DEMOGRAPHIC HISTORY AND POPULATION STRUCTURE OF POLISTES FUSCATUS PAPER WASPS

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Sarah Elena Bluher

August 2018

ABSTRACT

Limited dispersal and cooperation are predicted to coevolve in cooperatively breeding systems [24,33,49,50]. In *Polistes* paper wasps, mark-recapture studies and genetic analyses have shown that both cooperative and solitary nesting gynes can exhibit natal philopatry [4,18,23,31,38]. Male dispersal has not been studied in *Polistes*, therefore the potential effects of sex-biased dispersal on population structure are unknown. In this study I used whole genome resequencing data from *P. fuscatus* wasps collected across the Eastern US to assess demographic history and population structure within a well-studied cooperatively breeding system. In order to deduce the influence of sex-specific dispersal patterns on overall population structure, I analyzed population structure using nuclear as well as mitochondrial genomes. While low pairwise F_{ST} values indicated overall depressed levels of genetic differentiation, I found a consistent and significant pattern of spatial genetic clustering and isolation by distance in both the nuclear and mitochondrial genomes, supporting the hypothesis that limited female dispersal contributes towards population structure in *P. fuscatus* wasps.

BIOGRAPHICAL SKETCH

Sarah Bluher was raised in Maryland by a couple of mathematicians. She was first introduced to Biology research in 2008, as a high school intern at the Walter Reed Army Institute of Research in Silver Spring, researching the effects of Army Ranger training on immune response gene suppression. Sarah received a Bachelor of Science with a major in Geosciences from Princeton University in June 2014. She received the Geosciences Department Chairman's Award for her Senior Thesis research on reconstructing the ancient carbon cycle through stable isotope analysis. After graduating from Princeton she worked as a honey bee researcher with the US Department of Agriculture, before joining the Cornell Department of Neurobiology and Behavior in 2016. Sarah has enjoyed her time at Cornell and plans to pursue further education in Applied Statistics to expand her data analysis toolkit.

This document is dedicated to Froim Blyukher.

ACKNOWLEDGEMENTS

I would like to thank the Sheehan Lab and the Department of Neurobiology and Behavior for their support and contribution to this study. A special thank you to Sara Miller for her bioinformatics mentorship. I would also like to thank the Neurobiology and Behavior administrative staff for making the department a great place to work. Last but not least, I would like to thank my family and friends for their support. I have grown personally during my time in Ithaca and I have many to thank for that.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

0.1 Introduction

0.1.1 Cooperation and Dispersal

One of the key predictions arising from population-structured models of cooperation is that limited dispersal should facilitate the evolution of cooperative behavior [20, 24, 33, 44, 49]. The link between cooperation and dispersal was first proposed by Hamilton; under kin selection theory, limited dispersal should promote cooperation by increasing relatedness between potential beneficiaries within kin neighborhoods [20]. More recent models that take into account the costs of competition in kin neighborhoods nonetheless predict the coevolution of cooperation and dispersal [33, 44, 49].

Limited dispersal is favored in social systems when the benefits of cooperation and reduced dispersal risk outweigh the costs of inbreeding, kin competition and missed opportunities for independent reproduction [49]. Theoretical models predict the coevoution of cooperation and dispersal in a positive feedback, in which cooperative social systems favor limited dispersal, and the population viscosity resulting from limited dispersal makes cooperation more advantageous [24, 33]. More general models predict the evolution of linkage disequilibrium between cooperative traits and any traits that serve to enhance population structure [50].

Primitively eusocial wasps have long been a model system for studies of cooperative behavior. In the temperate zone, *Polistes* nests are initiated each spring by either a single foundress or a small group of cooperating foundresses [14, 17, 19, 30, 41, 43]. Individuals that attempt but fail to found nests independently often join associations later in the season [3, 37, 81]. Within cooperative associations, reproductive shares follow a linear dominance heirarchy, in which the dominant foundress lays the majority of eggs and often consumes those laid by her subordinates [26, 51, 57, 68, 77]. Subordinate foundresses help to provision the offspring of the dominant, contributing disproportionately to risky foraging efforts [12, 17].

Studies across cooperatively breeding taxa report philopatry [21, 53, 61, 63], fine-scale genetic structure [8, 46, 54, 67] and isolation by distance [7, 69], however similar evidence for philopatry and genetic structure has also been found in non-cooperative taxa [2, 47, 48, 64].

The relationship between cooperative breeding, dispersal and population structure may be expected to differ between cooperatively breeding systems. For example, in many cooperatively breeding species of birds, helpers are the offspring of the recipient [22]. The mechanism for generating greater population structure in these cases is delayed dispersal. Offspring that help on the nest may choose to disperse eventually, but they stay on nest throughout more of the breeding season to help rear their siblings [22]. In contrast, *Polistes* helpers (subordinates) disperse prior to the nesting season. The question for *Polistes* foundresses is therefore not whether or not to disperse, but how far to disperse. *Polistes* gynes undergo two predicatbly-timed dispersal events. One occurs in the fall to mate, the second in the spring to nest. Because there are two discrete decision periods, the link between dispersal and population structure is not as straight-forward as in taxa that make single, binary dispersal decisions. For instance, even if *Polistes* gynes exhibit a high degree of philopatry during the spring dispersal to nest, they may be dispersing longer distances to mate, thereby contributing to greater gene flow.

Mark-recapture studies and genetic analyses of foundress associations have shown that cooperating wasps tend to be philopatric and preferentially join former nestmates, with an average relatedness $r > 0.6$ [4, 18, 31]. In a study of *P. fuscatus* nesting behavior, Klahn found that individuals within larger foundress associations tended to be located closer to their natal nest sites, when compared to their less cooperative counterparts [31]. In spite of the evidence for natal philopatry in cooperative species, the relationship between cooperation and population structure in *Polistes* wasps is contentious. Lengronne *et al.* found similar degrees of fine-scale genetic structure in populations of *P. canadensis* and *P. dominula*, which vary significantly in their rates and degrees of cooperative founding [34]. Even species of *Polistes* that are exclusively solitary-founding can show high degrees of female philopatry [23, 38]. Thus, the relationship between cooperation and dispersal in *Polistes* is unknown and requires further study.

0.1.2 Population Structure

Population genetic structure is shaped by selection, genetic drift and gene flow. Localized directional selection and random genetic drift cause divergence between subpopulations, while stabilizing selection and migration between subpopulations induce homogeneity.

Wright developed a method of inferring population structure based on the hierarchical distribution of genetic variance within and between various defined groups [79]. At the scale of a subpopulation, fine-scale genetic structure can be inferred from the value of the inbreeding coefficient, F_{IS}, a measure of the degree of genetic variance within individuals relative to their subpopulation. F_{IS} is used to infer the extent to which individuals within a subpopulation assortatively mate with closer relatives, yielding higher than expected levels of homozygosity in individuals relative to the subpopulation. The statistic F_{ST}, and its haploid counterpart Φ_{ST}, provide a measure of diffferentiation between subpopulations, or the genetic varaince between subpopulations relative to the total variance.

Demographic fluctuations alter the rate and magnitude of genetic drift and the genetic diversity available for selection, and therefore exert considerable influence on resultant patterns of population structure. Local population bottlenecks increase the influence of local genetic drift and lead to increased divergence between subpopulations. Rapid population expansion leads to greater homogeneity across subpopulations emitting from recent shared ancestors, yilelding lower genetic diversity and decreased selection potential. In order to properly compare population structure across disparate populations and species, one must take into account the effects of demographic history on shaping local population structure.

Isolation by Distance

Sewall Wright coined the term ″isolation by distance″ (IBD) to describe the pattern of spatial genetic differentiation that emerges due to locally limited dispersal at drift-dispersal equilibrium [78]. Under IBD, local random genetic drift isolates subpopulations along a gradient of geographic distance due to restricted gene flow. Non-equilibrium models [25,66] have since expanded upon the earlier theoretical foundations of Wright and others [29,39,40]. All models of IBD predict a correlation between genetic and geographic distances arising from spatial limits to gene flow [59]. Empirical studies across several systems have revealed correlations between genetic and geographic distance consistent with model predictions [1,76].

Rousset devised a method to quantify the extent of IBD in order to infer dispersal rates [59]. By Rousset's method, average dispersal rates are inferred using the slope of the linear matrix regression of pairwise F_{ST} / (1 - F_{ST}) on geographic distance, in the one-dimensional case, and the natural logarithm of geographic distance, in the two-dimensional case. Rousset determined that the one-dimensional regression is appropriate for analyzing data along elongated transects, and yields the following predicted relationship between the mean squared parent-offspring axial dispersal distance (σ^2) and the regression slope (b):

$$\sigma^2 = \frac{1}{4\pi D_e b} \tag{1}$$

where D_e is the effective population density of breeders. Dispersal reconstructions calculated using the above equation should be interpreted with caution, however, as they are based on a simple model which assumes constant mutation rate, migration rate, and effective population density through time.

Sex-biased Dispersal

Due to the restriction of reproduction to specialized breeder castes, social insects are especially vulnerable to the costs of limited dispersal wrought by inbreeding depression [5]. In haplodiploid Hymenoptera, inbreeding leads to the production of infertile diploid males. Sex-biased dispersal can, however, mitigate these costs. In certain species of dependent colony founding social insects, increased male dispersal compensates for limited female dispersal [10]. Under haplodiploidy, male-biased dispersal can spur the evolution of reproductive altruism by increasing average relatedness of female helpers to their beneficiaries [28].

Male-biased dispersal and female philopatry result in a genetic signature of greater population viscosity in maternally-inherited mitochondrial genes than in the nuclear genome [52]. Studies of mitochondrial and nuclear population structure have revealed male-biased gene flow patterns in several species of eusocial Hymenoptera [6,13,58,60,70].

In order to assess differential contribution toward gene flow by males and females, workers compare the nuclear and mitochondrial genetic divergence statistics F_{ST} and Φ_{ST} [9,73,74]. Under equal dispersal by both sexes, differences in effective population size of the nuclear versus mitochondrial genome yield the following expected relationship between Φ_{ST} and F_{ST} [9]:

$$\Phi_{ST} = \frac{4F_{\mathrm{ST}}}{1 + 3F_{\mathrm{ST}}} \tag{2}$$

Given observed F_{ST} values, the above equation can be used to calculate expected Φ_{ST} values under equal male and female dispersal. If observed Φ_{ST} values are higher than those expected, the nuclear and mitochondrial data lend support towards a male-biased dispersal regime.

Ref.	Species	f_{avg}	Loc.	r_{P} (km)	r_{SP} (km)	Month(s)	Loci	$\mathrm{N_s}$	$\mathrm{N_n}$	$\mathrm{N_i}$	F_{IS}	F_{ST}	Φ_{ST}	IBD
[11]	*P. exclamans*	1.2	*SE.* US	5	0.25	Aug, Mar	6P, 1RS	5	84	315	0.19	0.77		
[11]	*P. exclamans*	1.2	*SE.* US	5	0.25	Mar	2P	5	49	188	0.52	0.47		
[11]	*P. metricus*	1.1	*SE.* US	5	0.25	Aug	6P, 1RS	5	22	82	0.22	-0.04		
[11]	*P. bellicosus*	1.6	*SE.* US	5	0.25	Aug, Mar	3P	5	60	206	0.25	-0.04		
[11]	*P. bellicosus*	1.6	*SE.* US	5	0.25	Mar	3P	5	38	116	0.21	-0.01		
[11]	*P. carolina*	2.1	*SE.* US	5	0.25	Aug	6P, 1RS	5	13	52	-0.12	0.03		
[72]	*P. jadwigae*	1.0	C. Japan	1	1	July-Aug	1P	1	67	941	0.04			
[45]	*P. chinensis*	1.3	C. Japan	7.5	7.5	June	3P	1	20	NR	-0.17			
[62]	*P. biglumis*	1.0	S. France	2.5	2.5	June-July	6M	1	15	160	0.06			
[62]	*P. biglumis*	1.0	*N.* Italy	0.5	0.5	June-July	6M	1	9	205	0.13			
[34]	*P. canadensis*	9.3	Panama	0.3	0.3	July	7M	1	26	129	0.03			NS
[34]	*P. dominula*	2.8	Spain	0.2	0.2	Mar	7M	1	26	129	0.04			S
[27]	*P. dominula*	2.8	*NE.* US	224	NR	Aug	12M	4	79	79	0.01	0.110		NS
[73]	*P. chinensis*	1.3	Japan	500	0.1	June-Aug	10M, COI	10	403	403		0.145	0.27	S
[73]	*P. chinensis*	1.3	NI, NZ	400	0.1	Dec-Feb	10M, COI	10	397	397		0.057	0.13	S
[74]	*P. olivaceus*	2.1	BGD	150	NR	Sep	7M, COI	14	208	208	0.006	0.003	0.22	NS
[32]	*P. nimpha*	1.3	POL	0.5	0.5	June	8M	1	15	59				S

Table 1: Previous studies on population structure in *Polistes* wasps. Mean foundress number (f_{avg}) as reported in Miller *et al.* [43], location (Loc.), radius of total study population in kilometers (r_{P}), average radius of defined subpopulations in kilometers (r_{SP}), month(s) sampled, types and numbers of genetic loci used (P = protein-coding, RS = ribosomal DNA restriction site, M = microsatellite, COI = mitochondrial gene), number of subpopulations ($\mathrm{N_s}$), number of nests ($\mathrm{N_n}$), number of individuals ($\mathrm{N_i}$), mitochondrial nucleotide diversity (π_{mt}), inbreeding coefficient (F_{IS}), fixation index (F_{ST}), mitochondrial fixation index (Φ_{ST}), and detection of IBD (NS = no signifcant IBD, S = significant IBD).

Genetic Structure in *Polistes* Wasps

The extent of inbreeding found in *Polistes* populations studied to date has been variable (Table 1). Davis *et al.* found inbreeding coefficients significantly higher than zero in populations of *P. exclamans*, *P. metricus* and *P. bellicosus* [11]. In contrast, studies of *P. carolina*, *P. jadwigae*, *P. chinensis*, *P. biglumis*, *P. canadensis* and *P. dominula* did not detect any significant genetic signature of inbreeding [11, 27, 34, 45, 62, 72]. Given that the value of the inbreeding coefficient F_{IS} depends on how authors define subpopulations, comparisons across studies must take into account differences in the spatial scales employed. Assuming equal degrees of population structure, larger defined subpopulation scales are expected to yield higher values of F_{IS}, due to fine-scale structure nested within the broadly defined subpopulations. Surprisingly, however, the studies reporting lower inbreeding coefficients defined subpopulations at scales ranging from 2 to 30 times larger than those defined in Davis *et al.* Variation in reported inbreeding coefficients may therefore reflect true differences in assortative mating and population structure between the studied species. However, given the restriction of positive F_{IS} results to one study, further investigation is warranted to build capacity for cross-species comparisons.

Mean pairwise genetic differentiation between subpopulations, as measured by F_{ST}, was found to be significantly higher than zero in studied populations of *P. exclamans*, *P. dominula* and *P. chinensis*. Excluding the high outlier values of F_{ST} found in *P. exclamans* populations, the mean F_{ST} across studies was 0.04, indicating overall low levels of divergence. No population structure was detected in *P. metricus*, *P. bellicosus*, *P. carolina*, or *P. olivaceus* across the areas studied. Under isolation by distance, larger distances between studied subpopulations are predicted to yield higher values of F_{ST}, therefore it is again important to consider the scale of total study areas. Across studies of *Polistes* population structure to date, there does not appear to be a relationship between reported values of F_{ST} and the scale at which total populations were defined, again indicating that reported variation in F_{ST} may reflect true differences in population viscosity between studied populations and species.

Of the six studies that have investigated isolation by distance to date, half have detected significant patterns of IBD (in *P. dominula*, *P. chinensis* and *P. nimpha*), while the others detected no significant IBD (in *P. dominula*, *P. canadensis*,and *P. olivaceus*) [27, 32, 34, 73, 74]. The IBD findings for *P. dominula* differ between the two studies reported, which also vary dramatically in scale. In the small-scale study ($r_P = 0.2km$), significant IBD was detected [34], whereas in the large-scale study ($r_p = 240km$), there was no significant IBD [27]. Notably, however, significant IBD patterns across all species were detected across extremely disparate scales, ranging from a radius of 0.2 to 500 kilometers. The ranges of F_{ST} values reported in studies that found significant versus no significant IBD are similarly low ($0.003 < F_{\mathrm{ST}} < 0.145$), indicating that patterns of IBD in *Polistes* can arise under low to moderate conditions of overall subpopulation divergence.

Only one study of *Polistes* has attempted to reconstruct dispersal using population genetic data. Applying Rousset's method (Equation 1) and using field-based estimates of breeder population density (D_e), Tsuchida *et al.* inferred mean axial parent-offspring dispersal distances of 33-60 meters for *P. chinensis* populations in Japan and New Zealand (Table 2).

Few studies have investigated genetic signatures of sex-biased gene flow in

Ref.	Species	Loc.	b	P	D_e (nests/ km^2)	$\sigma(m)$
[73]	*P. chinensis*	Japan	0.029	< 0.01	$2,865$	33
[73]	*P. chinensis*	NI, NZ	0.010	0.04	$2,317$	60

Table 2: Applying Rousset's method (Equation 1) and using field-based estimates of breeder population density (D_e), Tsuchida *et al.* inferred mean axial parent-offspring dispersal distances of 33-60 meters for *P. chinensis* populations in Japan and New Zealand.

Polistes wasps. One study of population structure in *P. olivaceus* found average pairwise F_{ST} values to be lower than effective-population-size-corrected Φ_{ST} estimated from mitochondrial haplotype sequences, supporting male-biased gene flow [74]. However, applying the same analysis (using equation 2) to results from Tsuchida *et al.* yields expected values of Φ_{ST} comparable with those observed, suggesting equal dispersal of the sexes.

In the present study, I infer sex-specific dispersal patterns through examination of population structure in cooperatively breeding *P. fuscatus* wasps. Reconstructed demographic history serves as a grounding from which to interpret population genetic statistics and evidence for population structure in *P. fuscatus* across the Eastern US.

0.2 Methods

0.2.1 Sampling and Sequencing

Female *P. fuscatus* individuals were collected both from nests and on the wing in Northern and Central New York, Massachusetts, Maryland and North Carolina. DNA was extracted from a single leg for each individual using the Qiagen

Puregene Core Kit A. DNA was sheared using Covaris S2 Adaptive Focused Acoustic Disruptor (Covaris, Inc.). We prepared 550 bp insert libraries with the Nextera library preparation kit. Library sizes were quantified prior to sequencing using a bioanalyzer. Libraries were sequenced on the Illumina HiSeq at Novogene (Davis, CA). Average coverage of re-sequenced genomes was 9.6x.

Paired-end reads were first processed with Trimmomatic (v0.36) to remove adaptors and poor-quality sequence. Trimmed reads for all species were mapped to the P. fuscatus reference genome using the Burrows-Wheeler Aligner (v0.7.13) [35]. SNPs were identified using Picard (v2.8.2) and GATK (v3.8) HaplotypeCaller following the best practice recommendations [75]. After alignment, SNPs were hard filtered using the parameters: FS > 60 || SOR > 3.0 || MQ < 40.0.

To avoid confounding the effects of nestmate relatedness with patterns of broad population structure, only one individual per nest was sequenced. We further filtered sequenced individuals by removing one individual from each pair with kinship coefficients (vcftools relatedness2) exceeding 0.1. The full set of available filtered samples was then downsampled to yield an even distribution of samples across all sites and regions. A filtered and downsampled set of 63 *P. fuscatus* females was used for all population genetic and population structure analyses. Demographic analyses were conducted using the same downsampled set of 63 samples, with the exception of the Central New York region, for which a larger available set of 170 samples was used.

0.2.2 Demographic History

Demographic history was reconstructed using the open source software SMC++, which takes as input unphased whole-genome sequencing data and generates population size history up to one thousand years before present using a sequential Markov coalescent model [71]. The SMC++ model was run with unphased whole-genome sequences of five to eight *P. fuscatus* females from each of the regionally-defined subpopulations across the Eastern US. A total of 170 Central NY samples were randomly resampled in groups of seven for twenty model runs to generate a 95% CI for regional analyses. The downsampled continental set of 63 samples was randomly resampled in groups of 40 for five model runs to generate a 95% CI for the complete data set. We assumed a generation time of one year and a mutation rate of 6.8e-9 per diploid genome per generation, as found in *Bombus terrestris* [36] and *Apis mellifera* [80].

0.2.3 Population Genetics

Mitochondrial gene nucleotide diversity (π_{mt}), mean whole-genome nucleotide diversity (π), and inbreeding coefficients (F_{IS}) were calculated using vcftools for the entire Eastern US population and each of the regionally-defined subpopulations. Pairwise nuclear divergence (F_{ST}) and mitochondrial divergence (Φ_{ST}) between subpopulations were calculated using vcftools and ARLEQUIN, respectively. All mitochondrial analyses were conducted using concatenated mitochondrial sequences from the COI, COII, ATP6, ATP8, ND4, ND5, 12S, 16S and CytB genes. Workers using a similar set of mitochondrial genes in *Apis mellifera* showed that concatenated mitochondrial genes are as effective as whole mitochondrial genome sequences for detecting population structure [15].

0.2.4 Population Structure

Population clustering at continental, regional and site scales were visualized using plink multidimensional scaling (MDS) plots. Structure at the continental scale was further analyzed using the software fastSTRUCTURE to visualize clustering and determine the optimal number of demes, K, based on model runs with K varying from 1 to 6 [55]. IBD was assessed for both the nuclear genome and mitochondrial genes by linear regression of pairwise geographic distance between regions versus pairwise linearized F_{ST} and Φ_{ST}, respectively. The strength and significance of observed patterns of IBD were analyzed using Mantel tests with the maximum potential 719 permutations.

0.3 Results

0.3.1 Demographic History

The SMC++ estimated population size histories for all regions across the Eastern US show similar patterns of expansion and contraction (Figure 1). The broad 95% CI for Central NY reflects variation due to the small sample size of seven samples per run.

Model outputs for all regions suggest a period of population expansion from 400,000 to 200,000 years before present, followed by a contraction from 200,000 to 100,000 years ago. The 300,000 year time span of this initial expansion and contraction encompasses three ice ages, visually identifiable as global temperature minima. The demographic history in this time period does not seem to correspond to periods of glacial expansion or retreat. A second, more gradual period of population contraction coincides with the period leading up to and following the Last Glacial Maximum in the Northeast US region from 21,750 to 13,000 years before present [42, 56, 65]. Population decline continued at a relatively steady rate from 100,000 to 5,000 years before present, at which point the effective population size was an estimated 175,000 individuals.

Between 5,000 and 1,000 years before present, certain regions' populations appear to have remained relatively stable or declined while others increased. Results from the models run with all samples across the Eastern US point towards a population increase from 175,000 to 240,000 individuals. Tellingly, the Central NY model run with n=16 samples shows a population increase over this time span, in contrast to the stability observed in models of the same region with n=8. It therefore seems probable that the apparent population stability in some regions is an artifact of the small sample sizes available, and that with greater sample sizes we might observe the same pattern of recent expansion in all regions.

Coalescence-based demographic models do not accurately reconstruct history more recent than one thousand years, therefore we cannot exclude the possibility of further population contraction or expansion from one thousand years ago to the present [71].

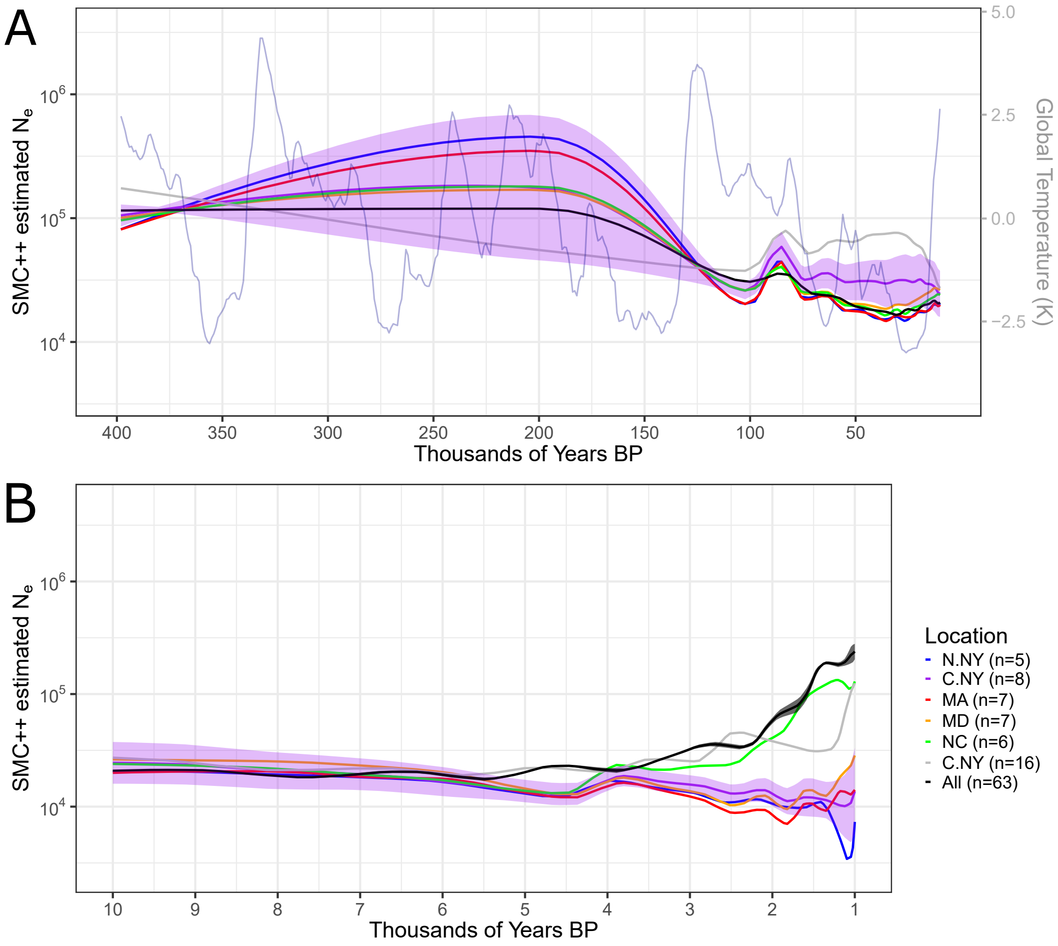

Figure 1: Demographic history reconstructed using SMC++ for each Eastern US region and across all regions shown for the distant (A) and recent past (B). The 95% CI for Central NY effective population size across twenty model runs is shown in shaded purple. The 95% CI across five model runs using 40-sample subsets of the full sample set is shown in shaded dark gray. The light gray line shows global temperature estimates based on earth system model simulation data published by Friedrich *et al.*, available through the NOAA Paleoclimatology Database [16].

0.3.2 Population Genetics

All studied regions had inbreeding coefficients F_{IS} significantly greater than zero, indicating a consistently high degree of inbreeding in *P. fuscatus* populations across the Eastern US. The average F_{IS} across all regions was 0.46. Mitochondrial and nuclear mean nucleotide diversity are reported for each region (Table 3).

Location	Latitude	Longitude	r (km)	N	π_{mt}	π	F_{IS}
NC	36	-79	3.5	6	0.0077	0.0019	0.3972
MD	39	-77	67	7	0.0072	0.0011	0.5665
MA	42	-71	15	7	0.0038	0.0014	0.4332
C. NY	42	-77	20	8	0.0038	0.0014	0.5417
N. NY	43	-74	2.2	5	0.0070	0.0019	0.3064
All regions	NA	NA	750	33	0.0039	0.0016	0.5085

Table 3: Population genetic statistics for each study region and across all regions in Eastern US, calculated using vcftools: r = radius of specified population in kilometers; N = number of nests (= number of individuals); π_{mt} = mitochondrial gene nucleotide diversity; π = mean whole-genome nucleotide diversity; F_{IS} = inbreeding coefficient.

0.3.3 Population Structure

MDS plots show clustering of genetically similar individuals at the scale of regions across the Eastern US, sites within a region, and buildings within a site (Figure 2). Clusters at the continental scale appear roughly oriented by latitude along the C2 axis, while the C1 axis encodes variation within regions. At all three scales, individuals sampled across multiple years clustered together according to location.

The fastSTRUCTURE model that maximized marginal likelihood included

K=2 demes (Figure 3). The model output for K=2 demes indicates near panmictic levels of homogeneity, with only slight variation occuring in Central NY. All model runs (K=1-6) yielded no more than two clusters present in the population at proportions significantly above zero. The two predominant clusters were consistent across multiple model runs. Both clusters co-occur in all regions apart from Northern NY, indicating that clustering reveals within-region structure but no significant divergence between regions.

0.3.4 Isolation by Distance

Whole-genome and mitochondrial analyses revealed significant patterns of IBD in *P. fuscatus* wasps across the Eastern US (Figure 4). Mantel tests showed significant correlation between linearized genetic and geographic distance for whole-genome ($y = 8.20 * 10^{-4}x - 0.086$; $R^2 = 0.366$, $P = 0.036$) and mitochondrial ($y = 1.14 * 10^{-4}x - 0.0041$; $R^2 = 0.454$, $P = 0.018$) analyses. The mean pairwise divergence across all comparisons was $F_{\mathrm{ST}} = 0.0463$, $\Phi_{ST} = 0.198$.

0.4 Discussion

0.4.1 Demographic History

The reconstructed demographic history results suggests that *P. fuscatus* in the Eastern US experienced a population expansion and subsequent decline over the past 400,000 years. Given the lack of correspondence between the occurence of global temperature minima and the pattern of demographic history, it does

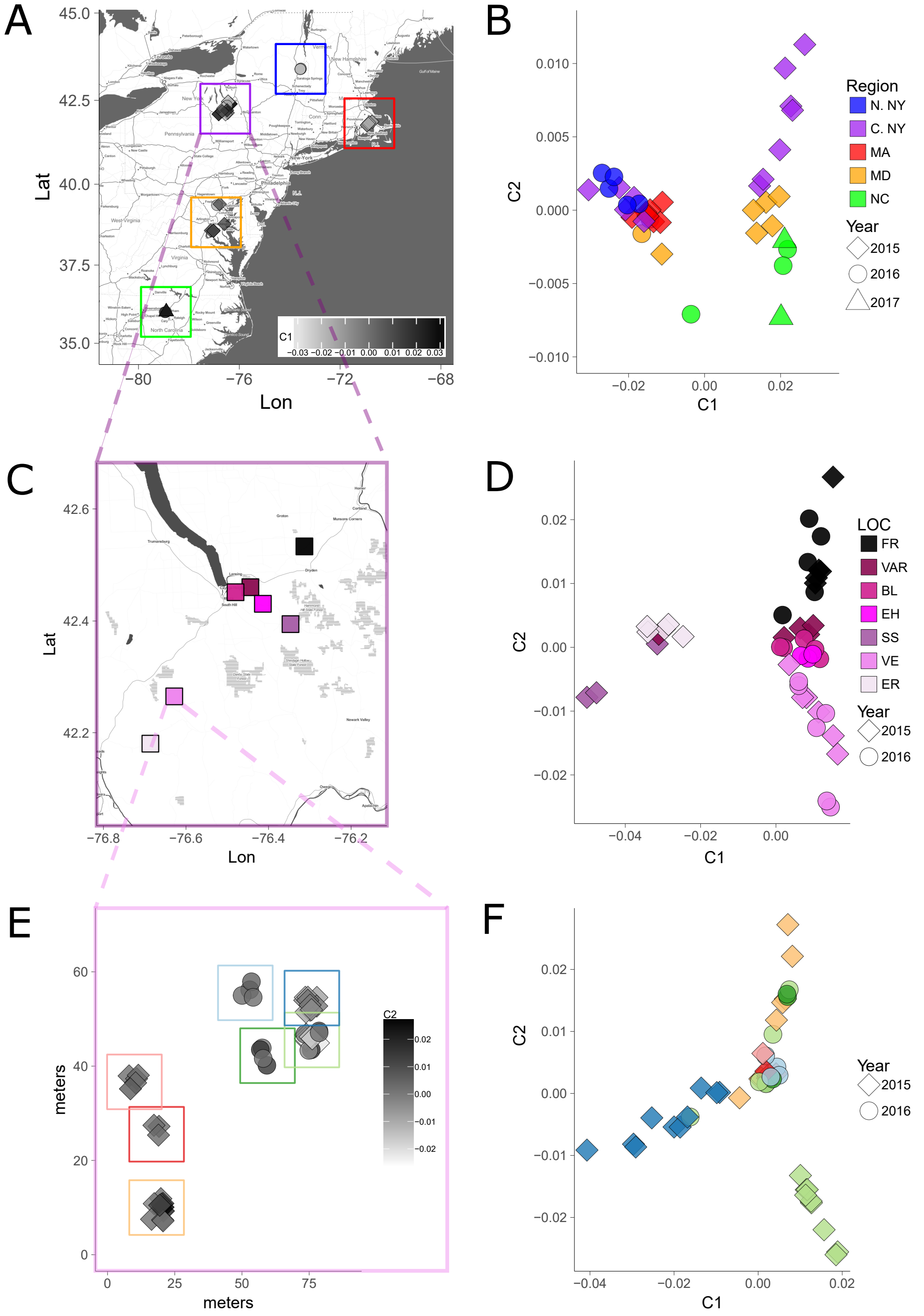

Figure 2: MDS plots of individual whole-genome sequences reveal population genetic structure at a continental scale between regions (A-B), at a regional scale between sites (C-D) and at a local scale between buildings within a site (E-F). Locations: Eastern US (A-B); Central NY (C-D); Van Etten, NY (E-F).

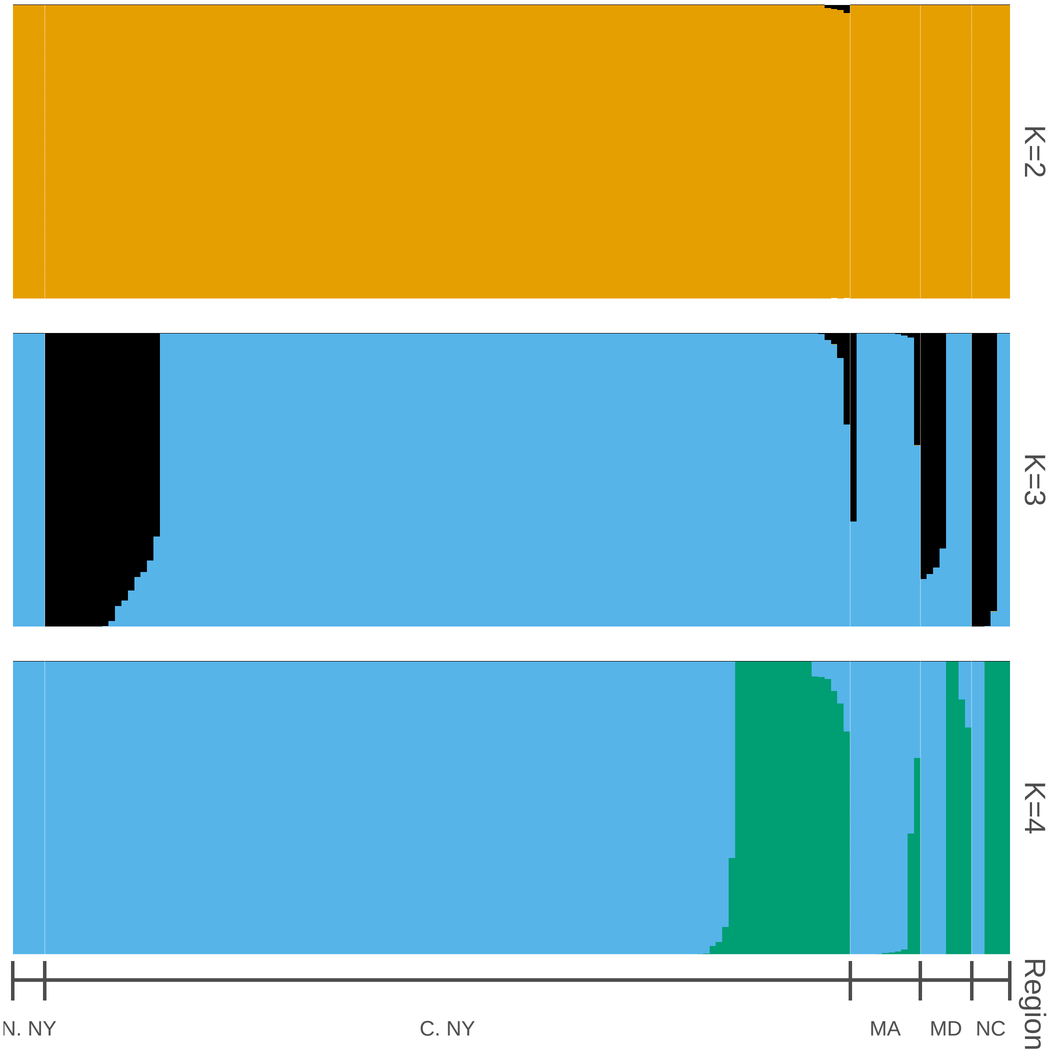

Figure 3: fastSTRUCTURE output for model runs with the number of demes K=2-4. The model complexity that maximizes marginal likelihood is K=2. For all models tested (K=2-6), no more than two clusters were represented in the population at proportions significantly above zero.

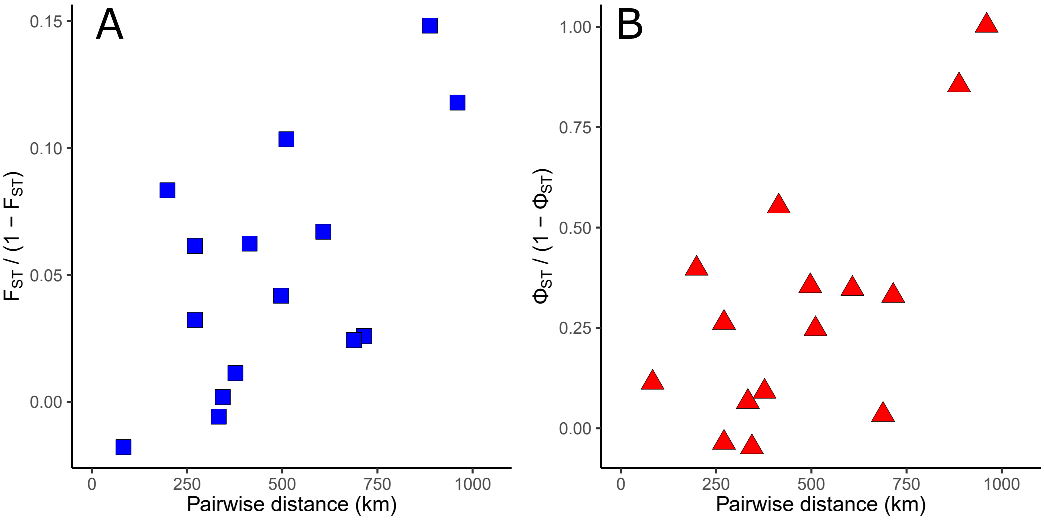

Figure 4: Linearized pairwise (A) whole-genome and (B) mitochondrial divergence versus geographic distance between Eastern US regions. Mantel tests showed significant correlation between geographic and genetic distance for both whole-genome ($y = 8.20 * 10^{-}4x - 0.086$; $R^2 = 0.366$, $P = 0.036$) and mitochondrial ($y = 1.14 * 10^{-}4x - 0.0041$; $R^2 = 0.454$, $P = 0.018$) analyses.

not seem likely that population expansion and retraction resulted from glacial advance and retreat.

The most recent reconstructed effective population size of roughly 200,000 individuals fits into the picture of a small effective breeding population with high inbreeding and low overall divergence across the continent. However, coalescence-based demographic reconstruction cannot be used to probe the most recent demographic history. Modern deforestation, which has increased the availability of *Polistes* wasps' preferred open habitat, may have led to an increase in effective population sizes over the past thousand years.

Demographic history is critical to interpreting population structure-based inference of dispersal, which relies on the assumption of constant population size on the timescales relevant to the drift-dispersal equilibrium. If the past thousand years saw a continued period of relative population stability of the previous hundred thousand, then population structure inference should yield accurate estimates of dispersal. In order to appropriately account for the effects of the population bottleneck and subsequent expansion, models to infer dispersal should be built around this background demographic history. Future comparative work should aim to include these demographic terms, to allow for more robust comparisons of dispersal-driven population structure between populations and species of *Polistes* wasps.

0.4.2 Spatial Genetic Structure

The results of fastSTRUCTURE model runs indicate an overall low degree of differentiation between regions, with nearly panmictic clustering observed for the optimal model with K=2. Interestingly, in models with K>2, two consistent clusters emerge but do not correspond to delineation between regions. One possible explanation for the clustering of samples into distinct groups independent of region is introgression from a separate lineage. The co-ocurrence of the two predominant clusters within multiple regions suggests the presence of fine-scale structure within regions, in spite of overall low divergence between regions.

High inbreeding coefficients observed across studied regions lend further support to the presence of fine-scale genetic structure at the level of sites within regions. MDS clustering reveals spatial genetic structure at all scales analyzed, ranging from the continental scale to the scale of hundreds of meters within a site. Given the arbitrary nature of MDS axes, the significance of distances between clusters cannot be interpreted. Rather, the MDS plots provide visual evidence of a degree of spatial structuring at all scales. Clustering of females at the level of buildings within a site supports a highly fine-tuned degree of natal philopatry, consistent with previous mark-recapture findings [31]. Notably, at the site scale, individuals sampled from separate nests across multiple years cluster together on the same buildings, indicating year-to-year fidelity to natal nest locations at the scale of buildings within a site. The consistency of spatial clustering across sampled years indicates that genetic similarity between neighboring individuals is not a transient pattern.

The reported population structure findings support the hypothesis that cooperative breeding is linked to fine-scale population structure. However, the methods employed in this study do not adequately test the alternative hypothesis that fine-scale population structure is unrelated to the presence or degree of cooperative breeding. Future analyses will probe evident fine-scale structure further to understand how cooperative nesting behavior relates to spatial genetic structure. If cooperative nesters have a greater tendency towards philopatry, we expect to find a pattern of greater fine-scale genetic structure between cooperatively nesting versus solitary nesting individuals. Similarly, if cooperatively breeding systems lead to a higher degree of natal philopatry, we expect to find correlations between foundress rates and population structure in future comparative studies between species of *Polistes* wasps that vary in their degree of cooperative nesting.

Isolation by Distance

In spite of low overall F_{ST} values, we detected a significant pattern of IBD in both nuclear and mitochondrial genomes. The low overall genetic divergence between populations (mean pairwise $F_{\mathrm{ST}} = 0.046$) may be due to a combination of recent population expansion and stabilizing selection. Further investigation of patterns of pairwise F_{ST} across the genome will address the extent to which stabilizing selection may play a role in limiting divergence. If regions of the genome are under selection to remain constant, we predict that these regions will show consistently low divergence across multiple pairwise comparisons.

The IBD regression for whole-genome analysis allows for estimation of dispersal distances, following Rousset's method [59]. We did not conduct systematic assays of nest density, however it is probable that *P. fuscatus* nesting density is comparable to that of *P. chinensis*, as estimated by Tsuchida *et al.* using a field searching algorithm. Applying Rousset's method of dispersal inference (Equation 1), and using Tsuchida *et al.* estimated density $D_e = 2500$ nests$/km^2$, yields an estimated mean parent-offspring axial dispersal distance of $\sigma = 530$ meters.

This absolute term should be interpreted with caution, however given similar datasets it may be possible to make comparative inferences about variation in dispersal patterns among populations and species of *Polistes* exhibiting varying degrees of cooperative breeding behavior.

Sex-biased Dispersal

Applying equation 2, with mean pairwise nuclear differentiation $F_{\mathrm{ST}} = 0.0463$, yields an expected $\Phi_{ST} = 0.163$, corresponding closely to the observed $\Phi_{ST} = 0.198$. Given the match between the observed Φ_{ST} and the expected value under equal dispersal of the sexes, the data does not support a model of sex-biased dispersal. It does not appear based on our analyses that the observed pattern of IBD is driven primarily by either sex.

While female nesting philopatry has been confirmed through mark-recapture studies and is supported by the results of this study, male dispersal has yet to be studied in the field. Further, the degree to which females and males disperse to mate is unknown. The contribution towards population structure of the two female dispersal decisions, one in the fall and the second in spring, remains to be analyzed. Future work will assess the degree of relatedness between males and females found mating in leks relative to the genetic background of nearby nests.

0.5 Conclusion

The results presented in this study provide genetic evidence of female philopatry that, when taken in context with local population attributes, could be quantified across populations and species in order to test theoretical predictions on the coevolution of cooperation and dispersal. In keeping with the predictions of theoretical models, evidence from previous work on cooperatively breeding taxa, and mark-recapture studies of *Polistes* wasps, we found that *P. fuscatus* female foundresses exhibit fine-tuned natal philopatry, frequently returning in the spring to nest on the same building from which they emerged the previous fall. In spite of low levels of overall genetic divergence, we found measurable population structure at the scale of buildings within sites, sites within regions, and regions across the continent.

Based on the observed pattern of IBD across the Eastern US, we inferred a mean axial parent-offspring dispersal distance of 530 meters, indicating that while some individuals renest in close proximity to their natal nests, others disperse on the scale of hundreds to thousands of meters. Future work will further explore the relationship between this inferred variation in dispersal distances and the variation in nesting strategies adopted between individuals. Additionally, the overall low genetic divergence in spite of clear patterns of spatial genetic structuring at multiple scales warrants further investigation. Future comparative methods should incorporate demographic history and control for or take into account differences in the spatial scales of analyses.

BIBLIOGRAPHY

[1] Stepfanie M Aguillon, John W Fitzpatrick, Reed Bowman, Stephan J Schoech, Andrew G Clark, Graham Coop, and Nancy Chen. Deconstructing isolation-by-distance: the genomic consequences of limited dispersal. *PLoS genetics*, 13(8):e1006911, 2017.

[2] Amanda Bretman, ROLANDO RODRÍGUEZ-MUÑOZ, Craig Walling, Jon Slate, and Tom Tregenza. Fine-scale population structure, inbreeding risk and avoidance in a wild insect population. *Molecular Ecology*, 20(14):3045–3055, 2011.

[3] R Cervo and Maria Cristina Lorenzi. Behaviour in usurpers and late joiners ofpolistes biglumis bimaculatus (hymenoptera, vespidae). *Insectes Sociaux*, 43(3):255–266, 1996.

[4] R Cervo and S Turillazzi. Associative foundation and nesting sites in polistes nimpha. *Naturwissenschaften*, 72(1):48–49, 1985.

[5] Roselle E Chapman and Andrew FG Bourke. The influence of sociality on the conservation biology of social insects. *Ecology Letters*, 4(6):650–662, 2001.

[6] Kylea E Clarke, Thomas E Rinderer, Pierre Franck, Javier G Quezada-Euán, and Benjamin P Oldroyd. The africanization of honeybees (apis mellifera l.) of the yucatan: a study of a massive hybridization event across time. *Evolution*, 56(7):1462–1474, 2002.

[7] Johanna Clémencet, Barbara Viginier, and Claudie Doums. Hierarchical analysis of population genetic structure in the monogynous ant cataglyphis cursor using microsatellite and mitochondrial dna markers. *Molecular Ecology*, 14(12):3735–3744, 2005.

[8] Rita Covas, Ambroise Dalecky, Alain Caizergues, and Claire Doutrelant. Kin associations and direct vs indirect fitness benefits in colonial cooperatively breeding sociable weavers philetairus socius. *Behavioral Ecology and Sociobiology*, 60(3):323–331, 2006.

[9] P-A Crochet. Genetic structure of avian populations—allozymes revisited. *Molecular Ecology*, 9(10):1463–1469, 2000.

[10] Adam L Cronin, Mathieu Molet, Claudie Doums, Thibaud Monnin, and Christian Peeters. Recurrent evolution of dependent colony foundation across eusocial insects. *Annual Review of Entomology*, 58, 2013.

[11] Scott K Davis, Joan E Strassmann, Colin Hughes, L Susan Pletscher, and Alan R Templeton. Population structure and kinship in polistes (hymenoptera, vespidae): an analysis using ribosomal dna and protein electrophoresis. *Evolution*, 44(5):1242–1253, 1990.

[12] André Rodrigues de Souza, Isabella Lopes Rodrigues, Ilinia V Assuncao Rocha, Walkiria A Acácio Reis, Juliane FS Lopes, and Fabio Prezoto. Foraging behavior and dominance hierarchy in colonies of the neotropical social wasp polistes ferreri (hymenoptera, vespidae) in different stages of development. *Sociobiology*, 52(2):293–303, 2008.

[13] Claudie Doums, Humberto Cabrera, and Christian Peeters. Population genetic structure and male-biased dispersal in the queenless ant diacamma cyaneiventre. *Molecular Ecology*, 11(11):2251–2264, 2002.

[14] MJ West Eberhard. The social biology of polistine wasps. 1969.

[15] Amin Eimanifar, Rebecca T Kimball, Edward L Braun, and James D Ellis. Mitochondrial genome diversity and population structure of two western honey bee subspecies in the republic of south africa. *Scientific reports*, 8(1):1333, 2018.

[16] Tobias Friedrich, Axel Timmermann, Michelle Tigchelaar, Oliver Elison Timm, and Andrey Ganopolski. Nonlinear climate sensitivity and its implications for future greenhouse warming. *Science Advances*, 2(11):e1501923, 2016.

[17] George J Gamboa, Bradley D Heacock, and Susanna L Wiltjer. Division of labor and subordinate longevity in foundress associations of the paper wasp, polistes metricus (hymenoptera: Vespidae). *Journal of the Kansas Entomological Society*, pages 343–352, 1978.

[18] George J Gamboa, Hudson K Reeve, and David W Pfennig. The evolution and ontogeny of nestmate recognition in social wasps. *Annual review of entomology*, 31(1):431–454, 1986.

[19] David L Gibo. The selective advantage of foundress associations in polistes fuscatus (hymenoptera: Vespidae): a field study of the effects of predation on productivity. *The Canadian Entomologist*, 110(5):519–540, 1978.

[20] William D Hamilton. The genetical evolution of social behaviour. ii. *Journal of theoretical biology*, 7(1):17–52, 1964.

[21] LJ Lawson Handley and N Perrin. Advances in our understanding of mammalian sex-biased dispersal. *Molecular ecology*, 16(8):1559–1578, 2007.

[22] Ben J Hatchwell. The evolution of cooperative breeding in birds: kinship, dispersal and life history. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1533):3217–3227, 2009.

[23] Yoshimi HIROSE and Masatoshi YAMASAKI. Dispersal of females for colony founding in polistes jadwigae dalla torre (hymenoptera, vespidae). *Journal of Environmental Entomology*, 52(1):65–71, 1984.

[24] Michael E Hochberg, Daniel J Rankin, and Michael Taborsky. The coevolution of cooperation and dispersal in social groups and its implications for the emergence of multicellularity. *BMC Evolutionary Biology*, 8(1):238, 2008.

[25] Delbert W Hutchison and Alan R Templeton. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, 53(6):1898–1914, 1999.

[26] JM Jandt, EA Tibbetts, and AL Toth. Polistes paper wasps: a model genus for the study of social dominance hierarchies. *Insectes Sociaux*, 61(1):11–27, 2014.

[27] Rebecca N Johnson and Philip T Starks. A surprising level of genetic diversity in an invasive wasp: Polistes dominulus in the northeastern united states. *Annals of the Entomological Society of America*, 97(4):732–737, 2014.

[28] Rufus A Johnstone, Michael A Cant, and Jeremy Field. Sex-biased dispersal, haplodiploidy and the evolution of helping in social insects. *Proc. R. Soc. B*, 279(1729):787–793, 2012.

[29] Motoo Kimura and George H Weiss. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, 49(4):561, 1964.

[30] J Klahn. Intraspecific comb usurpation in the social wasp polistes fuscatus. *Behavioral Ecology and Sociobiology*, 23(1):1–8, 1988.

[31] Jeff E Klahn. Philopatric and nonphilopatric foundress associations in the social wasp polistes fuscatus. *Behavioral Ecology and sociobiology*, 5(4):417–424, 1979.

[32] KB Kozyra, I Melosik, and E Baraniak. Genetic diversity and population structure of polistes nimpha based on dna microsatellite markers. *Insectes sociaux*, 62(4):423–432, 2015.

[33] Jean-François Le Galliard, Régis Ferriere, and Ulf Dieckmann. Adaptive evolution of social traits: origin, trajectories, and correlations of altruism and mobility. *The American Naturalist*, 165(2):206–224, 2004.

[34] Thibault Lengronne, Ellouise Leadbeater, Solenn Patalano, Stephanie Dreier, Jeremy Field, Seirian Sumner, and Laurent Keller. Little effect of seasonal constraints on population genetic structure in eusocial paper wasps. *Ecology and Evolution*, 2(10):2615–2624, 2012.

[35] Heng Li and Richard Durbin. Fast and accurate short read alignment with burrows–wheeler transform. *bioinformatics*, 25(14):1754–1760, 2009.

[36] Haoxuan Liu, Yanxiao Jia, Xiaoguang Sun, Dacheng Tian, Laurence D Hurst, and Sihai Yang. Direct determination of the mutation rate in the bumblebee reveals evidence for weak recombination-associated mutation and an approximate rate constancy in insects. *Molecular biology and evolution*, 34(1):119–130, 2016.

[37] Maria Cristina Lorenzi and R Cervo. Usurpations and late associations in the solitary founding social wasp, polistes biglumis bimaculatus (hymenoptera: Vespidae). *Journal of insect behavior*, 8(4):443–451, 1995.

[38] Shun'ichi Makino, Seiki Yamane, Tsukasa Sunose, and Shigeyuki Aoki. Dispersion distance of queens from natal sites in the two haplometrotic paper waspspolistes riparius andp. snelleni (hymenoptera: Vespidae). *Researches on population ecology*, 29(1):111–117, 1987.

[39] Gustave Malécot et al. Mathematics of heredity. *Mathematics of heredity.*, 1948.

[40] Takeo Maruyama. Rate of decrease of genetic variability in a two-dimensional continuous population of finite size. *Genetics*, 70(4):639–651, 1972.

[41] Robert A Metcalf and Gregory S Whitt. Intra-nest relatedness in the social wasp polistes metricus. *Behavioral Ecology and Sociobiology*, 2(4):339–351, 1977.

[42] David M Mickelson and Patrick M Colgan. The southern laurentide ice sheet. *Developments in Quaternary Sciences*, 1:1–16, 2003.

[43] Sara Miller et al. Waspnest: A worldwide assessment of social polistine nesting behavior. (in press).

[44] Joshua Mitteldorf and David Sloan Wilson. Population viscosity and the evolution of altruism. *Journal of Theoretical Biology*, 204(4):481–496, 2000.

[45] S Miyano and E Hasegawa. Genetic structure of the first brood of workers and mating frequency of queens in a japanese paper wasp, polistes chinensis antennalis. *Ethology Ecology & Evolution*, 10(1):79–85, 1998.

[46] JN Painter, RH Crozier, A Poiani, RJ Robertson, and MF Clarke. Complex social organization reflects genetic structure and relatedness in the cooperatively breeding bell miner, manorina melanophrys. *Molecular Ecology*, 9(9):1339–1347, 2000.

[47] Emmanuel Paradis, Stephen R Baillie, William J Sutherland, and Richard D Gregory. Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology*, 67(4):518–536, 1998.

[48] Stuart B Piertney, Xavier Lambin, Andrew DC Maccoll, Kerry Lock, Philip J Bacon, John F Dallas, Fiona Leckie, Francois Mougeot, Paul A Racey, Steve Redpath, et al. Temporal changes in kin structure through a population cycle in a territorial bird, the red grouse lagopus lagopus scoticus. *Molecular Ecology*, 17(10):2544–2551, 2008.

[49] Thomas G Platt and James D Bever. Kin competition and the evolution of cooperation. *Trends in ecology & evolution*, 24(7):370–377, 2009.

[50] Simon T Powers, Alexandra S Penn, and Richard A Watson. The concurrent evolution of cooperation and the population structures that support it. *Evolution*, 65(6):1527–1543, 2011.

[51] Michel Pratte. Foundress association in the paper wasp polistes dominulus christ.(hymen. vesp.). effects of dominance hierarchy on the division of labour. *Behaviour*, 111(1):208–219, 1989.

[52] Franck Prugnolle and Thierry De Meeûs. Inferring sex-biased dispersal from population genetic tools: a review. *Heredity*, 88(3):161, 2002.

[53] Ute Radespiel, Heike Lutermann, Barthel Schmelting, Michael W Bruford, and Elke Zimmermann. Patterns and dynamics of sex-biased dispersal in a nocturnal primate, the grey mouse lemur, microcebus murinus. *Animal Behaviour*, 65(4):709–719, 2003.

[54] Ute Radespiel, Zübeyde Sarikaya, Elke Zimmermann, and Michael W Bruford. Sociogenetic structure in a free-living nocturnal primate population: sex-specific differences in the grey mouse lemur (microcebus murinus). *Behavioral Ecology and Sociobiology*, 50(6):493–502, 2001.

[55] Anil Raj, Matthew Stephens, and Jonathan K Pritchard. Variational inference of population structure in large snp datasets. *Genetics*, pages genetics–114, 2014.

[56] John C Ridge, BA Canwell, MA Kelly, SZ Kelley, T Weddle, and M Retelle. Atmospheric 14c chronology for late wisconsinan deglaciation and sea-level change in eastern new england using varve and paleomagnetic records. In *Deglacial history and relative sea-level changes, northern New England and adjacent Canada*, volume 351, pages 173–191. 2001.

[57] P-F Röseler, I Röseler, and A Strambi. Studies of the dominance hierarchy in the paper wasp, polistes gallicus (l.)(hymenoptera vespidae). *Monitore Zoologico Italiano-Italian Journal of Zoology*, 20(3):283–290, 1986.

[58] Kenneth G Ross and D Dewayne Shoemaker. Nuclear and mitochondrial genetic structure in two social forms of the fire ant solenopsis invicta: insights into transitions to an alternate social organization. *Heredity*, 78(6):590, 1997.

[59] François Rousset. Genetic differentiation and estimation of gene flow from f-statistics under isolation by distance. *Genetics*, 145(4):1219–1228, 1997.

[60] Olav Rüppell, Michaela Strätz, Bernd Baier, and Jürgen Heinze. Mitochondrial markers in the ant leptothorax rugatulus reveal the population genetic consequences of female philopatry at different hierarchical levels. *Molecular Ecology*, 12(3):795–801, 2003.

[61] Andrew F Russell. *Ecological constraints and the cooperative breeding system of the long-tailed tit Aegithalos caudatus*. PhD thesis, 2000.

[62] Perttu Seppä, Johan Fogelqvist, Niclas Gyllenstrand, and MC Lorenzi. Colony kin structure and breeding patterns in the social wasp, polistes biglumis. *Insectes sociaux*, 58(3):345–355, 2011.

[63] Stuart P Sharp, Mitchell B Baker, Jarrod D Hadfield, Michelle Simeoni, and Ben J Hatchwell. Natal dispersal and recruitment in a cooperatively breeding bird. *Oikos*, 117(9):1371–1379, 2008.

[64] Lisa Shorey, Stuart Piertney, Jon Stone, and Jacob HoÈglund. Fine-scale genetic structuring on manacus manacus leks. *Nature*, 408(6810):352, 2000.

[65] Les Sirkin and Robert Stuckenrath. The portwashingtonian warm interval in the northern atlantic coastal plain. *Geological Society of America Bulletin*, 91(6):332–336, 1980.

[66] Montgomery Slatkin. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, 47(1):264–279, 1993.

[67] Jay F Storz. Genetic consequences of mammalian social structure. *Journal of Mammalogy*, 80(2):553–569, 1999.

[68] Joan E Strassmann. Wasp reproduction and kin selection: reproductive competition and dominance hierarchies among polistes annularis foundresses. *The Florida Entomologist*, 64(1):74–88, 1981.

[69] Liselotte Sundström, Laurent Keller, and Michel Chapuisat. Inbreeding and sex-biased gene flow in the ant formica exsecta. *Evolution*, 57(7):1552–1561, 2003.

[70] Liselotte Sundström, Perttu Seppä, and Pekka Pamilo. Genetic population structure and dispersal patterns in formica ants—a review. In *Annales Zoologici Fennici*, pages 163–177. JSTOR, 2005.

[71] Jonathan Terhorst, John A Kamm, and Yun S Song. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nature genetics*, 49(2):303, 2017.

[72] Koji Tsuchida. Genetic relatedness and the breeding structure of the japanese paper wasp, polistes jadwigae. *Ethology ecology & evolution*, 6(2):237–242, 1994.

[73] Koji Tsuchida, Kazuyuki Kudô, and Norio Ishiguro. Genetic structure of

an introduced paper wasp, p olistes chinensis antennalis (h ymenoptera, v espidae) in n ew z ealand. *Molecular ecology*, 23(16):4018–4034, 2014.

[74] Md Mesbah Uddin and Koji Tsuchida. Genetic population structure of the paper wasp polistes olivaceus (hymenoptera: Vespidae) in bangladesh. *Population ecology*, 54(1):103–114, 2012.

[75] Geraldine A Van der Auwera, Mauricio O Carneiro, Christopher Hartl, Ryan Poplin, Guillermo Del Angel, Ami Levy-Moonshine, Tadeusz Jordan, Khalid Shakir, David Roazen, Joel Thibault, et al. From fastq data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Current protocols in bioinformatics*, 43(1):11–10, 2013.

[76] Xavier Vekemans and Olivier J Hardy. New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular ecology*, 13(4):921–935, 2004.

[77] Mary Jane West. Foundress associations in polistine wasps: dominance hierarchies and the evolution of social behavior. *Science*, 157(3796):1584–1585, 1967.

[78] Sewall Wright. Isolation by distance. *Genetics*, 28(2):114, 1943.

[79] Sewall Wright. The genetical structure of populations. *Annals of eugenics*, 15(1):323–354, 1949.

[80] Sihai Yang, Long Wang, Ju Huang, Xiaohui Zhang, Yang Yuan, Jian-Qun Chen, Laurence D Hurst, and Dacheng Tian. Parent–progeny sequencing indicates higher mutation rates in heterozygotes. *Nature*, 523(7561):463, 2015.

[81] Lorenzo RS Zanette and Jeremy Field. Founders versus joiners: group formation in the paper wasp polistes dominulus. *Animal behaviour*, 82(4):699–705, 2011.