Spatially structured eco-evolutionary dynamics in a host-pathogen interaction render isolated populations vulnerable to disease

Layla Höckerstedt[1*], Elina Numminen[1*], Ben Ashby[2,3*], Mike Boots[2,4], Anna Norberg[5] and Anna-Liisa Laine[1,5]

*These authors contributed equally

[1]Organismal and Evolutionary Biology Research Program, Faculty of Biological and Environmental Sciences, 00014 University of Helsinki, Finland

[2]Department of Integrative Biology, University of California, Berkeley, CA, 94720 USA

[3]Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK

[4]Biosciences, University of Exeter, Penryn, TR10 9EZ, UK

[5]Department of Evolutionary Biology and Environmental Studies, University of Zürich CH-8057 Zurich, Switzerland

Corresponding author: anna-liisa.laine@uzh.ch

Keywords: Disease biology, coevolution, connectivity, host-parasite interactions, metapopulation ecology, natural populations, pathogen-imposed selection

Abstract

While the negative effects that pathogens have on their hosts are well-documented in humans and agricultural systems, direct evidence of pathogen-driven impacts in wild host populations is scarce and mixed. In particular, theory predicts that both ecological and evolutionary outcomes are shaped by the spatial structure of the interaction, yet comprehensive spatio-temporal data from nature to test this are scarce. Here, to determine how the strength of pathogen-imposed selection depends on spatial structure, we analyse growth rates across approximately 4000 host populations of a perennial plant through time coupled with data on pathogen presence-absence. We find that infection has the most devastating effect on population growth in isolated as opposed to connected host populations. Our inoculation study reveals isolated populations to be highly susceptible to the pathogen while connected host populations support the highest levels of resistance diversity, regardless of their disease history. A spatial eco-evolutionary model predicts that non-linearity in the costs to resistance may be critical in determining this pattern. Our results show that evolutionary feedbacks may define the ecological impacts of disease in spatially structured wild populations with host gene-flow more important than disease history in determining the outcome.

Main text

According to coevolutionary theory, hosts may evolve resistance under pathogen-imposed negative frequency-dependent selection (NFDS), whereby rare host genotypes have an advantage over the common ones[1,2]. The underlying assumptions of coevolutionary theory are the strong negative fitness effect of infection, with disease-free individuals outperforming infected ones[3], and costs of resistance that are central to maintenance of polymorphism within populations[4]. While consistent

negative effects of pathogens on their host populations are well documented in humans and agricultural systems[5,6], direct evidence of pathogen-driven ecological and evolutionary change in the wild is scarce and mixed [3,7–11]. The theoretical expectation is that the selective importance of diseases is directly correlated with the frequency and severity of epidemics [12]. However, our ability to quantify the strength of pathogen-imposed selection in natural populations is limited by few available systematic spatio-temporal data on pathogen occurrence across a sufficient number of host populations.

Spatial structure and heterogeneity supported by natural host populations is in stark contrast to human-managed systems that are typically highly conductive to disease transmission due to large population sizes, high densities and low genetic variability [13]. Not surprisingly, studies focusing on wild pathosystems have revealed highly variable disease prevalence levels. Moreover, local pathogen populations are typically ephemeral, persisting regionally as metapopulations through extinction and colonization events of local host populations [14–17]. Even when infection takes place, the fitness consequences – and the coevolutionary outcomes [18] – may vary depending on the genetic composition of the host and pathogen populations and their environment, either directly or via $\text{Genotype}^{\text{HOST}}$ x $\text{Genotype}^{\text{PATHOGEN}}$ x Environment –interactions [19,20]. Moreover, hosts in wild populations may suffer increased mortality or reduced reproduction irrespective of their infection status due to other factors such as extreme weather [21]. Hence, remarkably little is understood of how pathogens impact the fitness of their host populations in the wild.

There is increasing evidence that host-pathogen dynamics, both epidemiological and evolutionary, may be shaped by the spatial structure of the interaction [13,22–24]. Encounter rates between hosts and their pathogens are expected to be heavily influenced by connectivity to other populations, and the key metapopulation processes - gene flow, extinction, and colonization dynamics - are expected to contribute to the genetic structure of both the colonization dynamics, and the arrival of novel genetic variation into local populations [13]. As long as rates of migration are

low enough to not homogenize local populations, increasing immigration is expected to increase the diversity and evolutionary potential of both host and pathogen populations [25]. While measuring migration rates in natural populations is difficult [26], population connectivity, measured as the Euclidian distances separating populations and calibrated by the species dispersal capacity, provides a powerful proxy for migration rates [27]. Consequently, spatially structured eco-evolutionary feedback dynamics may emerge, with diversity accumulating in the well-connected populations. In line with this, there is evidence of spatial structure strongly influencing how resistance is distributed, with higher resistance observed in host populations that experience higher rates of gene flow[16,28,29]. To date, it has not been established what the relative roles of gene flow vs. pathogen-imposed selection are – and how they may vary in space - in generating spatially variable patterns of resistance that have been empirically observed[16,28,29].

Here, we combine a spatial analysis of a wild host-pathogen populations with an inoculation experiment, and a simulation model to understand how the ecological and evolutionary impacts of disease on host resistance vary in spatially structured populations. Specifically we ask: 1) Is there evidence of pathogen-imposed selection on its host populations across a large, naturally fragmented host-pathogen metapopulation; 2) Does host population resistance structure, measured through an inoculation assay, reflect variable selection pressure indicated by the spatial analysis; and 3) Using a coevolutionary metapopulation model we explore how gene flow, selection and costs of resistance contribute to the spatial structure of resistance detected with our empirical approach.

Our analysis is focused on annually recorded population size data from some ~ 4000 locations of host plant *Plantago lanceolata*, and the presence-absence dynamics of its obligate fungal pathogen, *Podosphaera plantaginis*, in this host population network in the Åland islands, South-Western Finland. *Plantago lanceolata* is a perennial that produces wind-dispersed pollen, while seeds typically drop close to the mother plant. During the epidemic season, *P. plantaginis*

96 disperses via clonally produced conidial spores that typically land within close proximity of the
97 infected source plant (REF). The visually conspicuous symptoms caused by *P. plantaginis* enable
98 accurate tracking of infection in the wild. Long-term epidemiological data have demonstrated this
99 pathogen to persist as a highly dynamic metapopulation with frequent extinctions and
100 (re)colonizations of local populations[16]. These data allow us to study whether the extent of
101 pathogen-imposed selection depends on host population connectivity (S^H) and hence, evolutionary
102 potential governed by gene flow, and whether resistance level and diversity vary among host
103 populations depending on their degree of connectivity and disease history. Previous metapopulation
104 models[30,31] have demonstrated the existence of overall higher resistance in well-connected
105 populations. To better understand the mechanisms that lead to the significant interaction between
106 population connectivity, infection history and resistance in our inoculation study, we built a host-
107 pathogen coevolutionary metapopulation model, where we examine how different trade-off
108 relationships impact the outcome. The model is not intended to be a replica of an empirical
109 metapopulation, but rather is used to reveal the key factors which lead to qualitatively similar
110 distributions of resistance and disease incidences observed in the study of the Åland islands. Hence,
111 the purpose of the model is to determine which biological factors are likely to be crucial to the
112 patterns observed herein.

113

114 **Results**

115 We used Spatial Bayesian modelling (Integrated Nested Laplace Approximation; INLA[32]) to
116 analyze how changes in host population size are influenced by the pathogen. To assess whether this
117 depends on host population connectivity, we estimated the separate effects of pathogen
118 presence/absence in the previous year for connectivity categories - high-, low and intermediate –
119 that were based on the 0.2 and 0.8 quantiles of the host-connectivity values (Supplementary Fig. 1).

The model controls for spatio-temporal autocorrelation characteristics of spatial ecological data, that may be due to unmeasured variables, thereby providing a conservative estimate of the model parameters (Supplementary Table 1)[32].

Infection by *P. plantaginis* had a negative effect on the growth of its host populations. All the estimated mean effects of pathogen presence were smaller than the effects with pathogen absence within the same connectivity category, suggesting an overall negative effect of the pathogen on host-population change (Fig. 1A, Supplementary Table 1). Furthermore, the estimated mean effects of the pathogen within the connectivity categories supports the interpretation that the relative effect of the pathogen on population growth is most negative in the isolated host populations (Fig. 1A, Supplementary Table 1). The posterior uncertainty in the effects of pathogen on the population growth (indicated by the confidence intervals in Fig. 1A) are due to the nature of observational data: pathogen infections were rare at the metapopulation level in studied years, thus there is considerably more pathogen absence observations in these data (See supplementary Table 2). The temporal autocorrelation in growth in *P. lanceolata* populations between consecutive years was estimated to be negative (Supplementary Table 1), indicating that local populations exhibit oscillatory dynamics, such that growth in one year is typically followed by a decline in the next year and vice versa. As many of the populations are well-established, these fluctuations could result from populations oscillating around their carrying capacities, dictated by the space and resources available for their growth. The estimated median effects for rainfall in July and August suggest that host population changes are not strongly driven by these effects, although the August rainfall had a slight positive effect on population growth (posterior mean effect 0.03, confidence interval -0.06, 0.12, Fig. 1B, Supplementary Table 1). The proportion of plants expressing drought symptoms in the previous year was significantly associated with a decline in host population size (posterior mean effect -0.38, confidence interval -0.43, -0.33, Fig. 1C, Supplementary Table 1).

144 To examine whether the diversity and level of resistance vary among host populations
145 depending on their degree of connectivity (S^H) and disease history (measured as infection status in
146 years 2001-2014), we performed an inoculation assay to characterize resistance phenotypes in
147 selected 19 *P. lanceolata* populations against four strains of *P. plantaginis*. These populations occur
148 in different locations of the host network, and were selected to represent both isolated and well-
149 connected populations. Our inoculation study confirmed that host plants varied in their resistance
150 against the tested powdery mildew strains (Table 1, Fig. 2A). We were able to identify all 16
151 possible resistance phenotypes in the sample of 190 plants (Fig. 2A). In the connected populations,
152 we found a greater diversity of different phenotypes, while isolated populations hosted fewer
153 resistance phenotypes (Fig. 2A). Both the Shannon diversity index (Table 1, Fig. 2B), and the
154 average level of resistance (Table 1, Fig. 2C), were higher in the well-connected than in the isolated
155 host populations (Table 1, Figs. 2B and C).

156 However, while disease history had no direct effect on phenotypic diversity nor the
157 level of resistance, we found a significant interaction between population connectivity and infection
158 history for both Shannon's diversity index and level of resistance (Table 1, Figs. 2B and C). The
159 highest diversity of phenotypes and highest resistance was measured in well-connected populations
160 without any history of disease. In contrast, in isolated populations, we found greater diversity of
161 resistance phenotypes and higher resistance in populations with a history of infection (Figs. 2B and
162 C).

163 We modeled both the ecological and coevolutionary dynamics of host and pathogen
164 metapopulations by constructing the network in two stages to account for relatively well and poorly
165 connected demes (see methods). We modeled the genetics of the system using a multilocus gene-
166 for-gene framework[33] with haploid host and pathogen genotypes characterised by L biallelic loci,
167 where 0 and 1 represent the presence and absence, respectively, of resistance and infectivity alleles.
168 Hosts and pathogen with more resistance or infectivity alleles are assumed to pay higher fitness

169 costs, as defined in the methods. We ran 200 simulations for each of the parameter sets described in
170 Supplementary Table 3 (example simulation dynamics are shown in Fig. 3D-F). On average,
171 disease prevalence (D), resistance (R) and infectivity (I) were always higher in well-connected than
172 in poorly connected populations regardless of metapopulation structure, transmissibility of the
173 pathogen, or the nature of the trade-offs (Supplementary Table 3). However, the difference between
174 well and poorly connected populations was generally greater when: (1) the metapopulation structure
175 was assortative (i.e. well connected populations are more likely to be connected to other well
176 connected populations than by chance) than random; (2) the pathogen was more transmissible; or
177 (3) host resistance was associated with fitness costs that diminish as resistance increases (i.e. costs
178 of resistance decelerate, $c_H^2 < 0$) (Supplementary Table 4). Overall, we found that the pattern of the
179 empirical results shown in Fig. 2C was most likely to occur when host resistance is associated with
180 diminishing fitness costs and is more likely for transient (Fig. 3B) than long-term dynamics (Fig.
181 3C).

182

183 **Discussion**

184 Here we show, to our knowledge for the first time, that the negative effect of pathogens on their
185 wild host populations depends on spatial structure. This finding suggests that the strength of
186 pathogen-imposed selection may vary across space in a predictable manner. Overall, finding a
187 consistent negative effect of infection on host population growth is noteworthy given the myriad
188 ecological factors that may hamper our ability to quantify costs of infection in wild populations[34].
189 The effect of infection on host population growth was the least negative in well-connected host
190 populations, while isolated host populations were most vulnerable to infection, suggesting that they
191 lack resistance diversity to effectively counter pathogen attack. Indeed, results of the inoculation
192 study confirmed that both the diversity and the average level of resistance were higher in the well-

193 connected than in the isolated host populations. When the interaction is characterized by strain-
194 specific resistance such as in the interaction between *P. lanceolata* and *P. plantaginis*, resistance
195 diversity will reduce the probability of establishment by an immigrant pathogen strain, and slow
196 down the spread of established strains due to a mismatch between the specific avirulence alleles of
197 pathogen and resistance alleles of host[35]. In agriculture, even slight additions of diversity to
198 monocultures have been shown to reduce disease levels significantly[36,37].

199 Theory predicts that pathogens maintain resistance polymorphism in their host
200 populations[38–40]. As described above, our spatial statistical population model demonstrated that the
201 isolated populations went through the strongest reductions in size - most likely through increased
202 mortality of infected individuals[41] - which could lead to selection increasing in the frequency of
203 resistant phenotypes locally[42]. Accordingly, in the isolated populations we measured higher
204 resistance diversity in host populations with a history of infection than in host populations that had
205 not been infected in the past. The effect of infection on host population growth rates in the well-
206 connected populations was much weaker, and hence, may explain why we did not detect signs of
207 past selection in these populations. The resulting differences in resistance among host populations is
208 in line with previous studies that have measured higher resistance levels in well-connected host
209 populations[16,28,29]. Jointly our results reveal that this pattern is generated by eco-evolutionary
210 feedback resulting from spatial differences in how gene flow *vs.* selection drive host-pathogen
211 dynamics in the in the wild - In the well-connected populations gene flow appears more important
212 than pathogen-imposed selection in maintaining resistance diversity.

213 In theory, polymorphism in resistance within populations is maintained by costs of
214 resistance in the absence of the pathogen, whereas under pathogen attack, the resistant hosts
215 outperform the susceptible ones[4]. Hence, finding high levels of resistance diversity where pathogen
216 impact has recently been negligible may appear contrary to expectations, and suggests dispersal to
217 be critical for maintaining variation within host populations. Our metapopulation model explored

218 scenarios under which spatial structure, disease dynamics and life-history trade-offs could yield
219 similar outcomes. We find that the shape of the host trade-off was the critical predictor of whether
220 the simulations would qualitatively match the empirical results. Our results suggest that the costs of
221 resistance are most likely to diminish as resistance increases. Diminishing costs mean that there is
222 an initial large cost associated with resistance and therefore it is less beneficial when disease is rare.
223 While fitness costs associated with resistance have been widely identified across many plant species
224 {REFS} and other taxa {REFS}, determining the shape of trade-offs from empirical data is
225 challenging, especially when trade-offs are close to linear or vary with environment, and it is
226 impossible to determine trade-off shapes when only two host phenotypes are compared (as is often
227 the case). However, experimental evolution of bacteria and phages has demonstrated that
228 decelerating costs of resistance are possible {REF}. In addition, our simulations suggest that the
229 pattern detected in the empirical results is most likely to occur prior to the system reaching
230 equilibrium and when metapopulation connectivity is assortative. The fact that the transient
231 simulations dynamics tend to provide a better qualitative match to the empirical results does not
232 imply that the resistance patterns detected in the archipelago will necessarily fade in the long-term
233 (many simulations were qualitative matches at equilibrium), although our model indicates that this
234 is a possibility. We think that it is interesting to note that the patterns we see are found for a wider
235 range of parameter values under transient dynamics, but we get the same inference of the key
236 characteristics that lead to the patterns we see. Whether or not the patterns are only transient is an
237 empirical question.

238 Together, our results show how spatial fragmentation leading to isolation of host
239 populations drives the loss of diversity and increases host vulnerability to infectious diseases. To
240 our knowledge this is the first empirical demonstration of how spatial structure generates variation
241 in the strength of pathogen-imposed selection, and thus provides a compelling example of how
242 landscape fragmentation drives epidemiological and coevolutionary processes in nature.

ACKNOWLEDGEMENTS

We would like to acknowledge Krista Raveala and Niko Vilenius for their assistance during the experimental work and all students who participated in annual metapopulation surveys. This work was funded by grants from the Academy of Finland (334276), and the European Research Council (Consolidator Grant RESISTANCE 724508) and SNF (310030_192770/1) to A-LL, and LUOVA Doctoral Programme funding to LH. MB acknowledges the Natural Environment Research Council (NE/J009784/1), NIH/R01-GM122061-03 and NSF-DEB- 2011109 for support. BA is supported by the Natural Environment Research Council (grant no. NE/N014979/1).

AUTHOR CONTRIBUTION

A-LL, EN, MB and LH conceived the ideas and designed the assay; LH conducted the experimental work and EN LH, and AN analyzed the data. MB and BA developed and analysed the simulation model. A-LL and MB wrote the first draft of the manuscript. All the authors contributed to the writing of the manuscript and approved the final draft.

METHODS

MATERIALS AND METHODS

The pathosystem

Plantago lanceola L. is a perennial monoecious ribwort plantain that reproduces both clonally via the production side rosettes, and sexually via wind pollination. Seeds drop close to the mother plant and usually form a long-term seed bank[43]. *Podospharea plantaginis* (Castagne; U. Braun and S.

265 Takamatsu) (*Erysiphales*, Ascomycota) is an obligate biotrophic powdery mildew that infects only
266 *P. lanceolata* and requires living host tissue through its life cycle[44]. It completes its life cycle as
267 localized lesions on host leaves, only the haustorial feeding roots penetrating the leaf tissue to feed
268 nutrients from its host. Infection causes significant stress for host plant and may increase the host
269 mortality[41]. The interaction between *P. lanceolata* and *P. plantaginis* is strain-specific, whereby the
270 same host genotype may be susceptible to some pathogen genotypes while being resistant to
271 others[45]. The putative resistance mechanism includes two steps. First, resistance occurs when the
272 host plant first recognizes the attacking pathogen and blocks its growth. When the first step fails
273 and infection takes place, the host may mitigate infection development. Both resistance traits vary
274 among host genotypes[45].

275 Approximately 4000 *P. lanceolata* populations form a network covering an area of 50
276 x 70 km in the Åland Islands, SW of Finland. Disease incidence (0/1) in these populations has been
277 recorded systematically every year in early September since 2001 by approximately 40 field
278 assistants, who record the occurrence of the fungus *P. plantaginis* in the local *P. lanceolata*
279 populations[46]. At this time, disease symptoms are conspicuous as infected plants are covered by
280 white mycelia and conidia. The coverage (m^2) of *P. lanceolata* in the meadows was recorded
281 between 2001-2008 and is used as an estimate of host population size. The proportion of *P.*
282 *lanceolata* plants in each population suffering from drought is also estimated annually in the survey.
283 Data on average rainfall (mm) in July and August was estimated separately for each population
284 using detailed radar-measured rainfall (obtained by Finnish Meteorological Institute) and it was
285 available for years 2001-2008.

286 Host population connectivity (S^H)[27] for each local population *i* was computed with the
287 formula that takes into account the area of host coverage (m^2) of all host populations surveyed,
288 denoted with (A_j), and their spatial location compared to other host populations. We assume that
289 the distribution of dispersal distances from a location are described by negative exponential

distribution. Under this assumption, the following formula quantifies for a focal population *i*, the effect of all other host populations, taking into account their population sizes and how strongly they are connected through immigration to it:

$$S_i^H = \sum_{j \neq i} e^{-\alpha d_{ij}} \sqrt{A_j}.$$

Here, d_{ij} is the Euclidian distance between populations *i* and *j* and $1/\alpha$ equals the mean dispersal distance, which was set to be two kilometers based on results from a previous study[16].

The annual survey data has demonstrated that *P. plantaginis* infects annually 2-16% of all host populations and persists as a highly dynamic metapopulation through extinctions and re-colonizations of local populations[16]. The number of host populations has remained relatively stable over the study period[45]. The first visible symptoms of *P. plantaginis* infection appear in late June as white-greyish lesions consisting of mycelium supporting the dispersal spores (conidia). Six to eight clonally produced generations follow one another in rapid succession, often leading to local epidemic with substantial proportion of the infected hosts by late summer within the host local population. *Podosphaera plantaginis* produces resting structures, chasmothecia, that appear towards the end of growing season in August-September[41]. Between 20-90 % of the local pathogen populations go extinct during the winter, and thus the recolonization events play an important role in the persistence of the pathogen regionally[16].

Inoculation assay: Effect of connectivity and disease history on phenotypic disease resistance

Host and pathogen material for the experiment

To examine whether the diversity and level of resistance vary among host populations depending on their degree of connectivity (S^H) and disease history, we selected 20 *P. lanceolata* populations for an inoculation assay. These populations occur in different locations in the host network, and were

314 selected based on their connectivity values (S^H of selected populations was 37-110 in isolated and
315 237-336 in highly connected category, Fig. 1). We did not include host populations in the
316 intermediate connectivity category that was used in the population dynamic analyses (plese see
317 above) in the icoculation assay due to logistic cnostraints. *Podosphaera plantaginis* is an obligate
318 biotrophic pathogen that requires living host tissue throughout its life cycle, and obtaining sufficient
319 inoculum for experiments is extremely time and space consuming. In both isolated and highly
320 connected categories, half of the populations (IDs 193, 260, 311, 313, 337, 507, 1821, 1999, 2818,
321 5206) were healthy during the study years 2001-2014, while half of the populations (IDs 271, 294,
322 309, 321, 490, 609, 1553, 1556, 1676, 1847) were infected by *P. plantaginis* for several years
323 during the same period. We collected *P. lanceolata* seeds from randomly selected ten individual
324 plants around the patch area from each host population in August 2014.

325 To acquire inoculum for the assay, we collected the pathogen strains as infected leaves,
326 one leaf from ten plant individuals from four additional host populations (IDs 3301, 4684, 1784 and
327 3108) in August 2014. None of the pathogen populations were same as the sampled host
328 populations and hence, the strains used in the assay all represent allopatric combinations. Both host
329 and pathogen populations selected for the study were separated by at least two kilometers. The
330 collected leaves supporting infection were placed in Petri dishes on moist filter paper and stored at
331 room temperature until later use.

332 Seeds from ten mother plants from each population were sown in 2:1 mixture of potting
333 soil and sand, and grown in greenhouse conditions at 20±2 °C (day) and 16 ±2 °C (night) with 16:8
334 L:D photoperiod. Due to the low germination rate of collected seeds, population 260 (isolated and
335 healthy population) was excluded from the study. Seedlings of ten different mother plants were
336 randomly selected among the germinated plants for each population (n=190), and grown in
337 individual pots until the plants were eight weeks old.

338 The pathogen strains were purified through three cycles of single colony inoculations and
339 maintained on live, susceptible leaves on Petri dishes in a growth chamber 20±2 °C with 16:8 L:D
340 photoperiod. Every two weeks, the strains were transferred to fresh *P. lanceolata* leaves. Purified
341 powdery mildew strains (M1-M4), one representing each allopatric population (3301, 4684, 1784
342 and 3108), were used for the inoculation assay. To produce enough sporulating fungal material,
343 repeated cycles of inoculations were performed before the assay.

344

345 *Inoculation assay quantifying host resistance phenotypes*

346 In order to study how the phenotypic resistance of hosts varies depending on population
347 connectivity and infection history, we scored the resistance of 190 host genotypes, ten individuals
348 from each study populations (n=19), in an inoculation assay. Here, one detached leaf from each
349 plant was exposed to a single pathogen strain (M1-M4) by brushing spores gently with a fine
350 paintbrush onto the leaf. Leaves were placed on moist filter paper in Petri dishes and kept in a
351 growth chamber at 20±2 with a 16/8D photoperiod. All the inoculations were repeated on two
352 individual Petri plates, leading to 760 host genotype – pathogen genotype combinations and a total
353 of 1520 inoculations and (19 populations * 10 plant genotypes * 4 pathogen strains * 2 replicates).
354 We then observed and scored the pathogen infection on day 12 post inoculation, under dissecting
355 microscope. The resulting plant phenotypic response was scored as 0 = susceptible (infection) when
356 mycelium and conidia were observed on the leaf surface, and as 1 = resistance (no infection), when
357 no developing lesions could be detected under a dissecting microscope. A genotype was defined
358 resistant only if both inoculated replicates showed similar response (1), and susceptible if one or
359 both replicates became infected (0).

360

361 **Statistical analyses**

Bayesian spatio-temporal INLA model of the changes in host population size

To study how the pathogen infection influences on host population growth, we analyzed the relative change in host population size (m2) (defined as population size (*t)* - population size (*t-1)*) / population size (*t-1)*) between consecutive years utilizing data from 2001-2008 in response to pathogen presence-absence status at *t-1* (Supplementary Table 2). To assess whether this depends on host population connectivity, we estimated the separate effects of pathogen presence/absence in the previous year for connectivity categories - high-, low and intermediate – that were based on the 0.2 and 0.8 quantiles of the host-connectivity values (Fig. 1A, Supplementary Fig.1).

As covariates, we included the proportion (0-100%) of dry host plants measured each year within each local population as well as data on the amount of rainfall at the summer months (June, July, August) obtained from the satellite images, as these were suggested be relevant for this pathosystem in an earlier analysis[16]. Observations where the change in host population size, or the host population coverage had absolute values larger than their 0.99 quantiles in the whole data, were regarded as outliers and omitted from the analysis. Before the analyses, all the continuous covariates were scaled and centered, and the categorical variables were transformed into binary variables.

The relative changes in local host population size between consecutive years was analyzed by a Bayesian spatio-temporal statistical model that simultaneously considers the effects of a set of biologically meaningful predictors. The linear predictor thus consists of two parts:

1) $\beta X_t + z_t A_t$

where β represents the correlation coefficients corresponding to the effects of environmental covariates, z_t corresponds to the spatiotemporal random effect, and X_t and A_t project these to the

observation locations. For z_t we assume that the observations from a location in consecutive time points (t-1) and t are described by 1st order autoregressive process:

$$2) \qquad z_t = \varphi z_{t-1} + w_t$$

where w_t corresponds to spatially structured zero-mean random noise, for which a Matern covariance function is assumed. Statistical inference then targets jointly the covariate effects β, the temporal autocorrelation φ, and the hyperparameters describing the spatial autocorrelation in w_t. From these the overall variance, as well as spatial range, a distance after which spatial autocorrelation ceases to be significant, can be inferred, Supplementary Fig. 3). For more detailed description of the structure of the statistical model and how to do efficient inference with it using R-INLA, we refer to[16,47].

Identification of resistance phenotypes

The phenotype composition of each study population was defined by individual plant responses to the four pathogen strains, where each response could be "susceptible = 0" or "resistant = 1". For example, a phenotype "1111" refers to a plant resistant to all four pathogen strains. The diversity of distinct resistance phenotypes within populations was estimated using the Shannon diversity index as implemented in the *vegan* software package [48]. The Shannon diversity index for all four study groups was then analyzed using a linear model with class predictors population type (well-connected or isolated), infection history (healthy or infected), and their interaction.

Analysis of population connectivity and infection history effects on host resistance

405 To test whether host population resistance varied depending on connectivity (S^H) and infection
406 history, we analyzed the inoculation responses (0=susceptible , 1=resistant) of each host-pathogen
407 combination by using a logit mixed-effect model in the *lme4* package[49]. The model included the
408 binomial dependent variable (resistance-susceptible; 1/0), and class predictors population type
409 (well-connected or isolated), infection history (healthy or infected), mildew strain (M1, M2, M3,
410 M4) and their interactions. Plant individual and population were defined as random effects, with
411 plant genotype (sample) hierarchically nested under population. Model fit was assessed using chi-
412 square tests on the log-likelihood values to compare different models and significant interactions,
413 and the best model was selected based on AIC-values. *P*-values for regression coefficients were
414 obtained by using the *car* package [50]. We ran all the analyses in R software [51].

415

416 **The metapopulation model**

417 We model the ecological and co-evolutionary dynamics of host and pathogen metapopulations. We
418 construct the metapopulations in two stages to account for relatively well and poorly connected
419 demes. All demes are identical in quality (i.e. no differences in intrinsic birth or death rates between
420 demes) and only differ in their connectivity. Our metapopulation consists of an outer network of 20
421 demes, equally spaced around the unit square (0.2 units apart), and a 7×7 inner lattice of demes at a
422 minimum distance of 0.2 units from the outer network (Fig. 3A), giving a total of 69 demes. Demes
423 that are separated by a Euclidean distance of at most 0.2 are then connected to each other. This
424 means that populations near the centre of the metapopulation are highly connected, while those on
425 the boundary of the metapopulation are poorly connected. This also has the effect of making
426 connections between well and poorly connected demes assortative (i.e. well/poorly connected
427 demes tend to be connected to well/poorly connected demes). We relax the assumption of
428 assortativity in a second type of network by randomly reassigning connections between demes,

429 while maintaining the same degree distribution. (i.e. the probability of two demes being connected
430 is proportionate to their degree). While well connected demes still have more connections to other
431 well connected demes than to poorly connected demes, they are not more likely to be connected to a
432 well connected deme than by chance based on the degree distribution. In both types of network
433 structure, we classify a deme as well-connected if it is in the top 20% of the degree distribution and
434 poorly connected if it is in the bottom 20%.

435 We model the genetics using a multilocus gene-for-gene framework with haploid host
436 and pathogen genotypes characterised by L biallelic loci, where 0 and 1 represent the presence and
437 absence, respectively, of resistance and infectivity alleles. Host genotype i and pathogen genotype j
438 are represented by binary strings: $x_i^1 x_i^2 \dots x_i^L$ and $y_j^1 y_j^2 \dots y_j^L$. Resistance acts multiplicatively such
439 that the probability of host i being infected when challenged by pathogen j is $Q_{ij} = \sigma^{d_{ij}}$, where σ is
440 the reduction in infectivity per effective resistance allele and $d_{ij} = \sum_{k=1}^{L} x_i^k (1 - y_j^k)$ is the number
441 of effective resistance alleles (i.e. the number of loci where hosts have a resistance allele but
442 pathogens do not have a corresponding infectivity allele). Hosts and pathogens with more resistance
443 or infectivity alleles are assumed to pay higher fitness costs, $c_H(i)$ and $c_P(j)$, with:

444

445 $$c_H(i) = c_H^1 \left(\frac{1 - e^{\frac{c_H^2}{L} \sum_{k=1}^{L} x_i^k}}{1 - e^{c_H^2}} \right)$$

446 and

447 $$c_P(j) = c_P^1 \left(\frac{1 - e^{\frac{c_P^2}{L} \sum_{k=1}^{L} y_j^k}}{1 - e^{c_P^2}} \right)$$

448 where $0 < c_H^1, c_P^1 \leq 1$ control the overall strength of the costs (i.e. the maximum proportional
449 reduction in reproduction (hosts) or transmission rate (pathogens)) and $c_H^2, c_P^2 \in \mathbb{R}_{\neq 0}$ control the

shape of the trade-off. When $c_H^2, c_P^2 < 0$ the costs decelerate (increasing returns) and when $c_H^2, c_P^2 > 0$ the costs accelerate the costs accelerate (decreasing returns). This formulation therefore allows for a wide-range of trade-off shapes that may occur in nature.

The dynamics of the (finite) host and pathogen populations are modelled stochastically using the tau-leap method with a fixed step size of $\tau = 1$. For population p, the mean host birth rate at time t for host i is

$$B_i^p(t) = \left(a\left(1 - c_H(i)\right) - qN_p(t)\right) S_i^p(t)$$

where a is the maximum per-capita birth rate, q is the strength of density-dependent competition on births, $N_p(t) = S_i^p(t) + I_{i\circ}^p(t)$ is the local host population size, $S_i^p(t)$ and $I_{i\circ}^p(t) = \sum_{j=1}^n I_{ij}^p(t)$ are the local sizes of susceptible and infected individuals of genotype i, and $I_{ij}^p(t)$ is the local size of hosts of genotype i infected by pathogen j. Host mutations occur at an average rate of μ_H per loci (limited to at most one mutation per time step), so that the mean number of mutations from host type i to i' is $\mu_H m_{ii'} B_i^p(t)$, where $m_{ii'} = 1$ if genotypes i and i' differ at exactly one locus, and is 0 otherwise.

The mean local mortalities for susceptible and infected individuals are $bS_i^p(t)$ and $(b + \alpha)I_{ij}^p(t)$, respectively, where b is the natural mortality rate and α is the disease-associated mortality rate. The average number of infected hosts that recover is $\gamma I_{ij}^p(t)$, where γ is the recovery rate.

The mean number of new local infections of susceptible host type i by pathogen j is:

$$INF_{ij}^p(t) = \beta\left(1 - c_P(j)\right) Q_{ij} S_i^p(t) Y_j^p(t)$$

where β is the baseline transmission rate and $Y_j^p(t)$ is the local number of pathogen propagules following mutation and dispersal. Pathogen mutations occur in a similar manner to host mutations,

with mutations from type j to j' occurring at rate $\mu_P m_{jj'} I_{\circ j}^p(t)$ where μ_P is the mutation rate per loci (limited to at most one mutation per timestep) and $I_{\circ j}^p(t) = \sum_{i=1}^n I_{ij}^p(t)$ is the local number of pathogen j. Following mutation, the local number of pathogens of type j is:

$$W_j^p(t) = I_{\circ j}^p(t)(1 - \mu_P L) + \mu_P m_{jj'} I_{\circ j}^p(t)$$

Pathogen dispersal occurs following mutation at a rate of ρ between connected demes, given by the adjacency matrix G_{pr}, with $G_{\Sigma p}$ the total number of connections for deme p. The mean local number of pathogen propagules following mutation and dispersal is therefore:

$$Y_j^p(t) = W_j^p(t)\left(1 - \rho G_{\Sigma p}\right) + \rho \sum_{r=1}^{M_\Sigma} G_{pr} W_j^r(t)$$

We focus our parameter sweep on: (i) the structure of the network (assortative or random connections); (ii) the strength (c_H^1, c_P^1) and shape (c_H^2, c_P^2) of the trade-offs; (iii) the transmission rate (β); and (iv) the dispersal rate (ρ), fixing the remaining parameters as described in Supplementary Table 1 (preliminary investigations suggested they had less of an impact on the qualitative outcome) and conducting 100 simulations per parameter set. For each simulation we initially seed all populations with the most susceptible host type and place the least infective pathogen type in one of the well-connected populations to minimise the risk of early extinction. We then solve the dynamics for 10,000 time steps (preliminary investigations indicated this was a sufficient period for the metapopulations to reach a quasi-equilibrium in terms of overall resistance). We calculate the average level of resistance (proportion of loci with a resistance allele) between time steps 4,001 and 5,000 (transient dynamics) and over the final 1,000 time steps (long-term dynamics) for well and poorly connected demes, categorised according to whether disease is present in (infected) or absent from (uninfected) the local population at a given time point and discarding simulations where the pathogen is driven globally extinct.

We compare the mean level of resistance in infected/uninfected poorly/well-connected populations across all simulations to the empirical results. We say that a simulation is a qualitative 'match' for the empirical findings if: (i) in poorly connected demes, the infected populations are on average at least 5% more resistant than uninfected populations; and (ii) in well-connected demes, the uninfected populations are on average at least 5% more resistant than infected populations. In other words, if R_{CS} is the mean resistance for a population with connectivity C ($C = W$ and $C = P$ for well and poorly connected demes, respectively) and infection status S ($S = U$ and $S = I$ for uninfected and infected populations, respectively), then a parameter set is a qualitative 'match' for the empirical findings if $R_{WU} > 1.05R_{WI}$ and $1.05R_{PI} > 1.05R_{PU}$. If these criteria are not met, then the parameter set is a qualitative 'mismatch' for the empirical findings.

Data and code availability

Data and code will be made available upon acceptance.

References

1. Haldane, J. Disease and evolution. *Ricerca Scientifica Supplemento A* **19**, 68–75 (1949).
2. Hamilton, W. D. Sex versus Non-Sex versus Parasite. *Oikos* **35**, 282–290 (1980).
3. Horns, F. & Hood, M. E. The evolution of disease resistance and tolerance in spatially structured populations. *Ecol Evol* **2**, 1705–1711 (2012).
4. Antonovics, J. & Thrall, P. H. The cost of resistance and the maintenance of genetic polymorphism in host—pathogen systems. *Proc. Royal Soc. B* **257**, 105–110 (1994).
5. Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control*. (Oxford University Press, 1992).

517 6. Savary, S. *et al.* The global burden of pathogens and pests on major food crops. *Nat Ecol Evol*
518 **3**, 430–439 (2019).
519 7. Fowler, N. L. & Clay, K. Environmental heterogeneity, fungal parasitism and the demography
520 of the grass Stipa leucotricha. *Oecologia* **103**, 55–62 (1995).
521 8. Gulland, F. M. The role of nematode parasites in Soay sheep (Ovis aries L.) mortality during a
522 population crash. *Parasitology* **105 (Pt 3)**, 493–503 (1992).
523 9. Hudson, P. J., Dobson, A. P. & Newborn, D. Prevention of Population Cycles by Parasite
524 Removal. *Science* **282**, 2256–2258 (1998).
525 10. Penczykowski, R. M., Walker, E., Soubeyrand, S. & Laine, A.-L. Linking winter conditions to
526 regional disease dynamics in a wild plant–pathogen metapopulation. *New Phytol* **205**, 1142–
527 1152 (2015).
528 11. Prendeville, H. R., Tenhumberg, B. & Pilson, D. Effects of virus on plant fecundity and
529 population dynamics. *New Phytol* **202**, 1346–1356 (2014).
530 12. Gillespie, J. H. Natural Selection for Resistance to Epidemics. *Ecology* **56**, 493–495 (1975).
531 13. Parratt, S. R., Numminen, E. & Laine, A.-L. Infectious Disease Dynamics in Heterogeneous
532 Landscapes. *Annu Rev Ecol Evol Syst* **47**, 283–306 (2016).
533 14. Ericson, L., Müller, W. J. & Burdon, J. J. 28-year temporal sequence of epidemic dynamics in a
534 natural rust–host plant metapopulation. *J Ecol* **105**, 701–713 (2017).
535 15. Grenfell, B. & Harwood, J. (Meta)population dynamics of infectious diseases. *Trends Ecol Evol*
536 **12**, 395–399 (1997).
537 16. Jousimo, J. *et al.* Disease ecology. Ecological and evolutionary effects of fragmentation on
538 infectious disease dynamics. *Science* **344**, 1289–1293 (2014).
539 17. Keeling, M. J. & Gilligan, C. A. Metapopulation dynamics of bubonic plague. *Nature* **407**, 903–
540 906 (2000).

541 18. MacPherson, A., Keeling, M. J. & Otto, S. P. Feedback between coevolution and epidemiology
542 can help or hinder the maintenance of genetic variation in host-parasite models. *Evolution* **75**,
543 582–599 (2021).
544 19. Mitchell, S. E., Rogers, E. S., Little, T. J. & Read, A. F. Host-parasite and genotype-by-
545 environment interactions: temperature modifies potential for selection by a sterilizing pathogen.
546 *Evolution* **59**, 70–80 (2005).
547 20. Wolinska, J. & King, K. C. Environment can alter selection in host-parasite interactions. *Trends*
548 *Parasitol* **25**, 236–244 (2009).
549 21. Bergen, E. van *et al.* The effect of summer drought on the predictability of local extinctions in a
550 butterfly metapopulation. *Conserv. Biol.* **34**, 1503–1511 (2020).
551 22. Boots, M. & Sasaki, A. 'Small worlds' and the evolution of virulence: infection occurs locally
552 and at a distance. *Proc Biol Sci* **266**, 1933–1938 (1999).
553 23. Gandon, S. & Michalakis, Y. Local adaptation, evolutionary potential and host–parasite
554 coevolution: interactions between migration, mutation, population size and generation time. *J*
555 *Evol Biol* **15**, 451–462 (2002).
556 24. Real, L. A. & Biek, R. Spatial dynamics and genetics of infectious diseases on heterogeneous
557 landscapes. *J R Soc Interface* **4**, 935–948 (2007).
558 25. Lion, S. & Gandon, S. Evolution of spatially structured host-parasite interactions. *J Evol Biol*
559 **28**, 10–28 (2015).
560 26. Bonte, D. *et al.* Costs of dispersal. *Biol Rev Camb Philos Soc* **87**, 290–312 (2012).
561 27. Hanski, I. *Metapopulation Ecology*. (Oxford University Press, 1999).
562 28. Carlsson-Granér, U. & Thrall, P. H. Host resistance and pathogen infectivity in host populations
563 with varying connectivity. *Evolution* **69**, 926–938 (2015).
564 29. Höckerstedt, L. M., Siren, J. P. & Laine, A. Effect of spatial connectivity on host resistance in a
565 highly fragmented natural pathosystem. *J Evol Biol* **31**, 844–852 (2018).

30. Damgaard, null. Coevolution of a plant host-pathogen gene-for-gene system in a metapopulation model without cost of resistance or cost of virulence. *J Theor Biol* **201**, 1–12 (1999).
31. Thrall, P. H. & Burdon, J. J. Evolution of gene-for-gene systems in metapopulations: the effect of spatial scale of host and pathogen dispersal. *Plant Pathol* **51**, 169–184 (2002).
32. Lindgren, F. & Rue, H. Bayesian Spatial Modelling with R-INLA. *J Stat Softw* **63**, 1–25 (2015).
33. Ashby, B. & Boots, M. Multi-mode fluctuating selection in host–parasite coevolution. *Ecol Lett* **20**, 357–365 (2017).
34. Hudson, P., Rizzoli, A., Grenfell, B., Heesterbeek, H. & Dobson, A. *The Ecology of Wildlife Diseases*. (Oxford University Press, 2002).
35. Dodds, P. & Thrall, P. Recognition events and host–pathogen co-evolution in gene-for-gene resistance to flax rust. *Funct Plant Biol* **36**, 395–408 (2009).
36. Mundt, C. C. Use of multiline cultivars and cultivar mixtures for disease management. *Annu Rev Phytopathol* **40**, 381–410 (2002).
37. Zhu, Y. *et al.* Genetic diversity and disease control in rice. *Nature* **406**, 718–722 (2000).
38. Boots, M. & Haraguchi, Y. The Evolution of Costly Resistance in Host-Parasite Systems. *Am Nat* **153**, 359–370 (1999).
39. Boots, M., White, A., Best, A. & Bowers, R. How Specificity and Epidemiology Drive the Coevolution of Static Trait Diversity in Hosts and Parasites. *Evolution* **68**, 1594–1606 (2014).
40. Taylor, R. L. The genus Lithophragma (Saxifragaceae). *Univ. Calif. Publ. Bot.* **37**, 1–89 (1965).
41. Laine, A.-L. Resistance variation within and among host populations in a plant–pathogen metapopulation: implications for regional pathogen dynamics. *J Ecol* **92**, 990–1000 (2004).
42. Laine, A.-L. Evolution of host resistance: looking for coevolutionary hotspots at small spatial scales. *Proc Royal Soc B* **273**, 267–273 (2006).

43. Bos, M. Plantago: A Multidisciplinary Study Vol. 89 Ecological Studies. in 222–231 (Springer-Verlag, 1992).
44. Bushnell, W., R. The powdery mildews - a comprehensive treatise. in 1–12 (The American Phytological Society, 2002).
45. Laine, A.-L. Pathogen fitness components and genotypes differ in their sensitivity to nutrient and temperature variation in a wild plant-pathogen association. *J Evol Biol* **20**, 2371–2378 (2007).
46. Ojanen, S. P., Nieminen, M., Meyke, E., Pöyry, J. & Hanski, I. Long-term metapopulation study of the Glanville fritillary butterfly (Melitaea cinxia): survey methods, data management, and long-term population trends. *Ecol Evol* **3**, 3713–3737 (2013).
47. Cameletti, M., Lindgren, F., Simpson, D. & Rue, H. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *Adv Stat Anal* **97**, 109–131 (2012).
48. Oksanen, J., Blanchet, G. & Kindt, R. Vegan: Community 711 Ecology Package. R package version 2.4-3. (2017).
49. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw* **67**, 1–48 (2015).
50. Fox, J. & Weisberg, S. *An R Companion to Applied Regression, Second Edition*. (SAGE Publications, 2011).
51. R Core Team. A Language and Environment for Statistical Computing. (2016).

613 **Tables and Figures**

614 **Table 1. The effects and effect sizes of connectivity and disease history on resistance diversity**
615 **(Shannon diversity), and the average level of resistance in the 19 studied *Plantago lanceolata***
616 **populations.** Statistics for minimum adequate models with smallest AIC values are reported.
617 Significant values are highlighted in bold.

Source (Shannon diversity)	**d.f.**	***F***	***P***
Connectivity	1	14.95	**0.001**
Disease history	1	1.61	0.2
Connectivity x Disease History	1	7.68	**0.01**
Shannon diversity coefficients		***Estimate***	***sd.***
Intercept		*1.85*	*0.13*
History (Infected)		*-0.18*	*0.18*
Connectivity (Isolated)		*-0.93*	*0.19*
*History (Infected) * Connectivity (Isolated)*		*0.76*	*0.27*
Source (Resistance)	**d.f.**	***X²***	***P***
Connectivity	1	16.55	**<0.0001**
Disease history	1	0.01	0.9
Connectivity x Disease History	1	9.91	**0.001**
Mildew strain	3	36.34	**<0.0001**
Random		**Variance**	**sd.**
Population		0.227	0.477
Sample (Population)		1.206	1.09
Resistance fixed effects		***Estimate***	***sd.***
Intercept		*0.5*	*0.34*
Connectivity (Isolated)		*-2.67*	*0.53*
History (Infected)		*-0.95*	*0.44*
Mildew_strain2		*-0.86*	*0.27*
Mildew_Strain3		*-0.6*	*0.26*
Mildew_strain4		*0.65*	*0.25*
*History (Infected) * Connectivity (Isolated)*		*2.17*	*0.69*

618

619

620

621

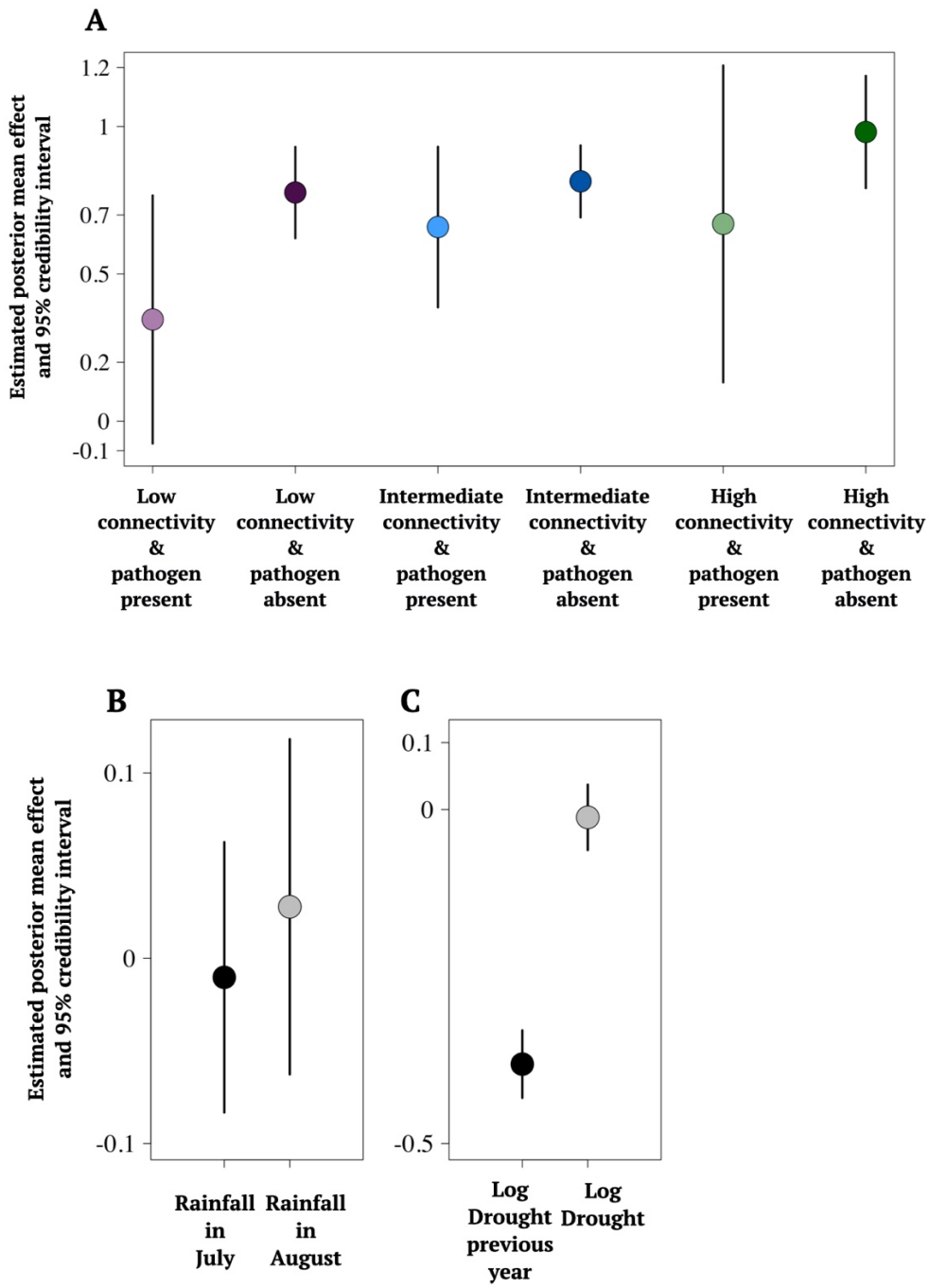

Fig. 1: Model estimated effects on *Plantago lanceolata* population size changes in the Åland islands in 2001-2008. The estimated median effects for host population growth with 95% credibility intervals of the fixed effects of the Bayesian INLA model: **A** The effect of pathogen presence and absence in the host populations in the three connectivity categories, **B** The effect of rainfall in July and August; and **C** The effect of detected drought symptoms in the host populations in the previous and current year.

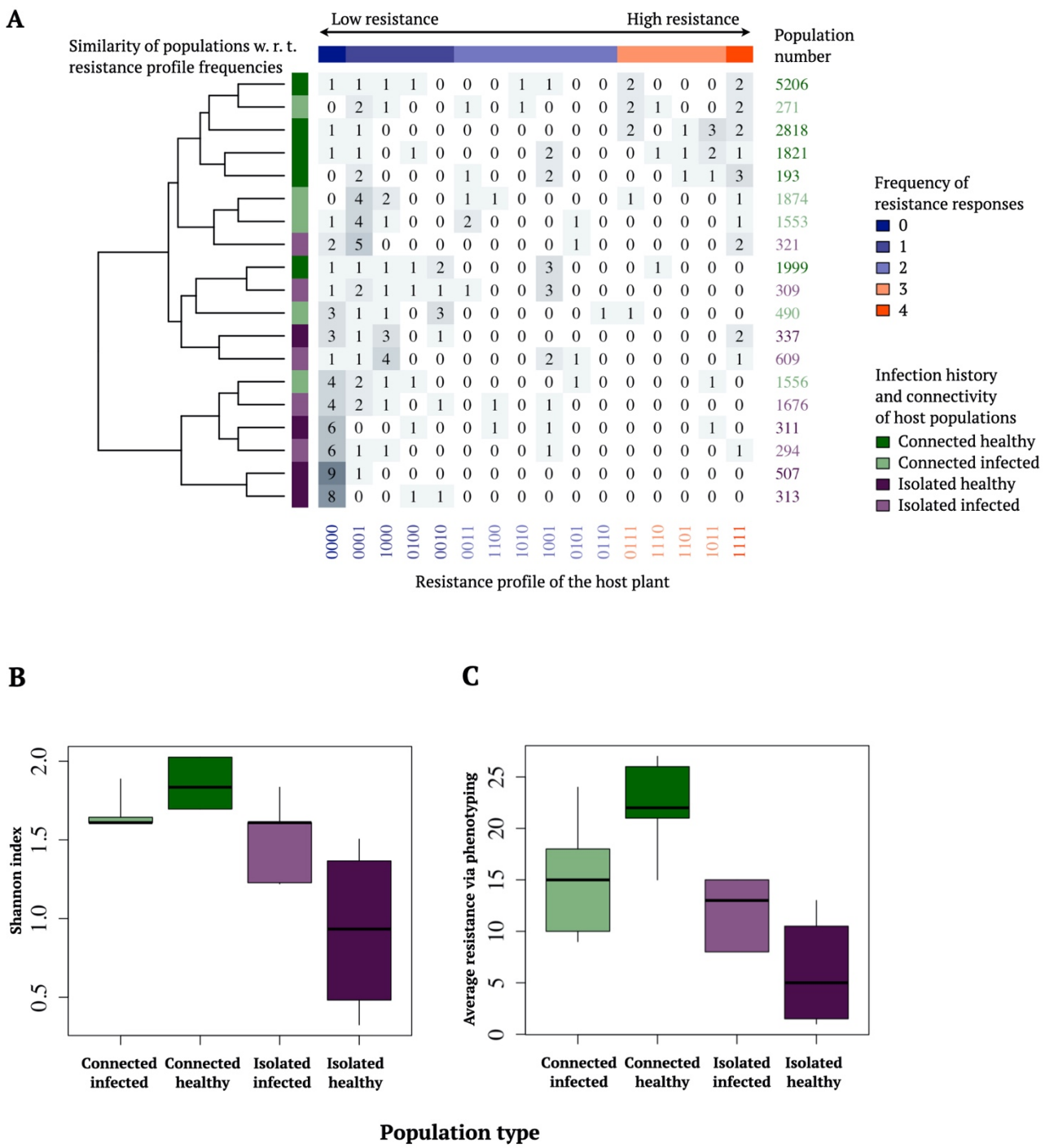

Fig. 2: Resistance of *Plantago lanceolata* populations depends on connectivity (S^H) and disease history. **A** The matrix of detected resistance phenotypes in the inoculation study shows clustering of similar phenotypic profiles detected in populations in each of the four connectivity (S^H)–infection history categories. The columns of the matrix correspond to resistance phenotypes, where the i'th

element of the vector is 1, if resistance to pathogen strain I was detected, and zero otherwise. The rows of the matrix encode the observed frequencies of resistance phenotypes within the studied populations. The dendrogram visualizes the similarity structure between the populations, distance along the tree encoding for the degree of similarity between the populations. It is based on a hierarchical clustering (implemented with complete linkage method, aiming to find similar clusters), applied to Euclidean distances between the phenotype profiles within the populations. **B** The average Shannon diversity index of host populations in each connectivity (S^H)-disease history category, and **C** the average resistance of the same populations in each category. The centre lines of the boxplots (**B-C**) show the medians, box limits show the 25% and 75% quantiles, and the whiskers span to the data extremes. Purple colours depict isolated populations, and green colours well-connected populations.

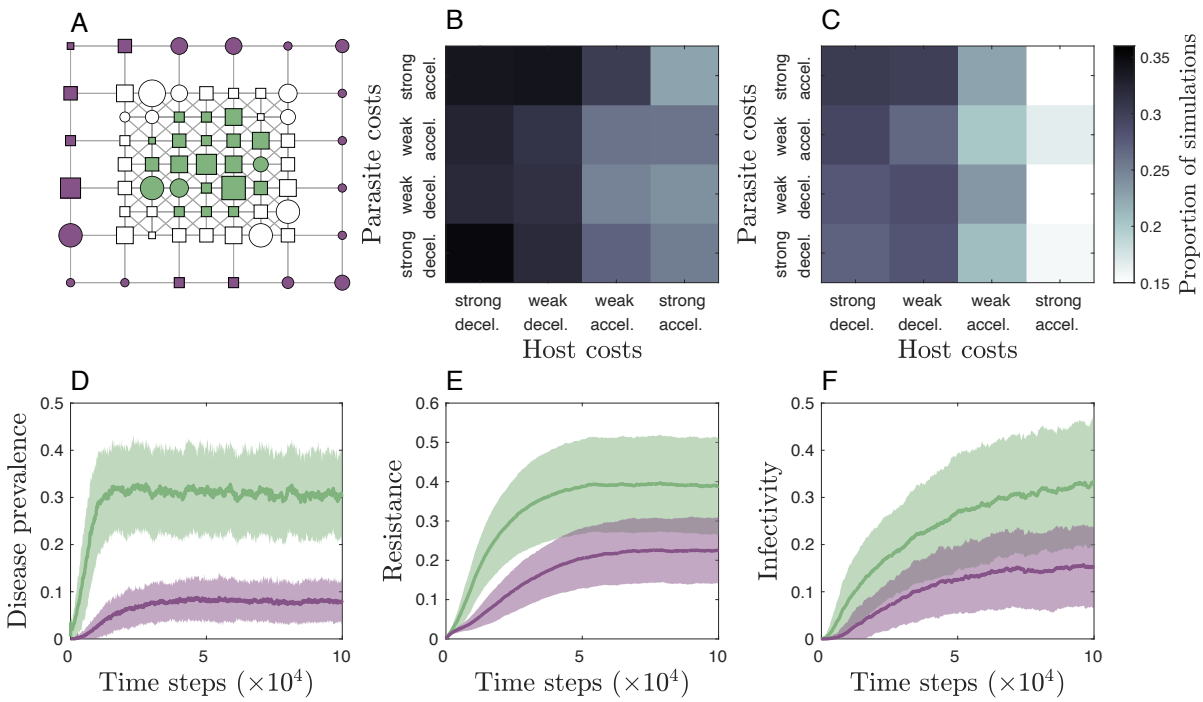

Fig. 3: Metapopulation simulation results. A Example snapshot of the simulation dynamics at t=10,000 across a metapopulation with assortative connectivity, highlighting well (green) and poorly (purple) connected populations (unshaded populations are neither well nor poorly connected) that are currently infected (squares) and uninfected (circles). The size of each node corresponds to the mean resistance of the local population. **B-C** Proportion of simulations which qualitatively match the empirical results as the shape of the host and pathogen cost functions are varied for transient (**B**) and long-term (**C**) dynamics: (strong decel. (decelerating): $c_H^2, c_P^2 = -10$; weak decel.: $c_H^2, c_P^2 = -3$; weak accel. (accelerating): $c_H^2, c_P^2 = 3$; strong accel.: $c_H^2, c_P^2 = 10$). **D-F** Example simulation results, showing mean (bold line) and standard deviations (shading) for disease prevalence (**D**), resistance (**E**), and infectivity (**F**) in well (green) and poorly (purple) connected populations ($c_H^2 = -3$, $c_P^2 = 10$, $\beta = 0.01$, with assortative network structure). Fixed parameters as defined in Supplementary Table 3.