A Two-Locus System with Strong Epistasis Underlies Rapid Parasite-Mediated Evolution of Host Resistance

Camille Ameline,1,2 Yann Bourgeois,1,2 Felix Vögtli,1 Eevi Savola,1,3 Jason Andras,1,4 Jan Engelstädter,5 and Dieter Ebert1

1Department of Environmental Sciences, Zoology, University of Basel, Basel, Switzerland
2School of Biological Sciences, University of Portsmouth, Portsmouth, United Kingdom
3Institute of Evolutionary Biology, Ashworth Laboratories, University of Edinburgh, Edinburgh, United Kingdom
4Department of Biological Sciences, Clapp Laboratory, Mount Holyoke College, South Hadley, MA, USA
5School of Biological Sciences, The University of Queensland, Brisbane, QLD, Australia

*Corresponding author: E-mail: cameline8@gmail.com.
Associate editor: Deepa Agashe

Abstract

Parasites are a major evolutionary force, driving adaptive responses in host populations. Although the link between phenotypic response to parasite-mediated natural selection and the underlying genetic architecture often remains obscure, this link is crucial for understanding the evolution of resistance and predicting associated allele frequency changes in the population. To close this gap, we monitored the response to selection during epidemics of a virulent bacterial pathogen, Pasteuria ramosa, in a natural host population of Daphnia magna. Across two epidemics, we observed a strong increase in the proportion of resistant phenotypes as the epidemics progressed. Field and laboratory experiments confirmed that this increase in resistance was caused by selection from the local parasite. Using a genome-wide association study, we built a genetic model in which two genomic regions with dominance and epistasis control resistance polymorphism in the host. We verified this model by selfing host genotypes with different resistance phenotypes and scoring their F1 for segregation of resistance and associated genetic markers. Such epistatic effects with strong fitness consequences in host–parasite coevolution are believed to be crucial in the Red Queen model for the evolution of genetic recombination.

Key words: parasite-mediated selection, zooplankton, resistance, genetic architecture, epistasis, dominance, multilocus genetics, Daphnia magna, Pasteuria ramosa.

Introduction

Darwinian evolution is a process in which the phenotypes that are best adapted to the current environment produce more offspring for the next generation. Genetic variants that code for these phenotypes are thus expected to increase in frequency in the population. Although this concept is fundamental in evolutionary biology, it remains difficult to connect the phenotype under selection with the underlying changes in the gene pool of natural populations (Ellegren and Sheldon 2008; Whitlock and Lotterhos 2015; Hoban et al. 2016). Although single-gene effects have been shown to explain the phenotype–genotype interplay in some naturally evolving populations (Daborn 2002; Cao et al. 2016; van’t Hof et al. 2016), the genetic architecture underlying a phenotype is often complex. In addition, the way the environment influences the expression of a trait, and genotype × environment interactions may further obscure the link between phenotype and genotype. It is, thus, often impossible to predict genetic changes in a population that result from selection on specific phenotypes. Among the most potent drivers of evolutionary change in host populations are parasites; parasite-mediated selection can raise the frequency of resistant phenotypes rapidly (Schmid-Hempel 2011; Kurtz et al. 2016; Morgan and Koskella 2017; Koskella 2018) and is thought to contribute to many biological phenomena, such as biodiversity (Laine 2009), speciation (Schlesinger et al. 2014), and the maintenance of sexual recombination in the host (Lively 2010; Gibson et al. 2018).

To link patterns produced by parasite-mediated selection with evolutionary theory, we need to know the genetic architecture that underlies resistance; this includes the number of loci, their relative contribution to the phenotype, and the interaction between loci (epistasis) and alleles (dominance). In this way, we may be able to predict the outcome of selection, test theoretical models, and understand epidemiological dynamics (Hamilton 1980; Galvani 2003; Schmid-Hempel 2011). In a few cases, resistance to parasites has been found to be determined by single loci with strong effects, for example, in plants (Gómez-Gómez et al. 1999; Li and Cowling 2003; Li et al. 2017), invertebrates (Juneja et al. 2015; Xiao et al.
Rapid Parasite-Mediated Evolution of Host Resistance.

In recent years, water fleas of the genus *Daphnia* (Crustacea, Cladocera) and their microparasites have become one of the best understood systems for studying the evolution and ecology of host–parasite interactions (Ebert 2005; Vale et al. 2011; Izhar and Ben-Ami 2015; González-Tortuero et al. 2016; Strauss et al. 2017; Shocket et al. 2018; Turko et al. 2018; Rogalski and Duffy 2020). Parasite selection in natural *Daphnia* populations has been shown to alter the phenotypic distribution of resistance (Little and Ebert 1999; Decaestecker et al. 2007; Duffy and Sivars-Becker 2007; Duncan and Little 2007), and genetic mapping studies identified loci involved in host resistance (Luijckx et al. 2012, 2013; Routtu and Ebert 2015; Metzger et al. 2016; Bento et al. 2017, 2020) and parasite infectivity (Andras et al. 2020); however, because studies on host resistance largely involved crosses among populations, the results may not reflect genetic variation within populations. Genetic changes in natural host populations have been observed but so far it was not possible to link this change to parasite resistance loci (Mitchell et al. 2004; Duncan and Little 2007). Understanding the link between parasite-mediated selection on host resistance and the underlying genetic architecture would enable us to determine and predict the tempo and mode of evolution in natural populations and to link observed phenotypic changes to frequency changes of alleles under selection. This study provides such a phenotype–genotype link. We quantified the change in frequency of resistance phenotypes over time in a natural *D. magna* population and, through experiments, showed that the locally dominant, virulent parasite genotype of *P. ramosa* played a major role in the observed phenotypic changes. A genome-wide association study (GWAS) and genetic crosses revealed the underlying genetic architecture of resistance in our study population and provided a genetic model for inheritance of resistance. This genetic model comprises two resistance loci presenting distinct dominance patterns and strongly linked with epistasis. These results strongly support the Red Queen model of host–parasite coevolution and the maintenance of genetic recombination.

**Results**

**Parasite-Mediated Selection Explains Phenotypic Dynamics**

**Monitoring**

We monitored the Aegelsee *D. magna* population from fall 2010 to fall 2015, whereas the present study focuses on the 2014 and 2015 planktonic seasons. In this population, *D. magna* diapausa during winter as resting eggs, whereas the active season spans from early April to early October. Each summer, we observed a *P. ramosa* epidemic that typically started in early May, about a month after *Daphnia* emerged from diapause, and lasted throughout the summer (fig. 1A) with peak prevalence of 70–95%. *Pasteuria ramosa* infection in the host is characterized by gigantism, a reddish-brownish opaque coloration, and castration, that is, an empty brood pouch. *Pasteuria ramosa* is a virulent parasite, stripping the host of 80–90% of its residual reproductive success and

2017), and vertebrates (Samson et al. 1996). However, the genetic architecture is often obscured by intrinsic complexity and confounding factors that may influence the phenotype. Resistance might be determined by multiple loci with qualitative or quantitative effects, present distinct dominance patterns, and display interactions with other genes or the environment. Indeed, multilocus genetic architecture of resistance can create more diversity, and is thus thought to be more common than single loci (Sasaki 2000; Tellier and Brown 2007; Wilfert and Schmid-Hempel 2008). Multilocus architecture was described in *Drosophila melanogaster*, for example, where resistance was found to be determined mostly by a few large-effect loci (Bingham et al. 2008; Magwire et al. 2012) and some additional small-effect loci (Cogni et al. 2016; Magalhães and Sucena 2016). Quantitative resistance has also been found in crops where it may be used as a pathogen control strategy (Pilet-Nayel et al. 2017). In the water flea *Daphnia magna*, resistance has been found to be quantitative to a microsporidian parasite, but qualitative to a bacterial pathogen (Routtu and Ebert 2015). Although resistance tends to be dominant (Hooker and Saxena 1971; Carton et al. 2005), resistant alleles have been found to be both dominant and recessive in plants (Gómez-Gómez et al. 1999; Li and Cowling 2003; Li et al. 2017) and invertebrates (Luijckx et al. 2012; Juneja et al. 2015; Xiao et al. 2017). Epistasis between resistance loci has also been found in diverse plants and animals (Kover and Caicedo 2001; Wilfert and Schmid-Hempel 2008; Jones et al. 2014; González et al. 2015; Metzger et al. 2016), emphasizing its crucial role in the evolution of resistance. The link between genetic architecture and natural selection for resistance remains weak, however, mainly limited to the theoretical extrapolation of results from laboratory experiments.

Dominance and epistasis describe the nonadditive interaction among alleles of the same or different loci, respectively, making them crucial for the evolutionary response to selection. Epistasis among resistance genes could contribute to the maintenance of genetic diversity by reducing fixation rates at individual loci, and is thus thought to be pervasive (Tellier and Brown 2007). In the Red Queen model for the evolution of sex, thus, epistasis among resistance loci helps maintain genetic diversity and recombination in the host (Hamilton 1980; Hamilton et al. 1990; Howard and Lively 1998; Salathé et al. 2008; Engelstädter and Bonhoeffer 2009; Kouyos et al. 2009). Important with regard to the role of epistasis for the evolution of host–parasite interactions is furthermore, that the interacting loci must be polymorphic within the same natural populations. However, the importance of genetic architecture for understanding the evolution of resistance stands in stark contrast to the limited amount of available data on natural populations (Alves et al. 2019). In this study, we investigate the evolution of resistance in a natural population of the planktonic crustacean *D. magna* as it experiences epidemics of the virulent bacterial pathogen *Pasteuria ramosa*. We link parasite-mediated selection to its associated allele frequency change by resolving the underlying genetic architecture of host resistance.
In 2011, we sampled a subset of infected animals \((n = 113)\) to characterize \(P. \) ramosa diversity among infected hosts throughout the active season and found that the P20 genotype represented about 50% of the parasite diversity among infected hosts when the epidemics began. This proportion decreased to zero during the epidemic, as other \(P. \) ramosa genotypes took over (supplementary fig S1, Supplementary Material online).

The temporal dynamics revealed an increase in animals resistant to P20 (RRSR and RRRR, in short: RR- \(\rightarrow\) R) soon after the onset of the epidemics, whereas animals susceptible to P20 (RRSS and SSSS, or \(\ldots\)S) declined accordingly (fig. 18) in both study years. Resistance to C1, C19, and P15 did not seem to play a strong role in the selection process during the epidemics. In the result described next, we tested the hypothesis that selection by \(P. \) ramosa isolate P20 is the main driver of natural resistotype dynamics in our study population during the early planktonic season. As a reminder, P20 has been isolated from a spring sample of the here-studied population.

**Experimental and Field Infections**

First, we tested the impact of the parasite on the different resistotypes to associate disease phenotype with resistotype. To do this, we obtained a sample of the spring cohort of the \(D. \) magna population by hatching resting eggs collected in February 2014. These animals represented, in total, 70 clones of the four most common resistotypes (RRSR, RRSS, SSSS, RRRR), with each clone replicated five times \((n = 350)\). We exposed these clonal offspring to a mixture of \(P. \) ramosa spores that represented the diversity of the parasite population during the early phase of the epidemic. We then monitored the hosts for infection (looking for visible signs) and fecundity (counting the number of produced clutches). Sixteen animals died before we could test their infection status, resulting in a total sample size of \(n = 334\). Individuals with resistotypes RRSR and SSSS (susceptible to P20) were infected far more frequently than RRSR and RRRR (resistant to P20) individuals (fig. 2A; null deviance = 461.3 on 333 df, residual deviance = 358.4 on 329 df, \(P < 0.001\)). The analysis also compared P20-susceptible and P20-resistant resistotypes, confirming the high susceptibility of the P20-susceptible animals (fig. 2A; null deviance = 461.3 on 333 df, residual deviance = 360.7 on 331 df, \(P < 0.001\)). Infected P20-susceptible individuals produced on an average about one less clutch before parasitic castration \((n = 136, 1.83 \pm 0.07 \) clutches\) than did infected P20-resistant individuals \((n = 19, 2.53 \pm 0.3 \) clutches\) (fig. 2B; null deviance = 76.9 on 154 df, residual deviance = 74.0 on 152 df, \(P = 0.023\)). Accordingly, the average time period until visible infection was shorter in P20-susceptible clones \((15.7 \pm 0.2 \) days\) than in P20-resistant clones \((19.4 \pm 1 \) days\) (fig. 2C; null deviance = 94.4 on 154 df, residual deviance = 85.1 on 152 df, \(P = 0.0018\)). These results clearly support the hypothesis that early season \(P. \) ramosa strains from the field select on the P20 resistotype.

In the following year, we looked at the relationship of disease phenotype and P20 resistotype only in the field by
rates (size corrected) for P20-susceptible than for P20-

clones for each), and RRRR (trophotypes (resistance phenotype). Resistotypes RRSR, RRSS, SSSS (magna
among the four resistotypes. (A) Proportion of infected D. magna among the four resistotypes. (B) Number of clutches produced before parasitic castration in the infected P20-resistant (-----R) and susceptible (----S) animals (n = 115). (C) Time before visible infec-
in P20-resistant and P20-susceptible individuals (n = 115).

measuring the parasite’s impact on P20-resistant and P20-

susceptible hosts. We collected animals in the field during the early half of the P. ramosa epidemic and raised them indi-

vidually in the laboratory, recording their disease symp-
toms. We then cured infected animals with antibiotics,
allowed them to produce clonal offspring, and deter-
determined their P20 resistotype. Our analysis revealed higher infection rates (size corrected) for P20-resistant than for P20-

resistant individuals in these natural conditions (fig. 3A; Fitted model: glm [Infected (1/0) ~ P20 resistotype + Body_size + Sampling_date], family = quasibinomial(), n = 331; null deviance = 415.1 on 330 df, residual devi-
ance = 209.1 on 327 df, P = 0.025). Field-caught infected P20-susceptible individuals also produced, on an average,

fewer offspring before parasitic castration than infected P20-resistant ones (fig. 3B; Fitted model: glm.nb [Fecundity ~ P20 resistotype + Body_size * Sampling_date], n = 224; null deviance = 127.9 on 223 df, residual deviance = 92.9 on 219 df, P = 0.014). In both models, the sampling date also had a significant effect. Parasite preva-

lence on the two sampling dates differed strongly (31% on

June 7 and 96% on June 28, 2015). We observed on the first

sampling date that larger individuals were more infected and consequently produced less offspring. This size difference is not visible anymore on the second sampling date, where almost all individuals were infected. The overall pattern in relation to the P20 resistotype remained the same, even though the difference in infection and fecundity between

field-collected P20 resistotypes was less pronounced than in the controlled infection experiment (compare figs. 2 and 3).

In summary, the results of the two experiments clearly sup-
port the hypothesis that early season P. ramosa from the field select on the P20 resistotype.

Linking Resistance Phenotypes to Genotypes

Excluding the P15 resistotype, which has very low variability because most animals are P15-susceptible, the study popu-

lation was composed mainly of three resistotypes: RR-R, RR-S, and SS-S. A supergene for resistance to C1 and C19 has been described in D. magna using QTL mapping (Roultt and Ebert 2015; Bento et al. 2017), and the genetic architecture of resistance at this so-called ABC-cluster, or P. ramosa resistance (PR) locus, has been further resolved using genetic crosses among host genotypes (Metzger et al. 2016). According to this genetic model, an SS--- resistotype (susceptible to C1 and C19) has an “aabccc” genotype (lower case letters indicate recessive alleles), whereas RS--- individuals (resistant to C1 and susceptible to C19) are “A---cc” (upper case letters indicate dominant alleles and a dash “-” indicates alleles that do not influence the phenotype); SR--- individuals are “aB-cc,” and R--- individuals are “---C-.” In other words, allele A epistatically nullifies variation at the B locus, and allele C nullifies variation at the A and B loci (Metzger et al. 2016; Bento et al. 2017). See also supplemen-
tary figure S2, Supplementary Material online. Considering this genetic model, we assume that the recessive allele at the A locus is fixed in our study population (“aa” genotype)
and that the dominant allele at the B locus is very rare, as we never observed RS_... individuals and only found SR_... in very low proportions. In our study population, the SS_.../RR_... polymorphism can therefore be best described by the C-locus polymorphism, that is, genotypes “aabbcc” and “aabbCc,” respectively, with C being the dominant allele for resistance. Given this, we assume, in the following sections, that variation at the C locus underlies the resistance polymorphism for C1 and C19.

Genomic Regions of Resistance to the Parasite

We sequenced the genomes of 16, 10, and 11 clones with resistotypes RR_R, RR_S, and SS_S, respectively and conducted a GWAS comparing five pairs of these resistotypes to identify candidates for resistance to C1, C19, and P20: (i) SS__ versus RR__, (ii) SS_S versus RR__S, (iii) __S versus __R, (iv) RR__S versus RR__R and (v) SS_S versus RR__R. Comparisons (i) and (ii) (variation at C1 and C19 resistotypes) revealed a strong signal on linkage group (LG) 3 (fig. 4A and B). This region encompasses the super gene described earlier by Routtu and Ebert (2015) and Bento et al. (2017), the so-called ABC-cluster, or PR locus. Comparisons (iii) and (iv) (variation at P20 resistotype) revealed a strong signal on LG 5 (fig. 4C and D), hereafter called the E-locus region. In the present host–parasite system, the D locus determines resistance to P15 and is not considered here (Bento et al. 2020). The E-locus region has not yet been associated with resistance, and no PR gene has been described on the same linkage group in D. magna. Finally, comparison (v) (variation at C1, C19, and P20 resistotypes) indicated a strong signal at both the ABC cluster and the E-locus region (fig. 4E). The genomic regions associated with resistotypes in our GWAS were not sharp peaks, but rather table-like blocks of associated SNPs (fig. 4). This structure was expected for the C locus, which is a known supergene—a large block of genome space with apparently little or no recombination that contains many genes (Bento et al. 2017). Figure 4 indicates that the same may be the case for the E-locus region, where the block of associated SNPs makes up nearly half of the linkage group. A few single SNPs also showed significant association in all the comparisons (fig. 4), but because of the strength of the observed pattern and because we expected a large region to be associated with resistance, we do not consider these single SNPs further.

The E-locus region encompassed 22 scaffolds and one contig on version 2.4 of the D. magna reference genome, with a cumulative length of more than 3 Mb (3,101,076 bp) (supplementary table S1, Supplementary Material online). We found 485 genes on all associated scaffolds. The strongest signals of association were found on scaffolds 2,167 and 2,560, which harbored 82 genes. Some of these genes were similar to genes identified in a previous study of the ABC cluster on LG 3 (Bento et al. 2017), with a glucosyltransferase found on scaffold 2,167. Three other sugar transferases (galactosyltransferases) were identified, two of them on scaffold 2,560 (supplementary table S2, Supplementary Material online).

Genetic Model of Resistance Inheritance

Mean allele frequencies at associated SNPs showed that SS__ individuals (susceptible to C1 and C19) display a single allele at the C locus, whereas RR__ individuals display two distinct alleles at the C locus. This suggests that SS__ individuals are homozygous at the C locus, whereas RR__ individuals comprise homo- and heterozygotes. At the E locus, __S__R individuals (resistant to P20) are presumably homozygous, whereas __S__S individuals (susceptible to P20) comprise homo- and heterozygous individuals (fig. 4, right panel). These results indicate that resistance to C1 and C19 is governed by a dominant allele (“C-” genotype), as shown before (Metzger et al. 2016). In contrast, resistance to P20 is determined by a recessive allele (“ee” genotype), as was shown before for a different resistance locus (D locus, Bento et al. 2020). Screening individual genomes revealed that some SS__S individuals (susceptible to C1 and C19, and P20) present the “ee” genotype at the E locus (underlying P20 resistotype), although this genotype should confer resistance to P20. This was not observed in RR__S individuals (resistant to C1 and C19, but susceptible to P20) (supplementary table S3, Supplementary Material online), which we hypothesize to be explained by an epistatic relationship linking the C and the E loci. This epistasis confers P20 susceptibility to individuals susceptible to C1 and C19, that is, presenting the “cc” genotype, regardless of their genotype at the E locus. This genetic model is presented in figure 5 (without variation at the B locus, see below). In the present study, we mostly considered variation at the C and E loci, as they seem to play a major role in the diversity of resistotypes in our study population.

To test the genetic model derived from the GWAS, we investigated segregation of resistotypes among selfed offspring of D. magna genotypes with diverse resistotypes. Daphnia magna reproduces by cyclical parthenogenesis, in which asexual eggs produce clonal lines and sexual reproduction allows to perform genetic crosses. Our genetic model allowed us to predict the segregation of genotypes and phenotypes, which can then be compared with the observed segregation patterns among selfed offspring. From 24 host genotypes (F0 parent clones), we produced 24 groups of selfed F1 offspring. Twenty-two F0 clones included animals with all possible combinations of alleles at the C and E loci, whereas two F0s showed the rare variation at the B locus and variation at the E locus. Expected and observed resistotype frequencies are presented in tables 1 and 2 and detailed for each F1 group in supplementary tables S4–S15, Supplementary Material online. In the 22 F1 groups showing variation at the C and E loci, segregation of offspring followed the predictions of our genetic model (fig. 5), that is, we observed all expected resistotypes and saw no significant deviations from the expected frequencies based on the model. These data clearly support the genetic model for resistance at the C and E loci.

As described above, earlier research (Metzger et al. 2016; Bento et al. 2017) has shown that the dominant allele at the C locus interacts epistatically with the A and B loci (all are part of the ABC cluster), such that variation at the A and B loci...
individuals present the "ee" genotype at the E locus. In these groups, the frequency of the dominant allele can be lower than 0.5.

Screening individual genomes revealed that some SS individuals (dominant allele frequency between 0.5 and 1). This shows that, in contrast with the C locus, susceptibility is dominant at the E locus.

Resistotype groups resistant to C1 and C19 did contain the dominant allele (frequency between 0.5 and 1), showing that resistance is dominant at the E locus.

The figure 4 shows the GWAS analysis comparing the most common resistance phenotypes (resistotypes) in the Aegelsee Daphnia magna population. The resistotype depicts resistance (R) or susceptibility (S) to Pasteuria ramosa isolates C1, C19, P15, and P20. (i) SS— versus RR--; (ii) SS—S versus RR--; (iii) S—S versus S—R; (iv) RR—S versus RR--; and (v) SS—S versus RR--;. Comparisons (i) and (ii) (variation at C1 and C19 resistotypes) revealed a strong signal on linkage group (LG) 3 corresponding to the C locus. Comparisons (iii) and (iv) (variation at P20 resistotype) revealed a strong signal on LG 5 corresponding to the E locus. Comparison (v) (variation at C1 and C19, and P20 resistotypes) revealed a strong signal on both regions. Left panel: Manhattan plots of relationships between different resistotype groups (showing only SNPs with P_corrected < 0.01). The x axis corresponds to SNP data mapped on the 2.4 D. magna reference genome (Rouutu et al. 2014), representing only SNPs, not physical distance on the genome. Middle panel: Quantile-quantile plots of noncorrected P values excluding SNPs from linkage groups 3 and 5, since these scaffolds displayed an excess of strongly associated markers. Right panel: Comparison of allele frequencies between resistotype groups at the C and the E loci. Significant SNPs on LG 3 or LG 5 were used (SNPs with P_corrected < 10^-5). (i) SS— versus RR--; (ii) SS—S versus RR--; (iii) S—S versus S—R; (iv) RR—S versus RR--; and (v) SS—S versus RR--;. Labels attached to peaks describe the inferred possible resistance (R) or susceptibility (S) to Pasteuria ramosa isolates C1, C19, P15, and P20. (i) SS— versus RR--; (ii) SS—S versus RR--; (iii) S—S versus S—R; (iv) RR—S versus RR--; and (v) SS—S versus RR--;. For each SNP, the allele with the minor allele frequency (MAF) within resistotype groups that presented only one allele at the C or E loci (all homozygous individuals) was used for comparisons. Hence, the x axis represents allele frequency of the dominant allele within resistotype groups (considering total allele number, or chromosome number: 2n). Labels attached to peaks describe the inferred possible genotypes at the C or the E locus within resistotype groups. In comparisons at the C locus on LG 3, resistotype groups susceptible to C1 and C19 presented only one allele, that is, they contained only homozygous recessive individuals at the C locus (dominant allele frequency of zero). Resistotype groups resistant to C1 and C19 did contain the dominant allele (frequency between 0.5 and 1), showing that resistance is dominant at the C locus, as the resistant group contains heterozygous individuals. Similarly, in comparisons at the E locus on LG 5, resistotype groups resistant to P20 do not present the dominant allele (frequency of zero), whereas resistotype groups susceptible to P20 do, that is, contain heterozygous individuals (dominant allele frequency between 0.5 and 1). This shows that, in contrast with the C locus, susceptibility is dominant at the E locus. Screening individual genomes revealed that some SS—S individuals (susceptible to C1 and C19, and P20) presented the “ee” genotype at the E locus (resistance to P20), although susceptibility is dominant at the E locus. This was not observed in RR—S individuals (resistant to C1 and C19 but susceptible to P20) (supplementary table S3, Supplementary Material online). This observation can be explained by an epistatic relationship linking the C and the E loci. This epistasis confers susceptibility to P20 to individuals resistant to C1 and C19, that is, presenting the “cc” genotype regardless of the genotype at the E locus. In contrast, with groups containing SS—S individuals, that is, comparisons (iii) and (iv), some SS—S individuals present the “ee” genotype at the E locus. In these groups, the frequency of the dominant allele can be lower than 0.5.
becomes neutral when a C allele is present. We assume that the a allele is fixed in the Aegelsee D. magna population, so that only variation at the B locus influences the C1 and C19 resistotypes in individuals with the “cc” genotype (see above). As variation at the B locus is very rare in our D. magna study population and could not be included in the GWAS analysis, we selfed two D. magna genotypes that presented the very rare SR...S resistotype, whose underlying genotype at the ABC cluster was expected to be “B-cc” (probably “Bbcc,” considering the B allele is rare in the population). In the F1 offspring of the two F0 parents with the SR...S resistotype and the “Bbcc...” genotype, we observed SR...R individuals. We speculate that SR...R animals have the genotype “B-ccce,” indicating that the epistatic relationship previously described between the C and the E loci (“cc” acts epistatically on the E locus) should also include the B locus. If this is the case, “bbcc” acts epistatically on the E locus (Fig. 5). The two groups of selfed F1 offspring involving “Bb” heterozygotes showed a good fit between this expectation in the expanded model and the observed phenotypic segregation. We observed one SR...R offspring produced from a SR...S parent with the inferred “aab-ccEE” genotype, which is not expected in our model (Table 2B, lower panel), but typing mistakes cannot be fully ruled out.

Linking the Genomic Regions and the Genetic Model of Resistance

To test whether the segregation of the genomic regions, we discovered in the GWAS and the segregation of resistotypes in our crosses agreed with each other, we designed size-polyorphic markers in the genomic regions of the C and the E loci (two for each locus). We tested whether these markers cosegregated with the resistotypes as predicted by our genetic model. Of the four markers, DMPR1 (C locus) and DMPR3 (E locus) showed better linkage with their respective resistance loci (99.6% and 94.8% match, respectively) compared with DMPR2 (C locus) and DMPR4 (E locus) (91.4% and 69.4% match, respectively) (supplementary tables S16–S18, Supplementary Material online). These numbers reflect that our markers are close to the actual resistance loci, but that recombination between them is possible, leading to nonperfect association. We further based our scoring of resistance genotypes on the more predictive marker genotypes of DMPR1 and DMPR3. In 20 of 22 F1 groups representing all possible combinations of alleles at the C and E loci (Table 1), the segregation of marker genotypes in the F1 offspring followed our genetic model predictions, that is, all expected genotypes were observed, with no statistically significant deviations from the expected frequencies. In two F1 groups from “CCee” and “Cc” F0 parents, the E-locus markers appeared not to be linked to the E locus (supplementary tables S5 and S8, Supplementary Material online). Based on the genotype markers results, we had assigned the “EE” genotype to the F0 parent and F1 offspring, but phenotypic segregation in the F1 offspring indicated the parent would have the “Ee” genotype. We speculate that recombination had uncoupled the genetic marker and the resistance loci in these two-parent genotypes. Together, these results show that the genomic regions found in the GWAS are indeed associated with the segregation of resistotype in the F1 selfed offspring, supporting our genetic model for the segregation of resistance (Fig. 5).

Discussion

This present study aims to assess how annual epidemics by a parasitic bacterium, P. ramosa, influence resistance and the frequencies of the underlying genes in a natural host population of the crustacean D. magna. Over the course of epidemics in two consecutive years, we observed drastic changes in resistance phenotype (resistotype). Using experimental infections and fitness measurements on wild-caught...
individuals, we showed that these changes in resistotype frequency were caused by a local parasite type common during the early phases of the epidemics. A GWAS and laboratory crosses enabled us to locate the resistance genes that responded to this selection and to uncover their mode of inheritance. We pinpointed the genetic architecture of resistance to two genomic regions with dominance and epistasis, thus bridging the gap between natural selection on phenotypes and the underlying genetics.

Parasite-Mediated Selection
Over the two consecutive years of this study, resistotype frequencies in the host population changed drastically during the parasite epidemics, but remained stable outside of the epidemics (fig. 1)—a pattern consistent with the host population being under strong selection for resistance to *P. ramosa*. The P20 *P. ramosa* isolate, collected during the early epidemic, turned out to be representative of the parasite population during the early part of the two epidemics studied here: Host genotypes characterized by their susceptibility to *P. ramosa* P20 drastically decreased in proportion during the epidemics and were much more susceptible to the local parasite than P20-resistant individuals in experimental infections (fig. 2A).

Table 1. Genetic Crosses of Resistance Phenotypes (resistotypes) from the Aegelsee *Daphnia magna* Population, Where Only the C and the E Loci Are Considered.

| F0 parent | Selfed F1 offspring | Proportion | Expected | Repeat (F1 groups from different parent clones) | C-M test on counts |
|-----------|---------------------|------------|----------|-----------------------------------------------|-------------------|
|           |                     |            | a        | b                 | c        | d               |                     |
|          |                     |            |          |                   |          |                 |                     |
| CCEE      | RR_S                | CCEE       | RR_S     | 1                 | 1        | 1               | NA                 |
| CCEe      | RR_S                | CCEe- CCEe | RR_S     | 0.75              | 0.79     | 0.74            | 0.84               |
|           |                     |            |          |                   |          |                 | X²=0.85, p>0.36    |
| CCEe      | RR_R                | CCEE       | RR_R     | 0.25              | 0.21     | 0.26            | 0.16               |
|           |                     |            |          |                   |          |                 |                     |
| CCEE      | RR_R                | CCEE- CCEE | RR_R     | 1                 | 1        | 1               | 1                  |
|           |                     |            |          |                   |          |                 |                     |
| CCEE      | RR_S                | CcEE       | SS_S     | 0.75              | 0.74     | 0.26            | p=1                |
|           |                     |            |          |                   |          |                 |                     |
| CCEe      | RR_S                | C-cE- C-cE | SS_S     | 0.56              | 0.47     | 0.56            | 0.65               |
|           |                     |            |          |                   |          |                 | 0.51               |
|           |                     |            |          |                   |          |                 | M^2=4.61, p=0.10   |
| CcEe      | RR_S                | c-cE- c-cE | SS_S     | 0.25              | 0.30     | 0.06            | 0.13               |
|           |                     |            |          |                   |          |                 | 0.21               |
|           |                     |            |          |                   |          |                 |                     |
| CcEe      | RR_R                | c-ccE- c-cE | SS_S     | 0.75              | 0.75     | 0.80            | 0.62               |
|           |                     |            |          |                   |          |                 | df=1, p=0.16       |
|           |                     |            |          |                   |          |                 |                     |
| ccEE      | SS_S                | ccEE       | SS_S     | 1                 | 1        | 1               | NA                 |
|           |                     |            |          |                   |          |                 |                     |
| ccEe      | SS_S                | c-cE- c-cE | SS_S     | 1                 | 1        | 1               | NA                 |
|           |                     |            |          |                   |          |                 |                     |
| ccee      | SS_S                | ccee       | SS_S     | 1                 | 1        | 1               | NA                 |

**NOTE.**—A, Punnett square for all possible gamete combinations according to our genetic model of resistance inheritance. The table shows the resistotypes (grouped by background color) from the 16 combinations of gametes from a double heterozygote for the C and the E loci. The bottom right cell (red font, italics) represents offspring individuals where the epistatic interaction between the C and the E loci is revealed (fig. 5); B, Results from selfing of *D. magna* clones. Resistotypes of F0 mothers and F1 offspring groups were obtained using the attachment test, and resistance genotypes of F0 parents at the C and E loci were inferred from their resistotypes and the segregation patterns of resistotypes in their F1 offspring. Expected resistotype proportions within F1 groups were calculated using the genetic model presented in the Punnett square and the R package “peas” (fig. 5 and supplementary doc. S1, Supplementary Material online). Detailed results and statistical analyses for each cross are presented in supplementary tables S4–S12 and S15, Supplementary Material online. Segregation of offspring is presented as proportions, although statistical tests were run on counts. One to four crosses using distinct mother clonal lines (repeats a to d) were conducted for each F0 mother resistance genotype at the C and E loci. No variation at the B locus was observed (all F0 mothers are inferred to have the “bb” genotype according to F1 resistotype segregation).
infection by the local parasite. Field data confirmed this result, as wild-caught P20-susceptible individuals were infected more frequently and produced fewer offspring than infected P20-resistant individuals (fig. 3), again showing the higher virulence of the parasite in these P20-susceptible individuals. This effect of the parasite seemed less strong in field-collected animals than for those infected in the laboratory. Multiple factors may contribute to this, including differences among field and laboratory, and differences in the host and parasite populations from the two study years (2014 and 2015). Our findings reveal nevertheless a strong and rapid response to parasite-mediated selection on host resistotypes, that are characterized by their interaction with the P20

---

**Table 2.** Genetic Crosses of Resistance Phenotypes (resistotypes) from the Aegelsee *Daphnia magna* Population Considering the B and the E Loci, with the C Locus Fixed for Genotype “cc.”

| A | BcE | Bce | bcE | bce |
|---|-----|-----|-----|-----|
| **SR₅** | **SR₅** | **SR₅** | **SR₅** | **SR₅** |
| **BccEe** | **SR₅** | **SR₅** | **SR₅** | **SR₅** |
| **Sr₅** | **SR₅** | **SR₅** | **SS₅** | **SS₅** |
| **Sr₅** | **SR₅** | **SR₅** | **SS₅** | **SS₅** |

**B**

| Inferred genotype | Selfed F1 offspring |
|-------------------|---------------------|
|                  | ** Expected ** | ** Observed ** | ** Fisher test on counts ** |
| **BccE** SR₅      | 0.75              | 0.74          | p=1                              |
| **BccE** SS₅      | 0.25              | 0.23          |                                  |
| **BccE** SR₅      | 0.00              | 0.03          |                                  |
| **BccE** SS₅      | 0.56              | 0.56          | p=0.58                           |
| **BccE** SS₅      | 0.25              | 0.33          |                                  |
| **BccE** SR₅      | 0.19              | 0.11          |                                  |

---

Although the parasite P20 was isolated during the early phase of the yearly epidemics, previous research also shows other parasite genotypes in the Aegelsee population (Andras and Ebert 2013) that, as we observed in an earlier year, become more common in infected hosts later in the epidemics (supplementary fig. S1, Supplementary Material online). We speculate that these later-season isolates may represent different parasite infectotypes (infection phenotypes). Consistent with this, we observed that animals resistant to P20 did, in fact, become infected, both in the field and in the laboratory (figs. 1–3), which cannot be explained with P20-infectotype parasites alone. The present study focuses on natural selection during the early part of the epidemics, which, as our data and data from other years shows, has a fairly consistent selection pattern (Ameline C, Vögli F, Andras J, Engelstädter J, Ebert D, unpublished data), being mainly defined by a drastic increase in P20-resistant individuals from around 50% to almost 100% within a period of 2–3 months (fig. 1).

The composition of the resistotypes at the beginning of the two seasons (2014 and 2015) in which we monitored this system was strikingly similar, which is surprising given that selection increased resistance over the course of the summer 2014. Although answering this question is not part of the current study, there are a few tentative explanations for this observation. First, part of the yearly resting eggs yield, which form the basis of the new population in the following
spring, are produced as early as mid-June before selection has diminished some of the resistotypes. Second, epistasis and dominance can protect alleles from natural selection, thus slowing down the response to selection (Feldman et al. 1975; Otto 2009). Our study, as well as earlier studies on this system (Luijckx et al. 2012, 2013; Metzger et al. 2016), all indicate strong epistasis and dominance for resistance loci. Further studies are needed to understand how much resting egg production and the genetic architecture of resistance explain the slow response to selection observed across seasons in the Aegelsee D. magna population.

Genetic Architecture of Resistance

To understand the genetic architecture of resistance loci under selection in our study population, we combined a GWAS using D. magna genotypes with different resistotypes together with a series of genetic crosses. We found that the most diversity in host resistance to the bacteria is determined by variation at the C locus, situated in a previously described supergene, the PR locus containing the ABC cluster (Bento et al. 2017), and at a newly discovered locus on a different chromosome, the E locus (fig. 4). Taken alone and in the right genetic background, that is, when there is no epistatic relationship, each of these two loci show Mendelian segregation with resistance being dominant (C locus) or recessive (E locus) (fig. 4, right panel). The two loci interact epistatically with each other, resulting in a complex pattern of inheritance (fig. 5). Balancing selection is hypothesized to maintain diversity at resistance genes (Llaurens et al. 2017; Wittmann et al. 2017; Connallon and Chenoweth 2019), and these genes are often found to have different dominance patterns and epistatic interactions (Saavedra-Rodriguez et al. 2008; González et al. 2015; Conlon et al. 2018).

The E locus is situated on linkage group (LG) 5 (genome version 2.4: Routtu et al. 2014) and appears as a large region of 3.1 Mb (fig. 4). In this regard, the E locus is similar to the ABC cluster, a well-characterized, nonrecombining, and extremely divergent region on LG 3 (Bento et al. 2017). Nonrecombining genomic structures, that is, supergenes, are suggested to facilitate adaptation via association of advantageous alleles in host–parasite coevolution (Joron et al. 2011; Llaurens et al. 2017). Such large, diverse genomic regions are difficult to study because the absence of recombination hampers fine mapping (Bento et al. 2017). Therefore, we do not know where the actual resistance loci lie within the ABC- and E-loci regions. This may also explain why our genetic markers are not perfectly linked to the resistance loci (supplementary tables S17 and S18, Supplementary Material online). Supergenes may also harbor several resistance loci, thus variation at the C or the E locus could actually represent variation at several loci physically very close to each other. Within the E-locus region, we find four sugar transferases. Glycosylation genes are candidates to explain variation of resistance in this system (Bento et al. 2017; Bourgeois et al. 2017).

In the D. magna–P. ramosa system, the ABC cluster has been shown to play a major role in host resistance and the evolutionary dynamics of resistance (Routtu and Ebert 2015; Bento et al. 2017; Bourgeois et al. 2017). Our results confirm the role of this cluster in a natural population and describe a new resistance region in the D. magna genome that is polymorphic in the Aegelsee population. Multilocus polymorphisms have been shown to underlie parasite resistance in host–parasite coevolution (Sasaki 2000; Tellier and Brown 2007; Cerqueira et al. 2017). In the Aegelsee D. magna population, there seems to be no variation at the A locus and little variation at the B locus. The observed variation at the B and C loci is consistent with the genetic model of resistance at the ABC cluster described in Metzger et al. (2016). In addition, resistance to P. ramosa isolate P15 (influenced by the D locus, Bento et al. 2020) remains fairly consistent, with the vast majority of animals being susceptible to P15 (fig. 1). Resistance to P. ramosa P21, also isolated from our study population, varies only toward the end of the summer epidemic (Ameline C, Vöglli F, Andras J, Engelstädt J, Ebert D, unpublished data). In summary, the ability to resist P20 plays a major role in the early epidemics and most resistotype diversity we measured in the Aegelsee D. magna population is well explained by genotypic variation at these loci. As we use more parasite isolates in further research, we might find other resistance regions in the D. magna genome. This is likely to be of importance in the later phase of the epidemics in the Aegelsee population.

Resistance segregation in D. magna is currently best explained by a genetic model where each locus contains just two alleles. This model was compiled by studies that used either mapping panels created from a few D. magna genotypes or, as here, host genotypes from one focal population. Additional resistance alleles may be revealed instead of new resistance regions if we test the genetic model on a larger panel of host and parasite genotypes.

We created 22 F1 offspring groups from the three common resistotypes in our study population. Segregation of resistance phenotypes and genotypes among the selfed F1 strongly supported the genetic model of resistance, consisting of the C and E loci, each with two alleles, and their epistatic interaction, produced by the GWAS (tables 1 and 2; supplementary tables S4–S15, Supplementary Material online). Two F1 offspring groups showed rare variation at the B locus, suggesting yet an additional epistatic interaction in this model besides the previously described role of the B locus for the P. ramosa C1 and C19 resistotypes. This consisted of the “bbc” genotype that causes susceptibility to P20, irrespective of the genotype at the E locus (fig. 5). However, this modified model needs to be further investigated and verified with more genetic crosses.

Conclusion

In this study, we demonstrate rapid parasite-mediated selection in a natural plankton population. We find the genomic regions associated with resistance under selection and describe their mode of inheritance. This knowledge will allow us to conduct direct measurements of resistance allele frequency changes over time and to test theories on the dynamics of host and parasite coevolution, for example by
tracing genetic changes in the resting stages of *D. magna* derived from the layered sediments in ponds and lakes (Decaestecker et al. 2007). Pinpointing resistance loci can also be used to infer mechanisms of selection in the host with the molecular evolution tool box (Charlesworth 2006; Fijarczyk and Babik 2015; Hahn 2018). Our model of resistance consists of a few loci linked with epistasis and different dominance patterns, characteristics that have been shown to be relevant in coevolution, in particular when balancing selection maintains diversity at resistance genes (Tellier and Brown 2007; Engelstäder 2015; Conlon et al. 2018). The genomic regions we pinpoint can now be further studied, for example, by testing for genomic signatures for balancing selection (Charlesworth 2006; Ebert and Fields 2020). Hence, a precise knowledge of the genetic architecture of resistance opens the door to addressing wider evolutionary questions. For example, the Red Queen theory states that host–parasite interactions may explain the ubiquity of sex and recombination (Salathé et al. 2008).

**Materials and Methods**

**Study Site**

Our study site is the Aegelsee, a pond near Frauenfeld, Switzerland (code: CH-H for Hohliberg coordinates: 47.557769 N, 8.862783 E, about 30,000 m² surface area) where *D. magna* is estimated to have a census population size over ten million individuals and an overwintering resting egg bank of about the same size. Every year from early October, the pond is used as a waste repository by a sugar factory; they progressively lower the water level from May to September and from October, warm ammoniacal condensation water is released into the pond, warming the water temporarily to 40–60 °C (Seefeldt and Ebert 2019) and killing all zooplankton, but not the resting eggs. In winter the pond usually freezes over, and in April, *Daphnia* and other invertebrates hatch from resting eggs. We sampled the pond in February 2014 and March 2015 and did not find *D. magna*, suggesting little or no overwintering. Besides *D. magna*, the plankton community includes *D. pulex*, *D. curvirostis* and a diverse array of other invertebrates, among them copepods, ostracods, rotifers, and corixids. The waste-water treatment prevents fish from invading the pond. The *D. magna* population experiences strong yearly epidemics of *P. ramosa*, reaching prevalence of 70–95%. Infections by other parasites were only rarely observed. The other *Daphnia* species in the pond were never observed to be infected by *P. ramosa*.

**Temporal Monitoring**

In 2014 and 2015, we sampled the host population every 2–3 weeks from early April to early October to study the impact of the pathogen epidemics. For each sampling date, we aimed to obtain about 100 cloned host lines (produced as iso-female lines). To achieve this, we randomly took about 200–300 female *D. magna* from the sample, placed them in 80-ml jars filled with ADaM (Artificial *Daphnia* Medium, Klüttgen et al. 1994, as modified by Ebert 1998) and let them reproduce asexually. Oversampling was necessary during the hot summer months, as many animals would die for unknown reasons within 48 h under laboratory conditions. This mortality was, to the best of our knowledge, not disease related. Over the following 3 weeks, we screened animals for *P. ramosa* infections by checking for the typical signs of disease: gigantism, reddish-brownish opaque body coloration, and empty brood pouch. Infected animals that had not yet reproduced asexually were treated with tetracycline (50 mg l⁻¹) (an antibiotic which kills Gram-positive bacteria) until an asexual clutch was observed, usually after about 2 weeks. They were fed 25 million cells of the unicellular green algae *Scenedesmus* sp. three times a week, and the medium was renewed every 2 weeks. Feeding and fresh medium protocols were adapted according to the size and number of animals in a jar when necessary.

**Resistotype Assessment: The Attachment Test**

We assessed resistance phenotype (resistotype) for all cloned hosts using four *P. ramosa* isolates (C1, C19, P15, and P20). We isolated the parasite, P20, from our study population at the start of the epidemic on May 13, 2011 and subsequently passaged it three times through a susceptible *D. magna* host clone from the same population. The three other *P. ramosa* clones or isolates had been previously established in the laboratory: C1 (clone), originated from a *D. magna* population in Russia (Moscow), C19 (clone) from Germany (Gaarzerfeld) and P15 (isolate) from Belgium (Heverlee) (Luijckx et al. 2011; Bento et al. 2020). We used these three *P. ramosa* allotypic isolates in the present study to implement our working genetic model for resistance (Luijckx et al. 2012, 2013; Metzger et al. 2016). Parasite transmission stage (≈ sporulation) were conducted in the laboratory followed the protocol by Luijckx et al. (2011).

The resistotypes of *D. magna* clones were assessed using a spor attachment test (Duneau et al. 2011). Bacterial spores attach to the foregut or the hindgut of susceptible host clones. Attachment is a prerequisite for subsequent infection. We call these genotypes susceptible, otherwise they are considered resistant. A genotype allowing attachment and penetration of the parasite into the host, may sometimes still resist infection, based on subsequent immune defense (Hall et al. 2019). To test for attachment, we exposed each individual *Daphnia* to 8,000 (C1, C19) or 10,000 (P15, P20) fluorescently labeled spores following the protocol of Duneau et al. (2011). We used higher spore doses for P15 and P20 because previous observations had shown that fewer of these isolate spores attach to the host esophagus, resulting in a weaker fluorescent signal. Three repeats were used for C1, C19, and P15, whereas six to nine repeats were used for P20. A clone was considered susceptible to the bacterial isolate when more than half of its replicates showed clear attachment. Its overall resistotype is the combination of its resistotypes to the four individual *P. ramosa* isolates in the following order: C1, C19, P15, and P20, for example, a clone susceptible to all four isolates would have resistotype SSSS. Since resistance to P15 had low variability in our study population, this isolate was only considered in the first experiment presented here and was otherwise represented with the placeholder “_,” for example, “RR__R resistotype.”
Experimental Infections of Resistotypes
As an initial assessment of the parasite’s fitness impact on the host population, we conducted experimental infections on a representative sample of the spring 2014 host population. We collected surface sediment from five different points in the pond in February 2014, before onset of the natural hatching season and placed 100 D. magna ephippia from each replicate in 80-l containers with 301 ADaM. The five containers were placed outdoors under direct sunlight and checked for hatchlings every 2 days. We recorded hatching dates and cloned hatchlings in the laboratory where we then scored their resistotypes. For the infection experiment, we used parasites collected from the ongoing epidemic in the pond. We collected three pools of 20 randomly chosen infected individuals during the first phase of the epidemic in early June 2014. These field-infected animals were kept in the laboratory under ad libitum feeding conditions. Shortly before their expected death, we pooled all 60 animals, homogenized them to produce a spore suspension, and froze it at −20 °C. A placebo suspension was produced from 60 homogenized uninfected D. magna. The parasite spore mixture was not passaged before we used it, so, in contrast to the isolates used for the attachment test, it represents a population sample of the parasite.

Among the four predominant resistotypes, we observed in the cloned cohort of spring hatchings (SSSS, RRSS, RRSR, and RRRR), we used 20 clones each from the more common resistotypes SSSS, RRSS, and RRSR and ten of the less common resistotype RRRR for an infection experiment, due to limited availability. From each of these 70 clones, we produced five replicate lines, and these 350 lines were maintained individually in 80-ml jars. To reduce maternal effects before the experiment, we kept all lines for three generations in the laboratory where we then scored their resistotypes. For the infection experiment, we used parasites collected from the ongoing epidemic in the pond. We collected three pools of 20 randomly chosen infected individuals during the first phase of the epidemic in early June 2014. These field-infected animals were kept in the laboratory under ad libitum feeding conditions. Shortly before their expected death, we pooled all 60 animals, homogenized them to produce a spore suspension, and froze it at −20 °C. A placebo suspension was produced from 60 homogenized uninfected D. magna. The parasite spore mixture was not passaged before we used it, so, in contrast to the isolates used for the attachment test, it represents a population sample of the parasite.

For the infection experiment, we used parasites collected from the ongoing epidemic in the pond. We collected three pools of 20 randomly chosen infected individuals during the first phase of the epidemic in early June 2014. These field-infected animals were kept in the laboratory under ad libitum feeding conditions. Shortly before their expected death, we pooled all 60 animals, homogenized them to produce a spore suspension, and froze it at −20 °C. A placebo suspension was produced from 60 homogenized uninfected D. magna. The parasite spore mixture was not passaged before we used it, so, in contrast to the isolates used for the attachment test, it represents a population sample of the parasite.

Infection Phenotypes of Field-Collected Hosts
As a second assessment of the impact of the local parasite on the host population, we measured fitness traits of animals caught during the epidemics. Because the infection experiment described above (carried out in the previous year) indicated that P20 played a strong role, we focused on this parasite isolate. On June 7 and 28, 2015, we collected large D. magna samples from our study site and measured body length, from the top of the head through the eye to the base of the tail spine. We kept all females (n = 331) individually under ad libitum feeding conditions, each in about 80 ml medium. We recorded clutches (time and size) and the onset time of disease symptoms over the following 3 weeks. After parasitic castration was evident, we cured animals with tetracycline. These data have also been reported in a paper describing the disease phenotype under natural conditions (Savola and Ebert 2019). The current data set is however smaller than the published data, as we report here only those animals for which we were able to score the resistotypes.

Using generalized linear models, we tested the effect of the P20 resistotype on infection and fecundity, taking body size into account. Sampling date was included as a fixed effect since there are only two sampling dates. Interaction terms were excluded from the model when not significant (P > 0.1). We fitted a general linear model using quasi-binomial data family type for infection, and a negative binomial generalized linear model for total fecundity (R packages used: MASS: Venables and Ripley 2002, lme4: Bates et al. 2015).

Genome-Wide Association Study
Because our experiments revealed that resistance to P20 plays a major role in the disease dynamics in both laboratory experiments and the field, we used a genome-wide association approach to investigate the genetic architecture of resistance with 37 clones that presented the three most common resistotypes in our study population, excluding P15 resistotype (n = 16 RR,...,R, 10 RR,...,S, and 11 SS,...,S). All 37 clones were derived directly from our study population (supplementary table S3, Supplementary Material online).

Whole-Genome DNA Extraction, Sequencing, and Bioinformatics
To remove microbial DNA, individuals were treated for 72 h with three antibiotics (streptomycin, tetracycline, ampicillin at a concentration of 50 mg l−1 each in filtered water) and fed twice daily with 200 μl of a dextran bead solution (Sephadex G-25 Superfine by Sigma Aldrich: 20–50 μm diameter at a concentration of 5 g l−1) to remove algae from the gut. DNA was extracted from 15 to 20 adult animals using an isopropanol precipitation protocol (QIAGEN DNeasy Blood
Assessment of Resistotype Segregation

The genetic model that resulted from the GWAS analysis allowed us to make predictions about the segregation of resistotypes in sexually reproducing *D. magna* lines. To test these predictions, we selfed *D. magna* clones with different resistotypes. Selfing is possible with *D. magna* because the same clonal line can produce sons (asexual production) as well as eggs by sexual production. The latter need fertilization by males. The resulting sexual eggs must undergo an obligatory resting phase before they can hatch (Ślusarczyk et al. 2019). The resistotypes of the selfed offspring (F1) were examined to assess whether their segregation matched expectations from the genetic model derived from the GWAS.

All clones used for the genetic crosses derived from the study population. We selfed five to ten *D. magna* clones of the three common resistotypes (RR → R, RR → S, and SS → S) and two clones of a rare resistotype (SR → S), following the protocol from Luijkx et al. (2012). Hatching of selfed offspring is not always successful, resulting in uneven sample sizes. We obtained between 19 and 89 selfed offspring from each of 22 parent clones (supplementary table S19, Supplementary Material online). Their resistotypes were assessed with the attachment test. Samples from each clonal line were stored at −20 °C for future DNA extraction and genotyping.

Predictions of Segregation Patterns

We compared the resistotype segregation patterns in the selfed offspring with predictions in our genetic model. To calculate proportions of expected phenotypes, we developed an R package called “peas” (https://github.com/IanEngelstaedter/peas, last accessed December 2020) that enables the user to predict distributions of offspring genotypes and phenotypes in complex genetic models with Mendelian inheritance (supplementary docs. S1 and S2, Supplementary Material online). We compared these predictions to the segregation patterns from our selfed offspring using the Cochran–Mantel–Haenszel (C–M–H) test for repeated tests of independence. The C–M–H test is applied to 2 × 2 tables. We ran a Fisher’s test of independence on each comparison (expected vs. observed for each repeat, Bonferroni corrected) to detect differences in opposite directions across repeats, which would have resulted in a nonsignificant C–M–H test, but no such differences in direction were detected (see supplementary table S15, Supplementary Material online for detailed results of statistical analyses). Tests were run on counts, but for better illustration, we present here segregation of offspring as proportions.

Raw reads were aligned using BWA MEM (Li and Durbin 2009) on the *D. magna* draft genome (v.2.4) and a genetic map (Routtu et al. 2014). BAM alignment files were then filtered for quality, and PCR duplicates were removed using PICARD tools (http://broadinstitute.github.io/picard/, last accessed December 2020). Variant calling was performed using freebayes (v. 0.9.15-1). VCF files were then filtered using VCFTOOLS v. 0.1.12b (Danecek et al. 2011) to include SNPs with a minimum quality of 20, a minimum genotype quality of 30, and a mean sequencing depth between 10 × and 50 ×. Only SNPs that passed filters in every clone sample were included in subsequent analyses, resulting in a data set of 510,087 SNPs. Association analyses were performed using the command “assoc” in PLINK (Purcell et al. 2007), which compares allele counts between cases and controls and outputs a *P* value from a χ² test with one degree of freedom. Five pairwise comparisons were performed to identify possible candidates for resistance to C1, C19, and P20: (i) SS → S versus RR → S, (ii) SS → S versus RR → S, (iii) SS → S versus RR → R, (iv) RR → S versus RR → R, and (v) SS → S versus RR → R. We corrected for the genomic inflation of *P* values (λ) that we have corrected with relatedness between samples using the R package GenABEL (Aulchenko et al. 2007). Lambda was calculated excluding SNPs from linkage groups 3 and 5, since these scaffolds displayed an excess of strongly associated markers. We divided raw χ² scores by λ to obtain corrected *P* values using R commands “pvalue” and “estlambda.” For each SNP:

Histograms of corrected *P* values were examined to confirm their uniform distribution. We estimated the minimum false discovery rate incurred when a given *P* value was identified as significant (so-called q-value) from the set of corrected *P* values using the R package “qvalue” (Storey et al. 2015).

The minimum significant threshold for a given association was then calculated as the maximum corrected *P* value with a q-value < 5%.

The “gg.manhattan” function in R was used to display manhattan plots of the comparisons between different resistotypes (https://github.com/timknut/gg.manhattan/, last accessed December 2020). We used BEDTOOLS (v 2.25.0) to extract genes found in the associated candidate regions, using the 2011 annotation of the genome (available at: w sleekbase.org, last accessed December 2020).

Assessment of Resistotype Segregation

The genetic model that resulted from the GWAS analysis allowed us to make predictions about the segregation of
Linking the Phenotype to the Genotype
We designed PCR-based diagnostic markers physically linked to the resistance loci that the GWAS identified (DMPR1 to 4 for “Daphnia magna–Pasteuria ramosa” markers, supplementary table S20, Supplementary Material online) and tested if these markers (and their corresponding resistance regions) are indeed associated with the resistotypes, by comparing expected and observed association between marker genotypes and resistotypes (supplementary tables S16–S18, Supplementary Material online). We then used these markers to confirm genotyping of the selfed parents.

DNA Extraction and PCR-Based Markers Analysis
DNA of parents and selfed offspring was extracted on 96-well PCR plates using a 10% Chelex bead solution (Bio-Rad) adapted from Walsh et al. (1991). First, individuals were crushed in the wells with 20 μl of deionized water using a customized rack of metallic pestles. We added 150 μl of 10% Chelex solution and 10 μl of proteinase K and incubated samples for 2 h at 55 °C followed by 10 min at 99 °C. Fragment amplification, genotyping, and allele scoring were performed using R software version 3.6.1 (R Core Team 2019). Packages used in R for package installation, data manipulation, and graphics are the following: package development, documentation, and installation: devtools v. 2.2.1 (Wickham, Hester, et al. 2019) and roxygen2 v. 6.1.1 (Wickham et al. 2018), data manipulation: dplyr v. 0.8.3 (Wickham, Francois, et al. 2019), tidyverse v. 1.0.0 (Wickham and Henry 2019), tidyquant v. 0.5.8 (Dancho and roxygen2 v. 6.1.1 (Wickham et al. 2018), data manipulation: dplyr v. 0.8.3 (Wickham, Francois, et al. 2019), tidyverse v. 1.0.0 (Wickham and Henry 2019), tidyquant v. 0.5.8 (Dancho and Vaughan 2019), tidyverse v. 1.2.1 (Wickham 2017), xlsx v. 0.6.1 (Dragulescu and Arendt 2018), graphics: ggplot2 v. 3.3.0 (Wickham 2016), extrafont v. 0.17 (Chang 2014), scales v. 0.6.1 (Wickham 2018), cowplot v. 1.0.0 (Wilke 2019), gridExtra v. 2.3 (Augie 2017), ggpubr v. 0.2.3 (Kassambara 2019), ggplotify v. 0.0.4 (Yu 2019), magick v. 2.2 (Ooms 2019), egg v. 0.4.5 (Augie 2019), ggsci v. 2.9 (Xiao 2018), and png v. 0.1-7 (Urbanek 2013).

Statistical Software
Unless otherwise stated, all statistical analyses and graphics were performed using R software version 3.6.1 (R Core Team 2019). Graphics were edited in Inkscape v. 1.0.1 (https://inkscape.org/), last accessed December 2020. Mean values are presented with standard error: mean ± SE (package RVAideMemoire v. 09-45-2, Hervé 2015). Packages used in R for package installation, data manipulation, and graphics are the following: package development, documentation, and installation: devtools v. 2.2.1 (Wickham, Hester, et al. 2019) and roxygen2 v. 6.1.1 (Wickham et al. 2018), data manipulation: dplyr v. 0.8.3 (Wickham, Francois, et al. 2019), tidyverse v. 1.0.0 (Wickham and Henry 2019), tidyquant v. 0.5.8 (Dancho and Vaughan 2019), tidyverse v. 1.2.1 (Wickham 2017), xlsx v. 0.6.1 (Dragulescu and Arendt 2018), graphics: ggplot2 v. 3.3.0 (Wickham 2016), extrafont v. 0.17 (Chang 2014), scales v. 0.6.1 (Wickham 2018), cowplot v. 1.0.0 (Wilke 2019), gridExtra v. 2.3 (Augie 2017), ggpubr v. 0.2.3 (Kassambara 2019), ggplotify v. 0.0.4 (Yu 2019), magick v. 2.2 (Ooms 2019), egg v. 0.4.5 (Augie 2019), ggsci v. 2.9 (Xiao 2018), and png v. 0.1-7 (Urbanek 2013).

Supplementary Material
Supplementary data are available at Molecular Biology and Evolution online.

Acknowledgments
We thank Jürgen Hottinger, Urs Stiefel, Kristina Müller, Michelle Krebs, Samuel Pichon, Dita Vizoso, and Jelena Rakov for help in the field and laboratory. Sequencing for the GWAS analysis was performed at the Genomics Facility at the Department of Biosystem Science and Engineering (D-BSSE, ETH) in Basel. The AB3130xl sequencer used for the markers analysis was operated by Nicolas Boileau. Members of the Ebert group, Jonathon Stillman and Luís Teixeira provided valuable feedback on the study and the manuscript. Suzanne Zweizig improved the language of the manuscript. Two anonymous reviewers provided a thoughtful and thorough review of the manuscript. This work was supported by the Swiss National Science Foundation (SNSF) (Grant No. 310030B_166677 to D.E.); the Freiwillige Akademische Gesellschaft Basel (FAG) to C.A.; the University of Basel to D.E. and C.A.; and the Australian Research Council through a Future Fellowship (FT140100907 to J.E.).

Author Contributions
D.E. and J.A. designed the overall study. F.V. and D.E. designed the infection experiment. F.V. conducted the infection experiment and analyzed the data. E.S. and D.E. designed the fitness measurements. E.S. conducted the fitness measurements and analyzed the data. Y.B. and D.E. designed the GWAS analysis. Y.B. conducted the GWAS analysis. C.A. and D.E. designed the crossings. C.A. conducted the crossings and analyzed the data. J.E. developed the “peas” R package. C.A. analyzed the data, wrote the manuscript, and designed the figures. All authors reviewed the manuscript.

Data Availability
The data underlying this article and analysis scripts are available in the Figshare Repository https://doi.org/10.6084/m9.figshare.13259828.v1. Raw sequence data from the GWAS analysis are deposited in the NCBI Bioproject PRJNA680821.

References
Alves JM, Cameiro M, Cheng JY, Lemos de Matos A, Rahman MM, Loog L, Campos PF, Wales N, Eriksson A, Manica A, et al. 2019. Parallel adaptation of rabbit populations to myxoma virus. Science 363(6433):1319–1326.
Andras JP, Ebert D. 2013. A novel approach to parasite population genetics: experimental infection reveals geographic differentiation, recombination and host-mediated population structure in Pasteuria ramosa, a bacterial parasite of Daphnia. Mol Ecol. 22(4):972–986.
Andras JP, Fields PD, Du Pasquier L, Fredericksen M, Ebert D. 2020. Genome-wide association analysis identifies a genetic basis of infectivity in a model bacterial pathogen. Mol Biol Evol. Augie B. 2017. gridExtra: miscellaneous functions for “Grid” graphics. Available from: https://CRAN.R-project.org/package=gridExtra. Last accessed on December 2020.
Augie B. 2019. egg: extensions for “ggplot2”: custom geom, custom themes, plot alignment, labelled panels, symmetric scales, and fixed panel size. Available from: https://CRAN.R-project.org/package=egg. Last accessed on December 2020.
Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics 23(10):1294–1296.
Bangham J, Knott SA, Kim K-W, Young RS, Jiggins FM. 2008. Genetic variation affecting host-parasite interactions: major-effect quantitative trait loci affect the transmission of sigma virus in Drosophila melanogaster. Mol Ecol. 17(17):3800–3807.
Bates D, Maechler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. J Stat Softw. 67(1):1–48.
Ben-Ami F. 2019. Host age effects in invertebrates: epidemiological, ecological, and evolutionary implications. Trends Parasitol. 35(6):466–480.
disease correlated with *Daphnia* clonal turnover. *Evolution* 72(3):619–629.

Urbanek S. 2013. png: read and write PNG images. Available from: https://CRAN.R-project.org/package=png. Last accessed on December 2020.

Vale PF, Wilson AJ, Best A, Boots M, Little TJ. 2011. Epidemiological, evolutionary, and coevolutionary implications of context-dependent parasitism. *Am Nat.* 177(4):510–521.

van’t Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ. 2016. The industrial melanism mutation in British peppered moths is a transposable element. *Nature* 534(7605):102–105.

Venables WN, Ripley BD. 2002. *Modern applied statistics with S.* 4th ed. New York: Springer.

Walsh PS, Metzger DA, Higuchi R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10(4):506–513.

Whitlock MC, Lotterhos KE. 2015. Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of $F_{ST}$. *Am Nat.* 186(13):S24–S36.

Wickham H. 2016. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag.

Wickham H. 2017. tidyr: easily install and load the "Tidyverse." Available from: https://CRAN.R-project.org/package=tidyr. Last accessed on December 2020.

Wickham H. 2018. scales: scale functions for visualization. Available from: https://CRAN.R-project.org/package=scales. Last accessed on December 2020.

Wickham H, Danenberg P, Eustler D. 2018. roxygen2: in-line documentation for R. Available from: https://CRAN.R-project.org/package=roxygen2. Last accessed on December 2020.