Insights into the population genetics of an extreme habitat specialist, the wood ant commensal *Formicoxenus nitidulus*

Accepted by **Ecological Entomology**, April 2025

Elva J.H. Robinson[1], Jenni A. Stockan[2], Stuart A'Hara[3], Sandy J. Macdonald[1] & Joan Cottrell[3]

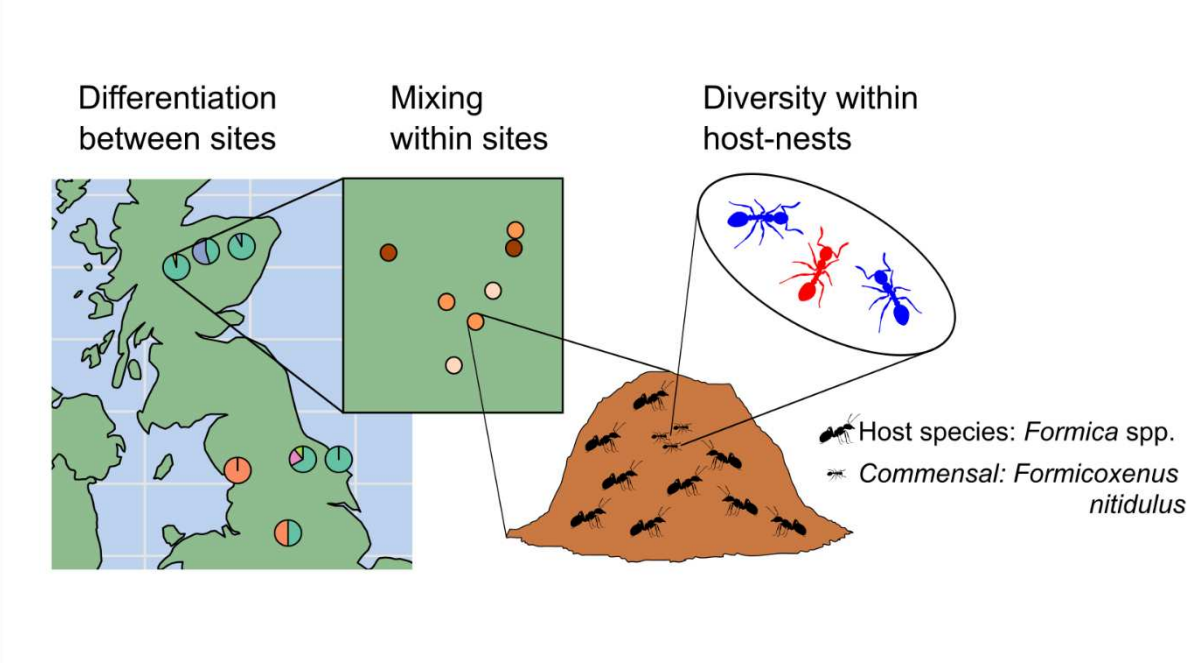
1. University of York, York, England, YO10 5DD
2. The James Hutton Institute, Craigiebuckler Aberdeen, Scotland, AB15 8QH
3. Forest Research, Roslin, Midlothian, Scotland, EH25 9SY

Abstract

1. Habitat specialists, particularly those that are poor dispersers, are highly susceptible to population isolation as a result of habitat fragmentation. Population isolation can lead to inbreeding, resulting in reduced genetic diversity and an increased risk of local extinction.
2. The shining guest ant, *Formicoxenus nitidulus*, lives only within the nests of its wood ant hosts. It is thus an extreme habitat specialist, dependent on patchy nesting sites within an already fragmented woodland habitat. We aimed to generate the first data on the population genetics of this species, to characterize its genetic diversity and degree of population isolation.
3. We developed eight novel nuclear microsatellite markers and generated mitochondrial DNA sequence data of the COI region to characterize samples from seven UK *F. nitidulus* populations collected from nests of three wood ant hosts: *Formica aquilonia*, *Formica lugubris* and *Formica rufa*. These novel nuclear microsatellite markers can be used in future studies of this species across a wider geographic range and may have utility in other closely related species.
4. We found clear differentiation between Scottish and English *F. nitidulus* populations. The six largest study populations were genetically diverse and showed signs of effective within-site dispersal. Our data show that wood ant nests often host multiple *F. nitidulus* colonies.
5. We found genetic diversity has been maintained in this extreme habitat specialist at risk of population isolation. We also demonstrate that a single wood ant host nest can have high conservation importance for the multiple *F. nitidulus* colonies it supports.

Keywords: fragmentation; population genetics, microsatellites, wood ants, commensals

Graphical Abstract



Introduction

Genetic diversity provides the basis of local adaptability and its maintenance reduces the risk of inbreeding depression. Habitat fragmentation can result in genetic isolation and consequent loss of local genetic diversity throughout a species' range (Haddad *et al.* 2015, Wilson *et al.* 2016). Habitat specialists are particularly susceptible to the impacts of habitat fragmentation, especially those that are less able to colonize distant fragments of suitable habitat due to poor dispersal abilities (Martinson and Raupp 2013, Rossetti *et al.* 2017). Globally, woodlands are becoming increasingly fragmented, leaving populations of woodland specialists isolated due to intervening inhospitable habitat (Bergerot *et al.* 2012, Püttker *et al.* 2020).

Woodland specialists differ in their degree of specialism. Forest generalist species can disperse through a variety of woodland types (Gillies and St. Clair 2008, Procter *et al.* 2015), whereas other species may be restricted to moving through forests of a certain age class or type (Gillies and St. Clair 2008, Smith *et al.* 2011). Ants from the *Formica rufa* group occupy a range of forest types, but several species have lost the ability to disperse long-distances, making their population distributions patchy (Risch *et al.* 2016). *Formica rufa* group ants make large, long-lasting nests which host numerous other invertebrate species (Parmentier *et al.* 2014, Robinson *et al.* 2016). Among these is the shining guest ant, *Formicoxenus nitidulus*, which lives exclusively in the thatched nests of ants belonging to the *Formica rufa* and *Formica exsecta* groups, collective here termed 'wood ants', where it is an opportunistic, harmless associate (Francoeur *et al.* 1985, Busch 2001). The survival of *F.*

nitidulus is therefore entirely dependent on queens finding not only suitable woodland, but also host wood ant nests within that woodland. The distribution of the *F. nitidulus* habitat is thus patchy and fragmented at three levels, firstly, at the level of the woodlands, secondly, because only a small proportion of woodlands are occupied by wood ants, and thirdly, even where the host wood ants are present, their nests are often sparsely and unevenly distributed within a woodland (Freitag *et al.* 2016, Risch *et al.* 2016). This extreme fragmentation of habitat suitable for nesting poses considerable challenges to the species' ability to spread to new sites.

In addition to woodland habitat loss and fragmentation, and the patchiness of host nest availability, *F. nitidulus* face yet another challenge to population connectivity. Unusually for ants, male *F. nitidulus* are wingless, making them very poor dispersers. Long-distance dispersal relies instead upon the winged queens, although ergatoid queens lacking wings also occur, further reducing *F. nitidulus* dispersal potential (Buschinger and Winter 1976, Francoeur *et al.* 1985). Taken together with the species habitat context, these aspects of social organization make the species at high risk of having low effective population sizes and high levels of inbreeding and thus it would be expected to be vulnerable to local extinction. As a result, *F. nitidulus* has been identified as a conservation priority, listed as 'Vulnerable' on the IUCN Red List (Social-Insects-Specialist-Group 1996) and appearing on several national and regional lists of species at risk (Lenna and Songia 2008, Seifert 2011, NatureScot 2020).

Despite their vulnerability, nothing is known about the genetic diversity and genetic structure of *F. nitidulus* populations, nor of closely related species. *Formicoxenus* belongs to the *Formicoxenus* genus-group within the Crematogastrini tribe, and is the sister genus to *Leptothorax* (Ojeda *et al.* 2023); these two genera have a Holarctic distribution and comprise around 28 species characterized by small colony size and xenobiosis or social parasitism. Population genetic information is key for accurate assessment of the degree of risk a species is under at both local and national level, and for appropriate conservation action plans to be developed. We therefore aimed to develop and characterize novel species-specific nuclear DNA microsatellite markers, to enable the first population genetic study of *F. nitidulus.* We also aimed to complement this with analysis of mitochondrial DNA sequence data from the Cytochrome Oxidase I (COI) region. The wingless males of *F. nitidulus* mate on the surface of their natal host nest and *F. nitidulus* colonies are functionally monogynous (Buschinger and Winter 1976), so if we detect multiple mitochondrial haplotypes among the males collected from a single wood ant host nest, this indicates that more than one *F. nitidulus* colony is present within the host nest.

Using both the nuclear microsatellite and mitochondrial data, we aimed to determine (i) the level of genetic differentiation between populations of *F. nitidulus* to provide an insight into their degree of genetic isolation (i.e. population fragmentation); (ii) whether there is any between host-species genetic differentiation, by testing for a relationship between host species and genetic similarity of *F. nitidulus* and; (iii) within-site genetic structure and differentiation, by comparing diversity within a host nest to the diversity between host nests, and providing a minimum estimate for the number of *F. nitidulus* colonies per host nest.

Materials and Methods

Sample collection

Within the British Isles, *F. nitidulus* occupies the nests of three host species, the wood ants *Formica aquilonia*, *Formica lugubris* and *Formica rufa* (Francoeur *et al.* 1985, Orledge 2002, Collingwood 2012). We focused primarily on the nests of *Formica aquilonia* and *Formica lugubris*, because these species co-occur at some sites in Scotland, potentially enabling us to identify host associations within the *F. nitidulus* populations at these sites (Stockan *et al.* 2017). We sampled populations at seven sites: two sites where these two hosts occur together, and five sites in which only one of the three host species was present (see Table 1 for site and host species details). Samples were collected on warm days during July-September 2021 by taking *F. nitidulus* males from the surface of their hosts' nests (Härkönen and Sorvari 2017, Stockan *et al.* 2017). Host ant workers were also collected for morphological verification of species identity. At one of the mixed host species sites, Abernethy Forest, *F. nitidulus* were located in several wood ant nests, but unfortunately only in nests of one of the potential hosts, *Formica lugubris.* At the other mixed host species site, Feshiebridge, we were able to collect samples from nests of both the host species present, but although we sought to make a balanced collection from both hosts in terms of sampled nests, only one *Formica aquilonia* nest hosting *F. nitidulus* could be found. The single-host *Formica rufa* site, Gaitbarrows, is a site where *F. nitidulus* has been recorded as abundant in the past in an area of mixed native woodland on limestone pavement (Robinson 2005a, Robinson and Robinson 2013). The wood ant population at this site is in decline (EJHR pers. obs.); *F. nitidulus* colonies are much less numerous than previously observed and could only be sampled from one wood ant nest for this study. Sample locations were recorded using Garmin eTrex GPS with a location accuracy of ~8m under woodland canopy.

In addition to the samples collected in 2021 for this study, a small number of samples collected in 2020 as part of a preliminary feasibility study for this work were also included in the final analyses to increase the overall dataset. Two of these specimens from a single host nest were ergatoid (wingless) females, rather than males (Table 1); all other *F. nitidulus* specimens collected in 2020 and 2021 were males. As is the case for the majority of ant species, *F. nitidulus* are haplodiploid, with diploid females and haploid males, therefore the majority of our samples were haploid individuals.

Molecular analysis

Genomic DNA extractions from whole ants stored in molecular grade ethanol were carried out as follows: the ants were dried at room temperature on a paper towel for 30 mins to allow the ethanol to evaporate and then placed individually in a 1.5 ml microfuge tube with a 3 mm steel ball-bearing. The tubes were then frozen in liquid nitrogen and the samples ground to a fine powder using a Retch 300 mixer-mill at 25 beats per second for 30 secs. A Qiagen Blood and Tissue DNA extraction kit (Qiagen, Germany) was then used to extract DNA following the manufacturer's protocol, although the final elution volume was reduced to 50 µl and the elution was loaded twice onto the extraction column to increase the final DNA yield. A 5 µl aliquot of the extraction was electrophoresed on a 1.2 % ethidium bromide-stained agarose gel and checked under UV light to confirm that the extraction was successful.

The DNA analysis was based on two approaches:

1) Nuclear microsatellite markers
2) Mitochondrial DNA sequence variation

No microsatellite markers were available in the literature for the shining guest ant. Recent next-generation sequencing based approaches have reduced the cost and time required for microsatellite discovery and genotyping (Lepais *et al.* 2020). Marker development was carried out in collaboration with INRAE, Bordeaux by generating a shallow sequencing dataset based on a DNA pool of eight individuals using the QIAseq FX DNA library kit. Illumina pair-end sequencing was carried out on a MiSeq sequencer at 2 x 250 bp read length. The R1 and R2 fastq files obtained were contigged with the bbmap bbmerge tool (version 38.87) with a minimum quality of 25, an overlap of R1 and R2 greater than 100 bp with no mismatch allowed. The QDD pipeline (version 3.1.2) was used to obtain a set of 60 SSR markers which had conserved priming sites encompassing a variable microsatellite repeat.

Of the 60 primer pairs discovered *in silico*, 25 candidate markers amplified loci of the expected size. Results from eight of these markers that produced the most complete

dataset across the samples were taken forward into the final analysis (primer sequences below, see also Table 2). Sequences for the remainder of the markers have been deposited in Genbank (accession numbers OQ384177-193) and can be further optimized for future study.

SSR	Forward primer	Reverse primer
GA13	5'-TTCCTCCTTCCTCAACCGCACCGAT-3'	5'-CGCTTAGGTAACCGAATTGGCGAGT-3'
GA14	5'-TCCCGACAGAATCTTACGAACCTCA-3'	5'-GATGCTGTTTCATGGACGAGCGCGA-3'
GA35	5'-TGGAGTTGCCTGTACACATTACGTGC-3'	5'-GCGATGCCTTGTGTACATATACGCGT-3'
GA45	5'-GCACGTGCTTGGATCTGAGTGGCTG-3'	5'-CGCGCTTCTCATCGTTATTATTGCCA-3'
GA48	5'-AAATGTCCGACTTGGTTGAATCCCA-3'	5'-CGATAAAGTCGTTCGCAGGTGCAAC-3'
GA53	5'-ACCGCGATGCTCTCCCTCTGATCAA-3'	5'-ACGCAATCACTCTCCGAACAAGGCAC-3'
GA54	5'-CGTCTTCACGGAGATTAATGGTGGA-3'	5'-GATGCAAGCCCGTGTGAGCTGAACG-3'
GA58	5'-TCGCGTCTCGAAGTTAATGCGT-3'	5'-CGTGCCATCTCCTCCCATCATGCTT-3'

The samples were genotyped with eight microsatellite markers (Table 2). Each forward primer had a 5'-AGGGTTTTCCCAGTCACGACGTT-'3 M13 sequence attached for detection purposes.

In brief, the 20 µL PCR reactions consisted of 1.2 µL template DNA, 2 µL of 10 x PCR buffer ($(NH)_2SO_4$, pH 8.8, 0.1% Tween 20, 20 mM $MgCl_2$) (Bioron, Germany), 5 pmol of each primer, 0.2 mM of each dNTP, 0.25 µM M13 oligo, labelled with either 700 or 800 nm fluorescent dye, (EUROFINS/MWG Biotech, Germany) and 0.25 U Superhot Taq DNA polymerase (Bioron, Germany). The PCR protocol consisted of an initial denaturation step at 94 ^{o}C for 3 mins, followed by 10 cycles of 94 ^{o}C for 30 secs, 55 ^{o}C for 1 min and 72 ^{o}C for 30 secs. This was followed by 28 cycles of 94 ^{o}C for 30 secs, 53 ^{o}C for 1 min then 72 ^{o}C for 30 secs. A final elongation step at 72 ^{o}C for 6 mins was then carried out. The labelled PCR products were run on a Li-Cor 4300 (Li-Cor Biosciences, Lincoln, NE, USA) and allele sizes scored using reference size standards.

A 533 bp sequence fragment of the Cytochrome Oxidase I (COI) region of the mitochondrial genome was generated using the universal primer pair LCO 1490 (5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3') and HCO 2198 (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') (Folmer *et al.* 1994), for 80 of the 95 sampled individuals, representing 30 host nests distributed across the seven sites.

The COI PCR amplification was carried out as follows: total reaction volume of 20 µl which consisted of: 1.5 µl template DNA, 2 µl 10 x PCR buffer ($(NH)_2SO_4$, pH 8.8, 0.1% Tween 20, 20 mM $MgCl_2$) (Bioron, Germany), 5 pmol of each primer, 0.2 mM of each dNTP, and 0.5 U Superhot Taq DNA polymerase (Bioron, Germany). The PCR conditions were as follows: 95

°C for 4 mins initial denaturation, followed by 30 cycles of 94 °C for 30 secs, 50 °C for 50 secs and 72 °C for 1 min, and a final extension step of 72 °C for 7 mins.

Following amplification, a 5 µl aliquot of PCR product was electrophoresed on a 1.4% agarose gel and, if a single band of the correct size was present, an ExoSAP-IT (Affymetrix, USA) PCR clean-up was carried out as per manufacturer's instructions and sent for Sanger sequencing at James Hutton Institute, Dundee. Sequences were edited and aligned using Sequencher 5.4 (Gene Codes Corporation, USA).

Data analysis

For the microsatellite SSR data, the initial dataset comprised 95 individuals at eight microsatellite loci. After filtering out 15 individuals due to missing data (>25% of loci) the available dataset comprised 80 individuals, representing 29 host nests, genotyped across eight loci. As our samples comprised a mix of haploid (n=77, all male) and diploid (n=3, 1 male, 2 female) individuals, we ran two versions of the analysis, one treating haploid individuals as homozygous diploids, and one treating all as haploid, but excluding the three diploid individuals. This had little to no effect on the resulting F_{ST} and DAPC analysis of the whole dataset, so we present the results from the n=77 version excluding the three diploid individuals.

Pairwise Fixation index (F_{ST}) values between the populations from the seven sites were calculated using the WC84 measure of F_{ST} (Weir and Cockerham 1984). Principal Component Analysis (PCA) and Discriminant Analysis of Principal Components (DAPC) were used to identify groupings within the data, implemented using the 'ade4' version 1.7-22 (Dray and Dufour 2007) and 'adegenet' version 2.1.10 (Jombart and Ahmed 2011) R packages respectively. PCA groups multidimensional data based upon variance between the samples, with each principal component containing a (successively decreasing) proportion of the total variance. DAPC reduces the principal components to fewer dimensions and thus accounts for a larger proportion of the variance in the data. Population structure analysis was performed using 'STRUCTURE' version 2.3.4 (Pritchard *et al.* 2000) with MAXPOPS=1, BURNIN=10000, NUMREPS=20000, and LOCPRIOR=1 and values of K from 1 to 10. StructureSelector (Li and Liu 2018) was used to estimate the optimal value/s of K from which to estimate the populations' ancestry, using Evanno's delta K method (Evanno *et al.* 2005). For each value of K, 5 replicate STRUCTURE analyses were run, and the resulting ancestry proportions were averaged across each value of K using the CluMPAK online server (Kopelman *et al.* 2015).

Detection of population bottlenecks was carried out using the M-ratio test from the 'strataG' R package version 2.5.01 (Archer *et al.* 2017), which is suitable for haploid datasets. Spatial distance between nests was calculated from GPS data using the 'distm' function of the 'geosphere' R package version 1.5-18 (Hijmans *et al.* 2017). Correlations between linearised F_{ST} ($F_{ST*}(F_{ST}/ 1-F_{ST})$ and spatial distance were tested with the Mantel test function from the R package 'vegan' version 2.6-8 (Oksanen *et al.* 2007). The Mantel test was applied using the Spearman correlation method at the between-site level, and also to check for within-site correlation between genetic and spatial distances for sites where at least five host nests were sampled. An AMOVA was carried out to assess within and between population genetic variance using the R package 'poppr' version 2.9.6 (Kamvar *et al.* 2014) with a posthoc permutation test using the 'randtest' function of 'ade4' version 1.7-22 (Dray and Dufour 2007). Further spatial analysis was performed using SPAGeDi 1.5d 2017 (Hardy and Vekemans 2002) to calculate the overall and site-level gene diversity corrected for sample size (*He*) (Nei 1978) and to estimate kinship and its relationship with spatial distribution of samples. The Loiselle kinship estimator (Loiselle *et al.* 1995) was used, because it is suitable for haploid data and is robust to the presence of low frequency alleles. SPAGeDi was run using the 'pairwise', 'within pairs', and 'whole sample reference allele frequencies' options.

To visualize relationships between mitochondrial haplotypes, a haplotype network was constructed using our new data and the eight additional *F. nitidulus* COI sequences available in the BOLD database (two from Finland, two from Switzerland, three from Spain and one from Norway). Sequences were aligned using MUSCLE version 5.2.osxarm64 (Edgar 2004) and the haplotype network was constructed using POPART version 1.7 (Leigh *et al.* 2015) with the TCS network type (Clement *et al.* 2002).

Results

Microsatellite markers

Eight microsatellite markers were developed and tested, and showed sufficient inter-individual variability as to be valuable in investigating population genetics. We found 44 alleles in total, with an average of 5.5 alleles per locus (Table 2). The allelic richness is similar across most of the sites (Figure 1a), but the number of private alleles per site ranged from 0 to 4 (Figure 1b). Allele frequency distributions show a strong left skew (Figure 2a), indicating that rare alleles make up the largest component of the total alleles; it is expected that rare alleles are most likely to be lost during population bottlenecks (Sanllorente *et al.*

2010), therefore this pattern of left skewed distribution is consistent with a lack of a recent bottleneck. This same high abundance of low frequency alleles is seen at the site level; Figure 2b shows Feshiebridge as an illustrative example with the highest sample size; the other sites with large enough sample size to plot distributions (N≥10) showed the same pattern. This visual evidence is supported by M-ratio tests: site-level M-ratio ranged from 0.85-0.98 (Table S1); all these values are well above the threshold value of 0.68, below which a recent reduction in population size can be inferred (Garza and Williamson 2001).

Genetic structure of *F. nitidulus* populations

The Scottish sites are highly differentiated from those in England with an average F_{ST} value of 0.328. This is supported by the STRUCTURE analysis which clearly separates the Scottish and English populations (Figure S4), Gene diversity is lower among the Scottish sites (*He* 0.21-0.27) than the English sites (*He* 0.37-0.44). The two PCA axes of Figure 3a account for 27.8% of the total variation in all genotypes included in the analysis. Axis 1 clearly separates the Scottish sites from those in England; however, within both countries there is considerable overlap in the positioning of samples. Average F_{ST} between all Scottish sites is low (0.146). Average F_{ST} is greater between all English sites (0.294), though this is largely due to the most differentiated site, Gaitbarrows with an average F_{ST} of 0.372. If Gaitbarrows is omitted, the average F_{ST} across all the other English sites is 0.186 which is more similar to that recorded for the Scottish sites. Overall, there is a positive relationship between F_{ST} and spatial distance at the site level (Mantel test r=0.46, n=77, p=0.01, Figure 4a, S1).

These patterns are replicated in the discriminant analysis of principal components (DAPC, Figure 3b) which accounts for 62% of the total variation, albeit with greater discrimination between the sites. The host species from which each individual was collected is also indicated in Figure 3: the two samples from the *F. aquilonia* nest (orange circles) at the mixed host site (Feshiebridge) do not cluster differently from the samples from the *F. lugubris* nests at the same site (orange triangles).

Within-site population structure

For the six sites for which multiple host nests were sampled, we investigated the within-site population structure. The populations appear to be well-mixed within sites as the AMOVA results (Table 3) indicate that very little of the population level genetic differentiation occurs between host nests within a site (1.7%, p=0.607). In contrast, there is significant

variation between sites (23%, p=0.001) and within host nests (76%, p=0.001). Population mixing is also supported by highly similar kinship estimates between individuals within a host nest (0.24) and between host nests within sites (0.23) (Table S2). Comparing F_{ST} values with spatial distance between host nests, there was no within site correlation between genetic and spatial distance (Figure 4b, S2). Abernethy shows some within-site variation in terms of how well-mixed the *F. nitidulus* population is, with F_{ST} between host nests ranging from 0 to 0.46), but there is no positive relationship between genetic differentiation and spatial distance (Mantel test: r=-0.62, n=8, p=0.88). For example, despite its proximity to other host nests, the *F. nitidulus* in AB10 were the most differentiated from other nests on the site, whereas the *F. nitidulus* in AB2020, which was further away from the other host nests, was only moderately differentiated (Figure 4b, Figure S2k,l, S3f). Similarly, at Feshiebridge, (F_{ST}=0-0.41) five of the host nest samples show no differentiation from each other; the other two (FB2020 and FB6) are more distinct, but this does not correlate with distance (Mantel test: r=-0.32, n=21, p=0.86, Figure 4b, S2g,h, S3d). Interestingly, one of the five samples that shows very low differentiation from the others within that site is the sample from the single *F. aquilonia* host nest, FB1. Little to no within-site population differentiation is evident at Longshaw (F_{ST}=0-0.25, Mantel test: r=0.31, n=8, p=0.19, Figure 4b, S2i,j) or at Aviemore (F_{ST}=0-0.09, Figure S2a,b) or Cropton Forest (F_{ST}=0-0, Figure S2e,f). There is some within-site variation in the level of differentiation at the Broxa site (F_{ST}=0-0.33, Figure S2c,d) but with only three host nests sampled, it is not clear whether this relates to distances between host nests.

Other evidence of genetic diversity

In this study we sampled males, which are typically haploid in ants. The presence of diploid males is generally considered to be a sign of inbreeding, and diploid males are usually sterile (Cook and Crozier 1995). Of the 78 males included in the microsatellite analysis only one diploid male (i.e. generating two alleles at a proportion of the markers) was found in our dataset. As expected, both the samples identified morphologically as ergatoid females were recorded as diploid based on the microsatellite results.

Mitochondrial DNA

We identified six haplotypes within the 533 bp sequence fragment of the mitochondrial COI region and examples of these sequences have been deposited in GenBank (Table 4). Four of

these haplotypes differed from the other COI sequences previously deposited in GenBank for this species.

The most common haplotype (haplotype 1) occurred in all sites except for Gaitbarrows and was the haplotype in 58 out of the 78 samples (Figure 5a). Four of the six British haplotypes (haplotypes 2, 3, 5 and 6) each differ from haplotype 1 by only one base. Haplotype 4 differs by three bases, and was found only in Cropton Forest. Multiple haplotypes were present at five of the seven sites. Haplotypes 3 and 6 were only found in the Scottish sites and haplotypes 2, 4 and 5 only occurred in the English samples (Table 4, Figure 5a). Haplotype 2 was restricted to Gaitbarrows and Longshaw, the two most westerly sampling sites in England. Haplotype 3 occurred only in two Scottish sites i.e. Aviemore and Abernethy. Haplotypes 4 and 5 were restricted to Cropton Forest and the only occurrence of haplotype 6 was in Feshiebridge. At Feshiebridge, the only site where we have data for the two co-occurring host species, almost all the ants were haplotype 1, irrespective of wood ant nest host species. The single representative of haplotype 6 was found in the more common of the two hosts, *F. lugubris*. Both ergatoid females found at the Longshaw Estate were haplotype 2.

Two of the UK haplotypes were also found in other European samples (Figure 5b). Haplotype 5 was also found in two Spanish accessions to GenBank, and haplotype 4 in a Swiss accession. The single Norwegian accession differed from haplotype 4 by one base, and the two Finnish accessions differed from haplotype 5 by four and five bases.

At five of our sites, we found host nests containing multiple *F. nitidulus* haplotypes. As males of this species mate on their natal nests, and *F. nitidulus* colonies are functionally monogynous (Buschinger and Winter 1976), the likeliest explanation for the presence of multiple haplotypes in the males from a host nest, is that more than one *F. nitidulus* colony is present within the host nest. Of the 27 host nests for which more than one *F. nitidulus* specimen provided a COI sequence, we detected multiple haplotypes in 10, i.e. 37%. Multiple haplotypes per host nest were detected at all sites for which multiple haplotypes appeared in the dataset (Table 5). By the same logic, we examined the proportion of host nests which contained three different alleles of a given microsatellite locus; as haploid males receive genetic material only from their mothers, again, the likeliest explanation for this diversity is that more than one *F. nitidulus* colony is present within the host nest. Of the 19 host nests for which at least three *F. nitidulus* specimens yielded microsatellite data, 14 (42%) contained three alleles in at least one locus (Table 5). These were not always the same nests as those identified by the mitochondrial haplotype analysis, meaning that overall, 47% of the host nests had at least one source of evidence for the presence of multiple *F. nitidulus* matrilines. These data provide a *minimum* indication of the prevalence

of multiple colonies of *F. nitidulus* present in a host nest, because we would not detect multiple colonies founded by queens that share alleles at our eight loci or share a haplotype, particularly given the frequency of haplotype 1 is so high.

Discussion

Genetic outputs

We developed eight novel microsatellite markers which are sufficiently variable to be informative in investigating the population genetics of *F. nitidulus*. In addition, a further 17 markers were identified which are available for further optimization and testing. The microsatellite markers are the first available in for this branch of the Crematogastrini, and may have wider utility for studying the population genetics of related species including other species of *Formicoxenus* and the twig-dwelling *Leptothorax acervorum*. We also identified six haplotypes in the COI region of the mitochondrial genome. In conjunction with other available *F. nitidulus* COI sequences (two from Finland, two from Switzerland, three from Spain and one from Norway) these sequence data will help inform future work on this conservation priority species, in the UK and throughout its European range if additional individuals from a broader geographical range can be sampled and analysed.

Landscape-level population genetic structure

Our population-level genetic data indicate that the *F. nitidulus* populations at different sites are genetically distinct from each other. In particular there is clear differentiation between English and Scottish sites, both in microsatellite data, and in the mitochondrial DNA with English and Scottish *F. nitidulus* populations each containing multiple unique haplotypes. Within England, Gaitbarrows appears highly differentiated from all other sites, although this could be in part due to the lower sample size from this site skewing the analysis. Among the other three English sites, the differentiation is less than between Scotland and England. Similarly, among the three Scottish sites, the differentiation is lower than between countries, in particular between Abernethy and Feshiebridge, suggesting that *F. nitidulus* is not genetically isolated between sites at this geographic scale. There is also within-site mitochondrial diversity, with five of the seven sites revealing multiple haplotypes. This indicates either that sites were originally colonised by multiple queens, or that queens are able to move between sites via long-distance dispersal, supporting the microsatellite evidence that indicates a lack of recent genetic bottlenecks.

Numerous invertebrate species are specialist myrmecophiles (Hölldobler and Kwapich 2022), or specialise in occupying particular plant structures e.g. galls or domatia (Sanver and Hawkins 2000, Palmer and Young 2017), and thus face the challenges of inhabiting highly patchy and disjunct habitats. Dispersal to find new patches is high risk, but also potentially high reward, if the disperser is the first to find the new habitat patch. Staying local is lower risk, especially if habitat patches are stable (Travis and Dytham 1999) but may result in kin competition and/or inbreeding, as seen in gall thrips (McLeish *et al.* 2006). Little is known about the long-range dispersal patterns of winged ants in general (Helms 2018). Winged *F. nitidulus* queens could potentially disperse long distances over unsuitable habitat, especially if aided by the wind, as has been reported for closely related species (Seifert and Hagman 2015), consequently gene flow between sites may be ongoing. Further work would be required to confirm this.

Within-site population genetic structure

Extreme specialists in patchy habitats can become trapped in a local minimum. This is because most dispersers will not find suitable habitat and long-distance dispersal ability is often lost in such species (Travis and Dytham 1999, Sanllorente *et al.* 2010). Habitat specialism in a stable but patchy landscape is the likely driver for the loss of dispersal ability in male *F. nitidulus* (Heinze and Tsuji 1995) and also for the occurrence of ergatoid (wingless, physically worker-like) queens in this species (Buschinger and Winter 1976), as observed at one site in this study. Reduced dispersal increases risk of local inbreeding, which in turn can lead to a loss of genetic diversity. We were unable to calculate an inbreeding coefficient (F_{IS}), because we focused our sampling on males which are usually haploid, but some indirect evidence of low inbreeding is available. The single occurrence of a diploid male in our samples could indicate that inbreeding is likely to be low. Hymenopteran males are usually haploid, but diploid individuals which are homozygous at the sex-determining locus also develop as male. These diploid males can occur at a low level in any hymenopteran population, but their frequency is higher when genetic variation within the population is low, e.g. due to a genetic bottleneck and/or inbreeding (Collet *et al.* 2016). Diploid males are costly to colonies that invest in rearing them as they are sterile or sub-fertile (Cook and Crozier 1995). Diploid male production has not been quantified in *F. nitidulus* populations, but closely-related ant species produce around 10% diploid males (Herbers and Grieco 1994, Foitzik and Heinze 2001). While we did not test directly for numbers of chromosomes in our samples, it would be expected that a diploid individual would show some heterozygosity across eight polymorphic loci. Only one male (1.3%) within our samples was considered to be diploid on the basis that it showed heterozygosity, indicating that the proportion of diploid males in our study populations is likely to be very

low; however, we cannot rule out the possibility that diploid males are produced at a greater rate but do not survive to maturity.

Our detection of multiple mitochondrial haplotypes per host nest is also indicative of within-site population mixing. This is supported by the within-nest kinship estimate of 0.23, which indicates that multiple males present within a nest were not always siblings (expected kinship of 0.5 for haploid brothers). Colonies of *F. nitidulus* are functionally monogynous and up to 8 colonies have been found within a host nest (Buschinger and Winter 1976). The presence of multiple same-haplotype *F. nitidulus* colonies within a host nest might indicate a complete failure to disperse even locally within a site*, i.e.* daughter queens staying in their natal host nest, mating and establishing a colony there, rather than seeking a new host nest. In contrast, our mitochondrial and microsatellite data together indicate that multiple genetically distinct *F. nitidulus* colonies are present within almost half of the host nests we studied, and this is supported by the similar estimates of within and between nest kinship, suggesting that mated *F. nitidulus* queens do not preferentially establish new colonies in their natal host nest. Together, these lines of evidence indicate successful dispersal between wood ant host nests by queens, and thus shows that within-site population mixing, necessary to avoid inbreeding, is occurring. Relatively little is known about the genetic diversity of within-nest myrmecophile communities, though where it has been reported, it is usually high, for example in the myrmecophilous spider *Masoncus pogonophilus* each ant nest hosts many unrelated spiders (Cushing 1998) and a single ant nest can harbour eggs from different myrmecophilous hoverfly *Microdon myrmicae* females (Scarparo *et al.* 2021). It is likely that evolving mechanisms to find new nests is as key an adaptation for a myrmecophilous lifestyle as avoidance of attack by the host species, though the former has received much less attention than the latter.

Our study was not designed specifically to characterize within-site population structure; however, our microsatellite data do provide some information on this. The relatively high within-site differentiation in the microsatellite data between certain pairs of host nests may show that there is a degree of isolation between some host nests at some of the sites, but it is notable that there is no sign of a positive correlation between genetic differentiation and distance between pairs of host nests, within the scale of our sampling ranges at the sites we studied (Figure 4b) and no significant overall within-site genetic differentiation between host nests (Table 3). Specifically, within host nest relatedness is similar to between host nest relatedness, and males from a certain host nest are, in many cases, as closely related to males from adjacent host nests as to males from host nests >70 m away. Other invertebrates that live in patchy but stable habitats show local dispersal resulting in spatially structured populations (Cushing 1998, Schönrogge *et al.* 2006). In contrast, our results suggest that *F. nitidulus* disperses widely within a site, breeding with ants from other host

nests, similar to in the behaviour of *M. myrmicae* (Scarparo *et al.* 2021); however, a more systematic within-site sampling process would be required to substantiate this interpretation. With relatively little data available on population viscosity in myrmecophilous species, it is unclear what drivers influence these different dispersal patterns; a comparative study across multiple species inhabiting the same ant nests, or other similar biological 'islands', would be very illuminating.

Host-species

While *F. nitidulus* is an obligate commensal of wood ants, it is a host generalist within this group of ants, enabled by its strategy of generalized chemical repellence, rather than mimicking host-specific cuticular hydrocarbons (Martin *et al.* 2007). We found no evidence for isolation by host nest species, neither with the microsatellite nor mitochondrial data. At Feshiebridge, the one site at which *F. nitidulus* samples were collected from two host nest species (*Formica aquilonia* and *Formica lugubris*) there was no differentiation between the occupants of different host nests. We are unable to draw any more general conclusions about the impact of the host species on *F. nitidulus* population genetics, because our mixed *Formica aquilonia* and *Formica lugubris* host sites did not yield a balanced set of samples of *F. nitidulus*: all but one of the wood ant nests in which *F. nitidulus* were found at these mixed sites were *Formica lugubris*, despite active efforts to collect from both host species. This contrasts with *F. nitidulus* distribution across mixed host Finnish sites, where although it was found most commonly in the nests of *F. polyctena* (a species not found in the UK) it occurred more frequently in nests of *Formica aquilonia* than those of *Formica lugubris*. *Formica lugubris* is a woodland edge specialist species that requires a more open canopy than *Formica aquilonia* (Risch *et al.* 2016) and is more numerous at both our mixed sites (JAS unpublished data). The bias in *F. nitidulus* distribution between hosts observed here could arise from a preference for host nests in sunnier locations or could simply be an artefact of the small population sizes. Future work targeted in more detail at mixed host sites, including those where a wider range of wood ant hosts are available, would be fruitful in identifying whether the lack of differentiation by host species we see here is a general phenomenon.

Population origins

We recorded six mitochondrial haplotypes in our dataset, only one of which occurred in both Scotland and England. This was the most common of the haplotypes, suggesting the Scottish and English populations may share a single origin; however, we would need to conduct a study of samples from a much broader geographic range across the species

distribution range to construct a phylogeographic tree to generate informative mitochondrial DNA lineages. This could inform on the historic colonization routes taken by *F. nitidulus* across Europe and identify which refugial sources arrived in Britain. Our study has revealed that UK mitochondrial haplotypes 4 and 5 are also present in Swiss and Spanish *F. nitidulus* respectively, so further study with a broader geographic coverage would be valuable to place the UK *F. nitidulus* populations in the wider European context.

Implications for conservation

Conservation of *F. nitidulus* and their wood ant hosts can be informed by our findings. Surveys recording this elusive species usually record presence/absence data at the level of the host nest (Robinson 2005b, Härkönen and Sorvari 2017, Stockan *et al.* 2017). Our data indicate that the effective population size of *F. nitidulus* at a given site is likely to be somewhat higher than the number of wood ant nests in which it is present, as there can be several colonies within a single host nest. In addition, this means that the conservation value of each wood ant nest that is hosting *F. nitidulus* can be considered to be even higher than previously thought (Balzani *et al.* 2022). The distribution of *F. nitidulus* colonies within a wood ant population can be uneven, with some wood ant nests hosting multiple colonies, while the majority of wood ant nests at the same site have none (Rare Invertebrates of the Cairngorms, unpublished data). This means that disruption to just one single hosting wood ant nest could have a disproportionate effect on *F. nitidulus* populations. The Aviemore population is particularly of note in this context, as the only site where we were able to detect multiple *F. nitidulus* colonies in every host nest included in our study. While the host nests are sparser at this site they are also larger than at other sites. *Formicoxenus nitidulus* is found more commonly in larger wood ant nests (Härkönen and Sorvari 2017) possibly due to the warmer and more stable conditions they provide (Dietrich 1997, Chen and Robinson 2015).

For the Scottish sites at least, our data suggest that higher *F. nitidulus* genetic diversity is found in areas with larger wood ant populations (Feshiebridge c.60 host nests, Abernethy 20-30 host nests, Aviemore <10 host nests) although more extensive site-level comparisons would be required to draw firm conclusions about this, as we were also able to generate a larger sample size of *F. nitidulus* at more populous host sites. Higher density of wood ant nests is also associated with greater abundance of *F. nitidulus* (Härkönen and Sorvari 2017). Wood ant nests occur at higher densities when colonies are polydomous, i.e. occupy multiple socially connected nests and reproduce by local budding (Ellis and Robinson 2014, Maeder *et al.* 2016). Maintaining suitable habitat for polydomous wood ant colonies (Sudd 1983, Robinson and Stockan 2016, Sorvari 2016) could therefore be particularly beneficial to *F. nitidulus* too. In polydomous nest networks, the trails connecting their host nests are used

by *F. nitidulus* to aid local dispersal (Boer *et al.* 1995, van Hengel 2011); future comparison between the population genetics of *F. nitidulus* within monodomous and polydomous host populations would be valuable to demonstrate how significant local dispersal via host trails is to within-population gene flow.

Our data indicate that the *F. nitidulus* populations at the Scottish sites were no less genetically diverse than the English sites, despite the fact that several of the English sites represent more substantial wood ant populations (Chen and Robinson 2015, Procter *et al.* 2015) with more *F. nitidulus* records (Orledge 2002). One possible explanation is that *F. nitidulus* are actually more abundant in Scottish wood ant populations than has been previously recorded. The established technique of recording *F. nitidulus* on the surface of their hosts' nests relies on the assumption that *F. nitidulus* are consistent between host nests in their tendency to emerge. Differences were evident during this study between English and Scottish populations, with *F. nitidulus* in English populations being visible on their host nests earlier in the year (from July) and in greater numbers compared to the Scottish populations (EJHR & JAS pers. obs.). It is possible that due to either climatic factors or local adaptation, Scottish *F. nitidulus* emerge onto host nest surfaces less readily and are thus more under-recorded than previously thought.

In conclusion, our results show that *F. nitidulus* populations at six sites within Scotland and England contain genetic diversity and maintain effective within-site dispersal, and that Scottish and English populations of *F. nitidulus* are genetically distinct. We find that wood ant nests that host *F. nitidulus* often contain more than one of these cryptic colonies, and thus are of high conservation value. Managing woodlands sensitively for the needs of wood ants benefits not just the wood ants themselves, but also all the commensal species that they host, including these obligate guest ants which are of conservation concern.

Acknowledgements

James Silvey and Peter Ashton provided advice at the planning stage. Assistance with field sampling was provided by Ross Johnston and Jacob Podesta. We are grateful to the Forestry Commission, Forestry and Land Scotland, the National Trust, Natural England, the RSPB and the Seafield Estate for site access. Bioinformatics analyses were supported by the Bioscience Technology Facility Data Science Group. The work was funded by Future Woodlands Scotland, with in-kind support from Forest Research, The University of York and the James Hutton Institute.

Contribution of authors

EJHR: Conceptualization; Data Curation; Funding Acquisition; Investigation; Project Administration; Resources, Supervision; Writing – original draft; Writing – review and editing. JAS Conceptualization; Funding Acquisition; Investigation; Resources; Writing – review and editing. SA'H: Conceptualization; Funding Acquisition; Investigation; Resources; Writing – review and editing. SJM: Data Curation; Formal Analysis; Investigation; Visualization; Writing – review and editing. JC: Conceptualization; Funding Acquisition; Investigation; Resources; Writing – review and editing.

Ethics statement

Formicoxenus nitidulus is listed as Vulnerable by the IUCN (Social-Insects-Specialist-Group 1996). The sampling carried out was approved by the Department of Biology Ethics Committee, University of York (Ref: ER202106). The negative impact of our sampling was minimized by taking no more than 4 specimens per host nest, and by collecting specimens from the host nest surface so that *F. nitidulus* colonies were not disturbed and queens (breeding females) were not removed.

Conflict of Interest Statement

All authors declare that they have no conflict of interests.

Data Availability

The data that support the findings of this study have been deposited in Dryad (Robinson *et al.* 2025). The data, code used to analyse the data, and figures, are available at GitHub: https://github.com/uoy-research/guest-ant-population-genetics.

References

Archer, F. I., P. E. Adams and B. B. Schneiders (2017). "stratag: An r package for manipulating, summarizing and analysing population genetic data." Molecular Ecology Resources **17**(1): 5-11.

Balzani, P., W. Dekoninck, H. Feldhaar, A. Freitag, F. Frizzi, J. Frouz, A. Masoni, E. Robinson, J. Sorvari and G. Santini (2022). "Challenges and a call to action for protecting European red wood ants." Conservation Biology **36**(6): e13959.

Bergerot, B., T. Merckx, H. Van Dyck and M. Baguette (2012). "Habitat fragmentation impacts mobility in a common and widespread woodland butterfly: do sexes respond differently?" BMC Ecology **12**(1): 1-11.

Boer, P., P. Boting, P. Dijkshra and H. Vallenduuk (1995). "*Formicoxenus nitidulus* in the Netherlands as guest in *Formica* nests (Hymenoptera: Formicidae, Myrmicinae)." Entomologische Berichten, Amsterdam **55**(1): 1-3.

Burns, D. D. R., D. W. Franks, C. L. Parr, C. Hawke, S. Ellis and E. J. H. Robinson (2020). "A longitudinal study of nest occupancy, trail networks and foraging in a polydomous wood ant population." Insectes Sociaux **67**: 419–427.

Busch, T. (2001). "Verbreitung der Gastameise *Formicoxenus nitidilus* (Nyl.) in Mecklenburg-Vorpommern (Norddeutschland) sowie bemerkenswerte Beobachtungen (Hymenoptera, Formicidae)." Ameisenschutz Aktuell **15**(3): 69-86.

Buschinger, A. and U. Winter (1976). "Funktionelle Monogynie bei der Gastameise *Formicoxenus nitidulus* (Nyl.) (Hym., Form.)." Insectes Sociaux **23**(4): 549-558.

Chen, Y. H. and E. J. H. Robinson (2015). "The relationship between canopy cover and colony size of the wood ant *Formica lugubris* - implications for the thermal effects on a keystone ant species." PLoS One **9**(2): e116113.

Clement, M., Q. Snell, P. Walke, D. Posada and K. Crandall (2002). TCS: estimating gene genealogies. Parallel and distributed processing symposium, international, IEEE Computer Society.

Collet, M., C. Vayssade, A. Auguste, L. Mouton, E. Desouhant, T. Malausa and X. Fauvergue (2016). "Diploid male production correlates with genetic diversity in the parasitoid wasp *Venturia canescens*: a genetic approach with new microsatellite markers." Ecology and Evolution **6**(18): 6721-6734.

Collingwood, C. A. (2012). "The Formicidae of North Europe." Entomologist's Monthly Magazine **148**: 1-13.

Cook, J. M. and R. H. Crozier (1995). "Sex determination and population biology in the Hymenoptera." Trends in Ecology & Evolution **10**(7): 281-286.

Cushing, P. E. (1998). "Population structure of the ant nest symbiont *Masoncus pogonophilus* (Araneae: Linyphiidae)." Annals of the Entomological Society of America **91**(5): 626-631.

Dietrich, C. O. (1997). "Quantifizierungsversuch des Vorkommens der Glänzenden Gastameise, *Formicoxenus nitidulus* (Nyl.), bei der Gebirgswaldameise *Formica lugubris* Zett. am Muttersbergmas-siv (Österreich: Vorarlberg, Lechtaler Alpen)." Verhandlungen der Zoologisch-Botanischen Gesellschaft in Österreich **134**: 119-132.

Dray, S. and A.-B. Dufour (2007). "The ade4 package: implementing the duality diagram for ecologists." Journal of statistical software **22**: 1-20.

Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic acids research **32**(5): 1792-1797.

Ellis, S. and E. J. H. Robinson (2014). "Polydomy in red wood ants." Insectes Sociaux **61**(2): 111-122.

Evanno, G., S. Regnaut and J. Goudet (2005). "Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study." Molecular ecology **14**(8): 2611-2620.

Foitzik, S. and J. Heinze (2001). "Microgeographic genetic structure and intraspecific parasitism in the ant *Leptothorax nylanderi*." Ecological Entomology **26**(5): 449-456.

Folmer, O., M. Black, W. Hoeh, R. Lutz and R. Vrijenhoek (1994). "DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates." Molecular Marine Biology and Biotechnology **3**(5): 294-299.

Francoeur, A., R. Loiselle and A. Buschinger (1985). "Biosystématique de la tribu Leptothoracini (Formicidae, Hymenoptera). 1. Le genre *Formicoxenus* dans la région holarctique." Le Naturaliste Canadien **112**: 343-403.

Freitag, A., J. A. Stockan, D. Cherix, C. Bernasconi and A. Maeder (2016). Sampling and monitoring wood ants. Wood Ant Ecology and Conservation. J. A. Stockan and E. J. H. Robinson. Cambridge, Cambridge University Press.

Garza, J. C. and E. G. Williamson (2001). "Detection of reduction in population size using data from microsatellite loci." Molecular Ecology **10**(2): 305-318.

Gillies, C. S. and C. C. St. Clair (2008). "Riparian corridors enhance movement of a forest specialist bird in fragmented tropical forest." Proceedings of the National Academy of Sciences **105**(50): 19774-19779.

Haddad, N. M., L. A. Brudvig, J. Clobert, K. F. Davies, A. Gonzalez, R. D. Holt, T. E. Lovejoy, J. O. Sexton, M. P. Austin and C. D. Collins (2015). "Habitat fragmentation and its lasting impact on Earth's ecosystems." Science Advances **1**(2): e1500052.

Hardy, O. J. and X. Vekemans (2002). "SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels." Molecular Ecology Notes 2 **2**: 618-620.

Härkönen, S. K. and J. Sorvari (2017). "Effect of host species, host nest density and nest size on the occurrence of the shining guest ant *Formicoxenus nitidulus* (Hymenoptera: Formicidae)." Journal of Insect Conservation **21**: 477-485.

Heinze, J. and K. Tsuji (1995). "Ant reproductive strategies." Population Ecology **37**(2): 135-149.

Helms, J. A. (2018). "The flight ecology of ants (Hymenoptera: Formicidae)." Myrmecological News **26**(2): 19-30.

Herbers, J. M. and S. Grieco (1994). "Population structure of *Leptothorax ambiguous* a facultatively polygynous and polydomous ant species." Journal of Evolutionary Biology **7**: 581-598.

Hijmans, R. J., E. Williams, C. Vennes and M. R. J. Hijmans (2017). "Package 'geosphere'." Spherical trigonometry **1**(7): 1-45.

Hölldobler, B. and C. L. Kwapich (2022). The guests of ants: how myrmecophiles interact with their hosts, Harvard University Press.

Jombart, T. and I. Ahmed (2011). "adegenet 1.3-1: new tools for the analysis of genome-wide SNP data." Bioinformatics **27**(21): 3070-3071.

Kamvar, Z. N., J. F. Tabima and N. J. Grünwald (2014). "Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction." PeerJ **2**: e281.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg and I. Mayrose (2015). "Clumpak: a program for identifying clustering modes and packaging population structure inferences across K." Molecular ecology resources **15**(5): 1179-1191.

Leigh, J. W., D. Bryant and S. Nakagawa (2015). "POPART: full-feature software for haplotype network construction." Methods in Ecology & Evolution **6**(9).

Lenna, P. and A. Songia (2008). Legge Regionale 10/2008: Disposizioni per la tutela e la conservazione della piccola fauna, della flora e della vegetazione spontanea. R. L. S. V. e. Paesaggio. Lombardy, Autochthonous Flora Center of the Lombardy Region (CFA).

Lepais, O., E. Chancerel, C. Boury, F. Salin, A. Manicki, L. Taillebois, C. Dutech, A. Aissi, C. F. E. Bacles and F. Daverat (2020). "Fast sequence-based microsatellite genotyping development workflow." PeerJ **8**: e9085.

Li, Y. L. and J. X. Liu (2018). "StructureSelector: a web based software to select and visualize the optimal number of clusters using multiple methods." Molecular Ecology Resources **18**: 176–177.

Loiselle, B. A., V. L. Sork, J. Nason and C. Graham (1995). "Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae)." American Journal of Botany **82**: 1420-1425.

Maeder, A., D. Cherix, C. Bernasconi, A. Freitag and S. Ellis (2016). Wood ant reproductive biology and social systems. Wood Ant Ecology and Conservation. J. A. Stockan and E. J. H. Robinson. Cambridge, Cambridge University Press.

Martin, S. J., E. A. Jenner and F. P. Drijfhout (2007). "Chemical deterrent enables a socially parasitic ant to invade multiple hosts." Proceedings of the Royal Society of London Series B **274**: 2717-2721.

Martinson, H. M. and M. J. Raupp (2013). "A meta‐analysis of the effects of urbanization on ground beetle communities." Ecosphere **4**(5): 1-24.

McLeish, M. J., T. W. Chapman and B. Crespi (2006). "Inbreeding ancestors: the role of sibmating in the social evolution of gall thrips." Journal of Heredity **97**(1): 31-38.

NatureScot (2020). "Scottish Biodiversity List."

Nei, M. (1978). "Estimation of average heterozygosity and genetic distance from a small number of individuals." Genetics **89**(3): 583-590.

Ojeda, D. I., M. John, R. L. Hammond, R. Savolainen, K. Vepsäläinen and T. Kvamme (2023). "Phylogeny of the *Formicoxenus* genus-group (Hymenoptera: Formicidae) reveals isolated lineages of *Leptothorax acervorum* in the Iberian Peninsula predating the Last Glacial Maximum." Biodiversity Journal **14**(4): 547-570.

Oksanen, J., R. Kindt, P. Legendre, B. O'Hara, M. H. H. Stevens, M. J. Oksanen and M. Suggests (2007). "The vegan package." Community ecology package **10**(631-637): 719.

Orledge, G. M. (2002). *Formicoxenus nitidulus* (Nylander, 1846). Provisional Atlas of Aculeate Hymenoptera of Britain and Ireland, Part 4. R. Edwards and M. Telfer. Cambridgeshire, Centre for Ecology and Hydrology**:** 26-27.

Palmer, T. M. and T. P. Young (2017). Integrating ecological complexity into our understanding of ant-plant mutualism: ant-acacia interactions in African Savannas. Ant-Plant Interactions. P. S. Oliveira and S. Koptur, Cambridge University Press**:** 200-222.

Parmentier, T., W. Dekoninck and T. Wenseleers (2014). "A highly diverse microcosm in a hostile world: a review on the associates of red wood ants (*Formica rufa* group)." Insectes Sociaux **61**(3): 229-237.

Pritchard, J. K., M. Stephens and P. Donnelly (2000). "Inference of population structure using multilocus genotype data." Genetics **155**(2): 945-959.

Procter, D. S., J. Cottrell, K. Watts and E. J. H. Robinson (2015). "Do non-native conifer plantations provide benefits for a native forest specialist, the wood ant *Formica lugubris*?" Forest Ecology and Management **357**: 22-32.

Püttker, T., R. Crouzeilles, M. Almeida-Gomes, M. Schmoeller, D. Maurenza, H. Alves-Pinto, R. Pardini, M. V. Vieira, C. Banks-Leite and C. R. Fonseca (2020). "Indirect effects of

habitat loss via habitat fragmentation: A cross-taxa analysis of forest-dependent species." Biological Conservation **241**: 108368.

Risch, A. C., S. Ellis and H. Wiswell (2016). Where and why? Wood ant population ecology. Wood Ant Ecology and Conservation. J. A. Stockan and E. J. H. Robinson. Cambridge, Cambridge University Press**:** 81-105.

Robinson, E. J. H. and J. A. Stockan (2016). Future directions for wood ant ecology and conservation. Wood Ant Ecology and Conservation. J. A. Stockan and E. J. H. Robinson. Cambridge, Cambridge University Press**:** 287-299.

Robinson, E. J. H., J. A. Stockan, S. A'Hara, S. J. Macdonald and J. Cottrell (2025). Data for: Insights into the population genetics of an extreme habitat specialist, the wood ant commensal *Formicoxenus nitidulus*. Dryad 10.5061/dryad.9zw3r22rt.

Robinson, E. J. H., J. A. Stockan and G. R. Iason (2016). Wood ants and their interaction with other organisms. Wood Ant Ecology and Conservation. J. A. Stockan and E. J. H. Robinson. Cambridge, Cambridge University Press**:** 177-206.

Robinson, N. A. (2005a). "The Red Wood Ant - the AONB's Special Ant." Keer to Kent **57**: 10-11.

Robinson, N. A. (2005b). "The 'Uninvited Guest Ant' *Formicoxenus nitidulus* (Nylander) in North West England." Bulletin of the Amateur Entomologists' Society **64**: 126-128.

Robinson, N. A. and E. J. H. Robinson (2008). "The population of the red wood ant *Formica rufa* L. (Hymenoptera: Formicidae) at Gait Barrows National Nature Reserve, Lancashire, England over the 20 year period 1986 – 2006: nest longevity, reproduction and the effect of management." British Journal of Entomology and Natural History **21**(4): 225-241.

Robinson, N. A. and E. J. H. Robinson (2013). "Myrmecophiles and other invertebrate nest associates of the red wood ant *Formica rufa* L. (Hymenoptera: Formicidae) in north-west England." The British Journal of Entomology and Natural History **26**: 67-88.

Rossetti, M. R., T. Tscharntke, R. Aguilar and P. Batáry (2017). "Responses of insect herbivores and herbivory to habitat fragmentation: a hierarchical meta‐analysis." Ecology Letters **20**(2): 264-272.

Sanllorente, O., R. Hammond, F. Ruano, L. Keller and A. Tinaut (2010). "Extreme population differentiation in a vulnerable slavemaking ant with a fragmented distribution." Conservation Genetics **11**: 1701-1710.

Sanver, D. and B. A. Hawkins (2000). "Galls as habitats: the inquiline communities of insect galls." Basic and Applied Ecology **1**(1): 3-11.

Scarparo, G., P. Rugman‐Jones, M. Gebiola, A. Di Giulio and J. Purcell (2021). "Social parasite distancing: RADseq reveals high inbreeding in the social parasite *Microdon myrmicae* but low philopatry for host ant nest." Ecological Entomology **46**(1): 89-99.

Schönrogge, K., M. G. Gardner, G. W. Elmes, E. K. V. Napper, D. J. Simcox, J. C. Wardlaw, J. Breen, B. Barr, J. J. Knapp and J. A. Pickett (2006). "Host propagation permits extreme local adaptation in a social parasite of ants." Ecology Letters **9**(9): 1032-1040.

Seifert, B. (2011). Rote Liste und Gesamtartenliste der Ameisen (Hymenoptera: Formicidae) Deutschlands. Rote Liste der gefährdeten Tiere, Pflanzen und Pilze Deutschlands. B. . In: Binot-Hafke, Becker, Gruttke, Haupt, Hofbauer, Ludwig, Matzke-Hajek & Strauch (Eds.), Rote Liste der gefährdeten Tiere, Pflanzen und Pilze Deutschlands.

Naturschutz und Biologische Vielfalt 70: 469-487., Naturschutz und Biologische Vielfalt. **70:** 469-487.

Seifert, B. and A. Hagman (2015). "Colonization overseas by long-range aerial drift in a Formicoxenine ant (Hymenoptera, Formicidae)." Entomologisk Tidskrift **136**(1-2): 5-15.

Smith, M. J., M. G. Betts, G. J. Forbes, D. G. Kehler, M. C. Bourgeois and S. P. Flemming (2011). "Independent effects of connectivity predict homing success by northern flying squirrel in a forest mosaic." Landscape Ecology **26**: 709-721.

Social-Insects-Specialist-Group (1996). *Formicoxenus nitidulus*. The IUCN Red List of Threatened Species 1996.

Sorvari, J. (2016). Wood ants: threats, conservation and management. Wood Ant Ecology and Conservation. J. A. Stockan and E. J. H. Robinson. Cambridge, Cambridge University Press**:** 264-286.

Stockan, J. A., E. J. H. Robinson and N. A. Littlewood (2017). "Surveys for *Formicoxenus nitidulus* (Nylander) (Hymenoptera: Formicidae) in northern Scotland." British Journal of Entomology and Natural History **30**(4): 245-246.

Sudd, J. H. (1983). "The distribution of foraging wood-ants (*Formica lugubris* Zett) in relation to the distribution of aphids." Insectes Sociaux **30**(3): 298-307.

Travis, J. M. and C. Dytham (1999). "Habitat persistence, habitat availability and the evolution of dispersal." Proceedings of the Royal Society of London. Series B: Biological Sciences **266**(1420): 723-728.

van Hengel, R. (2011). II. A Het leven van een dwerg tussen de reuzen *Formicoxenus nitidulus* (Nyl 1846). Forum Formicidarum.

Weir, B. S. and C. C. Cockerham (1984). "Estimating F-Statistics for the Analysis of Population Structure." Evolution **38**(6): 1358-1370.

Wilson, M. C., X.-Y. Chen, R. T. Corlett, R. K. Didham, P. Ding, R. D. Holt, M. Holyoak, G. Hu, A. C. Hughes and L. Jiang (2016). "Habitat fragmentation and biodiversity conservation: key findings and future challenges." Landscape Ecology **31**(2): 219-227.

Table 1. Summary of sites and sampling of *F. nitidulus* Shining Guest Ant (SGA). All SGAs were male unless indicated otherwise.

Site Name, woodland type, references for hosting *F. nitidulus*	**Host species present**	**Sampling details** Year of sampling, Number of host nests sampled, host nest species and number of SGA sampled per host nest	**Total Shining Guest Ants Collected**
1. Abernethy, Inverness-shire, native pine wood heath bordering native mixed (birch-pine) woodland	*Formica lugubris* *Formica aquilonia* (G. Tompkins pers. comm.)	2021: 4x *F. lugubris* nests: 2-3 SGA/nest 2020: 1x *F. lugubris* nest: 3 SGA	17
2. Feshiebridge, Inverness-shire Habitat mosaic: native (birch) woodland, (pine) plantation, acid grassland and dwarf shrub heath	*Formica lugubris* *Formica aquilonia* (Stockan *et al.* 2017)	2021: 5x *F. lugubris* nests: 3 SGA/nest 1x *F. aquilonia* nest: 4 SGA 2020: 1x *F. lugubris* nest: 4 SGA	23
3. Aviemore, Inverness-shire, native (birch) woodland bordered by development	*Formica aquilonia* (G. Tompkins pers. comm.)	2021: 3x *F. aquilonia* nest: 3 SGA 2020: 1x *F. aquilonia* nest: 3 SGA	12
4. Longshaw Estate, Derbyshire, mixed woodland pasture, primarily oak, birch and pine	*Formica lugubris* (Chen and Robinson 2015, Burns *et al.* 2020)	2021: 3x *F. lugubris* nests: 1-3 SGA/nest 2020: 4x *F. lugubris* nests: 1-3 SGA/nest, including 2 ergatoid females	15 (12 male, 2 female)
5. Cropton Forest, North York Moors plantation, mostly non-native conifer	*Formica lugubris* (Procter *et al.* 2015)	2021: 4x *F. lugubris* nests: 3-4 SGA/nest	13

6. Broxa, North York Moors, plantation, mostly non-native conifer	*Formica lugubris* (Procter *et al.* 2015)	2021: 4x *F. lugubris* nests: 4 SGA/nest	12
7. Gaitbarrows, Lancashire, mixed native deciduous woodland on limestone pavement	*Formica rufa* (Robinson and Robinson 2008, Robinson and Robinson 2013)	2021: 1x *F. rufa* nest: 3 SGA	3
Total			**95**

Table 2. Allele frequencies for the 8 microsatellite loci used in this study, and GenBank accession numbers.

Locus	**GenBank**	**Number of alleles**
GA 13	OQ384169	5
GA 14	OQ384170	6
GA 35	OQ384171	4
GA 45	OQ384172	4
GA 48	OQ384173	3
GA 53	OQ384174	3
GA 54	OQ384175	13
GA 58	OQ384176	6

Table 3. Analysis of molecular variation (AMOVA) results

Source of variation	Df	Sum Sq	Mean Sq	Variance component	Total variance	p-value
Between sites	6	59.8	9.96	0.70	22.7	0.001
Between nests within sites	24	59.0	2.46	0.05	1.7	0.485
Within host nests	46	107.4	2.97	2.33	75.6	0.001

Table 4. GenBank accession numbers for the six mitochondrial haplotypes and their distribution across the study sites in Scotland (SC) and England (EN).

	Hap1	Hap2	Hap3	Hap4	Hap5	Hap6
	OQ376286	OQ376285	OQ376283	OQ376282	OQ376281	OQ376284
Abernethy, SC	12	0	1	0	0	0
Feshiebridge, SC	21	0	0	0	0	1
Aviemore, SC	5	0	6	0	0	0
Longshaw, EN	6	6	0	0	0	0
Cropton Forest, EN	6	0	0	2	1	0
Broxa, EN	8	0	0	0	0	0
Gaitbarrows, EN	0	3	0	0	0	0

Table 5. Evidence for multiple matrilines present in a host nest, per site. *Multi-haplotype host nests* indicates the number of host nests in which multiple *F. nitidulus* mitochondrial haplotypes were present, as a proportion of the total host nests from which multiple *F. nitidulus* specimens were sequenced. *Three allele per locus host nests* indicates the number of host nests in which ≥ 3 alleles were present in at least one locus, as a proportion of the total host nests for which microsatellite data were available for ≥ 3 *F. nitidulus* specimens, per site. Overall *evidence for multiple matrilines* gives the number of host nests in which there is either mitochondrial or microsatellite (or both) evidence for multiple matrilines, out of the total for which at least one source of evidence was available.

Site	Multi-haplotype host nests	Three allele per locus host nests	Evidence for multiple matrilines
Abernethy	1/5	0/2	1/7
Feshiebridge	1/7	3/5	3/7
Aviemore	4/4	1/4	4/4
Longshaw	2/4	1/2	2/4
Cropton Forest	2/4	2/4	3/4
Broxa	0/2	1/2	1/3
Gaitbarrows	0/1	0/0	0/1
All sites	10/27 (37%)	8/19 (42%)	14/30 (47%)

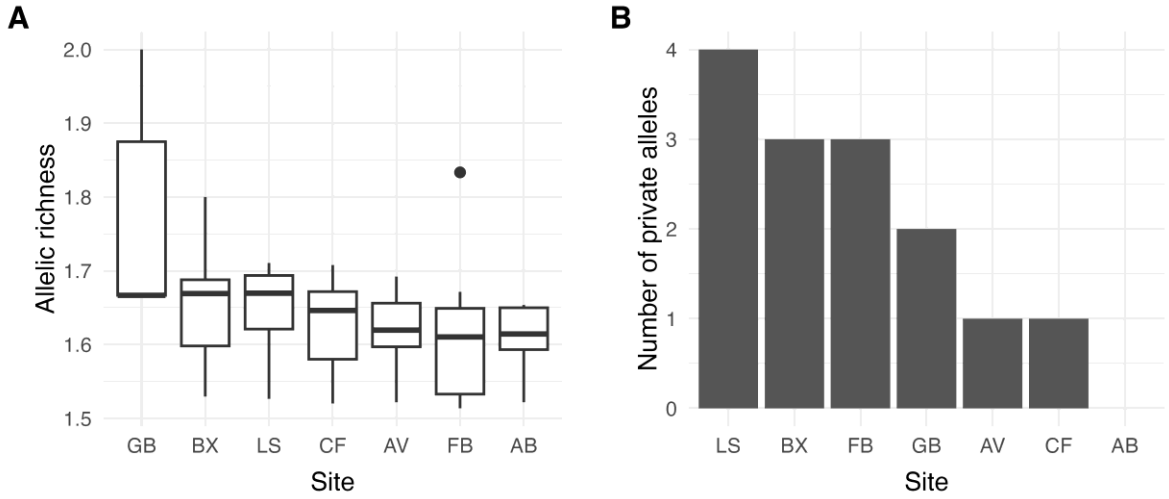


Figure 1: Distribution of alleles across sites. A) Allelic richness; B) Number of private alleles. Site codes and number of individuals used per site: Bx= Broxa, N=9; CF = Cropton Forest, N=13; LS = Longshaw, N=10; AV = Aviemore, N=12; FB = Feshiebridge, N=19; AB = Abernethy, N=12; GB = Gaitbarrows, N=2.

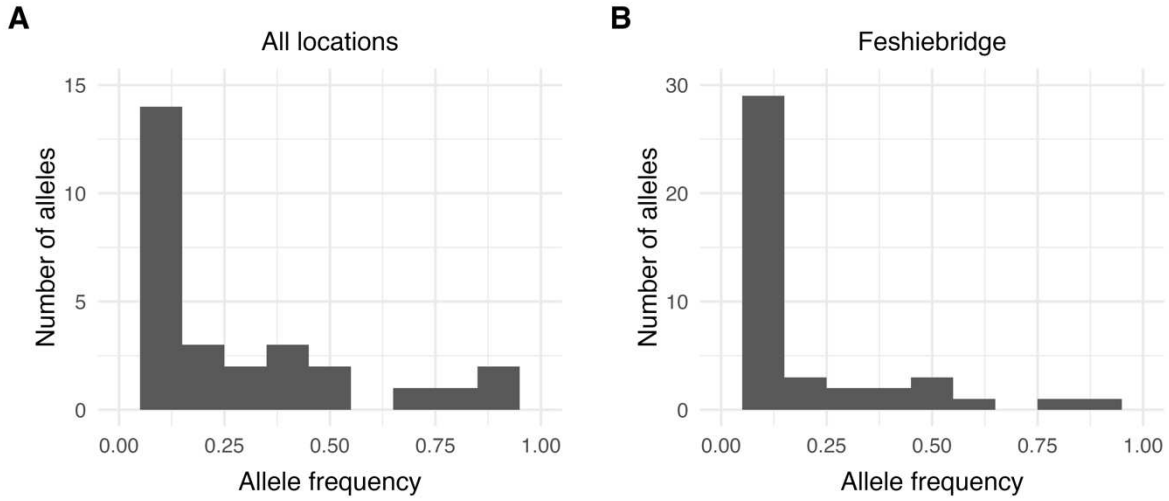


Figure 2. Allele frequency histograms for A) all sites combined; B) Feshiebridge, the site at which the highest number of individuals were sampled (N=19).

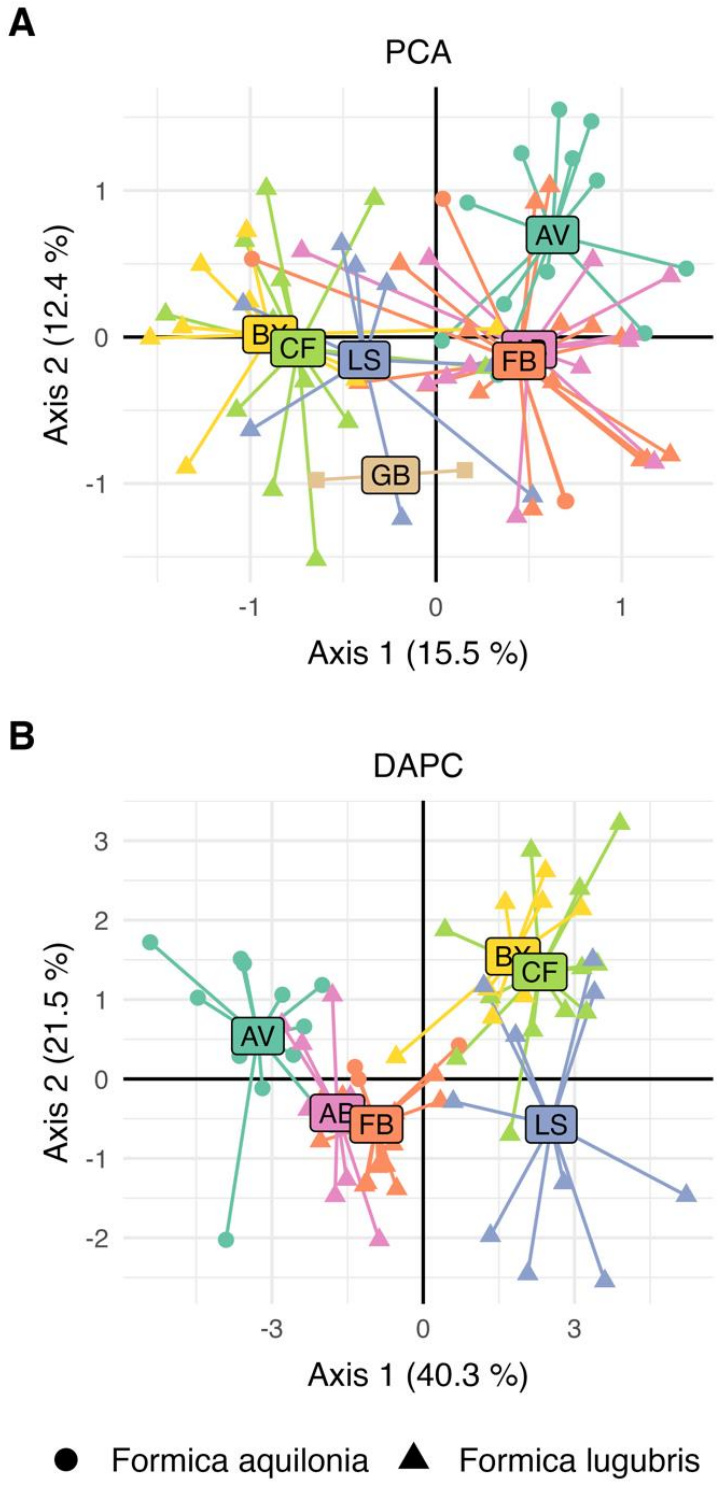


Figure 3. Genotypic groupings of individual ants by site and host species. Labels indicate centroids of each site's points. A) Principal Components Analysis (PCA). B) Discriminant Analysis of Principal Components (DAPC) reduces the dimensions of the PCA. BX= Broxa; CF = Cropton Forest; LS = Longshaw; AV = Aviemore, FB = Feshiebridge; AB = Abernethy. GB (Gaitbarrows) was very distant from the other six sites and is excluded from the DAPC plot for visual clarity.

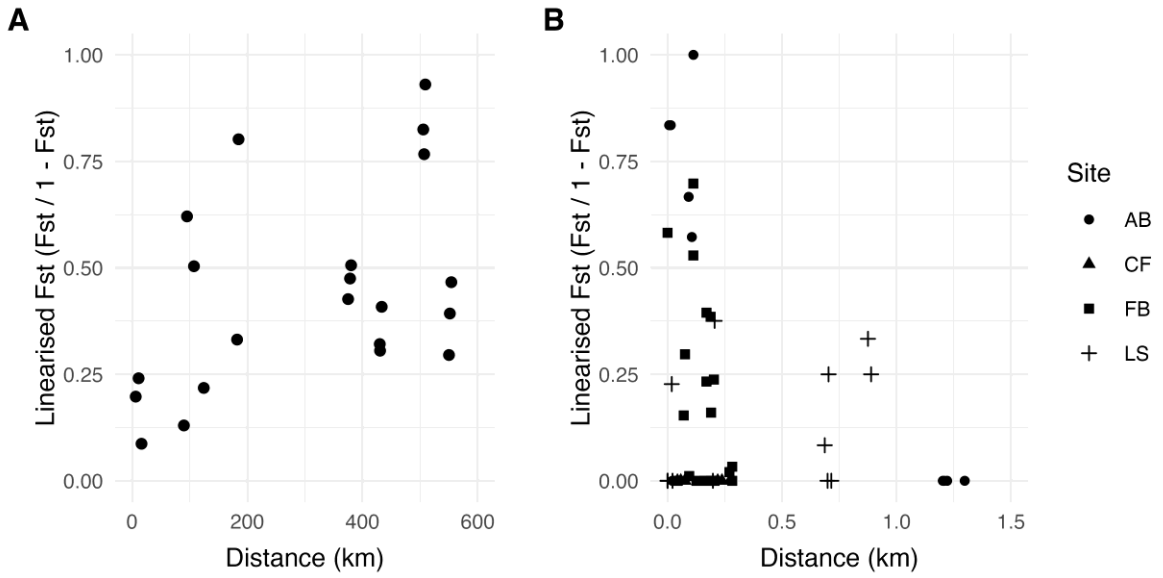


Figure 4. Relationship between spatial distance (km) and genetic distance (Linearised F_{ST}) in *Formicoxenus nitidulus* populations. a) Between site comparison: each point represents a pair of sites. Mantel test: r=0.46, n=21, p=0.01. b) Within-site comparison: each point represents a pair of host nests. The four sites with ≥5 sampled host nests are plotted. Mantel tests: AB: r=-0.62, n=8, p=0.88; FB: r=-0.32, n=21, p=0.86; LS: r=0.31, n=8, p=0.19.

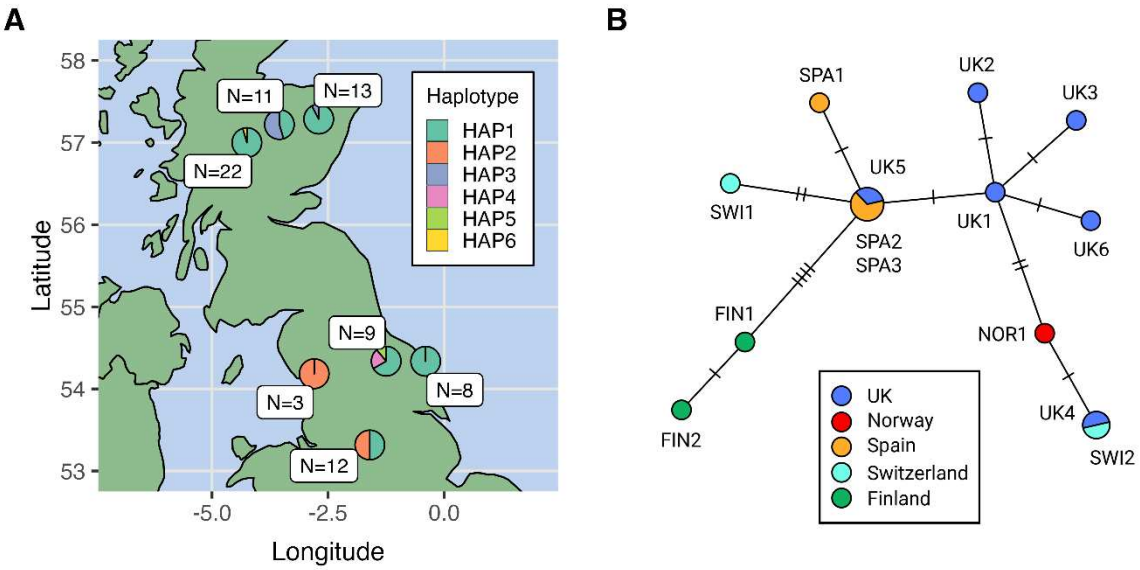


Figure 5. Mitochondrial haplotype data. A) Map of sites with frequency pie charts showing proportion of mitochondrial haplotypes at each site. N is number of individual ants used in each analysis. Scottish sites from west to east: Feshiebridge, Aviemore, Abernethy. Northern English sites from west to east: Gaitbarrows, Longshaw, Cropton Forest, Broxa. B) Haplotype network (TCS) of the 6 UK haplotypes and 8 European haplotypes from the BOLD database. NOR = Norway; SPA = Spain; SWI = Switzerland; FIN = Finland. Pie chart diameter is relative to the number of haplotypes at each node (n=1-3), as indicated by haplotype labels. Mappings of these labels to BOLD IDs are provided in Table S3.