

Rockefeller University

Digital Commons @ RU

Student Theses and Dissertations

2020

Hybridization in Ants

Ian Butler

Follow this and additional works at: https://digitalcommons.rockefeller.edu/student_theses_and_dissertations



Part of the [Life Sciences Commons](#)



HYBRIDIZATION IN ANTS

A Thesis Presented to the Faculty of
The Rockefeller University
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

by

Ian Butler

June 2020

HYBRIDIZATION IN ANTS

Ian Butler, Ph.D.
The Rockefeller University 2020

Interspecific hybridization is a relatively common occurrence within all animal groups. Two main factors make hybridization act differently in ants than in other species: eusociality and haplodiploidy. These factors serve to reduce the costs of interspecific hybridization in ants while simultaneously allowing them to take advantage of certain benefits. Eusociality may mitigate the effects of hybridization by allowing hybrids to be shunted into the worker caste, potentially reducing the effects of hybrid sterility. In haplodiploid species, males do not have a father. They instead develop from unfertilized eggs as haploid clones of their mother. This means that interspecifically mated queens do not completely sacrifice reproductive potential even if all hybrids are sterile because they can still produce fertile males. These factors in turn suggest that hybridization should be more common among the social Hymenoptera than other animal groups. Nevertheless, current data suggest that ants hybridize at rates similar to other animal groups, although these data are limited. Furthermore, there is a large amount of overlap between cases of interspecific hybridization and cases of genetic caste determination. A majority of the cases in ants where caste is determined primarily by genotype are associated with hybridization. However, it is not clear how these two phenomena are related, and more research is needed to answer this question.

As a first step in answering these questions, I designed a set of microsatellite markers for use in African driver ants in the genus *Dorylus*. Additionally, to facilitate population genetics research in all ant species I aimed to develop a set of primers that are broadly applicable to most ant species, since PCR primers for microsatellite loci are often not useful outside the species for which they were designed. I identified 45 conserved microsatellite loci based on the eight ant

genomes that were available at the time and designed primers for PCR amplification. Among these loci, I chose 24 for in-depth study in six species covering six different ant subfamilies. On average, 11.16 of these 24 loci were polymorphic and in Hardy-Weinberg equilibrium in any given species. The average number of alleles for these polymorphic loci within single populations of the different species was 4.59. This set of genetic markers will thus be useful for population genetic and colony pedigree studies across a wide range of ant species, supplementing the markers available for previously studied species and greatly facilitating the study of the many ant species lacking genetic markers. This work shows that it is possible to develop microsatellite loci that are both conserved over a broad range of taxa, yet polymorphic within species, and should encourage researchers to develop similar tools for other large taxonomic groups.

After the development of these microsatellites, I used them to investigate a system of hybridization between two species of African driver ants. All driver ants belong to the subgenus (*Anomma*) in the genus *Dorylus*. They are swarm-raiding army ants with colonies that can have as many as 12 million individual ants. Colonies frequently migrate to new nest sites and conduct daily swarm-raids, capturing and eating any invertebrates or even small vertebrates in their path. Colonies are monogynous, and the queens are highly multiply mated, mating with as many as 20 males. A previous study suggested that hybridization occurs between *Dorylus molestus* and *Dorylus wilverthi* at a site in western Kenya. However, the extent and exact pattern of hybridization have remained unclear, and its possible effect on caste determination has not been investigated. I aimed to determine the extent and direction of hybridization by measuring how frequently hybrids occur in colonies of both species, and to investigate the possibility of genetic caste determination. I show that hybridization is bidirectional and occurs at equal rates in both species. Hybrid workers make up only 1–2% of the population, and successful interspecific matings represent

approximately 2% of all matings in both species. This shows that, although interspecific matings that give rise to worker offspring occur regularly, they are much rarer than intraspecific matings. Finally, I find no evidence of an association between hybridization and genetic caste determination in this population. Genetic caste determination may be associated with hybridization, but it is not a necessary outcome of it in ants.

Although there was no evidence of genetic caste determination, studying this *Dorylus* system has uncovered the potential for a novel project. After viewing collection data from a collaborator, Caspar Schöning, I hypothesized that *Dorylus* ants in the subgenus *Anomma* would constitute a good system for addressing an unanswered question in evolutionary biology: what is the relationship between the permissibility of the genome to introgression between two species and divergence time? *Dorylus* (*Anomma*) is a good system for this study because it has multiple species with different areas of allopatry and areas of sympatry with other species in the group. This project would involve sequencing multiple samples of each species from both allopatric and sympatric areas and comparing the genomes of samples from areas of allopatry to those from areas of sympatry to measure the amount of introgression between multiple species pairs. A model is then fit to a plot of the amount of introgression versus divergence time to determine the shape of the relationship.

Acknowledgements

I would like to give special thanks to my advisor Daniel Kronauer and the rest of the Kronauer Lab. Thank you also to the other members of my committee, Cori Bargmann and Fred Cross, to my external examiner Tim Linksvayer, and to my various collaborators over the past several years: Marcell Peters, Kim Siletti, Peter Oxley and Caspar Schöning. Special thanks to my partner Ingrid Fetter Pruneda and our son Julian for supporting me through the end of the thesis writing process. Many thanks to Gotham Knights Rugby Football Club for giving me a physical outlet for all the stress over the last seven years.

Table of Contents

Acknowledgements	iii
Table of Contents	iv
List of Figures	v
List of Tables	vi
Chapter 1: Introduction	1
Chapter 2: Development of broadly applicable microsatellite markers in ants	24
Chapter 3: Hybridization between two species of African swarm raiding army ants	39
Chapter 4: Future directions	72
Appendices	82
References	100

List of Figures

Figure 1.1. Cartoon depicting the different systems of genetic caste determination.....	16
Figure 1.2. Cartoon Venn diagram depicting the high amount of overlap between cases of genetic caste determination in ants and interspecific hybridization.....	21
Figure 2.1. Phylogeny of the ants showing the phylogenetic distribution of the species used to develop universal ant microsatellite markers.....	30
Figure 2.2. Overview results of genotyping 24 microsatellite loci for six different ant species....	37
Figure 3.1. Photographs of <i>Dorylus molestus</i> and <i>Dorylus wilverthi</i> from allopatric populations and from Kakamega Forest.....	44
Figure 3.2. Haplotype network of driver ant samples from East Africa.....	51
Figure 3.3. Principal components analysis of nuclear microsatellite genotypes.....	55
Figure 3.4. Assignment of driver ant workers to two populations using the clustering program STRUCTURE.....	58
Figure 3.5. Assignment of reconstructed driver ant queen genotypes to two populations using the clustering program STRUCTURE.....	60
Figure 3.6. Assignment of reconstructed driver ant male genotypes to two populations using the clustering program STRUCTURE.....	62
Figure 4.1. Illustration of the history of substitutions between two independently evolving lineages.....	73
Figure 4.2. Illustration of the possible shapes of the relationship between divergence time and the proportion of the genome that is permissible to introgression.....	74
Figure 4.3. Map of collection localities for 9 species of <i>Dorylus</i> (<i>Anomma</i>).....	76

List of Tables

Table 1.1. Cases of hybridization in ants.....	13-14
Table 2.1. PCR amplification success across six ant species for 21 microsatellite loci that were tested with unlabeled primers.....	32
Table 2.2. Characteristics of 24 microsatellite loci tested in six different ant species using labeled primers.....	33
Table 2.3. Overview of number of alleles and expected and observed heterozygosity in eight studies of species-specific microsatellite primers in ants.....	37
Table 3.1. List of unique <i>cytochrome oxidase II (COII)</i> mitochondrial haplotypes from Kakamega Forest.....	53-54
Table 3.2. Number of alleles, size range of alleles, observed heterozygosity and expected heterozygosity estimates from 12 microsatellite loci measured for <i>Dorylus molestus</i> and <i>Dorylus wilverthi</i> at Kakamega Forest.....	55
Table 3.3. List of putative hybrid workers at three different proportion of ancestry cut-offs and species determinations of their colony, mother queen and father.....	59
Table 3.4. Mating frequencies of 44 reconstructed queen genotypes.....	63

Chapter 1: Introduction

Speciation is the process by which two populations diverge and begin to accumulate genetic differences that contribute to reproductive isolation, and ultimately results in these populations becoming two reproductively isolated species. Hybrid zones occur when two populations encounter each other before reproductive isolation is complete and produce interspecific hybrids. In many cases of animal hybridization, the most noticeable consequences are sterility or inviability of hybrid offspring due to incompatibilities between the genomes of the parental species (Coyne & Orr, 2004). For example, many instances of hybridization in Lepidoptera result in higher rates of inviability or sterility in the hybrid offspring compared to pure species (Presgraves, 2002). Extrinsic consequences are less immediately severe but are equally important for limiting interspecific mating and hybridization. These may include situations where hybrids are unable to utilize resources commonly exploited by the parental species or unusual courtship behaviors displayed by the hybrid that make mate seeking less successful (Arnold, 1997; Arnold, 2006; Coyne & Orr, 2004). To take another example from Lepidoptera, hybrids between two species of *Heliconius* butterflies have reduced success when attempting to court either of the parental species (Naisbit et al., 2001).

Consequences of hybridization in ants

Ants can suffer the same consequences of hybridization as most other species. For example, F1 hybrid queens between *Temnothorax nylander* and *T. crassispinus* had reduced viability, were smaller than pure lineage queens, and suffered extremely reduced colony founding success (Pusch et al., 2006a). Hybrid populations of *Solenopsis invicta* x *S. richteri* showed higher fluctuating

asymmetry (random deviations from perfect bilateral symmetry in individuals) than pure populations (Ross & Robertson, 1990), and in lab crosses of various social parasitic *Temnothorax* species, hybrid males were less vigorous in their mating attempts and copulations were less likely to result in insemination (Jessen & Klinkicht, 1990).

Ants can suffer additional consequences of hybridization not experienced by most other species due to two aspects of their biology: haplodiploidy and eusociality. In haplodiploid species, males are haploid clones of their mother (usually the queen) and are produced from unfertilized eggs in a process called arrhenotokous parthenogenesis. Females, on the other hand, are produced via normal sexual reproduction. This means that a queen that has mated with a heterospecific male will still produce purebred sons while producing hybrid female workers. Hybrid males can only be produced in the F2 generation as sons of hybrid queens. A heterospecifically mated queen therefore does not completely sacrifice her reproductive potential even if all hybrid offspring are sterile. This is the case in the socially parasitic ant genus *Temnothorax* where laboratory crosses did not produce hybrid males in the F1 generation, but hybrid queens did produce them in the F2 generation (Jessen & Klinkicht, 1990). F2 hybrid males are therefore equivalent to the F1 generation of males in species where they are produced sexually.

Eusocial species are characterized by reproductive division of labor, having one or several reproductive individuals and a large number of non-reproductive workers. Some of the consequences of hybridization can be mitigated as long as hybrids are viable. In most species, workers far outnumber reproductives, so the majority of hybrids will be part of the worker caste, which normally does not reproduce, thereby minimizing the effect of hybrid sterility. This can be taken to an extreme in some cases, when all hybrids are shunted into the worker caste, and none ever appear in the reproductive caste (Anderson et al., 2008a; Schwander et al., 2010).

Many biologists argue that the gene is the fundamental unit of selection, but genes can act at multiple levels to maximize their own fitness (Dawkins, 1976), and in many eusocial species, selection appears to act at both the level of the colony and the individual (Bourke, 2011; Breed, 1989; Keller & Reeve, 1999; Marshall, 2015; Okasha, 2006). This can be particularly true in some ant species that have an irreversible worker caste (e.g. Crespi & Yanega, 1995; Boomsma, 2007) where workers are incapable of transforming into a reproductive form or otherwise initiating reproduction. Colony level selection can make hybridization beneficial in some situations. There are no studies showing that hybridization directly contributes to an increase in fitness, but increased intracolony genetic variance is hypothesized to confer numerous colony level benefits (Boomsma & Ratnieks, 1996; Bourke & Franks, 1995; Crozier & Fjerdingstad, 2001; Crozier & Page, 1985; Crozier & Pamilo, 1996; Nonacs, 2017; Oldroyd & Fewell, 2007). In honeybees, different polyethisms based on different worker subfamilies within a colony have been identified (Calderone et al., 1989; Calderone & Page, 1988; Calderone & Page, 1991; Dreiller et al., 1995; Fewell & Page, 1993; Frumhoff & Baker, 1988; Fuchs & Moritz, 1999; Oldroyd et al., 1991; Oldroyd et al., 1992a; Oldroyd et al., 1992b; Oldroyd et al., 1993; Page et al., 1989; Page & Robinson, 1991). Although direct evidence is lacking, if worker behavior is in part under genetic control, then higher genetic diversity among the worker caste may allow colonies to exploit more resources or environmental conditions (Crozier & Page 1985; Oldroyd et al., 1995; Oldroyd et al., 1996; Robinson & Page, 1989; Page et al., 1989; Page et al., 1995).

Increased genetic diversity can also mitigate against the effects of parasitism by increasing intracolony resistance (Cremer et al., 2007; Hughes & Boomsma, 2004; Schmid-Hempel, 1995; Schmid-Hempel, 1997; Schmid-Hempel, 1998; Schmid-Hempel & Crozier, 1999; Sherman et al., 1988; Shykoff & Schmid-Hempel, 1991a; Shykoff & Schmid-Hempel, 1991b). A colony with

lower genetic diversity has a large number of genetically similar individuals living in close proximity. A pathogen or parasite that is adapted to the particular genotypes within the colony can spread rapidly and, if the infection is severe enough, may cause the colony to die completely (Schmid-Hempel, 1997). Increased genetic diversity would prevent such a pathogen from spreading quickly within the colony (Schmid-Hempel, 1995). In honeybees, increased genetic diversity has been associated with a greater ability to thermoregulate the colony environment (Jones et al., 2004), and in *Pogonomyrmex* harvester ants, low within-colony relatedness correlates with increased colony growth rate (Cole & Wiernasz, 1999).

Hybridization may additionally provide an increase in colony genetic diversity as an extreme form of outbreeding (Boomsma et al., 2009). Outbreeding can reduce production of diploid males (Page & Metcalf, 1982; Pamilo et al., 1994), which are sterile and impose a metabolic cost if they are reared instead of workers (Crozier & Pamilo, 1996), and a high diploid male load can be fatal to a colony (Tarpy & Page, 2001; Tarpy & Page, 2002). In Hymenoptera, sex is determined by a complementary sex-determination (CSD) locus. Females are produced when the locus is heterozygous, and males are produced when the locus is homozygous or hemizygous. If a queen mates with a male that carries one of her CSD alleles, half of the offspring produced from that mating will be sterile diploid males (Beye et al., 2003; Whiting, 1933). Outbreeding via hybridization may increase the number of alleles the queen can potentially be exposed to and reduce the chances of producing diploid males, although this is only likely to provide a benefit in polyandrous or polygynous colonies where the negative consequences of hybridization can be diluted among the worker population, and fertile reproductives can be produced in sufficient numbers.

Polyandry (queens having multiple mates) and polygyny (multiple queens in a single colony) are further ways to increase intracolony genetic diversity and have the possibility of contributing to the factors listed above. Additionally, they may dilute the negative consequences of hybridization in another way. High mating frequencies may reduce the potential deleterious effects of occasional interspecific hybridization because in each case only a small proportion of the workers in a colony will be affected. Because in many ant species workers normally do not reproduce, common and often significant consequences of hybridization, such as hybrid sterility, will incur no or little additional cost. Polyandry reduces the relative genetic contribution of each individual male, so as long as the queen has mated with a sufficient number of conspecific males, the cost of interspecific mating is low.

Polygyny may similarly reduce the colony level costs of interspecific matings if most queens mate conspecifically with only a few mating interspecifically. Some queens in polygynous colonies forego production of new queens in favor of workers (Helms Cahan & Vinson, 2003), so interspecific matings would bear little additional cost if those queens were already destined to produce workers. Furthermore, polyandry and polygyny can have additional benefits that counteract any negative consequences of hybridization. In bumblebees and honeybees, colonies headed by polyandrous queens had a lower rate of disease infection (Baer & Schmid-Hempel, 1999; Liersch & Schmid-Hempel, 1998; Palmer & Oldroyd, 2000; Palmer & Oldroyd, 2003; Seeley & Tarpy, 2007; Tarpy & Seeley, 2006). Multiply mated honeybee queens founded colonies more quickly (Matilla & Seeley, 2007), and multiply mated bumblebee queens produced more reproductive offspring (Baer & Schmid-Hempel, 1999) when compared to singly mated queens.

Hybrid vigor (also called heterosis or outbreeding enhancement) is the improvement of biological traits in hybrid offspring and has been observed in a number of species (Chen, 2010).

This phenomenon is often accompanied by deleterious hybrid traits such as sterility. However, as previously explained, the colony can be seen as the unit of selection in eusocial species, and hybrid vigor can confer colony level benefits. Hybrid workers can receive any benefits of hybrid vigor without the cost of sterility because workers are normally sterile anyway. This has been observed at least one time in ants. Hybrid workers between *Solenopsis invicta* and *S. richteri* were more tolerant of low temperatures than either pure species workers (James et al., 2002), which could provide a selective benefit to the colony in the introduced range in Mississippi, USA compared to their native range in South America.

How common is hybridization in ants?

On the level of the individual hybridization is rare, affecting very few individuals within a population. This statement is necessarily true because if hybridization were common at this level, gene flow between the involved populations would make the species indistinguishable from one another. Under these circumstances, they would not be classified as distinct species under most species concepts (Mallet, 2005). On the other hand, zoologists have come to recognize that natural hybridization is relatively common at the species level among all animal groups (Mallet, 2005; Mallet, 2007). Although hybridization is rare within a population, a relatively small number of hybrids can facilitate introgression between different species, and even low rates of hybridization can have significant evolutionary consequences (Arnold, 1997; Arnold, 2006).

Numerous cases of natural hybridization have been identified among nearly all animal groups, but there have been few attempts to estimate the frequency of hybridization for large taxonomic groups in the wild, (i.e. the number of species that naturally hybridize and the proportion of all species they represent) (Mallet, 2005). One such attempt was in North American

fishes where the percent of species forming hybrids with at least one other species ranged from ~3% in the perch family to 17% in Pacific minnow species (Hubbs, 1955). Another attempt estimated that, worldwide, 9.2% of all bird species hybridize (Grant & Grant, 1992). Combining these estimates with those from smaller taxonomic groups, Mallet (2005) estimates that hybridization occurs in approximately 10% of all animal species. That is, 10% of animal species naturally form hybrids with at least one other species.

Feldhaar et al. (2008) claim that hybridization should be more common in ants than in other groups. They make this claim based on two aspects of ant biology. First, ants are eusocial, so the negative consequences of hybridization can be mitigated if the effects, namely hybrid sterility, are felt predominantly by the worker caste. Colonies produce far more workers than queens, so as long as some fertile queens are produced the consequences for colony growth and maintenance are minimal. Second, male ants are haploid and are produced from unfertilized eggs laid by the queen. This means that heterospecifically mated queens do not completely forego reproduction. As long as they can still lay viable eggs, they can produce haploid males, even if all diploid offspring are sterile.

Seifert & Goropashnaya (2004) estimate that 12% of all ant species hybridize. This level is comparable to the overall estimate for animals, as well as with the estimates for other large taxonomic groups. This estimate is for ants overall, but there is a large amount of variation among the lower level taxa that this estimate is drawn from. For example, 60% of the ant species in the *Formica rufa* group of central Europe hybridize with at least one other species (Seifert & Goropashnaya, 2004), whereas only 10% of all central European ants do (Seifert, 1999). Similar to the overall estimate for ants, this is comparable with hybridization rates in other animals as well. In birds, 76% of British duck species hybridize and 43% of birds of paradise hybridize while no

hybridization events have been found among Warblers of the western Palearctic (Mallet, 2005). Although empirical research into hybridization rates in ants is scant, the limited data presented in this thesis suggest that ants are consistent with respect to broader patterns of hybridization, and they do not hybridize at a higher or lower rate than any other group of organisms. More studies are needed to determine if ants hybridize more readily than other groups.

Where does ant hybridization occur?

Feldhaar et al. (2008) compiled a list of all known instances of hybridization between various ant species. Several new cases have been identified since that publication, and their list has been expanded here to include these (Table 1.1). In total, there are 30 cases in ants where there is evidence that F1 hybrids of any caste are present in the wild population.

Europe represents the best-studied geographical area with 19 of 30 cases occurring there. North America and Asia have only five and four cases each, respectively. There are two in South America, one in Africa, and no cases have been described from Australia or any Pacific islands. This discrepancy in the number of cases of hybridization by continent is undoubtedly due to study bias. Europe, particularly central Europe, is the most extensively studied area, and it has by far the highest number of described cases of interspecific hybridization. More research in less well-studied areas will surely uncover more cases of hybridization in ants. In fact, several studies have uncovered cases of likely hybridization, but further work is needed to confirm this (Eyer et al., 2017; Feldhaar et al., 2003; Feldhaar et al., 2010; Pringle et al., 2012; Schlick-Steiner et al., 2005; Seifert, 1999).

Genetic caste determination (GCD)

Environmental factors have long been the main focus of caste determination research (Anderson et al., 2008a), but a genetic bias to caste determination has been claimed in many different eusocial species. The most common bias detected is that between different worker subcastes. There is a clear genetic component to size in all animal species (Conlon & Raff, 1999), and in many ant species this manifests in a genetic component to caste development with workers of different sizes taking different roles in the colony (Schwander et al., 2005). These differences can involve differences in size or morphology (Hughes et al., 2003; Jaffé et al., 2007; Rheindt et al., 2005) or predisposition to different behavioral tasks (Fraser et al., 2000; Julian & Fewell, 2004; Schwander et al., 2005; Smith et al., 2008; Stuart & Page, 1991).

There can also be a genetic component to different queen morphs. In populations of the fire ant *Solenopsis invicta*, there is a locus with two alleles, *B* and *b*, that controls queen size and colony monogyny or polygyny. Being homozygous for the *B* allele makes a queen larger and head of a monogynous colony, whereas being heterozygous makes them smaller and gives rise to polygynous colonies. Workers that have the *b* allele apparently recognize the presence of the same allele in queens, and attack and remove *BB* queens that initiate reproduction from polygynous colonies. Workers in monogynous colonies are aggressive to foreign queens regardless of genotype. Being homozygous for the *b* allele is lethal (Keller & Ross, 1998).

Another example comes from the slave-making ant *Harpagoxenus sublaevis*. Queens of this species can be gynomorphic (winged) or ergatomorphic (wingless). This polymorphism is under the control of a single locus with two alleles. Queens that are homozygous for the recessive allele, *e*, can be either gynomorphic or ergatomorphic, while queens with the genotypes *EE* or *Ee*

are always ergatomorphs. The dominant allele, E, likely increases some inhibitory effect on the likelihood of a larva developing into a gynomorphic queen. The two morphs have no apparent difference in fecundity but show a difference in caste ratios so that the EE and Ee genotypes produce more worker offspring. A likely explanation for the maintenance of this system is that balancing selection keeps both morphs present in the population. Gynomorphic queens have a fitness advantage by producing more reproductive offspring while ergatomorphic queens produce a higher proportion of workers. More workers presumably increase the chance of successful slave raids, which are necessary for the survival of these colonies (Buschinger & Winter, 1975; Winter & Buschinger, 1986).

One more example comes from *Leptothorax species A* from Quebec, Canada. As in *H. sublaevis*, there are two queen morphs that appear to be under the control of a single locus with two alleles. These alleles are similarly called E and e, but in neither species has the specific locus been identified. The dominant allele is hypothesized to suppress the development of ocelli, wings, and queen-like thoracic structures in larvae that are destined to become queens, causing the genotypes EE and Ee to become intermorphic queens (so-called because they show intermediate morphology between queens and workers). Queens with the ee genotype become gynomorphic queens. The different morphs are further hypothesized to be maintained by trade-offs in mating strategy that are suited to different environments. Gynomorphic queens fly several meters away to mate and then fly even farther to found a new colony, while intermorphic queens mate near their home nest and tend to live in patchier habitats where flying queens may become lost and unable to found a new colony (Heinze & Buschinger, 1986; Heinze & Buschinger, 1989).

Perhaps the more interesting form of genetic caste determination is that between queens and workers because this gives rise to genetic conflict within the population. Patrilineal that give

rise to reproductive queens should have higher fitness than those that give rise to non-reproductive workers (Anderson et al., 2008a, Linksvayer et al., 2006). Males who produce workers should be selected against because their effective fitness is zero when workers do not reproduce. However, workers are required for the normal function of eusocial colonies, so a reduction in the number of workers produced would lead to death of the colony. This creates a genetic conflict between queens and males. Queens need males that give rise to workers for the normal function of the colony, while males gain a significant fitness advantage by contributing primarily or exclusively to the reproductive caste. Several cases have been observed in ants where there is a genetic component to differential caste development between workers and queens. While the contributions of genetics and environment almost certainly lie on a continuum ranging from completely environmentally determined to completely genetically determined caste, Schwander et al. (2010) and Anderson et al. (2008a) provide a useful framework for discussing GCD by dividing the observed cases into several classes; those with a small genetic component to caste determination, and those with a strong genetic component where caste is primarily or exclusively under genetic control.

Weak genetic caste determination

Several studies have found that some patrilineages were more likely to be represented in the reproductive offspring, giving the appearance of a genetic component to caste determination. In *Acromyrmex echinator*, 20% of patrilineages were overrepresented among new queens (Hughes & Boomsma, 2008). In *Pogonomyrmex badius* (Smith et al., 2008) and *Formica truncorum* (Keller et al., 1997), 4% and 12.5% of patrilineages were similarly overrepresented, respectively. In *Pogonomyrmex rugosus*, different crosses of males and queens in field colonies produced different proportions of queens and workers, suggesting that the apparent bias results from the interaction

of maternal and paternal genomes rather than some lineages being predisposed to queen development (Schwander & Keller, 2008). Another study compared the reproductive outputs of different queens of the ant *Cardiocondyla kagutsuchi*. New queens from two stock colonies were mated with males from a third colony. The two stock colonies produced similar numbers of offspring over their lifetimes, but the ratio of queens to workers differed, suggesting that eggs in either colony differ in their propensity to become queens due to maternal or genetic effects (Frohschammer & Heinze, 2009). While it is possible that some patriline may be genetically biased towards production of queens over workers, other factors may also play a role. In *Pogonomyrmex occidentalis*, an apparent genetic component to caste determination may be more easily explained by patriline shifting (Wiernasz & Cole, 2010) where colonies raise genetically distinct cohorts at different times that differ in caste composition. The difference in patriline contribution to different castes may not be the result of differential allocation of patrilines, but of which ejaculate is used to fertilize any given cohort. Furthermore, colonies may adjust the caste ratio of each brood cohort according to the needs of the colony. In *Pheidole pallidula*, soldier production was observed to increase when a colony was presented with nearby competition from conspecific colonies (Passera et al., 1996). Colonies of *Pheidole flavens* were observed to alter their worker caste ratios in response to food availability (McGlynn & Owen, 2002). *Monomorium pharaonis* was observed to alter production of sexuals to meet the current colony requirements for growth and fitness (Warner et al., 2018). Reproductive cohorts could be produced only at certain times, such as during mating season or in response to changing colony conditions, and the overrepresentation of certain patrilines within in the reproductive caste could be explained by temporal variation in sperm use. Furthermore, sperm clumping, which makes the sperm of various

Table 1.1 Cases of hybridization in ants

Species	Evidence	Caste affected	Continent	Specific locality	Reference
<i>Crematogaster claudiae</i> ^a x					
<i>C. capiosa</i>	Morphology, nuclear markers, mtDNA	workers, queens	Asia	Malaysian Borneo	Feldhaar et al., 2003; Feldaar et al., 2010
<i>Cataglyphis israelensis</i> x					
<i>C. savignyi</i>	Nuclear markers, mtDNA	workers	Asia	Israel	Eyer et al., 2017
<i>Dorylus wilverthi</i> x					
<i>D. molestus</i>	Morphology, nuclear markers, mtDNA	workers	Africa	Kenya	Kronauer et al., 2011b
<i>Formica bruni</i> x					
<i>F. pressilabris</i>	Morphology	workers	Europe	Switzerland	Seifert, 1999
<i>Formica cinerea</i> x	Nuclear markers, cuticular hydrocarbon profiles, aggression assays	workers	Europe	Switzerland	Purcell et al., 2016
<i>F. selysi</i>			Europe	France; Yekaterinburg, Russia	Seifert & Goropashnaya, 2004
<i>F. pratensis</i>	Morphology, mtDNA	workers	Europe, Asia	Russia	Bernasconi et al., 2011; Seifert & Goropashnaya, 2004
<i>Formica lugubris</i> x			Asia	Lake Baikal, Russia	Kulmuni et al., 2010; Seifert & Goropashnaya, 2004; Beresford et al., 2017
<i>F. aquilonia</i>	Morphology, mtDNA	workers	Europe	Yekaterinburg, Russia	
<i>Formica polyctena</i> x					
<i>F. aquilonia</i>	Morphology, mtDNA	workers	Europe	Finland; Sweden	Seifert & Goropashnaya, 2004
<i>Formica rufa</i> x					
<i>F. lugubris</i>	Morphology, mtDNA	workers	Europe	Germany	Czechowski, 1996; Goropashnaya et al., 2004; Seifert, 1991; Seifert et al., 2010
<i>Formica rufa</i> x					
<i>F. polyctena</i>	Morphology, mtDNA	workers	Europe	England	Pearson, 1983a; Pearson, 1983b
<i>Lasius alienus</i> x					
<i>L. niger</i>	Morphology, nuclear markers	workers	Europe	England	Umphrey, 2006; Umphrey & Danzman, 1998; Wing, 1968
<i>Lasius latipes</i> x			North America	Ontario, Canada	Pearson, 1983a; Seifert, 1999; van der Have et al., 2011
<i>L. claviger</i>	Morphology, nuclear markers	workers, queens	Europe	Denmark; England	Steiner et al., 2011
<i>Lasius niger</i> x					
<i>L. psammophilus</i>	Morphology, nuclear markers	workers, queens, males	Europe	Italy	Bagherian Yazdi et al., 2012
<i>Messor cf. wasmanni</i> x					
<i>M. capitatus</i>	Morphology, nuclear markers, mtDNA	workers	Europe	Germany	Anderson et al., 2006; Julian et al., 2002; Schwander et al., 2007a; Schwander et al., 2007b
<i>Myrmica scabrinodis</i> x					
<i>M. vandeli</i>	Morphology	workers	Europe	North America	Anderson et al., 2007b
<i>Pogonomyrmex barbatus</i> x	Morphology, nuclear markers, mtDNA	workers, queens	North America	Arizona; New Mexico	
<i>P. rugosus</i>			North America	Arizona; Utah	Anderson et al., 2008b
<i>Pogonomyrmex occidentalis</i> x					
<i>P. maricopa</i>	Morphology, nuclear markers, mtDNA	workers	America		

Table 1.1 continued

Species	Evidence	Caste affected	Continent	Specific locality	Reference
<i>Solenopsis geminata</i> x	Morphology, nuclear markers	workers	North America	Texas	Helms Cahan & Vinson, 2003; Hung & Vinson, 1977
<i>S. xyloni</i>					
<i>Solenopsis invicta</i> x	Venom composition, cuticular hydrocarbon profiles, morphology,	workers, queens	North America	Mississippi	Ross et al., 1987; Ross & Robertson, 1990; Vander Meer et al., 1985
<i>S. richteri</i>					
<i>Solenopsis invicta</i> x			South America	Northern Argentina	Ross & Shoemaker, 2005; Shoemaker et al., 2006
<i>S. quinquecupis</i>	Morphology, nuclear markers, mtDNA	workers	South America	Northern Argentina	Ross & Shoemaker, 2005; Shoemaker et al., 2006
<i>Solenopsis richteri</i> x	Morphology, nuclear markers, mtDNA	workers	South America	Northern Argentina	
<i>S. quinquecupis</i>					
<i>Tapinoma darwini</i> x					
<i>T. magnum</i>	Morphology, mtDNA	workers	Europe	Spain; France	Seifert et al., 2017
<i>Temnothorax unifasciatus</i> x		workers, queens	Europe	Sweden; Germany	Douwes & Stille, 1991; Siefert, 1999
<i>T. nigriceps</i> ^b	Morphology, nuclear markers				
<i>Temnothorax albipennis</i> x		workers	Europe	Sweden; Germany	Douwes & Stille, 1991; Siefert, 1999
<i>T. nigriceps</i> ^b	Morphology, nuclear markers	workers, queens	Europe	Sweden; Germany	Douwes & Stille, 1991; Siefert, 1999
<i>Temnothorax unifasciatus</i> x					
<i>T. albipennis</i> ^b	Morphology, nuclear markers	workers	Europe	Sweden; Germany	Douwes & Stille, 1991; Siefert, 1999
<i>Temnothorax tuberculatus</i> x					
<i>T. albipennis</i> ^b	Morphology, nuclear markers	workers	Europe	Sweden; Germany	Douwes & Stille, 1991; Siefert, 1999
<i>Temnothorax parvulus</i> x					
<i>T. lichtensteini</i> ^b	Morphology, laboratory experiment	queens	Europe	France	Plateaux, 1979
<i>Temnothorax crassispinus</i> ^c x	Morphology, nuclear markers, mtDNA, cuticular hydrocarbon profiles	workers, queens	Europe	Germany	El-Shehaby et al., 2011; Pusch et al., 2006a; Pusch et al., 2006b; Siefert, 1995
<i>T. caespitum</i> ^b		workers, likely			
<i>Tetramorium immigrans</i> x		queens	Europe	France	Cordonnier et al., 2019
<i>T. caespitum</i>	Nuclear markers, mtDNA				
<i>Tetramorium alpestre</i> x					
<i>T. impurum</i>	Morphology, mtDNA, biogeography	workers	Europe	Switzerland	Steiner et al., 2010

^a *Crematogaster claudiae* is referred to as *Crematogaster* msp. 10 in listed reference.

^b The genus *Temnothorax* was previously considered a subgenus of *Leptothorax*. These species are in the genus *Leptothorax* in earlier publications.

^c *Temnothorax crassispinus* is referred to as *Leptothorax sodidulus* in earlier publications.

males unevenly distributed within the queen's spermatheca, is necessary for patriline shifting to occur, and has been observed in several species of *Eciton* (Whelden, 1963) and in *Formica truncorum* (Sundström & Boomsma, 2000), and suggests that patriline shifting can occur in many other species that vary the caste composition of different brood cohorts.

Strong genetic caste determination

Strong genetic effects on caste determination have been observed at least eight times in ants. These systems are described below (Figure 1.1).

Vollenhovia emeryi

In *Vollenhovia emeryi*, there are two distinct queen morphs: the S morph and the L morph, corresponding to colonies that produce only short-winged and long-winged queens, respectively (Ohkawara et al., 2006). In both morphs, new queens are almost exclusively homozygous, and are produced via thelytokous parthenogenesis, while workers are almost exclusively heterozygous and produced via sexual reproduction. Males of the S morph often contain alleles that are present in S workers but not in the queen, indicating that they share alleles with the queen's mate (Ohkawara et al., 2006). Furthermore, in the nuclear genome, both S and L males are genetically more similar to L queens than to S queens, but in the mitochondrial genome, males are more similar to queens of their own morph (Kobayashi et al., 2008; Kobayashi et al., 2011). This evidence indicates that males emerge from eggs laid by the queen, but as clones of the queen's mate (androgenesis) rather than haploid clones of the queen. The workers of the two morphs are not morphologically

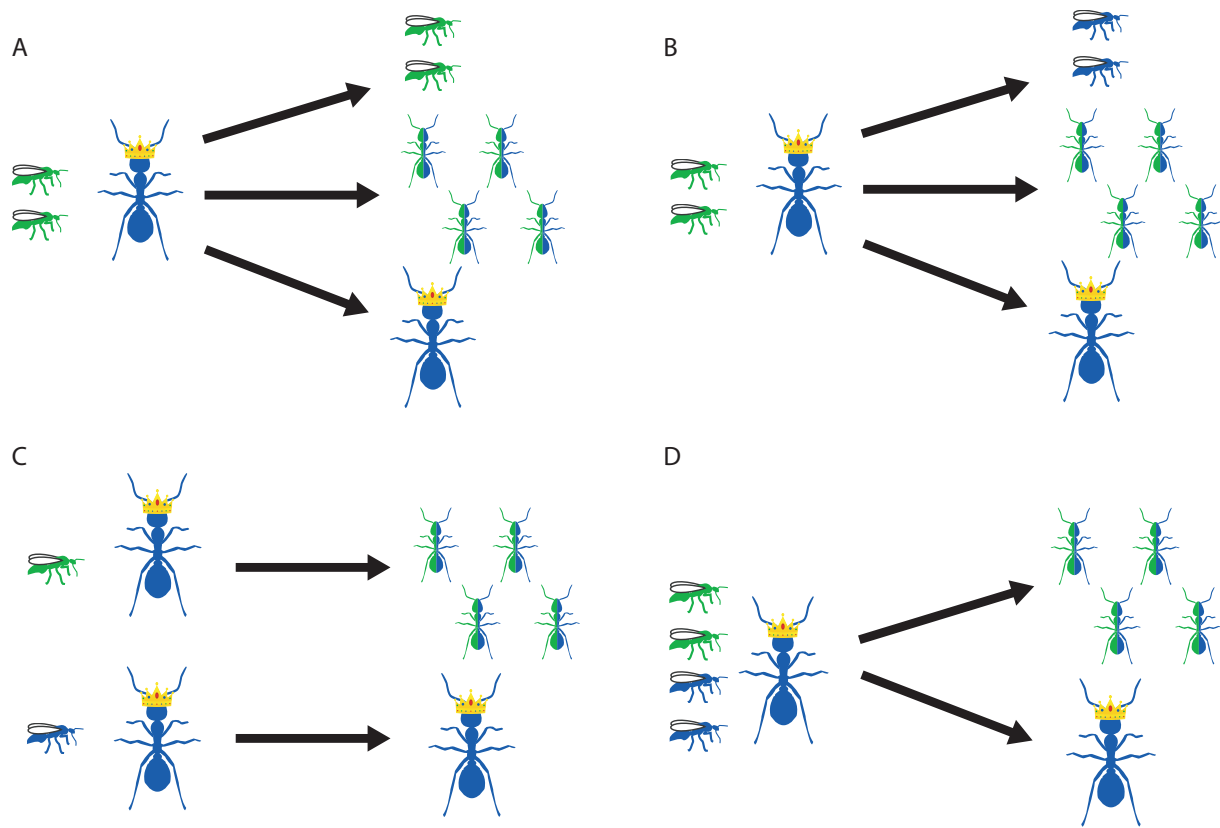


Figure 1.1 Cartoon depicting the different systems of genetic caste determination. A) GCD system in *Vollenhovia emeryi*, *Wasmannia auropunctata* and *Paratrechina longicornis*. Males are produced via androgenesis and queens are produced via thelytokous parthenogenesis. Workers are produced via normal sexual reproduction. B) GCD in *Cataglyphis cursor* and *C. hispanica*. Workers are produced via normal sexual reproduction, and males are produced via arrhenotokous parthenogenesis as is normal for ants. Queens are clones of their mother queen. C) Genetic caste determination in *Solenopsis xyloni*. Colonies have multiple singly-mated queens in each colony. Some queens mate with *S. geminata*, and others with their own species. New queens are produced from conspecific matings and workers are produced from heterospecific matings. D) Genetic caste determination in *Pogonomyrmex* and *Messor*. These species have monogynous colonies, whose queens are multiply mated. Colony survival depends on queens mating with both interlineage and intralineage males. New queens are produced from intralineage matings, and workers are produced from interlineage matings.

distinguishable, but there is clearly no gene flow between them since queens are produced via parthenogenesis, and any alleles transmitted to the offspring via hybridization between the lineages appear only in the sterile worker caste (Ohkawara et al., 2006). The similarity of S males to L queens in the nuclear genome is likely the result of a past hybridization event (Kobayashi et al., 2011), although it is not clear that the separate lineages constitute different species.

Wasmannia auropunctata

In *Wasmannia auropunctata*, analysis of 11 microsatellite loci in samples from 34 colonies showed that queens were produced clonally while workers were produced via normal sexual reproduction. Further analysis revealed that males were also produced clonally, but not via arrhenotokous parthenogenesis as is normal for ants (Foucaud et al., 2007; Foucaud et al., 2010; Fournier et al., 2005a). Pupal male genotypes were identical to the genotypes of sperm found in the queens spermathecae indicating that males were clones of the queen's mate. Like in *Vollenhovia emeryi*, males are produced via androgenesis, and there is no gene flow between the two sexes, which form independent lineages (Foucaud et al., 2010; Fournier et al., 2005a). In experimental crosses, queens from clonally reproducing colonies produced haploid males almost exclusively via androgenesis, regardless of whether they mated with males from clonal or sexual populations. This indicates that androgenesis in *W. auropunctata* is not a male trait; rather it is a trait of parthenogenetic females (Rey et al., 2013). Unlike other similar cases, queen parthenogenesis and male androgenesis do not appear to result from historic or current hybridization between two distinct genetic lineages (other than those formed by the different sexes). Instead, there are multiple independent origins of clonality arising out of sexual

populations, a pattern that is consistent in distantly located parts of the species range (Foucaud et al., 2007).

Cataglyphis cursor

In a monogynous population of *Cataglyphis cursor* 97.3% of workers had alleles that were not present in the queen. In contrast, 96.4% of gynes produced in these colonies had only alleles that could be attributed to the queen. These data suggest that queens are mostly produced parthenogenetically while workers are mostly produced sexually. It is unlikely that queens are produced sexually because the probability of a male having no diagnostic alleles at several highly polymorphic microsatellite loci is extremely low. In contrast to other cases of genetic caste determination, males and queens appear to come from the same gene pool (Pearcy et al., 2004).

Cataglyphis hispanica

Colonies of *Cataglyphis hispanica* are monogynous and queens are usually singly mated. In field-collected colonies, all new gynes produced were identical to the queen at all loci that were genotyped indicating that they were produced clonally. As in *C. cursor*, the probability that the queen mated with a male with no diagnostic alleles is extremely low. Pedigree analysis of the workers revealed that all were produced sexually. All males produced in these colonies were haploid and carried alleles of the colony queen indicating that they were produced via arrhenotokous parthenogenesis. In all colonies queens and their mates belonged to different genetic groups, and all workers were interlineage hybrids (Leniaud et al., 2012), a pattern that holds across the entire range of the species (Darras et al., 2014).

Solenopsis

Species of the genus *Solenopsis* are polygynous and their queens are singly mated. The ranges of two species, *S. xyloni* and *S. geminata*, overlap in central Texas (Hung & Vinson, 1977; Vinson, 1997). Colonies of *S. geminata* from sympatric populations are morphologically and genetically indistinguishable from colonies of the same species in areas of allopatry. In colonies of *S. xyloni* sympatric with *S. geminata*, all workers display some degree of intermediate morphology and are also genetically intermediate between the two species. Nearly all workers in these colonies are F1 hybrids, while nearly all queens are pure species *S. xyloni* (Helms Cahan & Vinson, 2003). This case is clearly an example of interspecific hybridization, but how it relates to genetic caste determination is not clear.

Pogonomyrmex

Pogonomyrmex colonies are monogynous and queens are multiply mated. *P. rugosus* and *P. barbatus* are two closely related species that live in the south western United States and Mexico. Their ranges overlap, and in some localities they are found near each other (Parker & Rissing, 2002; Volny & Gordon, 2002). In some of these areas of sympatry, hybrid lineages have been identified based on deviations from Hardy-Weinberg equilibrium as well as differences in morphology (Julian et al., 2002; Volny & Gordon, 2002). Within two of these areas of sympatry, there are multiple pairs of interdependent lineages that resemble either *P. rugosus* or *P. barbatus* morphologically, but that are reproductively isolated from each other as well as from their parental species (Anderson et al., 2006; Helms Cahan & Keller, 2003; Schwander et al., 2007a). At the locality called Hidalgo, H1 and H2 lineages are found, and at Junction, J1 and J2 lineages are found. The lineage pairs at each locality are interdependent such that each cannot exist without the

presence of the other because foundress queens must mate with males of both lineages. Mating between individuals belonging to two different lineages (H1 x H2, or J1 x J2) is necessary to produce workers. Meanwhile intralineage matings are necessary for the production of virgin queens. Thus, during their nuptial flight, new queens must mate with at least one male from each lineage to ensure successful colony founding. There is evidence of historical gene flow indicating that there is a complex history of hybridization between the two parental species, but it is not clear that interspecific hybridization directly gave rise to the system of interdependent lineages observed (Helms Cahan & Keller, 2003; see also Anderson et al., 2006).

Messor

A system of genetic caste determination remarkably similar to that seen in *Pogonomyrmex* is observed in *Messor barbarus*. Queens of this species are multiply mated, and colonies are headed by a single queen. Two independent genetic lineages are observed, and queens must mate with males of both lineages to successfully found a colony. Interlineage matings produce workers while intralineage matings produce queens (Norman et al., 2016; Romiguier et al., 2017). A similar system is likely also occurring in two other *Messor* species, *M. structor* and *M. ebeninus* (Romiguier et al., 2017).

Paratrechina longicornis

Paratrechina longicornis uses a similar mode of reproduction to *W. auropunctata* and *V. emeryi*. Queens are produced clonally via thelytokous parthenogenesis, while workers are produced sexually. Males are produced via androgenesis i.e. they are clones of the queen's mate. Although queens and males appear to form distinct genetic lineages, the origin of the unusual

reproductive system in this species is unclear, and it is not known if it is connected to interspecific or interlineage hybridization (Pearcy et al., 2011).

Hybridization and genetic caste determination

Of the eight cases of genetic caste determination described thus far, those in *Pogonomyrmex* and *Solenopsis* show clear evidence of hybridization between distinct species. The *Pogonomyrmex* system involves a complex history of hybridization between *P. barbatus* and *P. rugosus* that may have given rise to the system of interdependent lineages observed. However, there is also hybridization between two other *Pogonomyrmex* species in nearby areas but no evidence of strong GCD (Anderson et al., 2008b). In *Solenopsis*, the two species exist in a current hybrid zone that may ultimately be responsible for genetic caste determination. Furthermore, the

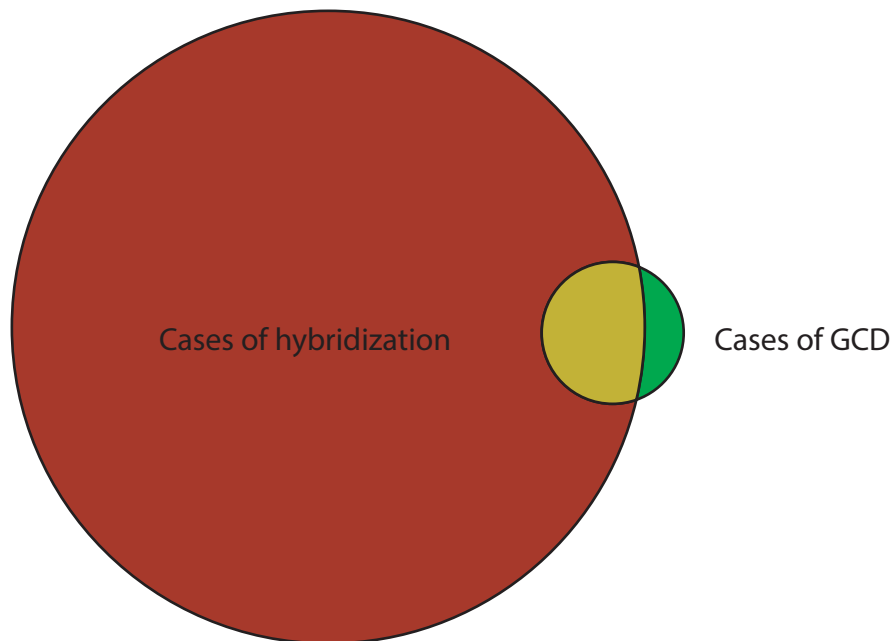


Figure 1.2 Cartoon Venn diagram depicting the high amount of overlap between cases of genetic caste determination in ants and interspecific hybridization.

systems in *Messor*, *Cataglyphis hispanica*, *Vollenhovia emeryi*, and *Paratrechina longicornis* may indeed involve interspecific hybridization in the normal sense, although further research is needed to clarify if the involved lineages constitute true species. Only two cases, those in *Wasmannia auropunctata* and *Cataglyphis cursor*, show multiple independent origins of GCD from within otherwise normal sexually reproducing populations. However, their independent lineages, as well as the lineages in all strong GCD cases, may be considered distinct species in the sense that they are reproductively isolated from each other (Queller, 2005). It is evident that there is a connection between the two phenomena of hybridization and genetic caste determination although how they are related is not clear (Figure 1.2).

Under polygyny or polyandry, effects of hybrid sterility can be averted by shunting hybrids out of the reproductive caste, so there may be selective pressures for the queen to reduce the metabolic cost of producing sterile daughter queens. Additionally, as outlined previously, the cost of keeping hybrid workers in the colony may be relatively low, and some benefits may even be realized if hybrids display any form of hybrid vigor. A proposed explanation for the evolution of strong genetic caste determination is that it is the result of hybridization between two independent genetic lineages. Genes that bias offspring to become queens are selected for because they increase the number of sexual offspring. However, workers are required for colony maintenance and brood care, so colony level selection acts on different loci to counteract these caste-biasing genes. Within a non-hybridizing population, these two competing forces counteract each other to create a stable evolutionary strategy. However, when two different species interbreed, the genes causing and counteracting queen bias are decoupled revealing the underlying genetic conflict and allowing a strong caste bias to rise in frequency (Anderson et al., 2008a). Another possible explanation is that because hybridization can have an effect on body size, growth-stunted hybrids do not reach the

size threshold during development required to become queens (Trible & Kronauer, 2017). Cases where interspecific matings result in larger hybrids would cause some to surpass this threshold and become queens. However, these situations would be selected against if hybrid queens have reduced fecundity or are completely sterile.

Chapter 2: Development of broadly applicable microsatellite markers in ants

Introduction

To facilitate looking for cases of natural hybridization and to determine if any constitute cases of genetic caste determination, I developed a set of microsatellite primers that are universal to all ants. Newly discovered and less well-studied species have limited genetic tools available, and there is a high cost to developing them for a new species. I have developed a set of 45 microsatellite primer pairs that have conserved binding sites but variable microsatellite sequences. I developed them using eight ant genomes that were available at the time and tested a subset of them on six different ant species from six different ant subfamilies. This means that for any given ant species, a subset of these primers can be used for population genetics studies without the need to spend time or money developing new loci.

Microsatellites, also called short tandem repeats (STRs) or simple sequence repeats (SSRs), are sequential repeats of 1 to 6 base pair motifs that have been used as genetic markers for more than 20 years (Litt & Luty, 1989; Tautz, 1989; Weber & May, 1989). Often found in noncoding regions, they are common in the genomes of eukaryotes (Ellegren, 2004; Molnar et al., 2012; Schlötterer, 2004). An important feature of these sequences is their high degree of length polymorphism within populations of single species, which has been attributed to DNA polymerase slippage during replication (Buschiazzo & Gemmell, 2006; Leclercq et al., 2010). This can result in a large number of alleles per locus that differ from one another in the number of repeats, making them distinguishable by size alone. This high degree of polymorphism and the ease of genotyping make them particularly suitable for studies in population genetics and pedigree analyses (Jarne & Lagoda, 1996; Selkoe & Toonan, 2006). For example, microsatellites have been used to measure

population differentiation and hybridization (Hansson et al, 2012; Kronauer et al., 2011b), to investigate ploidy levels (Jungman et al., 2010; Mishina et al, 2014), and to reconstruct parentage and pedigrees in wild and domestic populations (Kuo et al., 2014; Wang & Scribner, 2014). Microsatellites are comparatively cheap to genotype and can be used with low concentrations of DNA. Furthermore, they typically have more alleles per locus than single nucleotide polymorphisms (SNPs) and thus provide more information per locus (Gärke et al., 2011). Although they often have a high degree of polymorphism within species, some microsatellite loci can be conserved across species that diverged 100 million years ago or more (Buschiazzi & Gemmell, 2009; Buschiazzi & Gemmell, 2010; Ezenwa et al., 1998; FitzSimmons et al., 1995; Moore et al., 1991; Moore et al., 1998; Stolle et al., 2013).

More recently, next generation sequencing (NGS) techniques have risen in popularity, mainly because of the large number of marker loci they can generate at relatively low per locus cost. For example, restriction site-associated DNA (RAD) tags can generate thousands of markers and have proven instrumental for measuring gene flow between populations (Gagnaire et al., 2013), as well as for reconstructing shallow phylogenies (Rubin et al., 2012). However, the data generated from these techniques can be complex and difficult to analyze. Furthermore, although NGS has a low per locus cost, it has a much higher overall cost than older sequencing methods. Variants of the approach have been developed to reduce the complexity of DNA libraries, such as double digest RADseq (ddRAD) (Peterson et al, 2012) 2b-RAD (Wang et al., 2012), or genotyping by sequencing (GBS) (Elshire et al., 2011), but these still require expensive NGS platforms. At the same time, for many studies a smaller number of markers is sufficient, and markers such as microsatellites can be more attractive.

Despite their utility, a significant impediment to the use of microsatellites is the cost and effort associated with identifying a set of loci and developing PCR primers. Although the same loci can sometimes be useful for studying closely related species, loci that are polymorphic in one species are often not informative in another, and primers quickly lose affinity as species become more divergent. This usually requires new microsatellite loci to be characterized for each studied species. Depending on the research question, studies typically require a set of five to ten or more independent microsatellite loci. Paying a commercial service to develop these markers can be costly, and developing markers independently can be labor intensive and time consuming. Nevertheless, the utility of microsatellites in determining pedigree structures, relatedness and mating systems makes them particularly useful for social insect research because they can be used to address important questions related to inclusive fitness theory, including social organization (e.g. Leniaud et al., 2013), worker caste determination (e.g. Huang et al., 2013), and the evolution of supercolonies (e.g. Seppä et al., 2012). Of the social insects, ants are a particularly speciose and ecologically diverse group being intensively studied. Current estimates place the ant family Formicidae at 115 to 158 million years of age (Moreau & Bell, 2013; Moreau et al., 2006; Brady et al., 2006), and more than 14,000 species have been described, according to the Hymenoptera Name Server (v. 1.5, available from <https://hol.osu.edu>, accessed 1 September 2019). The genomes of 27 ant species have currently been sequenced (Boomsma et al., 2017; Dhaygude et al., 2019; Lau et al., 2019), although only eight were available at the time of this work. However, these eight represent most major ant clades, allowing highly conserved regions to be identified over most of the family. To help overcome the constraints of narrowly applicable primers and to make microsatellites broadly available as population genetic markers, we aimed to develop a set of

microsatellite markers that would be conserved across a wide range of species, yet polymorphic within species.

Materials and Methods

Specimen collection

All specimens of *Ectatomma ruidum* and *Paraponera clavata* were collected at the Organization for Tropical Studies field station in La Selva, Costa Rica. *Simopelta pentadentata* specimens were collected in Monteverde, Costa Rica. *Dorylus molestus* specimens were collected in Kakamega Forest, Kenya. *Lasius nearcticus* specimens were collected at the Rockefeller University Center for Field Research in Millbrook, New York, USA, and specimens of *Solenopsis invicta* were collected in Tallahassee, Florida, USA. Collection permits were acquired for all samples where necessary. A permit for specimens from Kakamega National Park, Kenya was granted by the National Council for Science and Technology (permit number NCST/RCD/12B/012/37B). A permit for specimens from Costa Rica was granted by Ministerio de Ambiente, Energia y Telecomunicaciones (permit number 192-2012-SINAC). Permits were not required for specimens collected in the United States. No protected species were sampled.

Bioinformatics

Bioinformatics were performed by Peter Oxley and Kimberly Siletti

Seven available ant genomes were downloaded from Ant Genomes Portal (hymenopteragenome.org/ant_genome), and our lab has previously published the *O. biroi* genome (Oxley et al., 2014). The genome versions for each species were *At. cephalotes* v1.0, *Ac. echinator* v2.0, *C. floridanus* v3.3, *O. biroi* v2.0, *H. saltator* v3.3, *L. humile* v1.0, *P. barbatus* v3.0, *S. invicta*

v1.0. Microsatellites in the *O. biroi* genome were located using Tandem Repeats Finder (‘TRF’; v. 4.04) (Benson, 1999), which utilizes Smith-Waterman style local alignment. Tandem repeats are reported only if they exceed a minimum alignment score, specified as 50 (Minscore =50). Alignment mismatches were assigned a weight of five (Mismatch =5). Additionally, the size of the repeat pattern was limited to five bases (Maxperiod =5). The microsatellite indices returned were used to generate a masked BLAST query for each microsatellite, extended to include 200-bp flanking regions. The query sequence was used to search all eight sequenced ant genomes, including *O. biroi*, using BLAST (v. 2.2.26+) (Altschul et al., 1990). The results were filtered to remove matches with less than 60% identity. Microsatellite flanking regions that generated unique BLAST hits in all eight genomes were aligned using MUSCLE (Edgar, 2004). To confirm that these conserved flanking regions indeed contained microsatellite sequences, TRF was used to search for microsatellites in all database genomes at the indices returned by BLAST for each hit (settings as stated above). Primer3 software (v. 2.3.4; <http://primer3.sourceforge.net/releases.php>) (Untergasser et al., 2012) generated primers from the consensus sequence in each flanking region. A maximum of four unknown bases were allowed in any primer set (PRIMER_MAX_NS_ACCEPTED =4). All unspecified parameters used the default or recommended settings. Custom Python scripts were used to parse TRF and Primer3 outputs, prepare files for BLAST and Primer3, and filter the BLAST results. These scripts are available upon request from the corresponding author. Initially, 176 loci were identified across all genomes with the described bioinformatics pipeline, from which we chose 45 loci for further study. These 45 loci were chosen subjectively based on the number of perfect repeats in different species and the presence of a microsatellite motif in as many ant genomes as possible.

DNA extraction, PCR amplification and genotyping

DNA was extracted by first homogenizing the tissue in a Qiagen TissueLyser II and then heating the sample at 96°C for 15 minutes in 200 µl of 10% Chelex in TE solution. The samples were then centrifuged at 9100 rpm for three minutes, and the supernatant containing the DNA was removed and used as the template for PCR amplification. The PCR cocktail (10 µl total volume) for all reactions contained 1 µl PCR Gold Buffer (10x), 0.5 µl MgCl₂ (25 mM), 0.5 µl dNTPs (10 mM total, 2.5 mM each), 0.1 µl of each forward and reverse primer (10 mM), 0.1 µl AmpliTaq Gold (5 U/µl), 1 µl DNA template and 6.7 µl H₂O. PCR reactions were run on an Eppendorf Mastercycler Pro S under the following conditions: 10 min at 95°C followed by 40 cycles of 15 s at 94°C, 30 s at 55°C and 30 s at 72°C, and a final extension of 10 min at 72°C. PCR products were sent to a commercial facility (Genewiz, Inc.) for genotyping. Analysis of chromatograms was performed using PeakScanner (Applied Biosystems). Calculations of observed and expected heterozygosity, as well as tests for linkage disequilibrium and deviations from Hardy-Weinberg equilibrium were performed using F-STAT (v2.9.3.2) (Goudet, 1995).

Results

To design a set of broadly applicable microsatellite primers I searched the eight available ant genomes for conserved microsatellite motifs with conserved flanking regions. The eight available ant genomes are from the red harvester ant *Pogonomyrmex barbatus* (subfamily Myrmicinae) (Smith et al., 2011b), Jerdon's jumping ant *Harpegnathos saltator* (subfamily Ponerinae), the Florida carpenter ant *Camponotus floridanus* (subfamily Formicinae) (Bonasio et al., 2010), the leaf-cutting ants *Atta cephalotes* (subfamily Myrmicinae) (Suen et al., 2011) and

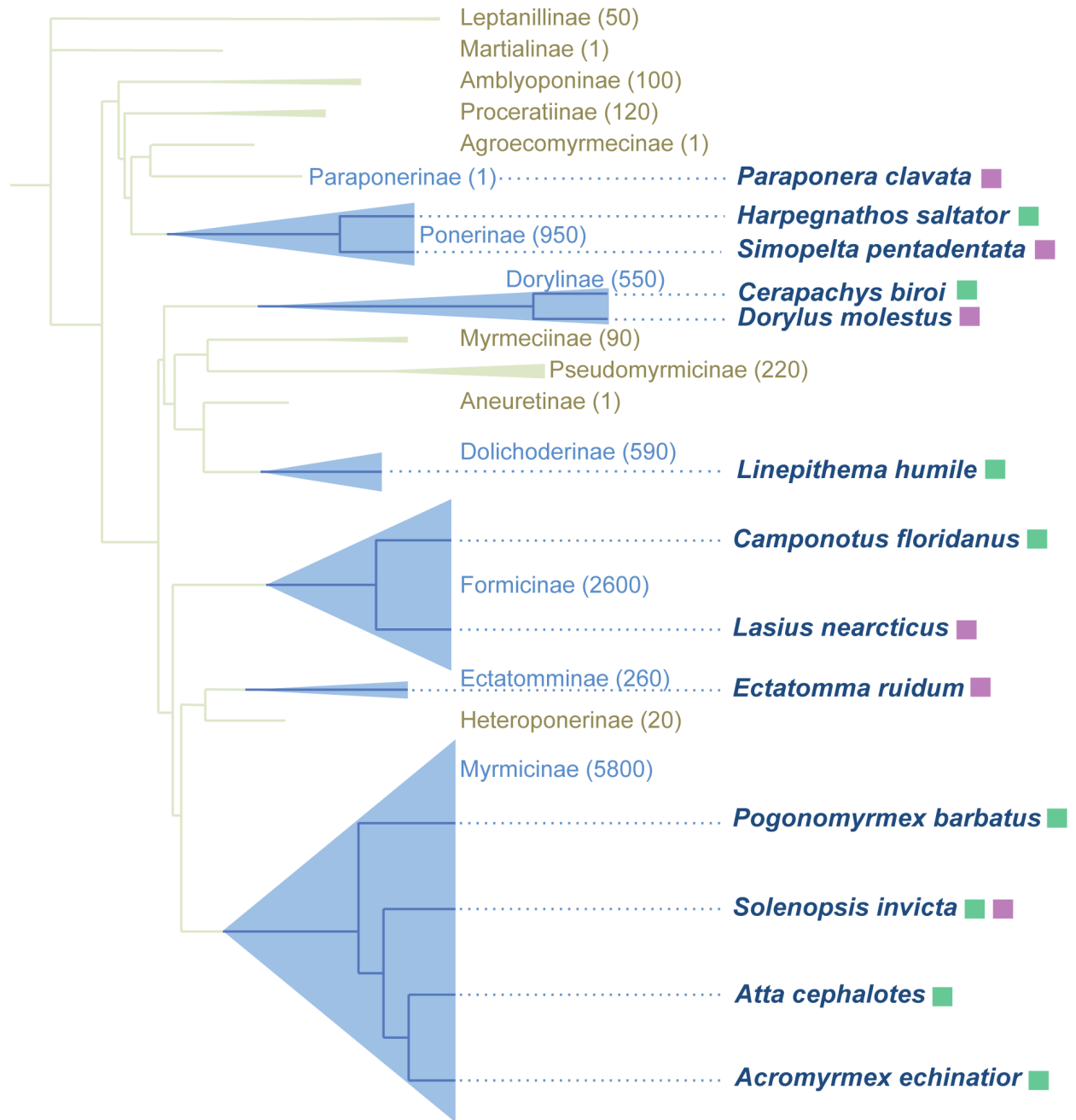


Figure 2.1 Phylogeny of the ants showing the phylogenetic distribution of the species used in this study. The size of each triangle is proportional to the number of species in each group, and the approximate number of species is given in parentheses next to the group name. Boxes next to species names indicate whether that species' genome was used to design (green) or test (purple) the PCR primers. Figure adapted with permission from Libbrecht et al., 2013.

Acromyrmex echinator (subfamily Myrmicinae) (Nygaard et al., 2011), the Argentine ant *Linepithema humile* (subfamily Dolichoderinae) (Smith et al., 2011a), the red imported fire ant *Solenopsis invicta* (subfamily Myrmicinae) (Wurm et al., 2011), and the clonal raider ant *Ooceraea biroi* (subfamily Dorylinae) (Oxley et al., 2014). The available genomes represent five of the 21 recognized extant ant subfamilies, allowing us to select primer sequences that are conserved in a wide range of species across the ants (Figure 2.1).

I identified 176 potential microsatellite loci with conserved flanking regions across all eight genomes, and among those selected 45 that had a repeat motif in most or all of the available genomes (Appendix A). To demonstrate their usefulness in species other than those with available genomes, I tested these primers for amplification in six species from six different subfamilies, only one of which was also used for primer design (*Solenopsis invicta*, subfamily Myrmicinae) (Figure 2.1).

The other five species in which the markers were tested were the bullet ant *Paraponera clavata* (subfamily Paraponerinae), the army ants *Simopelta pentadentata* (subfamily Ponerinae) and *Dorylus molestus* (subfamily Dorylinae), *Lasius nearcticus* (subfamily Formicinae), and *Ectatomma ruidum* (subfamily Ectatomminae). The success of PCR amplification varied by locus and species (Tables 2.1 & 2.2). From those 45 loci, I selected 24 that amplified well in all or most of the six species tested and had at least ten consecutive repeats of their motif in the genomes of more than one of the species with available genome sequences (Appendix A). I genotyped those 24 loci across all six species using fluorescently labeled primers (Applied Biosystems). PCR amplification was successful for all 24 loci in *L. nearcticus* and *D. molestus*, for 23 loci in *S. invicta*, for 22 loci in *P. clavata* and *E. ruidum*, and for 21 loci in *S. pentadentata* (Table 2.2, Figure 2.2). To determine which of the microsatellite loci were polymorphic in any given species,

Table 2.1 PCR amplification success across six ant species for the 21 microsatellite loci that were only tested with unlabeled primers.

Locus	Primer sequence (5'-3')	<i>P. clavata</i>	<i>S. pentadentata</i>	<i>D. molestus</i>	<i>L. nearcticus</i>	<i>E. rudum</i>	<i>S. invicta</i>
Ant21	F-TTCTCGGGAGCAACGCTGGT R-CCATCAGCACTCCACCTCG	Yes	Yes	Yes	Yes	Yes	Yes
Ant508	F-AGCGGATCTAGTGTCTTGG R-ATGGAGGGGAGTAAGAGCGA	Yes	Yes	Yes	Yes	Yes	Yes
Ant1049	F-GAGGATGGGTGGTGGCGGA R-CTGGCGCTCCGTGTGTAT	Yes	Yes	No	Yes	Yes	Yes
Ant1052	F-GGACCTTCTGTCAGGATATC R-CTTTTAGTTCAGACGCACGCG	Yes	Yes	No	Yes	Yes	Yes
Ant1387	F-ATAGGTCCACATACGCGTG R-CACAGCGACTCCCTCTCC	Yes	Yes	Yes	Yes	Yes	Yes
Ant1732	F-ATGATACGATCGAGTGCC R-GCCAGCTCTCCGAGCCTAT	Yes	No	No	Yes	Yes	Yes
Ant2409	F-ATCAGCGTACGATCGAGTT R-CGTGATTCTCTGACGCGAC	Yes	Yes	No	Yes	Yes	Yes
Ant3362	F-CCCCAAAACCTACTCTGTCC R-GTCTCAAGCTCGGATGGA	Yes	Yes	No	Yes	Yes	Yes
Ant3395	F-CCRACGGGCGTGGCGATCC R-CCGGCACTTGGTACAGGTA	Yes	Yes	Yes	Yes	Yes	Yes
Ant3411	F-GCGGACGACGATACACCC R-TGCAGCAGGACCGCGTRGT	Yes	Yes	Yes	Yes	Yes	Yes
Ant3452	F-TGTGAGTGGCGGCGARTGGGA R-ATCGACACAATCGTGGGC	No	Yes	No	Yes	Yes	Yes
Ant3505	F-TTACCGGACAATCGTGTGG R-TGAGCAGACGACGACATCT	Yes	No	Yes	Yes	Yes	Yes
Ant3541	F-TGCAACAAGTGTCTGTAGGT R-TCACATGTTCCGGCGYGAT	No	No	No	No	No	No
Ant4709	F-ACGGGGTAAAGGTTAGGGA R-AGCGATGGGAGATTGGAGAG	No	No	MP	Yes	Yes	Yes
Ant5033	F-TTCCCTCTCCCTGACACC R-TAAGACAAGAACGTCGCG	Yes	No	Yes	Yes	No	Yes
Ant7204	F-GCCCAATCTCTGCATTCT R-CCGCGAAAAGTCCATTTCG	Yes	Yes	MP	Yes	Yes	Yes
Ant8544	F-GGGGTGGTGGCAGTCTCGT R-CAATGCGATCTAGGTACCA	Yes	Yes	Yes	MP	No	Yes
Ant9564	F-TTAGAGGCGCAGSCTGCT R-AGCGAGCACTTCGATGACT	Yes	Yes	No	Yes	Yes	Yes
Ant10290	F-CGTTTCAAAATTAAAGTTTTC R-ACGCGGCTTCGCGCTCGG	No	No	No	No	No	No
Ant10427	F-AATCAGCTTAGCGGCTAA R-ATCCACGCTCTGGGATTC	Yes	Yes	No	Yes	No	Yes
Ant11610	F-GGATAYTGGGCGCGGTCAA R-GCCGAAAGTGTGGATACCTC	No	No	No	No	No	No

See Table 2.2 for details on the remaining 24 loci that were tested using labeled primers. “Yes” indicates clear amplification of a single product. “No” indicates no amplification of any product. “MP” indicates that there were multiple products from which the desired product could not be determined.

Table 2.2 Characteristics of 24 microsatellite loci tested in six different ant species using labeled primers.

Locus	Paragonera clavata					Sinopelta pentadentata					Dorylus molestus					Lasius neorcticus					Ectatomma ruidum					Solenopsis invicta										
	n	A	Size range (bp)	H ₀	H _E	Deviates from HWE	n	A	Size range (bp)	H ₀	H _E	Deviates from HWE	n	A	Size range (bp)	H ₀	H _E	Deviates from HWE	n	A	Size range (bp)	H ₀	H _E	Deviates from HWE	n	A	Size range (bp)	H ₀	H _E	Deviates from HWE						
primer sequence (5'-3')																																				
Ant20	10	1	137	0	0	no	10	3	171-177	0.3	0.54	no	10	1	150	0	0	no	10	1	74	0	0	no	10	1	140	0	0	no	10	1	153	0	0	no
Ant575	10	4	370-379	0.9	0.63	no	10	11	334-375	0.4	0.96	yes	10	2	230-250	0.2	0.19	no	10	4	209-234	0.3	0.37	no	10	1	248	0	0	no	10	3	218-239	0.7	0.63	no
Ant859	10	5	184-206	0.7	0.77	no	10	3	180-184	0.3	0.54	no	10	2	197-199	0.2	0.19	no	9	11	175-204	1	0.94	no	10	1	158	0	0	no	10	1	191	0	0	no
Ant1343	10	4	229-235	0.6	0.53	no	10	1	186	0	0	no	10	4	263-269	0.6	0.76	no	10	3	206-211	0.4	0.58	no	10	1	221	0	0	no	10	4	252-272	0.9	0.71	no
Ant1368	10	1	251	0	0	no	7	6	266-313	0.14	0.93	yes	10	8	299-322	0.9	0.85	no	10	5	278-309	0.6	0.62	no	10	1	269	0	0	no	10	1	280	0	0	no
Ant2341	no amplification						10	5	345-359	0.7	0.76	no	10	4	256-267	0.2	0.55	no	10	2	212-215	0.4	0.51	no	10	1	184	0	0	no	10	2	245-251	0.3	0.27	no
Ant2794	10	3	241-251	0.5	0.42	no	9	9	280-336	0.67	0.9	no	10	5	246-270	0.4	0.77	no	10	10	240-268	1	0.89	no	10	1	218	0	0	no	9	1	258	0	0	no
Ant2936	no amplification						10	7	314-336	0.3	0.92	yes	10	9	352-390	0.4	0.9	no	no amplification					10	5	349-365	0.1	0.81	yes							
Ant3648	9	1	337	0	0	no	10	10	368-410	0.5	0.94	yes	10	7	376-421	0.5	0.83	no	10	3	332-343	0.6	0.57	no	10	4	393-401	0.3	0.67	no	10	1	337	0	0	no
Ant3653	10	1	273	0	0	no	10	9	238-254	0.8	0.85	no	10	3	255-259	0.6	0.62	no	10	9	261-319	0.4	0.9	yes	10	4	357-363	0.4	0.74	no	10	2	254-256	0.4	0.33	no
Ant3993	8	7	368-387	0.88	0.88	no	10	5	368-379	0.2	0.81	yes	10	2	311-317	0.5	0.48	no	10	7	379-419	0.7	0.77	no	8	1	454	0	0	no	10	3	375-363	0.7	0.47	no
Ant4155	10	1	162	0	0	no	8	3	206-211	0.38	0.64	no	10	4	176-195	0.8	0.61	no	10	1	170	0	0	no	10	1	158	0	0	no	10	2	200-203	0.2	0.19	no
Ant5035	10	2	340-342	0.4	0.33	no	10	9	412-442	0.3	0.94	yes	10	6	365-384	0.8	0.77	no	10	8	284-341	0.8	0.9	no	10	1	331	0	0	no	10	1	311	0	0	no
Ant7249	10	1	425	0	0	no	10	7	320-359	0.4	0.86	yes	10	6	369-398	0.5	0.68	no	10	5	345-368	0.6	0.74	no	10	1	325	0	0	no	10	1	358	0	0	no
Ant7680	10	1	306	0	0	no	9	11	332-386	0.56	0.97	yes	10	6	310-328	0.6	0.74	no	10	1	257	0	0	no	10	1	219	0	0	no	10	1	264	0	0	no
Ant8424	10	2	262-265	0.2	0.19	no	10	8	894-318	0.5	0.82	no	10	3	232-238	0.5	0.48	no	10	4	193-240	0.4	0.36	no	10	2	266-275	0.4	0.44	no	10	3	235-259	0.4	0.35	no
Ant8498	10	2	214-218	0.4	0.51	no	10	1	181	0	0	no	10	1	147	0	0	no	10	1	145	0	0	no	10	1	172	0	0	no	10	1	201	0	0	no
Ant9181	10	1	280	0	0	no	no amplification					10	4	355-371	0.3	0.62	no	10	1	271	0	0	no	no amplification					no amplification							
Ant9218	10	1	335	0	0	no	5	6	500-564	0.6	0.93	no	10	4	311-322	0.7	0.74	no	10	5	336-343	0.4	0.44	no	10	1	383	0	0	no	10	1	360	0	0	no
Ant10878	10	1	302	0	0	no	10	7	358-377	0.6	0.88	no	10	3	292-298	0.6	0.51	no	10	8	280-321	0.8	0.86	no	10	2	283-285	0.1	0.1	no	10	1	320	0	0	no
Ant11315	10	1	358	0	0	no	10	1	380	0	0	no	10	1	317	0	0	no	10	1	355	0	0	no	10	1	322	0	0	no	10	1	343	0	0	no
Ant11400	10	1	258	0	0	no	10	9	259-336	0.7	0.85	no	10	2	234-238	0.3	0.53	no	9	1	242	0	0	no	10	1	294	0	0	no	10	1	251	0	0	no
Ant11893	10	9	375-392	1	0.89	no	10	5	377-412	0.2	0.82	yes	10	4	390-398	0.5	0.73	no	10	1	336	0	0	no	10	1	321	0	0	no	10	4	343-358	0.6	0.66	no
Ant12220	10	1	378	0	0	no	no amplification					10	3	274-280	0.3	0.28	no	10	1	226	0	0	no	10	2	306-360	0.3	0.4	no	10	1	327	0	0	no	

n is the number of individuals successfully genotyped for each locus, A is the number of alleles, H_0 is observed heterozygosity, H_E is expected heterozygosity, and the last column indicates whether each locus deviates from Hardy-Weinberg equilibrium in each species.

I genotyped ten individuals from ten different colonies from the same population of each species for each locus. On average, 12.83 (± 6.15 SD) of the 24 loci were polymorphic in a given species, and 11.16 (± 5.27 SD) were polymorphic and in Hardy-Weinberg equilibrium (Table 2.2, Figure 2.2). If the alleles at a locus deviate from Hardy-Weinberg equilibrium, they are not sorting randomly within the population. If some loci are in Hardy-Weinberg equilibrium within a population and others are not, those that are not may have technical problems that reduce their usefulness in population genetics, such as null alleles or multiple similar-sized PCR products. Across those polymorphic loci in Hardy-Weinberg equilibrium, the average number of alleles per locus per species was 4.59 (± 2.41 SD). The average observed heterozygosity was 0.534 (± 0.22 SD), and the average expected heterozygosity was 0.61 (± 0.22 SD). Most of the loci were monomorphic for multiple species. However, in all cases the monomorphic allele at a given locus was different for each species. I found no statistical linkage disequilibrium (at $p < 0.00003$ after Bonferroni correction, Appendix C) between any pair of loci in any species, but this is likely due to small sample sizes and reduced power due to the large number of tests performed. In fact, in all eight genomes, there are scaffolds containing multiple loci, i.e. these loci occur on the same chromosome and are therefore physically linked (Appendix B).

Discussion

To reduce the time and cost associated with developing microsatellite primers for a large number of different species, I designed a set of 45 primer pairs for potential use in a broad range of ant species spanning many millions of years of evolution. I tested 24 of these primer pairs in detail across six distantly related ant species from six different subfamilies. The number of useful polymorphic loci ranged from 5 to 20 for the six species we tested, although those loci were not

always the same across species. Although I found no statistical linkage between any loci, some loci were located on the same scaffold in the genome assemblies of the reference species, and the location of the loci in the reference genomes should be considered when selecting primers from this set (Appendices A and B). In assessing the utility of these markers in other species, it may be initially beneficial to test the entire set using inexpensive unlabeled primers. Then fluorescently labeled primers can be used for genotyping only those loci that amplify and yield clean PCR products. To further reduce costs, the primers described here could be used as unlabeled locus-specific primers in combination with universal labeled-tail primers (Schuelke, 2000). Microsatellites have been an important tool for studies in population genetics for more than 20 years (Litt & Luty, 1989; Tautz, 1989; Weber & May, 1989). They are excellent markers for many types of studies including pedigree analyses and mating system studies, but their applicability has previously been limited by the narrow range of taxa in which each locus can be used. Researchers usually develop sets of primers specifically for their study species or a group of closely related species, and ants are no exception in this respect (e.g. Ascunce et al., 2009; Azuma et al., 2004; Dalecky et al., 2002; Debout et al., 2006; Debout et al., 2007; Fournier et al., 2005b; Frizzi et al., 2009; Gyllenstrand et al., 2002; Kakazu et al., 2013; Kronauer et al., 2007a; Kronauer et al., 2011a; Qian et al., 2011; Rubin et al., 2009; Steiner et al., 2007; Suefuji et al., 2011). For example, I found 32 publications of microsatellite primer notes for ants in the journal *Molecular Ecology Resources*, a leading outlet for the publication of population genetic markers. These primer notes represented 31 species and 28 genera. Looking only at those studies that described more than ten polymorphic loci per species, the number of alleles per locus ranged from 2 to 21 (Table 2.3). Species-specific primers often had more alleles per locus than we report here. The average number of alleles per locus across all species and loci from Table 2.3 is 7.58(± 4.57 SD) while the average for the loci

described here is 4.59 (± 2.41 SD). One possible explanation for the smaller number of alleles per locus found here is that this reflects a tradeoff between sequence variability within species and sequence conservation across species. On the other hand, this trend is probably at least partly attributable to the small sample size of specimens per species here. The number of alleles per locus will likely increase as more samples are genotyped, especially if these come from different populations. Many microsatellite primers are effective at amplification in congeners, and some microsatellite primers have been successfully used across genera within the same ant subfamily (e.g. Kronauer et al., 2007a; Pol et al., 2008; Steinmeyer et al., 2012). However, to my knowledge, none have successfully amplified polymorphic microsatellite loci across multiple subfamilies. Here I characterize conserved microsatellite markers that are broadly useful across the ants and that will open opportunities for research on the many ant species lacking established genetic markers. These markers, like other microsatellites, will be especially useful for addressing questions in social insect research related to parentage, mating system and colony pedigree structure, i.e. questions for which it is preferable to maximize the number of samples genotyped while fewer markers are generally sufficient. The markers will also be useful in standard population genetic analyses, e.g. of population structure and gene flow. For questions that require a large number of markers such as genomic mapping, NGS data will generally be preferable. However, the loci presented here can readily be used to supplement NGS data. There is demand for broadly applicable microsatellite primers outside the ants as well. Attempts to use microsatellite primers far outside of the species for which they were designed have had varying success. For example, primers designed for use in cattle have proven useful in other closely related mammals (Maudet et al., 2001; Moore et al., 1991; Moore et al., 1998), and microsatellite primers designed for several different legumes have amplified polymorphic loci in the legume genus

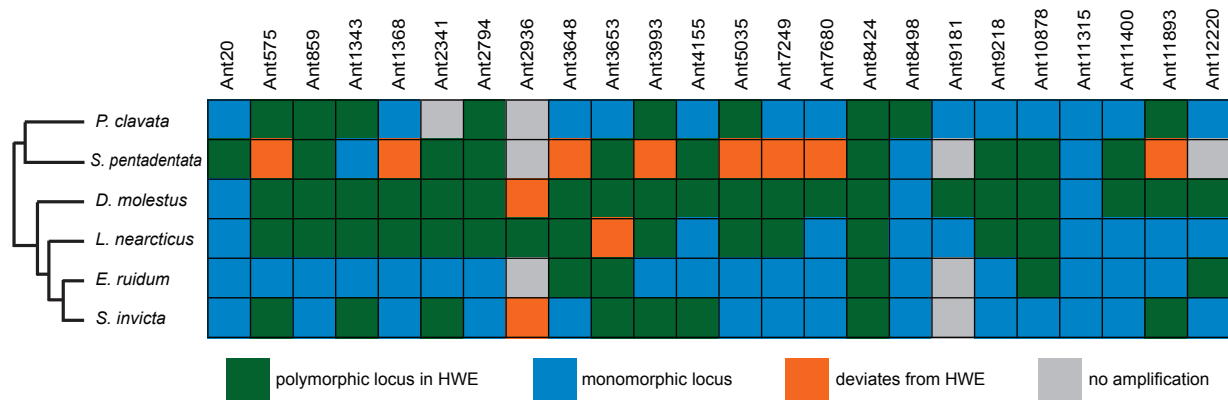


Figure 2.2 Overview results of genotyping 24 microsatellite loci for six different ant species. Green indicates loci that were polymorphic and in Hardy-Weinberg equilibrium, blue indicates monomorphic loci, orange indicates loci that were polymorphic but deviated from Hardy-Weinberg equilibrium, and grey indicates loci that did not amplify. The phylogeny to the left of the figure shows the evolutionary relationships of the species tested.

Table 2.3 Overview of number of alleles and expected and observed heterozygosity in eight studies of species-specific microsatellite primers in ants.

Species	Number of loci	Average number of alleles	Range of allele numbers	Average H_E	H_E range	Average H_O	H_O range	Study
<i>Allomerus octoarticulatus</i>	15	7.03	2-21	0.647	0.185-0.954	0.650	0.200-0.900	Debout et al., 2006
<i>Oecophylla smaragdina</i>	13	5.00	2-14	0.58	0.10-0.89	0.30	0.00-0.60	Azuma et al., 2004
<i>Petalomyrmex phylax</i>	14	7.43	2-15	0.678	0.050-0.925	0.658	0.050-1.000	Dalecky et al., 2002
<i>Formica exsecta</i>	14	8.07	3-18	0.723	0.358-0.965	0.599	0.370-0.826	Gyllenstrand et al., 2002
<i>Wasmannia auropunctata</i>	12	6.48	2-14	0.632	0.305-0.880	0.707	0.233-0.967	Fournier et al., 2005
<i>Azteca ulei</i>	12	10.21	4-18	0.806	0.271-0.965	0.658	0.200-1.000	Debout et al., 2007
<i>Lasius austriacas</i>	11	9.91	4-19	0.782	0.191-0.929	0.700	0.200-0.900	Steiner et al., 2007

Number of loci is the number of polymorphic loci described in that study, mean A is the average number of alleles per locus, A range is the range of allele numbers in each study, mean H_E and mean H_O are the average expected and observed heterozygosity respectively, and H_E range and H_O range are the ranges of expected and observed heterozygosity, respectively.

Glycyrrhiza (Erayman et al., 2014). Some primers designed for the paper wasp genus *Polistes* have also successfully amplified polymorphic loci in other polistine wasps and even in the related subfamilies Vespinae and Stenogastrinae (Ezenwa et al., 1998). In marine turtles, primers have successfully amplified polymorphic microsatellites in species that diverged 300 MYA (FitzSimmons et al., 1995). Additionally, a set of primers similar to those described here has been designed for birds using the genomes of the chicken, *Gallus gallus*, and the zebra finch, *Taenipygia guttata* (Dawson et al., 2010; Dawson et al., 2013). These conserved microsatellite loci also span a long evolutionary distance, as these species have diverged approximately 100 to 120 MYA (Brown et al., 2008; Paton et al., 2002). This work in ants and those in birds (Dawson et al., 2010; Dawson et al., 2013) present sets of primers designed explicitly for use in a broad range of species spanning a long evolutionary distance rather than testing species-specific primers in other distantly related species. Together, they set a precedent for identifying similar sets of markers in other diverse groups of comparable ages. This suggests that, with the availability of genomic information across an ever-increasing range of taxa, conserved microsatellites will become available as powerful population genetic tools for a wide variety of organisms.

Chapter 3: Hybridization between two species of African swarm raiding army ants

I used the microsatellites I previously developed, along with four other previously developed microsatellites to look at a hybridizing population of army ants in Kenya. *Dorylus wilverthi* is primarily a western Congo basin rainforest species, and *D. molestus* is primarily an eastern coastal plain species. The two species encounter each other in Kakamega Forest where hybrid workers are found regularly. I discovered that the hybridization rate was much lower than previously thought, and I found no evidence of genetic caste determination in the population. However, the hybridization rate was too low to conclusively determine whether hybrids were more likely to become workers than queens with my sample size or with any sample size that could reasonably be collected in these two species.

Introduction

In recent years, zoologists have recognized that hybridization between closely related animal species is relatively common, with natural hybrid zones occurring in a wide range of taxa (Coyne & Orr, 2004; Mallet, 2007). Perhaps the most significant consequences of interspecific hybridization are sterility or inviability of hybrid individuals due to deleterious epistatic interactions between the divergent genomes. These unfit hybrids tend to be removed from the population by natural selection (Burke & Arnold, 2001), which may then favor individuals that mate within their own species and reduce the occurrence of interspecific matings in a process known as reinforcement (Ortiz-Barrientos et al., 2004; Servedio & Noor, 2003). Furthermore, premating barriers often prevent interspecific encounters from happening in the first place (Gröning & Hochkirk, 2008). Therefore, even though hybridization as a general phenomenon can

be common (Abbott et al., 2013; Mallet, 2007), hybrid individuals within a population should be relatively rare.

Although interspecific hybridization occurs in a variety of animal groups, its consequences in ants can be very different from those in most other animals due to two aspects of ant biology: haplodiploidy and eusociality. In haplodiploid species, haploid males are produced from unfertilized eggs, a process called arrhenotokous parthenogenesis, while diploid females are produced via normal sexual reproduction. This means that a queen that has mated with a heterospecific male will still produce purebred sons, and hybrid males can only be produced in the F2 generation as the sons of hybrid queens. Because ants are eusocial, the negative effects of hybrid sterility and inviability can be further mitigated when the majority of hybrid individuals become non-reproductive workers rather than queens (Feldhaar et al., 2008; Schwander et al., 2010). These deleterious effects can be completely avoided when hybrids develop exclusively into workers, leading to interesting evolutionary novelties such as genetic caste determination which has been observed in several ant genera (Helms Cahan et al., 2002; Helms Cahan & Keller, 2003; Helms Cahan & Vinson, 2003; Schwander et al., 2010).

Strong genetic effects on caste determination have been observed at least eight times in ants, and in each case they go hand in hand with unusual reproductive systems. Three of these systems (*Wasmannia auropunctata*, *Vollenhovia emeryi*, and *Cataglyphis hispanica*) involve two distinct genetic lineages where sexual reproduction produces interlineage workers and thelytokous parthenogenesis produces reproductive queens, i.e. queens are produced from unfertilized diploid eggs. In *W. auropunctata* (Foucaud et al., 2007; Fournier et al., 2005a) and *V. emeryi* (Kobayashi et al., 2008; Ohkawara et al., 2006), males are produced via androgenesis, i.e. they are clonal copies of the queen's mate. In *C. hispanica*, on the other hand, males are produced via

arrhenotokous parthenogenesis as in most other ant species (Leniaud et al., 2012). *Cataglyphis cursor* is similar to *C. hispanica* in that it displays strong genetic effects on caste determination and males are produced arrhenotokously. However, in this species, populations do not consist of two distinct gene pools, and genetic caste determination does not involve interlineage hybridization (Pearcy et al., 2004).

The other five cases of strong genetic caste determination involve sexual reproduction of both queens and workers. In a hybrid zone between *Solenopsis geminata* and *Solenopsis xylene*, colonies have multiple queens (they are polygynous) and queens are singly mated (they are monandrous; Helms Cahan & Vinson, 2003). In *Pogonomyrmex* harvester ants, colonies have a single queen (they are monogynous) and queens mate with several males (they are polyandrous), mating with males of two genetically distinct lineages (Helms Cahan et al., 2002; Helms Cahan & Keller, 2003). *Messor barbarus* displays a similar system to that seen in *Pogonomyrmex* with monogynous colonies and polyandrous queens, and its congenics *M. structor* and *M. ebeninus* might also display similar mating systems (Romiguier et al., 2017). In all of these cases, reproductive females are produced exclusively from intralineage matings, and workers are produced nearly exclusively from interlineage matings. Both types of matings are required for a newly founded colony to be successful and to complete a full reproductive cycle. Thus, genetic caste determination can occur in species in which new queens are produced parthenogenetically, or in species that are either polyandrous or polygynous (Schwander et al., 2010). All currently known instances of strong genetic caste determination in ants involve matings between males and females that belong to different gene pools and can thus be considered cases of hybridization.

The only case of hybridization in ants yet observed outside of the two largest ant subfamilies (*Formicinae* and *Myrmicinae*) is in the genus *Dorylus* (subfamily *Dorylinae*), where

introgression has occurred between two species of driver ants at Kakamega Forest in Kenya (Kronauer et al., 2011b). Colonies of driver ants are monogynous and queens are highly multiply mated. The observation of hybridization between two species in this group thus raises the possibility of genetic caste determination similar to that found in *Pogonomyrmex*. Driver ants are a group of nomadic swarm-raiding army ants that are restricted to sub-Saharan Africa (Gotwald, 1995; Wilson, 1964). The group is comprised of approximately nine species (Kronauer et al., 2007b; Schöning, 2008) in the subgenus *Anomma*. Driver ants are top invertebrate predators with extremely large colonies that can contain more than 10 million workers (Gotwald, 1995). Unlike the queens of most other ants, army ant queens are permanently wingless. Instead of going on a mating flight, new queens mate within their natal nest with approximately 10-30 unrelated males that disperse on the wing. In driver ants, mating probably occurs within a two to three-week period before the new queen assumes regnancy of her colony (Kronauer et al., 2004b; Kronauer & Boomsma, 2007a), and the mother queen disperses on foot, taking a portion of the existing workers with her (Gotwald, 1995).

Two species of driver ants are found in Kakamega Forest (Garcia et al., 2009; Kronauer et al., 2011b; Peters et al., 2011; Peters & Okalo, 2009) where they occupy distinct ecological niches. *Dorylus wilverthi* (Emery) mainly inhabits intact rainforest habitat while *D. molestus* (Wheeler) occurs in open agricultural habitat and grasslands (Peters et al., 2009; Peters & Okalo, 2009; Schöning et al., 2006). Previous work showed evidence of historical mitochondrial introgression as well as hybridization in the nuclear genome within the worker caste. Both species at Kakamega also display intermediate morphologies when compared to allopatric populations (Kronauer et al., 2011). However, the study by Kronauer et al. (2011b) left many important questions open, mostly because of relatively small sample sizes and the lack of colony pedigree information. For example,

it is currently unclear whether hybridization at Kakamega is purely historical or ongoing, what the relative frequency of interspecific matings is, and whether interspecific matings are unidirectional or bidirectional. Furthermore, it is not known whether genetic caste determination plays a role in the Kakamega population. Here I present the first population genetic study based on large sample sizes and detailed colony pedigree information of this population, which allows me to provide a detailed characterization of the mating systems of the two driver ant species at Kakamega Forest, including reliable estimates of the occurrence of hybrids and an evaluation of the potential for genetic caste determination.

Materials and methods

Sample collection

I collected 27 colonies of *D. wilverthi* and 23 colonies of *D. molestus* in Kakamega Forest, Kenya in August and September 2012. Workers were collected from emigration trails, foraging trails, swarm raids, or nest sites encountered while hiking through the forest. All samples were identified to species morphologically. In allopatric populations, the two species are distinguishable primarily by the horn-like protrusions on the posterior corners of the head of *D. wilverthi* (Emery, 1899), as well as tubercles on the lateral edges of the petiole in *D. molestus* (Gotwald & Schaefer, 1982). At Kakamega Forest, many workers of *D. molestus* lack petiolar tubercles, but workers of *D. wilverthi* maintain their horn like protrusions although they are much reduced (Kronauer et al., 2011b) (Figure 3.1).

DNA extraction, amplification, sequencing, and microsatellite genotyping

DNA for sequencing the mitochondrial gene *cytochrome oxidase II* (*COII*) was extracted from one leg of one worker from each colony using QIAGEN DNeasy kits. The *COII* gene was amplified using the primers tRNA^{Leu} (Kronauer et al., 2007b) and Barbara (Simon et al., 1994) using a previously described protocol (Kronauer et al., 2007b). PCR products were sent to a commercial facility for purification and sequencing (Eton Bioscience, Charlestown MA). PCR products were sequenced in both directions for a final sequence length of 609 base pairs.

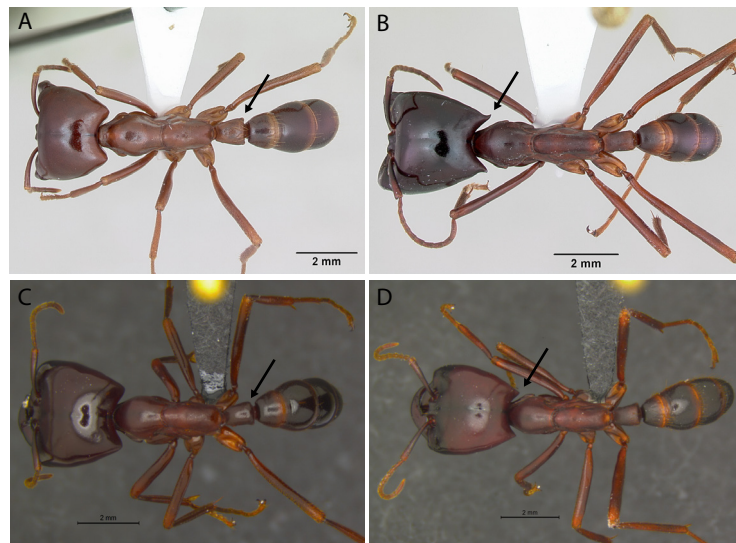


Figure 3.1 Photographs of *Dorylus molestus* and *Dorylus wilverthi* from allopatric populations and from Kakamega Forest. (a) *Dorylus molestus* from Chogoria, Kenya (allopatric with *D. wilverthi*). (b) *Dorylus wilverthi* from Salonga National Park, DR Congo (allopatric with *D. molestus*). (c) *Dorylus molestus* from Kakamega Forest. (d) *Dorylus wilverthi* from Kakamega forest. Arrows indicate lack of petiolate tubercles in *D. molestus* and reduced hornlike protrusions in *D. wilverthi* from Kakamega Forest. (a) and (b) courtesy of April Nobile and AntWeb at <http://www.antweb.org>.

DNA for microsatellite genotyping was extracted by boiling one leg of each worker in 100 μ L of 10% Chelex 100 (BioRad) for 15 minutes, centrifuging for 3 minutes at 9000 rpm and pipetting off the supernatant containing the DNA. We genotyped 1,128 workers at 12 different microsatellite loci. Three of these loci (DmoB, DmoD, and DmoG) were previously developed for *Dorylus molestus* (Kronauer et al., 2004a). The other 9 loci (Ant4155, Ant2341, Ant5035, Ant7248, Ant7680, Ant8424, Ant1343, Ant9218, and Ant10878) are from the set of universal ant microsatellite primers designed for use in diverse ant taxa described in the previous chapter.

The PCR cocktail (10 μ L total volume) for all reactions contained 1 μ L PCR Gold Buffer (10x), 0.5 μ L $MgCl_2$ (25mM), 0.5 μ L dNTPs (10mM total, 2.5mM each nucleotide), 0.1 μ L AmpliTaq Gold (5 U/ μ L) (Applied Biosystems), and 1 μ L DNA template. Primers were all used at 10 μ M concentration and were multiplexed in the following sets with the following volumes: Multiplex 1: Ant4155 (0.2 μ L), Ant2341 (0.2 μ L), Ant5035 (0.3 μ L). Multiplex 2: Ant7680 (0.1 μ L), Ant7249 (0.3 μ L). Multiplex 3: Ant9218 (0.1 μ L), Ant10878 (0.1 μ L). Multiplex 4: DmoB (0.1 μ L), DmoG (0.2 μ L). Ant1343, Ant8424, and DmoD were each amplified separately using 0.1 μ L of each primer in the PCR mix. Volumes listed are for each of the forward and reverse primers. Water was added to bring the total volume of each reaction to 10 μ L. PCR reactions were run on an Eppendorf Mastercycler Pro S under the following conditions: 10 min at 95°C followed by 40 cycles of 15 s at 94°C, 30 s at 55°C, and 30 s at 72°C, and a final extension of 10 min at 72°C. PCR products were sent to a commercial facility for genotyping (Genewiz, South Plainfield, NJ). The resulting chromatograms were scored using Peak Scanner software (Applied Biosystems).

Reconstruction of queen genotypes

I genotyped twenty workers from each colony at twelve microsatellite loci. Ten of the workers from each colony were majors and ten were minors. Majors and minors were determined using head width as a proxy for size. Major workers had a head width at the widest point of at least 3.20mm and minors had a maximum head width of 0.90mm. For two colonies (IB83 and IB124) we genotyped an additional fifty workers each, and for colony IB117 I genotyped an additional 20 workers. For colony IB67 I genotyped an additional six workers at all 12 loci and 2 workers at 11 loci to increase the confidence in the reconstructed queen genotypes (see below). These additional workers were also included in the STRUCTURE analyses (see below). This resulted in a total of 1,128 workers across all colonies. These additional workers were of variable sizes and did not necessarily fall within either of the two size classes mentioned above.

To reconstruct colony queen genotypes, I used the software MateSoft (Moilanen et al., 2004), which is designed for the analysis of mating systems in haplodiploid organisms. It gives all possible queen genotypes for the given set of offspring of a monogynous colony and assigns a probability to each one based on the allele frequencies in the population, and how closely alleles in the offspring correspond to Mendelian ratios. All workers except those from colony IB117 (see below) were used to calculate allele frequencies using the program FSTAT (Goudet, 1995). MateSoft returned queen genotypes for 36 of 49 colonies. The average probability of the queens predicted was 0.70 (+/- 0.24 SD). To increase this probability, we genotyped 547 additional workers from 25 different colonies for only those loci that were contributing to the uncertainty of the queen genotypes. Except for the previously mentioned eight workers from colony IB67, these additional workers were not used in the STRUCTURE (Pritchard et al., 2000) analysis or in mating frequency estimates, and are not part of the 1,128 total workers with complete or nearly complete

genotypes. However, they were used to calculate allele frequencies. These additional workers and genotyped loci are listed in Appendix D. This increased the average probability of queen genotypes to 0.91 (+/- 0.10 SD).

Workers from the 13 remaining colonies could not be explained as being derived from a single queen, despite the fact that *Dorylus* colonies are known to be monogynous, i.e. contain only a single reproductively active female. However, there can be offspring of both mother and daughter queens present if colonies are sampled after a recent fission event (see Introduction in this chapter). Furthermore, there can be drifters present that originated in a foreign colony (Kronauer et al., 2010). For the remaining 13 colonies, I input the worker genotypes into COLONY (Jones & Wang, 2010), a software package designed for pedigree analysis that accounts for multiple queens and drifters, as well as genotyping errors (Wang, 2004). COLONY constructs a pedigree assigning each worker to an inferred queen genotype and male genotype. The “Sibship Prior” setting was set to very strong, paternal sibship size was set to 1.2, and maternal sibship size was set to 20 so that all workers would be assigned to as few queens as possible. Queen genotypes that could account for 17 or more workers were used for further analyses, and the rest were discarded. In this way, I obtained queen genotypes for an additional 8 colonies. These 8 queen genotypes and the 36 queen genotypes obtained from MateSoft total the 44 colonies and associated queen genotypes used for further analyses. In attempting to recover queen genotypes for all putative hybrid workers, I also genotyped an additional 50 workers per colony for two colonies (IB83 and IB124) that contained hybrids but could not be explained by a single queen in the initial MateSoft analysis. I also tested for linkage disequilibrium between all pairs of loci in the queen dataset using FSTAT (Goudet, 1995).

Male genotypes

Male genotypes were reconstructed using MateSoft, by inputting the 44 reconstructed queen genotypes and the workers associated with those queens. MateSoft collapses redundant patriline and gives the smallest number of patriline that can explain all of the worker genotypes. It then gives all possible genotypes for each male and assigns a probability based on the allele frequencies in the population. I took the highest probability male genotype per patriline and used these genotypes for population structure analyses and mating frequency statistics.

Principle components analysis

A principle components analysis was performed by assigning each allele from each microsatellite locus to a variable. I included all 1,128 workers with complete or nearly complete genotypes. Each variable was then scored for each individual as 1 (homozygous), 0.5 (heterozygous), or 0 (allele not present). Missing values were imputed using the R package missMDA. Principle components analysis (PCA) was then performed using the R package FactoMineR.

Population structure and determination of hybrid genotypes

I used the statistical clustering program STRUCTURE to cluster workers based on their genotypes. Markov chains were run for 20^5 generations, with the first 10^5 generations being discarded as burn-in. The presence of family groups may bias the clustering to favor colonies that are represented by more genotypes in the dataset. Indeed, at higher numbers of clusters K (see below for how K was determined), STRUCTURE preferentially classified colonies with larger sample sizes as discrete clusters. To avoid this, colonies with extra workers (IB67, IB83 and

IB124), were input in groups of 20 in the following seven sets: 1) workers 1 to 20 of each of these three colonies, 2) workers 21 to 40 of IB83 and IB124, and workers 1 to 12 and 29 to 36 of IB67, 3) workers 41 to 60 of IB83 and IB124, and workers 1 to 4, 13 to 20 and 29 to 36 of IB67, 4) workers 1 to 10 and 61 to 70 of IB83 and IB124, and workers 5 to 20 and 29 to 32 of IB67, 5) workers 11 to 30 of IB83 and IB124, and workers 1 to 15 and 32 to 36 of IB67, 6) workers 31 to 50 of IB83 and IB124, and workers 1 to 7, 16 to 20 and 29 to 36 of IB67, 7) workers 51 to 70 of IB83 and IB124, and workers 8 to 20 and 29 to 35 of IB67. This resulted in each worker being run in at least two different sets, and all runs were repeated five times. Workers from colony IB117 were not included in this analysis (see below).

I used the proportion of ancestry predicted by STRUCTURE of each individual to assess whether a worker was of hybrid ancestry. If an individual had a proportion of ancestry in the species of their natal colony less than our cutoff I considered it a putative hybrid. I set three cutoffs for the maximum proportion of ancestry. I used cutoffs at 0.9 and 0.75 in order to make a direct comparison to the hybridization rate estimated by Kronauer et al. (2011b). I set a third cutoff at 0.65. I used this value because this was the highest proportion of ancestry where all hybrids had reconstructed queen and male genotypes from different populations (see below). This is therefore a conservative cutoff at which it seems unlikely that false positives are included in the identification of putative hybrid workers. At the same time, at this cutoff we might be missing actual F1 hybrids, so inferences based on this cutoff constitute minimum estimates of the proportion of interspecific matings.

To determine the correct number of K groups, I varied the number of assumed populations from 1 to 15 ($K=1-15$) and performed 5 replicates for each value of K for the worker, queen and male datasets. To control for potential confounding effects of the pedigree structure in our worker

dataset, I also conducted an analysis on a subsampled dataset that contained only a single worker per colony. Replicate runs were highly consistent for all values of K. The optimal values of K for workers, queens, and males were estimated using the ΔK method (Evanno et al., 2005).

Mating Frequencies

Mating frequencies were calculated in MateSoft using the previously gathered pedigree information for all workers, queens and males. Observed mating frequencies simply give the number of males that have contributed to the sampled female offspring of a given queen. However, the actual number of males a queen has mated with may be much higher due to limited sampling, and especially in cases where a subset of males do not contribute to the offspring at all due to ineffective mating or inviable offspring. Effective mating frequencies, on the other hand, also take into account the relative contributions of each male to the offspring. Effective mating frequencies thus decrease with increasing reproductive skew among the different mates of a given queen. Estimates of effective mating frequencies were corrected for limited sample sizes following Nielsen et al. (2003).

Results

Population structure

Mitochondrial COII sequences

The results I obtained were consistent with the pattern found by Kronauer et al. (2011b). I found four haplotypes that were previously reported (Kronauer et al., 2011b) in addition to eight new haplotypes (Figure 3.2, Table 3.1). The haplotypes that were shared between both datasets were also the most frequent in both, and the new haplotypes from this dataset were relatively rare,

found on average 1.25 times. Similar to what has been previously reported (Kronauer et al., 2011b), in a haplotype network constructed from the *COII* sequences of both datasets, samples from Kakamega Forest form a cluster, i.e. Kakamega samples of both species are more closely related to each other than to samples from any other locality, including other samples of the same species. However, there are two exceptions to this finding. First, in this dataset I found haplotype 8 three times in Kakamega Forest, while the previous dataset found it once at the same locality and once

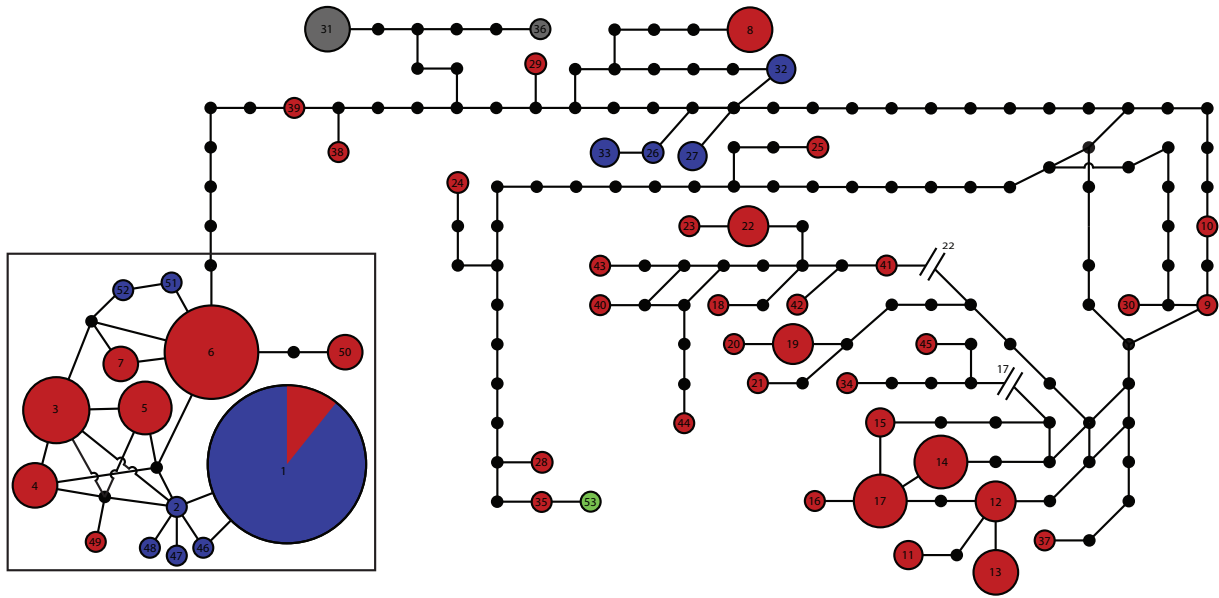


Figure 3.2 Haplotype network of driver ant samples from East Africa compiled from this data set and from Kronauer et al. (2011b). Each labelled circle indicates a single haplotype, and the size of the circle is proportional to the frequency of the haplotype. Each line between circles represents a single nucleotide change, and black dots represent missing haplotypes. Red indicates *Dorylus molestus*, blue indicates *Dorylus wilverthi*, and grey indicates *Dorylus terrificus*. The green circle represents colony IB117, which does not cluster with any of the other samples from Kakamega Forest. The box encloses samples from Kakamega Forest. GenBank accession numbers are listed in Table 1 for samples from Kakamega Forest. Haplotype numbers are the same as in Kronauer et al. (2011b).

at Mt. Elgon. This haplotype is more similar to haplotypes from Mt. Lole in Kenya, and Kibale and Budongo Forest in Uganda, than to other haplotypes from Kakamega Forest. Second, one new haplotype from this dataset (haplotype 53) is more similar to *D. molestus* haplotypes from Nakuru, Kenya and Semliki, Uganda than to other haplotypes from Kakamega. Both of these exceptions could represent recent migration events into the Kakamega Forest population.

Kronauer et al. (2011b) found that the most frequently occurring haplotype at Kakamega forest was shared by both *D. wilverthi* and *D. molestus*. The same haplotype was also the most frequently occurring in this dataset (haplotype 1), but all occurrences were from *D. wilverthi* and none were from *D. molestus*. No other haplotypes were shared by the two species in either dataset. Importantly, all haplotypes of *D. wilverthi* at Kakamega Forest are nested within *D. molestus* haplotypes suggesting that interspecific hybridization, either currently or in the recent past, has resulted in genetic introgression between the two species (see also Kronauer et al., 2011b).

Nuclear microsatellite markers

At none of the microsatellite markers did either species deviate from expected heterozygosity under Hardy-Weinberg equilibrium (Table 3.2), and I found no evidence of linkage disequilibrium between any pair of loci within either of the two species. A principle components analysis of these data found two distinct clusters corresponding to the two species *D. molestus* and *D. wilverthi* (Figure 3.3). A third smaller cluster corresponded to colony IB117, which is the same colony represented by haplotype 53, further supporting the idea that this colony is a recent immigrant to the population. This finding, along with the mitochondrial data, lead me to believe that IB117 is genetically distinct, and may not yield reliable results when examining genetic interactions between the two species. Therefore, I excluded it from further analyses.

Table 3.1 List of unique *cytochrome oxidase II (COII)* mitochondrial haplotypes from Kakamega Forest. Haplotype numbers are the same as those listed in Kronauer et al. (2011b).

Haplotype number	Genbank accession number	Species	Locality	Number of colonies in this dataset	Number of colonies in Kronauer <i>et al.</i> (2011)	Total number of colonies
1	GQ999016, KY701978	<i>D. wilverthi</i>	Kakamega Forest, Kenya	22	33	54
1	GQ999023	<i>D. molestus</i>	Kakamega Forest, Kenya	0	8	9
2	GQ999019	<i>D. wilverthi</i>	Kakamega Forest, Kenya	0	1	1
3	GQ999017, KY701977	<i>D. molestus</i>	Kakamega Forest, Kenya	5	6	11
4	GQ999020	<i>D. molestus</i>	Kakamega Forest, Kenya	0	5	5
5	GQ999022	<i>D. molestus</i>	Kakamega Forest, Kenya	0	7	7
6	GQ999018, KY701980	<i>D. molestus</i>	Kakamega Forest, Kenya	10	12	22
7	GQ999021	<i>D. molestus</i>	Kakamega Forest, Kenya	0	3	3
8	GQ999024, KY701981	<i>D. molestus</i>	Kakamega Forest, Kenya	3	1	4
8	GQ999043	<i>D. molestus</i>	Mt. Elgon, Kenya	0	1	1
46	KY701973	<i>D. wilverthi</i>	Kakamega Forest, Kenya	1	0	1
47	KY701974	<i>D. wilverthi</i>	Kakamega Forest, Kenya	1	0	1
48	KY701975	<i>D. wilverthi</i>	Kakamega Forest, Kenya	1	0	1
49	KY701976	<i>D. molestus</i>	Kakamega Forest, Kenya	1	0	1
50	KY701979	<i>D. molestus</i>	Kakamega Forest, Kenya	3	0	3
51	KY701983	<i>D. wilverthi</i>	Kakamega Forest, Kenya	1	0	1
52	KY701982	<i>D. molestus</i>	Kakamega Forest, Kenya	1	0	1
53	KY701972	Unknown	Kakamega Forest, Kenya	1	0	1
9	GQ999042	<i>D. molestus</i>	Mt. Kenya (West), Kenya	0	1	1
10	EF413797	<i>D. molestus</i>	Mt. Kenya (East), Kenya	0	1	1
11	GU065701	<i>D. molestus</i>	Mt. Kenya (East), Kenya	0	2	2
12	GU065703	<i>D. molestus</i>	Mt. Kenya (East), Kenya	0	4	4
13	GU065698	<i>D. molestus</i>	Mt. Kenya (East), Kenya	0	5	5
14	GU065699	<i>D. molestus</i>	Mt. Kenya (East), Kenya	0	7	7
15	GU065702	<i>D. molestus</i>	Mt. Kenya (East), Kenya	0	2	2
16	GU065704	<i>D. molestus</i>	Mt. Kenya (East), Kenya	0	1	1
17	GU065700	<i>D. molestus</i>	Mt. Kenya (East), Kenya	0	7	7

Table 3.1 Continued

18	GQ999037	<i>D. molestus</i>	Mt. Kenya (South), Kenya	0	1	1
19	GQ999050	<i>D. molestus</i>	Tana River (East), Kenya	0	4	4
20	GQ999052	<i>D. molestus</i>	Tana River (East), Kenya	0	1	1
21	GQ999049	<i>D. molestus</i>	Tana River (West), Kenya	0	1	1
22	GQ999032	<i>D. molestus</i>	Tana River (West), Kenya	0	4	4
23	GQ999051	<i>D. molestus</i>	Tana River (West), Kenya	0	1	1
24	GQ999027	<i>D. molestus</i>	Addis Ababa, Ethiopia	0	1	1
25	GQ999028	<i>D. molestus</i>	Bale Mts., Ethiopia	0	1	1
26	GQ999040	<i>D. wilverthi</i>	Budongo Forest, Uganda	0	1	1
27	GQ999035	<i>D. wilverthi</i>	Budongo Forest, Uganda	0	2	2
28	GQ999031	<i>D. molestus</i>	Semliki, Uganda	0	1	1
29	GQ999053	<i>D. molestus</i>	Mt. Lole, Kenya	0	1	1
30	GQ999045	<i>D. molestus</i>	Mt. Warges, Kenya	0	1	1
31	EF413804	<i>D. molestus</i>	Kibale Forest, Uganda	0	5	5
32	EF413798	<i>D. wilverthi</i>	Kibale Forest, Uganda	0	2	2
33	GQ999034	<i>D. wilverthi</i>	Mabira Forest, Uganda	0	2	2
34	GQ999030	<i>D. molestus</i>	Kalindu, Uganda	0	1	1
35	GQ999033	<i>D. molestus</i>	Nakuru, Kenya	0	1	1
36	GQ999039	<i>D. molestus</i>	Mgahinga, Uganda	0	1	1
37	GQ999038	<i>D. molestus</i>	Ngong Hills, Kenya	0	1	1
38	GQ999046	<i>D. molestus</i>	Kilimanjaro, Tanzania	0	1	1
39	GQ999047	<i>D. molestus</i>	Kilimanjaro, Tanzania	0	1	1
40	GQ999048	<i>D. molestus</i>	North Pare, Tanzania	0	1	1
41	GQ999029	<i>D. molestus</i>	Taita Hills, Kenya	0	1	1
42	GQ999044	<i>D. molestus</i>	Taita Hills, Kenya	0	1	1
43	GQ999041	<i>D. molestus</i>	Arabuko Soko, Kenya	0	1	1
44	GQ999025	<i>D. molestus</i>	Shimba Hills, Kenya	0	1	1
45	GQ999026	<i>D. molestus</i>	Gombe, Tanzania	0	1	1

Table 3.2 Number of alleles (N_a), size range of alleles in base pairs, observed heterozygosity (H_o) and expected heterozygosity (H_e) estimates from 12 microsatellite loci measured for the two species *Dorylus molestus* and *Dorylus wilverthi* at Kakamega Forest.

<i>D. molestus</i>					<i>D. wilverthi</i>				
locus	N_a	allele size range (bp)	H_o	H_e	N_a	allele size range (bp)	H_o	H_e	total alleles
Ant1343	6	261-271	0.732	0.724	8	261-274	0.789	0.733	8
Ant2341	5	261-273	0.352	0.394	4	261-270	0.443	0.447	5
Ant4155	5	185-194	0.607	0.585	4	185-194	0.408	0.442	5
Ant5035	11	361-385	0.706	0.685	8	365-385	0.753	0.779	12
Ant7249	8	359-398	0.525	0.521	7	361-398	0.669	0.657	8
Ant7680	11	269-328	0.671	0.716	8	310-328	0.755	0.744	11
Ant8424	8	223-250	0.732	0.693	7	223-250	0.733	0.755	8
Ant9218	7	305-326	0.727	0.725	6	305-326	0.721	0.74	7
Ant10878	7	287-297	0.535	0.506	5	287-297	0.64	0.601	7
DmoB	6	220-230	0.642	0.663	8	220-235	0.593	0.631	8
DmoD	9	148-166	0.757	0.773	8	148-164	0.699	0.706	10
DmoG	10	201-221	0.875	0.836	10	203-223	0.802	0.789	12

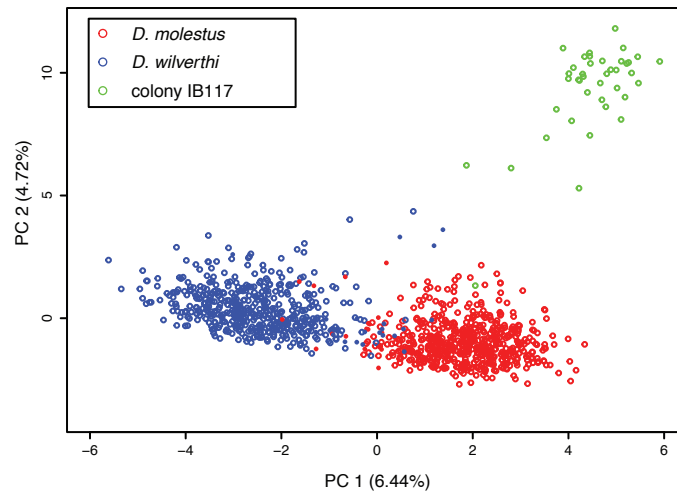


Figure 3.3 Principal components analysis of nuclear microsatellite genotypes. Red indicates *Dorylus molestus*, blue indicates *Dorylus wilverthi*, and green indicates colony IB117, which is genetically distinct from both species in Kakamega Forest at nuclear markers as well as mitochondrial markers. Filled circles represent putative hybrid workers at the 0.90 cut-off and are colored corresponding with the species of the colony they were found in. The two species clearly form distinct clusters, and hybrid individuals occur intermediately between the two clusters.

For colonies IB83 and IB124, I could not resolve the queen genotypes even after genotyping 50 additional workers for these two colonies. For colony IB83, COLONY predicted one queen genotype as the mother of 66 workers and a second queen genotype as the mother of the remaining 4 workers. However, the two genotypes were not consistent with a mother/daughter pair of queens and the four workers predicted as daughters of the second queen were all putative hybrids. This suggests that the presence of these hybrids in the colony is biasing COLONY to predict unrelated queen genotypes. For colony IB124, COLONY predicted three queens, one as the mother of 35 workers, one as the mother of 28 workers, and one as the mother of 7 workers. No combination of these three queens was consistent with a mother/daughter pair. For both colonies, we were also unable to recover one queen or a mother/daughter pair by deducing the queen genotypes manually following Kronauer et al. (2004b) even while accounting for genotyping errors. These colonies were also excluded from further analyses.

I then ran the microsatellite data for the remaining 1,088 workers (after excluding colony IB117) through the statistical clustering program STRUCTURE (Pritchard et al., 2000). This method has the advantage of testing for distinct genetic clusters without *a priori* assumptions about the origin population of each individual. The ΔK method clearly showed a change in the rate of increase of log-likelihood after $K=2$ for workers and males indicating that $K=2$ is in fact the correct number of groups for these datasets. For workers, this was true both for the entire worker dataset and the subsampled dataset with a single worker per colony (Appendix J). For queens, this method showed two modes in the value of ΔK at $K=3$ and $K=7$ (Appendix J). The uncertainty here may be due to a smaller dataset for the queens ($n=44$), where the ΔK method is less reliable in predicting the correct value of K (Evanno et al., 2005). This is further evidenced by the magnitude of the change in ΔK at the two modes. The peaks at $K=3$ and $K=7$ are only 2 to 3 times higher than the

neighboring values, whereas the peaks at K=2 in the worker and male datasets are two orders of magnitude higher than the neighboring values (Appendix J). Furthermore, rather than splitting queen genotypes into three distinct clusters, the STRUCTURE analyses at K=3 still consistently recovered two clusters that corresponded to the species assignments based on morphology, while adding contributions from the third presumed group to all individuals (Appendix I). This further suggests that K=2 is indeed the correct value for the queen dataset as well. Finally, previous taxonomic, ecological, and population genetic work has convincingly shown that only two driver ant species occur at the study site (Garcia et al., 2009; Peters & Okalo, 2009; Kronauer et al., 2011b; Peters et al., 2011), providing a strong expectation for K=2 being the correct number of groups. Based on our results, in combination with the previous work conducted on the Kakamega population, I am confident in the assessment that there are two populations present at Kakamega Forest corresponding to the two species. This is also consistent with the results of the principle components analysis showing two distinct clusters. Setting K=2 for all further STRUCTURE analyses therefore seemed justified. At K=2 STRUCTURE clustered all colonies as predicted based on morphological identification (Figure 3.4). However, some individuals did not cluster with the rest of their colony, i.e. were classified as having a low probability of belonging to the same species as their nestmates, even though they were offspring of the same colony queen (see below). These individuals are putative hybrids. We defined putative hybrids as having a proportion of ancestry in the same cluster (species) as the majority of the colony less than one of our three cutoffs. At the 0.9 cutoff, I found 57 putative hybrid workers, 31 of which were from 9 colonies of *D. molestus* and 26 from 10 colonies of *D. wilverthi* (Table 3.3, Appendix E). At the 0.75 cutoff, I found 17 putative hybrids from 5 colonies of *D. molestus* and 13 from 5 colonies of *D. wilverthi*.

At the 0.65 cutoff, I found 14 putative hybrids from 4 colonies of *D. molestus* and 10 from 4 colonies of *D. wilverthi*.

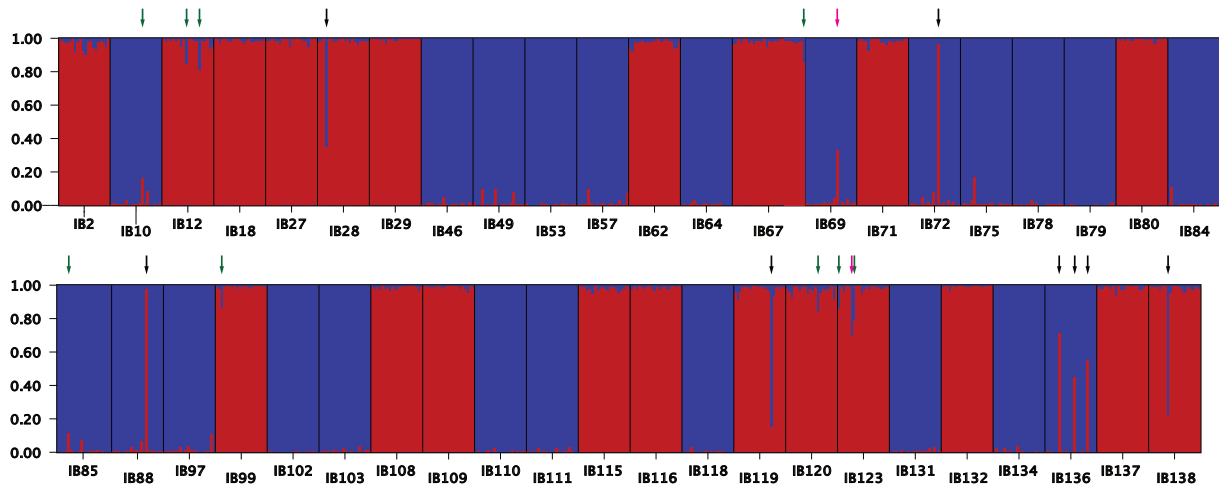


Figure 3.4 Assignment of workers to two populations using the clustering program STRUCTURE. Each vertical bar represents a single individual, and the y-axis represents the proportion of each multilocus genotype that is attributable to each of the two populations. Workers are grouped according to their colony, and the colony number is given underneath each group. Red corresponds to colonies of *D. molestus*, and blue corresponds to colonies of *D. wilverthi*. Arrows above the plot indicate putative hybrids at three different proportion of ancestry cut-offs. Black arrows indicate putative hybrids at the 0.65 cut-off, pink arrows indicate the two additional hybrids included at the 0.75 cut-off, and green arrows indicate the further additional hybrids included at the 0.90 cut-off. Individual hybrids are listed in Table 3.3. This is one representative run of a total of 35 replicate runs. Colonies without reconstructed queens were included in this sample but were excluded from the figure. STRUCTURE results for workers without reconstructed queen genotypes are given in Appendix H.

Table 3.3 List of putative hybrid workers at three different proportion of ancestry cut-offs and species determinations of their colony, mother queen and father.

Hybrids at 0.65 cutoff	Worker subcaste	Colony species	Queen species	Male	Male species
IB28.4	major	<i>D. molestus</i>	<i>D. molestus</i>	IB28M5	<i>D. wilverthi</i>
IB72.12	minor	<i>D. wilverthi</i>	<i>D. wilverthi</i>	IB72M19	<i>D. molestus</i>
IB88.14	minor	<i>D. wilverthi</i>	<i>D. wilverthi</i>	IB88M17	<i>D. molestus</i>
IB119.15	minor	<i>D. molestus</i>	<i>D. molestus</i>	IB119M12	<i>D. wilverthi</i>
IB136.6	major	<i>D. wilverthi</i>	<i>D. wilverthi</i>	IB136M2	<i>D. molestus</i>
IB136.12	minor	<i>D. wilverthi</i>	<i>D. wilverthi</i>	IB136M2	<i>D. molestus</i>
IB136.17	minor	<i>D. wilverthi</i>	<i>D. wilverthi</i>	IB136M2	<i>D. molestus</i>
IB138.8	major	<i>D. molestus</i>	<i>D. molestus</i>	IB138M8	<i>D. wilverthi</i>
Hybrids at 0.75 cutoff					
IB69i13	minor	<i>D. wilverthi</i>	<i>D. wilverthi</i>	IB69M9	<i>D. wilverthi</i>
IB123i6	major	<i>D. molestus</i>	<i>D. molestus</i>	IB123M7	<i>D. molestus</i>
Hybrids at 0.90 cutoff					
IB10i13	minor	<i>D. wilverthi</i>	<i>D. wilverthi</i>	IB10M11	<i>not assignable</i>
IB12i10	major	<i>D. molestus</i>	<i>D. molestus</i>	IB12M10	<i>D. wilverthi</i>
IB12i15	minor	<i>D. molestus</i>	<i>D. molestus</i>	IB12M12	<i>D. molestus</i>
IB67i28	neither	<i>D. molestus</i>	<i>D. molestus</i>	IB67M24	<i>D. molestus</i>
IB85i5	major	<i>D. wilverthi</i>	<i>D. wilverthi</i>	IB85M5	<i>D. molestus</i>
IB99i3	major	<i>D. molestus</i>	<i>D. molestus</i>	IB99M4	<i>D. molestus</i>
IB120i13	minor	<i>D. molestus</i>	<i>D. molestus</i>	IB120M12	<i>D. wilverthi</i>
IB123i1	major	<i>D. molestus</i>	<i>D. molestus</i>	IB123M5	<i>D. molestus</i>
IB123i7	major	<i>D. molestus</i>	<i>D. molestus</i>	IB123M3	<i>D. molestus</i>

All putative hybrids at the 0.65 cut-off are implicitly included in the 0.75 cut-off, and similarly, all hybrids at the 0.75 cut-off are implicitly included in the 0.90 cut-off. This table only includes hybrids used to calculate the hybridization rate (from colonies that have a reconstructed queen genotype). Putative hybrids from other colonies are listed in Appendix D. The species assignment of each colony is based on morphology. The species assignments of queens and males are based on the results of multiple STRUCTURE runs.

Queen genotypes

To strengthen the assessment that intermediate worker genotypes are true hybrids I reconstructed the parental genotypes of as many workers as possible to identify the population of origin of the queen and each of her mates. If putative hybrids are false positives, then the parental genotypes should not necessarily come from different populations. If the mothers and fathers of putative hybrid workers come from different species in a high proportion of cases, I can be more certain that I have identified true hybrids in the population. I was able to reconstruct the queen genotypes for 44 of 49 colonies (colony IB117 already being excluded). Running the queen genotypes through STRUCTURE gave results consistent with the worker species assignments. At $K=2$, the two clusters corresponded to the same two clusters as the worker genotypes, and all queen cluster assignments were consistent with the species assignment of their colony (Figure 3.5). Included in these queen genotypes were the mothers of 19, 10 or 8 putative hybrids according to the 0.9, 0.75, or 0.65 cutoffs, respectively. Finding maternal genotypes indicates that these workers were collected from their natal nest and are not heterospecific drifters in a foreign colony. Colonies we could not reconstruct queen genotypes for (IB40, IB63, IB83, IB87, IB124) were excluded when estimating hybridization rates because we cannot rule out the possibility of putative hybrids from these colonies being heterospecific drifters instead.



Figure 3.5 Assignment of reconstructed queen genotypes to two populations using the clustering program STRUCTURE. Queen genotypes are grouped by species, and all queen genotypes clustered with the same species as their colony. This is one representative run of five replicates.

Male genotypes and confirmation of putative hybrid workers

Using the predicted queen genotypes and the observed worker genotypes, I reconstructed a total of 720 male genotypes. STRUCTURE results ($K=2$) clustered most males with the same species as their mates (Figure 3.6). Out of 348 mates reconstructed for *D. molestus* queens, 332 were classified as *D. molestus* and 7 were classified as *D. wilverthi* based on our 0.65 cutoff. Nine males could not be assigned to either species, i.e. their estimated proportion of ancestry was less than 0.65 in either species. Out of 372 mates reconstructed for *D. wilverthi* queens, 358 were classified as *D. wilverthi*, 8 were classified as *D. molestus*, and 6 could not be assigned. I identified the fathers of all hybrid workers for which I had queen genotypes. At the 0.65 proportion of ancestry cutoff for workers, all putative hybrids had parental genotypes from different species. At the 0.75 cutoff, parental genotypes of 2 of the 10 putative hybrids were from the same population. At the 0.9 cutoff, 7 of the 19 putative hybrids had parental genotypes from the same population, and one had a paternal genotype that could not be assigned to either species (Table 3.3). The 0.65 proportion of ancestry cutoff for workers thus seems to return very small numbers and potentially no false positives. At the 0.75 cutoff, 20% (2 of 10) of the putative hybrids appeared to be false positives, and at the 0.9 cutoff, the rate was 42%. There was no statistical difference between the hybridization rates in the two possible directions (*D. molestus* queen and *D. wilverthi* male vs. *D. wilverthi* queen and *D. molestus* male) at any of the cutoffs (0.65 cutoff $X_1^2 = 0.37$, $p = 0.54$; 0.75 cutoff $X_1^2 = 0.27$, $p = 0.60$; 0.90 cutoff $X_1^2 = 1.23$, $p = 0.27$). There was also no statistical difference between the hybridization rates giving rise to major and minor worker subcastes in either species (*D. wilverthi*: 0.65 cutoff $X_1^2 = 1.81$, $p = 0.17$; 0.75 cutoff $X_1^2 = 2.10$, $p = 0.10$; 0.90 cutoff $X_1^2 = 2.04$, $p = 0.15$; *D. molestus*: 0.65 cutoff $X_1^2 = 0.34$, $p = 0.56$; 0.75 cutoff $X_1^2 =$

1.01, $p = 0.31$; 0.90 cutoff $X_1^2 = 1.64$, $p = 0.20$) or overall (0.65 cutoff $X_1^2 = 0.50$, $p = 0.48$; 0.75 cutoff $X_1^2 = 0.40$, $p = 0.52$; 0.90 cutoff $X_1^2 = 0$, $p = 1$).

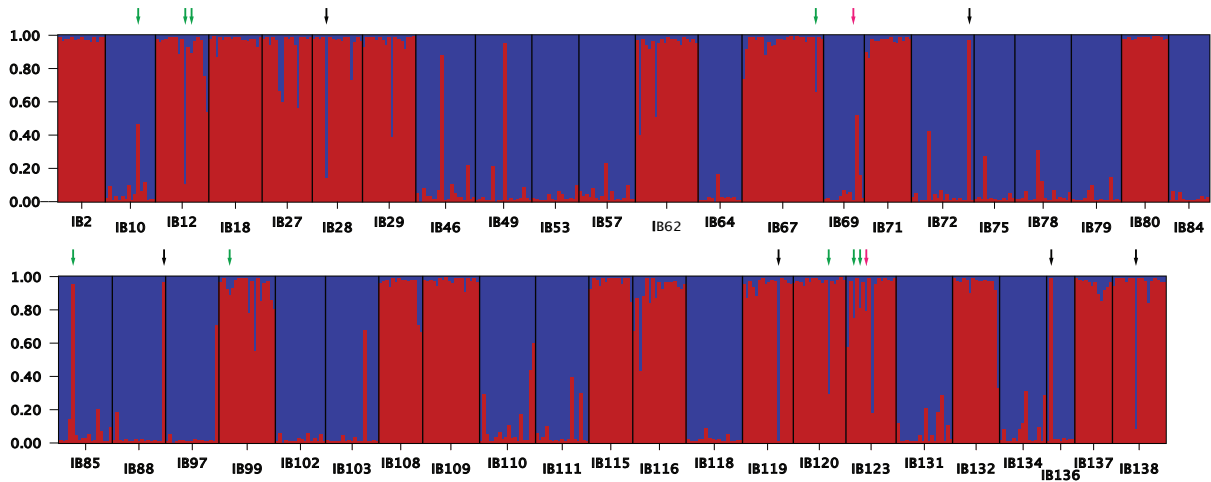


Figure 3.6 Assignment of reconstructed male genotypes to two populations using the clustering program STRUCTURE. Male genotypes are grouped by queen they mated with. Arrows above the plot indicate fathers of hybrid workers at three different proportion of ancestry cut- offs. Black arrows indicate putative hybrids at the 0.65 cut-off, pink arrows indicate the two additional hybrids included at the 0.75 cut-off, and green arrows indicate the further additional hybrids included at the 0.90 cut-off. Red indicates proportion of ancestry from *D. molestus*, and blue indicates proportion of ancestry from *D. wilverthi*. This is one representative run of five replicates.

Mating frequencies

As an additional check that the population structure is consistent with previous measurements, I measured the average mating frequencies of queens in the population using the predicted queen and male genotypes. The mean observed mating frequencies (observed number of mates) of *D. molestus* and *D. wilverthi* queens were 16.57 ± 0.60 (arithmetic mean, SE), and 16.22 ± 0.51 (arithmetic mean, SE), respectively. The effective mating frequencies were 35.96 ± 4.15 (harmonic mean, SE) for *D. molestus* and 35.42 ± 5.17 (harmonic mean, SE) for *D. wilverthi* (Table 3.4). There was no statistical difference in the observed mating frequencies (two

sample $t_{20} = 0.45$, $p = 0.65$, two tailed) or effective mating frequencies (Mann-Whitney $U = 215.5$, $Z = 0.60$, $p = 0.55$, two tailed) between the two species.

Table 3.4 Mating frequencies of 44 reconstructed queens.

<i>D. molestus</i>				<i>D. wilverthi</i>			
Queen	n	K_{obs}	m_e	Queen	n	K_{obs}	m_e
IB2	20	20	15 27.265861	IB10	20	16	31.806167
IB12	20	17	47.688243	IB46	20	19	95.25066
IB18	20	17	47.688243	IB49	20	18	69.323092
IB27	18	16	49.50339	IB53	20	15	21.210341
IB28	20	16	31.806167	IB57	20	18	63.556338
IB29	20	17	47.688243	IB64	20	14	23.859881
IB62	20	20	190	IB69	20	13	12.729196
IB67	27	26	120.4114	IB72	20	20	190
IB71	20	15	28.308118	IB75	17	13	22.692662
IB80	20	15	27.265861	IB78	20	18	63.556338
IB99	20	18	63.556338	IB79	20	16	35.105517
IB108	20	14	21.210341	IB84	17	14	27.229871
IB109	20	18	63.556338	IB85	20	17	38.160677
IB115	19	14	24.390501	IB88	20	17	47.688243
IB116	20	17	38.160677	IB97	20	17	47.688243
IB119	20	16	38.160677	IB102	17	16	66.363709
IB123	20	16	38.160677	IB103	19	17	55.642378
IB132	20	15	31.806167	IB110	20	18	63.556338
IB137	20	12	14.686737	IB111	20	17	47.688243
IB138	20	17	38.160677	IB118	20	18	63.556338
				IB131	20	18	63.556338
				IB134	20	15	31.806167
				IB136	20	9	12.729196
Overall	404	16.55±2.82	35.67±0.02		450	16.22±2.38	35.42±0.03

n is the number of offspring sampled for each queen. K_{obs} is the observed number of males mated with each queen (arithmetic mean \pm SE), and m_e is the sample size corrected estimate of the effective mating frequency per Nielsen et al., (2003) (harmonic mean \pm SE).

Discussion

While previous work had shown that historical mitochondrial introgression is rampant in East African driver ants (Kronauer et al., 2011b), the extent of ongoing interspecific mating and its possible effects on nuclear genetic admixture and genetic caste determination has remained unknown. This study analyzes a large population genetic dataset from driver ants at Kakamega Forest in Kenya, which for the first time allowed the reconstruction of within-colony pedigree relationships as well as queen and male genotypes for this population. Information on colony pedigree relationships and the genotypes of reproductives is crucial to infer mating systems, patterns of hybridization, and the potential for GCD in social insects. I find that the mating systems of the two species at Kakamega Forest are similar to those of other army ants in that the queens mate with many males (e.g. Kronauer et al., 2004b; Kronauer et al., 2006; Kronauer et al., 2007b). However, in the Kakamega Forest population ca. 1-2% of workers in colonies of both species are derived from interspecific matings, i.e. are F1 hybrids. This represents the first clear case of ongoing hybridization in an army ant. Cases of interspecific or interlineage matings in ants are frequently associated with GCD, and hybridization in ants is commonly studied in this context. This study, on the other hand, presents a case of hybridization in ants that is more typical for what is commonly observed in solitary species, despite the fact that the polyandrous mating system of driver ants should be particularly conducive to GCD.

Mating Frequencies

A higher than expected mating frequency could indicate errors in the dataset or biological oddities in the mating system normally observed in driver ants. The observed mating frequency of both species is consistent with previous measurements of queen mating frequency in *Dorylus*

populations (Kronauer et al., 2004b; Kronauer et al. 2006). However, the effective mating frequency is higher than previous measurements (Kronauer et al., 2004b), which is likely simply an artifact of the fact that smaller numbers of workers per colony were sampled compared to previous studies whose main goal it was to estimate mating frequencies (e.g. Kronauer et al., 2004b). I therefore conclude that, in terms of mating frequencies, the mating system of the two driver ant species at Kakamega Forest is probably typical for army ants.

Recent immigrants to Kakamega Forest

One of the colonies I sampled from Kakamega Forest had the signature of being a recent immigrant to the population. The mitochondrial haplotype for colony IB117 was more closely related to two *D. molestus* colonies from other localities than to any other haplotype from Kakamega Forest (haplotype 53 in Figure 3.2, Table 3.1). These haplotypes are from Nakuru, Kenya and Semliki, Uganda, and are 43 and 44 nucleotide changes from the closest sample from Kakamega Forest, respectively, and 52 and 53 nucleotide changes from the main cluster of Kakamega Forest haplotypes, respectively. This suggests that this haplotype did not originate in Kakamega Forest, and IB117 may be a recent immigrant to the population. The nuclear microsatellite data support this. In the PCA of the worker genotypes, IB117 forms a separate cluster distinct from either of the two species *D. molestus* and *D. wilverthi*. Given that queens mate in their natal nest and have not been shown to re-mate later in life (Kronauer & Boomsma, 2007a), this suggests that the queen did not mate with local males. Instead, she may have mated with males from her native population who sired the worker cohort I sampled, and subsequently migrated to Kakamega Forest, although I was unable to reconstruct a queen genotype for this colony to confirm this.

Historical hybridization and mitochondrial introgression

As shown in previous work (Kronauer et al., 2011b) the mitochondrial haplotypes of both species at Kakamega Forest cluster more closely together than with conspecific haplotypes from other localities. The observation that all haplotypes from *D. wilverthi* are nested entirely within a cluster of *D. molestus*, rather than clustered with *D. wilverthi* haplotypes from other populations (Figure 3.2) suggests that there have been historical hybridization events between the two species, although it is not clear whether haplotypes moved from *D. wilverthi* to *D. molestus* or *vice versa*. I was unable to replicate the previous finding of a shared haplotype between the two species, which would provide further evidence of historical introgression. A possible explanation for this is that sampling for Kronauer et al. (2011b) was geographically more extensive (samples were taken from both northern and southern areas of Kakamega Forest) and sampled a greater number of colonies. This study focused the majority of its sampling on the southern area of Kakamega Forest (see Appendix F) and sampled a much larger number of workers from fewer colonies. Indeed, all *D. molestus* colonies with haplotype 1 found previously came from the northern areas of the Kakamega National Reserve, Kisere National Reserve, Malava Forest and the farmland lying between these areas (Kronauer et al., 2011b; see map in Peters et al., 2011). This and the observation that *D. wilverthi* haplotypes are nested within *D. molestus* haplotypes indicate that this population has a complex history of introgression that may best be explained by multiple introgression events.

Ongoing hybridization

The mitochondrial data show that there has been introgression between the two species, and previous work also suggested that hybridization currently occurs in the population Kronauer

et al. (2011b). However, it did little to show the extent of hybridization or establish the presence or absence of hybrid queens, which would mediate introgression. The nuclear microsatellite data presented here show that hybridization is clearly happening by identifying specific individuals as F1 hybrids. However, the hybrid workers are relatively rare. Kronauer et al. (2011b) estimated that between 7 and 18 percent of the worker population at Kakamega Forest are hybrids between the two species. My data show that the hybridization rate is in fact much lower. Using the same proportion of ancestry cutoffs as Kronauer et al. (2011b), I estimated the frequency of hybrids to be between 1.13% at the 0.75 cutoff and 2.16% at the 0.90 cutoff. However, as I have shown, these measurements have a high incidence of false positives, and a more accurate measurement may be at the 0.65 cutoff. At this level, the frequency of hybrids is 0.90%. Additionally, I show that hybridization goes in both directions at approximately similar frequencies. This is consistent with a complex history of hybridization between the two species observed in the mitochondrial DNA sequences and also with the intermediate morphologies of the two species at Kakamega Forest (Kronauer et al., 2011b). The higher estimate previously reported is likely due to using fewer microsatellite markers (5 markers in Kronauer et al. (2011b) versus 12 markers used in this study). A small number of markers can impede the ability of STRUCTURE to accurately assign individuals to populations, and 10 or more microsatellite markers can significantly increase this accuracy (Rosenberg et al., 2001; Manel et al., 2002). Kronauer et al. (2011b) also used a smaller number of samples (110 workers from Kakamega Forest), but except in extreme cases, sample sizes have a very small effect on cluster assignment accuracy using STRUCTURE (Turakulov & Easteal, 2003).

To confirm that the increased number of loci rather than the sample size is the cause of difference in estimated hybridization rate I reran the STRUCTURE analysis on the workers, first

by reducing the number of samples and then by reducing the number of microsatellite loci. To reduce the number of samples I divided the dataset used to calculate hybridization rate in half and ran STRUCTURE analyses on each half separately. I then pooled the results and calculated the overall hybridization rate. The first half of the data contained 20 workers per colony from colonies IB2 to IB84, and the second half contained the same number of workers from colonies IB85 to IB138 (colonies IB40, IB63, IB83, IB87, IB117 and IB124 were excluded) I repeated this analysis dividing the dataset into quarters (colonies IB2 to IB57, IB62 to IB84, IB85 to IB115, and IB116 to IB138). In both cases, the number of hybrids recovered was not significantly different from that found running the whole dataset in a single run at any of our cutoffs. I then ran STRUCTURE analyses with reduced numbers of microsatellite loci: two different sets of six loci and one set of five loci. All three analyses significantly increased the number of hybrids detected in the data at all cutoffs (Appendix G), and for two of these runs the hybridization rate was almost identical to that found by Kronauer et al. (2011b). This confirms that the difference in hybridization rate between the two studies is due to the number of microsatellite markers used and highlights the importance of a sufficient number of markers in population genetics studies.

Another concern may be that the pattern of apparent admixture observed in the output of a clustering algorithm such as STRUCTURE may not actually be due to ongoing admixture. For example, the same pattern might also be observed under other scenarios involving shared ancestral alleles that give the superficial appearance of admixture (Lawson et al., 2018). In the case of my data, if the pattern observed in the workers was due to the presence of shared ancestral alleles, we would expect the same pattern to be apparent in the queen dataset, which is not the case. Furthermore, if the observed pattern was due to hidden shared ancestral alleles we would expect the majority of the workers in each affected colony to show signs of admixture by virtue of sharing

the same mother. However, only very few workers in any given colony show signs of admixture. The conclusion that the observed pattern of admixture in the worker caste is the result of the presence of F1 hybrids rather than of shared ancestral alleles therefore seems robust.

Although the frequency of hybrids in the population is very low, 15% to 32% of the colonies we sampled contained hybrid individuals. This high proportion makes sense considering the extremely high mating frequency observed in driver ants. Queens regularly mate with 20 or more males (Kronauer et al., 2004b; Kronauer et al., 2006; this study), and an occasional interspecific mating could explain the observed pattern. This is also consistent with the observation that, in some other *Dorylus* (*Anomma*) species, heterospecific males are occasionally found inside colonies (Raignier et al., 1974). Furthermore, high mating frequencies may reduce the potential deleterious effects of occasional interspecific hybridization because in each case only a small proportion of the workers in a colony will be affected. Because army ant workers normally do not reproduce (Kronauer et al., 2010), common and often significant consequences of hybridization, such as hybrid sterility, will incur no or little additional costs. At the same time, the increase in genetic diversity among the worker force derived from hybridization might in fact be beneficial in contexts such as disease resistance or task allocation (e.g. Boomsma et al. 2009).

Because interspecific hybridization is rare, mating between the two species is clearly not random, and there must be some isolating barriers preventing interspecific matings. A feature of wingless queens and reproduction by colony fission is that males must “run the gauntlet” of workers in order to mate with a queen (Franks & Hölldobler, 1987), and workers may be capable of detecting and eliminating heterospecific males before they have an opportunity to mate. The differences in habitat preference may also be contributing to reproductive isolation (Peters et al., 2009; Peters & Okalo, 2009; Schöning et al., 2006). Although males disperse on the wing, they

may be more likely to mate with conspecific queens due to similar habitat preferences or simply due to proximity. There are also many other factors that could contribute to reproductive isolation. If the two species had different preferences for seasonal mating conditions, the chances of interspecific matings would be reduced (Coyne & Orr, 2004; Rice, 1987; Rundle & Nosil, 2005). Furthermore, I cannot exclude the possibility that interspecific matings are more common, but in many cases do not result in viable offspring (Coyne & Orr, 2004; Stebbins, 1958). Such cases would have gone undetected in this study.

Hybridization and genetic caste determination

In the eight ant systems where genetic caste determination has been observed, the association between genotype and caste is always nearly perfect (i.e. interlineage genotypes become workers and intralinear genotypes become queens close to 100% of the time). In *Pogonomyrmex* and *Solenopsis*, more than 90% of workers are hybrids, and interlineage gynes are extremely rare (Helms Cahan & Vinson, 2003; Schwander et al., 2007b). In *Vollenhovia*, 95% of workers are produced from interlineage matings within the same species (Ohkawara et al., 2006), and in *Cataglyphis* and *Wasmannia*, genetic caste determination is perfect with 100% of workers being produced from interlineage matings, and queens being produced parthenogenetically (Fournier et al., 2005a; Leniaud et al., 2012). Less evidence exists for the three *Messor* species, but the existing data point to similarly high rates of interlineage workers and intralinear queens with little if any leakage (Romiguier et al., 2017). The connection between genetic caste determination and hybridization in ants is not clear, but in more recently discovered cases of hybridization in ants, it warrants examination to decide whether caste is determined at least in part by genetics. Our analysis found that the hybridization rate is low, with more than 97% of workers

sampled being produced from intralineage matings. Hybridization between *D. wilverthi* and *D. molestus* therefore clearly does not result in a strong association between caste and genotype as seen for example in *Pogonomyrmex*.

Chapter 4: Future directions

I did not find evidence of genetic caste determination in the hybrid zone between *D. molestus* and *D. wilverthi*, so I was unable to make a strong conclusion about the connection between hybridization and genetic caste determination. However, in the course of this work I noticed that the particular distribution of species in this group (*Dorylus* subgenus *Anomma*) could constitute an ideal model to study the prevalence of genetic exchange (introgression) between closely related species, and to determine how introgression is influenced by the time since divergence of closely related species.

The potential for introgression to act as an evolutionary force between distinct populations, whether these are species, subspecies, ecotypes or other distinct forms, has been recognized for more than 60 years (Anderson, 1949). An early definition of introgression was “the infiltration of the germ plasm of one species into that of another” via repeated backcrosses to a parental population (Anderson & Hubricht, 1938). It has been most frequently described in plants, although its importance in animals has been increasingly recognized (Harrison, 1993). Introgression historically was thought to be a rare phenomenon, with species boundaries being strictly maintained (Coyne & Orr, 2004; Dobzhansky, 1937; Mayr 1963), but has been discovered to be far more common than previously thought, and can be pervasive throughout the genome in some cases (Arnold, 2006; Arnold & Martin, 2009; Baack & Rieseberg, 2007; Hedrick, 2013; Vallejo-Marín et al., 2016; Wu, 2001).

An important question about introgression that is still unanswered is the relationship between divergence time and the amount of introgression permissible in the genome (Rieseberg & Wendel, 1993). It seems clear that the likelihood of detecting introgression between two taxa

should decrease with increasing divergence time, with more distantly related species having fewer potential introgressive loci and more closely related species having more such loci, but the shape of the function is still unknown.

One clue to the shape of this function may come from studies that attempted to measure the relationship between the number of genetic incompatibilities between different species and divergence time. Hybrid incompatibilities arise from deleterious interactions between loci in independently evolving lineages that have not been “tested” against each other in a hybrid individual (Bateson, 1909; Dobzhansky, 1936; Muller, 1942). Because the lineages are evolving independently, each new mutation has not been “tested” against all mutations in the other lineage up to that point in time, so as divergence time increases, the number of potential interactions grows quadratically (Figure 4.1) (Orr, 1995). Studies that have measured this effect suggest that, as divergence time increases, the number of actual incompatibilities between any two species “snowballs”, so that the number of incompatibilities causing sterility or inviability of hybrids should increase at a rate that is faster than linear with respect to divergence time (Matute et al.,

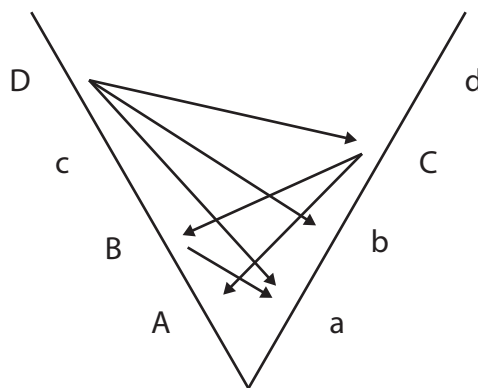


Figure 4.1 Illustration of the history of substitutions between two independently evolving lineages. Both populations are initially fixed for lower case alleles. Time moves upwards. Arrows indicate “untested” allele combinations that may give rise to hybrid incompatibilities.

2010; Moyle & Nakazato, 2010; Sherman et al., 2014; Wang et al., 2015). Because these incompatibilities cause hybrids to be sterile or inviable, there is no hybrid intermediate by which they can move from one species to another, and the regions containing them should be unable to introgress. This suggests that the shape of the function of introgressed regions versus divergence time should be an upside down version of that of the number of hybrid incompatibilities versus

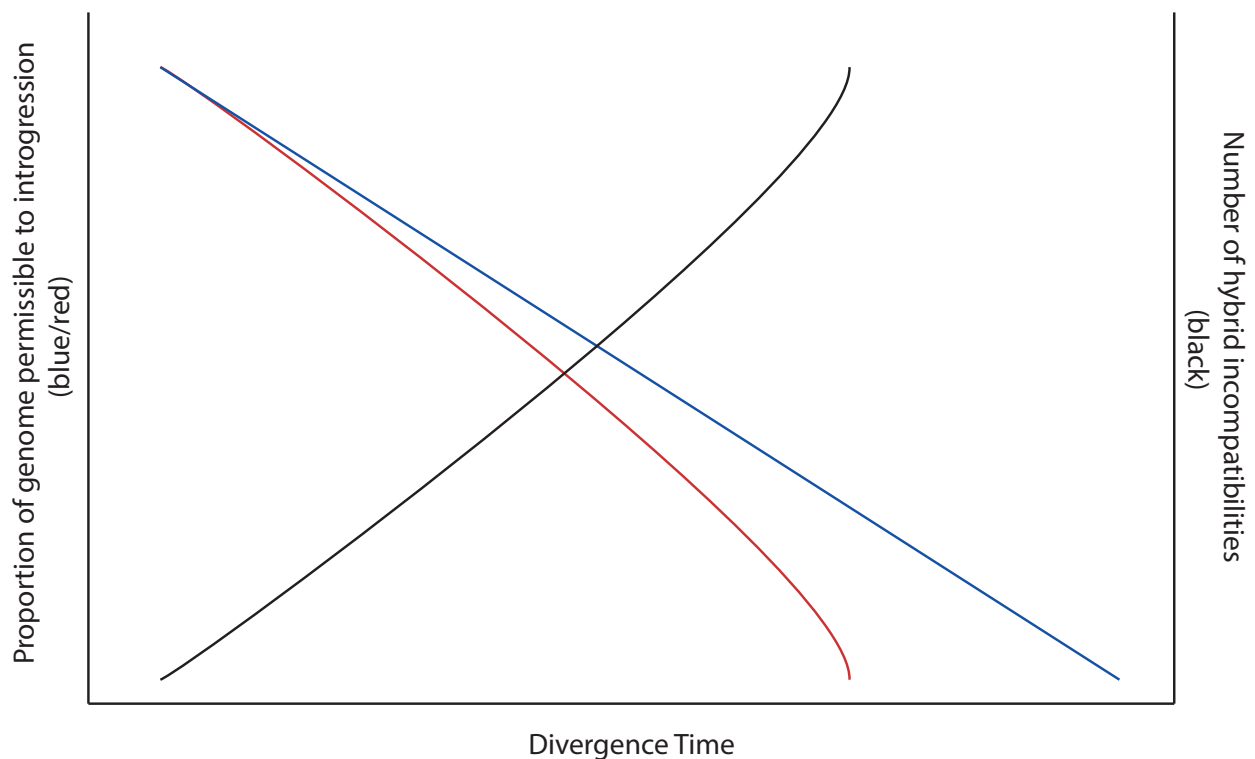


Figure 4.2 Illustration of the possible shapes of the relationship between divergence time and the proportion of the genome that is permissible to introgression and the relationship between divergence time and the number of hybrid incompatibilities. The black line represents the accumulation of hybrid incompatibilities over time as predicted (Orr, 1995) and as measured in previous studies (Matute et al., 2010; Moyle & Nakazato, 2010; Sherman et al., 2014; Wang et al., 2015). The blue line represents the case where the proportion of the genome that is permissible to introgression decays linearly over time. The red line represents the case where the relationship is faster than linear and is the upside down version of the black line. The red line is predicted to be reflective of the true relationship.

divergence time, i.e. the proportion of the genome that is permissible to introgression should decrease faster than linearly with respect to divergence time (Figure 4.2).

Measuring the shape of this function entails measuring the amount of introgression between multiple pairs of species and plotting the proportion of the genome that has introgressed between each species pair against divergence time. A model can then be fit to this chart to determine the shape of the function. Using wild populations has several advantages over using laboratory studies. The first is that introgression is measured in a natural setting. This means that all types of hybrid incompatibilities (intrinsic and extrinsic) are taken into account. Any locus that prevents the formation of hybrids or otherwise prevents backcrossing to the parental populations will not be permissible to introgression. Second, using natural populations precludes the requirement for currently hybridizing species. Previous studies that measured the “snowball effect” were restricted to identifying hybrid incompatibilities between species that interbreed and produce viable hybrids in the lab. This means that the number of potential points that can be used to fit a model is severely limited. Matute et al. (2010), Sherman et al. (2014), and Wang et al. (2015) were each only able to plot three points with one of them being assumed, and Moyle & Nakazato (2010) were similarly only able to plot 4 points including one assumed point. Furthermore, Moyle & Nakazato (2010) and Sherman et al. (2014) were both done in tomato plants, so cannot be considered independent tests of the snowball effect. Measuring the amount of introgression in the genomes means that any pair of species that has experienced hybridization and introgression in the past can be used, even if reinforcement has since strengthened the species boundaries so that hybridization no longer occurs today.

A suitable group in which to measure the shape of this function has several attributes. 1) A sufficient number of species that can potentially interbreed. 2) Species ranges that include areas of

sympatry, but also have large areas where only one species of the group is present. Significant areas of allopatry from all other species in the group will reduce the potential for introgression from other sources, which will make it easier to identify regions that have introgressed from the species of interest. After viewing collection data from a collaborator, Caspar Schöning, I hypothesized that *Dorylus* would be a good system in which to test this. Caspar has collected *Dorylus* samples from across Sub-Saharan Africa. He has records from nine species, several of which appear to have significant areas of allopatry where they are the only driver ant species, as well as areas of overlap with neighboring species (Figure 4.3).

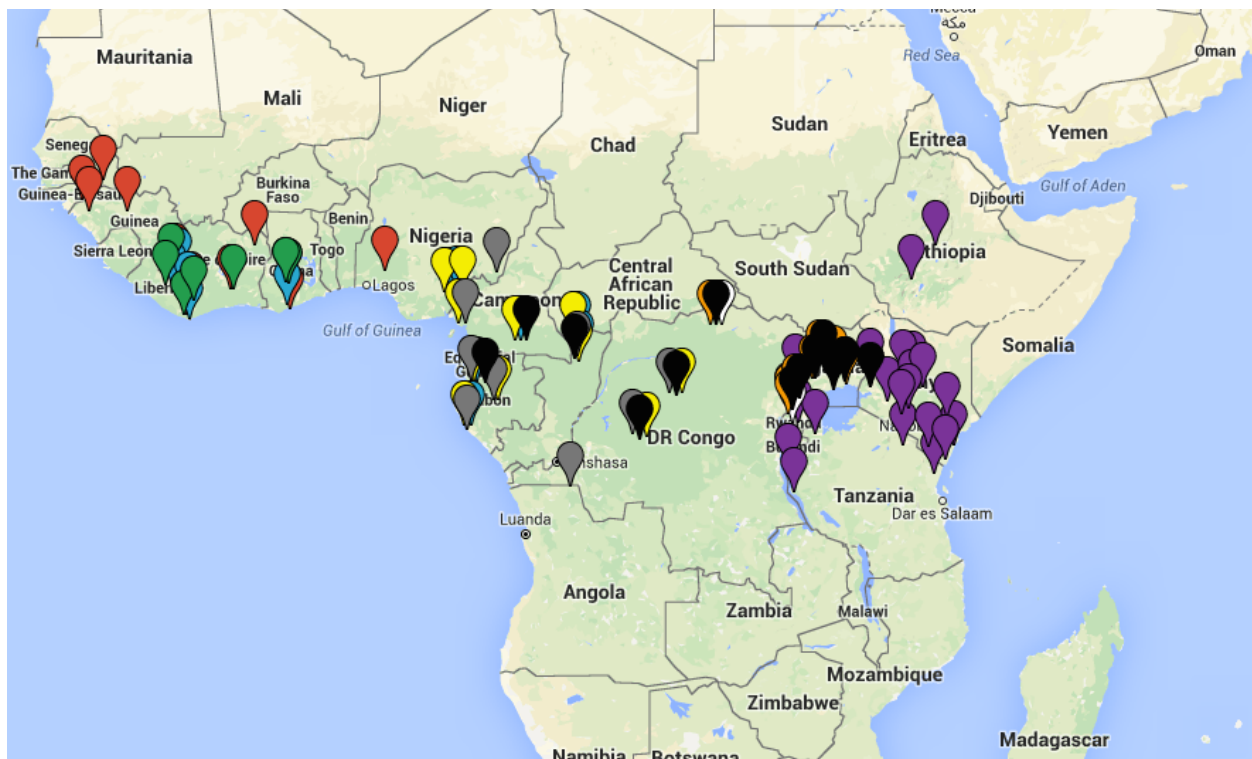


Figure 4.3 Map of collection localities for nine species of *Dorylus* (*Anomma*). There are significant areas of allopatry for several of the species. This map is not exhaustive. It represents the collection of a single researcher. More research is needed to identify potential localities for further collection. Red = *Dorylus burmeisteri*, blue = *D. mayri*, purple = *D. molestus*, green = *D. nigricans*, yellow = *D. sjoestedti*, orange = *D. terrificus*, grey = *D. rubellus*, black = *D. wilverthi*, white = *D. niarembensis*.

This project requires four steps that I have outlined below.

1. Phylogeny of the driver ant subgenus *Anomma*.

There is currently not a species level phylogeny of the genus *Dorylus*, which is necessary in order to estimate divergence times and to identify species pairs that may have experienced introgression. According to AntWiki.org, there are 15 described species in the subgenus *Dorylus* (*Anomma*) not including subspecies (AntWiki.org accessed 24 July 2019), but they do not appear to form a monophyletic group. All swarm-raiding species within the subgenus do, and there is a clear phylogenetic separation between the surface foraging swarm-raiders and the leaf-litter foragers or subterranean species (Kronauer et al., 2007b). A phylogeny could be constructed using available samples from collections or collaborators. The phylogeny could be constructed using reduced coverage next generation sequencing such as genotyping by sequencing (GBS) (Elshire et al., 2011).

2. Identify potential hybrid zones among the species and collect samples of allopatric and sympatric populations for all species pairs.

Use the phylogeny constructed in step one in conjunction with range information to determine areas where hybridization is likely to have occurred. These areas would be where pairs of closely related species currently have overlapping or adjacent ranges. Hybridization has been observed in species that diverged as long as 10-12 million years ago (Matute et al., 2009), so even more divergent species could show evidence of past introgression events. The four-taxon D-statistic (Durand et al., 2011) is a metric for detecting admixture between divergent lineages based on SNP

frequencies that are discordant with a hypothesized species tree topology and could be used to identify species or populations that have exchanged genetic information (Eaton et al., 2015). This method would not identify all potentially hybridizing populations, but it could quickly identify some areas where hybridization and introgression has most likely occurred. Samples would then need to be collected from the populations identified as potentially hybridizing. Samples should be collected from the area of sympatry, where both species occur and where hybridization has most likely happened, and also from areas of allopatry. To reduce complications due to confounding genetic information, it is preferable to collect from areas of allopatry where only one *Dorylus* (*Anomma*) species occurs.

3. Sequence the samples collected in step two and compare the genomes of samples from hybridizing sympatric populations to those from allopatric populations to identify regions of the genome that are a result of introgression.

The purpose of this step is to identify which regions of the genome originated in the same species and which originated in the species it has been hybridizing with. To get an accurate estimate of the proportion of the genome that originated in another species, a large portion of the genome of all samples needs to be sequenced. This requires a sequencing method that is suitable for populations with high genetic diversity and that can cover a relatively large proportion of the genome, and genotyping by sequencing (GBS) (Elshire et al., 2011) meets these requirements. Once the genomes of a sufficient number of samples have been sequenced, a comparison between genomes from the area of sympatry and populations of both species from areas of allopatry must be made. There are several methods that have been recently developed to detect introgression on a genomic scale, and to identify specific regions that originated in different species (Larson et al., 2013;

Schumer et al., 2016). One advantage to doing the project in ants is that they are haplodiploid. Sequencing these populations from haploid males therefore removes the uncertainty in determining whether a particular SNP is heterozygous in an individual or a sequencing error. Furthermore, Hymenoptera genomes tend to be moderately sized with most being between 180 and 340 Mb (Tsutsui et al., 2008) and have a low content of repetitive and transposable elements making them highly tractable for genome sequencing (Branstetter et al., 2017).

4. Plot the proportion of the genome from each species that is permissible to introgression from its hybridizing sister species against divergence time and fit a curve to the set of points to find the shape of the graph.

Once all samples are sequenced, measure the proportion of the sequenced portion of the genome of each species pair that originated from the other species, and plot this number against divergence time. Since estimates of divergence time can vary widely, a proxy for divergence time, such as K_s , should be used. K_s is the average number of synonymous substitutions per gene between each species pair. Each pair of species corresponds to one point that is plotted on the graph. It can be assumed that at the point in the past when any two current species were one species, 100% of their genomes were permissible to introgression. This allows me to place an assumed point at 100% introgression and zero divergence time. Then Akaike Information Criterion (AIC) is used to determine the best fit model of the points.

These results will shed light on how permissibility to introgression is related to divergence time, a currently unanswered question in evolutionary biology. Furthermore, if the hypothesized shape proves correct, it will provide more evidence supporting the “snowball effect” of the

accumulation of hybrid incompatibilities, which is currently supported by only four studies with relatively little data.

Conclusions

There are important differences in how hybridization acts between haplodiploid social insects and other animals. However, important questions remain unresolved. First, despite predictions, it is not clear how eusociality and haplodiploidy affect the propensity to form hybrids in nature, and second, the relationship between interspecific or interlineage hybridization and genetic caste determination remains murky. I attempted to address portions of these questions in this thesis. To facilitate future research into ant hybridization and population genetics, I developed a set of universal ant microsatellite markers. These tools will be useful for population genetics or pedigree studies in newly discovered or little studied ant species where no microsatellites have been previously developed and no genomic data are available. I then used these tools to investigate a hybridizing population of *Dorylus* army ants in western Kenya. I discovered that hybridization in this population is much lower than expected from previous estimates. Because of the low hybridization rate, I was unable to identify a genetic bias to caste determination, although the possibility cannot be excluded. Although there may be a connection between hybridization and genetic caste determination, strong genetic caste determination, where caste is determined primarily by genotype, was not observed in this case and is clearly not a necessary consequence of hybridization.

Lastly, while in the pursuit of this study, I identified the possibility to answer a standing question in speciation research: what is the relationship between divergence time and the permissibility of the genome to introgression between closely related hybridizing species? Upon

viewing collection data from a collaborator, I speculated that the genus *Dorylus* would be a good candidate to answer this question. It has relatively few species that are broadly distributed across sub-Saharan Africa. There appear to be large areas where only one *Dorylus* species is present, and multiple sites where the ranges of two or more species overlap. The proposal would involve sequencing the genomes of multiple samples of each species from areas of allopatry and areas of sympatry with other species and comparing the genomes of allopatric and sympatric populations to determine how much of the genome in the area of sympatry resulted from interspecific introgression. The last step would be to make a plot of the proportion of the genome permissible to introgression in each species versus divergence time and fitting a model to these points. This project has the potential to answer an open question in evolutionary biology and may provide further support for the snowball effect in the accumulation of hybrid incompatibilities.

Appendix A Details of microsatellite loci in eight ant genomes. Numbers in parentheses indicate there is an unknown number of bases inserted into the sequence of in the genome that is included in the size of the targeted fragment. The number inside the parentheses is the estimated number of inserted bases. Total base pairs (number of N bases inserted). Some loci have multiple motifs listed. All motifs are in the same region and are included in the size of the targeted fragment.

Appendix A continued

[illegible]

[illegible]

Appendix A continued

Locus	Safinavir motif (S)	Safinavir motif		Acronimax echovirus	
		Position on scaffold	Size of targeted fragment in base pairs	Position on scaffold	Size of targeted fragment in base pairs
Am10	FAGTCTTACAGGTACACTTGG	711867-71435	132	(AG) ¹⁰	156
Am175	FCTAGTTCGAAATGTTGGC	60262-60433	218	(TTG) ¹ (CTG) ¹	288
Am189	FACAGGCGGCTGATGCTGGA	377870-377913	193	(GAG) ₁ (GATGATGA) ₁	No motif
Am1343	FCTGGTCTGCTCTCTGAT	91379-914093	256	(AG) ₁ (A) ₁ (E) ₁ (A) ₁	282
Am1388	RAGGGGCGCTCGAATTCTCT	429571-429588	281	(TTG) ¹ (TTC)	472 (87)
Am1388	RCTAGGAGTGGGGGTGAT	429571-429588	281	(TTG) ¹ (TTC) ₁ (N ₁ CT) ₁	472 (87)
Am1341	F-AAGAGACTTCTCGG-AAG	564608-564612	255	(CAG) ₁ (N ₁ CA) ₁	208
Am13794	F-TGGTGTGGCTGTGTGCAAGG	286134-286193	361	(AG) ₁ (AGGAGCA) ₁ (AG) ₁	236
Am13916	RTGGCTCCGATGATATCTGT	628765-628798	375	(AG) ₁ (N ₁ AG) ₁	341
Am1648	FCTCTGGTCTGATCTGA	1301497-1301975	341	No motif	396
Am1648	RTAGAGATGCTGCTGTGG	1301497-1301975	341	No motif	396
Am1648	R-GGATATCGGACGGT	228955-23279	255	No motif	321
Am1993	F-TATCCGCTCTAATAATTAGATGGG	306234-306716	361	(GAG) ₁	387
Am1155	FAGGATAGTTTGGGGTTATGG	546923-547374	199	(GT) ₁ (N ₁ CT) ₁	173
Am1035	R-ACTGACTGAGTGTATTGAGGT	784837-785368	313	No motif	299
Am1249	F-AGTGTCAAGGGGACTGAG	668715-669171	363	No motif	373
Am18424	RTAGGAGAAATGAGACCGG	782011-782443	269	No motif	244
Am18424	R-GGGATATACAAATGGGAC	791397-791967	232	(TTC) ₁	No motif
Am1898	RTGTTCGAGCTGAGTGGCTTC	2248099-2248972	200	(TTG) ₁	196
Am1181	F-TCCACTTACGCTGTGCAAC	218778-218823	293	(ACG) ₁	291
Am10318	F-AGCCATTTCGCTGTGA	886529-887015	357	1. (TTG) ¹ (CTT) ₁ 2. (TGG) ₁ (N ₁ CT) ₁	395
Am10878	R-GATCAATCCGCACTGAT	604226-604616	325	(GAGG) ₁ (GGAG) ₁	330
Am13345	F-AGGTTTGGGAGGTTGAGG	107497-108852	342	(ATC) ₁ (N ₁ ATC) ₁ (N ₁ ATC) ₁ (GAT) ₁ (ATC) ₁	276
Am11400	R-GGAAATATGAGTGTGGG	516339-516800	253	(AG) ₁	257
Am11400	R-GGAACTGTATGAATCTCAC	2131034-2131506	346	(AGAGAAAG) ₁ (N ₁ AG) ₁	388
Am118220	F-AGGTTTGGGAGGCTGATG	1142107-1142611	327	(GGAG) ₁ (GGAG) ₁	363
Am121	R-CATCAATGCACTGCACTGG	714737-71955	234	(AG) ¹⁰	234
Am1808	F-AGGAGTCTAGTGTCTTGG	2211837-2292	193	(TTCG) ₁ (TTT) ¹ (TCC) ₁	198
Am11049	R-CTGGGCGCTCTGTGTAT	921994-922462	246	(GAA) ₁ (N ₁ GGAA) ₁	325
Am1052	R-CTTTTGTGAGGCTGGG	908910-909383	272	(AGC) ₁	263
Am1387	R-CAAGGCGACTGCTCTCC	2854102-285874	221	(GTCTG) ¹ (TTCG) ¹ (GGT) ₁	197
Am1732	F-ATGATACGATGCAAGTGGC	771723-77201	181	(CAC) ₁ (N ₁ CA) ₁	No motif
Am1809	F-ATGAGTCTGAGTGTGAGT	2527095-2527569	396	(CGA) ₁ (TGA) ₁ (GGTGG) ₁ (G) ₁	220
Am1382	F-CCGAACTCTACTCTCC	907120-907596	205	(GTG) ¹ (ATGG) ₁	208
Am1395	F-CGAGGGGCGTGGGCACTCC	136518-136957	236	(CAA) ₁ (CAG) ₁ (AA) ₁	232
Am13111	RTGAGAGAGACGCTGTGT	362486-363023	209	(CAG) ₁ (CAACAG) ₁	227
Am1452	F-TGTGGATGCTGGGATGGGA	835966-836440	220	(TTG) ₁	238
Am1505	R-ATGGAGGAGTGTGGGAG	2359827-2360354	300	No motif	303
Am1541	F-TGCAAGATGTTCTGAGGT	341419-341876	275	1. (E) ₁ (A) ₁ 2. (E) ₁ (A) ₁ (AG) ₁	262
Am1709	F-AGGGGTAAAGGTTTAAAGGA	956509-956893	319	(GGC) ¹ (T) ₁	493
Am1739	F-TGCTCTCTCTGACATC	839298-839346	769 (595)	(TTCG) ₁	180
Am17204	F-CCCATCTCTGTCACTCT	251074-251485	213	No motif	270
Am1544	R-CTAGGATCTTAAAGTACCA	40229-46225	878 (473)	No motif	316
Am1564	F-TTAAAGGCCGCCGCTCT	1856731-1857155	237	(AG) ₁	196
Am10427	F-ATCAGCTTACGCGCTTAA	5855948-5856937	323	(TTC) ¹ (GGTCT) ¹ (CTT) ¹ (CT) ₁	290
Am11610	R-ATCAGGATGAGGCGGTGAA	907131-91191	268	(AG) ₁	288
Am11610	R-CTGAAATGCTGATGCTCT	1479-2026	389	1. (TTC) ₁ GGCGCGG(CGT) ₁ (N ₁ CT) ₁ 2. (E) ₁	323

Appendix B Physical linkage of microsatellite loci. When assembling a draft genome from next generation sequence data, sections of sequence are assembled into scaffolds, which represent large sections of the genome on the same chromosome. If two microsatellite loci are on the same scaffold, they are therefore on the same chromosome and are physically linked. Linked loci will not necessarily sort independently, and using such loci together in a study should be avoided. In the table on the following pages, each rectangular box represents a comparison between two microsatellite loci and is subdivided into eight sections representing the eight ant species for which genomes were available at the time of this work. Only the top half of the table (shaded in blue and white) is used. The bottom half (shaded in brown) represents redundant comparisons. Blue and white shading is used to make reading the table easier and does not carry any information. X indicates where two loci are on the same scaffold in that species. The order of species left to right in every box is *Pogonomyrmex barbatus*, *Harpegnathos saltator*, *Atta cephalotes*, *Ooceraea biroi*, *Linepithema humile*, *Camponotus floridanus*, *Solenopsis invicta*, *Acromyrmex echinator*.

Appendix B continued

[illegible]

Appendix B continued

Locus	Ant4155	Ant5035	Ant7249	Ant7680	Ant8424	Ant8498	Ant9181	Ant9218	Ant10878	Ant11315	Ant11400
Ant20	X	X
Ant575
Ant859
Ant1343
Ant1368
Ant2341
Ant2794
Ant2936
Ant3648
Ant3653
Ant3993
Ant4155
Ant5035
Ant7249
Ant7680
Ant8424
Ant8498
Ant9181
Ant9218
Ant10878
Ant11315
Ant11400
Ant11893
Ant12220
Ant21
Ant608
Ant1049
Ant1052
Ant1387
Ant1732
Ant2409
Ant3362
Ant3395
Ant3411
Ant3452
Ant3505
Ant3541
Ant4709
Ant5033
Ant7204
Ant8544
Ant9564
Ant10290
Ant10427
Ant11610

Appendix B continued

Locus	Ant11893	Ant12220	Ant21	Ant608	Ant1049	Ant1052	Ant1387	Ant1732	Ant2409	Ant3362	Ant3395
Ant20			X X X X X X X								
Ant575				X							
Ant859											
Ant1343							X X X X X X X				
Ant1368											
Ant2341											
Ant2794					X						
Ant2936				X X X X X X X X X X X X X							
Ant3648											
Ant3653			X								
Ant3993											
Ant4155											
Ant5035											
Ant7249											
Ant7680						X					
Ant8424					X						
Ant8498				X							
Ant9181			X								
Ant9218			X X X X X								
Ant10878											
Ant11315											
Ant11400											
Ant11893											
Ant12220											
Ant21											
Ant608											
Ant1049						X X X X X X X X					
Ant1052											
Ant1387											
Ant1732											
Ant2409											
Ant3362											
Ant3395											
Ant3411											
Ant3452											
Ant3505											
Ant3541											
Ant4709											
Ant5033											
Ant7204											
Ant8544											
Ant9564											
Ant10290											
Ant10427											
Ant11610											

Appendix B continued

Locus	Ant3411	Ant3452	Ant3505	Ant3541	Ant4709	Ant5033	Ant7204	Ant8544	Ant9564	Ant10290	Ant10427	Ant11610
Ant20												
Ant575									X		X	
Ant859												
Ant1343												
Ant1368												
Ant2341		X								X X X		X
Ant2794												
Ant2936												
Ant3648												
Ant3653											X X X X X	
Ant3993												
Ant4155						X X X X X X X						
Ant5035												
Ant7249												
Ant7680												
Ant8424			X									
Ant8498												
Ant9181												
Ant9218											X	
Ant10878												
Ant11315												
Ant11400												
Ant11893												
Ant12220												
Ant21											X	
Ant1049												
Ant1052												
Ant1387												
Ant1732												
Ant2409												
Ant3362							X X		X X			
Ant3395								X				
Ant3411				X								
Ant3452												
Ant3505												
Ant3541												
Ant4709												
Ant5033												
Ant7204								X				
Ant8544												
Ant9564												
Ant10290												
Ant10427												
Ant11610												

Appendix C P-values for tests of statistical linkage disequilibrium in universal ant microsatellite loci based on 33120 permutations. Adjusted P-value for 5% nominal level is 0.000030. NA indicates that one of the loci involved either did not amplify or was not polymorphic.

	Simopelta	Paraponera	Ecatomma	Solenopsis	Lasius	Dorylus	All
Ant11893 X Ant3648	1	NA	NA	NA	NA	0.20984	0.29734
Ant11893 X Ant5035	0.11241	1	NA	NA	NA	0.31072	0.11162
Ant11893 X Ant11315	NA	NA	NA	NA	NA	NA	NA
Ant11893 X Ant10878	1	NA	NA	NA	NA	0.9452	0.96208
Ant11893 X Ant1343	NA	1	NA	1	NA	0.21066	0.94182
Ant11893 X Ant2794	1	1	NA	NA	NA	1	1
Ant11893 X Ant3653	1	NA	NA	0.30737	NA	0.59405	0.31087
Ant11893 X Ant575	1	1	NA	0.80969	NA	0.3753	0.71105
Ant11893 X Ant20	0.05245	NA	NA	NA	NA	NA	0.05245
Ant11893 X Ant11400	0.3154	NA	NA	NA	NA	0.37893	0.17557
Ant11893 X Ant8424	0.39574	0.46751	NA	0.57132	NA	0.06561	0.04103
Ant11893 X Ant8498	NA	0.734	NA	NA	NA	NA	0.734
Ant11893 X Ant2341	0.39505	NA	NA	1	NA	1	0.96211
Ant11893 X Ant1368	1	NA	NA	NA	NA	0.29716	0.36374
Ant11893 X Ant9218	1	NA	NA	NA	NA	1	1
Ant11893 X Ant7249	1	NA	NA	NA	NA	1	1
Ant11893 X Ant2936	NA	NA	NA	1	NA	1	1
Ant11893 X Ant9181	NA	NA	NA	NA	NA	0.7186	0.7186
Ant11893 X Ant12220	NA	NA	NA	NA	NA	1	1
Ant11893 X Ant7680	1	NA	NA	NA	NA	1	1
Ant11893 X Ant859	0.05332	1	NA	NA	NA	1	0.34997
Ant11893 X Ant4155	0.20607	NA	NA	0.69055	NA	0.49535	0.18527
Ant11893 X Ant3993	0.00556	1	NA	0.96295	NA	0.55163	0.09354
Ant3648 X Ant5035	1	NA	NA	NA	1	0.13548	0.15501
Ant3648 X Ant11315	NA	NA	NA	NA	NA	NA	NA
Ant3648 X Ant10878	1	NA	1	NA	0.40082	1	0.841
Ant3648 X Ant1343	NA	NA	NA	NA	0.25447	0.08596	0.07421
Ant3648 X Ant2794	1	NA	NA	NA	0.3994	0.17687	0.07274
Ant3648 X Ant3653	1	NA	0.94925	NA	0.85112	1	0.92431
Ant3648 X Ant575	1	NA	NA	NA	0.48406	0.3766	0.23554
Ant3648 X Ant20	1	NA	NA	NA	NA	NA	1
Ant3648 X Ant11400	1	NA	NA	NA	NA	1	1
Ant3648 X Ant8424	1	NA	0.36709	NA	1	0.07343	0.37696
Ant3648 X Ant8498	NA	NA	NA	NA	NA	NA	NA
Ant3648 X Ant2341	1	NA	NA	NA	1	1	1
Ant3648 X Ant1368	1	NA	NA	NA	1	1	1
Ant3648 X Ant9218	1	NA	NA	NA	0.67953	1	0.73427
Ant3648 X Ant7249	1	NA	NA	NA	0.17322	1	0.28134
Ant3648 X Ant2936	NA	NA	NA	NA	0.65595	1	0.67207
Ant3648 X Ant9181	NA	NA	NA	NA	NA	0.36818	0.36818
Ant3648 X Ant12220	NA	NA	1	NA	NA	0.77624	0.92271
Ant3648 X Ant7680	1	NA	NA	NA	NA	1	1
Ant3648 X Ant859	1	NA	NA	NA	0.44444	0.37467	0.28662
Ant3648 X Ant4155	1	NA	NA	NA	NA	1	1
Ant3648 X Ant3993	1	NA	NA	NA	0.51941	0.58699	0.4622
Ant5035 X Ant11315	NA	NA	NA	NA	NA	NA	NA
Ant5035 X Ant10878	1	NA	NA	NA	1	1	1
Ant5035 X Ant1343	NA	1	NA	NA	1	0.13312	0.25492
Ant5035 X Ant2794	1	1	NA	NA	1	0.24707	0.67868
Ant5035 X Ant3653	1	NA	NA	NA	1	0.43246	0.44481
Ant5035 X Ant575	1	0.11739	NA	NA	1	0.4612	0.06072
Ant5035 X Ant20	0.26887	NA	NA	NA	NA	NA	0.26887
Ant5035 X Ant11400	1	NA	NA	NA	NA	1	1
Ant5035 X Ant8424	1	0.13445	NA	NA	1	0.68478	0.14931
Ant5035 X Ant8498	NA	0.00921	NA	NA	NA	NA	0.00921
Ant5035 X Ant2341	1	NA	NA	NA	1	1	1
Ant5035 X Ant1368	1	NA	NA	NA	1	1	1
Ant5035 X Ant9218	1	NA	NA	NA	1	1	1
Ant5035 X Ant7249	1	NA	NA	NA	1	1	1
Ant5035 X Ant2936	NA	NA	NA	NA	1	0.06896	0.06896
Ant5035 X Ant9181	NA	NA	NA	NA	NA	0.51836	0.51836
Ant5035 X Ant12220	NA	NA	NA	NA	NA	0.88237	0.88237
Ant5035 X Ant7680	1	NA	NA	NA	NA	1	1
Ant5035 X Ant859	0.26567	0.21745	NA	NA	1	0.46561	0.02739
Ant5035 X Ant4155	0.15731	NA	NA	NA	NA	1	0.71908
Ant5035 X Ant3993	1	1	NA	NA	1	0.11652	0.11652
Ant11315 X Ant10878	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant1343	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant2794	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant3653	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant575	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant20	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant11400	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant8424	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant8498	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant2341	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant1368	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant9218	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant7249	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant2936	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant9181	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant12220	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant7680	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant859	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant4155	NA	NA	NA	NA	NA	NA	NA
Ant11315 X Ant3993	NA	NA	NA	NA	NA	NA	NA

	Simopelta	Paraponera	Ectatomma	Solenopsis	Lasius	Dorylus	All
Ant10878 X Ant1343	NA	NA	NA	NA	1	1	1
Ant10878 X Ant2794	1	NA	NA	NA	1	0.90389	0.90616
Ant10878 X Ant3653	1	NA	1	NA	1	0.8324	0.89801
Ant10878 X Ant575	1	NA	NA	NA	1	0.55942	0.79127
Ant10878 X Ant20	0.63306	NA	NA	NA	NA	NA	0.63306
Ant10878 X Ant11400	1	NA	NA	NA	NA	0.37153	0.41144
Ant10878 X Ant8424	1	NA	1	NA	1	1	1
Ant10878 X Ant8498	NA	NA	NA	NA	NA	NA	NA
Ant10878 X Ant2341	1	NA	NA	NA	0.29022	0.63255	0.39614
Ant10878 X Ant1368	1	NA	NA	NA	1	0.2946	0.40752
Ant10878 X Ant9218	1	NA	NA	NA	1	0.94438	0.95842
Ant10878 X Ant7249	0.20042	NA	NA	NA	0.08545	0.45879	0.05782
Ant10878 X Ant2936	NA	NA	NA	NA	1	1	1
Ant10878 X Ant9181	NA	NA	NA	NA	NA	1	1
Ant10878 X Ant12220	NA	NA	0.10097	NA	NA	1	0.27141
Ant10878 X Ant7680	1	NA	NA	NA	NA	0.82947	0.82947
Ant10878 X Ant859	0.63228	NA	NA	NA	1	0.55824	0.39435
Ant10878 X Ant4155	0.433	NA	NA	NA	NA	0.21398	0.09526
Ant10878 X Ant3993	0.36534	NA	NA	NA	1	1	0.6099
Ant1343 X Ant2794	NA	0.76612	NA	NA	0.26585	1	0.46739
Ant1343 X Ant3653	NA	NA	NA	0.35293	1	1	0.67029
Ant1343 X Ant575	NA	0.67796	NA	0.09526	0.60338	0.91271	0.21621
Ant1343 X Ant20	NA	NA	NA	NA	NA	NA	NA
Ant1343 X Ant11400	NA	NA	NA	NA	NA	1	1
Ant1343 X Ant8424	NA	1	NA	0.45106	0.75202	0.50809	0.57757
Ant1343 X Ant8498	NA	0.72515	NA	NA	NA	NA	0.72515
Ant1343 X Ant2341	NA	NA	NA	0.5513	0.72406	0.44795	0.4561
Ant1343 X Ant1368	NA	NA	NA	NA	0.17941	1	0.20211
Ant1343 X Ant9218	NA	NA	NA	NA	0.33527	1	0.38955
Ant1343 X Ant7249	NA	NA	NA	NA	0.77304	1	0.83952
Ant1343 X Ant2936	NA	NA	NA	1	0.06636	1	0.52428
Ant1343 X Ant9181	NA	NA	NA	NA	NA	1	1
Ant1343 X Ant12220	NA	NA	NA	NA	NA	0.77699	0.77699
Ant1343 X Ant7680	NA	NA	NA	NA	NA	1	1
Ant1343 X Ant859	NA	1	NA	NA	1	0.37787	0.71259
Ant1343 X Ant4155	NA	NA	NA	1	NA	0.59179	0.80749
Ant1343 X Ant3993	NA	1	NA	0.6548	0.01781	0.12255	0.0128
Ant2794 X Ant3653	1	NA	NA	NA	1	1	1
Ant2794 X Ant575	1	0.85531	NA	NA	1	0.46938	0.74212
Ant2794 X Ant20	1	NA	NA	NA	NA	NA	1
Ant2794 X Ant11400	1	NA	NA	NA	NA	0.76821	0.76821
Ant2794 X Ant8424	1	0.45151	NA	NA	1	0.79864	0.6468
Ant2794 X Ant8498	NA	1	NA	NA	NA	NA	1
Ant2794 X Ant2341	1	NA	NA	NA	1	0.18007	0.33225
Ant2794 X Ant1368	1	NA	NA	NA	1	1	1
Ant2794 X Ant9218	1	NA	NA	NA	1	1	1
Ant2794 X Ant7249	1	NA	NA	NA	1	0.5173	0.56147
Ant2794 X Ant2936	NA	NA	NA	NA	1	1	1
Ant2794 X Ant9181	NA	NA	NA	NA	NA	0.62772	0.62772
Ant2794 X Ant12220	NA	NA	NA	NA	NA	0.6013	0.6013
Ant2794 X Ant7680	1	NA	NA	NA	NA	1	1
Ant2794 X Ant859	1	0.53059	NA	NA	1	0.24399	0.15405
Ant2794 X Ant4155	1	NA	NA	NA	NA	0.83216	0.83216
Ant2794 X Ant3993	1	1	NA	NA	1	0.90755	0.92455
Ant3653 X Ant575	1	NA	NA	0.73838	0.29106	0.80051	0.48188
Ant3653 X Ant20	1	NA	NA	NA	NA	NA	1
Ant3653 X Ant11400	1	NA	NA	NA	NA	0.08511	0.09293
Ant3653 X Ant8424	1	NA	1	0.71434	0.73409	0.24894	0.63986
Ant3653 X Ant8498	NA	NA	NA	NA	NA	NA	NA
Ant3653 X Ant2341	1	NA	NA	0.49771	0.06449	0.5103	0.06908
Ant3653 X Ant1368	1	NA	NA	NA	1	0.40743	0.71975
Ant3653 X Ant9218	1	NA	NA	NA	0.56673	1	0.82769
Ant3653 X Ant7249	1	NA	NA	NA	0.24755	0.22443	0.09432
Ant3653 X Ant2936	NA	NA	NA	1	1	0.15498	0.74888
Ant3653 X Ant9181	NA	NA	NA	NA	NA	0.04481	0.04481
Ant3653 X Ant12220	NA	NA	1	NA	NA	1	1
Ant3653 X Ant7680	1	NA	NA	NA	NA	1	1
Ant3653 X Ant859	1	NA	NA	NA	0.06624	0.80223	0.61972
Ant3653 X Ant4155	1	NA	NA	1	NA	0.78505	0.84278
Ant3653 X Ant3993	1	NA	NA	1	1	0.82835	0.93925
Ant575 X Ant20	1	NA	NA	NA	NA	NA	1
Ant575 X Ant11400	1	NA	NA	NA	NA	0.46649	0.50287
Ant575 X Ant8424	1	0.66392	NA	0.52633	0.4131	0.22047	0.18056
Ant575 X Ant8498	NA	0.22041	NA	NA	NA	NA	0.22041
Ant575 X Ant2341	0.0869	NA	NA	0.39873	0.73527	0.7779	0.36108
Ant575 X Ant1368	1	NA	NA	NA	0.21718	0.73451	0.1869
Ant575 X Ant9218	1	NA	NA	NA	0.08216	1	0.21289
Ant575 X Ant7249	1	NA	NA	NA	0.6	0.82168	0.65619
Ant575 X Ant2936	NA	NA	NA	0.80921	0.78046	0.64626	0.60223
Ant575 X Ant9181	NA	NA	NA	NA	NA	0.0692	0.0692
Ant575 X Ant12220	NA	NA	NA	NA	NA	1	1
Ant575 X Ant7680	1	NA	NA	NA	NA	1	1
Ant575 X Ant859	1	1	NA	NA	0.49224	0.37856	0.6054
Ant575 X Ant4155	1	NA	NA	0.42467	NA	1	0.67612
Ant575 X Ant3993	1	1	NA	0.73653	0.56492	1	0.76658
Ant70X Ant11400	0.63457	NA	NA	NA	NA	NA	0.63457

Appendix C continued

Ant20 X Ant8424	0.75184	NA	NA	NA	NA	NA	0.75184
Ant20 X Ant8498	NA	NA	NA	NA	NA	NA	NA
Ant20 X Ant2341	0.75254	NA	NA	NA	NA	NA	0.75254
Ant20 X Ant1368	0.26295	NA	NA	NA	NA	NA	0.26295
Ant20 X Ant9218	1	NA	NA	NA	NA	NA	1
Ant20 X Ant7249	0.63644	NA	NA	NA	NA	NA	0.63644
Ant20 X Ant2936	NA	NA	NA	NA	NA	NA	NA
Ant20 X Ant9181	NA	NA	NA	NA	NA	NA	NA
Ant20 X Ant12220	NA	NA	NA	NA	NA	NA	NA
Ant20 X Ant7680	1	NA	NA	NA	NA	NA	1
Ant20 X Ant859	0.01114	NA	NA	NA	NA	NA	0.01114
Ant20 X Ant4155	0.15495	NA	NA	NA	NA	NA	0.15495
Ant20 X Ant3993	0.07068	NA	NA	NA	NA	NA	0.07068
Ant11400 X Ant8424	0.251	NA	NA	NA	NA	0.39336	0.24345
Ant11400 X Ant8498	NA	NA	NA	NA	NA	NA	NA
Ant11400 X Ant2341	0.25124	NA	NA	NA	NA	1	0.54206
Ant11400 X Ant1368	1	NA	NA	NA	NA	1	1
Ant11400 X Ant9218	1	NA	NA	NA	NA	0.82606	0.82606
Ant11400 X Ant7249	1	NA	NA	NA	NA	0.12832	0.21187
Ant11400 X Ant2936	NA	NA	NA	NA	NA	1	1
Ant11400 X Ant9181	NA	NA	NA	NA	NA	0.03711	0.03711
Ant11400 X Ant12220	NA	NA	NA	NA	NA	1	1
Ant11400 X Ant7680	1	NA	NA	NA	NA	1	1
Ant11400 X Ant859	0.63089	NA	NA	NA	NA	1	0.72926
Ant11400 X Ant4155	0.4298	NA	NA	NA	NA	1	0.49496
Ant11400 X Ant3993	0.36404	NA	NA	NA	NA	1	0.65223
Ant8424 X Ant8498	NA	0.28632	NA	NA	NA	NA	0.28632
Ant8424 X Ant2341	0.32518	NA	NA	0.63146	0.88439	0.65347	0.50275
Ant8424 X Ant1368	1	NA	NA	0.10782	1	1	0.46135
Ant8424 X Ant9218	1	NA	NA	0.81153	1	1	0.95347
Ant8424 X Ant7249	1	NA	NA	0.86643	0.9625	0.94457	0.94457
Ant8424 X Ant2936	NA	NA	NA	0.45163	0.09909	1	0.16567
Ant8424 X Ant9181	NA	NA	NA	NA	0.03361	0.03361	0.03361
Ant8424 X Ant12220	NA	NA	1	NA	NA	0.8327	0.91196
Ant8424 X Ant7680	1	NA	NA	NA	NA	1	1
Ant8424 X Ant859	0.21969	0.90966	NA	NA	1	1	0.68729
Ant8424 X Ant4155	0.52243	NA	NA	0.12974	NA	0.69381	0.1401
Ant8424 X Ant3993	0.46078	1	NA	0.54167	0.08916	1	0.21896
Ant8498 X Ant2341	NA	NA	NA	NA	NA	NA	NA
Ant8498 X Ant1368	NA	NA	NA	NA	NA	NA	NA
Ant8498 X Ant9218	NA	NA	NA	NA	NA	NA	NA
Ant8498 X Ant7249	NA	NA	NA	NA	NA	NA	NA
Ant8498 X Ant2936	NA	NA	NA	NA	NA	NA	NA
Ant8498 X Ant9181	NA	NA	NA	NA	NA	NA	NA
Ant8498 X Ant12220	NA	NA	NA	NA	NA	NA	NA
Ant8498 X Ant7680	NA	NA	NA	NA	NA	NA	NA
Ant8498 X Ant859	NA	0.49245	NA	NA	NA	NA	0.49245
Ant8498 X Ant4155	NA	NA	NA	NA	NA	NA	NA
Ant8498 X Ant3993	NA	1	NA	NA	NA	NA	1
Ant2341 X Ant1368	1	NA	NA	NA	0.85918	0.62252	0.71389
Ant2341 X Ant9218	1	NA	NA	NA	0.47464	0.82092	0.55954
Ant2341 X Ant7249	1	NA	NA	NA	0.26132	0.02228	0.01667
Ant2341 X Ant2936	NA	NA	NA	0.55079	1	1	0.76256
Ant2341 X Ant9181	NA	NA	NA	NA	NA	0.97135	0.97135
Ant2341 X Ant12220	NA	NA	NA	NA	NA	0.30217	0.30217
Ant2341 X Ant7680	1	NA	NA	NA	NA	0.57144	0.57144
Ant2341 X Ant859	0.22183	NA	NA	NA	0.28877	0.55722	0.07428
Ant2341 X Ant4155	0.52343	NA	NA	1	NA	0.26232	0.20833
Ant2341 X Ant3993	0.45679	NA	NA	0.37594	1	1	0.8452
Ant1368 X Ant9218	1	NA	NA	NA	0.14333	1	0.21896
Ant1368 X Ant7249	1	NA	NA	NA	0.63297	0.42388	0.38197
Ant1368 X Ant2936	NA	NA	NA	NA	1	1	1
Ant1368 X Ant9181	NA	NA	NA	NA	NA	1	1
Ant1368 X Ant12220	NA	NA	NA	NA	NA	0.9321	0.9321
Ant1368 X Ant7680	1	NA	NA	NA	NA	1	1
Ant1368 X Ant859	0.26984	NA	NA	NA	1	0.73336	0.43433
Ant1368 X Ant4155	1	NA	NA	NA	NA	0.7907	0.8433
Ant1368 X Ant3993	1	NA	NA	NA	0.05248	1	0.2817
Ant9218 X Ant7249	1	NA	NA	NA	0.70966	0.62198	0.54354
Ant9218 X Ant2936	NA	NA	NA	NA	1	1	1
Ant9218 X Ant9181	NA	NA	NA	NA	NA	1	1
Ant9218 X Ant12220	NA	NA	NA	NA	NA	0.16848	0.16848
Ant9218 X Ant7680	1	NA	NA	NA	NA	1	1
Ant9218 X Ant859	1	NA	NA	NA	1	1	1
Ant9218 X Ant4155	1	NA	NA	NA	NA	0.18315	0.18315
Ant9218 X Ant3993	1	NA	NA	NA	0.57929	0.07403	0.10601
Ant7249 X Ant2936	NA	NA	NA	NA	1	1	1
Ant7249 X Ant9181	NA	NA	NA	NA	NA	0.85571	0.85571
Ant7249 X Ant12220	NA	NA	NA	NA	NA	0.3782	0.3782
Ant7249 X Ant7680	1	NA	NA	NA	NA	0.39985	0.39985
Ant7249 X Ant859	0.6362	NA	NA	NA	1	0.68653	0.56954
Ant7249 X Ant4155	0.3971	NA	NA	NA	NA	0.4538	0.28421
Ant7249 X Ant3993	0.35833	NA	NA	NA	0.57524	0.80975	0.3516
Ant2936 X Ant9181	NA	NA	NA	NA	NA	0.19846	0.19846
Ant2936 X Ant12220	NA	NA	NA	NA	NA	1	1
Ant2936 X Ant7680	NA	NA	NA	NA	NA	1	1

Ant2936 X Ant859	NA	NA	NA	NA	1	0.64502	0.65972
Ant2936 X Ant4155	NA	NA	NA	0.04354	NA	1	0.07198
Ant2936 X Ant3993	NA	NA	NA	0.65613	0.32835	0.35643	0.18122
Ant9181 X Ant12220	NA	NA	NA	NA	NA	0.78783	0.78783
Ant9181 X Ant7680	NA	NA	NA	NA	NA	1	1
Ant9181 X Ant859	NA	NA	NA	NA	NA	0.32926	0.32926
Ant9181 X Ant4155	NA	NA	NA	NA	NA	0.85516	0.85516
Ant9181 X Ant3993	NA	NA	NA	NA	NA	1	1
Ant12220 X Ant7680	NA	NA	NA	NA	NA	0.88279	0.88279
Ant12220 X Ant859	NA	NA	NA	NA	NA	1	1
Ant12220 X Ant4155	NA	NA	NA	NA	NA	0.49517	0.49517
Ant12220 X Ant3993	NA	NA	NA	NA	NA	1	1
Ant7680 X Ant859	1	NA	NA	NA	NA	1	1
Ant7680 X Ant4155	1	NA	NA	NA	NA	0.66818	0.66818
Ant7680 X Ant3993	1	NA	NA	NA	NA	0.82947	0.82947
Ant859 X Ant4155	0.01359	NA	NA	NA	NA	1	0.08922
Ant859 X Ant3993	0.46963	1	NA	NA	1	1	0.75082
Ant4155 X Ant3993	0.68702	NA	NA	0.66857	NA	0.53107	0.60181

Appendix D Number of extra workers genotyped in each colony and at each locus to reconstruct queen genotypes. Colonies that are not listed did not have additional workers genotyped.

Colony	Ant4155	Ant2341	Ant5035	Ant7680	Ant7249	DmoG	DmoB	Ant8424	Ant1343	DmoD	Ant10878	Ant9218
IB2							40	36				
IB10		39										
IB12	7						7					
IB18	24											
IB28	16	8	32						20			
IB29	16			15								
IB46					8					8		
IB53		32							24		15	
IB57		24								24	16	
IB62			22									
IB67	16	8	8	8	8	6	24	16	16	8	8	8
IB69	16	15					16					
IB72		8										
IB78	16					8						
IB80	32	16					6				9	
IB85	8	7							8			
IB88					8					40		
IB97		38										
IB108					7		8					8
IB109					8				6			
IB123							31					
IB132	6				32				33			
IB134									7	24		
IB137	8											
IB138			8					8				7

Appendix E List of additional putative hybrid workers for which I could not reconstruct parental genotypes at three different proportion-of-ancestry cutoffs. In the STRUCTURE analysis, these workers had a proportion of ancestry attributable to the same species as most of the other workers in their colony lower than the given cutoff. However, I was unable to reconstruct parental genotypes for their colonies, so I was not able to confirm their hybrid status based on the clustering of the parental genotypes. Putative hybrid workers in the 0.65 proportion-of-ancestry cutoff are implicitly included in the 0.75 cutoff, and workers in both lower cutoff values are implicitly included in the 0.90 cutoff.

0.65 cutoff		0.75 cutoff		0.9 cutoff	
Hybrid	Colony Species	Hybrid	Colony Species	Hybrid	Colony Species
IB83.9	<i>D. wilverthi</i>	IB83.5	<i>D. wilverthi</i>	IB75i6	<i>D. wilverthi</i>
IB83.22	<i>D. wilverthi</i>	IB83.21	<i>D. wilverthi</i>	IB83.4	<i>D. wilverthi</i>
IB83.23	<i>D. wilverthi</i>	IB83.30	<i>D. wilverthi</i>	IB83.5	<i>D. wilverthi</i>
IB83.24	<i>D. wilverthi</i>	IB124.23	<i>D. molestus</i>	IB83i43	<i>D. wilverthi</i>
IB83.37	<i>D. wilverthi</i>	IB124.39	<i>D. molestus</i>	IB83i50	<i>D. wilverthi</i>
IB124.15	<i>D. molestus</i>			IB83i53	<i>D. wilverthi</i>
IB124.22	<i>D. molestus</i>			IB83i57	<i>D. wilverthi</i>
IB124.28	<i>D. molestus</i>			IB83i64	<i>D. wilverthi</i>
IB124.32	<i>D. molestus</i>			IB84i2	<i>D. wilverthi</i>
IB124.36	<i>D. molestus</i>			IB124i6	<i>D. molestus</i>
IB124.44	<i>D. molestus</i>			IB124i9	<i>D. molestus</i>
IB124.45	<i>D. molestus</i>			IB124i14	<i>D. molestus</i>
IB124.47	<i>D. molestus</i>			IB124i17	<i>D. molestus</i>
IB124.50	<i>D. molestus</i>			IB124i24	<i>D. molestus</i>
IB124.57	<i>D. molestus</i>			IB124i31	<i>D. molestus</i>
IB124.70	<i>D. molestus</i>			IB124i34	<i>D. molestus</i>
				IB124i35	<i>D. molestus</i>

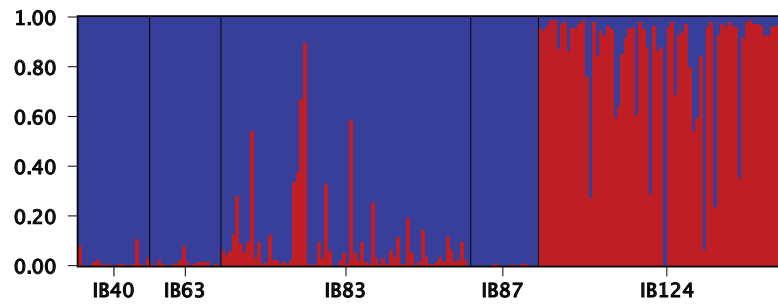
Appendix F GPS coordinates for driver ant colonies collected in Kakamega Forest.

Colony	Species	Latitude	Longitude
IB2	<i>D. molestus</i>	00°14'57"N	34°52'22"E
IB10	<i>D. wilverthi</i>	00°12'51"N	34°54'40"E
IB12	<i>D. molestus</i>	00°12'52"N	34°55'44"E
IB18	<i>D. molestus</i>	00°14'38"N	34°52'00"E
IB27	<i>D. molestus</i>	00°12'50"N	34°51'28"E
IB28	<i>D. molestus</i>	00°12'57"N	34°51'33"E
IB29	<i>D. molestus</i>	00°04'26"N	35°01'51"E
IB40	<i>D. wilverthi</i>	00°14'16"N	34°51'53"E
IB46	<i>D. wilverthi</i>	00°16'33"N	34°50'00"E
IB49	<i>D. wilverthi</i>	00°16'23"N	34°55'12"E
IB53	<i>D. wilverthi</i>	00°16'47"N	34°56'04"E
IB57	<i>D. wilverthi</i>	00°13'50"N	34°51'42"E
IB62	<i>D. molestus</i>	00°14'13"N	34°52'07"E
IB63	<i>D. wilverthi</i>	00°14'20"N	34°52'09"E
IB64	<i>D. wilverthi</i>	00°14'11"N	34°51'04"E
IB67	<i>D. molestus</i>	00°13'36"N	34°52'39"E
IB69	<i>D. wilverthi</i>	00°14'00"N	34°52'07"E
IB71	<i>D. molestus</i>	00°11'54"N	34°52'54"E
IB72	<i>D. wilverthi</i>	00°13'23"N	34°53'29"E
IB75	<i>D. wilverthi</i>	00°14'13"N	34°52'13"E
IB78	<i>D. wilverthi</i>	00°13'33"N	34°53'57"E
IB79	<i>D. wilverthi</i>	00°13'20"N	34°54'09"E
IB80	<i>D. molestus</i>	00°13'14"N	34°53'46"E
IB83	<i>D. wilverthi</i>	00°12'30"N±200m	34°52'46"E±200m
IB84	<i>D. wilverthi</i>	00°12'38"N	34°52'46"E
IB85	<i>D. wilverthi</i>	00°12'45"N	34°52'48"E
IB87	<i>D. wilverthi</i>	00°14'51"N	34°52'05"E
IB88	<i>D. wilverthi</i>	00°14'58"N	34°52'05"E
IB97	<i>D. wilverthi</i>	00°14'05"N	34°51'52"E
IB99	<i>D. molestus</i>	00°14'07"N	34°51'51"E
IB102	<i>D. wilverthi</i>	00°13'33"N	34°51'20"E
IB103	<i>D. wilverthi</i>	00°14'04"N	34°51'53"E
IB108	<i>D. molestus</i>	00°12'44"N	34°55'36"E
IB109	<i>D. molestus</i>	00°12'24"N	34°55'26"E
IB110	<i>D. wilverthi</i>	00°12'22"N	34°55'26"E
IB111	<i>D. wilverthi</i>	00°16'23"N	34°51'18"E
IB115	<i>D. molestus</i>	00°16'24"N	34°50'52"E
IB116	<i>D. molestus</i>	00°16'21"N	34°50'53"E
IB117	unknown	00°17'09"N	34°50'58"E
IB118	<i>D. wilverthi</i>	00°17'35"N	34°48'43"E
IB119	<i>D. molestus</i>	00°17'07"N	34°47'58"E
IB120	<i>D. molestus</i>	00°17'04"N	34°47'44"E
IB123	<i>D. molestus</i>	00°16'51"N	34°46'53"E
IB124	<i>D. molestus</i>	00°17'11"N	34°47'41"E
IB131	<i>D. wilverthi</i>	00°17'36"N	34°48'41"E
IB132	<i>D. molestus</i>	00°17'26"N	34°48'42"E
IB134	<i>D. wilverthi</i>	00°13'57"N	34°51'46"E
IB136	<i>D. wilverthi</i>	00°13'57"N	34°51'46"E
IB137	<i>D. molestus</i>	00°12'18"N	34°54'27"E
IB138	<i>D. molestus</i>	00°12'22"N	34°55'20"E

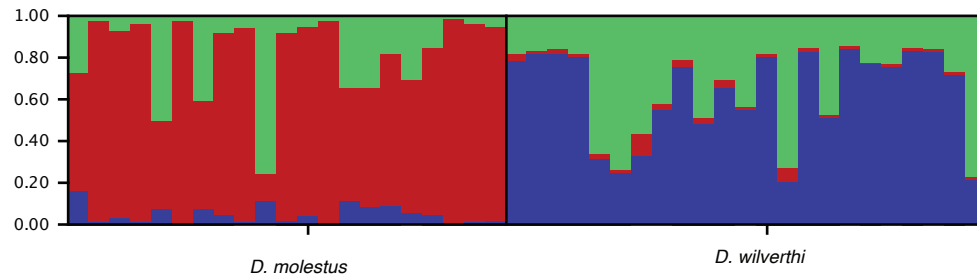
Appendix G Number of hybrids found and hybridization rates at three different cutoffs using the whole dataset and different subsets of the data.

	0.65 cutoff	0.75 cutoff	0.90 cutoff
Whole dataset as single run	8(0.0090)	10(0.011)	19(0.0216)
Data ran as two halves	8(0.0090)	9(0.010)	24(0.0273)
Data ran as four quarters	5(0.0057)	7(0.0080)	30(0.0341)
Six loci 1	50(0.0568)	77(0.0875)	163(0.185)
Six loci 2	32(0.0364)	44(0.0500)	89(0.101)
Five loci	50(0.0568)	69(0.0784)	170(0.193)

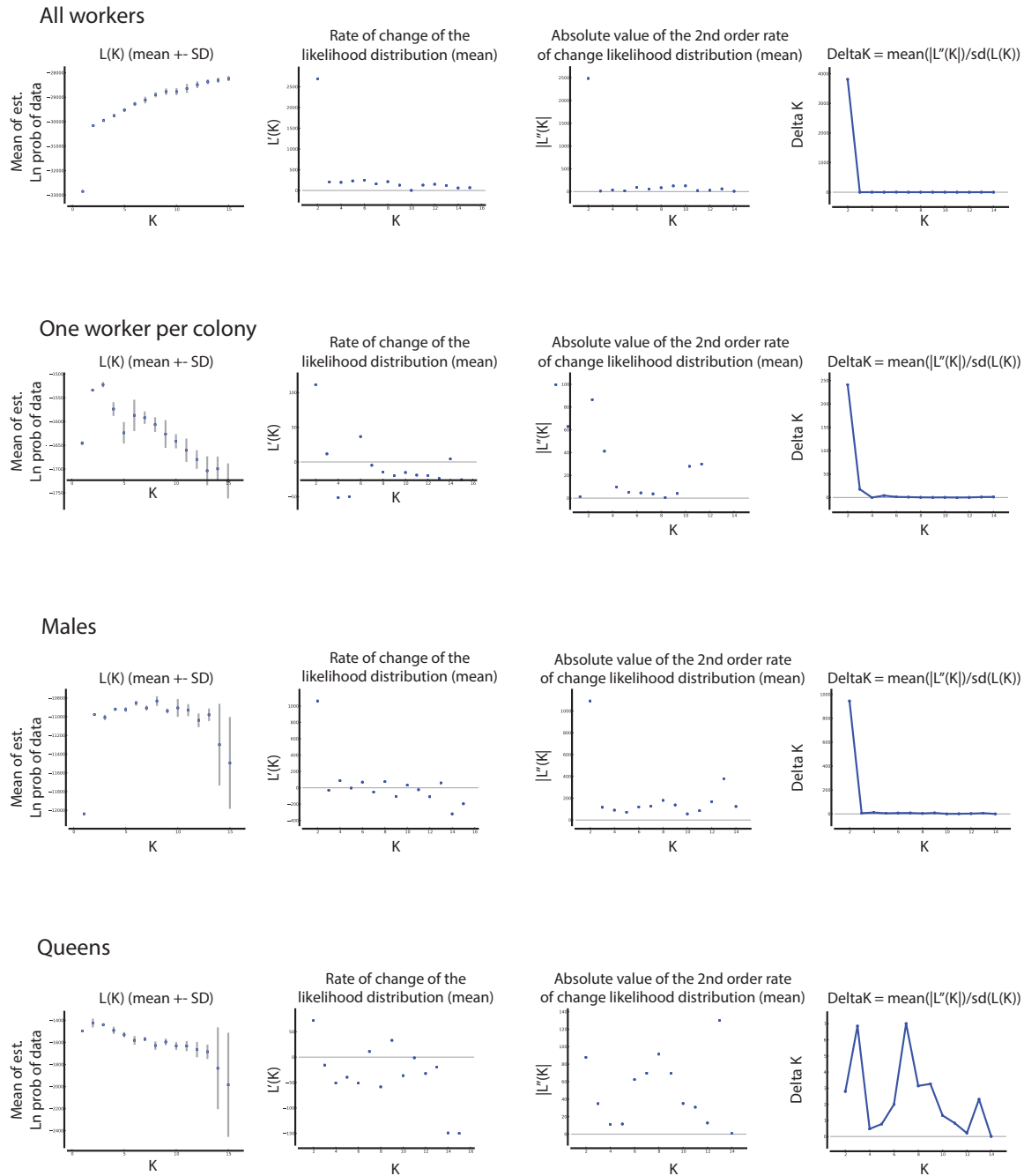
Appendix H STRUCTURE plot of worker genotypes from colonies without reconstructed queen genotypes. Blue is proportion of ancestry attributable to *Dorylus wilverthi*, and red is proportion of ancestry attributable to *D. molestus*.



Appendix I STRUCTURE plot of reconstructed queen genotypes at $K = 3$. Blue is proportion of ancestry attributable to *Dorylus wilverthi*, and red is proportion of ancestry attributable to *D. molestus*. Green is proportion of ancestry from an assumed third species or group.



Appendix J Output of ΔK measurements for values of K from 1 to 15 for all worker genotypes, one worker genotype per colony, all reconstructed male genotypes, and all reconstructed queen genotypes.



References

- Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., ... & Butlin, R.K. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, 26(2), 229-246.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., & Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403-410.
- Anderson, E. (1949). *Introgressive Hybridization*. New York, New York: John Wiley and Sons.
- Anderson, E., & Hubricht, L. (1938). Hybridization in *Tradescantia*. III. The evidence for introgressive hybridization. *American Journal of Botany*, 25(6), 396-402.
- Anderson, K.E., Gadau, J., Mott, J.B.M., Johnson, R.A., Altamirano, A., Strehl, C., & Fewell, J.H. (2006). Distribution and evolution of genetic caste determination in *Pogonomyrmex* seed-harvester ants. *Ecology*, 87(9), 2171-2184.
- Anderson, K.E., Linksvayer, T.A., & Smith, C.A. (2008a). The causes and consequences of hybridization in ants (Hymenoptera: Formicidae). *Myrmecological News*, 11(11), 119-132.
- Anderson, K.E., Novak, S.J., & Smith, J.F. (2008b). Populations composed entirely of hybrid colonies: bidirectional hybridization and polyandry in harvester ants. *Biological Journal of the Linnean Society*, 95(2), 320-336.
- Arnold, M.L. (1997). *Natural Hybridization and Evolution*. Oxford, United Kingdom: Oxford University Press.
- Arnold, M.L. (2006). *Evolution Through Genetic Exchange*. Oxford, United Kingdom: Oxford University Press.
- Arnold, M.L., & Martin, N.H. (2009). Adaptation by introgression. *Journal of Biology*, 8(9), 82.
- Ascunce, M.S., Bouwma, A.M., & Shoemaker, D. (2009). Characterization of 24 microsatellite markers in 11 species of fire ants in the genus *Solenopsis* (Hymenoptera: Formicidae). *Molecular Ecology Resources*, 9(6), 1476-1479.
- Azuma, N., Takahashi, J., Higashi, S., & Sasaki, M. (2004). Microsatellite loci for the weaver ant *Oecophylla smaragdina*. *Molecular Ecology Notes*, 4(4), 608-610.
- Baack, E.J., & Reiseberg, L.H. (2007). A genomic view of introgression and hybrid speciation. *Current Opinion in Genetics & Development*, 17(6), 1-6.
- Baer, B., & Schmid-Hempel, P. (1999). Experimental variation in polyandry affects parasite loads and fitness in a bumble-bee. *Nature*, 397(6715), 151-154.

- Bagherian Yazdi, A., Münch, W., & Seifert, B. (2012). A first demonstration of interspecific hybridization in *Myrmica* ants by geometric morphometrics (Hymenoptera: Formicidae). *Myrmecological News*, 17, 121-131.
- Bateson, W. (1909). Heredity and variation in modern lights. In *Darwin and Modern Science. Essays in Commemoration of the Centenary of the Birth of Charles Darwin and of the Fiftieth Anniversary of the Publication of the "Origin of Species."* (ed. Seward, A.C.). Cambridge, United Kingdom: University Press.
- Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research*, 27(2), 573–580.
- Bernasconi, C., Cherix, D., Seifert, B., & Pamilo, P. (2011). Molecular taxonomy of the *Formica rufa* group (red wood ants) (Hymenoptera: Formicidae): a new cryptic species in the Swiss Alps? *Myrmecological News*, 14, 37-47.
- Beresford, J., Elias, M., Pluckrose, L., Sundström, L., Butlin, R.K., Pamilo, P., & Kulmuni, J. (2017). Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown. *Molecular Ecology*, 26(15), 4013-4026.
- Beye, M., Hasselmann, M., Fondrk, M.K., Page, R.E., & Omholt, S.W. (2003). The gene *csd* is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell*, 114(4), 419– 429.
- Bonasio, R., Zhang, G., Ye, C., Mutti, N.S., Fang, X., Qin, N., ... & Zhang, P. (2010). Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science*, 329(5995), 1068–1071.
- Boomsma, J.J. (2007). Kin selection versus sexual selection: why the ends do not meet. *Current Biology*, 17(16), R673-R683.
- Boomsma, J.J., Brady, S.G., Dunn, R.R., Gadau, J., Heinze, J., Keller, L., ... & Sundström, L. (2017). The global ant genomics alliance (GAGA). *Myrmecological News*, 25, 61-66.
- Boomsma, J.J., Kronauer, D.J.C., & Pedersen, J.S. (2009). The evolution of social insect mating systems. In *Organization of Insect Societies—From Genome to Sociocomplexity* (eds. Gadau, J., & Fewell, J.). Cambridge, Massachusetts: Harvard University Press.
- Boomsma, J.J., & Ratnieks, F.L. (1996). Paternity in eusocial Hymenoptera. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1342), 947-975.
- Bourke, A.F.G. (2011). *Principles of Social Evolution*. New York, New York: Oxford University Press.

- Bourke, A.F.G., & Franks, N.R. (1995). *Social Evolution in Ants*. Princeton, New Jersey: Princeton University Press.
- Brady, S.G., Schultz, T.R., Fisher, B.L., & Ward, P.S. (2006). Evaluating alternative hypotheses for the early evolution and diversification of ants. *Proceedings of the National Academy of Sciences of the United States of America*, 103(48), 18172–18177.
- Branstetter, M., Childers, A.K., Cox-Foster, D., Hopper, K.R., Kapheim, K.M., Toth, A.L., & Worley, K.C. (2017). Genomes of the Hymenoptera. *Current Opinion in Insect Science*, 25, 65-75.
- Breed, M. (1989). Introduction: theory and empirical results in sociogenetics. In *The Genetics of Social Evolution* (eds. Breed, M.D., & Page, R.E.). Boulder, Colorado: Westview Press.
- Brown, J.W., Rest, J.S., García-Moreno, J., Sorenson, M.D., & Mindell, D.P. (2008). Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. *BMC Biology*, 6(1), 6.
- Burke, J.M., & Arnold, M.L. (2001). Genetics and the fitness of hybrids. *Annual Review of Genetics*, 35(1), 31-52.
- Buschiazzo, E., & Gemmell, N.J. (2006). The rise, fall and renaissance of microsatellites in eukaryotic genomes. *BioEssays*, 28(10), 1040–1050.
- Buschiazzo, E., & Gemmell, N.J. (2009). Evolution and phylogenetic significance of platypus microsatellites conserved in mammalian and other vertebrate genomes. *Australian Journal of Zoology*, 57(4), 175–184.
- Buschiazzo, E., & Gemmell, N.J. (2010). Conservation of human microsatellites across 450 million years of evolution. *Genome Biology and Evolution*, 2, 153–165.
- Buschinger, A., & Winter, U. (1975). Der Polymorphismus der sklavenhaltenden Ameise *Harpagoxenus sublaevis*. *Insectes Sociaux*, 22, 333–362.
- Calderone, N.W., & Page, R.E. (1988). Genotypic variability in age polyethism and task specialization in the honey bee, *Apis mellifera* (Hymenoptera: Apidae). *Behavioral Ecology and Sociobiology*, 22(1), 17-25.
- Calderone, N.W., & Page, R.E. (1991). Evolutionary genetics of division of labor in colonies of the honey bee (*Apis mellifera*). *The American Naturalist*, 138(1), 69-92.
- Calderone, N.W., Robinson, G.E., & Page, R.E. (1989). Genetic structure and division of labor in honeybee societies. *Experientia*, 45(8), 765-767.
- Chen, Z.J. (2010). Molecular mechanisms of polyploidy and hybrid vigor. *Trends in Plant Science*, 15(2), 57-71.

- Cole, B.J., & Wiernasz, D.C. (1999). The selective advantage of low relatedness. *Science*, 285(5429), 891-893.
- Conlon, I., & Raff, M. (1999). Size control in animal development. *Cell*, 96(2), 235-244.
- Cordonnier, M., Gayet, T., Escarguel, G., & Kaufmann, B. (2019). From hybridization to introgression between two closely related sympatric ant species. *Journal of Zoological Systematics and Evolutionary Resources*.
- Coyne, J.A., & Orr, H.A. (2004). *Speciation*. Sunderland, Massachusetts: Sinauer Associates.
- Cremer, S., Armitage, S.A.O., & Schmid-Hempel, P. (2007). Social immunity. *Current Biology*, 17(16), R693-R702.
- Crespi, B.J., & Yanega, D. (1995). The definition of eusociality. *Behavioral Ecology*, 6(1), 109-115.
- Crozier, R. H., & Fjerdingstad, E. J. (2001). Polyandry in social Hymenoptera—disunity in diversity? *Annales Zoologici Fennici*, 38, 267-285.
- Crozier, R.H., & Page, R.E. (1985). On being the right size: male contributions and multiple mating in the social Hymenoptera. *Behavioral Ecology and Sociobiology*, 18(2), 105–115.
- Crozier R.H., & Pamilo, P. (1996). *Evolution of Social Insect Colonies*. Oxford, United Kingdom: Oxford University Press.
- Czechowski, W. (1996). Colonies of hybrids and mixed colonies; interspecific nest takeover in wood ants (Hymenoptera, Formicidae). *Memorabilia Zoologica*, 50.
- Dalecky, A., Debout, G., Mondor, G., Rasplus, J.-Y., & Estoup, A. (2002). PCR primers for polymorphic microsatellite loci in the polymorphic facultatively polygynous plant-ant *Petalomyrmex phylax* (Formicidae). *Molecular Ecology Notes*, 2(4), 404–407.
- Darras, H., Kuhn, A., & Aron, S. (2014). Genetic determination of female castes in a hybridogenetic desert ant. *Journal of Evolutionary Biology*, 27(10), 2265-2271.
- Dawkins, R. (1976). *The Selfish Gene*. Oxford, United Kingdom; Oxford University Press.
- Dawson, D.A., Ball, A.D., Spurgin, L.G., Martín-Gálvez, D., Stewart, I.R.K., Horsburgh, G.J., ... & Ekblom, R. (2013). High-utility conserved avian microsatellite markers enable parentage and population studies across a wide range of species. *BMC Genomics*, 14(1), 176.

- Dawson, D.A., Horsburgh, G.J., Kupper, C., Stewart, I.R.K., Ball, A.D., Durrant, K.L., ... & Krupa, A.P. (2010). New methods to identify conserved microsatellite loci and develop primer sets of high cross-species utility – as demonstrated for birds. *Molecular Ecology Resources*, 10(3), 475–494.
- Debout, G.D.G., Pereyra, R., Emerson, B.C., & Yu, D.W. (2006). Characterization of polymorphic microsatellites in the castration plant-ant *Allomerus octoarticulatus* cf. *demerarae* (Formicidae: Myrmecinae). *Molecular Ecology Notes*, 6(1), 182–184.
- Debout, G.D.G., Ventolen-Debout, M., Emerson, B.C., & Wu, D.W. (2007). PCR primers for polymorphic microsatellite loci in the plant ant *Azteca ulei cordiae* (Formicidae: Dolichoderinae). *Molecular Ecology Notes*, 7(4), 607–609.
- Dhaygude, K., Nair, A., Johansson, H., Wurm, Y., & Sundström L. (2019). The first draft genomes of the ant *Formica exsecta*, and its *Wolbachia* endosymbiont reveal extensive gene transfer from endosymbiont to host. *BMC Genomics*, 20(1), 301.
- Dobzhansky, T.H. (1936). Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics*, 21(2), 113.
- Dobzhansky, T.H. (1937). *Genetics and the Origin of Species*. New York, New York: Columbia University Press.
- Douwes, P., & Stille, B. (1991). Hybridization and variation in the *Leptothorax tuberosus* group (Hymenoptera, Formicidae). *Journal of Zoological Systematics and Evolutionary Research*, 29(3), 165–175.
- Dreller, C., Fondrk, M.K., & Page, R.E. (1995). Genetic variability affects the behavior of foragers in a feral honeybee colony. *Naturwissenschaften*, 82(5), 243–245.
- Durand, E.Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution*, 28(8), 2239–2252.
- Eaton, D. A., Hipp, A. L., González-Rodríguez, A., & Cavender-Bares, J. (2015). Historical introgression among the American live oaks and the comparative nature of tests for introgression. *Evolution*, 69(10), 2587–2601.
- Edgar, R. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797.
- Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, 5(6), 435–445.
- El-Shehaby, M., Salama, M.S., Brunner, E., & Heinze, J. (2011). Cuticular hydrocarbons in two parapatric species of ants and their hybrid. *Integrative Zoology*, 6(3), 259–265.

- Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., & Mitchell, S.E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6(5), e19379.
- Emery, C. (1899). Formiche dell'ultima spedizione Bottego. *Annali del Museo Civico di Storia Naturale Giacomo Doria (Genova)*, 2, 499-501.
- Erayman, M., Ilhan, E., Guzel, Y., & Eren, A.H. (2014). Transferability of SSR markers from distantly related legumes to *Glycyrrhiza* species. *Turkish Journal of Agriculture and Forestry*, 38(1), 32–38.
- Evanno, G., Regnaut, S., & Gaudet J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14(8), 2611–2620.
- Eyer, P.A., Seltzer, R., Riener-Brodetzki, T., & Hefetz, A. (2017). An integrative approach to untangling species delimitation in the *Cataglyphis bicolor* desert ant complex in Israel. *Molecular Phylogenetics and Evolution*, 115, 128-139.
- Ezenwa, V.O., Peters, J.M., Zhu, Y., Arévalo, E., Hastings, M.D., Seppä, P., ... & Strassmann, J.E. (1998). Ancient conservation of trinucleotide microsatellite loci in polistine wasps. *Molecular Phylogenetics and Evolution*, 10(2), 168–177.
- Feldhaar, H., Fiala, B., Gadau, J., Maryati, M., & Maschwitz, U. (2003). Molecular phylogeny of *Crematogaster* subgenus *Decacrema* ants (Hymenoptera: Formicidae) and the colonization of *Macaranga* (Euphorbiaceae) trees. *Molecular Phylogenetics and Evolution*, 27(3), 441–452.
- Feldhaar, H., Foitzik, S., & Heinze, J. (2008). Lifelong commitment to the wrong partner: hybridization in ants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1505), 2891-2899.
- Feldhaar, H., Gadau, J., & Fiala, B. (2010). Speciation in obligately plant-associated *Crematogaster* ants: Host distribution rather than adaptation towards specific hosts drives the process. In *Evolution in Action*. (ed. Glaubrecht, M.) Heidelberg, Germany: Springer.
- Fewell, J.H., & Page, R.E. (1993). Genotypic variation in foraging responses to environmental stimuli by honey bees, *Apis mellifera*. *Experientia*, 49(12), 1106-1112.
- FitzSimmons, N.N., Moritz, C., & Moore, S.S. (1995). Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. *Molecular Biology and Evolution*, 12(3), 432–440.
- Fournier, D., Estoup, A., Orivel, J., & Foucaud, J. (2005a). Clonal reproduction by males and females in the little fire ant. *Nature*, 435(7046), 1230–1234.

- Fournier, D., Foucaud, J., Loiseau, A., Cros-Arteil, S., Jourdan, H., Orivel, J., ... & Estoup, A. (2005b). Characterization and PCR multiplexing of polymorphic microsatellite loci for the invasive ant *Wasmannia auropunctata*. *Molecular Ecology Notes*, 5(2), 239–242.
- Foucaud, J., Fournier, D., Orivel, J., Delabie, J.H.C., Loiseau, A., Le Breton, J., ... & Estoup, A. (2007). Sex and clonality in the little fire ant. *Molecular Biology and Evolution*, 24(11), 2465–2473.
- Foucaud, J., Estoup, A., Loiseau, A., Rey, O., & Orivel, J. (2010). Thelytokous parthenogenesis, male clonality and genetic caste determination in the little fire ant: new evidence and insights from the lab. *Heredity*, 105(2), 205–212.
- Franks, N.R., & Hölldobler, B. (1987). Sexual competition during colony reproduction in army ants. *Biological Journal of the Linnean Society*, 30(3), 229–243.
- Fraser, V.S., Kaufmann, B., Oldroyd, B.P., & Crozier, R.H. (2000). Genetic influence on caste in the ant *Camponotus consobrinus*. *Behavioral Ecology and Sociobiology*, 47(3), 188–194.
- Frizzi, F., Santini, G., Natali, C., Chelazzi, G., & Ciofi, C. (2009). Characterization of polymorphic microsatellite loci in the ant *Crematogaster scutellaris*. *Conservation Genetics Resources*, 1(1), 425–428.
- Frohschammer, S., & Heinze, J. (2009). A heritable component in sex ratio and caste determination in a *Cardiocondyla* ant. *Frontiers in Zoology*, 6(1), 27.
- Frumhoff, P.C., & Baker, J. (1988). A genetic component to division of labour within honey bee colonies. *Nature*, 333(6171), 358–361.
- Fuchs, S., & Moritz, R.F.A. (1999). Evolution of extreme polyandry in the honeybee *Apis mellifera* L. *Behavioral Ecology and Sociobiology*, 45(3–4), 269–275.
- Gagnaire, P.-A., Pavey, S.A., Normandeau, E., & Bertnatchez, L. (2013). The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. *Evolution*, 67(9), 2483–2497.
- Garcia, F.H., Fischer, G., Peters, M.K., Snelling, R.R., & Wägele, J.W. (2009). A preliminary checklist of the ants (Hymenoptera: Formicidae) of Kakamega Forest (Kenya). *Journal of East African Natural History*, 98(2), 147–165.
- Gärke, C., Ytounel, F., Bed'hom, B., Gut, I., Lathrop, M., Weigend, S., & Simianer, H. (2011). Comparison of SNPs and microsatellites for assessing the structure of chicken populations. *Animal Genetics*, 43(4), 419–428.
- Grant, P.R., & Grant, B.R. (1992). Hybridization of bird species. *Science*, 256(5054), 193–197.

- Goropashnaya, A.V., Fedorov, V.B., & Pamilo, P. (2004). Recent speciation in the *Formica rufa* group ants (Hymenoptera, Formicidae): inference from mitochondrial DNA phylogeny. *Molecular Phylogenetics and Evolution*, 32(1), 198–206.
- Gotwald, W.H. (1995). *Army Ants: The Biology of Social Predation*. Ithaca, New York: Cornell University Press.
- Gotwald, W.H., & Schaefer, R.F. (1982). Taxonomic implications of doryline worker ant morphology: *Dorylus* subgenus *Anomma* (Hymenoptera: Formicidae). *Sociobiology*, 7(2), 187–204.
- Goudet, J. (1995). FSTAT Version 1.2: a computer program to calculate F-statistics. *Journal of Heredity*, 86(6), 485–486.
- Gröning, J., & Hochkirk, A. (2008). Reproductive interference between animal species. *The Quarterly Review of Biology*, 83(3), 257–282.
- Gyllenstrand, N., Gertsch, P.J., & Pamilo, P. (2002). Polymorphic microsatellite DNA markers in the ant *Formica exsecta*. *Molecular Ecology Notes*, 2(1), 67–69.
- Hansson, B., Tarka, M., Dawson, D.A., & Horsburgh, G.J. (2012). Hybridization but no evidence for backcrossing and introgression in a sympatric population of great reed warblers and clamorous reed warblers. *PLoS One*, 7(2), e31667.
- Harrison, R.G. (1993). Hybrids and hybrid zones: historical perspective. In *Hybrid Zones and the Evolutionary Process* (ed. Harrison, R.G.). New York, New York: Oxford University Press.
- Heinze, J., & Buschinger, A. (1986). Queen polymorphism in a non-parasitic *Leptothorax* species (Hymenoptera: Formicidae). *Insectes Sociaux*, 34(1), 28–43.
- Heinze, J., & Buschinger, A. (1989). Queen polymorphism in *Leptothorax* sp. A: its genetic and ecological background. *Insectes Sociaux*, 36(2), 139–155.
- Helms Cahan, S., & Keller, L. (2003). Complex hybrid origin of genetic caste determination in harvester ants. *Nature*, 424(6946), 306–309.
- Helms Cahan, S., Parker, J.D., Rissing, S.W., Johnson, R.A., Polony, T.S., Weiser, M.D., & Smith, D.R. (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(1503), 1871–1877.
- Helms Cahan, S., & Vinson, S.B. (2003). Reproductive division of labor between hybrid and nonhybrid offspring in a fire ant hybrid zone. *Evolution*, 57(7), 1562–1570.

- Hedrick, P.W. (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, 22(18), 4606–4618.
- Huang, M.H., Wheeler, D.E., & Fjerdingstad, E.J. (2013). Mating system evolution and worker caste diversity in *Pheidole* ants. *Molecular Ecology*, 22(7), 1998–2010.
- Hubbs, C.L. (1955). Hybridization between fish species in nature. *Systematic Zoology*, 4(1), 1–20.
- Hughes, W.O.H., & Boomsma, J.J. (2004). Genetic diversity and disease resistance in leaf-cutting ant societies. *Evolution*, 58(6), 1251–1260.
- Hughes, W.O.H., & Boomsma, J.J. (2008). Genetic royal cheats in leaf-cutting ant species. *Proceedings of the National Academy of Sciences of the United States of America*, 105(13), 5150–5153.
- Hughes, W.O.H., Sumner, S., Van Borm, S., & Boomsma, J.J. (2003). Worker caste polymorphism has a genetic basis in *Acromyrmex* leaf-cutting ants. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16), 9394–9397.
- Hung, A.C.F., & Vinson, S.B. (1977). Interspecific hybridization and caste specificity of protein in fire ant. *Science*, 196(4297), 1458–1460.
- Jaffé, R., Kronauer, D.J.C., Kraus, F.B., Boomsma, J.J., & Moritz, R.F.A. (2007). Worker caste determination in the army ant *Eciton burchellii*. *Biology Letters*, 3(5), 513–516.
- James, S.S., Pereira, R.M., Vail, K.M., & Ownley, B.H. (2002). Survival of imported fire ant (Hymenoptera: Formicidae) species subjected to freezing and near-freezing temperatures. *Environmental Entomology*, 31(1), 127–133.
- Jarne, P., & Lagoda, P.J.L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, 11(10), 424–429.
- Jessen, K., & Klinkicht, M. (1990). Hybridization in the social parasitic ant genus *Epymyrma* (Hymenoptera: Formicidae). *Insectes Sociaux*, 37(4), 273–293.
- Jones, J.C., Myerscough, M.R., Graham, S., & Oldroyd, B.P. (2004). Honey bee nest thermoregulation: diversity promotes stability. *Science*, 305(5682), 402–404.
- Jones, O.R., & Wang, J. (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, 10(3), 551–555.
- Julian, G.E., & Fewell, J.H. (2004). Genetic variation and task specialization in the desert leaf-cutter ant, *Acromyrmex versicolor*. *Animal Behavior*, 68(1), 1–8.

- Julian, G.E., Fewell, J.H., Gadau, J., Johnson, R.A., & Larrabee, D. (2002). Genetic determination of the queen caste in an ant hybrid zone. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12), 8157–8160.
- Jungman, L., Vigna, B.B.Z., Boldrini, K.R., Sousa, A.C.B., do Valle, C.B., Resende, M.S., ... & de Souza, A.P. (2010). Genetic diversity and population structure analysis of the tropical pasture grass *Brachiaria humidicola* based on microsatellites, cytogenetics, morphological traits and geographical origin. *Genome*, 53(9), 698–709.
- Kakazu, S., Sanches, A., & Bacci, M. (2013). Microsatellite loci characterized in the leaf-cutter ant *Atta laevigata*. *BMC Research Notes*, 6(1), 328.
- Keller, L., & Reeve, H.K. (1999). Dynamics of conflicts within insect societies. In *Levels of selection in evolution*. (ed. Keller, L.). Princeton, New Jersey: Princeton University Press.
- Keller, L., & Ross, K.G. (1998) Selfish genes: a green beard in the red fire ant. *Nature*, 394(6693), 573–575.
- Keller, L., Sundström, L., & Chapuisat, M. (1997). Male reproductive success: paternity contribution to queens and workers in *Formica* ants. *Behavioral Ecology and Sociobiology*, 41(1), 11–15.
- Kobayashi, K., Hasegawa, E., & Ohkawara, K. (2008). Clonal reproduction by males of the ant *Vollenhovia emeryi* (Wheeler). *Entomological Science*, 11(2), 167–172.
- Kobayashi, K., Hasegawa, E., & Ohkawara, K. (2011). No gene flow between wing forms and clonal reproduction by males in the long-winged form of the ant *Vollenhovia emeryi*. *Insectes Sociaux*, 58(2), 163–168.
- Kronauer, D.J.C., & Boomsma, J.J. (2007a). Do army ant queens re-mate later in life? *Insectes Sociaux* 54(1), 20–28.
- Kronauer, D.J.C., & Boomsma, J.J. (2007b). Multiple queens means fewer mates. *Current Biology*, 17(17), R753–R755.
- Kronauer, D.J.C., Boomsma, J.J., & Gadau, J. (2004a). Microsatellite markers for the driver ant *Dorylus (Anomma) molestus*. *Molecular Ecology Notes*, 4(2), 289–290.
- Kronauer, D.J.C., Boomsma, J.J., & Pierce, N.E. (2011a). Nine novel microsatellite loci for the army ant *Simopelta pergandei* (subfamily Ponerinae). *Conservation Genetics Resources*, 3(1), 61–63.
- Kronauer, D.J.C., Gadau, J., & Boomsma, J.J. (2007a). Characterization of 12 new microsatellite loci in *Aenictus* and *Neivamyrmex* army ants. *Molecular Ecology Notes*, 7(4), 688–690.

- Kronauer, D.J.C., Peters, M.K., Schöning, C., & Boomsma, J.J. (2011b). Hybridization in East African swarm-raiding army ants. *Frontiers in Zoology*, 8(1), 20.
- Kronauer, D.J.C., Schöning, C., & Boomsma, J.J. (2006). Male parentage in army ants. *Molecular Ecology*, 15(4), 1147-1151.
- Kronauer, D.J.C., Schöning, C., d'Ettorre, P., & Boomsma, J.J. (2010). Colony fusion and worker reproduction after queen loss in army ants. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 277(1682), 755-763.
- Kronauer, D.J.C., Schöning, C., Pedersen, J.S., Boomsma, J.J., & Gadau, J. (2004b). Extreme queen-mating frequency and colony fission in African army ants. *Molecular Ecology*, 13(8), 2381-2388.
- Kronauer, D.J.C., Schöning, C., Vilhelmsen, L., & Boomsma, J.J. (2007b). A molecular phylogeny of *Dorylus* army ants provides evidence for multiple evolutionary transitions in foraging niche. *BMC Evolutionary Biology*, 7(1), 56.
- Kulmuni, J., Seifert, B., & Pamilo, P. (2010). Segregation distortion causes large-scale differences between male and female ant genomes in hybrid ants. *Proceedings of the National Academy of Sciences of the United States of America* 107(16), 7371-7376.
- Kuo, H.-C., Hsu, H.-H., Chua, C.S., Wang, T.-Y., Chen, Y.-M., & Chen, T.-Y. (2014). Development of pedigree classification using microsatellite and mitochondrial markers for giant grouper broodstock (*Epinephelus lanceolatus*) management in Taiwan. *Marine Drugs*, 12(5), 2397-2407.
- Larson, E.L., Andrés, J.A., Bogdanowicz, S.M., & Harrison, R.G. (2013). Differential introgression in a mosaic hybrid zone reveals candidate barrier genes. *Evolution*, 67(12), 3653-3661.
- Lau, M.K., Ellison, A.M., Nguyen, A., Penick, C., DeMarco, B., Gotelli, N.J., ... & Helms Cahan, S. (2019). Draft *Aphaenogaster* genomes expand our view of ant genome size variation across climate gradients. *PeerJ*, 7, e6447.
- Lawson, D.J., van Dorp, L., & Falush, D. (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications*, 9, 3258.
- Leclercq, S., Rivals, E., & Jarne, P. (2010). DNA slippage occurs at microsatellite loci without minimal threshold length in humans: a comparative genomic approach. *Genome Biology and Evolution*, 2, 325-335.
- Leniaud, L., Darras, H., Boulay, R., & Aron, S. (2012). Social hybridogenesis in the clonal ant *Cataglyphis hispanica*. *Current Biology*, 22(13), 1188-1193.

- Leniaud, L., Percy, M., & Aron, S. (2013). Sociogenetic organisation of two desert ants. *Insectes Sociaux*, 60(3), 337–344.
- Libbrecht, R., Oxley, P.R., Kronauer, D.J.C., & Keller, L. (2013). Ant genomics sheds light on the molecular regulation of social organization. *Genome Biology*, 14(7), 212.
- Liersch, S., & Schmid-Hempel, P. (1998). Genetic variation within social insect colonies reduces parasite load. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1392), 221-225.
- Linksvayer, T.A, Wade, M.J., & Gordon D.M. (2006). Genetic caste determination in harvester ants: possible origin and maintenance by cyto-nuclear epistasis. *Ecology*, 87(9), 2185-2193.
- Litt, M., & Luty, J.A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics*, 44(3), 397–401.
- Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, 20(5), 229-237.
- Mallet, J. (2007). Hybrid speciation. *Nature*, 446(7133), 279-283.
- Manel, S., Berthier, P., & Luikart, G. (2002). Detecting wildlife poaching: identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. *Conservation Biology*, 16(3), 650-659.
- Marshall, J.A.R. (2015). *Social Evolution and Inclusive Fitness Theory: An Introduction*. Princeton, New Jersey: Princeton University Press.
- Matilla, H.R., & Seeley, T.D. (2007). Genetic diversity in honeybee colonies enhances productivity and fitness. *Science*, 317(5836), 362-364.
- Matute, D.R., Butler, I.A., & Coyne, J.A. (2009). Little effect of the *tan* locus on pigmentation in female hybrids between *Drosophila santomea* and *D. melanogaster*. *Cell*, 139(6), 1180-1188.
- Matute, D.R., Butler, I.A., Turissini, D.A., & Coyne, J.A. (2010). A test of the snowball theory for the rate of evolution of hybrid incompatibilities. *Science*, 329(5998), 1518-1521.
- Maudet, C., Luikart, G., & Taberlet, P. (2001). Development of microsatellite multiplexes for wild goats using primers designed from domestic Bovidae. *Genetics Selection and Evolution*, 33, S293–S203.
- Mayr, E. (1963). *Animal Species and Evolution*. Cambridge, Massachusetts: Harvard University Press.

- McGlynn, T.P., & Owen, J.P. (2002). Food supplementation alters caste allocation in a natural population of *Pheidole flavens*, a dimorphic leaf-litter dwelling ant. *Insectes Sociaux*, 49(1), 8-14.
- Mishina, T., Takada, M., Takeshima, H., Nakano, M., Tabata, R., Nishida, M., & Watanabe, K. (2014). Molecular identification of species and ploidy of *Carassius* fishes in Lake Biwa, using mtDNA and microsatellite multiplex PCRs. *Ichthyological Research*, 61(2), 169–175.
- Moilanen, A., Sundström, L., & Pedersen, J.S. (2004). MATESOFT: a program for deducing parental genotypes and estimating system statistics in haplodiploid species. *Molecular Ecology Notes*, 4(4), 795-797.
- Molnar, R.I., Witte, H., Dinkelacker, I., & Sommer, R.J. (2012). Tandem-repeat patterns and mutation rates in microsatellites of the nematode model organism *Pristionchus pacificus*. *G3: Genes, Genomes, Genetics*, 2(9), 1027–1034.
- Moore, S.S., Hale, P., & Byrne, K. (1998). NCAM: a polymorphic microsatellite locus conserved across eutherian mammal species. *Animal Genetics*, 29(1), 33–36.
- Moore, S.S., Sargeant, L.L., King, T.J., Mattick, J.S., Georges, M., & Hetzel D.J.S. (1991). The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics*, 10(3), 654–660.
- Moreau, C.S., & Bell, C.D. (2013). Testing the museum versus cradle tropical biological diversity hypothesis: phylogeny, diversification and ancestral biogeographic range evolution of the ants. *Evolution*, 67(8), 2240–2257.
- Moreau, C.S., Bell, C.D., Vila, R., Archibald, S.B., & Pierce, N.E. (2006). Phylogeny of the ants: diversification in the age of angiosperms. *Science*, 312(5770), 101–104.
- Moyle, L.C., & Nakazato, T. (2010). Hybrid incompatibility “snowballs” between *Solanum* species. *Science*, 329(5998), 1521-1523.
- Muller, H. (1942). Isolating mechanisms, evolution, and temperature. *Biological Symposia*, 6, 71-125.
- Naisbit, R.E., Jiggins, C.D., & Mallet, J. (2001). Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius Melpomene*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1478), 1849-1854.
- Nielsen, R., Tarpay, D.R., & Reeve, H.K. (2003). Estimating effective paternity number in social insects and the effective number of alleles in a population. *Molecular Ecology*, 12(11), 3157-3164.

- Nonacs, P. (2017). Go high or go low? Adaptive evolution of high and low relatedness societies in social Hymenoptera. *Frontiers in Ecology and Evolution*, 5, 87.
- Norman, V., Darras, H., Tranter, C., Aron, S., & Hughes, W.O. (2016). Cryptic lineages hybridize for worker production in the harvester ant *Messor barbarus*. *Biology Letters*, 12(11), 20160542.
- Nygaard, S., Zhang, G., Schiøtt, M., Li, C., Wurm, Y., Hu, H., ... & Pan, H. (2011). The genome of the leaf-cutting ant *Acromyrmex echinatior* suggests key adaptations to advanced social life and fungus farming. *Genome Research*, 21(8), 1339–1348.
- Ohkawara, K., Nakayama, M., Satoh, A., Trindl, A., & Heinze, J. (2006). Clonal reproduction and genetic caste differences in a queen-polymorphic ant, *Vollenhovia emeryi*. *Biology Letters*, 2(3), 359–363.
- Okasha, S. (2006). *Evolution and the Levels of Selection*. Oxford, United Kingdom: Oxford University Press.
- Oldroyd, B.P., & Fewell, J.H. (2007). Genetic diversity promotes homeostasis in insect colonies. *Trends in Ecology and Evolution*, 22(8), 263–281.
- Oldroyd, B.P., Rinderer, T.E., & Buco, S.M. (1991). Intracolony variance in honey bee foraging behaviour: the effects of sucrose concentration. *Journal of Apicultural Research*, 30(3-4), 137–145.
- Oldroyd, B.P., Rinderer, T.E., & Buco, S.M. (1992a). Intra-colony foraging specialism by honey bees (*Apis mellifera*) (Hymenoptera: Apidae). *Behavioral Ecology and Sociobiology*, 30(5), 291–295.
- Oldroyd, B.P., Rinderer, T.E., Buco, S.M., & Beaman, L.D. (1993). Genetic variance in honey bees for preferred foraging distance. *Animal Behaviour*, 45(2), 323–332.
- Oldroyd, B.P., Rinderer, T.E., Harbo, J.R., & Buco, S.M. (1992b). Effects of intracolony genetic diversity on honey bee (Hymenoptera: Apidae) colony performance. *Annals of the Entomological Society of America*, 85(3), 335–343.
- Oldroyd, B.P., Smolenski, A.J., Cornuet, J.M., Wongsiri, S., Estoup, A., Rinderer, T.E., & Crozier, R.H. (1995). Levels of polyandry and intracolony genetic relationships in *Apis florea*. *Behavioral Ecology and Sociobiology*, 37(5), 329–335.
- Oldroyd, B.P., Smolenski, A.J., Cornuet, J.M., Wongsiri, S., Estoup, A., Rinderer, T.E., & Crozier, R.H. (1996). Levels of polyandry and intracolony genetic relationships in *Apis dorsata* (Hymenoptera: Apidae). *Annals of the Entomological Society of America*, 89(2), 276–283.

- Orr, H.A. (1995). The population genetics of speciation; the evolution of hybrid incompatibilities. *Genetics*, 139(4), 1805-1813.
- Ortiz-Barrientos, D., Counterman, B.A., & Noor, M.A.F. (2004). The genetics of speciation by reinforcement. *PLoS Biology*, 2(12) e416.
- Oxley, P.R., Lu, J., Fetter-Pruneda, I., McKenzie, S.K., Li, C., Hu, H., ... & Kronauer, D.J.C. (2014). The genome of the clonal raider ant *Cerapachys biroi*. *Current Biology* 24(4), 451–458.
- Page, R.E., & Metcalf, R.A. (1982). Multiple mating, sperm utilization, and social evolution. *The American Naturalist*, 119(2), 263-281.
- Page, R.E., & Robinson, G.E. (1991). The genetics of division of labour in honey bee colonies. *Advances in Insect Physiology*, 23, 117-169.
- Page, R.E., Robinson G.E., Calderone N.W., & Rothenbuhler, W.C. (1989). Genetic structure, division of labor, and the evolution of insect societies. In *The Genetics of Social Evolution* (eds. Breed, M.D., & Page, R.E.). Boulder, Colorado: Westview Press.
- Page, R.E., Robinson G.E., Fondrk, M.K., & Nasr, M.E. (1995). Effects of worker genotypic diversity on honey bee colony development and behavior (*Apis mellifera* L). *Behavioral Ecology and Sociobiology*, 36(6), 387-396.
- Palmer, K.A., & Oldroyd, B.P. (2000). Evolution of multiple mating in the genus *Apis*. *Apidologie*, 31(2), 235-248.
- Palmer, K.A., & Oldroyd, B.P. (2003). Evidence for intra-colonial genetic variance in resistance to American foulbrood of honey bees (*Apis mellifera*): Further support for the parasite/pathogen hypothesis for the evolution of polyandry. *Naturwissenschaften*, 90(6), 265-268.
- Pamilo, P., Sundström, L., Fortelius, W., & Rosengren, R. (1994). Diploid males and colony-level selection in *Formica* ants. *Ethology, Ecology and Evolution*, 6(2), 221-235.
- Parker, J.D., & Rissing, S.W. (2002). Molecular evidence for the origin of workerless social parasites in the ant genus *Pogonomyrmex*. *Evolution*, 56(10), 2017–2028.
- Passera, L., Roncin, E., Kaufmann, B., & Keller, L. (1996). Increased soldier production in ant colonies exposed to intraspecific competition. *Nature*, 397(6566), 630-631.
- Paton, T., Haddrath, O., & Baker, A.J. (2002). Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1493), 839–846.

- Pearcy, M., Aron, S., Doums, C., & Keller, L. (2004). Conditional use of sex and parthenogenesis for worker and queen production in ants. *Science*, 306(5702), 1780-1783.
- Pearcy, M., Goodisman, M.A.D., & Keller, L. (2011). Sib mating without inbreeding in the longhorn crazy ant. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 278(1718), 2677-2681.
- Pearson, B. (1983a). Hybridization between the ant species *Lasius niger* and *Lasius alienus*—the genetic evidence. *Insectes Sociaux*, 30(4), 402–411.
- Pearson, B. (1983b). The classification of morphologically intermediate ants in the *Lasius alienus/Lasius niger* taxon. *Insectes Sociaux*, 30(1), 100-105.
- Peters, M.K., Fischer, G., Schaab, G., & Kraemer, M. (2009). Species compensation maintains abundance and raid rates of African swarm-raiding army ants in rainforest fragments. *Biological Conservation*, 142(3), 668-675.
- Peters, M.K., Lung, T., Schaab, G., & Wägele, J.W. (2011). Deforestation and the population decline of the army ant *Dorylus wilverthi* in western Kenya over the last century. *Journal of Applied Ecology*, 48(3), 697-705.
- Peters, M.K., & Okalo, B. (2009). Severe declines of ant-following birds in African rainforest fragments are facilitated by a subtle change in army ant communities. *Biological Conservation*, 142(10), 2050-2058.
- Peterson, B.K., Wever, J.N., Kay, E.H., Fisher, H.S., & Hoekstra, H.E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in non-model species. *PLoS One*, 7(5), e37135.
- Plateaux, L. (1979). Ovarian polymorphism of queens in *Leptothorax* ants and interspecific variations: inferiority of interspecific hybrids. *Archives de Zoologie Experimentale et Generale*, 120(4), 381–398.
- Pol, R.G., Lopez de Casenave, J., Feldhaar, H., Milesi, F.A., & Gadau, J. (2008). Polyandry in two South American harvester ants. *Insectes Sociaux*, 55(1), 91–97.
- Presgraves, D.C. (2002). Patterns of postzygotic isolation in Lepidoptera. *Evolution*, 56(6), 1168-1183.
- Pringle, E.G., Ramirez, S.R., Bonebrake, T.C., Gordon D.M., & Dirzo, R. (2012). Diversification and phylogeographic structure in widespread *Azteca* plant-ants from the northern Neotropics. *Molecular Ecology*, 21(14), 3576-3592.
- Pritchard, J.K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945-959.

- Purcell, J., Zahnd, S., Athanasiades, A., Turler, R., Chapuisat, M., & Brelsford, A. (2016). Ants exhibit asymmetric hybridization in a mosaic hybrid zone. *Molecular Ecology*, 25(19), 4866-4874.
- Pusch, K., Heinze, J., & Foitzik, S. (2006a). The influence of hybridization on colony structure in the ant species *Temnothorax nylander* and *T. crassispinus*. *Insectes Sociaux*, 53(4), 439-445.
- Pusch, K., Seifert, B., Foitzik, S., & Heinze, J. (2006b). Distribution and genetic divergence of two parapatric sibling ant species in Central Europe. *Biological Journal of the Linnean Society*, 88(2), 223-234.
- Qian, Z.-Q., Ceccarelli, F.S., Carew, M.E., Schlüns, H., Schlick-Steiner, B.C., & Steiner, F.M. (2011). Characterization of polymorphic microsatellites in the giant bulldog ant, *Myrmecia brevinoda* and the jumper ant, *M. pilosula*. *Journal of Insect Science*, 11(1), 1-8.
- Queller, D. (2005). Evolutionary biology: males are from mars. *Nature*, 435(7046), 1167-1168.
- Raignier, A., Van Boven, J.K.A., & Ceusters, R. (1974). Der Polymorphismus der afrikanischen Wanderameisen unter biometrischen und biologischen Gesichtspunkten. In *Sozialpolymorphismus bei Insekten. Probleme der Katenbildung im Tierreich* (ed. Schmidt, G.H.). Stuttgart, Germany: Wissenschaftliche Verlagsgesellschaft.
- Rey, O., Facon, B., Foucaud, J., Louiseau, A., & Estoup, A. (2013). Androgenesis is a maternal trait in the invasive ant *Wasmannia auropunctata*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 280(1766), 20131181.
- Rheindt, F.E., Strehl, C.P., & Gadau, J. (2005). A genetic component in the determination of worker polymorphism in the Florida harvester ant *Pogonomyrmex badius*. *Insectes Sociaux*, 52(2), 163-168.
- Rice, W.M. (1987). Speciation via habitat specialization: the evolution of reproductive isolation as a correlated character. *Evolutionary Ecology*, 1(4), 301-314.
- Rieseberg, L.H., Wendel, J.F. (1993). Introgression and its consequences in plants. In *Hybrid Zones and the Evolutionary Process* (ed. Harrison, R.G.). New York, New York: Oxford University Press.
- Robinson, G.E., & Page E.R. (1989). Genetic basis for division of labor in an insect society. In *The Genetics of Social Evolution* (eds. Breed, M.D., & Page, R.E.). Boulder, Colorado: Westview Press.
- Romiguier, J., Fournier, A., Yek, S.H., & Keller, L. (2017). Convergent evolution of social hybridogenesis in *Messor* harvester ants. *Molecular Ecology*, 26(4), 1108-1117.

- Rosenberg, N.A., Burke, T., Elo, K., Feldman, M.W., Freidlin, P.J., Groenen, M.A.M., ... & Wimmers, K. (2001). Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics*, 159(2), 699-713.
- Ross, K.G., & Robertson, J.L. (1990). Developmental stability, heterozygosity and fitness in two introduced fire ants (*Solenopsis invicta* and *S. richteri*) and their hybrid. *Heredity*, 64(1), 93-103.
- Ross, K.G., & Shoemaker, D.D. (2005). Species delimitation in native South American fire ants. *Molecular Ecology*, 14(11), 3419-3438.
- Ross, K.G., Vander Meer, R.K., Fletcher, D.J.C., & Vargo, E.L. (1987). Biochemical phenotypic and genetic studies of two introduced fire ants and their hybrid (Hymenoptera: Formicidae). *Evolution*, 41(2), 280-293.
- Rundle, H.D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, 8(3), 336-352.
- Rubin, B.E., Makarewich, C.A., Talaba, A.L., Stenzler, L., Bogdanowicz, S.M. & Lovette, I.J. (2009). Isolation and characterization of microsatellite markers from the acacia-ant *Crematogaster mimosae*. *Molecular Ecology Resources*, 9(4), 1212-1214.
- Rubin, B.E.R., Ree, R.H. & Moreau, C.S. (2012). Inferring phylogenies from RAD sequence data. *PLoS One*, 7(4), e33394.
- Schlick-Steiner, B.C., Steiner, F.M., Sanetra, M., Heller, G., Stauffer, C., Christian, E., & Seifert, B. (2005). Queen size dimorphism in the ant *Tetramorium moravicum* (Hymenoptera, Formicidae): Morphometric, molecular genetic and experimental evidence. *Insectes Sociaux*, 52(2), 186-193.
- Schlötterer, C. (2004). The evolution of molecular markers-just a matter of fashion? *Nature Reviews Genetics*, 5(1), 63-69.
- Schmid-Hempel, P. (1995). Parasites and social insects. *Apidologie*, 26(3), 255-271.
- Schmid-Hempel, P. (1997). Infection and colony variability in social insects. In *Infection, Polymorphism and Evolution* (eds. Hamilton, W.D., & Howard, J.C.). London, United Kingdom: Chapman & Hall.
- Schmid-Hempel, P. (1998). *Parasites in Social Insects*. Princeton, New Jersey: Princeton University Press.
- Schmid-Hempel, P., & Crozier, R.H. (1999). Polyandry versus polygyny versus parasites. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 354(1382), 507-515.

- Schöning, C. (2008). Driver ants (*Dorylus* subgenus *Anomma*) (Hymenoptera: Formicidae). In: *Encyclopedia of Entomology* (ed. Capinera, J.L.). Dordrecht, Netherlands: Springer.
- Schöning, C., Kinuthia, W., & Boomsma, J.J. (2006). Does the afrotropical army ant *Dorylus* (*Anomma*) *molestus* go extinct in fragmented forests? *Journal of East African Natural History*, 95(2), 163-179.
- Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, 18(2), 233-234.
- Schumer, M., Cui, R., Powell, D.L., Rosenthal, G.G., & Andolfatto, P. (2016). Ancient hybridization and genomic stabilization in a swordtail fish. *Molecular Ecology*, 25(11), 2661-2679.
- Schwander, T., Helms Cahan, S., & Keller, L. (2007a). Characterization and distribution of *Pogonomyrmex* harvester ant lineages with genetic caste determination. *Molecular Ecology*, 16(2), 367-387.
- Schwander, T., Keller, L., & Helms Cahan, S. (2007b). Two alternate mechanisms contribute to the persistence of interdependent lineages in *Pogonomyrmex* harvester ants. *Molecular Ecology*, 16(17), 3533-3543.
- Schwander, T., & Keller, L. (2008). Genetic compatibility affects queen and worker caste determination. *Science*, 332(5901), 552-552.
- Schwander, T., Lo, N., Beekman, M., Oldroyd, B.P., & Keller, L. (2010). Nature versus nurture in social insect caste differentiation. *Trends in Ecology and Evolution*, 25(5), 256-383.
- Schwander, T., Rosset, H., & Chapuisat, M. (2005). Division of labor and worker size polymorphism in ant colonies: the impact of social and genetic factors. *Behavioral Ecology and Sociobiology*, 59(2), 215-221.
- Seeley, T.D., & Tarpy, D.R. (2007). Queen promiscuity lowers disease within honeybee colonies. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 274(1606), 67-72.
- Selkoe, K.A., & Toonan, R.J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, 9(5), 615-629.
- Seppä, P., Johansson, H., Gyllenstrand, N., Pålsson, S., & Pamilo, P. (2012). Mosaic structure of native ant supercolonies. *Molecular Ecology*, 21(23), 5880-5891.
- Servedio, M.R., & Noor, M.A.F. (2003). The role of reinforcement in speciation: theory and data. *Annual Review of Ecology, Evolution, and Systematics*, 34(1), 339-364.

- Seifert, B. (1991). The phenotypes of the *Formica rufa* complex in East Germany. *Abhandlungen und Berichte des Naturkindemuseums Görlitz*, 65, 1–27.
- Seifert, B. (1995). Two new Central European subspecies of *Leptothorax nylanderi* (Förster, 1850) and *Leptothorax sodidulus* (Müller, 1923) (Hymenoptera: Formicidae). *Abhandlungen und Berichte des Naturkindemuseums Görlitz*, 68, 1–18.
- Seifert, B. (1999). Interspecific hybridisations in natural populations of ants by example of regional fauna (Hymenoptera: Formicidae). *Insectes Sociaux*, 46(1), 45–52.
- Seifert, B., D'Eustacchio, D., Kaufmann, B., Centorame, M., Lorite, P., & Modica, M.V. (2017). Four species within the supercolonial ants of the *Tapinoma nigerrimum* complex revealed by integrative taxonomy (Hymenoptera: Formicidae). *Myrmecological News*, 24, 123–144.
- Seifert, B., & Goropashnaya, A.V. (2004). Ideal phenotypes and mismatching haplotypes - errors of mtDNA treeing in ants (Hymenoptera: Formicidae) detected by standardized morphometry. *Organisms, Diversity & Evolution*, 4(4), 295–305.
- Seifert, B., Kulmuni, J., & Pamilo, P. (2010). Independent hybrid populations of *Formica polychteta* X *rufa* wood ants (Hymenoptera: Formicidae) abound under conditions of forest fragmentation. *Evolutionary Ecology*, 24(5), 1219–1237.
- Sherman, N.A., Victorine, A., Wang, R.J., & Moyle, L.C. (2014). Interspecific tests of allelism reveal the evolutionary timing and pattern of accumulation of reproductive isolation mutations. *PLoS Genetics*, 10(9), e1004623.
- Sherman, P.W., Seeley, T.D., & Reeve, H.K. (1988). Parasites, pathogens, and polyandry in social Hymenoptera. *The American Naturalist*, 131(4), 602–610.
- Shoemaker, D.D., Ahrens, M.E., and Ross, K.G. (2006). Molecular phylogeny of fire ants of the *Solenopsis saevissima* species-group based on mtDNA sequences. *Molecular Phylogenetics and Evolution*, 38(1), 200–215.
- Shykoff, J.A., & Schmid-Hempel, P. (1991a). Parasites and the advantage of genetic variability within social insect colonies. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 243(1306), 55–58.
- Shykoff, J.A., & Schmid-Hempel, P. (1991b). Genetic relatedness and eusociality: parasite-mediated selection on the genetic composition of groups. *Behavioral Ecology and Sociobiology*, 28(5), 371–376.
- Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. & Flook, P. (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(6), 651–701.

- Smith, C.D., Zimin, A., Holt, C., Abouheif, E., Benton, R., Cash, E., ... & Fave, M.J. (2011). Draft genome of the globally widespread and invasive Argentine ant (*Linepithema humile*). *Proceedings of the National Academy of Sciences of the United States of America*, 108(14), 5673–5678.
- Smith, C.R., Anderson, K.E., Tillberg, C.V., Gadau, J., & Suarez, A.V. (2008). Caste determination in a polymorphic social insect: nutritional, social and genetic factors. *The American Naturalist*, 172(4), 497–507.
- Smith, C.R., Smith, C.D., Robertson, H.M., Helmkamp, M., Zimin, A., Yandell, M., ... & Cash, E. (2011). Draft genome of the red harvester ant *Pogonomyrmex barbatus*. *Proceedings of the National Academy of Sciences of the United States of America*, 108(14), 5667–5672.
- Stebbins, G.L. (1958). The inviability, weakness, and sterility of interspecific hybrids. *Advances in Genetics*, 9, 147–215.
- Steiner, F.M., Arthofer, W., Schlick-Steiner, B.C., Crozier, R.H., & Stauffer, C. (2007). Eleven microsatellite loci in the sociobiologically enigmatic ant *Lasius austriacus* (Hymenoptera: Formicidae). *Molecular Ecology Notes*, 7(3), 498–500.
- Steiner, F.M., Seifert, B., Grasso, D.A., Le Moli, F., Arthofer, W., Stauffer, C., ... & Schlick-Steiner, B.C. (2011). Mixed colonies and hybridization of *Messor* harvester ant species (Hymenoptera: Formicidae). *Organisms, Diversity & Evolution*, 11(2), 107–134.
- Steiner, F.M., Seifert, B., Moder, K., & Schlick-Steiner, B.C. (2010). A multisource solution for a complex problem in biodiversity research: Description of the cryptic ant species *Tetramorium alpestre* sp.n. (Hymenoptera: Formicidae). *Zoologischer Anzeiger-A Journal of Comparative Zoology*, 249(3–4), 223–254.
- Steinmeyer, C., Pennings, P.S., & Foitzik, S. (2012). Multiclonal population structure and nestmate recognition in an extremely dense population of the European ant *Lasius flavus*. *Insectes Sociaux*, 59(4), 499–510.
- Stolle, E., Kidner, J.H., & Moritz, R.F.A. (2013). Patterns of evolutionary conservation of microsatellites (SSRs) suggest a faster rate of genome evolution in Hymenoptera than in Diptera. *Genome Biology and Evolution*, 5(1), 151–162.
- Stuart, R.J., & Page, R.E. (1991). A genetic component to division of labor among workers of a Leptothoracine ant. *Naturwissenschaften*, 78(8), 375–377.
- Suefuji, M., Trindl, A., & Heinze, J. (2011). Characterization of 13 microsatellite markers for the threatened slave-making ant *Myrmoxenus ravouxi* (Formicidae: Myrmecinae). *Conservation Genetics Resources*, 3(2), 229–231.

- Suen, G., Teiling, C., Li, L., Holt, C., Abouheif, E., Bornberg-Bauer, E., ... & Denas, O. (2011). The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle. *PLoS Genetics*, 7(2), e1002007.
- Sundström, L., & Boomsma, J.J. (2000). Reproductive alliances and posthumous fitness enhancement in male ants. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1451), 1439-1444.
- Tarpy, D.R., & Page, R.E. (2001). The curious promiscuity of queen honey bees (*Apis mellifera*): evolutionary and behavioral mechanisms. *Annales Zoologici Fennici*, 38, 255– 265.
- Tarpy, D.R., & Page, R.E. (2002). Sex determination and the evolution of polyandry in honey bees (*Apis mellifera*). *Behavioral Ecology and Sociobiology*, 52(2), 143–150.
- Tarpy, D.R., & Seeley, T.D. (2006). Lower disease infections in honeybee (*Apis mellifera*) colonies headed by polyandrous vs monandrous queens. *Naturwissenschaften*, 93(4), 195-199.
- Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, 17(16), 6463–6471.
- Tribble, W., & Kronauer, D. J. (2017). Caste development and evolution in ants: it's all about size. *Journal of Experimental Biology*, 220(1), 53-62.
- Tsutsui, N.D., Suarez, A.V., Spagna, J.C., & Johnston, J.S. (2008). The evolution of genome size in ants. *BMC Evolutionary Biology*, 8(1), 64.
- Turakulov, R., & Easteal, S. (2003). Number of SNPs loci needed to detect population structure. *Human Heredity*, 55(1), 37-45.
- Umphrey, G.J. (2006). Sperm parasitism in ants: Selection for interspecific mating and hybridization. *Ecology*, 87(9), 2148-2159.
- Umphrey, G.J., & Danzman, R.G. (1998). Electrophoretic evidence for hybridization in the ant genus *Acanthomyops* (Hymenoptera: Formicidae). *Biochemical Systematics and Ecology*, 26(4), 431-440.
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., & Rozen, S.G. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Research*, 40(15), e115.
- Vallejo-Marín, M., & Hiscock, S. J. (2016). Hybridization and hybrid speciation under global change. *New Phytologist*, 211(4), 1170-1187.
- Van der Have, T.M., Pedersen, J.S., & Boomsma, J.J. (2011). Mating, hybridization and introgression in *Lasius* ants (Hymenoptera: Formicidae). *Myrmecological News*, 15, 109-115.

- Vander Meer, R.K., Lofgren, C.S., & Alvarez, F.M. (1985). Biochemical evidence for hybridization in fire ants. *Florida Entomologist*, 68(3), 501-506.
- Vinson, S.B. (1997). Insect Life: Invasion of the red imported fire ant (Hymenoptera: Formicidae) *American Entomologist*, 43(1), 23–39.
- Volny, V.P., & Gordon, D.M. (2002). Genetic basis for queen–worker dimorphism in a social insect. *Proceedings of the National Academy of Sciences of the United States of America*, 99(9), 6108-6111.
- Wang, J. (2004). Sibship reconstruction from genetic data with typing errors. *Genetics*, 166(4), 1963-1979.
- Wang, J., & Scribner, K.T. (2014). Parentage and sibship inference from markers in polyploids. *Molecular Ecology Resources*, 14(3), 541–553.
- Wang, R.S., White, M.A., & Payseur, B. (2015). The pace of hybrid incompatibility evolution in house mice. *Genetics*, 201(1), 229-242.
- Wang, S., Meyer, E., McKay, J.K., & Matz, M.V. (2012). 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nature Methods*, 9(8), 808–810.
- Warner, M.R., Lipponen, J., & Linksvayer, T.A. (2018). Pharaoh ant colonies dynamically regulate reproductive allocation based on colony demography. *Behavioral Ecology and Sociobiology*, 72(3), 31.
- Weber, J.L., & May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, 44(3), 388– 396.
- Whelden, R.W. (1963). The anatomy of the adult queen and workers of the army ants *Eciton burchelli* Westwood and *Eciton hamatum* Fabricius (continued). *Journal of the New York Entomological Society*, 71(2), 90-115.
- Whiting, P.W. (1933). Selective fertilization and sex determination in Hymenoptera. *Science*, 78(2032), 537-538.
- Wiernasz, D.C., & Cole, B.J. (2010). Patriline shifting leads to apparent genetic caste determination in harvester ants. *Proceedings of the National Academy of Sciences of the United States of America*, 107(29), 12958-12962.
- Wilson, E.O. (1964). The true army ants of the Indo-Australian area (Hymenoptera: Formicidae: Dorylinae). *Pacific Insects*, 6(3), 427-483.

- Wing, M.W. (1968). Taxonomic revision of the Nearctic genus *Acanthomyops* (Hymenoptera: Formicidae). Cornell University Agricultural Experiment Station, Ithaca, New York: Memoir 405, pp. 1-173.
- Winter, U., & Buschinger, A. (1986). Genetically mediated queen polymorphism and caste determination in the slave-making ant, *Harpagoxenus sublaevis* (Hymenoptera: Formicidae). *Entomologia Generalis*, 11(3-4), 125–137.
- Wu, C.I. (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology*, 14(6), 851-865.
- Wurm, Y., Wang, J., Riba-Grognuz, O., Corona, M., Nygaard, S., Hunt, B.G., ... & Dijkstra, M.B. (2011). The genome of the fire ant *Solenopsis invicta*. *Proceedings of the National Academy of Sciences of the United States of America*, 108(14), 5679–5684.