Resistance to gapeworm parasite has both additive and dominant genetic components in house sparrows, with evolutionary consequences for ability to respond to parasite challenge

Sarah L. Lundregan1 | Alina K. Niskanen1,2 | Stefanie Muff3 | Håkon Holand1 | Thomas Kvalnes1 | Thor-Harald Ringsby1 | Arild Husby1,4 | Henrik Jensen1

1Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway
2Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland
3Centre for Biodiversity Dynamics, Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway
4Evolutionary Biology, Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden

Abstract
Host–parasite relationships are likely to change over the coming decades in response to climate change and increased anthropogenic stressors. Understanding the genetic architecture of parasite resistance will aid prediction of species' responses to intensified parasite challenge. The gapeworm “Syngamus trachea” is prevalent in natural bird populations and causes symptomatic infections ranging from mild to severe. The parasite may affect ecological processes by curtailing bird populations and is important due to its propensity to spread to commercially farmed birds. Our large-scale data set on an insular house sparrow metapopulation in northern Norway includes information on gapeworm prevalence and infection intensity, allowing assessment of the genetics of parasite resistance in a natural system. To determine whether parasite resistance has a heritable genetic component, we performed variance component analyses using animal models. Resistance to gapeworm had substantial additive genetic and dominance variance, and genome-wide association studies to identify single nucleotide polymorphisms associated with gapeworm resistance yielded multiple loci linked to immune function. Together with genome partitioning results, this indicates that resistance to gapeworm is under polygenic control in the house sparrow, and probably in other bird species. Hence, our results provide the foundation needed to study any eco-evolutionary processes related to gapeworm infection, and show that it is necessary to use methods suitable for polygenic and nonadditive genetic effects on the phenotype.

Keywords
additive genetic variance, dominance variance, GWAS, heritability, parasite, resistance
1 | INTRODUCTION

Parasite prevalence and virulence are major drivers of ecological and evolutionary processes in natural populations, and are expected to shift over the coming decades in response to climate change (Altizer, Ostfeld, Johnson, Kutz, & Harvell, 2013; Harvell, Altizer, Cattadori, Harrington, & Well, 2009). Combined with an increase in anthropogenic stressors, such as habitat loss and environmental contaminants, these changes may have detrimental consequences for ecosystems (Marcogliese & Pirotte, 2008). In nature, parasites exert strong selective forces on their hosts and severe infections can reduce fitness dramatically (Graham et al., 2011). Environmental factors may influence parasite infection intensity, as poor conditions can make hosts more susceptible to parasite challenge. In this period of rapid environmental and ecological change, knowledge of the genetic architecture of host resistance to parasite infection is therefore crucial to understand individual variation in parasite resistance (Westerdahl, Asghar, Hasselquist, & Bensch, 2012) and to predict the evolutionary potential of natural populations to respond to such pathogens (Acevedo-Whitehouse & Cunningham, 2006).

A first step in quantifying the evolutionary potential of a population to respond to parasite challenge is to determine whether variation in resistance has a heritable genetic component. The phenotypic variance of a trait can be decomposed into additive genetic, nonadditive genetic (dominance, epistasis), environmental and residual components. Response to selection depends on additive genetic variance ($V_A$), whereas nonadditive variances may have important effects on population dynamics through individual effects on fitness but are often assumed to be unimportant in predicting trait evolution (Fisher, 1958; Lande, 1979; Visscher, Hill, & Wray, 2008). As a consequence, combined with model limitations due to small sample sizes, nonadditive variances are seldom estimated in ecological and evolutionary analyses of natural populations (Wolak & Keller, 2014). However, failure to include extant nonadditive effects in an animal model can bias $V_A$ estimators (Hill, Goddard, & Visscher, 2008; Waldmann, Hallander, Hoti, & Sillanpää, 2008; Wolak, 2012). The importance of dominance genetic variance ($V_D$) in the major histocompatibility complex (MHC)-related immune response is well documented (Bernatchez & Landry, 2003; Niskanen et al., 2014; Penn, Damjanovich, & Potts, 2002; Worley et al., 2010) and there is some support for the dominant action of alleles at non-MHC immune genes (Berghof et al., 2018; Koets et al., 2010; Psifidi et al., 2016; Stephan, Smirnova, Jacque, & Poltorak, 2007). Dominance effects also explain decline in pathogen resistance with inbreeding, and $V_D$ can affect $V_A$ recovery after bottlenecks (Barton, Turelli, & Hill, 2006; Goodnight, 1988). Therefore, estimation of $V_D$ alongside $V_A$ may be especially relevant to characterize the genetic architecture of parasite resistance in fragmented populations. There are few studies in natural populations to date that assess whether variation in resistance to parasites has a genetic component. However, heritable variation in resistance to gastrointestinal nematodes has been demonstrated in free-living ungulates (Coltman, Pillington, Kruuk, Wilson, & Pemberton, 2001; Smith, Wilson, Pillington, & Pemberton, 1999), and substantial additive genetic variance in immunocompetence has been found in the collared flycatcher (Cichoń, Sendacka, & Gustafsson, 2006). Assessment of the contribution of dominance variance to parasite resistance remains largely unexplored, but substantial dominance variance has been demonstrated for fitness-related traits in commercial breeds (Crnokrak & Roff, 1995; Lopes, Bastiaansen, Jans, Knol, & Bovenhuis, 2015; Misztal et al., 1998; Pante, Gjerde, McMillan, & Misztal, 2002). However, current lack of estimates of nonadditive genetic variance in natural populations make it challenging to predict the influence dominance variance may have on trait evolution in nature (Wolak & Keller, 2014).

Given that resistance to parasites isheritable, is such resistance polygenic or controlled by few major effect genes? This may influence durability of resistance (Telford, Cavers, Ennos, & Cottrell, 2015) and host parasite co-evolution dynamics, as virulent pathogen phenotypes can evolve rapidly where resistance is controlled by few major effect genes. There is evidence that vertebrate resistance to pathogens may be highly polygenic, especially when innate immunity is considered (Brown et al., 2013; Hayward, 2013; Woolhouse, Webster, Domingo, Charlesworth, & Levin, 2002). Innate immunity is integral to resistance against parasites (De Veer, Kemp, & Meeusen, 2007; McRae, Stear, Good, & Keane, 2015) and consists of diverse genetic pathways that contribute to total immune response (Hagai et al., 2018). Nonetheless, few studies to date have assessed the genomic architecture (i.e., monogenic, oligogenic or polygenic) of parasite resistance traits in natural populations (but see Brown et al., 2013; Sparks et al., 2019; Wenzel, James, Douglas, & Pierney, 2015). One way to accomplish this is by genome partitioning, which partitions genetic variance by chromosome (Yang, Manolio, et al., 2011). This method has recently been used to indicate an oligogenic to polygenic basis for gastrointestinal nematode burden in wild red grouse (Wenzel et al., 2015). Gene mapping in the form of genome-wide association studies (GWAS) can also be used to determine genomic architecture, and to identify putative causal loci for heritable traits. Limitations apply to the GWAS approach when it is used to map pathogen resistance genes: success is more likely where resistance is due to common, large-effect variants, and resistance genes identified in a given population may also depend on the genetic makeup of the pathogen community (MacPherson, Otto, & Nuismer, 2018). Despite these drawbacks, numerous genes relating to parasite resistance have been identified in livestock, including sheep (Benavides, Sonstegard, & Van Tassell, 2016; Berton et al., 2017; Periasamy et al., 2014; Pickering, Auvray, Dodds, & McEwan, 2015), cattle (May et al., 2019), goat (Silva et al., 2018; Zvirovova et al., 2016) and chicken (Boulton et al., 2018). GWAS on disease resistance traits in natural populations are limited. However, several genes associated with immunoglobulin (Ig) levels that contribute to nematode resistance have been identified in a natural population of Soay sheep (Sparks et al., 2019), and several genes involved in immune and physiological processes have been found in GWAS on gastrointestinal nematode burden in wild grousse (Wenzel et al., 2015).
The house sparrow (*Passer domesticus*) is a ubiquitous passerine bird that has been used as an ecological model species in previous genetic architecture studies (Andrew, Jensen, Hagen, Lundregan, & Griffith, 2018; Andrew et al., 2019; Lundregan et al., 2018; Silva et al., 2017). In Helgeland, northern Norway, a natural metapopulation of house sparrows has been studied at an individual level since 1993 and has a large sample archive with records of morphological and life history data, alongside individual genetic and pedigree information (Jensen et al., 2003, 2004; Pärn, Jensen, Ringsby, & Sæther, 2009). Data on infection by a parasitic nematode, *Syngamus trachea*, have also been systematically collected through faeces sampling. *S. trachea*, known colloquially as "gapeworm," is globally distributed and infects most terrestrial bird genera (Atkinson, Thomas, & Hunter, 2008; Campbell, 1935). The parasite is problematic in commercial bird farming because inflammation caused by mature worms that reside in the trachea often results in physical symptoms, such as coughing or gasping, that can lead to reduction in weight gain or increased mortality (Atkinson et al., 2008; Holand et al., 2014). Eggs from mature female worms are expectorated, swallowed and excreted in the faeces, and subsequent development of embryos to mature larvae occurs between 16 and 35°C (Barus, 1966). Transmission of the parasite between birds occurs through ingestion of mature larvae, or via ingestion of paratenic invertebrate hosts (Atkinson et al., 2008). Infection by *S. trachea* varies spatiotemporally within our study population (Holand, Jensen, Tufto, Saether, & Ringsby, 2013), reduces survival probability in severely infected adults (Holand et al., 2014) and negatively influences the reproductive success of female house sparrows (Holand et al., 2015). The prevalence of infections by *S. trachea* has also been shown to increase with temperature and following mild winters (Holand et al., 2019). Within our study system, house sparrows are rarely re-infected by *S. trachea* (Holand, Jensen, et al., 2013), suggesting development of adaptive immunity.

In the present study, we explore the relative contributions of additive genetic, dominance genetic and environmental variance to several measures of resistance to *S. trachea* in the house sparrow using variance component analysis. Narrow- and broad-sense heritabilities ($h^2$ and $H^2$ respectively) of these measures of parasite resistance were estimated, and the genomic architecture of parasite resistance was assessed using genome partitioning. Additive and dominance effect sizes of markers on our custom house sparrow genome-wide 200k single nucleotide polymorphism (SNP) array (Lundregan et al., 2018) were then estimated using GWAS, to identify genomic regions related to parasite resistance in the house sparrow and provide mechanistic insight into its evolutionary potential.

## METHODS

### 2.1 Data collection

We used data from our long-term study of an insular house sparrow metapopulation (66°30’N, 12°30’E) off the coast of northern Norway (Baalsrud et al., 2014). The data set includes adult birds and fledged juveniles with parasite data recorded between 2007 and 2013 on five of the 18 islands in the metapopulation (Table 1; Figure S1). Birds were ringed as nestlings in the nest, or captured by mist-netting after fledging. Birds were marked with a metal ring with a unique identification number and three coloured plastic rings to allow individual identification. Body weight to the nearest 0.1 g and tarsus length to the nearest 0.01 mm were measured. Phenotypic measurements for tarsus length taken by different fieldworkers were adjusted to T-H. Ringsby's measurements by using paired t tests and adding mean differences when $p < .05$ (Kvalnes et al., 2017). Physical symptoms of *S. trachea* infection, including wheezing or gasping for air (Holand et al., 2014), were recorded immediately upon capture, and birds were then placed in a paper bag for a maximum of 15 min for faecal sample collection prior to measuring. Faecal samples were stored in ~1 ml of MilliQ H2O in a 1.5-ml cryotube and refrigerated at 4°C until processed. *S. trachea* faecal egg count was quantified using the flotation method described by Holand, Jensen, et al. (2013). A small blood sample (25 μl) was collected from the brachial vein of each ringed bird. Island-specific microsatellite pedigrees are available for all birds (Araya-Ajoy et al., 2019; Billing et al., 2012). Birds that survived to adulthood were typed on our custom 200k Affymetrix Axiom SNP array (Lundregan et al., 2018).

### 2.2 Parasite resistance responses

Data on birds from five islands with *S. trachea* incidence (Table 1; Figure S1) were used in our analyses. Data were collected between June and October when the parasite was most prevalent, and multiple records with parasite data were available for a number of house sparrow individuals. Resistance to *S. trachea* infection was quantified by: (i) maximum faecal egg count (FEC) as a measure of infection intensity, (ii) a binary *S. trachea* infection variable (hereafter "*Syngamus* infection status"; with 1 denoting infected birds with FEC greater than zero and 0 denoting uninfected birds), and (iii) a binary variable indicating presence or absence of physical infection symptoms (hereafter "symptomatic infection"; with 1 denoting presence of symptoms and 0 denoting absence of symptoms). For FEC and *Syngamus* infection status, if multiple measurements within a year were available for an infected individual (401 individuals) the record with the maximum FEC was selected, whereas a random record was selected if birds were not infected. Records were selected in this manner to most accurately represent maximum infection intensity, based on *S. trachea* infection trajectory and unequal sampling of individuals (Holand et al., 2015). Within the study system there is very little incidence of re-infection by *S. trachea*: for the 1% of birds that were infected in multiple years, the record with maximum FEC in the first year of infection was selected. For the symptomatic infection response, the first record where a given individual displayed physical symptoms was selected, and for birds that were never symptomatic a random record was selected.

Predictors that have previously been shown to influence resistance to *S. trachea* in the Helgeland metapopulation of house sparrows were included in our models (Holand, Jensen, et al., 2013).
These predictors were year, island, season-day and life stage (juvenile, adult male or adult female). Spatiotemporal variation in environmental conditions and house sparrow microenvironment (Pärn, Ringsby, Jensen, & Sæther, 2012; Ringsby, Sæther, Tufto, Jensen, & Solberg, 2002), as well as in S. trachea prevalence (Holand, Jensen, et al., 2013) was controlled for by including year and island as random factors in all models. Year is the year the measurement was recorded, and island is the island on which the measurement was recorded. Season-day represents the day during the season on which the measurement was taken, with the first of May set as day 1. Season-day was mean centred across all years and islands as in Holand, Jensen, et al. (2013). Sex was determined using SNP data for birds typed on our 200k array, using a sex chromosomal marker for microsatellite genotyped birds (Husby, Saether, Jensen, & Ringsby, 2006), or as the phenotypic sex for any birds without information on genetic sex. Body condition index and population density were also included as model predictors because body condition can impact immune function (Reid et al., 2006), and host population density may affect parasite prevalence (Patterson & Ruckstuhl, 2013). The unstandardized residuals from linear regressions of body mass on tarsus length for each life stage were used as a body condition index (Husby et al., 2006). Relative population size proportional to the mean adult population size on each island between 2007 and 2013 was used as a measure of population density. Generalized linear mixed models (GLMMs) implemented in the R package GLMMTMB (Brooks et al., 2017) were compared using Akaike’s Information Criterion (AIC; Burnham & Anderson, 2002) to evaluate model predictors (Table S1). The full model was supported for all S. trachea resistance traits in the full data set. FEC was also analysed in the infected subset of individuals (individuals with FEC > 0) to directly estimate additive genetic variance in infection intensity of infected birds. All analyses were carried out on a subset of the data that contained only juvenile records (hereafter denoted the “juvenile data set”), as well as in the full data set, because the genetic architecture of parasite resistance may differ in juvenile birds due to their naive immune system. The juvenile data set included all individuals that were sampled for faeces or presence of physical symptoms as juveniles, and contained information on maximum FEC, Syngamus infection status and presence/absence of symptomatic infection. The proportion of affected individuals is given in parentheses for each measure of resistance to S. trachea.

2.3 SNP data set

A faecal sample was available for 916 of the 3,116 individuals from the Helgeland metapopulation of house sparrows that were successfully typed on our 200k SNP array (Lundregan et al., 2018), and 1,209 SNP-typed individuals had data on symptomatic infection by S. trachea. In this study, only markers ranked as PolyHigh or MonoHigh resolution by Affymetrix were used (185,587 SNPs). Quality control was carried out on these markers in plink (version 1.9; Purcell et al., 2007) using all SNP-typed individuals to remove SNPs with minor allele frequency (MAF) less than 0.01, genotype success rate
less than 0.9 or call rate less than 0.95 (183,145 SNPs remained after quality control). Prior to genetic analyses using the SNP data, any missing genotypes (0.76% of the remaining 570,679,820 genotypes) were imputed using LUNKMPUTE (Money et al., 2015) to improve power. Further quality control was carried out using GENABEL (Aulchenko, Ripke, Isaacs, & van Duijn, 2007) to remove SNPs with MAF less than 0.01 in our data set, and one from each pair of individuals with identity by state higher than 0.9 (see Table 2 for final numbers of individuals used in genetic analyses on resistance to S. trachea).

### 2.4 Variance component analysis

Animal models are GLMMs that use phenotypic similarity between relatives to determine the genetic basis of a trait. This is accomplished via inclusion of a relatedness matrix, estimated using a pedigree, that represents relatedness between individuals as a random effect (Henderson, 1976; Kruuk, 2004; Wilson et al., 2010). Animal models have long been used to calculate the breeding value of domestic animals (Lynch & Walsh, 1998), and have also been applied to variance component analysis in natural populations (Kruuk, 2004; Wilson et al., 2010). Here, Bayesian animal models were fitted in r-inla (Rue, Martino, & Chopin, 2009) to estimate the proportion of variation in our parasite resistance traits (Syngamus infection status, FEC, FEC in the infected subset of individuals and symptomatic infection) that was attributable to additive genetic, dominance genetic, maternal environmental (dam) and permanent environmental (year, island) variance in our study metapopulation (see Table 1 for numbers of individuals used in variance component and heritability analyses). Bayesian inference ameliorates the uncertainty quantification and hypothesis testing issues with frequentist methods (Ovaskainen, Cano, & Merilä, 2008), and the inla approach is accurate and computationally efficient because it avoids any sampling to obtain the posterior marginal distributions (Rue et al., 2009). Model predictors were those shown to be important for resistance to S. trachea in previous studies (Holand, Jensen, et al., 2013) and by model testing (Table S1). These predictors included mean centred season-day, relative population density and body condition index as covariates, with life stage as a fixed factor (life stage was not included in analyses of the juvenile data set). Year, Island and dam were included as independent random effects in all models. Inverse additive and dominance relatedness matrices that were created in R package Nadiv (Wolak, 2012) using our microsatellite pedigree (Billing et al., 2012) were used as the covariance matrices of the random effects to estimate Va and Vd respectively. The inverse of the dominance relatedness matrix was calculated using the function ‘makeDsim’ in Nadiv with 50,000 Monte Carlo iterations. This method traces alleles through the pedigree to incorporate inbreeding effects and is accurate for complex pedigrees (Ovaskainen et al., 2008). Animal models were run in inla using the simplified Laplace approximation and central composite design integration. A penalized complexity prior for the precision with parameters u = 2, α = 0.02 was used for

| Island          | Number of individuals (proportion affected) | FEC | FEC – infected subset | Symptomatic infection |
|-----------------|---------------------------------------------|-----|-----------------------|-----------------------|
|                 | Syngamus infection status                   |     |                       |                       |
| Full data set   |                                             |     |                       |                       |
| Træna           | 104 (0.298)                                 | 104 (0.298) | 31 (1) | 140 (0.071) |
| Gjerøy          | 201 (0.189)                                 | 201 (0.189) | 38 (1) | 286 (0.059) |
| Hestmannøy      | 399 (0.331)                                 | 399 (0.331) | 132 (1) | 510 (0.112) |
| Indre Kvarøy    | 135 (0.178)                                 | 135 (0.178) | 24 (1) | 171 (0.053) |
| Aldra           | 56 (0.482)                                  | 56 (0.482) | 27 (1) | 77 (0.221)  |
| Total           | 895 (0.282)                                 | 895 (0.282) | 252 (1) | 1,184 (0.093) |
| Juvenile data set |                                           |     |                       |                       |
| Træna           | 72 (0.347)                                  | 72 (0.347) | –              | 89 (0.067)  |
| Gjerøy          | 117 (0.068)                                 | 117 (0.068) | –              | 190 (0.053) |
| Hestmannøy      | 208 (0.269)                                 | 208 (0.269) | –              | 312 (0.051) |
| Indre Kvarøy    | 78 (0.179)                                  | 78 (0.179) | –              | 105 (0.048) |
| Aldra           | 19 (0.316)                                  | 19 (0.316) | –              | 41 (0.098)  |
| Total           | 494 (0.221)                                 | 494 (0.221) | –              | 737 (0.056) |

Note: The full data set contains both adult and juvenile records, whereas the juvenile data set contains only juvenile records, and only juveniles that recruited to the adult population were typed on our 200k SNP array. GWA analyses were performed on three measures of resistance to S. trachea: faecal egg count (FEC, FEC in the infected subset of individuals). Syngamus infection status, and presence/absence of symptomatic infection. Syngamus infection status is a binary variable where individuals with FEC greater than zero were categorized as infected. Genome partitioning and GWAS were not done on the infected subset of individuals in the juvenile data set because the very low number of individuals (n = 125) led to model convergence issues.
random effects with a starting value of 2 for environmental components and 4 for genetic components. Penalized complexity priors were used as they have high density at zero distance from the base model, hence preventing model overfitting (Simpson, Rue, Riebler, Martins, & Sørbye, 2017).

Non-Gaussian traits are inherently nonadditive on the scale on which they are expressed. In GLMMs, a link function is used to model traits on a latent scale that has a Gaussian distribution and where effects are expected to be additive (de Villemereuil, 2018; de Villemereuil, Schielzeth, Nakagawa, & Morrissey, 2016). This latent scale is statistically convenient and is the scale on which heritabilities for binomial and count data are commonly reported. However, it is more biologically relevant to give quantitative genetic parameters on the expected data scale, which is the scale on which selection and evolution occur. Therefore, variance component estimates were also given on the expected data scale using the R package gglmm and the transformation methods described by de Villemereuil et al. (2016). Because fixed effects can dramatically influence quantitive genetic parameters in GLMMs due to nonlinearity caused by the link function, we averaged over the predicted values of fixed effects to obtain more accurate data-scale estimates (de Villemereuil et al., 2016).

2.5 | Heritability estimation

Parasite resistance trait heritabilities were estimated in R-INLA using the same animal models as for variance component estimation. Heritability was estimated both on the latent scale and on the expected data scale. On the latent scale, narrow-sense heritability (h²[A,lab]) can be estimated as an intraclass correlation coefficient from GLMM variance components, although GLMMs do not directly estimate the residual variance (σ²[R]), which is crucial for calculation of h². In the Poisson and negative binomial models, σ²[A] can be approximated by (1/λ), where λ is the average count and is a function of the intercept (Holand, Steinsland, Martino, & Jensen, 2013; Nakagawa & Schielzeth, 2010). The latent scale heritability of FEC and FEC in the infected subset of individuals was estimated using

\[
h²_{[A,lab]} = \frac{V_{A,lab}}{V_{A,lab} + V_{year} + V_{island} + (1/λ)}\]

where \(V_{A,lab}\) is the latent-scale additive genetic variance and \(V_{year}\) and \(V_{island}\) are latent-scale variances due to year and island respectively. To obtain a sample from the posterior distribution of \(h²_{[A,lab]}\) samples from the joint posterior distributions for \(h²_{A}\) and all variance components were generated using the function “inla.hyperpar.sample()” in R-INLA, and thus for each sample an estimate for \(h²_{[A,lab]}\) could be calculated.

In binomial GLMMs, \(h²_{[A]}\) can be approximated by \(x²/3\). Latent-scale heritability was estimated for binary response variables (Syngamus infection status and symptomatic infection) according to Nakagawa and Schielzeth (2010), using

\[
h²_{[A,lab]} = \frac{V_{A,lab}}{V_{A,lab} + V_{year} + V_{island} + (x²/3)}\]

In contrast to \(h²\) estimates on the latent scale, expected data-scale estimates of narrow sense heritability \(h²_{[A,obs]}\) from GLMMs cannot be calculated as intraclass correlation coefficients using Equation (1) or Equation (2). Instead, \(h²_{[A,obs]}\) was estimated for binary and count data as the observed data-scale additive genetic variance \((V_{A,obs})\) divided by the observed data-scale phenotypic variance \((V_{p,obs})\) (de Villemereuil et al., 2016). Broad-sense heritability \(h²_{[C,obs]}\) was also estimated on the expected data scale using

\[
h²_{[C,obs]} = \frac{V_{A,obs} + V_{D,obs}}{V_{p,obs}}\]

where \(V_{D,obs}\) is the data-scale dominance variance.

2.6 | Genome partitioning

Genome partitioning was carried out using gcta (Yang, Lee, Goddard, & Visscher, 2011), to partition the proportion of additive genetic variance explained by chromosome and evaluate genomic architecture (monogenic, oligogenic or polygenic) of parasite resistance. A positive relationship between variance explained and chromosome size is expected for polygenic traits because larger chromosomes are likely to contain more protein-coding genes (Kemppainen & Husby, 2018a; Yang, Manolio, et al., 2011). Chromosome-specific genetic relationship matrices (GRMs) were fitted simultaneously as random effects (–mrgm) in AI-REML models, to estimate the proportion of variation in each parasite response explained by each chromosome. Due to convergence issues when including model predictors and because gcta does not allow fitting of random factors, residuals from models implemented in glmmTMB (Table S1) were used as the response for all S. stercoralis resistance traits to improve consistency with variance component and GWA analyses. For FEC these residuals were approximately normally distributed (Benavides et al., 2015). Any further issues with model nonconvergence were addressed by successively excluding the smallest chromosomes. To account for heteroscedasticity and censoring, p-values for ordinary least square regressions of variance explained on chromosome size (number of SNPs) were obtained by “HC-correction” (Kemppainen & Husby, 2018b). Briefly, to give a uniform distribution of p-values under the null hypothesis of no association between phenotype and genotype, chromosomal heritability estimates \(h²_{C}\) obtained from gcta were resampled from a normal distribution with mean 0 and standard deviation (SD) equal to the standard error (SE) for each \(h²_{C}\) estimate with 10⁶ replicates.

2.7 | Genome-wide association analyses

GWA analyses using additive association tests were carried out using the R package repeated (Rönneås et al., 2016). Models used for
binary traits included mean centred season-day, relative population density and body condition index as covariates, life stage as a fixed factor (not included in analyses on the juvenile data set), and year and island as independent random effects. For count traits (FEC, FEC in the infected subset of individuals), approximately normally distributed residuals from models implemented in glmmTMB (Table S1) were used as the response to enable computationally efficient GWAS (Benavides et al., 2015). GWA analyses using nonadditive association tests were also performed for FEC and FEC in the infected subset of individuals because these traits showed substantial dominance variance in variance component analyses. Again, approximately normally distributed residuals from models implemented in glmmTMB (Table S1) were used as the response. Nonadditive GWA analyses were carried out in the R package genabel (Aulchenko et al., 2007) using the “qtscore” function to obtain p-values for a two degrees of freedom test of difference of means. Only markers for which a minimum of four individuals had the least common genotype were included. Although gene flow occurs between islands in the metapopulation, with the proportion of recruits that are dispersers ranging from 15% to 20% annually (Saatoglu et al., in prep.), there is low to moderate genetic differentiation between subpopulations (Jensen et al., 2013; Saatoglu et al., in prep.). Population stratification and relatedness were controlled for in all GWA analyses by including the breeding value that is correlated according to the GRM as a random effect.

In additive GWA analyses, Benjamini–Hochberg correction and a false discovery rate (FDR) of 0.05 were used to identify significant SNPs. In nonadditive GWA analyses, Bonferroni correction and a family wise error rate (FWER) of 0.05 were used to identify significant SNPs. The annotated house sparrow genome (Elgvin et al., 2017) was used to determine whether these markers were within or flanked by genes previously linked to immune function. To search for genes not yet annotated in the house sparrow genome, 10-kbp sequences flanking significant markers were also blasted against the chicken (Gallus gallus Gallus_gallus-5.0 (GenBank: AADN00000000.5), flycatcher (Ficedula albicollis) FicAlb1.5 (Kawakami et al., 2014; GenBank: AGTO00000000.2) and great tit (Parus major) Parus_major1.1 (Laine et al., 2016; GenBank: JRXX00000000.1) genome assemblies, using ensemblBLASTN queries with search sensitivity set to “normal” (https://www.ensembl.org/Multi/Tools/Blast?db=core).

2.8 | GO analysis

Gene ontology (GO) enrichment analyses were carried out using PANTHER 15.0 (Mi et al., 2019), using binomial tests and Bonferroni correction of p-values. All genes within 100 kbp of SNPs from GWA analyses that were significantly associated with S. trochea resistance at an FWER of 0.2 were included in input gene lists. Table S10 provides a list of genes within 100 kbp of resistance-associated markers for each trait.

3 | RESULTS

3.1 | Variance component analysis

In the analysis on the full data set (Table 3), FEC and FEC in the infected subset of individuals had a large additive genetic component. Dominance variance contributed substantially to variance in these traits, leading to reduction in $h^2_C$ estimates when the dominance relatedness matrix was included in the animal model. Additive genetic variance also contributed to variation in Syngamus infection status, and to a lesser extent variation in symptomatic infection, whereas dominance variance contributed less to variation in these traits. For Syngamus infection status and symptomatic infection, island was the most important contributor to environmental variance, whereas for FEC and FEC in the infected subset of individuals, year had a greater effect.

3.2 | Heritability estimation

In our study metapopulation, heritability of Syngamus infection status was estimated to be $h^2_C = 0.124, h^2_{liab} = 0.075$, heritability was highest for FEC in the infected subset of individuals ($h^2_{liab} = 0.351, h^2_{obs} = 0.143$), high for FEC ($h^2_{liab} = 0.239, h^2_{obs} = 0.088$), and lowest for symptomatic infection ($h^2_{liab} = 0.041, h^2_{obs} = 0.010$). Table 3 gives the highest posterior density intervals of the heritabilities. When the dominance relatedness matrix was included in models, heritabilities of FEC ($h^2_{liab} = 0.150, h^2_{obs} = 0.031$) and FEC in the infected subset of individuals ($h^2_{liab} = 0.146, h^2_{obs} = 0.019$) were reduced substantially. This reduction in heritability was also observed when the inbreeding coefficient estimated using the microsatellite pedigrees ($F_{PED}$) was included as a covariate in the models, and thus a global inbreeding effect cannot account for the observed $V_C$ in infection intensity (see Table S3). Heritability estimates on the latent scale were higher than those on the expected data scale (Table 3), which is anticipated because of irreducible noise introduced by the error process and overdispersion variance in GLMMs, as well as amplification of this noise by the inverse link function (de Villemereuil et al., 2016). Heritability estimates were similar when considering only juveniles (Tables S4 and S5).

3.3 | Genome partitioning

After HC-correction, regression of $h^2_C$ on chromosome size was nonsignificant for all Syngamus resistance traits ($p = 0.104$ for Syngamus infection status, $p = 0.540$ for FEC and $p = 0.225$ for symptomatic infection; Figure 1; Table S6). The slope of the regression for faecal egg count and for symptomatic infection was shallower than the 95% quantile from HC-resampling. This may be due to low power and wide confidence intervals on $h^2_C$ estimates, or genes related to these traits may be clustered by chromosome (Kemppainen & Husby, 2018a). Nonetheless, chromosomes 7 and
1A appeared to explain a larger than expected proportion of the variation in *Syngamus* infection status, and chromosome 7 explained a larger than expected proportion of the variation in FEC, suggesting multiple genes that contribute to variation in resistance to *S. trachea* may be clustered on these chromosomes. Genome partitioning was performed only on the full data set because of
Figure 2. Results from genome-wide association studies on *S. trachea* resistance, using additive association tests. Plots show the q-value of markers against their chromosomal position in the house sparrow genome, and horizontal dashed lines give the FDR q-value equal to 0.05. For (a) faecal egg count (FEC) two markers, SNPa421564 on chromosome 1 and SNP172501 on chromosome 1A, had a q-value <0.05. For (b) FEC in the infected subset of individuals two markers, SNPa480095 on chromosome 11 and SNPa273203 on chromosome 9, had a q-value <0.05. For (c) *Syngamus* infection status in the juvenile data set two markers, SNPa123665 and SNPa123651 on chromosome 4 had a q-value <0.05. In addition, there is a near significant peak on chromosome 5, the top marker, SNPa161700, in this peak has a q-value of 0.075. For (d) FEC in the juvenile data set two markers, SNPa518359 on chromosome 1 and SNPa194449 on chromosome 7, had a q-value <0.05. Body condition index, season-day and proportional relative population density were included as model covariates, life stage as a fixed factor (full data set only), and year, island and GRM as random effects. Markers on chromosome 16 were not on the genotyping array, and those without a position were not included. [Colour figure can be viewed at wileyonlinelibrary.com]
model convergence issues due to fewer individuals in the juvenile data set (Table 2).

### 3.4 GWAS

GWAS analyses on resistance to *S. trachea* using additive association tests (Figure 2; Table S7) revealed several genomic regions that have previously been linked to adaptive or innate immune regulation, membrane barrier function, or physiological processes relating to immunity (Table S8). Additive GWAS analyses in the full data set revealed two markers associated with FEC. In the house sparrow genome, the top marker, SNPa421564, is within a gene duplication of Dystrophin that has been linked to inflammatory response in muscular dystrophy (Porter, 2002). SNPa1723501 is located 124 kbp away from B-cell translocation gene 1 (BTK1) that is involved in cell proliferation, including proliferation of tissue-resident macrophages. For FEC in the infected subset of individuals, two markers were identified. The top SNP, SNPa480095 on chromosome 11, is located 88 kbp away from CCAAT/Enhancer-Binding Protein Alpha (C/EBPα) that negatively regulates interferon-γ (IFN-γ) expression in T-cells (Tanaka et al., 2014) and regulates mast cell and basophil lineage commitment (Huang, Li, & Liu, 2016). SNPa272203 is 459 bp away from Mab-21 Domain Containing 2 (MB21D2) that predominantly governs cell fate determination, but all MB21 proteins display a ligand binding site similar to the site in MB21D1 that activates innate immune signalling upon detection of foreign DNA (De Oliveira Mann, Kiefersauer, Witte, & Hopfner, 2016). No significant markers were detected in GWAS analyses on *Syngamus* infection status or symptomatic infection in the full data set (Table S7; Figure S2), although this may be a consequence of low power to detect markers that explain only a small proportion of variance in the phenotype (Figure S11).

Additive GWAS analyses in the juvenile data set revealed different regions and putative genes associated with *S. trachea* resistance (Figure 2; Tables S7 and S8). For *Syngamus* infection status, the top marker, SNPa123665 on chromosome 4, is closest to Gamma-aminobutyric Acid Receptor Subunit Gamma-1 (Gabrg1; 177 kbp away). Gabrg1 is a subunit of the GABA-A channel that is expressed on T-cells and may have immunomodulatory function (Kim et al., 2018). For FEC, the top marker, SNPa518539 on chromosome 1, lies between Translatonally-controlled Tumor Protein Homolog (TPT1, alias p23, 40 kbp away) and BTB/POZ Domain-containing Protein KCDT4 (Kcdt4, 22 kbp away) in the house sparrow genome. Tpt1 is highly expressed in tick-infected cattle, and may be involved in the inflammatory response (Nascimento et al., 2010). The second marker associated with FEC in juveniles, SNPa194449 on chromosome 7, is between Retinol Dehydrogenase 3 (Rdh3, 12 kbp away) and ABCB11 (2 kbp away) in the house sparrow genome. Rdh3 is involved in synthesis of retinoic acid (vitamin A) that is integral to immune function (Broadhurst et al., 2012; Hall, Grainger, Spencer, & Belkaid, 2011). ABCB11 encodes the bile salt export pump that transports bile acids, which are important for innate immunity (Fiorucci, Biagioli, Zampella, & Distrotti, 2018) and are essential during development of some parasites (Smyth, 1962). Chromosome 5 also showed a notable peak for *Syngamus* infection status in the GWAS analysis in the juvenile data set (Figure 2c). The top marker in this peak, SNPa161700 (Table S7), is closest to Calmodulin (43 kbp away) in the house sparrow genome. Calmodulin promotes T-helper cell immunity (Bikah, Pogue-Caley, McHeyzer-Williams, & McHeyzer-Williams, 2000) that is important in protection against helminth parasites (Anthony, Rutitzky, Urban, Stadecker, & Gause, 2007). Although the MAF of several markers detected in additive GWAS analyses was low, the island-specific distribution of genotypes in relation to infection dynamics in the study population suggested that population structure did not drive associations between these markers and resistance to *S. trachea* (see Figures S3–S6 for genotype phenotype plots for SNPs detected in additive GWAS analyses). Furthermore, simulation tests using null markers indicated that SNPs associated with parasite resistance in this study are likely to represent genuine trait associations (see Tables S12 and S13; Figure S10). No significant markers were identified in GWAS analyses on symptomatic infection in the juvenile data set (Table S7; Figure S2), although power to detect markers that did not explain a large proportion of the phenotypic variance was low (Figure S11). A GWAS was not performed on the infected subset of individuals in the juvenile data set, due to model convergence issues caused by small sample size.

GWAS analyses on faecal egg count using nonadditive association tests also revealed multiple genomic regions linked to immune regulation, membrane barrier and repair processes, or immune-related physiological processes (Figure 3; Table S9). Several genes that have previously been implicated in the regulation of the immune response were detected. These included: Suppressor of Cytokine Signaling 6 (SOCS6; Duncan, Baganzi, Sahu, Singh, & Dennis, 2017), and the matrix metalloproteinases Matrilysin and Matrix Metalloproteinase-27 (MMP7 and MMP27 respectively; Madala et al., 2010) that were associated with FEC in the full data set, Phosphatidylinositol 4-Phosphate 3-Kinase C2 Domain-Containing Subunit Alpha (PIKC32A; Hawkins & Stephens, 2015) that was associated with FEC in the infected subset of individuals, and E3 Ubiquitin-Protein Ligase HECW1 (HECW1; Ebner, Versteeg, & Ikeda, 2017), Polypeptide N-Acetylgalactosaminyltransferase 11 (Galnt1; Tenno et al., 2007) and SH3 Domain-Containing YSC84-like Protein 1 (Sh3yl1; Fernandes et al., 2019) that were associated with FEC in the juvenile data set. The genes Iroquois-Class Homeodomain Protein IRX-5 (IRX-5; Tuyt et al., 2006), Family with Sequence Similarity 92 Member A (FAM92A11; Li et al., 2016), MMP7 and MMP27 (Madala et al., 2010) were associated with FEC in the full data set and have been respectively linked to lung and foregut morphogenesis, mucus membrane ciliogenesis, and tissue repair. Hydroxyacid Oxidase 2 (HAO2) was linked to FEC in the juvenile data set and is involved in production of reactive oxygen species that are important in defence against parasites (Fransen, Nordgren, Wang, & Apanasets, 2012). Table S9 gives a full list of genes linked to *S. trachea* resistance in nonadditive GWAS analyses, as well as an overview of their functions. As several
of the detected dominance associations are driven by comparatively few individuals with high faecal egg count, these results should be interpreted with caution (see Figures S7–S9 for genotype phenotype plots for significant SNPs). However, low frequencies of genotypes that reduce resistance to *S. trachea* are expected, due to the negative impacts of the parasite on survival and reproductive success (Holand et al., 2014, 2015).

### 3.5 GO analysis

Functional analysis based on grouping of genes by GO-term enrichment identified three functional groups linked to genes from nonadditive GWA analyses (Table S11). One functional group (GO:0030574) was identified for faecal egg count in the full data set that related to collagen catabolic process. The matrix metalloproteinases (MMP1, MMP3, MMP7, MMP27) enriched for this GO term play a role in collagen degradation during extracellular matrix remodelling, and have been linked to the inflammatory response to gastrointestinal nematode infection (Chitneedi, Suárez-Vega, Martínez-Valladares, Arranz, & Gutiérrez-Gil, 2018) and helminth-induced lung fibrosis (Madala et al., 2010). Two functional groups (GO:0071805, GO:0006813) that related to potassium ion transport were identified for FEC in the infected subset of individuals. Epithelial ion transport may contribute to parasite expulsion (Baird & O’malley, 1993), and several parasites have been shown to inhibit membrane ion transport in rat models (Kapczuk et al., 2018). No functional groups were identified for genes detected in additive GWA analyses (Table S11), due to the small number of genes that were significant at an FWER of 0.2 (Table S10).
4 | DISCUSSION

The current study is one of the first to assess additive and dominance genetic variance in parasite resistance in natural populations. We provide evidence that S. trachea infection status and infection intensity (FEC and FEC in the infected subset of individuals) have substantial additive genetic variance leading to moderate heritability estimates (Table 3; $h^2_{ab} = 0.124, 0.239$ and 0.351 respectively) that are in line with those from previous studies. A genetic basis for parasite infection intensity has previously been established in commercially farmed animals, with heritability estimates for faecal egg count in the region of 0.30 for cattle and pig (Gasbarre, Leighton, & Davies, 1990; Leighton, Murrell, & Gasbarre, 1989; Nejsun et al., 2009), 0.15–0.40 for sheep (Gauly, Kraus, Vervelde, Van Leeuwen, & Erhardt, 2002; Stear et al., 2007; Woolastont & Windon, 2001) and 0.35–0.41 for chicken (Psifidi et al., 2016). Comparatively few studies have assessed the genetic basis of parasite resistance in natural populations. In Soay sheep infected by gastrointestinal nematodes, heritability estimates for faecal egg count range from 0.11 to 0.26 (Colman et al., 2001; Smith et al., 1999), and a study on gastrointestinal nematode burden in red grouse estimated a heritability of 0.29 for the same trait (Wenzel et al., 2015). In natural populations, a substantial proportion of the variation in parasite resistance may be due to environmental conditions, and hence heritability estimates may be lower than for similar traits in commercial breeds. In our house sparrow metapopulation, heritability of symptomatic infection by S. trachea was low ($h^2_{ab} = 0.041$), which may be a consequence of low power as only 9% of individuals were symptomatic at the time of assessment (Table 1). However, symptomatic infection is a fitness-related trait (Holand et al., 2014), and theory and empirical studies suggest that genes involved in parasite resistance are likely to display high genetic variation due to balancing selection (Hedrick, 2002). Although additive genetic effects contributed to variation in symptomatic infection (Table 3), environmental effects along with any measurement error outweigh the detectable influence of additive genetics on this trait.

Non-additive genetic variance can be an important source of variation for fitness-related immune traits (Bernatchez & Landry, 2003; Mukherjee, Sarkar-Roy, Wagener, & Majumder, 2009; Psifidi et al., 2016). In our house sparrow metapopulation, dominance variance contributed little to Syngamus infection status or symptomatic infection. Estimates of additive genetic variance for these measures of parasite resistance were similar regardless of whether the dominance relatedness matrix was included in the animal models. However, there was evidence for substantial dominance variance in S. trachea infection intensity, and estimates of $V_D$ were reduced when the dominance relatedness matrix was included (Table 3; Tables S5–S5). For traits with a large dominant genetic component, it is known that a proportion of $V_D$ can be expressed as $V_A$ when dominance is not included in the model (Heidaritabar et al., 2016; Ovaskainen et al., 2008) or if the data do not allow complete separation of the variance components (Wolak & Keller, 2014). Although the 95% credible intervals for $V_A$ and $V_D$ estimates in the present study are wide (Table 3; Tables S3 and S4), large error margins are common for non-Gaussian data and infection intensity is nonetheless likely to have a large dominance component. Studies in commercial breeds show that $V_D$ estimates are often high for fitness-related traits under strong selection (Lopes et al., 2015; Misztal et al., 1998), and that $V_D$ may greatly exceed $V_A$ for such traits (Cronk and Roff, 1995; Pante et al., 2002). However, estimates of dominance variance in natural populations are largely lacking, and thus the present study is one of the first to demonstrate dominance variance in a fitness-related trait in nature. Our results indicate that accounting for dominance in variance component analyses in natural populations may be important, especially for fitness-related traits such as disease resistance, as failure to do so may lead to upwardly biased $V_A$ and $h^2$ estimates and overestimation of evolutionary response to selection (Wolak & Keller, 2014). However, power to estimate $V_D$ accurately may be limited in most data sets from natural populations.

In genome partitioning analyses, regression slopes were shallow (Figure 1), which may indicate clustering of genes related to resistance to S. trachea on specific chromosomes. Therefore polygenic inheritance of resistance to S. trachea is likely, but with marker effect sizes that are distributed nonuniformly throughout the genome (Kemppainen & Husby, 2018a; Thompson et al., 2015), although this result should be interpreted with caution due to the wide confidence intervals on $h^2_C$ estimates. This is in line with expectations for host immunity, which is governed by a complex, genome-wide network of genetic pathways that make up the total immune response (Benavides et al., 2016). Genome partitioning has previously been used to identify a polygenic basis for gastrointestinal nematode infection intensity in a natural population of red grouse (Wenzel et al., 2015), although the strength of the result may be overstated because heteroscedasticity and censoring were not accounted for (Kemppainen & Husby, 2018b).

In our GWA analyses on S. trachea resistance, multiple genomic regions that have previously been linked to parasite resistance or immune function were detected (Tables S8 and S9). GWAS in livestock have found numerous genes linked to parasite resistance and support the idea that resistance is polygenic. Several of these studies used Ig antibody levels (Boulton et al., 2018; Psifidi et al., 2016) or other immunity indexes (Berton et al., 2017; Boulton et al., 2018; Thompson-Crispi et al., 2014) as a phenotype to reveal genes that were often strongly linked to adaptive immunity, suggesting a major role for the MHC in parasite resistance. Ig levels in a natural population of Soay sheep have been linked in a GWAS to the MHC, or genes that govern the MHC (Sparks et al., 2019), confirming the importance of adaptive immunity to gastrointestinal nematodes in nature. Therefore, it is important to note that SNPs on chromosome 16, which is the genomic location of the MHC in the house sparrow, are not present on our 200k SNP array because of sequence assembly challenges directly related to the repetitive nature of the MHC gene family (Elgvin et al., 2017). Nevertheless, GWA analyses on raw or transformed faecal egg counts in livestock have frequently identified genes linked to adaptive and innate immunity, as well as diverse physiological processes, which indicates genes influencing
infection intensity are distributed throughout the genome (Berton et al., 2017; May et al., 2019; Pickering et al., 2015; Psifidi et al., 2016; Silva et al., 2018).

Several genes that were associated with resistance to S. trachea in our GWA analyses (Tables S8 and S9) have previously been linked to immunity against parasites. In additive GWA analyses in this study, the top marker for FEC was intronic to a gene duplication of Dystrophin in the house sparrow genome. The Dystrophin–glycoprotein complex that is important for membrane stability has previously been linked to FEC in a GWAS study on New Zealand sheep (Pickering et al., 2015), and previous studies have shown that genes governing mucous membrane barrier function are involved in physical expulsion of parasites (Benavides et al., 2016; De Veer et al., 2007). Furthermore, CEBPα that was associated with FEC in the infected subset of individuals in the present study is related to CEBPβ (CCAAT/Enhancer Binding Protein β) that has been linked to nematode burden in wild grouse (Wenzel et al., 2015). In our juvenile data set, the top marker associated with FEC was 40 kb away from TPT1 in the house sparrow genome. TPT1 is highly expressed in tick-infected cattle and may be involved in the inflammatory response to parasite challenge (Nascimento et al., 2010). A second marker associated with FEC in the juvenile data set was 12 kb away from Rdh3, which aids synthesis of vitamin A that is important for immune function (Hall et al., 2011). The LRAT gene involved in vitamin A synthesis has been linked by GWAS to FEC in creole goats (Silva et al., 2018), and vitamin A deficiency has been shown to diminish immunity to gapeworm in chicken (Clapham, 1934). Finally, nonadditive GWA analyses revealed two genes that have previously been linked to parasite resistance, Alpha-Ketoglutarate-Dependent Dioxygenase FTO (FTO) that has been associated with trematode-induced liver damage in GWA analyses on dairy cattle (Twomey et al., 2019), and Costars Family Protein ABRACL (ABRACL) that is abundant in resistant mice after nematode infection (Deslyper, Colgan, Cooper, Holland, & Carolan, 2016).

In summary, our GWA analyses on S. trachea resistance revealed genes linked to mucus membrane barrier function and physiological processes, as well as genes linked to the innate immune response (Tables S8 and S9), which is consistent with results from previous GWA analyses on parasite resistance traits. Although chromosome 16 containing the MHC complex is not included on our 200k SNP array, many of the regions associated with parasite resistance are regulators of adaptive immune components such as B1-cells and T-cells. This may indicate a role of adaptive immunity for resistance to S. trachea in the house sparrow that supports previously reported differences in prevalence between life stage groups (Holand, Jensen, et al., 2013). Acquired immunity to S. trachea (Atkinson et al., 2008) can explain both the predisposition to infection as juveniles and the low re-infection rates we observe in our study metapopulation. Thus, physiological processes such as vitamin A synthesis, production of reactive oxygen and nitrogen species, and the initial inflammatory response may represent a first line of defence against S. trachea in juvenile house sparrows, where the immune system is not yet fully developed.

The substantial additive genetic variance (Table 3; Table S2) observed in S. trachea infection intensity (FEC), and to a lesser extent in Syngamus infection status, suggests that genetic differences between individuals contribute to previously reported spatiotemporal variation in S. trachea prevalence within our study system (Holand, Jensen, et al., 2013). High levels of additive genetic variance allow populations to respond to challenge by rapidly evolving parasites (Acevedo-Whitehouse & Cunningham, 2006; MacPherson et al., 2018). This may have important eco-evolutionary consequences in our study metapopulation because FEC is a fitness-related trait that negatively impacts lifetime reproductive success of female house sparrows (Holand et al., 2015). The presence of substantial standing genetic variation may buffer populations against negative consequences of factors that lead to increased parasite challenge, such as changes in abiotic environment, greater paratenic host availability (Holand, Jensen, et al., 2013) or higher temperatures (Altizer et al., 2013; Harvell et al., 2009; Holand et al., 2019). Furthermore, the observed polygenic nature of the immune response to S. trachea could lead to durable resistance phenotypes because resistance results from cumulative, genome-wide allele frequency shifts that represent a diverse target for parasite co-evolution (Woolhouse et al., 2002). However, in polygenic resistance, severity of infection is often reduced rather than infection being prevented, which could select for increased parasite virulence (Gandon & Michalakis, 2000; Gates, Valletta, Bonneau, & Recker, 2018), or increase infection prevalence (Westerdahl et al., 2012). The large dominance variance contribution to S. trachea infection intensity should increase the strength of selection at resistance alleles (Conner & Hartl, 2004). However, V_D can reduce V_A recovery when genetic drift is strong (Barton et al., 2006), which poses a particular problem in small or fragmented populations. These considerations are especially relevant in light of results in Holand et al. (2019) that show S. trachea prevalence increases with temperature in our study metapopulation of house sparrows, as any intensified parasite challenge increases evolutionary pressure on the host and may influence co-evolutionary dynamics. Together, the polygenic genetic architecture of S. trachea resistance and contribution of V_A to infection intensity suggest that preservation of genetic diversity may be particularly important for natural bird populations, if they are to overcome changes in parasite prevalence and virulence that are predicted with rising temperatures (Altizer et al., 2013; Harvell et al., 2009; Holand et al., 2019) and increased anthropogenic stressors (Marcogliese & Pietrock, 2008).

ACKNOWLEDGEMENTS

We thank students and fieldworkers for help with fieldwork, students and laboratory technicians for assistance with genotyping and genetic parentage analyses, and Grete Stavik Eggen, Randi Rasbak and Henriette Vaagland who provided valuable help with analysing faecal samples. We are also grateful to the inhabitants of our study area at Helgeland, whose hospitality made this study possible. We thank Christophe Pélabon, Yimen Araya-Ajoy and others at the Centre for Biodiversity Dynamics (CBD), NTNU, for fruitful discussions, and Anna Santure and two anonymous reviewers for helpful
feedback on a previous version of the manuscript. This study was funded by the Research Council of Norway (RCN grant nos. 214553, 221956, 274930 and 302619), the RCN’s Centres of Excellence funding scheme (grant no. 223257), the Academy of Finland (grant no. 295204 to A.K.N.), and the Norwegian University of Science and Technology. The empirical research was carried out in accordance with permits from the Ringing Centre at Stavanger Museum, Norway.

AUTHOR CONTRIBUTIONS
S.L.L., H.J. and A.K.N. designed the study. H.J., H.H., T.H.R. and T.K. collected field data. H.H. quantified parasite faecal egg count. H.J., H.H. and T.K. developed the microsatellite pedigree. S.L.L. analysed the data with guidance from H.J., A.K.N., S.M., H.H., T.K. and A.H. The article was written by S.L.L. with input from all authors.

DATA AVAILABILITY STATEMENT
Phenotypic and genetic data used in this study are available on Dryad (https://doi.org/10.5061/dryad.5dv41ns3g).

ORCID
Sarah L. Lundregan https://orcid.org/0000-0002-4971-6208
Alina K. Niskanen https://orcid.org/0000-0003-2017-2718
Arild Husby https://orcid.org/0000-0003-1911-8351
Henrik Jensen https://orcid.org/0000-0001-7804-1564

REFERENCES
Acevedo-Whitehouse, K., & Cunningham, A. A. (2006). Is MHC enough for understanding wildlife immunogenetics? Trends in Ecology & Evolution, 21(8), 433–438. https://doi.org/10.1016/j.tree.2006.05.010
Altizer, S., Ostfeld, R. S., Johnson, P. T. J., Kutz, S., & Harvell, C. D. (2013). Climate change and infectious diseases: From evidence to a predictive framework. Science, 341(6145), 514–519. https://doi.org/10.1126/science.1239401
Andrew, S. C., Jensen, H., Hagen, I. J., Lundregan, S. L., & Griffith, S. C. (2018). Signatures of genetic adaptation to extremely varied Australian environments in introduced European house sparrows. Molecular Ecology, 27(22), 4542–4555. https://doi.org/10.1111/mec.14897
Andrew, S. C., Taylor, M. P., Lundregan, S. L., Lien, S., Jensen, H., & Griffith, S. C. (2019). Signs of adaptation to trace metal contamination in a common urban bird. Science of the Total Environment, 650(1), 678–686. https://doi.org/10.1016/j.scitotenv.2018.09.052
Anthony, R. M., Rutitzky, L. I., Urban, J. F., Stadecker, M. J., & Gause, W. C. (2007). Protective immune mechanisms in helminth infection. Nature Reviews Immunology, 7(12), 975–987. https://doi.org/10.1038/nri2199
Araya-Ajoy, Y. G., Ranke, P. S., Kvalnes, T., Ranneing, B., Holand, H., Myhre, A. M., ... Wright, J. (2019). Characterizing morphological (co)variation using structural equation models: Body size, allometric relationships and evolvability in a house sparrow metapopulation. Evolution, 73(3), 452–466. https://doi.org/10.1111/evol.13668
Atkinson, C. T., Thomas, N. J., & Hunter, D. B. (2008). Parasitic diseases of wild birds. Ames, IA: Wiley-Blackwell.
Aulchenko, Y. S., Ripke, S., Isaacs, A., & van Duijn, C. M. (2007). GenABEL: An R library for genome-wide association analysis. Bioinformatics, 23(10), 1294–1296. https://doi.org/10.1093/bioinformatics/btm108
Baalsrud, H. T., Sæther, B.-E., Hagen, J. J., Myhre, A. M., Ringsby, T. H., Pärn, H., & Jensen, H. (2014). Effects of population characteristics and structure on estimates of effective population size in a house sparrow metapopulation. Molecular Ecology, 23(11), 2653–2668. https://doi.org/10.1111/mec.12770
Baird, A. W., & O’Malley, K. E. (1993). Epithelial ion transport – Possible contribution to parasite expulsion. Parasitology Today, 9(4), 141-142. https://doi.org/10.1016/0169-4758(93)90180-N
Barton, N. H., Turelli, M., & Hill, W. G. (2006). Prediction of effects of genetic drift on variance components under a general model of epistasis. Theoretical Population Biology, 70(1), 56–62. https://doi.org/10.1016/j.tpb.2005.10.001
Barus, V. (1966). The effect of temperature and air humidity on the development and the resistance of eggs of the nematode Syngamus trachea. Helminthologica (Breslavia), 7(2), 103–106.
Benavides, M. V., Sonstegard, T. S., Kemp, S., Mugambi, J. M., Gibson, J. P., Baker, R. L., ... Van Tassell, C. (2015). Identification of novel loci associated with gastrointestinal parasite resistance in a red Maasai x Dorper backcross population. PLoS ONE, 10(4), e0122797. https://doi.org/10.1371/journal.pone.0122797
Benavides, M. V., Sonstegard, T. S., & Van Tassell, C. (2016). Genomic regions associated with sheep resistance to gastrointestinal nematodes. Trends in Parasitology, 32(6), 470–480. https://doi.org/10.1016/j.pt.2016.03.007
Berghoff, T. V. L., Visker, M. H. P. W., Arts, J. A. J., Parmentier, H. K., van der Poel, J. J., Vereijken, A. L. J., & Bovenhuis, H. (2018). Genomic region containing toll-like receptor genes has a major impact on total IGM antibodies including KLH-binding IGM natural antibodies in chickens. Frontiers in Immunology, 8, 1879. https://doi.org/10.3389/fimmu.2017.01879
Bernatchez, L., & Landry, C. (2003). MHC studies in nonmodel vertebrates: What have we learned about natural selection in 15 years? Journal of Evolutionary Biology, 16(3), 363–377. Retrieved from http://www.blackwell-synergy.com/doi/abs/10.1046/j.1400-9101.2003.00531.x
Berton, M. P., de Oliveira Silva, R. M., Periopoli, E., Stafuzza, N. B., Martin, J. F., Álvarez, M. S., ... Ferraz, J. B. S. (2017). Genomic regions and pathways associated with gastrointestinal parasites resistance in Santa Inês breed adapted to tropical climate. Journal of Animal Science and Biotechnology, 8(1). https://doi.org/10.1186/s4104-017-0190-4
Bikah, G., Pogue-Caley, R. R., McHeyzer-Williams, L. J., & McHeyzer-Williams, M. G. (2000). Regulating T helper cell immunity through antigen responsiveness and calcium entry. Nature Immunology, 1(5), 402–412. https://doi.org/10.1038/80841
Billing, A. M., Lee, A. M., Skjelseth, S., Borg, Å. A., Hale, M. C., Slate, J., ... Jensen, H. (2012). Evidence of inbreeding depression but not inbreeding avoidance in a natural house sparrow population. Molecular Ecology, 21(6), 1487–1499. https://doi.org/10.1111/j.1365-294X.2012.05490.x
Boulton, K., Nolan, M. J., Wu, Z., Riggio, V., Matika, O., Harman, K., ... Psifidi, A. (2018). Dissecting the genomic architecture of resistance to Eimeria maxima parasitism in the chicken. Frontiers in Genetics, 9, 528. https://doi.org/10.3389/fgene.2018.00528
Broadhurst, M. J., Leung, J. M., Lim, K. C., Girgis, N. M., Gundra, U. M., Fallon, P. G., ... Loke, P. (2012). Upregulation of retinal dehydrogenase 2 in alternatively activated macrophages during retinoid-dependent type-2 immunity to helminth infection in mice. PLoS Pathogens, 8(8), e1002883. https://doi.org/10.1371/journal.ppat.1002883
Brooks, M. E., Kristensen, K., van Bentham, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R Journal, 9(2), 378–400. https://doi.org/10.3299/ETHZ-B-000240890
