and Govindan, V. (2011). The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. Annu. Rev. Phytopathol. *49*, 465–481.
6. Saunders, D.G.O., Pretorius, Z.A., and Hovmøller, M.S. (2019). Tackling the re-emergence of wheat stem rust in Western Europe. Commun. Biol. *2*, 51.
7. Bhattacharya, S. (2017). Deadly new wheat disease threatens Europe's crops. Nature *542*, 145–146.
8. Islam, M.T., Croll, D., Gladieux, P., Soanes, D.M., Persoons, A., Bhattacharjee, P., Hossain, M.S., Gupta, D.R., Rahman, M.M., Mahboob, M.G., et al. (2016). Emergence of wheat blast in Bangladesh was caused by a South American lineage of Magnaporthe oryzae. BMC Biol. *14*, 84.
9. Giraud, T., Gladieux, P., and Gavrilets, S. (2010). Linking the emergence of fungal plant diseases with ecological speciation. Trends Ecol. Evol. *25*, 387–395.
10. Gladieux, P., Feurtey, A., Hood, M.E., Snirc, A., Clavel, J., Dutech, C., Roy, M., and Giraud, T. (2015). The population biology of fungal invasions. Mol. Ecol. *24*, 1969–1986.
11. Levadoux, L. (1956). Les populations sauvages et cultivees de *Vitis vinifera* L. Ann. Amelior. Plant. *6*, 59–117.
12. Terral, J.-F., Tabard, E., Bouby, L., Ivorra, S., Pastor, T., Figueiral, I., Picq, S., Chevance, J.-B., Jung, C., Fabre, L., et al. (2010). Evolution and history of grapevine (*Vitis vinifera*) under domestication: new morphometric perspectives to understand seed domestication syndrome and reveal origins of ancient European cultivars. Ann. Bot. *105*, 443–455.
13. Zohary, D., and Hopf, M. (1993). Domestication of Plants in the Old World: The Origin and Spread of Domesticated Plants in Southwest Asia, Europe, and the Mediterranean Basin (Clarendon Press).
14. Myles, S., Boyko, A.R., Owens, C.L., Brown, P.J., Grassi, F., Aradhya, M.K., Prins, B., Reynolds, A., Chia, J.-M., Ware, D., et al. (2011). Genetic structure and domestication history of the grape. Proc. Natl. Acad. Sci. USA *108*, 3530–3535.
15. De Lorenzis, G., Mercati, F., Bergamini, C., Cardone, M.F., Lupini, A., Mauceri, A., Caputo, A.R., Abbate, L., Barbagallo, M.G., Antonacci, D., et al. (2019). SNP genotyping elucidates the genetic diversity of *Magna Graecia* grapevine germplasm and its historical origin and dissemination. BMC Plant Biol. *19*, 7.

16. Grassi, F., and Arroyo-Garcia, R. (2020). Editorial: origins and domestication of the grape. Front. Plant Sci. *11*, 1176.
17. Galet, P. (1977). Les Maladies et les Parasites de la Vigne, t. 1: Les Maladies dues à des Végétaux: Champignons, Bactéries, Viroses et Phanérogames (Imprimerie du Paysan du Midi).
18. Millardet, A. (1881). Notes sur les Vignes Américaines et Opuscules Divers sur le Même Sujet (Ferret et fils).
19. Gessler, C., Pertot, I., and Perazzolli, M. (2011). *Plasmopara viticola*: a review of knowledge on downy mildew of grapevine and effective disease management. Phytopathol. Mediterr. *50*, 3–44.
20. Lafon, R., and Clerjeau, M. (1994). Downy mildew. In Compendium of Grape Diseases, R.C. Pearson, and A.C. Goheen, eds. (APS Press), pp. 11–13.
21. Viennot-Bourgin, G. (1949). Les Champignons Parasites des Plantes Cultivées (Masson).
22. Cornu, M. (1882). II. Le Peronospora des Vignes, Observations sur le Phylloxera et sur les Maladies de la Vigne (Institut de France, Académie des Sciences).
23. Dussert, Y., Legrand, L., Mazet, I.D., Couture, C., Piron, M.-C., Serre, R.-F., Bouchez, O., Mestre, P., Toffolatti, S.L., Giraud, T., and Delmotte, F. (2020). Identification of the first oomycete mating-type locus sequence in the grapevine downy mildew pathogen, *Plasmopara viticola*. Curr. Biol. *30*, 3897–3907.e4.
24. Brilli, M., Asquini, E., Moser, M., Bianchedi, P.L., Perazzolli, M., and Si-Ammour, A. (2018). A multi-omics study of the grapevine-downy mildew (*Plasmopara viticola*) pathosystem unveils a complex protein coding- and noncoding-based arms race during infection. Sci. Rep. *8*, 757.
25. Dussert, Y., Mazet, I.D., Couture, C., Gouzy, J., Piron, M.-C., Kuchly, C., Bouchez, O., Rispe, C., Mestre, P., and Delmotte, F. (2019). A high-quality grapevine downy mildew genome assembly reveals rapidly evolving and lineage-specific putative host adaptation genes. Genome Biol. Evol. *11*, 954–969.
26. Mestre, P., Carrere, S., Gouzy, J., Piron, M.C., Tourvieille de Labrouhe, D., Vincourt, P., Delmotte, F., and Godiard, L. (2016). Comparative analysis of expressed CRN and RXLR effectors from two *Plasmopara* species causing grapevine and sunflower downy mildew. Plant Pathol. *65*, 767–781.
27. Liu, Y., Lan, X., Song, S., Yin, L., Dry, I.B., Qu, J., Xiang, J., and Lu, J. (2018). In planta functional analysis and subcellular localization of the oomycete pathogen *Plasmopara viticola* candidate RXLR effector repertoire. Front. Plant Sci. *9*, 286.
28. Yin, L., An, Y., Qu, J., Li, X., Zhang, Y., Dry, I., Wu, H., and Lu, J. (2017). Genome sequence of *Plasmopara viticola* and insight into the pathogenic mechanism. Sci. Rep. *7*, 46553.
29. Camargo, M.P., Hong, C.F., Amorim, L., and Scherm, H. (2019). Cryptic species and population genetic structure of *Plasmopara viticola* in São Paulo State, Brazil. Plant Pathol. *68*, 719–726.
30. Delmas, C.E., Dussert, Y., Delière, L., Couture, C., Mazet, I.D., Richart Cervera, S., and Delmotte, F. (2017). Soft selective sweeps in fungicide resistance evolution: recurrent mutations without fitness costs in grapevine downy mildew. Mol. Ecol. *26*, 1936–1951.
31. Fontaine, M.C., Austerlitz, F., Giraud, T., Labbé, F., Papura, D., Richard-Cervera, S., and Delmotte, F. (2013). Genetic signature of a range expansion and leap-frog event after the recent invasion of Europe by the grapevine downy mildew pathogen *Plasmopara viticola*. Mol. Ecol. *22*, 2771–2786.
32. Gobbin, D., Pertot, I., and Gessler, C. (2003). Identification of microsatellite markers for *Plasmopara viticola* and establishment of high throughput method for SSR analysis. Eur. J. Plant Pathol. *109*, 153–164.
33. Gobbin, D., Pertot, I., and Gessler, C. (2003). Genetic structure of a *Plasmopara viticola* population in an isolated italian mountain vineyard. J. Phytopathol. *151*, 636–646.
34. Gobbin, D., Rumbou, A., Linde, C.C., and Gessler, C. (2006). Population genetic structure of *Plasmopara viticola* after 125 years of colonization in European vineyards. Mol. Plant Pathol. *7*, 519–531.
35. Koopman, T., Linde, C.C., Fourie, P.H., and McLeod, A. (2007). Population genetic structure of *Plasmopara viticola* in the Western Cape Province of South Africa. Mol. Plant Pathol. *8*, 723–736.
36. Li, X., Yin, L., Ma, L., Zhang, Y., An, Y., and Lu, J. (2016). Pathogenicity variation and population genetic structure of *Plasmopara viticola* in China. J. Phytopathol. *164*, 863–873.
37. Taylor, A.S., Knaus, B.J., Grünwald, N.J., and Burgess, T. (2019). Population genetic structure and cryptic species of *Plasmopara viticola* in Australia. Phytopathology *109*, 1975–1983.
38. Zhang, W., Manawasinghe, I.S., Zhao, W., Xu, J., Brooks, S., Zhao, X., Hyde, K.D., Chethana, K.W.T., Liu, J., Li, X., and Yan, J. (2017). Multiple gene genealogy reveals high genetic diversity and evidence for multiple origins of Chinese *Plasmopara viticola* population. Sci. Rep. *7*, 17304.
39. Gobbin, D., Jermini, M., Loskill, B., Pertot, I., Raynal, M., and Gessler, C. (2005). Importance of secondary inoculum of *Plasmopara viticola* to epidemics of grapevine downy mildew. Plant Pathol. *54*, 522–534.
40. Matasci, C.L., Jermini, M., Gobbin, D., and Gessler, C. (2010). Microsatellite based population structure of *Plasmopara viticola* at single vine scale. Eur. J. Plant Pathol. *127*, 501–508.
41. Rumbou, A., and Gessler, C. (2006). Particular structure of *Plasmopara viticola* populations evolved under Greek island conditions. Phytopathology *96*, 501–509.
42. Rouxel, M., Mestre, P., Baudoin, A., Carisse, O., Delière, L., Ellis, M.A., Gadoury, D., Lu, J., Nita, M., Richard-Cervera, S., et al. (2014). Geographic distribution of cryptic species of *Plasmopara viticola* causing downy mildew on wild and cultivated grape in eastern North America. Phytopathology *104*, 692–701.
43. Rouxel, M., Mestre, P., Comont, G., Lehman, B.L., Schilder, A., and Delmotte, F. (2013). Phylogenetic and experimental evidence for host-specialized cryptic species in a biotrophic oomycete. New Phytol. *197*, 251–263.
44. Rouxel, M., Papura, D., Nogueira, M., Machefer, V., Dezette, D., Richard-Cervera, S., Carrere, S., Mestre, P., and Delmotte, F. (2012). Microsatellite markers for characterization of native and introduced populations of *Plasmopara viticola*, the causal agent of grapevine downy mildew. Appl. Environ. Microbiol. *78*, 6337–6340.
45. Maddalena, G., Delmotte, F., Bianco, P.A., De Lorenzis, G., and Toffolatti, S.L. (2020). Genetic structure of Italian population of the grapevine downy mildew agent, *Plasmopara viticola*. Ann. Appl. Biol. Published online January 9, 2020. https://doi.org/10.1111/aab.12567.
46. Anderson, K., and Pinilla, V. (2018). Wine Globalization: A New Comparative History (Cambridge University Press).
47. Zecca, G., Abbott, J.R., Sun, W.B., Spada, A., Sala, F., and Grassi, F. (2012). The timing and the mode of evolution of wild grapes (*Vitis*). Mol. Phylogenet. Evol. *62*, 736–747.
48. Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. *11*. Published online October 15, 2010. https://doi.org/10.1186/1471-2156-11-94.
49. Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics *155*, 945–959.
50. Falush, D., Stephens, M., and Pritchard, J.K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics *164*, 1567–1587.
51. Pudlo, P., Marin, J.-M., Estoup, A., Cornuet, J.-M., Gautier, M., and Robert, C.P. (2016). Reliable ABC model choice via random forests. Bioinformatics *32*, 859–866.
52. Beaumont, M.A., Zhang, W., and Balding, D.J. (2002). Approximate Bayesian computation in population genetics. Genetics *162*, 2025–2035.
53. Beaumont, M.A. (2010). Approximate Bayesian computation in evolution and ecology. Annu. Rev. Ecol. Evol. Syst. *41*, 379–406.

54. Csilléry, K., Blum, M.G.B., Gaggiotti, O.E., and François, O. (2010). Approximate Bayesian Computation (ABC) in practice. Trends Ecol. Evol. *25*, 410–418.

55. Ryan, S.F., Lombaert, E., Espeset, A., Vila, R., Talavera, G., Dincă, V., Doellman, M.M., Renshaw, M.A., Eng, M.W., Hornett, E.A., et al. (2019). Global invasion history of the agricultural pest butterfly *Pieris rapae* revealed with genomics and citizen science. Proc. Natl. Acad. Sci. USA *116*, 20015–20024.

56. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc. B *57*, 289–300.

57. Rispe, C., Legeai, F., Nabity, P., Fernández, R., Arora, A.K., Baa-Puyoulet, P., Banfill, C.R., Bao, L., Barberà, M., Bouallègue, M., et al. (2020). Genome sequence of the grape phylloxera: insights into the genome evolution and invasion routes of an iconic pest. BMC Biol. Published online July 23, 2020. https://doi.org/10.1186/s12915-020-00820-5.

58. Wolkovich, E.M., García de Cortázar-Atauri, I., Morales-Castilla, I., Nicholas, K.A., and Lacombe, T. (2018). From Pinot to Xinomavro in the world's future wine-growing regions. Nat. Clim. Chang. *8*, 29–37.

59. Nguyen, V.C., Villate, L., Gutierrez-Gutierrez, C., Castillo, P., Van Ghelder, C., Plantard, O., and Esmenjaud, D. (2019). Phylogeography of the soil-borne vector nematode Xiphinema index highly suggests Eastern origin and dissemination with domesticated grapevine. Sci. Rep. *9*, 7313.

60. Rius, M., and Darling, J.A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? Trends Ecol. Evol. *29*, 233–242.

61. Tello, J., Mammerler, R., Čajić, M., and Forneck, A. (2019). Major outbreaks in the nineteenth century shaped grape phylloxera contemporary genetic structure in Europe. Sci. Rep. *9*, 17540.

62. Delmas, C.E., Fabre, F., Jolivet, J., Mazet, I.D., Richart Cervera, S., Delière, L., and Delmotte, F. (2016). Adaptation of a plant pathogen to partial host resistance: selection for greater aggressiveness in grapevine downy mildew. Evol. Appl. *9*, 709–725.

63. Peressotti, E., Wiedemann-Merdinoglu, S., Delmotte, F., Bellin, D., Di Gaspero, G., Testolin, R., Merdinoglu, D., and Mestre, P. (2010). Breakdown of resistance to grapevine downy mildew upon limited deployment of a resistant variety. BMC Plant Biol. *10*, 147.

64. Delmotte, F., Mestre, P., Schneider, C., Kassemeyer, H.-H., Kozma, P., Richart-Cervera, S., Rouxel, M., and Delière, L. (2014). Rapid and multiregional adaptation to host partial resistance in a plant pathogenic oomycete: evidence from European populations of *Plasmopara viticola*, the causal agent of grapevine downy mildew. Infect. Genet. Evol. *27*, 500–508.

65. Chen, W.-J., Delmotte, F., Richard-Cervera, S., Douence, L., Greif, C., and Corio-Costet, M.-F. (2007). At least two origins of fungicide resistance in grapevine downy mildew populations. Appl. Environ. Microbiol. *73*, 5162–5172.

66. Blum, M., Gamper, H.A., Waldner, M., Sierotzki, H., and Gisi, U. (2012). The cellulose synthase 3 (CesA3) gene of oomycetes: structure, phylogeny and influence on sensitivity to carboxylic acid amide (CAA) fungicides. Fungal Biol. *116*, 529–542.

67. Giresse, X., Ahmed, S., Richard-Cervera, S., and Delmotte, F. (2010). Development of new oomycete taxon-specific mitochondrial cytochrome *b* region primers for use in phylogenetic and phylogeographic studies. J. Phytopathol. *158*, 321–327.

68. Delmotte, F., Chen, W.J., Richard-Cervera, S., Greif, C., Papura, D., Giresse, X., Mondor-Genson, G., and Corio-Costet, M.F. (2006). Microsatellite DNA markers for *Plasmopara viticola*, the causal agent of downy mildew of grapes. Mol. Ecol. Notes *6*, 379–381.

69. Excoffier, L., and Lischer, H.E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. *10*, 564–567.

70. Gouy, M., Guindon, S., and Gascuel, O. (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. *27*, 221–224.

71. Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics *25*, 1451–1452.

72. Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.

73. Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol. Ecol. Resour. *8*, 103–106.

74. Earl, D.A., and vonHoldt, B.M. (2011). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. *4*, 359–361.

75. Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., and Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol. Ecol. Resour. *15*, 1179–1191.

76. Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics *24*, 1403–1405.

77. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/.

78. Langella, O. (2011). Populations 1.2.32. CNRS UPR9034. http://www.bioinformatics.org/project/?group_id=84.

79. Rambaut, A., and Drummond, A.J. (2018). FigTree version 1.4.4. Molecular evolution, phylogenetics and epidemiology. http://tree.bio.ed.ac.uk/software/figtree/.

80. Goudet, J. (2001). FSTAT, A program to estimate and test gene diversities and fixation indices (version 2.9.3). Lausanne University. https://www2.unil.ch/popgen/softwares/fstat.htm.

81. Szpiech, Z.A., Jakobsson, M., and Rosenberg, N.A. (2008). ADZE: a rarefaction approach for counting alleles private to combinations of populations. Bioinformatics *24*, 2498–2504.

82. Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F. (2004). GENETIX 4.05, logiciel sous WindowsTM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier. https://kimura.univ-montp2.fr/genetix/.

83. Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M., and Estoup, A. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics *30*, 1187–1189.

84. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

85. Pecrix, Y., Buendia, L., Penouilh-Suzette, C., Maréchaux, M., Legrand, L., Bouchez, O., Rengel, D., Gouzy, J., Cottret, L., Vear, F., and Godiard, L. (2019). Sunflower resistance to multiple downy mildew pathotypes revealed by recognition of conserved effectors of the oomycete *Plasmopara halstedii*. Plant J. *97*, 730–748.

86. Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. *14*, 2611–2620.

87. Vercken, E., Fontaine, M.C., Gladieux, P., Hood, M.E., Jonot, O., and Giraud, T. (2010). Glacial refugia in pathogens: European genetic structure of anther smut pathogens on *Silene latifolia* and *Silene dioica*. PLoS Pathog. *6*, e1001229.

88. Cavalli-Sforza, L.L., and Edwards, A.W. (1967). Phylogenetic analysis. Models and estimation procedures. Am. J. Hum. Genet. *19*, 233–257.

89. Weir, B.S., and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. Evolution *38*, 1358–1370.

90. Leberg, P.L. (2002). Estimating allelic richness: effects of sample size and bottlenecks. Mol. Ecol. *11*, 2445–2449.

91. Fraimout, A., Debat, V., Fellous, S., Hufbauer, R.A., Foucaud, J., Pudlo, P., Marin, J.-M., Price, D.K., Cattel, J., Chen, X., et al. (2017). Deciphering the

routes of invasion of Drosophila suzukii by means of ABC random forest. Mol. Biol. Evol. *34*, 980–996.

92. Raynal, L., Marin, J.M., Pudlo, P., Ribatet, M., Robert, C.P., and Estoup, A. (2019). ABC random forests for Bayesian parameter inference. Bioinformatics *35*, 1720–1728.
93. Estoup, A., Jarne, P., and Cornuet, J.M. (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol. Ecol. *11*, 1591–1604.
94. Estoup, A., Lombaert, E., Marin, J.-M., Guillemaud, T., Pudlo, P., Robert, C.P., and Cornuet, J.-M. (2012). Estimation of demo-genetic model probabilities with approximate Bayesian computation using linear discriminant analysis on summary statistics. Mol. Ecol. Resour. *12*, 846–855.
95. Cornuet, J.-M., Ravigné, V., and Estoup, A. (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). BMC Bioinformatics *11*, 401.

Please cite this article in press as: Fontaine et al., Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.03.009

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
CTAB	Sigma-Aldrich	CAT#H6269-250 g
Chloroform-Isoamyl alcohol	Sigma-Aldrich	CAT#C0549-1PJ
Deposited data		
Haplotype sequences for *Cytochrome-b* gene	This paper	GenBank: MW053146-MW053196
Haplotype sequences for *β-tubulin* gene	This paper	GenBank: MW053069-MW053145
Haplotype sequences for *28S* ribosomal RNA	This paper	GenBank: MW077425-MW077433
Sequence alignments for the *Cytochrome-b*, *β-tubulin, 28S* ribosomal RNA, and the two microsatellite genotype datasets	This paper	https://doi.org/10.15454/FD86M2
Oligonucleotides		
Forward and reverse primers and PCR conditions for *28S* ribosomal RNA and *β-tubulin*	[43]	See Data S3
Forward and reverse primers and PCR conditions for *Cytochrome-b*	[65,67]	See Data S3
Forward and reverse primers and PCR conditions for the eight microsatellite loci (Dataset 1): ISA, Pv7, Pv13, Pv14, Pv16, Pv17, Pv31, Pv39	[32,68]	See Data S3
Forward and reverse primers and PCR conditions for the 34 microsatellite markers (Dataset 2): ISA, Pv7, Pv14, Pv16, Pv17, Pv39, Pv65, Pv67, Pv74, Pv76, Pv83, Pv87, Pv88, Pv91, Pv93, Pv101, Pv103, Pv104, Pv124, Pv126, Pv127, Pv133, Pv134, Pv135, Pv137, Pv138, Pv139, Pv140, Pv141, Pv142, Pv143, Pv146, Pv147, Pv148	[44,68]	See Data S3
Experimental models: Organisms/strains		
Plasmopara viticola	See Figure S1 and Tables S1–S3	See Figure S1 and Tables S1–S3
Software and algorithms		
Arlequin v3.5	[69]	http://cmpg.unibe.ch/software/arlequin3513/Arlequin35.html
Seaview v.4.6.2	[70]	http://doua.prabi.fr/software/seaview
DnaSP v5.10.1	[71]	http://www.ub.edu/dnasp/
PhyML v.3.0	[72]	http://www.atgc-montpellier.fr/phyml/
GENEPOP v.4.2	[73]	https://kimura.univ-montp2.fr/~rousset/Genepop.htm
STRUCTURE v.2.3.4	[49,50]	https://web.stanford.edu/group/pritchardlab/structure.html
STRUCTURE HARVESTER v.0.6.94	[74]	http://taylor0.biology.ucla.edu/structureHarvester/
CLUMPAK v.1.1	[75]	http://clumpak.tau.ac.il/
ADEGENET v.2.0.1 R package	[76]	http://adegenet.r-forge.r-project.org/
R statistical environment version 3.6.3	[77]	https://www.r-project.org/
POPULATIONS v.1.2.32	[78]	https://bioinformatics.org/populations/
FigTree v.1.4.4	[79]	http://tree.bio.ed.ac.uk/software/figtree/
FSTAT v.2.9.4	[80]	https://www2.unil.ch/popgen/softwares/fstat.htm
ADZE v.1.0	[81]	https://rosenberglab.stanford.edu/adze.html

(*Continued on next page*)

Please cite this article in press as: Fontaine et al., Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.03.009

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
GENETIX v.4.05.2	[82]	https://kimura.univ-montp2.fr/genetix/
DIYABC v.2.1.0	[83]	http://www1.montpellier.inra.fr/CBGP/diyabc/
ABCRF v1.8.1 R package	[51]	https://cran.r-project.org/web/packages/abcrf/index.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Michael C. Fontaine (michael.fontaine@cnrs.fr).

Materials availability

This study did not generate new unique materials or reagents.

Data and code availability

Unique haplotypes from the β-tubulin, cytochrome *b*, and ribosomal 28S sequence data were deposited in GenBank under the accession codes: [MW053146-MW053196] for cytochrome-b; [MW053069-MW053145] for β-tubulin; and [MW077425-MW077433] for ribosomal 28S. The two microsatellite datasets, sequence alignments and analysis files are available via the INRAE Dataverse (https://doi.org/10.15454/FD86M2).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Sample collection and DNA extraction

Plasmopara viticola isolates were collected as sporulating lesions from 163 sites in the native (North America) or invasive (Europe, China, South Africa, Australia, and Argentina; Figures 2B, 2C, and S1; Tables S1–S3) range. This sampling covers the main grape-growing regions in which the climate is favorable for the disease. In Northeast America, samples were collected from five wild *Vitis* species (*V. riparia, V. labrusca*, *V. aestivalis, V. vulpina* and *P. quinquefolia*), and cultivated grapevines (*V. vinifera* and interspecific hybrids). In invaded areas, all samples were collected from diseased *V. vinifera* cultivars as previously described.[31] Briefly, samples were collected from separate grapevines within each plot early in the season (i.e., before the asexual cycles began). All samples were collected from commercial vineyards, in which fungicides were applied, except for those from Champagne (France), which were collected from small untreated plots. Oil spots were freeze-dried overnight, and DNA was extracted with standard CTAB and phenol-chloroform methods.[65,68] As the pathogen can only be grown on grape leaves, we extracted DNA from sporulating lesions collected in the field, which yielded DNA from both host tissue and pathogen mycelium and spores. However, the markers used were specific to the pathogen (see below).

METHOD DETAILS

DNA sequencing

For the identification of cryptic *P. viticola* species, three DNA fragments, from the 28S ribosomal RNA (*r28S*), β-tubulin (*tub*) and mitochondrial cytochrome-*b* (*cytb*) genes, were amplified by polymerase chain reaction (PCR) and sequenced (Table S1; Data S3A). DNA amplification, sequencing and assembly were performed as described by Rouxel et al.[43] for *r28S* and *tub*, and as described by Chen et al.[65] and Giresse et al.[67] for *cytb*. PCRs were carried out in a final volume of 15 μL containing 1 μL of 1:3 dilution of genomic DNA, 2 mM $MgCl_2$, 150 μM of each deoxynucleoside triphosphate (dNTP), 4 pmol of each primer (Data S3A) and 0.3 U Taq Silverstar DNA polymerase (Eurogentec, Belgium) in reaction buffer. Thermocycling conditions were as follows: 95°C for 4 min, 40 cycles of 95°C for 40 s, 58°C for 45 s, 72°C for 90 s, followed by 72°C for 10 min. Sequencing of PCR amplicons was performed at Genoscope (French National Sequencing Center, Evry, France). Forward and reverse sequences were imported into CodonCode Aligner v.2.0.6 (Centerville, MA, USA), assembled into contigs and visually checked for errors. All the polymorphic sites were confirmed by manual examination of the electropherogram. For sequences that presented heterozygote sites, gametic phase estimation was performed using the ELB algorithm implemented into Arlequin v3.5.[69] Using this method, we have determined for each isolate the two haplotypes at each locus.

Microsatellite genotyping

We constructed two different microsatellite datasets for analysis of the genetic structure and diversity of *P. viticola* f. sp. *aestivalis* populations across the main vineyards worldwide, identification of the most likely host-of-origin, and elucidation of the invasion routes followed. The first dataset included an extensive sampling of 1,974 strains genotyped for eight microsatellite loci (ISA, Pv7,

Please cite this article in press as: Fontaine et al., Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*, Current Biology (2021), https://doi.org/10.1016/j.cub.2021.03.009

Pv13, Pv14, Pv16, Pv17, Pv31 and Pv39)[32,68] (Table S2; Data S3B). Microsatellite PCR amplification and genotyping were conducted as previously described.[68] Briefly, PCR amplifications were carried out in a final 15 μL reaction volume including 1.5 μL of 10 × buffer (Eurogentec, Belgium), 0.45 μL of 50 mM MgCl2, 0.4 μL of 10 mM dNTPs, 0.3 μL of a dye-labeled forward primer and an unlabeled reverse primer (10 mM) (Data S3B), and 0.2 U of Taq Silverstar DNA polymerase (Eurogentec, Belgium). PCR cycles were performed in an Eppendorf Mastercycler ep gradient with the following conditions: an initial denaturation at 94°C for 4 min and 38 cycles of 30 s at 94°C, 30 s at the appropriate annealing temperature (Data S3B), and 35 s at 72°C, ending with a 5-min extension at 72°C. PCR products were analyzed as follows: 1 μL of PCR products (diluted at 1:100) was mixed with 8.86 μL of formamide and 0.14 μL of ABI GeneScan 600 LIZ (internal lane size standard) and analyzed in an Applied Biosystems 3130 capillary sequencer. Alleles were scored using the GeneMapper v4.0 software (Applied Biosystems, Foster City, CA). The repeatability of genotype scoring was checked by genotyping 5% of the samples twice. For the analyses, genotypes were built by combining the eight microsatellite markers and the three DNA sequences. Genotypes with more than 60% missing data were discarded. When putative clone-mates (i.e., repeated multilocus genotypes or MLGs within a given vineyard) were identified, a single individual of that genotype was retained. The final first genotype dataset encompassed 1,314 MLGs for the eight microsatellite markers and 1,383 MLGs for the 11 markers (combining the eight microsatellite markers with haplotypes for the three DNA fragments) from 105 localities (Tables S1 and S2).

The second dataset consisted of 181 samples genotyped for 34 microsatellite markers: ISA, Pv7, Pv14, Pv16, Pv17, Pv39, Pv65, Pv67, Pv74, Pv76, Pv83, Pv87, Pv88, Pv91, Pv93, Pv101, Pv103, Pv104, Pv124, Pv126, Pv127, Pv133, Pv134, Pv135, Pv137, Pv138, Pv139, Pv140, Pv141, Pv142, Pv143, Pv146, Pv147, Pv148.[44,68] The PCR amplification and genotyping procedures were as described above for the first microsatellite dataset (see also Data S3B). The repeatability of genotype scoring was checked by genotyping 5% of the samples twice. The mean error rate was $\leq 0.020 \pm 0.005$. Genotypes with missing data for more than 16 markers ($\geq$50%) were discarded. Significant linkage disequilibrium (LD) was detected between two pairs of microsatellite markers (Pv87 and Pv104; Pv124 and Pv133) with a permutation test (1,000 permutations) implemented in GENEPOP v.4.2.[73] A single marker for each pair was therefore retained for data analyses (Pv104 and Pv124). After filtering for missing data and LD, the final dataset included 174 MLGs for 32 markers from 35 sites (Table S3).

QUANTIFICATION AND STATISTICAL ANALYSIS

DNA fragment phylogenetic analyses and diversity

For each DNA fragment, sequences were aligned with *Muscle*[84] implemented in Seaview v.4.6.2.[70] Final alignment lengths after cleaning were 499 base pairs (bp) for *tub*, 685 bp for *cytb* and 703 bp for *r28S*. Distinct haplotypes were identified with DnaSP v5.10.1.[71] For each sequence alignment, genetic relationships between haplotypes (Figures 2 and S2) were inferred with a maximum likelihood (ML) method and a GTR substitution model implemented in PhyML v.3.0.[72] Node support was estimated by calculating 1,000 bootstraps. Each phylogeny was rooted with *Plasmopara hasltedii* sequences obtained from a recent genome assembly[85]. Nucleotide (π, per site) and haplotype (H) genetic diversities were estimated with the DnaSP program (Table S1).

Inference of population genetic structure and diversity from microsatellite data

We used the two datasets to investigate genetic structure: one with 11 markers (combining the eight microsatellite markers with haplotypes for the three DNA fragments), and the other with 32 microsatellite markers. We ran similar analyses on the two datasets. We used two complementary individual-based methods to estimate the population structure: the Bayesian model-based clustering method implemented in STRUCTURE v.2.3.4[49,50] (Figure 4), and a discriminant analysis on PCA (DAPC)[48] (Figures 3 and S3). We ran STRUCTURE with an admixture model with correlated allele frequencies, and uniform priors for the individual cluster of origin.[49,50] We performed simulations with a number of putative clusters (*K*) ranging from 1 to 10. For each *K* value, we conducted 10 independent replicates and checked for convergence. Each analysis included a burn-in period of 50,000 Markov chain Monte Carlo (MCMC) iterations followed by 500,000 MCMC iterations for the estimation of model parameters. We determined the most relevant number of clusters (*K*) using (i) the log-likelihood of the data for each *K* value,[49] (ii) the rate of change of the log-likelihood of the data with increasing *K*,[86] and (iii) the visual inspection of clusters newly generated with increases in *K*.[87] We used STRUCTURE HARVESTER v.0.6.94[74] to visualize the likelihood and its rate of change across *K* values and replicated runs. For the identification of potentially different clustering solutions, results were summarized and displayed with CLUMPAK v.1.1.[75] This analysis was performed on the two datasets combining native and introduced populations, and also on each area separately, to explore finer genetic structure.

DAPC[48] (Figures 3 and S3) does not make any specific assumptions about mating system or mode of reproduction. It provides a visualization of the genetic structure complementary to that provided by STRUCTURE, by summarizing the variance in allele frequencies summarized using principal components (PCs) and partitioning it between populations relative to the within-group variance. This analysis was performed on the two datasets, and on the native and introduced populations separately, to explore the finer genetic structure in the two areas. We used the ADEGENET v.2.0.1 R package[76] for conducting the DAPC. In accordance with the recommendations of the user guide, missing data were replaced by the mean value, the number of PCs used in the analysis was identified by the alpha-score optimization procedure and was set to 20 and 10 for the 11- and 32-marker datasets, respectively. *A priori* groups were defined according to host plant or continent, and further partitioning was implemented in accordance with the findings of STRUCTURE or previous studies (e.g., split between Eastern and Western Europe[31]). We assessed how distinct or admixed the clusters of the DAPC were, using scatterplots along the top discriminant axes (DA) and bar plots of individual membership probabilities for each group.

Genetic relationships between clusters were estimated with NJ population trees (Figure 3E) based on pairwise Cavalli-Sforza chord genetic distances[88] between populations computed with the program POPULATIONS v.1.2.32.[78] The node supports were estimated with 1,000 bootstraps over loci. The trees were drawn with FigTree v.1.4.4.[79]

Allelic frequency differentiation between clusters was estimated with the Weir and Cockerham F_{ST} estimator[89] (Data S1A) implemented in FSTAT v.2.9.4[80]. The significance of pairwise F_{ST} values was assessed with 100 random data permutations in FSTAT. We assessed the levels of microsatellite genetic diversity within clusters (Tables S2 and S3; Data S1), using allelic richness (A_r) and private allelic richness (P_{Ar}), both standardized by a rarefaction method,[90] to account for differences in sample size. We also calculated observed and expected heterozygosities (H_o and H_e, respectively), and the F_{IS} fixation index.[89] These statistics were calculated with ADZE v.1.0,[81] GENETIX v.4.05.2,[82] and FSTAT. Wilcoxon signed-rank tests were performed in the R statistical environment,[77] to assess the significance of differences in genetic diversity between clusters (Data S1).

ABC-RF-based inferences of global invasion history

We reconstructed the worldwide invasion history of *P. viticola* f. sp. *aestivalis* using an approximate Bayesian computation (ABC)[52–54] random forest (RF) analysis.[51,91,92] This ABC-RF analysis used the 32-microsatellite dataset and included eight populations: the seven genetic clusters identified outside the native range on the basis of the 32-microsatellite dataset and the wild population sampled in the native range on the summer grape *V. aestivalis,* which was the population genetically closest to the invasive clusters (Figures 2, 3, and 4). ABC-RF can estimate posterior probabilities of historical scenarios, based on historical data and massive coalescent simulations of genetic data. We used historical information (i.e., dates of first observation of invasive populations, Figure 6) to define five sets of competing introduction scenarios, which were analyzed sequentially (Figure S4; Table S4; Data S2). Step-by-step, each analysis made use of the results of the previous analyses, until the most recent invasive populations were considered. The scenarios for each analysis are detailed in Figure S4 and Table S4.

The scenario parameters (i.e., effective population size *N*, effective number of founders *NF*, admixture rate R_a, duration of the bottleneck *DB*, and time of population split and admixture *T*) were considered as random variables drawn from prior distributions (Data S2A). We assumed a generalized stepwise mutation process with possible single-nucleotide insertion, to model a realistic mutation process of microsatellite loci in the coalescent simulations.[93] We used DIYABC v.2.1.0[83] to simulate genetic data for ABC-RF analyses. Simulated and observed datasets were summarized using the whole set of summary statistics proposed by DIYABC for microsatellite markers, describing the genetic variation for each population (e.g., mean number of alleles per locus, and mean genetic diversity), pair of populations (e.g., pairwise genetic diversity, mean F_{ST} across loci between two populations, and shared allele distance), or trio of populations (e.g., maximum likelihood admixture estimates) (see the full list and details of the summary statistics in Data S2C). Linear discriminant analysis (LDA) components were also used as additional summary statistics.[94] The total number of summary statistics ranged from 70 to 256, depending on the analysis (Table S4).

We used the random forest (RF) classification procedure to compare the likelihood of the competing scenarios at each step with the R package *abcrf* v1.8.1.[51] RF is a machine-learning algorithm that uses hundreds of bootstrapped decision trees to perform classification, using the summary statistics as a set of predictor variables. Some simulations are not used in tree building at each bootstrap (i.e., the out-of-bag simulations), and are used to compute the "prior error rate," which provides a direct method for estimating the cross-validation error rate. We built a training set ranging from 10,000 to 50,000 simulated microsatellite datasets for each scenario, with the same number of loci and individuals as the observed dataset (Table S4). We then grew a classification forest of 500 or 1,000 trees based on the simulated training datasets. The size of the training set and number of decision trees was increased until the results converged over ten independent replicated RF analyses. The RF computation applied to the observed dataset provides a classification vote (i.e., the number of times a model is selected from the decision trees). We selected the scenario with the highest classification vote as the most likely scenario, and we estimated its posterior probability.[51] We assessed the global performance of our chosen ABC-RF scenario, by calculating the prior error rate based on the available out-of-bag simulations and we repeated the RF analysis 10 times to ensure that the results converged. Posterior distribution values of all parameters and some composite parameters for the best model identified (Data S2B) were estimated using a regression by RF methodology,[92] with classification forests of 1,000 decision trees.

We then performed a posterior model checking analysis on the final scenario, including all eight populations, to determine whether this scenario matched the observed genetic data well. Briefly, if a model fits the observed data correctly, then data simulated under this model with parameters drawn from their posterior distribution should be close to the observed data.[95] The lack of fit of the model to the data with respect to the posterior predictive distribution can be measured by determining the frequency at which the observed summary statistics are extreme with respect to the simulated summary statistics distribution (hence, defining a tail-area probability, or p value, for each summary statistic, Table S4; Data S2C). We simulated 100,000 datasets under the full final scenario (256 summary statistics), and obtained a "posterior sample" of 10,000 values of the posterior distributions of parameters through a rejection step based on Euclidean distances and a local regression post-treatment.[52] We simulated 10,000 new datasets with parameter values drawn from this "posterior sample," and each observed summary statistic was compared with the distribution of the 10,000 simulated test statistics (Figure S5); its p value was computed and corrected for multiple comparisons.[56] The simulation steps, the computation of summary statistics, and the model checking analysis were performed in DIYABC v2.1.0. All scenario comparisons were carried out in R, with the *abcrf* v1.8.1 package.[51]

Current Biology, Volume *31*

Supplemental Information

Europe as a bridgehead in the worldwide invasion history of grapevine downy mildew, *Plasmopara viticola*

Michael C. Fontaine, Frédéric Labbé, Yann Dussert, Laurent Delière, Sylvie Richart-Cervera, Tatiana Giraud, and François Delmotte

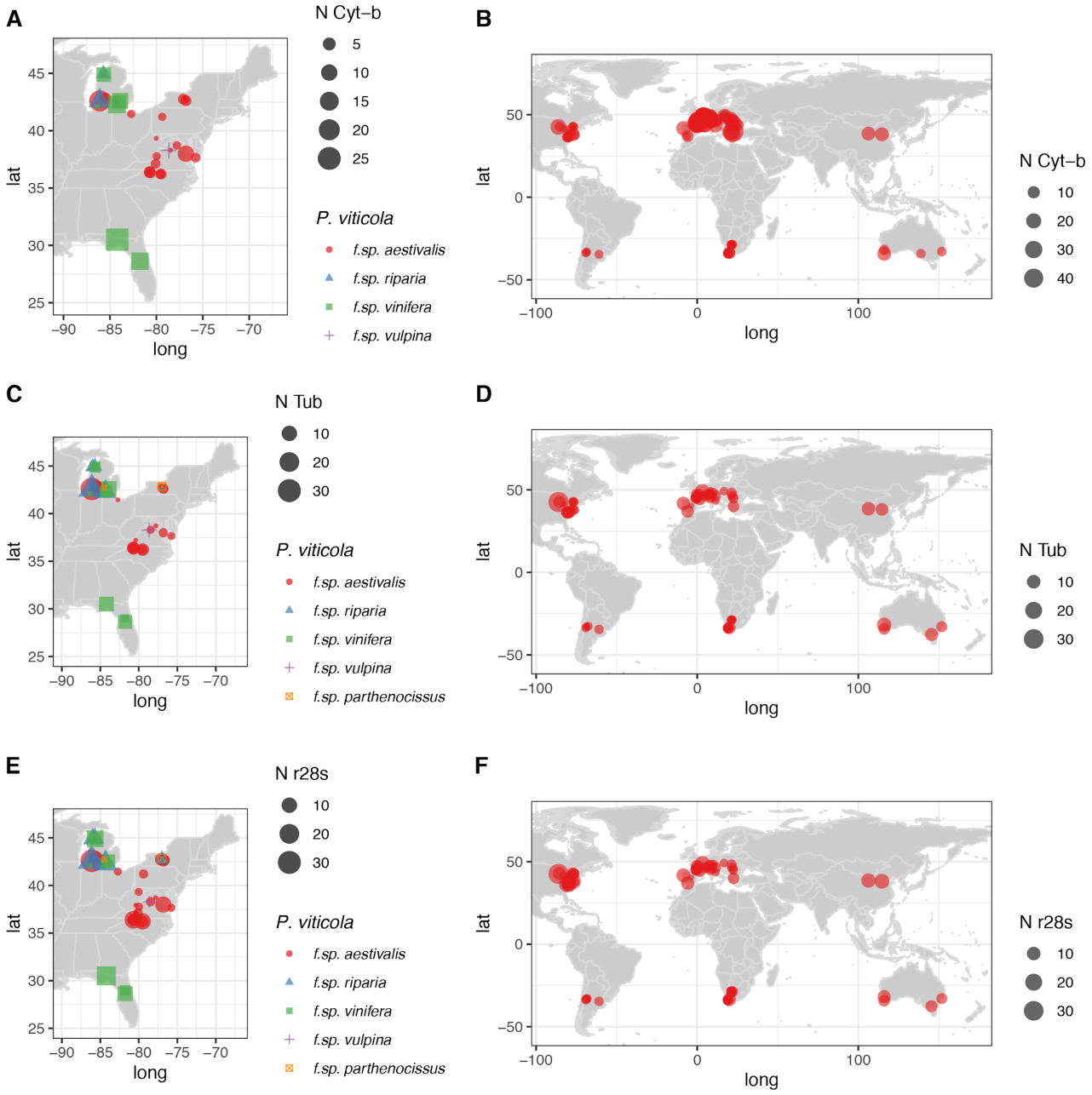


Figure S1. Location, *formae speciales* and number of samples collected for the three genes sequenced in this study, related to Figure 1, 2, Table S1, and to the section "Sample collection and DNA extraction" of the STAR Methods.

(**A**, **B**) cytochrome-*b* (*cytb*, n = 1,294), (**C**, **D**) *β*-tubulin (*tub*, n = 420), and (**E**, **F**) ribosomal 28S (*r28S*, n = 534). The size of the point is proportional to the sample size. The shape and color of the points refer to the *P. viticola* formae speciales (f. sp.).

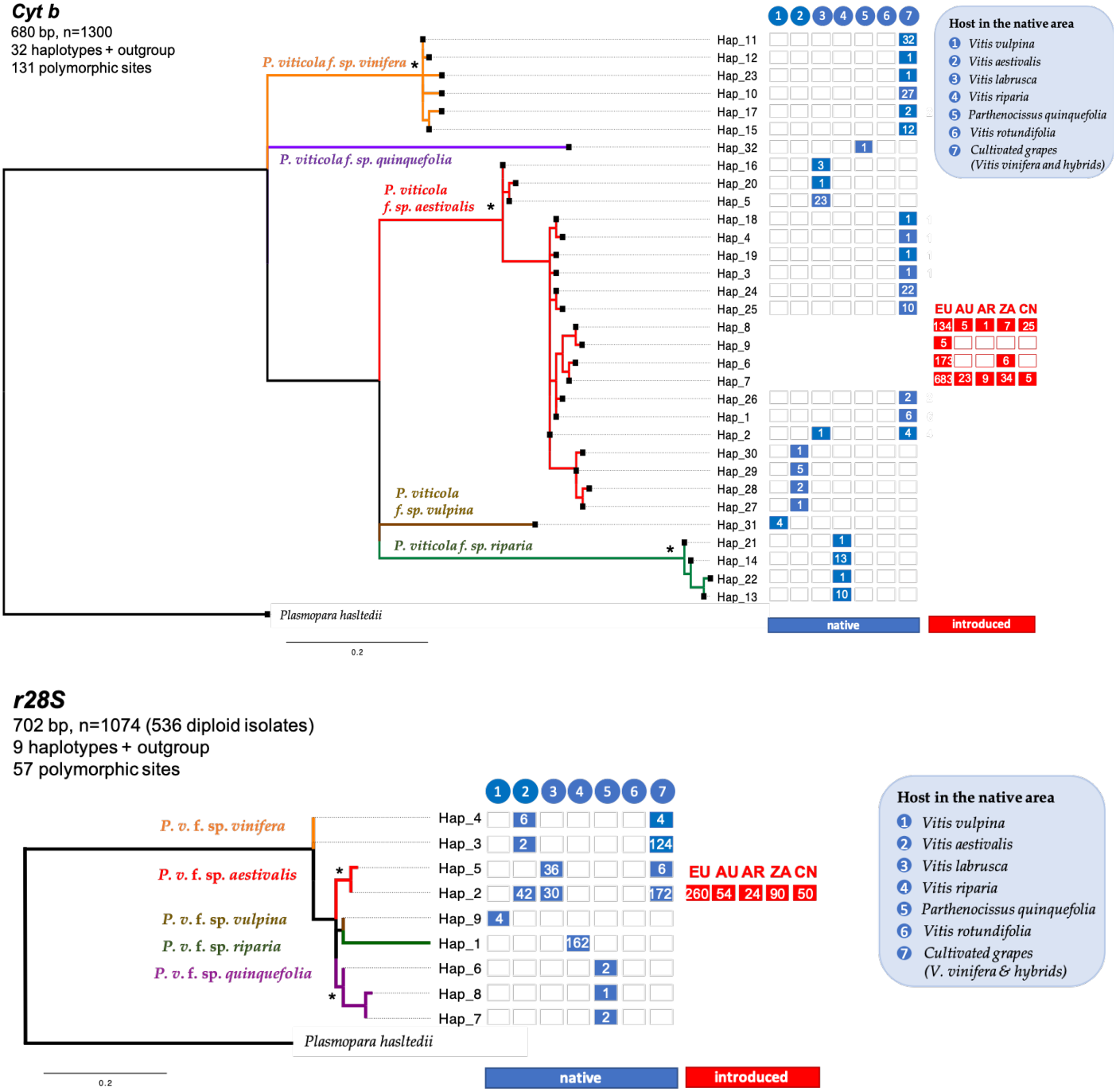


Figure S2. Maximum likelihood phylogenetic relationships of *Plasmopara viticola* haplotypes for the mitochondrial *cytb* gene (top) and nuclear ribosomal *28S* (*r28S*, bottom), related to Figure 2 and to the section "DNA fragment phylogenetic analyses and diversity" of the STAR Methods.

The tree is rooted with *Plasmopara hasltedii* sequences. Nodes with a star (*) are supported with a bootstrap value greater than 90%. The branches of the tree are color-coded according to the five *P. viticola formae speciales* (ff. spp.). Colored and empty boxes on the right side of the tree show the host plant on which the haplotype was found in the native area or its geographic location on the introduced areas. Geographic codes include Europe (EU), Australia (AU), Argentina (AR), South Africa (ZA) and China (CN). The number of isolates carrying a given haplotype is indicated by the number within the boxes.

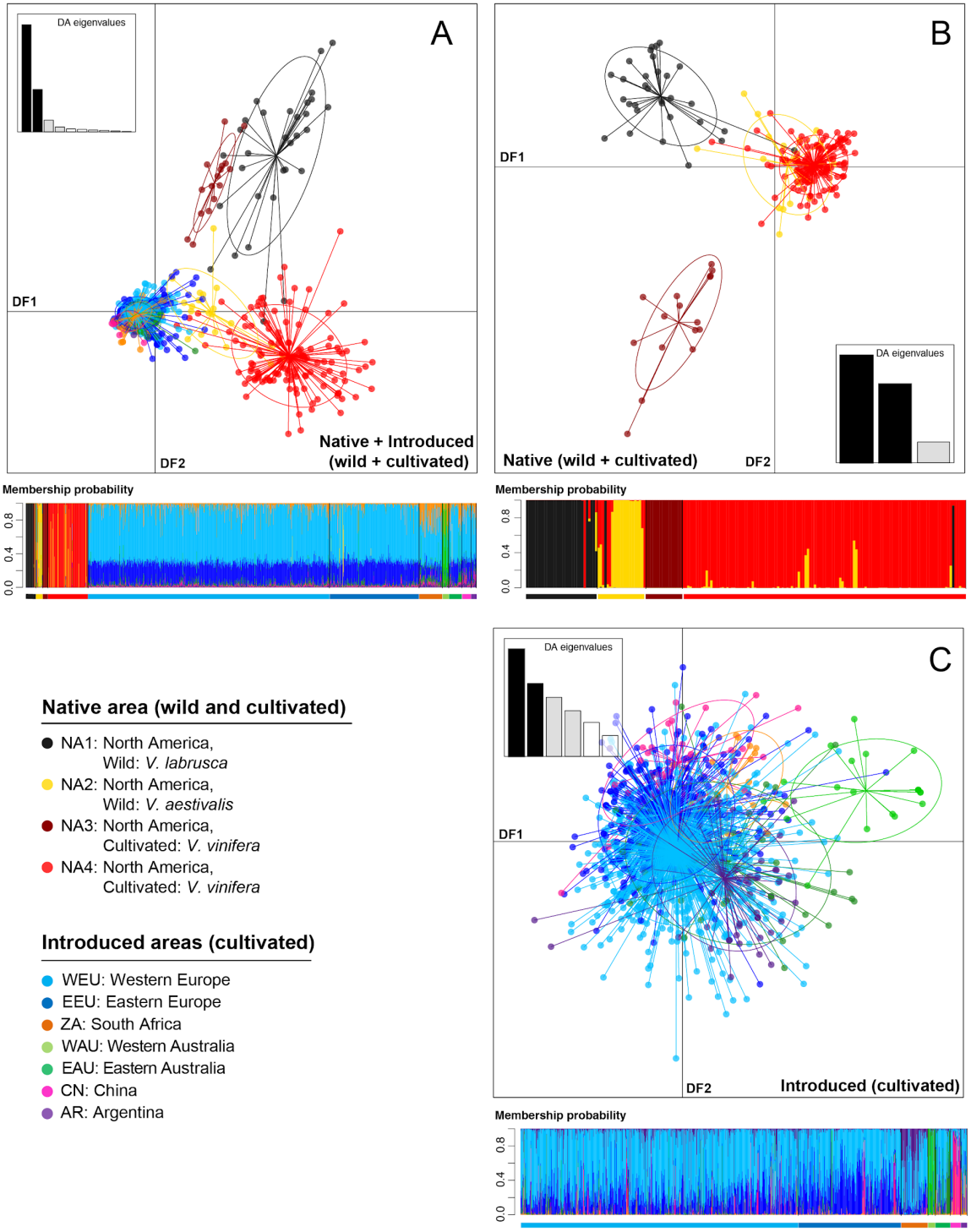
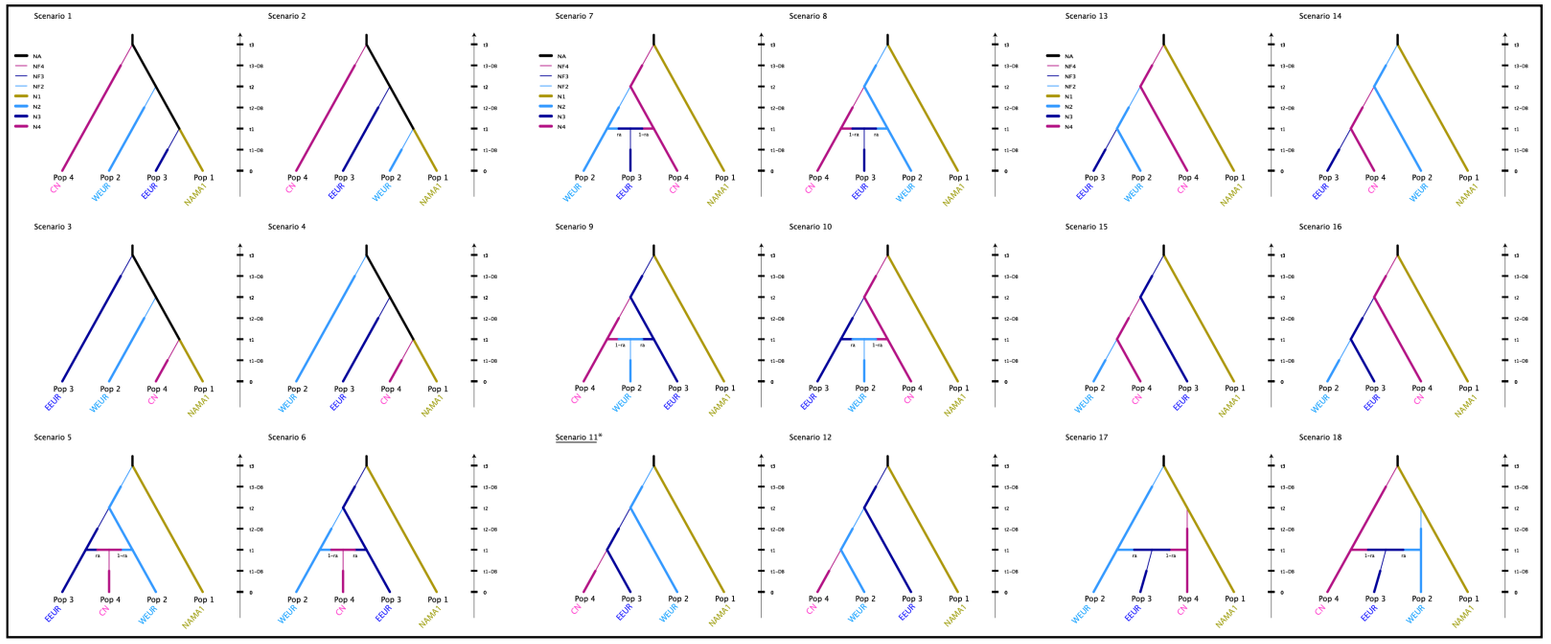


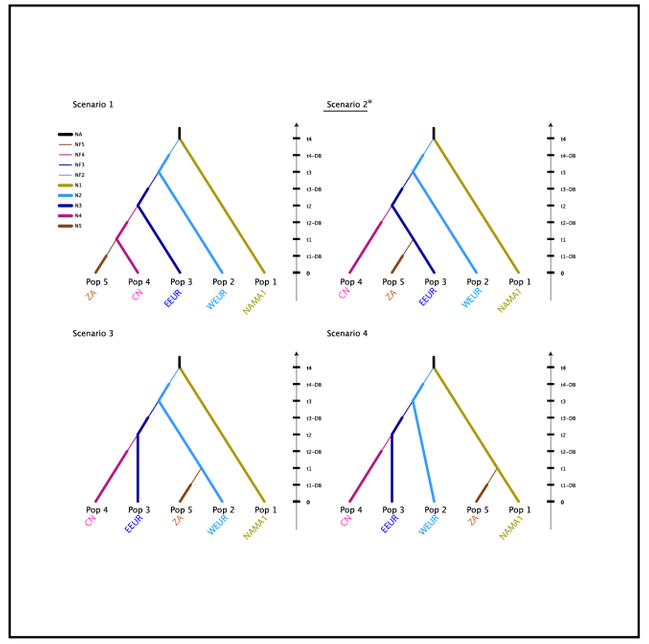
Figure S3. Worldwide population genetic structure of the downy mildew *Plasmopara viticola* f. sp. *aestivalis*, inferred from a DAPC using genotypes of 1,383 strains based on a dataset combining eight microsatellite markers and sequences from three genes (cytochrome-*b*, β-tubulin, and ribosomal 28S); related to Figure 3 and to the section "Inference of population genetic structure and diversity" of the STAR Methods.

DAPC scatterplots for the first two discriminant functions (DF1 and DF2) show the clustering of the genotypes for the whole dataset (**A**), a subset focusing on the strains collected in the native North American range on wild and cultivated plants (**B**), and another subset in the introduced range on cultivated grapes (**C**). The inset in each panel shows the eigenvalues of the DFs. The barplot below each scatterplot displays the membership probabilities estimated from the DAPC for each genotype (vertical bars) to *a priori* groups defined by the host plant and locality, as indicated by the bars below the barplot, with the color indicating the country of collection, as indicated in the legend.

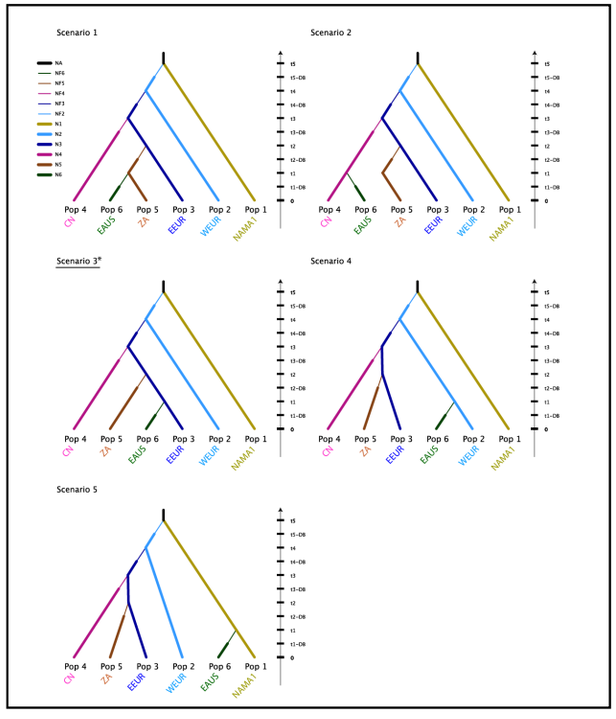
STEP 1: Evolutionary relationships among *P. v.* f. sp. *aestivalis* sampled on the wild *V. aestivalis* in Northeast America (**NAMA1**, yellow cluster in Figure 2), Western (**WEUR**) and Eastern European (**EEUR**) and Chinese (**CN**) clusters.



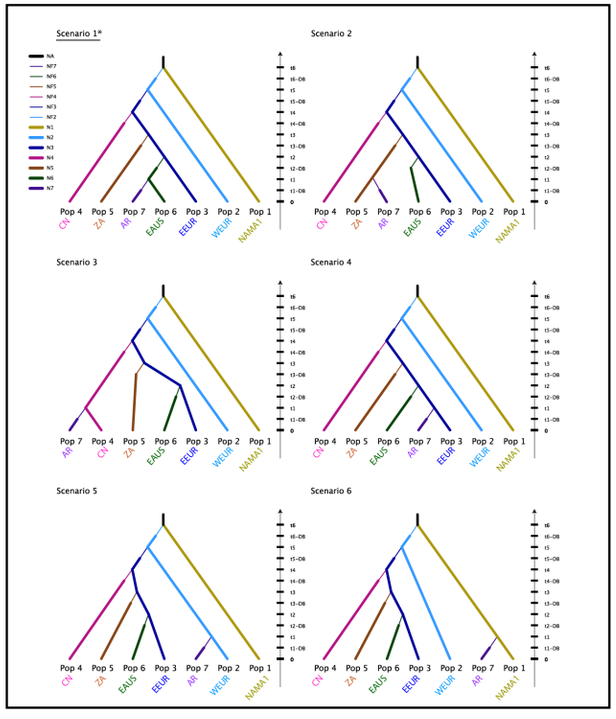
STEP 2: Introduction of South Africa (**ZA**) from…



STEP 3: Introduction of East Australia (**EAUS**) from …



STEP 4: Introduction of Argentina (**AR**) from …



STEP 5: Introduction of West Australia (**WAUS**) from …

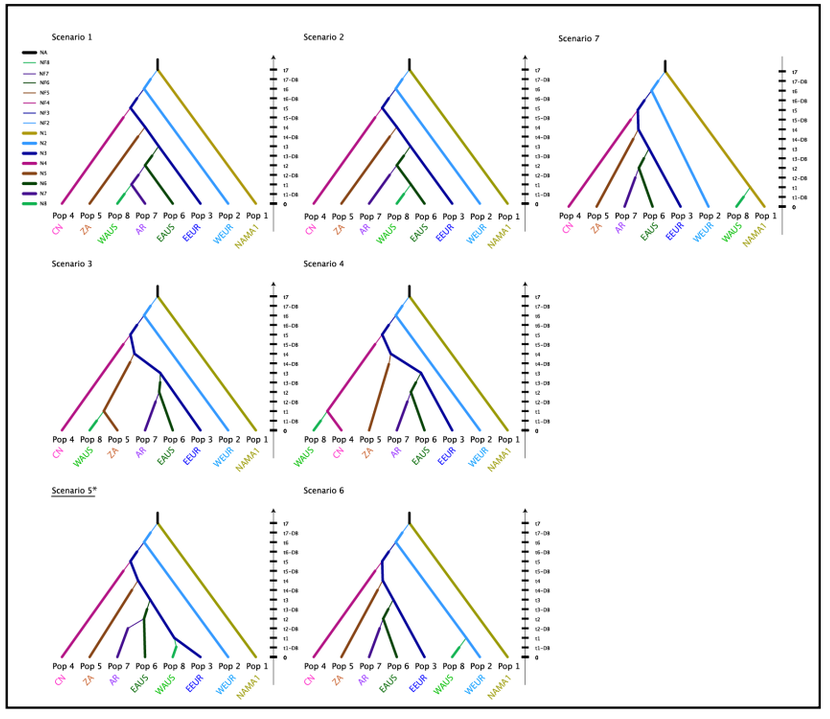


Figure S4. Schematic representation of the scenarios tested in the ABC-RF analyses to decipher the worldwide invasion routes of *Plasmopara viticola* f. sp. *aestivalis* populations; related to Figure 6, Table S4, Data S2, and to the section "ABC-RF-based inferences of global invasion history" of the STAR Methods.

For each step, the most likely scenario is underlined and marked with an asterisk (see Table S4 for the scenarios ranking, RF votes, posterior probability, and further details). Each analysis after step 1 is based on the best-fitting scenario identified in the previous one. Thin lines indicate bottlenecks. For parameter descriptions and priors, see Data S2A. Time is not to scale.

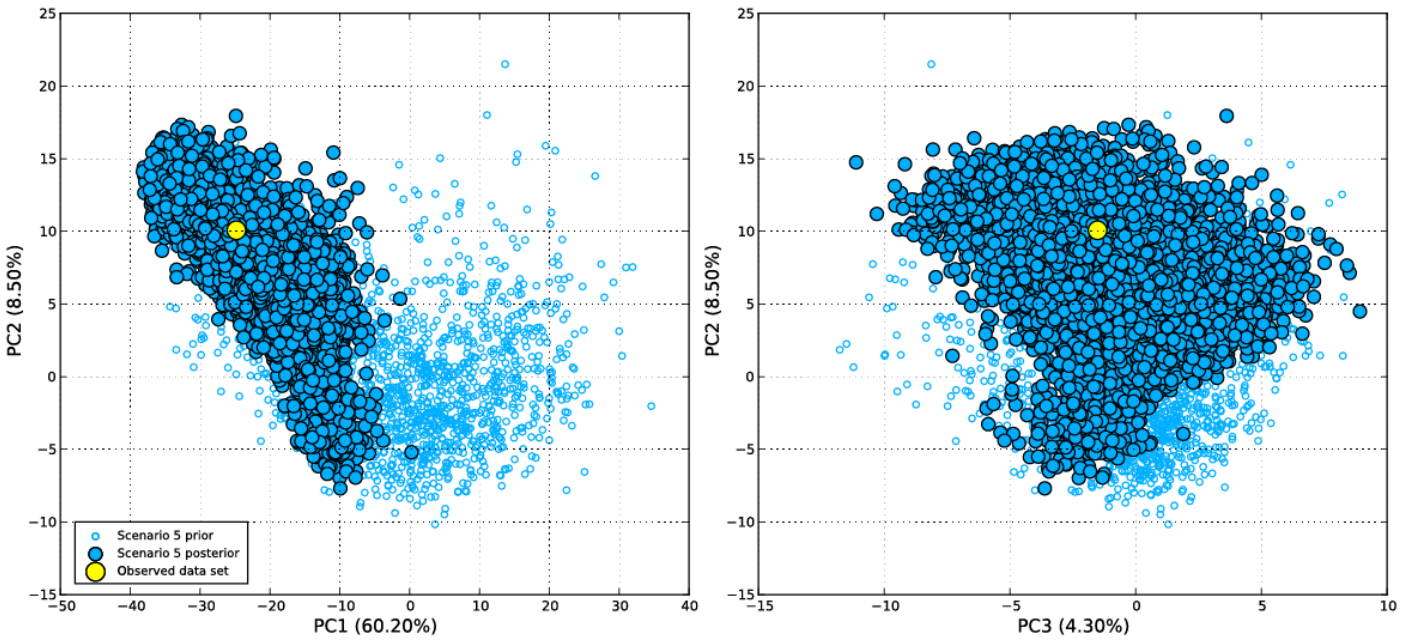


Figure S5. Goodness-of-fit for the final worldwide invasion scenario identified by the ABC-RF analysis, related to Figure 6, Data S2, and to the section "ABC-RF-based inferences of global invasion history" of the STAR Methods.

Principal component analysis (PCA) conducted on the values obtained for 256 summary statistics (Data S2C) used here as test quantities when processing ABC model-posterior checking for the final worldwide invasion scenario detailed in Figure 6 and S4. The two panels display the scatter plots obtained for PC1 and 2 (left panel) and for PC2 and 3 (right panel) from the PCA accounting for up to 73% of the total variation. Small open blue circles show simulated datasets from the prior parameter distributions (10,000 simulations). Large filled blue circles show simulated datasets from posterior parameter distributions (10,000 simulations). The large yellow circle is the observed dataset.

Formae speciales or population		**n**	**S**	**n_hap**	**H**	**π (x 1e-3)**	**k**
P. v. f. sp. vulpina	Native	8 / 4 / 4	3 / 0 / 0	3 / 1 / 1	0.71 / 0.00 / 0.00	3.15 / 0.00 / 0.00	1.57 / 0.00 / 0.00
P. v. f. sp. vinifera	Native	98 / 75 / 136	28 / 10 / 1	14 / 6 / 2	0.89 / 0.67 / 0.32	9.18 / 2.80 / 0.46	4.58 / 1.92 / 0.32
P. v. f. sp. quinquefolia	Native	6 / 1 / 6	1 / - / 3	2 / 1 / 3	0.53 / – / 0.73	1.07 / - / 1.99	0.53 / – / 1.40
P. v. f. sp. riparia	Native	112 / 24 / 162	13 / 4 / 0	12 / 4 / 1	0.44 / 0.60 / 0.00	2.83 / 1.76 / 0.00	1.41 / 1.20 / 0.00
P. v. f. sp. aestivalis							
All populations	–	624 / 1195 / 764	28 / 26 / 1	25 / 20 / 2	0.71 / 0.56 / 0.07	8.98 / 2.73 / 0.09	4.48 / 1.85 / 0.07
North America	Native	172 / 85 / 286	26 / 22 / 1	20 / 16 / 2	0.80 / 0.84 / 0.17	12.19 / 8.09 / 0.24	6.08 / 5.50 / 0.17
Europe	Introduced	258 / 995 / 260	6 / 4 / 0	5 / 4 / 1	0.56 / 0.48 / 0.00	3.24 / 1.48 / 0.00	1.62 / 1.02 / 0.00
South Africa	Introduced	72 / 47 / 90	2 / 4 / 0	3 / 3 / 1	0.39 / 0.45 / 0.00	1.46 / 1.47 / 0.00	0.73 / 1.01 / 0.00
Australia	Introduced	64 / 28 / 54	5 / 3 / 0	3 / 2 / 1	0.23 / 0.30 / 0.00	1.40 / 1.33 / 0.00	0.70 / 0.91 / 0.00
China	Introduced	36 / 30 / 50	2 / 3 / 0	2 / 2 / 1	0.36 / 0.29 / 0.00	1.43 / 1.26 / 0.00	0.71 / 0.86 / 0.00
Argentina	Introduced	22 / 10 / 24	2 / 3 / 0	2 / 2 / 1	0.17 / 0.20 / 0.00	0.69 / 0.88 / 0.00	0.35 / 0.60 / 0.00

Table S1. Genetic variation in different groups of *Plasmopara viticola* for each of the three DNA sequence fragments: β-tubulin (*tub,* n = 424, 499 bp), mitochondrial cytochrome-b (*cytb,* n = 1,299, 685 bp) and ribosomal 28S (r28S, n = 536, 703 bp); related to Figures 2, 5, and to the section "DNA sequencing" and "DNA fragment phylogenetic analyses and diversity" of the STAR Methods.

Values for each gene are separated by a slash (/); n*:* sample size; S: number of segregating sites; n_hap: number of haplotypes; H: haplotype diversity; π: per site nucleotide diversity; k: average number of differences among pairs of sequences.

Continent	Group	Host	n	A_r	P_{Ar}	H_e	H_o	F_{IS}
North America (Native)	NA1	*V. labrusca*	27	2.835 ± 1.222	0.581 ± 0.670	0.405 ± 0.289	0.187 ± 0.136	0.546***
	NA2†	*V. aestivalis*	21	3.676 ± 1.493	0.600 ± 0.987	0.589 ± 0.204	0.399 ± 0.238	0.329***
	NA3	*V. vinifera*	16	2.167 ± 1.011	0.273 ± 0.411	0.310 ± 0.266	0.306 ± 0.279	0.016
	NA4	*V. vinifera*	111	3.607 ± 1.036	0.516 ± 0.398	0.568 ± 0.195	0.386 ± 0.169	0.322***
Europe (Introduced)	WEU	*V. vinifera*	715	2.316 ± 0.633	0.035 ± 0.049	0.3910 ± 0.199	0.356 ± 0.218	0.09***
	EEU	-	265	2.391 ± 0.630	0.030 ± 0.032	0.356 ± 0.184	0.318 ± 0.224	0.108***
South Africa (Introduced)	ZA	-	67	2.250 ± 0.598	0.078 ± 0.186	0.381 ± 0.219	0.339 ± 0.279	0.11
Australia (Introduced)	WAU	-	14	1.727 ± 1.142	0.000 ± 0.000	0.203 ± 0.286	0.232 ± 0.323	-0.148
	EAU	-	36	2.333 ± 0.753	0.001 ± 0.002	0.411 ± 0.182	0.376 ± 0.235	0.09
China (Introduced)	CN	-	25	2.042 ± 0.428	0.012 ± 0.035	0.324 ± 0.178	0.366 ± 0.217	-0.131
Argentina (Introduced)	AR	-	17	1.679 ± 0.457	0.000 ± 0.000	0.285 ± 0.230	0.347 ± 0.373	-0.222

Table S2: Genetic diversity for *Plasmopara viticola forma specialis aestivalis* populations based on eight microsatellite markers; related to Figure 5, Data S1, and to the section "Inference of population genetic structure and diversity" of the STAR Methods.

n: sample size; A_r: allelic richness standardized for a minimum sample size of 14; P_{Ar}: private allelic richness standardized for a minimum sample size of 14; H_o: observed heterozygosity; H_e: expected heterozygosity; F_{IS}: fixation index (*: $P < 0.05$; ** < 0.01; *** < 0.001). †Population in the native range closest to the introduced cultivated populations of the rest of the world vineyards. See Data S1 for pairwise populations comparisons of A_r, P_{Ar} and *He* statistics.

Continent	Group	Range	Host	n	$A_r \pm SD$	$P_{Ar} \pm SD$	H_o	H_e	F_{IS}
North America	NA1	Native	*V. labrusca*	15	1.23 ± 0.23	0.30 ± 0.40	0.30 ± 0.31	0.30 ± 0.23	-0.034
-	NA2†	Native	*V. aestivalis*	5	1.49 ± 0.28	0.26 ± 0.34	0.40 ± 0.34	0.47 ± 0.30	0.156
-	NA3	Native	*V. aestivalis*	4	1.30 ± 0.26	0.16 ± 0.27	0.46 ± 0.46	0.32 ± 0.26	-0.600***
-	NA4	Native	*V. vinifera*	27	1.38 ± 0.19	0.19 ± 0.25	0.33 ± 0.24	0.39 ± 0.21	0.161***
Europe	WEU	Introduced	*V. vinifera*	48	1.35 ± 0.23	0.03 ± 0.05	0.30 ± 0.23	0.34 ± 0.23	0.121***
-	EEU	Introduced	*V. vinifera*	19	1.37 ± 0.24	0.03 ± 0.06	0.36 ± 0.27	0.36 ± 0.24	0.004
South Africa	ZA	Introduced	*V. vinifera*	12	1.32 ± 0.25	0.04 ± 0.10	0.40 ± 0.39	0.32 ± 0.25	-0.279***
Australia	WAU	Introduced	*V. vinifera*	7	1.16 ± 0.25	0.01 ± 0.01	0.25 ± 0.43	0.15 ± 0.25	-0.833***
-	EAU	Introduced	*V. vinifera*	5	1.25 ± 0.25	0.03 ± 0.10	0.31 ± 0.42	0.24± 0.26	-0.318
China	CN	Introduced	*V. vinifera*	24	1.29 ± 0.22	0.02 ± 0.05	0.36 ± 0.31	0.28 ± 0.22	-0.278***
Argentina	AR	Introduced	*V. vinifera*	8	1.22 ± 0.22	0.02 ± 0.04	0.25 ± 0.31	0.23 ± 0.22	-0.100

Table S3: Genetic diversity in *Plasmopara viticola forma specialis aestivalis* populations based on the 32 microsatellite marker dataset; related to Figure 5, Data S1, and to the section "Inference of population genetic structure and diversity" of the STAR Methods.

n: clone-corrected sample size; A_r: allelic richness standardized for a minimum sample size of two; P_{Ar}: private allelic richness standardized for a minimum sample size of two; H_o: observed heterozygosity; H_e: expected heterozygosity; F_{IS}: fixation index (*: $P < 0.05$; ** < 0.01; *** < 0.001). Population acronyms are as follows: North American *P. viticola* f. sp. *aestivalis* strains collected on wild *Vitis labrusca* (NA1), *V. aestivalis* group 1 (NA2) and group 2 (NA3), and North American strains collected on cultivated *V. vinifera* (NA4). Strains in the rest of the world were collected on cultivated *V. vinifera* including Western and Eastern Europe (WEU and WEU), South Africa (ZA), Western and Eastern Australia (WAU and EAU), China (CN), and Argentina (AR). †Population in the native range closest to the introduced cultivated populations of the rest of the world vineyards. See Data S1 for pairwise populations comparisons of A_r, P_{Ar} and *He* statistics.

Scenarios tested in 5 steps	Training set	Decision trees	Prior error rate (%)	RF votes (%)	Post. Prob. ± SD
Step 1 (18 scenarios, 70 summary statistics): North America *aestivalis 1 (NA2) to WEUR / EEUR / CN*	50,000	1,000	38%		
Scenario 11*: Introduction from NA2 > WEUR > EEUR > CN				**20.1%**	**0.42 ± 0.01**
Scenario 8†: Introduction from NA2 > WEUR > CN, EEUR admixed pop CN-WEUR				17.5%	–
Scenario 7†: Introduction from NA2 > CN > WEUR, EEUR admixed pop CN-WEUR				16.0%	–
Scenario 16: Introduction from NA2 > CN > EEUR > WEUR				13.5%	–
Scenario 5: Introduction from NA2 > WEUR > EEUR, CN admixed pop WEUR-EEUR				8.6%	–
Scenario 10: Introduction from NA2 > CN > EEUR, WEUR admixed pop CN-EEUR				6.2%	–
Scenario 9: Introduction from NA2 > EEUR > CN, WEUR admixed pop CN-EEUR				5.0%	–
Scenario 6: Introduction from NA2 > EEUR > WEUR, CN admixed pop WEUR-EEUR				3.9%	–
Scenario 14: Introduction from NA2 > WEUR > CN > EEUR				3.4%	–
Scenario 13: Introduction from NA2 > CN > WEUR > EEUR				2.9%	–
Scenario 18: Independent introduction of CN, then WEUR, EEUR admixed WEUR-CN				0.8%	–
Scenario 17: Independent introduction of WEUR, then CN, EEUR admixed WEUR-CN				0.7%	–
Scenario 12: Introduction from NA2 > EEUR > WEUR > CN				0.5%	–
Scenario 15: Introduction from NA2 > EEUR > CN > WEUR				0.4%	–
Scenario 3: Independent introduction from NA2 to EEUR > WEUR > CN				0.2%	–
Scenario 4: Independent introduction from NA2 to WEUR > EEUR> CN				0.1%	–
Scenario 2: Independent introduction from NA2 to CN> EEUR > WEUR				0.1%	–
Scenario 1: Independent introduction from NA2 to CN > WEUR > EEUR				0.1%	–
Step 2 (4 scenarios based on S11 of step 1, 100 summary statistics): Introduction of South Africa (ZA) from …	10,000	500	9%		
Scenario 2*: EEUR				**50%**	**0.78 ± 0.03**
Scenario 3†: WEUR				42%	–
Scenario 1: CN				8%	–
Scenario 4: NA2				0%	–
Step 3 (5 scenarios based on S2 of step 2, 144 summary statistics): Introduction of East Australia (EAUS) from …	20,000	500	12%		
Scenario 3*: EEUR				**39%**	**0.71 ± 0.02**
Scenario 4†: WEUR				29%	–
Scenario 1: ZA				11%	–
Scenario 2: CN				11%	–
Scenario 5: NA2				10%	–
Step 4 (6 scenarios based on S3 of step 3, 195 summary statistics): Introduction of Argentina (AR) from …	10,000	500	12%		
Scenario 1*: EAUS				**48%**	**0.70 ± 0.02**
Scenario 4: EEUR				23%	–
Scenario 2: ZA				12%	–
Scenario 5: WEUR				9%	–
Scenario 3: CN				5%	–
Scenario 6: NA2				2%	–
Step 5 (7 scenarios based on S1 of step 4, 256 summary statistics): Introduction of West Australia (WAUS) from …	10,000	1,000	7%		
Scenario 5*: EEUR				**23%**	**0.73 ± 0.02**
Scenario 4†: CN				18%	–
Scenario 2†: EAUS				18%	–
Scenario 3: ZA				17%	–
Scenario 6: WEUR				14%	–
Scenario 1: AR				6%	–
Scenario 7: NA2				4%	–

Table S4. Description of the competing scenarios and results of the five successive steps of random forest classification analysis (ABC-RF) to infer the invasion history of *Plasmopara viticola* based on the genetic variation at 32 microsatellite

markers; related to Figures 6, S4, Data S2, and to the section "ABC-RF-based inferences of global invasion history" of the STAR Methods.

For each ABC analysis, an ABC-RF analysis was run 10 times using a training set ranging from 10 to 50 thousand simulations and a forest of 500 or 1,000 trees. The size of the training set and number of trees used in each step was increased until the 10 ABC-RF replicates all converged. The best (most likely) scenario identified at each step is shown in bold and marked with a (*) symbol. Alternative scenario(s) receiving significant support (≥15% of the votes) are indicated with the symbol (†). Prior error rates, proportion of votes, and posterior probability values are averaged over the 10 replicates. Stepwise introduction is indicated as [source] > [derived] populations.

SUPPLEMENTAL REFERENCES

S1. Garza, J.C., and Williamson, E.G. (2001). Detection of reduction in population size using data from microsatellite loci. Mol. Ecol. *10*, 305-318.

S2. Weir, B.S., and Cockerham, C.C. (1984). Estimating F-Statistics for the Analysis of Population Structure. Evolution *38*, 1358-1370.

S3. Rannala, B., and Mountain, J.L. (1997). Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. U. S. A. *94*, 9197-9201.

S4. Pascual, M., Chapuis, M.P., Mestres, F., BalanyÀ, J., Huey, R.B., Gilchrist, G.W., Serra, L., and Estoup, A. (2007). Introduction history of *Drosophila subobscurain* the New World: a microsatellite-based survey using ABC methods. Mol. Ecol. *16*, 3069-3083.

S5. Chakraborty, R., and Jin, L. (1993). A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. EXS *67*, 153-175.

S6. Goldstein, D.B., Ruiz Linares, A., Cavalli-Sforza, L.L., and Feldman, M.W. (1995). An evaluation of genetic distances for use with microsatellite loci. Genetics *139*, 463-471.

S7. Choisy, M., Franck, P., and Cornuet, J.M. (2004). Estimating admixture proportions with microsatellites: comparison of methods based on simulated data. Mol. Ecol. *13*, 955-968.