Title: Comparing rates of introgression in parasitic feather lice with differing dispersal capabilities

Authors: Jorge Doña[1, 2]*, Andrew D. Sweet[1, 3], and Kevin P. Johnson[1]*

Affiliations:

[1]Illinois Natural History Survey, Prairie Research Institute, University of Illinois at Urbana-Champaign, 1816 S. Oak St., Champaign, Illinois 61820, USA.

[2]Departamento de Biología Animal, Universidad de Granada, Granada, 18001, Spain.

[3]Department of Entomology, Purdue University, 901 W. State St., West Lafayette, Indiana 47907, USA.

*Correspondence to: jorged@illinois.edu, kpjohnso@illinois.edu.

Running title: Dispersal and introgression rate.

Author contributions: J.D., and K.P.J. conceived the study. J.D., A.D.S., and K.P.J. designed the study. A.D.S. collected the data. J.D. and A.D.S. analysed the data. K.P.J. obtained financial support for the project. J.D. wrote the manuscript and all authors contributed to editing the manuscript.

Acknowledgments: This study was supported by the US National Science Foundation (DEB-1239788, DEB-1342604, DEB-1926919, and DEB-1925487 to K.P.J). We declare that we have no conflict of interest.

Data availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Supplementary Materials: Supplementary data are available at GitHub (https://jorge-dona.github.io/Comparing-rates-of-introgression-in-parasitic-feather-lice-with-differing-dispersal-capabilities/supplementary.html) and at Figshare (DOI: 10.6084/m9.figshare.9176204). Individual gene trees are available at Figshare (DOI: 10.6084/m9.figshare.9176204).

Abstract: Organisms vary in their dispersal abilities, and these differences can have important biological consequences, such as impacting the likelihood of hybridization events. However, there is still much to learn about the factors influencing hybridization, and specifically how dispersal ability affects the opportunities for hybridization. Here, using the ecological replicate system of dove wing and body lice (Insecta: Phthiraptera), we show that species with higher dispersal abilities exhibited increased genomic signatures of introgression. Specifically, we found a higher proportion of introgressed genomic reads and more reticulated phylogenetic networks in wing lice, the louse group with higher dispersal abilities. Our results are consistent with the hypothesis that differences in dispersal ability might drive the extent of introgression through hybridization.

Introduction

Dispersal is the permanent movement of organisms away from their place of origin. It is a fundamental process in biology with significant implications at multiple scales of organization (Barton 1992; Clobert et al. 2001; Nathan 2001; Matthysen 2012), including the reproduction of individuals, the composition of populations and communities, and the geographical distribution of species (Clobert et al. 2001, 2012).

Organisms differ in their dispersal abilities, and these differences have an impact on their biology, such as on the distributional range of a species or gene flow between populations (Bohonak 1999). For example, organisms with lower dispersal abilities tend to have smaller distributional ranges and populations that are genetically more structured (Bohonak 1999; Dawson et al. 2014; DiBlasi et al. 2018).

Dispersal ability might also affect the opportunities for hybridization between species because the rates at which individuals encounter different species are likely to be higher in organisms with higher dispersal capabilities. Indeed, recent evidence supports this prediction by demonstrating that range expansion is associated with the extent of introgression (Currat et al. 2008; Nussberger et al. 2018). Similarly, dispersal differences explain more than 30% of the variation in the width of hybrid zones across animals (McEntee et al. 2018). However, overall there is still much to learn about the factors influencing hybridization (Randler 2006; Arnold 2015; Taylor and Larson 2019), and, in particular, the influence of dispersal ability on the rate of hybridization remains understudied.

Testing for the effect of dispersal on hybridization should ideally hold constant most factors other than dispersal. The ecological replicate system of wing and body lice (Insecta: Phthiraptera) of pigeons and doves (Aves: Columbidae) has proven to be an ideal system for

61 comparing the impact of dispersal differences on other aspects of biology, such as population
62 structure and codivergence (Clayton and Johnson 2003; Johnson and Clayton 2004; Clayton et
63 al. 2015; DiBlasi et al. 2018; Sweet and Johnson 2018). Specifically, this is an excellent system
64 in which to assess the effect of differences in dispersal capabilities on levels of introgression
65 because both of these lineages of feather lice: 1) significantly differ in their dispersal ability
66 (Harbison et al. 2008, 2009; Bartlow et al. 2016), 2) co-occur across the diversity of pigeons and
67 doves, and 3) have the same basic life history and diet (Clayton et al. 2015; Sweet and Johnson
68 2015, 2018). Both wing and body lice disperse vertically between parents and offspring in the
69 nest. However, wing lice can also attach to and hitchhike on hippoboscid flies to disperse
70 "phoretically" between host individuals or host species (Harbison et al. 2008, 2009; Bartlow et
71 al. 2016). Indeed, this hitch-hiking dispersal mechanism profoundly influences their degree of
72 population structure and cophylogenetic history (Clayton and Johnson 2003; Sweet et al. 2017b;
73 DiBlasi et al. 2018; Sweet and Johnson 2018). In addition, wing lice have a higher rate of host-
74 switching (Clayton and Johnson 2003; Clayton et al. 2015; Sweet et al. 2017b) (i.e., successful
75 colonization of new host species) and of straggling (Whiteman et al. 2004) (i.e., dispersal to new
76 host species without reproduction on that new host).

77 To compare differences in the extent of introgression between wing and body lice, we
78 used whole-genome data from 71 louse individuals belonging to five species of wing lice
79 (*Columbicola*) and seven species of body lice (*Physconelloides*) that occur across the same suite
80 of host species and have highly comparable patterns of diversification (Sweet and Johnson 2015,
81 2018). Specifically, both lineages within these two groups of lice that are the focus of this study
82 originated on the common ancestor of *Metriopelia* doves (11.3-14.9 mya) and have a correlated
83 pattern of codiversification within the same group of hosts (including a shared cospeciation event

which occurred within the *Metriopelia* genus 5.2-7.4 mya; Sweet and Johnson 2015, 2018). We predicted that wing lice, which have higher dispersal abilities and thus higher odds of encountering individuals of a different louse species on the same host, should show more extensive evidence of introgression (Fig. 1).

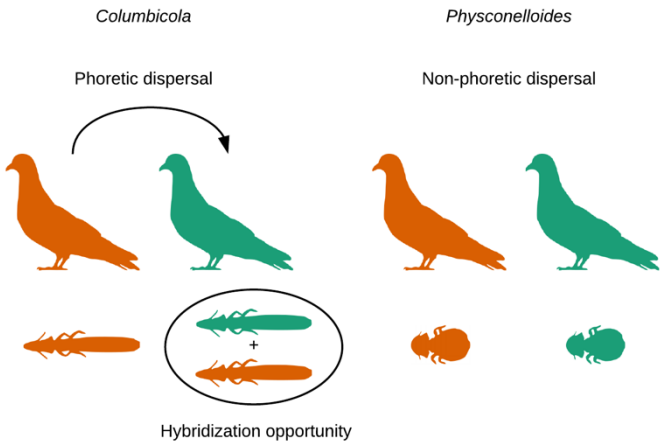


Figure 1. Diagram depicting the ecological replicate system and the hypothesis of this study. Wing lice (*Columbicola*) have higher dispersal abilities than body lice (*Physconelloides*), and thus higher odds of encountering individuals of a different louse sp.

Materials and Methods

Data

We analyzed Illumina whole genome sequence data (150 or 160bp paired-end reads) from 71 louse individuals belonging to five and seven taxa of *Columbicola* and *Physconelloides*, respectively (Supplementary Table S2) hosted by the monophyletic clade of small New World ground doves. This paper's taxonomic classification of lice is based on Sweet and Johnson (2018) species delimitation analyses. In particular, they found most *Columbicola* OTUs matched known species, and some *Physconelloides* OTUs were yet to be formally described as species

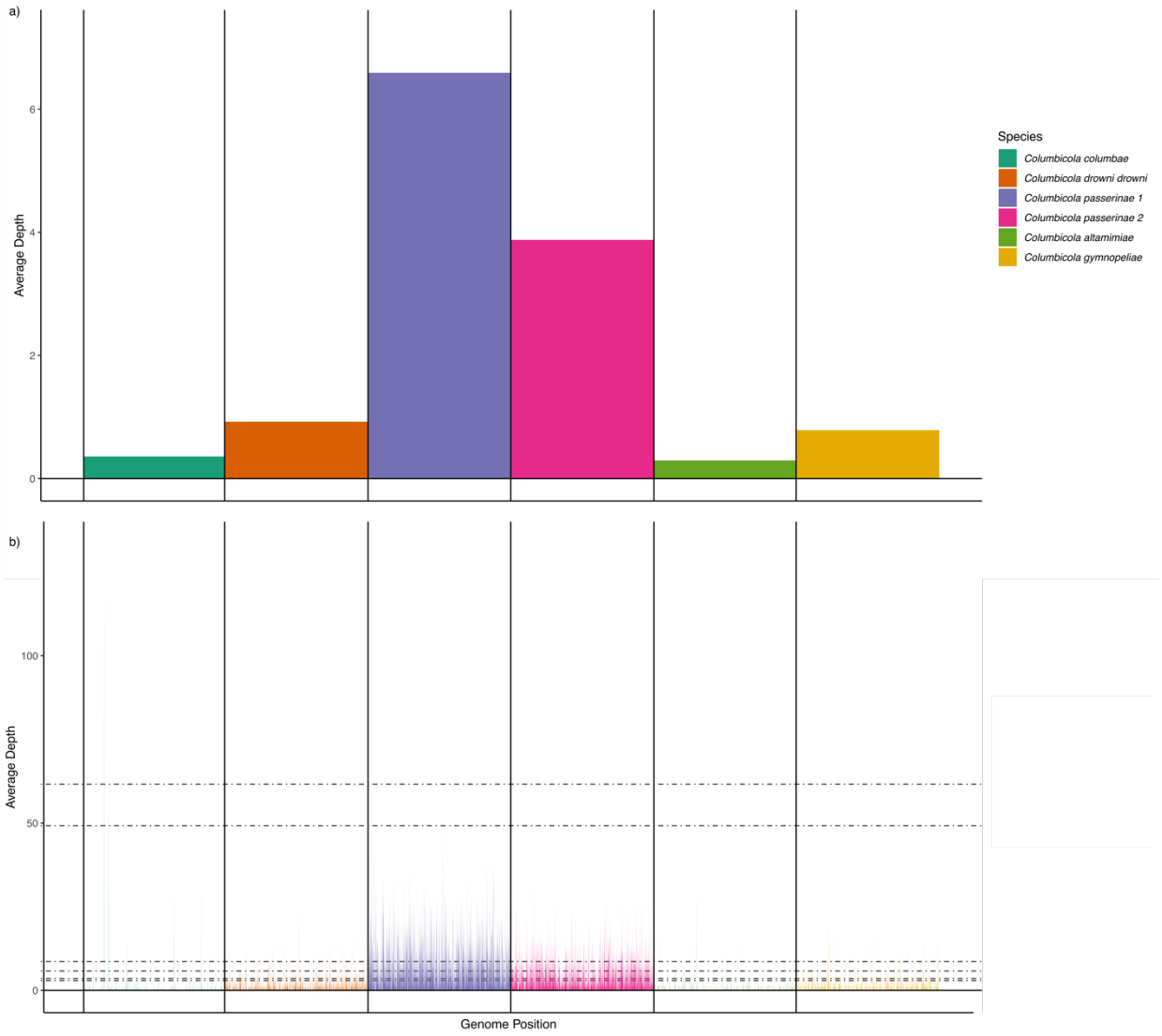
(and are named here following Sweet & Johnson, 2018). All raw sequence data used were available from previous studies (Boyd et al. 2017; Sweet et al. 2017a; Sweet and Johnson 2018) and represent all described New World ground-dove wing and body louse species (sampled from most host species in this group) including lice samples across multiple biogeographic areas within species of hosts (Sweet and Johnson 2018; Supplementary Table S2). Illumina genome sequence data pre-processing included several steps (Sweet and Johnson 2018). First, we discarded duplicate read pairs using the fastqSplitDups script (https://github.com/McIntyre-Lab/mcscriptand https://github.com/McIntyre-Lab/mclib). We then eliminated the Illumina sequencing adapters with Fastx_clipper v0.014 from the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit). Also, we removed the first 5 nt from the 5' ends of reads using Fastx_trimmer v0.014 and trimmed bases from the 3' ends of reads until reaching a base with a phred score ≥28 (which is equivalent to a base call accuracy higher than 99.8 %) using Fastq_quality_trimmer v0.014. Finally, we removed any reads less than 75 nt and analyzed the cleaned libraries with Fastqc v0.11.5 to check for additional errors. We assembled nuclear loci in aTRAM following previous studies (Allen et al. 2015; Boyd et al. 2017; Sweet and Johnson 2018). In particular, we mapped modest coverage (25-60X), multiplexed genomic data to reference loci from a closely related taxon. For our reference set of nuclear loci for wing lice, we used 1,039 exons of *Columbicola drowni* (Boyd et al. 2017) (raw data: SRR3161922). This data set was assembled de novo (Allen et al. 2015) using orthologous protein-coding genes from the human body louse genome (*Pediculus humanus humanus* (Kirkness et al. 2010) as a set of target sequences. We mapped *Columbicola* reads to the *C. drowni* references using Bowtie2 (Langmead and Salzberg 2012). For body lice, we obtained nuclear data using the same pipeline and software parameters, except that we used 1,095 loci from *Physconelloides emersoni* as the

reference for mapping. To generate the input ultrametric gene trees for Phylonet v3.6.8 (Than et al. 2008; Yu and Nakhleh 2015; Wen et al. 2018), we first aligned each nuclear locus in MAFFT (Katoh 2002) (--auto) and removed columns with only ambiguous sequences ("N"). Then, we estimated gene trees in RAxML v8.1.3 (Stamatakis 2006) with a GTR + Γ substitution model for each gene alignment. Finally, we made trees ultrametric using the nnls method in the force.ultrametric function within the "phytools" R package (Revell 2011).

Quantifying introgression

We used two different approaches to quantify differences in the extent of introgression (i.e., ancient plus recent) between the two louse genera. We employed methods suitable to detect introgression between species and between individuals from the same species (i.e., we did not employ methods aimed to detect differences at the population level, e.g., TreeMix; Pickrell and Pritchard 2012). First, we used sppIDer (Langdon et al. 2018) to quantify the genomic contributions of different louse species in an individual louse genome (Fig. 2). We built our reference for each genus using all the nuclear loci from a single individual per species. For the reference, we selected those individuals for which we assembled the highest number of genes for each genus. We estimated the extent of introgression as the sum of the mean coverages of reads mapped from all the species excluding the focal louse species, divided by the mean coverage of the focal louse species (Fig. 2). Note that these mean coverage values are calculated using only those reads that mapped with a mapping quality (MQ) > 3 (Li et al. 2009; Langdon et al. 2018; Figs. S1-S12). Plots of coverage across the genomes suggested that reads mapping to other species were not artificial mappings (e.g., high coverage mappings to short repetitive regions;

144 Figs. S1-S12). In addition, we used SPAdes v3.12.0 (default parameters) to perform a de-novo
145 assembly of the putatively introgressed reads detected by sppIDer).



146

147 **Figure 2**. Illustrative example of sppIDer results with an individual of *Columbicola passerinae* 1 (Cosp.Copas.11.7.2016.8).
148 Panel a) shows the average mean coverages of reads mapping to every species (i.e., the values that were used for the calculations
149 of introgression levels), and panel b) shows the mean coverages of reads mapping to each species across the whole set of loci.
150 These and additional visualizations of all the individuals can be found in the supplementary material (Figs. S1-S12).

151

152 Second, we quantified introgression at the species level, while accounting for incomplete lineage
153 sorting (ILS) using a maximum pseudo-likelihood framework with PhyloNet 3.6.1 (Than et al.
154 2008; Yu and Nakhleh 2015; Wen et al. 2018). Reticulations in this method can be attributed to

hybridization events. We trimmed the unrooted gene trees to the same individuals used as reference taxa in sppIDer, and performed eleven independent analyses with a differing maximum number of reticulation nodes (i.e., from zero to ten). We conducted ten runs per analysis. We then selected the optimal network for each genus based on AIC values.

Statistics and reproducibility

We compared the sppIDer results using generalized linear models (GLMs). We used a Gaussian distribution of errors and an identity link function. We performed one GLM for each simulation iteration using the glm function of the "stats" R package (R Core Team 2013). The extent of introgression for each louse genus was the dependent variable, the genus identity was the independent variable, and we accounted for the introgression differences between louse species including louse identity as a fixed factor. We confirmed assumptions underlying GLMs by testing the normality of regression residuals for normality against a Q-Q plot. We also considered the possibility that some of the reads mapping to other species were technical contaminations, i.e., due to index-swapping (Carlsen et al. 2012; Esling et al. 2015; Schnell et al. 2015; Sinha et al. 2017). Previous studies have found that the misassignment of reads generally ranges from 1 to 9% (Carlsen et al. 2012; Esling et al. 2015; Schnell et al. 2015; Sinha et al. 2017). Thus, to account for possible contaminants, we wrote a simulation in R that randomly subtracted 9% (Carlsen et al. 2012; Esling et al. 2015; Schnell et al. 2015; Sinha et al. 2017) from the mean coverage value of a particular sample (i.e., we subtracted a random proportion of the mean coverage value for each sample until reaching 9%). We ran 100 iterations of the simulation and ran a GLM for each iteration (Table S1). Finally, we used the χ^2 test to compare the number of species in pairwise comparisons of each genus with the number of reticulations found in each optimal phylogenetic network. Because we had an *a priori* prediction that

178 *Physconelloides* should exhibit less evidence of reticulation than *Columbicola*, we used a one-
179 tailed test; however, we also report the results of the two-tailed test equivalent.

180 **Results**

181 Both approaches revealed highly concordant results: higher levels of introgression among
182 species of wing lice compared to body lice. In particular, using a read-mapping based method,
183 the genomic signature of introgression was significantly higher in wing louse species than in
184 body louse species (GLM with the mean values of the simulations; $F = 21.0705$, $df = 69$, $P = 2.367 \times 10^{-5}$, $R^2=0.58$; Fig. 3, Supplementary Table S1, Figs. S1-S12). The contigs assembled
186 from reads mapping to the non-focal species were in the size range of the loci used as reference
187 (mean max contig length= 1214 bp; mean contig length=292 bp; Table S3). Even though wing
188 lice showed more evidence for introgression, one body louse individual (included in the GLMs)
189 exhibited the highest level of introgression (Fig. 3, Figs. S1-S12). However, the other individual
190 from the same taxon, inhabiting the same host species and collected in the same geographic
191 region, did not show these elevated levels of introgression (Supplementary Table S2, Figs. S1-
192 S12).

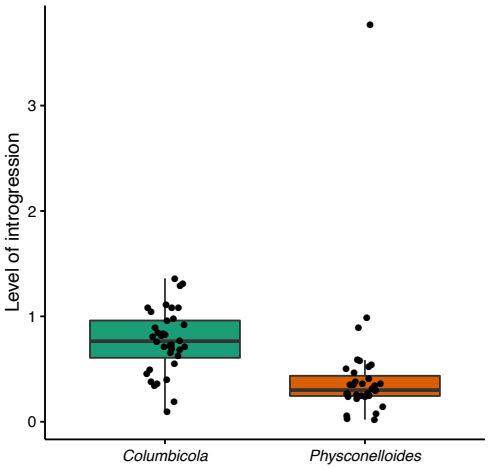
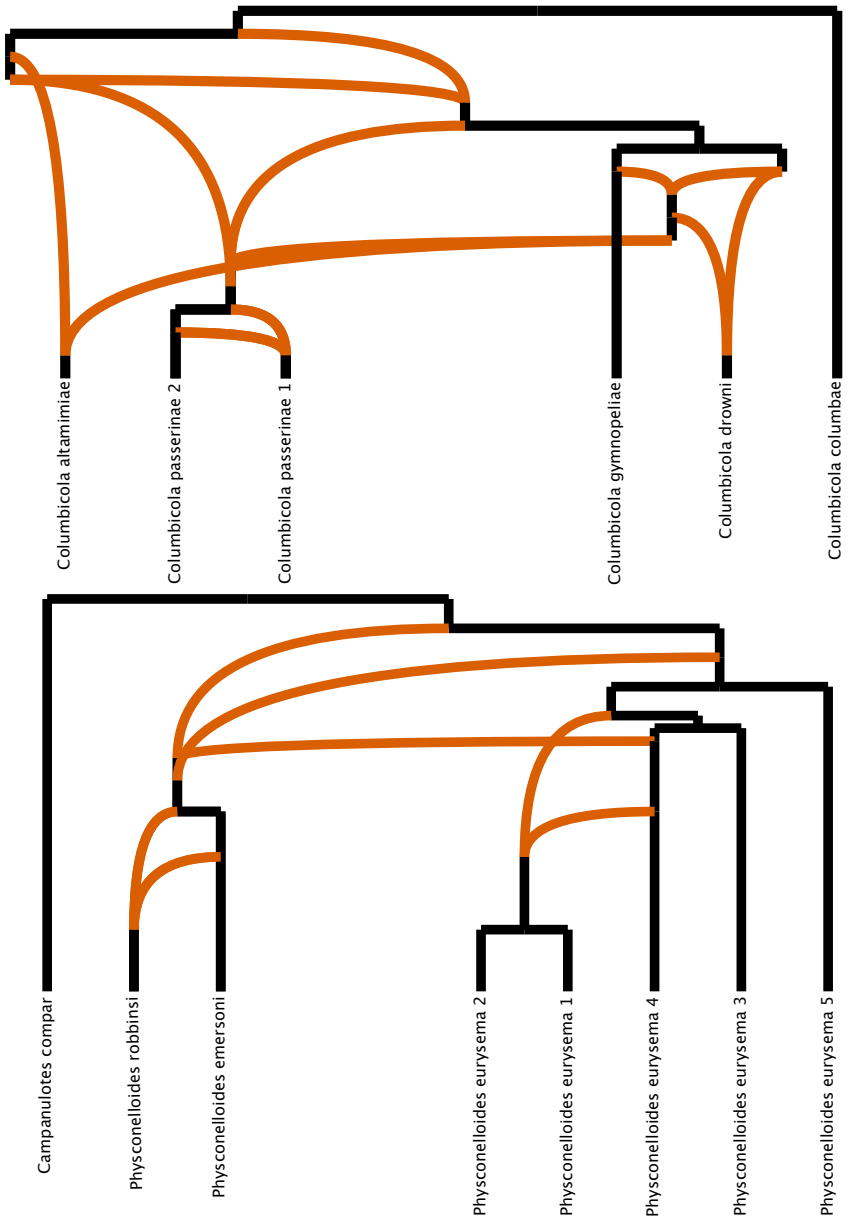


Figure 3. Boxplot showing the differences in levels of introgression between wing (green) and body (orange) lice. Level of introgression represents the sum of the mean coverage of reads mapped from all the species excluding the focal louse species, divided by the mean coverage of the focal louse species (see Methods). Black dots show the levels of introgression (i.e., resulting from the equation) for each individual sample (horizontally jittered values).

Secondly, in a phylogenetic network framework, the optimal networks of wing lice were more reticulated than those of body lice even though the number of taxa included in the networks was lower (seven reticulations in *Columbicola* vs. four in *Physconelloides*, Fig. 4). Accordingly, the number of reticulations given the number of potential combinations under a one-tailed test was significantly higher in *Columbicola* (One-sided: χ^2= 3.8132; df=1; *P*= 0.03; CI=0.03–1), and the p-value was still near 0.05 with a two-tailed test (Two-sided: χ^2= 3.8132; df=1; *P*= 0.05; CI=-0.01–0.66). Also, the specific lineages involved in the reticulations were generally congruent with signatures of introgression from the read-mapping based approach (Fig. S1-S12).



206

207 Figure 4. Optimal phylogenetic networks of feather lice genera. Orange branches depict reticulations: seven in *Columbicola* and

208 four in *Physconelloides*.

209 **Discussion**

210 Estimates of introgression in two groups of ectoparasites that differ in their dispersal abilities,

211 wing and body lice of doves, indicate that the lineage with higher dispersal ability (wing lice)

212 shows more evidence of introgression. This evidence from wing and body louse genomes is

consistent with the hypothesis that dispersal differences might drive differences in the level of introgression in this system of parasites. Admittedly, there may be some unknown factor, other than dispersal, differing between these two groups of lice that causes the difference in the level of introgression, but prior work on these groups of parasites points to dispersal as a crucial factor underlying many of the ecological and evolutionary patterns in these parasites. Further research on other taxa is needed to confirm the generality of these findings. This work is among the first studies of introgression in a host-symbiont system (Detwiler and Criscione 2010). Notably, recent studies have found that straggling and host-switching are relatively common processes in host-symbiont systems (De Vienne et al. 2013; Bourguignon et al. 2018; Nylin et al. 2018; Doña et al. 2019). Our study suggests that in a straggling/host-switching scenario, hybridization can provide further genetic variation with important ecological and evolutionary consequences (e.g., facilitating adaptation to current hosts or facilitating the colonization of new ones) (Barton 2018). Indeed, we may have found a potential recent hybridization event (i.e., the *Physconelloides* individual showing the highest level of introgression), though this requires further study to rule out methodological issues (e.g., wet-lab contamination).

In this vein, a careful examination of the introgression history of these taxa (and symbionts as a whole) is needed to better understand the patterns of introgression that we found. Questions such as how much introgression can be expected or how the introgressed regions are retained in parasite/symbiont genomes across time, among many others, require further attention. For instance, in this study, the levels of introgression detected by the sppIDer analyses (i.e., the magnitude but not the comparative pattern) may be unrealistic. It may be that some fraction of the level of introgression detected by sppIDer may be due to ILS, and not introgression. However, both louse genera are expected to have relatively similar rates of ILS (if any). It is also

236 possible that taxon age and interspecific divergence might affect introgression rates.
237 Nevertheless, Sweet and Johnson (2018) compared the degree of genetic divergence of two pairs
238 of species of both genera that inhabits the same host species and share a cospeciation event. In
239 this case of taxa of the same age, the pair of *Columbicola* species had lower interspecific genetic
240 distances than *Physconelloides*. This could be as a result of mutation rate differences between the
241 two genera. However, this could also be due to higher gene flow among host infrapopulations
242 due to higher dispersal capabilities of wing lice. The same pattern can be found overall across the
243 species studied here, i.e., on average, lower uncorrected interspecific genetic distances among
244 *Columbicola* than among *Physconelloides* species, though the range of the interspecific distances
245 does overlap (Table S4). Thus, if present, ILS could potentially be more prevalent in
246 *Columbicola* species (Pamilo & Nei, 1988; Maddison, 1997). The PhyloNet analysis, however,
247 does control for ILS, and showed highly congruent results. In addition, some individual gene
248 trees exhibit signatures suggestive of introgression with highly similar sequences shared by some
249 individuals of different species and much less likely to be a consequence of ILS (Figs. S13-S14).
250 Overall, the species of *Columbicola* and *Physconelloides* are from the same group of hosts and
251 thus are overall comparatively similar in levels of divergence, so it seems unlikely that these
252 small differences are driving the results.

253 Another caveat is that sppIDer can detect introgression from species that are not included
254 in the reference data. In those cases, the reads may map to the closest taxon available in the
255 reference set, and thus could artificially increase the level of introgression from a given species;
256 Langdon et al. 2018). Accordingly, the levels of introgression detected by sppIDer in certain
257 species could be an aggregate of introgression events from more than one species. Indeed, our
258 PhyloNet analysis supports this scenario, with several reticulations from ghost lineages and

259 species (Fig. 4). However, in this system we have nearly complete sampling of host taxa and are
260 missing few, if any, extant species making this concern less likely.

261

262 **References:**

263 Allen, J. M., D. I. Huang, Q. C. Cronk, and K. P. Johnson. 2015. aTRAM - automated target restricted assembly
264 method: a fast method for assembling loci across divergent taxa from next-generation sequencing data.
265 BMC Bioinformatics 16.
266 Arnold, M. L. 2015. Divergence with genetic exchange. OUP Oxford, Oxford, UK.
267 Bartlow, A. W., S. M. Villa, M. W. Thompson, and S. E. Bush. 2016. Walk or ride? Phoretic behaviour of
268 amblyceran and ischnoceran lice. Int J Parasitol 46:221–227.
269 Barton, N. H. 2018. The consequences of an introgression event. Mol Ecol 27:4973–4975.
270 Barton, N. H. 1992. The genetic consequences of dispersal. In: Animal Dispersal. Springer, Netherlands.
271 Bohonak, A. J. 1999. Dispersal Gene Flow, and Population Structure. Q Rev Biol 74:21–45.
272 Bourguignon, T., N. Lo, C. Dietrich, J. Šobotník, S. Sidek, Y. Roisin, A. Brune, and T. A. Evans. 2018. Rampant
273 host switching shaped the termite gut microbiome. Curr Biol 28:649–654.
274 Boyd BM, Allen JM, Nguyen N, Sweet AD, Warnow T, Shapiro MD, and et al. 2017. Phylogenomics using Target-
275 restricted Assembly Resolves Intra-generic Relationships of Parasitic Lice (Phthiraptera: Columbicola).
276 Syst Biol 66:896–911.
277 Carlsen, T., A. B. Aas, D. Lindner, T. Vrålstad, T. Schumacher, and H. Kauserud. 2012. Dont make a mista(g)ke: is
278 tag switching an overlooked source of error in amplicon pyrosequencing studies? Fungal Ecol 5:747–749.
279 Clayton, D. H., S. E. Bush, and K. P. Johnson. 2015. Coevolution of life on hosts: integrating ecology and history.
280 University of Chicago Press, Chicago, IL.
281 Clayton, D. H., and K. P. Johnson. 2003. Linking coevolutionary history to ecological process: doves and lice.
282 Evolution 57:2335–2341.
283 Clobert, J., M. Baguette, T. G. Benton, and J. M. Bullock (eds). 2012. Dispersal Ecology and Evolution. Oxford
284 University Press.
285 Clobert, J., E. Danchin, A. A. Dhondt, and N. J. D. 2001. Dispersal. Oxford Univ. Press.

286 Currat, M., M. Ruedi, R. J. Petit, and L. Excoffier. 2008. The hidden side of invasions: massive introgression by
287 local genes. Evolution 62:1908–1920.
288 Dawson, M. N., C. G. Hays, R. K. Grosberg, and P. T. Raimondi. 2014. Dispersal potential and population genetic
289 structure in the marine intertidal of the eastern North Pacific. Ecol Monogr 84:435–456.
290 De Vienne, D. M., G. Refrégier, M. López-Villavicencio, A. Tellier, M. E. Hood, and T. Giraud. 2013. Cospeciation
291 vs host-shift speciation: methods for testing, evidence from natural associations and relation to coevolution.
292 New Phytol 198:347–385.
293 Detwiler, J. T., and C. D. Criscione. 2010. An infectious topic in reticulate evolution: introgression and
294 hybridization in animal parasites. Genes (Basel) 1:102–23.
295 DiBlasi, E., K. P. Johnson, S. A. Stringham, A. N. Hansen, A. B. Beach, and D. H. Clayton. 2018. Phoretic dispersal
296 influences parasite population genetic structure. Mol Ecol 27:2770–2779.
297 Doña, J., D. Serrano, S. Mironov, A. Montesinos-Navarro, and R. Jovani. 2019. Unexpected bird–feather mite
298 associations revealed by DNA metabarcoding uncovers a dynamic ecoevolutionary scenario. Mol Ecol
299 28:379–390.
300 Esling, P., F. Lejzerowicz, and J. Pawlowski. 2015. Accurate multiplexing and filtering for high-throughput
301 amplicon-sequencing. Nucleic Acids Res 43:2513–24.
302 Folk, R. A., P. S. Soltis, D. E. Soltis, and R. Guralnick. 2018. New prospects in the detection and comparative
303 analysis of hybridization in the tree of life. Am J Bot 105:364–375.
304 Harbison, C. W., S. E. Bush, J. R. Malenke, and D. H. Clayton. 2008. Comparative transmission dynamics of
305 competing parasite species. Ecology 89:3186–3194.
306 Harbison, C. W., M. V. Jacobsen, and D. H. Clayton. 2009. A hitchhiker's guide to parasite transmission: The
307 phoretic behaviour of feather lice. Int J Parasitol 39:569–575.
308 Johnson, K. P., and D. H. Clayton. 2004. Untangling coevolutionary history. Syst Biol 53:92–4.
309 Katoh, K. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform.
310 Nucleic Acids Res 30:3059–3066.
311 Kirkness, E. F., B. J. Haas, W. Sun, H. R. Braig, M. A. Perotti, and J. M. Clark. 2010. Genome sequences of the
312 human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle.
313 Proc Natl Acad Sci U S A 107:12168–73.

314 Langdon, Q. K., D. Peris, B. Kyle, and C. T. Hittinger. 2018. sppIDer: A Species Identification Tool to Investigate
315 Hybrid Genomes with High-Throughput Sequencing. Mol Biol Evol 35:2835–2849.
316 Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–9.
317 Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. The
318 sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. Oxford University Press.
319 Maddison, W. P. 1997. Gene trees in species trees. Syst Biol, 46(3), 523-536.
320 Matthysen, E. 2012. Multicausality of dispersal: a review. In: Dispersal Ecology and Evolution. Oxford University
321 Press.
322 McEntee, J. P., J. G. Burleigh, and S. Singhal. 2018. Dispersal predicts hybrid zone widths across animal diversity:
323 Implications for species borders under incomplete reproductive isolation. bioRxiv, doi: 10.1101/472506.
324 Nathan, R. 2001. The challenges of studying dispersal. Trends Ecol Evol 16:481–483.
325 Nussberger, B., M. Currat, C. S. Quilodran, N. Ponta, and L. F. Keller. 2018. Range expansion as an explanation for
326 introgression in European wildcats. Biol Conserv 218:49–56.
327 Nylin, S., S. Agosta, S. Bensch, W. A. Boeger, M. P. Braga, D. R. Brooks, M. L. Forister, P. A. Hambäck, E. P.
328 Hoberg, and T. Nyman. 2018. Embracing colonizations: a new paradigm for species association dynamics.
329 Trends Ecol Evol 33:4–14.
330 Pamilo, P., & Nei, M. 1988. Relationships between gene trees and species trees. Mol Biol Evol, 5(5), 568-583.
331 Pickrell, J. K., and J. K. Pritchard. 2012. Inference of Population Splits and Mixtures from Genome-Wide Allele
332 Frequency Data. PLoS Genetics 8.
333 R Core Team. 2013. R: A language and environment for statistical computing.
334 Randler, C. 2006. Behavioural and ecological correlates of natural hybridization in birds. Ibis 148:459–467.
335 Revell, L. 2011. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol
336 Evol 3:217–223.
337 Schnell, I. B., K. Bohmann, and M. T. Gilbert. 2015. Tag jumps illuminated–reducing sequence-to-sample
338 misidentifications in metabarcoding studies. Mol Ecol Resour 15:1289–303.
339 Sinha, R., G. Stanley, G. S. Gulati, C. Ezran, K. J. Travaglini, and E. Wei. 2017. Index Switching Causes Spreading-
340 Of-Signal Among. bioRxiv, doi: 10.1101/125724.

341 Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa
342 and mixed models. Bioinformatics 22:2688–2690.
343
344 Sweet, A. D., B. M. Boyd, J. M. Allen, S. M. Villa, M. P. Valim, and J. L. Rivera-Parra. 2017a. Integrating
345 phylogenomic and population genomic patterns in avian lice provides a more complete picture of parasite
346 evolution. Evolution 72:95–112.
347 Sweet, A. D., R. T. Chesser, and K. P. Johnson. 2017b. Comparative cophylogenetics of Australian phabine pigeons
348 and doves (Aves: Columbidae) and their feather lice (Insecta: Phthiraptera). Int J Parasitol 47:347–356.
349 Sweet, A. D., and K. P. Johnson. 2015. Patterns of diversification in small New World ground doves are consistent
350 with major geologic events. Auk 132:300–312.
351 Sweet, A. D., and K. P. Johnson. 2018. The role of parasite dispersal in shaping a host–parasite system at multiple
352 evolutionary scales. Mol Ecol 27:5104–5119.
353 Taylor, S. A., and E. L. Larson. 2019. Insights from genomes into the evolutionary importance and prevalence of
354 hybridization in nature. Nat Ecol Evol 3:170.
355 Than, C., D. Ruths, and L. Nakhleh. 2008. PhyloNet: a software package for analyzing and reconstructing reticulate
356 evolutionary relationships. BMC Bioinformatics 9:322.
357 Wen, D., Y. Yu, J. Zhu, and L. Nakhleh. 2018. Inferring Phylogenetic Networks Using PhyloNet. Syst Biol 67:735–
358 740.
359 Whiteman, N. K., D. Santiago-Alarcon, K. P. Johnson, and P. G. Parker. 2004. Differences in straggling rates
360 between two genera of dove lice (Insecta: Phthiraptera) reinforce population genetic and cophylogenetic
361 patterns. Int J Parasitol 34:1113–1119.
362 Yu, Y., and L. Nakhleh. 2015. A maximum pseudo-likelihood approach for phylogenetic networks. BMC Genomics
363 16:Suppl 10: S10.
364
365