

# Characterization and distribution of *Pogonomyrmex* harvester ant lineages with genetic caste determination

TANJA SCHWANDER,\* SARA HELMS CAHAN† and LAURENT KELLER\*

\*Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland, †Department of Biology, University of Vermont, Burlington, VT 05405, USA

## Abstract

Genetic caste determination has been described in two populations of *Pogonomyrmex* harvester ants, each comprising a pair of interbreeding lineages. Queens mate with males of their own and of the alternate lineage and produce two types of diploid offspring, those fertilized by males of the queens' lineage which develop into queens and those fertilized by males of the other lineage which develop into workers. Each of the lineages has been shown to be itself of hybrid origin between the species *Pogonomyrmex barbatus* and *Pogonomyrmex rugosus*, which both have typical, environmentally determined caste differentiation. In a large scale genetic survey across 35 sites in Arizona, New Mexico and Texas, we found that genetic caste determination associated with pairs of interbreeding lineages occurred frequently (in 26 out of the 35 sites). Overall, we identified eight lineages with genetic caste determination that always co-occurred in the same complementary lineage pairs. Three of the four lineage pairs appear to have a common origin while their relationship with the fourth remains unclear. The level of genetic differentiation among these eight lineages was significantly higher than the differentiation between *P. rugosus* and *P. barbatus*, which questions the appropriate taxonomic status of these genetic lineages. In addition to being genetically isolated from one another, all lineages with genetic caste determination were genetically distinct from *P. rugosus* and *P. barbatus*, even when colonies of interbreeding lineages co-occurred with colonies of either putative parent at the same site. Such nearly complete reproductive isolation between the lineages and the species with environmental caste determination might prevent the genetic caste determination system to be swept away by gene flow.

**Keywords:** genetic caste determination, *Pogonomyrmex*, genetic lineage, hybridization, social insect

Received 5 May 2006; revision accepted 7 August 2006

## Introduction

In ants and other social Hymenoptera forming large and complex societies, reproductive division of labour is generally associated with marked morphological and physiological differences between the queen and worker castes (Hölldobler & Wilson 1990). These differences typically result from the combined effects of environmental and social factors inducing female larvae to initiate distinct queen or worker developmental pathways (Brian 1957; Wheeler 1994). Two exceptions to such environmentally determined caste differentiation have been

described in *Pogonomyrmex* harvester ants in Arizona and New Mexico (Helms Cahan *et al.* 2002; Julian *et al.* 2002; Volny & Gordon 2002b; Helms Cahan & Keller 2003). In each case, a single population is composed of two genetically distinct lineages [the H1 and H2 lineages at the site Hidalgo and the J1 and J2 lineages at the site Junction (Helms Cahan & Keller 2003)]. Queens and males of both lineages form a single mating aggregation; queens mate multiply and lay female eggs fathered by both types of males (Schwander *et al.* 2006; Volny *et al.* in press). Pure-lineage female offspring develop into queens while interlineage females almost invariably develop into workers (Helms Cahan *et al.* 2002; Julian *et al.* 2002; Volny & Gordon 2002b; Helms Cahan & Keller 2003). This mode of caste determination is considered to be genetically

Correspondence: Tanja Schwander, Fax: +41 21 692 41 65; E-mail: [tanja.schwander@unil.ch](mailto:tanja.schwander@unil.ch)

hardwired, as females mated exclusively to males of their own lineage largely fail to produce workers and cannot successfully found new colonies (Helms Cahan *et al.* 2004).

The four lineages with genetic caste determination have been proposed to derive from a historical hybrid zone of two well-studied species of harvester ants, *Pogonomyrmex rugosus* and *Pogonomyrmex barbatus* (Helms Cahan & Keller 2003, but see Anderson *et al.* 2006). Both parental species have typical environmental caste determination, but the trajectory which led from environmental caste determination in the parental species to genetic caste determination in the hybrid lineages remains speculative (Helms Cahan *et al.* 2002; Julian *et al.* 2002; Helms Cahan & Keller 2003, Linksvayer *et al.* 2006). Despite the evidence for historical gene flow, all four lineages are genetically isolated from both putative parents (Helms Cahan & Keller 2003). A detailed study in a contact zone between *P. rugosus* and the lineages H1 and H2 revealed complete lack of ongoing gene exchange even when colonies of all three types co-occurred at the same site (Helms Cahan *et al.* in press).

The genetic diversity and biogeography of lineages with genetic caste determination can provide important information on when, where and why these lineages became reproductively isolated from the parental species. A recent study investigated the geographical range of genetic caste determination by assessing levels of heterozygosity in workers at an allozyme locus which differs diagnostically between interbreeding lineages in both the H1-H2 and the J1-J2 lineage pairs and is thus always heterozygous in their workers (Anderson *et al.* 2006). This study revealed that, far from being geographically limited, colonies characterized with genetic caste determination occur at a number of other locations over a wide geographical range in southeastern Arizona, southern New Mexico and west Texas (Anderson *et al.* 2006). However, the limited number of molecular markers and the worker-only sampling regime could not identify the genetic composition of individual lineages or the extent of gene flow among sites. There is thus no information available on whether the populations across this range are similar or distinct from the four lineages described at Hidalgo and Junction. Finally, except for Hidalgo and Junction, it is unknown whether populations characterized with genetic caste determination contain two or more genetic lineages.

In this study, we conducted a large-scale genetic survey of sites across southeastern Arizona, southern New Mexico and western Texas to investigate the population genetic structure and biogeography of genetic caste determining lineages. We analysed a total of 35 sites with workers morphologically similar to either *P. rugosus* or *P. barbatus* and determined the mode of caste determination in each colony (genetic or environmental) by genotyping both virgin queens (gynes) and workers at multiple microsatellite loci. We then investigated the genetic structure of populations

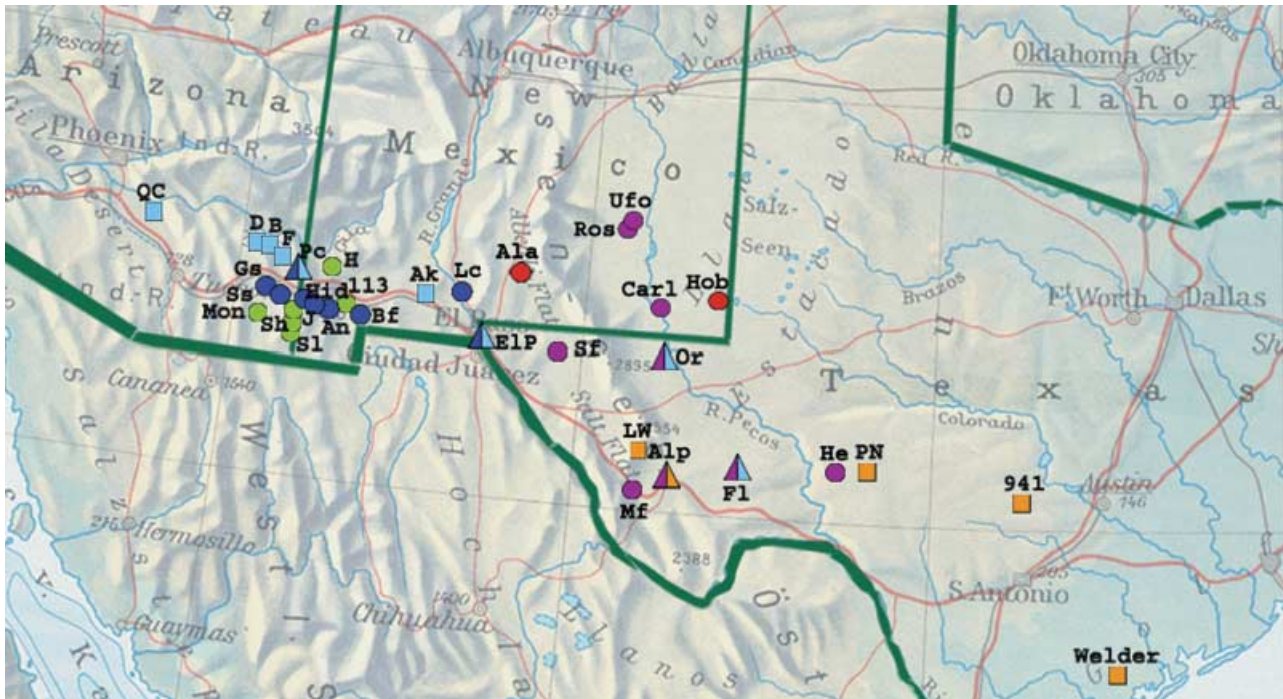
across the region by using microsatellite and mitochondrial DNA data from gynes, whose pure-lineage ancestries allowed us to genetically characterize each lineage separately. We assessed the levels of gene flow occurring between lineages within sites, between sites with the same system of caste determination, and between sites with different systems of caste determination.

## Methods

At each of 31 sites across eastern Arizona and western to central Texas (Fig. 1, for details on sites see Table 1), we collected workers, gynes and males from multiple colonies, ranging from 14 to over 100 colonies per site. Workers of these colonies were visually identified as either *Pogonomyrmex rugosus*-like or *Pogonomyrmex barbatus*-like in appearance. We also included gyne and worker samples from the sites 'Hidalgo' and 'Junction' which are both known to comprise two isolated and interbreeding lineages with genetic caste determination (H1 and H2, and J1 and J2, respectively) and worker samples from one previously characterized site of each of their putative parental species with environmental caste determination [sites 'QC' and '941' for *P. rugosus* and *P. barbatus*, respectively (Helms Cahan *et al.* 2002; Helms Cahan & Keller 2003)].

### Within-site analyses

To assess the mode of caste determination (genetic vs. environmental) at each of the 33 sites with both gyne and worker samples, we checked for the occurrence of four features which are characteristic of the lineages with genetic caste determination described at Junction and Hidalgo. First, we tested whether a given site comprised gynes of different genetic lineages. When caste determination is environmental, gynes should belong to a single genetic group, whereas sites with genetic caste determination should contain gynes of highly differentiated genetic lineages. To test for the presence of genetically distinct groups, one gyne per colony was genotyped at nine microsatellite markers: Pb-5, Pb-7 and Pb-10 (Volny & Gordon 2002a), Pr-1 (Gadua *et al.* 2003), L-18 (Foitzik *et al.* 1997), Myrt-3 (Evans 1993) and PO-3, PO-7 and PO-8 (Wiernasz *et al.* 2004). The number of genetic groups was then determined using maximum-likelihood methods implemented in the population assignment program STRUCTURE 2.1 (Pritchard *et al.* 2000) with a burn-in of 25 000 and run lengths of 50 000 under an admixture ancestry model. This program assesses the likelihood of an observed frequency distribution of genotypes when assuming that individuals come from a given number of genetically differentiated groups. In addition, it computes for each genotype the probability of belonging to each of these groups, with all probabilities summing to 1. Thus, if there



**Fig. 1** Sampling sites from southeastern Arizona to southwestern Texas. Details on sites are given in Table 1. Circles are for sites with colonies in which caste determination is genetic, squares for sites with colonies in which caste determination is environmental, and triangles for sites where both caste determination mechanisms occur at a high frequency. Different colours distinguish sites with *Pogonomyrmex rugosus* (light blue), *Pogonomyrmex barbatus* (orange) and the lineage pairs J1–J2 (green), F1–F2 (purple), G1–G2 (red) and H1–H2 (blue). (See results).

are several well-differentiated genetic groups, the assignment probability of a given individual would be close to 1 for the group to which it belongs and close to 0 for all others. An F1 cross between two groups should yield assignment probabilities close to 0.5 for each of them. For each site, we assessed the likelihood of the observed frequency distribution of gyne genotypes by considering the presence of one to five distinct genetic groups.

Second, we tested for an excess of heterozygosity in workers of each genetic group determined for gynes. Under genetic caste determination as described in *Pogonomyrmex* (Helms Cahan *et al.* 2002; Julian *et al.* 2002; Volny & Gordon 2002b; Helms Cahan & Keller 2003), all workers are F1 interlineage hybrids so that the observed level of heterozygosity greatly exceeds the level expected from the overall worker allele frequencies. By contrast, the observed level of heterozygosity should be in accordance with the expected level under classical environmental caste determination. To estimate the level of heterozygosity in workers, we genotyped one worker per colony at three to nine microsatellite loci. We used only a subset of the loci used to genotype gynes because our aim was not to conduct a population genetic study with workers but to detect deviation from Hardy–Weinberg distributions. For each site, we identified the most informative loci on the basis of the allele frequencies of gynes; we did not use loci which

were fixed for the same allele in both lineages or for which the genetic groups had similar allele frequencies. The observed heterozygosity levels of workers were compared to expected levels by computing the statistic  $F_{IS}$  in the program FSTAT (Goudet 2001).

Third, we tested for genetic differentiation of gynes and their worker nestmates in each genetic group by considering the two castes as two distinct populations and computing the extent of differentiation between them using the statistic  $F_{ST}$ . Under environmental caste determination, there is no genetic differentiation between gynes and workers, thus resulting in low and nonsignificant  $F_{ST}$  values. By contrast, under genetic caste determination as described in *Pogonomyrmex*, workers are interlineage hybrids and should thus exhibit allele frequencies that differ from those of gynes which originate from intralineage matings.  $F_{ST}$  values between workers and gynes from each of the genetic groups present at a given site were computed with the program FSTAT (Goudet 2001).

Finally, to determine the caste determination mechanism in each colony, we directly assessed the lineage ancestries of individual gynes and workers using the allele frequencies of all genetic groups identified at a given site. For each of the genetic groups, we selected the eight gyne genotypes with the highest assignment probabilities (reference samples) to develop a multilocus genetic profile of each

**Table 1** Details on sampling sites from south-eastern Arizona to southwestern Texas

Site	State	Coordinates	Year collected	Morphology*	Male colour (thorax)	Colonies	Genetic groups	Caste determination	Name
Ak	NM	32°14.857N, 107°23.435W	2002	<i>P. rugosus</i>	Red	20	1†	Environmental	<i>P. rugosus</i>
F	AZ	32°17.773N, 109°21.036W	2001/2002	<i>P. rugosus</i>	Red	29	1‡	Environmental	<i>P. rugosus</i>
B	AZ	32°18.913N, 109°29.054W	2004	<i>P. rugosus</i>	Red	60	1	Environmental	<i>P. rugosus</i>
D	AZ	32°22.608N, 109°38.173W	2001	<i>P. rugosus</i>	Red	28	1	Environmental	<i>P. rugosus</i>
QC	AZ	33°15.978N, 111°21.190W	1995	<i>P. rugosus</i>	Red	40§	1	Environmental	<i>P. rugosus</i>
Welder	TX	28°06.801N, 097°20.208W	2004	<i>P. barbatus</i>	Orange	16	1	Environmental	<i>P. barbatus</i>
941	TX	30°12.963N, 098°44.354W	1995	<i>P. barbatus</i>	Orange	20§	1	Environmental	<i>P. barbatus</i>
PN	TX	30°41.463N, 101°16.436W	2004	<i>P. barbatus</i>	Orange	29	1	Environmental	<i>P. barbatus</i>
LW	TX	30°42.291N, 104°06.433W	2004	<i>P. barbatus</i>	Orange	23	1	Environmental	<i>P. barbatus</i>
Hob	TX	32°48.990N, 103°14.250W	2004	<i>P. rugosus</i>	Red or black	28	2	Genetic	G1-G2
Ala	NM	32°55.049N, 105°57.809W	2004	<i>P. rugosus</i>	Red or black	25	2	Genetic	G1-G2
He	TX	30°41.606N, 101°27.610W	2004	<i>P. rugosus</i>	Red or black	31	2	Genetic	F1-F2
Mf	TX	30°19.205N, 104°05.183W	2004	<i>P. rugosus</i>	Red or black	22	2	Genetic	F1-F2
Sf	TX	31°45.820N, 104°55.615W	2004	<i>P. rugosus</i>	Red or black	24	2¶	Genetic	F1-F2
Carl	NM	32°23.391N, 104°13.214W	2004	<i>P. rugosus</i>	Red or black	33	2	Genetic	F1-F2
Ros	NM	33°23.619N, 104°41.240W	2004	<i>P. rugosus</i>	Red or black	41	2	Genetic	F1-F2
Ufo	NM	33°26.000N, 104°39.000W	2002	<i>P. rugosus</i>	Red or black	18	2	Genetic	F1-F2
Lc	NM	32°24.384N, 106°39.318W	2004	<i>P. rugosus</i>	Red or black	31	2	Genetic	H1-H2
Ent	NM	31°59.757N, 109°02.362W	2001	<i>P. rugosus</i>	Red or black	26	2	Genetic	H1-H2
Bf	NM	31°55.696N, 108°22.284W	2002	<i>P. rugosus</i>	Red or black	15	2	Genetic	H1-H2
An	NM	31°57.775N, 108°48.923W	2003	<i>P. rugosus</i>	Red or black	37	2	Genetic	H1-H2
Hid	NM	31°55.121N, 109°02.969W	2003	<i>P. rugosus</i>	Red or black	40§	2	Genetic	H1-H2
Ss	AZ	32°12.887N, 109°10.573W	2001	<i>P. rugosus</i>	Red or black	18	2**	Genetic	H1-H2
Gs	AZ	32°16.214N, 109°14.300W	2001	<i>P. rugosus</i>	Red or black	31	2	Genetic	H1-H2
113	NM	32°13.522N, 108°32.230W	2002	<i>P. barbatus</i>	Orange	18	2	Genetic	J1-J2
J	NM	31°50.725N, 109°02.956W	2001	<i>P. barbatus</i>	Orange	35§	2	Genetic	J1-J2
Mon	AZ	32°00.067N, 109°23.106W	2002	<i>P. barbatus</i>	Orange	19	2	Genetic	J1-J2
H	NM	32°24.971N, 109°39.930W	2002	<i>P. barbatus</i>	Orange	16	2	Genetic	J1-J2
Sh	NM	31°44.618N, 109°05.717W	2002	<i>P. barbatus</i>	Orange	14	2	Genetic	J1-J2
Sl	NM	31°00.001N, 108°58.000W	2002	<i>P. barbatus</i>	Orange	15	2	Genetic	J1-J2
Pc	AZ	32°17.555N, 109°20.235W	2004	<i>P. rugosus</i>	Red or black	141	3	Environmental/genetic	<i>P. rugosus</i> /H1-H2
EIP	TX	31°50.061N, 106°32.704W	2004/2005	<i>P. rugosus</i>	Red or black	59	3	Environmental/genetic	<i>P. rugosus</i> /H1-H2
Fl	TX	30°40.221N, 101°40.571W	2005	<i>P. rugosus</i>	Red or black	31	3	Environmental/genetic	<i>P. rugosus</i> /F1-F2
Or	TX	31°49.393N, 103°54.551W	2005	<i>P. rugosus</i>	Red or black	63	3	Environmental/genetic	<i>P. rugosus</i> /F1-F2
Alp	TX	30°21.678N, 103°38.683W	2005	<i>P. rugosus</i> / <i>P. barbatus</i>	Orange, red or black	50	3	Environmental/genetic	<i>P. barbatus</i> /F1-F2

\*Following Cole (1968).

†One colony in which caste determination was genetic.

‡Two colonies in which caste determination was genetic.

§Samples from reference populations described in Helms Cahan &amp; Keller (2003).

¶Two colonies in which caste determination was environmental.

\*\*One colony in which caste determination was environmental.

group. For all other individuals from the same site, we then calculated the relative assignment probabilities to the reference samples with STRUCTURE 2.1. Under environmental caste determination, nestmate workers and gynes should be assigned to the same reference sample with a probability close to 1. By contrast, in colonies with genetic caste determination gynes should be assigned to one reference sample with a probability close to 1 but workers should be assigned to two different reference samples with equal probability because they are of F1 interlineage ancestry.

In all analyses, we also checked all genotypes individually to avoid assignment errors when there were colonies from very rare groups. Indeed, the algorithm used by STRUCTURE 2.1 is not sensitive to the presence of individuals from a distinct but very rare genetic group as the presence of such individuals would only marginally affect the overall likelihood of the observed frequency distribution of genotypes. Such rare colonies were found at four sites, and in each case, the mode of caste determination differed from that in most other colonies. At two sites, most colonies were characterized by environmental caste determination, but in one colony at one site and in two at the other, caste determination was genetic (see Results, Table 1). Inversely, two sites contained mostly colonies in which caste determination was genetic, but at the first site one colony was characterized by environmental caste determination while at the second, two such colonies were found (see Results, Table 1). These rare colonies were not used for the analyses of gene flow within and between sites because a sample size of one or two individuals would not allow confident calculation of allele frequencies.

#### Among-site analyses

To determine the amount of gene flow within and between the 35 sites, we excluded all F1 hybrids and for all but two sites, we used only gynes. Workers were used for the sites QC and 941 since no gynes were available for these sites and because a previous study revealed that these sites only contained colonies with environmental caste determination, so the caste used should not affect the allele frequencies observed (Helms Cahan *et al.* 2002; Helms Cahan & Keller 2003). All the analyses of genetic structure were based on the same nine microsatellites used to determine the number of genetic groups per site. In addition, for each genetic group at each site, we sequenced a 433-bp portion of the cytochrome *c* oxidase I (COI) mitochondrial sequence of 5–10 individuals as described in Helms Cahan & Keller (2003).

Sites in which multiple genetic groups were detected in the within-site analyses were subdivided according to population assignment probabilities, and each subset was included as a separate genetic group. To determine

the number of distinct lineages with genetic caste determination across all sampled sites, we first computed the pairwise  $F_{ST}$  values between all genetic groups with genetic caste determination. If several genetic groups belonged to the same lineage, the  $F_{ST}$  value between these groups should be small and usually statistically non-significant. In contrast, the pairwise  $F_{ST}$  values between genetic groups belonging to different lineages should be relatively high and significant (Helms Cahan & Keller 2003). We indeed found a bimodal distribution of  $F_{ST}$  values, allowing us to cluster genetic groups belonging to the same lineage and determine the total number of lineages occurring across the region (see Results). The genetic distinctiveness of the lineages identified by this method was then verified by analysing the distribution of genetic variance at nuclear and mitochondrial markers within and between lineages to assess the robustness of the classification.

Finally, we used two methods to investigate whether each identified lineage was of hybrid origin between *P. rugosus* and *P. barbatus* as proposed by (Helms Cahan & Keller 2003). First, we computed for each lineage the proportion of nuclear alleles which were absent in both *P. rugosus* and *P. barbatus*. Second, we quantified more precisely the admixture proportions of each lineage using the maximum-likelihood methods implemented in LEADMIX2 (Wang 2003) with 500 integration points for the likelihood function and accounting for differentiation between parental populations. This program also estimates the extent of genetic drift that occurred to each population since the admixture event relative to effective population sizes (Wang 2003). For example, a computed value of  $10^{-6}$  would be equivalent to the amount of drift that occurred to a population with  $N_e$  of  $10^6$  over one generation. Because Anderson *et al.* (2006) argued that the lineage J2 was the parental population for the H1, H2 and J1 lineages instead of *P. barbatus*, we further compared the relative contributions of *P. barbatus* and J2 to all identified lineages using LEADMIX2. For this, we computed the admixture proportion for each lineage using only *P. barbatus* and J2 as parental populations. If *P. barbatus* was a more likely parental population for the lineages than J2, we expected to find a significantly higher contribution of *P. barbatus* than of J2 to all other lineages. In contrast, the admixture proportion for J2 should be significantly higher than the proportion for *P. barbatus* if J2 was more likely to be the parental population than *P. barbatus*. For each lineage, the contributions of J2 and *P. barbatus* were evaluated by comparing their 95% confidence intervals.

DNA was extracted in 150  $\mu$ L of 5% Chelex (Sigma-Aldrich) at 95 °C for 20 min. Polymerase chain reactions (PCR) were performed as described in (Helms Cahan & Keller 2003) except for PR1 where the hybridization temperature in the PCR cycles was set to 54 °C. Amplified

fluorescent fragments were visualized on 5% polyacrylamide/6 m urea sequencing gels using an automated 377 ABI sequencer. Gels were analysed with GENESCAN version 3.1.2. software (Applied Biosystems).

## Results

### *Within-site analyses*

The analysis of the gyne genotypes at the 33 sites revealed three different categories of sites. In the first category (seven of the 33 sites), maximum-likelihood analysis performed with STRUCTURE 2.1 indicated that gynes belonged to a single genetic group. In the second category (21 of the 33 sites), gynes clustered into two genetically distinct groups. Finally, the third category was composed of sites where gynes belonged to three different genetic groups (five out of 33 sites). None of the analysed sites comprised gynes from more than three genetic groups.

In the seven sites where all gynes belonged to a single genetic group, the vast majority of the colonies (202 out of 205, Table 1) were characterized by an environmental system of caste determination, as indicated by the small and nonsignificant levels of genetic differentiation between gynes and workers ( $F_{ST}$  -0.01 to 0.02 all  $P > 0.2$ ) and by the levels of heterozygosities in workers which were as predicted by their allele frequencies ( $F_{IS}$  -0.04 to 0.03, all  $P > 0.2$ ). Gynes from three remaining colonies were most likely from a genetically distinct group with genetic caste determination that could not be detected by STRUCTURE 2.1 because of small sample size. The genetic profiles of the three gynes and their worker nestmates were very similar to the genetic profiles of their counterparts from the closest site with colonies characterized by genetic caste determination. In addition, these gynes had between four and seven alleles that were absent in gynes from other colonies at their site but present at other sites characterized by genetic caste determination.

In the second category of sites ( $n = 21$ ), where gynes belonged to two different genetic groups, caste determination was genetic in the vast majority of colonies (534 out of 537, Table 1). Genetic caste determination for both genetic groups was revealed by significant genetic differentiation between gynes and workers ( $F_{ST}$  0.19–0.30, all  $P < 0.05$ ) and the significant excess of heterozygosity in workers ( $F_{IS}$  -0.15 to -0.43, all  $P < 0.05$ ). Workers from virtually all colonies (532 out of 534) were F1 interlineage hybrids between the two genetic groups occurring at the site as indicated by similar assignment probabilities to the two genetic groups obtained with STRUCTURE 2.1 (0.44–0.56) and confirmed by an inspection of individual genotypes. The two remaining workers were of pure-lineage ancestry. Five additional workers sampled and genotyped from each of the two colonies were all of

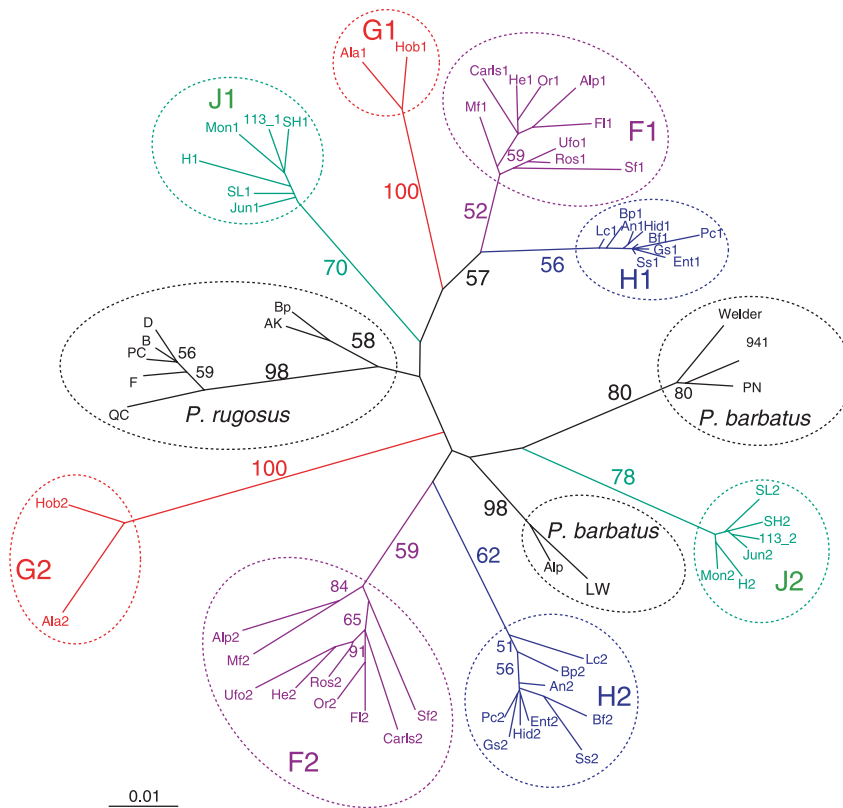
interlineage ancestry, indicating that although a low proportion of pure-lineage individuals apparently developed into workers, caste determination in these two colonies was largely genetic. Gynes from three remaining of the 537 colonies were most likely from a genetically distinct group with environmental caste determination that could not be detected due to small sample size. The genetic profiles of the three gynes and their worker nestmates were similar and corresponded to the genetic profile of individuals from the closest site with colonies characterized by environmental caste determination. These gynes had between 9 and 10 alleles which were absent in the gynes collected from the same site, but present in gynes from colonies characterized by environmental caste determination collected at other sites. None of the sites from the second category comprised two different groups of gynes from colonies with an environmental system of caste determination.

In the third and last category of sites ( $n = 5$ ), gynes belonged to three distinct genetic groups. The first group comprised gynes from 147 out of the 344 colonies (43%) and all these colonies were characterized by an environmental system of caste determination (Table 1), as revealed by the small and nonsignificant levels of genetic differentiation between gynes and workers ( $F_{ST}$  -0.01–0.03 all  $P > 0.2$ ) and by levels of worker heterozygosities in close agreement with the values predicted from the allele frequencies ( $F_{IS}$  -0.06–0.04, all  $P > 0.2$ ). Moreover, all the workers from the 147 colonies with environmental caste determination were unambiguously assigned by STRUCTURE 2.1 to the same genetic group as their nestmate gynes (assignment probabilities 0.82–0.99). The two other genetic groups of gynes (from 197 out of the 344 colonies; 57%) were both characterized by genetic caste determination as revealed by the high genetic differentiation between gynes and workers ( $F_{ST}$  0.18–0.35, all  $P < 0.05$ ) and the significant excess of heterozygosity in workers ( $F_{IS}$  -0.15 to -0.43, all  $P < 0.05$ ). All the workers from these 197 colonies were assigned with similar probabilities to the two lineages with genetic caste determination identified at their site (0.41–0.59), indicating that they were F1 hybrids between the two genetic groups. We did not find a single F1 hybrid between groups with environmental and genetic caste determination out of the 688 individuals analysed.

### *Among-site analyses*

#### *Sites characterized by environmental caste determination*

Overall, 412 colonies from 16 different sites were characterized by an environmental system of caste determination [including colonies from the sites QC and 941 from Helms Cahan & Keller (2003)]. To quantify the level of gene flow among sites, we used gynes from the 14 sites (409 colonies)



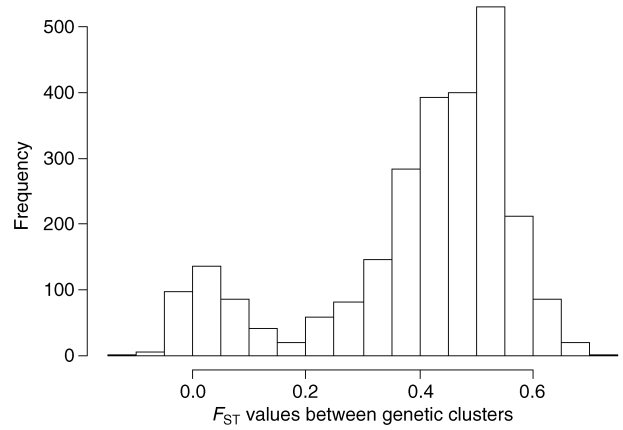
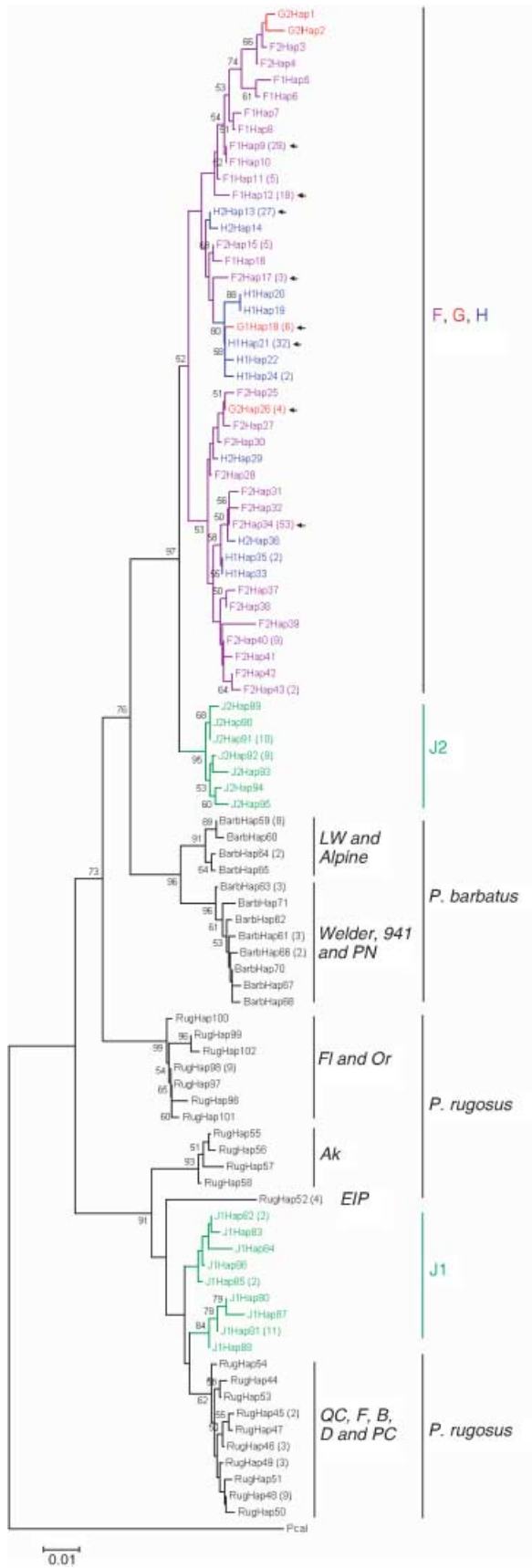
**Fig. 2** Neighbour-joining tree based on Cavalli-Sforza genetic distance (Cavalli-Sforza & Edwards 1967) computed with the program POPULATIONS (Langella 1999). Numbers on nodes indicate bootstrap percentages (1000 replicates on loci). Bootstrap values smaller than 50% are not indicated. Samples from the same geographical site have the same prefix but a different index for each genetic group (e.g. Alp, Alp1 and Alp2, which are all colonies sampled at the site Alp but with environmental caste determination for Alp, and colonies with genetic caste determination of lineage F1 and lineage F2 for Alp1 and Alp2, respectively). Lineages belonging to the same lineage pair have the same prefix (F, G, H or J) and are of the same colour.

which included at least eight colonies with environmental caste determination. These 14 sites comprised seven sites where caste determination was environmental in almost all colonies, five sites with a large number of colonies with both systems of caste determination, and the two previously characterized sites (QC and 941). Visual inspection of workers revealed that five of these sites comprised colonies of *Pogonomyrmex barbatus* while the nine others were *Pogonomyrmex rugosus* (Table 1). The degree of among-site genetic differentiation at nuclear and mitochondrial markers was low to moderate for *P. barbatus* sites (pairwise  $F_{ST}$  based on microsatellites: 0.01–0.15, Fig. 2; maximum haplotype divergence between sites: 0.04, Fig. 3). For seven of the nine *P. rugosus* sites, we also found low to intermediate levels of genetic differentiation between sites at the nuclear markers (pairwise  $F_{ST}$  values: 0.002–0.16, Fig. 2) and low levels for mitochondrial haplotypes (divergence between sites 0.001–0.05). For the two remaining sites (Orla and FL), the nuclear genotypes strongly suggested introgression from *P. barbatus* (data not shown). The ambiguous nuclear composition of these sites made it impossible to put them onto the genetic distance tree at a particular location with any degree of confidence. The mitochondrial haplotypes of these two sites also clustered into a distinct clade differing by 6% from other *P. rugosus* sites and *P. barbatus* sites (Fig. 3).

*Sites characterized by genetic caste determination*

Overall, 734 colonies from 28 sites were characterized by a genetic system of caste determination. To analyse levels of gene flow within and between sites, we used only the 26 sites (731 colonies) which had at least 14 such colonies and a minimum of four per genetic group. These comprised 21 sites where caste determination was genetic in almost all colonies, and five sites with a large number of colonies with each having a system of caste determination. The majority of gynes from these colonies (675 out of 731) were of pure-lineage ancestry (assignment probabilities 0.87–0.99). Fifty-two of the 56 remaining gynes were F1 interlineage hybrids (0–12 per site, assignment probabilities to both lineages 0.45–0.56). Finally, four gynes were assigned to one of the two genetic groups found at their site with a relatively low assignment probability (0.68–0.78) and their genotypes suggested some level of backcrossing. To determine the number of independent genetic lineages across all sites, we used the 679 gynes which were not F1 interlineage hybrids. As there were 26 sites each with two genetic groups with genetic caste determination, we included a total of 52 genetic groups.

*Classification and description of lineages with genetic caste determination* An analysis of the pairwise  $F_{ST}$  values



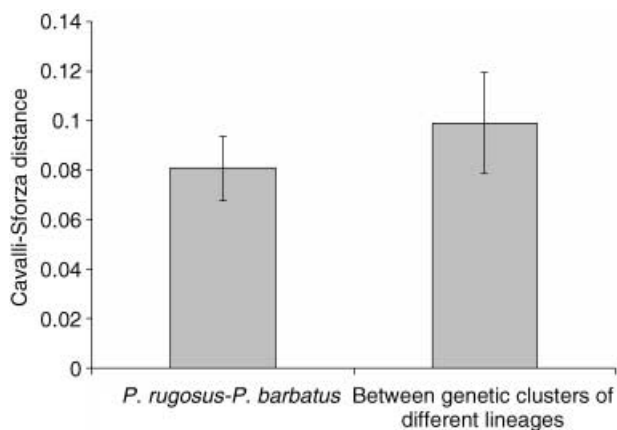
**Fig. 4** Bimodal frequency distribution of pairwise  $F_{ST}$  values between genetic groups of lineages with genetic caste determination within and between sites.

between the 52 groups with genetic caste determination from the 26 sites revealed two classes of pairwise distances: one class of relatively small and nonsignificant values ( $0.04 \pm 0.06$ ) and a class of relatively large and significant values ( $0.46 \pm 0.09$ ) (Fig. 4). Such a bimodal frequency distribution of pairwise  $F_{ST}$  values indicates the presence of several genetically differentiated lineages, each composed of genetic groups from different sites. Based on this bimodal frequency distribution, we identified a total of eight highly differentiated genetic lineages. The large majority (> 99%) of the pairwise  $F_{ST}$  values exceeding the intermode minimum (0.20) were between genetic groups belonging to different lineages. Inversely, 98% of the pairwise  $F_{ST}$  values smaller than the intermode minimum were between genetic groups of the same lineage. In accordance with the bimodal distribution of the pairwise  $F_{ST}$  values, an AMOVA [ARLEQUIN 2.000 (Schneider *et al.* 2000)] revealed that 60.2% of genetic variance at nuclear markers was among the eight lineages, while the variance between genetic groups belonging to the same lineage represented only 2.6% of the variance.

Two lines of evidence further support the existence of eight lineages. First, a general analysis of Cavalli-Sforza chord distances of the microsatellite data from all genetic

**Fig. 3** Kimura 2-parameter distance tree of a 430-bp portion of the mitochondrial sequence COI, computed with the program MEGA (Kumar *et al.* 2004), with *Pogonomyrmex californicus* as outgroup. Numbers at nodes represent bootstrap values (1000 replicates). Bootstrap values smaller than 50% are not indicated. Haplotypes sampled in the F lineages are in purple, haplotypes sampled in the G lineages in red, and haplotypes sampled in the H lineages in blue. The numbers in brackets after the haplotype names indicate the frequency of the haplotype. Arrows indicate haplotypes that are shared between several lineages, shared haplotypes have the colour of the most represented lineage.





**Fig. 5** Cavalli-Sforza genetic distances between sites of *Pogonomyrmex rugosus* and *Pogonomyrmex barbatus* and between genetic groups of different lineages with genetic caste determination (within and between sites).

groups revealed the presence of eight genetically differentiated clusters with genetic caste determination in the neighbour-joining tree (Fig. 2). These eight groups corresponded exactly to the eight lineages identified with the pairwise  $F_{ST}$  frequency distribution analysis. At nuclear markers, the level of differentiation between lineages even exceeded the differentiation between the two taxonomically well-defined species *P. rugosus* and *P. barbatus* as indicated by significantly larger Cavalli-Sforza chord distances between genetic groups of different lineages than between *P. rugosus* and *P. barbatus* populations (Wilcoxon tests on random subsamples of pairwise distances involving each population only once:  $W = 364.5-402$ ,  $P = 0.026-0.001$ , Fig. 5). The higher levels of differentiation were not due to different overall levels of genetic diversity, as Cavalli-Sforza chord distances are not affected by levels of observed heterozygosity (Cavalli-Sforza & Edwards 1967).

Second, the analysis of genetic variance at mitochondrial haplotypes also supported the sorting of genetic groups into the eight lineages identified with nuclear markers. Overall, 66.7% of the genetic variance was between lineages and only 14.5% of the variance was between genetic groups within lineages [including distances between haplotypes for computation; ARLEQUIN 2.000 (Schneider *et al.* 2000)]. The level of variance was relatively low between genetic groups within lineages mostly because genetic groups of the same lineage frequently shared identical haplotypes absent in other lineages (Fig. 3). Only eight out of the 42 different haplotypes were found in more than one lineage (Fig. 3) and five of them were very common in one lineage (haplotype frequency:  $40.7 \pm 11.5\%$ ) and rare in the other lineages.

The eight genetic lineages invariably occurred in the same lineage pairs across all 26 sites. There were thus four pairs of interbreeding lineages. Each of the four pairs will

be referred to by a capital letter (F, G, H and J) with the numbers 1 and 2 distinguishing the co-occurring lineages. One of these lineage pairs, J1 and J2, has been previously described at Junction and was found at six different sites across eastern Arizona and western New Mexico (Fig. 1, Table 1). Workers, gynes and males in colonies headed by queens of the lineages J1 and J2 were similar in appearance to their *P. barbatus* counterparts. Another lineage pair, H1 and H2, was previously described at Hidalgo (Helms Cahan & Keller 2003). The H1-H2 lineage pair occurred at nine sites in a range similar to that of the J1-J2 lineage pair (Fig. 1, Table 1). The two remaining lineage pairs (F and G) have not been described before. The F1-F2 lineage pair was found at nine sites across eastern New Mexico and western to central Texas while the G1-G2 lineage pair only occurred at two sites in central New Mexico (Fig. 1). In all six of these lineages (F, G and H), individuals were similar in appearance to *P. rugosus* with the exception of males of the lineages F2, G2 and H2, which had a black thorax, while *P. rugosus* males as well as males of all the other lineages had a red thorax (Table 1).

*Origin of lineages with genetic caste determination* Our data corroborate the idea that the lineages and hence genetic caste determination have evolved in small populations after complex hybridization events between *P. rugosus* and *P. barbatus* (Helms Cahan & Keller 2003). Each lineage at each site (52 genetic groups) exhibited a lower allelic diversity than the 14 populations of *P. rugosus* and *P. barbatus* (allelic richness at microsatellites: 1000 permutations,  $P = 0.001$ ). Similarly, the diversity of mitochondrial haplotypes was lower in the populations of the lineages than for *P. rugosus* and *P. barbatus* (Wilcoxon test on haplotype diversity:  $W = 364.5$ ,  $P = 0.026$ ). The program LEADMIX2 also estimated much higher levels of genetic drift at nuclear markers in each of the eight lineages than in *P. rugosus* and *P. barbatus* (lineages:  $0.134 \pm 0.036$ , *P. rugosus*:  $0.002 \pm 0.001$ , *P. barbatus*:  $0.007 \pm 0.002$ ). Only a small proportion of nuclear alleles ( $8.3 \pm 2.0\%$ , Appendix) present in any of the eight lineages did not occur in at least one of the seven pure *P. rugosus* or five *P. barbatus* sites. The admixture proportions of each of the lineages estimated with LEADMIX2 revealed a higher contribution of *P. rugosus* than *P. barbatus* to the lineages F1, G1, H1 and J1 and an almost exclusive contribution of *P. barbatus* to the lineages F2, G2, H2 and J2 (Table 2). They also suggest that *P. barbatus* is much more likely to be a parental population for the lineages F, G, H and J1 than is J2. Indeed, the nuclear contribution of J2 to the other lineages was virtually absent and always significantly smaller than the contribution estimated for *P. barbatus* as indicated by nonoverlapping 95% confidence intervals (Table 2).

Despite the evidence for the hybrid origin of the lineages, the nuclear genetic distances between each of the

**Table 2** Admixture proportions with their 95% confidence intervals estimated with LEADMIX2. For each lineage, we estimated the contribution of *Pogonomyrmex rugosus* and *Pogonomyrmex barbatus* (left side of the table) and compared the contribution of *P. barbatus* relative to the J2 lineage (right side of the table)

Lineage	<i>P. rugosus</i>	+95% CI	-95% CI	<i>P. barbatus</i>	+95% CI	-95% CI	<i>P. barbatus</i>	+95% CI	-95% CI	J2	+95% CI	-95% CI
F1	0.741	0.742	0.741	0.259	0.259	0.258	0.999	1.000	0.980	0.001	0.020	0.000
F2	0.040	0.169	0.029	0.960	0.971	0.831	0.999	1.000	0.953	0.001	0.047	0.000
G1	0.749	0.824	0.749	0.251	0.465	0.176	0.987	1.000	0.870	0.013	0.130	0.000
G2	0.005	0.198	0.044	0.995	0.956	0.802	0.999	1.000	0.807	0.001	0.193	0.000
H1	0.754	0.901	0.754	0.246	0.469	0.099	0.899	0.999	0.744	0.101	0.256	0.001
H2	0.002	0.040	0.020	0.998	0.980	0.960	0.854	0.999	0.736	0.146	0.264	0.001
J1	0.745	0.863	0.745	0.255	0.452	0.137	0.939	1.000	0.789	0.061	0.211	0.000
J2	0.005	0.081	< 0.001	0.995	> 0.999	0.919	—	—	—	—	—	—

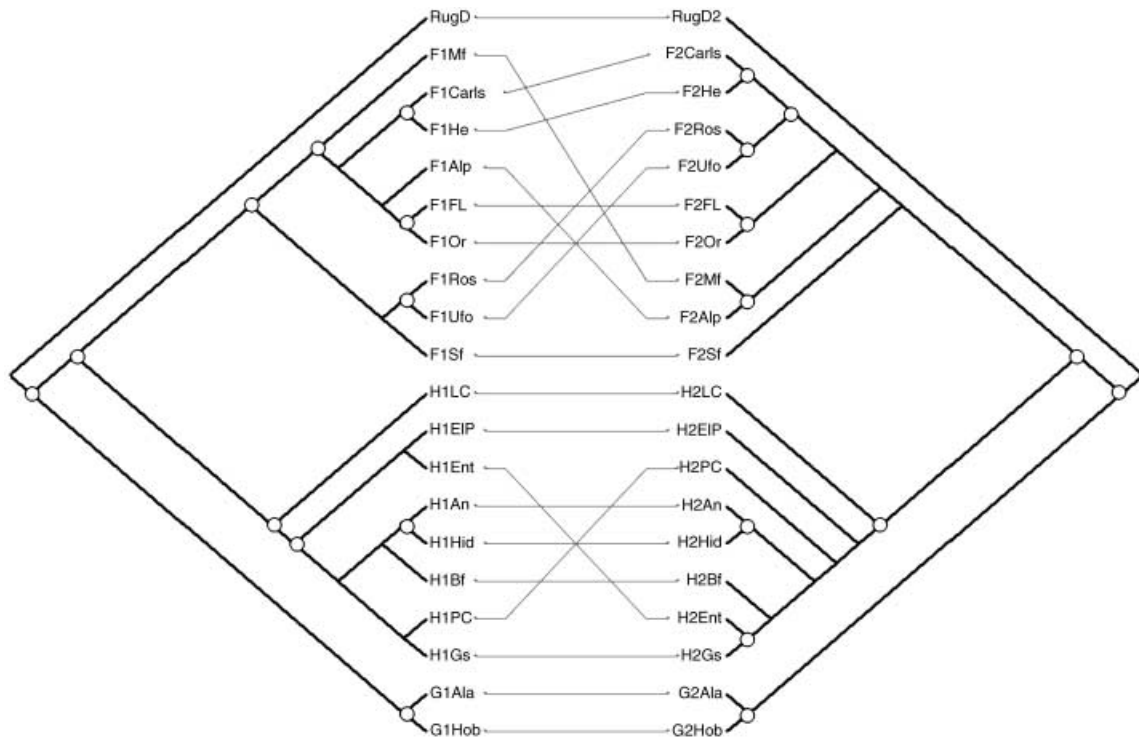
eight lineages and *P. rugosus* or *P. barbatus* were invariably high (Cavalli-Sforza distance on nuclear markers: 0.05–0.12, mean = 0.09, Fig. 2). Diagnostic alleles occurred at seven of the nine microsatellite loci (Appendix), although the vast majority of the alleles ( $92.0 \pm 6.7\%$ , range: 78.1–100.0%, Appendix) present in any of the lineages also occurred in either *P. rugosus* or *P. barbatus*. None of the lineages shared any mitochondrial haplotypes with either *P. rugosus* or *P. barbatus* (Fig. 3). One lineage (J1) had haplotypes that were similar (2.0% differentiation) to *P. rugosus* haplotypes. In contrast, the seven other lineages had haplotypes that were very divergent from those found in *P. barbatus* (differentiation of 4.2–4.9%, Fig. 3) and even more from *P. rugosus* (differentiation of 7.4%–8.1%).

In addition to the morphological similarities between queens, males and workers of the three lineages F1, G1 and H1, and the lineages H2, G2 and F2, three lines of evidence suggest that these three lineage pairs are historically linked. First, as mentioned before, lineages always occurred in the same lineage pairs across sites. Second, mitochondrial haplotypes of all six lineages form a single clade, suggesting they have a common ancestor (Fig. 6). Finally, when each set of lineages with similar male types were considered separately, the genetic distance tree of the 'set 1' lineages (F1, G1 and H1) was very similar to the genetic tree of the 'set 2' lineages (F2, G2 and H2). Overall, there were at least 11 events of parallel differentiation (Fig. 6), a pattern unlikely to occur by chance [probability for 11 or more parallel differentiations for random trees generated with the program TREEMAP 2.0b (Charleston & Page 2002) < 0.0001]. All instances of differentiation not occurring in parallel on both trees were between genetic clusters of the same lineage which were invariably characterized by a low level of genetic differentiation and thus had low power to resolve within-lineage relationships (Fig. 6). The length of the corresponding branches in both trees were also highly correlated ( $r = 0.80$ ,  $P = 0.0001$ ), further supporting the view of a single origin of the lineage

pairs H1-H2, G1-G2 and F1-F2. As for the J1-J2 lineage pair, the resolution of the genetic markers and the number of populations sampled do not allow us determine whether or not it had a common origin with the three other lineage pairs. At the nuclear markers, the grouping of the lineage J1 with the lineages F1, G1 and H1 is not robust (bootstrap values < 50%, Fig. 2). The lineage J2 is nested within *P. barbatus*, and is thus not particularly associated with any of the other lineages (Fig. 2). The J1 haplotypes were highly differentiated from the haplotypes found in all other lineages (Fig. 3), while the haplotypes of J2 had a basal position in the haplotype clade of the lineages F, G and H (Fig. 3). Any scenario linking the evolution of J1 and J2 with the evolution of the other lineage pairs would thus be highly speculative.

There were three systematic differences between the F1, G1 and H1 lineages and the F2, G2 and H2 lineages. First, the branches of the distance trees were systematically shorter in the F1, G1, H1 tree than the corresponding branches in the F2, G2 and H2 tree (Wilcoxon signed rank test:  $V = 111$ ,  $P = 0.011$ ). Second, the genetic groups in the F1, G1 and H1 lineages had a lower level of genetic diversity at microsatellite markers than the genetic groups in the F2, G2 and H2 lineages (paired Wilcoxon test on allelic richness per site:  $V = 28$ ,  $P = 0.005$ ). Finally, the F1, G1 and H1 lineages had fewer 'novel' alleles, absent in both *P. barbatus* and *P. rugosus*, than the F2, G2 and H2 lineages (paired Wilcoxon test:  $V = 42$ ,  $P < 0.0001$ ).

Several lines of evidence suggest that the lineage pairs G1-G2 and H1-H2 went through more severe bottlenecks than the lineage pair F1-F2 and hence that the allele distribution in the latter pair is more similar to the ancestral lineage pair than the lineage pairs G1-G2 and H1-H2. First, when considering only the three lineages F1, G1 and H1, the lineage F1 had significantly more private alleles per microsatellite locus than the two other lineages [Kruskal-Wallis test on number of specific alleles per marker, corrected for sample size as described in (Petit *et al.* 1998):



**Fig. 6** Matching of independent Cavalli-Sforza distance trees for sites with the lineages H1, G1, F1, respectively, F2, G2, and H2 computed with TREEMAP (Charleston & Page 2002). Both trees are rooted with a randomly chosen site of *Pogonomyrmex rugosus* (site D). Branch lengths are not indicated. Samples from the same site have the same prefix. Circles indicate events of parallel differentiation; lines connect interbreeding lineages collected at the same site.

$X_{22} = 8.30$ ,  $P = 0.02$ , posthoc  $P < 0.05$ ]. Similarly, when considering the lineages F2, G2 and H2 only, the lineage F2 had significantly more private alleles than the two other lineages ( $X_{22} = 5.87$ ,  $P = 0.05$ , posthoc  $P = 0.05$ ). Second, the pattern of allele sharing between the F lineages and the other two lineage pairs was not symmetrical. The majority of nuclear alleles sampled in the lineages G1 (79%) and H1 (90%) were also found in the lineage F1, while a significantly smaller proportion of alleles sampled in the lineage F1 also occurred in the lineages G1 (44%) and H1 (53%) (Fig. 7a, paired  $t$ -tests with Bonferroni corrections: F1 vs. G1:  $t_8 = -6.04$ ,  $P = 0.001$ ; F1 vs. H1:  $t_8 = -6.13$ ,  $P = 0.001$ ). Similarly, the majority of the nuclear G2 (64%) and H2 (82%) alleles also occurred in F2, whereas significantly fewer alleles sampled in F2 also occurred in the lineages G2 (21%) and H2 (61%) (paired  $t$ -tests with Bonferroni corrections: F2 alleles found in G2 vs. G2 alleles found in F2:  $t_8 = -4.51$ ,  $P = 0.004$ , F2 alleles found in H2 vs. H2 alleles found in F2:  $t_8 = -2.91$ ,  $P = 0.040$ , Fig. 7b). These results were not due to differences in sample sizes across lineages as we obtained qualitatively the same results when using random subsets of sites of the lineages F and H to equal the sample size of the rarer G lineages (lineages of set 1: all tests  $P < 0.05$ , lineages of set 2: all tests  $P < 0.05$ ). A similar

pattern was found for mitochondrial haplotypes, although the small numbers of different haplotypes per lineage precluded a statistical comparison. The lineage F1 had more unique haplotypes ( $n = 7$ ) than the lineages G1 ( $n = 0$ ) and H1 ( $n = 4$ , Fig. 7c) and F2 had more unique haplotypes ( $n = 18$ ) than G2 ( $n = 4$ ) and H2 ( $n = 5$ , Fig. 7d). A high proportion of haplotypes found in the lineages G1 (100%) and H1 (56%) also occurred in F1, while the proportion of F1 haplotypes that also occurred in G1 and H1 were only 31% and 38%, respectively (Fig. 7c). Similarly, the proportions of G2 (20%) or H2 (29%) haplotypes which also occurred in F2 were higher than the proportions of F2 haplotypes which also occurred in G2 (5%) or H2 (10%, Fig. 7d).

## Discussion

Kin selection predicts that caste determination should result from environmentally induced differences in gene expression (Seeger 1981), a prediction largely supported by empirical data (Wheeler 1986). The occurrence of genetic caste determination in two distinct populations of *Pogonomyrmex* harvester ants in southwestern New Mexico (Helms Cahan *et al.* 2002; Julian *et al.* 2002; Volny & Gordon 2002b) has thus raised questions about how general it is in

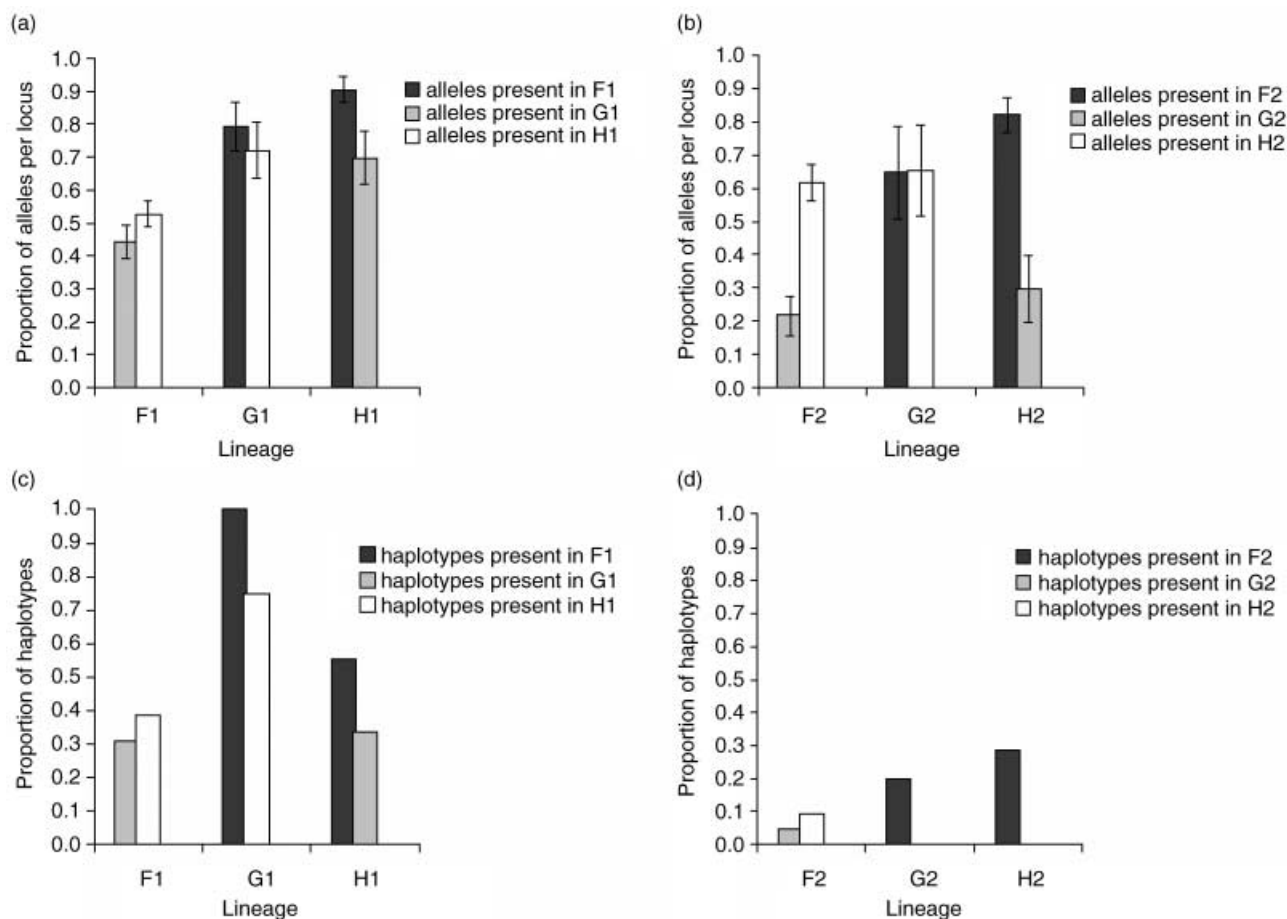


Fig. 7 The proportion of alleles sampled within a given lineage also occurring the other lineages of the same set for nuclear (a and b) and mitochondrial markers (c and d).

this genus, where, under which conditions, and how many times it has evolved.

The results of this study indicate that genetic caste determination is widespread in populations with workers morphologically identified as *Pogonomyrmex rugosus* or *Pogonomyrmex barbatus* across southeastern Arizona, southern New Mexico and west Texas. Sixty-four per cent of 1145 colonies sampled across 35 sites were characterized by genetic caste determination, while only 36% of colonies in the surveyed range were characterized by classical, environmentally determined caste differentiation. The species with genetic and environmental caste determination seem to be ecological competitors, as at the majority of sites (30 out of 35) we found a single or a highly dominant caste determination system. Colonies of both types co-occurred at high frequencies only at five out of the 35 sites (Table 1). The majority of sites characterized by environmental caste determination were mainly at the eastern (*P. rugosus*) and western (*P. barbatus*) extremes of the sampled range (Fig. 1). These patterns corroborate the results from Anderson *et al.* (2006) which also included populations

outside the range surveyed in the present study. From Texas to western Arizona, there is a east–west climatic gradient with mean annual precipitation varying from over 50 cm in central Texas to as little as 2 to 5 cm in western Arizona. Although we did not survey the entire distribution range described for *P. rugosus* and *P. barbatus*, this pattern suggests possible ecological differences between populations with genetic and environmental caste determination with the former occurring in more intermediate habitats relative to the latter. The abundance and wide distribution of the genetic caste determination mechanism in *Pogonomyrmex* is surprising because this mechanism of caste determination has been shown to impose significant costs for colony survival (Helms Cahan *et al.* 2004) and productivity (Schwander *et al.* 2006), and its potential advantages remain speculative Julian & Helms Cahan 2006.

The ancestry-based pattern of genetic caste determination was remarkably uniform across the entire studied range, suggesting similar underlying mechanisms for determination of the queen and worker developmental

pathways in these populations. Genetic caste determination was always associated with the interbreeding of two distinct lineages; we never found more than two lineages involved at a given site. The level of genetic divergence between two co-occurring lineages was comparable to divergence at the interspecific level (e.g. Morjan & Rieseberg 2004), and the association of genotype and caste, whereby intralinesage females develop into queens and interlinesage females develop into workers, was very consistent.

Despite the uniformity of the mode of genetic caste determination across populations, the populations themselves were not genetically homogeneous. Previous studies had identified two different pairs of lineages, H1-H2 and J1-J2, which could be distinguished morphologically as well as genetically (Helms Cahan *et al.* 2002; Julian *et al.* 2002; Helms Cahan & Keller 2003). Our broader geographical survey reveals the presence of additional lineage pairs within morphologically similar groups of populations. Both nuclear and mitochondrial analyses of gynes reveal four distinct pairs of interbreeding lineages with genetic caste determination. Three (F1-F2, G1-G2, and H1-H2) are similar to *P. rugosus* in body colour and morphologically distinct from the fourth lineage pair, J1-J2, which resembles *P. barbatus* in colour. The J1-J2 lineage pair was found in the most western part of the surveyed range along with the H1-H2 lineage pair (Fig. 1). The F1-F2 lineage pair has the widest distribution, as it occurs in eastern New Mexico and west Texas; the G1-G2 lineage pair is geographically restricted in southeastern New Mexico (Fig. 1). The identification of multiple lineage pairs within a colour morph is not simply a case of isolation by distance, as some lineage pairs are located in close geographical proximity yet are strongly differentiated at both classes of genetic marker (Fig. 1).

Several lines of evidence suggest that the F, G, and H lineages are derived from a single ancestral lineage pair. In addition to their overall morphological similarity, there is a consistent colour difference between the males of the two interbreeding lineages in each case. The lineages of set 1 (F1, G1, and H1) possess all-red males ['red males' (Helms Cahan *et al.* 2002)] while the lineages of set 2 (F2, G2, and H2) have males with a black-and-red pattern similar to gynes and workers ['black males' (Helms Cahan *et al.* 2002)]. The two sets of lineages with alternate male types show parallel patterns of nuclear differentiation (Fig. 6), and all six lineages cluster in a single mitochondrial clade (Fig. 3). The distribution of nuclear alleles among the lineages F, G, and H further suggests that these lineages derived from a single ancestral lineage pair which was most similar to the F1-F2 lineage pair. In both the lineages of set 1 and of set 2, the G and H lineages were significantly less diverse, contained significantly fewer private alleles, and the majority of their alleles were a subsample of alleles found in the corresponding F lineage (Fig. 7a, b). Taken together, these results suggest that gen-

etic caste determination may have originated in eastern New Mexico or west Texas, and small founder populations subsequently radiated west, with both members of the pair comigrating as a unit because genetic caste determination cannot function in the absence of continual interbreeding. The relative distributions of mitochondrial haplotypes among the F, G, and H lineages were similar to the pattern for nuclear alleles but were less evident because overall sharing of haplotypes among these lineages was very low (Fig. 7c, d). Each lineage was largely characterized by a few common haplotypes, the identities of which differed across lineages. However, none of the lineages F, G, or H formed statistically well-supported, monophyletic groups. This can be partially explained by the low levels of informative variation within this group (1.5% divergence), which limits the power to resolve evolutionary relationships with any certainty. If the apparent intermixing of haplotypes does reflect their real relationships, two mechanisms are possible. First, the lineages may have not achieved reciprocal monophyly and have retained the original ancestral haplotypes. Alternatively, the intermixed relationships and the few incidences of sharing of identical haplotypes across lineages may be evidence for gene flow, although the distribution of haplotypes suggests that genetic exchange between lineages, if it occurred, is only historical and not ongoing. All haplotypes shared between two lineages were never found in both lineages at the same site, as would be expected if introgression were currently occurring, and were not even always found in currently interbreeding lineages (e.g. F1 and G2, Haplotype 9). The lack of current gene flow is further supported by the extremely low incidence of backcross genotypes and the nuclear genetic distances, which are significantly higher between interbreeding lineages than between the species *P. rugosus* and *P. barbatus*. Overall, this suggests that genetic isolation between lineages is virtually complete in this system. All eight lineages with genetic caste determination should thus be considered as independent species or an independent species complex rather than as variants of *P. rugosus* or *P. barbatus*.

Two hypotheses have been proposed to explain the evolution of genetic caste determination in *Pogonomyrmex*. The first hypothesis suggests that the lineage pairs are derived from a historical hybrid zone between *P. rugosus* and *P. barbatus* (Helms Cahan & Keller 2003). Inter-specific hybridization and high levels of genetic drift resulted in the fixation of two alternate sets of genetically incompatible alleles that cannot trigger worker development independently. The second hypothesis states that genetic caste determination first evolved in *P. barbatus*, giving rise to an ancestral J2 lineage (and possibly additional cryptic lineages) that then spread genetic caste determination into *P. rugosus* through introgressive hybridization (Anderson *et al.* 2006).

The results from the present study are more in accordance with the first hypothesis. First, as previously noted for the H1-H2 and the J1-J2 lineage pair (Helms Cahan & Keller 2003), only a small proportion of nuclear alleles sampled in each of the eight lineages do not occur in *P. rugosus* or in *P. barbatus*. Using additional microsatellite markers and analysing additional populations and lineages in the present study did not alter this general pattern. For the lineages F1, G1, H1 and J1, maximum likelihood analyses strongly support the admixture of *P. rugosus* and *P. barbatus* (Table 2). The contribution of *P. rugosus* to the lineages F2, G2, H2, and J2 is less evident, but the *P. rugosus*-like morphology of individuals from the lineages F2, G2 and H2 also indicates their hybrid origin. Strong evidence for hybrid origin is only lacking for the J2 lineage (see also Anderson *et al.* 2006). Together, these results suggest that the lineages with genetic caste determination have evolved within an introgressive hybrid zone between *P. rugosus* and *P. barbatus*. Second, populations characterized by genetic caste determination have significantly lower levels of genetic diversity than the populations characterized by environmental caste determination at both nuclear and mitochondrial markers, indicating that they went through periods of low population sizes. The second hypothesis would predict that the source lineage (J2) is the genetically most diverse lineage and also more similar to other lineages than populations of *P. barbatus*. J2 might also be composed of additional cryptic lineages or co-occur with more than one recipient lineage, and possibly without an obvious second lineage (Anderson *et al.* 2006). Neither of these patterns was observed. We found no evidence for differences in genetic diversity between J2 and the other lineages with genetic caste determination. J2 is also unlikely to be the parental population for the F, G and H, and J1 lineages instead of other *P. barbatus* populations, as revealed by maximum-likelihood analyses. The lineages always occurred in fixed pairs and the presence of one of the eight genetic lineages, including the J2 lineage, invariably predicted the second co-occurring lineage. Finally, there was no evidence for the J2 lineage comprising multiple cryptic lineages.

As the two lineages of each lineage pair are completely interdependent, they should have evolved in parallel. However, in the lineage pair J1-J2, haplotypes found in the J1 lineage are less differentiated from *P. rugosus* haplotypes than the haplotypes found in the J2 lineage are differentiated from *P. barbatus* haplotypes. This thus suggests that the currently known *P. barbatus* populations are rather distinct from the *P. barbatus* population that historically contributed to the lineages with genetic caste determination. Anderson *et al.* (2006) documented a *P. barbatus* haplotype in Mexico which was similar to the F, G, H and J2 haplotypes. The population where this haplotype was sampled would thus be a likely candidate for the original parental *P. barbatus* population. Unfortunately however, no nuclear

data are yet available for this population so that the caste determination mechanism remains speculative. The contribution of an unsampled population to the lineages F2, G2 and H2 would also explain why these lineages have more alleles which were not present in the *P. rugosus* or *P. barbatus* populations sampled in this study than their co-occurring lineages, F1, G1, and H1. Admixture analyses at nuclear markers show that *P. barbatus* contributed more to the lineages of set 2 than to the lineages of set 1. As a consequence, if the novel alleles were contributed by an unsampled population of *P. barbatus*, they should occur predominantly in set 2 lineages. We also found that the lineages of set 2 had higher overall levels of genetic diversity at nuclear markers and were more differentiated across sites than the lineages of set 1. The origin of these differences remains enigmatic, but it might indicate that the lineages from set 2 became reproductively isolated from the parental species after lineages from set 1. Lineages of set 2 at different sites might thus have received alleles from different isolated parental populations.

In both *P. rugosus* and *P. barbatus*, we found evidence for relatively high levels of genetic substructure, which, in combination with results from other studies (Cole & Wiernasz 1997; Strehl & Gadau 2004), suggests that genetic structuring and local dispersal tend to be common in the genus *Pogonomyrmex*. The surveyed region is an area of patchy habitat, which may additionally restrict genetic exchange between populations. Such patchy habitat might be a prerequisite for the independent evolution of hybridizing populations and the evolution of lineages with genetic caste determination because the homogenizing effects of gene flow from parental species are interrupted (Buerkle *et al.* 2000).

The existence of two *P. rugosus* populations that show evidence of introgression from *P. barbatus* without displaying genetic caste determination is interesting, as it suggests that hybridization between the two species does not always lead to this type of system. The level of genetic diversity in these two populations was comparable to the levels found in populations of *P. rugosus* or *P. barbatus* which do not have a signature of interspecific hybridization. This suggests that these populations experienced little historical bottleneck, which could impact the probability of the evolution of genetic caste determination. Continual backcrossing with dispersing parental forms from allopatric populations might never allow for incompatibilities to appear. Strong skew in initial frequencies of hybridizing populations may also influence the likelihood that a genetic caste determination system might evolve. If one species (e.g. *P. rugosus*) were numerically very dominant, hybridization would have little overall impact on its genetic composition, while the rarer species (e.g. some variant of *P. barbatus*) may be swamped by hybridization and be lost as an independent entity.

In conclusion, this study shows that genetic caste determination is a widespread and consistent mechanism in *Pogonomyrmex*. It occurs in a genetically heterogeneous group of lineages, which apparently evolved after complex hybridization events between *P. barbatus* and *P. rugosus*. The current lack of gene flow among the lineages and between the lineages and the two parental species suggests that the differences in the caste determination mechanism may persist indefinitely and that the taxonomic status of this species complex may have to be reconsidered.

### Acknowledgements

We would like to thank the Earthwatch Institute's Student Challenge Awards Program teams from 2001 to 2003 for their tireless efforts collecting samples from across the desert southwest. We also thank G.E. Julian and J.-Y. Humbert for great help in the field and J.D. Parker for allowing us to use samples from his personal collection. J. Wang provided valuable information on the optimal use of LEADMIX2. This study was supported by several grants from the Swiss NSF to L.K., by a grant of the SANW to T.S., and by grants from the Earthwatch Institute to S.H.C. and G. Julian.

### References

- Anderson KE, Gadau J, Mott BM *et al.* (2006) Distribution and evolutionary history of genetic caste determination in the seed-harvester ants *Pogonomyrmex barbatus* and *P. rugosus*. *Ecology*, **87**, 2171–2184.
- Brian MV (1957) Caste determination in social insects. *Annual Review of Entomology*, **2**, 107–120.
- Buerkle CA, Morris JM, Asmusse MA, Rieseberg LH (2000) The likelihood of homoploid hybrid speciation. *Heredity*, **84**, 441.
- Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, **32**, 550–570.
- Charleston MA, Page RE (2002) TREEMAP 2.0B: A Macintosh® program for the analysis of how dependent phylogenies are related, by cophylogeny mapping. <http://taxonomy.zoology.gla.ac.uk/rod/treemap.html>.
- Cole AC (1968) *Pogonomyrmex Harvester Ants; a Study of the Genus in North America*. University of Tennessee Press, Knoxville, Tennessee.
- Cole BJ, Wiernasz DC (1997) Inbreeding in a lek-mating ant species, *Pogonomyrmex occidentalis*. *Behavioral Ecology and Sociobiology*, **40**, 79–86.
- Evans JD (1993) Parentage analyses in ant colonies using simple sequence repeat loci. *Molecular Ecology*, **2**, 393–397.
- Foitzik S, Haberl M, Gadau J, Heinze J (1997) Mating frequency of *Leptothorax nylanderi* ant queens determined by microsatellite analysis. *Insectes Sociaux*, **44**, 219–227.
- Gadau J, Strehl C-P, Oettler J, Hölldobler B (2003) Determinants of intracolony relatedness in *Pogonomyrmex rugosus* (Hymenoptera; Formicidae): mating frequency and brood raids. *Molecular Ecology*, **12**, 1931–1938.
- Goudet J (2001) *ESTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices Version 2.9.3*. Available from <http://www.unil.ch/izea/software/fstat.html>.
- Helms Cahan S, Parker JD, Rissing SW *et al.* (2002) Extreme genetic differences between queens and workers in hybridizing *Pogonomyrmex* harvester ants. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **269**, 1871–1877.
- Helms Cahan S, Julian GE, Rissing SW *et al.* (2004) Loss of phenotypic plasticity generates genotype–caste association in harvester ants. *Current Biology*, **14**, 2277–2282.
- Helms Cahan S, Julian GE, Schwander T, Keller L (2006) Reproductive isolation between the harvester ant *Pogonomyrmex rugosus* and two lineages of hybrid origin (H1/H2). *Ecology*, **87**, 2143–2147.
- Helms Cahan S, Keller L (2003) Complex hybrid origin of genetic caste determination in harvester ants. *Nature*, **424**, 306–309.
- Hölldobler B, Wilson EO (1990) *The Ants*. Springer-Verlag, Berlin.
- Julian GE, Fewell JH, Gadau J, Johnson RA, Larrabee D (2002) Genetic determination of the queen caste in an ant hybrid zone. *Proceedings of the National Academy of Sciences, USA*, **99**, 8157–8160.
- Julian GE, Helms Cahan S (2006) Behavioral differences between hybridogenetic and parental *Pogonomyrmex* harvester ants. *Ecology*, **87**, 2207–2214.
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.
- Langella O (1999) POPULATIONS 1.2.28. Available from <http://www.cnrs-gif.fr/pge/bioinfo/populations/index.php>.
- Linksvayer TA, Wade MJ, Gordon DM (2006) Genetic caste determination in harvester ants: possible origin and maintenance by cyto-nuclear epistasis. *Ecology*, **87**, 2185–2193.
- Morjan CL, Rieseberg LH (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology*, **13**, 1341–1356.
- Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN, version 2.000: a software for population genetics data analysis. *Genetics and Biometry Laboratory*. Department of Anthropology, University of Geneva.
- Schwander T, Helms Cahan S, Keller L (2006) Genetic caste determination in *Pogonomyrmex* harvester ants imposes costs during colony founding. *Journal of Evolutionary Biology*, **19**, 402–409.
- Seger J (1981) Kinship and covariance. *Journal of Theoretical Biology*, **91**, 191–213.
- Strehl CP, Gadau J (2004) Cladistic analysis of Paleo-island populations of the Florida harvester ant (Hymenoptera: Formicidae) based upon divergence of mitochondrial DNA sequences. *Florida Entomologist*, **87**, 576–581.
- Volny VP, Gordon DM (2002a) Characterization of polymorphic microsatellite loci in the red harvester ant, *Pogonomyrmex barbatus*. *Molecular Ecology Notes*, **2**, 302–303.
- Volny VP, Gordon DM (2002b) Genetic basis for queen-worker dimorphism in a social insect. *Proceedings of the National Academy of Sciences, USA*, **99**, 6108–6111.
- Volny VP, Green MJ, Gordon DM (2006) Brood production and lineage discrimination in a two-lineage population of the red harvester ant *Pogonomyrmex barbatus*. *Ecology*, **87**, 2194–2200.
- Wang J (2003) Maximum likelihood estimation of admixture proportions from genetic data. *Genetics*, **164**, 747–765.
- Wheeler DE (1986) Developmental and physiological determinants of caste in social hymenoptera — evolutionary implications. *American Naturalist*, **128**, 13–34.

Wheeler DE (1994) Nourishment in ant societies: patterns in individuals and societies. In: *Nourishment and the Evolution in Insect Societies*. (eds Hunt JH, Nalepa CA), Westview Press, Boulder, Colorado.

Wiernasz DC, Perroni CL, Cole BJ (2004) Polyandry and fitness in the western harvester ant, *Pogonomyrmex occidentalis*. *Molecular Ecology*, **13**, 1601–1606.

---

This work is part of T. Schwander's PdD thesis on the evolution and consequences of genetic caste determination in *Pogonomyrmex* harvester ants. S. Helms Cahan's research focuses on the consequences of sociality on the evolution of individual, population, and species characteristics. L. Keller works on various aspects of evolutionary ecology and social behaviour in ants.

---



Appendix

Allele frequencies at nine microsatellite loci in the eight lineages of *Pogonomyrmex* displaying genetic caste determination, as well as *P. rugosus* and *P. barbatus*. Microsatellite allele numbers indicate size as measured on an ABI 377 automated sequencer. Sample sizes are indicated for each locus in italics under the allele frequencies, samples are pooled across sites.

Locus	Allele	<i>P. barbatus</i>	<i>P. rugosus</i>	<i>P. rugosus</i> (Or, Fl)	F1	F2	G1	G2	H1	H2	J1	J2
L-18	137	0.004	—	—	—	—	—	—	—	—	—	—
	143	0.004	—	—	—	—	—	—	—	—	—	—
	144	—	0.002	—	—	—	—	—	—	—	—	—
	145	—	—	—	—	0.005	—	—	—	—	—	—
	147	0.013	—	—	—	0.005	—	—	—	—	—	—
	148	—	0.002	—	—	—	—	—	—	—	—	—
	150	0.013	—	—	—	—	—	—	—	—	—	—
	152	0.143	0.002	0.048	—	—	—	—	—	—	—	—
	153	—	—	—	—	—	—	—	—	0.003	—	—
	154	0.109	—	0.119	—	—	—	—	—	—	—	—
	155	0.057	—	—	—	—	—	—	—	—	—	—
	156	0.057	0.004	0.024	—	—	—	0.806	—	—	—	0.031
	157	0.030	0.004	—	—	—	—	—	—	—	—	—
	158	0.026	0.015	0.238	—	—	—	0.111	—	—	—	0.016
	159	0.017	0.002	—	—	—	—	—	—	—	—	0.008
	160	0.022	0.006	0.024	—	—	—	0.083	—	—	—	—
	161	0.013	—	—	—	—	—	—	—	0.003	—	0.117
	162	0.004	0.021	0.095	0.100	—	—	—	—	—	—	—
	163	0.052	—	—	—	0.041	—	—	—	0.003	—	0.289
	164	0.022	0.015	0.024	0.030	—	—	—	0.385	—	—	—
	165	0.043	—	—	—	0.021	—	—	—	0.031	—	0.148
	166	—	0.008	—	0.004	—	—	—	0.095	—	—	—
	167	0.026	0.002	—	—	0.010	—	—	—	0.306	—	0.102
	168	0.009	0.082	0.119	0.065	0.010	—	—	0.460	—	0.200	—
	169	0.104	—	—	—	0.062	—	—	—	0.065	—	0.195
	170	—	0.148	0.119	0.004	—	0.036	—	0.035	—	—	—
	171	0.030	—	—	0.004	0.046	—	—	—	0.017	—	0.07
	172	0.004	0.084	—	0.013	—	0.071	—	0.005	—	—	—
	173	0.039	—	—	—	0.041	—	—	—	0.020	—	—
	174	0.004	0.110	0.024	0.009	—	0.268	—	—	—	—	—
	175	0.017	0.002	—	—	0.015	—	—	—	—	—	0.008
	176	0.004	0.057	—	0.030	—	0.107	—	—	—	—	—
	177	0.048	—	—	—	0.026	—	—	—	0.003	—	—
	178	0.009	0.044	—	0.100	—	0.250	—	—	—	—	—
	179	—	0.002	—	—	0.046	—	—	—	0.014	—	0.008
	180	—	0.078	0.024	0.096	—	0.071	—	—	—	0.150	—
	181	0.035	—	—	—	0.057	—	—	—	0.031	0.017	—
	182	—	0.091	—	0.052	—	0.071	—	—	—	0.100	—
	183	0.009	—	—	—	0.155	—	—	—	0.065	—	—
	184	0.004	0.072	—	0.061	—	0.054	—	0.005	—	0.050	—
	185	0.004	—	—	—	0.155	—	—	—	0.099	—	—
	186	—	0.038	0.048	0.078	0.015	—	—	0.010	—	0.150	—
	187	0.009	—	—	—	0.077	—	—	—	0.126	—	0.008
	188	0.004	0.027	0.048	0.07	—	—	—	—	—	0.200	—
	189	—	—	—	—	0.098	—	—	—	0.122	—	—
	190	0.004	0.025	0.024	0.083	—	—	—	—	—	0.100	—
	191	—	—	—	—	0.015	—	—	—	0.051	—	—
	192	—	0.032	0.024	0.052	—	0.018	—	0.005	—	0.017	—
	193	—	0.002	—	—	0.036	—	—	—	0.027	—	—
	194	0.004	0.011	—	0.078	—	0.054	—	—	—	—	—
	195	—	—	—	—	0.041	—	—	—	0.003	—	—

## Appendix Continued

Locus	Allele	<i>P. barbatus</i>	<i>P. rugosus</i>	<i>P. rugosus</i> (Or, F1)	F1	F2	G1	G2	H1	H2	J1	J2
	196	—	0.006	—	0.052	—	—	—	—	—	0.017	—
	197	—	—	—	—	0.015	—	—	—	0.007	—	—
	198	—	0.004	—	0.013	—	—	—	—	—	—	—
	199	—	—	—	—	—	—	—	—	0.003	—	—
	200	—	—	—	0.004	—	—	—	—	—	—	—
	201	—	—	—	—	0.005	—	—	—	—	—	—
	<i>N</i>	115	237	21	115	97	28	18	100	147	30	64
Pr-1												
	370	—	—	—	—	—	—	—	—	—	0.016	—
	374	0.004	—	—	—	—	—	—	—	—	—	—
	376	0.004	0.01	—	—	—	—	—	—	—	—	—
	378	0.009	—	—	—	—	—	—	—	—	—	—
	382	0.155	—	—	0.008	0.005	—	—	—	—	—	—
	383	—	0.004	—	—	—	—	—	—	—	—	—
	384	0.004	—	—	—	—	—	—	—	—	—	—
	386	0.039	—	—	—	—	—	—	—	—	—	—
	388	0.228	0.323	0.119	0.004	—	—	—	—	—	0.891	—
	392	0.009	—	—	—	—	—	—	—	—	—	—
	394	0.142	0.191	0.024	0.864	—	0.733	0.083	0.757	0.007	0.063	—
	396	—	0.002	0.167	—	—	—	—	—	—	—	—
	398	0.004	—	—	—	—	—	—	—	—	—	—
	399	0.013	—	—	—	—	—	—	—	—	—	1.000
	400	0.004	0.173	—	0.120	0.035	0.250	—	0.233	0.003	0.031	—
	402	—	0.002	—	—	—	—	—	—	—	—	—
	406	0.306	0.165	0.095	—	0.015	0.017	0.917	0.010	0.877	—	—
	408	—	0.002	—	—	—	—	—	—	0.003	—	—
	412	0.009	0.082	0.310	—	0.005	—	—	—	0.003	—	—
	418	0.004	0.032	0.048	—	0.187	—	—	—	—	—	—
	424	0.017	0.004	0.167	—	0.010	—	—	—	—	—	—
	426	—	—	0.024	—	—	—	—	—	—	—	—
	430	—	0.004	0.024	—	0.404	—	—	—	—	—	—
	436	0.026	0.006	—	—	0.066	—	—	—	0.106	—	—
	442	—	—	—	—	0.111	—	—	—	—	—	—
	444	—	—	0.024	—	—	—	—	—	—	—	—
	448	0.013	—	—	0.004	0.056	—	—	—	—	—	—
	454	—	—	—	—	0.045	—	—	—	—	—	—
	456	—	—	—	—	0.030	—	—	—	—	—	—
	465	0.004	—	—	—	—	—	—	—	—	—	—
	466	0.004	—	—	—	0.025	—	—	—	—	—	—
	480	—	—	—	—	0.005	—	—	—	—	—	—
	<i>N</i>	116	251	21	121	99	30	18	103	151	32	68
Myrt-3												
	173	0.004	—	—	—	—	—	—	—	—	—	—
	177	—	0.002	0.024	—	—	—	—	—	—	—	—
	179	—	0.008	—	—	—	—	—	—	—	—	—
	181	0.232	0.838	0.881	0.996	0.894	1.000	1.000	1.000	1.000	0.500	—
	183	0.601	0.002	0.048	0.004	0.106	—	—	—	—	—	1.000
	185	0.145	0.006	0.048	—	—	—	—	—	—	—	—
	187	0.009	—	—	—	—	—	—	—	—	—	—
	189	0.009	0.145	—	—	—	—	—	—	—	0.141	—
	191	—	—	—	—	—	—	—	—	—	0.359	—
	<i>N</i>	114	262	21	123	99	30	18	105	159	32	68

Appendix Continued

Locus	Allele	<i>P. barbatus</i>	<i>P. rugosus</i>	<i>P. rugosus</i> (Or, Fl)	F1	F2	G1	G2	H1	H2	J1	J2
Pb-10												
	259	—	—	—	0.061	0.005	—	—	—	—	—	—
	262	0.005	—	—	—	—	—	—	—	—	—	—
	265	0.005	0.065	0.024	—	—	—	—	—	—	—	—
	268	0.027	0.036	0.024	0.906	—	0.967	—	0.841	0.003	—	—
	271	0.345	0.405	0.286	0.025	0.965	—	—	0.159	0.707	0.031	0.963
	274	0.236	0.352	0.357	—	0.030	0.017	1.000	—	0.003	0.906	0.037
	277	0.345	0.137	0.262	—	—	0.017	—	—	0.287	0.063	—
	280	0.027	0.002	0.048	—	—	—	—	—	—	—	—
	283	0.009	—	—	—	—	—	—	—	—	—	—
	289	—	0.002	—	0.008	—	—	—	—	—	—	—
	N	110	237	21	122	100	30	18	104	150	32	68
Pb-5												
	208	—	—	0.167	—	—	—	—	—	—	—	—
	210	0.034	—	—	—	—	—	—	—	—	—	—
	212	0.005	—	—	—	—	—	—	—	—	—	—
	214	0.038	—	0.024	—	—	—	—	—	0.089	—	—
	216	0.192	—	—	0.004	0.020	—	—	—	0.041	—	0.015
	218	0.154	—	0.310	—	0.010	—	—	—	—	—	—
	220	0.168	—	0.024	—	0.010	—	—	—	—	—	—
	222	0.106	—	0.024	—	0.040	—	0.972	—	0.006	—	0.015
	224	0.038	0.062	0.048	—	0.066	—	—	—	0.029	—	0.100
	226	0.024	0.004	—	—	0.258	—	—	—	0.809	—	0.031
	228	0.053	0.006	—	0.004	0.505	0.100	0.028	0.010	0.022	—	0.092
	230	0.005	0.006	—	0.004	0.056	—	—	0.014	—	0.016	0.015
	232	0.034	0.020	—	0.975	0.010	0.020	—	0.967	0.003	0.688	0.600
	234	0.005	0.120	0.048	0.004	0.015	—	—	0.005	—	—	0.015
	236	0.034	0.016	—	0.008	—	0.340	—	—	—	0.016	0.062
	238	0.034	0.384	—	—	—	0.540	—	0.005	—	—	0.023
	240	0.014	0.199	0.048	—	—	—	—	—	—	0.016	0.023
	242	0.005	0.062	0.262	—	—	—	—	—	—	0.250	—
	244	0.005	0.088	0.048	—	0.005	—	—	—	—	0.016	0.008
	246	0.005	0.024	—	—	0.005	—	—	—	—	—	—
	248	0.014	0.002	—	—	—	—	—	—	—	—	—
	250	0.005	0.006	—	—	—	—	—	—	—	—	—
	251	0.005	—	—	—	—	—	—	—	—	—	—
	252	0.005	—	—	—	—	—	—	—	—	—	—
	254	0.014	—	—	—	—	—	—	—	—	—	—
	266	0.005	—	—	—	—	—	—	—	—	—	—
	N	104	249	21	120	99	25	18	105	157	32	65
Pb-7												
	139	—	—	—	—	0.005	—	—	—	0.003	—	—
	141	—	0.522	—	—	0.005	—	—	—	0.003	—	—
	143	0.151	0.169	0.119	0.223	0.318	—	—	—	0.962	—	—
	145	0.014	0.105	0.048	0.013	0.172	—	—	—	0.006	—	—
	146	0.018	0.004	—	0.021	—	—	—	—	—	—	—
	147	0.009	—	—	—	0.010	—	—	—	—	—	—
	148	0.009	0.135	0.214	0.05	0.465	—	—	—	0.025	0.063	—
	150	0.128	0.016	0.381	0.088	0.010	0.017	—	0.005	—	—	—
	151	0.009	—	—	—	—	—	—	—	—	—	—
	152	0.115	0.008	0.095	0.055	—	0.103	—	0.714	—	0.094	0.082
	153	0.028	—	—	—	—	—	—	—	—	—	—
	154	0.142	0.018	0.024	0.416	0.015	0.466	—	0.029	—	0.703	0.455
	156	0.110	—	0.119	0.050	—	0.276	—	0.248	—	0.109	0.022
	157	0.073	—	—	—	—	—	—	—	—	—	—

## Appendix Continued

Locus	Allele	<i>P. barbatus</i>	<i>P. rugosus</i>	<i>P. rugosus</i> (Or, F1)	F1	F2	G1	G2	H1	H2	J1	J2
	158	0.037	—	—	0.004	—	0.138	—	0.005	—	—	0.007
	159	0.005	—	—	—	—	—	—	—	—	—	—
	160	0.009	—	—	0.008	—	—	—	—	—	0.016	0.007
	162	—	—	—	0.050	—	—	0.139	—	—	—	0.022
	163	0.005	—	—	—	—	—	—	—	—	—	—
	164	0.005	—	—	0.017	—	—	0.028	—	—	—	—
	165	0.005	—	—	—	—	—	—	—	—	—	—
	166	0.005	—	—	—	—	—	0.028	—	—	—	0.007
	168	—	—	—	0.004	—	—	0.028	—	—	0.016	0.134
	169	0.009	—	—	—	—	—	—	—	—	—	—
	170	0.023	—	—	—	—	—	0.500	—	—	—	0.060
	171	0.005	—	—	—	—	—	—	—	—	—	—
	172	0.005	—	—	—	—	—	0.028	—	—	—	0.075
	174	0.018	—	—	—	—	—	—	—	—	—	0.015
	176	0.014	0.020	—	—	—	—	—	—	—	—	—
	178	0.005	0.004	—	—	—	—	—	—	—	—	0.007
	180	0.014	—	—	—	—	—	0.056	—	—	—	—
	182	0.009	—	—	—	—	—	—	—	—	—	0.030
	184	0.005	—	—	—	—	—	—	—	—	—	0.052
	186	—	—	—	—	—	—	0.056	—	—	—	0.015
	188	0.009	—	—	—	—	—	0.056	—	—	—	0.007
	190	0.005	—	—	—	—	—	0.083	—	—	—	—
	194	0.005	—	—	—	—	—	—	—	—	—	—
	<i>N</i>	109	252	21	119	99	29	18	103	157	32	67
PO-3												
	143	0.005	—	0.025	—	—	—	—	—	—	—	—
	145	—	0.032	—	—	0.005	—	—	—	—	0.019	—
	147	—	0.038	—	—	—	—	—	—	—	—	—
	149	0.009	0.096	—	0.004	—	—	—	—	—	0.019	—
	151	—	0.034	—	—	—	—	—	—	—	—	—
	153	0.027	0.034	0.050	—	0.026	—	—	—	—	—	—
	155	0.095	0.062	0.075	—	—	—	0.083	—	0.028	—	—
	156	—	—	0.075	—	—	—	—	—	—	—	—
	157	0.095	0.032	0.025	0.013	0.010	0.379	0.028	—	0.014	—	0.108
	159	0.050	0.058	0.100	0.089	—	0.155	0.028	0.006	—	0.019	0.058
	161	0.081	0.030	0.025	0.025	0.005	0.069	0.083	0.048	—	—	0.167
	163	0.045	0.014	—	0.085	0.062	0.034	0.167	0.217	—	—	0.40
	165	0.081	0.026	0.025	0.174	0.124	0.017	0.111	0.488	0.052	0.130	0.158
	167	0.054	0.096	0.275	0.153	0.093	—	0.056	0.133	0.339	0.407	0.033
	169	0.077	0.127	0.125	0.195	0.113	—	0.333	0.072	0.133	0.222	0.008
	171	0.050	0.074	0.075	0.114	0.057	—	0.083	0.024	0.038	0.111	0.008
	173	0.077	0.072	0.050	0.064	0.057	0.017	—	0.006	0.049	0.037	0.033
	175	0.041	0.028	0.050	0.034	0.098	0.103	—	0.006	0.140	0.037	0.008
	177	0.050	0.072	0.025	0.017	0.088	0.103	—	—	0.035	—	0.008
	179	0.063	0.040	—	0.021	0.057	0.052	—	—	0.017	—	0.008
	181	0.023	0.022	—	0.013	0.046	0.052	—	—	0.014	—	—
	183	0.045	0.010	—	—	0.031	0.017	—	—	0.010	—	—
	185	0.014	—	—	—	0.021	—	—	—	0.056	—	—
	187	0.023	—	—	—	0.010	—	—	—	0.024	—	—
	189	—	—	—	—	0.062	—	0.028	—	0.021	—	—
	191	—	—	—	—	0.026	—	—	—	0.021	—	—
	193	—	—	—	—	—	—	—	—	0.003	—	—
	195	—	—	—	—	0.010	—	—	—	0.003	—	—
	<i>N</i>	111	249	20	118	97	29	18	83	143	27	60

Appendix Continued

Locus	Allele	<i>P. barbatus</i>	<i>P. rugosus</i>	<i>P. rugosus</i> (Or, Fl)	F1	F2	G1	G2	H1	H2	J1	J2
PO-8												
	213	—	0.002	—	—	—	—	—	—	—	—	—
	216	—	0.002	—	—	—	—	—	—	—	—	—
	219	0.005	0.012	—	—	—	—	—	—	—	—	—
	225	0.023	—	—	—	—	—	—	—	—	—	—
	228	0.131	0.004	0.100	—	—	—	—	—	—	0.019	—
	231	0.037	—	0.025	—	—	—	0.306	—	—	0.019	—
	234	0.093	0.002	0.150	0.008	0.660	—	—	0.016	0.205	—	0.008
	237	0.150	0.064	0.100	—	0.072	—	0.694	0.479	0.010	0.148	0.034
	240	0.145	0.116	0.150	0.012	0.232	—	—	—	0.064	—	0.619
	243	0.019	0.010	0.050	—	0.005	—	—	—	0.708	—	0.297
	246	0.042	0.134	—	0.004	—	0.067	—	0.011	—	—	0.025
	249	0.047	0.312	0.250	0.004	0.005	0.917	—	0.059	0.007	0.815	—
	251	—	0.004	—	—	—	—	—	—	—	—	—
	252	0.248	0.040	0.175	—	0.005	—	—	—	0.003	—	—
	255	0.014	0.164	—	—	—	—	—	—	—	—	0.017
	257	—	0.002	—	—	—	—	—	—	—	—	—
	258	0.009	0.024	—	0.033	—	—	—	0.005	—	—	—
	261	0.019	0.050	—	0.930	0.015	—	—	0.431	—	—	—
	263	—	—	—	0.004	—	—	—	—	—	—	—
	264	0.005	0.054	—	0.004	0.005	—	—	—	—	—	—
	267	—	0.004	—	—	—	—	—	—	0.003	—	—
	270	—	—	—	—	—	0.017	—	—	—	—	—
	285	0.005	—	—	—	—	—	—	—	—	—	—
	286	0.009	—	—	—	—	—	—	—	—	—	—
	N	107	250	20	121	97	30	18	94	149	27	59
PO-7												
	215	—	0.004	—	—	—	—	—	—	—	—	—
	221	—	—	—	—	—	0.034	—	—	—	—	—
—	233	—	—	—	0.004	—	—	—	—	—	—	—
	236	0.030	0.009	—	—	—	—	—	—	—	—	—
	239	0.025	0.009	—	—	—	—	—	0.005	—	0.019	0.017
	242	0.183	0.038	0.025	0.013	0.005	0.017	—	0.005	0.007	—	—
	245	0.233	0.353	0.500	0.059	0.189	0.069	—	0.121	—	0.426	0.595
	248	0.124	0.271	0.075	0.114	0.250	0.672	—	0.838	0.863	0.463	0.336
	249	—	0.002	—	—	—	—	—	—	—	—	—
	251	0.193	0.075	0.050	0.042	0.184	0.207	—	0.025	0.033	0.019	0.052
	254	0.084	0.12	0.250	0.568	0.122	—	0.028	0.005	0.003	0.074	—
	257	0.025	0.049	0.075	0.195	0.219	—	—	—	0.016	—	—
	260	0.064	0.045	—	0.004	0.020	—	0.889	—	0.065	—	—
	263	0.025	0.024	0.025	—	0.005	—	0.083	—	—	—	—
	266	—	—	—	—	—	—	—	—	0.010	—	—
	269	0.005	—	—	—	—	—	—	—	0.003	—	—
	272	0.005	0.002	—	—	—	—	—	—	—	—	—
	275	0.005	—	—	—	—	—	—	—	—	—	—
	278	—	—	—	—	0.005	—	—	—	—	—	—
	N	101	234	20	118	98	29	18	99	153	27	58