Southern Illinois University Carbondale
OpenSIUC

Publications | Department of Zoology

1-1-2010

Quantifying Male-biased Dispersal among Social Groups in the Collared Peccary (Pecari tajacu) Using Analyses Based on mtDNA Variation.

J D Cooper

R Vitalis

P M Waser

D Gopurenko

E C Hellgren

See next page for additional authors

Follow this and additional works at: http://opensiuc.lib.siu.edu/zool_pubs
Copyright 2010 Macmillan Publishers.
Published in *Heredity*, Vol. 104, No. 1 (January 2010), at 10.1038/hdy.2009.102.

Recommended Citation
Cooper, J D, Vitalis, R, Waser, P M, Gopurenko, D, Hellgren, E C, Gabor, T M and DeWoody, J A. "Quantifying Male-biased Dispersal among Social Groups in the Collared Peccary (Pecari tajacu) Using Analyses Based on mtDNA Variation.." 104, No. 1 (Jan 2010).

This Article is brought to you for free and open access by the Department of Zoology at OpenSIUC. It has been accepted for inclusion in Publications by an authorized administrator of OpenSIUC. For more information, please contact opensiuc@lib.siu.edu.

Authors

J D Cooper, R Vitalis, P M Waser, D Gopurenko, E C Hellgren, T M Gabor, and J A DeWoody

This article is available at OpenSIUC: http://opensiuc.lib.siu.edu/zool_pubs/32

1 **Quantifying male-biased dispersal among social groups in the collared peccary (*Pecari*
2 *tajacu*) using analyses based on mtDNA variation**
3

4 J.D. Cooper[1], R. Vitalis[2], P.M. Waser[1], D. Gopurenko[3], E.C. Hellgren[4], T.M. Gabor[5], J.A.
5 DeWoody[6]
6 [1]Biological Sciences, Purdue University, West Lafayette, IN, 47907, USA, (765) 494-8129,
7 pwaser@bilbo.bio.purdue.edu;
8 [2]Muséum National d'Histoire Naturelle – Centre National de la Recherche Scientifique UMR
9 7206 – Université Paris 7, Éco-Anthropologie et Ethnobiologie, Musée de l'Homme, 17 place du
10 Trocadéro, 75116 Paris, France; +33 1 44057268, vitalis@mnhn.fr;
11 [3]NSW Dept. Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW 2650,
12 AUSTRALIA, +011 (61 2) 6938-1946, david.gopurenko@dpi.nsw.gov.au
13 [4]Cooperative Wildlife Research Laboratory, Department of Zoology, Southern Illinois
14 University, Carbondale, IL, 62901, USA, (618) 453-6941, hellgren@siu.edu;
15 [5]Biology, TRF # 265E, Northland Community & Technical College, 1101 Hwy One East, Thief
16 River Falls, MN 56701, (218) 683-8730, tim.gabor@northlandcollege.edu;
17 [6]Forestry and Natural Resources, Purdue University, West Lafayette, IN, 47907, USA, (765) 496-
18 6109, dewoody@purdue.edu
19 Corresponding author: J. Andrew DeWoody
20 dewoody@purdue.edu; (765) 496-6109; fax: (765) 496-2422

21 Keywords: F-statistics, hierarchical analyses, mitochondrial DNA, philopatry, population
22 genetics, sex-biased dispersal
23 Running head: Genetic detection of sex-biased dispersal

24 Word count: 4532

25 **Abstract**

26 Recent advances in the statistical analysis of microsatellite data permit calculation of sex-specific
27 dispersal rates through sex-and age-specific comparisons of genetic variation. This approach,
28 developed for analysis of data derived from co-dominant autosomal markers, should be
29 applicable to a sex-specific marker such as mitochondrial DNA. To test this premise, we
30 amplified a 449bp control region DNA sequence from the mitochondrial genome of the collared
31 peccary (*Pecari tajacu),* and estimated intra-class correlations among herds sampled from 3
32 Texas populations. Analyses on data partitioned by breeding group showed a clear signal of
33 male-biased dispersal; sex-specific fixation indices associated with genetic variation among
34 social groups within populations yielded values for females (F_{GP} = 0.91) which were
35 significantly larger than values for males (F_{GP} = 0.24; p = 0.0015). The same general pattern
36 emerged when the analyses were conducted on age classes (albeit nonsignificantly), as well as
37 categories of individuals which were predicted *a posteriori* to be dispersers (adult males) and
38 philopatric (adult females and all immatures). By extending a previously published methodology
39 based on bi-parentally-inherited markers to matrilineally-inherited haploid data, we calculated
40 sex-specific rates of contemporary dispersal among social groups within populations ($m_{♂}$ = 0.37).
41 These results support the idea that mitochondrial DNA haplotype frequency data can be used to
42 estimate sex-specific instantaneous dispersal rates in a social species.

43 **Introduction**

44 Sex bias in natal dispersal is common; in most mammalian species, males are dispersers
45 while females are philopatric, and the opposite trend is exhibited in birds (Greenwood 1980).
46 Exploring why the sexes differ in their dispersal patterns can shed light on the evolutionary
47 causes of dispersal in general (Goudet *et al.* 2002), and accurate characterization of dispersal
48 behavior is integral to our understanding of the social structure, mating system, and population
49 genetic structure of a species. Yet detection of sex-biased dispersal can be tricky because a
50 dispersal event may occur once in an animal's lifetime, and such events can be difficult to
51 observe directly.
52

53 *Measuring dispersal*

54 In the last few decades molecular genetics has provided a means of investigating sex-biased
55 dispersal within and among populations (reviewed in Lawson-Handley and Perrin 2007).
56 Several powerful approaches have been developed to detect individual dispersers through
57 assignment tests or to characterize general patterns of dispersal through summary statistics of
58 population genetic structure (*F*-statistics, relatedness). Most of these approaches utilize
59 autosomal microsatellites as molecular markers, either alone (Goudet *et al.* 2002; Mossman and
60 Waser 1999; Petit *et al.* 2001; Waser *et al.* 2001) or in tandem with a uni-parentally inherited
61 marker such as mitochondrial DNA (mtDNA) or a Y chromosome locus (Escorza-Trevino and
62 Dizon 2000; Girman *et al.* 1997). The expectation inherent to all these approaches is that greater
63 genetic structure will be evident in the philopatric sex compared to the dispersing sex, thus
64 comparisons of sex-specific F_{ST} estimates should reveal the direction (and suggest the relative
65 strength) of sex-bias in dispersal (Goudet *et al.* 2002).

66 Because mitochondrial DNA is matrilineally inherited, it is commonly used to infer female-
67 biased dispersal rates (Prugnolle and de Meeus 2002). When mtDNA haplotype distribution
68 patterns are examined in isolation, inferences can be made about female dispersal behavior
69 without respect to males, but this approach is qualitative and not widely applied (Hoelzer *et al.*
70 1994). However, it is possible to use mtDNA alone to infer the relative dispersal of both sexes
71 by extending methods developed for autosomal, bi-parentally inherited markers. For instance,
72 the comparisons of sex-specific population differentiation from haplotype frequency data can
73 indicate which sex disperses more (Escorza-Trevino and Dizon 2000; Yang *et al.* 2003).
74

75 *Using sex-specific fixation indices to estimate instantaneous dispersal rates*

76 Vitalis (2002) developed a method to quantitatively measure sex bias in instantaneous
77 dispersal rates using data from bi-parentally inherited markers such as microsatellites. This
78 approach allows the inference of sex-specific dispersal rates by comparing sex-specific estimates
79 of genetic differentiation (F_{ST}) measured before and after dispersal. This intuitive method can be
80 further extended to incorporate the hierarchical structure within social species (Fontanillas *et al*.
81 2004), as it has been recognized that social organization can strongly influence correlations of
82 gene frequencies (Chesser 1991; Chesser and Baker 1996; Slatkin and Voelm 1991; Sugg and
83 Chesser 1994; Vigouroux and Couvet 2000). Herein we develop and use an extension of the
84 Vitalis' (2002) method to estimate instantaneous dispersal rates through analyses of mtDNA
85 haplotype distribution patterns in a social mammal, the collared peccary (*Pecari tajacu*, family
86 *Tayassuidae).*

87 We sampled extensively within three populations separated by long distances, with the goal
88 of quantifying local dispersal among breeding groups within populations. We then compared

89 sex- and age-specific estimates of population differentiation based solely on mtDNA haplotype
90 frequencies, using probability based estimates of intra-class correlations of gene frequencies
91 among social groups within populations. We used a resampling approach to test for the
92 significance of the observed age and sex-bias in dispersal. Last, the fixation indices generated by
93 these analyses were used to calculate single-generation sex-specific dispersal rates. Heretofore
94 mtDNA has been used primarily to infer female dispersal patterns, but we demonstrate that this
95 matrilineally-inherited genetic marker can be used to quantify male dispersal rates in the absence
96 of nuclear population genetic data.
97

98 **Materials and Methods**

99 *Study species*

100 The collared peccary is a socially complex, pig-like ungulate that forms stable, mixed sex
101 herds of 3 to 30 individuals (Sowls, 1978). These groups associate throughout the year and
102 vigorously defend territories against other social groups (Bissonette 1982; Hellgren *et al.* 1984;
103 Ellisor and Harwell 1969). Herds are socially cohesive and attempts to immigrate may be met
104 with aggression, although direct observational data on dispersal behavior are still scarce. Male
105 exchange between groups and solitary wandering of both sexes has been observed but natal
106 dispersal has not been adequately described (Day 1985; Ellisor and Harwell 1969; Gabor and
107 Hellgren 2000). Heretofore little population genetic data existed for *P. tajacu* (but see Gongora
108 *et al*. 2006). Theimer and Keim (1994) utilized mtDNA variation to measure sequence
109 divergence and geographic partitioning in Arizona populations, but their samples were not
110 associated with social groups. There was sufficient heterogeneity in mtDNA haplotype
111 distribution to indicate limited female dispersal across regions (rather than among neighboring

112 herds as is considered here), although it was not clear if the patterns observed were also a
113 signature of founding events (Theimer and Keim 1994).
114
115 *Sampling*
116 Data were collected from three wild populations of *P. tajacu* in Texas. In the mid-1990s,
117 102 whole blood samples were collected from the Chaparral Wildlife Management Area
118 (CWMA) in south Texas (Gabor and Hellgren 2000). These samples were taken from live-
119 trapped animals from 13 social groups, but not all group members were sampled. In 2005, we
120 collected 31 ear snip tissue samples from live-trapped animals from 4 groups in the Welder
121 Wildlife Refuge (WWR) in south Texas. In 2006-2007 we similarly sampled 134 animals from
122 13 groups in Big Bend Ranch State Park (BB) in west Texas, along the Texas-Mexico border.
123 The WWR and BB populations were sampled extensively; every social group at these locations
124 was identified through direct and remote camera observation and trapped in large corrals over
125 several sessions. Groups ranged in size from 2 to 18 animals and mean group size was 8.9.
126 Individuals were uniquely marked with numbered ear tags and the strongest possible effort was
127 made to trap and sample every unmarked individual. All samples include associated data on age
128 class (adult, subadult, juvenile, infant), sex, territory location and social group affiliation. Age
129 class was assigned according to behavior and morphological traits such as pelage, body size and
130 testicular development. Individuals exhibiting immature characteristics such as ginger or spotted
131 pelage, undescended or partially descended testicles, adult-oriented following behavior, or
132 estimated body size of less than 9 kg were classed as infants or juveniles, while individuals
133 which weighed 10-13 kg were classed as "subadults" that were on the cusp of sexual maturity.
134 Whole blood samples were frozen at -20°C, and tissue samples were stored in lysis buffer at

135 room temperature until DNA was extracted for long-term storage at 4°C.

136

137 *Genetic analysis*

138 Blood clot samples (~ 0.5 g) were digested by rotating for 12 hours at 55°C in 750 µL of
139 lysis buffer (100 mM Tris-Cl pH 8, 10 mM EDTA, 1% SDS, ddH_2O), 40 µL of proteinase K (10
140 mg/mL) and 2 µL of streptokinase (10 U/µL). Tissue samples (~ 5 X 5 mm) were digested by
141 rotating for 24 hours at 55°C in 750 µL of lysis buffer and 20 µL of proteinase K (10 mg/mL).
142 Genomic DNA was extracted from blood using a standard phenol-chloroform method, and from
143 tissue samples using either a phenol-chloroform-isopropanol method or ammonium acetate
144 method (Sambrook and Russell 2001). All DNA precipitations were washed twice in 70%
145 ethanol, and DNA pellets were resuspended in 250 µL of TLE (10 mM Tris-Cl, 0.1 mM EDTA).
146 A 449 bp region between sites 15,390 and 15,900 of the collared peccary mtDNA D-loop was
147 amplified from genomic DNA using porcine primers (Alves *et al.* 2003). This sequence lies in
148 the hypervariable 5' end of the mitochondrial control region and does not code for any known
149 protein product. PCR volumes were 25 µL and contained final concentrations of the following
150 reagents: 1.5 mM $MgCl_2$; 0.5 µM each primer; 0.21 mM dNTPs; 1.25 U *Taq* polymerase (NEB).
151 PCRs were performed in an Eppendorf MasterCycler using the following temperature profile:
152 denaturation for 3 min. at 94° C, followed by 30 cycles of 94° C for 4 s, 55° C for 4 s, and 72° C
153 for 12 s; finishing with a 15 min. extension step at 72° C. PCR products were cleaned using a
154 low sodium protocol; 28 µL of a mixture containing 500 ml of absolute ethanol and 20 µL of 3M
155 NaOAc (pH 5.2) was added to each sample, shaken for 15 min, and centrifuged at 2051 g for 35
156 min. This step was followed by 70% ethanol precipitation under centrifugation (twice) and
157 resuspension in 20 µL ddH_20.

158 PCR products were then directly sequenced in both directions using Big Dye 3.1
159 chemistry. Sequencing products were purified using the low sodium protocol described above,
160 and then electrophoresed using an AB Prism 3730XL sequencer (Applied Biosystems).
161 Sequence data were aligned and edited with Sequencher 4.5 (Gene Codes). Nuclear copies of
162 mtDNA genes (numts) can greatly confound evolutionary analyses, and we avoided numts using
163 methods described in Triant and DeWoody (2007). For example, a few individuals (<5%),
164 harbored apparently heterozygous sites so we reamplified their DNA and completely
165 resequenced the amplicons in both directions. In every case, this procedure completely resolved
166 the mismatch and suggested the initial discrepancy was probably a result of *Taq* error.
167 We converted sequences into NEXUS format and imported them into PAUP* 4.0
168 (Swofford 2003) for haplotype assignment. Haplotypes were determined through reconstruction
169 of unrooted phylogenetic trees using a neighbor-joining algorithm. Direct sequencing of a sub-
170 set of the CWMA population revealed that some of the mtDNA haplotypes could be
171 discriminated by restriction digest with the *Mbo*I enzyme, but all individuals from WWR and BB
172 were typed by direct sequencing.
173
174 *Statistical analyses*
175 **Among-populations differentiation**
176 MtDNA haplotype frequencies were calculated by hand for all three populations. Genetic
177 differentiation among populations was inferred from F_{ST} estimates (Weir and Cockerham 1984)
178 and exact tests of population differentiation (Raymond and Rousset 1995) using the software
179 package Arlequin Version 3.1 (Excoffier *et al.* 2005). For the latter, *p*-values were estimated
180 from a Markov chain set to 110,000 steps including 10,000 dememorization steps. All analyses

181 were based on pure haplotype frequency data rather than nucleotide differences.

182

183 **Within-populations differentiation**

184 Because *P. tajacu* populations are subdivided into breeding groups, we incorporated

185 breeding group as a hierarchical level. We calculated identity probabilities by simple counting of

186 identical pairs of genes at different hierarchical levels (Q_1 for pairs of genes within groups, Q_2

187 for pairs of genes sampled among groups within populations, and Q_3 for pairs of genes sampled

188 in different populations). We then estimated the intra-class correlations by taking appropriate

189 ratios of identity probabilities, weighted according to the number of pairs in each sample (see

190 Rousset 2007), following the definitions of *F*-statistics as functions of identity probabilities

191 between pairs of genes (see Appendix). Since the distances among populations are large in this

192 study (range of 225 km to 945 km between the three sampling sites), we considered the three

193 populations as independent replicates in the analysis, and we restricted our analyses to estimate

194 within-population dispersal. We focused on the level of genetic differentiation among social

195 groups within populations as measured by the parameter F_{GP}. The notation is adapted from

196 Wright (1965). This approach is different from that of Fontanillas *et al*. (2004) who considered

197 dispersal both among populations and among breeding groups. Although the samples from each

198 site were collected in different years, F_{GP} estimates do not depend upon identity between pairs of

199 genes from different populations and temporally discontinuous sampling is therefore unlikely to

200 undermine the approach. We employed a bootstrapping procedure to calculate confidence limits

201 around estimates of F_{GP} for each class of individuals. Using the statistical software package R

202 (R Development Core Team 2008), we generated 25,000 bootstrap samples, with each sample

203 being produced by random resampling (with replacement) of the 255 nucleotide sites from the

204 mtDNA haplotypes (254 sites + 1 indel). This allowed us to calculate F_{GP} estimates for each
205 sample and generate a distribution; confidence intervals endpoints were then calculated as the
206 2.5% and the 97.5% percentiles of this distribution. This procedure is strictly equivalent to that
207 implemented in the software package Arlequin Version 3.1 (Excoffier *et al.* 2005) to generate
208 95% confidence limits by bootstrapping genetic differentiation values in a locus-by-locus
209 AMOVA (see, e.g., Langergraber et al. 2007).

210

211 **Class-specific analyzes**

212 Dispersal is a trait that can be partitioned into pre- and post-dispersal conditions,
213 therefore our first analysis partitioned the data by age. We performed independent analyses on
214 data partitioned into two age sets, respectively for adults and immatures (the latter including both
215 juveniles and infants). Subadults were classed as immatures and then as adults in sequential
216 analyses. Each age-specific data set was composed of individuals assigned to their respective
217 populations and social groups, and intra-class correlations (F_{GP}) were calculated among social
218 groups within populations from identity probabilities of pairs of genes (see above). Only those
219 social groups containing a representative individual from each treatment were included in the
220 analysis (e.g. in the independent analyses on adult and immature data sets, a social group must
221 have contained at least 1 adult and 1 immature to be included). We then duplicated the analysis
222 with the data partitioned by sex rather than age. From these results, we were able to distinguish a
223 putative class of dispersing individuals, from a putative class of non-dispersers. We therefore
224 performed *a posteriori*, independent analyses on data sets of putative dispersers and non-
225 dispersers.

226

227 We used a resampling scheme after Goudet *et al.* (2002) to test whether the estimated
228 fixation indices among social groups within replicate populations (F_{GP}) for specific classes (age,
229 sex, or putative dispersal class) departed significantly from the null hypothesis that dispersal is
230 independent from the class of individuals. Resampling tests were all performed with the
231 statistical software package R (R Development Core Team 2008). For each class, we generated
232 25,000 randomized datasets, by re-assigning the age (or sex, or dispersal class) of each haplotype
233 randomly within each breeding group. By doing so, we kept the number of individuals from
234 each class constant within each breeding group. We calculated the probabilities of identities
235 between pairs of genes for each resampled dataset, and obtained the distribution of class-specific
236 F_{GP} estimates under the null hypothesis that dispersal behavior or capability is independent from
237 age, sex, or dispersal class. We then calculated *p*-values as the proportion of times where F_{GP}
238 from the randomized datasets was larger than or equal to the observed F_{GP} on the original
239 dataset.

240

241 **Estimating dispersal**

242 To calculate a sex-specific dispersal rate within a single generation, we adapted Vitalis'
243 (2002) approach and extended it to mtDNA data. In Vitalis (2002), the ratio of the sex-specific
244 differentiation evaluated after juvenile dispersal ($\hat{F}_{GP}^{XX}$) divided by the differentiation evaluated
245 before dispersal ($\hat{F}_{GP}^{*}$) gives the sex-specific dispersal rate. Appendix 1 shows that this
246 relationship also applies to uni-parentally inherited markers, and:

247

248 $$\hat{m}_X \approx 1-\sqrt{\frac{\hat{F}_{GP}^{XX}}{\hat{F}_{GP}^{*}}} \qquad \text{for all } X \in \{♂,♀\} \qquad (1)$$

249

250 gives the sex-specific dispersal rate. Here we use this simple model to compare fixation indices
251 before and after dispersal at the within-population level, focusing on dispersal of individuals
252 among breeding groups. This equation assumes that the number of breeding groups, n, is large
253 (infinite); by considering an infinitely large n, we slightly overestimate dispersal rate m_x (e.g.,
254 10% relative bias with $n = 10$). We estimated instantaneous sex-specific dispersal rates for *P.*
255 *tajacu* by applying equation 1, using fixation indices estimates for adult males, adult females and
256 all immatures of both sexes (Table 2). Confidence intervals for dispersal rates were obtained by
257 means of a bootstrap procedure, similar to that used for F_{GP} (see above), modified as follows.
258 For each bootstrap sample, F_{GP} estimates were calculated for adult males (resp. adult females)
259 and all immatures, and male- (resp. female-) specific migration rates were calculated using
260 equation (1). Confidence intervals for sex-specific dispersal rates were then derived from the
261 0.025 and 0.975 percentiles of the bootstrap distribution.

262

263 **Results**

264 *mtDNA haplotype distribution patterns*

265 A total of 18 nucleotide sites were variable (17 substitutions and a single indel) over 449
266 bp. We recovered 6 mtDNA haplotypes from 267 individual collared peccaries among the 3 sites
267 sampled (Table 1). Haplotype A was observed in all sampling sites, but haplotype B was unique
268 to the CWMA, and haplotype C was found in both the WWR and the CWMA. The BB
269 population was almost fixed for haplotype E (96%). Haplotypes F and G were only found in the
270 CWMA, and were represented by single individuals (both males).

271

272 We overlaid mtDNA haplotype distribution onto the social group territory distribution for
273 all populations. At the local level, haplotype distribution did not exhibit geographic structuring
274 in the CWMA or the WWR; all haplotypes present at each sampling site were found distributed
275 throughout that site. In the BB population, haplotype A was found only in the eastern portion of
276 the sampling site. At the regional level across Texas, we observed significant population
277 differentiation. Pairwise F_{ST} estimates ranged from 0.31 to 0.86 between populations and
278 pairwise exact tests of population differentiation were highly significant ($p = 0.001$), indicating
279 that these populations are significantly divergent from one another.
280

281 *Patterns of genetic variation revealed by F-statistics as functions of identity probabilities*

282 Because dispersal status is often dependent upon age, we tested for an age bias in
283 dispersal. To that end, we pooled infants and juveniles (categorized hereafter as “immatures”) in
284 one class, and adults in another class. It was not clear if individuals categorized as subadults
285 were sufficiently developed to be considered as adults, therefore we performed a preliminary
286 analysis on adult-only and immature-only data sets partitioned into social groups, which revealed
287 a decrease in F_{GP} when subadults were included in the adult class (not shown). This result
288 indicates that individual genetic variation in the subadult class is apportioned among rather than
289 within social groups, and therefore subadults were classed as adults in all subsequent analyses.
290 We estimated fixation indices among social groups for each sex, with individuals partitioned into
291 known breeding groups (Table 2). It is clear that F_{GP} for adults (0.30 [0.03, 0.38]) is much
292 smaller than that for immatures (0.60 [0.60, 1.00]), as would be expected if the adult class
293 included dispersed individuals. To test for significance of these quantitative differences, we used
294 a randomization approach, and generated randomized data sets by assigning an age randomly to

295 each mtDNA haplotype. Under the null hypothesis that dispersal is not age-biased, we expect
296 the observed F_{GP} of adults and immatures not to depart significantly from the null distribution.
297 For adults, there was a large proportion of randomized data sets with a differentiation among
298 groups within populations (F_{GP}) larger than the observed, although this proportion did not
299 achieve significance ($p = 0.79$; Fig. 1A). In contrast, for immatures of both sexes, there was only
300 a small proportion of randomized data sets giving a F_{GP} larger than the observed, although the
301 test was not significant ($p = 0.20$; Fig. 1B). In general terms, these results clearly indicate a
302 greater amount of dispersal among social groups for adults when compared to immatures.

303 To test for a signal of sex-biased dispersal, intra-class correlations (F_{GP}) were estimated
304 for each sex with individuals partitioned into known breeding groups (Table 2). It can be seen
305 from these results that F_{GP} among social groups is much smaller for males (0.23 [0.09, 0.36])
306 than it is for females (0.90 [0.87, 1.00]), which indicates that even when pre-dispersal age
307 individuals are included in the male class the sex difference is still apparent. To test the
308 significance of the sex difference, we used a randomization approach identical to the one
309 described for age bias, and generated randomized data sets by assigning a sex randomly to each
310 mtDNA haplotype. For males, there was a very large proportion of randomized data sets with a
311 larger F_{GP} than the observed ($p = 0.98$; Fig. 1C). In contrast, for females, there was only a very
312 small proportion of randomized data sets giving a F_{GP} larger than the observed, and the test was
313 therefore highly significant ($p < 0.001$; Fig. 1D). These results suggest that dispersal is strongly
314 male-biased in *P. tajacu*.

315 The inferred dispersal pattern of *P. tajacu* being of adult male dispersal, we conducted a
316 further analysis, *a posteriori*, on data partitioned by putative dispersal condition: the data were
317 partitioned into “dispersers” (adult males) and “philopatrics” (immature males and all females)

318 and separate analyses performed on individuals assigned to breeding groups. As expected, F_{GP}
319 for the philopatric class was much larger (0.76 [0.73, 0.99]) than was seen for adult males (0.24
320 [0.07, 0.36]). For adult males, there was a very large proportion of randomized data sets with a
321 F_{GP} larger than the observed ($p = 0.99$; Fig. 1E). In contrast, for putative non-dispersers, the test
322 was highly significant, with very few datasets giving a F_{GP} larger than the observed ($p < 0.001$;
323 Fig. 1F).

324

325 *Dispersal rate estimates*

326 The instantaneous sex-specific dispersal rate among social groups within populations was
327 estimated using equation (1). We used the F_{GP} estimates among social groups within populations
328 (Table 2) for adult males (dispersers) ($F_{GP} = 0.24$), for adult females ($F_{GP} = 0.91$), and for pre-
329 dispersal individuals of both sexes (also categorized as "immature"; $F_{GP} = 0.60$). This yielded a
330 male-specific dispersal rate estimate ($m_{♂}$) of 0.37 [0.32, 0.65]. Equation (1) only makes sense if
331 there is a significant difference between F_{GP} measured after dispersal and before dispersal. Since
332 the confidence limits of F_{GP} for adult females ([0.88; 1.00]) and immatures ([0.60; 1.00]) largely
333 overlap, we were unable to calculate a female-specific dispersal rate from equation (1).

334

335 **Discussion**

336 We have demonstrated that maternally-inherited genes can be used to describe the
337 contemporary dispersal patterns of males (and the overall dispersal patterns of females) within an
338 analytical framework based on intra-class genetic correlations. This was accomplished through
339 comparisons of age- and sex-specific intra-class correlations partitioned hierarchically within
340 populations. A second aim was to show that instantaneous sex-specific dispersal rates can be

341 calculated from sex-specific estimates of differentiation using single-locus haplotypic data.

342

343 *Dispersal in* Pecari tajacu

344 In our study we quantitatively demonstrated that dispersal in collared peccaries is

345 strongly biased toward males, and that approximately one-third of males dispersed from their

346 natal groups in this single generation. This is a minimum estimate, as some individuals die

347 before or during dispersal, and the lack of mtDNA variation undoubtedly prevented our detection

348 of dispersal between some groups. Moreover, the pronounced local genetic structure indicates

349 that males preferentially disperse over short distances, perhaps into neighboring herds; this is

350 congruent with trapping data (Gabor and Hellgren 2000). The results from the age-based

351 analysis indicate that dispersal in this species is usually accomplished by subadults (18-24

352 months). At this age, they have not reached their full body mass and may be forced out by larger,

353 resident males.

354

355 *Measuring dispersal biases*

356 Our approach allowed us to organize data into age classes, sex classes, social groups, and

357 populations and then test hypotheses about the dispersal rate of each class. For example, by

358 performing separate analyses on sex-specific datasets, we were able to both detect a sex-bias in

359 dispersal and also determine which sex contributed to the pattern. Because the method relies on

360 contrasts of sex-specific estimates of population differentiation, rather than absolutes, the power

361 to detect differences among hierarchies is limited only by the intensity of the bias (Vitalis 2002).

362 In this study there was sufficient contrast between pre- and post-dispersal age classes in males to

363 provide a direct estimate of the instantaneous dispersal rate.

364 The method presented here should be applicable to any species in which there is a bias in
365 dispersal, whether that bias is conditional on sex, age, or some other phenotype, so long as trait
366 variation can be readily distinguished and assigned to different hierarchical levels. This
367 approach does not impose spatial distance (or a distance proxy) onto the analysis, as is seen in
368 other approaches such as spatial autocorrelation (Smouse and Peakall 1999). Such approaches
369 force investigators to make assumptions about how distance interacts with social organization
370 when it may be inappropriate or irrelevant (e.g., when sampling a highly mobile species, or at a
371 scale where an individual is equally likely to disperse to any location under consideration). Our
372 approach removes metric distance and location from the equation, and shifts the focus onto how
373 the genetic variation is distributed across space irrespective of distance, which is especially
374 useful for addressing questions of how sociality influences dispersal.
375

376 *Measuring sex-biased dispersal with uni-parentally inherited markers*

377 The approach discussed herein relies upon contrasts: we compared the genetic structure
378 of the pre-dispersal class to the sex-specific genetic structure of the post-dispersal class to
379 estimate instantaneous dispersal within a single generation (Lawson-Handley and Perrin 2007).
380 When autosomal markers are used, the expectation is that genetic structure will be more apparent
381 in the pre-dispersal class compared to the post-dispersal class as a whole, and even more
382 apparent in the non-dispersing sex (whichever sex it may be). When a uni-parentally inherited
383 marker is used the expectation is similar, but not identical, to what is seen for bi-parentally
384 inherited markers. For instance, under a system of male-biased dispersal mtDNA haplotypes are
385 carried within males into breeding groups, but males do not contribute mtDNA to the subsequent
386 generation and thus the contrast between pre-dispersal individuals and adult males is substantial.

387 However, under a system of female-biased dispersal haplotypes would be re-distributed within
388 and among populations each generation. Thus a contrast between genetic differentiation for pre-
389 and post-dispersal individuals would be difficult to detect. As a result, this approach is most
390 useful for deriving instantaneous sex-specific dispersal rates with mtDNA data under a system of
391 male-biased dispersal, or a double uniparental system of mitochondrial DNA inheritance (e.g.
392 *Mytilus* mussels). Here we use mtDNA haplotypes as a tag, but any physical or genetic tag that
393 could be identified in males and females before and after dispersal may play the same role as
394 mtDNA markers in this context.

395 We have demonstrated that mtDNA can be used in isolation to estimate sex-specific dispersal
396 in the current generation,. The main caveat is that mtDNA is, in effect, a single genetic marker
397 that might be biased by selection (Bazin et al. 2005). Yet, because we based our analyzes upon
398 differences of variation in male and female within a single generation, it is difficult to imagine a
399 pattern of selection that would undermine the approach.

400 **Acknowledgements**

401 This research project was supported by a Grant in Aid of Research from the American
402 Society of Mammalogists to J.D. Cooper and by Purdue University funding to J.A. DeWoody.
403 J.D. Cooper is a Welder Wildlife Refuge Fellow, and many thanks are offered to the faculty and
404 staff at the Welder for their encouragement and assistance. We also thank Jim Cathey and Texas
405 Parks and Wildlife Department personnel for their help and cooperation during sample
406 collection. Special thanks goes to volunteers Edmund Harriss, Rachel Bonner and Raul Martinez
407 for their dedication and hard work in trapping collared peccaries. Finally, we thank the members
408 of the DeWoody lab and Peter Smouse for their insightful comments during the preparation of
409 this manuscript. Part of this work was carried out by using the resources of the Computational
410 Biology Service Unit from the Museum National d'Histoire Naturelle (MNHN – CNRS UMS
411 2700) which was partially funded by Saint Gobain.

412

413 **References**

414 Alves E, Ovilo C, Rodriguez MC, Silio L (2003) Mitochondrial DNA sequence variation
415 and phylogenetic relationships among Iberian pigs and other domestic and wild pig
416 populations. Animal Genetics 34, 319-324.
417 Bazin E, Glemin S, Galtier N (2006) Population size does not influence mitochondrial genetic
418 diversity in animals. Science 312, 570-572.
419 Bissonette JA (1982) Social behavior and ecology of the collared peccary in Big Bend National
420 Park. Pp. 85 in National Park Service Science Monograph Number 16.
421 Chesser RK (1991) Gene diversity and female philopatry. Genetics 127, 437-447.
422 Chesser RK, Baker RJ (1996) Effective sizes and dynamics of uniparentally and
423 diparentally inherited genes. Genetics 144, 1225-1235.
424 Day GI (1985) Javelina research and management in Arizona. Arizona Game and Fish
425 Department, Phoenix.
426 Ellisor JE, Harwell WF (1969) Mobility and home range of collared peccary in southern
427 Texas. Journal of Wildlife Management 33, 425-427.
428 Escorza-Trevino S, Dizon AE (2000) Phylogeography, intraspecific structure and sex-
429 biased dispersal of Dall's porpoise, *Phocoenoides dalli*, revealed by mitochondrial and
430 microsatellite DNA analyses. Molecular Ecology 9, 1049-1060.
431 Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software
432 package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47-
433 50.
434 Fontanillas P, Petit E, Perrin N (2004) Estimating sex-specific dispersal rates with
435 autosomal markers in hierarchically structured populations. Evolution 58, 886-894.

436 Gabor TM, Hellgren EC (2000) Variation in peccary populations: landscape composition
437 or competition by an invader? Ecology 81, 2509-2525.
438 Girman DJ, Mills MGL, Geffen E, Wayne R (1997) A molecular genetic analysis of
439 social structure, dispersal and interpack relationships of the African wild dog (*Lycaon*
440 *pictus*). Behavioural Ecology and Sociobiology 40, 187-198.
441 Gongora J, Morales S, Bernal JE, Moran C (2006) Phylogenetic divisions among Collared
442 peccaries (*Pecari tajacu*) detected using mitochondrial and nuclear sequences. Molecular
443 Phylogenetics and Evolution 41, 1-11.
444 Goudet J, Perrin N, Waser PM (2002) Tests for sex-biased dispersal using bi-parentally
445 inherited genetic markers. Molecular Ecology 11, 1103-1114.
446 Hellgren EC, Lochmiller RL, Grant WE (1984) Demographic, morphologic, and
447 reproductive status of a herd of collared peccaries (*Tayassu tajacu*) in south Texas.
448 American Midland Naturalist 112, 402-407.
449 Hoelzer GA, Dittus WPJ, Ashley MV, Melnicks DJ (1994) The local distribution of
450 highly divergent mitochondrial DNA haplotypes in toque macaques *Macaca sinica* at
451 Polonnaruwa, Sri Lanka. Molecular Ecology 3, 451-458.
452 Langergraber KE, Siedel H, Mitani JC, Wrangham RW, Reynolds V, et al (2007) The genetic
453 signature of sex-biased migration in patrilocal chimpanzees and humans. PLoS ONE
454 2(10): e973. doi:10.1371/journal.pone.0000973.
455 Lawson-Handley LJ, Perrin N (2007) Advances in our understanding of mammalian sex-
456 biased dispersal. Molecular Ecology 16, 1559-1578.
457 Mossman CA, Waser PM (1999) Genetic detection of sex-biased dispersal. Molecular
458 Ecology 8, 1063-1067.

459 Petit E, Balloux F, Goudet J (2001) Sex-biased dispersal in a migratory bat: a
460 characterization using sex-specific demographic parameters. Evolution 55, 635-640.
461 Prugnolle F, de Meeus T (2002) Inferring sex-biased dispersal from population genetic
462 tools: a review. Heredity 88, 161-165.
463 R Development Core Team (2008) R: A language and environment for statistical computing. R
464 Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL
465 http://www.R-project.org.
466 Rousset F (2007) Inferences from spatial population genetics. In Handbook of Statistical
467 Genetics, third edition (eds. Balding D, Bishop M, Cannings C) Pp. 945–979. John Wiley
468 & Sons, Chichester.
469 Sambrook J, Russell DW (2001) Molecular Cloning: A Laboratory Manual, 3rd edition. Cold
470 Spring Harbor Laboratory Press, New York.
471 Scott, D.W. (1992) Multivariate Density Estimation: Theory, Practice, and Visualization. John
472 Wiley, New York
473 Sowls LK (1978) Collared peccary. In: Big game of North America. (eds. Schmidt JL, Gilbert
474 DL) Pp. 191-205. Stackpole Books, Harrisburg, Pennsylvania.
475 Slatkin M, Voelm L (1991) F_{ST} in a hierarchical island model. Genetics 127, 627-629.
476 Sugg DW, Chesser RK (1994) Effective population sizes with multiple paternity.
477 Genetics 137, 1147-1155.
478 Swofford DL (2003) PAUP*: Phylogenetic analysis using parsimony (* and other methods),
479 version 4.0b 10. Sinauer Associates, Sunderland, Massachusetts.
480 Theimer TC, Keim P (1994) Geographic patterns of mitochondrial-DNA variation in
481 collared peccaries. Journal of Mammalogy 75, 121-128.

482 Triant DA, DeWoody JA (2007) The occurrence, detection, and avoidance of mitochondrial
483 DNA translocations in mammalian systematic and phylogeography. Journal of
484 Mammalogy 88:908-920.
485 Vigouroux V, Couvet D (2000) The hierarchical island model revisited. Genetics Selection
486 Evolution. 32, 395-402.
487 Vitalis R (2002) Sex-specific genetic differentiation and coalescence times: estimating sex-
488 biased dispersal rates. Molecular Ecology 11, 125-138.
489 Waser PM, Strobeck C, Paetkau D (2001) Estimating interpopulation dispersal rates. In:
490 Carnivore Conservation. (eds. Gittleman JL, Funk SM, MacDonald DW, Wayne RK) Pp.
491 484-497. Cambridge University Press, Cambridge.
492 Weir B, Cockerham C (1984) Estimating F-Statistics for the analysis of population structure.
493 Evolution 38, 1358-1370.
494 Wright S (1965) The interpretation of population structure by *F*-statistics with special regard to
495 systems of mating. Evolution 19, 395-420.
496 Yang G, Lui S, Ren WH, Zhou KY, Wei FW (2003) Mitochondrial control region
497 variability of baiji and the Yangtze finless porpoises, two sympatric small cetaceans in
498 the Yangtze river. Acta Theriologica 48, 469-483

499 **Table 1** Distribution of mtDNA haplotypes in three wild populations of *P. tajacu* in Texas, across sex and age classes.

				Sex		Age class		
Population	Haplotype	*n*	Freq.	F	M	I	J	A
WWR	C	24	0.77	10	14	1	6	17
	A	7	0.23	3	4	0	3	4
CWMA	B	43	0.43	24	19	1	10	32
	A	38	0.38	21	17	1	9	28
	C	19	0.19	9	10	3	2	14
	G	1	0.01	0	1	0	0	1
	F	1	0.01	0	1	0	0	1
BB	E	129	0.96	61	68	13	35	81
	A	5	0.04	3	2	0	0	5

500 *n*: sample size; F: females, M: males; I: infants; J: juveniles; A: adults (including subadults)

501 **Table 2.** Intra-class correlations for pairs of genes among social groups within replicate populations estimated by means of identity
502 probabilities. F_{GP} estimates used to calculate sex-specific migration rates from equation (1) are given in bold. Confidence intervals are
503 noted in brackets, and represent the 2.5% and the 97.5% quantiles produced by bootstrapping nucleotide sites over 25,000 samples. *P*-
504 values give the results of significance tests for differentiation among groups. They were calculated as the proportion of times where
505 F_{GP} from randomized datasets was larger than or equal to the observed F_{GP} on the original dataset. Randomized datasets were
506 obtained by by permuting haplotypes at random among groups within populations (25,000 permutations).

Category	F_{GP} estimate	*p*-value
All data	0.50 [0.46, 0.59]	$p < 0.001$
Adults	0.30 [0.03, 0.38]	$p < 0.001$
Immatures	0.60 [0.60, 1.00]	$p < 0.001$
Females	0.90 [0.87, 1.00]	$p < 0.001$
Males	0.23 [0.09, 0.36]	$p < 0.001$
Adult females	0.91 [0.88, 1.00]	$p < 0.001$
Adult males (dispersers)	0.24 [0.07, 0.36]	$p = 0.013$
Philopatrics	0.76 [0.73, 0.99]	$p < 0.001$

507 **Figure Legends**

508

509

510 Figure 1. Re-sampled data null distributions for each class of individuals. Observed F_{GP} for
511 each analysis represented by hatched vertical line. Significance tested over 25,000 permutations.
512 Histogram class heights are represented as black dots, and the smoothed density was obtained
513 using the Average Shifted Histogram (ASH) algorithm (Scott 1992) with smoothing parameter m
514 = 20.

515 **Appendix**

516 *Sex-specific differentiation before and after (instantaneous) dispersal*

517 *F*-statistics are defined as intra-class correlations for the probability of identity in state

518 (IIS correlations) (Cockerham and Weir 1987; Rousset 1996). Yet, the infinite allele model

519 (IAM) provides the value of the probability of identity by descent (IBD probabilities) and in the

520 low mutation rate limit, for two given classes, IIS and IBD correlations converge to the same

521 value (Slatkin 1991; Rousset 1996). Thus, the properties of *F*-statistics can be deduced from the

522 properties of intra-class correlations for IBD probabilities (Rousset 1996).

523 In the following, we consider mitochondrial DNA (mtDNA) markers, i.e. haploid

524 markers, transmitted by females only. Let Q_1^{XY}, Q_2^{XY} and Q_3^{XY} be the IBD probability of two

525 mtDNA gene copies sampled from two individuals of sex *X* and *Y* after dispersal among

526 individuals within the same breeding group, among breeding groups within the same

527 populations, and among distinct populations, respectively. These individuals may be two males,

528 two females, or one male and one female. Let Q_1^*, Q_2^* and Q_3^* be the corresponding IBD

529 probabilities for gene copies sampled before dispersal. For pairs of genes sampled before

530 dispersal there is no need to consider distinct coefficients for different pairs of individuals of the

531 same or opposite sex (Vitalis 2002).

532 Let us consider an infinite island model of population structure (Wright 1951) where each

533 population is isolated and further subdivided into *n* breeding groups, and where dispersal among

534 breeding groups is achieved by juveniles, before reproduction. Here, because we consider that

535 populations are independent, we restrict our analyses to estimate within-population dispersal

536 only. We therefore focus on the IBD probabilities among individuals within social groups (Q_1^{XY}

537 and Q_1^*), and among social groups within populations (Q_2^{XY} and Q_2^*), and further consider the IBD

538 probabilities between pairs of genes from distinct populations (Q_3^{XY} and Q_3^*) to be nil in the

539 model. A migrant individual in a breeding group is equally likely to come from any of the

540 $(n-1)$ other breeding groups. Generations do not overlap. Let m_X denote the probability that

541 an individual of sex *X* has immigrated. Each generation, after migration, the frequency of pairs

542 of individuals taken at random in one breeding group that come from a single group before

543 migration is $a_{XY} = (1-m_X)(1-m_y) + m_X \cdot m_Y/(n-1)$, for pairs of individuals of sex *X* and *Y* with

544 $X \in \{♂,♀\}$ and $Y \in \{♂,♀\}$. Conversely, the frequency of pairs of individuals taken at random

545 from two breeding group after migration that originate from the same group before migration is

546 $b_{XY} = (1-a_{XY})/(n-1)$ (see, e.g., Rousset 1996). We assume that the mutation rate is low (i.e., μ

547 $<< m_X$), so that virtually no mutation arises over a single generation.

548 The genes sampled among individuals after dispersal in a breeding group at any

549 generation come from the same breeding group before dispersal with probability a_{XY} and from

550 different breeding group with probability $(1 - a_{XY})$. These mtDNA gene copies are then IBD

551 with probabilities Q_1^* and Q_2^*, respectively. Thus,

552 $$Q_1^{XY} = a_{XY}Q_1^* + (1-a_{XY})Q_2^*. \quad \text{(A.1)}$$

553 Similarly, for genes sampled before dispersal,

554 $$Q_2^{XY} = b_{XY}Q_1^* + (1-b_{XY})Q_2^*. \quad \text{(A.2)}$$

555 We can rewrite equations 1 and 2 as:

556 $$Q_1^{XY} = a_{XY}\left(Q_1^* - Q_2^*\right) + Q_2^*, \quad \text{(A.3)}$$

557 and:

558 $$Q_2^{XY} = b_{XY}\left(Q_1^* - Q_2^*\right) + Q_2^*. \quad \text{(A.4)}$$

559 Subtracting equations A.3 and A.4, we obtain:

560 $$\left(Q_1^{XY} - Q_2^{XY}\right) = d_{XY}\left(Q_1^* - Q_2^*\right), \quad \text{(A.5)}$$

561 where $d_{XY} = \left(a_{XY} - b_{XY}\right)$. Rearranging equation A.4 gives:

562 $$\left(1 - Q_2^{XY}\right) = \left(1 - Q_2^*\right) - b_{XY}\left(Q_1^* - Q_2^*\right). \quad \text{(A.6)}$$

563 Taking the ratio of equation A.6 over A.5 gives:

564 $$\frac{1 - Q_2^{XY}}{Q_1^{XY} - Q_2^{XY}} = \frac{1}{d_{XY}} \frac{1 - Q_2^*}{Q_1^* - Q_2^*} - \frac{b_{XY}}{d_{XY}}. \quad \text{(A.7)}$$

565 Since we consider independent populations, the relevant parameter to infer dispersal among

566 breeding groups is F_{GP}. From the definition of sex-specific F-statistics as appropriate ratios of

567 sex-specific IBD probabilities (Vitalis 2002), we get:

568 $$F_{GP}^* = \frac{Q_1^* - Q_2^*}{1 - Q_2^*} \quad \text{(A.8)}$$

569 before dispersal, and:

570 $$F_{GP}^{XY} = \frac{Q_1^{XY} - Q_2^{XY}}{1 - Q_2^{XY}} \quad \text{(A.9)}$$

571 after dispersal. We get from equation A.7:

572 $$\frac{1}{F_{GP}^{XY}} = \frac{1}{d_{XY}} \frac{1}{F_{GP}^*} - \frac{b_{XY}}{d_{XY}}, \quad \text{(A.10)}$$

573 i.e., multiplying both sides by $\left(d_{XY} \cdot F_{GP}^{XY}\right)$ and rearranging:

574 $$\frac{F_{GP}^{XY}}{F_{GP}^*} = d_{XY} + b_{XY} F_{GP}^{XY} \quad \text{for all } (X, Y) \in \{♂, ♀\} \quad \text{(A.11)}$$

575 which is equation 15 in Vitalis (2002) except that here it holds for mtDNA markers, and it does

576 not assume equilibrium. This relation is valid for any generation, for samples taken before and

577 after dispersal. The last term in the right-hand side of equation A.11 is negligible compared to

578 d_{XY}. Thus, F_{GP}^{XX} differs from F_{GP}^* by a factor $d_{XX} = \left(1 - m_X\left[n/(n-1)\right]\right)^2 \approx \left(1 - m_X\right)^2$, for large n.

579 Therefore, taking the ratio of sex-specific F_{GP}^{XX} evaluated after juvenile dispersal, over F_{GP}^{*}

580 evaluated before dispersal gives the sex-specific migration rate:

581 $$m_X \approx 1-\sqrt{\frac{F_{GP}^{XX}}{F_{GP}^{*}}} \qquad \text{for all } X \in \{♂,♀\} \qquad (A.12)$$

582 Following Rousset (2007), we estimated the intra-class correlations by taking appropriate

583 ratios of identity probabilities, weighted according to the number of pairs in each sample. From

584 the definitions of F-statistics as functions of identity probabilities between pairs of genes

585 (equations A.8 and A.9), we get:

586

587 $$\hat{F}_{GP}^{*} = \frac{\hat{Q}_1^{*}-\hat{Q}_2^{*}}{1-\hat{Q}_2^{*}}, \text{ and} \qquad (A.13)$$

588 $$\hat{F}_{GP}^{XY} = \frac{\hat{Q}_1^{XY}-\hat{Q}_2^{XY}}{1-\hat{Q}_2^{XY}}, \qquad (A.14)$$

589 where $\hat{Q}_i^{*}$ and $\hat{Q}_i^{XY}$ denote the estimates of identity probabilities between pairs of genes,

590 calculated by simple counting of identical pairs of genes at the ith hierarchical level. An estimate

591 of the sex-specific migration rate therefore reads:

592 $$\hat{m}_X \approx 1-\sqrt{\frac{\hat{F}_{GP}^{XX}}{\hat{F}_{GP}^{*}}} \qquad \text{for all } X \in \{♂,♀\} \qquad (A.15)$$