This is a repository copy of *Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown.*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/117003/

Version: Accepted Version

Article:

Beresford, J., Elias, M., Pluckrose, L. et al. (4 more authors) (2017) Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown. Molecular Ecology, 26 (15). pp. 4013-4026. ISSN 0962-1083

https://doi.org/10.1111/mec.14183

This is the peer reviewed version of the following article: Beresford, J, Elias, M, Pluckrose, L, et al. Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown. Mol Ecol. 2017; 26: 4013–4026. which has been published in final form at https://doi.org/10.1111/mec.14183. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Accepted Article

DR. JONNA KULMUNI (Orcid ID : 0000-0002-8852-0793)

Article type : Original Article

Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown

Beresford J[1,2], Elias M[3], Pluckrose L[2], Sundström L[1,4], Butlin RK[2,5], Pamilo P[1], & Kulmuni J[1,2,4,6]

1) Centre of Excellence in Biological Interactions, Department of Biosciences, POB 65, FI-00014, University of Helsinki, Finland
2) Department of Animal and Plant Sciences, Alfred Denny Building, University of Sheffield, Western Bank, Sheffield, S10 2TN, UK
3) Institut de Systématique, Évolution, Biodiversité, ISYEB - UMR 7205 – CNRS MNHN UPMC EPHE, Muséum national d'Histoire naturelle, Sorbonne Universités, 57 rue Cuvier CP50 F-75005, Paris, France.
4) Tvärminne Zoological Station, University of Helsinki, J.A. Palménin tie 260, FI-10900 Hanko, Finland
5) Department of Marine Science, University of Gothenburg, S-405 30 Gothenburg, Sweden
6) Department of Biology and Biocenter Oulu, University of Oulu, 90014 Oulu, Finland

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.14183
This article is protected by copyright. All rights reserved.

Accepted Article

Keywords: Hybridization, Hymenoptera, Formica rufa wood ants, heterosis, hybrid breakdown

Corresponding author: Jonna Kulmuni, Department of Biosciences, University of Helsinki, 00014 Helsinki, Finland, jonna.kulmuni@helsinki.fi

Running title: Hybridization in *Formica* wood ants

Abstract

Hybridization and gene flow between diverging lineages is increasingly recognized as a common evolutionary process and its consequences can vary from hybrid breakdown to adaptive introgression. We have previously found a population of wood ant hybrids between *Formica aquilonia* and *F. polyctena* that shows antagonistic effects of hybridization: females with introgressed alleles show hybrid vigour, whereas males with the same alleles show hybrid breakdown. Here we investigate whether hybridization is a general phenomenon in this species pair, and analyze 647 worker samples from 16 localities in Finland using microsatellite markers and a 1200 bp mitochondrial sequence. Our results show that 27 sampled nests contained parental-like gene pools (six putative *F. polyctena* and 21 putative *F. aquilonia*) and all remaining nests (69), from nine localities, contained hybrids of varying degrees. Patterns of genetic variation suggest these hybrids arise from several hybridization events or, instead, have backcrossed to the parental gene pools to varying extents. In contrast to expectations, the mitochondrial haplotypes of the parental species were not randomly distributed among the hybrids. Instead, nests that were closer to parental-like *F. aquilonia* for nuclear markers preferentially had *F. polyctena*'s mitochondria and vice versa. This systematic pattern suggests there may be underlying selection favoring cytonuclear mismatch

This article is protected by copyright. All rights reserved.

and hybridization. We also found a new hybrid locality with strong genetic differences between the sexes similar to those predicted under antagonistic selection on male and female hybrids. Further studies are needed to determine the selective forces that act on male and female genomes in these newly discovered hybrids.

Introduction

Speciation is a continuum (Butlin et al. 2008) and a process dividing populations into reproductively isolated lineages, which usually happens over long evolutionary time scales. Gene flow during species divergence was once considered to be rare, especially in animals, but studies over the last decade suggest that hybridization may be more ubiquitous than we ever thought (Mallet 2007; Heliconius Genome Consortium 2012; Trigo et al. 2013; Abbott et al. 2013). As suggested by Abbott et al. (2013), "If hybridization is defined as reproduction between members of genetically distinct populations (Barton & Hewitt, 1985), producing offspring of mixed ancestry, then it occurs in almost all proposed processes of speciation". Hybridization, and the resulting gene flow, bring together genetic material from two independently evolving genomes. For this reason hybrids can reveal incompatibilities between the parental genomes resulting in reduced hybrid fitness or hybrid breakdown in later generations, but the new allelic combinations could also be adaptive.

The impacts of hybridization can range from hybrid inviability and / or sterility (Bateson 1909; Dobzhansky 1936; Muller 1942; Barreto et al. 2015) to reverse speciation (Rudman & Schluter 2016), through adaptive introgression (Song et al. 2011; Heliconius Genome Consortium 2012) and even the formation of entirely new species through hybrid speciation (Rieseberg 1997; Schwarz 2005; Mavarez et al. 2006; Mavarez & Linares 2008; Abbott et al.

This article is protected by copyright. All rights reserved.

Accepted Article

2013; Capblancq et al. 2015). For example, closely related *Heliconius* butterfly species have exchanged protective colour-pattern genes promiscuously, implying that hybridization has an important role in the evolution of major adaptive traits within this group (Heliconius genome consortium 2012). Similarly, in a laboratory study on *Cobitis* fish, hybridization between diverging populations triggered clonal reproduction in a small proportion of the crosses (Choleva et al. 2012). Thus, the outcomes of hybridization likely depend on the stage of the speciation continuum (Seehausen et al. 2014), but we are still poorly positioned to explain the conditions under which one outcome is more likely than another. Importantly, the consequences of hybridization may vary across the genome, even within a single hybridization event, in such a way that gene flow between diverging populations has both deleterious and adaptive consequences (Song et al 2011; Janoušek et al. 2012; Martin et al. 2013; Kulmuni & Pamilo 2014). With the available genome-scans in many model systems, it now becomes possible to characterize regions of reduced gene flow and adaptive introgression simultaneously (Payseur & Rieseberg 2016).

Hymenopteran insects offer a promising model system to study the role of hybridization in speciation due to being haplodiploid. Females develop from fertilized eggs and are diploid, but males are haploid and develop from unfertilized eggs. As a result the outcomes of hybridization among divergent populations of Hymenoptera are often different compared to diplodiploid organisms (Umphrey 2006; Feldhaar et al. 2008). Importantly, natural selection operates differently in haploid and diploid genomes (Avery 1984; Werren 1993; Muirhead & Presgraves 2016). All recessive alleles are exposed to selection in haploids, whereas the same alleles are masked in diploids. In the context of hybridization, this means that both beneficial and deleterious recessive allelic combinations are selected more strongly in haploid males than diploid females. In accordance with Haldane's rule (Haldane 1922), hybrid breakdown

This article is protected by copyright. All rights reserved.

should affect the hemizygous sex (i.e. haploid males) disproportionately, and this has been confirmed by recent studies in haplodiploids (Koevoets & Beukeboom 2009; Kulmuni et al. 2010; Beukeboom et al. 2015). Hybrid breakdown may be due to failing interactions between two or more nuclear loci (nucleo-nuclear) or between cytoplasmic and nuclear loci (cytonuclear). However, both types of incompatibilities should result in differential selection between males and females because of haplodiploidy.

The red wood ants of the *Formica rufa* group (or *Formica s. str.* Stockan et al. 2016) are dominant species in boreal coniferous forests. Genetic divergence within the *Formica rufa* group is very low (< 2% in mitochondria), and the species are estimated to have diverged only in the last 500ky, possibly while residing in different glacial refugia (Goropashnaya et al. 2004a). The species *Formica aquilonia* and *F. lugubris* have boreo-alpine distributions, whereas the distributions of *F. polyctena, F. pratensis* and *F. rufa* are more southern or at lower altitudes (Stockan et al. 2016). The aforementioned five species of *F. rufa* group occur sympatrically in Southern Finland, and they have been shown to hybridize at least occasionally in other parts of their range (Seifert et al. 2010; Goropashnaya & Seifert 2012). Societies of *F. polyctena* and *F. aquilonia* have many, up to hundreds of reproductive queens in a single nest (polygyny) and neighboring nest mounds are normally connected to form large colonial networks (Rosengren & Pamilo 1983).

We previously described a population of wood ant hybrids between *Formica polyctena* and *F. aquilonia* (Kulmuni et al. 2010; 2014) in Southern Finland at one locality called Långholmen. Hybridization has resulted in two genetic groups (previously named R and W) showing signs of past introgression with each other but very little contemporary gene flow, despite temporally overlapping breeding seasons. All these hybrid ants share *F. polyctena*-

This article is protected by copyright. All rights reserved.

like mitochondrial DNA. However, we have earlier suggested that one of the groups in this Långholmen population is more similar to *F. polyctena* and the other one closer to *F. aquilonia*, but we lacked the data to test this hypothesis (Kulmuni et al. 2010; Kulmuni & Pamilo 2014). Genomic fragments introgressed from the other group are maintained at relatively constant frequencies across generations. In each generation, however, the males with introgressed alleles die during development, likely because males are haploid and those genomic regions that are incompatible between the two original species are exposed to selection causing inviability (Kulmuni & Pamilo 2014). The incompatibilities are likely to be nucleo-nuclear since the two genetic groups in the population have identical mitochondrial haplotypes, yet different sets of nuclear alleles are selected against in the haploid males of the two genetic groups. Conversely, in heterozygous females the incompatibilities are recessive and are not selected against, because they are masked by the other allele at the locus (Kulmuni et al. 2010). In fact, the results suggest that hybrid females experience a fitness advantage (Kulmuni & Pamilo 2014). In the absence of such an advantage, the introgressed fragments should gradually disappear owing to recombination and selection against these fragments in haploids. Thus opposing selective forces operating in males and females result in large-scale genomic differences between the sexes within each genetic group. These results raise questions on the origin of these hybrid ants, and on the occurrence and outcome of possible hybridization in other localities where the distribution of the parental species overlap.

Here we investigate 16 new localities close to the above described hybrid population in Långholmen to examine the distribution of parental species and ask if there are signs of other hybrid populations. We use nuclear microsatellite and mitochondrial DNA markers to examine genetic structure and discordances between the two marker types that would be

This article is protected by copyright. All rights reserved.

Accepted Article

indicative of hybridization. In particular, we reason that hybridization could result in the following specific genetic patterns; 1) nuclear genotype of one species, but mitochondrial haplotype of another, 2) individuals with admixed nuclear genotypes. Mitochondrial haplotypes from two species present in a single nest and/or locality would suggest opportunities for hybridization, but does not in itself demonstrate the presence of hybrids. If new localities with hybrids are found, we ask if there are genetic differences between males and females, similar to those observed in the previously studied Långholmen hybrid population (Kulmuni et al. 2010). In particular, we search for significant allele frequency differences between the sexes and for alleles that do not occur in adult males but are present as heterozygotes in the females. Such patterns would be suggestive of recessive alleles that are incompatible between the two species and selected against in the haploid males.

Material and Methods

Genotyping

To map putative parental and hybrid populations of *F. aquilonia* and *F. polyctena* we sampled altogether 647 worker ants from 96 nests at 16 locations (Supplementary Table 1) within an area of approximately 3000 square kilometres in southern Finland (Figure 1) in the years 2001-2010. The previously analyzed Långholmen data (Kulmuni et al. 2010) were added as the 17th location in order to allow comparisons. We failed to record coordinates for one of the sampling locations (named Unknown; 1 nest sampled), but it represents an independent sample collected from nature. Some of the locations were within putative dispersal distance of each other (i.e. 500 meters apart but separated by the sea) or close to the Långholmen hybrid population studied earlier (Kulmuni *et al.* 2010; Kulmuni & Pamilo 2014). The sampling localities were here considered as separate populations that may be connected by some gene flow. As the nests have many (up to hundreds) reproductive queens

This article is protected by copyright. All rights reserved.

Accepted Article

that can be mated with several males (Pamilo 1993), the individuals sampled from the same nest are not necessarily relatives, but can be considered as a small population sample. The ants were examined under a microscope to determine the species morphologically according to the criteria described by Collingwood (1979) and Stockan et al. (2016). In our analysis, we included samples that were considered to be *F. aquilonia*- or *F. polyctena*-like. Based on the morphological criteria (hairs on the back of the head, and on the thorax), none of the nests could be unequivocally classified as either pure *F. aquilonia* or pure *F. polyctena*: The ants had more hairs than samples of *F. polyctena* from Central Europe but fewer than *F. aquilonia* from other parts of Finland. The samples formed a continuum between the two species, one end being close to *F. polyctena* and the other end close to *F. aquilonia*. However, without specialist techniques it is not possible to analyze morphology quantitatively. Taxonomic uncertainty concerning these two species in Finland has also been noted e.g. by Betrem (1964) and Sorvari (2006). Because of these difficulties with standard morphological identification, we decided to proceed with the genetic analysis without a priori classification of samples into putative parental species or hybrids.

DNA was extracted with the DNAeasy Tissue Kit (Qiagen) using the manufacturer's protocol designed for insects. Four to twelve individuals per nest were genotyped with microsatellite markers and two or four of these were analyzed for a 1200 bp fragment of the mitochondrial Cytochrome b gene. PCR amplification of a 1500 bp mitochondrial DNA fragment was performed with primers Cyt-Fe-F (Liautard & Keller 2001) and Trs (Jermiin & Crozier 1994) and amplicons were sequenced with these primers and primers CB1, CB11 and CT2F, as described previously (Goropashnaya et al. 2004b). The reference sequences from Genbank were missing the first 300 bp of the mitochondrial fragment we sequenced. Thus we use the overlapping 1200 bp in all of our analyses.

This article is protected by copyright. All rights reserved.

All individuals were genotyped at 9 unlinked (Kulmuni & Pamilo 2014) microsatellite loci (FE7, FE13, FE19, FE17 (Gyllenstrand et al. 2004), FL29 (Chapuisat 1996), FY12, FY13, FY15, and FY3 (Hasegawa & Imai 2004)). These loci were chosen because they contained alleles that were diagnostic for each of the two genetic groups present in the Långholmen hybrid population and alleles that had introgressed from one group to the females, but not to males of the other group (Kulmuni et al. 2010; Kulmuni & Pamilo 2014). By using the same loci our aim was to investigate whether similar hybrids or parental-like gene pools are present in other localities. Genotypes were assayed by polymerase chain reaction with fluorescent labeling using Peltier Thermal Cycler-200-PCR equipment (MJ Research) and the conditions described in Kulmuni & Pamilo (2014). Genotypes were resolved by capillary electrophoresis (3730 DNA Analyzer, Applied Biosystems) using 500 LIZ size standard (Applied Biosystems) and scored with Genemapper v.4.0 (Applied Biosystems).

Mitochondrial haplotype structure

In total, we sequenced the 1200 bp mitochondrial DNA fragment from 265 individuals. We first assigned the individuals to their putative parental groups (*F. aquilonia* or *F. polyctena*) based on their mitochondrial haplotype. For this purpose we downloaded the available unique mitochondrial cyt-b sequences of *F. rufa* group species from GenBank (Goropashnaya et al. 2004a; Goropashnaya et al. 2012) to be used as references and aligned them with the sequences produced in this study using Muscle (Edgar 2004) with the default parameters. We used this alignment to build a TCS haplotype network (Clement et al. 2002) with PopArt (v1.7, Leigh 2016). We drew the final haplotype network using only unique sequences. We compared the overall structure of the haplotype network to a phylogenetic tree of unique sequences with the program MrBayes (v3.2.6, Ronquist et al. 2011). MrBayes was run twice for a comparison using the standard settings (lset nst=6 rates=invgamma); the MCMC

This article is protected by copyright. All rights reserved.

Accepted Article

modeling chains ran for 1 000 000 generations sampling every 1 000th generation, the output of coalescence was assessed using Tracer (Beast v1.8.3, Drummond et al. 2012) and the burnin was set to 25 % (250000 generations). From the MrBayes runs, a consensus tree was edited in FigTree (v1.4.2, Rambaut 2014). We also presented the mitochondrial haplotypes on a map of sampling localities to search for an obvious clinal pattern or a possible hybrid zone, and to study whether haplotypes of different species occur in the same localities.

Genetic structure and admixture in the nuclear markers

To study the genetic structure based on nuclear markers and to identify putative parental gene pools, we analyzed the microsatellite markers with Principal Coordinates Analysis in GenAlEx (Peakall & Smouse 2006, 2012). For this we calculated the pairwise genetic distance between each nest, pooling individual worker genotypes within a nest and using the codominant genotypic distance option. To further study admixture and parental gene pools at the individual level, and to search for discordances between nuclear and mitochondrial DNA, we analyzed the nuclear genotypes with Structure 2.3.4 (Pritchard et al. 2000). To investigate the most likely number of groups in our data we ran Structure with K=1 to K=10, with the model allowing for admixture. No information on sampling locations was used and the model with independent allele frequencies was chosen. For each K value, we ran 5 iterations with a burn-in of 200 000 and 1 000 000 MCMC replicates after burnin. We used Structure Harvester (Earl & vonHoldt 2012) to calculate DeltaK statistics (Evanno et al. 2005), which detect the uppermost hierarchical level of genetic structure in the data based on the rate of change in the log-likelihood of the data between successive K values. We used StructurePlot (Ramasamy et al. 2014) to draw the plots with the most likely number of genetic groups.

This article is protected by copyright. All rights reserved.

Accepted Article

Finally, we tested for discordances between the mitochondrial and nuclear structure, indicative of hybridization, by comparing nuclear and mitochondrial groups. We used only the individuals for which we had genotyped both nuclear and mitochondrial markers and studied whether the mitochondrial haplotypes were randomly distributed in the nuclear groups by a contingency test using the Chi-Square statistic. We also studied whether the nests that were intermediate based on nuclear DNA (i.e. hybrids) had a non-random distribution of mitochondrial haplotypes from the two parental species. For this we excluded nests that represented putative parental gene pools based on Principal Coordinates and Structure analysis. We tested if the mean of the 1st principal coordinate (which represents the relative proportions of parental-like nuclear gene pools) differed between hybrid nests that had *F. aquilonia* mitochondrial haplotype and those that had *F. polyctena* mitochondrial haplotype by using a two-tailed t-test.

We included a total of 200 males and females from our earlier study of the Långholmen population (Kulmuni and Pamilo 2014) in the above analyses of nuclear genetic markers in order to compare the new localities with our previous hybrid population. In doing this, we aimed to test whether the two genetic groups detected among the Långholmen hybrids were situated in between the putative parental species as we would expect, with one group being genetically closer to parental-like *F. aquilonia* and the other closer to parental-like *F. polyctena*.

Testing for genetic differentiation between sexes in two new hybrid localities

Our previous hybrid population showed systematic genetic differences between males and females because hybrid males suffered hybrid breakdown but females enjoyed hybrid vigor (Kulmuni et al. 2010; Kulmuni & Pamilo 2014). This resulted in some alleles (the

Accepted Article

This article is protected by copyright. All rights reserved.

introgressed ones) being absent from males and appearing only as heterozygotes in the females. For comparison, we tested whether this pattern of differentiation was present in two of the hybrid localities that were discovered in this study (Halsholmen nests 1-4 and Tvärminne road nests 1-3) that had both sexes from multiple nests for comparison. First, we confirmed that there was no within-locality genetic structure that could hamper our analysis of genetic differentiation, by analyzing 10 worker genotypes per nest with PCoA in GenAlEx (Peakall & Smouse 2006, 2012). We also confirmed that the nests were polygynous, i.e. that our sampling represents offspring from several queens, by calculating the mean relatedness among workers (Queller & Goodnight 1989) using GenAlEx. Next, we genotyped a total of 18 males and 20 young queens from four nests in the locality Halsholmen. We included only nests from which we could sample both sexes. Second, we genotyped 10 males and 10 workers from each of the three nests in the locality Tvärminne Road. Workers could be used as representatives of females as new queens and workers did not differ genotypically in our original hybrid population (Kulmuni et al. 2010). In both hybrid localities we searched for significant allele frequency differences between males and females using an exact test in Genepop (Raymond & Rousset 1995) and especially searched for alleles that were lacking from males and present only as heterozygotes in the females. These patterns would be suggestive of recessive incompatibility alleles between the two species, which are selected against in the haploid males.

Results

Mitochondrial haplotype structure

Our samples contained 17 unique mitochondrial haplotypes. The mitochondrial haplotype network concurs with the earlier finding that the major division within the mound building red wood ants is between the lineage including the haplotypes of *F. rufa* and *F. polyctena* and

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

that including *F. aquilonia*, *F. lugubris* and *F. paralugubris* (Figure 2) (Goropashnaya et al. 2004a). The Bayesian tree also supports the existence of these two diverged haplotype groups with 99 % Bayesian posterior probability, with little resolution to tell apart species within these groups (Supplementary figure 1). When all the observed SNPs were taken into account, the two haplotype groups differed by a minimum of 22 positions.

As expected our 265 sampled mitochondrial haplotypes from the 16 study localities clustered with the haplotypes of *F. polyctena* and *F. aquilonia* from GenBank. One *F. aquilonia* haplotype was shared with *F. paralugubris*, but the nearest known location of the latter species is in Central Europe, so it is unlikely that it would occur in Finland. The two species *F. polyctena* and *F. rufa* share several mitochondrial haplotypes and cannot be distinguished based on the 1200 bp fragment sequenced here. Although we cannot exclude the possibility of some admixture with *F. rufa*, our samples were morphologically consistent with, and most likely represent *F. polyctena*. Thus, based on the 1200 bp mitochondrial fragment we could reliably separate the two parental species haplotypes: *F. aquilonia* (158 sampled haplotypes), and *F. polyctena* (107 sampled haplotypes).

Five of the 16 localities sampled for this study had only haplotypes representing *F. aquilonia*, and four localities only haplotypes representing *F. polyctena* (Figure 1). In addition, the previously studied Långholmen hybrid population contained only *F. polyctena* haplotype (Kulmuni et al. 2010). The remaining seven localities contained haplotypes from both *F. aquilonia* and *F. polyctena* (Figure 1). Within these seven localities, three nests (from the localities Tenala and Pikkala) contained mitochondrial haplotypes from both species, which indicates mixed-species nests or hybrids. Localities harbored from one to five, and nests from

This article is protected by copyright. All rights reserved.

one to two unique mitochondrial haplotypes, respectively. There was no clear geographical pattern in the occurrence of haplotypes of *F. aquilonia* and *F. polyctena* (Figure 1).

Genetic structure and hybridization

The Principal Coordinates Analysis (PCoA) was based on pairwise genetic distances between nests using DNA-microsatellite data. In the PCoA, the first axis explained 18.6 % of variation, and the combined evidence indicates that it captured the genetic distance between the gene pools of the two putative parental species (Figure 3, Structure results, see below). The suggested parental gene pools segregated clearly along the first axis, leaving a cluster of putative hybrids in the middle (Figure 3). This interpretation is corroborated by the distribution of mitochondrial haplotypes, as only *F. aquilonia* haplotypes were present in the samples at the left end of the first axis, and only *F. polyctena* haplotypes at the right end, whereas the nests in between (i.e., in the middle of the first coordinate axis) had either *F. aquilonia* or *F. polyctena* haplotype. The second coordinate axis explained 15.6 % of the variation, and separated the locality Pikkala from the rest of the samples. This separation was apparently affected by larger number of nest samples from Pikkala (34 nests sampled in Pikkala, compared to 1-10 nests in the other localities), because it disappeared when PCoA was performed with a reduced set of Pikkala nests (Supplementary Figure 2).

To assess whether the cluster of putative hybrid nests indicated by the PCoA indeed contained hybrid individuals, rather than being mixed-species nests, we analyzed the population structure at the individual level. We ran Structure on the microsatellite data using K-values from 1 to 10. The deltaK statistics suggested that the uppermost hierarchical level of genetic structure in the nuclear data was three genetic groups (deltaK=31.1, Supplementary Figure 3), with a strong signal of admixture in the majority of the sampled

This article is protected by copyright. All rights reserved.

Accepted Article

localities. We thus present the results for K=3 after dividing the data into the three genetic groups (Figure 4); the overall results concerning admixture remained the same even when the data were divided into two, or more than three genetic groups. With higher K-values the groups divided into subgroups suggesting the existence of hierarchical genetic structure in the data, rather than inconsistent group assignment.

When combining the information from nuclear grouping with the mitochondrial haplotypes, the three genetic groups formed by Structure corresponded to the following (Table 1): putative parental-like *F. aquilonia* (group 1), putative parental-like *F. polyctena* with introgression (group 2) and the previously studied hybrids of type W (Kulmuni et al. 2010; Kulmuni & Pamilo 2014) from Långholmen (group 3). In more detail, group 1 consisted of individuals from four localities (Pusula, Grabbskog, Solböle and Siuntio1). This group was significantly associated with the *F. aquilonia* mitochondrial DNA haplotype as all the individuals with over 90% of the nuclear genome from group 1 (35 individuals from 21 nests and 4 localities), had *F. aquilonia* mitochondrial haplotype ($X^2 = 27.3$, df = 1, $P < 0.0001$). Group 1 was consistent with the Principal Coordinates analysis as nests belonging to this group had negative scores and situate to the left end of the first principal coordinate axis (Figure 3), forming a group separated from the rest of the samples. Because these samples contained only *F. aquilonia* mitochondrial haplotype and form their own group both in the PCoA and in Structure analysis, these 21 nests likely represent the least admixed *F. aquilonia* gene pool within our sampling with little or no introgression from *F. polyctena* (Table 1).

Group 2 (individuals having over 90% of their nuclear genome assigned into group 2) was most numerous and consisted of individuals from all localities except Pusula, Siuntio 1 and Tvärminne Station. This group showed no significant association with mitochondrial

This article is protected by copyright. All rights reserved.

haplotype (Figure 4) as 91 individuals carried *F. aquilonia* mitochondrial DNA and 63 individuals *F. polyctena* mitochondrial DNA ($X^2 = 0.04$, df = 1, $P > 0.1$). However, all the individuals in this group were from nests with either low absolute or positive values on the 1st axis of PCoA (Figure 3). Thus, this nuclear group probably represents different degrees of admixture and putative *F. polyctena*. The least admixed *F. polyctena* individuals sampled in this study were present in six nests (Fiskars 1 and 2, Grabbskog 7, Joskär 1 and 3, Unknown), a result that is supported by three pieces of evidence. First, all individuals that were sampled for mitochondria from these nests had *F. polyctena* haplotypes. Second, the majority of the individuals sampled for nuclear markers from these nests (44 out of the 49 individuals) had ≥98 % of nuclear ancestry in group 2 indicating little admixture. Third, these nests had high positive scores and situate to the right end of the first coordinate axis in the PCoA (Figure 3). Thus, the PCoA, mitochondrial haplotypes and Structure results are consistent and suggest six nests in our data represent *F. polyctena*-like gene pool (Table 1). Nest Tenala had 0.98 – 0.97 nuclear ancestry in this group, but had mitochondrial haplotypes of both *F. aquilonia* and *F. polyctena*.

Group 3 consisted of individuals from the W group of the Långholmen hybrid population (Kulmuni et al. 2010). In the PCoA these samples were situated in between the putative *F. aquilonia* and *F. polyctena* gene pools (Figure 3), consistent with our previous results (Kulmuni et al. 2010) that this group is hybrid.

Patterns of admixture

The genetic data shows signatures of admixture in several locations (Figure 4). Based on the microsatellite loci used here, 195 of the 647 individuals genotyped had > 5 % of admixture between groups defined by Structure. A subset of these individuals (n=117) was sequenced in

This article is protected by copyright. All rights reserved.

mitochondrial DNA and found to harbor 10 unique haplotypes. However, only 76 individuals had more than 20 % admixture. These individuals were found from the localities Pikkala (46 out of 136 individuals sampled from this locality), Siuntio 2 (3 out of 30), Tvärminne Road (1 out of 30), Solböle (1 out of 30), Pusula (1 out of 44), Sammatti (4 out of 10), Svanvik (5 out of 56), Halsholmen (7 out of 50), Tvärminne Station (7 out of 10) and Hanko (1 out of 98). Fourtyone of these individuals were sequenced in mitochondrial DNA and they harbor seven unique haplotypes, including both *F. aquilonia* and *F. polyctena* haplotypes. Finally, 17 of these individuals, from five localities (Pikkala, Tvärminne Station, Halsholmen, Siuntio2 and Svanvik), show admixture proportions between 40 and 60 %.

Out of the 76 individuals identified as > 20 % admixed by Structure, all except one (from Pusula) came from the nests that were intermediate also in the PCoA. Altogether, the nests that fell in between putative parental gene pools in the PCoA formed a heterogeneous group (including Structure group 3) with varying degrees of admixture and either *F. polyctena* or *F. aquilonia* mitochondria (Table 1). In fact, the mitochondrial haplotypes of *F. polyctena* and *F. aquilonia* were not randomly distributed in these hybrids based on the PCoA. Instead, Figure 3 shows that nests with nuclear genomes closer to putative parental *F. aquilonia*, preferentially had *F. polyctena* mitochondrial haplotypes and vice versa. There was a significant difference along the 1st principal coordinate axis between the hybrid nests that had *F. aquilonia* mitochondrial haplotypes (mean of 1st principal coordinate 0.143, SE = 0.014) and those that had *F. polyctena* mitochondrial haplotypes (mean = 0.036, SE = 0.004) (t-test, df = 70, P = 0.0006). A similar pattern of significant cytonuclear mismatch can be seen in Supplementary Figure 2, which contains all the other samples but has a reduced number of nests from the Pikkala locality (t-test, df = 48, P = 0.015). Finally, significant cytonuclear mismatch is also found within a single locality with large sample size, that of Pikkala (t-test, df = 30, P = 0.002).

This article is protected by copyright. All rights reserved.

The samples from the previously studied hybrid population in Långholmen were included in the analysis and they confirmed the patterns described earlier (Kulmuni & Pamilo 2014); the population consisted of two genetic groups (Figure 4) that were both of hybrid origin, but they had only *F. polyctena* mitochondrial haplotype. As expected, both groups fell between the putative parental gene pools in the PCoA (Figure 3). However, one of the groups (previously called W, Kulmuni et al. 2010) is genetically closer to *F. aquilonia* and the other one closer to *F. polyctena* (previously called R, Kulmuni et al. 2010). In the Structure analysis the W individuals form Structure group 3 and R individuals are admixed between Structure groups 3 and 2.

Based on the Structure results, our new data had only two nests (Tvärminne Station 1 and Halsholmen 5), which had nuclear composition similar to the R group Långholmen hybrids and these were situated only some hundreds of meters from the Långholmen population (Figure 4). Yet, one of these new nests had an *F. aquilonia* mitochondrial haplotype indicating that these hybrids did not arise from the same maternal lineage as our previous hybrid population since that had an *F. polyctena* mitochondrial haplotype. None of the newly sampled localities was the same as the W group present in the Långholmen hybrids. Based on the results from Structure the rest of the hybrids discovered in this study (admixed mainly between Structure groups 1 and 2) are genetically different from our previous hybrid population (either group 3 or admixed between groups 2 and 3), which may indicate varying degrees of backcrossing.

This article is protected by copyright. All rights reserved.

Genetic differentiation between males and females in two newly discovered hybrid localities

Genetic differences between sexes were studied in two localities, Halsholmen (four nests) and Tvärminne Road (three nests). Workers showed average admixture of 3 % between the genetic groups defined by Structure in locality Halsholmen and 5% in locality Tvärminne. In both of these localities the majority of nuclear ancestry was from group 2, with admixture from groups 1 and 3. Nests from both of these localities fell between putative parental gene pools in the PCoA (Figure 3), suggesting they were hybrids, and both had *F. aquilonia* mitochondrial haplotypes. Neither of the localities (Supplementary Figure 4 and 5) showed within-locality substructure: Instead the worker genotypes from different nests were not clustered by the nest but were overlapping and widely scattered in the PCoA figures (Supplementary Figure 4 and 5). Within-nest relatedness in both localities was low, indicating multiple egg-laying queens per nest, and varied between - 0.098 and 0.007 in Tvärminne road and between - 0.058 and 0.083 on Halsholmen (Supplementary Table 2). Six diploid males were found on Halsholmen, and these were removed from further analyses. We found no significant genetic differences in microsatellite allele frequencies between males and females on Halsholmen (Table 2), but a significant differentiation between the sexes was found in locality Tvärminne road (5 loci, each $P < 0.01$, four of which are significant after Bonferroni correction). Altogether 8 alleles at 5 loci were present as heterozygotes in 24-54 % of the females, but were never found as homozygotes. These same alleles were absent from the males in this locality (Table 2). Seven of these female-specific alleles were found from all three nests.

Accepted Article

This article is protected by copyright. All rights reserved.

Discussion

Our previous studies have revealed an unusual pattern of hybridization in the Långholmen population in Southern Finland (Kulmuni et al. 2010; Kulmuni & Pamilo 2014). Males and females in this population show large-scale differentiation in their nuclear genomes, which is caused by antagonistic natural selection favoring introgressed alleles in females but eliminating them in males (Kulmuni & Pamilo 2014). In this study we sampled 16 new localities from nearby areas and compared them with the hybrids studied earlier, to estimate the prevalence of hybridization and determine whether the parental species exist in the area. We found that six of the studied localities had mitochondrial haplotypes from both parental species, indicating that opportunities for hybridization exist. With nuclear markers we were able to differentiate putative parental-like gene pools, but over half of the 647 sampled worker individuals are genetically intermediate. In contrast to expectations, these putative hybrids showed a systematic mismatch between nuclear and mitochondrial DNA, raising the question of which evolutionary forces drive this pattern. One newly discovered locality (Tvärminne Road) containing admixed individuals (out of two investigated) revealed patterns of genetic differentiation between the sexes similar to our previous Långholmen hybrid population. Males lacked alleles that occurred at high frequencies in female heterozygotes, a pattern consistent with hybrid breakdown in haploid males. Below we discuss these results in the context of hybridization and speciation.

Populations in Southern Finland can be characterized as a hybrid swarm between* F. aquilonia *and* F. polyctena *wood ants

Out of the 96 new nests studied here, 6 nests represent *F. polyctena*-like gene pool and 21 *F. aquilonia*-like gene pool. However, because we did not sample a wider geographical area and we used a limited number of markers, we cannot rule out the possibility that even these nests

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

contain some degree of introgression. The remaining 69 nests fall in between the putative parents in the PCoA. These represent likely hybrids with varying degrees of admixture with both *F. aquilonia* and *F. polyctena* mitochondrial haplotype, suggesting the hybrids do not form a homogenous group. Multiple matrilines among the hybrids indicate several origins for these hybrids. Only 20 individuals (from 4 localities) out of 647 showed admixture proportions between 40 and 60 % based on Structure, suggesting that there are few, if any, first generation hybrids. These individuals came from nests that harboured altogether 7 unique mitochondrial haplotypes, again suggesting multiple hybridization events. Together these results indicate introgression has been happening for many generations and is continuing in Southern Finland. In conclusion, the sampled area in Southern Finland can be characterized as a hybrid swarm.

Previous studies on *Formica* wood ants have documented some cases of hybridization between the close relatives *F. polyctena* and *F. rufa* in Central Europe (Seifert et al. 2010), but our data suggest extensive hybridization between the more differentiated clades within this group; *F. aquilonia* and *F. polyctena* are not closest relatives. In the past few years new cases of on-going gene flow and hybridization in ants have been documented, some between non-sister species (Steiner et al. 2011; Kronauer et al. 2011). In several cases hybridization has been largely unidirectional. For example, the hybrids between *Formica cinerea* and *F. selysi* from the subgenus *Serviformica* carry mainly mtDNA of *F. cinerea* (Purcell et al. 2016). Our results did not indicate such unidirectionality, as haplotypes of both parental species were present in the hybrids. Interestingly, Purcell et al. (2016) identified few genomic regions underlying reproductive barriers in the hybrids, while most genetic markers seemed to be mixing freely. In contrast to this, our earlier studies have indicated selection on several genomic regions scattered across the genome in the Långholmen hybrid population (Kulmuni et al. 2010; Kulmuni and Pamilo 2014).

This article is protected by copyright. All rights reserved.

Accepted Article

Ancestry of the Långholmen hybrids

The current data allow us to establish the ancestry of the previously studied Långholmen hybrids. The Långholmen population (Kulmuni et al. 2010; Kulmuni & Pamilo 2014) contains two distinct genetic groups (R and W) both of which have a hybrid ancestry and live side by side within the same nests. We previously anticipated that one of these hybrid groups is more similar to *F. aquilonia* and the other one to *F. polyctena* (Kulmuni et al. 2010) but lacked the genetic data to test this hypothesis. With the current data we are able to confirm that the group previously called W is closer to putative *F. aquilonia* in the PCoA and the group called R is closer to the putative *F. polyctena* (Figure 3). This conclusion is further supported when we look at the alleles that were earlier found to be diagnostic (i.e. not introgressing) between the R and W groups in Långholmen. The individuals in the six nests, here considered to represent parental-like *F. polyctena,* had on average 4.04 alleles diagnostic to the R group and 0.76 alleles diagnostic to the W group. In the 21 nests, considered here to represent *F. aquilonia*, the individuals had on average 2.12 alleles diagnostic to group W and 0.11 diagnostic to R. Thus, enrichment of R diagnostic alleles in the putative parental *F. polyctena* and W diagnostic alleles in the putative parental *F. aquilonia*, further supports the conclusion that the R group contains *F. polyctena*-like hybrids and W group contains *F. aquilonia*-like hybrids.

Based on the Structure results, the majority of the newly discovered hybrids were genetically distinct from the Långholmen hybrids. Only two hybrid nests (Tvärminne Station 1, Halsholmen 5) had a similar nuclear composition to the R group in the Långholmen hybrid population and none were exactly like the W group. In the current analysis Structure separated W individuals into their own group, in spite of their clearly being hybrids based on previous morphological analysis (Kulmuni et al. 2010) and the current PCoA (Figure 3). The

This article is protected by copyright. All rights reserved.

separation of this group may have arisen because Structure has a tendency to identify a higher number of clusters in cases where the majority of the samples are admixed (Pritchard et al. 2000b).

Underlying patterns of cytonuclear mismatch

Our data show a striking pattern of cytonuclear discordance in the hybrids: nests which are closer to parental-like *F. polyctena* for nuclear markers preferentially have *F. aquilonia* mitochondrial haplotypes and nests with nuclear genomes closer to parental-like *F. aquilonia* are likely to have *F. polyctena* mitochondrial haplotypes. This pattern is in strong contrast to the general findings of cytonuclear incompatibilities in several hybrid systems, where mismatches between mitochondrial and nuclear genes are a common cause of hybrid breakdown (King & Attardi 1989; Derr et al. 2012; For a review see Burton et al. 2013). This breakdown is likely to be due to the disruption of the co-evolutionary interactions between nucleus and mitochondria, because the majority of the proteins required for mitochondrial function are encoded by the nuclear genome in both plants and animals (Burton et al. 2013).

Despite potential for hybrid breakdown, introgression of mitochondrial DNA to sister species is not uncommon. Furthermore, recent modeling work has demonstrated that, all else being equal, the mitochondrial DNA introgresses more readily than the nuclear DNA in haplodiploids (Patten et al. 2015). Patten et al. (2015) model introgression comparing one nuclear locus and one mitochondrial locus in diplodiploid and haplodiploid species. They show that in a neutral case with no asymmetries (equal mating probabilities, survival etc. in both sexes) mitochondrial and nuclear introgression are equal in diplodiploid species. By contrast, haplodiploids are assured of biased mitochondrial compared to nuclear introgression. This is because the production of hybrid males (that promote nuclear

This article is protected by copyright. All rights reserved.

Accepted Article

introgression) always lags behind hybrid females (that promote both nuclear and mtDNA introgression) by one generation. However, the model by Patten et al. (2015) considers unidirectional migration from one species into another and does not explain our results, where we have a case comparable to a hybrid swarm and symmetric cytonuclear mismatch in both directions (i.e. hybrids close to parental-like *F. aquilonia* for nuclear markers have *F. polyctena* mitochondrial haplotypes and vice versa).

The pattern of cytonuclear mismatch observed in our hybrids is not likely to arise under random mating without selection. There are at least two possible explanations for the systematic cytonuclear mismatch observed in this study. First, if females systematically choose to mate with heterospecific rather than conspecific males, this could result in cytonuclear mismatch. This type of preference to mate with dissimilar males could evolve within a species in response to inbreeding depression, which is a potential problem for both single and multiple queen ant colonies (Trontti et al. 2005; Vitikainen et al. 2011). However, if preference for dissimilar males leads to heterospecific mating it is likely to have negative fitness consequences. Despite this potential for reduced fitness in hybrid offspring, preference for dissimilar males has been shown for example in copepods (Palmer & Edmands 2000) and flycatchers (Veen et al. 2001). The second alternative scenario is that females mate randomly, but offspring with heterospecific combinations of nuclear genome and maternal cytoplasm have higher fitness than individuals with conspecific combinations. This is consistent with our previous results that show selection for nuclear introgression in females (Kulmuni and Pamilo 2014). In this scenario is also possible that selection does not favor the introgressed mitochondria specifically, but rather heterospecific cytoplasm. Heterospecific cytoplasm could be favoured because of maternally inherited endosymbionts (Werren 1997; Sirviö & Pamilo 2012) that are known to bias reproductive output (Werren 2008), but also to confer adaptive traits (e.g. Xie et al. 2015).

This article is protected by copyright. All rights reserved.

Patterns of genetic differentiation between sexes in hybrids

One of the newly studied hybrid localities showed genetic differences between sexes similar to those observed in the Långholmen hybrid population. In locality Tvärminne road there were altogether 8 alleles at 5 loci that were lacking from the males but occurred with high frequencies in the female heterozygotes, the frequency differences between the sexes being highly significant at all five loci. Genetic differences between the sexes in the Långholmen population have been shown to arise due to antagonistic effects of hybridization: females heterozygous for introgressed alleles show hybrid vigour, whereas males with the same alleles show hybrid breakdown (Kulmuni & Pamilo 2014). To explain this pattern we suggested that haploid males suffer from recessive incompatibilities, which can be masked in diploid females as heterozygotes. In addition, the heterozygous hybrid females enjoy a fitness advantage. Although we have not specifically tested for selection in the current study, the genetic patterns observed in locality Tvärminne Road are consistent with such selection. We know that in the Långholmen hybrid population the ants reproduce sexually and their genomes recombine (Kulmuni & Pamilo 2014). If we assume these same conditions in the Tvärminne Road locality (which seem likely as relatedness among sampled individuals was low), antagonistic selection among sexes is a likely cause maintaining the observed sexes differences. In the current study we do not have the data to test whether incompatibilities causing putative hybrid male death are due to nucleo-nuclear or cytonuclear mismatches. However, the data from our previous hybrid population (Kulmuni & Pamilo 2014) is inconsistent with purely cytonuclear incompatibilities.

The five loci showing genetic differentiation between sexes and putative hybrid problems in locality Tvärminne Road also show these patterns in the Långholmen hybrid population. However, some of the alleles linked to putative hybrid problems were not the same in these

This article is protected by copyright. All rights reserved.

Accepted Article

two localities (e.g. FE13[198] shows sex differences in R group in Långholmen but does not in Tvärminne Road). Also, the mitochondrial haplotypes differ between these localities (mtDNA of *F. polyctena* in Långholmen and of *F. aquilonia* in Tvärminne Road), suggesting that independent hybridization events have led to a similar outcome. Even though some of the alleles under putative antagonistic selection between males and females differ between the two hybridization events, the loci linked to these alleles and causing the selection pressures may still be the same. Interestingly, the other hybrid locality tested here (Halsholmen) did not show genetic differentiation between the sexes for the markers used. This suggests that hybridization can have varied consequences even between the same pair of species or alternatively, that hybridization histories differ between the localities e.g. due to strong ancestral population differentiation (Pamilo et al. 2005).

Recent studies have revealed more cases in haplodiploids that show both positive and negative fitness consequences of hybridization within the same system (Knegt et al. 2017; König 2016). For example, haploid males from interspecific crosses of spider mites revealed that some interspecific marker combinations were under-represented indicating incompatibilities, while others were over-represented suggesting heterosis (Knegt et al. 2017). These emerging systems allow exciting opportunities to study the interplay between adaptive introgression and incompatibilities.

Life history characteristics that may promote hybridization

Several life history characteristics may predispose *F. aquilonia* and *F. polyctena,* along with other *F. rufa* group species, to hybridization. Both *F. aquilonia* and *F. polyctena* are highly polygynous (i.e. have multiple queens) (Rosengren & Pamilo 1983; Stockan & Robinson 2016). New males and queens from polygynous colonies tend to mate within a nest (Stockan

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

& Robinson 2016) but can disperse longer distances as well, in which case the dispersing queens are not able to establish new nests alone. Instead, they can enter existing nests of the subgenus *Serviformica*, killing the host queens and taking over the colony (Stockan & Robinson 2016). Alternatively, a dispersing queen can enter an existing nest of their own, or a related species, and become an additional queen. The outcome is a polygynous colony with a possibility of a mixed nest with related species if a foreign queen is allowed to enter. Mixed nests can also arise when a population of one species expands and takes over nearby nests of related species. Such mixed colonies could create opportunities for hybridization, when mating takes place within the nest. In summary, polygyny and social parasitism may create conditions that predispose ant species to hybridization, even though they are not necessary (Purcell et al. 2016). Eventually, the species that is less abundant may end up hybridizing to extinction (Rhymer 1996). This is because it is difficult to find conspecific individuals to mate with when a species is rare. This may be the case with *F. polyctena*, whose distribution in Finland is restricted to South and for which we could find only six parental-like nests.

In conclusion, our results suggest widespread hybridization and incomplete reproductive isolation between the two closely-related wood ant species, *F. aquilonia* and *F. polyctena,* in Southern Finland. Furthermore we document an unusual case of systematic cytonuclear mismatch, the causes of which should be studied in more detail. We also reveal a second hybrid locality showing genetic patterns consistent with selection favoring hybrid females but acting against hybrid males. Taken together, our current and previous studies (Kulmuni and Pamilo 2014) suggest introgression has both negative and positive consequences in terms of fitness. In the future the causes of these selection pressures should be investigated using a genome-wide perspective.

This article is protected by copyright. All rights reserved.

Accepted Article

Acknowledgements

We thank Riitta Jokela for work in the laboratory and Anne Hurskainen and Tiina Möllări for their help in sample collection. We thank Marita and the late Rainer Rosengren for giving us access to the Pikkala population, and for their help in sample collection. JK was supported by Biocenter Oulu Graduate School, Finnish foundations post doc pool, Finnish Cultural foundation and Human Frontier Science program (LT000614/2014). ME was supported by a postdoctoral grant of the EU 6th framework training network 'INSECT' with PP, and by a Marie Curie postdoctoral fellowship with LS.

References

Abbott R, Albach D, Ansell S et al. (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.

Avery PJ (1984) The population genetics of haplo-diploids and X-linked genes. *Genetics Research*, **44**, 321–341.

Bateson W (1909) Darwin and Modern Science. Cambridge University Press, Cambridge.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology, Systematics and Evolution*, **16**, 113–148.

Barreto FS, Pereira RJ, Burton RS (2015) Hybrid Dysfunction and Physiological Compensation in Gene Expression. *Molecular Biology and Evolution,* **32**, 613–622.

Betrem JG (1964) Einige Bemerkungen über Formica-material aus Nordostkarelien. *Waldhygiene,* **5**, 215-216

Beukeboom LW, Koevoets T, Morales HE, Ferber S, van de Zande L (2015) Hybrid incompatibilities are affected by dominance and dosage in the haplodiploid wasp Nasonia. *Frontiers in Genetics*, **6**, 1–16.

Burton RS, Pereira RJ, Barreto FS (2013) Cytonuclear Genomic Interactions and Hybrid Breakdown. *Annual Review of Ecology, Evolution and Systematics*, **44**, 281–302.

Butlin RK, Galindo J, Grahame J (2008) Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philosophical Transactions of the Royal Society B*, **363**, 2997–3007.

Capblancq T, Després L, Rioux D, Mavárez J (2015) Hybridization promotes speciation in Coenonympha butterflies. *Molecular Ecology*, **24**, 6209-6222.

This article is protected by copyright. All rights reserved.

Accepted Article

Chapuisat M (1996) Characterization of microsatellite loci in Formica lugubris B and their variability in other ant species. *Molecular Ecology*, **5**, 599–601.

Choleva L, Janko K, De Gelas K, Bohlen J, Šlechtová V, Rábová M, Ráb P (2012) Synthesis of clonality and polyploidy in vertebrate animals by hybridization between two sexual species. *Evolution*, **66**, 2191–203.

Clement M, Snell Q, Walke P, Posada D & Crandall K. (2002) TCS: estimating gene genealogies. *Proceedings of the 16th International Parallel and Distributed Processing Symposium*, **2**, 184.

Collingwood CA (1979) The Formicidae (Hymenoptera) of Fennoscandia and Denmark. Science Press, Klampenbor.

Derr JN, Hedrick PW, Halbert ND, Plough L, Dobson LK, et al. 2012. Phenotypic effects of cattle mitochondrial DNA in American bison. *Conservation Biology,* **26**, 1130–36

Dobzhansky T (1936) Studies on hybrid sterility. II. Localization of sterility factors in Drosophila pseudoobscura hybrids. *Genetics,* **21**, 113–135.

Douwes P (2012) Nationalnyckeln till Sveriges flora och Fauna. Steklar: Myror-getingar. Hymenoptera: Formicidae-Vespidae. ArtData-banken, SLU, Uppsala.

Drummond AJ. Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources,* **4**, 359–361

Edgar RC. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acid Research*, **32**, 1792–1797.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, **14**, 2611–2620.

Feldhaar H, Foitzik S, Heinze J (2008) Lifelong commitment to the wrong partner: Hybridization in ants. *Philosophical Transactions of the Royal Society B,* **363**, 2891–2899.

Goropashnaya AV, Fedorov VB & Pamilo P (2004a) Recent speciation in the Formica rufa group ants (Hymenoptera, Formicidae): inference from mitochondrial DNA phylogeny. *Molecular Phylogenetics and Evolution*, **32**, 198–206.

Goropashnaya AV, Fedorov VB, Seifert B, Pamilo P (2004b) Limited phylogeographical structure across Eurasia in two red wood ant species Formica pratensis and F. lugubris (Hymenoptera, Formicidae). *Molecular Ecology*, **13**, 1849–1858.

This article is protected by copyright. All rights reserved.

Goropashnaya AV, Fedorov VB, Seifert B & Pamilo P. (2012) Phylogenetic relationships of Palaearctic Formica species (hymenoptera, Formicidae) based on mitochondrial cytochrome b sequences. *PLoS ONE*, **7**, e41697.

Gyllenstrand N, Gertsch PJ, Pamilo P (2002) Polymorphic microsatellite DNA markers in the ant Formica exsecta. *Molecular Ecology Notes*, **2**, 67–69.

Haldane JBS (1922) Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, **12**, 101–109.

Hasegawa E, Imai S (2004) Characterization of microsatellite loci in red wood ants Formica (s.str.) spp. and the related genus Polyergus. *Molecular Ecology Notes*, **4**, 200–203.

Heliconius Genome Consortium. (2012) Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, *487*, 94–98.

Janoušek V, Wang L, Luzynski K, Dufková P, Vyskočilová MM, Nachman MW, Munclinger P, Macholán M, Piálek J, Tucker PK (2012) Genome-wide architecture of reproductive isolation in a naturally occurring hybrid zone between Mus musculus musculus and M. m. domesticus. *Molecular Ecology*, **21**, 3032–47.

Jermiin LS, Crozier RH. (1994) The cytochrome b region in the mitochondrial DNA of the ant Tetraponera rufonigra: sequence divergence in Hymenoptera may be associated with nucleotide content. *Journal of Molecular Evolution*, **38**, 282–94

King MP, Attardi G (1989) Human cells lacking mtDNA: repopulation with exogenous mitochondria by complementation. *Science*, **246**, 500–3.

Knegt B, Potter T, Pearson NA, Sato Y, Staudacher H, Schimmel BCJ, Kiers ET, Egas M (2017) Detection of genetic incompatibilities in non-model systems using simple genetic markers: hybrid breakdown in the haplodiploid spider mite Tetranychus evansi. *Heredity*, **118**, 311-321.

Koevoets T, Beukeboom LW (2009) Genetics of postzygotic isolation and Haldane's rule in haplodiploids. *Heredity*, **102**, 16–23.

Kulmuni J, Seifert B, Pamilo P (2010) Segregation distortion causes large-scale differences between male and female genomes in hybrid ants. *Proceedings of the National Academy of Sciences*, **107**, 7371–7376.

Kulmuni J & Pamilo, P. (2014) Introgression in hybrid ants is favored in females but selected against in males. *Proceedings of the National Academy of Sciences*, **111**, 12805–12810.

König K (2016) Mechanisms of speciation in the parasitoid wasp Lariophagus distinguendus (Föster). Thesis Chapter III. Unpublished PhD thesis, University of Hohenheim, Stuttgart.

Leigh JW. (2016) Popart. v.1.7. Department of Maths and Statistics, The University of Otego, Dunedin, New Zeeland. Available at: http://popart.otago.ac.nz [Last Accessed: 4th May 2016]

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

Liautard C, Keller L (2001) Restricted effective queen dispersal at a microgeographic scale in polygynous populations of the ant Formica exsecta. *Evolution,* **55**, 2484–2492.

Mallet J (2007) Hybrid speciation. *Nature,* **446**, 279–283.

Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, Blaxter ML, Manica A, Mallet J, Jiggins CD (2013) Genome-wide evidence for speciation with gene flow in Heliconius butterflies. *Genome Research,* **23**, 1817–1828.

Mavarez J, Salazar CA, Bermingham E, Salcedo C, Jiggins CD and Linares M (2006) Speciation by hybridization in *Heliconius* butterflies. *Nature*, **441**, 868–871

Mavárez J, Linares M (2008) Homoploid hybrid speciation in animals. *Molecular Ecology*, **17**, 4181–4185.

Muirhead CA, Presgraves DC (2016) Hybrid incompatibilities, local adaptation, and the genomic distribution of natural introgression between species. *American Naturalist*, **187**, 249–261.

Muller HJ (1942) Isolating mechanisms, evolution and temperature. *Biological Symposia*, **6**, 71–125.

Palmer CA, Edmands S. (2000) Mate choice in the face of both inbreeding and outbreeding depression in the intertidal copepod Tigriopus californicus. *Marine Biology*, **136**, 693–698.

Pamilo P (1993) Polyandry and allele frequency differences between the sexes in the ant *Formica aquilonia*. *Heredity* **70**, 472-480.

Pamilo P, Zhu D, Fortelius W, Rosengren R, Seppä P, Sundström L. (2005) Genetic patchwork of network-building wood ant populations. *Annales Zoologici Fennici* **42**, 179-187.

Patten MM, Carioscia SA, Linnen CR. Biased introgression of mitochondrial and nuclear genes: a comparison of diploid and haplodiploid systems. *Molecular Ecology*, **24**, 5200–5210.

Payseur BA, Rieseberg LH (2016) A genomic perspective on hybridization and speciation. *Molecular Ecology*, **25**, 2337–2360.

Peakall, R. and Smouse P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, **28**, 2537–2539.

Peakall, R. and Smouse P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000b) Association mapping in structured populations. *American Journal of Human Genetics*, **67**, 170–181.

This article is protected by copyright. All rights reserved.

Accepted Article

Purcell J, Zahnd S, Athanasiades A, Türler R, Chapuisat M, Brelsford A (2016) Ants exhibit asymmetric hybridization in a mosaic hybrid zone. *Molecular Ecology*, **25**, 4866-4874.

Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution*, **43**, 258–275.

Rambaut A. (2014) FigTree. v1.4.2 (2014) Last accessed 4/5/2016. Available from: http://tree/bio.ed.ac.uk/software/figtree/

Ramasamy RK, Ramasamy S, Bindroo BB, Naik VG (2014) STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. *Springerplus,* **13**, 431

Raymond M, Rousset F (1995) GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rhymer MR (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83–109.

Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology and Systematics,* **28**, 359–389.

Ronquist FM, Teslenko P, van der Mark D, Ayres A, Darling S, Höhna B, Larget L, Liu MA, Suchard, Huelsenbeck JP (2011) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539-42.

Rosengren R, Pamilo P (1983) The evolution of polygyny and polydomy in moundbuilding Formica ants. *Acta Entomologica Fennica*, **42**, 65–77.

Rudman R, Schluter D (2016) Ecological impacts of Reverse Speciation in Threespine Stickleback. *Current Biology*, **26**, 490-495.

Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, Peichel CL. Saetre G et al. (2014) Genomics and the origin of species. *Nature Reviews Genetics*, **15**, 176–92.

Seifert B (1999) Interspecific hybridisations in natural populations of ants by example of a regional fauna. (Hymenoptera, Formicidae). *Insectes Sociaux*, **46**, 45–52.

Seifert, B Kulmuni J & Pamilo, P. (2010) Independent hybrid populations of Formica polyctena X rufa wood ants (Hymenoptera:Formicidae) abound under conditions of forest fragmentation. *Evolutionary Ecology*, **24**, 1219–1237.

Schwarz D, Matta BM, Shakir-Botteri NL, McPheron BA (2005) Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature*, **436**, 546–549

Sirviö A, Pamilo P (2010) Multiple endosymbionts in populations of the ant Formica cinerea. *BMC Evolutionary Biology*, 10, 1–12.

This article is protected by copyright. All rights reserved.

Accepted Article

Song Y, Endepols S, Klemann N, Richter D, Matuschka F, Shih C, Nachman MW, Kohn MH (2011) Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Current Biology*, **21**, 1296–1301.

Sorvari, J (2006) Two distinct morphs in the wood ant *Formica polyctena* in Finland: a result of hybridization? *Entomologica Fennica* **17,** 1-17.

Steiner FM, Seifert B, Grasso DA, Le Moli F, Arthofer W, Stauffer C, Crozier RH, Schlick-Steiner BC (2011) Mixed colonies and hybridization of Messor harvester ant species (Hymenoptera: Formicidae). *Organisms Diversity and Evolution,* **11**, 107–134.

Stockan J, Robinson EJH (2016). Wood Ant Ecology and Conservation. Cambridge University Press, Cambridge.

Trigo TC, Schneider A, de Oliveira TG, Lehugeur LM, Silveira L, Freitas TRO, Eizirik E (2013) Molecular data reveal complex hybridization and a cryptic species of neotropical wild cat. *Current Biology*, **23**, 2528–33.

Trontti K, Aron S, Sundström L (2005) Inbreeding and kinship in the ant Plagiolepis pygmaea. *Molecular Ecology*, **14**, 2007–15.

Umphrey GJ (2006) Sperm parasitism in ants: Selection for interspecific mating and hybridization. *Ecology*, **87**, 2148–2159.

Veen T, Borge T, Griffith SC, Sætre G-P, Bures S, Gustafsson L, Sheldon BC (2001) Hybridization and adaptive mate choice in flycatchers. *Nature*, **411**, 45-50.

Vitikainen E, Haag-Liautard C, Sundström L. (2011) Inbreeding and reproductive investment in the ant Formica exsecta. *Evolution*, **65**, 2026–37.

Werren JH. (1993) The evolution of inbreeding in haplodiploid organisms. in: Thornhill NW (ed.) The Natural History of Inbreeding and Outbreeding. Theoretical and Empirical Perspectives. The University of Chicago Press, Chicago.

Werren JH (1997) Biology of Wolbachia. *Annual Review of Entomology,* **42**, 587–609.

Werren JH, Baldo L, Clark ME (2008) Wolbachia: master manipulators of invertebrate biology. *Nature Reviews Microbiology*, **6**, 741–751.

Xie JC, Winter L, Winter, Mateos M (2015) Rapid spread of the defensive endosymbiont Spiroplasma in Drosophila hydei under high parasitoid wasp pressure. *FEMS Microbiology Ecology*, **91**, 1–11.

Data accessibility

Genotype data and mitochondrial sequences have been deposited to Dryad

(doi:10.5061/dryad.576hg).

This article is protected by copyright. All rights reserved.

Accepted Article

Author contributions

JK designed the research. ME, PP and JK collected the samples. JB, ME, LP, PP and JK analyzed the data. RB contributed to analysis methods. JB, ME, LS, RB, PP and JK wrote the paper.

Figure 1. Sampling localities of ants in the Southern Finnish coastal region, to the west of Helsinki. The location of new sampling localities is represented by a number from 1 to 16, with the names outlined in the key. Our previous hybrid population (no. 17) is indicated within a red rectangle and contains a single mitochondrial haplotype, that of *F. polyctena*. Sampling localities are accompanied by a pie chart displaying the proportion of mitochondrial haplotypes sampled from that locality (orange for *F. polyctena*, blue for *F. aquilonia*), along with the number of individuals tested for mitochondrial DNA at each location (n). *) The Unknown locality does not have gps coordinates but had *F. polyctena* mitochondrial haplotype.

Figure 2. Haplotype network from 265 sampled mitochondrial sequences and reference sequences from Genbank. Reference sequences have been named according to the species identity reported in Genbank together with their ID number and new haplotypes found in this study are accompanied by a prefix Seq. The branch lengths are not to scale, but the number on each branch indicates the number of SNPs dividing the two haplotypes connected by that branch. The size of the haplotype circle is proportional to the number of haplotypes sampled in this study.

Figure 3. Principal Coordinates analysis of nuclear markers in 96 nest samples. The 1st axis explains 18.6 % of the variation and the 2nd axis 15.6 %. The data points are colored according to the mitochondrial haplotype that was sampled from those nests (orange for *F.*

This article is protected by copyright. All rights reserved.

Accepted Article

polyctena, blue for *F. aquilonia* and black for nests that contained both haplotypes). Previously studies (Kulmuni et al. 2010) hybrid groups R and W are indicated with arrows. The 1st axis divides the samples according to parental gene pools, where parental-like *F. polyctena* is situated to the right, parental-like *F. aquilonia* to the left and hybrids in between.

Figure 4. Results from the analysis of population structure with a model allowing for admixture. Here we present the uppermost genetic structure of three groups as indicated by the deltaK statistics (group 1 = green, group 2 = yellow, group 3 = blue). Sampling localities are separated by vertical red lines and numbered from 1 to 17. Several localities show signatures of admixture. Mitochondrial haplotypes found in each nest are indicated below the Structure plot with dots (orange for *F. polyctena*, light blue for *F. aquilonia* and black where both haplotypes were found from the same nest). Sampling localities are the following; 1=Hanko, 2=Tvärminne Station, 3= Halsholmen, 4= Joskär, 5= Svanvik, 6= Sammatti,7= Unknown, 8= Pusula, 9= Fiskars, 10= Tenala, 11= Grabbskog, 12= Solböle, 13= Siuntio1, 14= Tvärminne Road, 15= Siuntio2, 16= Pikkala, 17= Previously studied hybrid population on Långholmen (order of samples in this locality; W-males, W-females, R-males, R-females).

This article is protected by copyright. All rights reserved.

Table 1. Assignment* of nest samples into putative parental-like gene pools and hybrids.

	Putative parental-like *F. aquilonia*	**Putative parental-like *F. polyctena***	**Hybrids**
Localities found from	Pusula, Grabbskog, Solböle, Siuntio	Fiskars, Grabbskog, Joskär, Unknown	Pikkala, Sammatti, Hanko, Svanvik, Halsholmen, Siuntio 1, Siuntio 2, Solböle, Tvärminne Road, Tvärminne Station
No. nests	21	6	69
Structure assignment of individuals	Group 1	Group 2	Group 1, 2 and 3
Mean of the 1st principal coordinate axis over all nests	-0.497	0.516	0.075
Admixture score between the three Structure groups	<10% in all individuals	≤ 2% in 44 out of 49 sampled individuals, ≤10% in the remaining 5 individuals	varies between 1 and 50% per individual
Mt-haplotypes found	*F. aquilonia*	*F. polyctena*	*F. aquilonia* and *F. polyctena*

*) Nests are assigned into one of the three categories based on the combination of Structure, PCoA and mtDNA analysis (e.g. nest is assigned into putative parental-like *F. aquilonia* if it harbours only *F. aquilonia* mtDNA, has over 90% assignment into Structure group 1 and situates to the left on the PCoA).

Accepted Article

This article is protected by copyright. All rights reserved.

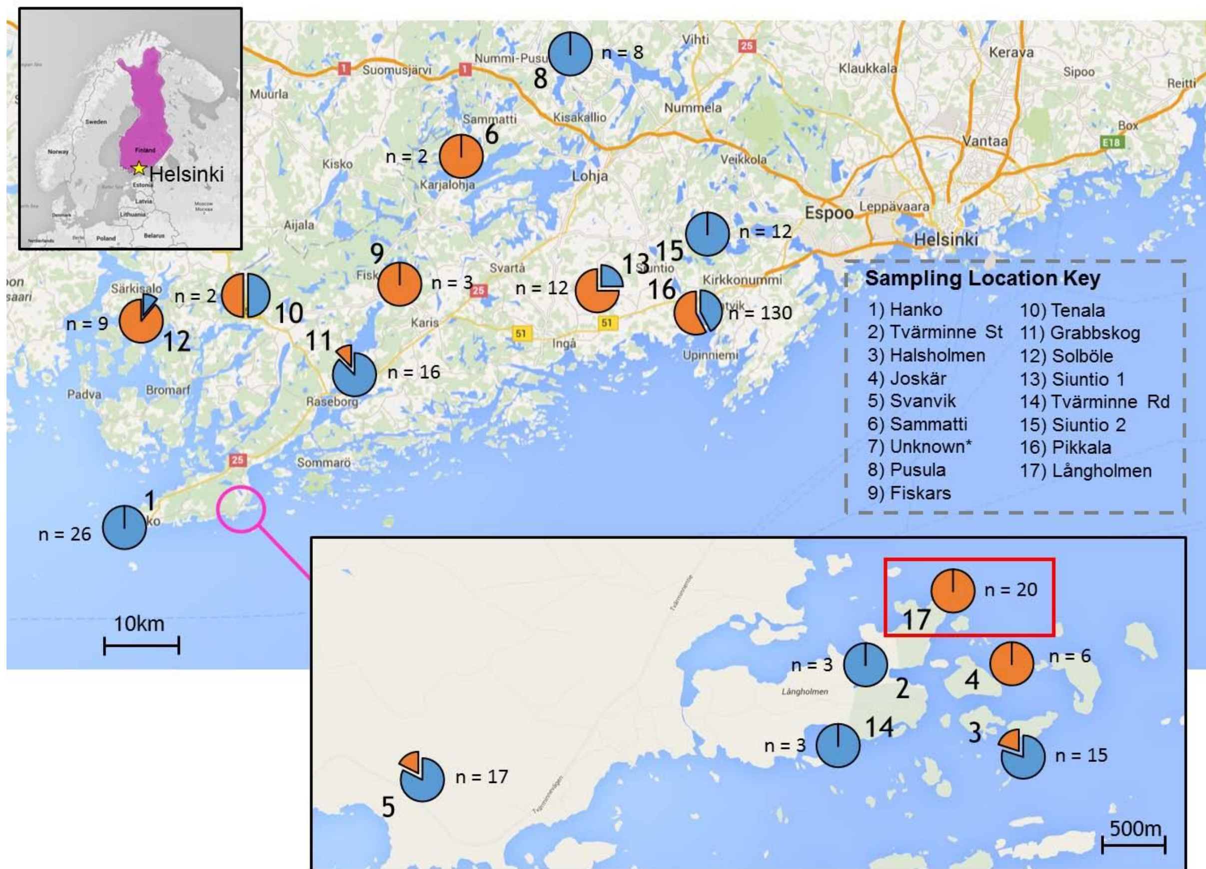
Accepted Article

Table 2. Allele frequencies in males and females in two new hybrid localitites Halsholmen and Tvärminne Road.

		Tvärminne Rd.			Halsholmen		
Locus	**Allele**	**males (N=30)**	**females (N=30)**	**p-value[1]**	**males (N=18)**	**females (N=54)**	**p-value[1]**
FE13	**186**	0,000	0,086	0,000	0,063	0,123	0,111
	189	0,167	0,190		-	-	
	192	0,000	0,172*		0,563	0,443	
	195	0,000	0,155*		-	-	
	198	0,833	0,397		0,375	0,434	
FE17	**110**	0,000	0,083	0.185	0,000	0,009	0,467
	116	0,367	0,300		0,104	0,204	
	118	0,633	0,617		0,896	0,787	
FE19	**184**	0,200	0,400	0.062	0,313	0,104	0,112
	186	0,800	0,600		0,542	0,689	
	193	-	-		0,146	0,208	
FE7	**56**	0,000	0,183*	0,000	0,313	0,122	0,010
	62	0,000	0,083		-	-	
	66	0,500	0,367		0,208	0,122	
	68	-	-		0,229	0,245	
	70	-	-		0,021	0,163	
	74	0,000	0,117*		0,229	0,347	
	89	0,500	0,250		-	-	
FY13	**195**	0,000	0,100*	0.129	0,543	0,415	0,326
	197	0,571	0,550		0,413	0,425	
	199	0,429	0,350		0,043	0,160	
FY15	**221**	0,233	0,117	0,000	0,250	0,337	0,203
	229	0,000	0,150*		0,354	0,194	
	234	0,000	0,067		0,250	0,214	
	236	0,467	0,300		-	-	
	240	0,000	0,267*		0,146	0,255	
	244	0,300	0,100		-	-	
FY3	**185**	-	-	0.005	0,167	0,104	0,723
	188	0,000	0,117*		0,125	0,189	
	190	0,400	0,567		0,333	0,387	
	192	-	-		0,000	0,019	
	194	0,600	0,317		0,146	0,189	
	196	-	-		0,229	0,113	
FY12	**185**	0,533	0,300	0.019	0,326	0,519	0,106
	187	0,000	0,083		0,522	0,333	
	189	0,200	0,250		-	-	
	191	0,267	0,250		0,152	0,148	
	193	0,000	0,117*		-	-	
FL29	**182**	0,600	0,667	0.663	1,000	1,000	1,000
	188	0,400	0,317		-	-	
	190	0,000	0,017		-	-	

This article is protected by copyright. All rights reserved.

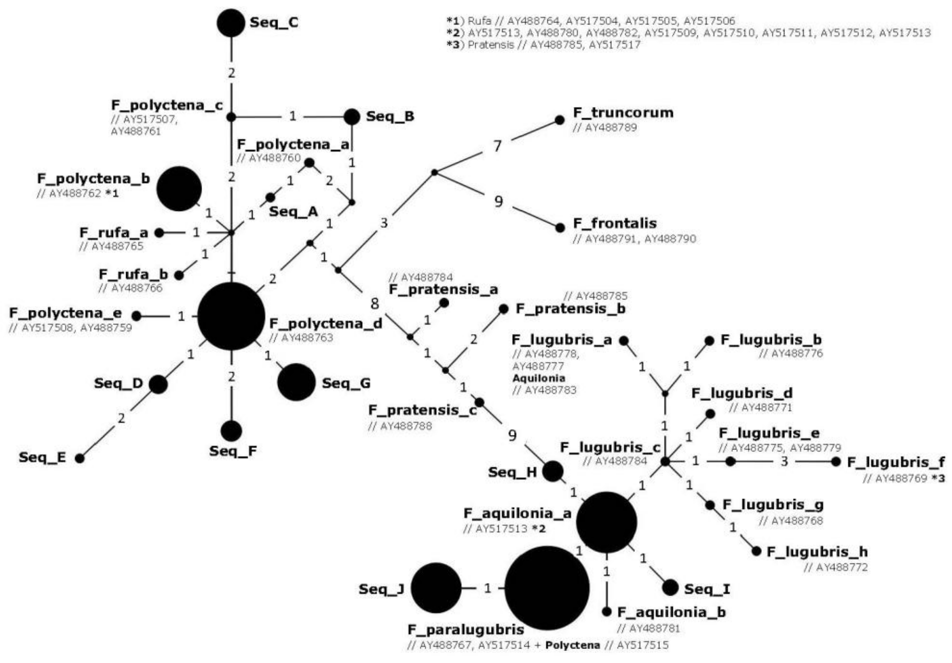
* Alleles that are absent from males, but occur in females only as heterozygotes and at frequencies above 0.1
[1] P-values test differentiation between male and female allele frequencies within a locality

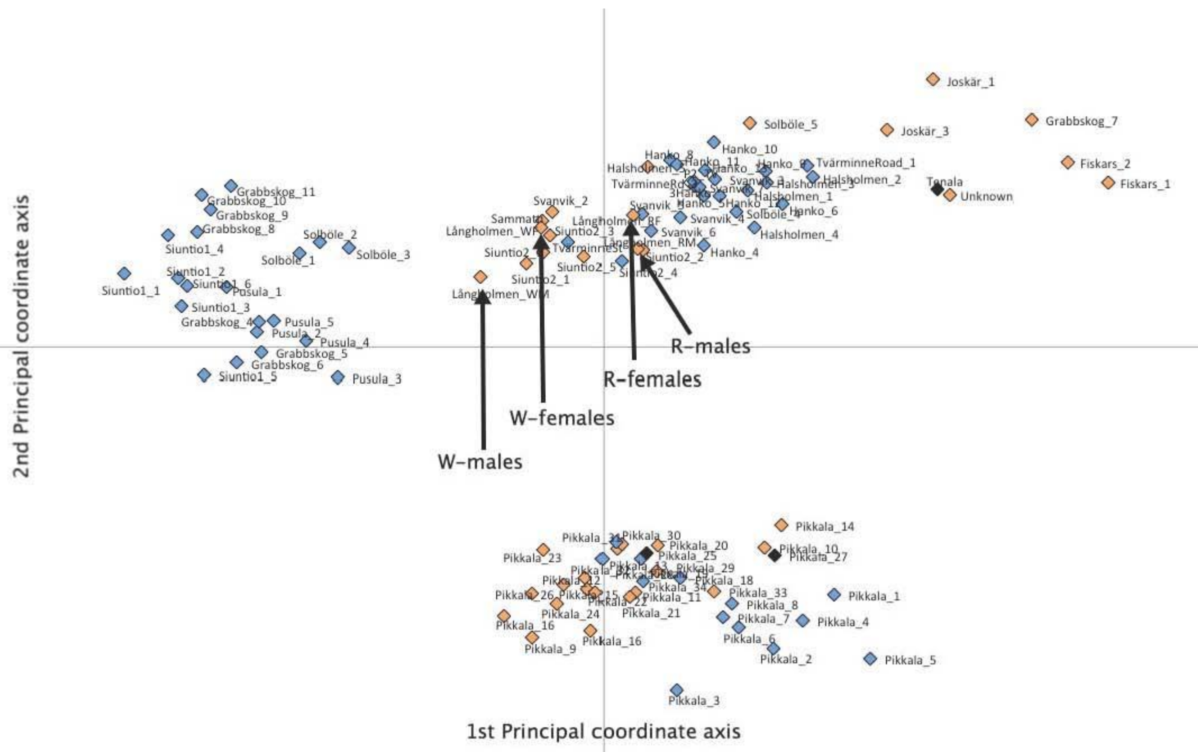


This article is protected by copyright. All rights reserved.

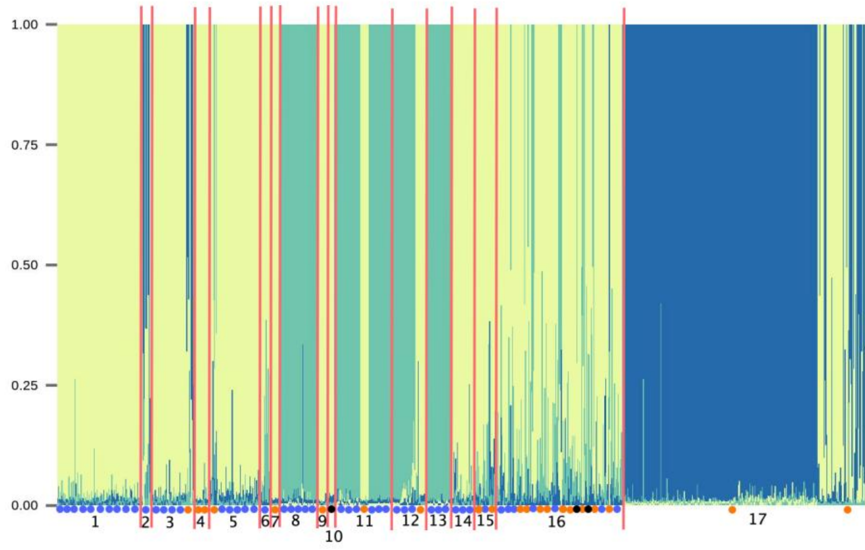
Accepted Article

Accepted Article





This article is protected by copyright. All rights reserved.



Accepted Article

This article is protected by copyright. All rights reserved.