Investigating sex differences in genetic relatedness in great-tailed grackles in Tempe, Arizona to infer potential sex biases in dispersal

A Sevchik[1]   CJ Logan[2]   MA Folsom[2]   L Bergeron[3]   AD Blackwell[4]
C Rowney[2]   D Lukas[2]*

29 November 2019

Affiliations:

1. Arizona State University, Tempe
2. Max Planck Institute for Evolutionary Anthropology Leipzig
3. University of California, Santa Barbara
4. Washington State University, Pullman

Correspondence: * dieter_lukas@eva.mpg.de

Peer Community In
Ecology

ABSTRACT

In most bird species, females disperse prior to their first breeding attempt, while males remain close to the place they were hatched for their entire lives (Greenwood and Harvey (1982)). Explanations for such female bias in natal dispersal have focused on the potential benefits that males derive from knowing the local environment to establish territories, while females search for suitable mates (Greenwood (1980)), however the exact factors shaping dispersal decisions appear more complex (Mabry et al. (2013), Végvári et al. (2018)). Here, we investigate whether females are the dispersing sex in great-tailed grackles, which have a mating system where the males hold territories and the females choose which territory to place their nest in (Johnson et al. (2000)). We will use genetic approaches to identify sex biases in the propensity to disperse. We will

first determine whether, for individuals caught at a single site in Arizona, the average relatedness among all female dyads is lower than that among all male dyads. If supported, this would suggest that females are less likely to be found close to genetic relatives, which indicates that females disperse away from relatives. Second, we will assess whether in males close relatives are most likely to be found within very short distances of each other; whereas, in females, relatives live both near and far from each other. Results will inform our long-term study on the relationship between behavioral flexibility and rapid geographic range expansion by elucidating which individuals are likely to experience similar conditions across their lives, and which are likely to face new conditions when they become breeders.

A. STATE OF THE DATA

The first version of this preregistration was written (June 2019) and submitted to Peer Community In Ecology (July 2019) after blood was collected and before processing the DNA to obtain the genetic data. The revised version of this preregistration following peer review at Peer Community in Ecology (October 2019) was written after obtaining the genotypes for the individuals in the sample and resubmitted in November 2019.

B. HYPOTHESIS

Hypothesis There are sex differences in the natal disperal rate and distance among individuals in great-tailed grackles (*Quiscalus mexicanus*) with males remaining close to where they hatched and females moving away from where they hatched. Males are expected to remain close to the area where they hatched, therefore a large number of the males on the Arizona State University (ASU) campus are expected to have hatched within the area of the study site and stay close to their relatives. In contrast, females are expected to move before their first breeding attempt (Greenwood (1980)), therefore females on campus are likely to come from areas outside of campus in the surrounding area, having moved away from relatives.

Alternative 1 Males disperse away from where they hatched, while females remain where they hatched.

Alternative 2 Individuals of both sexes remain close to where they hatched.

Alternative 3 Individuals of both sexes disperse away from where they hatched.

We predict that the movement of individuals will influence the spatial distribution of genetic relatives. Individuals of the sex who remain close to where they hatched are expected to be close to relatives while individuals of the sex who disperse are expected to not be close to relatives (see Fig 1 for a visualization). We also expect that the further the distance an individual moves, the less likely they are to be even distantly related to another individual within the study area. We will perform three analyses to investigate the spatial distribution of genetic relatives: the first two aim to detect whether there are sex biases in levels of average genetic relatedness among indivduals found within a certain distance of each other (analysis i: average levels of relatedness among individuals in our sample; analysis ii: geographic distances between individuals assumed to be genetic relatives) and the third aims to describe the genetic structure separately for each sex and whether relatives are predominantly found within certain distances from each other or whether relatives are not structured in geographic space (analysis iii: spatial autocorrelation).

Predictions for the hypothesis: dispersal males < females - **Analysis i (average relatedness)**: Both the mean level of and the variance in average genetic relatedness will be higher among males compared to females in our sample. - **Analysis ii (distance between relatives)**: Both the mean and variance of the geographic distances between pairs of individuals inferred to be genetic relatives (individuals estimated to be related at levels of cousins or closer, $r>0.125$) will be shorter among males compared to females. - **Analysis iii (spatial autocorrelation)**: There will be a spatial autocorrelation signal indicating a negative relationship between genetic relatedness and geographic distance for males. There will be no spatial autocorrelation signal indicating the absence of a relationship between genetic relatedness and geographic distance for females.

Predictions for alternative 1: dispersal males > females - **Analysis i (average relatedness)**: Both the mean level of and the variance in average genetic relatedness will be lower among males compared to females in our sample. - **Analysis ii (distance between relatives)**: Both the mean and variance of the geographic distances between pairs of individuals inferred to be genetic relatives (individuals estimated to be related at levels of cousins or closer, $r>0.125$) will be higher among males compared to females. - **Analysis iii (spatial autocorrelation)**: There will be no spatial autocorrelation signal indicating the

absence of a relationship between genetic relatedness and geographic distance for males. There will be a spatial autocorrelation signal indicating a negative relationship between genetic relatedness and geographic distance for females.

Predictions for alternative 2: neither males nor females disperse - **Analysis i (average relatedness)**: Both the mean level of and the variance in average genetic relatedness will be similar among males compared to females in our sample. - **Analysis ii (distance between relatives)**: Both the mean and variance of the geographic distances between pairs of individuals inferred to be genetic relatives (individuals estimated to be related at levels of cousins or closer, r>0.125) will be similar among males compared to females. - **Analysis iii (spatial autocorrelation)**: There will be a spatial autocorrelation signal indicating a negative relationship between genetic relatedness and geographic distance for males. There will be a spatial autocorrelation signal indicating a negative relationship between genetic relatedness and geographic distance for females.

Predictions for alternative 3: both males and females disperse - **Analysis i (average relatedness)**: Both the mean level of and the variance in average genetic relatedness among males will be similar to that among females. - **Analysis ii (distance between relatives)**: Both the mean and variance of the geographic distances between pairs of males inferred to be genetic relatives (individuals estimated to be related at levels of cousins or closer, r>0.125) will be similar to distances among female genetic relatives. - **Analysis iii (spatial autocorrelation)**: There will be no spatial autocorrelation signal indicating the absence of a relationship between genetic relatedness and geographic distance for males. There will be no spatial autocorrelation signal indicating the absence of a relationship between genetic relatedness and geographic distance for females.

C. METHODS

Planned Sample

DNA from 57 great-tailed grackles was obtained from wild adults (n=40 adult females, n=17 adult males, juvenile samples were excluded because they had not yet dispersed) caught in Tempe, Arizona, USA (see Fig 2 for a map). These individuals were either immediately released, or temporarily brought into aviaries for behavioral testing and then released back to the wild.

The larger number of females than males in our sample appears to reflect the adult sex ratio at this study site. To estimate the sex ratio at the field site, we counted the number of females and males that were trapped in mist nets since the beginning of our study (September 2017 - October 2019). This trapping method likely does not elicit a sex bias in terms of which sex is caught because the nets are invisible. Therefore, if one sex is more neophobic than the other, both sexes are likely to be trapped using this method. A total of 26 females and 11 males were trapped using mist nets (a ratio of 2.36 females per 1 male), which is very similar to the sex ratio in our sample consisting of 40 females and 17 males (2.35 females per 1 male).

Females were caught at all but one site, such that comparisons are possible of the genetic relatedness of pairs of females trapped at various distances from each other. Males were not caught at all trap sites, but there are several sites at which multiple males were caught and sufficient sites for comparisons of males that were caught close to each other, at intermediate, and at long distances (Figure 3).

Sample size rationale

The sample size presented was the largest one possible by July 2019 when the DNA were sequenced using ddRADseq.

Data collection stopping rule

We analyzed all blood samples that were collected through June 2019, which was the end of the trapping season.

Open data

When the study is complete, the data will be published in the Knowledge Network for Biocomplexity's data repository.
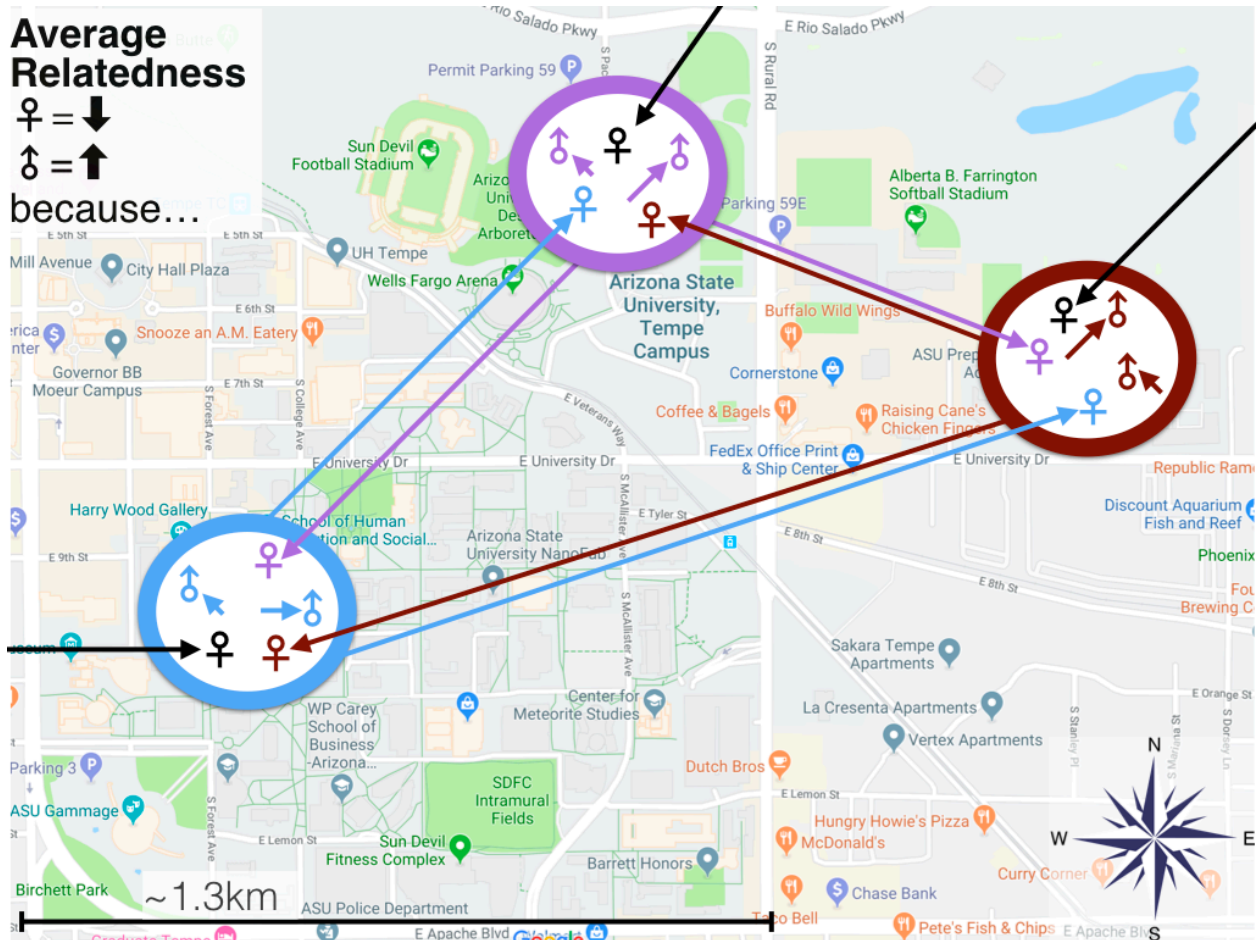
Figure 1: Figure 1. Visual representation of the hypothesis that males will have higher levels of genetic relatedness than females because females likely have larger dispersal distances.
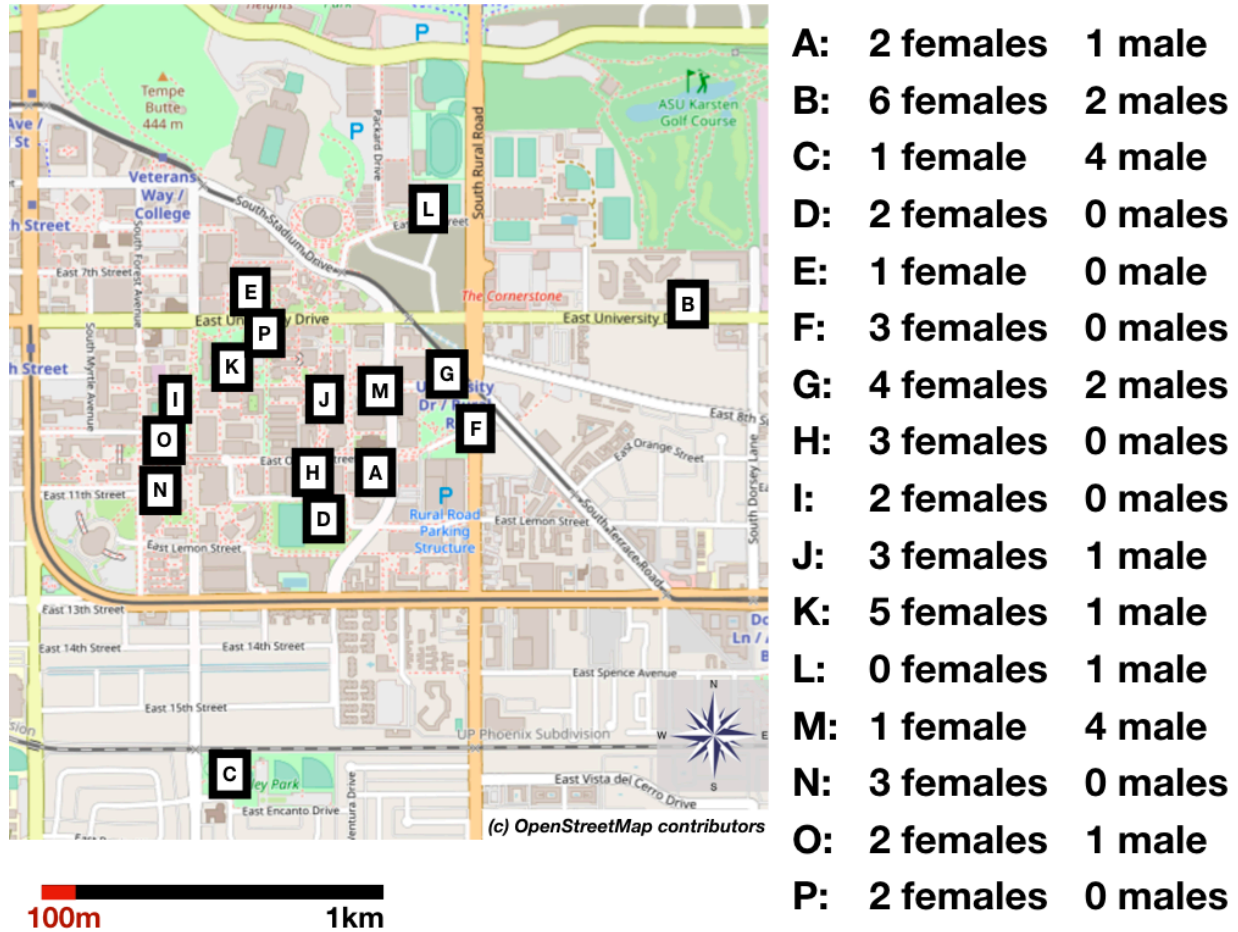
A: 2 females 1 male
B: 6 females 2 males
C: 1 female 4 male
D: 2 females 0 males
E: 1 female 0 male
F: 3 females 0 males
G: 4 females 2 males
H: 3 females 0 males
I: 2 females 0 males
J: 3 females 1 male
K: 5 females 1 male
L: 0 females 1 male
M: 1 female 4 male
N: 3 females 0 males
O: 2 females 1 male
P: 2 females 0 males

Figure 2: Figure 2. Map displaying the sampling locations of grackles on the ASU campus and the number of individuals trapped at each location as part of this project.

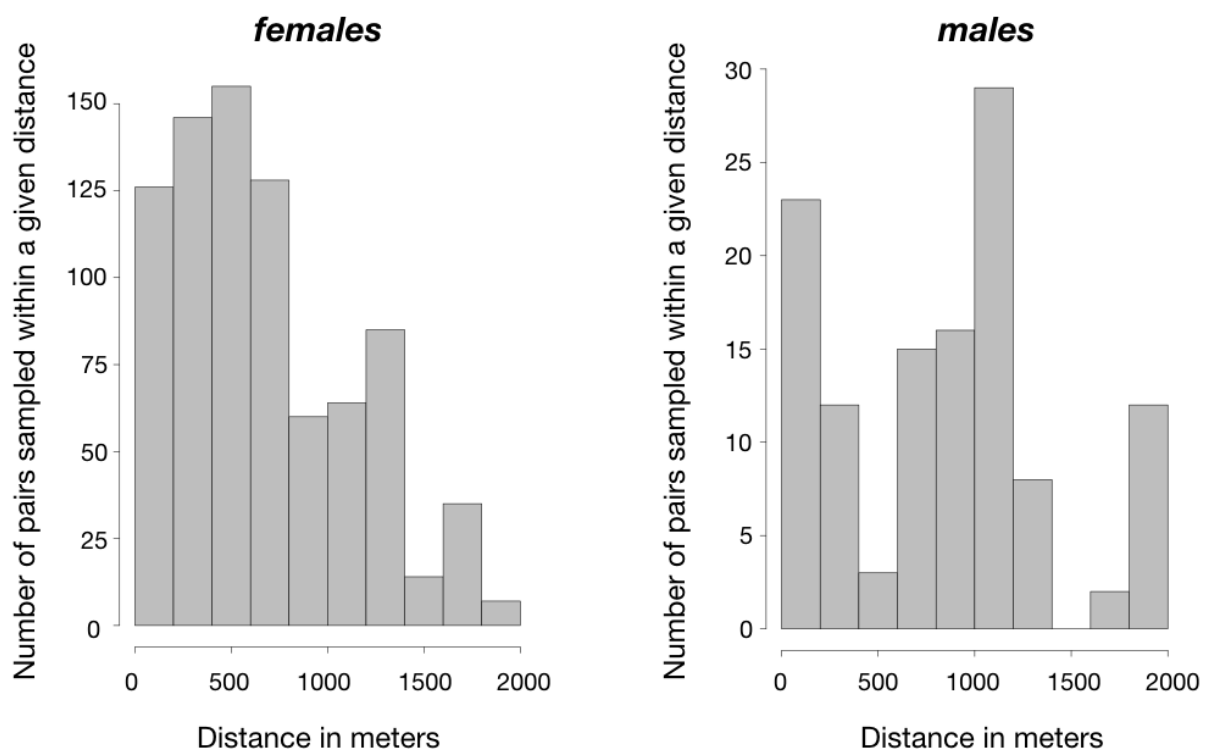# Distribution of geographic distances between pairs of individuals

Figure 3: Figure 3. Histograms displaying the geographic distances between all pairs of females (left) and all pairs of males (right) in our sample.

Randomization and counterbalancing

No randomization or counterbalancing is involved in this study.

Blinding of conditions during analysis

Experimenters are blind to the sex of the bird when processing samples using ddRADseq (only the alphanumeric bird ID was visible on the tube and no team member has memorized which ID goes with which bird because we give the birds names).

Blood collection

Blood was collected and stored in one of two ways until DNA extraction:

1) At the beginning of the project (2018), 70 uL of whole blood was added to silicone-coated micro-blood collection tubes containing 280 uL of lysis buffer (White and Densmore (1992), pp. 50-51) and stored at room temperature for up to a year before DNA extraction.

2) In 2018 a different method was implemented, using DNA from packed red blood cells: 150uL of blood was collected from trapped great-tailed grackles and stored for a minimum of 30 minutes and a maximum of 60 minutes at room temperature or 3 hours on ice. Samples were then centrifuged at 15x gravity for 10 minutes to separate the serum from the cellular fraction. After the serum layer was removed and stored, 600uL lysis buffer (White and Densmore (1992), pp. 50-51) was added to the remaining packed cells. Tubes containing packed cells and lysis buffer were stored at room temperature for up to 1 year before extraction.

DNA extraction and quantification

Some samples were extracted at Arizona State University by Rowney (samples through Dec 2018), while others were shipped with ice packs to Washington State University for extraction by Blackwell and his lab (samples collected Jan-Jun 2019). DNA was extracted from the above samples using the DNeasy Blood and Tissue kit (Qiagen) with slight modifications from the manufacturer's protocol (see details in Thrasher et al. (2018) Supporting Information, page 7; our slightly modified protocol is available here). Approximately 100ul of blood/lysis mixture was mixed with 20ul Proteinase K, 150ul PBS, and 200ul buffer AL, then incubated overnight at 64C while shaking. Samples were mixed with 200ul ethanol and added to spin columns. Columns were centrifuged and washed according to kit protocol using buffers AW1 and AW2. DNA was eluted into 50ul of RNAse and DNAse free water at 64C after a 5-10 min incubation on columns. DNA quantification was then performed on a Qubit 4.0 Fluorometer (Fisher Scientific) following the manufacturer's protocol for broad range dsDNA. The average yield of samples used for sequencing was 34ng/ul. Extracted DNA samples were shipped with ice packs to the Cornell Lab of Ornithology for ddRAD sequencing in July 2019.

ddRADseq

The DNA was processed using ddRADseq by Sevchik and Bronwyn Butcher (Cornell University) following methods in Thrasher et al. (2018). In brief, a sample is chosen for each individual with a concentration as close to between 5-50µl as possible and any sample that remains outside that range is concentrated or diluted as necessary. Concentrations are obtained using the Qubit Fluorometer. The DNA extracts are then run through a PCR thermocycle where the fragments are digested with a combination of two restriction enzymes and 20 different adapters attached to the end of the DNA pieces. A gel is then run to ensure the proper digestion and ligation of the DNA samples. The samples are then cleaned up using SPRI beads and size selected using BluePippin for a prespecified length (between 400-700 base pairs). After the samples return from size selection, they are amplified using a low-cycle PCR process and pooled together to be sent in to be sequenced. Sequencing was performed on an Illumina NextSeq500 (using a mid-output kit) to generate 150 bp single end reads. These data were post-processed to generate SNP data for relatedness analyses as in Thrasher et al. (2018). After filtering reads for quality and demultiplexing to assign sequences back to specific individuals, genetic loci were assembled *de novo* because no reference genome exists for great-tailed grackles. We will follow the cut-offs described in Thrasher et al. (2018) for single nucleotide polymorphism filtering, but might adjust settings (e.g., threshold of accepted minor allele frequencies) depending on the composition

of our sample (these decisions will have no influence on testing our predictions since they influence females and males equally and will be performed prior to any further analyses).

Relatedness analyses

Genetic relatedness between all pairs of individuals is calculated using the package "related" (Pew et al. (2015)) in R, following methods in Thrasher et al. (2018). We will use the function 'compareestimators' to assess which relatedness estimator appears to perform the best given the characteristics of our data.

Dependent variable

Average relatedness between all pairs of individuals within one sex

Independent variables

1) Sex (female, male)

2) Distance between trap sites (meters)

D. ANALYSIS PLAN

We do not plan to **exclude** any data. If any **missing** data occur for the genotyping, we will exclude individuals more than half of their genotype is unknown. Analyses will be conducted in R (current version 3.6.1; R Core Team (2017)).

Ability to detect actual effects

Birds the size of a grackle (~100-150 grams) are expected to show a median natal dispersal distance of about 250-300 meters (Sutherland et al. (2000)). Our 15 trap locations are located within a ~1000m radius circle, suggesting that if there are dispersers in our sample, these individuals will have most likely come from areas outside of the trapping circle. In turn, if individuals remain close to their natal area, they would only move distances much shorter than this, suggesting that the pairwise distances between non-dispersed relatives would be shorter than the random distance between any two birds we caught. However, we do not know the average distance that either philopatric or dispersing individuals move. The scale of our sampling area might be so small that individuals of the sex that disperses the least are likely to have hatched outside of this area. In addition, there could be variation among either females or males in the distances individuals move, with potentially also a small proportion of individuals of the predominantly philopatric sex dispersing, which could obscure patterns in the small sample of individuals in our study. Accordingly, we might not be able to detect differences in average relatedness between females and males (analysis i), but we still might expect a sex bias in the geographic distances among relatives (analysis ii).

We restricted our sample to adults to focus on the distribution of individuals after any potential natal dispersal (Goudet, Perrin, and Waser (2002)). We only have individuals from within a single site, so we cannot use methods that rely on assigning individuals to a source population or measure the relative distribution of genetic variation within versus among populations (Fst or similar measures). We therefore rely on measuring genetic relatedness between pairs of individuals. Approaches relying on spatial analyses of multi-locus genotypes have been shown to be able to detect even modest sex biased dispersal in fine-scale spatial distribution, in particular analyses of spatial autocorrelation (Banks and Peakall (2012)). However, our sample size is small, meaning that we might have only limited power to detect potential differences between females and males (Goudet, Perrin, and Waser (2002)). For the spatial distribution of relatives (Analysis ii), the number of related individuals in our sample might be too small to detect a strong pattern of the relatives of one sex being more geographically closer to each other than relatives of the other sex. For the isolation-by-distance leading to a change in relatedness within the range of our sampling locations, the signal might be too weak in either or both sexes to make inferences about sex differences (Analysis iii). However, for the comparison of average relatedness (Analysis i), given that we have a large number of SNP loci, we expect that we should have sufficient power to obtain a qualitative assessment of whether relatives are present in our sample (Wang and Santure (2009)) and accordingly whether dispersal is more prevalent in either females or males (examples of empirical studies that detected a signal with small sample sizes include Hofmann et al. (2012), Quaglietta et al. (2013), Gour et al. (2013), Botero-Delgadillo et al. (2017)).

Analysis i: average relatedness and sex

We will compare the average and variance in relatedness among all females to that among all males. Since average relatedness tends to decrease as the number of individuals in the sample increases (regression to the mean), we will perform a permutation analysis to investigate whether the average relatedness among the males or among the females in our sample is higher than what would be expected for a random sample of the same number of females or of individuals of both sexes. We will perform 10,000 random draws of 17 individuals either from among the females or from among all individuals and of 40 individuals from among all individuals, and generate distributions of average relatedness among these samples. We will assess whether the observed average relatedness among the 17 males or the 40 females in our sample is higher than what is observed in the majority of random samples. We will report the proportion of random samples with lower relatedness than the observed values and, for comparison with other approaches, assess whether the observed relatedness is higher than the relatedness calculated for 95% of all random draws.

```r
#Make code wrap text so it doesn't go off the page when Knitting to PDF
library(knitr)
opts_chunk$set(tidy.opts=list(width.cutoff=60),tidy=TRUE)
```

```r
input <- readgenotypedata("grackle_genepop_converted.txt")


# Analysis 1: Assess whether average relatedness is higher
# among females or among males Calculate pairwise
# relatedness, here choosing the relatedness method developed
# by Wang
outfile <- coancestry(input$gdata, wang = 1)


# extract the relevant information from the file
pairwise_r <- outfile$relatedness


# identify which individuals are female and which are male
individuals <- input$gdata$V1
females <- input$gdata$V1[grep("AF", input$gdata$V1)]
males <- input$gdata$V1[grep("AM", input$gdata$V1)]


# Calculate average relatedness among all individuals, all
# females, and all males
mean(filter(pairwise_r, ind1.id %in% females, ind2.id %in% females)$wang)
mean(filter(pairwise_r, ind1.id %in% males, ind2.id %in% males)$wang)
mean(pairwise_r$wang)



# Perform a simulation to assess whether average relatedness
# among males is different from what we would expect in a
# random subset of the same number of individuals
simulatedrelatedness <- matrix(ncol = 1, nrow = 10000)
for (i in 1:10000) {
    currentset <- sample(individuals, length(males))
    simulatedrelatedness[i, 1] <- mean(filter(pairwise_r, ind1.id %in%
        currentset, ind2.id %in% currentset)$wang)
}
hist(simulatedrelatedness)
# This value is similar to a p-value, it reflects the
# probability that the average relatedness observed among
# males would be expected in a random subsample
sum(simulatedrelatedness > mean(filter(pairwise_r, ind1.id %in%
```

```
    males, ind2.id %in% males)$wang))/10000
```

Analysis ii: distances among genetic relatives

Based on the calculations of pairwise genetic relatedness, we will select the subset of pairs who are estimated to be more closely related than cousins (r 0.125) or half-siblings (r 0.25). For this subset of individuals, we will determine whether the pairwise geographic distances are shorter for the males or the females in the sample (Coulon et al. (2006)). We will perform 10,000 random draws of pairs of males and of females matching the numbers of inferred closely related dyads, and calculate the difference between the average geographic distances for each sex. We will assess whether the observed difference in geographic distances is higher than the majority of random samples and, for comparison with other approaches, determine whether the observed distance is higher than that calculated for 95% of all random draws.

```
options(width = 60)
matrix(runif(100), ncol = 20)

# Analysis 2: Assess whether distances
# among closely related females are
# shorter than distances among closely
# related males

# Load relevant files
input <- readgenotypedata("grackle_genepop_converted.txt")
gracklelocations <- read.csv("GrackleSNPdata_Individuals_Location_Sex.csv")

# Calculate relatedness among pairs of
# individuals
outfile <- coancestry(input$gdata, wang = 1)
pairwise_r <- outfile$relatedness

# Calculate all pairwise distances
all_pairwise_distances <- distm(gracklelocations[,
    c("Lon", "Lat")], gracklelocations[,
    c("Lon", "Lat")], fun = distVincentyEllipsoid)
rownames(all_pairwise_distances) <- gracklelocations$Individual
colnames(all_pairwise_distances) <- gracklelocations$Individual
diag(all_pairwise_distances) <- NA

# Calculate pairwise distances among all
# the females
female_pairwise_distances <- distm(gracklelocations[gracklelocations$Sex ==
    "F", c("Lon", "Lat")], gracklelocations[gracklelocations$Sex ==
    "F", c("Lon", "Lat")], fun = distVincentyEllipsoid)
rownames(female_pairwise_distances) <- gracklelocations[gracklelocations$Sex ==
    "F", ]$Individual
colnames(female_pairwise_distances) <- gracklelocations[gracklelocations$Sex ==
    "F", ]$Individual
diag(female_pairwise_distances) <- NA

# Calculate pairwise distances among all
# the females
male_pairwise_distances <- distm(gracklelocations[gracklelocations$Sex ==
    "M", c("Lon", "Lat")], gracklelocations[gracklelocations$Sex ==
    "M", c("Lon", "Lat")], fun = distVincentyEllipsoid)
rownames(male_pairwise_distances) <- gracklelocations[gracklelocations$Sex ==
```

```
    "M", ]$Individual
colnames(male_pairwise_distances) <- gracklelocations[gracklelocations$Sex ==
    "M", ]$Individual
diag(male_pairwise_distances) <- NA


# First define close relatives as all
# pairs of individuals who are related by
# a level of 0.25 or higher
# (half-siblings or higher)
close_relatives_females <- filter(pairwise_r,
    wang > 0.2499, ind1.id %in% females,
    ind2.id %in% females)
close_relatives_females_individuals <- c(close_relatives_females$ind1.id,
    close_relatives_females$ind2.id)
# Next subset the the distance matrix to
# only include these individuals
close_relatives_females_pairwise_distances <- as.data.frame(female_pairwise_distances)
close_relatives_females_pairwise_distances <- select(close_relatives_females_pairwise_distances,
    close_relatives_females_individuals)
close_relatives_females_pairwise_distances <- close_relatives_females_pairwise_distances[rownames(close_
    close_relatives_females_individuals,
    ]
close_relatives_females_pairwise_distances <- as.matrix(close_relatives_females_pairwise_distances)
diag(close_relatives_females_pairwise_distances) <- NA
median(close_relatives_females_pairwise_distances,
    na.rm = T)


# repeat the same for the males
close_relatives_males <- filter(pairwise_r,
    wang > 0.2499, ind1.id %in% males, ind2.id %in%
        males)
close_relatives_males_individuals <- c(close_relatives_males$ind1.id,
    close_relatives_males$ind2.id)
close_relatives_males_pairwise_distances <- as.data.frame(male_pairwise_distances)
close_relatives_males_pairwise_distances <- select(close_relatives_males_pairwise_distances,
    close_relatives_males_individuals)
close_relatives_males_pairwise_distances <- close_relatives_males_pairwise_distances[rownames(close_rela
    close_relatives_males_individuals, ]
close_relatives_males_pairwise_distances <- as.matrix(close_relatives_males_pairwise_distances)
diag(close_relatives_males_pairwise_distances) <- NA
median(close_relatives_males_pairwise_distances,
    na.rm = T)


# calculate difference between the
# distances among males and among females
observeddifferenceindistances <- median(close_relatives_males_pairwise_distances,
    na.rm = T) - median(close_relatives_females_pairwise_distances,
    na.rm = T)


# perform simulation to generate random
# draws of matching numbers of
# individuals to assess whether the
```

```r
# sex-difference in the distance is more
# or less than what would be expected by
# chance
number_close_relatives_females <- nrow(close_relatives_females)
number_close_relatives_males <- nrow(close_relatives_males)

simulateddifferencesindistances <- matrix(ncol = 1,
    nrow = 10000)

for (i in 1:10000) {
    simulated_close_relatives_females <- sample_n(pairwise_r,
        number_close_relatives_females, replace = TRUE)
    close_relatives_females_individuals <- c(simulated_close_relatives_females$ind1.id,
        simulated_close_relatives_females$ind2.id)
    close_relatives_females_pairwise_distances <- as.data.frame(all_pairwise_distances)
    close_relatives_females_pairwise_distances <- select(close_relatives_females_pairwise_distances,
        close_relatives_females_individuals)
    close_relatives_females_pairwise_distances <- close_relatives_females_pairwise_distances[rownames(c)
        close_relatives_females_individuals,
        ]
    close_relatives_females_pairwise_distances <- as.matrix(close_relatives_females_pairwise_distances)
    diag(close_relatives_females_pairwise_distances) <- NA

    simulated_close_relatives_males <- sample_n(pairwise_r,
        number_close_relatives_males, replace = TRUE)
    close_relatives_males_individuals <- c(simulated_close_relatives_males$ind1.id,
        simulated_close_relatives_males$ind2.id)
    close_relatives_males_pairwise_distances <- as.data.frame(all_pairwise_distances)
    close_relatives_males_pairwise_distances <- select(close_relatives_males_pairwise_distances,
        close_relatives_males_individuals)
    close_relatives_males_pairwise_distances <- close_relatives_males_pairwise_distances[rownames(close_
        close_relatives_males_individuals,
        ]
    close_relatives_males_pairwise_distances <- as.matrix(close_relatives_males_pairwise_distances)
    diag(close_relatives_males_pairwise_distances) <- NA

    simulateddifferencesindistances[i, 1] <- median(close_relatives_males_pairwise_distances,
        na.rm = T) - median(close_relatives_females_pairwise_distances,
        na.rm = T)
}

sum(simulateddifferencesindistances > observeddifferenceindistances)/10000
```

Analysis iii: spatial autocorrelation

To test whether males and females show different patterns of genetic isolation by geographic distance, we will follow analyses as in Aguillon et al. (2017). For the analysis, we will initially create 11 distance bins separated by 200m between 0m-2000m (the maximum distance between trapping sites). The 200m bin size was chosen because there are roosting trees that are ~50m apart suggesting that dispersal might be occuring below this scale and also to maximize the number of pairs in each distance class. The individuals in our sample were caught at one of 15 trap sites, and the resulting 105 pairwise distances among individuals will be assigned to one of the 11 bins. In case the pairwise distances among individuals are too small in each bin to detect a signal, then we will reduce the number of bins. In addition, we might adjust the distances covered by each bin to have shorter distances for the first few bins to increase the chance to detect relatives within

the smallest bins (changing from 11 equally sized 200m bins to, for example, 9 bins at varying distances such as 0-50m, 50m-100m, 100m-150m, 150m-200m, 200m-500m, 500m-750m, 750m-1000m, 1000m-1500m, 1500m-2000m) (following Peakall, Ruibal, and Lindenmayer (2003)). For males and females separately, we will link the matrices of average relatedness and of geographic distance between all pairs of individuals by first plotting genetic relatedness against geographic distance and next by assessing the strength of their association using Mantel correlograms. We will use the function 'mantel.correlog' in the vegan package (Oksanen et al. (2013)) in R, performing 10,000 permutations to assess the strength of the association. This approach relies on the establishment of the multivariate Mantel correlogram by Legendre and Legendre (2012). The approach relies on partitioning the geographic locations into a series of discrete distance classes. The result of this set of analyses is a Mantel's correlogram, analogous to an autocorrelation function but performed on a set of distance matrices. For each distance class, a separate matrix is generated and codes whether a given geographic distance between a pair of individuals falls within that range or not. A normalized Mantel statistic is calculated using permutations for each distance class. The permutation statistics, plotted against distance classes, produce a multivariate correlogram. These analyses are performed separately for each sex to determine whether isolation-by-distance might occur and indicate dispersal of the individuals of that sex. A stronger negative correlation between genetic relatedness and spatial distance for males than for females would indicate that males disperse shorter distances than females, and in particular we expect that males captured at the same trapping site will be much more closely related to each other than females captured at the same trapping site.

```r
# Analysis 3: Correlogram to assess change of relatedness
# with distances

# have each value only once in the distance matrix
for (i in 1:ncol(all_pairwise_distances)) {
    all_pairwise_distances[i, i:ncol(all_pairwise_distances)] <- NA
}

all_pairwise_distances <- all_pairwise_distances[2:10, 1:9]

# turn pairwise_r$wang into a matrix
all_relatedness <- select(pairwise_r, ind1.id, ind2.id, wang)
relatedness_matrix <- spread(all_relatedness, "ind1.id", "wang")
relatedness_matrix[5, 1] <- "AZ_10"
relatedness_matrix <- arrange(relatedness_matrix, ind2.id)
relatedness_matrix <- relatedness_matrix[, 2:10]

# perform the correlogram analysis first way, defining the
# distance classes
mantel.correlog(D.eco = relatedness_matrix, D.geo = all_pairwise_distances,
    break.pts = c(0, 100, 200, 300, 400, 500, 750, 1250, 1550,
        2000, 2500))
# second way, setting the number of distance classes
mantel.correlog(D.eco = relatedness_matrix, D.geo = all_pairwise_distances,
    n.class = 11)
```

E. ETHICS

This research is carried out in accordance with permits from the:

1) US Fish and Wildlife Service (scientific collecting permit number MB76700A-0,1,2)
2) US Geological Survey Bird Banding Laboratory (federal bird banding permit number 23872)
3) Arizona Game and Fish Department (scientific collecting license number SP594338 [2017], SP606267 [2018], and SP639866 [2019])
4) Institutional Animal Care and Use Committee at Arizona State University (protocol number 17-1594R)
5) University of Cambridge ethical review process (non-regulated use of animals in scientific procedures:

zoo4/17 [2017])

F. AUTHOR CONTRIBUTIONS

Sevchik: Hypothesis development, sample processing, data analysis and interpretation, write up, revising/editing.

Logan: Hypothesis development, data analysis and interpretation, write up, revising/editing, materials/funding.

Folsom: Blood collection, revising/editing.

Bergeron: Blood collection, revising/editing.

Blackwell: Hypothesis development, DNA extraction, revising/editing.

Rowney: DNA extraction, write up, revising/editing.

Lukas: Hypothesis development, data analysis and interpretation, write up, revising/editing, materials/funding.

I. CONFLICT OF INTEREST DISCLOSURE

We, the authors, declare that we have no financial conflicts of interest with the content of this article. Corina Logan and Dieter Lukas are Recommenders at PCI Ecology and Corina Logan is on the Managing Board at PCI Ecology.

K. REFERENCES

Aguillon, Stepfanie M, John W Fitzpatrick, Reed Bowman, Stephan J Schoech, Andrew G Clark, Graham Coop, and Nancy Chen. 2017. "Deconstructing Isolation-by-Distance: The Genomic Consequences of Limited Dispersal." *PLoS Genetics* 13 (8). Public Library of Science: e1006911.

Banks, Sam C, and ROD Peakall. 2012. "Genetic Spatial Autocorrelation Can Readily Detect Sex-Biased Dispersal." *Molecular Ecology* 21 (9). Wiley Online Library: 2092–2105.

Botero-Delgadillo, Esteban, Verónica Quirici, Yanina Poblete, Élfego Cuevas, Sylvia Kuhn, Alexander Girg, Kim Teltscher, Elie Poulin, Bart Kempenaers, and Rodrigo A Vásquez. 2017. "Variation in Fine-Scale Genetic Structure and Local Dispersal Patterns Between Peripheral Populations of a South American Passerine Bird." *Ecology and Evolution* 7 (20). Wiley Online Library: 8363–78.

Coulon, A, J-F Cosson, N Morellet, J-M Angibault, B Cargnelutti, M Galan, S Aulagnier, and AJM Hewison. 2006. "Dispersal Is Not Female Biased in a Resource-Defence Mating Ungulate, the European Roe Deer." *Proceedings of the Royal Society B: Biological Sciences* 273 (1584). The Royal Society London: 341–48.

Goudet, Jérôme, Nicolas Perrin, and Peter Waser. 2002. "Tests for Sex-Biased Dispersal Using Bi-Parentally Inherited Genetic Markers." *Molecular Ecology* 11 (6). Wiley Online Library: 1103–14.

Gour, Digpal Singh, Jyotsna Bhagavatula, Maradani Bhavanishankar, Patlolla Anuradha Reddy, Jaya A Gupta, Mriganka Shekhar Sarkar, Shaik Mohammed Hussain, Segu Harika, Ravinder Gulia, and Sisinthy

Shivaji. 2013. "Philopatry and Dispersal Patterns in Tiger (Panthera Tigris)." *PLoS One* 8 (7). Public Library of Science: e66956.

Greenwood, Paul J. 1980. "Mating Systems, Philopatry and Dispersal in Birds and Mammals." *Animal Behaviour* 28 (4). Elsevier: 1140–62.

Greenwood, Paul J, and Paul H Harvey. 1982. "The Natal and Breeding Dispersal of Birds." *Annual Review of Ecology and Systematics* 13 (1). Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA: 1–21.

Hofmann, Sylvia, Peter Fritzsche, Torstein Solhøy, Tsering Dorge, and Georg Miehe. 2012. "Evidence of Sex-Biased Dispersal in Thermophis Baileyi Inferred from Microsatellite Markers." *Herpetologica* 68 (4). The Herpetologists' League, Inc. 1011 E. Ash Street, Springfield, IL 62703: 514–22.

Johnson, Kristine, Emily DuVal, Megan Kielt, and Colin Hughes. 2000. "Male Mating Strategies and the Mating System of Great-Tailed Grackles." *Behavioral Ecology* 11 (2). Oxford University Press: 132–41.

Legendre, Pierre, and Loic FJ Legendre. 2012. *Numerical Ecology.* Vol. 24. Elsevier.

Mabry, Karen E, Erin L Shelley, Katie E Davis, Daniel T Blumstein, and Dirk H Van Vuren. 2013. "Social Mating System and Sex-Biased Dispersal in Mammals and Birds: A Phylogenetic Analysis." *PloS One* 8 (3). Public Library of Science: e57980.

Oksanen, Jari, F Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R Minchin, RB O'hara, Gavin L Simpson, et al. 2013. "Package 'Vegan'." *Community Ecology Package, Version* 2 (9): 1–295.

Peakall, Rod, Monica Ruibal, and David B Lindenmayer. 2003. "Spatial Autocorrelation Analysis Offers New Insights into Gene Flow in the Australian Bush Rat, Rattus Fuscipes." *Evolution* 57 (5). Wiley Online Library: 1182–95.

Pew, Jack, Paul H Muir, Jinliang Wang, and Timothy R Frasier. 2015. "Related: An R Package for Analysing Pairwise Relatedness from Codominant Molecular Markers." *Molecular Ecology Resources* 15 (3). Wiley Online Library: 557–61.

Quaglietta, Lorenzo, Vania C Fonseca, Petra Hájková, António Mira, and Luigi Boitani. 2013. "Fine-Scale Population Genetic Structure and Short-Range Sex-Biased Dispersal in a Solitary Carnivore, Lutra Lutra." *Journal of Mammalogy* 94 (3). American Society of Mammalogists 810 East 10th Street, PO Box 1897, Lawrence …: 561–71.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org.

Sutherland, Glenn D, Alton S Harestad, Karen Price, and Kenneth P Lertzman. 2000. "Scaling of Natal Dispersal Distances in Terrestrial Birds and Mammals." *Conservation Ecology* 4 (1). JSTOR.

Thrasher, Derrick J, Bronwyn G Butcher, Leonardo Campagna, Michael S Webster, and Irby J Lovette. 2018. "Double-Digest Rad Sequencing Outperforms Microsatellite Loci at Assigning Paternity and Estimating Relatedness: A Proof of Concept in a Highly Promiscuous Bird." *Molecular Ecology Resources* 18 (5). Wiley Online Library: 953–65.

Végvári, Zsolt, Gergely Katona, Balázs Vági, Robert P Freckleton, Jean-Michel Gaillard, Tamás Székely, and András Liker. 2018. "Sex-Biased Breeding Dispersal Is Predicted by Social Environment in Birds." *Ecology and Evolution* 8 (13). Wiley Online Library: 6483–91.

Wang, Jinliang, and Anna W Santure. 2009. "Parentage and Sibship Inference from Multilocus Genotype Data Under Polygamy." *Genetics* 181 (4). Genetics Soc America: 1579–94.

White, P.S., and L.D. Densmore. 1992. "Mitochondrial Dna Isolation." In *Molecular Genetic Analysis of Populations: A Practical Approach*, 50–51. IRL Press at Oxford University Press.