Weak population structure in the ant *Formica fusca*

Helena Johansson [Corresp., 1] , **Perttu Seppä** [1, 2] , **Heikki Helanterä** [1, 2] , **Kalevi Trontti** [1] , **Liselotte Sundström** [1, 2]

[1] Centre of Excellence in Biological Interactions, Organismal and Evolutionary Biology Research Programme, University of Helsinki, Helsinki, Finland

[2] Tvärminne Research Station, University of Helsinki, Hangö, Finland

Corresponding Author: Helena Johansson
Email address: helena.z.johansson@helsinki.fi

Dispersal is a fundamental trait of a species' biology. High dispersal results in weakly structured or even panmictic populations over large areas, whereas weak dispersal enables population differentiation and strong spatial structuring. We report on the genetic population structure in the polygyne ant *Formica fusca* and the relative contribution of the dispersing males and females to this. We sampled twelve localities across a ~35 km^2 study area in Finland and generated mitochondrial (mtDNA) haplotype data and microsatellite data. First, we assessed queen dispersal by estimating population differentiation from mitochondrial (mtDNA) haplotype data. Second, we analyzed nuclear DNA microsatellite data to determine overall population genetic substructure in the study area with principal components analysis, Bayesian clustering, hierarchical *F* statistics and testing for evidence of isolation-by-distance. Third, we directly compared genetic differentiation estimates from maternally inherited mtDNA and bi-parentally inherited DNA microsatellites to test for sex-bias in dispersal. Our results showed no significant spatial structure or isolation by distance in neither mtDNA nor DNA microsatellite data, suggesting high dispersal of both sexes across the study area. However, mitochondrial differentiation was weaker (F_{st-mt} = 0.0047) than nuclear differentiation (F_{st-nuc} = 0.027), which translates into a six-fold larger female migration rate compared to that of males. We conclude that the weak population substructure reflects high dispersal in both sexes, and it is consistent with *F. fusca* as a pioneer species exploiting unstable habitats in successional boreal forests.

1 **Weak population structure in the ant *Formica fusca***

3 **Helena Johansson[1], Perttu Seppä[1, 2], Heikki Helanterä[1, 2], Kalevi Trontti[1] and Liselotte**

4 **Sundström[1, 2]**

6 **Affiliations:**

7 1. Centre of Excellence in Biological Interactions, Organismal and Evolutionary Biology

8 Research Programme, Faculty of Biological and Environmental Sciences, FIN-00014 University

9 of Helsinki, Finland

10 2. Tvärminne Zoological Station, University of Helsinki, Helsinki, Finland

12 **Corresponding author:**

13 Helena Johansson

14 E-mail address: helena.z.johansson@helsinki.fi

27

28

29

30

31

32

33

34

35 **Abstract**

36 Dispersal is a fundamental trait of a species' biology. High dispersal results in weakly structured

37 or even panmictic populations over large areas, whereas weak dispersal enables population

38 differentiation and strong spatial structuring. We report on the genetic population structure in the

39 polygyne ant *Formica fusca* and the relative contribution of the dispersing males and females to

40 this. We sampled twelve localities across a ~35 km^2 study area in Finland, and generated

41 mitochondrial (mtDNA) haplotype data and microsatellite data. First, we assessed queen

42 dispersal by estimating population differentiation from mitochondrial (mtDNA) haplotype data.

43 Second, we analyzed nuclear DNA microsatellite data to determine overall population genetic

44 substructure in the study area with principal components analysis, Bayesian clustering,

45 hierarchical *F* statistics and testing for evidence of isolation-by-distance. Third, we directly

46 compared genetic differentiation estimates from maternally inherited mtDNA and bi-parentally

47 inherited DNA microsatellites to test for sex-bias in dispersal. Our results showed no significant

48 spatial structure or isolation by distance in neither mtDNA nor DNA microsatellite data,

49 suggesting high dispersal of both sexes across the study area. However, mitochondrial

50 differentiation was weaker ($F_{ST-mt} = 0.0047$) than nuclear differentiation ($F_{ST-nuc} = 0.027$), which

51 translates into a six-fold larger female migration rate compared to that of males. We conclude

52 that the weak population substructure reflects high dispersal in both sexes, and it is consistent

53 with *F. fusca* as a pioneer species exploiting unstable habitats in successional boreal forests.

54

55

56 **Introduction**

57 Dispersal serves to move genes between natural populations, maintaining gene flow and

58 connectivity, thereby reducing the risks of local population extinction and genetic drift (Clobert

59 et al., 2001). High gene flow may lead to weak population substructure or even panmixia over

60 large areas (Peel et al., 2013), whereas reduced gene flow often leads to significant substructure

61 and divergence between populations, through genetic drift or through strong selection acting on

62 the diverging populations (Slatkin, 1987). Genetic population substructures may also arise due to

63 differential dispersal in e.g. sexes or morphs (Seppä et al., 2006) and reveal social relationships,

64 especially in highly social and eusocial species (Ross, 2001). The scale at which population

65 genetic patterns arise can yield information about the processes that leads to divergence, as well

66 as a species potential migratory or dispersal ability (Kovach et al., 2013).

67

68 Eusocial insects (ants, termites, some wasps and bees) form societies in which only specialised

69 sexual castes (queens and males) disperse and reproduce, whereas workers forgo their own

70 reproduction to help raise the offspring of the resident queen(s). In ants, queens and males are

71 commonly winged (Peeters & Ito, 2001) and flight is the primary medium for mating and gene

72 flow (Helms, 2018). However, ants are slow fliers (1.3-5 m/s) and there is great variation

73 between species in distance covered (30-32200 m), duration (25-140 min) and altitude of flight

74 (1-300 m) (Helms, 2018) These three factors are influenced by physiology of the flying

75 individuals, such as size, wing size, fat storage capacity, and behaviour. Ant species with small

76 queens that have relatively large wings and fat content (>40% of their body weight) tend to fly at

77 higher altitudes and cover the largest distances (Helms, 2018, and references therein).

78 Furthermore, dispersal capability and distance are associated with the nest-founding strategy of

79 queens. Queens able to found nests independently, through temporary parasitism, or by joining

80 unrelated queens in foundress associations, can usually disperse relatively far (Keller & Passera,

81 1989). Conversely, queens that rely on the aid of their own workers to found their nests

82 (dependent nest founding) or rely on re-adoption into established colonies are restricted in their

83 dispersal (Keller & Passera, 1989; Cronin et al., 2013). Dependent founding is commonly found

84 in species with several related queens in a single colony (polygyne species), and is thought to be

85 an adaptation to reduce dispersal risks (Keller & Passera, 1989; Cronin et al., 2013).

86   Queen founding strategy has consequences for hierarchical genetic substructure, i.e. partitioning
87   of genetic variation from nests to populations or species. With independently founding queens,
88   the resulting population substructure is expected to show strong structure at the nest level, yet
89   sub-structuring at higher hierarchical levels such as habitat patch or population may be absent if
90   queen dispersal is strong With dependently founding queens short-distance dispersal is expected
91   to lead to genetic viscosity (isolation-by-distance over small spatial scales) and discontinuous
92   population substructure (Ross & Shoemaker, 1997; Ross, 2001). Dependently founding species
93   may also form polydomous populations, i.e. comprise networks of related and interconnected
94   nests that sometimes occupy entire habitat patches (Pamilo et al., 2005). Consequently, strong
95   substructure at the level of the patch is expected, and substructure at the level of a population is
96   often observed (Pamilo et al., 2005; Holzer, Keller & Chapuisat, 2009). The strength of these
97   patterns is modified by the rate of male dispersal (see e.g. Sanllorente & al. 2015).
98
99   The population level consequences of the queens' strategy for nest founding can be elucidated
100   from patterns of population differentiation based on maternally inherited mitochondrial DNA.
101   Furthermore, by contrasting mtDNA differentiation with that estimated from nuclear markers
102   (e.g. DNA microsatellites) relative levels of female and male dispersal can be estimated. Such
103   research on ants has revealed that dispersal is usually but not overtly male biased, with
104   greatest bias in species with wingless queens, regardless of gyny level (Table S1; Sundström,
105   Seppä & Pamilo, 2005). To our knowledge, female-biased dispersal has not been inferred
106   with genetic data in any study on ants.
107
108   While different population genetic consequences between independently and dependently
109   founding species has support from empirical data (Ross & Shoemaker, 1997; Ross, 2001),
110   species with both modes of founding are more challenging. Unless queens lack wings entirely,
111   winged (alate) queens remain potentially important for dispersal even if dependent founding is
112   common, i.e. a few dispersing queens may be sufficient to allow efficient gene flow to
113   homogenize populations (Peeters & Ito, 2001). Both modes of founding may co-exist in the same
114   population, or one may become dominant in different populations within the same species, such
115   as in *Formica exsecta* (Sundström, Seppä & Pamilo, 2005). Population substructures in such
116   species are also less predictable; they may change with the number of reproductive queens in the

117   colony (i.e. one (monogyne) or several (polygyne) queens per nest)), as has been shown in *F.*

118   *truncorum* (Sundström, 1995)*,* or, like in the case of *F. polyctena*, spatial scale (Gyllenstrand,

119   Seppä & Pamilo, 2004). Studying such species may hence offer an insight into ant dispersal

120   evolution, and the social and ecological conditions under which dispersal risk adaptations are

121   likely to evolve.

122

123   European members of the genus *Formica* have long been an important model for population

124   genetics, dispersal and ecology in ants. The genus is socially diverse and includes dependently

125   founding species with obligate polygyny, polydomy and strong population genetic structuring,

126   such as *F. aquilonia* (Pamilo et al., 2005) and *F. paralugubris* (Holzer, Keller & Chapuisat,

127   2009), but also independently founding species (with or without queen re-adoption), such as *F.*

128   *rufa* (Gyllenstrand, Seppä & Pamilo, 2004), socially polymorphic species with both types, such

129   as *F. exsecta* (Pamilo & Rosengren, 1984) and *F. truncorum* (Sundström, 1993), and species

130   with continuous variation in queen numbers, such as *F. fusca* (Bargum, Helanterä & Sundström,

131   2007). The genus is also ecologically diverse, inhabiting a wide variety of habitats from sand

132   dunes (*F. cinerea*) to mature conifer forest (*F. rufa* group) (Czechowski, Radchenko &

133   Czechowska, 2002). Habitat stability has been proposed to promote evolution of dependent

134   founding and polydomy in *Formica* species (Rosengren & Pamilo, 1983), yet no study has

135   directly addressed the population genetic consequences of inhabiting temporally unstable habitat

136   in this group of ants. To date, male-biased dispersal has been confirmed in *F. paralugubris*

137   (Holzer, Keller & Chapuisat, 2009), *F. exsecta* (Seppä et al., 2004) and *F. lugubris* (Gyllenstrand

138   & Seppä, 2003) and thought to be a consequence of polygyny.

139

140   Here, we investigate patterns of population differentiation in the widespread generalist ant

141   *Formica fusca* (Czechowski, Radchenko & Czechowska, 2002) which is one of the first species

142   to arrive in clear-cut areas in boreal forests (Punttila et al., 1991). Both sexes are winged and

143   mature to flight in July, but patterns of nuptial flight and mating behaviour are unknown. It

144   founds new colonies independently (Pamilo, 1983), suggesting that dispersal distances can be

145   substantial. The majority of nests (55-68%) are polygyne (Hannonen, Helanterä & Sundström,

146   2004; Bargum & Sundström, 2007) and 26% of queens (regardless of gyny level) are multiply

147   mated, resulting in worker relatedness levels around 0.65 in monogyne nests (Hannonen,

148 Helanterä & Sundström, 2004; Bargum & Sundström, 2007). Nest-mate worker relatedness

149 values in polygyne nests range from 0-0.8 suggest that in some nests daughter queens are re-

150 adopted, and in other nests the queens are unrelated (Hannonen & Sundström, 2003; Hannonen,

151 Helanterä & Sundström, 2004). Previous small-scale studies on island populations (≤5 km^2)

152 using DNA microsatellites and allozymes suggest weak population substructure, no dispersal

153 limitation, monodomy (nests do not form interrelated networks), and no genetic differentiation

154 between social forms (Pamilo, 1983; Bargum, Helanterä & Sundström, 2007).

155

156 We investigated genetic population substructure in a continuous area with principal component

157 and Bayesian assignment methods, and complemented these with tests of hierarchical population

158 substructure, and an isolation-by-distance analysis. We were asking whether population

159 substructure remains weak also over larger spatial scales in a homogenous landscape, or if we

160 can observe a change towards greater genetic structure with increasing distance due restricted

161 dispersal. To test for sex-biased dispersal, and to assess the sex-specific contribution to

162 population genetic structure, we exploited the different properties of microsatellites (bi-parental,

163 larger effective population size) and mitochondrial DNA (maternal inheritance, smaller effective

164 population size) by estimating and comparing population differentiation in these two markers.

165

166 **Materials and methods**

167 ***Sampling and molecular methods***

168 Twelve localities were sampled in April-June 2005 in an area spanning ca. 10x35 km in Southern

169 Finland (Figure 1, Table 1). Based on vegetation characteristics the localities varied in age,

170 however no formal assessment of age was made. The locality altitudes, the monthly mean and

171 maximum temperatures and precipitation values, as well as the bioclimate values

172 (http://www.worldclim.org/bioclim, Table S3) suggest homogenous bioclimatic conditions

173 across this area. There are no major geographical barriers to dispersal in this region (Figure 1),

174 nor are there post-glacial contact zones potentially contributing to the estimated population

175 differentiation and dispersal estimates (Schmitt, 2007)**.** At each sampling locality, five workers

176 and all queens found were collected from between five and fifteen nests. For all localities, the

177 exact location for the first nest that was sampled was recorded with a Garmin GPS 12XL, and for

178 the subsequently sampled nests coordinates were calculated from the metric distance and angle**.**

179

180   All workers were genotyped at nine DNA microsatellite loci: FL12, FL20 (Chapuisat, 1996),

181   FE13, FE17, FE19, FE21(Gyllenstrand, Gertsch & Pamilo, 2002), FY4, FY7, and FY13

182   (Hasegawa & Imai, 2004), following protocols within. Genotyping was performed on a

183   MegaBACE 1000 with the internal size standard ET400-R (GE Healthcare) and scored using

184   Fragment Profiler v1.2 (GE Healthcare).

185

186   For one queen (preferentially) or one worker from each nest, a ca. 600 bp segment of the

187   mitochondrial COI gene was amplified with primers: coi_bar_F:5'-

188   ACTAGGATCTCCAGACATAGC-3' and coi_bar_R: 5'-GCTCGTGTATCAACATCTAA-3',

189   following protocols in Seppä et al., 2011. Sequencing of the PCR-products was carried out by the

190   Institute of Biotechnology DNA Sequencing and Genomics Laboratory (University of Helsinki,

191   Finland).

192

193   ***Analyses on the complete data set***

194   For the DNA microsatellite data we tested for population deviations from Hardy-Weinberg

195   Equilibrium (HWE) and Linkage Disequilibrium (LD) using ARLEQUIN (Excoffier & Lischer,

196   2010). Allele frequencies averaged over all populations, local and global inbreeding coefficients

197   (F_{IS}), gene diversity (H_S) and allelic richness (A_R) were calculated for each locality using FSTAT

198   2.9.3.2.(Goudet, 1995).

199

200   To test for occurrence and extent of spatial genetic sub-structuring we used an individual level

201   Principal Component Analysis (PCA) using SMARTPCA as implemented in EIGENSOFT 4.2.

202   (Patterson, Price & Reich, 2006). Principal components analysis summarises the variance in

203   microsatellite data into a smaller number of orthogonal components, and the analysis is free of

204   assumptions on the clustering of individuals, nests, localities or populations. Major genetic

205   structuring is captured in the first few principal components, and an implemented Tracy-Wisdom

206   statistic indicates if the components summarizes significant substructure.

207

208 Genetic differentiation among subpopulations was further examined by calculating global and

209 pairwise F_{ST}'s (Weir & Cockerham, 1984) for each locality and each colony using FSTAT. The R

210 package HIERFSTAT v. 0.04-10 (Goudet, 2005) was used to estimate hierarchical population

211 substructure at the following levels: individual, nest, locality and total population. HIERFSTAT

212 achieves this through the computation of variance components (hierarchical F-statistics) in a

213 fully random hierarchical design. Isolation-by-distance patterns (IBD) were tested with Mantel

214 tests (9999 permutations), with the package CULTEVO for R (Stadler, 2017). Isolation-by-

215 distance was tested both among localities within the total population and among nests within the

216 total population.

217

218 ***Analyses on a single individual per nest***

219 We ran a Bayesian clustering analysis on the data by using the software STRUCTURE (Pritchard,

220 Stephens & Donnelly, 2000), which clusters individual genotypes by probability of similarity,

221 regardless of the origin of the individuals. In case that nest kin structure violates the underlying

222 model assumptions, we used only a single individual per nest. The most likely number of clusters

223 (K) was estimated by running the analysis with K ranging from 1 to 12, using the admixture

224 model with **locprior** and correlated allele frequencies. For each K value, the analysis was run ten

225 times with a burn-in of 100 000 steps for a total run length of 300 000 steps. STRUCTURE

226 HARVESTER (Earl & Bridgett, 2012) was used to plot the mean and standard deviation of L(K)

227 for each run, and to apply the deltaK methods (Evanno, Regnaut & Goudet, 2005) to determine

228 the most likely number of clusters in the data set.

229 *Mitochondrial population structure*

230 Mitochondrial sequences were trimmed, edited and aligned using CODONCODE ALIGNER

231 (CodonCode Corporation, Dedham MA. USA). To visualize haplotype divergence, a Neighbour-

232 joining tree (rooted at the centre) using Jukes and Cantor distance and 1000 bootstraps was

233 constructed in SEAVIEW 4.2.12. (Gouy, Guindon & Gascuel, 2010). Genetic differentiation

234 among sub-populations was estimated by calculating overall F_{ST} (F_{ST-mt}) using haplotype

235 frequency data in ARLEQUIN v. 3.5.1.2 (Excoffier & Lischer, 2010). A test for isolation-by-

236 distance among localities was performed on pairwise F_{ST}'s calculated in ARLEQUIN as above.

237

Sex-biased dispersal

239 Mitochondrial haplotype and microsatellite profiles from one and the same individual per nest
240 were used to test for sex differences in dispersal. Because of haplo-diploidy and smaller effective
241 population size in the mitochondrial compared to nuclear genome, male and female gene flow
242 match when:

243

$$\frac{1}{4 \times \Phi_{ST}} = \frac{2}{3 \times F_{ST}}$$

244   (Equation 1)

245

246 assuming Wright's infinite island model (Berg, Lascoux & Pamilo, 1998). The infinite island
247 model makes several strict assumptions about dispersal, demographic structure and genetic
248 system (Wright, 1931), and its use has been strongly criticized (Whitlock & Mccauley, 1999).
249 However, for ants the use of these estimates appears reliable. In species where queens are
250 wingless (and disperse only by foot) and males have wings, the expected difference in dispersal
251 capability is accurately reflected in the results of the analysis (Seppä & Pamilo, 1995;
252 Fernández-Escudero, Seppä & Pamilo, 2001; Fernández-Escudero, Pamilo & Seppä, 2002;
253 Seppä et al., 2006; Berghoff et al., 2008).

254

255 **Results**

256 *Analyses using the complete data set*

257 Only individuals with microsatellite genotypes from a minimum of four loci were retained in the
258 analysis, in total ca. 95% of genotyped individuals. Locus FY7 was also excluded from analysis
259 due to inconsistent genotyping and the likely presence of null alleles. The remaining loci had on
260 average 6.25 alleles per locus (SD 2.82; range 2-11). Most loci were in HWE in most localities
261 after Bonferroni correction, and only a few significant deviations appeared randomly: FY13 in
262 BAC, FY7 in FOR and FL12 in MOS. Significant LD found appeared also random; locus FE19
263 was found to deviate from LD in six locations, but always with different loci, and some other

264   locus combinations were sporadically significant in different locations. We hence kept eight loci

265   for further analyses.

266

267   The global inbreeding value (F_{IS}) was estimated at -0.003, suggesting random mating at the level

268   of the total population. F_{IS} estimated by locality ranged from -0.172-0.196, with small deviations

269   from zero in eight of the 12 populations (Table 1).

270

271   PCA indicated no significant substructure (PC1: eigenvalue 26.78, Tracey-Wisdom 0.18, PC2:

272   eigenvalue 25.83, Tracey-Wisdom 0.07). A plot of the two first principal components clearly

273   shows the absence of a spatial pattern (Figure 2).

274

275   The overall genetic differentiation among localities was significantly greater than zero in nuclear

276   markers (F_{ST} = 0.048; 95% C.I. 0.030-0.074). When nest structure was taken into account in a

277   hierarchical F analysis, differentiation among localities within the total population was weaker

278   but still significant (F_{CT} = 0.015). A negative inbreeding value for individuals within nests (F_{IS} =

279   -0.38) was most likely a consequence of the non-independence of individuals sampled from the

280   same nest, i.e. nestmates are related. Finally, the large and similar inbreeding values for nests in

281   the total populations (F_{IT} = 0.31) and for nests in localities (F_{IC} = 0.30) are consistent with a

282   strong Wahlund effect, i.e. that nestmate workers are related and that nests act as if they were

283   randomly breeding subunits within the localities. These results hence suggest that nests are

284   highly differentiated within localities and in the total population, and that there is only very weak

285   substructure corresponding to geographic location, results which are in line with those obtained

286   from the principal components and Bayesian clustering analyses.

287   We found no isolation-by-distance pattern among localities (r = 0.13, P = 0.214), nor among

288   colonies (r = 0.0183, P=0.651). However, several pairwise F_{ST}'s between localities were

289   significant (Table S2).

290   ***Analyses of a single individual per nest***

291   The STRUCTURE analysis had the highest log likelihood for K=2 (Figure S1), and the DeltaK

292   method also suggested K=2 (Figure S2). On examination of the assignment plot (Figure S3) it

293   was apparent that every individual was assigned to both clusters approximately equally,

294   suggesting no substructure according to geography in the dataset.

295   A total of 555 bp of COII sequence data was recovered for mtDNA analysis. There were 25

296   informative differences in the dataset (position: 60 (C/T) 69: (C/T), 78: (A/C), 123:(T/C) 144

297   (C/T), 156 (A/G), 228(C/A), 319 (T/C), 351 (G/A), 354 (C/A) 537(T/C)), which resulted in

298   eleven different haplotypes. The NJ tree resolved two fairly well separated clades, joined by an

299   intermediate haplotype (Figure 1). The most common haplotype was A (66.8% of the sequenced

300   individuals), followed by haplotypes I (19.3%), D (4.6%) and L (4.0%). The remaining

301   haplotypes (B, C, F-H, J, K) were present in 1% or fewer individuals, but all haplotypes were

302   found in more than one individual. Several haplotypes (C, B, D, F, G and L) differed only in one

303   base from the A haplotype. More than one haplotype was present in all localities. The locality

304   MAT had the highest number of haplotypes (five), in its 13 sampled colonies, and the average

305   number of haplotypes across all localities was 2.75 (Table 1). The spatial distribution of

306   mitochondrial haplotypes among localities appeared random (Figure 1) and this pattern was

307   mirrored within localities (plots not show). Mitochondrial differentiation ($F_{ST\text{-}mt} = 0.018$) was not

308   significantly larger than zero, suggesting panmixia. The test for isolation by distance was not

309   significant (r = -0.0415, $P= 0.573$).

310

311   ***Sex-biased dispersal***

312   The overall genetic differentiation among localities estimated from a single individual per nest

313   was larger for the microsatellite genotype data ($F_{ST} = 0.040$; 95%, CI 0.011-0.076) than for the

314   mtDNA data (see above), and only the former was significantly greater than zero. After

315   adjustment (equation 1), the value indicating male gene flow (adjusted $F_{ST} = 0.027$) is six times

316   larger than that indicating female gene flow (adjusted $F_{ST\text{-}mt} = 0.0047$).

317

318   **Discussion**

319

320   Our study of the spatial genetic structure of populations of the pioneer ant species *F. fusca* in

321   southern Finland has produced two key results. First, both nuclear and mitochondrial

322   differentiation among populations are weak and there is no evidence for dispersal limitation. This

323   suggests that the total population is very weakly structured across the entire ca. 35km^2 study

324   area. Second, both female and male gene flow is very high, with a trend for higher female

325   dispersal.

326

327   In line with expectations for independently founding ant species, mitochondrial data showed that

328   the localities were colonized by several different and independent matrilines. Correspondingly,

329   there was neither significant differentiation in mtDNA nor any isolation-by-distance. These

330   results suggest high female dispersal and a panmictic population substructure (Oomen et al.,

331   2011). Our results may even underestimate queen dispersal because the estimates are based on

332   only one sequenced individual per nest. Nest-mate queens are sometimes unrelated (Hannonen,

333   Helanterä & Sundström, 2004; Bargum, Helanterä & Sundström, 2007) and by sampling and

334   sequencing more individuals per nests higher haplotype diversities per locality might have been

335   revealed. Nevertheless, queen dispersal is undoubtedly high in this species.

336

337   High dispersal was also evident from the nuclear data. Bayesian clustering and PCA suggested

338   that the total population is homogenous and overall F_{IS} suggested random mating in the total

339   population. Furthermore, the absence of IBD showed that genetic substructure did not arise with

340   increasing distances either; hence dispersal was not limited across the study area. Patterns arising

341   from nuclear DNA reflect male and queen substructure combined, both confirming high

342   dispersal in queens and suggesting high dispersal also in males. In ants, weak population

343   substructure or panmixia over similar spatial scales has earlier been observed in the monogyne

344   form of *Solenopsis invicta* (Ross & Shoemaker, 1997) and in several species of wasp (Queller,

345   Strassmann & Hughes, 1992; Strassmann, Queller & Solís, 1995; Hoffman, Kovacs &

346   Goodisman, 2008) showing similar or higher rates of dispersal in social Hymenopterans, with

347   independent nest founding. Our results for *F. fusca* probably do not extend to its entire

348   distribution, however, and dispersal barriers such as large water bodies or other dispersal

349   restriction may give rise to genetic differentiation at greater spatial scales.

350

351   By testing for sex-biased dispersal, we found that that gene flow was six times higher in the

352   mitochondrial than in the nuclear genome, suggesting female-biased dispersal in *F. fusca*. Of the

353   15 ant species where sex-biased dispersal has been earlier investigated with the same methods

354   (Table S1), only *F. fusca* clearly shows higher queen than male gene flow, and only seven

355   comparisons were greater in magnitude than those for *F. fusca*. Of the latter comparisons, four

356   were on species where queens have no or altered wings and normal winged males. We note that

357   the power of our analysis is not very high given the high dispersal rates in both queens and

358   males.

359

360   An earlier study on *F. fusca* found that colony fathers (genotyped spermathecal content) were

361   related in an older habitat, but unrelated in a younger habitat (Hannonen, Helanterä &

362   Sundström, 2004), without corresponding changes in queen relatedness. If queens mate before

363   they arrive at the habitat patch, they carry sperm from (unrelated) males in their spermathecae,

364   which would make the male population genetically diverse in young habitat. Previous research

365   has shown annual nest survival rates of about 65% (Bargum & Sundström, 2007), ensuring both

366   that genetic diversity from the founding queens is maintained for several years after the initial

367   colonization, and that there is a constant supply of vacant nest sites. Soon after colonization,

368   daughter queens are produced (Bargum, Helanterä & Sundström, 2007) that local males can mate

369   with. Local mating between males and daughter queens in polygyne nests is expected to increase

370   relatedness among colony fathers with time. This is in line with the findings in (Hannonen,

371   Helanterä & Sundström, 2004) and should result in a greater spatial genetic structure among

372   males than among females, as we report here. If vacated nests sites are predominately colonized

373   by immigrant queens, genetic diversity in the queen population may remain relatively stable

374   across time, despite local mating by males. When our present results and those from earlier

375   studies in *F. fusca* are considered together, female-biased dispersal seems likely in *F. fusca.*

376   Although the general pattern reported from population genetic studies on sex-biased dispersal

377   suggest that male-biased dispersal dominates in ants (Table S1), these studies are few, and show

378   considerable taxonomic and physiological bias. Female-biased dispersal is assumed for species

379   with wingless males, but its actual prevalence among ants with winged sexuals is unknown.

380

381   *Formica fusca* queens are physiologically capable of long distance mating flights. The fat

382   content of *F. fusca* is sufficiently high (ca 50% of bodyweight) to ensure good resources for

383   mating flights (Keller & Passera, 1989). The average weight of *F. fusca* queens is ca 22 mg

384 (range 10.19-30.54 mg, S. Hakala unpublished). Following Helms et al., (2016), ants within this

385 weight-range fly about 35-45 m above ground. *Formica fusca* flight altitude is not known, but

386 the similar-sized and closely related *F. lemani* (Goropashnaya et al., 2012) flies at an altitude of

387 40 m (Duelli, Näf & Baroni-Urbani, 1989). Since *F. fusca's* physiology apparently allows for

388 dispersal but flying potentially occurs at low speeds, wind conditions must be favourable for

389 longer distance dispersal (e. g. Markin et al., 1971). At an altitude of 50 m, the average wind

390 speed in the sampled area during the mating flight season ranges between 3.80-5.60 m/s

391 (http://www.tuuliatlas.fi/en/). Travelling the shortest distance between the localities sampled for

392 this study (1.4 km) would take 4-6 min on such wind currents, and the farthest distance (57.2

393 km) 170- 250 minutes. This is a simplification of insect flight and wind dynamics, but in short,

394 dispersing *F. fusca* sexuals could potentially reach considerable dispersal distances in our study

395 area given a flying altitude of ca 40 m and normal wind conditions.

396

397 The PCA analysis and the hierarchical F-statistics revealed nuclear genetic substructure only at

398 the level of the nest. This pronounced colony-level structure is consistent with monodomy and

399 significant within-colony relatedness, as reported from previous studies (Pamilo, 1983;

400 Hannonen, Helanterä & Sundström, 2004; Bargum, Helanterä & Sundström, 2007). Finally,

401 nuclear differentiation among sampling localities, as indicated by the low F_{ST} estimates, was

402 weak, but significantly greater than zero. Thus, the study area was not genetically entirely

403 homogenous, but there was no pattern of isolation-by-distance, neither among localities, nor

404 among nests across localities, in neither of the markers. Such a pattern is expected to arise when

405 dispersal is high in a spatially and temporally unstable habitat (Wade & McCauley, 1988;

406 Leblois et al., 2000; Yang, Bishop & Webster, 2008; Peel et al., 2013), such that the number of

407 long-distance colonizers founding new populations is at least twice the number of local

408 colonizers. The scenario is consistent with the ecology of *F. fusca,* since optimal habitats are

409 patchily distributed and rapidly colonized (Punttila et al., 1991). These habitats are also

410 temporally variable. Continuous colonization by *F. fusca* into clear-cut plots leads to a steady

411 increase in nest densities for several years (Punttila et al., 1991). As the forest succession

412 advances dispersal remains high but the habitat becomes less suitable for *F. fusca*, and more

413 suitable for mound-building wood ants (*Formica* sensu stricto) who arrive and outcompete them

414 (Punttila et al., 1991). Similar population genetic substructure has been reported for other

415   invasive and pioneer species that colonize plots in a similar fashion to *F. fusca* such as

416   *Vaccinium* plants (Yang, Bishop & Webster, 2008), the invasive cane toad (Leblois et al., 2000)

417   and the invasive yellow crazy ant *Anoplolepis gracilipes* (Drescher, Blüthgen & Feldhaar, 2007).

418

419   Conclusion

420   This study shows high male and queen dispersal, and as a consequence, near panmictic

421   population genetic structure in *F. fusca*. Our results are consistent with independent founding

422   and the winged status of both sexes, and with species' ecology. Temporally and spatially variable

423   habitat necessitates the maintenance of wings and independent founding for rapid colonization

424   and habitat escape, in both queens and males. Polygyny in *F. fusca* may nevertheless be a

425   dispersal adaptation. Re-adopted queens face no dispersal risks, and if unrelated queens can also

426   either be adopted into nests, or if queens can found nest through pleometrosis, dispersal risks are

427   reduced also for dispersing queens. The results in this study contrast with findings from other

428   polygyne European *Formica* species, for which stable habitat and weak dispersal are thought to

429   promote polygyny and polydomy (Rosengren & Pamilo, 1983) They also contrast with results

430   for highly invasive ant species, for which human-mediated dispersal and a highly unstable

431   habitat are closely associated with polygyny and supercoloniality (Holway et al., 2002). Relative

432   to these, *F. fusca's* habitat is intermediate in stability, and perfectly exploited by queen dispersal

433   strategies.

434

435   Acknowledgements

436   We are grateful to Katja Bargum, Stephen J. Martin, Riitta Ovaska and Anniina Mattila for help

437   in collecting samples, and Soile Kupiainen for genotyping. Sanja Hakala kindly shared her weight

438   measurements for *Formica fusca* queens.

439

440

441   **References**

442

443   Bargum K., Helanterä H., Sundström L. 2007. Genetic population structure, queen supersedure

444   and social polymorphism in a social Hymenoptera. *Journal of Evolutionary Biology*

445   20:1351–60. DOI: 10.1111/j.1420-9101.2007.01345.x.

446   Bargum K., Sundström L. 2007. Multiple breeders, breeder shifts and inclusive fitness returns in

447   an ant. *Proceedings. Biological sciences / The Royal Society* 274:1547–51. DOI:

448   10.1098/rspb.2007.0295.

449   Berg LM., Lascoux M., Pamilo P. 1998. The infinite island model with sex-differentiated gene

450   flow. *Heredity* 81:63–68. DOI: 10.1046/j.1365-2540.1998.00358.x.

451   Berghoff SM., Kronauer DJC., Edwards KJ., Franks NR. 2008. Dispersal and population

452   structure of a New World predator, the army ant *Eciton burchellii*. *Journal of Evolutionary*

453   *Biology* 21:1125–32. DOI: 10.1111/j.1420-9101.2008.01531.x.

454   Chapuisat M. 1996. Characterization of microsatellite loci in *Formica lugubris B* and their

455   variability in other ant species. *Molecular Ecology* 5:599–601. DOI: 10.1111/j.1365-

456   294X.1996.tb00354.x.

457   Clobert J., Danchin E., Dhondt A., Nichols J. 2001. *Dispersal*. Oxford University Press.

458   Cronin AL., Molet M., Doums C., Monnin T., Peeters C. 2013. Recurrent evolution of dependent

459   colony foundation across eusocial insects. *Annual Review of Entomology* 58:37–55. DOI:

460   10.1146/annurev-ento-120811-153643.

461   Czechowski W., Radchenko A., Czechowska W. 2002. *The ants (Hymenoptera, Formicidae) of*

462   *Poland*. Museum and Institute of Zoology.

463   Drescher J., Blüthgen N., Feldhaar H. 2007. Population structure and intraspecific aggression in

464   the invasive ant species *Anoplolepis gracilipes* in Malaysian Borneo. *Molecular ecology*

465   16:1453–65. DOI: 10.1111/j.1365-294X.2007.03260.x.

466   Duelli P., Näf W., Baroni-Urbani C. 1989. Flughöhen verschiedener Ameisenarten in der

467   Hochrheinebene. *Mitteilungen der Schweizerischen Entomologischen Gesellschaft =*

468   *Bulletin de la Société Entomologique Suisse* 62:1–4.

469   Earl DA., Bridgett M. 2012. STRUCTURE HARVESTER: a website and program for

470   visualizing STRUCTURE output and implementing the Evanno method. *Conservation*

471   *Genetics Resources* 4:359–361.

472   Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the

473   software STRUCTURE: A simulation study. *Molecular Ecology* 14:2611–2620. DOI:

474   10.1111/j.1365-294X.2005.02553.x.

475   Excoffier L., Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform

476   population genetics analyses under Linux and Windows. *Molecular Ecology Resources*

477   10:564–7. DOI: 10.1111/j.1755-0998.2010.02847.x.

478   Fernández-Escudero I., Pamilo P., Seppä P. 2002. Biased sperm use by polyandrous queens of

479   the ant *Proformica longiseta. Behavioral Ecology and Sociobiology* 51:207–213. DOI:

480   10.1007/s00265-001-0433-3.

481   Fernández-Escudero I., Seppä P., Pamilo P. 2001. Dependent colony founding in the ant

482   Proformica longiseta. *Insectes Sociaux* 48:80–82. DOI: 10.1007/PL00001750.

483   Goropashnaya A V., Fedorov VB., Seifert B., Pamilo P. 2012. Phylogenetic relationships of

484   palaearctic Formica species (Hymenoptera, Formicidae) Based on Mitochondrial

485   Cytochrome b Sequences. *PLoS ONE* 7:e41697. DOI: 10.1371/journal.pone.0041697.

486   Goudet J. 1995. FSTAT (Version 1.2): A computer program to calculate F-Statistics. *Journal of*

487   *Heredity* 86:485–486.

488   Goudet J. 2005. A package for R to compute and test hierarchical F-statistics. *Molecular Ecology*

489   *Notes* 5:184–186. DOI: 10.1111/j.1471-8278.

490   Gouy M., Guindon S., Gascuel O. 2010. SeaView version 4: A multiplatform graphical user

491   interface for sequence alignment and phylogenetic tree building. *Molecular Biology and*

492   *Evolution* 27:221–4. DOI: 10.1093/molbev/msp259.

493   Gyllenstrand N., Gertsch PJ., Pamilo P. 2002. Polymorphic microsatellite DNA markers in the

494   ant *Formica exsecta. Molecular Ecology Notes* 2:67–69. DOI: 10.1046/j.1471-

495   8286.2002.00152.x.

496   Gyllenstrand N., Seppä P. 2003. Conservation genetics of the wood ant, *Formica lugubris*, in a

497   fragmented landscape. *Molecular Ecology* 12:2931–2940. DOI: 10.1046/j.1365-

498   294X.2003.01975.x.

499   Gyllenstrand N., Seppä P., Pamilo P. 2004. Genetic differentiation in sympatric wood ants,

500   *Formica rufa* and *F. polyctena. Insectes Socieaux*:139–145.

501   Hannonen M., Helanterä H., Sundström L. 2004. Habitat age, breeding system and kinship in the

502   ant *Formica fusca. Molecular Ecology* 13:1579–88. DOI: 10.1111/j.1365-

503   294X.2004.02136.x.

504   Hannonen M., Sundström L. 2003. Sociobiology: Worker nepotism among polygynous ants.

505   *Nature* 421:910. DOI: 10.1038/421910a.

506   Hasegawa E., Imai S. 2004. Characterization of microsatellite loci in red wood ants Formica (s.

507   str.) spp. and the related genus Polyergus. *Molecular Ecology Notes* 4:200–203. DOI:

508   10.1111/j.1471-8286.2004.00614.x.

509   Helms JAI. 2018. The flight ecology of ants (Hymenoptera: Formicidae). *Myrmecological News*

510   26:19–30.

511   Helms JA., Godfrey AP., Ames T., Bridge ES. 2016. Predator foraging altitudes reveal the

512   structure of aerial insect communities. *Scientific Reports* 6:28670.

513   Hoffman EA., Kovacs JL., Goodisman MAD. 2008. Genetic structure and breeding system in a

514   social wasp and its social parasite. *BMC Evolutionary Biology* 8:239. DOI: 10.1186/1471-

515   2148-8-239.

516   Holway D., Lach L., Suarez A., Tsutsui ND., Case T. 2002. The Causes and Consequences of

517   Ant Invasions. *Annual Review of Ecology and Systematics* 33:181–233.

518   Holzer B., Keller L., Chapuisat M. 2009. Genetic clusters and sex-biased gene flow in a

519   unicolonial Formica ant. *BMC Evolutionary Biology* 9:69. DOI: 10.1186/1471-2148-9-69.

520   Keller L., Passera L. 1989. Size and fat content of gynes in relation to the mode of colony

521   founding in ants (Hymenoptera; Formicidae). *Oecologia* 80:236–240. DOI:

522   10.1007/BF00380157.

523   Kovach AI., Breton TS., Enterline C., Berlinsky DL. 2013. Identifying the spatial scale of

524   population structure in anadromous rainbow smelt (*Osmerus mordax*). *Fisheries Research*

525   141:95–106. DOI: 10.1016/J.FISHRES.2012.07.008.

526   Leblois R., Rousset F., Tikel D., Moritz C., Estoup A. 2000. Absence of evidence for isolation

527   by distance in an expanding cane toad (*Bufo marinus)* population: an individual-based

528   analysis of microsatellite genotypes. *Molecular Ecology* 9:1905–1909. DOI:

529   10.1046/j.1365-294x.2000.01091.x.

530   Markin GP., Dillier JH., Hill SO., Blum MS., Hermann HR. 1971. Nuptial flight and flight

531   ranges of the imported fire ant, *Solenopsis saevissima richteri* (Hymenoptera: Formicidae).

532   *Journal of Georgia Entomological Society* 6:145–156.

533   Oomen RA., Reudink MW., Nocera JJ., Somers CM., Green MC., Kyle CJ. 2011. Mitochondrial

534   evidence for panmixia despite perceived barriers to gene flow in a widely distributed

535   waterbird. *Journal of Heredity* 102:584–92. DOI: 10.1093/jhered/esr055.

536   Pamilo P. 1983. Genetic differentiation within subdivided populations of Formica ants. *Evolution*

537   37:1010–1022.

538   Pamilo P., Rosengren R. 1984. Evolution of nesting strategies of ants: genetic evidence from

539   different population types of Formica ants. *Biological Journal of the Linnean Society*

540   21:331–348. DOI: 10.1111/j.1095-8312.1984.tb00370.x.

541   Pamilo P., Zhu D., Fortelius W., Rosengren R., Seppä P., Sundström L. 2005. Genetic patchwork

542   of network-building wood ant populations. *Annales Zoologici Fennici* 42:179–187.

543   Patterson N., Price AL., Reich D. 2006. Population structure and eigenanalysis. *PLoS Genetics*

544   2:e190. DOI: 10.1371/journal.pgen.0020190.

545   Peel AJ., Sargan DR., Baker KS., Hayman DT., Barr JA., Crameri G., Suu-Ire R., Broder CC.,

546   Lembo T., Wang LF., Fooks AR., Rossiter SJ., Wood JL., Cunningham AA. 2013.

547   Continent-wide panmixia of an African fruit bat facilitates transmission of potentially

548   zoonotic viruses. *Nature Communications* 4:2770. DOI: 10.1038/ncomms3770.

549   Peeters C., Ito F. 2001. Colony dispersal and the evolution of queen morphology in social

550   Hymenoptera. *Annual Review of Entomology* 46:601–630.

551   Punttila P., Haila Y., Pajunen T., Tbkia H. 1991. Colonisation of clearcut forests by ants in the

552   southern Finnish taiga: a quantitative survey. *Oikos* 61:250–262.

553   Queller DC., Strassmann JE., Hughes CR. 1992. Genetic relatedness and population structure in

554   primitively eusocial wasps in the genus Mischocyttarus (Hymenoptera: Vespidae). *Journal*

555   *of Hymenopteran Research* 1:81–89.

556   Rosengren R., Pamilo P. 1983. The evolution of polygyny and polydomy in mound-building

557   Formica ants. *Acta Entomologica Fennica* 42:65–77.

558   Ross KG. 2001. Molecular ecology of social behaviour: analyses of breeding systems and

559   genetic structure. *Molecular Ecology* 10:265–284. DOI: 10.1046/j.1365-

560   294x.2001.01191.x.

561   Ross KG., Shoemaker DD. 1997. Nuclear and mitochondrial genetic structure in two social

562   forms of the fire ant *Solenopsis invicta:* insights into transitions to an alternate social

563   organization. *Heredity* 78:590–602. DOI: 10.1038/hdy.1997.98.

564   Schmitt T. 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends.

565   *Frontiers in Zoology* 4:11. DOI: 10.1186/1742-9994-4-11.

566 Seppä P., Fernández-Escudero I., Gyllenstrand N., Pamilo P. 2006. Obligatory female philopatry

567   affects genetic population structure in the ant *Proformica longiseta*. *Insectes Sociaux*

568   53:362–368. DOI: 10.1007/s00040-006-0882-9.

569 Seppä P., Gyllenstrand N., Corander J., Pamilo P. 2004. Coexistence of the social types: genetic

570   population structure in the ant *Formica exsecta. Evolution* 58:2462. DOI: 10.1554/04-295.

571 Seppä P., Helanterä HO., Trontti K., Punttila P., Chernenko A., Martin SJ., Sundström L. 2011.

572   The many ways to delimit species: hairs, genes and surface chemistry. *Myrmecological*

573   *News* 51:41–31.

574 Seppä P., Pamilo P. 1995. Gene flow and population viscosity in Myrmica ants. *Heredity*: 200–

575   209.

576 Slatkin M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236:787

577   LP-792.

578 Stadler K. 2017. cultevo: Tools, measures and statistical tests for cultural volution. R package.

579 Strassmann JE., Queller DC., Solís CR. 1995. Genetic relatedness and population structure in the

580   social wasp, *Mischocyttarus mexicanus* (Hymenoptera: Vespidae). *Insectes Sociaux*

581   42:379–383. DOI: 10.1007/BF01242166.

582 Sundström L. 1993. Genetic population structure and sociogenetic organisation in *Formica*

583   *truncorum* (Hymenoptera; Formicidae). *Behavioral Ecology and Sociobiology* 33. DOI:

584   10.1007/BF00172934.

585 Sundström L. 1995. Dispersal polymorphism and physiological condition of males and females

586   in the ant, *Formica truncorum. Behavioral Ecology* 6:132–139. DOI:

587   10.1093/beheco/6.2.132.

588 Sundström L., Seppä P., Pamilo P. 2005. Genetic population structure and dispersal patterns in

589   Formica ants — a review. *Annales Zoological Fennici*:163–177.

590 Wade M., McCauley D. 1988. Extinction and recolonization: Their effects on the genetic

591   differentiation of local populations. *Evolution*:995–1005.

592 Weir B., Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure.

593   *Evolution* 38:1358–1370.

594 Whitlock MC., Mccauley DE. 1999. Indirect measures of gene flow and migration : FST ≠1

595   /(4Nmn+1). *Heredity* 82:117–125. DOI: doi:10.1038/sj.hdy.6884960.

596   Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159.

597   Yang S., Bishop JG., Webster MS. 2008. Colonization genetics of an animal-dispersed plant

598   (*Vaccinium membranaceum*) at Mount St Helens, Washington. *Molecular Ecology* 17:731–

599   40. DOI: 10.1111/j.1365-294X.2007.03625.x.

600

601

Figure 1

Map of sampling localities in southern Finland. Haplotypes found and neighbour-joining tree of the haplotypes shown

Each square represents a nest; a group of squares represents a sampling locality. The colour of the square matches the colour in the NJ tree and shows which haplotype was found in the nest.
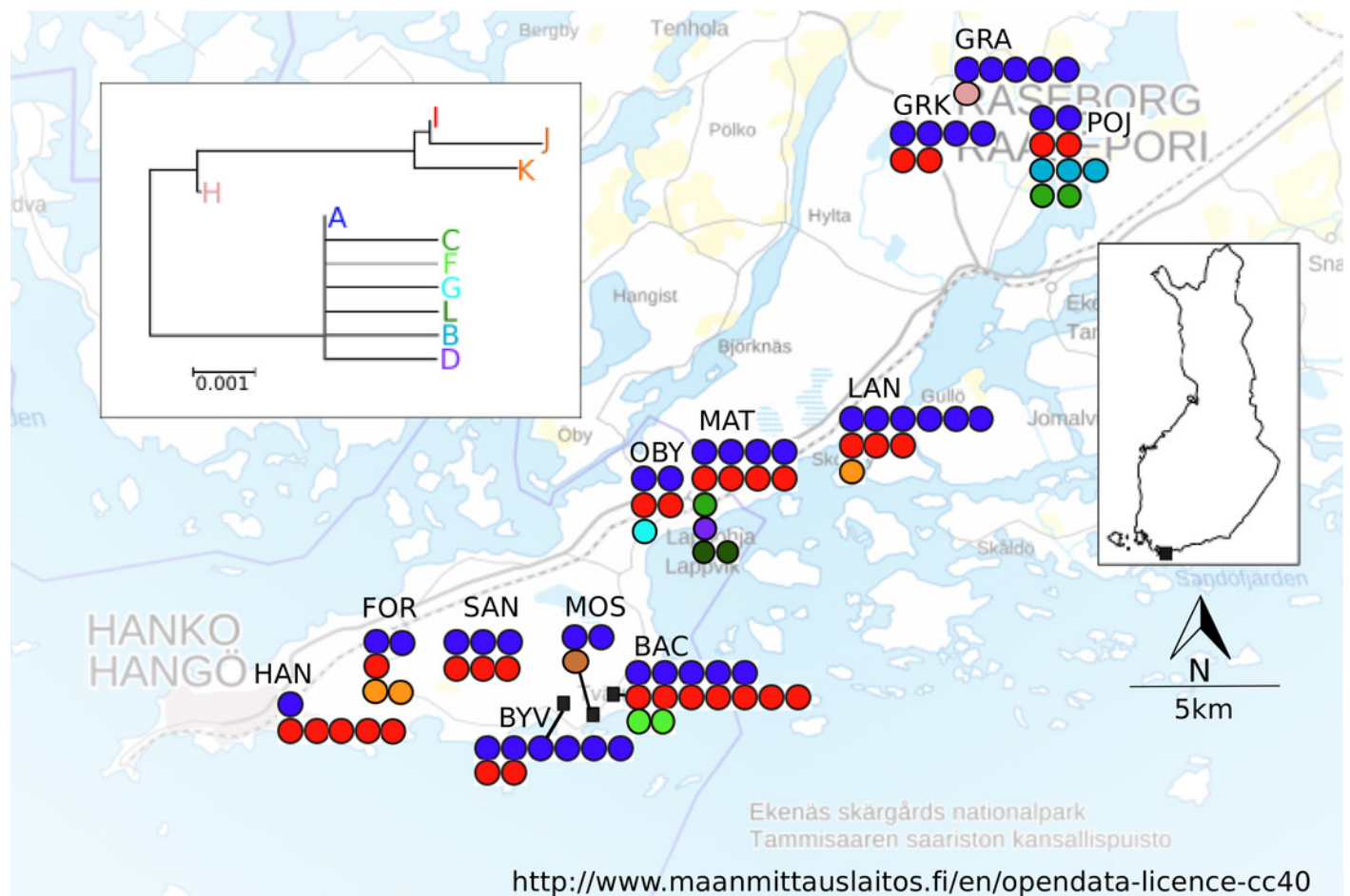
Figure 2

Factor scores from first two principal components from an individual-based principal component analysis on *F. fusca* genotype data

Each dot represents an individual colour-coded by locality.
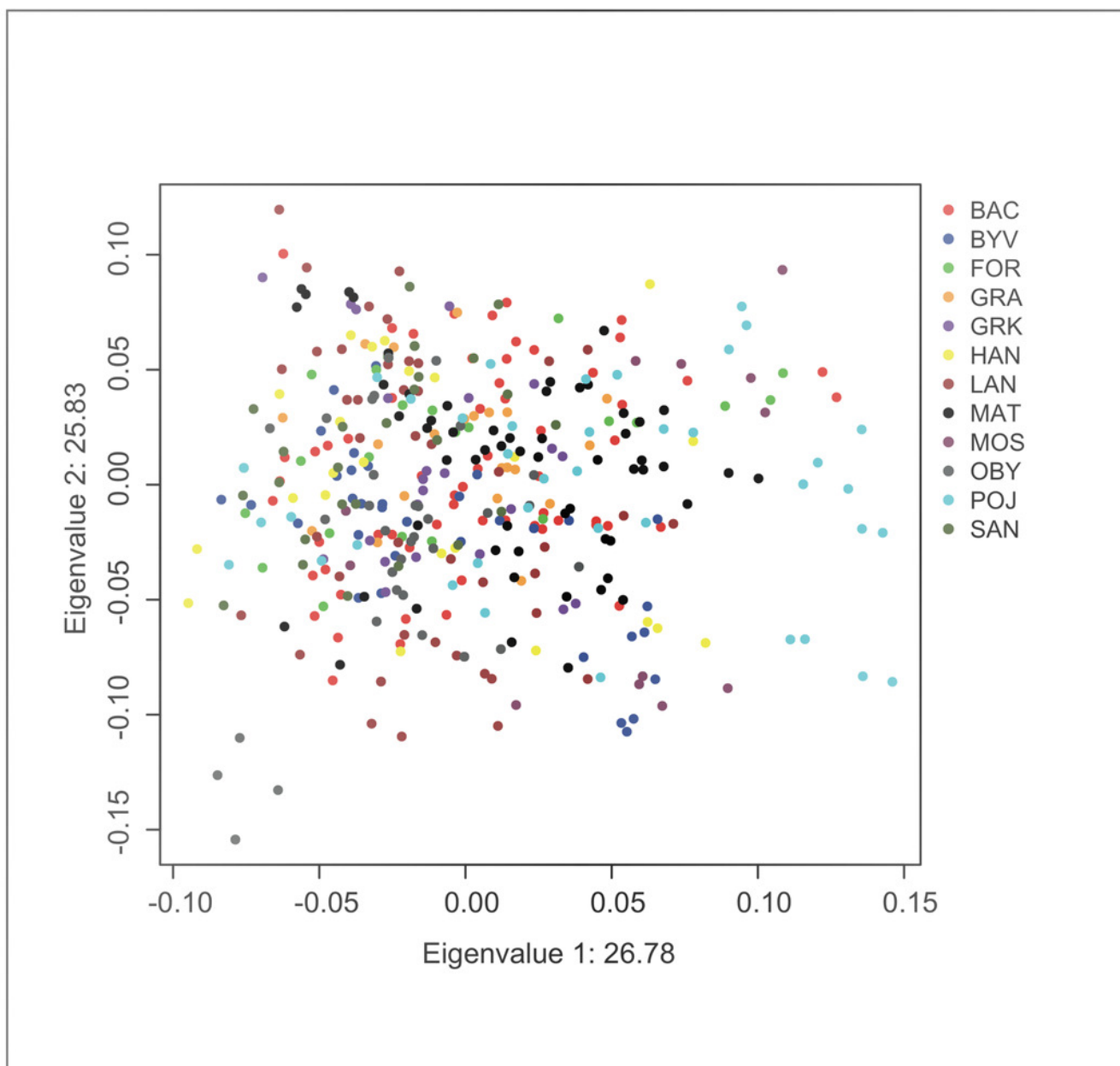
Table 1. Locality names, locality short names (LOC), latitude, longitude, numbers of individuals typed at a minimum of four microsatellite loci (NIND), numbers of colonies (NCOL) Genetic diversity HS, allelic richness AR, FIS and number of haplotypes (NHA

Totals are given for number of individuals and colonies, averages are given for the genetic diversity measures and number of haplotypes.

Name	LOC	latitude	longitude	N_{IND}	N_{COL}	H_S	A_R	F_{IS}	N_{HAP}
Bäckvägen	BAC	59.8438629	23.1559805	73	15	0.468	2.942	0.061	3
Byvägen	BYV	59.8447971	23.1896538	37	8	0.481	3.033	-0.071	2
Forcit	FOR	59.8609232	23.0380195	24	5	0.388	2.473	0.187	3
Grabbskog	GRA	60.0473469	23.4136252	25	6	0.408	2.604	-0.082	2
Gråkärr	GRK	60.0268888	23.3718594	26	6	0.455	2.833	-0.13	2
HankoP	HAN	59.8411238	22.9778762	26	6	0.538	3.367	-0.062	2
Långvik	LAN	59.9320293	23.3350602	47	10	0.505	3.146	0.056	3
Mattakärr	MAT	59.923116	23.2513624	60	13	0.522	3.206	-0.038	5
Mossåker	MOS	59.8460816	23.1682066	15	3	0.557	3.311	0.196	2
Öby	OBY	59.9147305	23.2125462	39	6	0.467	2.798	-0.069	3
Pojo	POJ	60.0328783	23.4600863	42	9	0.569	3.369	0.079	4
Sandö	SAN	59.8608072	23.0791277	29	6	0.386	2.703	-0.172	2
Totals/ averages	-	-	-	443	93	0.479	2.982	0.01	2.75