Strain Diversity and Spatial Distribution Are Linked to Epidemic Dynamics in Host Populations*

Eck, Jenalle L.

2022-01-01

FOCUSED TOPIC

Strain Diversity and Spatial Distribution Are Linked to Epidemic Dynamics in Host Populations*

Jenalle L. Eck,[1,2,†] Benoit Barrès,[1,‡] Samuel Soubeyrand,[3] Jukka Sirén,[4,§] Elina Numminen,[1,‖] and Anna-Liisa Laine[1,2]

1. Organismal and Evolutionary Biology Research Program, University of Helsinki, 00790 Helsinki, Finland;   2. Department of Evolutionary Biology and Environmental Studies, University of Zurich, 8057 Zurich, Switzerland;   3. INRAE, BioSP, 84914 Avignon, France;   4. Department of Computer Science, Helsinki Institute for Information Technology, Aalto University, 02150 Espoo, Finland

ABSTRACT: The inherently variable nature of epidemics renders predictions of when and where infection is expected to occur challenging. Differences in pathogen strain composition, diversity, fitness, and spatial distribution are generally ignored in epidemiological modeling and are rarely studied in natural populations, yet they may be important drivers of epidemic trajectories. To examine how these factors are linked to epidemics in natural host populations, we collected epidemiological and genetic data from 15 populations of the powdery mildew fungus, *Podosphaera plantaginis*, on *Plantago lanceolata* in the Åland Islands, Finland. In each population, we tracked spatiotemporal disease progression throughout one epidemic season and coupled our survey of infection with intensive field sampling of the pathogen. We found that strain composition varied greatly among populations in the landscape. Within populations, strain composition was driven by the sequence of strain activity: early-active strains reached higher abundances, leading to consistent strain compositions over time. Co-occurring strains also varied in their contribution to the growth of the local epidemic, and these fitness inequalities were linked to epidemic dynamics: a higher proportion of hosts became infected in populations containing strains that were more similar in fitness. Epidemic trajectories in the populations were also linked to strain diversity and spatial dynamics: higher infection rates occurred in populations containing higher strain diversity, while spatially clustered epidemics experienced lower infection rates. Together, our results suggest that spatial and/or temporal variation in the strain composition, diversity, and fitness of pathogen populations are important factors generating variation in epidemiological trajectories among infected host populations.

Keywords: pathogen diversity, strain composition, plant disease, disease epidemic, *Plantago lanceolata*, powdery mildew.

Introduction

Predicting the trajectory of disease progression is a central goal of epidemiology (Rivers and Scarpino 2018). Such predictions are important given the negative effects of disease on human and wildlife health (Gulland 1995; Yach et al. 2006), food security (Gregory et al. 2009; Chakraborty and Newton 2011; Leung and Bates 2013), and biodiversity and ecosystem services in ecological communities (Daszak et al. 2000; Boyd et al. 2013). However, accurate predictions of disease dynamics are challenging because epidemics arise from ecological and evolutionary processes occurring between hosts, pathogens, and the environment across multiple scales of biological organization (Penczykowski et al. 2015; Parratt et al. 2016). Within pathogen species, population dynamics can vary across a landscape because of factors such as the availability of hosts through space and time (Burdon et al. 1995; Sapoukhina et al. 2009), the genetic resistance of hosts against pathogens (Burdon and Jarosz 1991; DiLeone and Mundt 1994; Alexander et al. 1996; Thrall and Burdon 2000; Laine 2004; Ostfeld and Keesing 2012; Susi and Laine 2015) and environmental conditions (Laine 2007). Furthermore, it is well established

that epidemics can have very different trajectories even when host densities, genetic diversities, and environmental conditions are similar, such as in agricultural crops (Thrall and Jarosz 1994; Shaw 2006). This suggests that the composition of pathogen populations could play a key role in how epidemics unfold (Susi et al. 2015; Zhang et al. 2019). Yet for many pathogen species remarkably little is understood of how the diversity, fitness, or distribution of strains within pathogen populations varies through space or time (Carlsson-Granér 1997; Frank 1997; Orum et al. 1997; Carlsson-Granér and Thrall 2002; Barrett et al. 2008; Tack and Laine 2014b; Nagy et al. 2019) and influences epidemic dynamics (Ericson et al. 1999; Archie et al. 2009). As molecular identification tools advance the study of pathogens and other parasites (Giraud et al. 2008; Scholz et al. 2016; Truong et al. 2017), studies that compare pathogen population dynamics at multiple spatial and biological scales will allow us to better understand why epidemics vary.

The strain composition and diversity of pathogen populations are hypothesized to vary among populations in a landscape and within populations during an epidemic season, but studies at these spatiotemporal scales are rare (Bayman and Cotty 1991; Orum et al. 1997; Palmer et al. 2001; Barrett et al. 2008). Primary disease foci establish in populations by immigrating from existing populations or by surviving from the previous epidemic season in situ. Thus, variation in strain composition among pathogen populations is determined in part by differential abilities among strains to persist in the local environmental conditions within populations and in part by strain assembly processes from the regional genetic pool (akin to the processes governing species assembly in ecological communities; Chase 2005; Chase and Myers 2011; Shipley et al. 2012; Krasnov et al. 2015). Strain assembly and off-season survival are affected by a number of niche-based and stochastic factors, such as pathogen dispersal ability (Shaw 1995; Tack and Laine 2014b), filtering by local environmental conditions and host resistance (Marçais and Desprez-Loustau 2014), local demographic stochasticity (Gibson et al. 2004), and connectivity in the landscape (Jousimo et al. 2014; Tack et al. 2014). In some pathogen species this variability, along with the generation of novel strains within populations via sexual reproduction, can lead to very different strain composition and infection dynamics among populations or to metapopulation dynamics where colonization and extinction of populations is common in the landscape (Burdon et al. 1995; Thrall and Burdon 2003). Once pathogen populations are established, their composition may diversify over the course of an epidemic through a variety of processes. These include the arrival of novel strains from other populations (Hiremath et al. 2008; Biek and Real 2010), differential virulence or host specificity among strains leading to differential fitness (Ichielevich-Auster et al. 1985; Cotty 1989; Davies and

Donachie 1996; Barrett et al. 2009; Wang et al. 2011), and variation in the timing of life history traits among strains (Woodhams et al. 2008; Vaumourin and Laine 2018; Numminen et al. 2019). Selection pressure that results in adaptation to local host defenses and abiotic conditions (Gupta and Maiden 2001; Thrall and Burdon 2003; Laine 2005), spatial variation among host populations in their resistance structure and environmental conditions (Ebert et al. 1998; Thompson 2005; Laine et al. 2011; Jousimo et al. 2014; Höckerstedt et al. 2018), and proximity to other pathogen populations that may act as sources of new strains that colonize midepidemic (Susi et al. 2015; Bousset et al. 2018) could all influence further divergence in strain composition among populations. Altogether, variation in strain composition and diversity among and within pathogen populations arises from multiple processes occurring across regional and local scales (Penczykowski et al. 2015).

Pathogen populations can also be shaped by the biotic interactions among strains and hosts that determine strain abundance and fitness, and such variation in strain fitness within pathogen populations could affect epidemic dynamics in host populations. For example, priority effects, in which the relative timing or sequence of strain activity causes early-arriving strains to gain a competitive advantage over or to facilitate late-arriving strains (via changes in host susceptibility), could shape strain relative abundance (Johnson et al. 2015; Meester et al. 2016). Strain fitness could also be linked to strain abundance at certain spatial and temporal scales (as more fit strains reproduce more abundantly), but competition among strains and other environmental factors, such as host susceptibility, complicate this relationship (Kirchner and Roy 2002; Zhan and McDonald 2013). Variation in the relative fitness of co-occurring strains within populations has been observed in several pathogen species (Kaltz and Shykoff 2002; Laine 2008) and could affect the composition and outcome of local epidemics (Zhang et al. 2019). For example, variation in strain relative fitness may influence strain coexistence (Cobey and Lipsitch 2013) or rates of epidemic growth (Osnas et al. 2015). Understanding why some strains become abundant while others remain rare in populations, what determines strain relative fitness, and how variation in fitness among co-occurring strains affects epidemics are all of critical importance for predicting and responding to disease (Khanna et al. 2008).

Variation in the rate and magnitude of infections among and within host populations could also be explained in part by variation in strain diversity or the spatial distribution of infections (Twizeyimana et al. 2009; Tollenaere et al. 2012). Because pathogen strains typically exhibit some degree of specialization on host genotypes (Thompson and Burdon 1992; Salvaudon et al. 2008; Barrett and Heil 2012), the presence of many pathogen strains in diverse pathogen

populations could facilitate the infection of a broad range of host genotypes, leading to a higher proportion of hosts becoming infected. Within populations, strain diversity is also hypothesized to influence the rate of epidemic growth: more diverse pathogen populations are expected to have faster epidemic growth rates than less diverse populations (McDonald and Linde 2002). Finally, the spatial distribution of infections within populations can be highly variable, depending on pathogen dispersal ability and host distributions, and could also determine the rate of progression of epidemics (Mundt and Leonard 1985). For example, the growth rate of spatially clustered epidemics is highly limited by pathogen dispersal distances, as demonstrated by the early limitations on epidemic growth for initial disease foci establishing in new sites (McCartney and Fitt 1998). Insights into the drivers of epidemic trajectories can come from linking measured disease outcomes to spatial and/or temporal variation in pathogen strain diversity and distribution among and/or within populations.

In this study we examine how pathogen strain composition, diversity, fitness, and the spatial distribution of infections change over space and/or time and are linked to epidemic dynamics in 15 host populations in a wild plant-pathogen system. Specifically, we test the following. First, does strain composition vary among pathogen populations or within populations over time? Second, is strain abundance determined by the sequence of strain activity or by strain fitness? Third, do co-occurring strains vary in fitness, and are fitness inequalities linked to epidemic dynamics? Fourth, are pathogen strain diversity or the spatial distribution of infections linked to epidemic dynamics? To address these questions, we combine epidemiological surveys of fungal disease progression in 15 natural populations of a wild plant species with intensive genetic sampling of the pathogen populations. With this study we aim to illustrate how knowledge of spatial and/or temporal variation in pathogen strain composition, diversity, and fitness can help predict the dynamics of epidemics in host populations.

Methods

Study Populations and Species

We studied epidemic dynamics in 15 populations of the perennial herb *Plantago lanceolata* L. (Plantaginaceae) infected by the fungal pathogen *Podosphaera plantaginis* (Castagne) U. Braun and S. Takam. in the Åland Islands (60°08′53″N, 19°47′18″E), Finland. The populations are part of a larger network of >4,000 populations of *P. lanceolata* spread across the Åland archipelago that have been monitored for their size and location since the early 1990s (Hanski 1999). *Plantago lanceolata*, or ribwort plantain, is native to Åland and much of Eurasia and occurs mainly in small meadows and dis-

turbed areas in Åland. It is monoecious and self-incompatible and can reproduce sexually via seeds or asexually via clonally produced side rosettes (Sagar and Harper 1964). In Åland, *P. lanceolata* flowering and seed production occurs from June to August. The wind-dispersed pollen and seeds fall primarily near the maternal plant (Bos 1992).

Podosphaera plantaginis, a powdery mildew fungus (order Erysiphales), is an obligate pathogen of living foliar tissue. Fungal hyphae grow on the leaf surface and produce localized infections that inhibit plant growth and reproduction (Bushnell 2002) and may lead to mortality in the presence of other stressors, such as drought (Laine 2004; Susi and Laine 2015). Asexual reproduction occurs cyclically throughout a growing season, and infection is transmitted via wind-dispersed spores (conidia; Ovaskainen and Laine 2006). Resting structures (chasmothecia) that enable the pathogen to overwinter and initiate new infections in the next growing season (Tack and Laine 2014*a*) are produced via haploid selfing or outcrossing between strains (Tollenaere and Laine 2013). Infection on individual hosts can be cleared if infected leaves are dropped as a result of natural senescence or drought before pathogen reproduction occurs.

Infection of *P. lanceolata* by *P. plantaginis* has been studied in Åland since 2001 (Laine and Hanski 2006). In Åland, *P. plantaginis* persists as a metapopulation through frequent colonization and extinction events, infecting ~1%–20% of the *P. lanceolata* populations in a given year (Jousimo et al. 2014). For this study we selected 15 *P. lanceolata* populations that had been infected for at least three consecutive years before the study (Jousimo et al. 2014). Distances between pairs of study populations range from ~1 to 40 km. The interaction between *P. lanceolata* and *P. plantaginis* in Åland is characterized by high levels of diversity in resistance within and among populations of the host (Laine 2004, 2007), coupled with high levels of pathogen genetic diversity among populations (Tack et al. 2014). Infection is mediated by a high degree of specificity through a genotype (host) × genotype (pathogen) interaction (Laine 2007). Co-infection, in which more than one pathogen genotype contributes to a local infection, occurs commonly in this system (Tack et al. 2014; Susi et al. 2015).

Epidemiological Surveys and Pathogen Sampling

In each of the 15 study populations, *P. plantaginis* infection on *P. lanceolata* was surveyed periodically (every 1–2 weeks) throughout the 2014 epidemic season, from early July to late August, when signs of infection are visible to the eye (fig. S1; table S1; figs. S1–S3, tables S1–S9 are available online). During the first epidemiological survey in each population, up to 30 infected focal individuals were located and tagged by visually scanning plants for signs of the pathogen. Focal

individuals were located at least 3 m apart, and their locations were recorded by GPS. The infection severity of each focal individual was scored categorically on a scale of 0–3 (1 = one leaf lightly infected; 2 = several leaves lightly infected; 3 = one leaf heavily infected or many leaves infected; 0 = loss of infection). In a 1.5-m-radius circle surrounding each focal individual, we also visually estimated the following field measurements: (1) the percentage of cover by area of *P. lanceolata*, (2) the percentage of the *P. lanceolata* individuals showing signs of severe drought stress (i.e., brown, wilted leaves), and (3) the percentage of the non-drought-stressed *P. lanceolata* individuals showing signs of infection by *P. plantaginis* (the infection is not visible under drought stress). During each subsequent survey of each population, the survival and infection severity of each focal individual and the corresponding circle-level variables were remeasured. In populations containing <30 focal individuals, new focal individuals (and corresponding circles) could be established from new infections. Between four and eight surveys were completed per study population (table S1). From the field measurements, we calculated a_{ik}, the total number of *P. lanceolata* individuals in circle (i) during survey (k), by assuming that 1% cover by area (corresponding to field measurement 1 in the list above) equals 10 individuals (Penczykowski et al. 2018). We then calculated b_{ik}, the number of drought-stressed individuals in circle (i) during survey (k), as a_{ik} times the proportion of drought-stressed individuals (corresponding to field measurement 2 in the list above); a_{ik} minus b_{ik} yielded c_{ik}, the number of potential host individuals in circle (i) during survey (k). We then calculated f_{ik}, the number of infected individuals in circle (i) during survey (k), as c_{ik} times the proportion of non-drought-stressed individuals showing signs of infection (corresponding to field measurement 3 in the list above). To examine infection rates at the population level, we then calculated the infection rate among at-risk individuals (i.e., those within 1.5 m of an infected focal individual and thus in close proximity to an existing infection center), g_{kp}, in each survey (k) of each population (p), as the sum of the number of infected individuals in the circles in the population (f_{kp}) divided by the sum of the number of potential host individuals in the circles in the population (c_{kp}). Our decision to sample the progression of infection by visually locating infected individuals was based on the previous finding that infection within host populations is highly aggregated (Ovaskainen and Laine 2006); hence, surveying the entire host population would not be time effective, while using a standardized sampling scheme would likely miss most infections.

During each survey of each population, infection data were also collected on a random sample of 50 *P. lanceolata* per population (only 25 individuals were sampled during the final survey of each population). Each individual was visually inspected for powdery mildew infection and was considered infected if at least one leaf showed signs of infection. Overall host infection rates, h_{kp}, were then calculated for each survey (k) of each population (p) as the proportion of infected individuals in this sample. Because a new random sample of individuals was selected during each survey of a population, this provides an additional, independent measure of infection rates in each host population over time (that is independent of the design of the focal individual and circle-level measurements).

Pathogen samples were collected from focal individuals in each study population during the survey when the individual's infection score reached 3 (to allow collection of one infected leaf without disturbing the growth of the epidemic). If the infection score of a focal individual never reached 3, a sample was collected from it during the final survey of the population. Some focal individuals that had been sampled in an earlier survey were also resampled during the final survey (the second sample was collected from a different leaf than the first sample and thus represents a new infection). During the final survey in each population, samples were also collected from up to four additional infected individuals within each circle. Each infected leaf sample was placed in a paper envelope and stored temporarily in a cold, dry room to ensure rapid drying. Samples were then transported to the University of Helsinki (Helsinki, Finland) and stored at −20°C until DNA extraction.

Pathogen Strain Identification

To prepare samples for DNA extraction and genotyping, a small piece of infected tissue was cut from each sample, soaked in liquid nitrogen, and grinded. DNA was extracted from the samples using the E.Z.N.A. plant DNA kit (Omega Bio-tek, Norcross, GA) at the Institute of Biotechnology (Helsinki, Finland) following the manufacturer's instructions. To genotype the samples, we used a panel of 19 single-nucleotide polymorphism (SNP) markers developed for *P. plantaginis* (Tollenaere et al. 2012) to assign multilocus genotypes (hereafter referred to as "strains") to each sample (for additional details, see the supplemental PDF, available online). Genotyping was performed at the Finnish Institute for Molecular Medicine (Helsinki, Finland) using the Sequenom MassARRAY iPLEX platform. Unique strains were identified by variation at any SNP site. Coinfected pathogen samples (i.e., those comprised of two or more strains) were identified by the presence of heterozygosity at any SNP site (Susi et al. 2015). Coinfections were resolved to two parent strains using a computer algorithm (for details, see the supplemental PDF). The algorithm also identified some strains that occurred only in coinfections (i.e., that were not sampled singly), yielding a total of 106 strains for analysis. To most accurately represent the abundance of each

strain, coinfected samples are included as one count of each of the two parent strains in all analyses.

Statistical Methods

Does the Strain Composition of Pathogen Populations Vary over Space and Time? We tested whether pathogen strain composition (i.e., the identity and abundance of strains) is dissimilar among contemporaneous populations and within populations over time. For this test we compared initial pathogen composition (the set of samples collected on or before July 25, 2014, corresponding roughly to the first half of the study period) with final pathogen composition (the full set of samples collected during the study period). We simultaneously tested (1) whether strain composition is dissimilar over time within populations (between the initial and final composition of each population), (2) whether strain composition is dissimilar among contemporaneous initial populations, and (3) whether strain composition is dissimilar among contemporaneous final populations. We used a generalized Monte Carlo plug-in test with calibration (GMCPIC; Soubeyrand et al. 2017; for details, see the supplemental PDF) to test the equality of the vectors of probabilities of two multinomial draws (in this case, the composition of the two populations), using the StrainRanking package (Soubeyrand et al. 2014) in the R statistical environment (R Development Core Team 2019). We also calculated several strain diversity metrics for each survey of each population: strain rarefied richness (which was correlated with richness; fig. S2), Shannon diversity (H'), and Pielou's evenness (J'). Finally, we calculated dissimilarity in final strain composition for all pairwise combinations of populations via ordination-based procedures (Bray-Curtis dissimilarity and Jaccard index) using the vegan package (Oskanen et al. 2017) in the R statistical environment. Because ordination-based metrics produced the same qualitative results as the GMCPIC test (table S2), we focus here on GMCPIC.

Is Strain Abundance Predicted by the Sequence of Activity or by Strain Fitness? We used a linear mixed effects model to test whether the abundance of the strains in our study populations is explained by the sequence of strain activity or by strain fitness. In this model we explained the log-transformed abundance of each strain in each population it occurs in at the end of the epidemic (i.e., the number of times a strain was sampled in a population during the course of the epidemic) as a function of (1) the numbered day at which the strain was first sampled in the population (relative to the first survey of the population) and (2) the relative fitness estimate of the strain (estimated as described in the paragraph below). Population was included as a random effect in this model. For the (few) strains that occur in more than one population, each occurrence of the strain was included separately in the model. All linear (and generalized linear) mixed effects models in our study were performed using the lme4 package (Bates et al. 2015) in conjunction with the lmerTest package (Kuznetsova et al. 2017) in the R statistical environment.

Do Co-occurring Strains Vary in Fitness, and Are Fitness Inequalities Linked to Epidemic Dynamics? We estimated the fitness of each strain (i.e., the relative contribution of each strain to the local epidemic), then tested whether fitness varied among strains at two spatial scales. Strain fitness was estimated by linking the spatiotemporal data on strain occurrence (i.e., in which circles a strain was sampled and during which surveys) with the epidemiological time series data on changes in the number of infected hosts in the circles (using the StrainRanking package in the R statistical environment; for full details, see the supplemental PDF). Namely, local epidemic growth (Z_i) in a given circle (i; i.e., the change in the number of infected hosts in the circle between surveys relative to the time elapsed) was equated to the sum (over strains s) of the products of the strain fitnesses (z_s) and the local strain proportions ($p_{i,s}$) plus a centered Gaussian noise (ε_i):

$$Z_i = \sum_{i=1}^{S} z_s p_{i,s} + \varepsilon_i. \qquad (1)$$

Strain fitnesses were estimated from the coefficients of this equation, yielding a fitness estimate for each strain in each population it occurred in (representing strain fitness over the duration of the epidemic). We then tested whether pairwise combinations of strains varied in fitness at two spatial scales: (1) within each population (i.e., among co-occurring strains only) and (2) among all 15 populations simultaneously (i.e., among all strains detected in the study) using the StrainRanking package. More precisely, for any pair of strains x and y, we tested the null hypothesis $z_x = z_y$ using a permutation approach, in which the strain proportion vectors are randomly and uniformly reallocated to any sampling circle. To account for the multiplicity of tests, we assessed whether the proportion of rejected tests for each population is larger than 5%.

We then calculated strain fitness inequality in each of the study populations as the proportion of pairwise strain combinations that differed significantly in fitness relative to the total number of strain combinations in the population. To test whether the proportion of strain fitness inequality in the pathogen populations predicted maximum infection rates in the associated epidemics, we then used two generalized linear models to model (1) the maximum overall infection rates in the host populations and (2) the

maximum infection rates among at-risk individuals in the host populations.

Is Pathogen Strain Diversity or the Spatial Distribution of Infections Linked to Epidemic Dynamics? We used a series of generalized linear mixed effects models to explore whether (*a*) strain richness, diversity, or evenness and (*b*) the spatial distribution of infections within pathogen populations (i.e., the amount of infection clustering) explain infection rates in host populations. To quantify the spatial distribution of infections within each epidemic and capture variation in the degree of infection clustering among populations and within populations over time, we first constructed infection connectivity matrices for each population at each time point. The connectivity matrix, **M**, was constructed according to the circle-level infection and spatial data, and its (*i*, *j*)th entry is defined by

$$\mathbf{M}_{ij} = \frac{n_i}{n_{tot}} e^{-\alpha d_{ij}}. \tag{2}$$

In equation (2), n_i denotes the number of infected plants inside circle *i*, n_{tot} is the total number of infected plants in all the circles, and d_{ij} is the distance in meters between circles *i* and *j*. Following Penczykowski et al. 2018, we assume that the probability of pathogen dispersal from circle *i* to circle *j* declines exponentially with distance between the circles, according to an average dispersal distance ($1/\alpha$; we set $\alpha = 0.5$ m). We then calculated the amount of infection clustering during the epidemic for each population at each time point as the first eigenvalue of the connectivity matrix. The first eigenvalue characterizes the size of the largest cluster of infected individuals. Its values range from $\max_i(n_i/n_{tot})$ in a minimally clustered case (if the distances d_{ij} are very high so that the exponential terms in eq. [2] are close to zero) to 1 with maximal clustering (if the distances d_{ij} are very small so that the exponential terms in eq. [2] are close to one).

Within-population infection spatial clustering was then used as a fixed explanatory variable in models explaining host infection rates in the populations. We fitted three generalized logistic models, using three different metrics of strain diversity, as additional explanatory variables: (1) log-transformed richness, (2) Shannon diversity (*H'*), and (3) Pielou's evenness (*J'*). Population was included as a random effect in each model. First, we modeled the overall infection rates in the host populations, then built three additional models examining infection rates among at-risk individuals. These models were fitted using a beta-binomial distribution (to account for overdispersion) with the brms package (Bürkner 2017), and model fits were compared with leave-one-out cross-validation using the loo package (Vehtari et al. 2017) in the R statistical environment. Finally, we built a series of six generalized linear regression models examining the effect of the three strain diversity metrics on the maximum infection rates experienced overall and among at-risk individuals in each host population (for details, see the supplemental PDF). Data used in this study have been deposited in the Dryad Digital Repository (https://doi.org/10.5061/dryad.rjdfn2z94; Eck et al. 2021).

Results

Does the Strain Composition of Pathogen Populations Vary over Space and Time?

Pathogen strain richness, diversity, and evenness were variable among the 15 populations and over time within the populations (fig. 1*A*–1*C*; table 1; a map of the populations is displayed in fig. 2*A*). In total, 106 strains were identified in the populations. Strain richness ranged from only a single strain in three of 15 populations (20%) to 18 strains in the most strain-rich population (mean = 7.93 ± 5.61 strains; fig. 1*A*; table 1). We found that pathogen strain composition at the end of the epidemic was dissimilar among all pairwise combinations of the 15 study populations (fig. 2*B*, 2*C*; table S3). Only 7.5% of the strains (eight of 106 strains) were found in more than one population (seven of these eight strains were found in only two populations, while the remaining strain was found in five populations). We also found that initial pathogen strain composition in the early weeks of the epidemics was dissimilar among most pairwise combinations of the 15 study populations (in 81 of 105 combinations; for some combinations, similarity in strain composition could not be ruled out because of small sample sizes during the early epidemic; fig 2*C*; table S3). In contrast, initial and final strain composition were similar within each of the 15 populations over time (figs. 2*C*, S3; table S3).

Is Strain Abundance Predicted by the Sequence of Activity or by Strain Fitness?

We found that the abundance of pathogen strains within populations is predicted by an advantage to early-active strains rather than by strain fitness. Strains that were active earlier in an epidemic became more abundant by the end of the epidemic than later-active strains (fig. 3*A*, 3*C*; table S4; $F = 70.44$, $P < .001$, $n = 109$ strain occurrences). Abundance was variable among strains, with a few common strains and many rare strains (including several single-occurrence strains; fig. 3*D*). The timing of activity was also variable among strains: some were causing severe infections from the beginning of an epidemic, with others doing so only during the final survey (fig. 3*A*). In contrast, strain fitness was not a predictor of strain abundance
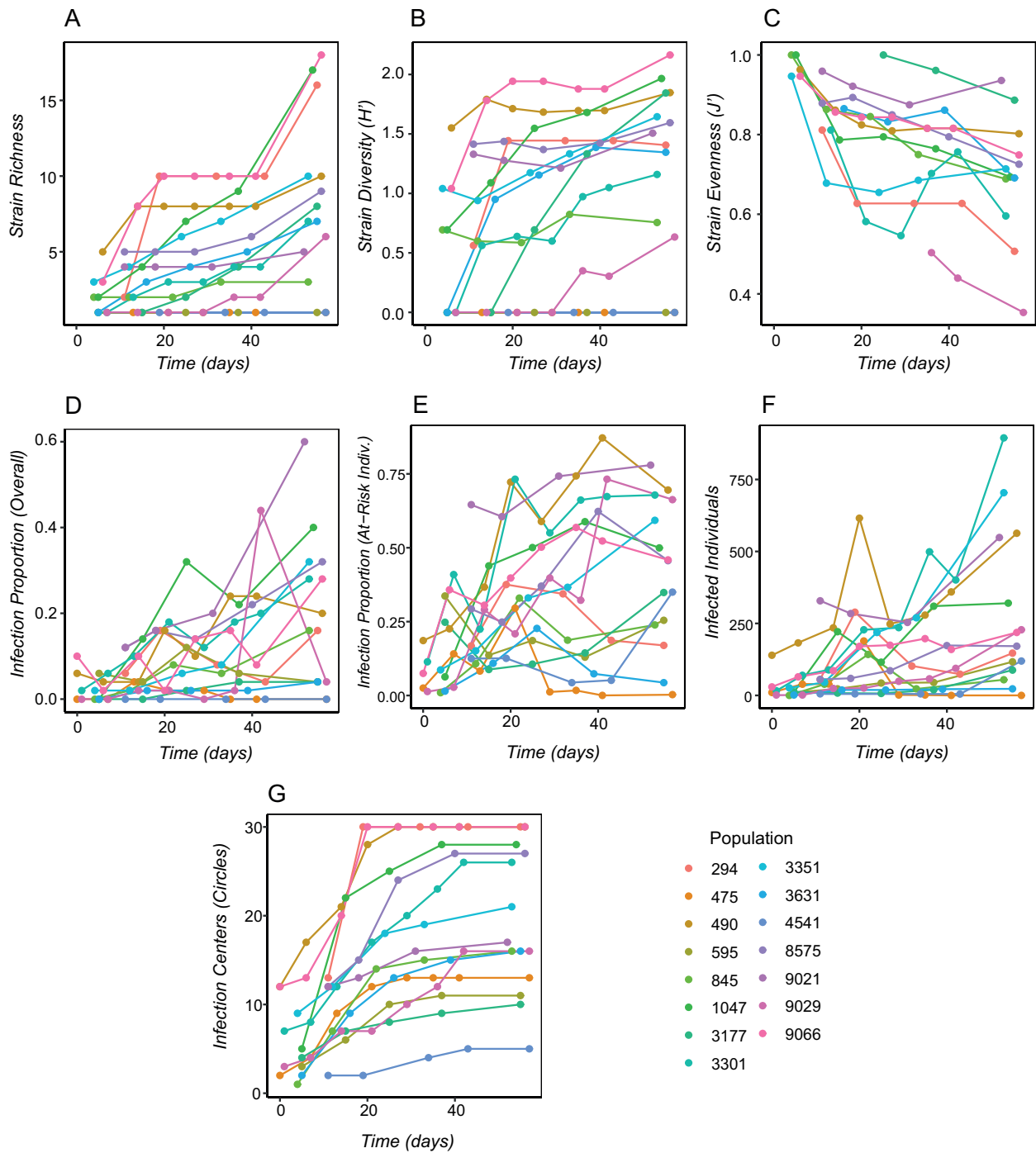
Figure 1: Pathogen strain diversity and epidemic dynamics among and within populations. In each of the 15 study populations, changes in pathogen strain richness (*A*), diversity (*H'*; *B*), and evenness (*J'*; *C*) as well as changes in overall infection proportion (*D*), infection proportion among at-risk hosts (*E*), the number of infected hosts (*F*), and the number of infection centers (i.e., infection circles; *G*) are shown over time (measured as the number of days elapsed since the study began).

Table 1: Strain richness, diversity, evenness, and fitness inequality among populations

Population	Strain richness	Strain richness (rarified)	Strain diversity (H')	Strain evenness (J')	Strains compared (fitness)	Strain combinations compared (fitness)	Strain combinations varying in fitness	Strain fitness inequality (proportion)
294	16	3.59	1.41	.51	14	91	18	.198
475	1	1	0	NA	NA	NA	NA	NA
490	10	4.88	1.85	.80	10	45	11	.244
595	1	1	0	NA	NA	NA	NA	NA
845	3	2.53	.76	.69	3	3	1	.333
1,047	17	4.90	1.97	.69	15	105	9	.086
3,177	8	5.62	1.84	.89	8	28	0	0
3,301	7	3.25	1.16	.60	6	15	0	0
3,351	10	4.37	1.64	.71	10	45	2	.044
3,631	7	3.95	1.34	.69	7	21	4	.191
4,541	1	1	0	NA	NA	NA	NA	NA
8,575	9	4.13	1.59	.73	9	36	3	.083
9,021	5	4.12	1.51	.94	5	10	1	.1
9,029	6	2.26	.63	.35	5	10	0	0
9,066	18	5.36	2.16	.75	17	136	1	.007
All	106	NA	3.79	.81	99	4,851	346	.071

Note: Strain richness, diversity, evenness, and fitness inequality in each of the 15 study populations at the end of the epidemics are compared. Metrics are cumulative, capturing strain accumulation through each epidemic's duration. Three populations contained only one strain, precluding calculation of some diversity- and fitness-related metrics. The number of strains compared for fitness differences is less than strain richness in some populations because of scarce data on some strains. The strain fitness inequality proportion is calculated as the number of strain combinations that vary in fitness divided by the number of strain combinations compared.

in the populations (fig. 3B; table S4; $F = 0.394$, $P = .532$, $n = 109$ strain occurrences).

Do Co-occurring Strains Vary in Fitness, and Are Fitness Inequalities Linked to Epidemic Dynamics?

We found that strain fitness varied among and within pathogen populations and that the level of fitness variation among strains within populations is linked to epidemic dynamics. Co-occurring strains varied in fitness in nine of 12 pathogen populations containing more than one strain (table 1). In populations where strains varied in fitness, levels of fitness inequality ranged from ~1% to 33% of the strain combinations varying in fitness (the overall level of fitness inequality among all strains in all populations was ~7%; table 1). Variation in strain fitness occurred more regularly in some populations (up to a maximum rate of ~1 in 3 combinations differing in fitness) than was common at the landscape scale (where ~1 in 14 combinations vary; table 1). Meanwhile, maximum overall host infection rates in the study populations ranged from 4% in the least infected population to 60% in the most infected population, while maximum infection rates among at-risk individuals ranged from 23% to 87% (fig. 1D–1G). We found that the proportion of fitness inequality among strains within the populations predicts maximum infection rates in the populations: populations with higher strain fitness inequality had lower maxi-

mum infection rates among at-risk individuals (fig. 4D; table S5; $z = -7.62$, $P < .01$, $n = 12$ populations).

Is Pathogen Strain Diversity or the Spatial Distribution of Infections Linked to Epidemic Dynamics?

We found that pathogen strain diversity and the spatial clustering of infections had strong, but opposite, relationships with infection rates in host populations. In general, metrics of strain diversity were positively linked to host infection rates (fig. 4A, 4C; tables S5–S8). Infection rates were higher (fig. 4A; tables S7, S8; posterior probabilities = 1.00, $n = 89$ surveys) and reached higher maxima (fig. 4C; tables S5, S6; richness: $z = 2.92$, $P < .01$; diversity [H']: $z = 3.69$, $P < .001$; $n = 15$ populations) in populations containing higher strain richness and diversity (H'). Models utilizing strain richness were best fitted to the data, but differences in fit were small (table S9). Strain evenness (J') was also positively linked to infection rates in the populations (tables S7, S8; posterior probabilities = 0.99, $n = 89$ surveys) but only predicted maximum infection rates among at-risk individuals (tables S5, S6). In contrast, the spatial clustering of infections was negatively linked to host infection rates: more spatially clustered epidemics had lower host infection rates than more spatially dispersed epidemics (fig. 4B; tables S7, S8; posterior probabilities = 0.99–1.00, $n = 89$ surveys). In general, strain diversity metrics and the
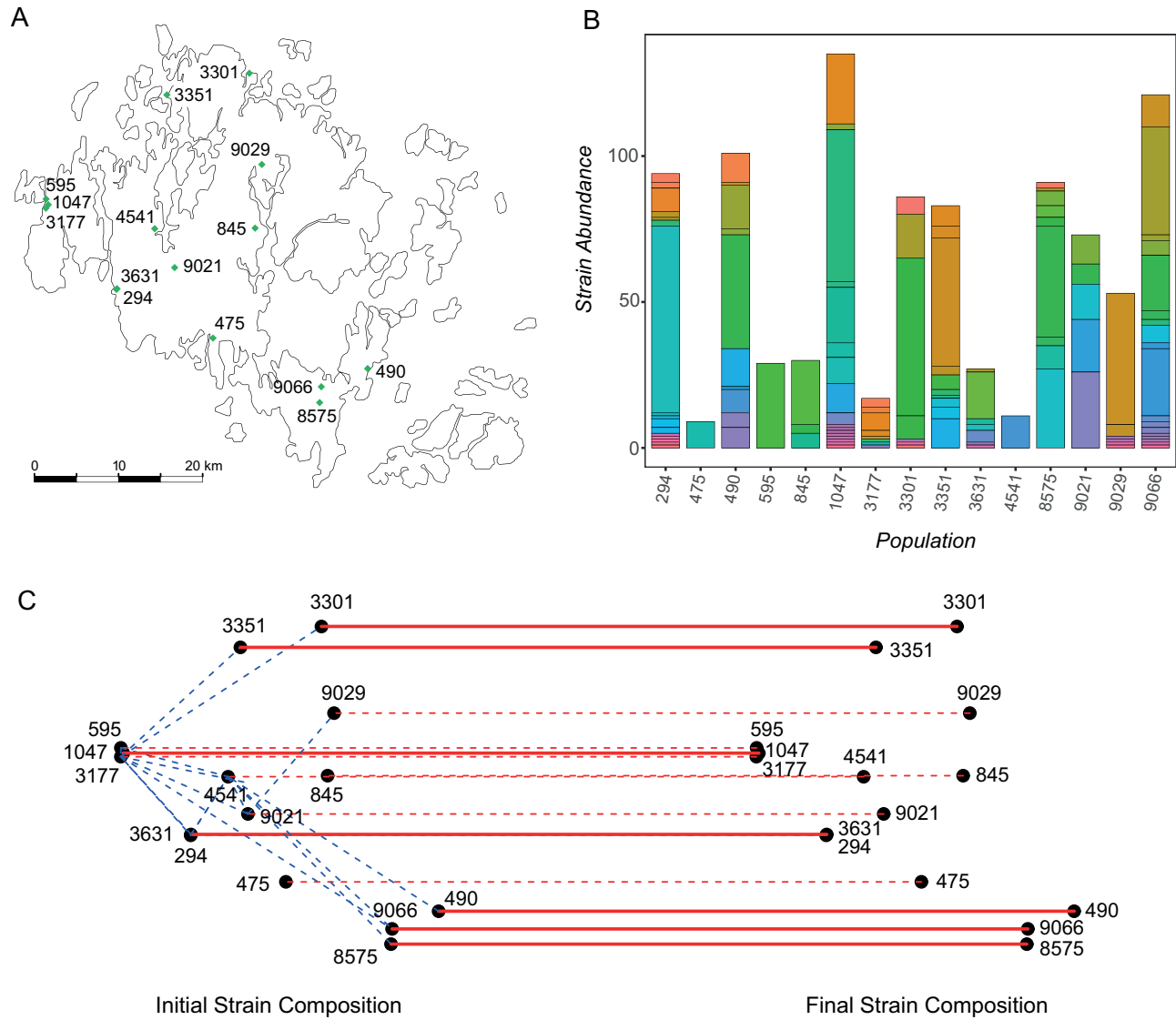
Figure 2: Pathogen strain composition is dissimilar among the 15 study populations. *A*, The 15 study populations were located on Fasta Åland (Åland Islands, Finland). Spatial distances between pairs of populations vary from <1 to ~40 km. *B*, Final strain composition of each of the 15 study populations is represented by a bar on the *x*-axis, with strain abundance on the *y*-axis. Strains are differentiated by color (because there are 106 strains, very similar shades may represent different strains). *C*, A spatiotemporal representation of the results of a test of dissimilarity in pathogen strain composition shows that final composition is dissimilar among populations but not dissimilar within populations over time (table S3). The 15 populations (depicted as points and corresponding to actual spatial relationships) are mapped twice to represent two time points: the left map represents initial strain composition, while the right map represents final strain composition. Red lines connecting a population to itself between maps indicate lack of dissimilarity in composition in the population over time. Blue lines connecting populations within a map indicate lack of dissimilarity in composition among populations at a given time. Lack of a connecting line between two populations within a map indicates dissimilarity in composition. Solid lines indicate that >10 samples were available for the test in each of the compared populations; dashed lines indicate that ≤10 samples were available in at least one of the populations.

spatial clustering of infections were more tightly linked to infection rates among at-risk individuals than in the host populations overall (tables S5–S8), though similar patterns emerged in both groups.

Discussion

The composition, diversity, and fitness of strains within pathogen populations are variable over space and time,
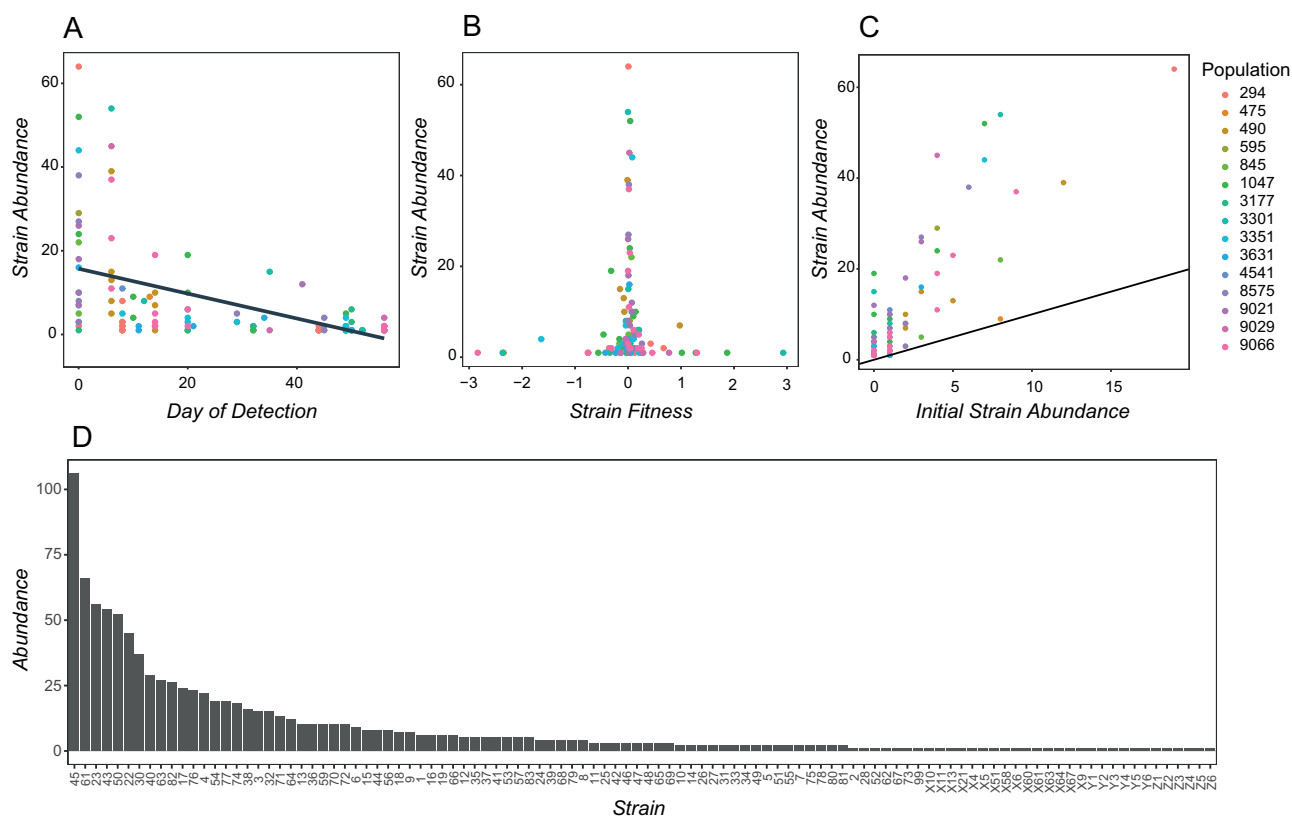
Figure 3: Abundance of pathogen strains was predicted by the timing of their activity but not by strain fitness. In *A–C*, each strain is a dot (colored by the population it occurs in), with the final abundance of the strain in the population on the *y*-axis. Mixed effects models show that strain abundance is predicted by the timing of strain activity (*A*) but not by strain fitness (*B*; table S4). In *C*, initial strain abundance in each population (measured in the first weeks of the epidemic) is depicted on the *x*-axis: strains that fall farther above the diagonal unity line had larger increases in abundance by the end of the epidemic. *D*, Histogram that shows the range of abundance of strains at the end of the study: some strains were abundant, while some strains occurred only rarely.

with potentially important consequences for epidemic trajectories in host populations, but studies at these scales are scarce (Bayman and Cotty 1991; Orum et al. 1997; Palmer et al. 2001; Barrett et al. 2008). We found that strain composition varied widely among 15 natural host-pathogen populations in a landscape but remained consistent within these populations over the course of their epidemics. Strain composition within populations was driven by a strong advantage to early-active strains, which reached higher abundances than their late-active counterparts. The fitness of strains within the pathogen populations also varied considerably, and more variability in strain fitness within populations was linked to lower maximum host infection rates during epidemics. Epidemic trajectories in the host populations were also linked to pathogen strain diversity and to the spatial distribution of infections with the epidemic: higher infection rates occurred in populations containing higher strain diversity, while spatially clustered epidemics

had lower infection rates. Together, our results suggest that spatial and/or temporal variation in the strain composition, diversity, and fitness of pathogen populations are important factors generating variation in epidemiological trajectories among and within infected host populations.

Variation in strain composition among pathogen populations or within populations over time and the factors that influence such variation are important if composition is linked to epidemic outcomes in host populations. Differences in composition among pathogen populations, such as those we found in our study populations, are characteristic of many pathogen species, and there is ample evidence that such differentiation is generated through coevolutionary interactions with host populations and results in local adaptation (Greischar and Koskella 2007; Hoeksema and Forde 2008). Dissimilarity in strain composition among pathogen populations is expected as a result of strain assembly processes from the regional landscape and differential survival of strains
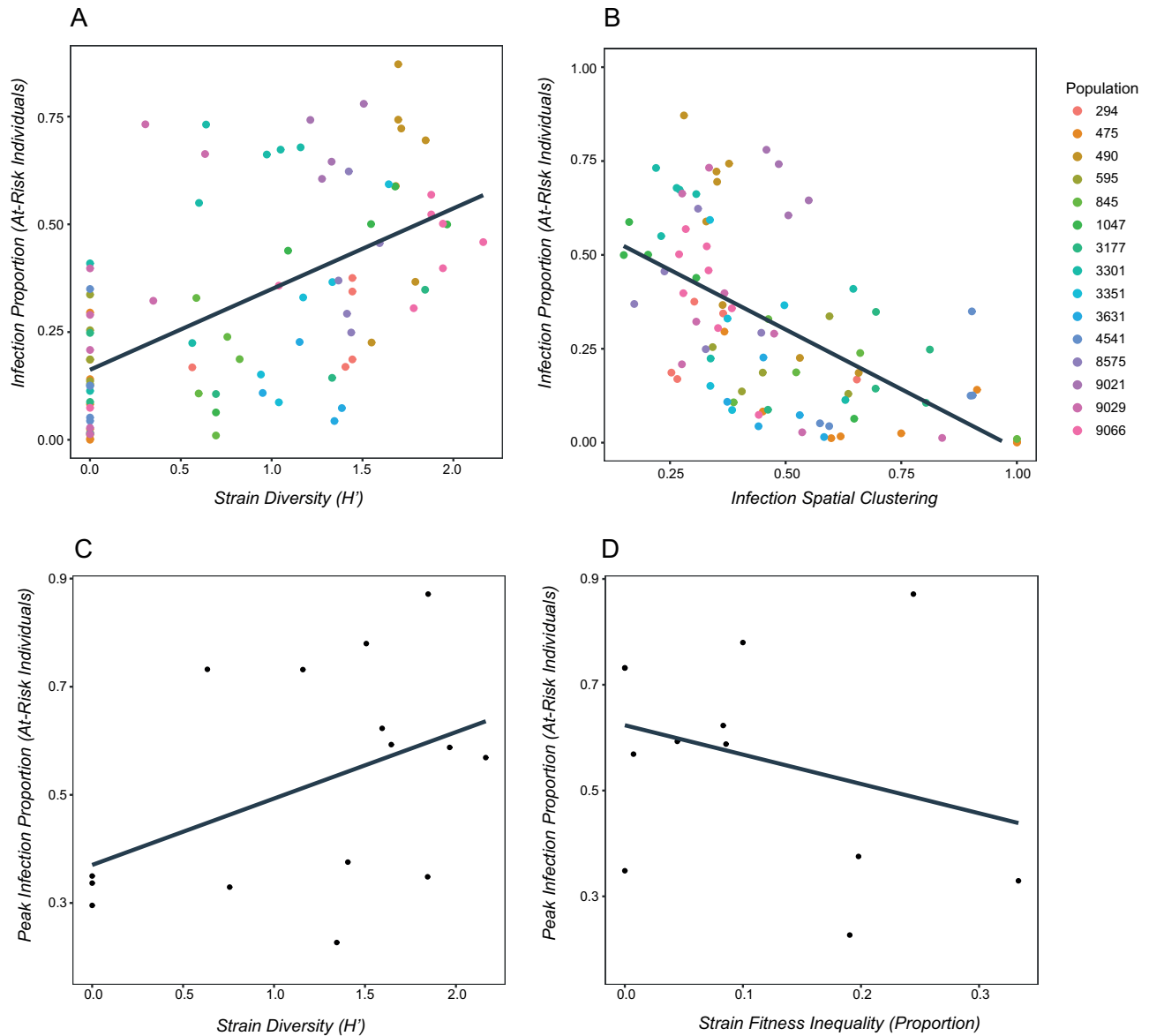
Figure 4: Pathogen strain diversity, infection spatial clustering, and strain fitness inequality are linked to infection rates in the host populations. Generalized linear regression models show that strain diversity (H'; *A*) and infection spatial clustering (*B*) in host populations predict contemporaneous infection rates ($n = 89$ surveys; table S7). Similar models show that strain diversity at the end of the epidemic (*C*), as well as strain fitness inequality during the epidemic (*D*), predicts peak infection rates in host populations containing more than one strain ($n = 12$ populations; table S6).

in their local environmental conditions (Krasnov et al. 2015). In line with previous findings in this pathosystem, more than 90% of the 106 strains we identified were unique to one population: the *Podosphaera plantaginis* pathogen metapopulation in Åland supports considerable strain diversity in each epidemic season, and the majority of strains are found in a single locality (Numminen et al. 2019). Spatial variation in strain composition may be explained to some degree by the life history features of *P. plantaginis*: at the end of

each epidemic, unique strains are generated through outcrossing and persist to the next epidemic season in resting spores (Laine et al. 2019). At the onset of the next epidemic season, the new strains are released locally, and during the growing season, infection spreads among hosts via winddispersed spores that typically disperse only a few centimeters (Ovaskainen and Laine 2006; Tack et al. 2013). In addition to dispersal limitation, heterogeneity in host resistance and environmental conditions in the landscape are

expected to generate variation in pathogen composition among populations by filtering arriving pathogen strains (Krasnov et al. 2015). Genotype-specific responses to host variation and environmental heterogeneity will further promote differentiation of pathogen populations across landscapes (Thompson 2005).

Although the composition of the pathogen populations diversified over the course of the epidemic season, pathogen strain composition did not significantly change within the study populations over time. This is likely explained by the strong effect of the sequence of strain activity on strain abundance that we see in our data: strains that were active earlier in an epidemic became more abundant by the end of the epidemic than later-active strains. Consequently, populations containing considerable strain diversity were often composed of a few abundant, early-active strains and many rare, late-active strains. An advantage of early arrival is a well-established fact in community ecology (Fukami 2015) and could potentially result from a priority effect in which early-active strains competitively exclude later-active strains. The timing and sequence of activity, in conjunction with strain-specific responses to host genotype, are receiving increasing support in disease biology as important determinants of pathogen community structure (Halliday et al. 2017, 2020; Clay et al. 2018).

We found that fitness varied widely among co-occurring pathogen strains in many populations, and this fitness variation was linked to epidemic outcomes. Interestingly, strain fitness did not predict strain abundance within the populations. As strain fitness is estimated from the epidemiological data (and some strains are found only in a few localities within the host populations), fitness may not be a good predictor of strains' ability to cope with the full range of host genetic and environmental variation that these populations support (Laine 2008). Indeed, strain-specific responses to host resistance and abiotic conditions have been broadly reported for pathogens (Salvaudon et al. 2008; Wolinska and King 2009) and could cause strain abundance at the population level to become unlinked from fitness at smaller spatial scales (e.g., if a strain is highly infective only on a relatively uncommon host genotype). As a consequence, the relative fitness of strains will change according to local conditions, and this context dependency is considered a powerful mechanism maintaining variation within pathogen populations (Thomas and Blanford 2003; Mitchell et al. 2005; Fels and Kaltz 2006; Laine 2007). For example, environmental factors, such as drought, may affect pathogen population dynamics by reducing the availability of host tissue. Variation in the fitness of strains within pathogen populations is also a potentially important, but understudied, factor affecting epidemic dynamics. Pathogen populations composed of strains that were more similar in fitness also had higher maximum infection rates,

suggesting that functional or trait-based diversity may also determine whether pathogen populations can overcome local conditions (Aguilar-Trigueros et al. 2014).

Variation in pathogen strain diversity and the spatial distribution of infections were also linked to epidemic outcomes in the populations. More diverse pathogen populations had more severe epidemics (i.e., higher host infection rates in general and higher maximum infection rates), consistent with the idea that diverse strain assemblages might allow pathogen populations to overcome a higher proportion of host defenses (Garrett and Mundt 1999; Thrall and Burdon 2000; Susi et al. 2015). When strains occurred in more even abundances within populations, this was positively correlated with maximum host infection rate, suggesting that strain dominance leads to less stable patterns of epidemic growth, which could result from competition among strains or increased virulence in diverse pathogen populations (Koskella et al. 2006). The level of relatedness within pathogen populations could also be expected to influence infection rates in host populations, as more closely related strain assemblages may infect a narrower range of host genotypes in interactions governed by specificity (e.g., in gene-for-gene resistance). Finally, epidemics are expected to progress depending on the spatial distribution of several factors, including the distribution of infected and uninfected individuals, host resistance, and strain fitness and diversity (Frank 1997; Carlsson-Granér and Thrall 2002; Penczykowski et al. 2015; Bousset et al. 2018). In addition to the influence of space on variation in strain composition among the populations at the regional scale, the spatial distribution of infections within populations also influenced epidemic dynamics, with more spatially clustered epidemics experiencing lower infection rates than more spatially dispersed epidemics. This is consistent with theoretical expectations that spatially clustered epidemics are more limited by pathogen dispersal distances, slowing epidemic growth relative to spatially dispersed epidemics (McCartney and Fitt 1998). Along with influencing epidemic dynamics, the amount of spatial clustering could also have longer-term effects on strain composition in pathogen populations, as a higher number of pathogen strains in close proximity should lead to increased opportunity for pathogen populations to diversify via sexual recombination (Laine et al. 2019). Pathogen strain diversity and the spatial distribution of infections are both important factors influencing infection rates and the trajectory of epidemics in host populations.

Altogether, we show that spatial and/or temporal variation in the composition, diversity, and fitness of strains within pathogen populations are generally important in determining epidemic trajectories. Such variation is an essential feature of epidemics and of biological populations in general, and important eco-evolutionary insights can be

gained by linking genotypic, phenotypic, and environmental variation in one population with realized outcomes in interacting populations across scales of biological organization (Lowe et al. 2017). In general, populations containing greater genetic or phenotypic diversity are expected to experience weaker density-dependent regulation and thus could reach larger population sizes (Johnson et al. 2016). In host-pathogen systems, the genetic, demographic, and trait composition of pathogen populations should also influence their coevolutionary dynamics, affecting their ability to overcome host resistance and adapt to local host genotypes or environmental conditions (McDonald and Linde 2002; Croll and McDonald 2017). In addition, disease is an important general factor shaping the abundance, diversity, and distribution of host populations (Scott 1988; Vredenburg et al. 2010). Our ability to predict and respond to epidemics, as well as to understand how pathogens and disease shape biological systems, will benefit from deeper knowledge about the factors that influence pathogen composition, diversity, and fitness in space and time.

Acknowledgments

Statement of Authorship

B.B. and A.-L.L. designed the field and molecular studies. B.B. conducted the field and molecular studies. J.L.E., S.S., J.S., E.N., and A.-L.L. analyzed the data. J.L.E. wrote the first draft of the manuscript. All authors contributed to and approved the final version.

Data and Code Availability

Data underlying this article have been deposited in the Dryad Digital Repository (https://doi.org/10.5061/dryad.rjdfn2z94; Eck et al. 2021). Code used to conduct the analyses and generate the figures is provided in a zip file, available online.[1]

Literature Cited

Aguilar-Trigueros, C. A., J. R. Powell, I. C. Anderson, J. Antonovics, and M. C. Rillig. 2014. Ecological understanding of root-infecting fungi using trait-based approaches. Trends in Plant Science 19:432–438.

Alexander, H. M., P. H. Thrall, J. Antonovics, A. M. Jarosz, and P. V. Oudemans. 1996. Population dynamics and genetics of plant disease: a case study of anther-smut disease. Ecology 77:990–996.

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

Barrett, L. G., and M. Heil. 2012. Unifying concepts and mechanisms in the specificity of plant-enemy interactions. Trends in Plant Science 17:282–292.

Barrett, L. G., J. M. Kniskern, N. Bodenhausen, W. Zhang, and J. Bergelson. 2009. Continua of specificity and virulence in plant host-pathogen interactions: causes and consequences. New Phytologist 183:513–529.

Barrett, L. G., P. H. Thrall, J. J. Burdon, A. B. Nicotra, and C. C. Linde. 2008. Population structure and diversity in sexual and asexual populations of the pathogenic fungus *Melamspora lini*. Molecular Ecology 17:3401–3415.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1–48.

Bayman, P., and P. J. Cotty. 1991. Vegetative compatibility and genetic diversity in the *Aspergillus flavus* population of a single field. Canadian Journal of Botany 69:1707–1711.

Biek, R., and L. A. Real. 2010. The landscape genetics of infectious disease emergence and spread. Molecular Ecology 19:3515–3531.

Bos, M. 1992. Gene flow characters and population structure in *Plantago lanceolata*. Pages 222–231 in P. J. C. Kuiper, and M. Bos, eds. Plantago: a multidisciplinary study. Springer, Berlin.

Bousset, L., S. J. Sprague, P. H. Thrall, and L. G. Barrett. 2018. Spatio-temporal connectivity and host resistance influence evolutionary and epidemiological dynamics of the canola pathogen *Leptosphaeria maculans*. Evolutionary Applications 11:1354–1370.

Boyd, I. L., P. H. Freer-Smith, C. A. Gilligan, and H. C. J. Godfray. 2013. The consequences of tree pests and diseases for ecosystem services. Science 342:1235773.

Burdon, J. J., L. Ericson, and W. J. Muller. 1995. Temporal and spatial changes in a metapopulation of the rust pathogen *Triphragmium ulmariae* and its host, *Filipendula ulmaria*. Journal of Ecology 83:979–989.

Burdon, J. J., and A. M. Jarosz. 1991. Host-pathogen interactions in natural populations of *Linum marginale* and *Melamspora lini*. I. Patterns of resistance and racial variation in a large host population. Evolution 45:205–217.

Bushnell, W. R. 2002. The role of powdery mildew research in understanding host-pathogen interaction: past, present, and future. Pages 1–12 in R. R. Bélanger, W. R. Bushnell, A. J. Dik, and L. W. Carver, eds. The powdery mildews: a comprehensive treatise. American Phytopathology Society, St. Paul, MN.

Bürkner, P. 2017. brms: an R package for Bayesian multilevel models using Stan. Journal of Statistical Software 80:1–28. https://doi.org/10.18637/jss.v080.i01.

Carlsson-Granér, U. 1997. Anther-smut disease in *Silene dioica*: variation in susceptibility among genotypes and populations, and patterns of disease within populations. Evolution 51:1416–1426.

Carlsson-Granér, U., and P. H. Thrall. 2002. The spatial distribution of plant populations, disease dynamics and evolution of resistance. Oikos 97:97–110.

Chakraborty, S., and A. C. Newton. 2011. Climate change, plant diseases and food security: an overview. Plant Pathology 60:2–14.

Chase, J. M. 2005. Towards a really unified theory for metacommunities. Functional Ecology 19:182–186.

Chase, J. M., and J. A. Myers. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. Philosophical Transactions of the Royal Society B 366:2351–2363.

Clay, P. A., M. H. Cortez, M. A. Duffy, and V. H. W. Rudolf. 2018. Priority effects within coinfected hosts can drive unexpected population-scale patterns of parasite prevalence. Oikos 128:571–583.

Cobey, S., and M. Lipsitch. 2013. Pathogen diversity and hidden regimes of apparent competition. American Naturalist 181:12–24.

Cotty, P. J. 1989. Virulence and cultural characteristics of two *Aspergillus flavus* strains pathogenic on cotton. Phytopathology 79:808–814.

Croll, D., and B. A. McDonald. 2017. The genetic basis of local adaptation for pathogenic fungi in agricultural ecosystems. Molecular Ecology 26:2027–2040.

Daszak, P., A. A. Cunningham, and A. D. Hyatt. 2000. Emerging infectious diseases of wildlife—threats to biodiversity and human health. Science 287:443–449.

Davies, R. L., and W. Donachie. 1996. Intra-specific diversity and host specificity within *Pasteurella haemolytica* based on variation of capsular polysaccharide, lipopolysaccharide and outer-membrane proteins. Microbiology 142:1895–1907.

DiLeone, J. A., and C. C. Mundt. 1994. Effect of wheat cultivar mixtures on populations of *Puccinia striiformis* races. Plant Pathology 43:917–930.

Ebert, D., C. D. Zschokke-Rohringer, and H. J. Carius. 1998. Within- and between-population variation for resistance of *Daphnia magna* to the bacterial endoparasite *Pasteuria ramosa*. Proceedings of the Royal Society B 265:2127–2134.

Eck, J. L., B. Barrès, S. Soubeyrand, J. Siren, E. Numminen, and A.-L. Laine. 2021. Data from: Strain diversity and spatial distribution are linked to epidemic dynamics in host populations. American Naturalist, Dryad Digital Repository, https://doi.org/10.5061/dryad.rjdfn2z94.

Ericson, L., J. J. Burdon, and W. J. Müller. 1999. Spatial and temporal dynamics of epidemics of the rust fungus *Uromyces valerianae* on populations of its host *Valeriana salina*. Journal of Ecology 87:649–658.

Fels, D., and O. Kaltz. 2006. Temperature dependent transmission and latency of *Holospora undulata*, a micronucleus-specific parasite of the ciliate *Paramecium caudatum*. Proceedings of the Royal Society B 273:1031–1038.

Frank, S. A. 1997. Spatial processes in host-parasite genetics. Pages 325–352 *in* I. A. Hanski, and M. E. Gilpin, eds. Metapopulation biology: ecology, genetics and evolution. Academic Press, New York.

Fukami, T. 2015. Historical contingency in community assembly: integrating niches, species pool and priority effects. Annual Review of Ecology, Evolution, and Systematics 46:1–23.

Garrett, K. A., and C. C. Mundt. 1999. Epidemiology in mixed-host populations. Phytopathology 89:984–990.

Gibson, G. J., A. Kleczkowski, and C. A. Gilligan. 2004. Bayesian analysis of botanical epidemics using stochastic compartmental models. Proceedings of the National Academy of Sciences of the USA 101:12120–12124.

Giraud, T., J. Enjalbert, E. Fournier, F. Delmotte, and C. Dutech. 2008. Population genetics of fungal diseases of plants. Parasite 15:449–454.

Gregory, P. J., S. N. Johnson, A. C. Newton, and J. S. I. Ingram. 2009. Integrating pests and pathogens into the climate change/food security debate. Journal of Experimental Botany 60:2827–2838.

Greischar, M. A., and B. Koskella. 2007. A synthesis of experimental work on parasite local adaptation. Ecology Letters 10:418–434.

Gulland, F. M. D. 1995. The impact of infectious diseases on wild animal populations—a review. Pages 20–51 *in* B. T. Grenfell and A. P. Dobson, eds. Ecology of infectious diseases in natural populations. Cambridge University Press, Cambridge.

Gupta, S., and M. C. J. Maiden. 2001. Exploring the evolution of diversity in pathogen populations. Trends in Microbiology 9:181–185.

Halliday, F. W., R. Penczykowski, B. Barrès, J. L. Eck, E. Numminen and A.-L. Laine. 2020. Facilitative priority effects drive parasite assembly under coinfection. Nature Ecology and Evolution 4:1510–1521. https://doi.org/10.1038/s41559-020-01289-9.

Halliday, F. W., J. Umbanhowar, and C. E. Mitchell. 2017. Interactions among symbionts operate across scales to influence parasite epidemics. Ecology Letters 20:1285–1294.

Hanski, I. 1999. Metapopulation ecology. Oxford University Press, Oxford.

Hiremath, S. J., A. Chowdhary, T. Kowshik, H. S. Randhawa, S. Sun, and J. Xu. 2008. Long-distance dispersal and recombination in environmental populations of *Cryptococcus neoformans* var. *grubii* from India. Microbiology 154:1513–1524.

Höckerstedt, L. M., J. P. Siren, and A.-L. Laine. 2018. Effect of spatial connectivity on host resistance in a highly fragmented natural pathosystem. Journal of Evolutionary Biology 31:844–852.

Hoeksema, J. D., and S. E. Forde. 2008. A meta-analysis of factors affecting local adaptation between interacting species. American Naturalist 171:275–290.

Ichielevich-Auster, M., B. Sneh, Y. Koltin, and I. Barash. 1985. Pathogenicity, host specificity and anastomosis groups of *Rhizoctonia* spp. isolated from soils in Israel. Phytoparasitica 13:103–112.

Johnson, D. W., J. Freiwald, and G. Bernardi. 2016. Genetic diversity affects the strength of population regulation in a marine fish. Ecology 97:627–639.

Johnson, P. T. J., J. C. de Roode, and A. Fenton. 2015. Why infectious disease research needs community ecology. Science 349:1069–1077.

Jousimo, J., A. J. M. Tack, O. Ovaskainen, T. Mononen, H. Susi, C. Tollenaere, and A.-L. Laine. 2014. Ecological and evolutionary effects of fragmentation on infectious disease dynamics. Science 344:1289–1293.

Kaltz, O., and J. A. Shykoff. 2002. Within- and among-population variation in infectivity, latency, and spore production in a host-pathogen system. Journal of Evolutionary Biology 15:850–860.

Khanna, M., P. Kumar, K. Choudhary, B. Kumar, and V. K. Vijayan. 2008. Emerging influenza virus: a global threat. Journal of Biosciences 33:475–482.

Kirchner, J. W., and B. A. Roy. 2002. Evolutionary implications of host-pathogen specificity: fitness consequences of pathogen virulence traits. Evolutionary Ecology Research 4:27–48.

Koskella, B., T. Giraud, and M. E. Hood. 2006. Pathogen relatedness affects the prevalence of within-host competition. American Naturalist 168:121–126.

Krasnov, B. R., G. I. Shenbrot, I. S. Khokhlova, M. Stanko, S. Morand, and D. Mouillot. 2015. Assembly rules of ectoparasite communities across scales: combining patterns of abiotic factors, host composition, geographic space, phylogeny and traits. Ecography 38:184–197.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. lmerTest package: tests in linear mixed effects models. Journal of Statistical Software 82:1–26. https://doi.org/10.18637/jss.v082.i13.

Laine, A.-L. 2004. Resistance variation within and among host populations in a plant-pathogen metapopulation: implications for regional pathogen dynamics. Journal of Ecology 92:990–1000.

———. 2005. Spatial scale of local adaptation in a plant-pathogen metapopulation. Journal of Evolutionary Biology 18:930–938.

———. 2007. Pathogen fitness components and genotypes differ in their sensitivity to nutrient and temperature variation in a wild plant-pathogen association. Journal of Evolutionary Biology 20:2371–2378.

———. 2008. Temperature-mediated patterns of local adaptation in a natural plant-pathogen metapopulation. Ecology Letters 11:327–337.

Laine, A.-L., B. Barrès, E. Numminen, and J. P. Siren. 2019. Variable opportunities for outcrossing result in hotspots of novel genetic variation in a pathogen metapopulation. eLIFE 8:e47091. https://doi.org/10.7554/eLife.47091.

Laine, A.-L., J. J. Burdon, P. N. Dodds, and P. H. Thrall. 2011. Spatial variation in disease resistance: from molecules to metapopulations. Journal of Ecology 99:96–112.

Laine, A.-L., and I. Hanski. 2006. Large-scale spatial dynamics of a specialist plant pathogen in a fragmented landscape. Journal of Ecology 94:217–226.

Leung, T. L. F., and A. E. Bates. 2013. More rapid and severe disease outbreaks for aquaculture at the tropics: implications for food security. Journal of Applied Ecology 50:215–222.

Lowe, W. H., R. P. Kovach, and F. W. Allendorf. 2017. Population genetics and demography unite ecology and evolution. Trends in Ecology and Evolution 32:141–152.

Marçais, B., and M.-L. Desprez-Loustau. 2014. European oak powdery mildew: impact on trees, effects of environmental factors, and potential effects of climate change. Annals of Forest Science 71:633–642.

McCartney, H. A., and B. D. L. Fitt. 1998. Dispersal of foliar plant pathogens: mechanisms, gradients and spatial patterns. Pages 138–160 in B. M. Cook, D. G. Jones, and B. Kaye, eds. The epidemiology of plant diseases. Springer, Dordrecht.

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

Meester, L. D., J. Vanoverbeke, L. J. Kilsdonk, and M. C. Urban. 2016. Evolving perspectives on monopolization and priority effects. Trends in Ecology and Evolution 31:136–146.

Mitchell, S. E., E. S. Rogers, T. J. Little, and A. F. Read. 2005. Host-parasite and genotype-by-environment interactions: temperature modifies potential for selection by a sterilizing pathogen. Evolution 59:70–80.

Mundt, C. C., and K. J. Leonard. 1985. A modification of Gregory's model for describing plant disease gradients. Phytopathology 75:930–935.

Nagy, T., W. P. Pfliegler, A. Takács, J. Tökölyi, and A. Molnár V. 2019. Distribution, infection rates and DNA barcoding of *Uromyces erythronii* (Pucciniaceae), a parasite of *Erythronium* (Liliaceae) in Europe. Willdenowia 49:13–20.

Numminen, E., E. Vaumourin, S. R. Parratt, L. Poulin, and A.-L. Laine. 2019. Variation and correlations between sexual, asexual and natural enemy resistance life-history traits in a natural plant pathogen population. BMC Evolutionary Biology 19:142. https://doi.org/10.1186/s12862-019-148-2.

Orum, T. V., D. M. Bigelow, and M. R. Nelson. 1997. Spatial and temporal patterns of *Aspergillus flavus* strain composition and propagule density in Yuma County, Arizona, soils. Plant Disease 81:911–916.

Oskanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, et al. 2017. vegan: community ecology package. R package version 2.4-5. https://CRAN.R-project.org/package=vegan.

Osnas, E. E., P. J. Hurtado, and A. P. Dobson. 2015. Evolution of pathogen virulence across space during an epidemic. American Naturalist 185:332–342.

Ostfeld, R. S., and F. Keesing. 2012. Effects of host diversity on infectious disease. Annual Review of Ecology, Evolution, and Systematics. 43:157–182.

Ovaskainen, O., and A.-L. Laine. 2006. Inferring evolutionary signals from ecological data in a plant-pathogen metapopulation. Ecology 87:880–891.

Palmer, G. H., F. R. Rurangirwa, and T. F. McElwain. 2001. Strain composition of the Ehrlichia *Anaplasma marginale* within persistently infected cattle, a mammalian reservoir for tick transmission. Journal of Clinical Microbiology 39:631–635.

Parratt, S. P., E. Numminen, and A.-L. Laine. 2016. Infectious disease dynamics in heterogeneous landscapes. Annual Review of Ecology, Evolution, and Systematics 47:283–306.

Penczykowski, R. M., A.-L. Laine, and B. Koskella. 2015. Understanding the ecology and evolution of host-parasite interactions across scales. Evolutionary Applications 9:37–52.

Penczykowski, R. M., S. R. Parratt, B. Barrès, S. K. Sallinen and A.-L. Laine. 2018. Manipulating host resistance structure reveals impact of pathogen dispersal and environmental heterogeneity on epidemics. Ecology 99:2853–2863.

R Development Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Rivers, C. M., and S. V. Scarpino. 2018. Predict trajectory, not emergence. Nature 559:477.

Sagar, G. R., and J. L. Harper. 1964. *Plantago major* L., *P. media* L. and *P. lanceolata* L. Journal of Ecology 52:189–221.

Salvaudon L., T. Giraud, and J. A. Shykoff. 2008. Genetic diversity in natural populations: a fundamental component of plant-microbe interactions. Current Opinion in Plant Biology 11:135–143.

Sapoukhina, N., C.-E. Durel, and B. Le Cam. 2009. Spatial deployment of gene-for-gene resistance governs evolution and spread of pathogen populations. Theoretical Ecology 2:229–238.

Scholz, M., D. V. Ward, E. Pasolli, T. Tolio, M. Zolfo, F. Asnicar, D. T. Truong, A. Tett, A. L. Morrow, and N. Segata. 2016. Strain-level microbial epidemiology and population genomics from shotgun metagenomics. Nature Methods 13:435–438. https://doi.org/10.1038/nmeth.3802.

Scott, M. E. 1988. The impact of infection and disease on animal populations: implications for conservation biology. Conservation Biology 2:40–56.

Shaw, M. W. 1995. Simulation of population expansion and spatial pattern when individual dispersal distributions do not

decline exponentially with distance. Proceedings of the Royal Society B 259:243–248.

———. 2006. Pathogen population dynamics. Pages 193–214 *in* B. M. Cooke, D. G. Jones, and B. Kaye, eds. The epidemiology of plant diseases. Springer, Dordrecht.

Shipley, B., C. E. T. Paine, and C. Baraloto. 2012. Quantifying the importance of local niche-based and stochastic processes to tropical tree community assembly. Ecology 93:760–769.

Soubeyrand, S., V. Garreta, C. Monteil, F. Suffert, H. Goyeau, J. Berder, J. Moinard, et al. 2017. Testing differences between pathogen compositions with small samples and sparse data. Phytopathology 107:1199–1208.

Soubeyrand, S., C. Tollenaere, E. Haon-Lasportes, and A.-L. Laine. 2014. Regression-based ranking of pathogen strains with respect to their contribution to natural epidemics. PLoS ONE 9:e86591.

Susi, H., B. Barrès, P. F. Vale, and A.-L. Laine. 2015. Co-infection alters population dynamics of infectious disease. Nature Communications 6:5975. https://doi.org/10.1038/ncomms6975.

Susi, H., and A.-L. Laine. 2015. The effectiveness and costs of pathogen resistance strategies in a perennial plant. Journal of Ecology 103:303–315.

Tack, A. J. M., J. Hakala, T. Petaja, M. Kulmala, and A.-L. Laine. 2014. Genotype and spatial structure shape pathogen dispersal and disease dynamics at small spatial scales. Ecology 95:703–714.

Tack, A. J. M., F. Horns, and A.-L. Laine. 2013. The impact of spatial scale and habitat configuration on patterns of trait variation and local adaptation in a wild plant parasite. Evolution 68:176–189.

Tack, A. J. M., and A.-L. Laine. 2014*a*. Ecological and evolutionary implications of spatial heterogeneity during the off-season for a wild plant pathogen. New Phytologist 202:297–308.

———. 2014*b*. Spatial eco-evolutionary feedback in plant-pathogen interactions. European Journal of Plant Pathology 138:667–677.

Thompson, J. N. 2005. The geographic mosaic of coevolution. University of Chicago Press, Chicago.

Thompson, J. N., and J. J. Burdon. 1992. Gene-for-gene coevolution between plants and parasites. Nature 360:121–125.

Thomas, M. B., and S. Blanford. 2003. Thermal biology in insect-parasite interactions. Trends in Ecology and Evolution 18:344–350.

Thrall, P. H., and J. J. Burdon. 2000. Effect of resistance variation in a natural plant host-pathogen metapopulation on disease dynamics. Plant Pathology 49:767–773.

———. 2003. Evolution of virulence in a plant host-pathogen metapopulation. Science 299:1735–1737.

Thrall, P. H., and A. M. Jarosz. 1994. Host-pathogen dynamics in experimental populations of *Silene alba* and *Ustilago violacea*. I. Ecological and genetic determinants of disease spread. Journal of Ecology 82:549–559.

Tollenaere, C., and A.-L. Laine. 2013. Investigating the production of sexual resting structures in a plant pathogen reveals unexpected self-fertility and genotype-by-environment effects. Journal of Evolutionary Biology 26:1716–1726.

Tollenaere, C., H. Susi, J. Nokso-Koivisto, P. Koskinen, A. Tack, P. Auvinen, L. Paulin, M. J. Frilander, R. Lehtonen, and A.-L. Laine. 2012. SNP design from 454 sequencing of *Podosphaera plantaginis* transcriptome reveals a genetically diverse pathogen metapopulation with high levels of mixed-genotype infection. PLoS ONE 7:e52492.

Truong, D. T., A. Tett, E. Pasolli, C. Huttenhower, and N. Segata. 2017. Microbial strain-level population structure and genetic diversity from metagenomes. Genome Research 27:626–638.

Twizeyimana, M., P. S. Ojiambo, K. Sonder, T. Ikotun, G. L. Hartman, and R. Bandyopadhyay. 2009. Pathogen variation of *Phakopsora pachyrhizi* infecting soybean in Nigeria. Phytopathology 99:353–361.

Vaumourin, E., and A.-L. Laine. 2018. Role of temperature and coinfection in mediating pathogen life-history traits. Frontiers in Plant Science 9:1670. https://doi.org/10.3389/fpls.2018.01670.

Vehtari, A., A. Gelman, and J. Gabry. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing 27:1413–1432.

Vredenburg, V. T., R. A. Knapp, T. S. Tunstall, and C. J. Briggs. 2010. Dynamics of an emerging disease drive large-scale amphibian population extinctions. Proceedings of the National Academy of Sciences of the USA 107:9689–9694.

Wang, Y., Y. Meng, M. Zhang, X. Tong, Q. Wang, Y. Sun, J. Quan, F. Govers, and W. Shan. 2011. Infection of *Arabidopsis thaliana* by *Phytopthora parasitica* and identification of variation in host specificity. Molecular Plant Pathology 12:187–201.

Wolinska, J., and K. C. King. 2009. Environment can alter selection in host-parasite interactions. Trends in Parasitology 25:236–244.

Woodhams, D. C., R. A. Alford, C. J. Briggs, M. Johnson, and L. A. Rollins-Smith. 2008. Life-history trade-offs influence disease in changing climates: strategies of an amphibian pathogen. Ecology 89:1627–1639.

Yach, D., D. Stuckler, and K. D. Brownell. 2006. Epidemiological and economic consequences of the global epidemics of obesity and diabetes. Nature Medicine 12:62–66.

Zhan, J., and B. A. McDonald. 2013. Experimental measures of pathogen competition and relative fitness. Annual Review of Phytopathology 51:131–153.

Zhang, X.-S., H. Zhao, E. Vynnycky, and V. Chalker. 2019. Positively interacting strains that co-circulate within a network structured population induce cycling epidemics of *Mycoplasma pneumoniae*. Scientific Reports 9:541. https://doi.org/10.1038/s41598-018-36325-z.

References Cited Only in the Online Enhancements

Agresti, A. 2007. An introduction to categorical data analysis. Wiley, Chichester, United Kingdom. https://doi.org/10.1002/0470114754.

Associate Editor: Bret D. Elderd
Editor: Daniel I. Bolnick