Check for updates

Molecular Ecology

WILEY

MOLECULAR ECOLOGY

ORIGINAL ARTICLE

Simulations of Life History Variation for Demographic Inference From Population Genomic Data

Rilquer Mascarenhas[1,2] | Michael J. Hickerson[1,2,3] | Ana Carolina Carnaval[1,2]

[1]The City College of New York, New York City, New York, USA | [2]The Graduate Center of the City University of New York, New York City, New York, USA | [3]American Museum of Natural History, Division of Invertebrate Zoology, New York City, New York, USA

Correspondence: Rilquer Mascarenhas (rmascarenhasdasilva@gradcenter.cuny.edu)

Received: 28 January 2025 | **Revised:** 30 June 2025 | **Accepted:** 7 October 2025

Handling Editor: Jason Bragg

Funding: R.M., M.J.H. and A.C.C. acknowledge funding by NSF award DBI 1926928 (IIBR RoL: Collaborative Research: A Rules of Life Engine (RoLE) Model to Uncover Fundamental Processes Governing Biodiversity), which partly funded Mascarenhas as he developed this work. R.M. also acknowledges support from the Graduate Center and The City College of the City University of New York.

Keywords: age-structured populations | Amazon forest | comparative phylogeography | nucleotide diversity | polygamy

ABSTRACT

Ecological differences among species, particularly dispersal capacity and life history strategies, influence population response to environmental changes. Genetic simulations now allow us to directly incorporate this variation into models of past demographic changes. However, the impact of life history strategies in demographic inference has been far less explored relative to that of dispersal capacity. Here, we utilise individual-based simulations of a non-Wright-Fisher population to ask whether differences in life history traits (the average age of first reproduction of individuals, the average adult mortality and the average number of mates per reproductive season) lead to consistent and predictable differences in the summary statistics of genetic diversity commonly used for simulation-based parameter estimation and demographic inference. Using a Random Forest model, we also estimate three population parameters (variance in reproductive success, generation time and effective population size) from genome-wide SNP variation for two bird species known to have distinct life history strategies. The results demonstrate that life history variation leads to predictable differences in patterns of genetic diversity: higher values of life history traits, representing extreme polygamy, long adult longevity and later onset of reproduction, are associated with higher variance in reproductive success, longer generation times, smaller effective population sizes and overall lower genetic diversity. Parameter estimates from empirical datasets also agree with the general expectation that polygamic species with later onset of reproduction and long adult longevity exhibit higher variance in reproductive success, longer generation times and smaller effective population sizes. Since the signal of life history differences is observed in the genetic summary statistics, we argue that simulation- and model-based multi-species demographic inference will gain from the incorporation of life history information.

1 | Introduction

Ecological characteristics influence how individuals interact with their surroundings and are therefore central to understanding species responses to environmental change (Comte et al. 2024; Germain et al. 2023). Species-specific morphological traits (Pabijan et al. 2012; Paz et al. 2015), unique microhabitat associations (Massatti and Knowles 2014) and behaviour (Burney and Brumfield 2009; Miller et al. 2021), for instance, impact individual dispersal capacity and can lead to differences among species in inferred present and past population connectivity. Similarly, differences in life history

© 2025 John Wiley & Sons Ltd.

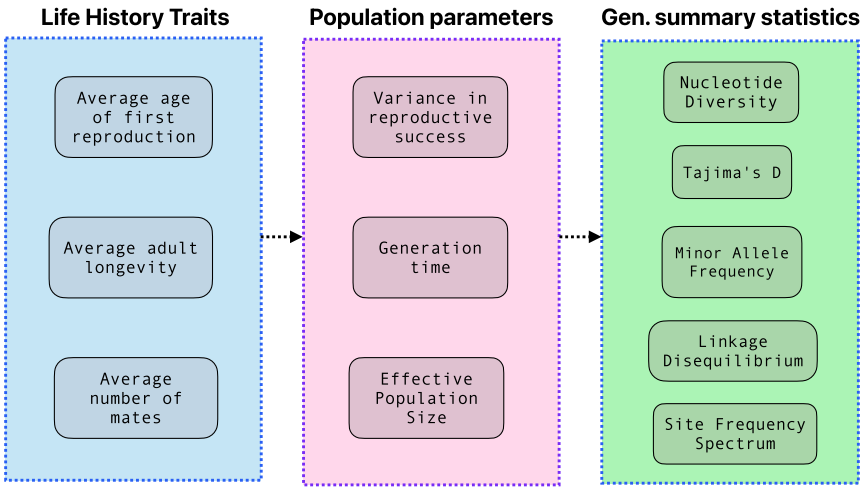

FIGURE 1 | Description of life history traits, population parameters and genetic summary statistics utilised in this simulation study. Life history traits are average values for traits measurable at the population level. Different values were provided to the model at the beginning of forward simulations (see main text). Population parameters are values that describe characteristics of the population and emerge from the simulations. Genetic summary statistics are values used to describe patterns of genomic variation at the end of simulations.

strategies across species are expected to lead to differences in inferred demographic changes over time, due to their effect on reproductive and mortality rates in a population (Sæther et al. 2004; Coulson et al. 2010). Collectively, these differences may result in idiosyncratic patterns of genetic variation across co-distributed species, even under a shared history of environmental shifts (Carvalho et al. 2021; Fan et al. 2016; García-Rodríguez et al. 2020; Zamudio et al. 2016). To accurately infer demographic responses to past environmental changes, it is therefore important to account for ecological variation when building multi-species inferential models. Incorporating such variation allows for a better interpretation of pattern comparison across species, where the expectation of pattern concordance can be refined based on known ecological differences (Papadopoulou and Knowles 2015, 2016).

Inference of former demographic change is generally made through the estimation of parameters that describe the history of a population (i.e., population parameters) from empirical genomic data, given a probabilistic model (Hickerson et al. 2010; Knowles and Maddison 2002). Because these models are often too complex for parameter estimation in an analytical approach (Sunnåker et al. 2013), genetic simulations are commonly used to explore population parameter space (i.e., all possible parameter values), and parameter estimation is made by comparing simulated and observed datasets using approaches such as Approximate Bayesian Computation and Statistical Machine Learning (Csilléry et al. 2010; Schrider and Kern 2018; Sunnåker et al. 2013). The flexibility of available simulation tools allows us to explore ways of incorporating ecological variation across species to add more ecological realism into genetic simulations (e.g., Blischak et al. 2020; Sackman et al. 2019; Tellier and Lemaire 2014; Xue and Hickerson 2020). Much of this effort has focused on the acknowledgement of ecological differences that impact spatial population connectivity, which have been incorporated either implicitly (e.g., when allowing wide priors for interspecific variation in migration rates in cross-species simulations; Xue and Hickerson 2020) or explicitly (e.g., by using ecological niche models as proxies of habitat preferences to inform landscape connectivity; Knowles and Alvarado-Serrano 2010).

One far less explored aspect, however, is age-dependent individual variation in the probability of reproduction and death, which is translated into rates of reproduction and death within a population (i.e., demographic rates; Saether et al. 2013). Differences across species in these demographic rates give rise to different life histories, which can be summarised as *life history traits* measurable at the population level: the average age at which individuals first reproduce, the average longevity of adult individuals and the average number of mates across individuals that successfully reproduce (Figure 1 left column; Lee et al. 2011; Waples et al. 2013). The combination of these life history traits governs two key *population parameters* (Figure 1, central column): the variance in reproductive success across individuals (Waples 2023) and the average age of parents in a reproductive cohort, usually referred to as the generation time (Ellner 2018). When individuals begin reproducing early in life (i.e., the average age at which individuals first reproduce is low), both the variance in reproductive success across individuals of different ages as well as the average age of reproducing individuals (i.e., generation time) tend to be small. This life history strategy characterises short-lived, fast-reproducing species (Sæther et al. 2004). Alternatively, if individuals begin reproducing later in life, reproductive success is achieved in older individuals, leading to higher variance in reproductive success across individuals and higher generation times, characterising a long-lived, slow-reproducing species. Variance in reproductive success and generation time are also influenced by the mating system exhibited by species (Nunney 1991), which is reflected in the average number of mates per reproducing individual. When this average number is high (as the case, for instance, in polygamic species), variance in reproductive success is increased since a small proportion of individuals in the current generation will be contributing to the next one (Gaiotti et al. 2020;

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College , Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

McDonald 1993). Simultaneously, if older individuals exhibit higher reproductive success (which is the case in many polygamic species; Hasselquist 1998; Ryder et al. 2009; Dubuc et al. 2014), the average age of reproducing individuals (i.e., generation time) tends to be higher.

These differences in generation time and reproductive success across individuals are relevant when inferring demographic change over time because they impact the number of reproducing individuals and consequently the effective population size (N_e)—a key population parameter that is directly linked to genetic diversity (Crow and Denniston 1988; Nomura 2002; Nunney 1991; Waples 2022b; Figure 1, central column). Populations with higher variance in reproductive success and longer generation time are expected to have fewer reproducing individuals and smaller effective population sizes (Nunney 1991, 1993; Waples 2023). As such, variation in life history traits across species is expected to impact (and be reflected in) the *summary statistics* used to describe patterns of genomic variation (Figure 1, right column). Nucleotide diversity, minor allele frequency and the site frequency spectrum, for instance, are all examples of summary statistics often employed in studies of historical demography (Hickerson et al. 2006; Knowles 2009; Tataru et al. 2017; Xue and Hickerson 2015), and which are all expected to be impacted by those life history traits. These summary statistics are commonly implemented in likelihood-free methods using simulations to infer the history of populations because they contain information on key demographic parameters, such as past and present effective population sizes, as well as divergence times (Hickerson et al. 2010; Knowles 2009). If cross-species differences in life history parameters, like generation time and individual variance in reproductive success, are also reflected in cross-species differences in genetic summary statistics, it may be possible to estimate those life history parameters from genomic data (Bradburd and Ralph 2019; Ianni-Ravn et al. 2024). However, it remains to be seen (i) whether differences in life history traits lead to consistent differences in genetic summary statistics across species, (ii) whether the use of this set of summary statistics results in accurate estimates of population parameters under life history variation, and (iii) whether predictions based on these same summary statistics from empirical data agree with expectations based on observed life history strategies. This study aims to fill these knowledge gaps.

Here, we use genetic simulations to explore if and how cross-species differences in three target life history traits (average age of first reproduction, average adult longevity and average number of mates) lead to species-specific differences in key summary statistics inferred from genetic data. We further verify if they lead to different estimates of population parameters from empirical data of co-occurring species with contrasting life history strategies. Incorporating life history traits into genetic simulations is not straightforward because most simulation tools used to date rely on the Wright-Fisher (WF) population model (Fisher 1922; Wright 1931; Kingman 1982), which makes simplifying ecological assumptions such as random mating across individuals and non-overlapping generations (Haller and Messer 2019). Although other population models, such as the Moran model, have relaxed some of these assumptions (Hahn 2018), they are still limited both in the incorporation of life history traits measurable at the population level and in the exploration of how these traits interact to yield patterns of genomic variation. Additionally, they do not allow for cases where age-dependent reproductive success varies across sex (as is the case for polygamic species). To circumvent this, we implement individual-based simulations under a non-Wright Fisher model to allow for age-structured populations where the probability of reproduction and death depends on individual age. This allows us to explore how variation in values of the aforementioned life history traits (average age of first reproduction, average adult longevity and average number of mates across reproducing individuals) impact population parameters (variance in reproductive success, generation time and effective population size). The values of population parameters emerging from each combination of life history traits are then combined into coalescent simulations for the calculation of genetic summary statistics, allowing us to map variation in life history traits to patterns of genomic variation (Figure 1). We expect that species whose individuals reproduce and die at an older age, or where the number of mates per individual is higher (i.e., species with more extreme forms of polygamy), will display higher variance in reproductive success, longer generation times, smaller effective population sizes and, consequently, lower levels of genetic diversity. Accordingly, we expect the opposite pattern for fast-reproducing species.

We additionally use our simulations to explore how accurately one can estimate the aforementioned population parameters (variance in reproductive success, generation time and effective population size; Figure 1) from genomic data under models that incorporate variation in the life history traits. For that, we implement a likelihood-free estimation method using supervised machine learning (Collin et al. 2021; Fonseca et al. 2021; Schrider and Kern 2018), allowing us to quantify our ability to accurately predict generation time, effective population size and variance in reproductive success using genetic summary statistics as predictor variables. To illustrate the applicability of this approach, we use it to estimate ecological parameters from published genomic information from two passerine birds: *Pipra filicauda* (the Wire-tailed Manakin, a slow-reproducing, long-lived species) and *Thamnophilus aethiops* (the White-shouldered Antshrike, a fast-reproducing, short-lived species). We expect that estimates of generation time and variance in reproductive success will be higher in the slow-reproducing *P. filicauda* relative to estimates for *T. aethiops*. The opposite, however, is expected for estimates of effective population size; we expect estimates for *P. filicauda* to be lower than those for *T. aethiops*.

2 | Methods

2.1 | Simulation Framework

To simulate age-structured populations with variation in reproductive success across age groups, we utilised SLiM (v 4.2.1), an individual-based forward-time simulator that allows us to control individual behaviour to explore complex ecological dynamics (Haller and Messer 2019, 2023). A model in SLiM consists of repeated cycles of reproduction and death (Figure 2), where model parameters pre-defined in a custom-made script control how individuals reproduce and die. A general description of the model is given below and the full script is available online (see Data Availability statement).

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College, Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

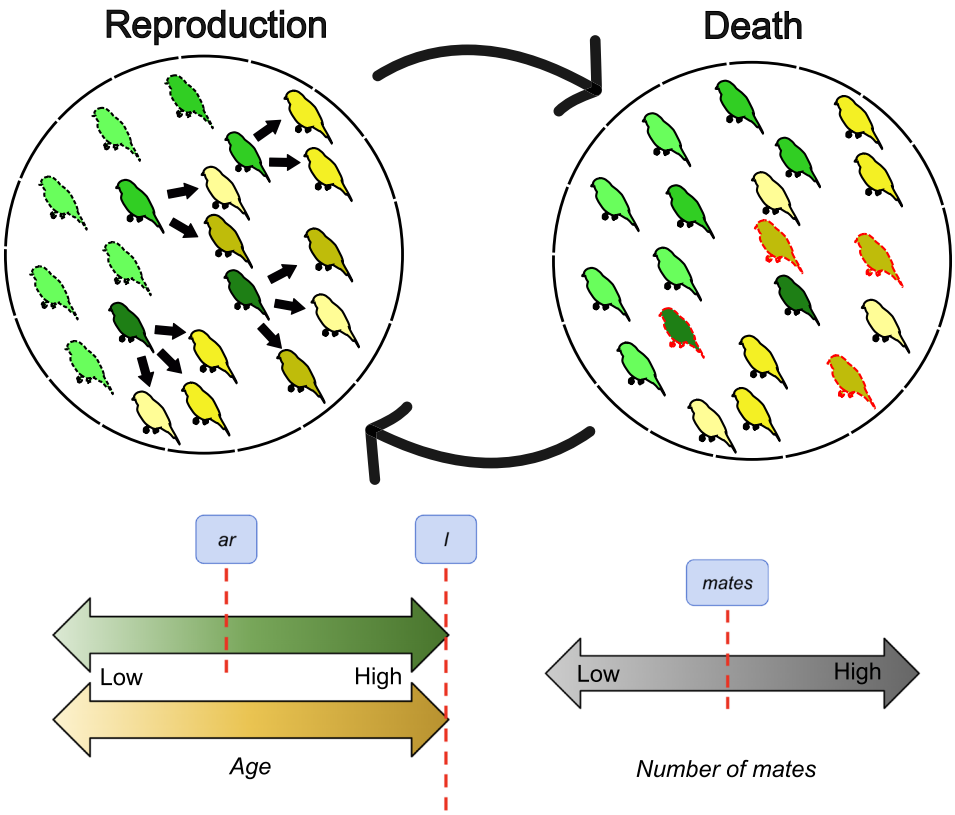

FIGURE 2 | Top: Description of one simulation cycle in the SLiM model, highlighting how the value for each life history trait is incorporated throughout the two phases of the cycle (reproduction and death). Individuals are coloured according to sex: males are green and females are yellow. Colour shades represent different ages: lighter shades indicate younger individuals, whereas darker shades indicate older individuals. Black arrows connecting individuals represent reproductive events. During the reproduction phase, only individuals above a certain age (*ar*) are able to reproduce (reproducing individuals are indicated by solid black contour lines). The number of female mates per male individual is determined by the average number of mates (*mates*). During the death phase, individuals that have reached the maximum average adult longevity (*l*) are removed from the population (indicated by the dashed red contour lines). The cycle is then repeated. Bottom: Double-ended arrows represent the gradient of life history traits that was explored in simulations. Arrows on the left represent the gradient on age-related traits, that is, average age of first reproduction and average adult longevity, whereas the arrow on the right represents the gradient in the average number of mates per male individual. Arrows are coloured according to the effect they have on the simulation: Age arrows follow the colour shading for young and old individuals observed in the simulation diagram at the top, whereas the arrow for average number of mates follow a grey gradient reflecting the number of reproductive events in the top diagram (i.e., black arrows) per male per cycle. Gradient arrows for age are coloured following the colours for each sex in the top diagram (i.e., green for males and yellow for females) to highlight that the model parameters affecting age-dependent rates (*ar* and *l*) impact sexes differently: *ar* impacts males only (red dashed line in green arrow), whereas *l* impacts both males and females (red dashed line in both green and yellow arrows).

The model simulates a population of individuals with varying ages. Three model parameters influence the distribution of ages across individuals (Figure 2): intrinsic fecundity, offspring mortality and adult mortality. At each cycle of the simulation, adult individuals reproduce and the number of offspring individuals generated at each reproduction event is randomly drawn from a Poisson distribution where the mean number of events (λ) equals the intrinsic fecundity. Offspring individuals are then assigned age 0 and, during the same cycle, a proportion of them dies according to the value of offspring mortality. Finally, a proportion of adult individuals dies at the end of the cycle, according to the value of adult mortality. In the model, offspring mortality does not interact with carrying capacity, that is, it is considered as an intrinsic rate of the population and interpreted as the fixed proportion of offspring that dies right after birth, without reproducing. However, adult mortality does interact with the carrying capacity (K) to allow control of the population size over time. We implement a ceiling carrying capacity (Gotelli and Ellison 2004) where the survival probability for each adult individual (derived from the model constant for adult mortality) is multiplied by the ratio between K and the number of individuals in the population above age 0 (N). When $N > K$, the ratio is below 1 and the survival probability is therefore reduced. The fecundity and mortality parameters control the degree of generation overlap in the population; in the simulations, they were kept constant at values that generate a population with a high degree of generation overlap and a uniform age distribution across individuals (see Supporting Information).

Life history traits—that is, the average age of first reproduction, average adult longevity and average number of mates per reproduction—are included in the simulations through model parameters that control when and how individuals reproduce (Figure 2). In this study, we focused on controlling the age of first reproduction and the average number of mates for male

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College , Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

individuals only, since polygyny (i.e., a reproductive system where a single male individual mates with several female individuals) is the most common type of polygamy in nature (Lindström and Kokko 1998). Polygyny leads to a higher variance in the age of first reproduction and the average number of mates across males than across females. In our simulations, at every cycle during reproduction, the average age of first reproduction (*ar*) controls male age-dependent fecundity (i.e., which male individuals are allowed to reproduce) based on their age. This model parameter therefore determines the average youngest age for reproducing males. Adult longevity (*l*) determines the maximum age that all individuals (both males and females) can reach; any individual reaching that threshold is removed from the population during the death cycle. As such, the probability that an individual male will reproduce is zero when its age is below *ar*, and non-zero between ages *ar* and *l* (Figure 2). This probability is further reduced by different levels of polygamy, dictated by a model parameter controlling the average number of female mates per male (*mates*). While in monogamous populations males have a similar probability of reproduction, in polygamous populations each male reproduces with more than one female. Due to the limited number of female individuals, the number of reproducing males is smaller than the total number of males. At every reproduction cycle, a number of potential female mates for each male is defined based on this model parameter, and individuals are randomly matched.

Throughout the simulation, the model records four emergent values: the actual number of mates per male individual, the total number of offspring per male individual, the number of all reproducing adults and the average age across all reproducing adults. These values are used to calculate three population parameters: variance in reproductive success, generation time and effective population size. Reproductive success is defined as the successful mating and generation of offspring; the value of reproductive success for each male individual is therefore calculated as the product between the number of mates and the total number of offspring per male individual. The distribution of individual reproductive success is used to calculate the total variance. To facilitate a biological interpretation of the results, we extracted the square root of that variance to obtain the standard deviation, a metric that is in the same measurement unit of the original values of individual reproductive success. Generation time is directly calculated as the average age of all reproducing individuals at each cycle (Ellner 2018). Finally, the number of reproducing individuals is used to calculate the effective population size, using the formula $N_e = \left(4N_fN_m\right) / \left(N_f N_m\right)$, where N_f and N_m are the number of female and male individuals that effectively reproduce at each cycle, respectively (Wright 1931). This estimation of N_e is independent from genetic data and incorporates solely the number of reproducing individuals per sex, therefore accounting for the variance in reproductive success across different mating systems (Nomura 2002). The final value for each of these three population parameters is obtained as the average of the values across all cycles of the simulation.

Each forward simulation in SLiM consists of a combination of values of life history traits (i.e., average age of first reproduction, average adult longevity and average number of mates per male individual), which yield a set of values for the three population parameters (i.e., variance in reproductive success, generation time and effective population size). For each forward simulation, we utilised the emergent values of population parameters to implement a corresponding coalescent simulation that generates a pattern of genomic variation for that simulated population. Because the rates of coalescence, mutation and recombination are implemented per generation in a coalescent simulation (Battey et al. 2020), we scaled parameters of the coalescent simulation by the generation time emerging from the forward simulation. Specifically, the effective population size used in coalescent simulations was derived from the product between the emergent values of effective population size and the generation time in forward simulations. This multiplication assures that inferred longer generation times are interpreted in coalescent simulations as relatively higher effective population sizes and therefore lower coalescence rates (Haller et al. 2019; Battey et al. 2020). Similarly, the rates of mutation and recombination of coalescent simulations were derived by dividing a pre-defined rate value by the generation time emerging from forward simulations, to represent a decrease in the rates of mutation and recombination over longer generation times (Battey et al. 2020). This scaling allows us to establish a correspondence between the parameter space and the genetic summary statistics explored in the coalescent simulations with the parameter space and the demographic traits explored by the forward simulations.

2.2 | Simulating Variation in Life History Traits

Within the simulation framework described above, we explored the impact of life history traits on population parameters and genetic summary statistics by simulating panmictic populations while varying, across models, the three model parameters corresponding to life history traits: *ar*, *l* and *mates*. We implemented 1000 replicates of forward simulations; at each replicate, values of these three model parameters were randomly sampled from a pre-defined range. The limits of this range were chosen to represent the range of variation in life history traits observed in passerine birds, to facilitate comparison with the empirical data utilised in this study (Bird et al. 2020; Saether 1988). Values for ages at first reproduction ranged from 1 to 7, values for average adult longevity ranged from 8 to 16, and values for average number of mates ranged from 1 to 5. In all simulations, carrying capacity (*K*) was set to 100,000 individuals. Populations were simulated for 100 cycles, and emergent population parameters (variance in reproductive success, generation time and effective population size) were calculated as an average across all cycles, after confirming that the values of these parameters were stable throughout the simulation.

For each of the 1000 populations simulated in SLiM, patterns of genomic variation were then simulated using the software *msprime* (Baumdicker et al. 2022). We simulated a single chromosome with 10 million base pairs (Mb), under a mutation rate of 2.5×10^{-10} substitutions per site per generation and a recombination rate of 1×10^{-8} per site per generation. This mutation rate is a conservative estimation of the average genomic mutation rate in birds (Nadachowska-Brzyska et al. 2015). The recombination rate was kept relatively low to minimise the memory overhead of simulations. Effective population size, mutation rate and recombination were scaled by generation time as described in the previous section. We utilised a sample size of 20

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College, Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

diploid individuals for easier comparison with the empirical data utilised in this study.

Simulated genomes were then exported using the *vcf* format and the following genetic summary statistics were calculated using the software *vcftools* (Danecek et al. 2011): nucleotide diversity (π), Tajima's D, linkage disequilibrium and minor allele frequencies (MAF). Nucleotide diversity, Tajima's D, and linkage disequilibrium were calculated using 10,000 base pairs (10 kb) windows across the simulated genome. Linkage disequilibrium was calculated as the squared correlation coefficient between genotypes within a 10 kb window (Sved 2009). We utilised the mean and the standard deviation of these three metrics calculated across the genome as the summary statistics describing genomic variation at each replicate. We additionally calculated the Site Frequency Spectrum from all variant sites across the simulated genome, utilising the Python script *easySFS* (http://github.com/isaacovercast/easySFS). To allow for comparisons across simulations and with the empirical data, we calculated the relative proportion of variant sites within each bin of the site frequency spectrum by dividing the number of variant sites observed per bin by the total number of variant sites observed in the genome. The calculation of the mean and the standard deviation for each summary statistic and the calculation of the relative proportion of variant sites for the bins of the site frequency spectrum were performed in the R platform (R Core Team 2024).

2.3 | Quantifying Accuracy in Estimating Population Parameters

To investigate the relationship between life history traits, population parameters and genetic summary statistics, we first visually explored the distribution of their values across all simulations and further calculated the correlation between each trait, population parameter and genetic summary statistics using a Spearman correlation. We then explored the relationship between the population parameters and summary statistics to investigate how accurately one may be able to estimate those parameters directly from genetic data. To quantify this accuracy, we used a supervised machine learning regression model based on all simulations previously generated, using the genetic summary statistics as predictor variables, and the three population parameters (variance in reproductive success, generation time and effective population size) as response variables. We trained the machine learning regression model using the Random Forest algorithm (Breiman 2001) implemented in the R package *ranger* (Wright and Ziegler 2015). We estimated the best set of Random Forest parameters (i.e., number of trees, number of variables per tree and minimum node size) using the R package *tune* (Kuhn 2023). We explored different combinations of values for the following parameters: number of trees (from 1 to 2000), number of variables per tree from (1 to 22) and minimum node size (from 2 to 40).

We explored prediction accuracy by generating 1000 replicates of the Random Forest model for each population parameter, resulting in a total of 3000 replicates. For each replicate, 70% of the simulations were retained as a training dataset, on which the model learns how to estimate parameters from the summary statistics, and 30% of the simulations were used as a testing dataset, on which predictions of parameter values were made. Accuracy was evaluated by calculating the coefficient of determination (r^2), which summarises differences between observed and predicted values of the population parameter for each replicate. The final distribution of r^2 values represents the degree of accuracy in the models used to estimate each of the population parameters from genetic summary statistics.

2.4 | Estimating Population Parameters From Empirical Data

To illustrate how simulations of different life history traits can be used to estimate population parameters, we utilised the Random Forest approach described previously to estimate our target population parameters (variance in reproductive success, generation time and effective population size) from empirical genomic data for two passerine birds with contrasting life history strategies: *Pipra filicauda* (a slow-reproducing species) and *Thamnophilus aethiops* (a fast-reproducing species; Figure 3). Males of *P. filicauda* commonly engage in cooperative reproductive displays, where a single dominant (and usually older) male in the cooperative group reproduces with the female visiting the display territory (Heindl 2002). Since different females can visit the same territory during the reproductive season, this reproductive behaviour leads to cases where few older male adults each mate with several female individuals; one study taking place during a reproductive season in eastern Ecuador reported that the three top-ranked males in each of six reproductive leks contributed 70%–90% of the offspring generated in that season (Ryder et al. 2009). This life history strategy results in populations where a small subset of males contributes to the next generation, and many younger male individuals do not reproduce in their first years, leading to high variance in reproductive success across males and a longer generation time. The opposite reproductive behaviour is observed in *T. aethiops*: male and female individuals are known to form monogamous pairs and defend a territory during their reproductive season (Lima et al. 2019). Both sexes are assumed to start reproducing at an early age, and pairs have been observed to remain the same for multiple reproductive seasons (Billerman et al. 2022). This leads to an expectation of lower variance in reproductive success across males and shorter generation times than what would be observed in *P. filicauda*.

To estimate population parameters, we utilised summary genetic statistics calculated from empirical genetic datasets for *P. filicauda* and *T. aethiops* as predictor variables in a Random Forest model previously trained on the simulated dataset. Empirical summary statistics were calculated from two restriction-site associated DNA sequences (RADSeq) datasets retrieved from published phylogeographic studies (Barrera-Guzmán et al. 2022; Musher et al. 2024). We utilised the results from individual admixture analyses presented in these studies to guide our selection of individuals and focused on sampling a single differentiated panmictic lineage within the distribution of each species. For *P. filicauda*, we retrieved sequences from individuals across the entire range of the species, since no signal of intraspecific genetic structure was observed in a previous study (Barrera-Guzmán et al. 2022). For *T. aethiops*, given the intraspecific

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College, Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

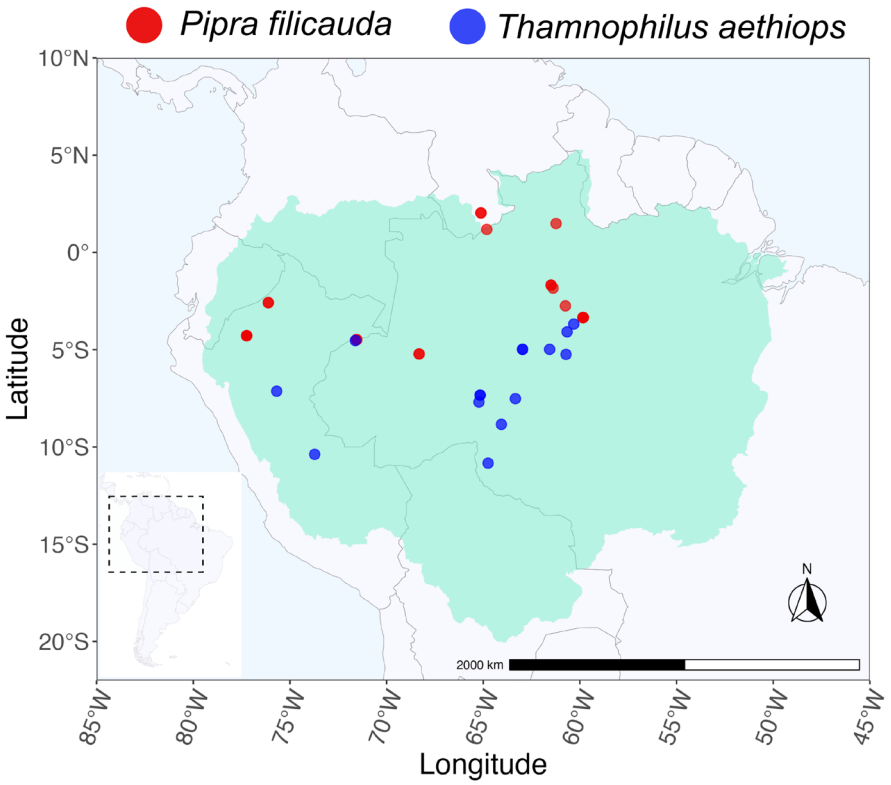

FIGURE 3 | Geographic distribution of the samples utilised for estimates of population parameters from empirical genomic data. Red points represent the localities sampled for *Pipra filicauda*, a slow-reproducing polygamic species, while blue dots represent the localities sampled for *Thamnophilus aethiops*, a fast-reproducing monogamic species (see text for more information on their life history differences).

genetic structure found in a previous study (Musher et al. 2024), we retrieved sequences from solely one panmictic population distributed in the westernmost portion of the species range. This approach minimised the differences between the two datasets by focusing on panmictic populations, allowing us to attribute differences in summary statistics to differences in their life history strategies. The final selected dataset consisted of 28 individuals across 11 localities for *P. filicauda*, and 15 individuals across 13 localities for *T. aethiops* (Figure 3; Table S1).

Raw sequences were downloaded and mapped into a reference genome utilising the iPyrad pipeline (Eaton and Overcast 2020). We utilised reference genomes available for species that are phylogenetically close to the focal species: the sequenced genome of *Lepidothrix coronata* (GenBank Assembly: GCA_001604755.1) was utilised for *P. filicauda*, and the sequenced genome of *Sakesphorus luctuosus* (GenBank Assembly: GCA_013396695.1) was utilised for *T. aethiops*. The parameters utilised in the assembly pipeline were kept at their default values, except for the clustering threshold and the percentage of missing data. We tested different values of clustering threshold (0.85, 0.9 and 0.95) to explore the impact of this parameter on the final dataset. We found no impact of this parameter on reference mapping and therefore set a strict value of 0.95. Additionally, we implemented a strict filter that allowed for no missing data, forcing all variant sites to be present in all individuals, to allow comparison of the empirical data with the simulated dataset on which the Random Forest model was created. We retrieved the *vcf* file exported by the assembly and utilised the approach described above to calculate, for the empirical dataset, the same genetic summary statistics used to train the Random Forest model. Summary statistics were calculated across all samples since previous studies confirm they belong to a single panmictic population. To allow for comparisons with the simulated dataset, we projected the sample size of the Site Frequency Spectrum calculated for each species to 20 diploid individuals, following Marth et al. (2004). Empirical summary statistics were then utilised as predictors in the Random Forest model to estimate the values for variance in reproductive success, generation time and effective population size for both species. A confidence interval around each value was calculated by estimating the prediction standard error using the jackknife procedure described by (Wager et al. 2014). That procedure estimates the standard error associated with a prediction based on the bootstrap replicates that are performed during the building of the Random Forest model (Efron 1992; Wager et al. 2014). We then compared the predicted values and their respective confidence intervals between both species to verify whether they suggested that the slow-reproducing long-lived *P. filicauda* has higher variance of reproductive success, longer generation time and smaller effective population size relative to the fast-reproducing short-lived *T. aethiops*—hence supporting our expectations given their natural history observations.

3 | Results

Forward simulations on SLiM revealed that population parameters (variance in reproductive success, generation time and effective population size) were sensitive to variation in life history traits (Figure 4; Table 1). The average age of first reproduction

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College , Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

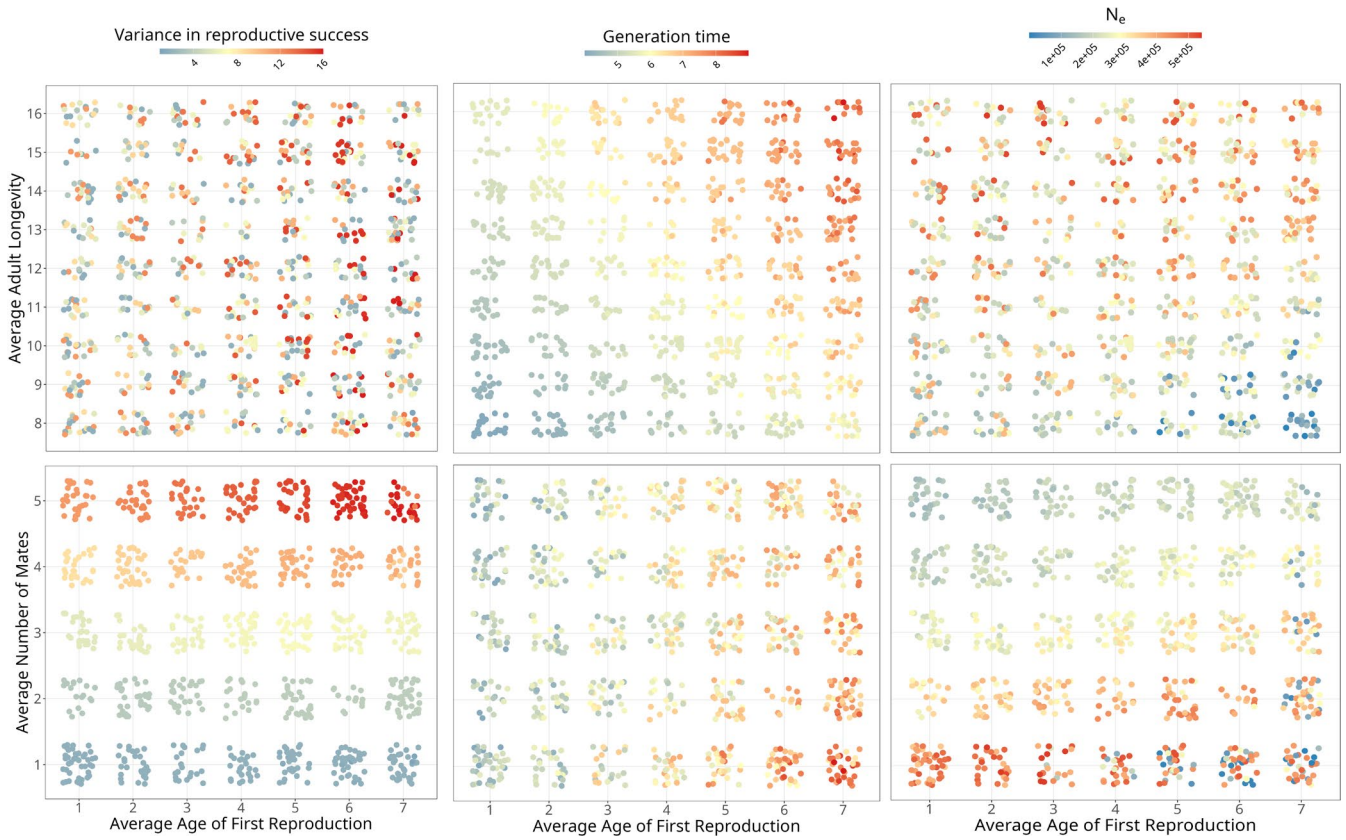

FIGURE 4 | Variation in values of population parameters (left column: Variance in reproductive success; central column: Generation time; right column: Effective population size) emerging from simulations varying life history traits (top row: combinations of age of first reproduction and average adult longevity; bottom row: combinations of average age of first reproduction and average number of mates).

TABLE 1 | Coefficient values of Spearman correlations between life history traits (columns) and population parameters (rows).

	Average age at first reproduction	Average adult longevity	Average number of mates per reproducing male individual
Variance in reproductive success	0.1362	0.0903	**0.9792**
Generation time	**0.7582**	**0.6136**	−0.0772
Effective population size	−0.2791	0.2649	**−0.6809**

Note: Coefficient values below −0.5 and above 0.5 are highlighted in bold.

showed a weak positive correlation with variance in reproductive success ($\rho = 0.136$) and a weak negative correlation with effective population size ($\rho = -0.279$), but a strong positive correlation with generation time ($\rho = 0.758$; Figure 4). Average adult longevity showed a weak positive correlation with both variance in reproductive success ($\rho = 0.09$) and effective population size ($\rho = 0.265$), but a strong positive correlation with generation time ($\rho = 0.614$; Figure 4, top row). The average number of mates showed a strong positive correlation with variance in reproductive success ($\rho = 0.979$) and a strong negative correlation with effective population size ($\rho = -0.681$), but a weak negative correlation with generation time ($\rho = -0.077$; Figure 4, bottom row).

Varying life history traits also impacted our target genetic summary statistics (Figure 5 and Figures S1–S3; Table 2), although the sensitivity of the summary statistics to changes in life history traits was more varied. Average nucleotide diversity showed a weak negative correlation with average age of first reproduction ($\rho = -0.286$) and a weak positive correlation with average adult longevity ($\rho = 0.175$), but a strong negative correlation with average number of mates ($\rho = -0.679$; Figure 5, left column). Similar correlation values were found for measurements of average linkage disequilibrium but in the opposite direction: average linkage disequilibrium showed a weak positive correlation with average age of first reproduction ($\rho = 0.251$) and a weak negative correlation with average adult longevity ($\rho = -0.099$), but a strong positive correlation with the average number of mates ($\rho = 0.676$; Figure 5, central column). Correlation of life history traits with average Tajima's D, average minor allele frequencies and the frequency classes of the site frequency spectrum were overall low, ranging between −0.06 and 0.062 (Figure 5, right column, for average minor allele frequencies; Figure S1 left for Tajima's D; Figure S2 for the site frequency spectrum). Finally, the average number of mates showed a moderate negative correlation with

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College, Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

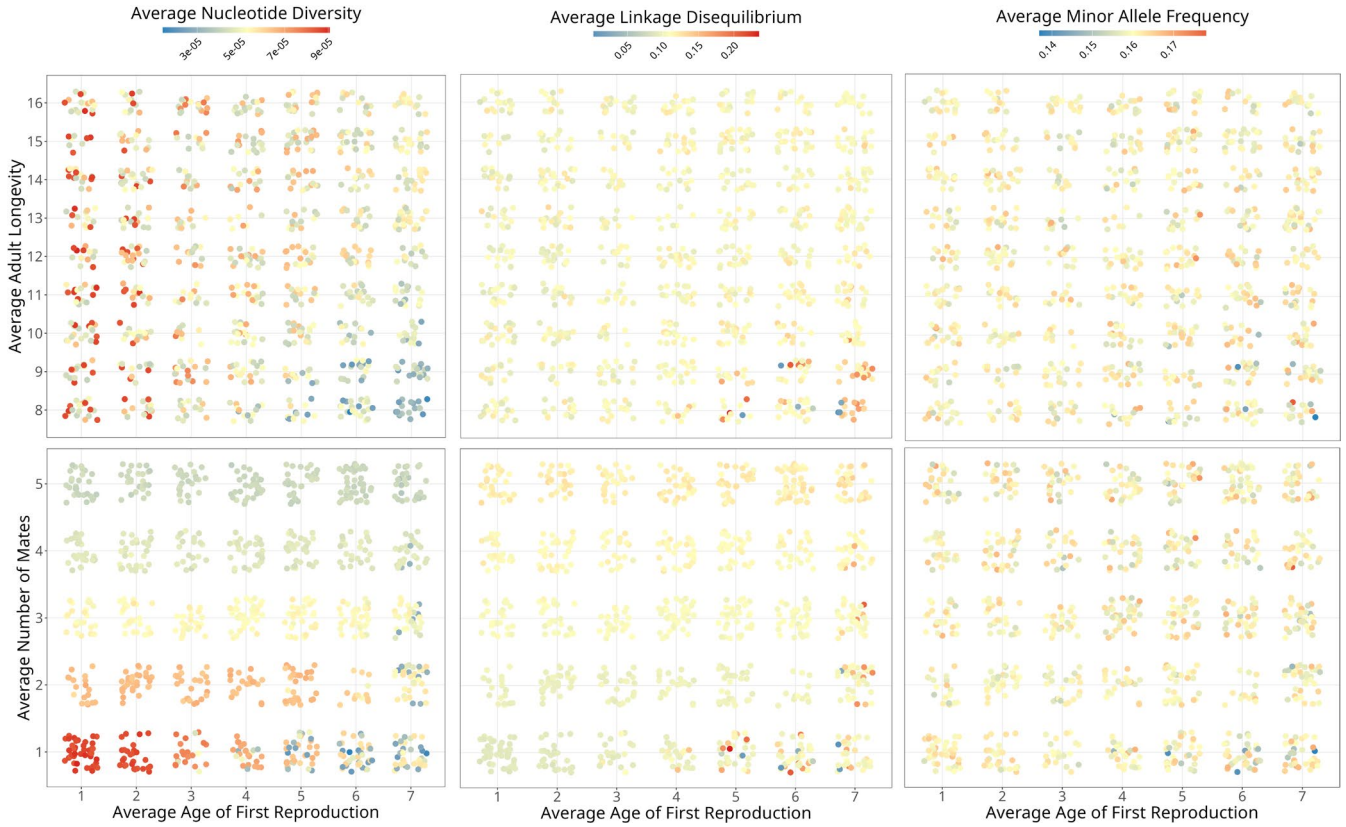

FIGURE 5 | Variation in the values of three genetic summary statistics (left column: average nucleotide diversity; central column: average linkage disequilibrium; right column: average minor allele frequencies) across different values of life history traits (top row: combinations of age of first reproduction and average adult longevity; bottom row: combinations of average age of first reproduction and average number of mates).

TABLE 2 | Coefficient values of Spearman correlations between life history traits (columns) and genetic summary statistics (rows).

	Average age at first reproduction	Average adult longevity	Average number of mates per reproducing male individual
Average nucleotide diversity	−0.286	0.1753	**−0.6788**
Standard deviation of nucleotide diversity	−0.2979	0.1651	**−0.6675**
Average minor allele frequency	−0.0165	0.007	0.0044
Standard deviation of minor allele frequency	0.0143	0.0102	−0.0137
Average Tajima's D	−0.0018	0.0075	0.0086
Standard deviation of Tajima's D	0.2779	−0.1495	**0.5281**
Mean linkage disequilibrium	0.2509	−0.0989	**0.6758**
Standard deviation of linkage disequilibrium	0.2563	−0.0963	**0.6703**
Site Frequency Spectrum	−0.0075	−0.0238	−0.005

Note: Coefficient values below −0.5 and above 0.5 are highlighted in bold. Values for the site frequency spectrum represent the average correlation coefficient across all frequency classes.

the standard deviation of nucleotide diversity across the genome ($\rho = -0.668$; Figure S3, left) and a moderate positive correlation with the standard deviation of both Tajima's D ($\rho = 0.528$; Figure S2, right) and measurements of linkage disequilibrium ($\rho = 0.67$; Figure S3, right) across the genome.

Random forest models of ecological population parameters, built with simulated genetic summary statistics as predictor variables, showed high accuracy in estimating variance in reproductive success (mean $R^2 = 0.8098 \pm 0.0405$) and effective population size (mean $R^2 = 0.8155 \pm 0.161$), and low accuracy

13652294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College, Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

in estimating generation time (mean $R^2 = 0.1713 \pm 0.0457$). However, parameter values estimated from a Random Forest model trained on the entire simulated dataset, and utilising empirical genetic data from our target bird species as predictor variables, show little overlap in the confidence interval of estimated values and agree with our original expectations: estimated values of the variance in reproductive success and generation time were higher for *P. filicauda* than *T. aethiops*, whereas estimated values of effective population size were lower for *P. filicauda* than for *T. aethiops* (Figure 6).

4 | Discussion

Our simulations showed that interspecific differences in measurable life history traits (average age of first reproduction, average adult longevity and average number of mates) do impact population parameters (variance in reproductive success, generation time and effective population size). Variation in life history traits also impacts genetic summary statistics with varying levels of correlation depending on the metric of choice. Our simulations support previous expectations that populations characterised by higher values of life history traits tend to exhibit higher variance in reproductive success, longer generation times and smaller effective population sizes (Figure 4 and Table 1; Waples 2023; Waples et al. 2013). However, the results from our simulations suggest that age-related traits (i.e., average age of first reproduction and average adult longevity) have a stronger impact on generation time than variance in reproductive success and effective population size (Figure 4, top row, central column; Table 1, central row). The opposite is observed for the average number of mates: although we found the expected influence of this life history trait on variance in reproductive success and effective population size (Figure 4, bottom row, left and right columns; Table 1, top and bottom rows), we found it to be weakly correlated with generation time (Figure 4, bottom row, central column; Table 1, central row), in contrast to suggestions from previous studies (Nunney 1991).

Through our simulations, we additionally observe that the average nucleotide diversity and estimates of linkage disequilibrium are influenced by life history traits, especially the average number of mates, which shows the highest correlation coefficients with those two summary statistics (Figure 5, bottom row; Table 2). Additionally, the standard deviation of three summary statistics (i.e., nucleotide diversity, Tajima's D and estimates of linkage disequilibrium) was highly correlated with the average number of mates (Figures S1 and S3; Table 2). Collectively, these results suggest that (1) mating systems (represented by the average number of mates) have the strongest impact on summaries of patterns of genomic variation; (2) the signal of life history variation is pervasive across different metrics that summarise genomic variation. While recent studies have shown that measurements of genetic diversity correlate with different life history strategies in both empirical and simulated datasets (Barry et al. 2022; Waples 2022a), in this study, we were able to look beyond nucleotide diversity and show that the signal of life history variation is additionally present across measurements of minor allele frequencies, linkage disequilibrium and traditional population genetic statistics such as Tajima's D. Following expectations from the literature, we found that polygamous populations have lower genetic diversity and higher levels of linkage disequilibrium due to higher variance in reproductive success leading to smaller population sizes. Surprisingly, age-related traits (which are the ones largely influencing generation time; Figure 4, central column; Table 1) have a relatively weaker impact on almost all genetic summary statistics: their influence is mostly observed in simulations where male individuals start reproducing late and die young (Figure 5, top row, central and right column; Table 2), that is, a combination of life history traits that lead to a largely reduced number of reproducing male individuals. An exception to this pattern is observed through the effect of age-related traits on average nucleotide diversity, which tends to decrease as the average age of first reproduction increases (Figure 5, top row, left column; Table 2). This reduced effect of age-related traits on genetic summary statistics is reflected in the overall lower predictive accuracy observed in Random Forest models when using simulated summary statistics to estimate generation time (mean $R^2 = 0.17$) than that observed for estimates of variance in reproductive success and effective population size (mean $R^2 = 0.8$). The lower predictive accuracy reflects the low correlation, observed within the simulated dataset, between age-related traits and genetic summary statistics. However, when utilising the same Random Forest model to estimate population parameters of two bird species with contrasting life histories, our estimates agree with theoretical expectations for all three parameters: estimates of variance in reproductive success and generation time were higher for *P. filicauda* than the estimates of these same parameters for *T. aethiops*, while

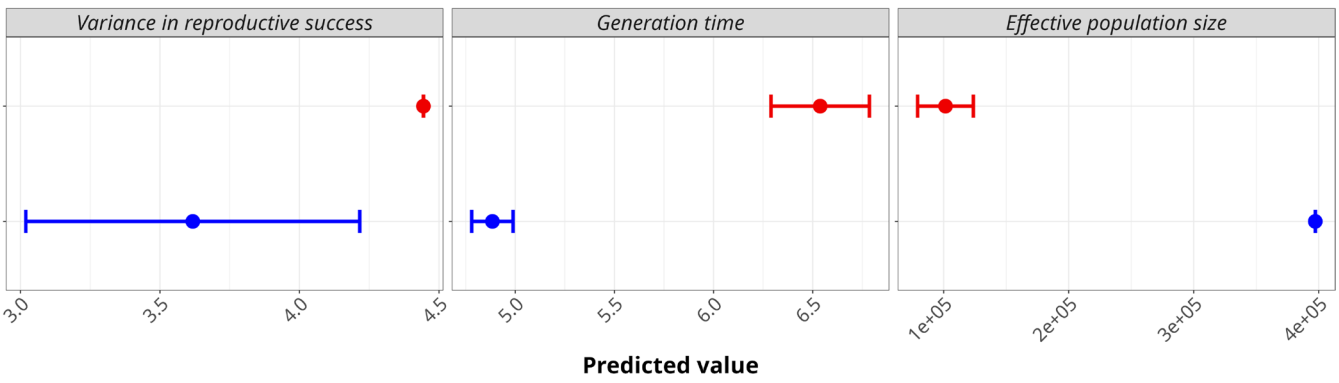

FIGURE 6 | Estimates and confidence intervals for variance in reproductive success (left), generation time (centre) and effective population size (right) for *Pipra filicauda* (red) and *Thamnophilus aethiops* (blue), based on summary statistic values calculated from empirical data.

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College, Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the opposite was observed for the effective population size (Figure 6). These results suggest that the relative difference between the two empirical genomic datasets follows the pattern of genomic variation found in our simulations; that is, lower genetic diversity is observed in both the simulated populations exhibiting higher values of life history traits as well as the empirical populations in species for which those same values were shown to be comparatively higher (in our case, populations of *P. filicauda*). Although the panmictic populations analysed here occur in different parts of western Amazon (i.e., samples of *P. filicauda* occur at lower latitudes than the samples of *T. aethiops*; Figure 3), both species inhabit similar habitat types in this region, consisting of secondary growth evergreen forest and often associated with forest gaps and streams (Snow 2020; Zimmer and Isler 2020). This suggests that historical changes in the availability of suitable habitat would have been similar for both species, and any relative differences in effective population size are likely to be the result of their differences in life history strategies. Despite the low correlation of life history traits with some of the summary statistics, our results indicate that a Random Forest model is able to combine the small effect of life history traits across all summary statistics to provide relative parameter estimates with enough resolution to distinguish between these different life history strategies.

Our findings highlight the importance of incorporating life history variation into simulation-based models for demographic inference across an assemblage of species with a shared climatic history. In any local community, genetic patterns across species will differ due to both stochastic and deterministic factors (Papadopoulou and Knowles 2016). Our power in identifying shared responses and accurately estimating demographic parameters in the presence of different genetic patterns depends on the extent to which our simulations incorporate the deterministic factors that have led to differences across species. If variation across species in life history strategies alone leads to measurable differences in genetic patterns, our findings suggest that, whenever life history variation is not incorporated in the simulation approach, differences in summary statistics that stem from it will be treated as unexplained variation. This may only be ideal when one is mostly interested in the signal of shared demography response to a historical event instead of the signal of ecological differences, in which cases these differences can effectively be treated as nuisance parameters (Prates et al. 2016; Xue and Hickerson 2020). When ecological differences are large, or when the goal is to specifically capture any (even if weak) signal of ecological differences, explicitly incorporating such differences may increase the power to infer historical demographic parameters (Massatti and Knowles 2014; Papadopoulou and Knowles 2015). Understanding the best simulation approach for a specific group therefore requires stating what ecological mechanisms, if any, may lead to predictable differences in genetic patterns.

Our study focused on one of those mechanisms: how differences in life history traits lead to consistent and predictable differences in genetic patterns through their impact on key population parameters: variance in reproductive success, generation time and effective population size. Life history traits impact variance in reproductive success and generation time through their influence on the distribution of reproductive success across individuals based on age and sex. Since variance in reproductive success, and to a lesser extent generation time, is negatively correlated with the number of reproducing individuals (Figure 4), life history traits ultimately influence the fundamental population genetic parameter θ ($=4N_e\mu$) through their influence on the effective population size (N_e). As such, variance in reproductive success and generation time are two key parameters to describe the deterministic way in which life history influences patterns of genetic diversity. Although many theoretical and simulation studies have highlighted the link between these parameters and effective population size (Waples et al. 2013; Waples and Yokota 2007), few empirical studies performing demographic inference utilise them to describe differences across species when defining demographic hypotheses, and/or explicitly incorporate them in simulations.

A remaining challenge to understand how life history variation across species impacts inference of comparative historical demography is investigating at what temporal and spatial scales the signal of life history variation is expected to be most important, and at what scales it is expected to be surpassed by the signal of other ecological factors, such as species-specific dispersal capacities or environmental tolerances (Papadopoulou and Knowles 2016). For instance, we observe that a slow-reproducing polygamic life history strategy leads to lower effective population size; however, if the same species also exhibits low intrinsic dispersal capacity, high genetic differentiation would be observed among localities within the species' range. In this scenario, inferred effective population size would be small when estimated at local spatial scales, with low genetic differentiation, but higher if estimated across larger spatial scales. Because the influence of space is ubiquitous in spatial patterns of genomic variation (Bradburd and Ralph 2019), future studies investigating the relevance of life history variation for historical demographic inference will benefit from exploring those differences under spatially explicit simulation approaches, where both variation in life history and variation in landscape connectivity can be investigated.

We understand that several methodological approaches exist to implement life history variation in the simulation-based inference framework, many of those based on classic coalescent methods. Under the coalescent, one can incorporate variation in effective population size, migration rates or generation time by integrating across a wide prior parameter space (Xue and Hickerson 2020). Other studies have relaxed the coalescence assumption that all individuals have equal reproductive success by explicitly stating a vector of relative probabilities of reproduction for the parental generation in a Wright-Fisher model (Waples 2022a). Extensions of coalescent methods have also been developed to account for life history complexity, such as the skewed distribution of offspring size across individuals, and other forms of complex demographic structure, using multiple merger coalescence (Lepers et al. 2021; Tellier and Lemaire 2014). Here, we followed a non-coalescence approach to estimate the parameters impacting effective population size directly from phenotypic ecological information (i.e., life history traits that are measurable at the population level). This approach differs from Wright-Fisher models in that the effective population size (along with other focal population parameters) emerges from the interaction of the different life history traits

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College , Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and informs the exploration of parameter space in coalescent simulations. This allowed us to explore the relationship between life history traits and effective population size in coalescent simulations, as well as map patterns of genomic variation arising from these simulations to parameter values for generation time and variance in reproductive success that are not explicitly incorporated in coalescent simulations. Because we implement life history traits that are measurable at the population level as values in our forward simulation, this approach additionally allowed us to incorporate available information on the life history of species as input in our models, making it similar to other measurable and available phenotypic traits such as body size or diet. Even though life history traits may not be as easily measurable in nature as other phenotypic traits, an attempt to directly incorporate them into simulations can bring important insights into how effective population size scales with life history, and therefore, how it can be better interpreted or modelled in traditional coalescent approaches.

Author Contributions

All authors designed the research. R.M. performed genetic simulations and analysed simulated and empirical data. All authors wrote the paper.

Acknowledgements

R.M. acknowledges the research group at The City College of New York and Prof. Matthew Aiello-Lammens, Prof. Robert P. Anderson and Prof. Brian T. Smith for providing valuable feedback and guidance during the design and writing of this project. R.M., M.J.H. and A.C.C. acknowledge funding by NSF award DBI 1926928 (IIBR RoL: Collaborative Research: A Rules of Life Engine (RoLE) Model to Uncover Fundamental Processes Governing Biodiversity), which partly funded Mascarenhas as he developed this work. R.M. also acknowledges support from the Graduate Center and The City College of the City University of New York.

Disclosure

Benefit-Sharing Statement: Benefits from this research stem from the sharing of detailed scripts and datasets on public databases (as described above) in order to increase reproducibility.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All scripts, results and empirical datasets utilised in this manuscript are available on https://github.com/rilquer-mascarenhas/lifehistory-models-files.

References

Barrera-Guzmán, A. O., A. Aleixo, M. Faccio, S. d. M. Dantas, and J. T. Weir. 2022. "Gene Flow, Genomic Homogenization and the Timeline to Speciation in Amazonian Manakins." *Molecular Ecology* 31, no. 15: 4050–4066.

Barry, P., T. Broquet, and P.-A. Gagnaire. 2022. "Age-Specific Survivorship and Fecundity Shape Genetic Diversity in Marine Fishes." *Evolution Letters* 6, no. 1: 46–62.

Battey, C. J., P. L. Ralph, and A. D. Kern. 2020. "Space is the Place: Effects of Continuous Spatial Structure on Analysis of Population Genetic Data." *Genetics* 215, no. 1: 193–214.

Baumdicker, F., G. Bisschop, D. Goldstein, et al. 2022. "Efficient Ancestry and Mutation Simulation With Msprime 1.0." *Genetics* 220, no. 3: iyab229.

Billerman, S. M., B. K. Keeney, P. G. Rodewald, and T. S. Schulenberg. 2022. *Birds of the World*. Cornell Laboratory of Ornithology.

Bird, J. P., R. Martin, H. R. Akçakaya, et al. 2020. "Generation Lengths of the World's Birds and Their Implications for Extinction Risk." *Conservation Biology* 34, no. 5: 1252–1261.

Blischak, P. D., M. S. Barker, and R. N. Gutenkunst. 2020. "Inferring the Demographic History of Inbred Species From Genome-Wide SNP Frequency Data." *Molecular Biology and Evolution* 37, no. 7: 2124–2136.

Bradburd, G. S., and P. L. Ralph. 2019. "Spatial Population Genetics: It's About Time." *Annual Review of Ecology, Evolution, and Systematics* 50: 427–449. https://doi.org/10.1146/annurev-ecolsys-110316-022659.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45, no. 1: 5–32.

Burney, C. W., and R. T. Brumfield. 2009. "Ecology Predicts Levels of Genetic Differentiation in Neotropical Birds." *American Naturalist* 174, no. 3: 358–368.

Carvalho, A. F., R. S. T. Menezes, E. A. Miranda, M. A. Costa, and M. A. Del Lama. 2021. "Comparative Phylogeography and Palaeomodelling Reveal Idiosyncratic Responses to Climate Changes in Neotropical Paper Wasps." *Biological Journal of the Linnean Society* 132, no. 4: 955–969.

Collin, F.-D., G. Durif, L. Raynal, et al. 2021. "Extending Approximate Bayesian Computation With Supervised Machine Learning to Infer Demographic History From Genetic Polymorphisms Using DIYABC Random Forest." *Molecular Ecology Resources* 21, no. 8: 2598–2613.

Comte, L., R. Bertrand, S. Diamond, et al. 2024. "Bringing Traits Back Into the Equation: A Roadmap to Understand Species Redistribution." *Global Change Biology* 30, no. 4: e17271.

Coulson, T., S. Tuljapurkar, and D. Z. Childs. 2010. "Using Evolutionary Demography to Link Life History Theory, Quantitative Genetics and Population Ecology." *Journal of Animal Ecology* 79, no. 6: 1226–1240.

Crow, J. F., and C. Denniston. 1988. "Inbreeding and Variance Effective Population Numbers." *Evolution* 42, no. 3: 482–495.

Csilléry, K., M. G. B. Blum, O. E. Gaggiotti, and O. François. 2010. "Approximate Bayesian Computation (ABC) in Practice." *Trends in Ecology & Evolution* 25, no. 7: 410–418.

Danecek, P., A. Auton, G. Abecasis, et al. 2011. "The Variant Call Format and VCFtools." *Bioinformatics* 27, no. 15: 2156–2158.

Dubuc, C., A. Ruiz-Lambides, and A. Widdig. 2014. "Variance in Male Lifetime Reproductive Success and Estimation of the Degree of Polygyny in a Primate." *Behavioral Ecology* 25, no. 4: 878–889.

Eaton, D. A. R., and I. Overcast. 2020. "Ipyrad: Interactive Assembly and Analysis of RADseq Datasets." *Bioinformatics* 36, no. 8: 2592–2594.

Efron, B. 1992. "Jackknife-After-Bootstrap Standard Errors and Influence Functions." *Journal of the Royal Statistical Society. Series B, Statistical Methodology* 54, no. 1: 83–111.

Ellner, S. P. 2018. "Generation Time in Structured Populations." *American Naturalist* 192, no. 1: 105–110.

Fan, D., W. Hu, B. Li, et al. 2016. "Idiosyncratic Responses of Evergreen Broad-Leaved Forest Constituents in China to the Late Quaternary Climate Changes." *Scientific Reports* 6, no. 1: 31044.

Fisher, R. A. 1922. "XXI—On the Dominance Ratio." *Proceedings of the Royal Society of Edinburgh* 42: 321–341.

Fonseca, E. M., G. R. Colli, F. P. Werneck, and B. C. Carstens. 2021. "Phylogeographic Model Selection Using Convolutional Neural Networks." *Molecular Ecology Resources* 21, no. 8: 2661–2675.

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College, Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Gaiotti, M. G., M. S. Webster, and R. H. Macedo. 2020. "An Atypical Mating System in a Neotropical Manakin." *Royal Society Open Science* 7, no. 1: 191548.

García-Rodríguez, A., C. E. Guarnizo, A. J. Crawford, A. A. Garda, and G. C. Costa. 2020. "Idiosyncratic Responses to Drivers of Genetic Differentiation in the Complex Landscapes of Isthmian Central America." *Heredity* 126: 251–265. https://doi.org/10.1038/s41437-020-00376-8.

Germain, R. R., S. Feng, G. Chen, et al. 2023. "Species-Specific Traits Mediate Avian Demographic Responses Under Past Climate Change." *Nature Ecology & Evolution* 7: 862–872. https://doi.org/10.1038/s41559-023-02055-3.

Gotelli, N. J., and A. M. Ellison. 2004. *A Primer of Ecological Statistics*. Sinauer Associates.

Hahn, M. W. 2018. *Molecular Population Genetics*. Oxford University Press.

Haller, B. C., J. Galloway, J. Kelleher, P. W. Messer, and P. L. Ralph. 2019. "Tree-Sequence Recording in SLiM Opens New Horizons for Forward-Time Simulation of Whole Genomes." *Molecular Ecology Resources* 19, no. 2: 552–566.

Haller, B. C., and P. W. Messer. 2019. "SLiM 3: Forward Genetic Simulations Beyond the Wright–Fisher Model." *Molecular Biology and Evolution* 36, no. 3: 632–637.

Haller, B. C., and P. W. Messer. 2023. "SLiM 4: Multispecies Eco-Evolutionary Modeling." *American Naturalist* 201, no. 5: E127–E139.

Hasselquist, D. 1998. "Polygyny in Great Reed Warblers: A Long-Term Study of Factors Contributing to Male Fitness." *Ecology* 79, no. 7: 2376–2390.

Heindl, M. 2002. "Social Organization on Leks of the Wire-Tailed Manakin in Southern Venezuela." *Condor* 104, no. 4: 772–779.

Hickerson, M. J., B. C. Carstens, and J. Cavender-Bares. 2010. "Phylogeography's Past, Present, and Future: 10 Years After." *Molecular Phylogenetics and Evolution* 54, no. 1: 291–301.

Hickerson, M. J., G. Dolman, and C. Moritz. 2006. "Comparative Phylogeographic Summary Statistics for Testing Simultaneous Vicariance." *Molecular Ecology* 15, no. 1: 209–223.

Ianni-Ravn, M. K., M. Petr, and F. Racimo. 2024. "Exploring the Effects of Ecological Parameters on the Spatial Structure of Genetic Tree Sequences." *Peer Community Journal* 4, no. e75: 439. https://doi.org/10.24072/pcjournal.439.

Kingman, J. F. C. 1982. "On the Genealogy of Large Populations." *Journal of Applied Probability* 19: 27–43.

Knowles, L. L. 2009. "Statistical Phylogeography." *Annual Review of Ecology, Evolution, and Systematics* 40, no. 1: 593–612.

Knowles, L. L., and D. F. Alvarado-Serrano. 2010. "Exploring the Population Genetic Consequences of the Colonization Process With Spatio-Temporally Explicit Models: Insights From Coupled Ecological, Demographic and Genetic Models in Montane Grasshoppers." *Molecular Ecology* 19, no. 17: 3727–3745.

Knowles, L. L., and W. P. Maddison. 2002. "Statistical Phylogeography." *Molecular Ecology* 11, no. 12: 2623–2635.

Kuhn, M. 2023. "Tune: Tidy Tuning Tools." https://tune.tidymodels.org/.

Lee, A. M., B.-E. Saether, and S. Engen. 2011. "Demographic Stochasticity, Allee Effects, and Extinction: The Influence of Mating System and Sex Ratio." *American Naturalist* 177, no. 3: 301–313.

Lepers, C., S. Billiard, M. Porte, S. Méléard, and V. C. Tran. 2021. "Inference With Selection, Varying Population Size, and Evolving Population Structure: Application of ABC to a Forward-Backward Coalescent Process With Interactions." *Heredity* 126, no. 2: 335–350.

Lima, J. M., D. P. Guimarães, and E. Guilherme. 2019. "Notes on Bird Breeding Activity in a Lowland Forest in South-West Brazilian Amazonia." *Bulletin of the British Ornithologists' Club* 139, no. 4: 338.

Lindström, J., and H. Kokko. 1998. "Sexual Reproduction and Population Dynamics: The Role of Polygyny and Demographic Sex Differences." *Proceedings. Biological sciences* 265, no. 1395: 483–488.

Marth, G. T., E. Czabarka, J. Murvai, and S. T. Sherry. 2004. "The Allele Frequency Spectrum in Genome-Wide Human Variation Data Reveals Signals of Differential Demographic History in Three Large World Populations." *Genetics* 166, no. 1: 351–372.

Massatti, R., and L. L. Knowles. 2014. "Microhabitat Differences Impact Phylogeographic Concordance of Codistributed Species: Genomic Evidence in Montane Sedges (*Carex* L.) From the Rocky Mountains." *Evolution; International Journal of Organic Evolution* 68, no. 10: 2833–2846.

McDonald, D. B. 1993. "Demographic Consequences of Sexual Selection in the Long-Tailed Manakin." *Behavioral Ecology* 4, no. 4: 297–309.

Miller, M. J., E. Bermingham, B. L. Turner, J. C. Touchon, A. B. Johnson, and K. Winker. 2021. "Demographic Consequences of Foraging Ecology Explain Genetic Diversification in Neotropical Bird Species." *Ecology Letters* 24: 13674. https://doi.org/10.1111/ele.13674.

Musher, L. J., G. Del-Rio, R. S. Marcondes, R. T. Brumfield, G. A. Bravo, and G. Thom. 2024. "Geogenomic Predictors of Genetree Heterogeneity Explain Phylogeographic and Introgression History: A Case Study in an Amazonian Bird (*Thamnophilus aethiops*)." *Systematic Biology* 73, no. 1: 36–52.

Nadachowska-Brzyska, K., C. Li, L. Smeds, G. Zhang, and H. Ellegren. 2015. "Temporal Dynamics of Avian Populations During Pleistocene Revealed by Whole-Genome Sequences." *Current Biology* 25, no. 10: 1375–1380.

Nomura, T. 2002. "Effective Size of Populations With Unequal Sex Ratio and Variation in Mating Success." *Journal of Animal Breeding and Genetics* 119, no. 5: 297–310.

Nunney, L. 1991. "The Influence of Age Structure and Fecundity on Effective Population Size." *Proceedings Biological Sciences/the Royal Society* 246, no. 1315: 71–76.

Nunney, L. 1993. "The Influence of Mating System and Overlapping Generations on Effective Population Size." *Evolution; International Journal of Organic Evolution* 47, no. 5: 1329–1341.

Pabijan, M., K. C. Wollenberg, and M. Vences. 2012. "Small Body Size Increases the Regional Differentiation of Populations of Tropical Mantellid Frogs (Anura: Mantellidae)." *Journal of Evolutionary Biology* 25, no. 11: 2310–2324.

Papadopoulou, A., and L. L. Knowles. 2015. "Species-Specific Responses to Island Connectivity Cycles: Refined Models for Testing Phylogeographic Concordance Across a Mediterranean Pleistocene Aggregate Island Complex." *Molecular Ecology* 24, no. 16: 4252–4268.

Papadopoulou, A., and L. L. Knowles. 2016. "Toward a Paradigm Shift in Comparative Phylogeography Driven by Trait-Based Hypotheses." *Proceedings of the National Academy of Sciences of the United States of America* 113, no. 29: 8018–8024.

Paz, A., R. Ibáñez, K. R. Lips, and A. J. Crawford. 2015. "Testing the Role of Ecology and Life History in Structuring Genetic Variation Across a Landscape: A Trait-Based Phylogeographic Approach." *Molecular Ecology* 24, no. 14: 3723–3737.

Prates, I., A. T. Xue, J. L. Brown, et al. 2016. "Inferring Responses to Climate Dynamics From Historical Demography in Neotropical Forest Lizards." *Proceedings of the National Academy of Sciences of the United States of America* 113, no. 29: 7978–7985.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College, Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Ryder, T. B., P. G. Parker, J. G. Blake, and B. A. Loiselle. 2009. "It Takes Two to Tango: Reproductive Skew and Social Correlates of Male Mating Success in a Lek-Breeding Bird." *Proceedings of the Royal Society B: Biological Sciences* 276, no. 1666: 2377–2384.

Sackman, A. M., R. B. Harris, and J. D. Jensen. 2019. "Inferring Demography and Selection in Organisms Characterized by Skewed Offspring Distributions." *Genetics* 211, no. 3: 1019–1028.

Saether, B. E. 1988. "Pattern of Covariation Between Life-History Traits of European Birds." *Nature* 331, no. 6157: 616–617.

Saether, B.-E., T. Coulson, V. Grøtan, et al. 2013. "How Life History Influences Population Dynamics in Fluctuating Environments." *American Naturalist* 182, no. 6: 743–759.

Sæther, B.-E., S. Engen, A. Pape Møller, et al. 2004. "Life-History Variation Predicts the Effects of Demographic Stochasticity on Avian Population Dynamics." *American Naturalist* 164, no. 6: 793–802.

Schrider, D. R., and A. D. Kern. 2018. "Supervised Machine Learning for Population Genetics: A New Paradigm." *Trends in Genetics* 34, no. 4: 301–312.

Snow, D. 2020. "Wire-Tailed Manakin (*Pipra filicauda*), Version 1.0." In *Birds of the World*, edited by J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana. Cornell Lab of Ornithology.

Sunnåker, M., A. G. Busetto, E. Numminen, J. Corander, M. Foll, and C. Dessimoz. 2013. "Approximate Bayesian Computation." *PLoS Computational Biology* 9, no. 1: e1002803.

Sved, J. A. 2009. "Correlation Measures for Linkage Disequilibrium Within and Between Populations." *Genetics Research* 91, no. 3: 183–192.

Tataru, P., M. Simonsen, T. Bataillon, and A. Hobolth. 2017. "Statistical Inference in the Wright-Fisher Model Using Allele Frequency Data." *Systematic Biology* 66, no. 1: e30–e46.

Tellier, A., and C. Lemaire. 2014. "Coalescence 2.0: A Multiple Branching of Recent Theoretical Developments and Their Applications." *Molecular Ecology* 23, no. 11: 2637–2652.

Wager, S., T. Hastie, and B. Efron. 2014. "Confidence Intervals for Random Forests: The Jackknife and the Infinitesimal Jackknife." *Journal of Machine Learning Research: JMLR* 15, no. 1: 1625–1651.

Waples, R. S. 2022a. "*TheWeight*: A Simple and Flexible Algorithm for Simulating Non-Ideal, Age-Structured Populations." *Methods in Ecology and Evolution* 13, no. 9: 2030–2041.

Waples, R. S. 2022b. "What Is Ne, Anyway?" *Journal of Heredity* 113, no. 4: 371–379.

Waples, R. S. 2023. "Partitioning Variance in Reproductive Success, Within Years and Across Lifetimes." *Ecology and Evolution* 13, no. 11: e10647.

Waples, R. S., G. Luikart, J. R. Faulkner, and D. A. Tallmon. 2013. "Simple Life-History Traits Explain Key Effective Population Size Ratios Across Diverse Taxa." *Proceedings of the Royal Society B: Biological Sciences* 280, no. 1768: 20131339.

Waples, R. S., and M. Yokota. 2007. "Temporal Estimates of Effective Population Size in Species With Overlapping Generations." *Genetics* 175, no. 1: 219–233.

Wright, M. N., and A. Ziegler. 2015. Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R arXiv [stat. ML]. arXiv. http://arxiv.org/abs/1508.04409.

Wright, S. 1931. "Evolution in Mendelian Populations." *Genetics* 16, no. 2: 97–159.

Xue, A. T., and M. J. Hickerson. 2015. "The Aggregate Site Frequency Spectrum for Comparative Population Genomic Inference." *Molecular Ecology* 24, no. 24: 6223–6240.

Xue, A. T., and M. J. Hickerson. 2020. "Comparative Phylogeographic Inference With Genome-Wide Data From Aggregated Population Pairs." *Evolution; International Journal of Organic Evolution* 74, no. 5: 808–830.

Zamudio, K. R., R. C. Bell, and N. A. Mason. 2016. "Phenotypes in Phylogeography: Species' Traits, Environmental Variation, and Vertebrate Diversification." *Proceedings of the National Academy of Sciences of the United States of America* 113, no. 29: 8041–8048.

Zimmer, K., and M. L. Isler. 2020. "White-Shouldered Antshrike (*Thamnophilus aethiops*), Version 1.0." In *Birds of the World*, edited by J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana. Cornell Lab of Ornithology.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Appendix S1:** mec70155-sup-0001-AppendixS1.docx.

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.70155 by Cuny - Baruch College , Wiley Online Library on [10/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License