Threat to Asian wild apple trees posed by gene flow from domesticated apple trees and their pestified pathogens

Alice Feurtey, Ellen Guitton, Marie de Gracia Coquerel, Ludovic Duvaux, Jason Shiller, Marie-noëlle Bellanger, Pascale Expert, Mélanie Sannier, Valérie Caffier, Tatiana Giraud, et al.

HAL Id: hal-02972460
https://hal.inrae.fr/hal-02972460

Submitted on 22 Oct 2020

DR. LUDOVIC JA DUVAUX (Orcid ID : 0000-0003-0960-0312)

DR. TATIANA GIRAUD (Orcid ID : 0000-0002-2685-6478)

DR. BRUNO LE CAM (Orcid ID : 0000-0001-9884-011X)

DR. CHRISTOPHE LEMAIRE (Orcid ID : 0000-0002-3078-5801)

Threat to Asian wild apple trees posed by gene flow from domesticated apple trees and their pestified pathogens

Alice Feurtey[1,2,a], Ellen Guitton[3,a], Marie De Gracia Coquerel[3], Ludovic Duvaux[3,4], Jason Shiller[3,5], Marie-Noëlle Bellanger[3], Pascale Expert[3], Mélanie Sannier[3], Valérie Caffier[3], Tatiana Giraud[1], Bruno Le Cam[3] and Christophe Lemaire[3*]

[1] Ecologie Systématique Evolution, CNRS, AgroParisTech, Université Paris-Saclay, 91400 Orsay, France

[2] Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany

[3] IRHS-UMR1345, Université d'Angers, INRAE, Institut Agro, SFR 4207 QuaSaV, 49071, Beaucouzé, France

[4] BIOGECO, INRAE, Université de Bordeaux, 33610 Cestas, France

[5] Noble Research Institute, Ardmore, OK 73401, USA

[a] Alice Feurtey and Ellen Guitton should be considered joint first authors

* Correspondence: Dr. Christophe Lemaire. IRHS, 42 rue G. Morel. 49071 Beaucouzé Cedex, France.

Tel: +33 2 41 22 57 27

E-mail: christophe.lemaire@univ-angers.fr

Abstract

Secondary contact between crops and their wild relatives poses a threat to wild species, not only through gene flow between plants, but also through the dispersal of crop pathogens and genetic exchanges involving these pathogens, particularly those that have become more virulent by indirect selection on resistant crops, a phenomenon known as "pestification". Joint analyses of wild and domesticated hosts and their pathogens are essential to address this issue, but such analyses remain rare. We used population genetics approaches, demographic inference and pathogenicity tests on host-pathogen pairs of wild or domesticated apple trees from Central Asia and their main fungal pathogen, *Venturia inaequalis,* which itself has differentiated agricultural and wild-type populations. We confirmed the occurrence of gene flow from cultivated *(Malus domestica)* to wild *(Malus sieversii)* apple trees in Asian forests, potentially threatening the persistence of Asian wild apple trees. Pathogenicity tests demonstrated the pestification of *V. inaequalis*, the agricultural-type population being more virulent on both wild and domesticated trees. SNP markers and the demographic modeling of pathogen populations revealed hybridization following secondary contact between agricultural and wild-type fungal populations, and dispersal of the agricultural-type pathogen population in wild forests, increasing the threat of disease in the wild apple species. We detected a SNP potentially involved in pathogen pestification, generating an early stop codon in a gene encoding a small secreted protein in the agricultural-type fungal population. Our findings, based on joint analyses of paired host and pathogen datasets, highlight the threat posed by cultivating a crop near its center of origin, in terms of pestified pathogen invasions in wild plant populations and introgression in the wild-type pathogen population.

Keywords: apple, crop-to-wild gene flow, hybridization, *Malus*, secondary contact, avirulence gene, *Venturia*

Introduction

Plant domestication involves genetic and phenotypic differentiation between crops and their wild relatives under human selection (Zeder, Emshwiller, Smith & Bradley, 2006). Secondary contacts between crops and their wild relatives can lead to introgression (Ellstrand et al., 2013; Wang, Viera, Crawford, Chu & Nielsen, 2017), and potentially drive wild lineages to extinction due to massive gene flow from the related crop (Todesco et al., 2016; Wolf, Takebayashi & Rieseberg, 2001). A much less widely studied, but just as alarming, consequence of secondary contact

between crops and their wild relatives, is the possible dispersal of their respective pathogens and/or gene flow between them. Indeed, crop pathogen populations have often been subjected to indirect selection for virulence following the introduction of resistant crops, a phenomenon known as pestification (Saleh, Milazzo, Adreit, Fournier & Tharreau, 2014), which is potentially highly damaging to wild species. Furthermore, gene flow between differentiated pathogen populations or species parasitizing wild and domesticated hosts can promote the emergence of new diseases or the breakdown of resistance, through the generation of pathogens with an expanded host range (Depotter, Seidl, Wood & Thomma, 2016) or enhanced virulence (i.e. degree of damage to the host caused by the pathogen) (Stukenbrock, 2016; Stukenbrock, Christiansen, Hansen, Dutheil & Schierup, 2012). This phenomenon poses a major threat to both crop health and the persistence of the wild host (Lemaire et al., 2016; Leroy, Lemaire, Dunemann & Le Cam, 2013; Leroy et al., 2016). However, the consequences of crop-to-wild gene flow for the evolution of wild pathogen populations and for the risk of disease on the wild host have been little investigated. Joint analyses of wild and agricultural hosts and pathogens are required to improve our understanding of the consequences of secondary contacts between crops and their wild relatives, and between the corresponding sets of pathogens, and to elucidate coevolutionary histories and local adaptation. However, such studies, including population genetic analyses of both host and pathogen, remain scarce (Croll & Laine 2016).

We addressed these questions for apple trees and their main fungal pathogen, *Venturia inaequalis* (Cooke) G. Winter, the causal agent of apple scab. This pathosystem constitutes a good model for investigations of the consequences of secondary contact between crops and their wild relatives. Both apple trees and *V. inaequalis* have differentiated wild and agricultural-type species/populations that have recently been brought together at their center of origin, where hybridization may occur (Figure 1). The cultivated apple was initially domesticated from *M. sieversii* Ledeb. M. Roem (Figure 1A), which forms quasi-monospecific natural wild apple forests in the Tian Shan Mountains of Central Asia (Cornille et al., 2012; Harris, Robinson & Juniper, 2002; Vavilov, 1931; Velasco et al., 2010), a site that has been declared a World Heritage Center by UNESCO. The cultivated apple was imported into Europe via the Silk Route. This led to major secondary contributions to the cultivated apple tree genepool from several crabapples (Cornille et al., 2012; Cornille, Gladieux & Giraud 2013; Nikiforova, Cavalieri, Velasco & Goremykin, 2013). In particular, in Western Europe, the European crabapple, *M. sylvestris* (L.) Mill, has made a

major contribution to the *M. domestica* genome (Figure 1A). Domestication and breeding led to a strong differentiation between *M. domestica* and its various wild progenitors (Cornille et al., 2012), although introgression can still occur between wild and cultivated forms (Cornille et al., 2012; Nikiforova et al., 2013). This situation raises serious concerns regarding possible gene flow and its deleterious consequences for the conservation of important wild genetic resources. However, most studies on gene flow between domesticated and wild apples have focused on European wild apple trees (Cornille et al., 2015; Feurtey, Cornille, Shykoff, Snirc & Giraud, 2017; Figure 1A). In Central Asia, at the end of the 19th century, *M. domestica* apple varieties from Western Europe were planted in orchards close to *M. sieversii* forests, leading to recent secondary contact between *M. domestica* and *M. sieversii*. A few studies have suggested that there is gene flow from *M. domestica* to the Asian crabapple *M. sieversii* (Cornille et al. 2012; Omasheva et al., 2017). Such gene flow could jeopardize the conservation of *M. sieversii*, increasing the threat to Asian wild-apple forest ecosystems, which are already endangered by forest destruction; *M. sieversii* is now included as a vulnerable organism in the International Union for Conservation of Nature's Red List of Threatened Species (IUCN Red List).

Apple domestication has also fostered divergence in *V. inaequalis*, a haploid ascomycete fungus with a cycle including an annual sexual reproduction event. This fungus causes apple scab disease, which is characterized by gray-brown lesions on leaves and fruits, leading to major economic losses (Belete & Boyraz, 2017) in cultivated apple crops. *Venturia inaequalis* parasitizes *M. domestica* and several wild apple tree species, including *M. sieversii*, *M. sylvestris* and *M. floribunda* (Siebold ex Van Houtte). Differentiated populations of *V. inaequalis* occur on these different *Malus* species (Figure 1B; Gladieux et al., 2010b). Both apple trees and *V. inaequalis* originated in Central Asia (Cornille et al., 2012; Gladieux et al., 2008), *M. sieversii* being the wild host of origin of the fungus (Gladieux et al., 2010b). In the Central Asian Mountains, in forests devoid of *M. domestica*, a wild-type *V. inaequalis* population is present on *M. sieversii* trees (Gladieux et al., 2010b). Agricultural-type populations of *V. inaequalis* also occur in Central Asia (Gladieux et al., 2010b) on *M. domestica* and *M. sieversii* in periurban and cultivated environments (Figure 1B). These wild and agricultural-type populations of *V. inaequalis* began diverging in Central Asia between 2,000 and 4,000 years ago (Gladieux et al., 2010b). The agricultural type then spread to Europe, together with the domesticated apple (Figure 1B; Gladieux et al., 2010b). It has yet to be determined whether the current co-occurrence of agricultural and

wild-type *V. inaequalis* populations represents a secondary contact between the two fungal populations, and its epidemiological consequences, through potential spillover and introgressions, remain unknown (Wang et al., 2016).

Despite the importance of wild Asian apple trees as an endangered wild species and as a valuable genetic resource for future breeding programs, the potential impact of the secondary contact between domesticated and wild apple trees in Central Asia on the fitness of both the plant and its associated pathogens has yet to be investigated. We performed joint analyses on apple trees and their pathogens, using microsatellite markers in apple trees and genome-wide single-nucleotide polymorphisms (SNPs) in *V. inaequalis*, together with pathogenicity assays, to assess the occurrence of secondary contact and its impact on both fruit trees and their fungal pathogens in their center of diversity. We investigated 1) whether introgressions from introduced domesticated apple trees in Central Asia threatened the genetic conservation of *M. sieversii* wild apple trees, and whether orchards planted in forested areas represented a particular threat; 2) whether agricultural-type fungal populations hybridized with the wild-type fungal population. We used demographic modeling and scenario comparisons to test the hypothesis that the co-occurrence and hybridization of agricultural and wild-type *V. inaequalis* populations resulted from secondary contact following the introduction of domesticated trees in orchards close to wild apple forests, after a period of allopatry; 3) whether the agricultural and wild-type *V. inaequalis* populations were specific to domesticated and wild apple trees, respectively; 4) whether particular genomic regions in *V. inaequalis* could be identified as good predictors of its ability to parasitize domesticated or wild apple trees; 5) whether fungal hybridization (question 2) increased the threats to wild apple trees by increasing the virulence of the pathogen populations, as assessed by experimental inoculations.

Materials and Methods

Sampling

Plant and fungal materials were sampled in 2012, from nine sites in the Tian Shan Mountains in Kazakhstan, the center of origin for apple tree domestication, where the natural forest is dominated by wild apple trees (Cornille, Giraud, Smulders, Roldán-Ruiz & Gladieux, 2014). Seven of the nine sampling sites were located in natural apple forests, more than a few tens of kilometers from any orchard. The two remaining sampling sites corresponded to a 20- to 30-year-old orchard and

the wild natural forest immediately surrounding it. We sampled apple trees at the orchard site and wild apple trees in the surrounding forest. This orchard was located in the middle of the natural forest, providing a rare opportunity for investigating the consequences of secondary contact for both the host and the pathogen. We obtained plant material from 245 apple trees at the nine sites. The 245 plant samples corresponded to 185 leaves and 60 pieces of fruit skin, all with visible scab lesions (Table S1). *Malus domestica* and *M. sieversii* have no marked morphological differences that can be used to distinguish them reliably in the field. We therefore assigned tree species and hybrids using genetic markers and reference genotypes (see assignment method below), *i.e.,* 50 apple genotypes previously characterized as being non-hybrid *M. domestica* genotypes and 25 apple genotypes previously characterized as being non-hybrid *M. sieversii* genotypes (Cornille et al., 2012).

Monospore isolations of the fungus were performed on the disease lesions of 205 of these 245 trees. The spore suspension was spread on malt agar medium in a Petri dish and incubated for 24 to 48 h. We then used a needle to remove a single germinated spore from the plate, under a stereomicroscope, for culture on a new Petri dish. We performed one monospore isolation per tree for 148 trees, and two to four monospore isolations per tree for 57 trees. In total, we obtained 269 *V. inaequalis* strains (Table S1). In addition to the fungal collection sampled for this study, our final dataset included 15 *V. inaequalis* strains previously assigned to the pure wild-type population and 18 *V. inaequalis* strains previously assigned to the agricultural-type population (Gladieux et al., 2010b). These 33 *V. inaequalis* reference strains, for which genome sequences were available (Le Cam et al., 2019), were sampled from *M. sieversii* in Kazakhstan, either from a periurban environment near Almaty (43°15000N-76°54000E) or from a natural forest site in the Tian Shan Mountains (43°13807N-77°16783E). Despite being sampled on *M. sieversii*, the 18 strains from the periurban environment clustered with strains sampled from domesticated apples in Kazakh orchards, and were therefore considered to be of the agricultural type (Gladieux et al., 2010b) (Figure 1B; Table S2).

Genotyping of the host apple tree with microsatellite markers
All apple trees were genotyped with microsatellite markers, to determine whether they should be assigned to the wild, domesticated or hybrid type (question 1). Apple tree DNA was extracted

from leaves with the NucleoSpin plant DNA extraction kit II® (Macherey & Nagel). Microsatellites were amplified by PCR with a multiplex PCR kit (QIAGEN, Inc.). We genotyped the 245 trees with 33 microsatellite markers spread across the 17 chromosomes, in 10 different multiplex reactions, as previously described (Cornille et al., 2012). Genotyping was performed at the *Gentyane* platform (INRA, Clermont-Ferrand, France). We used only microsatellite markers with missing data rates of less than 30%, and only individuals with less than 50% missing data for our analyses. The final dataset corresponded to 240 apple tree samples genotyped for 28 microsatellite markers. Reference genotypes with these 28 microsatellite markers were already available (Cornille et al., 2012).

Single nucleotide polymorphisms (SNPs) and genotyping for the apple scab fungus *Venturia inaequalis*

For the genotyping of fungal isolates (for question 2), we used the set of single nucleotide polymorphisms (SNPs) previously detected in the genomes of the 15 wild-type and 18 agricultural-type *V. inaequalis* reference strains (Gladieux et al., 2010b; Le Cam et al., 2019). SNPs were filtered to avoid transposable elements and AT-rich regions, and SNPs with more than 20% missing data were also eliminated (Le Cam et al., 2019). After filtering, the dataset contained 1,210,121 SNPs, with no difference in the frequency of missing data between wild and agricultural-type *V. inaequalis* populations. We defined a set of SNPs distinguishing between agricultural and wild-type *V. inaequalis* strains, by selecting 192 SNPs from the 1,210,121 available SNPs meeting the following three requirements: *i)* moderate ($F_{ST} = 0.21$) to high ($F_{ST} = 1$) differentiation between the wild-type and agricultural-type reference populations (Figure S1, Table S3), to ensure sufficient statistical power for the assignment of genotypes to the wild, agricultural and hybrid types, *ii)* sufficient physical distance to ensure an absence of linkage disequilibrium between them (correlation coefficient between markers within population r² < 0.1), and iii) found in predicted genes (Le Cam et al., 2019).

For each monospore *V. inaequalis* strain, DNA was extracted from 30 to 40 mg of mycelium, as previously described (Leroy et al., 2016). The 269 *V. inaequalis* strains were genotyped at the 192 SNP markers (Table S3) by the *KASPAR* method (KBioscience, competitive allele-specific polymerase chain reaction assay). *KASPAR* genotyping was performed at the *Gentyane* platform

(INRA, Clermont-Ferrand, France) with 10 ng of haploid DNA mixed with the KASP genotyping master mix (Catalogue number: KBS-1016-017, LGC Genomics, Hoddesdon, UK) and customized KASP SNP assays (LGC Genomics, Hoddesdon, UK). Only SNPs and individuals with less than 20% missing data were retained for the analysis. After filtering for missing data, the dataset corresponded to 255 *V. inaequalis* strains genotyped for 181 SNPs. The location of the SNPs on the *V. inaequalis* reference genome are indicated in Table S3. The alleles at the 181 chosen SNP markers were directly extracted from genomic SNP data for the 33 reference strains.

As fungi were sampled from the genotyped host trees, we had genotypes for both of the partners in host-pathogen pairs. We obtained genotypes for 249 host-pathogen pairs, with more than one fungal sample obtained per tree in some cases (132 trees with one fungal strain, 49 trees with two strains, five trees with three strains and one tree with four strains).

Analyses of diversity, differentiation and hybridization

Reference genotypes of *M. domestica* (N=50) and *M. sieversii* (N=25) (Cornille et al., 2012) were used to estimate the proportion of ancestry from these two species in the 240 apple trees genotyped (question 1). We calculated a hybrid index (*HI*) in the add-on R package *Introgress* (Gompert & Buerkle, 2010), by estimating the proportion of alleles inherited from one of the two parental populations, to determine the level of ancestry of the tree for the population concerned. We set *HI* to 1 for pure *M. domestica* trees, and to 0 for pure *M. sieversii* trees. This index therefore represents the proportion of ancestry in the *M. domestica* genepool (referred to here as P_{dom}). For *V. inaequalis* (question 2), we used the same package to calculate a hybrid index based on the reference genotypes of strains previously identified as belonging to the agricultural-type population (N=18) or to the wild-type (N=15) population (Gladieux et al., 2010b). We set *HI* to 1 for fungi of the agricultural-type, such that the hybrid index represented the proportion of ancestry in the agricultural-type fungal population (here called P_{agr}) genepools.

Wherever possible, the hybrid indices were analyzed as continuous values, rather than in discrete classes, to avoid potential biases due to the use of arbitrary thresholds. Only for analyses of the spatial distributions of apple and fungal populations or species, for analyses of the genetic determinants of adaptation and for pathogenicity tests was it necessary to assign apple and fungal genotypes to discrete classes. Apple tree genotypes with a P_{dom} index below 0.2 and those with a

P_{dom} index greater than 0.8 were considered to belong to *M. sieversii* and *M. domestica*, respectively; this threshold was based on the distribution of the hybrid index (Figure 2A). A fungal strain with a P_{agr} below 0.1 was considered to be a wild-type strain, whereas a strain with a P_{agr} greater than 0.9 was considered to be an agricultural-type strain; this threshold was based on the distribution of the hybrid index (Figure 2B). The threshold was more stringent for fungi than for trees, as SNP markers on fungi have been selected to distinguish between agricultural and wild types (*i.e.* high F_{ST}) while microsatellites on apple trees have not been chosen for their discriminative power. Our discrimination power between hybrids and parental species was therefore expected to be more powerful in fungi than in apple trees. For both pathogens and hosts, genotypes with intermediate hybrid index values were considered to be hybrids. It was not possible to assign hybrids to more precise hybrid classes (*e.g.* F1, F2 or backcrosses) for a haploid organism such as *V. inaequalis*. When using discrete classes for analyses, we first checked that the choice of the threshold had no effect on our inferences.

For both apple trees and *V. inaequalis* (questions 1 and 2, respectively), we also investigated genetic differentiation and species or genetic lineage assignment by principal component analysis (PCA) with the ade4 R package (Dray & Dufour, 2007), based on multilocus genotypes (microsatellite markers for trees and SNPs for fungi). We further investigated the existence of population structure within apple tree species and within the previously delimited *V. inaequalis* populations, with STRUCTURE v.4.3 software (Falush, Stephens & Pritchard, 2003), using 10 runs of 250,000 Monte Carlo Markov Chain (MCMC) iterations after a burn-in of 25,000 iterations, according to the admixture model and assuming a number of clusters ranging from 1 to 6. STRUCTURE output was processed with Clumpp (Jakobsson & Rosenberg, 2007), and we used the method described by Evanno, Regnaut and Goudet (2005) and implemented in STRUCTURE HARVESTER (Earl & vonHoldt, 2012) to detect the strongest level of population subdivision. We assessed the significance of correlations between the degrees of hybridization inferred from the values of the coordinates of the first PCA axis, the hybrid index and the probability of ancestry provided by STRUCTURE, by calculating Pearson's product-moment correlation coefficient in R.

For both apple trees and *V. inaequalis* populations (questions 1 and 2, respectively), we estimated classical population genetics statistics: expected heterozygosity (H_e, Weir & Cockerham, 1984), and global and pairwise differentiation, as assessed by F_{ST} (Weir & Cockerham, 1984), between

the nine sampling sites. All statistics were computed and tested for significance with GENEPOP v.4.7.5 (Rousset, 2008). We assessed the significance of global and pairwise F_{ST} with Fisher's exact tests of allele frequency differences between sampling sites, with 10,000 dememorizations and batches, and 10,000 iterations per batch. For question 3, we also performed Mantel tests of correlation (Mantel, 1967) between the pairwise F_{ST} matrices computed for apple and *V. inaequalis* populations on the one hand, and between each pairwise F_{ST} matrix and a matrix of geographic distance between sampling sites on other hand, using the *mantel.test* function of the ape v5.3 R package (Paradis & Schliep, 2019).

Genetic footprints of host adaptation in *Venturia inaequalis*

We searched for genetic signatures of adaptation to host species in *V. inaequalis* populations (question 4), by performing a discriminant analysis on principal components (DAPC; Jombart, Devillard & Balloux, 2010) implemented in the R package *adegenet v.2.12* (Jombart, 2008). For this analysis, we defined groups within *V. inaequalis* on the basis of the group to which the apple trees from which they were collected were assigned (*M. domestica*, *M. sieversii* and hybrids). This approach also aims to detect the genetic factors (here SNPs) making the largest contribution to this ecologically based clustering. If one of the 181 SNPs in *V. inaequalis* is closely linked to a genetic variant involved in adaptation to host species, its loading on the first discriminant functions would be expected to be greater than that of other SNPs. We also estimated monolocus pairwise F_{ST} between the three host-based fungal clusters, with Genepop v4.7.5 (Rousset, 2008).

We determined which alleles were ancestral or derived at the SNP found to predict the host species from which the fungus was collected, using five available genomes (Le Cam et al., 2019) belonging to the outgroup *V. inaequalis* lineage parasitizing pyracantha (Gladieux, Caffier, Devaux & Le Cam, 2010a; Le Cam, Parisi & Arene, 2002), with the following NCBI accession numbers SAMN07816617 to 07816621 (Le Cam et al., 2019).

Tests of *V. inaequalis* pathogenicity on *M. domestica* and *M. sieversii*

We investigated the effect of the agricultural-type, hybrid or wild-type status of *V. inaequalis* strains on their pathogenicity on *M. domestica* and *M. sieversii* (question 5), by performing

artificial inoculations with 57 strains from our sampling campaign, corresponding to 20 strains identified as of the agricultural type, 17 strains identified as wild-type and 20 hybrid strains (Table S1). We used one variety of *M. domestica* (GALA®, X4712) and one accession of *M. sieversii* (GMAL 3619.b) for the pathogenicity tests. Gala is a variety that has been planted worldwide despite being highly susceptible to apple scab. GMAL 3619.b was collected from the Tarbagatai mountain range by American apple breeders (Forsline, Aldwinckle, Dickson, Luby & Hokanson, 2003) and was subsequently introduced in France. We grafted the plants onto the MM106 rootstock to obtain vigorous growth favoring disease development, and the plants were inoculated with *V. inaequalis* strains while they were actively growing. The grafted plants were transferred to a quarantine chamber with controlled climatic conditions before inoculation. We used a quarantine chamber to reduce the risk of pathogen escape, as the strains were of Asian origin. Due to the lack of space available in such chambers, we were obliged to perform experiments on two different dates. In the first experiment, we used 10 agricultural-type strains, 10 hybrid strains and eight wild-type strains. In the second experiment, we used 10 agricultural-type strains, 10 hybrid strains and nine wild-type strains. For the purposes of normalization, we also included six of the strains (two strains of each type) from the first experiment in the second experiment.

Each *V. inaequalis* strain was grown on a cellophane sheet placed on malt-agar medium at 17°C to obtain spores (Caffier et al., 2014). The cellophane sheets were dried and stored at -20°C. A spore suspension was obtained by collecting spores from these cellophane sheets and mixing them with water to attain a final concentration of 1.5×10^5 spores mL^{-1} (Lê Van et al., 2012). The spore suspension obtained from each *V. inaequalis* strain was used to inoculate three Gala trees and three GMAL 3619.b trees with a mechanical pressurized sprayer. As the trees were actively growing during the experiment, we labeled the youngest fully deployed leaf on each tree as the F0 leaf, one or two days before inoculation, to facilitate disease scoring. Spore germination and fungal infection were favored by keeping the plants in darkness, at 100% humidity and at a temperature of 17°C for 48 hours after inoculation (Lê Van et al., 2012). Humidity levels were then decreased to 80% during the day and 90% at night, and the plants were supplied with 12 hours of light per day. These climatic conditions are highly favorable to scab infection and this protocol therefore provided us with a good indication of the ability of the strain tested to infect tree genotypes (MacHardy, 1996). For each host genotype, plants were randomized to three blocks, with one host replicate for each strain in each block. Disease severity (*i.e.*, the virulence of

the fungal strain) was measured when disease symptoms had ceased to worsen: 19 days post inoculation (dpi) for *M. sieversii* and 21 dpi for *M. domestica.* Disease severity was measured as the percentage of a leaf displaying sporulation, from 0% in the absence of disease symptoms to 100% when the whole leaf was covered with fungal spores. The F0 leaf being the last leaf deployed one or two days before inoculation, assessments were performed for the F0 leaf and the leaf immediately below F0, the F1 leaf.

Statistical analysis of phenotypic data

We standardized the data from the two independent inoculation experiments (question 5), based on the data for the six replicated strains used for calibration: disease severity values were corrected to obtain the same mean between two experiments for the six calibration strains. The virulence of each strain was estimated as the mean percentage of the leaf area bearing scabs over the two leaves assessed (F0 and F1), for the three inoculated plants of a given genotype (either *M. domestica* or *M. sieversii*). We therefore analyzed one value per fungal strain and per apple tree genotype. As the assumptions of normality were not met, we compared the median virulence between the three populations (agricultural-type, wild-type and hybrid) in Kruskal-Wallis and *post-hoc* Wilcoxon tests performed 21 dpi for *M. domestica* and 19 dpi for *M. sieversii* with R (version 3.4.4; R Core Team, 2018), considering the different strains in each population as replicates for assessment of the population effect. A Kendall test was also performed to analyze the correlation between virulence and hybrid index P_{agr} for the 20 hybrid strains.

Demographic inference for wild and agricultural types of the apple scab fungus *Venturia inaequalis*

We tested demographic scenarios (question 2), including the likelihood of a history of divergence without gene flow followed by secondary contact in the fungal pathogen, using the composite-likelihood, diffusion approximation-based approach for demographic inference implemented in ∂a∂i (Gutenkunst, Hernandez, Williamson & Bustamante, 2009). We used the 33 reference *V. inaequalis* genomes from 15 wild-type and 18 agricultural-type strains described above. We ensured that the markers used were independent, by thinning the 1,210,121 SNPs available for these genomes to retain only one SNP every 5 kb (a threshold based on the LD decay

curve, Figure S2) with vcftools (Danecek et al., 2011). This left us with 6,187 SNPs. Using $\partial a \partial i$ (Gutenkunst et al., 2009), we tested a set of four basic models (Tine et al., 2014): strict isolation (SI), in which the two lineages diverge without gene flow; isolation with migration (IM), in which the two lineages diverge with constant gene flow; ancient migration (AM), in which the two lineages diverge with gene flow and stop exchanging genes at a time denoted T_{AM}; and secondary contact (SC), in which the two lineages diverge without gene flow and then begin to exchange genes following a secondary contact at time T_{SC}. Each of these four models was evaluated and fitted with the observed joint allele frequency spectrum (jAFS), in 20 independent runs. Each run started with a perturbation of the starting parameters, followed by global optimization with a simulated annealing optimization procedure and an optimization phase with up to 100 iterations (https://popgensealab.wordpress.com/dadi-inference/; Christe et al., 2017; Tine et al., 2014). For nested models, likelihood ratio tests (LRTs) were used to identify the best model. For non-nested models, we used the Akaike criterion (AIC) (Christe et al., 2017).

Demographic parameters were estimated for the four models. These parameters included migration rates in the two directions, effective sizes in the two lineages, divergence time and time elapsed since other demographic events (T_{AM} in AM model and T_{SC} in SC model). The ancestral effective size before the split was calculated as $N_{ref} = \theta / (2 \times \mu \times L)$ (Gutenkunst et al., 2009), with θ estimated by $\partial a \partial i$, where μ is the mutation rate per nucleotide per generation, estimated at 2×10^{-8} from an analysis of divergence from the closely related species *Venturia pirina* parasitizing pear trees (Le Cam et al., 2019; data not shown). L is the effective genome length analyzed, calculated as G x s/S (Gutenkunst et al., 2009), G is the size of the genome used, S is the total number of SNPs called in G, and s is the actual number of SNPs used for inference. In our analysis, 1,210,121 SNPs were called on a portion of 34,942 Mb of the genome and we used 6,187 SNPs for the analysis, so L was calculated as 34,942,000 x (6,187 / 1,210,121) = 178,648 bp. All other parameters were scaled by N_{ref}. Times in years were calculated with a generation time of one year, as *V. inaequalis* undergoes an obligate sexual reproduction event once per year. The model parameters were estimated and their 95% confidence intervals were obtained with Godambe methods (Coffman, Hsieh, Gravel & Gutenkunst, 2015), with 1000 bootstraps across SNPs.

Results

Contrasting distributions of wild and agricultural-type host and pathogen populations in the Tian Shan Mountains

We explored the genetic makeup of apple trees and their scab pathogens in a wild environment, by sampling apple trees and fungal strains from the same trees, at nine sites in the forests of the Tian Shan Mountains in Kazakhstan. For apple trees, the PCA (Figure 2A) and STRUCTURE analysis (Figures S3A and B) confirmed the existence of two distinct genetic clusters, one corresponding to the orchard trees assigned to *M. domestica*, and the other to most of the wild forest trees assigned to *M. sieversii*. We detected no further population subdivision for the wild species (Figures 2A and S3). Both the first axis of the PCA and STRUCTURE analyses revealed footprints of admixture between the two tree species. We also estimated P_{dom} hybrid index values for the 240 successfully genotyped trees. All three methods gave consistent results (Figs. 2A and S3C). The correlation coefficient for the relationship between P_{dom} and the coordinates of the first PCA axis was high and significant (r=-0.97; *p*-value < 0.001; Figure S4A). The correlations between the proportion of ancestry inferred by STRUCTURE and P_{dom} on the one hand, and between the proportion of ancestry inferred by STRUCTURE and the coordinates of the first axis of the PCA on the other, were also strong and significant (r=0.96, *p*-value<0.001, and r=-0.97, *p*-value<0.001, respectively). Based on the P_{dom} hybrid index, we identified 78% pure wild *M. sieversii* trees, about 12% *M. domestica* trees and almost 10% hybrid trees (Figure 2A; Table S1). Genetic species identifications were consistent with our expectations based on sampling conditions: only one *M. domestica* was identified at natural forest sites (Figure 3), whereas most of the trees sampled in the orchard belonged to *M. domestica* (28 of the 30 trees sampled in the orchard, with 14 different genotypes). Three of the forest sites contained only *M. sieversii* trees, but the other four contained both *M. sieversii* and hybrid trees (with 11.1 to 27.6% hybrid trees, Figure 3), as did the forest site directly surrounding the orchard (in which the proportion of hybrid trees reached 27.3%).

For *V. inaequalis*, the PCA (Figure 2B) and STRUCTURE analysis (Figures S5A and B) confirmed the existence of two differentiated populations and an absence of further subdivision, and revealed admixture between the wild and agricultural-type populations (Figures 2B and S5B). The hybrid index P_{agr}, the ancestry coefficient given by STRUCTURE and the first axis of the PCA, separating individuals, according to their degree of ancestry, into the two divergent populations (Figure 2B), gave consistent results (Figure S5C). The correlation coefficient for the

relationship between P_{agr} and the first axis of the PCA was high and significant (r=0.99; p-value=0.001) (Figure S4B). The correlations between the proportion of ancestry inferred by STRUCTURE and P_{agr}, and between the proportion of ancestry inferred by STRUCTURE and the coordinates of the first axis of the PCA were also strong and significant (r=0.99, p-value<0.001 and r=0.98, p-value<0.001, respectively).

The P_{agr} hybrid index estimated for the 255 *V. inaequalis* strains followed a U-shaped distribution, with 33 fungal strains (12.9%) having P_{agr} values between 0.1 and 0.9 and therefore being considered to be hybrids, and high frequencies of pure agricultural-type (38.5%) and pure wild-type (48.6%) fungal genotypes (Figure 2B; Table S1). All the fungal strains sampled in the orchard were of the agricultural type (Figure 3), and the fungal strains sampled in the surrounding forest were mostly either pure agricultural or wild-type genotypes. A single hybrid strain was detected in the forest surrounding the orchard, whereas 3.4% to 37% of the fungal strains at the other forest sites were considered to be hybrids (Figure 3).

We explored the relationship between host and pathogen genotypes further, by comparing genetic diversity and differentiation across sampling sites. Genetic diversities per sampling site ranged from 0.703 (tuk) to 0.802 (kot) for apple trees and from 0.244 (tuk) to 0.487 (kot) for *V. inaequalis* (Table S4). Genetic diversities were minimal and maximal at these sampling sites for both host and the pathogen, but no significant correlation was observed between the diversity levels of the apple and the fungus across sites (Kendall's tau=0.389, p-value=0.18; Figure S6). The global F_{ST} between sites was 0.259 for apple trees and 0.243 for *V. inaequalis*. Pairwise F_{ST} estimates ranged from 0.000 to 0.207 for apple trees (Table S5) and from 0.000 to 0.721 for *V. inaequalis* (Table S6). For both host and pathogen, the orchard (esi_o) displayed the strongest differentiation from other sampling sites (Tables S5 and S6). Host and pathogen pairwise differentiation matrices were significantly correlated (Mantel-test *Z*-score=0.681, p-value=0.011; Figure S7). However, we found no significant correlation between pairwise F_{ST} and geographic distances matrices, for apple trees (Mantel test *Z*-score= 33.10, p-value=0.42) or for *V. inaequalis* (Mantel test *Z*-score=103.20, p-value=0.51), indicating a lack of isolation by distance.

All pathogen types occur on wild apple trees, but wild-type pathogens occur only on trees with no domesticated ancestry

We investigated the association between fungal pathogen type and *Malus* genotypes. We obtained 249 host-pathogen pairs from the 240 trees and 255 fungal strains successfully genotyped, several strains having been isolated from the same host tree. A sharp L-shaped pattern was observed when the hybrid index values of the host trees were plotted against those of their associated fungal genotypes (Figure 4A). Hybrids were found among both trees and fungi, but no hybrid pathogens were collected from hybrid trees. Hybrid and wild-type fungi were found only on wild apple trees, whereas agricultural-type fungi were found on all types of apple trees, across the whole range of hybrid index.

Genetic determinants of host adaptation in *Venturia inaequalis*

For the identification of SNPs predicting the ability of fungal genotypes to parasitize the host species, we performed a DAPC on the genotypes of the 249 *V. inaequalis* strains, with groups defined *a priori* on the basis of the assignment of the tree from which they were collected to the *M. domestica*, hybrid or *M. sieversii* group. The first linear discriminant function (LD1) clearly separated the strains sampled from domesticated apple trees from those sampled from wild and hybrid apple trees (Figure S8). The second linear function separated strains found on hybrid apple trees from those sampled from *M. sieversii* (LD2, Figure S8). An analysis of the contribution of alleles (loadings) to the first discriminant function revealed that the *V_08160_319* SNP was strongly implicated in genetic differentiation between the strains parasitizing *M. domestica* and those found on hybrid and wild hosts (loading of 0.123, Figure S9), whereas the other SNPs made much smaller contributions (from 8.723×10^{-8} to 0.024). The estimated global F_{ST} was, accordingly, much higher for the V_08160_319 locus ($F_{ST} = 0.887$) than for other loci (Figure S10).

The *V_08160_319* SNP was located in a gene identified as encoding a small secreted protein (SSP) as described by Le Cam et al. (2019). The two alleles of this SNP were A and G, and were non-synonymous. According to the allele sequences in the 33 reference genomes, the *V_08160_319* SNP was associated with a second SNP in the same codon, generating a TGA stop codon rather than a TAC codon, encoding a tyrosine, at position 140 in the 295 amino-acid long protein. For the 189 strains for which genotypes were available for this SNP, the stop codon allele had a frequency of 1.00 in *M. domestica* (*N*=28), 0.92 in hybrid apple trees (*N*=13) and 0.06 in *M.*

sieversii (*N*=148) (Figure 4B). Taking the ancestry of fungal strains into account, the stop codon allele had a frequency of 0.89 (*N*=56) in the agricultural-type population and 0 in the wild-type population (*N*=107) and in hybrids (*N*=26). The detection of only A (tyrosine) alleles in hybrids significantly deviated from expectations under the hypothesis of neutral segregation at this locus, given the allele frequencies in the pure populations (binomial-test, *p*-value=6.10×10^{-5}). No known avirulence gene colocalized with this new candidate gene.

We used five available genomes (Le Cam et al., 2019) belonging to the outgroup *V. inaequalis* lineage parasitizing pyracantha (Gladieux et al., 2010a; Le Cam et al., 2002) to identify the ancestral allele. These genomes carried the TAC codon, as did the wild-type *V. inaequalis* population on *M. sieversii*. The stop codon truncating the small secreted protein is, thus, a derived allele that evolved in the agricultural-type populations of *V. inaequalis* growing on *M. domestica*.

Agricultural-type and hybrid fungal strains are more virulent than wild-type strains

We then investigated experimentally whether the observed association between fungal pathogen type and host tree species resulted from differences in infectivity, by inoculating one genotype of *M. sieversii* and one genotype of *M. domestica* with 57 fungal strains (20 agricultural-type, 17 wild-type and 20 hybrid strains; Table S1). All the agricultural-type fungal strains were able to cause disease on *M. domestica* (*N*=20; Figures 5A and 6A), whereas only one wild-type and two hybrid strains were able to cause disease on *M. domestica*, and even then, with only very low levels of virulence (less than 2% of the leaf area diseased, Figure 5A). All agricultural-type strains (*N*=20), hybrids (*N*=20) and wild-type strains (*N*=17) of the fungus were able to cause disease on *M. sieversii* (Figure 5B). Both agricultural-type and hybrid fungal strains were significantly more virulent (i.e. caused larger and/or more numerous lesions) than wild-type strains on *M. sieversii* (Figures 5B and 6B; Wilcoxon tests: *P*=1.7×10^{-5} and *P*=1.2×10^{-4}, respectively). The *V. inaequalis* hybrids were less virulent on *M. sieversii* and appeared to be intermediate between the agricultural-type and wild-type populations in terms of virulence, although the difference between hybrids and agricultural-type populations was not significant (Figure 5B). Similar results were obtained when the same hybrid index threshold as for apple trees was used for assignment to species and hybrids (wild-type: *Pagr*<0.2; hybrids: $0.2 \leq Pagr \leq 0.8$; agricultural-type: *Pagr*>0.8; Figure S11). We found no significant correlation between the proportion of agricultural-type

ancestry and the degree of virulence of the hybrid fungal strains (Figure S12; Kendall test: τ=-0.047, P=0.7702).

The allele at the outlier locus *V_081690_319* in *V. inaequalis* partially explained virulence on *M. domestica*. Forty-seven genotypes were available for the 57 strains used in the experiment. Thirteen of the 15 agricultural-type strains for which genotypes were available (87%) carried the stop codon (TGA) allele. Conversely, all 15 hybrid strains for which genotypes were available (100%), and 17 wild-type strains (100%) carried the tyrosine (TAC) allele. A comparison of the two groups of strains defined on the basis of their alleles at the *V_081690_319* locus showed that strains carrying the stop codon (TGA) allele were more virulent than those carrying the tyrosine (TAC) allele, on both *M. domestica* and *M. sieversii* (Wilcoxon tests: P=1.4x10^{-8} and P=1.2x10^{-3}, respectively; Figure S13). Most of the fungal strains carrying the tyrosine allele were unable to cause disease on *M. domestica*, but a few induced mild symptoms (Figure S13A).

Demographic inference of the divergence history of wild and agricultural-type populations of *Venturia inaequalis*

We used demographic inference methods to determine whether the observed co-occurrence and hybridization of wild and agricultural-type *V. inaequalis* populations in Kazakhstan resulted from divergence with gene flow or originated from secondary contact following the introduction of *M. domestica*, about a century ago, in a few orchards in the forests of wild *M. sieversii*. As detailed below, we compared the likelihoods of four contrasting divergence scenarios involving strict isolation, isolation with continuous migration, isolation with initial migration or secondary contact (Figure 7A).

The models allowing gene flow were better supported than the strict isolation model (AIC= 2403.587; LogL= -1197.793). The model of isolation with migration (IM; AIC= 1975.638; LogL= -981.819) did not perform significantly better than the ancient migration model (AM; AIC= 1977.698; LogL= -981.849). The secondary contact model (SC) was the best-supported model (AIC= 1871.313, LogL= -928.656), indicating that the agricultural-type *V. inaequalis* population was probably absent from Asian wild apple forests until the recent introduction of domesticated apple trees in orchards (Figure 7B; Table 1). Using the θ estimate obtained with the best model (SC, θ=143.729), and a mutation rate μ of 2x10^{-8}, we found that the ancestral effective size before split, N_{ref}, was close to 20,000 individuals (Figure 7C). Parameter estimates for the secondary

contact model indicated that the divergence between the wild-type and agricultural-type *V. inaequalis* populations occurred 4,570 years ago (4,570 ± 324 SD), and was followed by a secondary contact 130 years ago (130 ± 9 SD). This scenario and estimates are consistent with fungal divergence triggered by apple domestication under strict genetic isolation, followed by recent secondary contact promoted by the introduction of modern orchards (Figure 7C).

Discussion

We show here that wild *M. sieversii* apple trees were affected by gene flow following the introduction of domesticated apple trees in orchards close to the natural apple forests of the Kazakh Tian Shan Mountains. Gene flow from *M. domestica* to *M. sieversii* was suggested in previous studies (Cornille et al. 2012; Omasheva et al., 2017), but we reveal here the strong impact of the presence of orchards close to wild forests. Gene flow currently appears to pose a less severe threat to *M. sieversii* in Asian than it does to *M. sylvestris* in Europe, where wild European crabapple tree populations are displaying massive introgression from the cultivated apple tree, to the extent that the European wild species is considered endangered (Cornille et al., 2015; Feurtey et al., 2017). The lower level of introgression into wild populations of *M. sieversii* in Central Asia probably reflects more recent and much less extensive secondary contact between wild and cultivated species than in Europe. The cultivated apple tree was introduced into Europe before the 3rd century BC (Cornille et al., 2014), whereas it was not introduced into the area sampled here until about 100 years ago, and its introduction in Asia has been much less extensive than that in Europe. Furthermore, the higher density of *M. sieversii* trees in the wild forests of Kazakhstan (Harris et al., 2002; Vavilov, 1931) than of *M. sylvestris* in the wild forests of Europe probably also restricts gene flow from cultivated apple trees. Indeed, pollination distances decrease with increasing density around pollinated trees in European wild apple trees, with longer pollination distances increasing the risk of interspecific mating with cultivated trees (Feurtey et al. 2017; Reim et al. 2015).

The overall percentage of apple tree hybrids in the wild forest remained low, but almost 25% of the trees were found to be hybrids in the forest immediately adjacent to a cultivated orchard, and at the forest site closest to the periurban environment of Almaty. These findings raise concerns about the consequences of increasing human-induced changes in this region. As indicated by the IUCN

Red List, *M. sieversii* is already threatened by the shrinking of its range in Central Asia (Eastwood, Lazkov & Newton, 2009), with the disappearance of more than 70% of the wild apple tree forest area over the last 30 years due to agricultural expansion and overgrazing. In addition to this range reduction, our results support the view that there is also a threat to the conservation of *M. sieversii*. Thus, unless preventive measures are taken to protect Central Asian forests, it may be only a question of time before the genetic integrity of *M. sieversii* deteriorates in Central Asia, as already observed for *M. sylvestris* in Europe (Cornille et al., 2015; Feurtey et al., 2017), and for other native plant species in Europe, including poplars (Bleeker, Schmitz & Ristow, 2007; Rhymer & Simberloff, 1996; Vanden Broeck, Villar, Van Bockstaele & Van Slycken, 2005).

We also reveal an additional threat to *M. sieversii* in Central Asia associated with cultivated apple tree orchards, due to the dispersal of the agricultural-type *V. inaequalis* apple scab pathogen into wild forests. Indeed, we show that agricultural and wild-type *V. inaequalis* populations recently came into secondary contact after having diverged in strict isolation during apple tree domestication, and that the agricultural-type fungal population is more virulent, and is spreading in wild forests, with introgression into the wild-type fungal population. Demographic inferences showed that the agricultural and wild-type fungal populations initially diverged without gene flow about 4,500 years ago, a time corresponding to apple tree domestication, and came into secondary contact about a century ago, when domesticated apple trees were introduced into Central Asia. This secondary contact has led to the recent invasion of wild *M. sieversii* forest by the agricultural-type population and to the production of hybrids.

We also demonstrate the pestification of the agricultural-type *V. inaequalis* population, with the agricultural-type *V. inaequalis* population and hybrids displaying greater virulence on wild apple trees than the endemic wild-type population. Wild-type and hybrid fungal strains were found only on *M. sieversii* trees, whereas agricultural-type fungal strains were found on trees of all ancestries. Pathogenicity tests in controlled conditions were consistent with the distribution observed in nature. Indeed, the agricultural-type fungal strains were able to cause disease on *M. domestica* and *M. sieversii*, whereas the wild-type and hybrid fungal strains were able to induce symptoms only on wild host trees. These results, obtained with high spore concentrations and in favorable climatic conditions, are conservative, and confirm the conclusions previously obtained with a limited

number of *V. inaequalis* strains (Lê Van et al., 2012). We also detected agricultural-type *V. inaequalis* strains at eight different sites in wild forests, on many different trees, suggesting that the agricultural-type population can parasitize a large range of *M. sieversii* genotypes. By contrast, no wild-type or hybrid fungal strain was detected on any of the *M. domestica* genotypes present in the orchard. This host-range pattern probably explains the significant correlation found between the apple tree and *V. inaequalis* pairwise F_{ST} matrices, which was not due to isolation by distance patterns in hosts and pathogens. Indeed, the proportion of *M. domestica* ancestry at each site probably drives the allele frequency distribution in *V. inaequalis* by selection, because *M. domestica* can be parasitized only by agricultural-type strains, and not by wild-type or hybrid pathogens.

The spread of the pestified agricultural-type fungal pathogen into forests may further increase the threat to the wild tree species. Indeed, scab disease on apple tree leaves reduces photosynthesis (Spotts & Ferree, 1979), potentially affecting the growth of the plants. The observed increase in virulence may have a limited impact on the growth of adult trees, but its effect on the growth of young seedlings may be severe. Furthermore, scab disease can lead to early fruit fall (MacHardy, 1996), decreasing seed number, thereby affecting host population dynamics. On *M. sieversii*, hybrid strains were more virulent than wild-type fungal strains but their virulence was not significantly different from that of the agricultural-type fungal strains. Further experiments are required to estimate precisely the effects of disease on the fitness of seedlings and on the number of seeds produced by adult trees, to determine the long-term impact of this pathogen population on natural populations of wild trees. Nevertheless, our results suggest that the invasion of wild forests by the agricultural-type *V. inaequalis* population may increase the threat to natural *M. sieversii* populations.

Our findings suggest that *V. inaequalis* populations are evolving by host tracking of the domesticated apple tree, with no loss of the ability to parasitize the wild host. Many resistance genes have introgressed into *M. domestica* during domestication and modern breeding (Cornille et al., 2012). During its tracking of apple tree evolution, the agricultural-type population of *V. inaequalis* has probably accumulated multiple alleles counteracting the resistance genes present in domesticated apple crops, possibly including the truncated allele of the gene encoding small

secreted protein identified in this study as a putative avirulence gene. Indeed, we found that the host lineage on which *V. inaequalis* strains were sampled was predicted by a SNP in this gene. In the fungal strains collected on *M. domestica*, the allele corresponded to a stop codon, whereas strains sampled from *M. sieversii* carried the full-length gene. Together with the presence of the full-length gene in an outgroup, the *V. inaequalis* lineage parasitizing pyracantha (Gladieux et al., 2010a; Le Cam et al., 2002), this suggests that a virulence allele has been acquired by the agricultural-type *V. inaequalis* population during crop host tracking, through the loss of a protein recognized by the domesticated apple tree selected for resistance to apple scab. Such an increase in the virulence of a pathogen following domestication of its host is consistent with a process of pestification (Saleh et al., 2014), and is consistent with the virulence evolution mechanisms described in many crop pathogens (Raffaele & Kamoun, 2012).

The identification of a putative avirulence gene from a small set of 181 SNPs can be explained by the initial choice, for the analysis, of the set of SNPs most strongly differentiated between the fungal populations and located within genes. However, we cannot exclude the possibility that the *V_081690_319* SNP is not directly involved in host adaptation, but is instead linked to one of its genomic determinants. Functional experiments are required to confirm that the identified gene is an avirulence gene. On the plant side, the identification of the receptor involved in recognizing this putative avirulent protein may lead to the discovery of a new resistance gene.

In conclusion, we found that natural *M. sieversii* populations are currently only mildly affected by gene flow from *M. domestica,* which was introduced into the area about a century ago but remains very rare in the Kazakh mountains. Further disturbances in this area might, however, lead to much higher levels of crop-to-wild gene flow, as already reported in the European crabapple in Europe (Cornille et al., 2013; 2015; Feurtey et al. 2017). In addition, the agricultural-type *V. inaequalis* population is invading wild forests and introgression into the wild-type *V. inaequalis* population is occurring, probably due to the greater virulence of the agricultural-type fungal population acquired during its tracking of apple tree domestication. This study presents one of the very few joint analyses of host and pathogen populations performed to date (Croll & Laine, 2016), despite the importance of such analyses for understanding the evolutionary mechanisms and histories leading to host specialization and local adaptation. Thanks to this study of host-pathogen pairs, we were able to identify a putative avirulence gene of potential importance in the cultivated crop. Similar approaches may prove fruitful in future studies on host and pathogen populations.

References

Belete, T., & Boyraz, N. (2017). Critical review on apple scab (*Venturia inaequalis*) biology, epidemiology, economic importance, management and defense mechanisms to the causal agent. *Journal of Plant Physiology and Pathology*, 5(2) doi: 10.4172/2329-955X.1000166

Bleeker, W., Schmitz, U., & Ristow, M. (2007). Interspecific hybridisation between alien and native plant species in Germany and its consequences for native biodiversity. *Biological Conservation*, *137*(2), 248-253. doi: 10.1016/j.biocon.2007.02.004

Caffier, V., Lasserre-Zuber, P., Giraud, M., Lascostes, M., Stievenard, R., Lemarquand, A., ... Le Cam, B. (2014). Erosion of quantitative host resistance in the apple × *Venturia inaequalis* pathosystem. *Infection, Genetics and Evolution*, *27*, 481-489. doi: 10.1016/j.meegid.2014.02.003

Christe, C., Stölting, K. N., Paris, M., Fraïsse, C., Bierne, N., & Lexer, C. (2017). Adaptive evolution and segregating load contribute to the genomic landscape of divergence in two tree species connected by episodic gene flow. *Molecular Ecology*, *26*(1), 59-76. doi: 10.1111/mec.13765

Coffman, A. J., Hsieh, P. H., Gravel, S., & Gutenkunst, R. N. (2015). Computationally efficient composite likelihood statistics for demographic inference. *Molecular Biology and Evolution*, *33*(2), 591-593. doi: 10.1093/molbev/msv255

Cornille, A., Feurtey, A., Gélin, U., Ropars, J., Misvanderbrugge, K., Gladieux, P., & Giraud, T. (2015). Anthropogenic and natural drivers of gene flow in a temperate wild fruit tree: a basis for conservation and breeding programs in apples. *Evolutionary Applications*, *8*(4), 373-384. doi: 10.1111/eva.12250

Cornille, A., Giraud, T., Smulders, M. J., Roldán-Ruiz, I., & Gladieux, P. (2014). The domestication and evolutionary ecology of apples. *Trends in Genetics*, *30*(2), 57-65. doi: 10.1016/j.tig.2013.10.002

Cornille, A., Gladieux, P., & Giraud, T. (2013). Crop-to-wild gene flow and spatial genetic structure in the closest wild relatives of the cultivated apple. *Evolutionary Applications*, *6*(5), 737-748. doi: 10.1111/eva.12059

Cornille, A., Gladieux, P., Smulders, M. J., Roldan-Ruiz, I., Laurens, F., Le Cam, B., ... Gabrielyan, I. (2012). New insight into the history of domesticated apple: secondary contribution of the European wild apple to the genome of cultivated varieties. *PLoS Genetics*, *8*(5), e1002703. doi: 10.1371/journal.pgen.1002703

Croll, D., Laine, A. L. (2016). What the population genetic structures of host and pathogen tell us about disease evolution. *New Phytologist, 212*(3):537-539. doi: 10.1111/nph.14203.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... McVean, G. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158. doi :10.1093/bioinformatics/btr330

Depotter, J. R. L., Seidl, M. F., Wood, T. A., Thomma, B. P. H. J. (2016). Interspecific hybridization impacts host range and pathogenicity of filamentous microbes. *Current Opinion in Microbiology*, *32*:7–13

Dray, S., & Dufour, A. B. (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4), 1-20. doi: 10.18637/jss.v022.i04

Earl, D. A. & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4* (2), 359-361. doi: 10.1007/s12686-011-9548-7

Eastwood, A., Lazkov, G., & Newton, A. (2009). *The red list of trees of Central Asia*. Fauna and Flora International.

Ellstrand, N. C., Meirmans, P., Rong, J., Bartsch, D., Ghosh, A., De Jong, T. J., ... Strasburg, J. L. (2013). Introgression of crop alleles into wild or weedy populations. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 325-345. doi: 10.1146/annurev-ecolsys-110512-135840

Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14, 2611-2620. doi:10.1111/j.1365-294X.2005.02553.x

Falush, D., Stephens, M. & Pritchard, J. K. (2003). Inference of population structure: Extensions to linked loci and correlated allele frequencies. *Genetics*, 164(8), 1567-1587.

Feurtey, A., Cornille, A., Shykoff, J. A., Snirc, A., & Giraud, T. (2017). Crop-to-wild gene flow and its fitness consequences for a wild fruit tree: Towards a comprehensive conservation strategy of the wild apple in Europe. *Evolutionary Applications*, *10*(2), 180-188. doi: 10.1111/eva.12441

Forsline, P. L., Aldwinckle, H. S., Dickson, E. E., Luby, J. J., & Hokanson, S. C. (2003). Collection, maintenance, characterization, and utilization of wild apples of Central Asia. *Horticultural Reviews*, *29*, 1-62.

Gladieux, P., Caffier, V., Devaux, M., Le Cam, B. (2010a). Host-specific differentiation among populations of *Venturia inaequalis* causing scab on apple, pyracantha and loquat. *Fungal Genetics and Biology, 47*(6), 511-521. doi: 10.1016/j.fgb.2009.12.007

Gladieux, P., Zhang, X. G., Afoufa-Bastien, D., Sanhueza, R. M. V., Sbaghi, M., & Le Cam, B. (2008). On the origin and spread of the scab disease of apple: out of central Asia. *PLoS One*, *3*(1), e1455. doi: 10.1371/journal.pone.0001455

Gladieux, P., Zhang, X. G., RóLdan-Ruiz, I., Caffier, V., Leroy, T., Devaux, M., ... Le Cam, B. (2010b). Evolution of the population structure of *Venturia inaequalis,* the apple scab fungus,

associated with the domestication of its host. *Molecular Ecology*, *19*(4), 658-674. doi :10.1111/j.1365-294X.2009.04498.x

Gompert, Z., & Alex Buerkle, C. (2010). INTROGRESS: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources*, *10*(2), 378-384. doi: 10.1111/j.1755-0998.2009.02733.x

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, *5*(10), e1000695. doi: 10.1371/journal.pgen.1000695

Harris, S. A., Robinson, J. P., & Juniper, B. E. (2002). Genetic clues to the origin of the apple. *Trends in Genetics*, *18*(8), 426-430. doi: 10.1016/S0168-9525(02)02689-6

Jakobsson, M. & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23(14), 1801-1806.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403-1405. doi: 10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics 11*: 94.

Le Cam, B., Parisi, L. & Arene, L. (2002). Evidence of two *formae speciales* in *Venturia inaequalis*, responsible for apple and pyracantha scab. *Phytopathology*, *92*, 314-320.

Le Cam, B., Sargent, D., Gouzy, J., Amselem, J., Bellanger, M. N., Bouchez, O., ... Lemaire, C. (2019). Population genome sequencing of the scab fungal species *Venturia inaequalis, Venturia pirina, Venturia aucupariae* and *Venturia asperata*. *G3: Genes, Genomes, Genetics*, *9*(8), 2405-2414. doi: 10.1534/g3.119.400047

Lê Van, A., Gladieux, P., Lemaire, C., Cornille, A., Giraud, T., Durel, C. E., ... Le Cam, B. (2012). Evolution of pathogenicity traits in the apple scab fungal pathogen in response to the domestication of its host. *Evolutionary Applications*, *5*(7), 694-704. doi: 10.1111/j.1752-4571.2012.00246.x

Lemaire, C., De Gracia, M., Leroy, T., Michalecka, M., Lindhard-Pedersen, H., Guerin, F., ... Le Cam, B. (2016). Emergence of new virulent populations of apple scab from nonagricultural disease reservoirs. *New Phytologist*, *209*(3), 1220-1229. doi: 10.1111/nph.13658

Leroy, T., Caffier, V., Celton, J. M., Anger, N., Durel, C. E., Lemaire, C., & Le Cam, B. (2016). When virulence originates from nonagricultural hosts: evolutionary and epidemiological consequences of introgressions following secondary contacts in *Venturia inaequalis*. *New Phytologist*, *210*(4), 1443-1452. doi: 10.1111/nph.13873

Leroy, T., Lemaire, C., Dunemann, F., & Le Cam, B. (2013). The genetic structure of a *Venturia inaequalis* population in a heterogeneous host population composed of different *Malus* species. *BMC Evolutionary Biology*, *13*(1), 64. doi: 10.1186/1471-2148-13-64

MacHardy, W. E. (1996). *Apple scab: biology, epidemiology, and management*. American Phytopathological Society (APS Press).

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research 27*, 209-20.

Nikiforova, S. V., Cavalieri, D., Velasco, R., & Goremykin, V. (2013). Phylogenetic analysis of 47 chloroplast genomes clarifies the contribution of wild species to the domesticated apple maternal line. *Molecular Biology and Evolution*, *30*(8), 1751-1760. doi: 10.1093/molbev/mst092

Omasheva, M. Y., Flachowsky, H., Ryabushkina, N. A., Pozharskiy, A. S., Galiakparov, N. N., & Hanke, M. V. (2017). To what extent do wild apples in Kazakhstan retain their genetic integrity? *Tree Genetics & Genomes*, *13*(3), 52. doi: 10.1007/s11295-017-1134-z

Paradis, E. & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R, *Bioinformatics*, 35(3), 526–528. https://doi.org/10.1093/bioinformatics/bty633

Raffaele, S., Kamoun, S. (2012). Genome evolution in filamentous plant pathogens: why bigger can be better. *Nature Reviews Microbiology 10*, 417-430. doi: 10.1038/nrmicro2790

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. URL: http://www.R-project.org/

Reim, S., Proft, A., Heinz, S., Lochschmidt, F., Höfer, M., Tröber, U., & Wolf, H. (2017). Pollen movement in a *Malus sylvestris* population and conclusions for conservation measures. *Plant Genetic Resources, 15*(1), 12-20. doi:10.1017/S1479262115000301

Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, *27*(1), 83-109. doi: 10.1146/annurev.ecolsys.27.1.83

Rousset, F. (2008). Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, *8*, 103-106. doi:10.1111/j.1471-8286.2007.01931.x

Saleh, D., Milazzo, J., Adreit, H., Fournier, E., & Tharreau, D. (2014). South-East Asia is the center of origin, diversity and dispersion of the rice blast fungus, *Magnaporthe oryzae. New Phytologist*, *201*(4), 1440-1456. doi: 10.1111/nph.12627

Spotts, R. A., & Ferree, D.C. (1979). Photosynthesis, transpiration, and water potential of apple leaves infected by *Venturia inaequalis*. *Phytopathology, 69*(7), 717-719. doi: 10.1094/Phyto-69-717

Stukenbrock, E. H., Christiansen, F. B., Hansen, T. T., Dutheil, J. Y., & Schierup, M. H. (2012). Fusion of two divergent fungal individuals led to the recent emergence of a unique widespread pathogen species. *Proceedings of the National Academy of Sciences*, *109*(27), 10954-10959. doi: 10.1073/pnas.1201403109

Stukenbrock, E. H. (2016). Hybridization speeds up the emergence and evolution of a new pathogen species. *Nature Genetics*, *48*(2), 113. doi: 10.1038/ng.3494

Tine, M., Kuhl, H., Gagnaire, P. A., Louro, B., Desmarais, E., Martins, R. S., ... Dieterich, R. (2014). European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature Communications*, *5*, 5770. doi: 10.1038/ncomms6770

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., ... Rieseberg, L. H. (2016). Hybridization and extinction. *Evolutionary Applications*, *9*(7), 892-908. doi: 10.1111/eva.12367

Vanden Broeck, A. V., Villar, M., Van Bockstaele, E., & Van Slycken, J. (2005). Natural hybridization between cultivated poplars and their wild relatives: evidence and consequences for native poplar populations. *Annals of Forest Science*, *62*(7), 601-613. doi: 10.1051/forest:2005072

Vavilov NI. (1931). The role of Central Asia in the origin of cultivated plants. *Bulletin of Applied Botany Genetics and Plant Breeding* 26(3), 3-44.

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., ... Salvi, S. (2010). The genome of the domesticated apple (*Malus×domestica* Borkh.). *Nature Genetics*, *42*(10), 833. doi: 10.1038/ng.654

Wang, H., Vieira, F.G., Crawford, J. E., Chu, C. & Nielsen, R. (2017). Asian wild rice is a hybrid swarm with extensive gene flow and feralization from domesticated rice. *Genome Research*, 27(6),1029-1038. doi: 10.1101/gr.204800.116.

Wang, Y., Arenas, C. D., Stoebel, D. M., Flynn, K., Knapp, E., Dillon, M. M., ... Cooper, T. F. (2016). Benefit of transferred mutations is better predicted by the fitness of recipients than by their ecological or genetic relatedness. *Proceedings of the National Academy of Sciences*, *113*(18), 5047-5052. doi: 10.1073/pnas.1524988113

Weir, B. S., & Cockerham, C. C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution 38*: 1358-70.

Wolf, D. E., Takebayashi, N., & Rieseberg, L. H. (2001). Predicting the risk of extinction through hybridization. *Conservation Biology*, *15*(4), 1039-1053. doi: 10.1046/j.1523-1739.2001.0150041039.x

Zeder, M. A., Emshwiller, E., Smith, B. D., & Bradley, D. G. (2006). Documenting domestication: the intersection of genetics and archaeology. *Trends in Genetics*, *22*(3), 139-155. doi: 10.1016/j.tig.2006.01.007

Data accessibility

All *Venturia* and apple genotype data, ancestry, and pathogenicity results were archived and available via the following link https://doi.org/10.5281/zenodo.3966994. All *Venturia* genome

sequencing data used in this study are accessible (assembly fasta files) in GenBank under the BioProject number PRJNA407103.

Author contributions

AF, EG, TG and CL wrote the manuscript with input from BLC, VC, JS and LD. Sampling in Kazakhstan was performed by CL and BLC. Apple microsatellite genotyping was performed by AF. Apple scab SNP design and genotyping was performed by MDG, MS and CL. SNP mapping and calling was performed by LD and JS. EG, MNB, PE and VC managed *in vitro* cultures of *Venturia inaequalis* and pathology assays. AF, EG and CL performed population genetics and statistical analyses. CL, BLC and TG conceived the project and obtained financial support.

Tables

Table 1: Comparison of four alternative demographic models of divergence between the reference wild and agricultural-type populations of *Venturia inaequalis,* performed with *∂a∂i* on the SNP dataset.

Model	k	logL	AIC	Model Comparison	N_{Wild}	N_{Agri}	M_{WA}	M_{AW}	T_S	T_{AM}	T_{SC}
SI	3	-1197.8	2403.587	-	0.094	1.360	-	-	0.092	-	-
IM	5	-981.8	1975.638	SI***	0.139	0.898	1.259	1.131	0.460	-	-
AM	6	-981.8	1977.698	SI***, IM[NS]	0.139	0.892	1.270	1.140	0.460	1.09×10^{-9}	-
SC	**6**	**-928.6**	**1871.313**	**SI***, IM***, AM[+++]**	**0.163**	**0.991**	**1.355**	**4.389**	**0.227**	**-**	**0.006**

SI, strict isolation model; IM, isolation with migration model; AM, ancient migration model; SC, secondary contact model; k, parameter numbers; logL, log-likelihood; AIC, Akaike information criterion; N_{Wild}, effective size of the wild-type lineage; N_{Agri}, effective size of the agricultural-type lineage; M_{WA}, migration rates from the wild to the agricultural-type lineage; M_{AW}, migration rates from the agricultural-type to the wild-type lineage; T_S, time of split; T_{AM}, time since ancient migration; T_{SC}, time since secondary contact.

The best supported model (SC) is indicated in bold. Model comparisons within and between classes of models are shown. Nested models were compared in likelihood ratio tests, with superscripts indicating significance levels (abbreviated ***P<0.001; **P <0.01; *P<0.05, NS non-significant). Non-nested models were compared on the basis of AIC, with the relative likelihood of each model versus the best model L(Mi|Mbest)=exp((AICmin-AICi)/2) (abbreviated +++L(Mi|Mbest)<0.001; ++L(Mi|Mbest) <0.01; +L(Mi|Mbest)<0.05, L(Mi|Mbest)> 0.05: AIC difference shown).

Figure legends

Figure 1: Summary of previous studies and questions addressed in this study concerning *Malus* apple trees (A) and their fungal pathogen *Venturia inaequalis* (B). **A.** The divergence and gene flow between the domesticated apple tree *M. domestica*, its initial progenitor *M. sieversii* from Asia and its secondary contributor the European crabapple tree *M. sylvestri*s (Cornille et al., 2012) are shown at the top. The distribution areas of the two wild apple trees are shown at the bottom, indicating (1) the initial domestication in Asia (Cornille et al., 2012), migration along the Silk Route (solid yellow line) and (2) secondary gene flow between *M. domestica* and *M. sylvestris* in Europe (Cornille et al., 2013; Feurtey et al., 2017), followed by the spread of modern cultivars worldwide (dashed yellow arrows). (3) A couple of previous studies have investigated the consequences of secondary contact between the domesticated apple and its Asian progenitor in terms of gene flow from *M. domestica* to *M. sieversii* (Cornille et al., 2012; Omasheva et al., 2017). Here, we consider the consequences of planting orchards in forests. **B.** The divergence between populations of the fungal pathogen *V. inaequalis* is shown at the top, and the distribution area of these populations is shown at the bottom, with colored trees symbolizing apple tree species and colored circles the fungal populations. (1) Gladieux et al. (2010b) demonstrated the existence of differentiated populations in Europe on *M. silvestris* and *M. domestica* (1A), in anthropic Central Asian areas on *M. sieversii* and *M. domestica* (1B) and in the wild Central Asian forests on *M. sieversii* (1C). (2) Lemaire et al. (2016) showed that the differentiated European *V. inaequalis* population on cultivated apples carrying Rvi6 resistance originated from *M. floribunda*. (3) Leroy et al. (2016) investigated the existence of gene flow between populations infecting *Riv6* and non-*Riv6* apple trees in Europe. (4) In this study, we investigated the occurrence of gene flow between Central Asian pathogen populations (i.e., the population in anthropic Central Asian areas on *M. sieversii* and *M. domestica* and the population in the wild Central Asian forests on *M. sieversii*).

Figure 2: Genetic diversities and ancestry assignments for apple trees and *Venturia inaequalis.* (A) Principal component analysis on the genotypes of 240 apple trees (*Malus spp.*) for 28 microsatellite markers, with their ancestry assignment (P_{dom}) to the cultivated apple tree *M. domestica* indicated by color. The distribution of the P_{dom} index is shown in the inset at the top

left. (B) Principal component analysis on the genotypes of 255 *V. inaequalis* strains for 181 SNPs, with their ancestry assignment (P_{agr}) to the agricultural type of *V. inaequalis* indicated by color. The distribution of the P_{agr} index is shown in the inset at the top left.

Figure 3: Distribution of the samples of apple trees and their fungal pathogen *Venturia inaequalis* at nine sites in the Tian Shan Mountains near Almaty (Kazakhstan). The black dot on the map in the upper left corner indicates the global sampling location. Proportions of wild (blue), domesticated/agricultural-type (yellow) and hybrid (red) genotypes are shown for each sampling site, for *Malus* trees (upper row of pie charts) and for their fungal scab pathogen (lower row of pie charts). The pie charts below the figure correspond to sampling sites in the natural forest. The pie charts above the figure correspond to a two-site sampling location with one site in an orchard and the other in the natural forest immediately surrounding it. Codes of the sampling locations: kot: Koturbulac; tag and tar: Talghar; lac, gec and esi (esi_o for the orchard and esi_f for the surrounding forest): Esik; tuk and tum: Turgen.

Figure 4: The proportion of *Malus domestica* ancestry in the apple tree (*y* axis) plotted against the proportion of agricultural-type ancestry in the *Venturia inaequalis* strain (*x* axis). A) For each of the 249 host-pathogen pairs, two sampling sites are highlighted: the orchard (esi_o) in orange and the surrounding forest (esi_f) in blue. B) For each of the 189 host-pathogen pairs for which the genotype at the *V_081690_319* locus was available, the *V. inaequalis* strains carrying the tyrosine (A) allele are shown in gray, and those carrying the stop codon (G) allele are shown in red.
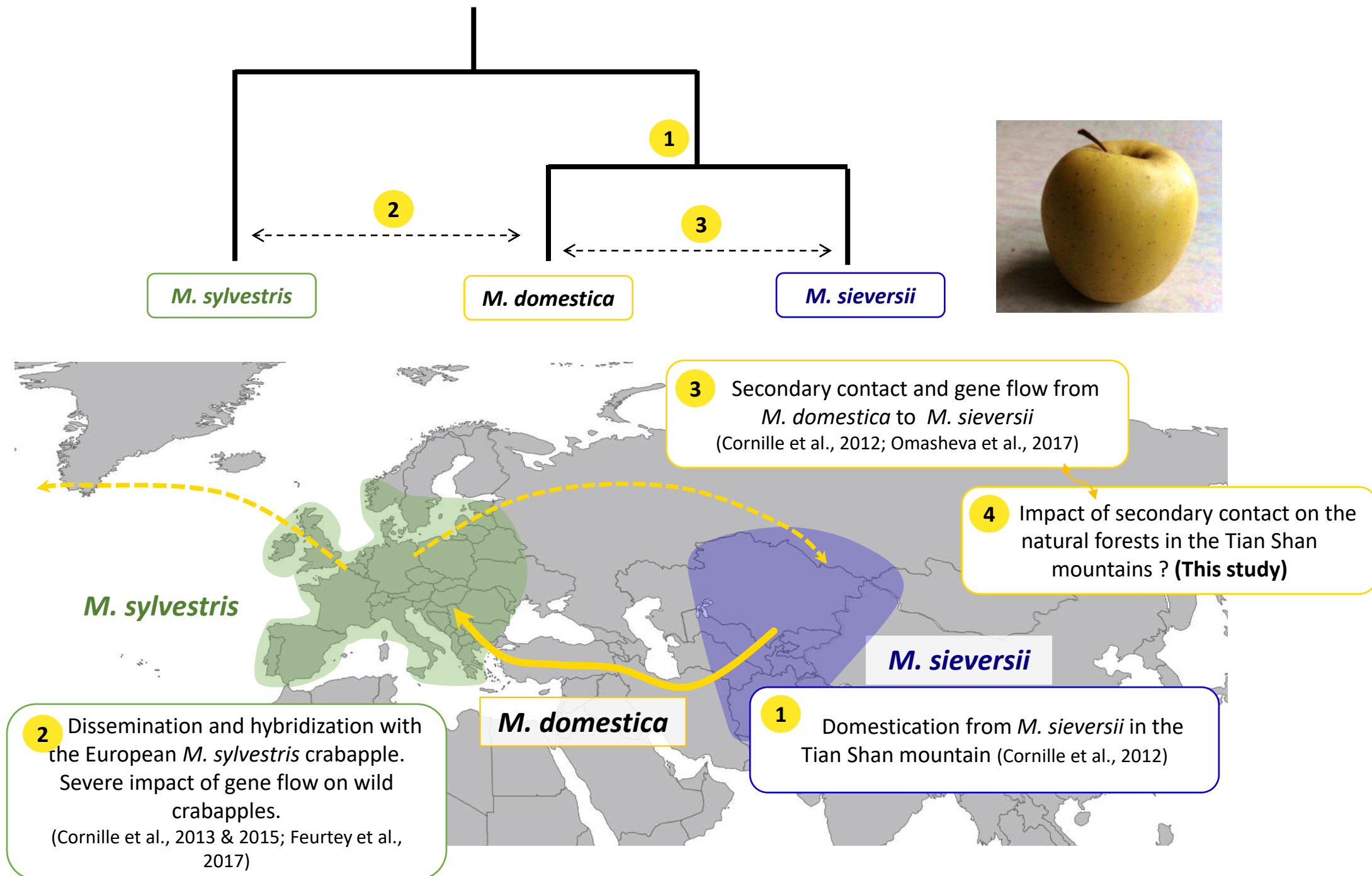
Figure 5: Results of pathogenicity experiments in controlled conditions involving the artificial inoculation of apple tree leaves (*Malus sieversii* and *M. domestica*) with the fungus *Venturia inaequalis* (57 fungal strains; mean of three replicates per strain, one genotype per apple species). Boxplots of the percentage of the leaf area displaying scabbing for wild-type, agricultural-type and hybrid *V. inaequalis* strains on A) *M. domestica* at 21 days post inoculation (dpi) and B) *M. sieversii* at 19 dpi. The box represents the lower and upper quartiles. The thick horizontal line

represents the median. The whiskers represent the highest and lowest values falling within 1.5 times the interquartile range. The points represent the mean values of the percentage of the leaf area presenting disease, for each strain, across the three replicates. Different letters indicate significant differences between populations ($P<0.05$; Wilcoxon's rank sum tests).

Figure 6: Images illustrating the percentage scabbed leaf area for the most virulent wild-type, agricultural-type and hybrid *Venturia inaequalis* strains on A) *Malus domestica* at 21 days post inoculation (dpi) and B) *M. sieversii* at 19 dpi.

Figure 7: Demographic inference concerning the divergence between wild-type and agricultural-type *Venturia inaequalis* strains in Kazakhstan. **A.** The four contrasting scenarios of divergence compared involved strict isolation (SI), isolation with continuous migration (IM), isolation with initial migration (AM) and secondary contact (SC), with effective populations sizes of the ancestral population (N_{ref}) and of the two daughter populations (N_w and N_A), the rate of gene flow in the two directions (M_w and M_A) and the time since the cessation of gene flow (T_{AM}) or secondary contact (T_{SC}) indicated. **B.** The joint allele-frequency spectra (jAFS) for the wild and agricultural-type fungal populations, showing the derived allele counts. The jAFS on the left corresponds to observed data and on the right to the maximum-likelihood model of secondary contact. Each entry in the joint AFS is colored according to the number of SNPs it contains, according to the scale shown. **C.** The secondary contact (SC) model, which was identified as the best-supported model, with parameter estimates.

A

M. sylvestris

M. domestica

M. sieversii

3 Secondary contact and gene flow from *M. domestica* to *M. sieversii* (Cornille et al., 2012; Omasheva et al., 2017)

4 Impact of secondary contact on the natural forests in the Tian Shan mountains ? **(This study)**

M. sylvestris

M. sieversii

M. domestica

2 Dissemination and hybridization with the European *M. sylvestris* crabapple. Severe impact of gene flow on wild crabapples. (Cornille et al., 2013 & 2015; Feurtey et al., 2017)

1 Domestication from *M. sieversii* in the Tian Shan mountain (Cornille et al., 2012)

orchard surrounding forest

Host   N=30   N=22

Pathogen   N=30   N=30

Longitude

77.00   77.25   77.50

43.4

Latitude

43.2

Energeticheskiy

Pervomayskiy

Turgen'

Esik

Talghar

Almaty

Host   N=16   N=33   N=30   N=16   N=31   N=30   N=32

Pathogen   N=29   N=27   N=25   N=29   N=27   N=28   N=30

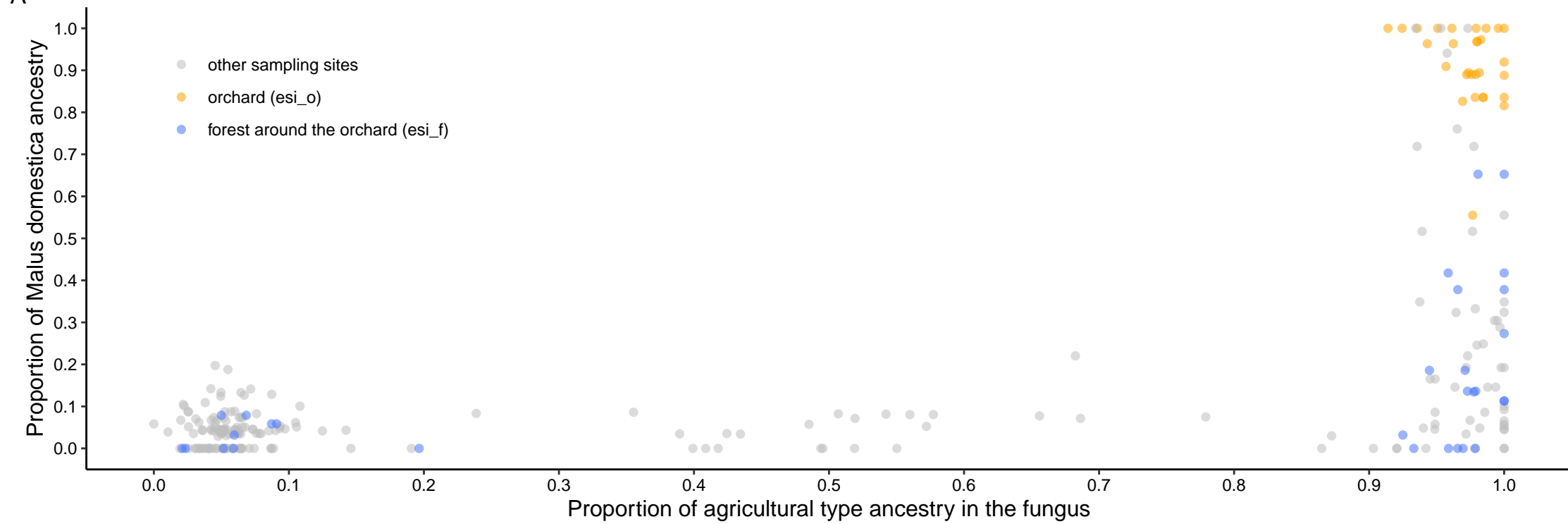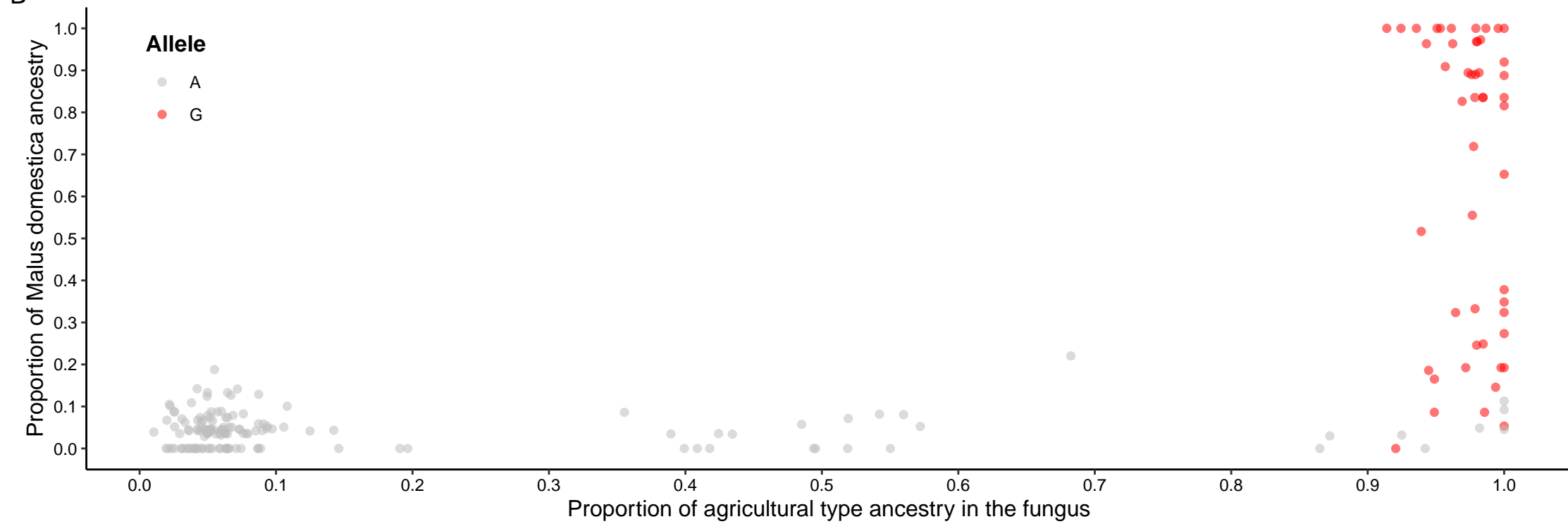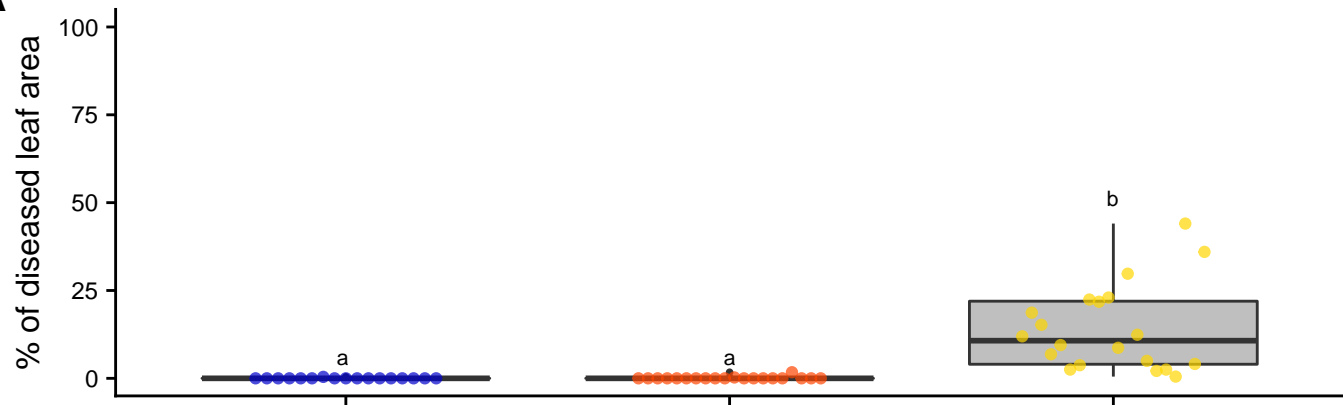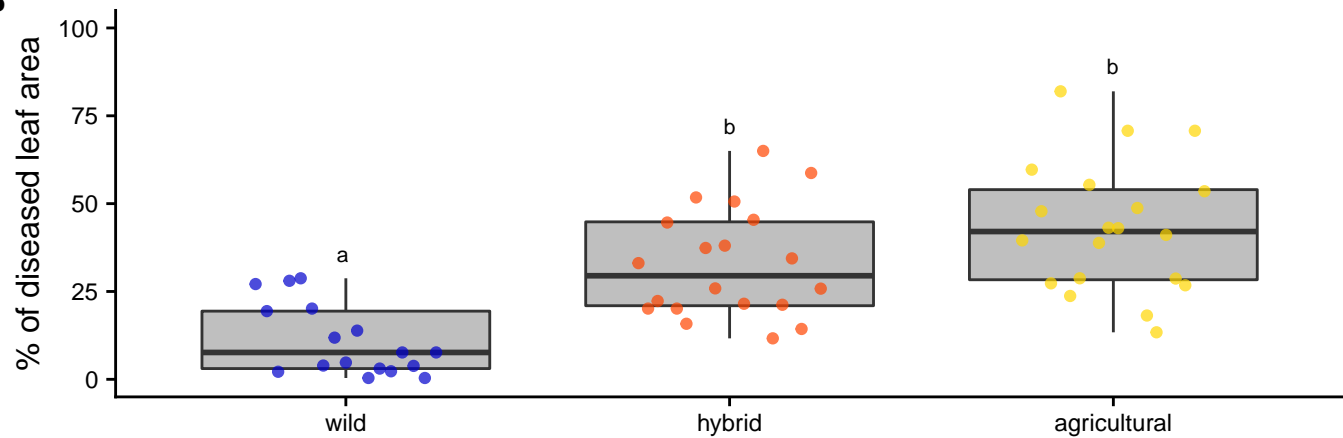kot   tag   tar   lac   gec   tum   tuk

A
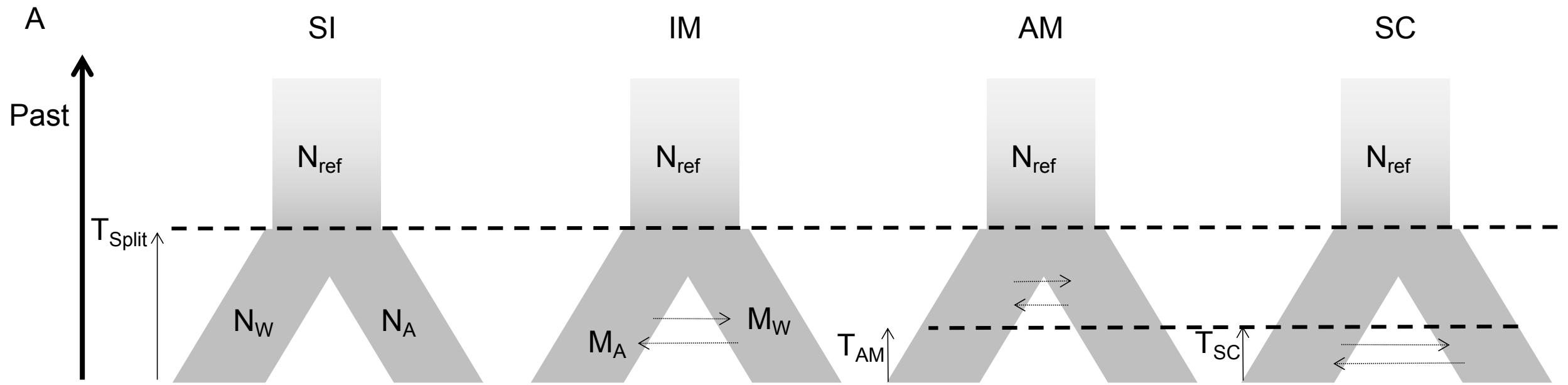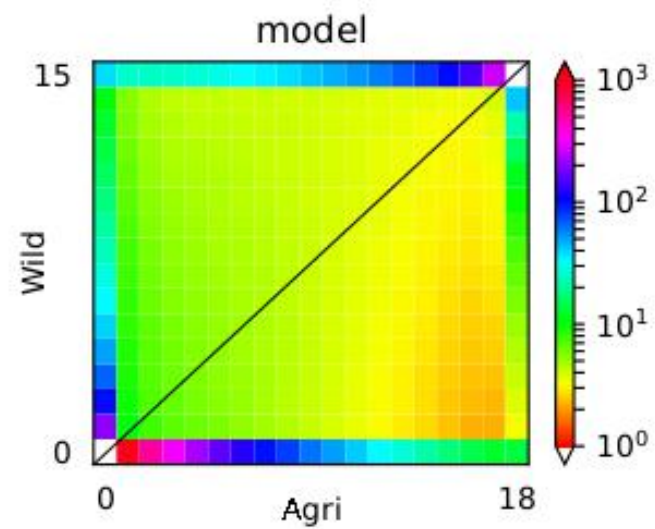
B
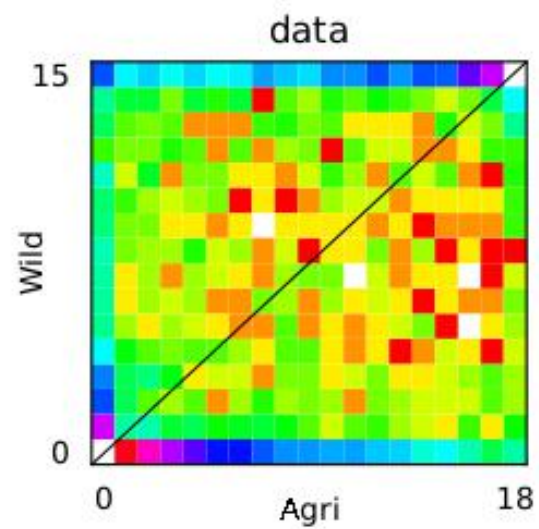
wild   hybrid   agricultural

A

SI

N_{ref}

T_{Split}

N_W N_A

IM

N_{ref}

M_A ← ─ ─ ─ → M_W

AM

N_{ref}

T_{AM}

SC

N_{ref}

T_{SC}

B

data

model

Wild

Agri

Wild

Agri

10^3
10^2
10^1
10^0

C   SC

N_{ref}

$T_{Split} \sim$ 4,500 yrs

$T_{SC} \sim$ 130 yrs

N_W ← ─ ─ ─ → N_A