

## Species mtDNA genetic diversity explained by intrapopulation size in a host-symbiont system

Jorge Doña<sup>1\*</sup>, Marina Moreno-García<sup>1\*</sup>, Charles D. Criscione<sup>2</sup>, David Serrano<sup>3</sup> & Roger Jovani<sup>1</sup>

<sup>1</sup>Department of Evolutionary Ecology, Estación Biológica de Doñana (CSIC), Avda. Americo Vespucio s/n, Sevilla, Spain

<sup>2</sup>Department of Biology, Texas A&M University, 3258 TAMU, College Station, Texas 77843

<sup>3</sup>Department of Conservation Biology, Estación Biológica de Doñana (CSIC), Avda. Americo Vespucio s/n, Sevilla, Spain

### Keywords

COI, demography, feather mites, genetic diversity, host-parasite interactions, mtDNA.

### Correspondence

Jorge Doña, Department of Evolutionary Ecology, Estación Biológica de Doñana (CSIC), Avda. Americo Vespucio s/n, Sevilla, Spain. Tel: (+34) 954 466 700; Fax: (+34) 954 621 125; E-mail: jdona@ebd.csic.es

### Funding Information

Consejo Superior de Investigaciones Científicas (Grant/Award Number: 'CGL2011-24466', 'RYC-2009-03967', 'SVP-2013-067939').

Received: 14 October 2015; Revised: 26 October 2015; Accepted: 28 October 2015

doi: 10.1002/ece3.1842

\*These authors contributed equally to this work.

## Introduction

Understanding what shapes among-species variation in genetic diversity remains a major challenge in evolutionary ecology (Leffler et al. 2012; Romiguier et al. 2014; Fujisawa et al. 2015). This is especially true for parasites (Criscione et al. 2005; Huyse et al. 2005; Poulin 2011), despite being a widespread lifestyle (Poulin 2011) and a key element of ecosystems (Lafferty et al. 2006). Knowledge of what impacts parasite genetic diversity also has implications for host health and conservation (Criscione et al. 2005; Whiteman and Parker 2005).

Parasites and other symbionts (including commensals and mutualists) live in patchy and ephemeral habitats (i.e., individual hosts) and Price (1980) predicted that

## Abstract

Understanding what shapes variation in genetic diversity among species remains a major challenge in evolutionary ecology, and it has been seldom studied in parasites and other host-symbiont systems. Here, we studied mtDNA variation in a host-symbiont non-model system: 418 individual feather mites from 17 feather mite species living on 17 different passerine bird species. We explored how a surrogate of census size, the median intrapopulation size (i.e., the median number of individual parasites per infected host individual), explains mtDNA genetic diversity. Feather mite species genetic diversity was positively correlated with mean intrapopulation size, explaining 34% of the variation. As expected from the biology of feather mites, we found bottleneck signatures for most of the species studied but, in particular, three species presented extremely low mtDNA diversity values given their intrapopulation size. Their star-like haplotype networks (in contrast with more reticulated networks for the other species) suggested that their low genetic diversity was the consequence of severe bottlenecks or selective sweeps. Our study shows for the first time that mtDNA diversity can be explained by intrapopulation sizes, and suggests that departures from this relationship could be informative of underlying ecological and evolutionary processes.

parasites, in contrast with free-living species, would have a lower genetic diversity. However, we currently know that parasite genetic diversity is comparable to that of free-living organisms and variable between parasite species (Nadler 1990; Bush et al. 1997, 2001; Blouin et al. 1999; Criscione and Blouin 2005; Poulin 2011). The question that remains then is what dictates the variation in genetic diversity among parasites species and other symbionts in host-symbiont interactions.

Population genetic theory predicts that neutral genetic diversity should increase proportionally with the effective population size,  $N_e$  (i.e., the size of a Wright-Fisher ideal population that experiences the same rate of drift as the population under consideration). In fact, recent correlative studies that used surrogates of census population

sizes of various free-living species have found support for this prediction (McCusker and Bentzen 2010; Romiguier et al. 2014; Fujisawa et al. 2015). For parasites, several factors such as aggregated distributions, crowding effects, and host immunity could influence the  $N_e$  of parasite component populations (i.e., all the parasites among all the infected hosts within the host population; Bush et al. 1997) (Criscione and Blouin 2005). As an initial null expectation, Criscione et al. (2005) hypothesized that parasite species with larger census sizes would have larger  $N_e$ , and thus higher genetic diversity. Interestingly, Criscione et al. (2005) found tentative support for this hypothesis by qualitatively correlating published mtDNA diversity data for seven nematode species and their mean intensities of infection, i.e., the mean number of parasites per infected host. Their prediction is based on the assumption that the census size of the parasite component population is reflected by the mean infrapopulation size (i.e., mean infection intensities) (Criscione et al. 2005). Mean intensities were used because an estimate of the census component population would require an estimate of the host population size as well. To our knowledge, there has yet to be an explicit statistical test to determine if infection intensities correlate with genetic diversity in parasites.

Here, we test the hypothesis that species-wide mitochondrial (mtDNA) genetic diversity in a group of ectosymbiont species is correlated with their median infrapopulation sizes. The finding of a significant relationship would indicate that infrapopulation sizes influence species-wide  $N_e$  (or more specifically the effective sizes of the mtDNA of each species). We test this hypothesis using 17 feather mite species living on 17 different passerine bird species. Feather mites (Acari: Astigmata: Analgoidea and Pterolichoidea) are among the most abundant ectosymbionts living on birds (Gaud and Atyeo 1996; Proctor 2003). Feather mite median intensity strongly and consistently differs among bird species when studied in different populations and habitats (Díaz-Real et al. 2014), so median infrapopulation size can be considered as an inherent characteristic of each host species. Within feather mites some families are parasites of the skin and nasal

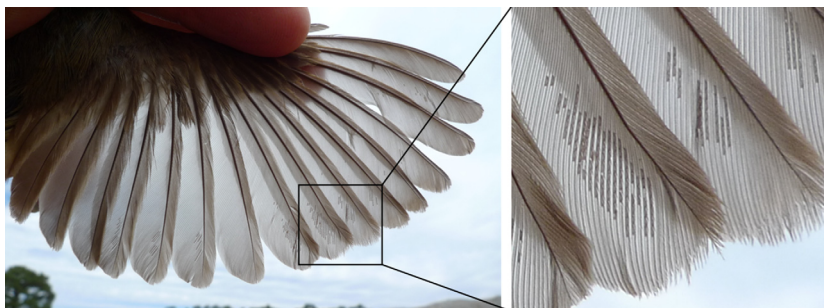
cavities of birds. However, most species are highly specialized to different plumage microhabitats (plumicolous mites, hereafter feather mites), where they cause no visible damage to their hosts (Blanco et al. 2001; Dowling et al. 2001; Galván et al. 2008; Fig. 1). Recent studies even suggest that these feather mites may be commensals or even mutualists of birds, feeding on keratinophilic fungi and feather degrading bacteria (Blanco et al. 2001; Galván et al. 2012). Nonetheless, they share commonalities to the parasite life style in that feather mites expend their whole life on their bird hosts, mainly transmitting from bird parents to offspring on the host nest (Dubinin 1951; Dabert and Mironov 1999; Proctor 2003; Dabert et al. 2015).

## Materials and Methods

### Data

418 individual mites from 17 bird species (one mite species for each bird species) were sampled by capturing birds with mist nets and collecting all present mites from the wing flight feathers using a cotton swab impregnated with ethanol and preserved at  $-20^{\circ}\text{C}$  in tubes with 96% ethanol for posterior lab separation. Sampling was done in two steps (Table S1, Supporting information).

First, in different localities of South Spain during 2014, we chose (see stereomicroscope separation below) five mite individuals from each of two to five individual birds from each of 14 bird species (one mite species per bird species). This sample consisted of a total of 297 individual mites. Sampling details (i.e., locality, date of sampling, and sample size) can be found in Table S1 (Supplementary information). From previous taxonomic (Atyeo and Braasch 1966; Gaud and Atyeo 1996) and molecular (Doña et al. 2015) information, we had a reasonable knowledge on the common feather mite species living on our study bird species. Thus, once in the laboratory, we identified mite specimens corresponding to the most abundant feather mite species living on each bird species under a stereomicroscope according to basic morphological traits. Subsequent molecular barcoding identification



**Figure 1.** *Dolichodectes hispanicus* feather mites on the wing flight feathers of a Melodious warbler *Hippolais polyglotta*. Infrapopulation size is measured by counting the number of mites exposing the wing against daylight.

(Doña et al. 2015) confirmed initial identifications in all cases (see also Díaz-Real et al. 2015).

Second, we included 91 mtDNA sequences (91 mite individuals) from the same 14 mite species, coming from Doña et al. (2015), where we sampled from 2011 to 2013 in Spain and Russia one mite individual from each bird individual (Table S1, Supporting information). Moreover, 30 sequences from three new mite species were included from Doña et al. (2015) (Table S1, Supporting information). [Other mite species with less than nine sequences in Doña et al. (2015) were not included to be conservative with the minimum sample size suggested in Goodall-Copestake et al. (2012)].

In addition, we included 34 individual mites (i.e., 25 new samples from the field and 9 from Doña et al. (2015)) from *Proctophylloides serini*, a species we a priori did not expect to fulfill our prediction due to its recent evolutionary history (see Discussion). In particular, we were interested on exploring whether the recent divergence of this species translated into a departure from the general relationship between intrapopulation size and genetic diversity. Therefore, we conducted two analyses: one that included *P. serini* (18 species total) and one that did not (17 species total).

Median intensity significantly differs among feather mite species and is moderately repeatable (Díaz-Real et al. 2014). Here, median intrapopulation size for each mite species was obtained from Díaz-Real et al. (2014), whose data set is the largest so far available for feather mite loads in European passerine birds (gathered by counting the number of mites exposing the wing against daylight; Fig. 1). For the bird species analysed here, this data set includes a total of 6,607 individual birds, with a mean (range) of 389 (11–2,197) birds per species.

## DNA analysis

Genomic DNA was extracted using HotSHOT (Truett et al. 2000). A segment of approximately 650 bp of the COI region was amplified. This marker has been shown to be useful to characterize the genetic diversity in other species (Delrieu-Trottin et al. 2014; Fujisawa et al. 2015). In addition, this marker offers within-species polymorphism in feather mites species (Doña et al. 2015). PCRs were done by using degenerate primers bcdF05 (50-TTTTCTACH AAYCATAAAGATATTGC-30) and bcdR04 (50-TATAAACYTCDGGATGNCCAA AAAA-30) (Dabert et al. 2008). PCRs were carried out in 20  $\mu$ L reaction volumes containing 1x (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> reaction buffer (Bioline), 2.5 mM MgCl<sub>2</sub>, 19 BSA, 0.25 mM DNTPs, 2  $\mu$ m of each primer, 1.25 U BIOTAQTM (Bioline), and 7  $\mu$ L of DNA template. The reaction followed a touchdown PCR profile: 95°C for 3 min, 20 cycles of 95°C for 1 min, 55°C

for 30 s with a decrease of 0.5°C every cycle, 72°C for 1 min, and 20 cycles of 95°C for 1 min, 45°C for 30 s and 72°C for 1 min, with a final extension step of 72°C for 5 min. PCR products were quantitatively assessed by electrophoresis on a 2% agarose gel, and visible bands corresponding to the COI fragment size were sequenced in both directions. COI sequencing was carried out using the Sanger method and performed by Macrogen, Europe (Holland).

Sequences, including those from Doña et al. (2015), were visually edited and manually trimmed to 507 bp using Geneious v4.7 (Drummond et al. 2009). New sequences were deposited in GenBank with the accession nos. KT025257–KT025578. Sequences were aligned using MUSCLE with default settings (Edgar 2004). Clustering of species was evaluated on Neighbor-joining trees done in MEGA 6 (Tamura et al. 2013).

## Data analysis

### Haplotype networks, nucleotide diversity, and repeatability of nucleotide diversity

We used the *haplotype* and *haplonet* functions from the PEGAS R package to construct haplotype networks (Paradis 2010). Nucleotide diversity ( $\pi$ , hereafter genetic diversity; Nei 1987) was calculated using the *nuc.div* function implemented in the R package PEGAS (Paradis 2010). This function computes the sum of the number of differences between pairs of sequences divided by the number of comparisons.

Our study is based on the assumption that our sampling was good enough to obtain a reasonable estimate of the overall genetic diversity for each mite species. This assumption is reasonably supported by current evidence. First, no genetic structure has been found in any species of feather mite studied to date in relation to the migratory status of the host species (Fernández et al. 2013), or in relation to the sampling locality in our previous study from distant geographic localities (Russia vs. Spain, Doña et al. 2015). Moreover, a visual inspection of haplotype networks of the sequences analyzed here showed no evidence of geographic structure (Figs. S3–S12 Supporting information), and no consistent difference was found within species in genetic diversity estimates when including or not the 42 Russian samples (paired *t* test  $t = -0.245$ ,  $df = 31.994$ ,  $P = 0.8$ ; Fig. S13, Supporting information).

Second, we used the *ICCbare* function of the ICC R package (Wolak et al. 2012) to calculate the intraclass correlation coefficient (i.e., repeatability, *R*), that is, to test how repeatable is genetic diversity among intrapopulations from the same species. For the 15 (including

*P. serini*) species for which we sampled more than two mites per individual bird, we calculated the genetic diversity of each infrapopulation. The genetic diversity was highly repeatable between infrapopulations of the same species, but differed between species ( $R = 0.79$ , 95% CI = 0.63–0.94). In addition, we explored if some infrapopulations were biasing estimates of species-wide genetic diversity. In particular, we tested if there was any difference between the full data set (i.e., five mites per bird) and a partial data set (i.e., one mite per bird) generated via a subsampling simulation (see “Simulation analysis”). There was a significant positive correlation between the average genetic diversity of all infrapopulations from each species and our species genetic diversity values calculated by picking up only one mite from each individual bird (Spearman’s correlation,  $\rho = 0.75$ ,  $P < 0.01$ ), thus supporting the robustness of the estimates.

Overall, we are confident that our estimates of genetic diversity at the species level are robust and thus we were able to test for the relationship between the typical infrapopulation size of each mite species (coming from the data set in Díaz-Real et al. 2014) and our estimate of feather mite genetic diversity at the species level.

### Neutrality test

Departures from mutation-drift equilibrium of a Wright-Fisher model were assessed by calculating Tajima  $D$  (Tajima 1989), using the *tajima.test* function of R package PEGAS (Paradis 2010); significance was obtained assuming that  $D$  follows a normal distribution with mean zero and variance one. We also investigated departures using the  $R_2$  test, known to be very powerful in detecting population expansions even for small sample sizes (Ramos-Onsins and Rozas 2002).  $R_2$  was calculated using the *R2.test* function implemented in the R package PEGAS (Paradis 2010) and its statistical significance was calculated through 1000 coalescent simulations.

### Simulations analysis

In order to use all mtDNA sequences available, and to take into account uncertainty on the estimated infrapopulation sizes, we tested the relationship between species infrapopulation size and genetic diversity by running 100 Generalized Linear Models (GLMs). For each run, we recalculated species genetic diversity and infrapopulation size of each species as explained below.

Feather mite infrapopulations are thought to be created mainly by vertical transmission of mites from bird parents to offspring (Proctor 2003). It is thus expected that mites within a bird individual would not be genetically independent (even probably including close relatives). In

fact, we found mites from the same bird individual often sharing the same haplotype (Figs. S3–S12 Supporting information), thus potentially underestimating species genetic diversity (see Simulation details). Therefore, we created 100 new data sets, each having a genetic diversity estimate ( $\pi$ ) for each mite species that was calculated by randomly sampling a single mite individual from each individual bird (see Data, Table S1 Supporting information). We used the data generated in each iteration to test the statistical relationship between genetic diversity and infrapopulation size (see below; “Statistical analysis”). In the same way, we also used this simulation to estimate Tajima’s  $D$  and  $R_2$  values for each mite species. The average of the estimates from all iterations was used as the representative estimate of genetic diversity, Tajima’s  $D$ , or  $R_2$  values for each species (Table 1).

We also used this simulation approach to explore the robustness of our results against uncertainty in infrapopulation size per mite species. The feather mite loads in Díaz-Real et al. (2014) are based on counts of all feather mites living on the wing of each bird host (i.e., counting all mites without differentiating mite species). It should be noted that two or three feather mite species often inhabit the wing of a same bird species (Proctor 2003). However, we studied the genetic diversity of the most abundant mite species occurring on each bird species (see above). Therefore, the median infrapopulation size reported by Díaz-Real et al. (2014) is either the real median of our focal mite species (if this was the only species occurring in the bird host) or a slightly overestimated value (i.e., if other species were in that bird). Therefore, to consider this uncertainty, for each of the 100 iterations (see above) we also calculated a new median infrapopulation size value for each mite species by extracting a pseudo-random value between half and the real median load of feather mites per bird species reported in Díaz-Real et al. (2014). Nevertheless, a similar positive relationship with the genetic diversity (see Results) was obtained when using the non-simulated infrapopulation size data (GLM:  $F = 9.00$ ;  $df = 15$ ,  $P < 0.01$ ).

### Statistical analysis

We used the most comprehensive phylogeny for European plumicolous feather mite species (Doña et al. 2015) to control for phylogenetic non-independence of the data (Felsenstein 1985; Freckleton et al. 2002). We calculated the phylogenetic signal for genetic diversity estimating the maximum likelihood value of Pagel’s  $\lambda$  (Pagel 1999). We used the function *pgls* of the R package Caper (Orme 2012) version 0.5.2 (<http://caper.r-forge.r-project.org/>). Mean  $\lambda$  was  $< 0.001$  for genetic diversity (Fig. S1), not

**Table 1.** Feather mite species and bird host names, number of mites ( $n$ ) and number of infrapopulations (lf) sampled, number of polymorphic sites ( $S$ ), and demographic statistics results ( $D$ ,  $R_2$ , and  $R_{2\%}$  sig.). Demographic statistic values reported are the average from 100 iterations (see Materials and Methods). The % sig. column shows the percentage of statistically significant iterations. Species with percentages of significant iterations higher than 95% or 75% were considered as significant (\*\*) or marginally significant (\*) respectively. Tajima's  $D$  only was statistically significant for some iterations in *P. rubeculinus* (33%) and *P. serini* (57%).

Feather mite species	Bird species	$n$	lf	$S$	$D$	$R_2$	$R_{2\%}$ sig.
<i>Dolichodectes hispanicus</i>	<i>Hippolais polyglotta</i>	30	10	22	-1.10	0.08	100**
<i>Monojoubertia microphylla</i>	<i>Fringilla coelebs</i>	35	15	46	-0.93	0.09	79*
<i>Proctophyllodes acanthicaulus</i>	<i>Muscicapa striata</i>	10	10	26	-1.50	0.09	100**
<i>Proctophyllodes ateri</i>	<i>Periparus ater</i>	29	9	10	-0.01	0.18	1
<i>Proctophyllodes cetti</i>	<i>Cettia cetti</i>	29	9	22	-0.29	0.13	21
<i>Proctophyllodes clavatus</i>	<i>Sylvia borin</i>	14	8	23	-0.41	0.12	52
<i>Proctophyllodes cotyledon</i>	<i>Phoenicurus ochruros</i>	24	8	11	-1.62	0.13	36
<i>Proctophyllodes doleophyes</i>	<i>Ficedula hypoleuca</i>	29	9	22	-1.19	0.09	100**
<i>Proctophyllodes lusciniae</i>	<i>Luscinia megarhynchos</i>	41	21	23	-1.51	0.06	100**
<i>Proctophyllodes motacillae</i>	<i>Motacilla flava</i>	11	11	43	-1.55	0.07	100**
<i>Proctophyllodes musicus</i>	<i>Turdus merula</i>	29	9	17	-0.52	0.12	42
<i>Proctophyllodes rubeculinus</i>	<i>Erithacus rubecula</i>	37	17	10	-1.88	0.08	76*
<i>Proctophyllodes schoenicli</i>	<i>Emberiza schoeniclus</i>	9	9	18	-1.80	0.07	100**
<i>Proctophyllodes serini</i>	<i>Serinus serinus</i>	34	14	22	-1.97	0.10	46
<i>Proctophyllodes stylifer</i>	<i>Parus major</i>	19	10	25	-0.67	0.09	100**
<i>Proctophyllodes sylviae</i>	<i>Sylvia atricapilla</i>	32	12	59	-0.71	0.09	85*
<i>Pteronyssoides parinus</i>	<i>Cyanistes caeruleus</i>	28	8	19	0.25	0.16	4
<i>Trouessartia rosterii</i>	<i>Sturnus unicolor</i>	12	6	21	-1.19	0.12	57

significantly different from 0 ( $P=1$ ), but significantly different from 1 ( $P<0.001$ ). Genetic diversity was always the response variable in our analysis. Thus, as response variables are the most influential on residuals, it is unlikely that our residuals could show phylogenetic signal.

Therefore, we used a non-phylogenetic approach (Revell 2010): we analysed the relationship between genetic diversity and infrapopulation size using GLMs with a Gaussian distribution of errors and an identity link function. We performed one GLM for each simulation iteration (see above, Table S2. Supporting information). We used the *glm* function of the STATS R package (R Development Core Team, 2013). The genetic diversity of feather mite species was the dependent variable and the median infrapopulation size was the independent variable. Sample size (number of analysed mites per bird species) varied among mite species, but it was uncorrelated with their genetic diversity (Spearman's correlation,  $\rho = -0.31$ ,  $P = 0.1962$ ). Even so, we used the "weights" option of the *glm* function to give more confidence to data coming from higher sample sizes. We confirmed assumptions underlying GLMs by exploring regression residuals for normality against a Q-Q plot.

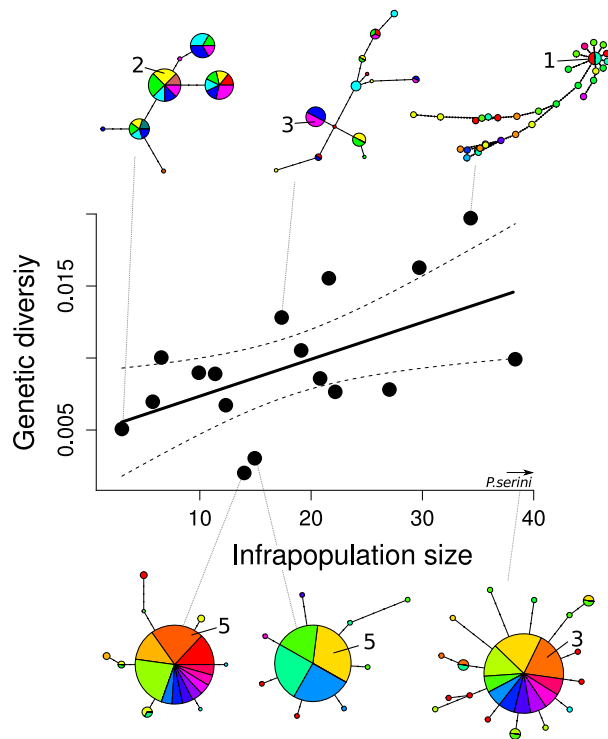
## Results

Non-significant departures from mutation-drift equilibrium were found for Tajima's  $D$ , but 7 out of 18 (including *P. serini*) species displayed positive and significant low

values in all iterations for the most conservative  $R_2$  statistic (a signature of population expansion). Three other species showed the same patterns but only with marginal statistical significance (Table 1). Despite so, species presented high and contrasting genetic diversity values (Table 1).

When excluding *P. serini*, we found a strong positive relationship between infrapopulation size and genetic diversity across species, and results were robust across the 100 GLMs performed:  $F$  mean (95%CI) = 8.66 (7.61–9.71),  $df = 15$ ,  $P$  mean (95% CI) = 0.03 (0.02–0.04), explaining a mean (95% CI) of 34% (31–36%) of the original deviance. Similarly, the relationship between the median infrapopulation size and the mean  $\pi$  calculated for each species across the 100 iterations (see "Simulation analysis") explained 37% of the original deviance (GLM:  $F = 8.98$ ;  $df = 15$ ,  $P < 0.01$ ; Fig. 2).

Despite being the species with the largest median infrapopulation size, *P. serini* showed, as predicted (see Materials and Methods and Discussion), an extremely low genetic diversity, and a star-like haplotype network. Including this species into the analyses lead to a statistically non-significant relationship between median infrapopulation size and mean genetic diversity (GLM:  $F = 0.74$ ;  $d.f. = 16$ ,  $P = 0.40$ ). Interestingly, not only this species, but also *P. rubeculinus* and *P. cotyledon* showed low values of genetic diversity according to their infrapopulation sizes, and the three species shared a strong star-like haplotype network structure, in contrast with the more



**Figure 2.** Relationship between intrapopulation size and nucleotide diversity ( $\pi$ ) across feather mite species. Haplotype network examples, from top left to right: *Proctophyllodes ateri*, *P. cetti* and *P. sylviae*; bottom left to right: *P. rubeculinus*, *P. cotyledon* and *P. serini*. Each color in pie charts represents an individual bird. Numbers indicate for each network the maximum number of mites sharing a given haplotype into an individual bird. *P. serini* intrapopulation size was 49, and its genetic diversity 0.003.

reticulated haplotype network structure of the other species (Figs. 2; S2–S12, Supporting information). The original deviance explained by the model increased to 42% when excluding these three star-like species from the analysis (GLM:  $F = 9.27$ ;  $df = 15$ ,  $P < 0.01$ ).

## Discussion

Our results show a strong positive relationship between median intrapopulation size and genetic diversity across 17 feather mite species. Therefore, our findings are in line with previous preliminary evidence suggesting that a simple ecological measure such as the median number of parasites per infected host individual is a relevant variable to understand evolutionary processes at the species level (Criscione et al. 2005). So, disentangling why symbionts such as feather mites differ so strongly in their intrapopulation sizes among species (Díaz-Real et al. 2014) is a necessary ingredient to understand the evolutionary ecology of these host-symbiont systems. In fact, while species richness has been a major topic in the study of the evolu-

tionary ecology of parasites, understanding intrapopulation size variation between parasite species and parasite load between host species is still a major challenge (Poulin and George-Nascimento 2007).

In spite of the importance of our results, there is still a considerable amount of variance in mtDNA diversity that remains unexplained by median infection intensities. There are several possible and not necessarily mutually exclusive explanations for this. First, departures from mutation-drift equilibrium can distort the median intrapopulation size  $\sim$  genetic diversity relationship (Crow and Kimura 1970). That is, historical demographic events (expansions and bottlenecks) may distort levels of diversity beyond what one might expect based on a strict census to effective size relationship. For example, Dabert et al. (2015) reported low genetic diversity for another feather mite species, *Zachvatkinia isolate*. They interpreted this as a consequence of bottlenecks and founder effects occurring during bird-to-bird feather mite transmission. Interestingly, our data reveal several signatures of population expansions based on the  $R_2$  tests (Table 1) that may have introduced noise into the data set. For instance, the low genetic diversity values coupled with the shapes of the networks for *P. rubeculinus* and *P. cotyledon* suggested extreme recent demographic changes or a selective sweep (Shaw 2002; Bensch et al. 2006; Irwin et al. 2009; Rato et al. 2011). In effect, if we exclude these two species from the analysis the model fitted the data better (the original deviance explained increases to 42%), providing further evidence of the importance of the demographic and evolutionary context on the relationship between genetic diversity and population size.

Species divergence history may also be important in influencing an intrapopulation size - genetic diversity relationship. Regardless of the intrapopulation sizes of the species, recently diverged species may show low genetic diversity because they would have retained only a fraction of mutations that were segregating in the parental species (Barton and Charlesworth 1984; Klicka and Zink 1997; Nichols 2001; De Brito et al. 2005). As an extreme example of this, we highlight the effect of *P. serini*. This species is part of the *Pinnatus* group. The *Pinnatus* group of the genus *Proctophyllodes* is the best known example of this sort of processes among feather mites from European passerines. This is a group identified by taxonomists as being constituted by close morphological species (Atyeo and Braasch 1966; Badek et al. 2008), and a multilocus phylogenetic study showed that this is a recent species complex with rapid diversification rate (Knowles and Klimov 2011). In addition, we recently found that COI barcoding was unable to delimitate between morphological species within this complex (Doña et al. 2015). Here, we have studied one of these species, *P. serini*, living on the feathers of European

serins (*Serinus serinus*), and we found an extremely low genetic diversity (Goodall-Copestake et al. 2012), regardless of its large infrapopulation sizes (see Fig. 2).

Another factor introducing noise on the correlation between mtDNA genetic diversity and infrapopulation size could be that infrapopulation size is not an accurate reflection of the total census size (component population) of feather mites. Further work is needed here. Lastly, the high genetic diversity values found for most species (Goodall-Copestake et al. 2012) could be explained by their high mutation rate (Hodgkinson and Eyre-Walker 2011), as reported in previous feather mite studies (Dabert et al. 2010; Štefka et al. 2011).

Overall, our study shows that infrapopulation size does indeed explain a significant amount of the among-species variation in genetic diversity of parasites. Infrapopulation size is easier to estimate than census size or  $N_e$ . In addition, the relationship between infrapopulation size and genetic diversity we report here supports the idea that ecological processes occurring at the within-individual host scale are key to understand the distribution of genetic variation among parasite species. However, how other factors such as prevalence, host census size, or other measures of infrapopulation size variation (such as the harmonic mean) may influence the genetic diversity of feather mites will require further studies. Moreover, we report that pervasive demographic processes in host-parasite systems such as bottlenecks or population growth events could strongly impact the species genetic diversity. This should encourage further research of underlying mechanisms of infrapopulation dynamics that will be relevant to achieve a complete picture of the factors shaping parasite species genetic diversity.

## Acknowledgments

Funding was provided by the Ministry of Economy and Competitiveness (Ramón y Cajal research contract RYC-2009-03967 to RJ and research project CGL2011-24466 to RJ). JD was also supported by the Ministry of Economy and Competitiveness (SVP-2013-067939). Special thanks to Javier Díaz-Real for the lab assistance; Alejandro Gonzalez-Voyer for his help with comparative analysis; Carolina Osuna, Carlos Gutiérrez Expósito, Pepe Ayala, Carlos Camacho, David Ochoa, Jaime Potti, José Luis Garzón Perez and Francisco Jiménez Cazalla and GOES for their collaboration during field work. The authors thank three anonymous reviewers for suggesting improvements to an earlier version of the manuscript.

## Conflict of Interest

None declared.

## Data Accessibility

DNA sequences: GenBank accessions: KT025257-KT025578.

Sequence alignment and simulation results are deposited in Dryad: doi:10.5061/dryad.93sh8. The R code used for simulation is available on the GitHub repository: [https://github.com/Jorge-Dona/Infrapopulation\\_genetics](https://github.com/Jorge-Dona/Infrapopulation_genetics).

## References

- Atyeo, W. T., and N. L. Braasch. 1966. The feather mite genus *Proctophyllodes* (Sarcoptiformes: Proctophyllodidae). *Bullet. Univ. Nebras. State Mus.* 5:1–354.
- Badek, A., M. Dabert, S. V. Mironov, and J. Dabert. 2008. A new species of the genus *Proctophyllodes* (Analgoidea: Proctophyllodidae) from Cetti's warbler *Cettia cetti* (Passeriformes: Sylviidae) with DNA Barcode Data. *Annal. Zool.* 58:397–402.
- Barton, N. H., and B. Charlesworth. 1984. Genetic revolutions, founder effects, and speciation. *Annu. Rev. Ecol. Evol. Syst.* 15:133–164.
- Bensch, S., D. E. Irwin, J. H. Irwin, L. Kvist, and S. Åkesson. 2006. Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Mol. Ecol.* 15:161–171.
- Blanco, G., J. L. Tella, and J. Potti. 2001. Feather mites on birds: costs of parasitism or conditional outcomes? *J. Avian Biol.* 32:271–274.
- Blouin, M. S., J. Liu, and R. E. Berry. 1999. Life cycle variation and the genetic structure of nematode populations. *Heredity* 83:253–259.
- Bush, A. O., K. D. Lafferty, J. M. Lotz, and A. W. Shostak. 1997. Parasitology meets ecology on its own terms: Margolis et al. revisited. *J. Parasitol.*, 83, 575–583.
- Bush, A. O., J. C. Fernandez, G. W. Esch, and J. R. Seed. 2001. *Parasitism: The Diversity and Ecology of Animal Parasites*. Cambridge University Press, Cambridge, UK.
- Criscione, C. D., and M. S. Blouin. 2005. Effective sizes of macroparasite populations: a conceptual model. *Trends Parasitol.* 21:212–217.
- Criscione, C. D., R. Poulin, and M. S. Blouin. 2005. Molecular ecology of parasites: elucidating ecological and microevolutionary processes. *Mol. Ecol.* 14:2247–2257.
- Crow, J. F., and M. Kimura. 1970. *An introduction to population genetics theory*. The Blackburn Press, Caldwell, New Jersey.
- Dabert, J., and S. V. Mironov. 1999. Origin and evolution of feather mites (Astigmata). *Exp. Appl. Acarol.* 23:437–454.
- Dabert, J., R. Ehrnsberger, and M. Dabert. 2008. *Glaucales tytonis* sp. n. (Analgoidea, Xolalgidae) from the barn owl *Tyto alba* (Strigiformes, Tytonidae): compiling morphology with DNA barcode data for taxon descriptions in mites (Acari). *Zootaxa* 1719:41–52.

- Dabert, M., W. Witalinski, A. Kazmierski, Z. Olszanowski, and J. Dabert. 2010. Molecular phylogeny of acariform mites (Acari, Arachnida): strong conflict between phylogenetic signal and long-branch attraction artifacts. *Mol. Phylogenet. Evol.* 56:222–241.
- Dabert, M., S. J. Coulson, D. J. Gwiazdowicz, B. Moe, S. A. Hanssen, E. M. Biersma, et al. 2015. Differences in speciation progress in feather mites (Analgoidea) inhabiting the same host: the case of *Zachvatkinia* and *Alloptes* living on arctic and long-tailed skuas. *Exp. Appl. Acarol.* 65:163–179.
- De Brito, R. A., L. S. Pletscher, and J. M. Cheverud. 2005. The evolution of genetic architecture. I. Diversification of genetic backgrounds by genetic drift. *Evolution* 59:2333–2342.
- Delrieu-Trottin, E., J. Maynard, and S. Planes. 2014. Endemic and widespread coral reef fishes have similar mitochondrial genetic diversity. *Proc. R. Soc. Lond. [Biol.]* 281:20141068.
- Díaz-Real, J., D. Serrano, J. Pérez-Tris, S. Fernández-González, A. Bermejo, J. Calleja, et al. 2014. Repeatability of feather mite prevalence and intensity in passerine birds. *PLoS One* 9:e107341.
- Díaz-Real, J., D. Serrano, A. Piriz, and R. Jovani. 2015. NGS metabarcoding proves successful for quantitative assessment of symbiont abundance: the case of feather mites on birds. *Exp. Appl. Acarol.* 67:209–218.
- Doña, J., J. Díaz-Real, S. Mironov, P. Bazaga, D. Serrano, and R. Jovani. 2015. DNA barcoding and minibarcoding as a powerful tool for feather mite studies. *Mol. Ecol. Resour.* doi:10.1111/1755-0998.12384.
- Dowling, D. K., D. S. Richardson, and J. Komdeur. 2001. No effect of a feather mite on body condition, survivorship, or grooming behavior in the Seychelles warbler *Acrocephalus sechellensis*. *Behav. Ecol. Sociobiol.* 50:257–262.
- Drummond, A.J., B. Ashton, M. Cheung, et al. (2009) Geneious v4. 7. Biomatters, Ltd., Auckland, New Zealand.
- Dubinin, V. B. 1951. Feather mites (Analgosoidea). Part I. Introduction to their study. *Fauna SSSR, Paukoobraznye* 6:1–363.
- Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.* 125:1–15.
- Fernandez, S. 2013. Ecology and evolutionary perspectives of feather mite coexistence on the black cap *Sylvia atricapilla*. Ph.D. thesis. Universidad Complutense de Madrid.
- Freckleton, R. P., P. H. Harvey, and M. Pagel. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.* 160:712–726.
- Fujisawa, T., A. P. Vogler, and T. G. Barraclough. 2015. Ecology has contrasting effects on genetic variation within species versus rates of molecular evolution across species in water beetles. *Proceed. Royal Soc. B: Biol. Sci.* 282:20142476.
- Galván, I., E. Barba, R. Piculo, J. L. Cantó, V. Cortés, J. S. Monrós, et al. 2008. Feather mites and birds: an interaction mediated by uropygial gland size? *J. Evol. Biol.* 21:133–145.
- Galván, I., E. Aguilera, F. Atiénzar, E. Barba, G. Blanco, J. L. Cantó, et al. 2012. Feather mites (Acari: Astigmata) and body condition of their avian hosts: a large correlative study. *J. Avian Biol.* 43:273–279.
- Gaud, J., and W. T. Atyeo. 1996. Feather mites of the World (Acarina, Astigmata): the supraspecific taxa. *Annal. du Musée Roy. de L'Afrique Centrale Sci. Zool.*, 277, 1–193 (Pt. 1, text), 1–436 (Pt. 2, illustrations).
- Goodall-Copestake, W. P., G. A. Tarling, and E. J. Murphy. 2012. On the comparison of population-level estimates of haplotype and nucleotide diversity: a case study using the gene *cox1* in animals. *Heredity* 109:50–56.
- Hodgkinson, A., and A. Eyre-Walker. 2011. Variation in the mutation rate across mammalian genomes. *Nat. Rev. Genet.* 12:756–766.
- Huysse, T., R. Poulin, and A. Theron. 2005. Speciation in parasites: a population genetics approach. *Trends Parasitol.* 21:469–475.
- Irwin, D. E., A. S. Rubtsov, and E. N. Panov. 2009. Mitochondrial introgression and replacement between yellowhammers (*Emberiza citrinella*) and pine buntings (*Emberiza leucocephalos*) (Aves: Passeriformes). *Biol. J. Linn. Soc.* 98:422–438.
- Klicka, J., and R. M. Zink. 1997. The importance of recent ice ages in speciation: a failed paradigm. *Science* 277:1666–1669.
- Knowles, L. L., and P. B. Klimov. 2011. Estimating phylogenetic relationships despite discordant gene trees across loci: the species tree of a diverse species group of feather mites (Acari: Proctophylloidea). *Parasitology* 138:1750–1759.
- Lafferty, K. D., A. P. Dobson, and A. M. Kuris. 2006. Parasites dominate food web links. *Proc. Natl Acad. Sci. USA* 103:11211–11216.
- Leffler, E. M., K. Bullaughey, D. R. Matute, W. K. Meyer, L. Segurel, A. Venkat, et al. 2012. Revisiting an old riddle: what determines genetic diversity levels within species. *PLoS Biol.* 10:e1001388.
- McCusker, M. R., and P. Bentzen. 2010. Positive relationships between genetic diversity and abundance in fishes. *Mol. Ecol.* 19:4852–4862.
- Nadler, S. A. 1990. Molecular approaches to studying helminth population genetics and phylogeny. *Int. J. Parasitol.* 20:11–29.
- Nei, M. 1987. *Molecular evolutionary genetics*. Columbia University Press, New York.
- Nichols, R. 2001. Gene trees and species trees are not the same. *Trends Ecol. Evol.* 16:358–364.
- Orme, C. D. L. 2012. Caper: comparative analyses of phylogenetics and evolution in R. (<http://cran.r-project.org/web/packages/caper/>).
- Pagel, M. 1999. Inferring the historical patterns of biological evolution. *Nature* 401:877–884.



- Paradis, E. 2010. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26:419–420.
- Poulin, R. 2011. *Evolutionary ecology of parasites*. Princeton University Press, New Jersey, USA.
- Poulin, R., and M. George-Nascimento. 2007. The scaling of total parasite biomass with host body mass. *Int. J. Parasitol.* 37:359–364.
- Price, P. W. 1980. *Evolutionary biology of parasites*. Princeton University Press, Princeton, New Jersey.
- Proctor, H. C. 2003. Feather mites (Acari: Astigmata): ecology, behavior, and evolution. *Annu. Rev. Entomol.* 48:185–209.
- R Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Ramos-Onsins, S. E., and J. Rozas. 2002. Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* 19:2092–2100.
- Rato, C., S. Carranza, and D. J. Harris. 2011. When selection deceives phylogeographic interpretation: the case of the Mediterranean house gecko, *Hemidactylus turcicus* (Linnaeus, 1758). *Mol. Phylogenet. Evol.* 58:365–373.
- Revell, L. J. 2010. Phylogenetic signal and linear regression on species data. *Methods Ecol. Evol.* 1:319–329.
- Romiguier, J., P. Gayral, M. Ballenghien, A. Bernard, V. Cahais, A. Chenuil, et al. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* 515:261–263.
- Shaw, K. L. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proc. Natl Acad. Sci. USA* 99:16122–16127.
- Štefka, J., P. E. Hoeck, L. F. Keller, and V. S. Smith. 2011. A hitchhikers guide to the Galápagos: co-phylogeography of Galápagos mockingbirds and their parasites. *BMC Evol. Biol.* 11:284.
- Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30:2725–2729.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Truett, G. E., P. Heeger, R. L. Mynatt, A. A. Truett, J. A. Walker, and M. L. Warman. 2000. Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques* 29:52–54.
- Whiteman, N. K., and P. G. Parker. 2005. Using parasites to infer host population history: a new rationale for parasite conservation. *Anim. Conserv.* 8:175–181.
- Wolak, M. E., D. J. Fairbairn, and Y. R. Paulsen. 2012. Guidelines for estimating repeatability. *Methods Ecol. Evol.* 3:129–137.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** Log likelihood profile of lambda estimation (here is shown only one iteration calculation).

**Figure S2–S13.** Each colour in pie charts represents an individual bird. Pie size represents haplotype frequency.

**Table S1.** Sampling details: locality, date of sampling, sample size, samples id, collectors and GenBank accessions.