Interplay between demography and selection: a forward model illuminates the temporal genetic signatures of rapid adaptation in plant pathogens

Méline Saubin[1], Solenn Stoeckel[2], Aurélien Tellier[3] and Fabien Halkett[1]

[1] Université de Lorraine, INRAE, IAM, F-54000 Nancy, France

[2] INRAE, Agrocampus Ouest, Université de Rennes, IGEPP, F-35653 Le Rheu, France

[3] Professorship for Population Genetics, Technical University of Munich, Freising, Germany

1 Abstract

The demography and genetic structure of a population are closely linked. The study of this interplay is crucial, especially for organisms with frequent demographic fluctuations such as pathogen species responsible for emerging diseases. The joint study of demographic variations and genetic evolution of a pathogen population can be achieved by coupling epidemiology and population genetics. This framework is particularly suited to the study of genetic signatures of strong and rapid selective events, such as pathogen adaptation to plant genetic resistances. We have limited knowledge of how these events shape the genetic evolution of pathogens and what would be the typical genetic signatures of strong demographic and selective events on the pathogen population. In this study, we model pathogen population dynamics and genetic evolution at neutral loci to draw a typology of scenarios of eco-evolutionary dynamics. Using a clustering method dedicated to time-series variations, we identify three main scenarios: 1) small variations in the pathogen population size and small changes in genetic structure, 2) a strong founder event on the resistant host that in turn leads to a genetic structure on the susceptible host, and 3) evolutionary rescue that results in a strong founder event on the resistant host, preceded by a bottleneck on the susceptible host. Beyond the selective event itself, the demographic trajectory imposes specific changes in the genetic structure of the pathogen population. Most of these genetic changes are transient, with a signature of resistance overcoming that vanishes in a few years only. Considering time series is therefore of utmost importance for accurate tracking of the evolutionary trajectories of pathogen populations.

2 Introduction

Disease outbreaks caused by pathogens impact both natural and human-managed ecosystems (*e.g.* agro-ecosystems) (Anderson et al., 2004; Tobin, 2015; Savary et al., 2019). The number of emerging diseases is increasing exponentially and unprecedentedly during the last decades (Fisher et al., 2012). Understanding pathogen evolution is essential to comprehend how they affect ecosystems (Fischhoff et al., 2020) and to establish relevant disease management programs (Bonneaud and Longdon, 2020). Yet, this task is particularly challenging due to the rapid adaptation of pathogen populations (McDonald and Linde, 2002; Saubin et al., 2022), and the high stochasticity of pathogen evolutionary trajectories (Parsons et al., 2018). In addition, pathogens depend on host abundance over time and space (Stukenbrock and McDonald, 2008) with intricate life cycles and reproductive modes (Agrios, 2005) that combine to shape distinctive patterns of genetic structure (*e.g.* Rouger et al., 2016).

Pathogen populations commonly endure large demographic variations, including repeated bottlenecks and founder events. Most pathogen species display seasonal growing patterns with marked differences between epidemic and dormant phases (Agrios, 2005). The latter phase corresponds to a surviving period to face harsh environmental conditions or host absence, during which the pathogen population commonly collapses. These extreme variations in population sizes make pathogen populations particularly vulnerable to local extinctions (Laine and Barrès, 2013). Moreover, pathogen populations are subjected to strong selection, in particular during host adaptation events (Gladieux et al., 2011). Selection affects the pathogen's demography, increasing mortality of the non-adapted part of pathogen populations and leading to a burst of adapted pathogens (McDonald and Linde, 2002). Selection pressures are all the more strengthened in human-managed ecosystems such as in agrosystems (Stukenbrock and McDonald, 2008; Zhan et al., 2015).

In some extreme cases, mechanisms can lead to the pathogen survival even if the decline of the population was approaching extinction. Three forms of such population 'rescue' are commonly described (Carlson et al., 2014): 1) the demographic rescue, when the population survival is only attributed to the increase in population size due to immigration of new individuals (Brown and Kodric-Brown, 1977); 2) the genetic rescue, when the survival of the population is attributed to the novel genetic variation brought by the immigration of new individuals, in a small population suffering genetic load (Thrall, 1998); 3) the evolutionary rescue, when adaptive evolution lead to population recovery from negative growth initiated by environmental change (Gonzalez et al., 2013; Bell, 2017). These forms of rescue closely link demography and selection. The probability of fixation of alleles is strongly impacted by changes in population size (Otto and Whitlock, 1997), with the effect of genetic drift accentuated by a reduction in population size. It is therefore all the more important to focus on the interplay between selection and genetic drift in a population with fluctuating size (Gokhale et al., 2013; Živković et al., 2019) to weigh up the balance between deterministic and stochastic processes that drive the evolutionary trajectories of pathogen populations.

The joint study of demographic variations and genetic evolution of the pathogen population can be achieved using population genetics (Day and Proulx, 2004; Day and Gandon, 2007; Saubin et al., 2022).

Coupling epidemiology and population genetics provides insights into the mechanisms underpinning pathogen evolution or host-parasite coevolution and acting at both short (ecological) and long (evolutionary) time scales (Milgroom and Peever, 2003; Archie et al., 2009). This framework is thus dedicated to the study of genetic signatures of strong and rapid selective events (Saubin et al., 2022). The interplay between demography and selection is captured by monitoring both selected and neutral loci. It allows in particular detailed analyses of transition periods (Day and Proulx, 2004; Day and Gandon, 2007; Bolker et al., 2010), through variables like the pathogen population size, affecting both the disease incidence in epidemiology and the impact of genetic drift in population genetics (McDonald, 2004; Živković et al., 2019).

Pathogen adaptation to plant genetic resistances is a typical example of rapid adaptation linked to variations in population size. In agrosystems, the breeding of resistant hosts based on the gene-for-gene interaction is crucial to mitigate the severity of epidemics of plant pathogens. Gene-for-gene interaction attracts a great deal of breeding efforts because in most cases it confers complete host immunity. In the simplest form, plant genetic resistance thus prevents infection from a class of pathogen genotypes called *avirulent* (or non-infective). In agrosystems, the deployment of pure resistant plants exerts a strong selection pressure on pathogen populations that favours any variant that can infect the resistant host (Zhan et al., 2015). This class of pathogen genotypes is called *virulent* (or infective). The infective phenotype is usually determined by changes at a single locus, the avirulence locus. Therefore, the issue of qualitative resistance deployments is often a resistance overcoming, *i.e.* the invasion of resistant hosts by virulent individuals (Johnson, 1984; Pink and Puddephat, 1999; Brown and Tellier, 2011; Burdon et al., 2016). The resistant variety is usually deployed widely across the landscape, and this intense and unidirectional selection pressure can result in severe epidemics, which impose rapid and drastic demographic changes for the pathogen population (Burdon et al., 2016; Persoons et al., 2017; Saubin et al., 2021). We have limited knowledge of how such events shape the genetic evolution of pathogens and what would be the typical genetic signatures of such strong demographic and selective events on the pathogen population. This theoretical knowledge would help to decipher the causes of current and past events of resistance overcoming and, ultimately, to predict when and how often such events occur.

In this study, we aim to understand the interplay between demography and genetic evolution at the avirulence locus and neutral markers during resistance overcoming, and to assess the impact of such rapid adaptive events on the genetic structure of the pathogen population. We model pathogen population dynamics and genetic evolution to investigate the impact of the pathogen life cycle on these selective and demographic dynamics using a demogenetic approach. We perform simulations under several realistic scenarios of resistance overcoming. Using a clustering method dedicated to time-series variations, we show that we are able to identify several main scenarios of eco-evolutionary dynamics. We conclude by commenting on the typology of these dynamics and the potential to use our simulation framework to analyse real datasets.

3 Materials and methods

3.1 Model description

3.1.1 Model overview

To study the effects of resistance overcoming on genetic indices in diploid pathogen populations along generations, we develop an individual-based, forward-time demogenetic model. It couples population dynamics and population genetics to follow through time the evolutionary trajectory of different genotypes at the selected avirulence locus and at neutral genetic markers scattered in the genome. The model is similar to the model described in Saubin et al. (2021), with the addition in this study of the evolution of neutral markers through time. We consider a life cycle commonly found in temperate pathogen species, which alternate rounds of clonal reproduction with an annual event of sexual reproduction (Agrios, 2005). This model is designed to simulate two distinct life cycles: with or without host alternation for the sexual reproduction. Without host alternation, the model simulates the evolution over time of a population of pathogens on two static host compartments of the same host species: a resistant compartment (R) and a susceptible compartment (S). Fixed carrying capacities of pathogens, K_R and K_S, are respectively assigned to R and S compartments and represent the maximum amount of pathogens that each compartment can sustain. With host alternation, another alternate host species (a third compartment, A) is added to the simulations. Pathogens annually gather on this additional compartment to achieve sexual reproduction. This static compartment is assumed to be larger than the two other compartments, with a fixed carrying capacity K_A. Hereafter, we refer to the three compartments as S (susceptible hosts), R (resistant hosts) and A (alternate hosts). We do not consider spatial substructure among host compartments. The way we model the life cycles is presented in Figure 1, and can be described as follow.

3.1.2 Reproduction events

Each discrete generation corresponds to a reproduction event, either sexual or clonal. Each year is composed of g non-overlapping generations, with one annual sexual reproduction event followed by a succession of $g-1$ clonal reproduction events. In our simulations, we consider a year composed of $g = 11$ generations, based on the case for the poplar rust pathogen with host alternation (*Melampsora larici-populina*, Hacquard et al., 2011) and the apple scab pathogen without host alternation (*Venturia inaequalis*, Le Cam, personal communication). At each reproduction event, successive generations do not overlap so that parents give way to offsprings and the new population is composed exclusively of the newly produced individuals.

Before each sexual reproduction event, a proportion τ of pathogen individuals is picked randomly to form the parental population. We fixed $\tau = 0.75$ to mimic pathogen life cycles displaying a moderate reduction in population size before and during sexual reproduction that usually takes place before or during the unfavourable season, such as winter in temperate climates.

For the sexual reproduction event itself, the population size is considered constant before and after the

reproduction event:

$$N_{n+1} = N_n \tag{1}$$

With N_{n+1} the population size at generation $n+1$ and N_n the population size at generation n. Sexual offspring genotypes are obtained by random association of gametes produced by the parental population.

For each clonal reproduction event on S, R, or A, the population growth has a logistic function, with the following relation:

$$N_{n+1} = N_n + (r-1) \times N_n \times (1 - \frac{N_n}{K}) \tag{2}$$

With r the growth rate of the pathogen population, K the carrying capacity of the compartment (K_S, K_R, or K_A for S, R, or A respectively). For all clonal reproduction events, offspring genotypes are drawn randomly with replacement from the parental population in which all parental pathogen individuals have the same uniform probability to reproduce.

3.1.3 Migration events

A regular two-way migration event takes place each generation before the reproduction event, between R and S. The number of migrants is determined by a fixed migration rate (m) multiplied by the number of individuals on the source compartment. In this effective migration scheme, migrants are always successfully established on the destination compartment, even if the number of individuals on this compartment reached its maximum carrying capacity. Thereby, this choice enables the immigration of new pathogens regardless of the size of the population, as it is observed in natural populations for plant pathogens.

For the life cycle with host alternation, the annual sexual reproduction event coincides with the obligate migration of the entire pathogen population to and from the alternate host. The first migration event takes place once per year after $g-2$ clonal reproduction events on R and S. For this migration event, an established proportion of individuals τ is picked randomly from R and S and migrates to A. R and S compartments are then emptied of any pathogen for the next two generations. After the two reproduction events (sexual and clonal) on A, the second migration event redistributes randomly all individuals from A to R and S, in proportion to the relative size of R and S compartments, leaving no living individuals on A.

3.1.4 Genetic evolution of the pathogen population

To simulate the gene-for-gene interaction between resistant and susceptible hosts and virulent and avirulent pathogens, we model allele frequency evolution at the avirulence locus assuming standing genetic variation (Barrett and Schluter, 2008). In diploids, pathogens are virulent only if the products of both alleles avoid detection by the host (Stukenbrock and McDonald, 2009; Thrall et al., 2016). In the following, we name *Avr* the dominant allele conferring avirulence and *avr* the recessive allele conferring virulence at the homozygous

state.

In addition to the selected avirulence locus, we simulate the evolution of 100 neutral genetic markers with a mutation rate of 10^{-3}. Each locus has four possible allelic states under a classical k-allele mutation model (Wright, 1949). Upon mutation, an allele changes into any of the three other allelic states with equal probability. At these 100 loci, we compute 19 classical population genetics indices each year before the sexual reproduction on both R and S compartments (Table 1).

These indices are chosen to 1) describe intra-population genetic and genotypic diversity, 2) measure overall linkage disequilibrium, and 3) assess genetic differentiation (F_{ST}). The differentiation is considered between populations on R and S at a given generation, and through time between a population on R or S at a given generation and the initial genotypic state. The variation in these indices captures the footprints of the different processes expected to occur during resistance overcoming: effects of reproductive modes, founder events followed by expansions, patterns of coancestry and admixture between R and S.

3.1.5 Implementation of model analyses

The model is implemented in Python (version 3.7, van Rossum, 1995) and Numpy (Harris et al., 2020). Each simulation starts with genotypes randomly drawn from the four possible alleles followed by a burn-in period of 11,000 generations under a constant population size of K_S individuals. In this way, we ensure that the pathogen population is at the drift-mutation equilibrium before the resistance overcoming. A proportion f_{avr} of virulent alleles is introduced (replacing avirulent Avr alleles) on S after the burn-in period as initial standing genetic variation. Homozygous avr/avr and heterozygous AVR/avr individuals are therefore initially present, based on the frequency f_{avr}. Simulations are run with a fixed total carrying capacity for the host population sizes of each host species, $K = K_A = K_R + K_S = 10{,}000$. The size of R is given by the fixed and constant proportion of resistant hosts $propR = \frac{K_R}{K}$ time K. The size of S is deduced from K and K_R.

We run a random design set of 30,000 simulations. Each combination of parameter values is drawn at random in defined distributions to investigate a wide range of genetic dynamics through resistance overcoming (Table 2). Simulations are run for 400 generations. During this period, nearly all replicates reach new steady states including invasion of R and loss of the avirulent Avr allele or extinction of the pathogen population. To focus on the genetic signatures of a resistance overcoming, only the simulations with at least 60% of virulent avr alleles at the end of the simulation are kept for this analysis. This threshold is chosen to ensure that the resistance overcoming and the invasion of the resistant compartment do occur during the simulated period.

To evaluate the effect of stochasticity on the dynamics, we perform also a regular simulation design composed of 100 replicates of 216 fixed combinations of parameter values (Table 2).

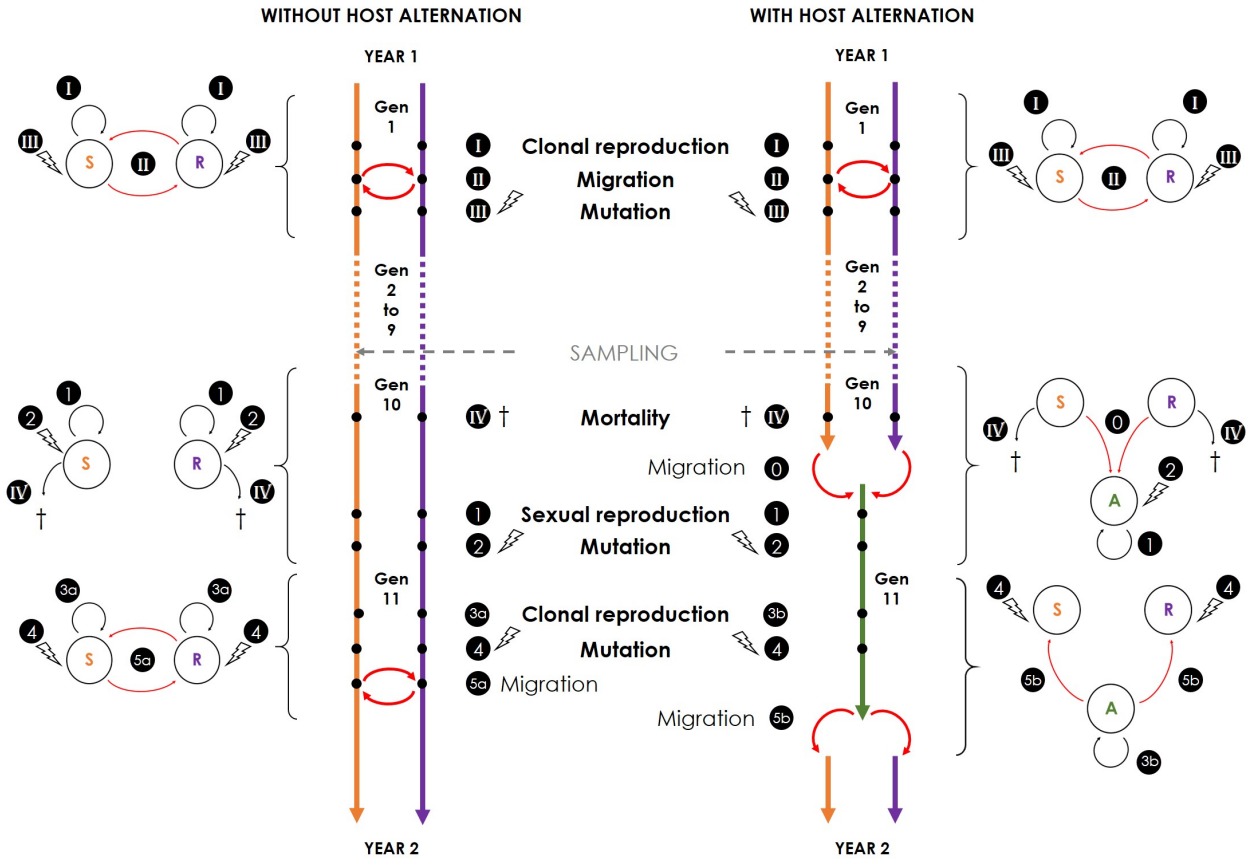


Figure 1: Modelling steps for each simulated year with the three host compartments S, R, and A. Each year is composed of 11 generations. During the clonal phase (generation 1 to 9), each generation is composed of three steps identical between both life cycles: (I) clonal reproduction (equation 2); (II) migration of a proportion m of each population between R and S; (III) mutation at all neutral markers with a mutation rate of 10^{-3}. The sampling takes place at the end of generation 9. At the end of the clonal phase, the pathogen overwinter as a dormant stage and is subjected to (IV) mortality of a proportion τ of each population. Then the sexual phase (generations 10 and 11) differs depending on the life cycle: (0) represents the migration of all individuals from R and S towards A; (1) sexual reproduction (equation 1); (2) and (4) mutation of all neutral markers with a mutation rate of 10^{-3}; (3a) and (3b) clonal reproduction (equation 2); (5a) migration of a proportion m of each population between R and S; (5b) migration of all individuals from A towards R and S.

Table 1: Description of population genetics indices computed in the model and calculated each year before the sexual reproduction on both R and S compartments.

Index	**Description**	**Reference**	**Clustering**	**Graphics**
$\beta\ of Pareto$	Genotypic diversity index	Arnaud-Haond et al., 2007	X	
$Simpson\ index$	Genetic diversity index			
$Shannon - Wiener\ index$	Genetic diversity index			
G/N	Proportion of unique genotypes	Dorken and Eckert, 2001		
G	Number of unique genotypes	Halkett et al., 2005		
$Mean\ H_E$	Mean expected heterozygosity overall loci	Nei, 1978		X
$Variance\ H_E$	Variance of the expected heterozygosity overall loci			X
$Mean\ H_O$	Mean observed heterozygosity overall loci			
$Variance\ H_O$	Variance of the observed heterozygosity overall loci			
$Mean\ LA$	Mean number of alleles by locus	Nielsen and Signorovitch, 2003		X
$VarianceLA$	Variance of the number of alleles by locus			
$Mean\ F_{IS}$	Mean inbreeding coefficient overall loci	Wright, 1931, 1949	X	X
$Variance\ F_{IS}$	Variance of the inbreeding coefficient overall loci	Stoeckel and Masson, 2014	X	
$\bar{r}_D$	Linkage disequilibrium index	Agapow and Burt, 2001	X	X
N_e	Effective population size	Wright, 1931		
F_k	Pollack's standardised variance of allele frequency	Pollak, 1983	X	
F_c	Nei and Tajima's standardised variance of allele frequency	Nei and Tajima, 1981		
$F_{ST}\ R - S$	Population differentiation between populations on R and S compartments	Wright, 1949, 1978		X
$Temporal\ F_{ST}$	Population differentiation between the initial population and the population at each recorded generation		X	X

Table 2: Input parameters and their range of variations for the two simulation designs run for 400 generations. For the regular simulation design, 100 replicates are run for each combination of parameters.

Variable	Description	Random design		Regular design
		Distribution	Interval	
f_{avr}	Initial frequency of the *avr* allele in the pathogen's population	Log-uniform	[log(0.0005), log(0.3)]	**4 levels**: 0.01; 0.03; 0.1; 0.32
m	Migration rate between R and S compartments	Log-uniform	[log(0.001), log(0.2)]	**3 levels**: 0.001; 0.01; 0.1
r	Growth rate of the pathogen	Uniform	[1.1, 2]	**3 levels** : 1.2; 1.6; 2.0
$propR$	Proportion of resistant hosts in the landscape	Uniform	]0.01, 0.99[	**3 levels**: 0.1; 0.5; 0.9
$Cycle$	Life cycle of the pathogen	Binomial	Without or with host alternation, probability 0.5	**2 levels**: Without host alternation; With host alternation

3.2 Comparisons of temporal dynamics

Analyses of changes in population genetics indices are performed using R (R, 2018). We present all results on S because it enables to compare the effect of life cycles, all else being equal, and to represent the genetic signatures expected without selection. When needed, we refer to the evolution of indices on R in supplementary data.

To analyse the dynamics of population genetics indices from the random simulation design, we performed clustering analyses regrouping simulations leading to similar dynamics on S. Distinct clustering analyses are performed for the two life cycles using the R package DTWCLUST (Sardá-Espinosa, 2019). To avoid redundancy in genetic information, clustering analyses are based on the temporal dynamics of the six indices that are orthogonal by their mathematical construction: $\beta\, of Pareto$ for genotypic diversity, $\bar{r}_D$ for overall linkage disequilibrium between loci, $Mean\ F_{IS}$ and $Variance\ F_{IS}$ for allele identity within individuals, F_k for the variation of allele identity between individuals of a same compartment, and $Temporal\ F_{ST}$ for the genetic differentiation. We perform a multivariate analysis by merging the normalised temporal dynamics of the six indices. We build a distance matrix between all simulations based on the dynamic time warping distance ('DTW_basic'), with the norm 'Euclidean distance' for the local cost matrix and the step pattern 'symmetric2'. We perform hierarchical clustering based on the distance matrix, with the agglomeration method 'Ward.D2'. We compare hierarchical clustering for a number of clusters between 2 and 8 included. For each life cycle, we select the number of clusters that maximises the Calinski-Harabasz index (Arbelaitz et al., 2013).

To understand how input parameters impact the clustering, we represent the distribution of input parameters for each cluster and both life cycles. We assess significant differences between distributions of parameter values using pairwise Kruskal-Wallis tests. To rank the impact of input parameters on the clustering, we perform a dominance analysis for each clustering and represent the estimated general dominance of each parameter on cluster assignments. Dominance analyses for both life cycles are performed with the R package DOMIR (Luchman, 2022).

For each cluster, we illustrate the evolution of population genetics indices through time. For the sake of clarity and to limit the amount of information, we focus in this article on the most informative four intra-population and two inter-population indices: $\bar{r}_D$, $Mean\, F_{IS}$, $Mean\, H_E$, $Mean\, LA$, $Temporal\, F_{ST}$, $F_{ST}\, R-S$ (Table 1). We display for each cluster and both life cycles the mean and standard error of these six population genetics indices calculated at each generation. These are not specific dynamics but mean changes over all simulations assigned to a given cluster.

To illustrate the realised dynamics of each population genetics index, we complement these results by displaying a representative realised simulation of each cluster (the medoid, *i.e.* the simulation that minimises the average distance to all other simulations in the same cluster). To highlight the effect of selection, we supplement each representation of the medoid dynamics with the corresponding null model dynamics, that is a simulation run under the same set of parameter values but without selection ($f_{avr} = 0$). The deviation

between medoid and null model dynamics highlights the specific signatures of selection. To interpret genetic changes with respect to resistance overcoming, we calculate the generation of resistance overcoming as the generation for which 1% of R is invaded by virulent individuals for the first time in a simulation. In addition, we define the generation of invasion as the first generation in which a virulent individual migrates to R.

Some simulations lead to evolutionary rescue, that is the invasion of R resulting in the recovery of the population collapse on S and preventing population extinction. To understand which input parameter combinations favour the occurrence of evolutionary rescue, we calculate for each simulation the growth rate threshold under which the population goes extinct if R is not accessible (population extinction in the 'null model'). We thus obtain the proportions of evolutionary rescue events in each cluster and perform a Fisher's exact test to assess the significance of the assignments to different clusters.

For the regular simulation design, we perform similar clustering analyses with and without host alternation to identify the proportion of simulations that result from the same combination of input parameters (replicates) but are assigned to different clusters.

4 Results

4.1 Influence of model parameters on the evolutionary dynamics

The clustering based on the random simulation design results in a partition into two clusters without host alternation and into three clusters with host alternation. For both life cycles, we name Cluster 1 the cluster that regroups the majority of the simulations (72 % and 78 % of the simulations without and with host alternation, respectively, Table 3). Cluster 1 displays small genetic changes through time for nearly all indices (Figure 2). Conversely, the other clusters display stronger changes in population genetics indices (Figure 2, Appendix Figures S1, S2).

Table 3: Cluster description for the clustering based on the random simulation design. This table contains the number of simulations assigned to each cluster, and the percentages of simulations leading to evolutionary rescue. Only simulations leading to resistance overcoming have been kept for the clustering analysis. The $P-value$ corresponds to the $P-value$ of the exact test of Fisher to determine if differences in proportions of evolutionary rescue events among clusters are significant.

Cycle	**Cluster**	**Nb of simulations**	**Evolutionary rescue (%)**	**Fisher** $P-value$
Without host alternation	1	4067	1.3	4.6×10^{-5}
	2	1575	3.0	
With host alternation	1	5858	3.7	$< 2.2 \times 10^{-16}$
	2	1581	10.6	
	3	81	97.5	

To understand the origin of the different types of dynamics, we jointly examine the influence of the input parameters on cluster delineation (Figure 3) and the difference in the distributions of parameter values for each cluster (Figure 4). Without host alternation, the most influential parameter for cluster delineation is the proportion of resistant hosts ($propR$), followed by the initial frequency of avirulent allele (f_{avr}), and,

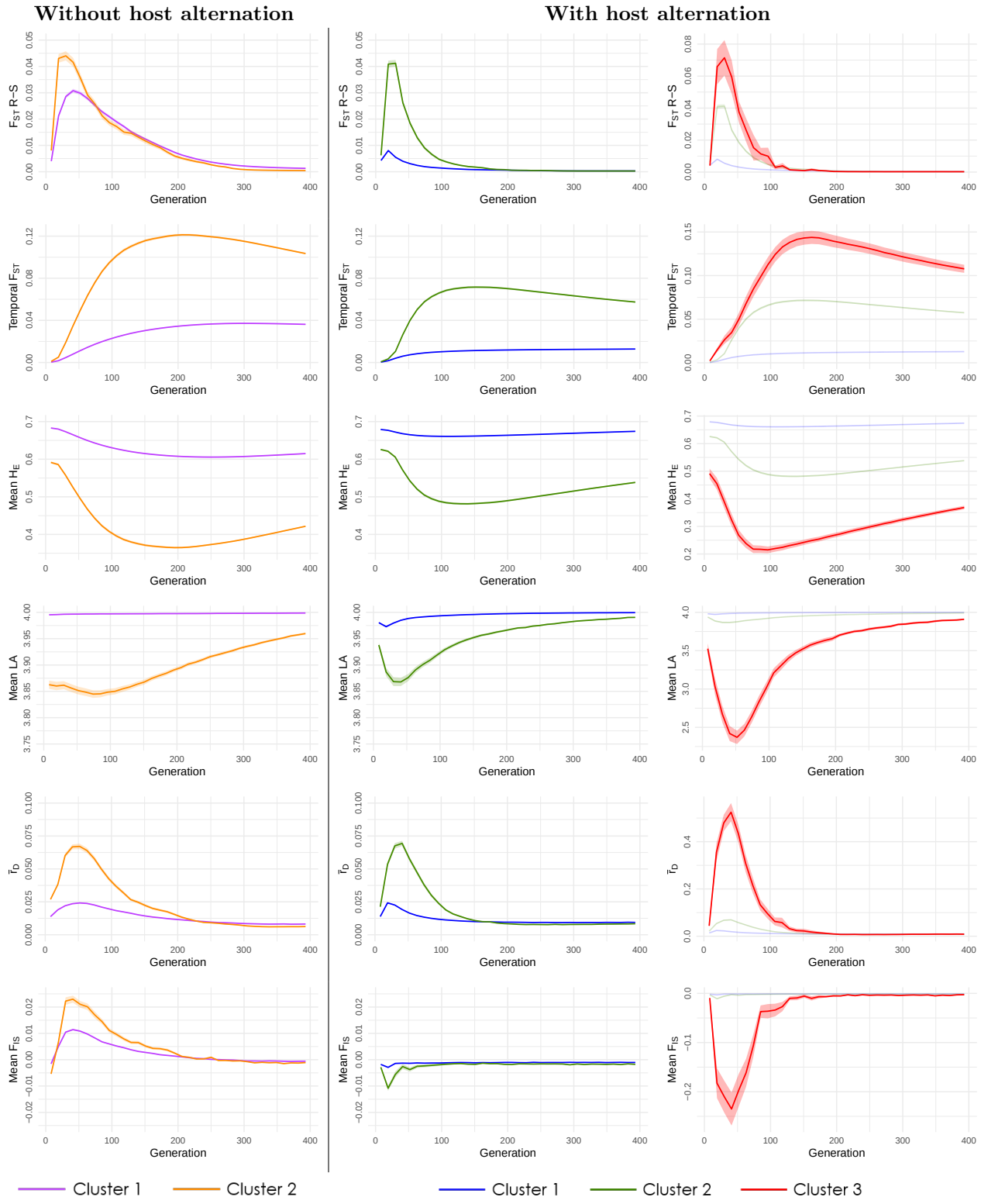
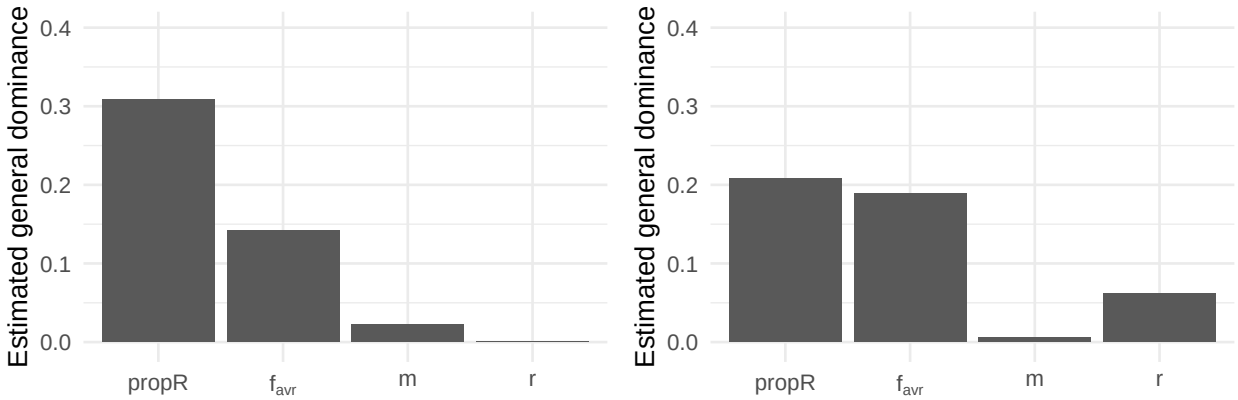


Figure 2: Temporal evolution of population genetics indices used for the clustering analysis. The plotted results correspond to the mean temporal dynamics for all simulations in the corresponding cluster, among all simulations of the random simulation design. Coloured intervals correspond to standard error intervals. For clarity, we apply the same scale for Cluster 1 and Cluster 2 of both life cycles, but we use a different scale to display the changes in population genetics indices for Cluster 3 with host alternation.

to a lesser extent, the migration rate (m) (Figure 3). As such, Cluster 2 is composed of simulations with significantly larger values of $propR$ and significantly smaller values of f_{avr} and slightly, albeit significantly, lower m. The growth rate (r) has no effect on cluster delineation (and consistently there is no significant difference in parameter distribution across clusters).



(a) Without host alternation

(b) With host alternation

Figure 3: Estimated general dominance calculated from a multinomial logistic regression applied to the cluster assignment for each simulation of the random design. The general dominance was estimated for the four input parameters: $propR$, f_{avr}, m, and r.

With host alternation, we observe the same differences between Cluster 1 and Cluster 2, with similar variations in the distribution of parameter values and ranking of parameter effects. The only difference is that r has an effect on cluster delineation (even more pronounced than m), with significantly higher growth rates for Cluster 2. Cluster 3 represents a particular case with skewed distributions towards very high values of $propR$ and low values of r (Figure 4). Cluster 3 displays rather high values of f_{avr} and higher values of m compared to the two other clusters.

4.2 Cluster delineation reflects different demographic scenarios

Cluster delineation, and thus the magnitude of genetic changes, reflects and distinguishes three demographic scenarios.

As seen above, the most influential parameter for cluster delineation is $propR$. As this parameter determines the maximal population sizes on R and S, it results in differences in population dynamics among clusters (Figure 5, Appendix Figure S3). The mean compartment size of S is higher for Cluster 1, irrespective of the life cycle, which leads to higher initial population sizes and less genetic changes through time on S.

For both life cycles, the second most influential parameter is the initial frequency of avirulent allele (f_{avr}), which plays a major role in the assignments to Cluster 2, with an over-representation of simulations with low values of f_{avr} (Figures 4, Appendix Figure S4). These low values of f_{avr} result in fewer virulent

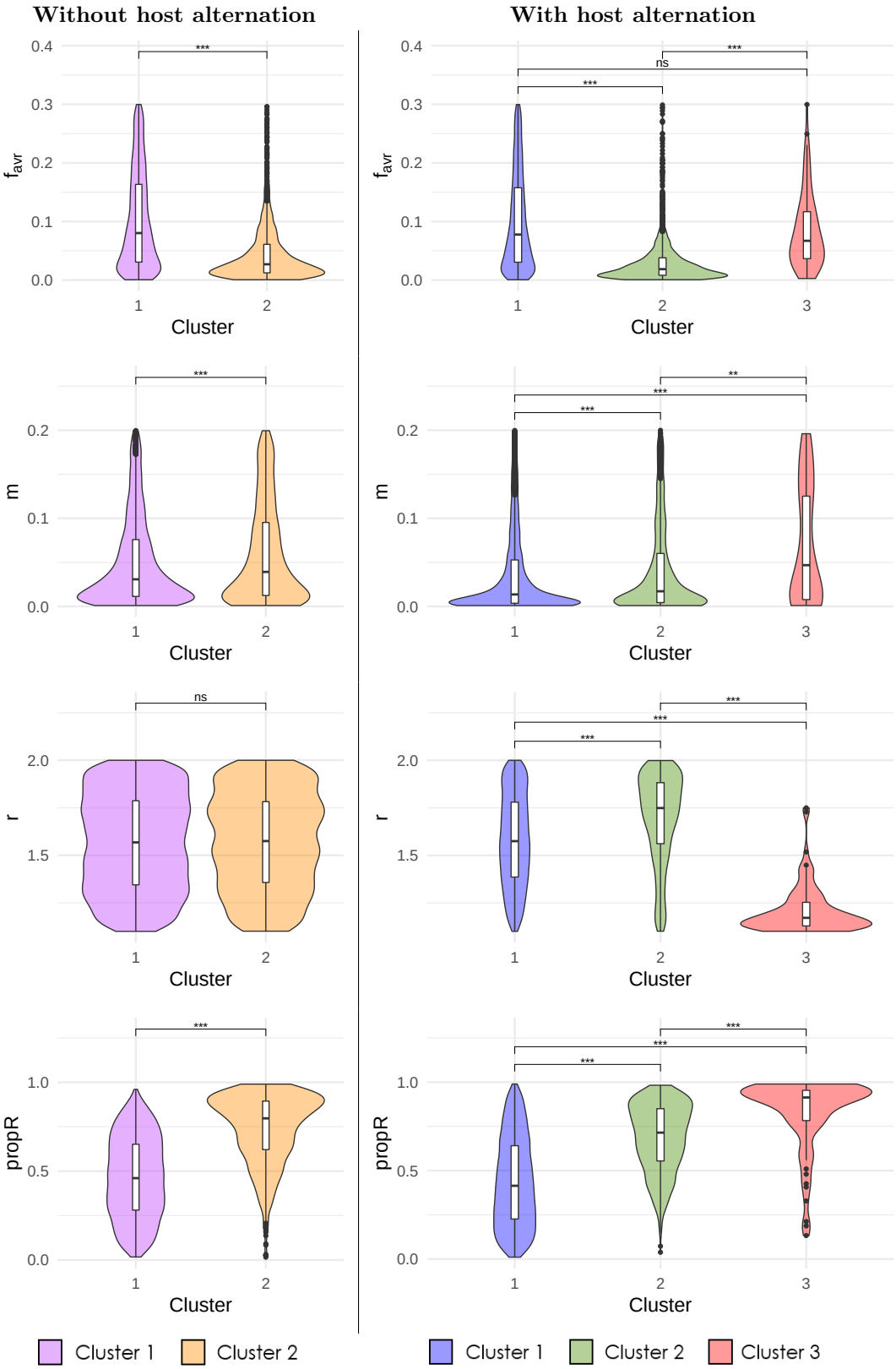


Figure 4: Distribution of epidemic parameters for each cluster obtained from the dynamics of population genetic indices. Difference between distributions were statistically assessed using pairwise Kruskal-Wallis tests : ns, non significant; ·, $P-value < 0.1$; *, $P-value < 0.05$; **, $P-value < 0.01$; ***, $P-value < 0.001$. Without host alternation, Cluster 1 and Cluster 2 are composed of 4067 and 1575 simulations, respectively. With host alternation, Cluster 1, Cluster 2 and Cluster 3 are composed of 5858, 1581, and 81 simulations, respectively.

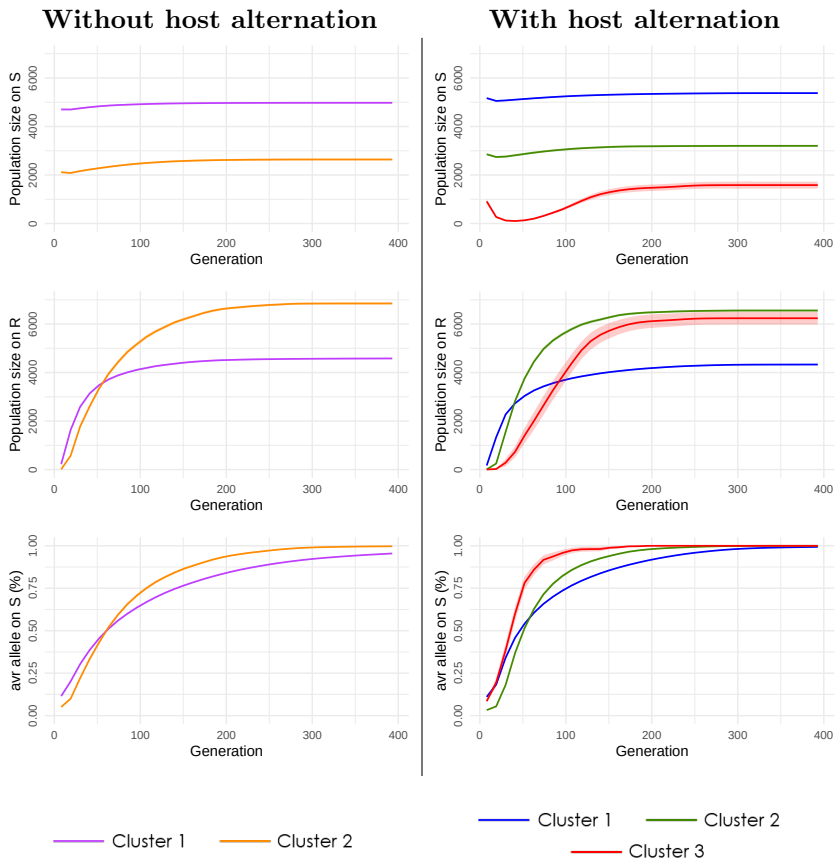


Figure 5: Temporal evolution of population sizes and virulent allele frequency depending on the cluster. The plotted results correspond to the mean temporal dynamics for all simulations in the corresponding cluster, among all simulations of the random simulation design. Coloured intervals correspond to standard error intervals.

individuals. It causes founder effects on R (Figure 5) and leads to more pronounced genetic changes through time (Appendix, Figures S1, S2).

A peculiar range of parameter values defines Cluster 3 observed with host alternation. It results in evolutionary rescue dynamics. The low growth rate associated with a restricted size of S (high $propR$) causes an initial decrease in population size until near extinction (Figure 5). Then, the emergence of virulent individuals leads to the establishment of a new population on R and these individuals, by migrating back to S, prevent population extinction (see example in Figure 6, Cluster 3). Proportions of evolutionary rescue events differ significantly among clusters, with Cluster 3 composed almost entirely of such simulations (Table 3).

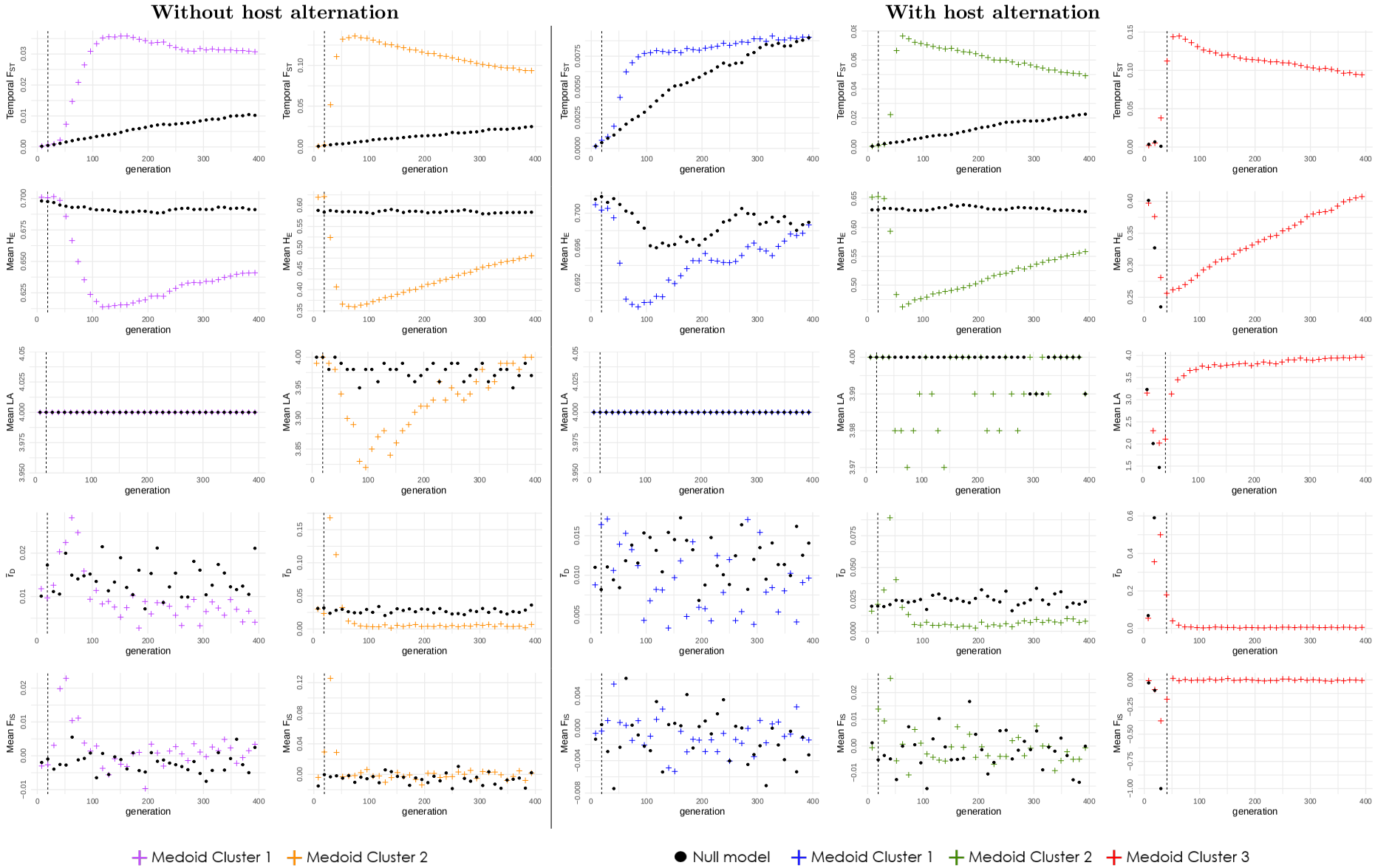


Figure 6: Comparison of the temporal evolution of population genetics indices used for the clustering analysis, between medoid simulations and null models. The plotted results correspond to the medoid dynamic of each cluster, with ('Medoid', cross shapes) or without ('Null model', round shapes) introducing virulent allele, that is, with or without resistance overcoming. For each of the plotted medoid, the vertical dashed line corresponds to the generation of resistance overcoming.

4.3 Inter- and intra-population genetic changes

To compare inter- and intra-population genetic changes associated with resistance overcoming, we focus on 1) mean dynamics, *i.e.* mean temporal evolution of population genetics indices on each cluster, calculated from neutral loci, and 2) realised dynamics, *i.e.* temporal evolution of population genetics indices for medoid or null model simulations of each cluster, calculated from neutral loci.

Inter-population genetic signatures

All clusters show a peak in the mean dynamics of genetic differentiation between populations on R and S ($F_{ST}R-S$, Figure 2). This peak in differentiation occurs within the first 50 generations and is concomitant with resistance overcoming (Appendix, Figure S5). The peak is more abrupt with host alternation, especially for Cluster 1, and the genetic differentiation persists for a shorter period with host alternation. This genetic differentiation decreases at the same rate for Cluster 1 and Cluster 2. The peak is higher for Cluster 2 than Cluster 1, which reflects stronger founder effects resulting from the invasion of R highlighted in Cluster 2. With host alternation, this peak in differentiation is magnified in Cluster 3 by the conjunction of the founder effect and the strong genetic drift on S resulting from the initial decrease in population size before evolutionary rescue.

For all clusters we observe a two-phase dynamics of temporal differentiation on S (from the initial population collected after burn-in) (Figures 2, 6). The first phase corresponds to an initial increase in temporal F_{ST} that coincides with resistance overcoming. Temporal F_{ST} reaches a maximum between 50 and 200 generations, depending on the cluster and the life cycle. This increase in temporal F_{ST} on S is delayed but similar to the temporal F_{ST} on R (see the example of the temporal F_{ST} on R for Cluster 2 with host alternation, Appendix, Figure S6). The second phase is more stable with a slight increase or decrease in temporal F_{ST}. This second phase is concomitant with the regain in genetic diversity on both compartments, which favours homogenisation of allele frequencies distorted by the founder event (Figure 2). With host alternation, simulations displaying strong genetic differentiation (realised dynamics of Cluster 2 and Cluster 3) show a steeper peak of differentiation (Figure 6). This is in accordance with the rapid decrease in differentiation between populations on R and S with host alternation, which shifts the temporal F_{ST} to the next phase of the dynamics. Conversely, without host alternation, differentiation between populations on R and S persists for a longer period, which leads to a slower change in temporal F_{ST} dynamics on S. Without selection, the temporal F_{ST} on S increases linearly through time at a slope depending on the relative importance of mutation and genetic drift forces (Realised dynamics of temporal F_{ST}, Figure 6). As the mutation rate is fixed, this slope depends on population size only. This increase without selection is much lower than that observed in the first phase of realised dynamics of resistance overcoming. The higher level of differentiation observed with selection is caused by genetic admixture with the differentiated population on R.

Intra-population genetic signatures

For all clusters, we observe a two-phase dynamics of temporal change in gene diversity (expected heterozygosity, H_E) on S. The first phase corresponds to a strong decrease in H_E, until a minimum value is reached between 50 and 200 generations, followed by a slower increase towards a new mutation-drift equilibrium (Figure 2). However, the timing of this change differs among clusters. For Cluster 1 and Cluster 2 of both life cycles, the decrease in H_E follows the generation of resistance overcoming, while the null model dynamics show no variations in H_E (Figure 6). The decrease in gene diversity results from the immigration of less diverse pathogen individuals from the founded population on R that overcame the resistance. For Cluster 3 (with host alternation), a strong decrease in H_E is observed for both the null model and the medoid realised dynamics, and the drop in H_E precedes resistance overcoming. Unlike other clusters, this drop in H_E is preceded for Cluster 3 by a strong decrease in the mean number of alleles $Mean\,LA$ (Figures 2, 6), indicating a bottleneck.

Simulations display a peak in linkage disequilibrium, with a maximum value of $\bar{r}_D$ being reached in the first 100 generations (Figures 2, 6). The variation in $\bar{r}_D$ is to be examined in relation to the variation in F_{IS}, which patterns differ between clusters and life cycles. Without host alternation, $\bar{r}_D$ and F_{IS} display similar variations but slightly delayed in time, with a maximum value of F_{IS} reached in the first 50 generations. This indicates admixture of genetically differentiated individuals on S. With host alternation, F_{IS} is null or displays slightly negative values for Cluster 1 and Cluster 2 while $\bar{r}_D$ remains positive. This indicates that the signature of the admixture on S is rapidly being erased with host alternation. In Cluster 3 with host alternation, the peak in $\bar{r}_D$ comes with very negative values of F_{IS}. The negative values in F_{IS} results from the decrease in H_E that is preceding the decrease in H_O (Figure 6, Appendix, Figure S7).

4.4 Stochasticity of dynamics

As for the clustering based on the random simulation design above, temporal changes in population genetics indices from the regular simulation design group into two clusters without host alternation, and three clusters with host alternation (Table 4). For both life cycles, Cluster 1 regroups the majority of the simulations (53 % and 83 % of the simulations without and with host alternation, respectively, Table 4). Notably, we identify simulation replicates (*i.e.* with identical combinations of parameter values) that show different outcomes in the temporal changes of population genetics indices. Hence, simulation replicates can be assigned to different clusters, demonstrating a certain degree of stochasticity. These replicates assigned to different clusters represent 6% and 12% of the simulations in Cluster 1 and Cluster 2 without host alternation, respectively, and 17%, 85% and 98% of the simulations in Cluster 1, Cluster 2, and Cluster 3 with host alternation, respectively (Table 4). Therefore, cluster delineation is not defined by specific combinations of parameter values, but rather by the inherent demographic scenarios underpinning the genetic dynamics.

Table 4: Cluster description for the clustering based on the regular simulation design. This table contains the number of simulations assigned to each cluster, and the percentages of simulations from a same set of parameters (100 replicates) which have been assigned to different clusters ('Different assignments (%)'). Only simulations leading to resistance overcoming have been kept for the clustering analysis.

Cycle	Cluster	Number of simulations	Different assignments (%)
Without host alternation	1	2966	5.9
	2	2605	12.4
With host alternation	1	6937	16.9
	2	1332	85.0
	3	42	97.6

5 Discussion

5.1 Genetic signatures of resistance overcoming

All the recorded population genetics indices were impacted by the simulated resistance overcoming. Overall, resistance overcoming leads to a founder effect on the resistant host, with a differentiated sub-sampled population settling and growing on resistant hosts. Migrations between susceptible and resistant hosts then homogenise the genetic coancestry over all pathogen populations at a pace that depends on the migration rate and pathogen population sizes. Resistance overcoming has a strong impact on the population genetic structure on susceptible hosts, with 1) a decrease in pathogen genetic diversity, 2) a peak in linkage disequilibrium, 3) a strong increase in temporal genetic differentiation between the initial and evolved populations, and 4) a peak in population differentiation between the susceptible and resistant compartments. The comparison of the evolution of the genetic indices through time with and without selection showed that these changes were signatures of evolution under selection and did not result from genetic drift only. Our first important result is thus that changes in population genetics of neutral markers allow identifying a selective event of resistance overcoming. We note that most of these genetic changes are transient, with a signature of resistance overcoming that vanishes in a few years only.

5.2 Typology of dynamics under resistance overcoming

Each evolutionary scenario represents distinct genetic signatures of resistance overcoming

Simulations can be distinguished by the magnitude of their genetic signatures and grouped into clusters that are indicative of different evolutionary scenarios. This clustering is strongly linked to variations in pathogen population sizes. Cluster 1 regroups simulations with the slightest genetic signatures, associated with a steady and slow demographic expansion. Hence, a large part of the simulations leads to signatures of particularly low magnitudes. Cluster 2 regroups simulations with stronger genetic signatures, associated with larger demographic expansions on the resistant compartment. In extreme cases with host alternation, the few simulations assigned to Cluster 3 present the strongest genetic signatures, mainly associated with a specific demographic scenario, namely evolutionary rescue. These simulations are characterised not only by

a strong demographic expansion on the resistant compartment (as for Cluster 2) but also by a demographic recovery on the susceptible compartment. Overall, founder effects lead to a demographic expansion that is all the more important that the resistant host is abundant, because of the logistic population growth. During the first generations following the invasion of the resistant compartment, the successive clonal reproductions lead to a large population of few genotypes largely repeated, thus strongly entangling demographic variations and genetic changes.

Differences due to input parameters

In this model, the proportion of resistant hosts is the main determinant of pathogen population sizes, and not the intrinsic demographic parameters (pathogen growth rate and migration rate). For both life cycles, the proportion of resistant hosts determines the initial pathogen population size and the size of the compartment available through adaptation. As such, it drives the strength of both population expansion and selection pressure exerted on the pathogen population. With host alternation, this proportion also shapes the population demography in determining the vulnerability of the population to evolutionary rescue (See Figure 4 in Saubin et al., 2021). Pathogen control often leads to high proportions of resistant hosts in agricultural landscapes (Stukenbrock and McDonald, 2008; Zhan et al., 2015). Here we show that in case of resistance overcoming, higher proportions of resistant hosts lead to more pronounced genetic changes in the pathogen population. The initial proportion of virulent allele in the pathogen population is also a strong driver of the strength of the genetic signatures. This extent of standing genetic variation determines the proportion of individuals that will be able to respond to selection, in other words, the adaptive potential of the initial population. In particular, the proportion of virulent allele impacts the number of individuals that settle the virulent population, hence the genetic diversity of the founded population. A small proportion leads to fewer virulent individuals, and therefore a less diverse and more differentiated population founded on resistant hosts, as observed for frequent turnover of extinction and recolonisation (McCauley, 1991).

Differences due to the pathogen life cycle

Without host alternation, genetic signatures remain detectable for a longer period of time, because of sustained genetic admixture. Under this life cycle, gene flow results from the movements of a limited number of individuals, namely a proportion of the population determined by the migration rate. Migration occurs at each generation, which leads to a regular but progressive homogenisation of populations evolving on resistant and susceptible hosts. This accounts for the delay between the maximum differentiation observed between compartments at the time of the founder event, and the return to a low differentiation after the homogenisation of the populations. At the time of maximum differentiation, the immigration on susceptible hosts of individuals from the newly founded population on resistant hosts causes a Wahlund effect (Wahlund, 1928), that is a distortion of allele frequencies caused by the admixture of genotypes originating from different subpopulations. This explains the positive values of both the inbreeding coefficient and the linkage disequilibrium.

With host alternation, genetic differentiation between compartments fades more rapidly because of the obligate mating event taking place in a common host (A), where alleles at each locus are reshuffled through sexual reproduction. It erases the Wahlund effect observed without host alternation as a single event of sexual reproduction among all individuals is sufficient for a return to Hardy-Weinberg equilibrium (Rouger et al., 2016). This explains the small or slightly negative values of the inbreeding coefficient. Yet, linkage disequilibrium remains positive because the associations of alleles across loci are still preserved for some time, as recombination occurs within individuals (whose allele frequencies are inherited from either population. Last, all individuals are redistributed randomly between the two compartments. The death of avirulent individuals on resistant hosts distorts allele frequencies and regenerates differentiation between pathogens evolving on resistant and susceptible hosts. This life cycle increases gene flow and leads to a fast homogenisation of pathogen populations.

These differences between life cycles explain the observed differences in demographic variations and genetic changes in pathogen populations. Examples of these different genetic outcomes can be found in the literature: from conservation of genetic structure without host alternation (*e.g.* in Leroy et al., 2013) to strong selective sweep and gene swamping with host alternation (*e.g.* in Persoons et al., 2017).

Stochasticity and genetic signatures

Beyond the influence of input parameters and life cycle on demographic variations and genetic changes, identical combinations of input parameters can lead to different outcomes. This variation among simulation replicates results from the inherent stochasticity of the model. In particular, migrants are picked randomly. The generation at which the first virulent individuals actually invade the resistant host drives the resulting demography and genetic structure. One year is composed of one sexual reproduction and ten clonal reproductions. Therefore, if the invasion occurs right after the sexual reproduction, several clonal reproductions follow, building a strongly differentiated population on resistant hosts. Conversely, if the invasion occurs at the end of the season of clonal reproductions, the population size does not change much before the next sexual reproduction reshuffles alleles in the population. Moreover, the stochasticity impacts not only the timing of invasion but also the number of successful migrations between comparments, hence the genetic diversity of the founded population. Beyond the inherent properties of the model, the life cycle with host alternation also adds stochasticity. Without host alternation, a first invasion of the resistant compartment settles a population, hence leading inevitably to resistance overcoming. Conversely, with host alternation, a first invasion event is necessary but not sufficient for a sustained settlement. The regular emptying of compartments before sexual reproduction and random individual redistribution can lead thereafter to stochastic invasions of resistant hosts as long as the virulent allele is rare (*e.g.* invaded compartment one year and cleared compartment the next year). As the virulent allele is recessive, it is more vulnerable to extinction

with host alternation (Saubin et al., 2021). Here we show that in addition to the stochasticity in the fate of the virulent allele, this life cycle also increases the stochasticity in the genetic signatures of resistance overcoming.

5.3 Genetic signatures characteristic of evolutionary rescue

In extreme cases where the pathogen growth rate is not sufficient to counteract the mortality of individuals on resistant hosts, the population can decline towards extinction. Through resistance overcoming, a population rescue can occur. Such rescue can be considered as demographic or evolutionary, depending on the definition of the population. If we consider distinctly pathogen populations on susceptible and resistant hosts, the adaptation of virulent individuals leads to their invasion of the resistant host, hence to the survival of the population on the susceptible host. The survival of the population on the susceptible host corresponds to demographic rescue resulting from the immigration of adapted individuals from the resistant host. If we consider a single population gathering all individuals evolving on both hosts, the survival of the population corresponds to an evolutionary rescue event, *via* the adaptation to the new environment, *i.e.* the newly deployed resistant hosts. Such evolutionary rescue events occur mostly for the life cycle with host alternation (Saubin et al., 2021). This is because of additional mortality that originates once a year from the massive redistribution of individuals after the sexual reproduction, with the death of all avirulent individuals that migrate to resistant hosts. Besides the additional mortality, the massive redistribution also increases the probability that a virulent individual migrates to the resistant host, and hence increases the probability of evolutionary rescue (Saubin et al., 2021). Here we demonstrate that such events can lead to strong and typical genetic signatures at neutral loci, that define a specific cluster with host alternation (Cluster 3) and can thus be uncovered using population genetics indices. These dynamics are characterised by a bottleneck with few possible genotypes combining the remaining alleles. This is evidenced by the changes in several indices, such as the drop in the mean number of alleles, followed by the drastic reduction in gene diversity and the increase in linkage disequilibrium (Cornuet and Luikart, 1996). Unexpectedly, this bottleneck comes along with a strikingly negative value of inbreeding coefficient. To understand this result, we investigate the cause of the discrepancy between expected and observed heterozygosity. This is due to the clonal reproduction events that maintain identical genotypes, hence the value of observed heterozygosity remains constant over generations, whereas expected heterozygosity steadily decreases because the population size is collapsing very fast. We note that as the sampling takes place at the end of the clonal phase, the difference between expected and observed heterozygosity is magnified by the small genotypic drift that happened between clonal lineages over the handful of clonal generations. At the time of the founder effect on the resistant host, these dynamics are characterised by a large differentiation between compartments. Yet this differentiation erases as quickly as the other genetic signatures of resistance overcoming. Overall, even for such drastic demographic events, the resulting genetic signatures remain transients, and thus can only be captured by using time sample data in the appropriate (short) time window.

5.4 Temporal changes of demogenetic signatures

Among indices, we observe different temporal dynamics of genetic signatures through resistance overcoming. The fastest changes are observed for the genetic differentiation between S and R and the mean inbreeding coefficient, during the 50 first generations after the resistance overcoming. This is consistent with the resistance overcoming of resistance RMlp7 in poplars by the poplar rust pathogen that led to drastic genetic changes in less than five years (Persoons et al., 2017). Changes in linkage disequilibrium and mean number of alleles are slower and detected within the 100 first generations. Finally, the slowest changes (and less sharp temporal signatures) are observed for the expected heterozygosity and the temporal genetic differentiation, with most of the detectable signal occurring between generations 50 and 200. Overall no peak of genetic signature occurs after generation 200, and only residual signal remains on population genetic indices.

For indices reflecting disequilibrium induced by the founder effect (F_{ST} $R-S$, $\bar{r}_D$, $Mean\ F_{IS}$), genetic changes stabilise rapidly, especially when gene flow is enhanced by a life cycle with host alternation. For other indices (temporal F_{ST}, H_E, $Mean\ LA$) and when it exists, the difference with the null model (*i.e.* without selection) persists for a much longer period (for at least 300 generations for $Mean\ LA$ and until the end of the simulated period for temporal F_{ST} and H_E). Following rapid adaption, we thus observe a return to genetic equilibrium in two steps. The first step involves the homogenisation of allele frequencies both within and between populations thanks to the convergence to Hardy-Weinberg and migration-drift equilibria. This occurs very fast because the modelled system is regularly subjected to the stabilising effect of sexual reproduction (Rouger et al., 2016). We hypothesise that this return would be slower for systems that deviate from this mode of reproduction, either because of an increased rate of clonality (Reichel et al., 2016) or because of selfing (Jullien et al., 2019). The second step is the return to allele numbers and frequencies expected under the migration-mutation-drift equilibrium. This evolution is slower because it relies on the progressive change in allelic states and potential recovery of the lost genetic diversity that happened during the upheaval caused by resistance overcoming.

Therefore, observing the first return to equilibrium requires the sampling of the population close to the time of resistance overcoming. Conversely, observing the second return to equilibrium requires a broader temporal sampling including the characterisation of the initial state. In the absence of this initial state, it is not possible to highlight these changes in allele frequencies at the virulence locus and neutral markers (Persoons et al., 2017).

The timing of the modelled events may appear rapid but is consistent with empirical observations. For example, in the poplar rust pathogen (*Melampsora larici-populina*) which alternates on larch every year, the overcoming of resistance RMlp7 in poplars led in 1994 to a strong genetic disequilibrium. This was followed by a quick return to Hardy-Weinberg equilibrium the following year and drastic changes in the population genetic structure occurring in less than four years (Persoons et al., 2017; Louet et al., 2021).

Overall, studies employing time-series remain rare compared to the amount of work focusing on the genetic analysis of one contemporary pathogen population (Buffalo and Coop, 2019, 2020; Pavinato et al.,

2021). Analyses of time-series are mostly used to analyse the speed and timing of selection for life history traits (Le Rouzic et al., 2011, 2015), loci under positive selection (Bergland et al., 2014; Foll et al., 2015), or coevolution between host and parasite (Decaestecker et al., 2007; Gandon et al., 2008). We note that studies including neutral markers mainly use them to draw a statistical inference of loci under selection rather than to document the epidemic. In the rare cases where time-series are used to study specifically neutral genetic evolution, the data typically exhibit a limited temporal range (*i.e.* two time points (Pavinato et al., 2021). However, as temporal data allow to trace the changes in allele frequency through time, the analysis of neutral markers can improve our inference and understanding of evolutionary processes (Dehasque et al., 2020; Feder et al., 2021; Saubin et al., 2022). In cases of rapid adaptation especially, we show here that the resulting genetic signatures may be very brief and require time samples around the selection event. Studies focusing on one or two time points may allow documenting only the second return to equilibrium (from the influence of evolutionary forces) and will fail to highlight transient dynamics, which provide the most relevant information regarding the demogenetic interplay. Conversely, epidemiological considerations lead to focus on the time when the invasion is detected, therefore when the genetic signature is the strongest. These considerations tend to neglect the initial state, as well as the return to a new equilibrium, which may nevertheless occur in a relatively short time scale. A key result of our study is to demonstrate that neutral markers can be used to uncover demogenetic processes due to selection events.

5.5 Concluding remarks

Our demogenetic analysis and new simulation framework enable us to study in depth the entire range of possible scenarios of resistance overcoming and to identify and discriminate between several eco-evolutionary dynamics using neutral genetic markers. To deal with the high amount of information from stochastic dynamics we performed clustering based on the time-series changes of population genetics indices. Using this method, we highlight three main types of genetic signatures of increasing magnitude that reflect different demographic processes through resistance overcoming: 1) small variations in population sizes, 2) a strong founder event, and 3) a strong founder event associated with a population bottleneck. We speculate that this simulation framework and our results pave the way for inferring demogenetic dynamics underpinning rapid adaptation using empirical genetic data.

Acknowledgements

We thank Alexis Sarda for insightful answers on the Dtwclust package. This work was supported by grants from the French National Research Agency (ANR-18-CE32-0001, Clonix2D project; ANR-11-LABX-0002-01, Cluster of Excellence Arbre). Méline Saubin was supported by a PhD fellowship from INRAE and the French National Research Agency (ANR-18-CE32-0001, Clonix2D project).

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Agapow, P. M. and Burt, A. (2001). Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, 1(1-2):101–102.

Agrios, G. N. (2005). *Plant pathology*. Elsevier edition.

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 in Ecology and Evolution*, 19(10):536–544.

Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., and Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognition*, 46(1):243–256.

Archie, E. A., Luikart, G., and Ezenwa, V. O. (2009). Infecting epidemiology with genetics: A new frontier in disease ecology. *Trends in Ecology and Evolution*, 24(1):21–30.

Arnaud-Haond, S., Duarte, C. M., Alberto, F., and Serrão, E. A. (2007). Standardizing methods to address clonality in population studies. *Molecular Ecology*, 16:5115–5139.

Barrett, R. D. H. and Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, 23(1):38–44.

Bell, G. (2017). Evolutionary rescue. *Annual Review of Ecology, Evolution, and Systematics*, 48:605–627.

Bergland, A. O., Behrman, E. L., O'Brien, K. R., Schmidt, P. S., and Petrov, D. A. (2014). Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in Drosophila. *PLoS Genetics*, 10(11).

Bolker, B. M., Nanda, A., and Shah, D. (2010). Transient virulence of emerging pathogens. *Journal of the Royal Society Interface*, 7.

Bonneaud, C. and Longdon, B. (2020). Using evolutionary theory to understand the fate of novel infectious pathogens. *Science Society*, 21(e51374).

Brown, J. H. and Kodric-Brown, A. (1977). Turnover rates in insular biogeography: Effect of immigration on extinction. *Ecology*, 58:445–449.

Brown, J. K. M. and Tellier, A. (2011). Plant-parasite coevolution: Bridging the gap between genetics and ecology. *Annual Review of Phytopathology*, 49(1):345–367.

Buffalo, V. and Coop, G. (2019). *The linked selection signature of rapid adaptation in temporal genomic data*, volume 213.

Buffalo, V. and Coop, G. (2020). Estimating the genome-wide contribution of selection to temporal allele frequency change. *Proceedings of the National Academy of Sciences of the United States of America*, 117(34):20672–20680.

Burdon, J. J., Zhan, J., Barrett, L. G., Papaix, J., and Thrall, P. H. (2016). Addressing the challenges of pathogen evolution on the world's arable crops. *Phytopathology*, 106(10):1117–1127.

Carlson, S. M., Cunningham, C. J., and Westley, P. A. H. (2014). Evolutionary rescue in a changing world. *Trends in Ecology and Evolution*, 29(9):521–530.

Cornuet, J. M. and Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, 144(4):2001–2014.

Day, T. and Gandon, S. (2007). Applying population-genetic models in theoretical evolutionary epidemiology. *Ecology Letters*, 10:876–888.

Day, T. and Proulx, S. R. (2004). A general theory for the evolutionary dynamics of virulence. *The American naturalist*, 163(4).

Decaestecker, E., Gaba, S., Raeymaekers, J. A. M., Stoks, R., Van Kerckhoven, L., Ebert, D., and De Meester, L. (2007). Host-parasite 'Red Queen' dynamics archived in pond sediment. *Nature*, 450(7171):870–873.

Dehasque, M., Ávila-Arcos, M. C., Díez-del Molino, D., Fumagalli, M., Guschanski, K., Lorenzen, E. D., Malaspinas, A. S., Marques-Bonet, T., Martin, M. D., Murray, G. G. R., Papadopulos, A. S. T., Therkildsen, N. O., Wegmann, D., Dalén, L., and Foote, A. D. (2020). Inference of natural selection from ancient DNA. *Evolution Letters*, 4(2):94–108.

Dorken, M. E. and Eckert, C. G. (2001). Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). *Journal of Ecology*, 89(3):339–350.

Feder, A. F., Pennings, P. S., and Petrov, D. A. (2021). The clarifying role of time series data in the population genetics of HIV. *PLoS Genetics*, 17(1):1–10.

Fischhoff, I. R., Huang, T., Hamilton, S. K., Han, B. A., LaDeau, S. L., Ostfeld, R. S., Rosi, E. J., and Solomon, C. T. (2020). Parasite and pathogen effects on ecosystem processes : A quantitative review. *Ecosphere*, 11(5).

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(7393):186–194.

Foll, M., Shim, H., and Jensen, J. D. (2015). WFABC: A Wright-Fisher ABC-based approach for inferring effective population sizes and selection coefficients from time-sampled data. *Molecular Ecology Resources*, 15(1):87–98.

Gandon, S., Buckling, A., Decaestecker, E., and Day, T. (2008). Host-parasite coevolution and patterns of adaptation across time and space. *Journal of Evolutionary Biology*, 21(6):1861–1866.

Gladieux, P., Byrnes, E. J., Aguileta, G., Fisher, M. C., Heitman, J., and Giraud, T. (2011). *Epidemiology and evolution of fungal pathogens in plants and animals*.

Gokhale, C. S., Papkou, A., Traulsen, A., and Schulenburg, H. (2013). Lotka – Volterra dynamics kills the Red Queen : population size fluctuations and associated stochasticity dramatically change host-parasite coevolution. *BMC Evolutionary Biology*, 13(254):3–10.

Gonzalez, A., Ronce, O., Ferriere, R., and Hochberg, M. E. (2013). Evolutionary rescue: An emerging focus at the intersection between ecology and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1610):1–8.

Hacquard, S., Petre, B., Frey, P., Hecker, A., Rouhier, N., and Duplessis, S. (2011). The poplar-poplar rust interaction: Insights from genomics and transcriptomics. *Journal of Pathogens*, 2011:1–11.

Halkett, F., Simon, J. C., and Balloux, F. (2005). Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology and Evolution*, 20.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825):357–362.

Johnson, R. (1984). A critical analysis of durable resistance. *Annual Review of Phytopathology*, 22.

Jullien, M., Navascués, M., Ronfort, J., Loridon, K., and Gay, L. (2019). Structure of multilocus genetic diversity in predominantly selfing populations. *Heredity*, 123(2):176–191.

Laine, A. L. and Barrès, B. (2013). Epidemiological and evolutionary consequences of life-history trade-offs in pathogens. *Plant Pathology*, 62:96–105.

Le Rouzic, A., Hansen, T. F., Gosden, T. P., and Svensson, E. I. (2015). Evolutionary time-series analysis reveals the signature of frequency-dependent selection on a female mating polymorphism. *American Naturalist*, 185(6):E182–E196.

Le Rouzic, A., Houle, D., and Hansen, T. F. (2011). A modelling framework for the analysis of artificial-selection time series. *Genetics Research*, 93(2):155–173.

Leroy, T., Lemaire, C., Dunemann, F., and 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).

Louet, C., Saubin, M., Andrieux, A., Persoons, A., Gorse, M., Pétrowski, J., Fabre, B., De Mita, S., Duplessis, S., Frey, P., and Halkett, F. (2021). A point mutation and large deletion at the candidate avirulence locus *AvrMlp7* in the poplar rust fungus correlate with poplar RMlp7 resistance breakdown. *Molecular Ecology*, (February):1–12.

Luchman, J. (2022). *domir: Tools to support relative importance analysis*.

McCauley, D. E. (1991). Genetic consequences of local population extinction and recolonization. *Trends in Ecology and Evolution*, 6(1):5–8.

McDonald, B. A. (2004). Population genetics of plant pathogens. *The Plant Health Instructor*.

McDonald, B. A. and Linde, C. (2002). Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, 40:349–379.

Milgroom, M. G. and Peever, T. L. (2003). Population biology of plant pathogens: The synthesis of plant disease epidemiology and population genetics. *Plant Disease*, 87(6):608–617.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89(3):583–590.

Nei, M. and Tajima, F. (1981). Genetic drift and estimation of effective population size. *Genetics*, 98(3):625–640.

Nielsen, R. and Signorovitch, J. (2003). Correcting for ascertainment biases when analyzing SNP data: Applications to the estimation of linkage disequilibrium. *Theoretical Population Biology*, 63(3):245–255.

Otto, S. P. and Whitlock, M. C. (1997). The probability of fixation in populations of changing size. *Genetics*, 146(2):723–733.

Parsons, T. L., Lambert, A., Day, T., and Gandon, S. (2018). Pathogen evolution in finite populations : slow and steady spreads the best. *Journal of royal society*, 15.

Pavinato, V. A. C., De Mita, S., Marin, J. M., and de Navascués, M. (2021). Joint inference of adaptive and demographic history from temporal population genomic data. *bioRxiv*, pages 1–18.

Persoons, A., Hayden, K. J., Fabre, B., Frey, P., De Mita, S., Tellier, A., and Halkett, F. (2017). The escalatory Red Queen: Population extinction and replacement following arms race dynamics in poplar rust. *Molecular Ecology*, 26(7):1902–1918.

Pink, D. and Puddephat, I. (1999). Deployment of disease resistance genes by plant transformation - a 'mix and match' approach. *Trends in Plant Science*, 4(2):71–75.

Pollak, E. (1983). A new method for estimating the effective population size from allele frequency changes. *Genetics*, 104(3):531–548.

R (2018). R Core Team, R: A language and environment for statistical computing.

Reichel, K., Masson, J. P., Malrieu, F., Arnaud-Haond, S., and Stoeckel, S. (2016). Rare sex or out of reach equilibrium? The dynamics of FIS in partially clonal organisms. *BMC Genetics*.

Rouger, R., Reichel, K., Malrieu, F., Masson, J. P., and Stoeckel, S. (2016). Effects of complex life cycles on genetic diversity: Cyclical parthenogenesis. *Heredity*, 117(5):336–347.

Sardá-Espinosa, A. (2019). Time-series clustering in R using the dtwclust package. *R Journal*, 11(1):1–22.

Saubin, M., De Mita, S., Zhu, X., Sudret, B., and Halkett, F. (2021). Impact of ploidy and pathogen life cycle on resistance durability. *Peer Community Journal*, 1(e8):1–31.

Saubin, M., Louet, C., Bousset, L., Fabre, F., Frey, P., Fudal, I., Grognard, F., Hamelin, F., Mailleret, L., Stoeckel, S., Touzeau, S., Petre, B., and Halkett, F. (2022). Improving sustainable crop protection using population genetics concepts. *Molecular Ecology*, 00:1–11.

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., Mcroberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature ecology evolution*, 3(3):430–439.

Stoeckel, S. and Masson, J. P. (2014). The exact distributions of FIS under partial asexuality in small finite populations with mutation. *PLoS ONE*, 9(1).

Stukenbrock, E. H. and McDonald, B. A. (2008). The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, 46:75–100.

Stukenbrock, E. H. and McDonald, B. A. (2009). Population genetics of fungal and oomycete effectors involved in gene-for-gene interactions. *Molecular Plant-Microbe Interactions*, 22(4):371–380.

Thrall, P. H. (1998). Metapopulation collapse: The consequences of limited gene flow in spatially structured populations. In *Modeling Spatiotemporal Dynamics in Ecology,*, pages 83–104. Springer–Verlag.

Thrall, P. H., Barrett, L. G., Dodds, P. N., and Burdon, J. J. (2016). Epidemiological and evolutionary outcomes in gene-for-gene and matching allele models. *Frontiers in Plant Science*, 6(1084).

Tobin, P. C. (2015). Ecological consequences of pathogen and insect invasions. *Forest pathology and entomology*, 1:25–32.

van Rossum, G. (1995). Python tutorial, Technical Report CS-R9526. *CWI*.

Wahlund, V. S. (1928). Zusammensetzung von populationen und korrelationserscheinungen vom standpunkt der vererbungslehre aus betrachtet. *Hereditas*, 11:65–106.

Wright, S. (1931). Evolution in mendelian populations. *Genetics*, 16.

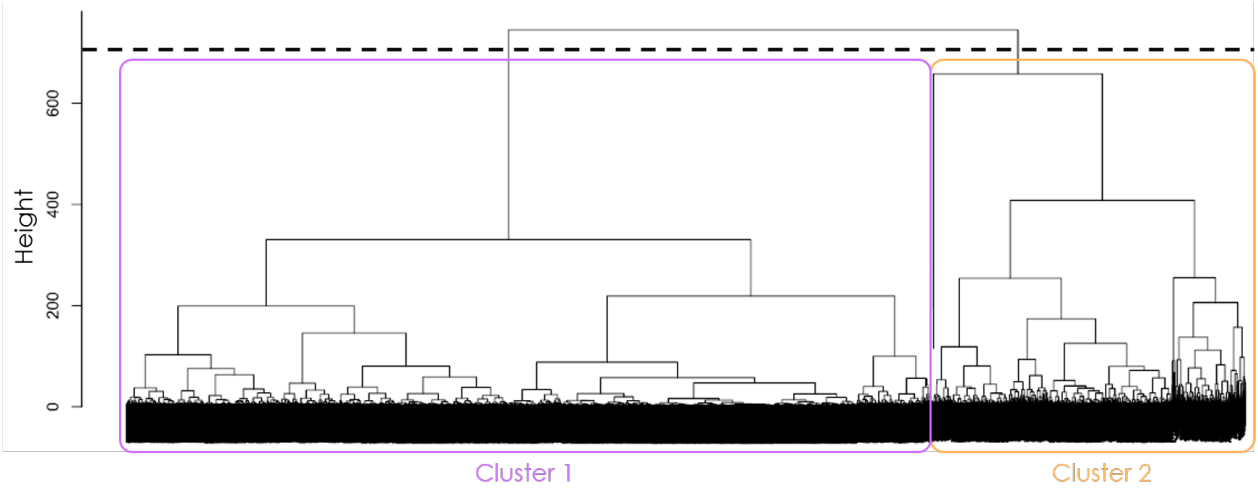
Wright, S. (1949). *Adaptation and selection*, chapter Adaptation, pages 365–389. Princeton Univ. Press, Princeton, NJ.

Wright, S. (1978). *Evolution and the genetics of populations - Variability within and among natural populations*. The University of Chicago Press.

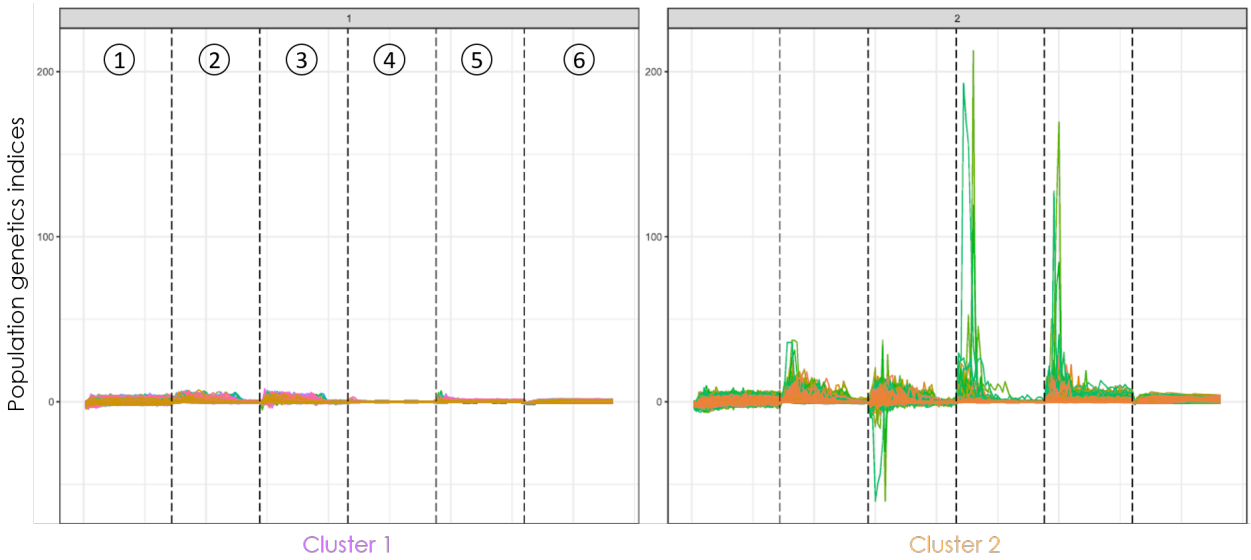
Zhan, J., Thrall, P. H., Papaïx, J., Xie, L., and Burdon, J. J. (2015). Playing on a pathogen's weakness: Using evolution to guide sustainable plant disease control strategies. *Annual Review of Phytopathology*, 53(1):19–43.

Živković, D., John, S., Verin, M., Stephan, W., and Tellier, A. (2019). Neutral genomic signatures of host-parasite coevolution. *BMC Evolutionary Biology*, 19(1):1–11.

Appendix

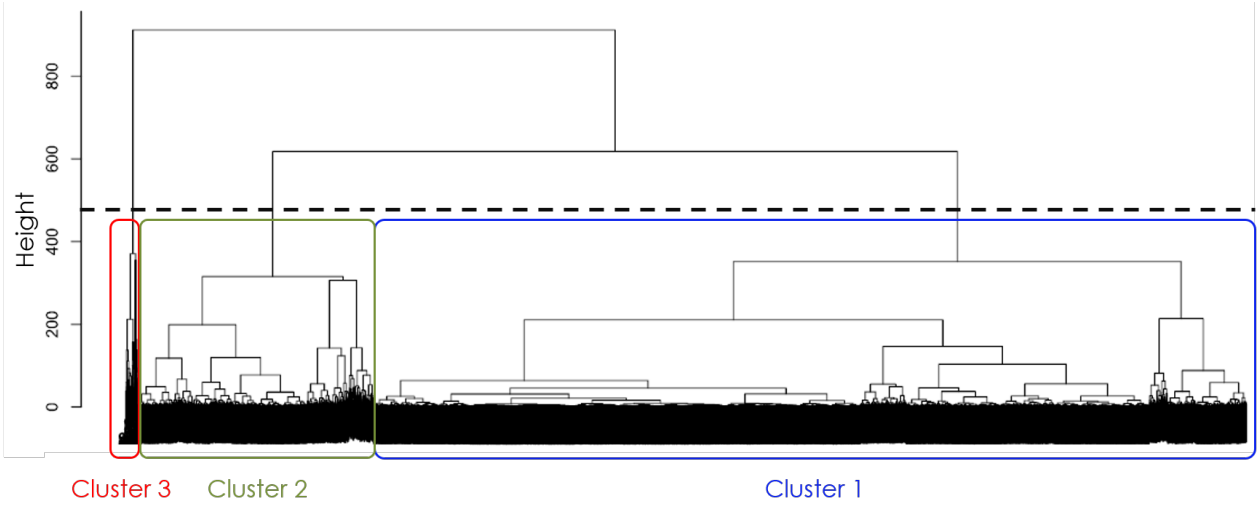


(a) Clustering of simulations

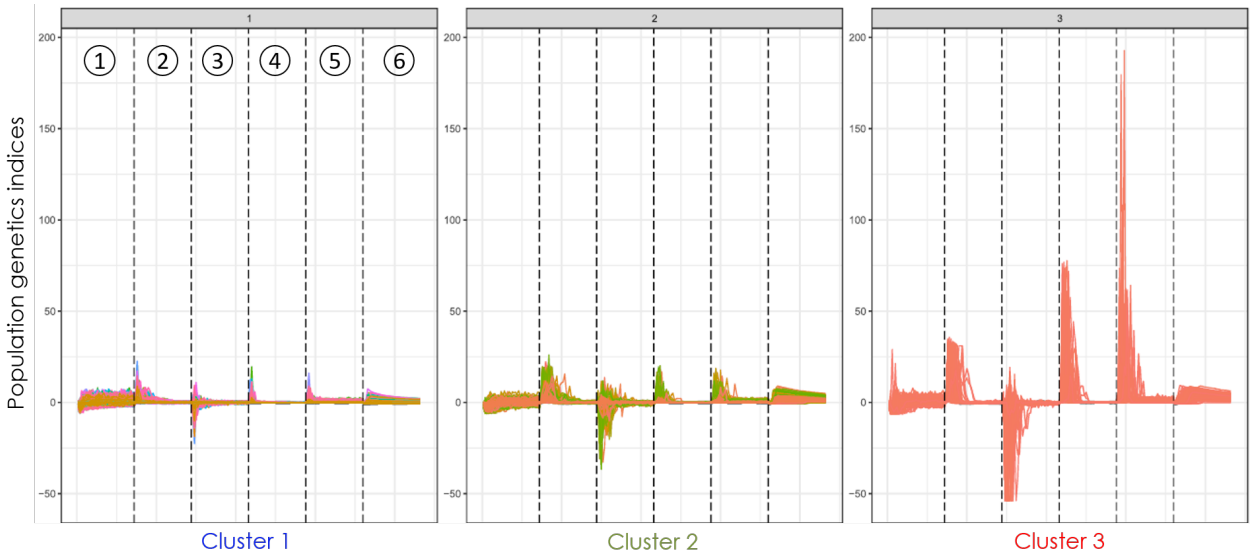


(b) Temporal evolution of population genetics indices

Figure S1: Clustering of simulations without host alternation. The clustering is based on the temporal evolution of six population genetics indices. (a) represents the dendrogram obtained from the clustering analysis. Distances between simulations were calculated based on the dynamic time warping distance, with the norm 'Eucliean distance' and the step pattern 'symmetric 2'. (b) represents for each cluster the concatenation of the temporal evolution of six normalised populations genetics indices, in the following order: (1) $\beta of Pareto$, (2) $\bar{r}_D$, (3) $Mean\ F_{IS}$, (4) $Variance\ F_{IS}$, (5) F_k, (6) $Temporal\ F_{ST}$.



(a) Clustering of simulations



(b) Temporal evolution of population genetics indices

Figure S2: Clustering of simulations with host alternation. The clustering is based on the temporal evolution of six population genetics indices. (a) represents the dendrogram obtained from the clustering analysis. Distances between simulations were calculated based on the dynamic time warping distance, with the norm 'Eucliean distance' and the step pattern 'symmetric 2'. (b) represents for each cluster the concatenation of the temporal evolution of six normalised populations genetics indices, in the following order: (1) $\beta of Pareto$, (2) $\bar{r}_D$, (3) $Mean\ F_{IS}$, (4) $Variance\ F_{IS}$, (5) F_k, (6) $Temporal\ F_{ST}$.

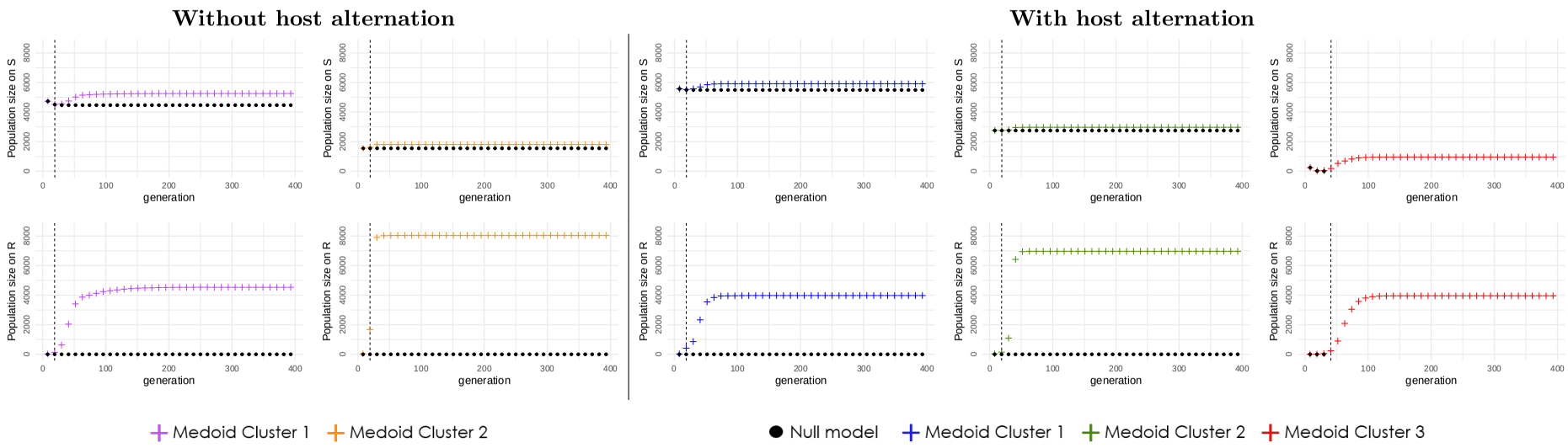
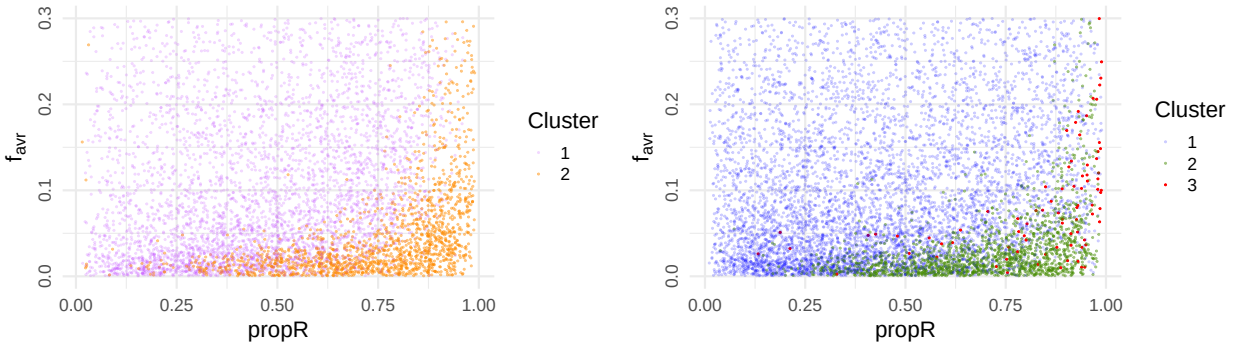


Figure S3: Comparison of the temporal evolution of population sizes, between medoid simulations and null models. The plotted results correspond to the medoid dynamic of each cluster, with ('Medoid', cross shapes) or without ('Null model', round shapes) virulent allele, that is, with or without resistance overcoming.



(a) Without host alternation

(b) With host alternation

Figure S4: Distribution of *propR* and f_{avr} input parameters for each simulations of both life cycles. Each point represents a simulation of the random simulation design that led to resistance overcoming, coloured by its cluster assignment.

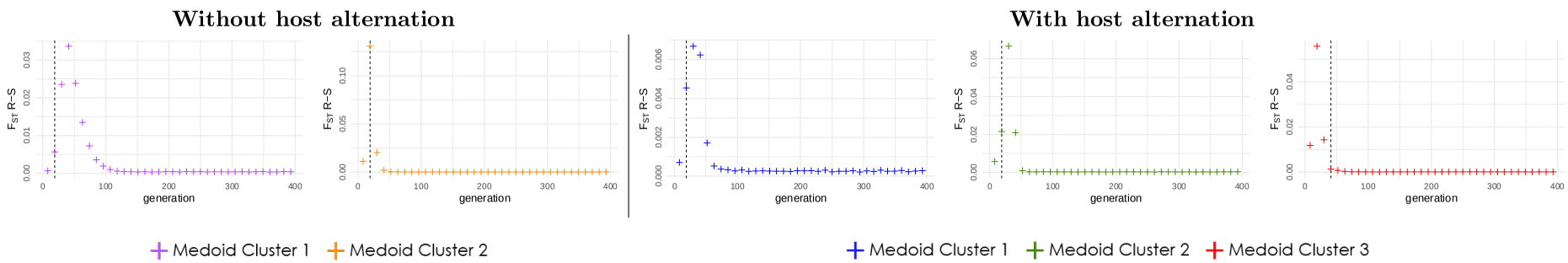


Figure S5: Temporal evolution of $F_{ST}\ R-S$. The plotted results correspond to the medoid dynamics for each cluster and each life cycle. For each of the plotted medoid, the vertical dashed line corresponds to the generation of resistance overcoming.

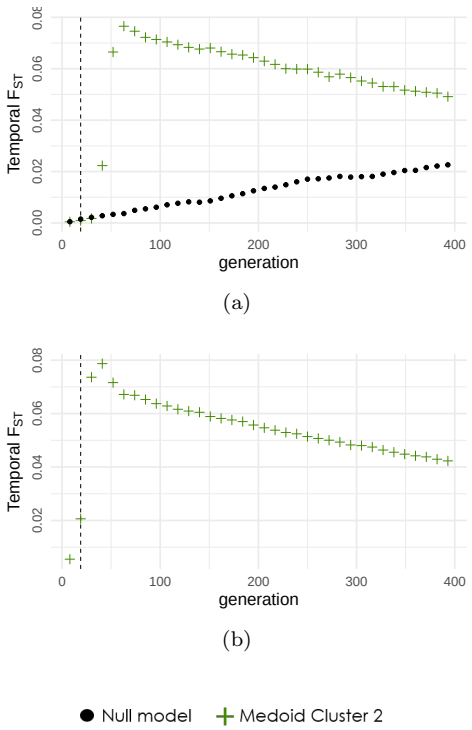


Figure S6: Comparison of the medoid dynamics of *Temporal* F_{ST} on (a) S and (b) R for Cluster 2 with host alternation. Subfigure (a) is already presented in the main document (Figure 6, with host alternation, Cluster 2).

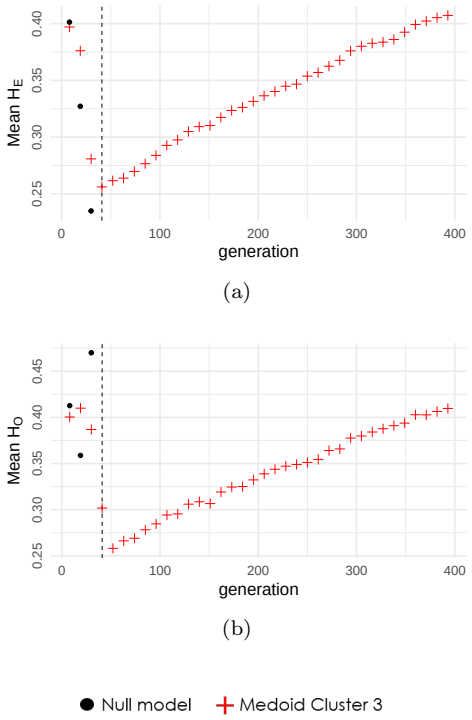


Figure S7: Comparison of the medoid dynamics of (a) $Mean\ H_E$ and (b) $Mean\ H_O$ on S for Cluster 3 with host alternation. Subfigure (a) is already presented in the main document (Figure 6, with host alternation, Cluster 3).