Ecology, 87(9), 2006, pp. 2171–2184

DISTRIBUTION AND EVOLUTION OF GENETIC CASTE DETERMINATION IN *POGONOMYRMEX* SEED-HARVESTER ANTS

KIRK E. ANDERSON,[1,3] JÜRGEN GADAU,[1,2] BRENDON M. MOTT,[1] ROBERT A. JOHNSON,[1] ANNETTE ALTAMIRANO,[1] CHRISTOPH STREHL,[2] AND JENNIFER H. FEWELL[1]

[1]*Department of Biology, Arizona State University, Tempe, Arizona 85287 USA*
[2]*Department of Behavioral Physiology and Sociobiology, Biozentrum, University of Würzburg, Würzburg D-97074, Germany*

Abstract. We examined the distribution and ancestral relationships of genetic caste determination (GCD) in 46 populations of the seed-harvester ants *Pogonomyrmex barbatus* and *P. rugosus* across the east-to-west range of their distributions. Using a mtDNA sequence and two nuclear markers diagnostic for GCD, we distinguished three classes of population phenotypes: those with GCD, no evidence of GCD, and mixed (both GCD and non-GCD colonies present). The GCD phenotype was geographically widespread across the range of both morphospecies, occurring in 20 of 46 sampled populations. Molecular data suggest three reproductively isolated and cryptic lineages within each morphospecies, and no present hybridization. Mapping the GCD phenotype onto a mtDNA phylogeny indicates that GCD in *P. rugosus* was acquired from *P. barbatus*, suggesting that interspecific hybridization may not be the causal agent of GCD, but may simply provide an avenue for GCD to spread from one species (or subspecies) to another. We hypothesize that the origin of GCD involved a genetic mutation with a major effect on caste determination. This mutation generates genetic conflict and results in the partitioning and maintenance of distinct allele (or gene set) combinations that confer differences in developmental caste fate. The outcome is two dependent lineages within each population; inter-lineage matings produce workers, while intra-lineage matings produce reproductives. Both lineages are needed to produce a caste-functional colony, resulting in two reproductively isolated yet interdependent lineages. *Pogonomyrmex* populations composed of dependent lineages provide a unique opportunity to investigate genetic variation underlying phenotypic plasticity and its impact on the evolution of social structure.

Key words: caste; cryptic species; dependent lineages; genetic caste determination; genetic conflict; heterozygosity; negative frequency dependent selection; Pogonomyrmex.

INTRODUCTION

Eusociality is viewed as cooperation among genetically related individuals with associated sterility in some or most colony members (Hamilton 1964). The reproductive strategies of many eusocial insects leave them prone to colony-level conflict, because most members of a colony are sterile and spend their lives helping another individual reproduce, rather than producing their own progeny (Hamilton 1964). The dominant accepted mechanism for differentiation of group members into reproductive and sterile castes is caste polyphenism, a developmentally plastic process of caste determination in which similar genotypes can develop into discrete phenotypes that lack intermediates (Nijhout 1994, 1999). In the social Hymenoptera, each female embryo theoretically has the genetic potential to develop either into a reproductive queen or a sterile worker, and does so according to environmental cues, either nutritional or hormonal (Nijhout and Wheeler 1982, Wheeler 1986, Evans and Wheeler 2001). Genetic caste determination (GCD) is an association between genotype and caste phenotype in which females of different genotypes have different probabilities of becoming a queen or worker. GCD is an unlikely alternative to environmental caste determination (ECD), because it suggests the evolution of sterility at the genetic-level. Theoretically, any genetic segment that results in sterility should be quickly eliminated from a normal breeding population (Brian 1965, Hölldobler and Wilson 1990).

While nutritional caste determination is clearly more prevalent, there is evidence for genetic influences on reproductive caste determination in multiple Hymenopteran species (Marchal 1897, Kerr 1950, Heinze and Buschinger 1989). In meliponine bees, genotype may influence the interaction between nutrition and induction of the queen developmental pathway (Kerr 1950). A genetic factor in the slave-making ant species *Harpogoxenus sublaevis* can prevent complete expression of queen characters even if larvae are well fed and uninhibited by a dominant queen (Buschinger and Winter 1975, Hölldobler and Wilson 1990). In the fire ant *Solenopsis* there is evidence of a recessive lethal that

Manuscript received 12 May 2005; revised 17 October 2005; accepted 31 January 2006; final version received 28 March 2006. Corresponding Editor: P. Nonacs. For reprints of this Special Feature, see footnote 1, p. 2141.
[3] E-mail: kirk.anderson@asu.edu

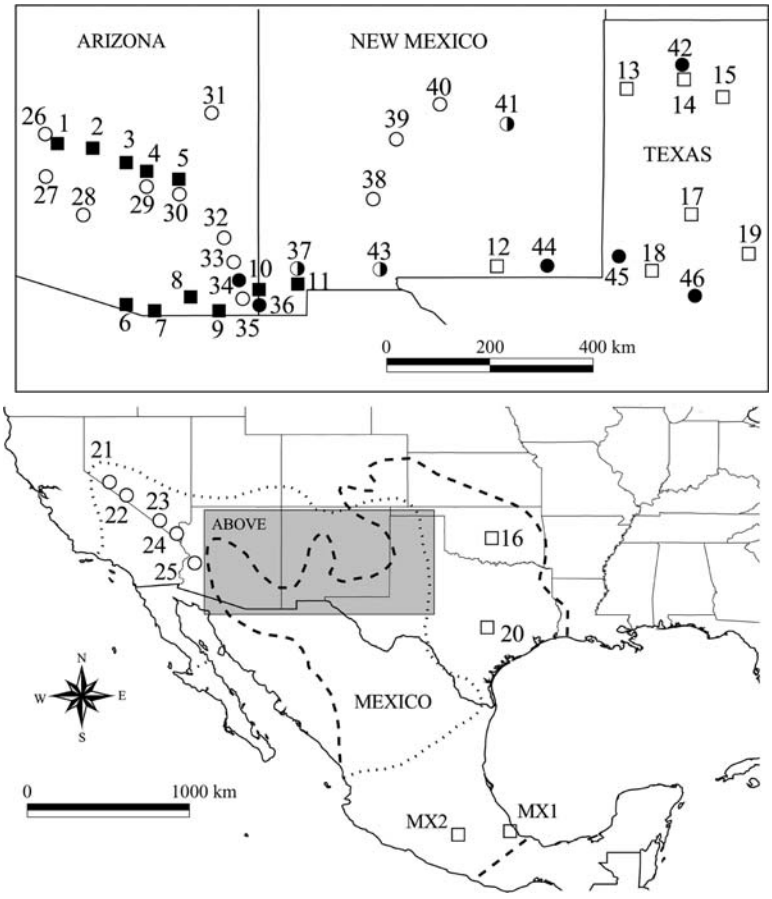

FIG. 1. Distribution of environmental caste determination (ECD) and dependent-lineage (DL) populations across the northern range of *Pogonomyrmex barbatus* and *P. rugosus*. The range of nominal *P. rugosus* is shown as a dotted line, and populations with *P. rugosus* morphology are denoted with circles. The range of nominal *P. barbatus* is shown as a dashed line, and populations with *P. barbatus* morphology are denoted with squares. Open symbols are populations with ECD, and solid symbols represent DL populations composed entirely of GCD colonies. The H1/H2 lineages are represented by solid circles, the J1/J2 lineages by solid squares. Half-open symbols are sites containing both ECD and GCD colonies. The mode of caste determination is unknown in *P. barbatus* MX1 and MX2.

causes premature death of caste-related genotypes (Ross 1997, Bourke 2002), and another that changes both queen behavior and worker tolerance of alternative genotypes (Ross and Keller 1998). Worker polymorphism in the leaf-cutting ant *Acromyrmex echinatior* was significantly associated with distinct patrilines suggesting that environmental response thresholds may be genetically determined (Hughes et al. 2003). These studies highlight the need to consider the evolution of the genome as well as environment when investigating the proximate mechanisms and ultimate causes of caste determination in social insects.

Each of the above examples involves allelic effects within a single lineage. However, a system of genetic caste determination occurs in some populations of the seed harvester ants, *Pogonomyrmex barbatus* and *P. rugosus*, in which multiple classes of molecular markers suggest a strong genetic system of caste determination that is associated with the maintenance of two distinct lineages within each population. Queens mate with a male of each lineage to produce inter-lineage workers composed genetically of both lineages, and intra-lineage reproductive females (new queens) composed genetically of a single lineage (see Fig. 1 in Nonacs 2006). Genetic caste determination in these populations relies on obligate polyandry, as queens must mate with a male of their own lineage to generate reproductive queens and also the alternate lineage to generate workers (Julian et al. 2002, Volny and Gordon 2002, Helms Cahan and Keller 2003). We label these dependent-lineage (DL) systems, because both lineages must be sustained in the population to generate functional GCD colonies.

Pogonomyrmex barbatus and *P. rugosus* are closely related species that form a complex within the genus (Cole 1968). Nuclear markers coupled with a mitochondrial DNA phylogeny indicate that the *P. rugosus–P. barbatus* complex is composed of at least six independently evolving lineages: one *P. rugosus* and one *P. barbatus* apparently with environmental caste determination (ECD), and two pairs of dependent lineages

which interbreed to form GCD colonies. The dependent lineages are also referred to as H lineages (H1/H2) or J lineages (J1/J2) based on discovery locations in Hidalgo and a highway junction, respectively (Helms Cahan and Keller 2003). The H lineages are morphologically indistinguishable from ECD *P. rugosus* and the J lineages are morphologically indistinguishable from ECD *P. barbatus*. Thus "*P. barbatus*" or "*P. rugosus*" are used here in reference to nominal morphospecies that lack genetic data.

Pogonomyrmex barbatus and *P. rugosus* have broad overlapping ranges throughout the western United States and reaching into Mexico (Cole 1968, Johnson 2000). However, recent investigations of GCD have focused on a small number of populations, limiting our ability to determine the geographic extent and evolutionary history of the GCD phenotype, and to test hypotheses on its origin and maintenance. One hypothesis for the origin of GCD is that recent hybridization between ancestral populations of the two species generated epistatic incompatibilities between two nuclear loci, producing reproductively isolated but interdependent lineages within each species. However, there are alternative pathways for the evolution of a GCD/DL system. These include: (1) ongoing hybridization between *P. barbatus* and *P. rugosus*, (2) hybridization generating an initial system of GCD that then introgressed into other lineages, and (3) GCD originating as a result of genetic caste bias generated within species.

Dissection of these alternate pathways requires a comprehensive phylogenetic and geographical analysis of GCD within this species complex. In this study, we extensively sample populations throughout the east-west ranges of both morphospecies, across areas of sympatry and allopatry. We use a diagnostic nuclear marker to assess the association of caste with zygosity in multiple colonies within each population as an indicator of the presence of GCD. We also sequence the *cox1* mitochondrial gene across the geographical range of both morphospecies to determine lineage membership and construct a more complete phylogeny of the species complex. Finally, we compare the pattern of GCD across populations, as indicated by nuclear markers, with their mitochondrial haplotypes to assess alternate scenarios for the evolution of GCD and the emergence of the dependent lineages.

Methods

Sampling

We collected $\geq$20 workers from 10–20 colonies from each of 46 populations of *P. barbatus* and *P. rugosus* across a transect spanning their east-west geographic range in the United States (Fig. 1, Appendix A). We sampled 20 populations of *P. barbatus* from south central Texas to central Arizona, USA and 26 populations of *P. rugosus* from central Texas to western Nevada, USA and alate virgin queens (winged reproductives) when available. Our collection sites included colonies of both species from areas of extreme allopatry outside their common range, allopatric areas within their common range, and six sympatric sites (Appendix A). We also collected *P. barbatus* workers from two allopatric populations in southern Mexico at sites 600 km south of the southernmost range of *P. rugosus*.

Populations were sampled from June through August of 2000–2003. Ants were preserved in 95% ethanol, or collected live, then transferred to an ultra-cold freezer (−72°C). We sorted samples according to accepted morphotypes based head and thorax sculpture and color (see Cole 1968). Voucher specimens have been deposited in the collections of Kirk E. Anderson, Robert A. Johnson (RAJC), and the Bohart Museum (University of California–Davis).

Allozyme analyses

Allozymes are allelic forms of enzymes inherited as Mendelian alleles. Allozyme polymorphisms are typically used as molecular markers for determining relationships at many levels of organization. When subject to electrophoresis, allozymes separate in a gel matrix according to the specific molecular properties of each allele, primarily net charge.

In female hymenoptera, one allele is inherited from each parent, so a diploid individual will possess two alleles at a single locus. The numbers designate the relative positions occupied by a particular allele following electrophoretic separation, and agree with previously published allele designations (Cahan and Keller 2003). Higher numbers indicate an allele that migrates faster (and therefore farther) across the electrical field than a lower numbered allele. The fractions represent individual genotypes. Heterozygous individuals (2/4) inherited two alleles that migrated different distances. Homozygous individuals (2/2 or 4/4) inherited two alleles that migrated the same distance.

We characterized nuclear differences among taxa by analyzing three allozymes (PGI, EST-1, and PGM-1) using standard extraction and staining protocols (Richardson 1982). In previous analyses, these allozyme markers were differentially associated with caste phenotype: PGI was diagnostic of caste, EST-1 was weakly associated, and PGM-1 was unassociated (Helms Cahan et al. 2002, Helms Cahan and Keller 2003). Detected alleles were numbered according to distance migrated on cellulose acetate gels, and correspond to published allele lengths for this species complex (Helms Cahan and Keller 2003). All 46 populations were characterized with the GCD diagnostic PGI locus, 19 populations with EST-1, and 12 populations with PGM-1. Frozen worker and reproductive ants were split in thirds. Gasters were used in allozyme analyses, heads were used for mtDNA sequencing, and thoraces were saved for morphological analysis. We ran six workers and six alate queens per colony from eight to 12 colonies from each of seven populations. In the remaining 39 populations, we ran six workers from each of four to 20 colonies.

mtDNA analyses

We determined ancestry and gene flow among lineages by analyzing 999 bp of the cytochrome oxidase 1 (*cox1*) mitochondrial gene for at least two colonies from all 46 populations. Colonies of both morphospecies are headed by one queen (Hölldobler 1976, Gordon and Kulig 1996), so the mitochondrial haplotype of one worker represents that of the queen and the colony. Haplotypes were determined by crushing the head in a 1.5-mL microcentrifuge tube, and isolating total genomic DNA using a standard phenol-chloroform extraction method (see Gadau et al. 1998). Amplification was achieved using the following profile: 3 min at 94°C, 40 cycles of 1 min at 94°C, 1 min at 45°C, 1.5 min at 72°C, and a final elongation step of 10 min at 72°C. Partial *cox1* fragments were amplified using two universal primer pairs in a PTC-100 MJ Research thermal cycler (Global Medical Instrumentation, Ramsey, Minnesota, USA). We used the primer pairs "Jerry" (Simon et al. 1994) and "Ben3R" (Brady et al. 2001), and "LCO" and "HCO" (Folmer et al. 1994). The latter primer pair produces a 630-bp DNA segment that includes 395 bps of a 433-bp sequence published in Genebank (Helms Cahan and Keller 2003), allowing comparisons with initial lineage designations.

Determination of GCD phenotype

Although we are characterizing populations and broad geographic patterns, genetic caste determination is a colony-level phenotype, initially characterized by heterozygosity in the worker caste and homozygosity in the alate queen caste at the same nuclear loci (Helms Cahan et al. 2002, Julian et al. 2002, Volny and Gordon 2002). Colonies of *Pogonomyrmex* produce sexual reproductives over a short period, making it difficult to collect both workers and alate queens from colonies across an extensive transect. Therefore, populations in which we sampled both workers and alate queens are used to infer genotypes for populations in which we sampled only workers.

We used allozymes to establish the presence of GCD in seven populations (76 colonies) from which we were able to collect both workers and alate queens. We confirmed the GCD phenotype when alate queens and workers from the same colonies violated Hardy-Weinberg equilibrium (HWE) due to complete fixation or a statistical excess of homozygous queens and heterozygous workers at the same loci. Associations between genotype and caste were assessed with a *G* test.

For populations in which we sampled only workers, we analyzed the diagnostic PGI locus for six workers per colony. Levels of worker heterozygosity at PGI are at or near 100% in colonies exhibiting GCD in this and previous studies (Helms Cahan et al. 2002, 2004, Helms Cahan and Keller 2003). Thus, GCD can be inferred with high confidence for colonies in which six randomly selected workers are heterozygous at the PGI locus. This colony level measure has biological significance as it increases the likelihood that field colonies are produced by a single queen that is homozygous at PGI. However, workers from the same colony have correlated genotypes and represent non-independent samples (Ross 1997). Thus, for each population we calculated statistical significance using a conservative approach to determine an excess of heterozygotes expected under HWE. We selected one worker genotype per locus at random from each colony in a population to estimate allele and genotype frequencies.

Dependent lineage confirmation

We expect that every population showing GCD on a colony level should possess two distinct mtDNA haplotypes, one that corresponds to each dependent lineage. Each of the four dependent lineages is defined by a particular association of morphology, mtDNA haplotype, and nuclear markers (Helms Cahan and Keller 2003). The nuclear marker PGI is most associated with GCD in this and previous studies (Helms Cahan et al. 2002, 2004) and is consequently most diagnostic of lineage. We confirmed lineage by determining the correspondence (ctyo-nuclear linkage) between the PGI alleles of alate queens and the mtDNA haplotype of their natal colony by sequencing 630 bp of the *cox1* mtDNA gene for all colonies ($n = 60$) from which we genotyped PGI of both workers and alate queens. This mtDNA gene fragment includes 395 bps of a *cox1* gene sequence published in Genebank, allowing comparisons with published results of H and J lineage compositions (Helms Cahan and Keller 2003). This comparison will resolve whether GCD across broad geography involves the interbreeding of two dependent lineages in every population. Associations between mtDNA haplotype and PGI alleles were determined with a *G* test.

For populations that lack alate queen genotypes but possess a statistical excess of workers heterozygous at the PGI locus, we confirmed the presence of two lineages by sequencing the same 630bp *cox1* fragment from progressively more colonies in that population ($n = 2$–6) until each dependent-lineage mtDNA haplotype was sampled. Nuclear alleles suggested co-occurrence of GCD and ECD colonies for a few populations in which some colonies showed all workers fixed as PGI heterozygotes and others showed variable worker genotypes. For these apparently "mixed populations" we determined lineage membership by sequencing 630 bps of the *cox1* gene for every colony in the population.

Phylogenetic inference

The maternal inheritance and non-recombining nature of mtDNA make it ideal for tracking the evolution of GCD. Understanding the number and nature of reticulate events recovered in our phylogenetic analysis requires a comparison of genetic divergence among dependent lineages and the most recent common ancestors expressing environmental caste determination. We generated a topology for sampled populations of the

TABLE 1. Correspondence between queen allele frequency for one nuclear marker (PGI allozyme) and mitochondrial haplotype for one population each of *Pogonomyrmex barbatus* and *P. rugosus* with environmental caste determination and five dependent lineage (DL) populations.

Population, lineage, and location	*N*	PGI allele	Mitotype 1	Mitotype 2	Mitotype 3	Worker genotype	χ^2	*P*
A. J1/J2 lineage			B	J1	J2			
ECD *P. barbatus* (B)								
14. Texas: Potter	48 (8)	2	0.00			0.00	†	†
		4	1.00					
DL *P. barbatus* (J1/J2)								
2. Arizona: Yavapai	30 (5)	2		1.00	0	1.00	12	0.0005
	42 (7)	4		0	1.00			
6. Arizona: Santa Cruz	18 (3)	2		1.00	0	1.00	11	0.0005
	48 (8)	4		0	1.00			
11. New Mexico: Grant	36 (6)	2		1.00	0	1.00	10	0.002
	24 (4)	4		0	1.00			
B. H1/H2 lineage			R	H1	H2			
ECD *P. rugosus* (R)								
28. Arizona: Pinal	60 (10)	3	0.98			0.03	0.03	NS
		4	0					
		5	0.02					
DL *P. rugosus* (H1/H2)								
37. New Mexico: Grant	48 (8)	3		0.98	0.08	0.96	12	0.0005
	24 (4)	4		0.02	0.92			
42. Texas: Potter	30 (5)	3		1.00	0.01	1.00	12	0.0005
	42 (7)	4		0	0.99			

Notes: The number before each population corresponds to locations in Appendix A; *N* indicates the number of alate queens (number of colonies) sampled for the PGI allozyme at each site; B and R represent ECD populations; J1/J2 and H1/H2 represent the dependent lineages. Mitotype was determined for one individual per colony using a 650-bp sequence from the *cox1* mtDNA gene (see *Methods*). Worker genotype gives the proportion of workers heterozygous for PGI in each population (see Appendix B). Chi-square (1 df) and *P* values were assessed by selecting one random worker per colony, then comparing the number of heterozygotes to that expected at Hardy-Weinburg equilibrium ("NS" indicates not significant).

† The population was fixed for a single allele, and no χ^2 test was performed.

two morphospecies, by sequencing 999 bps of *cox1* for at least one colony per population and using *P. californicus* as an outgroup ($n = 52$ sequences). Sequences were aligned by eye using BioEdit Version 6.0.7 (Hall 1999) with gaps removed. Phylogenetic and molecular analyses were conducted with MEGA version 3.1 (Kumar et al. 2004) using both maximum parsimony and neighbor-joining. For comparative purposes, the neighbor-joining topology was generated as described by Helms Cahan and Keller (2003), using the Kimura two-parameter distance model. The maximum parsimony phylogeny was obtained by branch and bound search with all sites weighted equally. For each topology, bootstrap analysis (500 replicates of heuristic searching) was used to determine strength of support for individual nodes. We then assessed current hypotheses concerning the evolutionary history of GCD, by overlaying the mtDNA topology with morphology and the GCD phenotype. To estimate genetic divergence among resulting groups we calculated the within and between group average genetic distance as the average *p* distance using the Kimura two-parameter method (Nei and Kumar 2000).

RESULTS

Association of GCD and PGI genotype

To confirm the PGI locus as diagnostic of GCD across widespread geography we determined if genotypes were distributed non-randomly with respect to caste within populations for which we collected both alate queens and workers. The PGI locus was significantly associated with caste, showing 98% homozygosity in queens ($n = 335$ of 342) and 99% heterozygosity in workers ($n = 340$ of 342) across five geographically distant populations (Table 1): (J1/J2 populations 2, 8, and 11, $n = 198$ queens from 33 colonies, $G = 136.04$, df $= 2$, $P < 0.0001$; and H1/H2 populations 37 and 42, $n = 144$ queens from 24 colonies, $G = 95.52$, df $= 2$, $P < 0.0001$). One population of each nominal morphospecies lacked an association between PGI genotype and caste, indicating environmental caste determination: (ECD population 28 of *P. rugosus,* $n = 60$ queens from 10 colonies, $G = 0.8$, df $= 2$, $P > 0.05$; and ECD population 14 of *P. barbatus*, $n = 48$ queens from eight colonies, $G = 2.1$, df $= 2$, $P > 0.05$).

Dependent lineage determination

Two distinct mtDNA haplotypes were detected in every population for which workers were significantly heterozygous at the PGI locus (Appendix B). Within each of the four lineages, there was complete concordance between morphology, mtDNA haplotype, and PGI alleles (Table 1). We compared the geographically distant populations 2, 8, and 11 with sequences published for the J lineages (Cahan and Keller 2003;

TABLE 2. Kimura two-parameter average *p* distance within (boldface on the diagonal) and between lineages based on a 999-bp sequence of the mtDNA gene cytochrome oxidase 1 (*cox1*).

Lineage	H1	H2	R	J1	J2	B
H1	**0.010**					
H2	0.011	**0.006**				
R	0.063	0.064	**0.023**			
J1	0.062	0.062	0.024	**0.004**		
J2	0.026	0.027	0.064	0.060	**0.023**	
B	0.056	0.054	0.076	0.071	0.054	**0.016**

Notes: B and R represent environmental caste determination (ECD) populations of *Pogonomyrmex barbatus* and *P. rugosus*, respectively; J1/J2 and H1/H2 represent the two sets of dependent lineages.

accession numbers AY542358, AY542362), and found that 17 of 395 *cox1* nucleotide sites distinguished lineage J1 from lineage J2. Within alate queens from populations 2, 8, and 11 there was a highly significant association between haplotype and PGI alleles confirming lineage ($G = 264.79$, df = 1, $P < 0.0001$; Table 1); J1 haplotypes were fixed at PGI allele 4 and J2 haplotypes were fixed at PGI allele 2. All workers ($n = 198$) from these populations were heterozygous (2/4) at the PGI locus (Table 1).

A comparison of geographically distant populations 37 and 42 with H1/H2 lineages (Helms Cahan and Keller 2003; accession numbers AY542355, AY542356) revealed that mtDNA haplotype distinguished lineage H1 from H2 at only one of 395 nucleotide sites (site 143; H1 = a, H2 = g). This diagnostic site had a highly significant association with distinct PGI alleles within alate queens from populations 37 and 42 confirming lineage (G test = 172.6, df = 1, $P < 0.0001$; Table 1); H1 haplotypes were fixed at PGI allele 3 and H2 haplotypes were fixed at allele 4. All but one worker ($n = 144$) from these two populations was heterozygous (3/4) at the PGI locus (Table 1). It is important to note that both lineages J2 and H2 were fixed at PGI allele 4, but each lineage possessed distinct morphology and haplotypes. In summary, each set of dependent lineages consisted of two distinct mtDNA haplotype groups, and each group showed complete concordance with a distinct PGI allele and/or morphology.

Worker heterozygosity and GCD

The significant association between genotype and caste extends previous results (Helms Cahan et al. 2002, Helms Cahan and Keller 2003) and confirms that the PGI locus is highly diagnostic of GCD across the east to west distribution of both morphospecies (Fig. 1, Table 1). For populations where only workers were sampled, we inferred GCD based on excess worker heterozygosity at the PGI locus. Across the sampled transect, GCD was detected in 11 of 20 populations of *P. barbatus* and nine of 26 populations of *P. rugosus* (Fig. 1, Appendix B). In *P. barbatus*, GCD occurred in western portions of its geographic range throughout Arizona and southwestern New Mexico, but was absent from central and eastern portions of its range (central New Mexico, Texas, and Oklahoma; Fig. 1). Eleven western populations of *P. barbatus* were significantly heterozygous at PGI ($f(2/4) = 1.0$, $n = 666$ of 666 workers from 111 colonies; Appendix B). Nine populations of *P. barbatus* from western portions of its range (central New Mexico to central Texas) were fixed at one allele or in HWE for PGI ($n = 582$ workers from 97 colonies). In *P. rugosus*, GCD occurred in the eastern portion of its range (southeastern Arizona, New Mexico, and Texas; Fig. 1). Nine populations were significantly heterozygous at PGI ($f(3/4) = 0.996$, $n = 556$ of 558 workers from 93 colonies; Appendix B). Seventeen *P. rugosus* populations in central and southeastern Arizona and areas further west were in HWE for PGI ($n = 900$ workers from 150 colonies). Patterns of GCD among *P. rugosus* populations also displayed rapid shifts over distances less than 30 km in southeastern Arizona (Fig. 1, Appendix B; see also Helms Cahan et al. 2006).

Dependent lineages and worker heterozygosity

Each DL population is composed of two distinct and reproductively isolated lineages that interbreed to produce the hybrid worker caste of GCD colonies. GCD within both pairs of dependent lineages (J1/J2 and H1/H2) was associated with a complete correspondence between excess worker heterozygosity at PGI and the presence of two distinctive mtDNA haplotypes in each population (Appendix B). This relationship reveals a nuclear-mitochondrial marker combination diagnostic of DL populations that can be inferred using only worker genotypes. The mtDNA haplotypes of GCD dependent lineages (J1/J2, H1/H2) were differentiated from ECD *P. rugosus* (R) and ECD *P. barbatus* (B) according to average *p* distance in the *cox1* sequence fragment (Table 2). Workers of *P. barbatus* from 11 western populations were significantly heterozygous at the PGI locus ($f(2/4) = 1.0$, $n = 666$ workers from 111 colonies) and all 11 populations possessed both J1 and J2 haplotypes (Appendix B). Eastern populations of *P. barbatus* either fixed at one allele or in HWE for PGI possessed ECD and halplotypes corresponding to lineage B, ($n = 582$ workers from 97 colonies; Appendix B).

Workers of *P. rugosus* from 9 east-central populations were significantly heterozygous at the PGI locus ($f(3/4) = 0.996$, $n = 556$ of 558 workers from 93 colonies) and all nine populations possessed both H1 and H2 haplotypes (Appendix B). Six of nine populations possessed the nucleotide site (number 143) established as diagnostic between H1 and H2 dependent lineages. In the remaining three H1/H2 populations (41, 43, and 45), one or the other lineage was variable at nucleotide site 143. However, haplotypes within each of these populations were distinctive (*p* distance, 0.08–0.17; Table 2), and these lineages were diagnosed using a different nucleotide site (site 307; H1 = a, H2 = g) that was fixed in seven

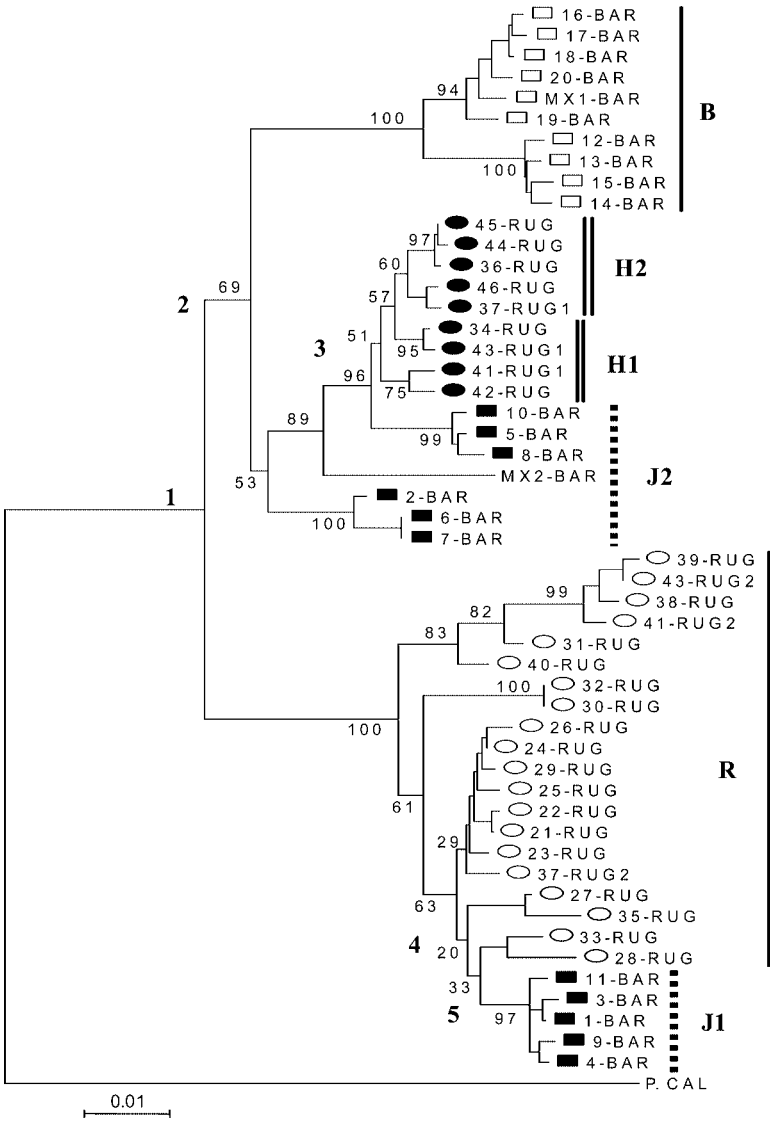

FIG. 2. A mtDNA (999-bp partial *cox1*) topology estimated under neighbor joining with 500 bootstraps and *P. californicus* (P. CAL) as the outgroup. Nodes relevant to the discussion are in bold (1–5). Terminal taxa are denoted by population number and morphology (*P. barbatus*, BAR; and *P. rugosus*, RUG). Caste determination is unknown for the Mexico populations (MX1 and MX2). Open symbols signify environmental caste determination (ECD), and solid symbols represent dependent-lineage (DL) populations. Groups of terminal taxa are labeled with capital letters according to lineage: ECD *P. barbatus*, B; DL *P. barbatus*, J1/J2 (dotted line); ECD *P. rugosus*, R; DL *P. rugosus*, H1/H2 (double line). Mixed populations (37, 41, 43) are represented by one DL (1) and one ECD (2) haplotype. Although DL populations are represented here by a single haplotype, two lineages were present in every DL population. The scale bar is substitutions/site according to the Kimura two-parameter distance method.

of nine (H1/H2) populations allowing lineage assignment. Seventeen *P. rugosus* populations in the central and western range were in HWE for PGI and colonies from these populations possessed ECD and corresponded to haplotype R ($n = 900$ workers from 150 colonies, Appendix B).

Ancestry of GCD

A total of 52 mtDNA sequences were used to construct a maternal phylogeny using both neighbor-joining and maximum parsimony methods. Both methods resulted in highly concordant topologies. We present the neighbor joining topology, as this method is consistent with that of Helms Cahan and Keller (2003). The strong concordance of *P. barbatus* morphology with mtDNA haplotypes of ECD group B, and *P. rugosus* morphology with mtDNA haplotypes of ECD group R suggests that this split (node 1) resulted from divergence of the ancestral species with environmental caste determination (Fig. 2). Assuming this morpholog-

ical-haplotype relationship represents ancestral forms, the overall structure of the topology reveals a pattern of non-correspondence between morphology and mtDNA haplotype indicative of introgression. MtDNA from morphological *P. rugosus* (H1/H2) is nested among morphological *P. barbatus* (B), making *P. barbatus* paraphyletic. Likewise, mtDNA from morphological *P. barbatus* (J1) is nested among morphological *P. rugosus* (R), making *P. rugosus* paraphyletic. At node 2, divergence results in one functionally monophyletic group with environmental caste determination (B), and a GCD group that includes lineage J2, and both H lineages. At node 3, H1 and H2 diverge from three basal groups: two highly distinct groups of lineage J2 and a *P. barbatus* from southern Mexico (MX2) of unknown caste determination. At node 4, two basal groups of lineage R root the J1 group with low bootstrap support, and at node 5, lineage J1 diverges to form a monophyletic group with high bootstrap support (Fig. 2).

The ECD groups of *P. barbatus* (B), and *P. rugosus* (R) each contain distinct sub-groups that correspond to geography. In group B, the distribution of ECD *P. barbatus* populations 12–15 is relatively northwestern, while 16–20 and MX1 are relatively southeastern (Figs. 1 and 2). Within group R, the first split results in a discrete group of ECD *P. rugosus* localized along the Rio Grande River and Colorado Plateau. The remaining members of the R group are relatively southwestern in distribution. The H lineages were localized in the east-central range of *P. rugosus*, (southeastern Arizona, New Mexico, and Texas), and the J lineages dominate the western range of *P. barbatus* from central to southeastern Arizona (Fig. 1).

The J1/J2 dependent-lineage groups show high sequence divergence between, but variable divergence within lineage, with J2 being the most variable of any lineage (Table 2). In contrast, both H lineages belong to a discrete group with similar levels of haplotype divergence both within and between lineages. *Pogomomyrmex barbatus* (B) is the most recent common ancestor with environmental caste determination to the GCD group H1/H2/J2 with an average between-group *p* distance of 0.055. The MRCA with environmental caste determination to GCD clade J1 is ECD *P. rugosus* from geographically adjacent populations 28 and 33 in Arizona with an average between group *p* distance of 0.017 (Table 2).

Sympatric and ECD/GCD populations

We compiled genotypes from six sympatric sites to test for present hybridization among ECD *P. barbatus*, ECD *P. rugosus,* and the dependent lineages that produce GCD colonies. At all six sites, both mtDNA haplotype and morphology agreed with lineage data from allopatric sites, suggesting no current gene flow between ECD species and the dependent lineages, or between J1/J2 and H1/H2 (Appendix C). The J and H lineages were sympatric at two locales near the southern border of Arizona and New Mexico: (populations 10 and 36, 11 and 37). At both locales, workers of the J lineages were fixed at the (2/4) genotype for PGI, and workers of the H lineages were fixed at the (3/4) genotype. At one of these sites, (11 and 37) the frequency and number of alleles sampled at EST-1 and PGM-1 also indicate reproductive isolation between H and J lineages (Appendix C). ECD *P. rugosus* and the J lineages co-occurred at three sites (1 and 26, 4 and 29, 5 and 30), and possessed no common PGI alleles ($n = 336$ workers, Appendix C). Where ECD *P. barbatus* and the H lineages co-occurred in Texas (populations 14 and 42), workers of ECD *P. barbatus* were fixed at one allele for PGI, and workers of the H lineages were 100% heterozygous (Appendices B and C). Across the sampled transect, ECD *P. barbatus* did not co-occur with the J lineages, and ECD *P. rugosus* did not co-occur with ECD *P. barbatus*.

At three sites (37, 41, and 43) *P. rugosus* was identified as H lineages interspersed with ECD *P. rugosus* (Appendix B). (For a detailed analysis of this lineage sympatry, see Helms Cahan et al. 2006.) At all three sites, colonies in which six randomly chosen workers were heterozygous at PGI also had haplotypes belonging to lineage H1 or H2. This association between fixed worker heterozygosity at PGI and DL haplotype occurred in 30 of 31 colonies. The one exception (population 37) showed four heterozygous and two homozygous workers and an H lineage haplotype.

Geography and nuclear alleles

Across DL populations, the pattern of heterozygosity produced by EST-1 reveals differences that correspond to intra-lineage variation. In contrast to fixed worker heterozygosity typical of PGI, EST-1 revealed levels of worker heterozygosity that differed within the J lineages, and between the J and H lineages. Within the J lineages there is a geographic cline of decreasing worker heterozygosity from central to southeastern Arizona (Fig. 3). The four J1/J2 populations (1, 2, 4, and 5) where EST-1 was highly heterozygous displayed a strong bias in workers for 2/3 and 2/2 genotypes (Table 3). We examined the within lineage (queen caste from populations 2, 8, and 11) contribution to this worker heterozygosity and determined that the geographic pattern was the result of variation within the J2 lineage at the EST-1 locus. Worker genotypes occur because lineage J1 is virtually fixed at allele 2 across geography, while lineage J2 has two EST-1 alleles (2 and 3) that result in a geographic cline (Fig. 3, Table 3).

In contrast to the geographic cline seen in the J2 lineage, four geographically distant H1/H2 populations were significantly out of HWE at the EST-1 locus due to near fixation of heterozygous workers (Appendix B, Fig. 3). Here the queen caste (H1/H2 populations 37 and 42; Table 3) revealed allelic variation in both contributing lineages; three alleles from each lineage in population 37, and two alleles from each lineage in population 42.

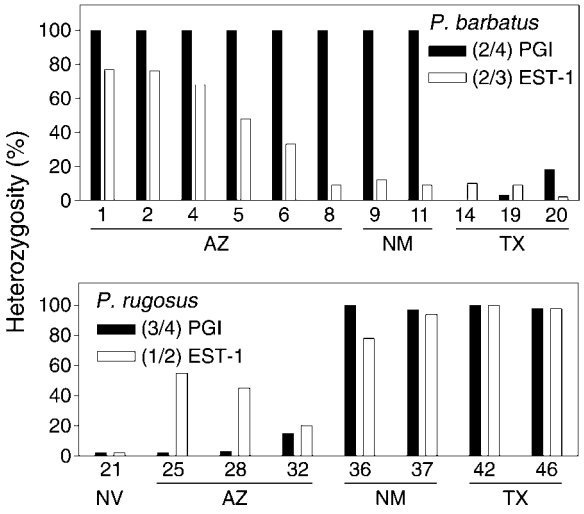

FIG. 3. Percentage of worker heterozygosity at two allozyme loci (PGI and EST-1) associated with GCD in populations of *P. barbatus* and *P. rugosus*. Numbers on the *x*-axis are populations, and the *y*-axis is percentage of heterozygosity in workers. State abbreviations are Nevada (NV), Arizona (AZ), New Mexico (NM), and Texas (TX) (refer to Fig. 1, Appendix B, and Table 3).

However, worker populations were significantly heterozygous for the 1/2 genotype (Appendix B, Fig. 3). Worker populations with R haplotypes in parts of central and southeastern Arizona and areas further west were in HWE for EST-1 (Appendix B, Fig. 3). Unlike PGI and EST-1, the PGM-1 locus was apparently unassociated with GCD in any of the 12 sampled populations (Appendix C).

DISCUSSION

Theoretically, environmental mediation of caste determining genes is required to maintain sociality (Queller and Strassman 1998). In contrast to this expectation, the colony level phenotype of genetic caste determination (GCD), in which queens and workers display different genotypes, is common and geographically widespread within both *P. barbatus* and *P. rugosus*, occurring in 20 of 46 sampled populations (Fig. 1). Populations in which we sampled queens and workers displayed complete association of GCD with both PGI alleles and lineage specific mtDNA haplotypes indicating that these haplotypes reliably diagnose dependent lineages, and infer GCD (Table 1). These data corroborate earlier results that the interbreeding of cryptic lineages within *P. rugosus* and *P. barbatus* generates GCD (Volny and Gordon 2002, Helms Cahan and Keller 2003, Parker 2004). Our study sampled individuals from large geographic areas compared to localized sampling used in previous studies. This necessitated collecting from numerous populations for which only workers were available. In the absence of queen data, we inferred the GCD phenotype if the population possessed two criteria: (1) two distinct lineage-specific mtDNA haplotypes, and (2) a significant excess of worker heterozygosity at PGI (Appendix B).

Modeling the origin of GCD

Two genetic models have been proposed to explain the origin of GCD. The first asserts that recent interspecific hybridization generated epistatic incompatibilities at two nuclear loci (Helms Cahan and Keller 2003). This model presents a parsimonious account for both the origin and maintenance of a DL system, but the generation of stable hybrid lineages is questionable as it requires that F_1 double heterozygotes become gynes and not workers (also see Linksvayer et al. 2006). A second model, similar to cytoplasmic male sterility, states that GCD occurs via interactions between the cytoplasm and nuclear genes such that some cyto-nuclear combinations develop into gynes while others develop into workers (Linksvayer et al. 2006). However, because both workers

TABLE 3. Variation at the EST-1 locus in alate virgin queens and workers from five populations of dependent-lineage *Pogonomyrmex*.

		EST-1 variation by caste				
		Alate queen alleles			Workers	
DL populations and localities	*L* (*N*)	1	2	3	Genotype	Heterozygous frequency
2. DL *P. bar* Yavapai, Arizona	J1 (30)	0	0.97	0.03	2/3	0.76
	J2 (42)	0	0.33	0.67		
8. DL *P. bar* Santa Cruz, Arizona	J1 (18)	0	1.00	0	2/2	0.09
	J2 (48)	0	0.88	0.12		
11. DL *P. bar* Hidalgo, New Mexico	J1 (36)	0	1.00	0	2/2	0.08
	J2 (24)	0	0.92	0.08		
37. DL *P. rug* Hidalgo, New Mexico	H1 (48)	0.27	0.69	0.04	1/2	0.94
	H2 (24)	0.88	0.10	0.02		
42. DL *P. rug* Potter, Texas	H1 (30)	0.03	0.97	0	1/2	1.00
	H2 (42)	0.98	0.02	0		

Notes: DL populations correspond to all figures and appendices. The second column shows the number of queens (*N*) from each lineage (*L*) used to calculate allele frequencies at EST-1 per lineage per population. The following columns list EST-1 frequency at three alleles in alate virgin queens; the final column presents the most common worker genotype, with frequency of workers heterozygous for this genotype (also see Figs. 1 and 3).

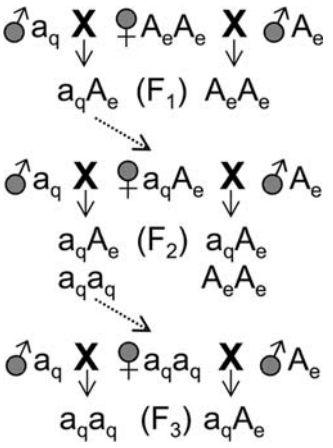

FIG. 4. A model for the origin and introgression of GCD. Solid arrows indicate the offspring of a particular mating cross; dashed arrows indicate generations. The wild-type regulatory allele is designated (A_e) and in the homozygous state results in either a worker or a queen via environmental caste determination. The mutant allele is designated (a_q) and in the F_1 heterozygote (a_q A_e) allows queen or worker expression. The F_2 generation produces (a_q a_q) homozygotes with genetically predetermined queen expression.

and queens within the same GCD colony possess the same cytoplasm and mitochondria, cyto-nuclear epistasis may appear indistinguishable from nuclear-nuclear epistasis.

A model of GCD maintenance postulates a single caste determining locus; individuals homozygous at this locus develop into gynes, and individuals heterozygous at this locus develop into workers (Volny and Gordon 2002). This model relies on the assumption that loci showing a classic GCD colony profile are physically linked to a locus with major caste influence. Our results confirm that the PGI locus is a prime candidate for such linkage, and in accord with the single locus model, may well infer the state of zygosity at an undetected locus with major caste influence. We expand upon the single locus model to discuss a potential origin of GCD via genetic mutation. Because this mutation precludes worker development and results in a queen phenotype, it becomes disproportionately represented in reproductive individuals. Thus it behaves as a "selfish" genetic element promoting its own survival at the expense of other parts of the genome. This mutation would result in intragenomic conflict through strong selection to retain the worker caste, and one stable resolution may be the evolution of dependent lineages as we observe in our system.

We suggest that GCD may have originated through mutation of a gene with a major influence on caste determination, e.g., a master gene in the caste regulatory network. The reproductive queen caste requires the full expression of many different structures like wings and ovaries. Recent comparative data show that the caste specific gene network for wing development is highly conserved while the suppression of this network that leads to worker development is evolutionarily labile (Abouheif and Wray 2002). Thus, a genetic mutation may result in a caste suppression network which responds poorly to environmental stimuli. An inefficient (mutant) caste suppressor would then bias the possessor toward queen development, and thereby bias its own representation in the reproductive caste. Because the worker caste is necessary to produce a colony, the mutant caste suppressor would result in strong selective pressure to retain an efficient caste suppressor (or suppression network) in the same population leading to antagonistic selection, and potentially generating genome evolution analogous to a general modification/rescue system (Werren 1997). One evolutionary stable outcome of this genetic conflict may be strict GCD and the system of dependent lineages in *P. barbatus* and *P. rugosus*.

The mutant caste locus (a_q) may be neutral in the heterozygous state or show an additive response that varies environmentally (Fig. 4). In the homozygous state (a_q a_q) this allele is incapable of suppressing the queen developmental pathway, resulting in individuals that can not develop into workers but are genetically predetermined to become queens. If heterozygotes (a_q A_e) are selected to become workers, a completely recessive gene could not spread within or invade an ECD population because (a_q) is continually shunted into the sterile worker caste. However, if we make the additional hypothesis that (a_q A_e) heterozygotes are neutral or generate a slight propensity toward queen development, the (a_q) allele may increase in frequency via drift, and after attaining some minimal frequency in the population increase rapidly by biasing its own representation in reproductive individuals (queens and males).

Non-hybrid origin

The contribution of hybridization to the origin and maintenance of GCD remains speculative. Previous studies suggest that dependent lineages emerged as a result of complex hybrid events between *P. rugosus* and *P. barbatus* (Helms Cahan and Keller 2003; recombinational speciation). The few cases for which we have genetic data on this mode of speciation suggest that it should occur rapidly, producing a similar set of surviving parental chromosomal blocks after a few generations of fertility selection (Rieseberg et al. 1995, 1996). Our phylogenetic data suggest that hybridization is not the most parsimonious explanation for the origin of GCD, primarily because levels of haplotype divergence between dependent lineages and their most recent common ancestor with environmental caste determination (ECD) are incongruent (Fig. 2, Table 2). The average sequence divergence of lineage J2 from its most recent common ancestor with ECD is more than three times that of lineage J1 from its most recent common ancestor with ECD. If GCD-associated dependent lineages originated via one hybridization event we would

expect each dependent lineage to quickly fix on a specific chromosomal combination, and attain rapid reproductive isolation from one another and the parental species. Assuming similar rates of mutation in mtDNA, each lineage should show similar degrees of mtDNA sequence divergence from its most recent common ancestor with ECD. The topology indicates that the J2 lineage evolved shortly following divergence with normal *P. barbatus*, suggesting that it separated from ECD species long before the evolution of lineage J1 (Fig. 2, Table 2).

To assess the hybrid nature of each lineage, we reanalyzed the supplementary data from Helms Cahan and Keller (2003). Although we found that the genetic character of lineage J1 is certainly due to introgression, lineage J2 has retained the morphology, mitotype, and nuclear genome of ECD *P. barbatus*. Lineage J2 possessed no allozyme loci specific to the putative parental *P. rugosus*, while all allozyme loci analyzed in J2 were represented in ECD *P. barbatus*. Of 63 total alleles, only four suggested hybridization between the putative parental species *P. barbatus* and *P. rugosus* (Helms Cahan and Keller 2003); all four were sampled at highly polymorphic microsatellite loci and occurred at very low frequencies in both J2 (mean = 0.126) and the putative parent *P. rugosus* (mean = 0.056). Given the rapid mutation of microsatellite loci and the time scale suggested by our study, the microsatellite results are best interpreted as chance convergence of alleles and not signatures of hybridization. This result, combined with the basal position of the J2 lineage in our topology, and paraphyly of J2 with the H lineages, strongly suggests that the J2 lineage is not of hybrid origin, and was established long before the H lineages and lineage J1 (Fig. 2). These results are consistent with the hypothesis that GCD evolved in *P. barbatus* and later introgressed into *P. rugosus*.

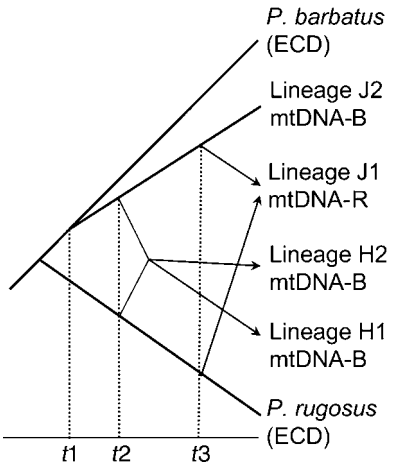

FIG. 5. Hypothesized origin and reticulation of the GCD phenotype based on the neighbor joining topology and morphology. *Pogonomyrmex rugosus* and *P. barbatus* possess environmental caste determination (ECD). The J lineages are morphologically indistinguishable from *P. barbatus*, and the H lineages are morphologically indistinguishable from *P. rugosus*. Capital letters B and R represent the mitochondrial (mtDNA) of *P. barbatus* and *P. rugosus*, respectively. The *x*-axis represents time; *t*1 is the origin of GCD, *t*2 and *t*3 represent the introgression of GCD (converging arrows) into another lineage.

Spread of GCD

Pogonomyrmex barbatus and *P. rugosus* are closely related species, raising the question of whether GCD originates via current hybridization. We found no evidence for current interspecific hybridization in sympatric populations among any of the lineages (including those with ECD), as morphology, allozymes, and mitochondrial haplotype indicated at least six reproductively isolated genomes (Appendices B and C). Three *P. rugosus* populations (37, 41, and 43) were mixed, containing both ECD and GCD colony phenotypes but even these populations showed no evidence of present introgression among lineages (Fig. 1, Appendices B and C).

The reproductive isolation and hybrid nature of lineages J1, H1, and H2 indicate that these lineages were produced via two separate introgression events. The timing and nature of hybridization events indicated by our mtDNA topology favor a non-hybrid origin of GCD within *P. barbatus*, followed by introgression into *P. rugosus* (Fig. 5). The first introgression was from lineage J2 of *P. barbatus* to normal *P. rugosus* lineages to produce the H lineages of *P. rugosus,* and the second introgression occurred much later to produce lineage J1. How was it possible for GCD to introgress from the J2 lineage into normal populations of *P. rugosus* and form stable dependent lineages? Initial hybrids generally encounter severe selection on the road to reproductive isolation (Arnold 1997). If they remain capable of interbreeding with parental species they would be assimilated by one parent and would not reach the status of biological species. Alternatively, hybrids showing strong postzygotic barriers would initially face a minority disadvantage (Abbott 2003) because most potential mates would be of parental type and these hybrids are likely to go extinct before they can form an independent lineage. Hybrids may overcome both problems if they quickly develop prezygotic isolation from their parent species, either by adapting to a new habitat (adaptive hybrid speciation, e.g., Rieseberg et al. 2003) or by selective mate choice. In *Pogonomyrmex*, there is no evidence that either of these prezygotic isolating mechanisms were in operation (but see Volny et al. 2006, Helms Cahan et al. 2006). Dependent lineages typically occupy the same habitats as the putative parental species and mating aggregations of mixed lineages are often synchronized, and appear to lack selective mate choice (K. Anderson, *personal observation*). During GCD introgression, resulting hybrids found a novel postzygotic mechanism that avoids the negative effects of both parental assimilation and

minority disadvantage. One possible mechanism is a gene or set of genes that result in queen development and thereby bias their representation in F_1 reproductive queens. Sperm parasitism, polyandry, and selection that favor interspecific mating (see Umphreys 2006) are all factors that would increase the probability and frequency of GCD introgression. When any relatively compatible lineage is encountered, the introgression of queen biasing genes (i.e., the a_q suppression mutant, Fig. 4) is also facilitated due to the rarity of the gene(s) which are subject to negative frequency dependent selection until attaining equilibrium within the new population.

Dependent lineage maintenance

The maintenance of dependent lineages relies on conserving specific allele combinations needed to produce both sterile workers and reproductive sexuals. However, reproductive isolation between lineages allows drift, selection, and mutation within each lineage, which may partially account for the geographical differences in the linkage disequilibrium between GCD and different nuclear markers (e.g., EST-1/PGI, Fig. 3, Tables 1 and 3). Markers that indicate a "classic" GCD colony profile (complete worker heterozygosity and complete queen homozygosity) may result from genome wide differences expected from reproductively isolated lineages including initial polymorphisms, fixation via drift, and hitchhiking of non-caste portions of chromosomes with selected caste determining loci (e.g., a selective sweep [Schlötterer 2003]).

In contrast, the loss of association of certain GCD markers may indicate the gradual loss of linkage disequilibrium as associated markers begin to segregate independent of caste determining loci. For these reasons, nuclear loci that indicate a GCD profile in one geographic region may show no association with GCD in another region (e.g., EST-1, Fig. 3). Our results and that of other studies (Helms Cahan et al. 2002, Helms Cahan and Keller 2003, Clark et al. 2006) indicate that the PGI locus is in strong linkage disequilibrium with caste genes of major effect. That PGI shows the greatest correspondence with both GCD colony phenotype and mitochondrial DNA lineage in both sets of dependent lineages (Fig. 3; Helms Cahan and Keller 2003, Helms Cahan et al. 2004) suggests it is tightly linked to highly conserved caste determining genes. To a lesser degree than PGI, the EST-1 locus is associated with caste in both morphospecies (Tables 1 and 3). Assuming that linkage disequilibrium, rather than functional relationship, is responsible for the difference in association between EST-1 and GCD, we propose that chromosomal segregation or intra-chromosomal recombination has decoupled the EST-1/GCD linkage disequilibrium in lineage J2 reflecting a longer history with GCD than lineage J1, H1, or H2.

Haplotype variability across our 46 populations also suggests a more recent evolution of the H1/H2 lineages because they retain ancestral polymorphism and possess highly similar mtDNA sequences (Table 2). Additionally, the bootstrap values for nodes defining H1 and H2 are small indicating that the H lineages are the result of a single invasion and may not be distinct lineages (Fig. 2). Alternatively, this haplotype similarity could reflect a recent transfer of mtDNA between lineages H1 and H2 (see Helms Cahan et al. 2006). While lineages H1 and H2 have very close mtDNA haplotypes, the nuclear genome of H1 is more similar to *P. rugosus* while H2 is more similar to *P. barbatus* supporting their origin via introgression (Helms Cahan and Keller 2003). Similarities of these mitochondrial sequences indicate that the cytoplasms of H1/H2 are functionally homologous, while those of J1/J2 are not. This suggests that the H lineages represent a different stage of dependent-lineage evolution relative to the J lineages. In support of this assertion, Helms Cahan et al. (2004) documented phenotypic plasticity within the H lineages (19% of intra-lineage matings produced viable workers) whereas phenotypic plasticity was completely absent within the J lineages (only inter-lineage matings produce workers).

Geographic patterns

Geographical data also support two independent introgressions of GCD. Populations of J1/J2 occur only at the western edge of their morphospecies range, and appear to be strictly allopatric from ECD *P. barbatus* in the east (Fig. 1). That the J lineages are highly disjunct with ECD *P. barbatus* indicates that they have not only the greatest mtDNA divergence (Table 2) from the most recent common ancestor with normal caste determination, but also the greatest geographic isolation (Fig. 1) Interestingly, the *P. barbatus* sample from southern Mexico (MX2) is nested within lineage J2 (Fig. 2). This result suggests that GCD may occur in southern Mexico where *P. barbatus* occurs several hundred kilometers south of the southernmost *P. rugosus* populations (Fig. 1).

A single and recent hybridization event causing GCD is expected to result in broad sympatry between the H and J lineages. However, the H lineages occur far into the range of ECD *P. barbatus* and show only slight parapatry with the J lineages at the southern-most border of New Mexico and Arizona. The H lineages occupy the central and eastern region of the *P. rugosus* range, and are sympatric with ECD *P. rugosus* throughout the central range, suggesting that the main concentration of H1/H2 populations is in Texas (Fig. 1). That the H lineages are not broadly sympatric with the most recent common ancestor lineage J2, but occur primarily in the range of ECD *P. barbatus,* suggests a relatively ancient introgression event. In contrast, the geographical proximity of lineage J1 with most recent common ancestor ECD *P. rugosus* is concordant with the phylogeny and suggests the relatively recent introgression of GCD from lineage J2 into ECD *P. rugosus* (R) to establish lineage J1 (Figs. 1 and 2).

Conclusions

The contribution of hybridization to the origin of GCD (Helms Cahan and Keller 2003) remains controversial. However, our results favor a relatively ancient and non-hybrid origin in *P. barbatus*. GCD may have originated through a mutation in the caste regulatory network which behaves as an egoistic element. To explain the mitochondrial phylogeny we need to assume at least two separate introgression events (Fig. 5). The first event was the introgression of GCD into *P. rugosus* to form dependent lineages H1 and H2. The second event formed the J1 dependent lineage. It remains to be determined if GCD exists outside the molecular criteria established by this study. We would be unable to detect a form of GCD that was unassociated with the PGI locus. If GCD does not originate with hybridization, it may be common and result in previously undetected cryptic lineages within any polyandrous ant species. It is conceivable that dependent lineages of *Pogonomyrmex* contain multiple lineages which vary in their degree of reproductive isolation. This study suggests that single genes of major effect (i.e., Gp-9 in *Solenopsis* [Krieger and Ross 2002]) can generate complex interactions at higher levels of biological organization. More detailed genetic and behavioral studies of dependent-lineage *Pogonomyrmex* will simultaneously provide insights into the genetic basis of both sociality and speciation.

Acknowledgments

The authors are particularly indebted to Belynda S. and Ariel K. Anderson for their patience with field-work and science in general. We thank Timothy Linksvayer, Sara Helms Cahan, Gary Umphreys, and two anonymous reviewers for comments on the manuscript. This project was supported by NSF award INT-0129319 to J. Fewell, DFG SFB 554-TP C5 to J. Gadau, and NSF DDIG award DEB-0508892 to K. Anderson. Study performed in partial fulfillment of a doctoral degree at Arizona State University.

Literature Cited

Abbott, R. J. 2003. Sex, sunflowers and speciation. Science **301**: 1189–1190.

Abouheif, E., and G. A. Wray. 2002. Evolution of the gene network underlying wing polyphenism in ants. Science **297**: 249–452

Arnold, M. L. 1997. Natural hybridization and Evolution. Oxford University Press, New York, New York, USA.

Bourke, A. F. G. 2002. Genetics of social behaviour in fire ants. Trends in Genetics **18**:221–223.

Brady, S. G., J. Gadau, and P. S. Ward. 2001. Systematics of the ant genus *Camponotus* (Hymenoptera: Formicidae): a preliminary analysis using data from the mitochondrial gene cytochrome oxidase I. Pages 131–139 *in* A. D. Austin and M. Dowton, editors. Hymenoptera: evolution, biodiversity and biological control. CSIRO, Collingwood, Victoria, Australia.

Brian, M. V. 1965. Studies of caste determination in *Myrmica rubra*: larval developmental sequences. Insectes Sociaux **12**: 347–362.

Buschinger, A., and U. Winter. 1975. Der Polymorphismus der sklavenhaltende Ameise *Harpagoxenus sublaevis*. Insectes Sociaux **22**:333–362.

Clark, R. M., K. E. Anderson, J. Gadau, and J. H. Fewell. 2006. Behavioral regulation of genetic caste determination in a *Pogonomyrmex* population with dependent lineages. Ecology **87**:2201–2206.

Cole, A. C. 1968. *Pogonomyrmex* harvester ants. A study of the genus in North America. First edition. University of Tennessee Press, Knoxville, Tennessee, USA.

Evans, J. D., and D. E. Wheeler. 2001. Gene expression and the evolution of insect polyphenisms. Bioessays **23**:62–68.

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

Gadau, J., P. J. Gertsch, J. Heinze, P. Pamilo, and B. Hölldobler. 1998. Oligogyny by unrelated queens in the carpenter ant, *Camponotus ligniperdus*. Behavioral Ecology and Sociobiology **44**:15–22.

Gordon, D. M., and A. W. Kulig. 1996. Founding, foraging and fighting: relationships between colony size and the spatial distribution of harvester ant nests. Ecology **77**:2393–2409.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Service **41**:95–98.

Hamilton, W. D. 1964. The genetic evolution of social behavior. I., II. Journal of Theoretical Biology **7**:1–52.

Heinze, J., and A. Buschinger. 1989. Queen polymorphism in *Leptothorax* sp. A: its genetic and ecological background. Insectes Sociaux **36**:139–155.

Helms Cahan, S., G. E. Julian, S. W. Rissing, T. Schwander, J. D. Parker, and L. Keller. 2004. Loss of phenotypic plasticity explains genotype-caste association in harvester ants. Current Biology **14**:2277–2282.

Helms Cahan, S., G. E. Julian, T. Schwander, and L. Keller. 2006. Reproductive isolation between *Pogonomyrmex rugosus* and two lineages with genetic caste determination. Ecology **87**:2160–2170.

Helms Cahan, S., and L. Keller. 2003. Complex hybrid origin of genetic caste determination in harvester ants. Nature **424**: 306–309.

Helms Cahan, S., J. D. Parker, S. W. Rissing, R. A. Johnson, T. S. Polony, M. D. Weiser, and D. R. Smith. 2002. Extreme genetic differences between queens and workers in hybridizing *Pogonomyrmex* harvester ants. Proceedings of the Royal Society of London B **269**:1871–1877.

Hölldobler, B. 1976. The behavioral ecology of mating in harvester ants (Hymenoptera: Formicidae: *Pogonomyrmex*). Behavioral Ecology and Sociobiology **1**:405–423.

Hölldobler, B., and E. O. Wilson. 1990. The ants. First edition. Harvard University Press, Cambridge, Massachusetts, USA.

Hughes, W. O. H., S. Sumner, S. Van Borm, and J. J. Boomsma. 2003. Worker caste polymorphism has a genetic basis in *Acromyrmex* leaf-cutting ants. Proceedings of the National Academy of Sciences (USA) **100**:9394–9397.

Johnson, R. A. 2000. Seed-harvester ants (Hymenoptera: Formicidae) of North America: an overview of ecology and biogeography. Sociobiology **36**:83–122.

Julian, G. E., J. H. Fewell, J. Gadau, R. A. Johnson, and D. Larrabee. 2002. Genetic determination of the queen caste in an ant hybrid zone. Proceedings National Academy of Sciences (USA) **99**:8157–8160.

Kerr, W. E. 1950. Evolution of the mechanisms of caste determination in the genus *Melipona*. Evolution **4**:7–13.

Krieger, M. J. B., and K. G. Ross. 2002. Identification of a major gene regulating complex social behavior. Science **295**: 328–332.

Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics **5**:150–163.

Linksvayer, T. A., M. J. Wade, and D. M. Gordon. 2006. Genetic caste determination in harvester ants: possible origin and maintenance by cyto-nuclear epistasis. Ecology **87**:2185–2193.

Marchal, P. 1897. La castration nutriciale chez les Hymenoptéres sociaux. Comptes Rendu de la Société de Biologie (Paris) **1897**:556–557.
Nei, M., and S. Kumar. 2000. Molecular evolution and phylogenetics. Oxford University Press, New York, New York, USA.
Nijhout, H. F. 1994. Insect hormones. Princeton University Press, Princeton, New Jersey, USA.
Nijhout, H. F. 1999. Control mechanisms of polyphenic development in insects. BioScience **49**:181–192.
Nijhout, H. F., and D. E. Wheeler. 1982. Juvenile-hormone and the physiological-basis of insect polymorphisms. Quarterly Review of Biology **57**:109–133.
Nonacs, P. 2006. Interspecific hybridization in ants: at the intersection of ecology, evolution, and behavior. Ecology **87**: 2143–2147.
Parker, J. D. 2004. A major evolutionary transition to more than two sexes? Trends in Ecology and Evolution **19**:83–86.
Queller, D. C., and J. Strassman. 1998. Kin selection and social insects. Journal of Bioscience **48**:165–175.
Richardson, B. J., P. R. Baverstock, and M. Adams. 1982. Allozyme electrophoresis. A handbook for animal systematics and population studies. Academic Press, New York, New York, USA.
Rieseberg, L. H., O. Raymond, D. M. Rosenthal, Z. Lai, K. Livingston, T. Nakazato, J. L. Durhy, A. E. Schwarzbach, L. A. Donovan, and C. Lexer. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. Science **301**:1211–1216.
Rieseberg, L. H., B. Sinervo, C. R. Linder, M. C. Ungerer, and D. M. Arias. 1996. Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. Science **272**:741–745.
Rieseberg, L. H., C. Van Fossen, and A. M. Desrochers. 1995. Hybrid speciation accompanied by genomic reorganization in wild sunflowers. Nature **375**:313–316.
Ross, K. G., and L. Keller. 1998. Genetic control of social organization in an ant. Proceedings of the National Academy of Sciences (USA) **95**:14232–14237.
Ross, K. L. 1997. Mulitlocus evolution in fire ants: effects of selection, gene flow and recombination. Genetics **145**:961–974.
Schlötterer, C. 2003. Hitchhiking mapping: functional genomics from the population from the population perspective. Trends in Genetics **19**:32–38.
Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America **87**:651–701.
Umprhey, G. J. 2006. Sperm parasitism in ants: selection for interspecific mating and hybridization. Ecology **87**:2148–2159.
Volny, V. P., and D. M. Gordon. 2002. Genetic basis for queen-worker dimorphism in a social insect. Proceedings of the National Academy of Sciences (USA) **99**:6108–6111.
Volny, V. P., M. J. Greene, and D. M. Gordon. 2006. Brood production and lineage discrimination in the red harvester ant (*Pogonomyrmex barbatus*). Ecology **87**:2194–2200.
Werren, J. H. 1997. Biology of *Wolbachia*. Annual Review of Entomology **42**:587–609.
Wheeler, D. E. 1986. Developmental and physiological determinants of caste in social Hymenoptera: evolutionary implications. American Naturalist **128**:13–34.

APPENDIX A

Location data (state, county, locale) and sample size for specimens of *Pogonomyrmex barbatus* and *P. rugosus* used for allozyme data (*Ecological Archives* E087-133-A1).

APPENDIX B

Summary of worker heterozygosity data for two allozymes (PGI and EST-1) sampled across populations of *Pogonomyrmex barbatus* and *P. rugosus* (*Ecological Archives* E087-133-A2).

APPENDIX C

Summary of allele frequencies at three allozymes (PGI, EST-1, and PGM-1) for workers and alate queens across populations of the nominal morphospecies *Pogonomyrmex barbatus* and *P. rugosus* (*Ecological Archives* E087-133-A3).