THE ROLE OF *GP-9* IN REGULATING SOCIAL BEHAVIOR IN FIRE ANTS

by

DIETRICH AXEL GOTZEK

(Under the Direction of Kenneth G. Ross)

ABSTRACT

This research presented here represents a modest contribution to the regulation of social organization in the red imported fire ant, *Solenopsis invicta* Buren, and close relatives. Colony queen number is associated with a certain allele at the nuclear protein-coding locus *general protein-9* (*Gp-9*). Multiple queen colonies (polygyny) always contain the *b* allele. Single queen colonies (monogyny) always and only contain the *B* allele. The research presented here seeks to begin to bridge the gap between molecular pattern and process in our understanding of the association between *Gp-9* and polygyny in fire ants. I have sought to show how a relatively small number of the *b* allele carrying workers are sufficient to elicit polygyne behavior and that queen effects do not influence the acceptance of supernumerary queens. This represents an important first step into understanding the ties between individual genotypes and colony level expression of social organization.

I have also attempted to show that single nucleotide substitutions at the *Gp-9* locus are not sufficiently associated with polygyne behavior, but that several such changes must occur for the expression of the polygyne phenotype. This raises interesting questions about the validity of the presumed causal role of the *b*-like alleles at *Gp-9* in inducing polygyne behavior in South American fire ants.

Finally, I discuss the state of our current understanding of *Gp-9* and its role in regulating social behavior, resulting in a simple, testable model, and suggestions for future avenues of research. Whether *Gp-9* remains the prime candidate gene for controlling social organization in fire ants remains to be seen. But even if it does not, it will deserve continued attention by students of social behavior.

INDEX WORDS: *Solenopsis invicta*, fire ants, polygyny, social organization

THE ROLE OF *GP-9* IN REGULATING SOCIAL BEHAVIOR IN FIRE ANTS

by

DIETRICH AXEL GOTZEK

Diplom-Biologe, Universität Bayreuth, Germany, 1999

M.S., Western Carolina University, 2000

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2006

THE ROLE OF *GP-9* IN REGULATING SOCIAL BEHAVIOR IN FIRE ANTS

by

DIETRICH AXEL GOTZEK

Major Professor: Kenneth G. Ross

Committee: Michael Arnold
Rodney Mauricio
Daniel Promislow
Michael Strand

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2006

TABLE OF CONTENTS

Page

LIST OF TABLES v

LIST OF FIGURES vi

CHAPTER

1 INTRODUCTION 1

2 GENETIC REGULATION OF COLONY SOCIAL ORGANIZATION IN FIRE ANTS: AN INTEGRATIVE OVERVIEW 9

3 EXPERIMENTAL CONVERSION OF COLONY SOCIAL ORGANIZATION IN FIRE ANTS (*SOLENOPSIS INVICTA*): WORKER GENOTYPE MANIPULATION IN THE ABSENCE OF QUEEN EFFECTS 77

4 MOLECULAR VARIATION AT A CANDIDATE GENE REGULATING COMPLEX SOCIAL BEHAVIOR IN THE FIRE ANT *SOLENOPSIS INVICTA* 108

5 CONCLUSIONS 160

LIST OF TABLES

Page

Table 4.1.: DIVERSITY STATISTICS FOR GP-9 IN *S. INVICTA*. 152

Table 4.2.: AMINO ACID VARIATION AT THREE GP-9 CODON POSITIONS..... 153

Table 4.3.: IDENTIFICATION OF SITE-SPECIFIC POSITIVE SELECTION ON GP-9 BY DIFFERENT METHODS..... 154

LIST OF FIGURES

Page

Figure 2.1.: ALLELE FREQUENCY DIFFERENTIATION BETWEEN SYMPATRIC MONOGYNE AND POLYGYNE *S. INVICTA* POPULATIONS FROM THE INTRODUCED AND NATIVE RANGES AS ASSESSED BY ESTIMATES OF F_{ST} 74

Figure 2.2.: PHYLOGENETIC POSITION OF *SOLENOPSIS* GP-9 SEQUENCES RELATIVE TO ODORANT BINDING PROTEIN (OBP) AND CHEMOSENSORY PROTEIN (CSP) SEQUENCES FROM OTHER INSECTS.. 75

Figure 2.3.: REGULATION OF QUEEN NUMBER IN A POLYGYNE *S. INVICTA* COLONY BY MEANS OF QUEEN-CASTE SIGNALS PRODUCED BY INDIVIDUAL ADULT FEMALES. 76

Figure 3.1.: SOCIAL BEHAVIOR EXPRESSED BY TREATMENT COLONIES DURING BIOASSAYS, AND PROPORTIONS OF *BB* AND *b*-BEARING WORKERS AT CONVERSIONS IN SOCIAL FORM. 106

Figure 3.2.: PROPORTIONS OF ADULT *b*-BEARING WORKERS IN TREATMENT COLONIES EXPRESSING MONOGYNE OR POLYGYNE SOCIAL BEHAVIOR. 107

Figure 4.1.: LOCATIONS OF SAMPLING SITES FOR *S. INVICTA* IN ITS NATIVE RANGE, ... 156

Figure 4.2.: DISTRIBUTIONS OF NUMBERS OF NUCLEOTIDE SUBSTITUTIONS BETWEEN PAIRS OF UNIQUE GP-9 HAPLOTYPES IN *S. INVICTA*. 157

Figure 4.3.: NUMBERS OF DIFFERENT AMINO ACIDS OCCURING AT EACH CODON POSITION OF GP-9 IN THE B-LIKE AND B-LIKE ALLELE CLASSES IN *S. INVICTA* 158

Figure 4.4.: SUMMARY PHYLOGENETIC HYPOTHESIS OF GP-9 HAPOLTYPE RELATIONSHIPS BASED ON FOUR INDEPENDENT MCMC RUNS USING BAYESIAN INFERENCE 159

CHAPTER 1

INTRODUCTION

Social organisms, especially the complex and highly integrated insect societies, have always held a certain fascination for man (Costa 2002). As a result, they have been intensely studied and have become important model systems for our understanding of ecology, evolution, and behavior (Wilson 1975; Oster and Wilson 1978; Grafen 1984; Hamilton 1996). By their very nature they constitute profound evolutionary transitions in the history of life (Wilson 1975; Maynard Smith and Szathmáry 1995; Keller 1999). The taxonomic distribution of social behavior is immense, not only indicating that social systems have repeatedly arisen independently, but also that significant adaptive advantages apply to group living (Wilson 1971, 1975; Costa 2006). Perhaps the best known examples involve the evolution of complex insect societies (Wilson 1971; Costa 2006), especially in the Hymenoptera (Michener 1974; Hölldobler and Wilson 1990; Ross and Matthews 1991; Bourke and Franks 1995; Crozier and Pamilo 1996).

A major goal in evolutionary biology is to elucidate the genetic architecture of complex adaptations (Orr and Coyne 1992; Orr 1998, 2005), i.e., to describe the number of genes involved in regulating a particular trait, how these genes interact with one another and the environment, and how they produce the phenotype. This holds true for social behavior as well, with the added level of how expression of certain genes influences relevant social phenotypes in other individuals and how this integrates to the

colony level (Grafen, 1984; Crozier and Pamilo 1996; Robinson et al. 1997; Robinson 1999; Linksvayer and Wade 2005; Robinson et al. 2005). Only with such information can researchers hope to adequately model the evolution of important social adaptations (e.g., Robinson 1999; Linksvayer and Wade 2005; Vasemägi and Primmer 2005).

However, the sole focus on the mechanistic and structural basis of complex traits ignores another fundamental component of evolutionary biology, namely the documentation of variation and its evolutionary history. Genetic variation is the raw fuel of evolution and any serious evolutionary study must sooner or later address this issue. Hence, one of the main goals in evolutionary biology is to document genetic variation and to understand its maintenance (Gillespie 1991; Hedrick 2000).

This tension between pattern (i.e., population genetic variation and molecular evolution) and process (i.e., genetic architecture and molecular function) is inherent to the study of molecular adaptation, and bridging this divide can greatly enrich any study into the genetics of adaptation (Golding and Dean 1998). And exactly this is the professed goal of "sociogenomics" (Robinson 1999), integrating molecular variation, biochemistry, physiology, individual phenotypes, and integrative social behavior into a complex whole to fully understand the emergent social phenotype of a colony.

Increasingly, genetic components to specific behaviors have been documented in social insects (reviewed in Ross and Keller 2002), but have mainly focused on the Hymenoptera (e.g., ants, bees, and wasps, with the exception of tent caterpillars [Costa and Ross 2003]). But perhaps not surprisingly, few complex group-level traits have been shown to have an apparently simple genetic architecture featuring one or few genes of major effect (reviewed in Ross and Keller 2002), possibly explaining the dearth of

knowledge on the level of genetic variation in these genes, or to link this variation to biochemical and physiological difference which affect the social phenotype.

Of the sparse examples of few genes of major effect, the best characterized system involves a single gene, *general protein-9* (*Gp-9*) thought to regulate colony queen number in fire ants. The system has been most clearly elucidated in *Solenopsis invicta* Buren, the red imported fire ant, in which it also was first described (Ross 1997). Some colonies in this and closely related species are headed by a single reproductive queen (monogyny), whereas others have multiple reproductive queens (polygyny). This phenotypic dimorphism in social organization is mirrored by a genetic biallelic polymorphism at the *Gp-9* locus. All members of monogyne colonies always and only carry the *B* allele. In contrast, colonies containing multiple queens (polygyny) always contain the *b* allele, alongside the *B* allele (Shoemaker and Ross 1996; Ross 1997; Krieger and Ross 2002, 2005). Coupled with the absence of any other gene linked to multiple queen societies, this observation led to the hypothesis that the *b* allele is necessary and sufficient to elicit polygyne behavior (Ross 1997; Ross and Keller 1998, 2002; Krieger and Ross 2002, 2005). While it is conceivable that *Gp-9* simply marks a tightly linked gene or genes that actually determine social organization, a major role for non-genetic factors such as prior worker social experience, queen fecundity, and the social origin of queens in the expression of this colony-level trait has been discounted (Ross and Keller 1998). Associated with variation in queen number are a whole suit of important reproductive and life history traits (Bourke and Franks 1995; Tschinkel 2005).

While the association between the *b* allele and polygyny has been carefully studied, to date two critical areas have received little attention. First, it is not completely clear how the *b* allele induces polygyny, even at the highest, most integrated (i.e., colony) level. In a recent study Ross and Keller (2002) concluded that a minimum percentage of workers were required to carry the *b* allele for the colony to express the polygyne phenotype, but they could not exclude the role of queens in eliciting supernumerary queen acceptance. Second, a clear and explicit model of how a single gene of major effect could regulate queen number has been lacking, hindering the formulation of testable hypotheses about *Gp-9* at the physiological, behavioral, and colony level. Third, and perhaps most importantly, a detailed catalogue of the existing genetic variation has been missing. Such information is absolutely crucial to undeniably link allelic differences to social organization or to dispel its involvement altogether and relegate it to (an incomplete) marker status. Also, should the association be upheld, hypotheses can be proposed about functional differences between the allelic gene products, which open the door to mechanistic studies of protein function and its role in the physiology, metabolism, and behavior, linked pattern and process, as envisioned by Golding and Dean (1999) and Robinson (1999).

Goals of this Thesis

The work presented here endeavors to address several shortcomings of our understanding of how *Gp-9* is associated with polygyny and how it potentially regulates its expression. As such, it seeks to tentatively, perhaps prematurely, gap the divide between pattern and process, illuminating the need for further research into combining

the observed patterns with functional and mechanistic studies of the genes effects at the molecular level and how this integrates up to the colony level.

First, I show that allelic variation at *Gp-9*, or a closely linked gene, in the worker force is necessary and sufficient to elicit polygyne behavior in *S. invicta* colonies. Second, I attempt to demonstrate that sequence-specific variation at *Gp-9* is consistent with its role in regulating social organization in fire ants. Finally, I review our current understanding of the system, propose an improved model of how GP-9 could conceivably affect acceptance of supernumerary queens in fire ant colonies, and suggest future avenues of research.

REFERENCES

Bourke AFG, Franks NR. 1995. *Social evolution in ants*. Princeton, NJ: Princeton University Press.

Costa JT. 2002. Scale models? What insect societies teach us about ourselves. Proc Am Phil Soc 146:170-180.

Costa JT. 2006. *The Other Insect Societies*. Cambridge: Belknap.

Costa JT, Ross KG. 2003. Fitness effects of group merging in a social insect. Proc R Soc Lond B 270:1697-1702.

Crozier R H, Pamilo P. 1996. *Evolution of social insect colonies; sex allocation and kin selection*. Oxford: Oxford University Press.

Gillespie JH. 1991. *The Causes of Molecular Evolution*. Oxford: Oxford University Press.

Golding GB, Dean AM. 1998. The structural basis of molecular adaptation. Mol Biol Evol 15:355-369.

Grafen A. 1984. Natural selection, kin selection and group selection. In Krebs JR, Davies NB. (eds.) *Behavioural Ecology; An Evolutionary Approach*, Sunderland: Sinauer, pp. 62-84.

Hamilton WD. 1996. *Narrow Roads of Gene Land. Vol. 1 Evolution of Social Behavior*. New York: Freeman.

Hedrick PW. 2000. *Genetics of Populations*. 2nd edition. Sudbury: Jones and Bartlett.

Hölldobler B, Wilson EO. 1990. *The Ants*. Cambridge: Belknap.

Keller L. 1999. *Levels of Selection in Evolution.* Princeton: Princeton University Press.

Krieger MJB, Ross KG. 2002. Identification of a major gene regulating complex social behavior. Science 295:328-332.

Krieger MJB, Ross KG. 2005. Molecular evolutionary analyses of the odorant-binding protein gene *Gp-9* in fire ants and other *Solenopsis* species. Mol Biol Evol 22:2090-2103.

Linksvayer TA, Wade MJ. 2005. The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: maternal effects, sib-social effects, and heterochrony. Quart Rev Biol 8:317-336.

Maynard Smith J, Szathmáry E. 1995. *The Major Transitions in Evolution.* New York: Freeman.

Michener CD. 1974. *The Social Behavior of the Bees: A Comparative Study*. Cambridge: Belknap.

Orr HA. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52:935–949.

Orr HA. 2005. The genetic theory of adaptation: a brief history. Nature Rev Gen 6:119-127.

Orr HA, Coyne JA. 1992. The genetics of adaptation: a reassessment. Am Nat 140:725-742.

Oster GF, Wilson EO. 1978. *Caste and Ecology in the Social Insects*. Princeton: Princeton University Press.

Robinson GE. 1999. Integrative animal behaviour and sociogenomics. Trends Ecol Evol 14:202-205.

Robinson GE, Fahrbach SE, Winston ML. 1997. Insect societies and the molecular biology of social behavior. Bioessays 19:1099-1108.

Robinson GE, Grozinger CM, Whitfield CW. 2005. Sociogenomics: social life in molecular terms. Nature Rev Genet 6:257-271.

Ross KG. 1997. Multilocus evolution in fire ants: effects of selection, gene flow, and recombination. Genetics 145:961-974.

Ross KG, Matthews RW. 1991. *The Social Biology of Wasps*. Ithaca: Comstock.

Ross KG, Keller L. 1998. Genetic control of social organization in an ant. Proc Natl Acad Sci USA 95:14232-14237.

Ross KG, Keller L. 2002. Experimental conversion of colony social organization by manipulation of worker genotype composition in fire ants (*Solenopsis invicta*). Behav Ecol Sociobiol 51:287-295.

Shoemaker DD, Ross KG. 1996. Effects of social organization on gene flow in the fire ant *Solenopsis invicta*. Nature 383:613-616.

Tschinkel W R. 2006. *The Fire Ants*. Cambridge: Belknap Press.

Vasemägi A, Primmer CR. 2005. Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. Mol Ecol 14:3623-3642.

Wilson EO. 1971. *The Insect Societies*. Cambridge, Belknap.

Wilson EO. 1975. *Sociobiology: the new synthesis*. Cambridge, Belknap.

CHAPTER 2

GENETIC REGULATION OF COLONY SOCIAL ORGANIZATION IN FIRE ANTS: AN INTEGRATIVE OVERVIEW[1]

[1] Dietrich Gotzek and Kenneth G. Ross, to be submitted to *Quarterly Review of Biology*

Abstract.—Expression of colony social organization in fire ants appears to be under the control of a single Mendelian factor of large effect. Variation in social organization (colony queen number) in *Solenopsis invicta* and its close relatives is completely associated with allelic variation at the nuclear gene *Gp-9* but not with variation at other unlinked genes, and workers regulate queen identity and number on the basis of *Gp-9* genotypic compatibility. Non-genetic factors such as prior social experience of queens and workers, reproductive status of queens, and local environmental conditions have negligible effects on social organization, illustrating the nearly complete penetrance of *Gp-9* under diverse conditions. As predicted, social organization can be manipulated experimentally by altering worker *Gp-9* genotype frequencies. The *Gp-9* allele lineage associated with polygyny in South American fire ants has been retained across multiple speciation events, which may signal the historical action of balancing selection to maintain social polymorphism in these species. Moreover, positive selection is implicated in driving the molecular evolution of *Gp-9* in association with the origin of polygyny. The identity of the product of *Gp-9* as an odorant binding protein suggests plausible scenarios for its direct involvement in the regulation of social organization via a role in chemical communication. While these and other lines of evidence show that *Gp-9* represents a legitimate candidate gene of major effect, studies aimed at determining the biochemical pathways in which GP-9 functions, the phenotypic effects of molecular variation at *Gp-9* and other pathway genes, and the potential involvement of other genes in linkage disequilibrium with *Gp-9* are needed to fully elucidate the genetic architecture underlying expression of social organization in fire ants. Detailed information revealing the links between molecular variation, individual phenotype, and colony-level behaviors,

combined with refined behavioral models incorporating details of the chemical communication involved in regulation of queen number, will yield a novel integrated view of the evolutionary changes underlying a key social adaptation.

INTRODUCTION

A major goal in evolutionary biology is to understand the genetic architecture of complex adaptations in wild populations (Orr and Coyne 1992; Orr 1998, 2005). Only with information on the numbers and types of genes influencing expression of key phenotypes, the spectra of mutational effects at these genes, the patterns of epistasis and degree of pleiotropy the genes exhibit, and the norms of reaction of the phenotypes can comprehensive models of the evolution of significant adaptations be developed (e.g., Robinson 1999; Baker et al. 2001; Linksvayer and Wade 2005; Nachman 2005; Phillips 2005; Storz 2005; Vasemägi and Primmer 2005; Greenberg and Wu 2006). In the study of social behavior, in particular, intense interest has centered on the genetic architectures underlying important social adaptations (de Bono and Bargmann 1998; Krieger and Ross 2002; Lim et al. 2005; Robinson et al. 2005; Linksvayer and Wade 2005; Nedelcu and Michod 2006). This interest stems in large part from curiosity about the numbers and types of genetic steps that convert a solitary animal into a group-living one and that, subsequently, can radically transform the structure of social groups. Such information is substantially enriched when accompanied by descriptions of the causal biochemical and physiological links between molecular variation, individual phenotype, and social behavior, an integrative view Robinson (1999) has termed "sociogenomics."

Increasingly, genetic components to specific behaviors have been documented in social insects. For example, individual reproductive roles in honey bees (Moritz and Hillesheim 1985; Page and Robinson 1994; Montague and Oldroyd 1998), worker task performance in ants, wasps, honey bees, and caterpillars (Snyder 1992; Hunt et al. 1995; O'Donnell 1996; Fewell and Bertram 2002; Costa and Ross 2003), and components of

the dance language of honey bees (Johnson et al. 2002) have been shown to have a heritable basis. At the level of the social group, genetic effects on colony social organization have been documented in ants and sweat bees (Cahan et al. 1998; Plateaux-Quénu et al. 2000; Julian et al. 2002; Schwander et al. 2006). Perhaps not surprisingly, few complex group-level traits have been shown to have an apparently simple genetic architecture featuring one or few factors of major effect (Moritz 1988; Hunt et al. 1995; Ross and Keller 1998; Johnson et al. 2002). Of these, the best characterized system involves the genetic regulation of colony queen number in fire ants.

The fire ant *Solenopsis invicta* exists in two distinct social forms that differ in the number of reproductive queens per colony. Monogyne colonies are always headed by a single reproductive (wingless, egg-laying) queen, whereas polygyne colonies contain multiple such queens (sometimes hundreds). The two social forms differ not only in colony queen number but in many other life history and reproductive traits (Ross and Keller 1995; Tschinkel 2006). For instance, colony founding in the monogyne form occurs when young virgin queens undertake nuptial flights, during which they mate and disperse widely, then attempt to start a colony without the help of workers (independently) and without foraging (claustrally). In preparation, young monogyne queens accumulate extensive fat reserves during a two-week maturation period in their natal nest (Keller and Ross 1993a, 1999; DeHeer 2002). Young polygyne queens, in contrast, often make only very limited nuptial flights or forgo them altogether and mate within their nest (Porter 1991; Ross et al. 1996a; DeHeer et al. 1999; Goodisman et al. 2000a). Queens of this form rarely try to found nests independently, instead opting to enter existing polygyne nests in attempts to gain admission as new reproductives

(Glancey and Lofgren 1988; Porter 1991; Goodisman and Ross 1999). Associated with these alternative reproductive strategies, most young polygyne queens acquire far less fat reserves than their monogyne counterparts (Keller and Ross 1993a; DeHeer 2002).

The contrasting dispersal and reproductive strategies in the two social forms lead to different colony and population structures. Mature monogyne colonies, with their single reproductive queens, are simple families that are highly territorial and aggressive towards non-nestmate conspecifics (Ross and Fletcher 1985; Ross et al. 1997; Vander Meer and Alonso 2002), leading to a highly overdispersed distribution of nests (Tschinkel 2006). Polygyne colonies contain multiple queens and families, and both nestmate relatedness and aggression between non-nestmates is reduced compared to the monogyne form (Ross and Fletcher 1985; Ross et al. 1997; Vander Meer and Alonso 2002). Consequently, polygyne ants move between nests, including newly budded and parent nests (Vargo and Porter 1989), nests frequently are clustered, and individual ants as well as nests are present at far higher densities than in the monogyne form (Porter 1992). Thus, a seemingly simple difference between the forms, colony queen number, affects population attributes such as colonizing potential, as well as higher order ecological attributes such as energy flow through food webs (Tschinkel 2006). Although some differences exist between polygyne populations in the native South American range and in the USA, where the species has been introduced, the fundamental dichotomy in colony and population structure associated with the two social forms is universal in *S. invicta* (Ross et al. 1996a; Porter et al. 1997; Ross et al. 1997).

The existence of a strong heritable component to the expression of colony queen number, featuring a single Mendelian factor of large effect, has been well documented in

S. invicta and its close relatives over the past two decades. In this paper, we review the evidence for this genetic architecture, discuss the influence of non-genetic factors on social organization, and describe a candidate gene implicated in regulation of this complex social trait. We then discuss the remaining important gaps in our knowledge of this system, emphasizing the most promising avenues for future research to fill these gaps. We conclude by presenting a new behavioral model for the regulation of colony queen number in fire ants that is consistent with available data. Future refinement of our model pending the advent of information linking molecular variation, individual phenotype, and colony-level behaviors promises to provide new insights into the molecular, physiological, and behavioral changes underlying the evolution of a key social adaptation.

WHAT HAS BEEN LEARNED ABOUT *GP-9*

A single Mendelian factor regulates social organization in S. invicta

Variation in colony queen number in invasive *S. invicta* in the USA appears to be determined by allelic variation at a single Mendelian factor. This has been established by virtue of the fact that variation in social organization is completely associated with allelic variation at a single nuclear, protein-coding gene. This gene, designated *general protein-9* (*Gp-9*), encodes electrophoretically distinct protein products that have been shown by family studies to be inherited as allelic variants (Ross 1997). The nature of the association of social organization with *Gp-9* variation is remarkably simple (Ross 1997; Ross and Keller 1998; Shoemaker et al. 2006). Without exception, the *b* allele of the

gene (*Gp-9*^{b}) is represented at relatively high frequency among the inhabitants of polygyne colonies, whereas this allele is never found among the inhabitants of monogyne colonies. Colonies of the monogyne form instead contain only individuals bearing the alternate, *B* allele (*Gp-9*^{B}), which occurs as well in the polygyne form along with the *b* allele. An important related difference in *Gp-9* composition between the forms concerns the genotypes of reproductive queens. While all monogyne colonies are headed by a single homozygous *BB* queen, virtually all reproductive queens in polygyne colonies are *Bb* heterozygotes. (Reasons for the near exclusive occurrence of these heterozygotes are discussed below.) Thus, there is no overlap in the genotypes of reproductive queens of the two social forms. Importantly, DNA sequence data have confirmed the allelic nature of the *B* and *b* variants of *Gp-9* initially characterized by protein electrophoresis (Krieger and Ross 2002).

Allelic variation at a second nuclear gene, *phosphoglucomutase-3* (*Pgm-3*), also has been shown to be associated with social organization in *S. invicta* (Ross 1986; Keller and Ross 1993b; Ross et al. 1996b), although the association is not complete as it is for *Gp-9*. Indeed, variation at *Pgm-3* does not remain significantly associated with traits diagnostic for the two forms, such as queen weight and acceptability to workers, once the effects of *Gp-9* genotype are taken into account (Keller and Ross 1999). The cause of the association of *Pgm-3* variation with social organization when the gene is considered alone presumably is its tight linkage and strong gametic disequilibrium with *Gp-9* and/or other genes linked to *Gp-9* that affect social behavior (Ross 1997, Keller and Ross 1999). Nonetheless, the association between *Pgm-3* allele frequencies and social form adds

further weight to the conclusion that colony social organization is regulated by a single Mendelian factor of large effect.

Other nuclear genes surveyed in *S. invicta* do not exhibit meaningful differentiation associated with social organization. Indeed, allele frequencies at a variety of polymorphic loci typically are highly similar between the two social forms where they occur in sympatry in the invasive (USA) and native (Argentina) ranges (Figure 2.1), consistent with extensive interform gene flow (Ross and Shoemaker 1993; Shoemaker and Ross 1996; Goodisman et al. 2000b). Evident conclusions to be drawn from these data are: i) most of the nuclear genome is shared between the two social forms by virtue of their recent ancestry and ongoing gene flow, ii) elements of the genome marked by *Gp-9* and *Pgm-3* are not shared as extensively between the forms, presumably because of incompatibilities they induce between the forms in their social and breeding habits (Ross and Shoemaker 1993; Shoemaker and Ross 1996; Ross and Keller 1998), and iii) these elements do not cover a large portion of the nuclear genome. This last conclusion also is consistent with the existence of a single Mendelian factor with a major effect on social organization.

The defining feature of social organization, the number and identity of the reproductive queens heading a colony, is under the collective control of workers, which tolerate and nurture queens judged to be acceptable as new reproductives and execute those deemed not to be so (Fletcher and Blum 1983a; Keller and Ross 1993b; Keller and Ross 1998). A corollary of the observation that colony *Gp-9* composition is completely predictive of social organization is that workers of each social form discriminate against queens of inappropriate genotype when these queens attempt to become supernumerary

or replacement reproductives. Indeed, a series of laboratory experiments have confirmed that i) colonies with workers bearing the *b* allele (typical polygyne colonies) accept multiple queens also bearing this allele but do not tolerate *BB* queens, and ii) colonies with only *BB* workers (typical monogyne colonies) accept single *BB* replacement queens but not queens bearing the *b* allele (Keller and Ross 1998; Ross and Keller 1998; Keller and Ross 1999). These results are important in explaining the characteristic genotype compositions of colonies in the wild.

Rarely, *Gp-9* genotype composition is not fully consistent with other evidence regarding the form of social organization of a colony. In a study of 1060 colonies sampled widely across the USA range, Shoemaker et al. (2006) found that 0.4% of colonies lacked the *b* allele despite evidence from other markers that they comprised multiple families, whereas 2.3% of colonies contained only one detectable matriline despite the presence of some females bearing the *b* allele. The former most likely were monogyne colonies in which queen turnover had recently occurred (DeHeer and Tschinkel 1998), while the latter most likely were polygyne colonies with very low effective queen numbers. Significantly, every polygyne *S. invicta* colony confirmed as such on the basis of recovery of multiple reproductive queens has been shown to contain individuals bearing the *b* allele (Ross 1997; Ross et al. 1999; Ross and Keller 1998, 2002; Fritz et al. 2006).

Non-genetic factors have little influence on social organization in S. invicta

Although population genetic data from *S. invicta* clearly implicate a single genetic factor in the expression of social organization, non-genetic factors conceivably could still

play some role. Therefore, the influence of such factors as prior worker social experience and queen reproductive phenotype in overriding the effects of colony *Gp-9* composition on social organization were tested in controlled laboratory trials (Ross and Keller 1998). Effectively, such trials assess the penetrance of *Gp-9* in controlling the essential features of social organization across environments (for colony-level traits such as social organization, penetrance can be defined as the proportion of colonies of a given genotype composition that display the expected colony phenotype). The trials employed in these and related studies tested the willingness of temporarily queenless colonies of varying history and genetic composition to accept multiple introduced queens bearing the *b* allele, a hallmark of polygyny (alternatively, they tested the willingness of colonies to accept a single *BB* queen, a hallmark of monogyny).

Ross and Keller (1998) speculated that workers that were reared and spent their entire adult lives in a colony with a single reproductive queen might be inclined to remain tolerant of only a single replacement queen. However, trials using a variety of such effectively monogyne test colonies revealed that only the presence or absence of the *b* allele in the worker force influences queen acceptance. That is, effectively monogyne colonies with workers bearing allele *b* expressed the polygyne social phenotype by accepting multiple *Bb* queens but rejecting *BB* queens, whereas colonies lacking this allele expressed the monogyne social phenotype by accepting single *BB* queens but rejecting *Bb* queens.

The social environment in which introduced queens were reared also was shown to have no effect on worker acceptance of queens (Keller and Ross 1998; Ross and Keller 1998). Polygyne or effectively monogyne test colonies with workers bearing the *b* allele

accepted multiple *Bb* queens but rejected *BB* queens, regardless of whether the latter originated from polygyne or monogyne colonies. Conversely, effectively monogyne test colonies lacking allele *b* accepted single *BB* queens originating from colonies of either social form. The social environment in which the reproductive queens heading each test colony were reared similarly was found to be unimportant. Altogether, these results indicate that the prior social experience of both workers and queens plays little role in the expression of social organization.

The reproductive status of sexually mature queens (whether they are virgin or mated, non–egg-layers or egg-layers) also seems not to play a role in their acceptability to workers in polygyne test colonies. Moreover, the degree of physogastry (fecundity) of egg-laying queens plays no role once *Gp-9* genotype is accounted for. This was tested by experimentally restricting the diets of monogyne *BB* queens and supplementing the diets of single polygyne *Bb* queens, thus reversing the normal phenotype/genotype associations found in the wild by producing low-fecundity *BB* queens and high-fecundity *Bb* queens (Ross and Keller 1998). Such queens were accepted or rejected by polygyne colonies solely on the basis of their *Gp-9* genotypes (all *Bb* queens were accepted while no *BB* queens were accepted; see also Ross [1988] for evidence that multiple highly physogastric *Bb* queens are tolerated in polygyne colonies). This finding is perhaps surprising because fecundity is an important trait affecting acceptability of replacement reproductive queens in the monogyne form (Fletcher and Blum 1983a), in which the *BB* genotype constitutes a uniform *Gp-9* background. Evidently, queen-derived cues associated with *Gp-9* genotype (presumably chemical) occupy a high position in the hierarchy of cues used by workers to regulate colony queen composition in *S. invicta*.

One non-genetic factor that interacts with *Gp-9* genotype to play a role in the acceptability of pre-reproductive queens is their age (degree of sexual maturity) (Keller and Ross 1998, 1999). Virgin queens possessing the *BB* genotype increasingly become targets of worker aggression in polygyne colonies as they mature sexually following emergence from the pupa. (This aggression towards queens lacking a *b* allele is perpetrated mainly by workers that possess it, suggesting that the *b* allele is a selfish genetic element that causes its bearers to act in ways that bias its transmission [Keller and Ross 1998; Mescher 2001]). The time period during which aggression escalates, between a few days and two weeks of age, coincides with a period of profound physiological change of queens in anticipation of the onset of reproduction (e.g., Goodisman and Ross 1999; Brent and Vargo 2003; Tian et al. 2004). Apparently, some chemical cue produced in association with reproductive development is used by workers to identify queens of different *Gp-9* genotype (Keller and Ross 1999).

Chemical cues that make reproductive queens individually recognizable to workers are well known from monogyne *S. invicta*. Once imprinted on the chemical signature of a single queen, monogyne workers generally will not accept any replacement queen unless they have been held queenless for at least one to several days (Fletcher 1986). Such imprinting on single queens, which apparently also occurs to some extent with polygyne workers (Ross and Keller 1998), can override the effects of *Gp-9* genotype in queen introduction experiments conducted within 24 h of dequeening—hence, the need to use temporarily dequeened colonies in the trials described above. Conversely, colonies of either form that have been held queenless for prolonged periods become increasingly tolerant of any conspecific reproductive queens encountered. For this reason,

experiments conducted with colonies held queenless for a week or longer (Vander Meer and Alonso 2002) probably reveal little useful information regarding the normal role of *Gp-9* in regulating colony queen number.

Evidence that the local physical environment exerts no influence on social organization comes from the fact that the two social forms of *S. invicta* are broadly distributed in both the native and introduced ranges and that they often occur in close proximity within single microhabitats (Porter et al. 1991, 1992; Ross et al. 1997; Shoemaker et al. 2006). This lack of social plasticity in response to local conditions contrasts with the apparent situation in some ants in which the form of social organization is associated with, and presumably induced by, the type of habitat occupied (Elmes and Keller 1993; Herbers 1993; Bourke and Franks 1995).

The combined results from the studies summarized above show conclusively that the interaction of queen and worker *Gp-9* genotypes is the single dominant factor involved in regulation of colony queen identity and number in invasive *S. invicta* in the USA. Normal queenright colonies containing workers with the *b* allele tolerate multiple reproductive queens also bearing the allele, whereas colonies containing only homozygous *BB* workers tolerate only a single queen, which also must bear the *BB* genotype. These experimental results illustrating the nearly complete penetrance of *Gp-9* under diverse conditions are fully consistent with the allele and genotype distributions observed in the two forms in the wild.

Worker Gp-9 *genotype composition predicts colony social organization*

The preceding data led to the development of a specific phenomenological hypothesis of the relation between colony *Gp-9* genotypic composition and social organization: the presence in a colony of workers bearing *b*-like alleles induces expression of the polygyne social phenotype, whereas the absence of such workers induces the monogyne phenotype. The resulting simple prediction that colony social organization can be altered by experimentally manipulating worker genotype frequencies has been tested in two studies on invasive *S. invicta*. In the first, queens of each social form were cross-fostered into previously dequeened colonies of the alternate form (Ross and Keller 2002). Over a period of months, the adult worker genotype compositions gradually changed to those characteristic of the alternate form; that is, monogyne colonies that adopted a polygyne queen acquired increasing frequencies of workers bearing allele *b*, whereas, polygyne colonies that adopted a monogyne queen gradually lost workers with the *b* allele. As predicted, all experimental colonies switched their social phenotype over the time course of the study (as assessed by introducing multiple *Bb* reproductive queens), and these changeovers consistently occurred at a threshold of 5-10% workers bearing the *b* allele. An alternative, non-genetic, explanation for the results, that workers became increasingly habituated to their adopted queens and thus more tolerant of introduced queens of the same *Gp-9* genotype (and intolerant of queens of the alternate genotype) is inconsistent with the results described above showing a lack of such queen influences. Moreover, it does not explain why the changeovers typically did not occur for several months, only when the frequencies of workers with allele *b* passed the 5-10% threshold.

In order to entirely eliminate a role for any such queen effects in inducing social changeovers, a second study fostered polygyne worker brood containing the *b* allele into monogyne test colonies that retained their original queens (Gotzek and Ross 2006). As controls, monogyne test colonies received brood lacking allele *b* pooled from many other monogyne colonies (this was done to match the genetic variation present in the fostered polygyne brood). All treatment colonies switched their social phenotype (behaved as polygyne colonies) at several weeks after brood introduction, when emergence from the added brood led to 12-42% of all adult workers bearing the *b* allele. No control colonies switched. Remarkably, treatment colonies generally reverted back to the monogyne phenotype once their frequencies of adult workers with allele *b* again dropped below 15% due to replacement of the adult foreign workers by offspring of the resident queen.

Both of these studies support the phenomenological hypothesis and, equally significantly, they suggest that only a relatively low proportion of workers (10-15%) need possess the *b* allele in order that polygyne behavior be expressed by a colony. Proportions of such workers always greatly exceed this threshold in polygyne colonies in the wild (46-80% [Ross and Keller 2002; Fritz et al. 2006]), as expected since all polygyne reproductive queens possess allele *b*. Nonetheless, a major unresolved issue surrounding this finding is how a relatively small fraction of workers can bring about a major change in this emergent, group-level behavior (Ross and Keller 2002, see also below).

A single Mendelian factor marked by Gp-9 regulates social organization in the close relatives of S. invicta

The strong association of colony *Gp-9* genotype composition with social organization is not limited to invasive *S. invicta*. In the native range, polygyne colonies of this species invariably possess *b*-like alleles (variants similar in their amino acid sequence to the *b* allele from the invasive range), whereas monogyne colonies lack them, instead possessing only *B*-like alleles (variants similar to the *B* allele from the invasive range) (Krieger and Ross 2002, 2005; Mescher et al. 2003). Moreover, this same association has been found in the four closest relatives of *S. invicta* in their native South American ranges (*S. macdonaghi*, *S. megergates*, *S. quinquecuspis*, and *S. richteri*), where polygyny again is associated with the presence of *b*-like alleles in a colony and monogyny is associated with the presence of only *B*-like alleles (Krieger and Ross 2002, 2005). The association of *b*-like alleles with polygyny in these other fire ant species has been particularly well documented in native *S. richteri* (Hallar et al. 2006). Significantly, *S. invicta* and these other species are the only members of a large group of fire ants centered in South America in which *b*-like alleles have been discovered, and they are also the only members of this group known to display both monogyny and polygyny. Thus, there is a strong phylogenetic correlation between the occurrence of social polymorphism and the presence of a major *Gp-9* allele polymorphism in South American fire ants.

Phylogenetic analyses of *Gp-9* nucleotide sequences from 21 New World *Solenopsis* species have revealed that the *b*-like alleles from the five socially polymorphic South American fire ants form an exclusive, relatively recently derived clade (Krieger and Ross 2002, 2005). Hence, the association of this class of alleles with polygyny apparently has

evolved only once in the genus, and its retention in the different species represents a trans-species polymorphism. As has been inferred for other such polymorphisms (Vekemans and Slatkin 1994; Klein et al. 1998; Bos and Waldman 2006), persistence of the *b*-like allele lineage across multiple speciation events may signal the historical action of balancing selection. Such selection presumably reflects the distinct demographic advantages of each form of social organization within each socially polymorphic species (see Ross and Keller 1995; Tschinkel 2006).

We note that *Gp-9* polymorphism is not linked to polygyny in all *Solenopsis* species. The monogyne and polygyne forms of *S. geminata*, a quite distant fire ant relative of *S. invicta* (Pitts et al. 2005), possess identical *Gp-9* alleles where they co-occur in northern Florida, USA (Ross et al. 2003). There, the polygyne form displays greatly reduced genetic variation compared to the monogyne form, suggesting that it underwent a recent bottleneck and that loss of allelic variation at genes encoding recognition cues may have contributed to the emergence of polygyny by eroding worker discrimination capabilities. Importantly, such loss of genetic diversity can be ruled out as a factor promoting polygyny in *S. invicta* (Ross et al. 2003). Making matters even more complicated in *S. geminata*, a polymorphism in the signal sequence distinguishes *Gp-9* alleles of polygyne ants from Chiapas, Mexico from all other sequences of this or any other *Solenopsis* species, suggesting yet another possible route to polygyny in fire ants (Krieger and Ross 2005). A common element in the various postulated genetic causes of polygyny in fire ants is changes in worker discrimination abilities that function in the regulation of colony queen number (see below).

Gp-9 *encodes an odorant-binding protein*

Complete *Gp-9* sequences have been obtained for 21 *Solenopsis* species, with several sequences available for each of the more common fire ant species (Krieger and Ross 2002, 2005) and many dozen sequences available for native *S. invicta* (Gotzek et al. 2007). The exon/intron structure and respective lengths of the five exons of *Gp-9* are conserved across *Solenopsis*, and there is no evidence of significant historical or recent intragenic recombination; thus, variation in the coding region seems to have evolved solely or primarily by means of point substitutions (Krieger and Ross 2005). The 1700 bp gene of the "true" fire ants encodes a protein of 153 amino acids that, when cleaved of its 19-residue signal peptide, yields a mature protein with an estimated molecular mass of 14.7 kD (Krieger and Ross 2002).

BLAST searches using amino acid sequences initially revealed that *Gp-9* from *S. invicta* most closely resembles tortricid moth genes of the insect odorant binding protein (OBP) family (Krieger and Ross 2002). With the advent of new sequences from this family, *Gp-9* now is judged to be most similar to different OBP genes, *OBP56e* from the mosquito and *OBP3* from the honey bee (October 2006). GP-9 shares several important structural features with other insect OBPs, including its size, presence of a signal sequence, and presence of six cystein residues arranged in a highly characteristic pattern (Vogt 2003, 2005). Most importantly with respect to its classification, GP-9 groups phylogenetically with other insect OBPs (Box 1). Because initial phylogenetic analyses failed to resolve the placement of GP-9 with respect to other OBPs (Vogt 2003, 2005), we added as outgroups to the original data set ten chemosensory proteins (CSPs), a group believed to be closely related to OBPs (Vogt et al. 1999; Nagnan-Le Meillour and

Jacquin-Joly 2003). Although the overall topologies of the resulting trees for all 169 sequences vary considerably depending on the alignment and tree-building methods used, *Gp-9* sequences from *Solenopsis* always form a monophyletic group placed within the OBP family (Figure 2). Indeed, *Gp-9* falls within a major clade including many of the OBPs regarded as the "gold standard" proteins of this family (based on their localization to chemosensilla and ability to bind ligands; Vogt 2003). Phylogenetic analyses of a reduced data set also strongly support the conclusion that *Gp-9* is an OBP gene (Figure 2.2).

OBPs that have been biochemically characterized are small, water soluble, extracellular carrier proteins that usually bind small hydrophobic ligands. Originally described from the antennae of Lepidoptera, proteins of this family subsequently were shown to be present in the lumen of the antennal olfactory sensilla and to be capable of binding and transporting pheromones or food odorants (Picimbon 2003; Leal 2003; Vogt 2003, 2005; Xu 2005); these "gold standard" OBPs thus appear to be crucial molecular components of insect chemoreception that function in transporting hydrophobic chemostimulants to receptors on the dendrites of sensory neurons. More recently, some OBP genes have been shown to be expressed in non-chemosensory organs (Vogt 2003, 2005), suggesting that the products in these cases function in transport roles other than those serving peripheral chemoreception. The sites of expression of *Gp-9* and identity of the ligand(s) to which its product binds are unknown, although the protein is abundant in the hemolymph of adult females (D. Gotzek and K. Ross, unpubl.). Thus, *Gp-9* may fall in the ranks of those OBP genes encoding products that do not function in the traditional manner assumed for the "gold standard" OBPs.

GP-9 *is functionally important*

Several lines of evidence suggest that *Gp-9* encodes a functionally important product. First, OBP homologs of *Gp-9* often play significant roles in the sensory lives of insects, enabling them to locate mates or appropriate food sources (Vogt 2003). Further, the sex specificity and time course of expression suggest that *Gp-9* has an important function, which may relate to chemical communication involved in regulating social organization. In *S. invicta*, the gene is not expressed in the larval or pupal stages of either sex, and it is expressed in adult males only at relatively low levels (Ross 1997; Liu and Zhang 2004). In adult workers and queens, the protein is undetectable in newly emerged individuals, increases in abundance to high levels by approximately 7-10 days of age, then remains abundant throughout adult life (Ross 1997; Liu and Zhang 2004). Mature workers exhibit relative mRNA levels four-fold higher than non-reproductive queens (Liu and Zhang 2004). These GP-9 expression patterns parallel the patterns of expression of social behaviors involved in regulating queen number; males are not involved in such activities and mature adult workers increasingly discriminate among young adult queens of different *Gp-9* genotype as these queens mature.

A further clue that *Gp-9* is functionally important comes from the fact that several individual traits are associated with *Gp-9* genotype in *S. invicta*. The best studied involve physiological and behavioral phenotypes of non-reproductive and young reproductive queens. Among samples from the introduced (USA) range, *Gp-9* genotype is strongly associated with the weight of young queens, an effect caused by differential accumulation of fat during adult maturation (Keller and Ross 1993a,b, 1995, 1999; Ross and Keller 1998; DeHeer et al. 1999; DeHeer 2002). Mature *BB* queens are the heaviest individuals

with the greatest fat reserves (regardless of social form of origin), *Bb* queens are of intermediate weight, and *bb* queens are of lowest weight (data from the native range suggest similar inhibitory effects on queen weight for all alleles of the *b*-like class [Mescher 2001]). Extensive reserves are needed to carry a young monogyne (*BB*) queen through independent, claustral colony founding (Markin et al. 1972; DeHeer et al. 1999; DeHeer 2002), whereas they are not needed by *b*-bearing queens, which typically seek adoption by an established colony before initiating reproduction.

Gp-9 genotype also is associated with the dispersal and oviposition behaviors of young queens. Homozygous *BB* queens of both forms disperse widely during mating flights, *Bb* queens disperse more locally (as expected if these queens seek out established polygyne nests to join), and *bb* queens rarely undertake mating flights at all (DeHeer et al. 1999). Under laboratory conditions mimicking those following dispersal, *BB* queens have earlier onset of oogenesis and higher fecundity than queens bearing the *b* allele (Keller and Ross 1993b; Ross and Keller 1998; DeHeer 2002). Rapid onset of reproduction is a trait suited to queens that found colonies independently (*BB* queens), because successfully founding a colony in this way depends on the rapid production of many workers (Markin et al. 1973). On the other hand, it is not imperative for young queens joining established colonies (most *Bb* queens) to begin rapid egg production soon after mating.

These genotypic differences among queens early in their reproductive lives apparently carry over to egg-layers in mature colonies. Monogyne (*BB*) queens are capable of producing more eggs per unit body weight than polygyne (*Bb*) queens (Vander Meer et al. 1992), suggesting some inherent limitation on the metabolic efficiency of *Bb*

queens related to egg production (Tschinkel 2006). Altogether, these results show that queen *Gp-9* genotype is associated with a suite of phenotypic characteristics contributing to the different reproductive syndromes of the two social forms.

Somewhat surprisingly, adult body weight in workers and males also has been shown to be associated with *Gp-9* genotype. Similar to queens, individuals of these castes bearing the *b* allele tend to be lighter than individuals lacking it (Goodisman et al. 1999). Presence of the *b* allele in polygyne workers thus partly explains the relatively smaller size of workers of this form (Greenberg et al. 1985; Goodisman et al. 1999).

An unusual phenotypic effect of *Gp-9* genotype suggesting functional importance of the gene is the reduced viability of *S. invicta* workers and queens bearing genotype *bb*. Adult workers with the genotype are under-represented or even absent in population samples, and *bb* queens rarely survive to become egg-layers (Ross 1997; DeHeer et al. 1999; DeHeer 2002; Fritz et al. 2006; Hallar et al. 2006). The recessive deleterious effects of the *b* allele of *S. invicta* apparently are not characteristic of other *b*-like alleles found in this species and in *S. richteri* (Hallar et al. 2006). Significantly, the *b* allele features a radical, charge-altering amino acid substitution not found in any other *Gp-9* alleles, including other *b*-like alleles. While it is conceivable that this mutation arose in association with a mutant allele at another gene producing the deleterious effect, and that the two mutations have since remained in complete disequilibrium, an equally reasonable conclusion is that the radical substitution influences the normal function of GP-9. Supporting evidence for this view is the inference that the altered residue occurs at a position involved either in ligand binding/ unloading or in the formation of a biologically active dimer (Krieger 2005; Hallar et al. 2006). In either case, a charge-changing

substitution at this position might be expected to confer deleterious effects on the *b* allele by altering the normal functioning of its product.

Perhaps the most important evidence for the functional importance of *Gp-9* comes from tests of historical selection acting directly on this gene (Krieger and Ross 2002, 2005). These tests reveal a significant excess of nonsynonymous (amino acid replacing) substitutions relative to synonymous (silent) substitutions in the stem lineage of the *b*-like allele clade of the socially polymorphic South American fire ants. The implication is that positive selection has driven the molecular evolution of *Gp-9* (selection at linked genes cannot produce such patterns), and that such selection acted specifically in the context of the joint emergence of *b*-like alleles and polygyne social organization in this group. Such selection can be hypothesized to have operated via the ligand-binding properties of GP-9, as one of the two amino acids uniquely shared by all *b*-like alleles is predicted to be a binding-pocket residue (Krieger and Ross 2005).

WHAT MUST BE LEARNED ABOUT *GP-9*

Gp-9*: Candidate gene or marker?*

Although much is known about *Gp-9* and its association with social traits in *S. invicta*, a great deal more remains to be learned about the potential direct involvement of the gene in regulation of social organization. Given the overwhelming evidence that social organization acts like a trait under simple genetic control (single Mendelian factor of large effect), the issue can be distilled to the question of whether *Gp-9* is a legitimate candidate gene or simply a marker for one or more other genes that produce the observed

effects on social behavior and with which *Gp-9* is in complete gametic disequilibrium. Although decisive functional and manipulative genetic studies have yet to be undertaken, there is circumstantial support for each viewpoint.

Direct Involvement of Gp-9 *in the Regulation of Social Organization*

GP-9 is an insect odorant binding protein (OBP). This family includes several proteins that have been implicated as molecular components in signal transduction during detection of chemostimulants (reviewed in Leal 2003; Plettner 2003; Vogt 2003, 2005). Consistent with a role in olfaction or gustation, many OBP genes are expressed solely in the antennae or other structures bearing chemosensilla (i.e., mouthparts, legs, wings). Given the evolutionary and structural affinities of GP-9 with the "gold standard" OBPs implicated in chemoreception, a possible role for the protein in fire ants is transport of pheromones employed in the regulation of colony queen number, an idea made plausible because the core feature of such regulation is discrimination among queens by workers based on specific chemical signals emanating from the queens (Keller and Ross 1998; Ross and Keller 1998). One hypothetical scenario is that GP-9 functions in the manner traditionally conceived for insect OBPs (as a tranporter of odorants from cuticular pores to receptors on sensory neurons in chemosensilla), and that the *B*-like and *b*-like protein forms differ in their binding properties. Such biochemical differences have been postulated to confer different recognition capabilities on workers of different genotypes, with different colony phenotypes of collective worker tolerance toward queens emerging from the different worker genotype compositions in each form (Krieger and Ross 2002).

Unfortunately, tests to confirm one basic premise of this scenario, that *Gp-9* is expressed in appropriate chemosensilla of females, have not yet been conducted.

A related scenario acknowledges the likely role of GP-9 as a pheromone transporter protein, but in a manner different from that traditionally conceived for OBPs. Evidence has mounted for considerable heterogeneity in expression patterns and, by implication, specific functions of insect OBPs (Vogt 2005). Several are expressed not in chemosensory organs but in such sites as the hemolymph (Graham et al. 2001; Paskewitz and Shi 2005), male accessory gland (Paesen and Happ 1995), and other tissues devoid of known chemosensory structures (Galindo and Smith 2001; Hekmat-Scafe et al. 2002). In *S. invicta* as well, GP-9 is present in the hemolymph of adult females (D. Gotzek and K. Ross, unpubl.), while it is not expressed in the antennae of males (Guntur et al. 2004) (antennae of females have not yet been examined). OBPs also are not present as major proteins in the antennae of other ant females (Ishida et al. 2002; Ozaki et al. 2005). These data suggest the possibility that GP-9, and perhaps other hemolymph OBPs, may be involved not in pheromone detection but in some other molecular component of chemical communication (Calvello et al. 2003). Indeed, Pelosi et al. (2005) speculate that GP-9 might act to transport a pheromone or its precursor from the site of production to relevant organs of activation or release. (Anatomical separation of the sites of pheromone production and release has been described in fire ant queens [Vargo and Hulsey 2000].) Flexibility and interchangeability of roles in chemical communication analogous to that suspected for OBPs has been proposed as well for the related CSP family in insects (Calvello et al. 2003; Pelosi et al. 2005).

A more speculative scenario posits GP-9 as a hemolymph carrier protein serving some primary function other than chemical communication, perhaps as a transporter of small hydrophobic endocrine factors (Pelosi et al. 2005). Again, CSPs may provide analogous examples if they are confirmed to function not only as transporters of chemostimulants but also as general lipid carriers involved in more wide-ranging functions, including regulation of development (e.g., Vogt et al. 1999; Nagnan-Le Meillour and Jacquin-Joly 2003; Wanner et al. 2005). One possibility is that GP-9 has some regulatory metabolic function, similar to lipocalin transport proteins (Åkerstrom et al. 2000), and may influence individual signal production or perception via indirect physiological routes. Regardless of the specific transport function of GP-9 in fire ants, its direct involvement in regulation of colony social structure is suggested by the fact that the protein is highly up-regulated in the individuals that participate in the process, mature queens and workers (Ross 1997; Liu and Zhang 2004).

Another line of evidence that *Gp-9* may be directly involved is the invariant association of *b*-like alleles with polygyny in all members of the socially polymorphic clade of South American fire ants (Krieger and Ross 2002, 2005; Hallar et al. 2006). The persistence of this association is difficult to explain if *Gp-9* does not directly affect social organization, because the inexorable pressure of recombination over evolutionary time should lead to the decay of gametic disequilibrium between variation at *Gp-9* and the actual genes involved, even if they are tightly linked to *Gp-9* (Hedrick et al. 1978; Langley et al. 2000; Remington et al. 2001), yielding a concomitant decay in the association between *b*-like alleles and polygyny. Considering the 600 Mb genome size (Li and Heinz 2000) and assuming the presence of 10,000-20,000 genes, a gene occurs on

average every 30-60 kb in *S. invicta*. Using conservative estimates of the extent of crossing over between linked human genes (Sabeti et al. 2006), and assuming a minimum age of 500,000 years for the *b*-like allele clade, anywhere from five to several dozen crossing over events are expected to have occurred between *Gp-9* and adjacent genes since the origin of the clade. This example is consistent with empirical findings from outbreeding species with large effective population sizes (like fire ants) that gametic disequilibrium often drops to insignificant levels even over distances as small as 1 kb (Vasemägi and Primmer 2005). Thus, in the presence of typical recombination rates the apparently ancient, obligate association between *b*-like alleles and polygyny is only likely to have been preserved if *Gp-9* directly influences social organization. Of course, some mechanism that completely suppresses recombination between *Gp-9* and other linked genes of direct effect over evolutionary time could exist, as discussed below.

Perhaps the most compelling evidence that *Gp-9* is directly involved in the expression of social organization is the inference that selection has acted both to drive the molecular divergence of *b*-like alleles from the ancestral *B*-like alleles and to maintain these derived alleles in all socially polymorphic South American fire ant species. The evident elevated rate of non-synonymous substitutions in the stem lineage of the *b*-like allele clade implicates positive selection as having acted specifically on *Gp-9* to propel the adaptive molecular evolution of the *b*-like lineage in the context of the origin of polygyny (Krieger and Ross 2002, 2005). Moreover, the maintenance of the *b*-like clade as a trans-species polymorphism through multiple speciation events implicates balancing selection in protecting these alleles from loss through drift (Vekemans and Slatkin 1994; Klein et al. 1998).

Involvement of Other Genes in the Regulation of Social Organization

Several arguments also can be made to support the view that *Gp-9* plays no direct part in control of social organization or, at best, is but one of many genes involved in this role. First, the mere idea that a single gene could have a large effect on a multifaceted, highly complex social phenotype is anathema to many biologists, as appears to be reflected in Wilson's (1975) view that "… social behavior comprises the set of phenotypes farthest removed from DNA." However, an apparent precedent of such a major effect of variation at a single gene on complex social behavior has been reported in voles (e.g., Lim et al. 2004, 2005; Nair and Young 2006; but see Fink et al. 2006).

The complex nature of social organization in fire ants is evident from several perspectives. The terms monogyny and polygyny simply refer to the occurrence of single or multiple reproductive queens in a colony. However, queen number is a group attribute that emerges from the collective actions of hundreds of genetically non-identical workers. Moreover, the terms monogyny and polygyny actually connote suites of diagnostic traits in addition to colony queen number, including differences in individual size, differences in dispersal and reproductive strategies, and differences in the structure of colonies and populations (Ross and Keller 1995; Tschinkel 2006). Remarkably, *Gp-9* genotype has been shown to be associated with many of these diagnostic traits, the sheer diversity of which would seem to argue against variation at any single gene being entirely responsible. On the other hand, individual phenotypes of highly social organisms depend on interactions with colony-mates, so that what appear to be independent traits may often be linked by virtue of common communication mechanisms. In the case of *Gp-9*, the product is an OBP that may function in chemical communication. If the variant proteins

function upstream in behavioral interactions requiring reciprocal production and perception of chemical signals, then downstream consequences of altered behaviors could be manifested in diverse physiological and behavioral traits. As a hypothetical example, if queens with altered chemosensory abilities due to their *Gp-9* genotype do not appropriately recognize and respond to workers attempting to feed them, this could lead to their loss of nourishment, failure to sequester fat during maturation, and inability to sustain flight sufficient for long-range dispersal.

One observation that would clearly seem to implicate involvement of other genes is the association between *Gp-9* genotype and weight in adult workers and males (Goodisman et al. 1999). This association very likely results from genotype-specific differences in larval developmental programs rather than differential accumulation of fat or other materials during adult maturation (Porter and Tschinkel 1985; Tschinkel 1993), yet *Gp-9* appears not to be expressed until the adult stage in any caste (Ross 1997; Liu and Zhang 2004). It is difficult to imagine any scenario for the involvement of *Gp-9* in the expression of these traits consistent with a complete absence of its product in larvae. This fact alone suggests that there are other, as yet unknown, genes in complete gametic disequilibrium with *Gp-9* that influence expression of at least some traits comprising the contrasting social syndromes. Indeed, theoretical expectations are for such genes to be recruited into the genomic segment containing both the gene(s) determining social form (marked by *Gp-9*) and factors suppressing recombination (Hey 1998; Mescher 2001), such that inheritance of this entire region becomes Mendelian. In this view, *Gp-9* is embedded in an epistatic network of genes that are resistant to being broken up and that regulate various aspects of the social syndrome (i.e., a coadapted gene complex or

supergene; Clarke and Sheppard 1960; Palopoli and Wu 1996; Munté et al. 2005). Again, such resistance to recombination presumably is achieved by some chromosomal feature such as an inversion or other mechanism of crossing-over suppression (Mathiopoulos and Lanzaro 1995; Kelly 2000; Schimenti 2000). The existence of such a non-recombining unit is at least minimally supported by the absence of evidence for historical recombination within a 2200 bp region including *Gp-9* (Krieger and Ross 2005).

Future research on Gp-9 *as a candidate gene*

Given the lack of specific evidence for involvement of genes other than *Gp-9* in regulating social organization, combined with at least minimally compelling evidence for a direct role of *Gp-9*, we advocate studies directed at this gene and the physiological and behavioral pathways in which it functions as the most promising avenues of future investigation. Such an approach is, of course, subject to change pending concrete data disproving direct involvement of *Gp-9* and must, in any case, be augmented by aggressive exploration for other candidate genes.

The prime obstacle to understanding the involvement of *Gp-9* in social regulation is the lack of knowledge of its functional role in individual ants and the colony as a whole. Thus, an important research priority is to describe the normal workings of the protein product at the biochemical, physiological, and behavioral levels, from discovery of details of the protein structure, to identification of its ligand(s), to determination of the effects of sequence variation on binding properties, to determination of phenotypically downstream effects on individual physiology and social interactions. Excessive

extrapolation from other OBPs seems risky, because in many respects GP-9 appears not to be typical of the "gold standard" OBPs about which most is known (Vogt 2005).

Although nucleotide and amino acid sequences have been determined for *Gp-9* alleles from many *Solenopsis* species (Krieger and Ross 2002, 2005; Gotzek et al. 2007), the secondary, tertiary, and quaternary protein structures have not been solved. Inferential information has come from structure prediction modeling based on other insect OBPs with solved structures (Krieger 2005; Krieger and Ross 2005), an approach that may be robust at a general level given the conservation of basic structural features among even highly divergent OBPs (Leal 2003; Picimbon 2003; Vogt 2003, 2005). As an example, the diagnostic Ile^{139} residue of all *b*-like alleles maps to the binding pocket in the predicted GP-9 structure; thus, this substitution may have changed the ligand-binding properties, which may in turn have led to the individual phenotypic differences associated with possession of a *b*-like allele (Krieger and Ross 2005). As another example, the charge-changing amino acid replacement characterizing the *b* allele of *S. invicta* (Lys151Glu) maps to a position in the C-terminus that, by inference, may play a role in ligand binding and/or unloading or may participate in dimer formation (Krieger 2005). Thus, plausible mechanistic explanations of how this radical substitution disrupts GP-9 function and causes the recessive deleterious effects of the *b* allele were formulated and tested following structure prediction analysis (Hallar et al. 2006). As useful as this approach may be, analysis of the solved structures of GP-9 variants could be additionally informative, for instance by directly linking structural to binding properties of the *b*-like and *B*-like protein forms. Such analyses may also reveal whether GP-9 forms biologically active dimers (a matter of controversy with regard to other insect OBPs;

Sandler et al. 2000; Leal 2000; Honson et al. 2003) and, if so, if all types of homodimers and heterodimers are equally likely to form and are equally active.

Of equal importance for understanding GP-9 function will be information on its natural ligand(s) and binding properties. Based on information from other insect OBPs, it is likely that GP-9 binds small, hydrophobic molecules (but see Rivière et al. 2003). High specificity for a single ligand apparently is not characteristic of OBPs implicated in chemoreception, so discrimination among chemostimulants may be facilitated at the molecular level by differential binding of related ligands (Picimbon 2003; Rivière et al. 2003; Vogt 2005). By analogy, it is possible that GP-9 transports an array of related compounds in varying contexts and so may serve multiple functions, perhaps explaining the diverse phenotypic effects in the different genotypes. A most relevant issue here is whether GP-9 normally binds compounds that act as pheromones or pheromone precursors in fire ants.

Expression studies revealing detailed caste-, tissue-, and age-specific expression patterns also will be required to help unravel GP-9 function. Liu and Zhang (2004) described general patterns of expression in the different life stages and castes, which seem to reflect relative protein abundances (Ross 1997), but complete information is needed on the adult tissues in which expression occurs and the changes in expression as individuals age or change reproductive status (queens). GP-9 protein routinely is extracted from whole-thorax or head homogenates of adult females and it is abundant in thoracic hemolymph (D. Gotzek and K. Ross, unpubl.), suggesting that it may circulate throughout the hemocoel. This simple information alone is of value in suggesting that

Gp-9 does not encode an OBP that functions exclusively in signal transduction in chemosensilla.

Ultimately, experiments involving genetic transformation by means of techniques such as precise allele substitution may be needed to definitively establish a causal connection between *Gp-9* variation and expression of social organization in fire ants (Glazier et al. 2002; Philips 2005; Wright and Gaut 2005; Greenberg and Wu 2006; also Lim et al. 2004). Such experiments allowing replacement of one natural allele by another while holding the genetic background constant may pose no special difficulties beyond those associated with any non-model organism (which are still substantial) for analyses of form-diagnostic individual traits. On the other hand, they will be especially challenging when applied to colony queen number, because this group-level social trait represents the collective outcome of decisions made by many individual workers. An obvious experiment would be the transgenic equivalent of the studies of Ross and Keller (2002) and Gotzek and Ross (2006), in which workers of a monogyne colony have a $Gp\text{-}9^{b}$ transgene substituted for one of their *B* alleles; however, such genetic transformation will need to be achieved for a sufficiently large number of workers to make up 10% or more of the population of each test colony. Colonies with fewer than several thousand workers may not behave "normally" in the context of such experiments.

Even were transformation experiments successful in implicating a role for *Gp-9* in control of social organization, a much broader integrative approach is required to fully establish the mechanistic, causal connections between molecular allelic variation and social organization (see Feder and Watt 1992; Watt and Dean 2000; Glazier et al. 2002; Vasemägi and Primmer 2005). Moreover, such a broad approach can yield a wealth of

complementary results that, by themselves, can compellingly establish causality in non-model organisms (Glazier et al. 2002; Vasemägi and Primmer 2005). The broad objective of this approach as applied to GP-9 will be to discern the position and role of the protein in specific metabolic or signaling pathways, and to learn how variation in the protein affects the behavior of the pathway(s) and, therefore, individual and colony-level phenotypes. By necessity, this approach focuses on determination of pathway architecture, including identification of other genes contributing products to the pathway. A promising point of entry to characterizing this architecture is through microarray gene expression analyses, because clusters of genes that are coordinately expressed with *Gp-9* may participate in the same pathway (Qu and Xu 2006; Whitehead and Crawford 2006). In fact, some coordinately expressed genes serving related pathway functions may reside in close genomic proximity to *Gp-9* (Hurst et al. 2004; Ranz and Machado 2006), a possibility that raises the important issue of what genes are tightly linked to *Gp-9*.

Given that social organization in fire ants comprises expression of alterate syndromes of traits involving worker behavior, queen reproductive physiology, and colony breeding strategies (Ross and Keller 1995), and that the *b* allele behaves in some respects as a selfish genetic element with detrimental individual-level effects (Keller and Ross 1998), there are theoretical reasons to expect that *Gp-9* may be one component in a complex of tightly linked, coadapted genes that do not undergo recombination (Mescher 2001). Thus, an important goal of future research should be to identify genes in the genomic region surrounding *Gp-9* and to learn whether they constitute a large non-recombining complex (supergene). A top-down approach of constructing a high resolution linkage map to delimit the boundaries of the chromosomal region in tight linkage with *Gp-9*

should be combined with a bottom-up approach of obtaining sequence data for the region surrounding *Gp-9*. Such sequence data will permit identification and annotation of additional candidate genes tightly linked to *Gp-9* that may be part of a supergene, and may also contribute toward resolution of the architecture of pathways in which GP-9 functions. An important subsequent step will be to determine the extent of gametic disequilibrium among allelic variants of these linked genes in native *S. invicta*; if a supergene exists, it is expected not only that recombination is low or absent among its members (including *Gp-9*) but that particular combinations of alleles have been preserved over evolutionary time by selection (Mescher 2001). Long haplotypes including the *b*-like alleles of *Gp-9* and specific variants of neighboring genes are expected invariably to be associated with polygyny.

A final important research objective will be to identify and characterize other OBP genes in *S. invicta*. These genes typically occur as members of multigene families within insect species (Vogt 2005), so it is likely that paralogs of *Gp-9* occur in the fire ant genome. Presumably, some of these function in the manner traditionally conceived for insect OBPs, as molecular components of signal transduction in chemosensilla, while others may play supplementary roles in the expression of social behavior, perhaps even forming part of a supergene including *Gp-9*.

With extensive information available on the function of *Gp-9* and the genes with which it interacts, it should be possible to develop explicit models of the causal links between molecular variation, individual phenotype, and colony social organization. Although conceptually and technically daunting, this enterprise fulfills the core aim of the field of sociogenomics, to understand the molecular genetic bases of alternative traits

that provide the raw material for social evolution (Robinson 1999; Amdam et al. 2004; Robinson et al. 2005; Nedelcu and Michod 2006). In the spirit of advancing progress on this front, we present a new behavioral model for the regulation of colony queen number in fire ants that is consistent with currently available evidence. Although we assume that variation at *Gp-9* is the major causative factor in the expression of colony social organization, our general results rely only on the fact that a single Mendelian factor playing this role is marked with complete fidelity by *Gp-9*.

A PROXIMATE MODEL FOR REGULATION OF COLONY QUEEN NUMBER IN FIRE ANTS

Several models have been proposed to explain features of the behavioral regulation of colony queen number in *S. invicta* (Fletcher and Blum 1983a; Keller and Ross 1999; Crozier 2002; Keller and Parker 2002; Ross and Keller 2002; Krieger 2005). A common element of many of these is the assumption of an optimal level of "queen reproductive pheromone" that must be maintained in a colony. Fletcher and Blum (1983a) hypothesized that the ability to produce such a pheromone varies among egg-laying queens in association with their fecundity (degree of physogastry), and that single monogyne queens, which are highly physogastric, produce enough pheromone to reach the optimal colony level and prevent workers from adopting supernumerary queens. Single polygyne queens, being far less physogastric, individually produce insufficient amounts of pheromone to reach the colony optimum, leading Fletcher and Blum (1983a) to speculate that polygyne workers accept additional queens until their combined pheromone production reaches this optimum. Models developed since the discovery of

Gp-9 assume that production of this hypothetical pheromone is linked more closely to a queen's *Gp-9* genotype than her fecundity, with *Bb* queens producing less pheromone than *BB* queens (Crozier 2002; Keller and Parker 2002; Krieger 2005). Polygyne workers thus require multiple *b*-bearing reproductive queens to reach a colony's optimal pheromone level while monogyne workers require only a single *BB* queen.

Several of these models also allow for an entirely different mechanism of regulation of queen number that features differences in worker response to or perception of the queen pheromone rather than differences in queen pheromone production. Thus, Keller and Parker (2002) suggested that *Bb* (polygyne) workers may have a higher optimum for colony queen pheromone than *BB* (monogyne) workers (also suggested by Fletcher and Blum [1983a] without regard to *Gp-9* genotype), while Crozier (2002) and Krieger (2005) suggested that individual *Bb* workers may suffer some impairment in their perception of the queen pheromone. However, these proposals fail to explain why the *BB* workers in a polygyne colony fail to function as normal "monogyne" workers and attack supernumerary *Bb* queens, even when they constitute as much as 90-95% of the colony's worker force (Ross and Keller 2002; Gotzek and Ross 2006). A general model of queen regulation must explain why polygyne workers of all genotypes are complicit in tolerating multiple *Bb* reproductive queens.

The model presented below expands on these previous behavioral models by positing two different sets of individual phenotypic traits in workers and queens that are influenced by their *Gp-9* genotypes and that elicit worker behaviors that in aggregate determine colony queen number (see also Keller and Ross 1998). The two trait sets constitute distinct components of the model that are presented as separate hypotheses.

Which queens are accepted – the smelly worker hypothesis

We postulate that workers and queens carrying a *b* allele produce a novel chemical label (odor) distinct from that produced by individuals lacking the allele (see also Keller and Ross 1999; Ross and Keller 2002). Workers encountering *b*-bearing individuals form a specific exclusionary template (memory) that they use as a reference when making decisions about whether to tolerate or attack incipient or actual reproductive queens attempting to gain acceptance as egg-layers. Workers in colonies with *b*-bearing individuals accept only *b*-bearing reproductive queens because the queens' labels are compatible with the workers' b^+ templates; conversely, workers in colonies with only *BB* individuals accept only *BB* queens, whose labels are compatible with these workers' b^- templates. This hypothesis is an improvement over others in the sense that it successfully explains how *b*-bearing queens can be tolerated when as few as 5-10% of workers in a colony bear the allele (presumably, worker contacts are sufficiently frequent that all nestmates contact even rare *b*-bearing workers within a brief period). We note that the hypothesis does not address the issue of whether any specific queen is actually added to the resident queen pool, only whether a queen is compatible with a colony's worker force and thus eligible for acceptance (the second hypothesis deals with queen addition). We note further that the hypothesized discrimination system likely differs from, but may interact with, another system ordinarily used in queen (nestmate) discrimination in monogyne colonies (see below; also Vander Meer and Morel 1998; Tschinkel 2006).

There are a number of plausible mechanisms by which an individual's *Gp-9* genotype could affect production of the hypothesized odor label. For instance, if GP-9 functions to transport one component of a "pheromone blend" from its site of production to relevant

organs of activation or release, and the *b* allele product is non-functional in this role (i.e., *b* is a null allele; Krieger 2005), then *BB* individuals would produce a label dominated by the component carried by GP-9 and *b*-bearing individuals would produce a label with far less of it. In this view, the label differs qualitatively in its composition among the genotypes. Alternatively, if GP-9 has some general regulatory metabolic function, it may influence individual label production indirectly by influencing flux through relevant biosynthetic pathways. Here, the compositional blend of the *Gp-9* label is similar among all individuals, but the label differs quantitatively among the genotypes.

There is reason to believe that this hypothesized label is cuticular in origin and that behavioral interactions are not required for its perception. Adding freshly killed adult *Bb* workers to monogyne colonies can induce conversion to polygyny (D. Gotzek and K. Ross, unpubl.). Also, significant aggression can be elicited in polygyne workers by rubbing nestmate workers against the cuticle of *BB* queens, whereas a similar reaction is not observed if *Bb* queens are used (Keller and Ross 1998). Evidently, the label is not ordinarily shared among colony members, because virgin polygyne *BB* queens would not be expected to meet escalating aggression as they mature in their natal colony if it were.

Two different types of *b*-like allele are associated with polygyny in native *S. invicta* (Krieger and Ross 2002, 2005). Designated as *b′* and *b*, these alleles rarely co-occur within single polygyne colonies (Hallar et al. 2006). It is possible that they induce production of labels that are sufficiently distinct as to be incompatible in terms of worker acceptance of queens; that is, workers in polygyne colonies with one of the *b*-like alleles form templates not sufficiently inclusive to induce tolerance of queens with the other *b*-like allele. Nonetheless, the two *b*-like alleles appear to induce production of similar

queen-caste signals postulated to affect the number of queens tolerated in a colony (see below).

How many queens are accepted – the wimpy queen hypothesis

We further postulate that *S. invicta* females produce different amounts or blends of a pheromone that signals their actual or potential reproductive status to workers (see also Fletcher and Blum 1981, 1983a,b; Willer and Fletcher 1986; Keller and Ross 1999; Vargo 1999). Production of this hypothetical "queen-caste signal" is predicted to vary quantitatively or qualitatively with caste, age of pre-reproductive queens, and reproductive potential of mature queens, with the latter trait associated with *Gp-9* genotype (Figure 2.3). Young pre-reproductive queens yet to attain sexual maturity, as well as workers of any age, have low or no potential fertility and are hypothesized to create pheromone profiles that signal this to workers. Mature pre-reproductive queens ready to initiate oogenesis, as well as egg-laying queens, vary in their fertility potential depending on their *Gp-9* genotype, as signified by differences in accumulation of fat reserves during maturation, time to onset of oogenesis, initial fecundity, and metabolic efficiency of egg production (Keller and Ross 1993b, 1999; Ross and Keller 1998; Mescher 2001; DeHeer 2002; Tschinkel 2006). This genotype-associated variation in reproductive potential is hypothesized to be communicated as well to workers via the queen-caste signal.

Specifically, we propose that possession of a *b*-like allele pleiotropically "down-regulates" expression of these fertility components as well as the queen-caste signal. Whereas pre-reproductive *BB* queens in polygyne colonies begin producing a strong

queen-caste signal once they approach sexual maturity, queens with allele *b* never produce this type or level of pheromone, even after having mated and commenced laying eggs (Figure 2.3). Once polygyne *BB* queens begin producing a signal by which workers recognize them as potential reproductive females, they are attacked because, in accordance with our first hypothesis, they lack the appropriate *Gp-9* odor label. Queens bearing *b*-like alleles are not perceived as being fully fertile by workers (they are pheromonal wimps), even after they become egg-layers, so workers always tolerate them as if they were immature pre-reproductives, leading to polygyny. In essence, polygyne colonies can be viewed as being pheromonally permanently queenless. The exact number of supernumerary egg-laying queens tolerated by a polygyne colony probably depends on a number of non-genetic factors such as worker/queen ratios or frequencies of attempted infiltrations by newly mated queens.

Additional considerations for the model

It is likely that the *Gp-9* label proposed in the first hypothesis is superimposed on another system that also functions in queen (nestmate) discrimination in the constant *Gp-9* background of the monogyne form of *S. invicta*. Workers in monogyne colonies clearly imprint on their mother queen (presumably via chemical labels) and will kill any substitute reproductive queen presented to them, regardless of her *Gp-9* genotype. However, this discrimination can be extinguished by removing the mother queen from her colony for a period of 2-4 days, at which point a single substitute queen bearing genotype *BB* becomes acceptable (Fletcher 1986; also DeHeer and Tschinkel 1998; Ross

and Keller 1998). Such imprinting apparently also can occur to some extent with polygyne workers held with single queens (Ross and Keller 1998).

Our model does not directly explain why numerous *BB* alates are tolerated in queenright monogyne colonies even after they attain sexual maturity. One possible reason is that workers with alternate *Gp-9* genotypes differ in their signal thresholds at which they perceive a maturing queen as a future reproductive, with *Bb* workers having a lower threshold than *BB* workers. In this view, the hatched band in Figure 2.3 depicting the collective worker threshold in a polygyne colony is a composite of a lower band for *Bb* workers and higher band for *BB* workers (the band in a monogyne colony would be narrowed and elevated). Queen-caste signals of *b*-bearing queens remain below the threshold of both *BB* and *Bb* workers in polygyne colonies, while mature *BB* queens' signals cross the collective threshold of *Bb* workers, inducing their execution by such workers (consistent with the findings of Keller and Ross 1998). Monogyne *BB* workers, with their higher thresholds, fail to perceive nestmate *BB* alates as potential reproductives, thus explaining the absence of aggression toward them. An alternative possibility is that the *BB* reproductive queens in monogyne colonies suppress the production of queen-caste signal in maturing alates, while the *Bb* queens in polygyne colonies are incapable of doing so. In this view, the purple ribbon in Figure 2.3 is depressed in a monogyne colony, so that few *BB* queens cross the collective worker recognition threshold.

Finally, our model can with little accommodation explain why a small minority of mature *BB* queens in polygyne colonies actually survive to take part in mating flights (DeHeer et al. 1999; Goodisman et al. 2000a; DeHeer 2002). These individuals are

hypothesized to be relatively delayed in developing the typical *BB* queen-caste signal; thus, they attain sexual maturity without crossing the collective worker threshold of recognition of reproductives and so escape aggression before departing the nest on mating flights (see Figure 2.3). Fletcher and Blum (1983b) describe apparently similar uncoupling of the queen-caste signal and reproductive development in overwintered virgin queens of monogyne *S. invicta*.

Several questions must be answered to test and extend our model. Can odor cues extracted from *b*-bearing workers cause social conversion when presented to monogyne colonies? How do workers form exclusionary templates on the basis of brief and infrequent contacts with *b*-bearing nestmates? Do definable predictors of "potential reproductive status," such as the numbers of ovarioles comprising the ovaries, differ between polygyne queens of different *Gp-9* genotypes? Are previously identified compounds from the queen poison sac thought to be components of "queen recognition pheromones" (Rocca et al. 1983a, b) ligands of GP-9?

CONCLUSIONS

Despite considerable recent progress in our understanding of the genetic regulation of colony social organization in fire ants, much work remains. It is clear that expression of social organization in *S. invicta* is regulated by a single Mendelian factor and that this genetic architecture has been conserved over evolutionary time in the South American fire ants. Selection appears to have been involved both in driving the molecular evolution and in preserving the variation of one candidate gene comprising this factor,

Gp-9. This gene is likely to be functionally important and, given the identity of its product as an odorant binding protein, plausible scenarios for its direct involvement in the regulation of social organization can be envisaged. However, our current lack of knowledge of the functional role of the protein in individual ants, combined with our ignorance of the precise behavioral interactions by which colony queen number is regulated, has hindered progress toward a fuller understanding of the importance of *Gp-9*. Future studies aimed at identifying the biochemical pathways in which the gene product functions, at discerning the involvement of other pathway genes or genes in linkage disequilibrium with *Gp-9*, and at revealing details of the chemical communication involved in queen acceptance or rejection by workers will go far toward painting a complete picture of the genetic architecture underlying regulation of colony queen number in fire ants. In so doing, such work will illuminate the major molecular, individual, and colony-level changes that transpired during the evolution of a key social adaptation.

ACKNOWLEDGMENTS

We thank Laurent Keller, Jürgen Liebig, Emily Matthews, Yannick Wurm, and Richard Vogt for comments on a previous version of this paper.

REFERENCES

Åkerstrom B, Flower DR, Salier J-P. 2000. Lipocalins: unity in diversity. Biochimica et Biophysica Acta 1482: 1-8.

Amdam GV, Norberg K, Fondrk MK, Page RE. 2004. Reproductive ground plan may mediate colony-level selection effects on individual foraging behavior in honey bees. Proc Natl Acad Sci USA 101: 11350-11355.

Baker BS, Taylor BJ, Hall JC. 2001. Are complex behaviors specified by dedicated regulatory genes? Reasoning from *Drosophila*. Cell 105: 13-24.

Bork P, Dandekar T, Diaz-Lazcoz Y, Eisenhaber F, Huynen M, Yuan Y. 1998. Predicting function: from genes to genomes and back. J Mol Biol 283: 707-725.

Bos DH, Waldman B. 2006. Evolution by recombination and transspecies polymorphism in the MHC Class I gene of *Xenopus laevis*. Mol Biol Evol 23: 137-143.

Bourke AFG, Franks NR. 1995. *Social Evolution in Ants*. Princeton: Princeton University Press.

Brent CS, Vargo EL. 2003. Changes in juvenile hormone biosynthetic rate and whole body content in maturing virgin queens of *Solenopsis invicta*. J Insect Physiol 49: 967-974.

Cahan S, Helms KR, Rissing SW. 1998. An abrupt transition in colony founding behaviour in the ant *Messor pergandei*. Anim Behav 55: 1583-1594.

Calvello M, Guerra N, Brandazza A, D'Ambrosio C, Scaloni A, Dani FR, Turillazzi S, Pelosi P. 2003. Soluble proteins of chemical communication in the social wasp *Polistes dominulus*. Cell Mol Life Sci 60: 1933-1943.

Clarke CA, Sheppard PM. 1960. Super-genes and mimicry. Heredity 14: 175–185.

Costa JT, Ross KG. 2003. Fitness effects of group merging in a social insect. Proc R Soc Lond B 270: 1697-1702.

Crozier RH. 2002. Pheromones and the single queen. Nature Genet 30: 4-5.

de Bono M, Bargmann CI. 1998. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. Cell 94: 679–689.

DeHeer CJ. 2002. A comparison of the colony-founding potential of queens from single- and multiple-queen colonies of the fire ant *Solenopsis invicta*. Anim Behav 64: 655-661.

DeHeer CJ Tschinkel WR. 1998. The success of alternative reproductive tactics in monogyne populations of the ant *Solenopsis invicta*: significance for transition in social organization. Behav Ecol 9:130-135.

DeHeer CJ, Goodisman MAD, Ross KG. 1999. Queen dispersal strategies in the multiple-queen form of the fire ant *Solenopsis invicta*. Am Nat 153:660-675.

Elmes GW, Keller L. 1993. Distribution and ecology of queen number in ants of the genus *Myrmica*. In Keller L. (ed.) *Queen Number and Sociality in Insects*. Oxford: Oxford University Press. Pp: 294-307

Feder ME, Watt WB. 1992. Functional biology of adaptation. In Berry RJ, Crawford TJ, Hewitt GM. (eds.) *Genes in Ecology*. Oxford: Blackwell Science. Pp: 365-392.

Fewell JH, Bertram SM. 2002. Evidence for genetic variation in worker task performance by African and European honey bees. Behav Ecol Sociobiol 52: 318-325.

Fink S, Excoffier L, Heckel G. 2006. Mammalian monogamy is not controlled by a single gene. Proc Natl Acad Sci USA 103: 10956–10960.

Fletcher DJC. 1986. Triple action of queen pheromones in the regulation of reproduction in fire ant (*Solenopsis invicta*) colonies. In Porchet M, Andries J-C, Dhainaut A. (eds.) *Advances in Invertebrate Reproduction 4*. London: Elsevier Academic. Pp: 305-316.

Fletcher DJC, Blum MS. 1981. Pheromonal control of dealation and oogenesis in virgin queen fire ants. Science 212: 73-75.

Fletcher DJC, Blum MS. 1983a. Regulation of queen number by workers in colonies of social insects. Science 219: 312-314.

Fletcher DJC, Blum MS. 1983b. The inhibitory pheromone of queen fire ants: effects of disinhibition on dealation and oviposition by virgin queens. J Comp Physiol A 153: 467-475.

Fritz GN, Vander Meer RK, Preston CA. 2006. Selective male mortality in the red imported fire ant, *Solenopsis invicta*. Genetics 173: 207-213.

Galindo K, Smith DP. 2001. A large family of divergent *Drosophila* odorant-binding proteins expressed in gustatory and olfactory sensilla. Genetics 159: 1059-1072.

Glancey BM, Lofgren CS. 1988. Adoption of newly-mated queens: a mechanism for proliferation and perpetuation of polygynous red imported fire ants, *Solenopsis invicta* Buren. Florida Entomol 71: 581-587.

Glazier AM, Nadeau JH, Aitman TJ. 2002. Finding genes that underlie complex traits. Science 298: 2345-2349.

Goodisman MAD, Ross KG. 1999. Queen recruitment in a multiple-queen population of the fire ant *Solenopsis invicta*. Behav Ecol 10: 428-435.

Goodisman MAD, Mack PD, Pearse DE, Ross KG. 1999. Effects of a single gene on worker and male body mass in the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae). Ann Entomol Soc Am 92: 563-570.

Goodisman MAD, DeHeer CJ, Ross KG. 2000a. Unusual behavior of polygyne fire ant queens on nuptial flights. J Insect Behav 13: 455-468.

Goodisman MAD, Ross KG, Asmussen MA. 2000b. A formal assessment of gene flow and selection in the fire ant *Solenopsis invicta*. Evolution 54: 606-616.

Gotzek D, Ross KG. 2006. Experimental conversion of colony social organization in fire ants (*Solenopsis invicta*): worker genotype manipulation in the absence of queen effects. J Insect Behav (submitted).

Gotzek D, Shoemaker DD, Ross KG. 2007. Molecular variation at a candidate gene regulating complex social behavior in the fire ant *Solenopsis invicta*. Mol Biol Evol (submitted).

Graham LS, Tang W, Baust JG, Liou YC, Reid TS, Davies PL. 2001. Characterization and cloning of a *Tenebrio molitor* hemolymph protein with sequence similarity to insect odorant-binding proteins. Insect Biochem Mol Biol 31: 691-702.

Greenberg AJ, Wu C-I. 2006. Molecular genetics of natural populations. Mol Biol Evol 23: 883-886.

Greenberg L, Fletcher DJC, Vinson SB. 1985. Differences in worker size and mound distribution in monogynous and polygynous colonies of the fire ant *Solenopsis invicta* Buren. J Kanses Entomol Soc 58: 9-18.

Guntur KVP, Velasquez D, Chadwell L, Carroll C, Weintraub S, Cassill JA, Renthal R. 2004. *Apolipophorin-III*-like protein expressed in the antenna of the red imported

fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae). Arch Insect Biochem Physiol 57: 101-110.

Hallar BL, Krieger MJB, Ross KG. 2006. Potential cause of lethality of an allele implicated in social evolution in fire ants. Genetica (in press).

Hedrick PW, Jain S, Holden L. 1978. Multilocus systems in evolution. Evol Biol 11: 101-182.

Hekmat-Scafe DS, Scafe CR, McKinney AJ, Tanouye MA. 2002. Genome-wide analysis of the odorant-binding protein gene family in *Drosophila melanogaster*. Genome Res 12: 1357-1369.

Herbers JM. 1993. Ecological determinants of queen number in ants. In Keller L. (ed.) *Queen Number and Sociality in Insects*. Oxford: Oxford University Press. Pp: 262-293.

Hey J. 1998. Selfish genes, pleiotropy and the origin of recombination. Genetics 149: 2089-2097.

Higgs PG, Attwood TK. 2005. *Bioinformatics and Molecular Evolution*. Oxford: Blackwell Publishing.

Honson N, Johnson MA, Oliver JE, Prestwich GD, Plettner E. 2003. Structure-activity studies with pheromone-binding proteins of the gypsy moth, *Lymantria dispar*. Chem Senses 28: 479-489.

Hunt GJ, Page RE, Fondrk MK, Dullum CJ. 1995. Major quantitative trait loci affecting honey bee foraging behavior. Genetics 141: 1537-1545.

Hurst LD, Pal C, Lercher MJ. 2004. The evolutionary dynamics of eukaryotic gene order. Nature Rev Genet 5: 299-310.

Ishida Y, Chiang V, Leal WS. 2002. Protein that makes sense in the Argentine ant. Naturwissenschaften. 89: 505-507.

Johnson RN, Oldroyd BP, Barron AB, Crozier RH. 2002. Genetic control of the honey bee (*Apis mellifera*) dance language: segregating dance forms in a backcrossed colony. J Hered 93: 170-173.

Julian GE, Fewell JH, Gadau J, Johnson RA, Larrabee D. 2002. Genetic determination of the queen caste in an ant hybrid zone. Proc Natl Acad Sci USA 99: 8157-8160.

Keller L, Ross KG. 1993a. Phenotypic plasticity and "cultural transmission" of alternative social organizations in the fire ant *Solenopsis invicta*. Behav Ecol Sociobiol 33: 121-129.

Keller L, Ross KG. 1993b. Phenotypic basis of reproductive success in a social insect: genetic and social determinants. Science 260: 1107-1110.

Keller L, Ross KG. 1998. Selfish genes: a green beard in the red fire ant. Nature 394: 573-575.

Keller L, Ross KG. 1999. Major gene effects on phenotype and fitness: the relative roles of *Pgm-3* and *Gp-9* in introduced populations of the fire ant *Solenopsis invicta*. J Evol Biol 21: 672-680.

Kelly JK. 2000. Epistasis, linkage, and balancing selection. In Wolf JB, Brodie ED, Wade MJ. (eds.) *Epistasis and the Evolutionary Process*. Oxford: Oxford University Press. Pp: 146-157.

Klein J, Sato A, Nagl S, O'hUigín C. 1998. Molecular trans-species polymorphism. Ann Rev Ecol Evol Syst 29: 1–21.

Krieger MJB. 2005. To *b* or not to *b*: a pheromone–binding protein regulates colony social organization in fire ants. BioEssays 27: 91-99.

Krieger MJB, Ross KG. 2002. Identification of a major gene regulating complex social behavior. Science 295: 328-332.

Krieger MJB, Ross KG. 2005. Molecular evolutionary analyses of the odorant-binding protein gene *Gp-9* in fire ants and other *Solenopsis* species. Mol Biol Evol 22: 2090-2103.

Langley CH, Lazzaro BP, Phillips W, Heikkinen E, Braverman JM. 2000. Linkage disequilibria and the site frequency spectra in the *su(s)* and *su(w^a)* regions of the *Drosophila melanogaster* X chromosome. Genetics 156: 1837-1852.

Leal WS. 2000. Duality monomer-dimer of the pheromone-binding protein from *Bombyx mori*. Biochem Biophys Red Comm 268: 521-529.

Leal W S. 2003. Proteins that make sense. In Blomquist GJ, Vogt RG. (eds.) *Insect pheromone biochemistry and molecular biology: the biosynthesis and detection of pheromones and plant volatiles*. London: Elsevier Academic. Pp: 447-476.

Leal WS. 2005. Pheromone Reception. Topics Curr Chem 240: 1-36.

Li J, Heinz KM. 2000. Genome complexity and organization in the red imported fire ant *Solenopsis invicta* Buren. Genome Res 75: 129-135.

Lim MM, Wang ZX, Olazabal DE, Ren XH, Terwilliger EF, Young LJ. 2004. Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. Nature 429: 754-757.

Lim MM, Wang ZX, Young LJ. 2005. The molecular basis for the evolution of monogamy in voles. Hormones Behav 48: 84-85.

Linksvayer TA, Wade MJ. 2005. The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: maternal effects, sib-social effects, and heterochrony. Quart Rev Biol 8: 317-336.

Liu N, Zhang L. 2004. *CYP4AB1*, *CYP4AB2*, and *Gp-9* gene overexpression associated with workers of the red imported fire ant, *Solenopsis invicta* Buren. Gene 327: 81-87.

Markin GP, Collins HL, Dillier JH. 1972. Colony founding by queens of the red imported fire ant, *Solenopsis invicta*. Ann Entomol Soc Am 65: 1053-1058.

Markin GP, Dillier JH, Collins HL. 1973. Growth and development of colonies of the red imported fire ant, Solenopsis invicta. Ann Entomol Soc Am 66: 803-808.

Mathiopoulos KD, Lanzaro GC. 1995. Distribution of genetic diversity in relation to chromosomal inversions in the malaria mosquito *Anopheles gambiae*. J Mol Evol 40: 578-584.

Mescher MC. 2001. *Levels of Selection and the Evolution of Social Organization*. PhD dissertation. Athens: University of Georgia.

Mescher MC, Ross KG, Shoemaker DD, Keller L, Krieger MJB. 2003. Distribution of the two social forms of the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae) in the native South American range. Ann Entomol Soc Am 96: 810-817.

Montague CE, Oldroyd BP. 1998. The evolution of worker sterility in honey bees: an investigation into a behavioral mutant causing failure of worker policing. Evolution 52: 1408-1415.

Moritz RFA. 1988. A reevaluation of the two-locus model for hygienic behavior in honeybees (*Apis mellifera* L.). J Hered 79: 257-262.

Moritz RFA, Hillesheim E. 1985. Inheritance of dominance in honeybees (*Apis mellifera capensis* Esch.). Behav Ecol Sociobiol 17: 87-89.

Munté, A., Rozas J, Aguadé M, Segarra C. 2005. Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *Drosophila subobscura*. Genetics 169:1573-1581.

Nachman MW. 2005. The genetic basis of adaptation: lessons from concealing coloration in pocket mice. Genetica 123: 125-136.

Nair HP, Young LJ. 2006. Vasopressin and pair-bond formation: Genes to brain to behavior. Physiology 21: 146-152.

Nagnan-Le Meillour P, Jacquin-Joly E. 2003. Biochemistry and diversity of insect odorant-binding proteins. In Blomquist GJ, Vogt RG. (eds.) *Insect pheromone biochemistry and molecular biology: the biosynthesis and detection of pheromones and plant volatiles*. London: Elsevier Academic. Pp: 509-537.

Nedelcu AM, Michod RE. 2006. The evolutionary origin of an altruistic gene. Mol Biol Evol 23: 1460-1464.

O'Donnell S. 1996. RAPD markers suggest genotypic effects on forager specialization in a eusocial wasp. Behav Ecol Sociobiol 38: 83-88.

Orr HA. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52: 935–949.

Orr HA. 2005. The genetic theory of adaptation: a brief history. Nature Rev Genet 6: 119-127.

Orr HA, Coyne JA. 1992. The genetics of adaptation: a reassessment. Am Nat 140: 725-742.

Ozaki M, Wada-Katsumata A, Fujikawa K, Iwasaki M, Yokohari F, Satoji Y, Nisimura T, Yamaoka R. 2005. Ant nestmate and non-nestmate discrimination by a chemosensory sensillum. Science 309: 311-314.

Paesen GC, Happ GM. 1995. The B proteins secreted by the tubular accessory sex glands of the male meal-worm beetle, *Tenebrio molitor*, have sequence similarity to moth phermonone-binding proteins. Insect Biochem Mol Biol 25: 401-408.

Page RE, Robinson GE. 1994. Reproductive competition in queenless honey bee colonies (*Apis mellifera* L.). Behav Ecol Sociobiol 35: 99-107.

Palopoli MF, Wu C-I. 1996. Rapid evolution of a coadapted gene complex: evidence from the segregation distorter (SD) system of meiotic drive in *Drosophila melanogaster*. Genetics 143: 1675–1688.

Paskewitz SM, Shi L. 2005. The hemolymph proteome of *Anopheles gambiae*. Insect Biochem Mol Biol 35: 815-824.

Pelosi P, Calvello M, Ban L 2005. Diversity of odorant proteins and chemosensory proteins in insects. Chem Senses 30: i291-i292.

Phillips P C. 2005. Testing hypotheses regarding the genetics of adaptation. Genetica 123: 15-24.

Picimbon J-F. 2003. Biochemistry and evolution of OBP and CSP proteins. In Blomquist GJ, Vogt RG. (eds.) *Insect pheromone biochemistry and molecular biology: the biosynthesis and detection of pheromones and plant volatiles*. London: Elsevier Academic. Pp: 539-566.

Pitts JT, McHugh JV, Ross KG. 2005. Cladistic analysis of the fire ants of the *Solenopsis saevissima* species-group (Hymenoptera: Formacidae). Zool Scripta 34: 493-505.

Plateaux-Quénu C, Plateaux L, Packer L. 2000. Population-typical behaviours are retained when eusocial and non-eusocial forms of *Evylaeus albipes* (F.) (Hymenoptera, Halictidae) are reared simultaneously in the laboratory. Insectes Soc 47: 263-270.

Plettner E. 2003. The peripheral pheromone olfactory system in insects: targets for species-selective insect control agents. In Blomquist GJ, Vogt RG. (eds.) *Insect pheromone biochemistry and molecular biology: the biosynthesis and detection of pheromones and plant volatiles*. London: Elsevier Academic. Pp: 477-507.

Porter SD. 1991. Origins of new queens in polygyne red imported fire ant colonies (Hymenoptera: Formicidae). J Entomol Sci 26: 474-478.

Porter SD. 1992. Frequency and distribution of polygyne fire ants (Hymenoptera: Formicidae) in Florida. Florida Entomol 75: 248-257.

Porter SD, Tschinkel WR. 1985. Fire ant polymorphism (Hymenoptera: Formicidae): factors affecting worker size. Ann Entomol Soc Am 78: 381-386.

Porter SD, Bhatkar A, Mulder R, Vinson SB, Clair DJ. 1991. Distribution and density of polygyne fire ants (Hymenoptera: Formacidae) in Texas. J Econ Entomol 54: 866-874.

Porter SD, Williams DF, Patterson RS, Fowler GH. 1997. Intercontinental differences in the abundance of *Solenopsis* fire ants (Hymenoptera: Formicidae): escape from natural enemies? Environ Entomol 26: 373-384.

Qu Y, Xu S. 2006. Quantitative trait associated microarray gene expression data analysis. Mol Biol Evol 23: 1558-1573.

Ranz JM, Machado CA. 2006. Uncovering evolutionary patterns of gene expression using microarrays. Trends Ecol Evol 21: 29-37.

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES. 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc Natl Acad Sci USA 98: 11479–11484.

Rivière S , Lartigue A, Quennedey B, Campanacci V, Farine J-P, Tegoni M, Cambillau C, Brossut R. 2003. A pheromone-binding protein from the cockroach *Leucophaea maderae*: cloning, expression and pheromone binding. Biochem J 371: 573–579.

Robinson GE. 1999. Integrative animal behaviour and sociogenomics. Trends Ecol Evol 14: 202-205.

Robinson GE, Grozinger CM, Whitfield CW. 2005. Sociogenomics: social life in molecular terms. Nature Rev Genet 6: 257-271.

Rocca JR, Tumlinson JH, Glancey GM, Lofgren CS. 1983a. The queen recognition pheromone of *Solenopsis invicta*, preparation of (E)-6-1-pentenyl)-2H-pyran-2-one. Tetrahedron Letters 24: 1889-1892.

Rocca JR, Tumlinson JH, Glancey GM, Lofgren CS. 1983b. Synthesis and stereochemistry of tetrahydro-3,5-dimethyl-6-1-methylbutyl)-2H-pyran-2-one, a component of the queen recognition pheromone of *Solenopsis invicta*. Tetrahedron Letters 24: 1839-1896.

Ross KG. 1986. Kin selection and the problem of sperm utilization in social insects. Nature 323: 798 – 800.

Ross KG. 1997. Multilocus evolution in fire ants: effects of selection, gene flow, and recombination. Genetics 145: 961-974.

Ross KG, Fletcher DJC. 1985. Comparative study of genetic and social structure in two forms of the fire ant, *Solenopsis invicta* (Hymenoptera: Formicidae). Behav Ecol Sociobiol 17: 349-356.

Ross KG, Keller L. 1995. Ecology and evolution of social organization: insights from fire ants and other highly eusocial insects. Ann Rev Ecol Evol Syst 26: 631-656.

Ross KG, Keller L. 1998. Genetic control of social organization in an ant. Proc Natl Acad Sci USA 95: 14232-14237.

Ross KG, Keller L. 2002. Experimental conversion of colony social organization by manipulation of worker genotype composition in fire ants (*Solenopsis invicta*). Behav Ecol Sociobiol 51: 287-295.

Ross KG, Shoemaker DD. 1993. An unusual pattern of gene flow between the two social forms of the fire ant *Solenopsis invicta*. Evolution 47: 1595-1605.

Ross KG, Vargo EL, Keller L. 1996a. Social evolution in a new environment: the case of introduced fire ants. Proc Natl Acad Sci USA 93: 3021-3025.

Ross KG, Vargo EL, Keller L. 1996b. Simple genetic basis for an important social trait in the fire ant *Solenopsis invicta*. Evolution 50: 2387-2399.

Ross KG, Krieger MJB, Shoemaker DD, Vargo EL, Keller L. 1997. Hierarchical analysis of genetic structure in native fire ant populations: results from three classes of molecular markers. Genetics 147: 643-655.

Ross KG, Shoemaker DD, Krieger MJB, DeHeer CJ, Keller L. 1999. Assessing genetic structure with multiple classes of molecular markers: a case study involving the introduced fire ant *Solenopsis invicta*. Mol Biol Evol 16: 525-543.

Ross KG, Krieger MJB, Shoemaker DD. 2003. Alternative genetic foundations for a key social polymorphism in fire ants. Genetics 165: 1853-1867.

Ross KG, Krieger MJB, Keller L, Shoemaker DD. 2006. Genetic variation and structure in native populations of the fire ant *Solenopsis invicta*: evolutionary and demographic implications. Biol J Linn Soc (in press).

Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D, Lander ES. 2006. Positive natural selection in the human lineage. Science 312: 1614-1620.

Sandler BH, Nikinova L, Leal WS, Clardy J. 2000. Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. Chem Biol 7: 143-151.

Schimenti J. 2000. Segregation distortion of mouse *t* haplotypes: the molecular basis emerges. Trends Genet 16: 240-243.

Schwander T, Helms Cahan S, Keller L. 2006. Genetic caste determination in *Pogonomyrmex* harvester ants imposes costs during colony founding. J Evol Biol 19: 402-409.

Shoemaker DD, Ross KG. 1996. Effects of social organization on gene flow in the fire ant *Solenopsis invicta*. Nature 383: 613-616.

Shoemaker DD, DeHeer CJ, Krieger MJB, Ross KG. 2006. Population genetics of the invasive fire ant *Solenopsis invicta* (Hymenoptera: Formicidae) in the United States. Ann Entomol Soc Am 99: 1213-1233.

Snyder LE. 1992. The genetics of social behavior in a polygynous ant. Naturwissenschaften 79: 525-527.

Storz JF. 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. Mol Ecol 14: 671-688.

Tian H, Vinson SB, Coates CJ. 2004. Differential gene expression between alate and dealate queens in the red imported fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae). Insect Biochem Mol Biol 34: 937-949.

Tschinkel W R. 1993. Sociometry and sociogenesis of colonies of the fire ant *Solenopsis invicta* during one annual cycle. Ecological Monographs 64: 425-457.

Tschinkel W R. 2006. *The Fire Ants*. Cambridge: Belknap Press.

Vander Meer RK, Morel L. 1998. Nestmate recognition in ants. In Vander Meer RK, Breed M, Espelie KE, Winston ML. (eds.) *Pheromone communication in social insects: ants, wasps, bees, and termites*. Boulder: Westview. Pp: 79-103.

Vander Meer RK, Alonso LE. 2002. Queen primer pheromone affects conspecific fire ant (*Solenopsis invicta*) aggression. Behav Ecol Sociobiol 51: 122-130.

Vander Meer RK, Morel L, Lofgren CS. 1992. A comparison of queen oviposition rates from monogyne and polygyne fire ant, *Solenopsis invicta*, colonies. Physiol Entomol 17: 384-390.

Vargo EL. 1999. Reproductive development and ontogeny of queen pheromone production in the fire ant *Solenopsis invicta*. Physiol Entomol 24: 370-376.

Vargo EL, Hulsey CD. 2000. Multiple glandular origins of queen pheromones in the fire ant *Solenopsis invicta*. J Insect Physiol 46: 1151-1159.

Vargo EL, Porter SD. 1989. Colony reproduction by budding in the polygyne form of *Solenopsis invicta* (Hymenoptera: Formicidae). Ann Entomol Soc Am 82: 307-313.

Vasemägi A, Primmer CR. 2005. Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. Mol Ecol 14: 3623-3642.

Vekemans X, Slatkin M. 1994. Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics 137: 1157–1165.

Vogt RG. 2003. Biochemical diversity of odor detection: OBPs, ODEs and SNMPs. In Blomquist GJ, Vogt RG. (eds.) *Insect pheromone biochemistry and molecular biology: the biosynthesis and detection of pheromones and plant volatiles.* London: Elsevier Academic. Pp: 391-445.

Vogt RG. 2005. Molecular basis of pheromone detection in insects. In Gilbert LI, Iatro K, Gill S. (eds.) *Endocrinology vol. 3. Comprehensive Insect Physiology, Biochemistry, Pharmacology and Molecular Biology*. London: Elsevier. Pp: 753-804.

Vogt RG, Callahan FE, Rogers ME, Dickens JC. 1999. Odorant binding protein diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug *Lygus lineolaris* (Hemiptera, Heteroptera). Chem Senses 24: 481-495.

Wanner KW, Isman MB, Feng Q, Plettner E, Theilmann DA. 2005. Developmental expression patterns of four chemosensory protein genes from the Eastern spruce budworm, *Chroistoneura fumiferana*. Insect Mol Biol 14: 289-300.

Watt WB, Dean AM. 2000. Molecular-functional studies of adaptive genetic variation in prokaryotes and eukaryotes. Ann Rev Genet 34: 593-622.

Whitehead A, Crawford DL. 2006. Variation within and among species in gene expression: raw material for evolution. Mol Ecol 15: 1197-1211.

Willer DE, Fletcher DJC. 1986. Differences in inhibitory capability among queens of the ant *Solenopsis invicta*. Physiol Entomol 11: 475-482.

Wilson EO. 1975. *Sociobiology: The New Synthesis*. Cambridge: Harvard University Press.

Wright SI, Gaut BS. 2005. Molecular population genetics and the search for adaptive evolution in plants. Mol Biol Evol 22: 506-519.

Xu P. 2005. A *Drosophila* OBP required for pheromone signaling. Science 310: 708-709.

BOX 1

Current consensus is that protein families be recognized on the basis of the evolutionary relationships of the molecules (e.g., Vogt 2003, 2005; Higgs and Attwood 2005) rather than other features such as their structures or presumed functions (e.g., Leal 2003, 2005). Thus, classifications of proteins serve as guides to the common evolutionary origins and divergence of the genes encoding them. Although similarity of function can be postulated from common membership in a family (and, indeed, is important for gene annotation and candidate gene discovery; Bork et al. 1998; Higgs and Attwood 2005), assignment of function necessarily remains provisional pending specific genetic, biochemical, and physiological experiments.

FIGURE LEGENDS

Figure 2.1.—ALLELE FREQUENCY DIFFERENTIATION BETWEEN SYMPATRIC MONOGYNE AND POLYGYNE *S. INVICTA* POPULATIONS FROM THE INTRODUCED AND NATIVE RANGES AS ASSESSED BY ESTIMATES OF F_{ST}

Estimates are shown for *Gp-9* and the linked gene *Pgm-3*, as well as for other nuclear genes of various classes (microsatellites and allozymes were surveyed in all populations; two classes of anonymous nuclear loci were surveyed as well in two introduced populations). *N* indicates the number of nests (monogyne, polygyne) from which genetic data were obtained for each population (one genotype per nest was used for the F_{ST} estimates). Data from Ross et al. (1996b, 1997, 1999, 2006), Ross (1997), Shoemaker et al. (2006), and K. Ross and C. DeHeer (unpubl.).

Figure 2.2.—PHYLOGENETIC POSITION OF *SOLENOPSIS* GP-9 SEQUENCES RELATIVE TO ODORANT BINDING PROTEIN (OBP) AND CHEMOSENSORY PROTEIN (CSP) SEQUENCES FROM OTHER INSECTS

The large phylogeny contains exemplar GP-9 sequences from ten *Solenopsis* species as well as 159 OBP and CSP sequences. The small phylogeny contains a GP-9^B sequence from *S. invicta* as well as eight exemplar OBP and three exemplar CSP sequences. Amino acid sequences were aligned using various algorithms (ClustalX and MUSCLE on the complete data, T-Coffee and DIALIGN on the data subset) then subjected to phylogenetic analysis using neighbor joining, maximum parsimony, and Bayesian inference. The depicted phylogenies resulted from Bayesian inference, with the larger one based on a MUSCLE alignment and the smaller one on a T-Coffee alignment.

Posterior probability support values greater than 50% are shown for the smaller phylogeny. The "gold standard" OBPs have been implicated in chemoreception.

Figure 2.3.—REGULATION OF QUEEN NUMBER IN A POLYGYNE *S. INVICTA* COLONY BY MEANS OF QUEEN-CASTE SIGNALS PRODUCED BY INDIVIDUAL ADULT FEMALES

Variation in queen-caste signal among individuals is represented by the width of the colored ribbons. Orange hatching indicates individual workers' threshold levels of signal at which a female is recognized as a potential or actual reproductive (variation among workers is represented by the width of the hatching). Increasing density of ribbon colors signifies the increased expression of *Gp-9* as adult females mature.

Figure 2.1

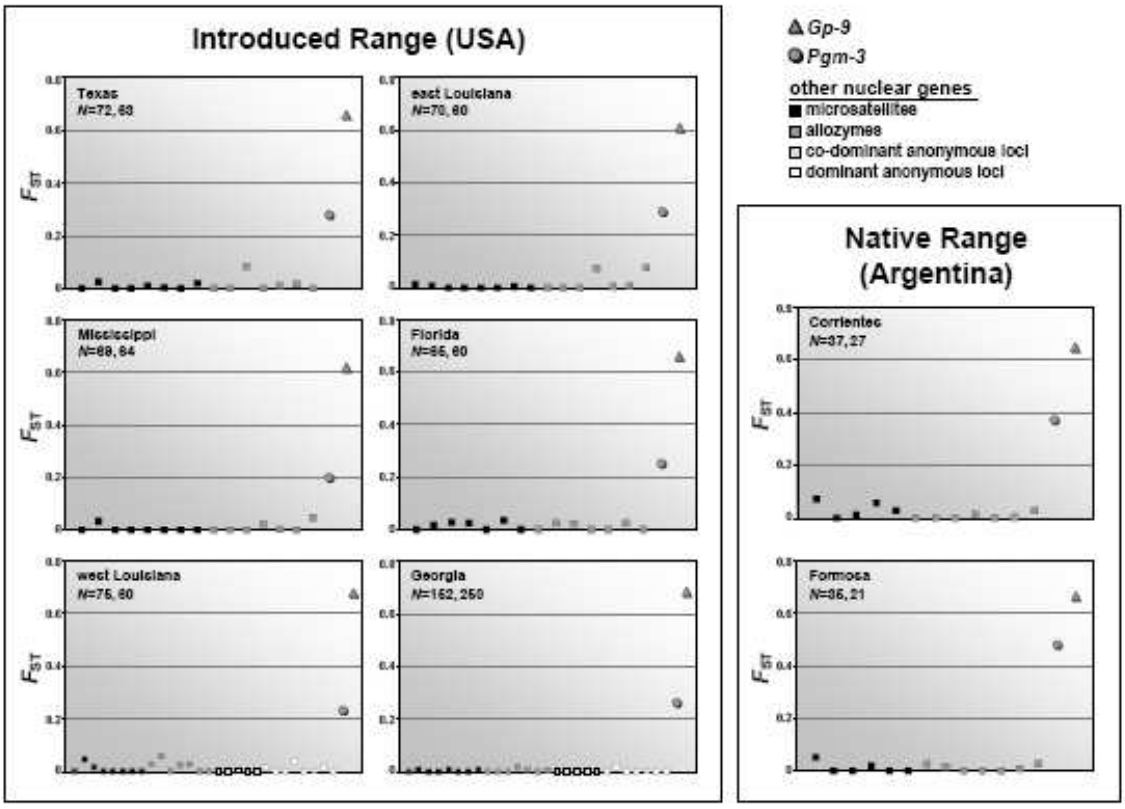

Introduced Range (USA)
Texas
N=72, 68
east Louisiana
N=70, 60
Mississippi
N=68, 64
Florida
N=85, 60
west Louisiana
N=75, 60
Georgia
N=152, 250
Native Range
(Argentina)
Corrientes
N=37, 27
Formosa
N=36, 21
Gp-9
Pgm-3
other nuclear genes
microsatellites
allozymes
co-dominant anonymous loci
dominant anonymous loci
FST

Figure 2.2

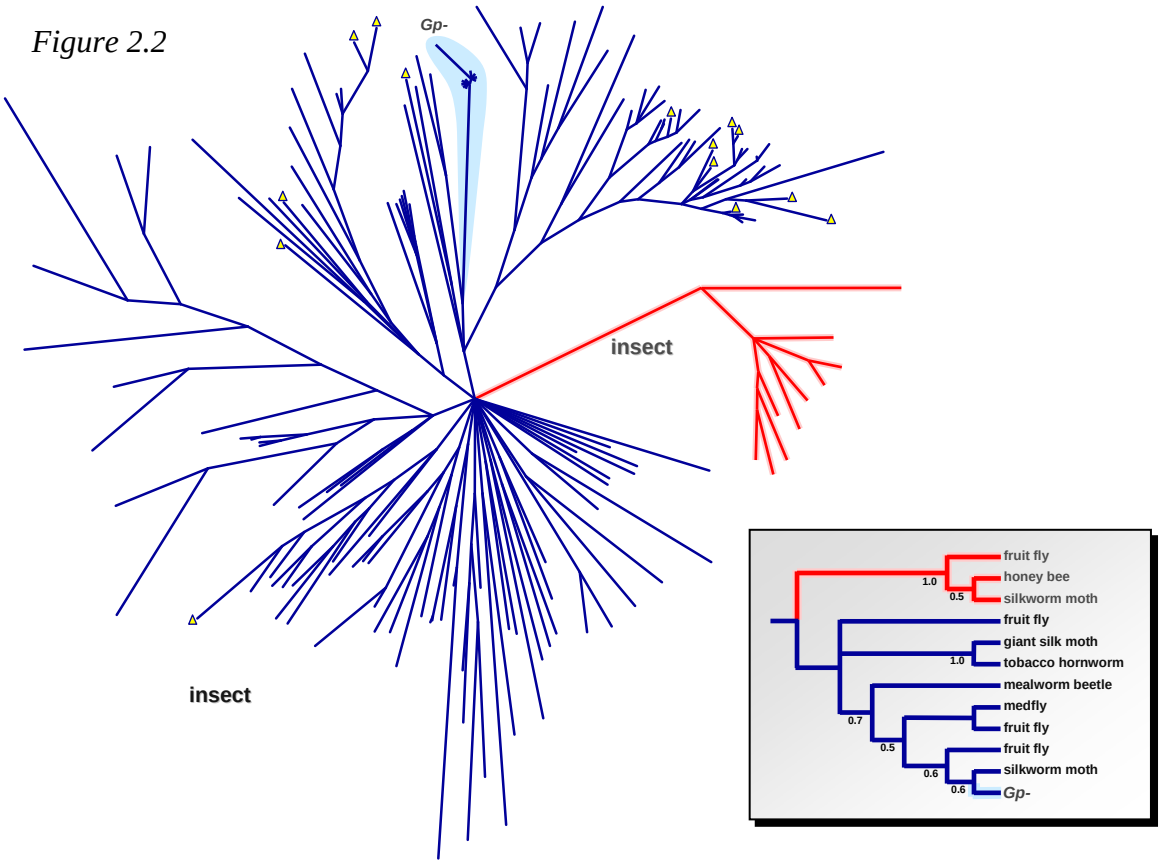
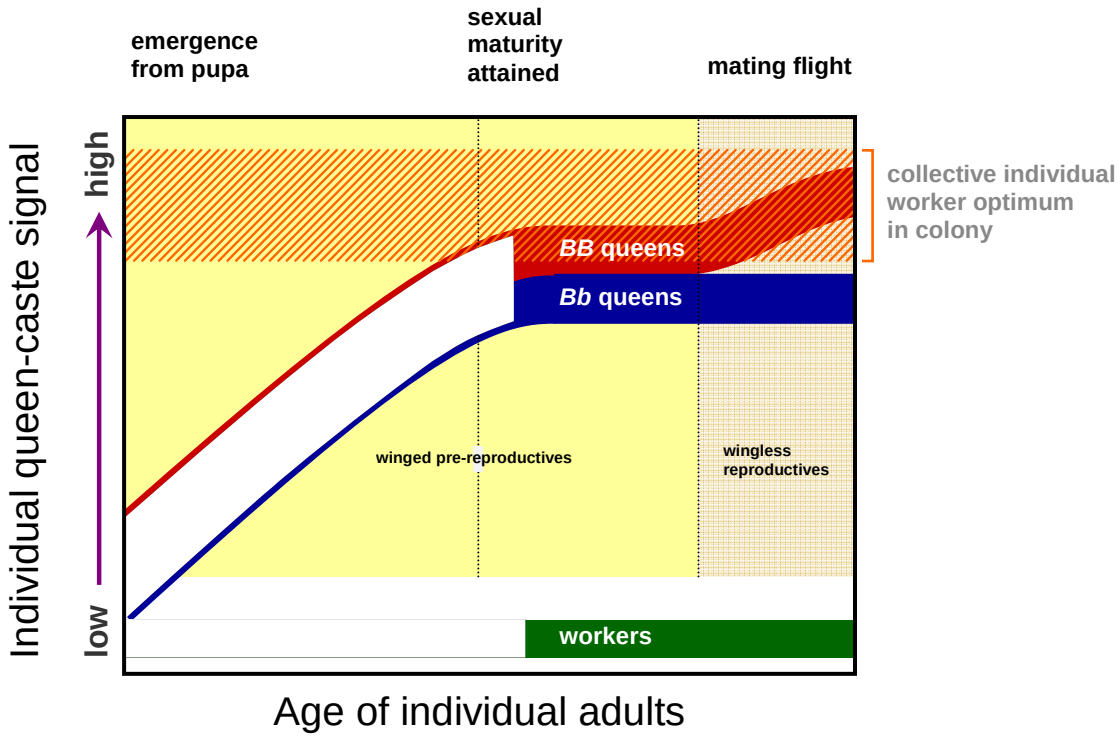

Gp-
insect
insect
fruit fly
honey bee
silkworm moth
fruit fly
giant silk moth
tobacco hornworm
mealworm beetle
medfly
fruit fly
fruit fly
silkworm moth
Gp-
1.0
0.5
1.0
0.7
0.5
0.6
0.6

Figure 2.3


emergence from pupa
sexual maturity attained
mating flight
Individual queen-caste signal
high
low
BB queens
Bb queens
winged pre-reproductives
wingless reproductives
workers
collective individual worker optimum in colony
Age of individual adults

CHAPTER 3

EXPERIMENTAL CONVERSION OF COLONY SOCIAL ORGANIZATION IN FIRE ANTS (*SOLENOPSIS INVICTA*): WORKER GENOTYPE MANIPULATION IN THE ABSENCE OF QUEEN EFFECTS[1]

[1] Dietrich Gotzek and Kenneth G. Ross, submitted to *Journal of Insect Behavior*, 10/15/06

Abstract.—Colony social organization in the fire ant *Solenopsis invicta* appears to be under strong genetic control. In the invasive USA range, polygyny (multiple queens per colony) is marked by the presence of the *Gp-9*^{b} allele in most of a colony's workers, whereas monogyny (single queen per colony) is associated with the exclusive occurrence of the *Gp-9*^{B} allele. Ross and Keller (2002) experimentally manipulated social organization by cross-fostering queens into colonies of the alternate form, thereby changing adult worker *Gp-9* genotype frequencies over time. Although these authors showed that social behavior switched predictably when the frequency of *b*-bearing adult workers crossed a threshold of 5-10%, the possibility that queen effects caused the conversions could not be entirely excluded. We addressed this problem by fostering polygyne brood into queenright monogyne colonies. All such treatment colonies switched social organization to become polygyne, coincident with their proportions of *b*-bearing workers exceeding 12%. Our results support the conclusion that polygyny *in S. invicta* is induced by a minimum frequency of colony workers carrying the *b* allele, and further confirm that its expression is independent of queen genotype or history, worker genotypes at genes not linked to *Gp-9*, and colony genetic diversity.

INTRODUCTION

A fundamental question in biology is how simple units are organized and integrated to ultimately emerge as a complex and unified whole. Such transitions have proved to be defining moments in the history of life (Maynard Smith and Szathmáry 1995). One such major transition has been the evolution of complex insect societies. Students of social insects recognize that the issue of how emergent colony-level phenotypes evolve and are expressed requires sufficiently complex behavioral models that capture salient features of the underlying genetic architecture. Important components of such models are the number of genes involved in regulating particular social traits, the relative impact of these genes, their interactions with one another (epistasis), their environmentally dependent phenotypic effects (reaction norm), and their contribution to the expression of relevant social phenotypes in other individuals (indirect genetic effects) (Grafen 1984; Crozier and Pamilo 1996; Robinson et al. 1997; Wolf et al. 1998; Robinson 1999; Linksvayer and Wade 2005).

Several studies have documented genetic contributions to individual traits of social relevance, including behaviors, in social insects. For example, caste determination in some ants and stingless bees (Kerr 1974; Winter and Buschinger 1986; Fersch et al. 2000; Fraser et al. 2000; Volny and Gordon 2002; Julian et al., 2002; Helms Cahan et al., 2002, 2004; Helms Cahan and Keller 2003) as well as individual reproductive roles in honey bees (Moritz and Hillesheim 1985; Page and Robinson 1994; Moritz et al. 1996; Montague and Oldroyd 1998) are under strong genetic control, and worker task performance in ants, wasps, honey bees, and caterpillars (Trump et al. 1967; Snyder 1992; Hunt et al. 1995, 1998; O'Donnell 1996; Robinson et al. 1997; Giray et al. 2000;

Fewell and Bertram 2002; Costa and Ross 2003) as well as the honey bee dance language (Rinderer and Beaman 1995; Johnson et al. 2002) have some heritable basis. At the colony level, genetic effects on social organization have been documented in ants and sweat bees (Cahan et al. 1998; Ross and Keller 1998; Plateaux-Quénu et al. 2000; Julian et al. 2002; Helms Cahan and Keller 2003; Helms Cahan et al. 2004). Perhaps not surprisingly, relatively few studies have shown a simple genetic architecture underlying complex colony-level traits (Moritz 1988; Hunt et al. 1995, 1999; Johnson et al. 2002). Of these, the best characterized system involves a single gene of major effect thought to regulate colony queen number in fire ants. The system has been most clearly elucidated in *Solenopsis invicta*, the red imported fire ant, in which it also was first described. Some colonies in this and related species are headed by a single reproductive queen (monogyny), whereas others have multiple reproductive queens (polygyny). In the invasive range of *S. invicta* in the USA, an allele dimorphism at the nuclear protein-coding gene *Gp-9* (*general protein-9*) is associated with this social dimorphism (Ross 1997; Ross and Keller 1998, 2002). Female ants of the monogyne form invariably are homozygous for the *B* allele of *Gp-9*. This is in contrast to the polygyne form, in which most females (and, indeed, all egg-laying queens) bear the alternate, *b* allele; those polygyne workers not bearing the *b* allele are homozygotes for the *B* allele, which segregates along with allele *b* in this form.

Because the *b* allele is only and always present in polygyne colonies, it is believed to be an indispensable component in the expression of this form of social organization in *S. invicta*. Significantly, variation at no other nuclear genes has been found to be associated with polygyny (Ross et al. 1997; Keller and Ross 1999; Ross et al. 1999; Shoemaker et

al. 2006). While it is conceivable that *Gp-9* simply marks a tightly linked gene or genes that actually determine social organization, a major role for non-genetic factors such as prior worker social experience, queen fecundity, and the social origin of queens in the expression of this colony-level trait has been discounted (Ross and Keller 1998). Based on these results, and the fact that it is the adult workers in fire ant societies who accept or reject supernumerary queens (Fletcher and Blum 1983; Keller and Ross 1993, 1998; Ross and Keller 2002), it has been hypothesized that the presence of the *b* allele at some minimal frequency in the worker caste induces the polygyne colony phenotype (Ross and Keller 2002). These latter authors conducted an experiment to test this hypothesis by cross-fostering queens between colonies of the different social forms. As predicted, colonies converted to the alternate social form over time as the original workers were replaced by the offspring of the new queens and frequencies of adult *b*-carrying workers steadily increased or decreased. Surprisingly, these social transitions consistently occurred at a relatively low threshold frequency of such workers (5-10%), which the authors interpreted as evidence that only a low proportion of *b*-bearing workers is required to elicit polygyne behavior by a colony.

Several potential shortcomings in the experimental design of Ross and Keller (2002) are remedied in the present study in order to test the robustness of these previous results. First, the earlier study did not control for possible queen effects, such as worker habituation to queens that might have induced the worker force as a whole to adopt the social behavior characteristic of the fostered queens' colony of origin. Ross and Keller (2002) concluded that such effects were unlikely, given that social conversion did not occur until 70-130 days following queen adoption and because no such queen effects had

been detected in a previous set of similar experiments (Ross and Keller 1998). Nonetheless, more subtle queen effects may have influenced the time course of social conversion to give unrealistic estimates of the threshold frequencies of *b*-bearing workers in the Ross and Keller (2002) study. Thus, we avoided potential involvement of queen effects altogether in the present study by fostering brood rather than queens between colonies. Also, the earlier study yielded only relatively coarse estimates of the threshold frequencies of *b*-bearing workers required for social transition. Those estimates are improved upon here by assaying social organization more frequently and by estimating frequencies of *b*-bearing workers from larger samples using a PCR-based rather than a protein-electrophoretic method. The former method is expected to be less biased because workers of all sizes can be scored (see Goodisman et al. [1999] for discussion of possible biases introduced by electrophoretic scoring of *Gp-9* in workers). Our new results confirm the hypothesis that polygyny in *S. invicta* is induced by the presence in a colony of workers bearing the *b* allele at even relatively low frequencies and refute a role for effects of queen genotype or history in experimental colony social conversion.

MATERIALS AND METHODS

Collection and Rearing of Colonies

Twenty queen-right monogyne colonies of *S. invicta* were collected in Oglethorpe and Clarke Counties, Georgia to serve as treatment and control colonies. Fifteen polygyne *S. invicta* colonies were collected from Clarke County to serve as donors of brood and test queens for the experiment. Social organization of the monogyne colonies

was confirmed by genotyping 50 workers per colony at *Gp-9* using horizontal starch gel electrophoresis (DeHeer et al. 1999); all 1000 workers so genotyped were determined to be *BB* homozygotes. Social organization of the polygyne colonies was confirmed by virtue of their possession of multiple reproductive (wingless, egg-laying) queens, large samples of which were subsequently confirmed to contain only *Bb* heterozygotes at *Gp-9* (see below). All colonies were returned to the laboratory following collection, separated from the soil, and reared under standard laboratory conditions (Jouvenaz et al. 1977; Ross 1988; Ross and Keller 2002). Sexual brood and winged adults were removed by sieving the colonies during initial establishment in the laboratory and periodically thereafter as brood became sexualized. Colonies and queenless test fragments were fed daily by alternating a high-protein diet (tuna / dog food / peanut butter mix) with a high-carbohydrate diet (assorted vegetables / granulated sugar mix). These diets were supplemented with frozen crickets and/or meal worms provided on a daily basis.

Addition of Brood

Fifteen monogyne colonies were randomly assigned to the treatment group, while the remaining five colonies were assigned to the control group. We took advantage of the fire ant habit of accepting alien conspecific brood (Tschinkel 1992; Stamps and Vinson 1991) to introduce such brood into the 20 queenright colonies. Treatment group colonies received only polygyne brood, whereas control colonies received only monogyne brood. Brood was introduced twice into each colony, at the start of the experiment and two weeks later. On each occasion, each recipient colony received a mixture of brood from several (5-11) donor colonies of the relevant social form. Mixing of fostered brood was

done for two reasons. First, it decreased genetic differentiation among the brood cohorts fostered into each colony both within the treatment and control groups and across the two groups; thus, the genetic background of the introduced brood was largely homogenized between the treatment and control colonies at all genes except those linked to *Gp-9*. Second, brood was mixed in order to match the genetic diversity in the monogyne brood introduced into the control colonies with the high genetic diversity in brood from even a single polygyne donor colony (e.g., Ross and Fletcher 1985). Brood collected from donor colonies was sieved repeatedly prior to mixing to exclude all sexual brood and as many adult workers as possible. The sieved brood of each recipient colony was weighed, and mixed donor brood weighing 50-100% of the original weight was then added. This typically amounted to 10-25 g of added brood, which corresponds roughly to 10,000-30,000 individuals (see Porter and Tschinkel 1985).

Successful acceptance by the recipient colonies of adult workers derived from the introduced foreign brood was monitored in the following ways. For the five control colonies and five randomly chosen treatment colonies, genotypic profiles at the allozyme loci *Est-4*, *G3pdh-1*, and *Pgm-1* (Ross 1993) were generated by genotyping ten adult workers per colony before brood addition; because monogyne colonies constitute simple families (Ross and Fletcher 1985), the limited range of variation occurring within each at these loci was uncovered with this protocol. Three weeks after final brood addition (at the time of the second bioassay), another 20 adult workers from each of these colonies were genotyped at the same loci. The proportions of these individuals with multilocus genotypes inconsistent with the original colony profile are minimum estimates of the proportion of foreign workers in the colony at that time, because foreign workers whose

genotypes matched those in the original colony were indistinguishable from the resident queen's offspring. The genetic relatedness of colony-mates sampled before and after brood addition was estimated from the allozyme data as a measure of colony genetic diversity using the program RELATEDNESS 5.0 (Queller and Goodnight 1989).

For all 15 treatment colonies, acceptance of the foreign brood and emerged adults was further confirmed by estimating the frequencies of adult workers bearing the *b* allele at *Gp-9* three weeks after the second (final) brood addition. Again, this provides a minimum estimate of the proportion of foreign workers in the colony, because on average about a third of the introduced polygyne brood are expected to bear the same *BB* genotype possessed by the original workers in the monogyne treatment colonies (e.g., Ross and Keller 2002).

Bioassay of Social Organization and Estimation of Frequencies of b*-bearing Workers*

Bioassays followed the general protocol of Ross and Keller (2002). For each, a large test colony fragment containing brood but no queen was prepared and held in isolation from the parent colony for 48 h; fragments were held queenless for this period because queenright monogyne colonies will not accept any foreign queen(s) (Fletcher 1986; Vander Meer and Alonso 2002). Next, four polygyne reproductive queens were introduced into the foraging arena of the queenless fragment. All surviving queens and any identifiable heads or thoraces of executed queens were removed 24 h after the queen introductions. These were genotyped electrophoretically at *Gp-9* to confirm that all introduced queens possessed a *b* allele. Survival of multiple introduced polygyne (*b*

allele-carrying) queens for 24 h indicates expression of the polygyne colony phenotype, whereas execution of all such queens signals expression of the monogyne phenotype (Ross and Keller 1998, 2002). Queenless fragments were re-combined with their parent colony immediately after completion of each assay.

The first bioassay was conducted just before initial brood addition to confirm that the monogyne test colonies behaved as expected. The second assay was conducted three weeks after the final addition of foreign brood. This approximates the developmental period from newly laid egg to eclosion of the adult minor worker in *S. invicta* under our rearing conditions (O'Neal and Markin 1975), so that most of the introduced brood probably had developed to mature (several day old) adult workers by the time of the second assay. Colonies subsequently were bioassayed at approximately two week intervals. Treatment colonies were terminated when they displayed a reversal to monogyne social behavior that was stable for at least three consecutive assays or, in the absence of such reversal, eight months after the final brood addition. Most adult workers probably live no more than four months under our rearing conditions (Calabi and Porter 1989), so few workers derived from foreign brood were expected to inhabit the latter colonies by the end of the experiment.

Estimation of the frequency of adult workers bearing the *b* allele was done at the time of the first two bioassays for each treatment colony. The procedure was suspended after a colony converted to polygyny, but was re-instituted once it reverted back to monogyny (or, in the few cases in which such reversal did not occur, at the last assay). For each sample, nests were disturbed by lifting their tops and a random sample of several hundred workers from the nests and foraging arena was collected. The DNA of

75 of the collected individuals was extracted using a simple method (Turelli and Hoffmann 1995) and used as the template in *b* allele-specific PCR reactions (Mescher et al. 2003). The microsatellite locus *Sol-42* (Krieger and Keller 1997) was multiplexed in the PCRs as a control for successful amplification. Amplification products were run out in 1.5% agarose gels stained with ethidium bromide and were visualized under UV illumination. Presence of the *b* allele was indicated by amplification of a 219 bp fragment (Mescher et al. 2003). Any ambiguous amplifications (i.e., those in which the microsatellite locus did not amplify) were redone; after a second failure the sample was substituted by another. A jackknife resampling procedure was used to estimate the frequency of *b*-bearing workers from each sample, with the 95% confidence intervals obtained by assuming the *t*-distribution (Weir 1996). For samples following brood addition in which only *BB* workers were collected, 95% confidence intervals were obtained after assuming that the next individual assayed (the 76^{th}) would have possessed allele *b*.

RESULTS

Acceptance of Foreign Brood and Emerged Adult Workers

Foreign brood was readily accepted and reared to adulthood by all treatment and control colonies. A minimum of 10-45% of the adult workers in each control colony and 5-10% of the workers in five randomly chosen treatment colonies are estimated to have been derived from foreign brood (monogyne and polygyne, respectively) three weeks after final brood addition, based on their possession of multilocus allozyme genotypes

inconsistent with the original colony profiles. Based on presence of the *b* allele of *Gp-9*, a minimum of 12-42% of adult workers in each treatment colony are estimated to have been derived from foreign polygyne brood at this juncture. Average colony-mate relatedness estimated from three allozyme loci fell from 0.69 (95% jackknife confidence interval 0.59-0.78) before brood addition to 0.38 (0.22-0.55) in the treatment colonies and 0.29 (0.18-0.39) in the control colonies at the three week juncture, suggesting comparable within-colony genetic diversity in the two classes of test colonies at the time of the second bioassay.

Conversion of Colony Social Organization

All colonies rejected all polygyne (*b*-carrying) queens introduced in the first bioassay, done immediately prior to the addition of foreign brood, and thus displayed the social behavior expected of monogyne colonies. The outcome of the second assay, conducted three weeks after final brood addition, was markedly different. All treatment colonies accepted all four introduced queens, thus invariably displaying polygyne social behavior. Associated with this change in social organization, frequencies of adult workers bearing the *b* allele in each colony rose from 0% to no less than 12% by the time of this second assay (Figure 3.1). In complete contrast to the treatment colonies, none of the control colonies accepted any introduced queens in the second assay or any of the subsequent 18 assays conducted on each during the course of the experiment. Thus, the controls behaved as typical monogyne colonies throughout the course of the experiment.

One of the 15 treatment colonies (M-04) was excluded from further analysis because it declined dramatically in size after the second assay and the queen eventually died. Of the remaining 14 colonies, nine reverted back to monogyne behavior by the end of the study period (see Figure 3.1), an event that occurred between 12 and 24 weeks after the second assay as frequencies of *b*-bearing workers inevitably declined. One of these nine colonies (M-15) transiently displayed ambiguous social behavior during this period of reversal; after ten weeks of polygyne behavior, it accepted only a single *Bb* queen in four consecutive assays (for 8 weeks) before finally rejecting all introduced queens and thus displaying genuine monogyne behavior for the ensuing four assays.

The proportion of *b*-carrying individuals in the 15 treatment colonies at the time that they first displayed polygyne behavior (second bioassay) ranged from 12% to 42% (Figure 3.1). Among the nine colonies that reverted back to monogyny, this proportion dropped to between 0% and 14.7% (mean=3.4%) at the sample point when reversal was first detected. Five treatment colonies never reverted back to monogyne behavior over the course of the experiment. Four of these still had detectable levels of *b*-carrying workers at the end of the experiment (estimated at 1%-11.8%), but we were unable to recover any such workers in colony M-14 (even in an expanded sample of 100 individuals). The proportion of *b*-bearing workers in the five permanently polygyne treatment colonies averaged 4.5% at the end of the experiment.

In Figure 3.2, each estimate of the proportion of *b*-bearing workers in the treatment colonies is grouped according to the type of social behavior displayed by the colony at the time the sample was taken. The distributions are quite discrete for the two forms of behavior, with an apparent discontinuity occurring at around 10% *b*-carrying workers.

Conspicuous exceptions are the single monogyne-behaving colony with a frequency close to 15% and the two polygyne-behaving colonies with frequencies below 4%. These results generally correspond with the finding of Ross and Keller (2002) that colonies with cross-fostered queens underwent social transitions when their frequencies of *b*-carrying workers crossed a threshold of 5-10%.

DISCUSSION

The results of this study support the hypothesis that colony social organization in *S. invicta* is determined by the presence or absence of adult workers bearing the *b* allele at the gene *Gp-9*. Specifically, polygyny is expressed when such workers are present in a colony at a minimal frequency of several percent, whereas monogyny typically is expressed when such workers make up fewer than several percent of the worker force or are absent altogether. To test this hypothesis, we introduced brood pooled from natural polygyne colonies, which contain a high frequency of *b*-carrying individuals, into monogyne colonies, which lack them. Shortly thereafter, adult workers bearing the allele appeared at moderate frequencies and these treatment colonies concurrently switched social organization from monogyny to polygyny. Control colonies, all of which failed to undergo such social conversion, received a mixture of brood from monogyne colonies. This protocol served to homogenize brood from the different donor types with respect to their levels of genetic diversity as well as their genetic background at genes not linked to *Gp-9*. By retaining each colony's original queen, we avoided any potential influence of direct queen effects on expression of social organization in the experiment.

Our results are important for several reasons. First, they strongly suggest that social organization in *S. invicta* is under worker control rather than queen control, and that worker genotype composition at *Gp-9* (or linked genes) is the dominant factor determining collective worker social behavior. The relative unimportance of queen control (and concomitant importance of worker *Gp-9* genotypes) was indicated also in experiments by Ross and Keller (1998). Our results further show that genes not linked to *Gp-9* are not involved in regulation of social organization, a conclusion supported as well by the overall similarity of sympatric populations of the two forms at numerous neutral genetic markers (Ross et al. 1987, 1997, 1999, 2006; Shoemaker et al. 2006). Finally, our demonstration that differences in levels of genetic diversity between colonies of the two forms are irrelevant to the expression of social organization stands in contrast to the view that polygyny may be induced by high genetic diversity and the correspondingly high odor cue diversity it engenders (see Hölldobler and Wilson 1971; Morel et al. 1990; Obin et al. 1993; Ross et al. 1996).

One benefit of our experimental design is that we were able to test the effect of *b*-carrying workers on expression of social organization twice for each colony — once as these workers increased in frequency shortly after brood addition and again later as they dropped in frequency due to their replacement by the resident queen's offspring (all of which possessed genotype *BB*). At our first bioassay after brood addition, each treatment colony had switched from monogyne to polygyne social behavior, coincident with the appearance of adult *b*-carrying workers at estimated frequencies of 12% to 42%. This result is fully consistent with the findings of Ross and Keller (2002) based on a fundamentally different experimental design (fostering queens rather than brood between

colonies), where conversion from monogyny to polygyny occurred as the frequency of adult *b*-bearing workers in a colony crossed a threshold of 5-10%. We note that polygyne colonies in the wild invariably have far greater frequencies of such workers owing to the fact that all reproductive queens of this social form carry the *b* allele (Ross 1997; Ross and Keller 1998, 2002; Fritz et al. 2006). Thus, in one elementary sense our combined laboratory experiments provide a robust proximate explanation of why naturally occurring colonies with *b*-carrying queens invariably display the polygyne social phenotype.

The second test of the effect of *b*-bearing workers on expression of social organization, as their frequencies returned to low levels due to worker turnover, also was largely consistent with the findings of Ross and Keller (2002). Nine of the 14 relevant treatment colonies reverted back to monogyny within several months after their initial conversion to polygyny; all nine exhibited frequencies of *b*-bearing workers at or near the 5-10% threshold at the time of their reversal (colony M-08 had the highest frequency of such workers−14.7%, with a lower jackknife confidence limit of 8%). Among the five colonies that did not revert to monogyny, four still had *b*-bearing workers present at sample frequencies of 1% to 12%, which could correspond to actual frequencies in the colonies ranging as high as 3% to 18% (based on the upper 95% confidence limits). We found no *b*-carrying workers in a sample of 100 from the fifth such colony (M-14), suggesting that such workers were exceedingly rare or absent in this polygyne-behaving colony. This lone example represents a seeming contradiction to the view that expression of polygyny requires the presence of adult workers bearing the *b* allele at the gene *Gp-9*.

On the other hand, there are potential sources of error or natural variation that may reconcile the case of colony M-14 late in the experiment with the great majority of other test colonies in this study and that of Ross and Keller (2002). Although our increased sample sizes, use of a PCR-based technique to score worker *Gp-9* alleles, and increased frequency of bioassays represent methodological improvements over the previous study, estimation errors attributable to finite sample sizes and the two-week intervals between each sample may still explain discrepancies in any single anomalous estimate. Perhaps more importantly, variation in any of a number of potential colony attributes (e.g., colony size, adult worker/brood ratio, queen vigor) may in some cases mitigate or even override the effect of worker *Gp-9* genotype composition in determining colony social organization; that is, the penetrance of *Gp-9* may not be complete under all circumstances in experimental laboratory colonies. As an extreme example, colonies of each form can be induced to accept queens they would normally reject based on the joint queen/worker *Gp-9* genotypes if they are held queenless for extended periods (Fletcher 1986; Ross and Keller 2002; Vander Meer and Alonso 2002). Given the complex, emergent nature of a colony-level phenotype such as social organization, it is remarkable that our combined experimental results are so congruent in pointing to a paramount role of worker *Gp-9* composition in its expression.

One potentially confounding attribute not considered in this study or that of Ross and Keller (2002) is colony size. Under some conceivable proximate models of worker regulation of queen number, the frequency of *b*-bearing workers required to elicit polygyny is not expected to scale linearly with colony population. For instance, if the cues introduced by *b*-bearing workers that induce a switch in colony social phenotype are

pheromonal rather than behavioral in nature, these cues may be more persistent or easily disseminated in smaller than larger colonies. If not readily degraded, they could persist in the presence of recently deceased *b*-carrying workers at levels effective in small but not large colonies. If colony M-14 and other apparent outliers were unusually small late in the experiment due to stress from continued manipulation and worker attrition, very low frequencies of *b*-bearing workers may still have been sufficient to maintain the prevailing polygyne social phenotype.

This point raises the more general issue of how a small fraction of workers can influence a colony-level phenotype that, at its core, constitutes collective tolerance of multiple reproductive queens bearing the same *Gp-9* allele as these minority workers (e.g., Keller and Ross 1998). Although there are several examples of such behaviorally dominant colony phenotypes in social insects, some of which can be readily explained by the cumulative actions of just a few individuals (e.g., hygienic behavior of honey bees [Trump et al. 1967] and sex ratio manipulation in ants [Aron et al. 1995; Passera et al. 2001]), development of proximate explanations of regulation of queen number in fire ants is hampered by an almost complete lack of knowledge of the individual behaviors involved (but see Keller and Ross 1998). Ross and Keller (2002) proposed two potential classes of mechanisms by which relatively few *b*-bearing workers could influence a large worker force to tolerate multiple *b*-bearing queens. The first is behavioral manipulation, which could occur, for instance, by *b*-bearing workers releasing an appeasement pheromone to forestall aggressive responses toward queens or by *b*-bearing workers transporting queens into the nest and physically protecting them from attack until aggression dissipates. The second is habituation of the entire colony to a unique odor

(pheromone) associated with *b*-carrying adults (workers and queens). Of the two, we favor this second alternative for several reasons. First, the likelihood of a small minority of workers finding and somehow protecting a queen seeking acceptance into a colony from other workers seems remote. On the other hand, an odor cue can perhaps more easily explain the dominant behavioral effect of rare individuals in a colony, as discussed above. Finally, *Gp-9* encodes a transport protein belonging to the odorant binding protein (OBP) gene family (Krieger and Ross 2002), so if directly involved in control of queen number it conceivably could exert its effects via a role in odor transport, emanation, or detection (see Vogt [2005] for discussion of the possible roles of OBPs in insect chemoreception).

Ideally, our experimental protocol would have included bioassays of the treatment colonies immediately after initial brood addition in order to completely rule out any role of the brood in inducing social conversion. For instance, maternal effects mediated through surface hydrocarbons placed on eggs by queens could be hypothesized to play such a role (e.g., Endler et al. 2006). However, the difficulty of completely removing all adult workers from the large volumes of added brood would have compromised such an early assay. Moreover, the experimental results of Ross and Keller (2002) appear to contradict any role for brood in influencing social organization, given the considerable delay between cross-fostering of queens and conversion in colony social phenotype found in that study. Similar long delays between the final addition of polygyne brood and reversion of treatment colonies from polygyny to monogyny in the present study also are inconsistent with the brood playing such a role.

In conclusion, this study confirms a strong genetic component to social organization in *S. invicta* that is mediated by colony worker genotype frequencies at *Gp-9* (or closely linked genes). We have shown that polygyne social behavior is induced by the presence of *b*-carrying adult workers, even at quite low frequencies, whereas monogyne behavior is expressed in the absence of such workers. Direct queen effects were eliminated as possible causative factors in our experimental design, as were effects of genes not linked to *Gp-9* and overall colony genetic diversity. The processes involved in worker regulation of colony queen number in *S. invicta*, which constitutes the core element of social organization, are presently poorly understood. Detailed understanding of queen acceptance and rejection behaviors, coupled with improved knowledge of the biochemical and physiological roles of *Gp-9* (and linked candidate genes), are required to explicitly model the proximate and evolutionary mechanisms underlying the seemingly simple yet evolutionarily important transition from one social form to the other in fire ants.

ACKNOWLEDGMENTS

We thank Chris DeHeer, Laurent Keller, Emily Matthews, and Yannick Wurm for comments on the manuscript.

REFERENCES

Aron S, Vargo E L, Passera L. 1995. Primary and secondary sex ratios in monogyne colonies of the fire ant. Anim Behav 49: 749-757.

Cahan S, Helms KR, Rissing SW. 1998. An abrupt transition in colony founding behaviour in the ant *Messor pergandei*. Anim Behav 55: 1583-1594.

Calabi P, Porter SD. 1989. Worker longevity in the fire ant *Solenopsis invicta*: ergonomic considerations of correlations between temperature, size and metabolic rates. J Insect Physiol 35: 643-649.

Costa JT, Ross KG. 2003. Fitness effects of group merging in a social insect. *Proc R Soc Lond B* 270: 1697-1702.

Crozier RH, Pamilo P. 1996. *Evolution of Social Insect Colonies; Sex Allocation and Kin Selection*, Oxford University Press, Oxford.

DeHeer CJ, Goodisman MAD, Ross KG. 1999. Queen dispersal strategies in the multiple-queen form of the fire ant *Solenopsis invicta*. Am Nat 153: 660-675.

Endler A, Liebig J, Hölldobler B. 2006. Queen fertility, egg marking and colony size in the ant *Camponotus fIoridanus*. Behav Ecol Sociobiol 59: 490-499.

Fersch R, Buschinger A, Heinze J. 2000. Queen polymorphism in the Australian ant *Monomorium* sp. 10. Insectes Soc 47: 280-284.

Fewell JH, Bertram SM. 2002. Evidence for genetic variation in worker task performance by African and European honey bees. Behav Ecol Sociobiol 52: 318-325.

Fletcher DJC. 1986. Triple action of queen pheromones in the regulation of reproduction in fire ant (*Solenopsis invicta*) colonies. In Porchet M, Andries J-C, Dhainaut A. (eds.) *Advances in Invertebrate Reproduction 4*, London: Elsevier. Pp. 305-316.

Fletcher DJC, Blum MS. 1983. Regulation of queen number by workers in colonies of social insects. Science 219: 312-314.

Fraser VS, Kaufmann B, Oldroyd BP, Crozier RH. 2000. Genetic influence on caste in the ant *Camponotus consobrinus*. Behav Ecol Sociobiol 47: 188-194.

Fritz GN, Vander Meer RK, Preston CA. 2006. Selective male mortality in the red imported fire ant, *Solenopsis invicta*. Genetics 173: 207-213.

Giray T, Guzmán-Novoa E, Aron CW, Zelinsky B, Fahrbach SE, Robinson GE. 2000. Genetic variation in worker temporal polyethism and colony defensiveness in the honey bee, *Apis mellifera*. Behav Ecol 11: 44-55.

Goodisman MAD, Mack PD, Pearse DE, Ross KG. 1999. Effects of a single gene on worker and male body mass in the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae). Ann Entomol Soc Am 92: 563-570.

Grafen A. 1984. Natural selection, kin selection and group selection. In Krebs JR, Davies NB. (eds.) *Behavioural Ecology; An Evolutionary Approach*, Sunderland: Sinauer. Pp. 62-84.

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

Helms Cahan S, Parker JD, Rissing SW, Johnson RA, Polony TS, Weiser MD, Smith DR. 2002. Extreme genetic differences between queens and workers in hybridizing *Pogonomyrmex* harvester ants. Proc R Soc Lond B 269: 1871-1877.

Helms Cahan S, Julian GE, Rissing SW, Schwander T, Parker JD, Keller L. 2004. Loss of phenotypic plasticity generates genotype-caste association in harvester ants. Curr Biol 14: 2277-2282.

Hölldobler B, Wilson EO. 1977. The number of queens: An important trait in ant evolution. Naturwissenschaften 64: 8-15.

Hunt GJ, Page RE, Fondrk MK, Dullum CJ. 1995. Major quantitative trait loci affecting honey bee foraging behavior. Genetics 141: 1537-1545.

Hunt GJ, Guzmán-Novoa E, Fondrk MK, Page RE. 1998. Quantitative trait loci for honey bee stinging behavior and body size. Genetics 148: 1203-1213.

Johnson RN, Oldroyd BP, Barron A B, Crozier RH. 2002. Genetic control of the honey bee (*Apis mellifera*) dance language: segregating dance forms in a backcrossed colony. J Hered 93: 170-173.

Jouvenaz DP, Allen GE, Banks WA, Wojcik DP. 1977. A survey for pathogens of fire ants, *Solenopsis* spp. in the southeastern United States. Fla Entomol 60: 275-279.

Julian GE, Fewell JH, Gadau J, Johnson RA, Larrabee D. 2002. Genetic determination of the queen caste in an ant hybrid zone. Proc Natl Acad Sci USA 99: 8157-8160.

Keller, L., Ross KG. 1993. Phenotypic basis of reproductive success in a social insect: genetic and social determinants. Science 260: 1107-1110.

Keller L, Ross KG. 1998. Selfish genes: a green beard in the red fire ant. Nature 394: 573-575.

Keller L, Ross KG. 1999. Major gene effects on phenotype and fitness: the relative roles of *Pgm-3* and *Gp-9* in introduced populations of the fire ant *Solenopsis invicta*. J Evol Biol 21: 672-680.

Kerr W. 1974. Advances in cytology and genetics of bees. Annu. Rev. Entomol. 19: 253-268.

Krieger MJB, Keller L. 1997. Polymorphism at dinucleotide microsatellite loci in fire ant *Solenopsis invicta* populations. Mol Ecol 6: 997-999.

Krieger MJB, Ross KG. 2002. Identification of a major gene regulating complex social behavior. Science 295: 328-332.

Linksvayer TA, Wade MJ. 2005. The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: maternal effects, sib-social effects, and heterochrony. Quart Rev Biol 8: 317-336.

Maynard Smith J, Szathmáry E. 1995. *The Major Transitions in Evolution.* New York: Freeman.

Mescher MC, Ross KG, Shoemaker DD, Keller L, Krieger MJB. 2003. Distribution of the two social forms of the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae) in the native South American range. Ann Entomol Soc Am 96: 810-817.

Montague CE, Oldroyd BP. 1998. The evolution of worker sterility in honey bees: an investigation into a behavioral mutant causing failure of worker policing. Evolution 52: 1408-1415.

Morel L, Vander Meer RK, Lofgren CS. 1990. Comparison of nestmate recognition between monogyne and polygyne populations of *Solenopsis invicta* (Hymenoptera: Formicidae). Ann Entomol Soc Am 83: 642-647.

Moritz RFA. 1988. A reevaluation of the two-locus model for hygienic behavior in honeybees (*Apis mellifera* L.). J Hered 79: 257-262.

Moritz RFA, Hillesheim E. 1985. Inheritance of dominance in honeybees (*Apis mellifera capensis* Esch.). Behav Ecol Sociobiol 17: 87-89.

Moritz RFA, Kryger P, Allsopp MH. 1996. Competition for royalty in bees. Nature 384: 31.

Obin MS, Morel L, Vander Meer RK. 1993. Unexpected, well-developed nestmate recognition in laboratory colonies of polygyne imported fire ants (Hymenoptera: Formicidae). J Insect Behav 6: 579-589.

O'Donnell S. 1996. RAPD markers suggest genotypic effects on forager specialization in a eusocial wasp. Behav Ecol Sociobiol 38: 83-88.

O'Neal J, Markin GP. 1975. Brood development of the various castes of the imported fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae). J Kansas Entomol Soc 48: 152-159.

Page RE, Robinson GE. 1994. Reproductive competition in queenless honey bee colonies (*Apis mellifera* L.). Behav Ecol Sociobiol 35: 99-107.

Passera L, Aron S, Vargo EL, Keller L. 2001. Queen control of sex ratio in fire ants. Science 293: 1308-1310.

Plateaux-Quénu C, Plateaux L, Packer L. 2000. Population-typical behaviours are retained when eusocial and non-eusocial forms of *Evylaeus albipes* (F.) (Hymenoptera, Halictidae) are reared simultaneously in the laboratory. Insectes Soc 47: 263-270.

Porter SD, Tschinkel WR. 1985. Fire ant polymorphism (Hymenoptera: Formicidae): factors affecting worker size. Ann Entomol Soc Am 78: 381-386.

Queller DC, Goodnight KF. 1989. Estimating relatedness using genetic markers. Evolution 43: 258-275.

Rinderer TE, Beaman LD. 1995. Genetic control of honey bee dance language dialect. Theor Appl Genet 91: 727-732.

Robinson GE. 1999. Integrative animal behaviour and sociogenomics. Trends Ecol Evol 14: 202-205.

Robinson GE, Fahrbach SE, Winston ML. 1997. Insect societies and the molecular biology of social behavior. Bioessays 19: 1099-1108.

Ross KG. 1988. Differential reproduction in multiple-queen colonies of the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae). Behav Ecol Sociobiol 23: 341-355.

Ross KG. 1993. The breeding system of the fire ant *Solenopsis invicta*: Effects on colony genetic structure. Am Nat 141: 554-576.

Ross KG. 1997. Multilocus evolution in fire ants: effects of selection, gene flow, and recombination. Genetics 145: 961-974.

Ross KG, Fletcher DJC. 1985. Comparative study of genetic and social structure in two forms of the fire ant, *Solenopsis invicta* (Hymenoptera: Formicidae). Behav Ecol Sociobiol 17: 349-356.

Ross KG, Keller L. 1998. Genetic control of social organization in an ant. Proc Natl Acad Sci USA 95: 14232-14237

Ross KG, Keller L. 2002. Experimental conversion of colony social organization by manipulation of worker genotype composition in fire ants (*Solenopsis invicta*). Behav Ecol Sociobiol 51: 287-295.

Ross KG, Vargo EL, Fletcher DJC. 1987. Comparative biochemical genetics of three fire ant species in North America, with special reference to the two social forms of *Solenopsis invicta* (Hymenoptera: Formicidae). Evolution 41: 979-990.

Ross KG, Vargo EL, Keller L. 1996. Social evolution in a new environment: the case of introduced fire ants. Proc Natl Acad Sci USA 93: 3021-3025.

Ross KG, Krieger MJB, Shoemaker DD, Vargo EL, Keller L. 1997. Hierarchical analysis of genetic structure in native fire ant populations: results from three classes of molecular markers. Genetics 147: 643-655.

Ross KG, Shoemaker DD, Krieger MJB, DeHeer CJ, Keller L. 1999. Assessing genetic structure with multiple classes of molecular markers: a case study involving the introduced fire ant *Solenopsis invicta*. Mol Biol Evol 16: 525-543.

Ross KG, Krieger MJB, Keller L, Shoemaker DD. 2006. Genetic variation and structure in native populations of the fire ant *Solenopsis invicta*: evolutionary and demographic implications. (in review)

Shoemaker DD, DeHeer CJ, Krieger MJB, Ross KG. 2006. Population genetics of the invasive fire ant *Solenopsis invicta* (Hymenoptera: Formicidae) in the United States. Ann Entomol Soc Am 99: 1213-1233.

Snyder LE. 1992. The genetics of social behavior in a polygynous ant. Naturwissenschaften 79: 525-527.

Stamps WT, Vinson SB. 1991. Raiding in newly founded colonies of *Solenopsis invicta* Buren (Hymenoptera: Formicidae). Environ Entomol 20: 137-141.

Trump RF, Thompson VC, Rothenbuhler WC. 1967. Behaviour genetics of nest cleaning in honeybees. V. Effect of previous experience and composition of mixed colonies on response to disease-killed brood. J Apic Res 6: 127-131.

Tschinkel WR. 1992. Brood raiding in the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae): laboratory and field observations. Ann Entomol Soc Am 85: 638-646.

Turelli M, Hoffmann AA. 1995. Cytoplasmic incompatibility in *Drosophila simulans*: dynamics and parameter estimates from natural populations. Genetics 140: 1319-1338.

Vander Meer RK, Alonso LE. 2002. Queen primer pheromone affects conspecific fire ant (*Solenopsis invicta*) aggression. Behav Ecol Sociobiol 51: 122-130.

Vogt RG. 2005. Molecular basis of pheromone detection in insects. In Gilbert LI, Iatro K, Gill S. (eds.), *Comprehensive Insect Physiology, Biochemistry, Pharmacology and Molecular Biology. Endocrinology vol. 3.* London: Elsevier. Pp. 753-804.

Volny VP, Gordon DM. 2002. Genetic basis for queen-worker dimorphism in a social insect. Proc Natl Acad Sci USA 99: 6108-6111.

Weir BS. 1996. *Genetic Data Analysis II*. Sunderland: Sinauer.

Winter U, Buschinger A. 1986. Genetically mediated queen polymorphism and caste determination in the slave-making ant, *Harpagoxenus sublaevis* (Hymenoptera: Formicidae). Entomol Gen 11: 125-137.

Wolf JB, Brodie ED, Cheverud JM, Moore AJ, Wade MJ. 1998. Evolutionary consequences of indirect genetic effects. Trends Ecol Evol 13: 64-69.

FIGURE LEGENDS

Figure 3.1.—SOCIAL BEHAVIOR EXPRESSED BY TREATMENT COLONIES DURING BIOASSAYS, AND PROPORTIONS OF *BB* AND *B*-BEARING WORKERS AT CONVERSIONS IN SOCIAL FORM.

The first assay was conducted prior to addition of foreign brood, when all colonies contained only *BB* workers. The second assay was conducted approximately three weeks after the final addition of foreign brood. All subsequent assays were conducted approximately every two weeks. Worker *Gp-9* frequencies were determined only at the first two assays and at the subsequent assay at which a colony reverted in social behavior (or at the end of the experiment, see text). The type of social behavior expressed is indicated also for each assay at which *Gp-9* frequencies were not determined, with the "X" indicating the number of consecutive assays yielding the identical form of behavior. **M**: monogyne behavior expressed (no supernumerary queens accepted), **P**: polygyne behavior expressed (multiple supernumerary queens accepted), **?**: ambiguous social behavior expressed (a single queen accepted) by colony M-15 (see text). Colony M-04 declined after the second assay and so was dropped from the experiment. White bars represent 95% jackknife confidence intervals.

Figure 3.2.—PROPORTIONS OF ADULT *B*-BEARING WORKERS IN TREATMENT COLONIES EXPRESSING MONOGYNE OR POLYGYNE SOCIAL BEHAVIOR.

The threshold proportion of *b*-bearing workers hypothesized by Ross and Keller (2002) to be required for conversion between monogyny and polygyny is highlighted with stippling.

Figure 3.1

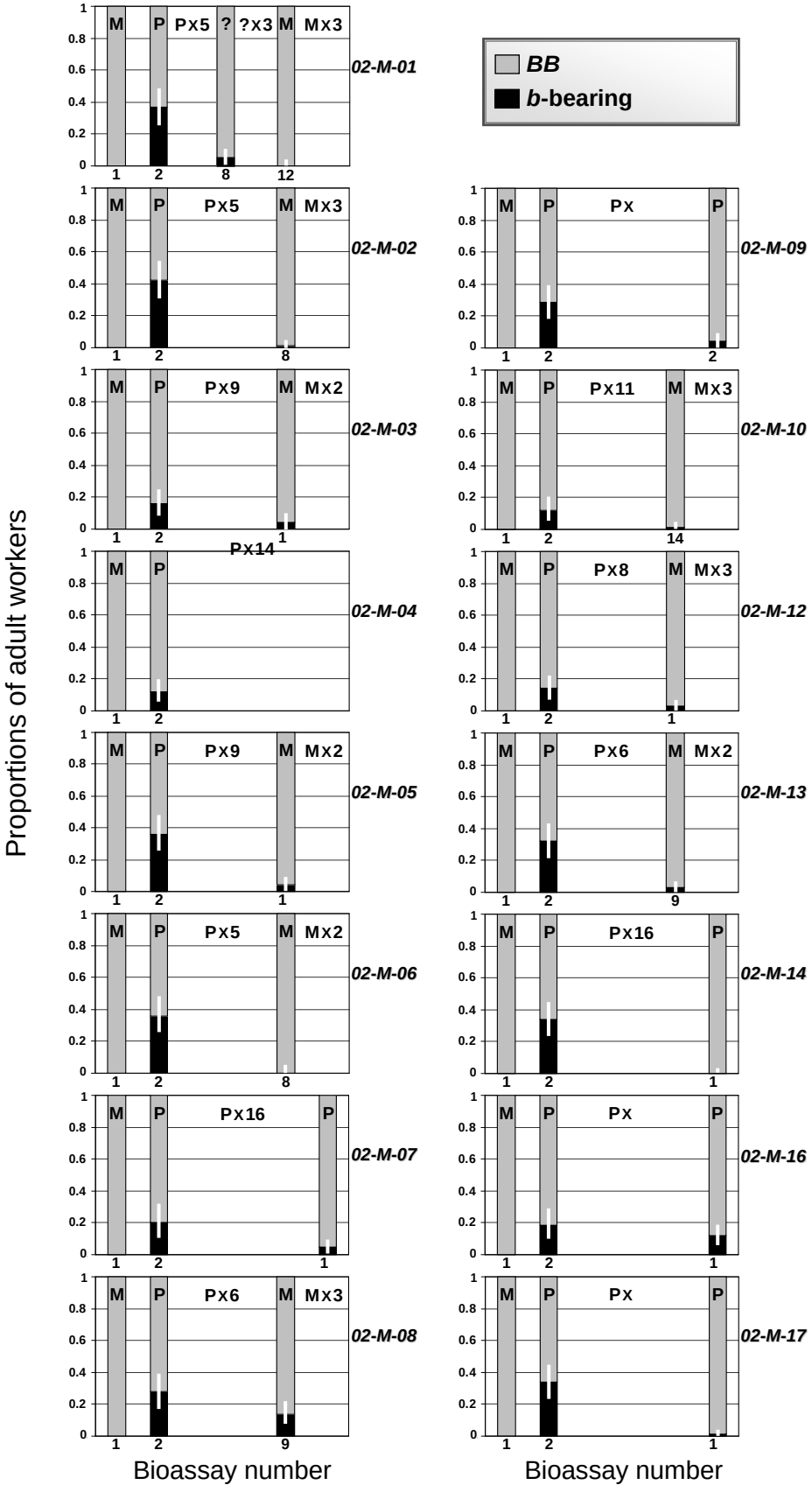
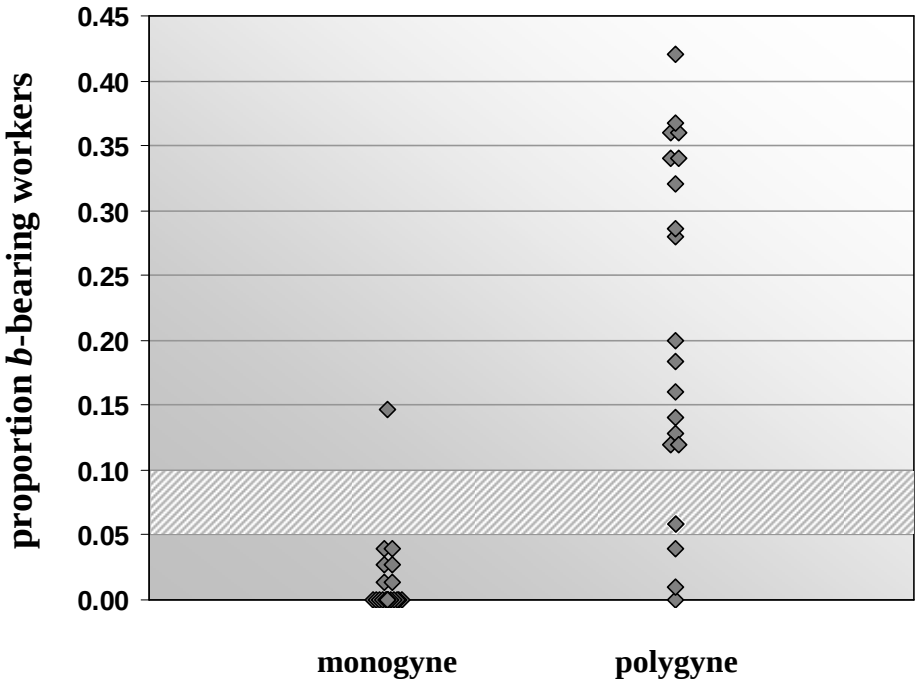

BB
b-bearing
Proportions of adult workers
Bioassay number
Bioassay number
02-M-01
M P Px5 ? ?x3 M Mx3
1 2 8 12
02-M-02
M P Px5 M Mx3
1 2 8
02-M-03
M P Px9 M Mx2
1 2 1
02-M-04
M P Px14
1 2
02-M-05
M P Px9 M Mx2
1 2 1
02-M-06
M P Px5 M Mx2
1 2 8
02-M-07
M P Px16 P
1 2 1
02-M-08
M P Px6 M Mx3
1 2 9
02-M-09
M P Px P
1 2 2
02-M-10
M P Px11 M Mx3
1 2 14
02-M-12
M P Px8 M Mx3
1 2 1
02-M-13
M P Px6 M Mx2
1 2 9
02-M-14
M P Px16 P
1 2 1
02-M-16
M P Px P
1 2 1
02-M-17
M P Px P
1 2 1

Figure 3.2


proportion b-bearing workers
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
monogyne
polygyne

CHAPTER 4

MOLECULAR VARIATION AT A CANDIDATE GENE REGULATING COMPLEX SOCIAL BEHAVIOR IN THE FIRE ANT *SOLENOPSIS INVICTA*[1]

[1] Dietrich Gotzek, DeWayne D. Shoemaker, and Kenneth G. Ross, to be submitted to *Molecular Biology and Evolution*

Abstract.—The fire ant *Solenopsis invicta* displays a profound social polymorphism involving differences in colony queen number. Colonies are headed either by a single reproductive queen (monogyne form) or by multiple queens (polygyne form). This variation in colony social organization is associated with variation at the gene *Gp-9*, with monogyne colonies harboring only *B*-like allelic variants and polygyne colonies always containing *b*-like variants as well. We describe naturally occurring variation at *Gp-9* in *S. invicta* based on the generation of 130 new sequences from samples collected over much of its native range. While there is little overall genetic variation between most of the observed haplotypes, a surprising amount of the variation is found in the coding regions of the gene, with such substitutions usually causing amino acid replacements. We obtained formal evidence that positive selection acted on the basal lineages of the *b*-like allele clade, suggesting the occurrence of episodes of diversifying selection on *Gp-9* coincident with the evolution of polygyny. While our extensive data set revealed considerable paraphyly and polyphyly of *S. invicta* sequences with respect to those of several congeners, monophyly of the *b*-like allele clade was recovered. An expanded analysis of colonies containing alleles of this clade confirmed the invariant link between their presence and the expression of polygyny. Finally, our discovery of several unique haplotypes bearing various combinations of *B*-like and *b*-like codons allowed us to provisionally identify the Met95Ile amino acid replacement as completely predictive of polygyne behavior and potentially causally involved in its expression.

INTRODUCTION

A main goal of evolutionary biology is to document genetic variation and to reconcile observed patterns with population history and demography, fitness consequences, and selection regimes at genes of interest (Gillespie 1991; Hedrick 2000). Study of the adaptive maintenance of molecular variation historically has followed one of two approaches, elucidation of the mechanistic and functional components of molecular adaptations at the biochemical level or description of the historical footprints of selection acting on sequence variants (Golding and Dean 1998). An important objective in modern studies of molecular adaption is to bridge the two approaches by means of comprehensive research integrating functional biochemical and phenotypic data with information on patterns of variation that implicate past selection (Wheat et al. 2006; see also Golding and Dean 1998; Nachman 2005; Phillips 2005; Vasemägi and Primmer 2005).

The fire ant *Solenopsis invicta* displays an important colony-level social polymorphism that is associated with variation at a single gene, *general protein-9* (*Gp-9*) (Ross 1997). Colonies with a single reproductive queen (monogyne colonies) always feature the exclusive presence of the *B* allele of *Gp-9* in all colony members. In contrast, colonies with multiple reproductive queens (polygyne colonies) always have an alternate class of alleles, designated *b*-like alleles, represented along with the *B* allele among colony members (Ross 1997; Ross and Keller 1998; Krieger and Ross 2002, 2005). Remarkably, all reproductive queens in colonies of this latter type invariably bear at least one *b*-like allele. These patterns, coupled with similar genetic compositions of monogyne and polygyne populations at numerous other nuclear genes, have led to the hypothesis that the presence of *b*-like alleles in a colony's workers is both necessary and sufficient to

elicit polygyne social behavior (Ross 1997; Ross and Keller 1998; Ross and Keller 2002, Krieger and Ross 2002; Gotzek and Ross 2007). Because variation in queen number represents a dramatic social polymorphism that is associated with a suit of important reproductive, demographic, and life history differences (Bourke and Franks 1995; Ross and Keller 1995; Tschinkel 2006), variation at *Gp-9* is hypothesized to underlie the expression of major alternative adaptive syndromes in *S. invicta*.

Our understanding of the association of variation at *Gp-9* with expression of colony social organization has been advanced by the production of sequence data, for *S. invicta* in its native (South American) and introduced (USA) ranges, as well as for numerous other *Solenopsis* species (Krieger and Ross 2002, 2005). These studies revealed several important patterns. First, the monophyletic *b*-like alleles are restricted to a clade of six South American fire ant species (including *S. invicta*) that display the monogyne-polygyne polymorphism (this group of species is informally termed the socially polymorphic clade [Pitts et al. 2005]; other South American fire ants appear to exhibit only monogyne behavior). Second, in all of the socially polymorphic species, polygyne colonies always contain *b*-like alleles, presumably because all reproductive queens of this form bear at least one copy of such an allele (Hallar et al. 2006). Third, the *b* alleles, a small clade of *b*-like alleles that apparently has arisen recently in *S. invicta*, feature a radical, charge-changing substitution (Glu151Lys) that may underlie their observed deleterious effects in homozygous condition (Hallar et al. 2006). Finally, the *b*-like alleles of the socially polymorphic species bear diagnostic amino acid residues at codon positions 42, 95, and 139 that distinguish them from all other *Gp-9* alleles (collectively known as *B*-like alleles). This latter finding prompted speculation that the substitutions at

one or more of these positions may alter the function of GP-9 protein with respect to its proposed role in modulating social behavior (Krieger and Ross 2002, 2005).

GP-9 protein is a member of the insect odorant-binding protein family (Krieger and Ross 2002). Several well studied proteins in this family have been implicated as important molecular components of chemoreception in insects, presumably effecting the transduction of pheromones or food chemostimulants to neuronal signals by tranporting these ligands through the chemosensillar lymph to neuronal receptors (Vogt 2005). Regulation of colony queen number in fire ants involves reciprocal chemical signalling and perception between workers and queens, with workers ultimately making decisions about which queens, and how many, are tolerated as reproductives in a colony based on queen pheromonal signatures (Keller and Ross 1998; Ross and Keller 1998). The role of some odorant-binding proteins in chemoreception, the invariant association of one class of *Gp-9* alleles with polygyny, and the restriction of this class of alleles to the socially polymorphic clade of South American fire ants have been viewed as evidence that *Gp-9* may directly influence colony social organization rather than merely being a marker for another gene or genes of major effect on this trait (reviewed in Gotzek and Ross 2007).

Further complexities in our understanding of *Gp-9* and social evolution have arisen as additional sequence data have continued to be generated. For instance, it is now apparent that variation at this gene is not invariably associated with expression of colony social organization within the genus *Solenopsis*, given that the fire ant *S. geminata*, a distant North American relative of *S. invicta*, does not exhibit allelic variation associated with colony social form (Ross et al. 2003). Also, discovery of a *Gp-9* allele with a *b*-like amino acid residue at position 95 but *B*-like residues at positions 42 and 139 in the

undescribed *S.* species "X" indicates that the two classes of alleles in the socially polymorphic species are not as uniformly divergent (internally homogeneous) as previously believed (Krieger and Ross 2005). Unfortunately, the social organization of the source colony for this sequence could not be determined, precluding a crucial test of the importance of residue 95 in the induction of polygyny. These recent results suggest that progress in our understanding of *Gp-9* in fire ants has been hampered to some extent by the limited numbers of samples available for sequencing; indeed, even in the best studied species, *S. invicta*, only a handful of individuals from a single site in the native range (Formosa, Argentina) have been sequenced.

To help remedy this shortcoming, the present study documents sequence variation at *Gp-9* using extensive samples of nominal *S. invicta* collected over a large portion of its vast native range. Our specific objectives were to fully characterize the molecular evolution of *Gp-9* in *S. invicta*, to test for effects of selection on the gene, to confirm the association between polygyny and the presence of variants encoding *b*-like amino acid residues, and to test for phylogeographic patterns in the distribution of the observed variation. We also were particularly interested in finding new variants encoding unique combinations of *B*-like and *b*-like residues at positions 42, 95, and 139, with the hope that their discovery might shed light on the role of each substitution in mediating the expression of social organization. In combination with the other analyses, information from such variant colonies is expected to aid progress in connecting the genetic and phenotypic variation underlying regulation of fire ant social behavior.

MATERIALS AND METHODS

Samples for Gp-9 Sequencing

We obtained samples from several fire ant species of varying phylogenetic relationship to *S. invicta* (Pitts et al. 2005) in order to root the *Gp-9* allele phylogeny, to assess the level of variation in *S. invicta* sequences, and to evaluate the monophyly of *S. invicta* sequences with respect to those of its closest relatives. These samples consisted of a single individual per colony of the following species (numbers of individuals in parentheses): *S. altipunctata* (1), *S. amblychila* (1), *S. aurea* (1), *S. daguerrei* (2), *S. electra* (1), *S. interrupta* (4), *S. macdonaghi* (4), *S. megergates* (6), *S. pusillignis* (3), *S. quinquecuspis* (6), *S. richteri* (8), *S. saevissima* (9), *S. xyloni* (1), and the undescribed *S.* species "X" (2).

Samples of nominal *S. invicta* were obtained from 44 sites distributed over much of the native range (Figure 4.1) as well as 4 sites (Florida, Georgia, Texas, California) in the introduced range in the USA (see Krieger and Ross 2002). Sites in the native range were chosen not only to maximize geographic coverage but also to include all of the genetically differentiated populations detected in earlier studies of neutral nuclear and mtDNA variation (Ross and Shoemaker 2005; Shoemaker et al. 2006; Ross et al. 2007). Multiple colonies (2-13) were sampled at many of the sites (see Fig 4.1), but only a single specimen (sequence) was used from any single colony. Samples were collected during several trips to Argentina and Brazil made between 1988 and 2004. Live specimens were collected directly from colonies then placed immediately on liquid nitrogen for transport back to the laboratory, where they were held in a -80ºC freezer. The social organization of many of the sampled colonies was determined previously by a combination of methods

including discovery of multiple reproductive queens, determination of the number of offspring matrilines using allozyme markers, and detection of *b*-like *Gp-9* alleles (Ross et al. 1997; Mescher et al. 2003; Ross and Shoemaker 2005).

Laboratory Methods

Laboratory methods followed the protocols of Krieger and Ross (2002, 2005), with some modifications. DNA was extracted using the Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN). Polymerase chain reaction (PCR) reactions were set up in 10 μl volumes using 1.1x high fidelity PCR-ready reaction mix (Bio-X-Act Short Mix, Bioline, Randolph, MA) and 0.2 μM primers, with a hotstart thermal cycling regime starting at 95ºC and followed by 35 cycles at 94ºC (20 s), 62ºC (30 s), and 68ºC (1 min 40 s), with a final elongation step at 68ºC (10 min). Primer sequences were those used by Krieger and Ross (2002) (Gp-9/-33 forward: 5'-CATTCAAAGTACAGTAGAATAACTGCC-3', Gp-9_2218 reverse: 5'-CAGGAGTTTGAGTTTGTCACTGC-3'). The approximately 2200-bp amplification products included the full length 1700-bp *Gp-9* gene (containing five exons and four introns) as well as a 500-bp segment of the 3′ flanking region. These products were gel purified (QIAquick Gel Extraction Kit, Qiagen, Valencia, CA) and cloned into pCR2.1 vectors (Invitrogen, Carlsbad, CA), which were then used to transfect competent TOP10F′ *E. coli* cells (Invitrogen). Blue-white screening was used to identify positive clones, which were picked and subjected directly to a hotstart PCR amplification using M13 or the *Gp-9* primers (same conditions as in the previous PCR but using Taq-Pro Complete, Denville Scientific Inc., Meutchen, NJ). The resulting PCR product was

checked for correct length by running it out on an ethidium bromide-stained agarose gel, then it was purified using PEG 8000 (Promega, Madison, WI).

In order to ensure a sufficiently large sample of *b*-like alleles, clones derived from suspected polygyne colonies were screened for such alleles using competitive allele-specific PCR (Imyanitov et al. 2002). Reaction mixes contained 0.13 μM primers, 0.33 μM complementary primer, 1x Taq-Pro Complete, and 0.5 μL of the clone PCR product; PCR was conducted using a cycling regime of 94ºC (2 min), 35 cycles at 94ºC (45 s), 64ºC (45 s), and 72ºC (1 min), with a final elongation step at 72ºC (5 min). The primers used in this allele-specific PCR recognize the single nucleotide substitutions at codons 95 and 139 considered to be diagnostic of all *b*-like alleles (Krieger and Ross 2002; Ross et al. 2003).

Methods for conducting DNA sequencing reactions using internal primers also followed the protocols of Krieger and Ross (2002, 2005). Reactions were performed using the ABI PRISM BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, Calif.), with the products run out in an ABI PRISM 3740xl DNA Sequencer (Applied Biosystems). In light of the considerable overlap of the internal sequencing reads, sequences were not determined in the reverse directions. Critical base calls in phylogenetically important sequences were resequenced to minimize the impact of sequencing errors.

Determination of Colony Social Organization

Colonies of unknown social organization yielding sequences that encoded one or more of the residues considered diagnostic for all *b*-like alleles (Gly^{42}, Ile^{95}, Ile^{139}) were

subjected to microsatellite analyses to learn whether these colonies were monogyne or polygyne. Genotypes at three loci (*Sol-42*, *Sol-49*, and *Sol-55*) were scored for ten workers from each colony (methods in Krieger and Keller 1997; Shoemaker et al. 2006). PCR products were visualized using an ABI PRISM 3740xl DNA Sequencer (Applied Biosystems). Queens of native *S. invicta* normally mate only once (Ross et al. 1993, 1997), so the presence of more than three alleles at a locus among a colony's workers indicates the presence of multiple offspring matrilines. The presence of multiple offspring matrilines signifies polygyny, whereas the presence of a single matriline signifies monogyny.

Genetic Analyses

All sequences were readily aligned by hand. The alignment was tested for evidence of recombination using a variety of methods (following the recommendations of Posada and Crandall 2001, Wiuf et al. 2001, and Posada 2002). The DSS and PDM (McGuire et al. 1997; McGuire and Wright 2000) methods were implemented using the program TOPALi (Milne et al. 2004).

Differentiation among *Gp-9* haplotypes in their nucleotide composition was tested using homogeneity χ^2 analysis (implemented in the program PAUP*; Swofford 2004) and visual inspection (implemented in the program SeqVis; Ho et al. 2006). Nonrandom codon usage was tested using the program DNASP 4.10.9 (Rozas et al. 2003), with Yates' correction for the observed G+C content employed.

Several measures of sequence variation were estimated for the *Gp-9* haplotypes of nominal *S. invicta*. The total uncorrected number of nucleotide substitutions (*d*) and

proportion of nucleotide substitutions (*p*) were calculated separately for the coding and non-coding regions of all unique haplotypes using the program MEGA 3.1 (Kumar et al. 2004). To measure the levels of observed nucleotide and amino acid variation in the coding region of *S. invicta* sequences, we estimated two additional diversity indices, the codon diversity and amino acid diversity (Krieger and Ross 2005). These two indices were compared across all variable codon positions in the mature GP-9 protein in order to determine the extent to which codon variation translates into amino acid replacements. All of these analyses were performed separately for the *b*-like alleles, *B*-like alleles, and all alleles combined.

Phylogenetic Analyses

Only non-identical *Gp-9* sequences were used in the phylogenetic analyses. Among-haplotype heterogeneity in nucleotide composition, the presence of which could cause errors in phylogeny estimation (Lockhart et al. 1994, Tarrio et al. 2001, Jermiin et al. 2004), was not statistically significant for our data set (see Results). We assessed the potential impact of among-species compositional differences by constructing preliminary phylogenies with the Neighbor-Joining method (Saitou and Nei 1987) using either the minimum evolution (ME) criterion with LogDet (Lockhart et al. 1994) or the maximum likelihood (ML) distances between haplotypes (with all parameters estimated from the data). The resulting two trees were compared using the SH test (Shimodaira and Hasegawa 1999).

Due to the prohibitive computational time required for even a single heuristic search under the maximum parsimony (MP) criterion, we employed the parsimony ratchet

method (Nixon 1999); this approach was implemented by means of the program PAUPRat (Sikes and Lewis 2001) using 500 repetitions and randomly perturbing 25% of the characters for each re-weighting. The analysis was repeated ten times to ensure that tree space had been adequately searched, then repeated another ten times while considering gaps as character states.

Because the MP and ME phylogenies did not differ significantly according to an SH test, we estimated the best fitting model of *Gp-9* nucleotide evolution using the better resolved ME tree with the program Modeltest 3.7 (Posada and Crandall 1998). We selected the most appropriate models for the complete data set and various partitions of it (non-coding regions; coding regions; 1^{st}, 2^{nd}, and 3^{rd} codon positions) using the Aikake information criterion (AIC, Aikake 1974) and Bayesian information criterion (BIC, Schwarz 1978) (Posada and Buckley 2004).

Finally, we conducted four independent Markov chain Monte Carlo (MCMC) tree searches under the Bayesian posterior probability optimality criterion (Bayesian inference; BI) using the program MrBayes 3.1 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). Multiple analyses were run to ensure adequate exploration of the optimal tree space (Larget 2005; Ronquist et al. 2005). Five parallel chains, four of which were heated incrementally (temperature=0.1), were started from random trees for each analysis, with initial parameter values based on the evolutionary model selected by Modeltest. The chains were run for two million generations, with sampling every 100 generations. Stationarity of the chains was ascertained visually by plotting sample log-likelihoods through the course of each run, as well as by examining the convergence diagnostics (the potential scale reduction factor for all parameters

approached 1.0 at stationarity) (Ronquist et al. 2005). Pre-stationarity MCMC samples were discarded as burn-in (usually, around the first 700 samples), and the model parameters and tree topology were estimated using the remaining samples. The log-likelihoods, substitution models, and tree topologies were compared among independent runs using an SH test. After ensuring that all runs had converged on the same area in tree/parameter space, the samples from the four different runs were combined for final analysis.

Selection Analyses

Two general types of analyses were employed to test for positive selection on *Gp-9*, random-effects and counting analyses (Kosakovsky Pond and Frost 2005a). Fixed-effects analyses were not employed because they are too computationally intensive for the number of sequences in our data set. Also, they require a priori designation of sites evolving under different selective regimes (Kosakovsky Pond and Frost 2005a); because we intended to use our data set for a largely independent test of the findings of Krieger and Ross (2002, 2005), we wished to avoid biasing the results by focusing on specific sites previously identified as being under selection.

We employed two different random-effects methods, which fit a distribution of substitution rates across sites and then infer the rate at which each site evolves (Nielsen and Yang 1998). Because of computational limitations, the Bayesian method (Huelsenbeck and Dyer 2004) was conducted on a greatly reduced data set (25 exemplar sequences representing all major clades in the BI phylogeny) using the program MrBayes. Coding and non-coding sites were separated into unlinked partitions for the

analysis. Selection on the coding partition was estimated according to the M3 codon model (Yang et al. 2000), which is less restrictive than the commonly used Nielsen and Yang (1998) model (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). MCMC chains were run for one million generations, with sampling every 100 generations. Otherwise, model specifications followed those used in the BI phylogeny estimation.

The second random-effects method used was a maximum likelihood branch-site-specific model (Yang and Nielsen 2002, Zhang et al. 2005). We removed all non-coding sites from the data set, then used the BI phylogeny as a guide to prune additional sequences that were sisters to sequences with identical coding regions or occurred in unresolved clades of identical coding sequences (a single exemplar was retained). This resulted in a reduced data set of 92 sequences. The CODEML program in the PAML 3.15 package (Yang 1997) was used to run the branch-site-specific model A, test 2 (Zhang et al. 2005), which infers selection on codons along specified branches. We used the pruned BI phylogeny (with branch lengths estimated under the M3 model) to test selection on the stem lineages of both the *b*-like allele clade and the *b* allele clade, as well as on the internal branches of each of these clades. We adopted the Bayes empirical Bayes (Yang et al. 2005) approach in place of the naïve empirical Bayes approach (Nielsen and Yang 1998) to identify sites under selection (Scheffler and Seoighe 2005; Yang 2005). The analysis incorporated the F3x4MG model of codon substitution (Muse and Gaut 1994), where the substitution rate is proportional to the frequency of the target nucleotide rather than the the target codon, because it yielded better likelihood scores

than the F3x4NY model (Nielsen and Yang 1998). We ran each analysis three times, each with different initial values for the model parameters ω and κ.

We employed three different counting methods based on the Suzuki-Gojobori approach (Suzuki and Gojobori 1999). These methods estimate the number of nonsynonymous and synonymous substitutions at each codon position, then test for significant differences between the number of nonsynonymous changes per nonsynonymous site (*dN*) and number of synonymous changes per synonymous site (*dS*), a hallmark of selection. The reduced data set created for the branch-site-specific random-effects method was used also for the counting methods. The first counting method was the Single Likelihood Ancestor Counting (SLAC) analysis implemented in the program Datamonkey (Kosakovsky Pond and Frost 2005b). This analysis used the pruned BI phylogeny and HKY model of nucleotide substitution. The global *dN*/*dS* ratio was estimated along with 95% confidence intervals, and ambiguous characters were resolved according to the most likely solution given by the model. We used a nominal α-level of 0.25 to identify statistically significant bouts of selection because the SLAC method tends to be conservative (Kosakovsky Pond and Frost 2005a).

The second counting method we used follows more closely the Suzuki and Gojobori (1999) approach. This analysis was conducted in association with the branch-site-specific random-effects analysis using the program CODEML (Yang 1997).

As a third counting method for detecting selection, we used the Zhang et al. (1997) method adapted from the approach of Messier and Stewart (1997). This approach uses reconstructed ancestral sequences to test the null hypothesis of neutral evolution (*dN*=*dS*) along each branch of the inferred phylogeny by means of Fisher's exact tests. To

maximize the statistical power of this test, we pooled non-coding sites with coding-region synonymous sites (e.g., Rooney and Zhang 1999) after determining that there was no significant difference in substitution rates between these partitions (Fisher's exact test, P=0.115). Ancestral sequences were reconstructed using the baseml program in PAML, with *dN* and *dS* for each branch estimated using the "free-ratio" model (Nielsen and Yang 1998). We did not test for negative selection using this method.

Phylogeographic Analyses

Evidence for geographical restriction of related *Gp-9* haplotypes in native *S. invicta* was examined by conducting a series of Analysis of Molecular Variance (AMOVA) analyses (Excoffier et al. 1992) using the program ARLEQUIN (Schneider et al. 2000). This procedure partitions total genetic variation among the different sampling sites or clusters of sites in order to reveal hierarchical patterns of spatial differentiation. Only *B*-like alleles were included in order to avoid any effect of spatial restriction of polygyny on the results (e.g., Mescher et al. 2003). Genetic distances between *Gp-9* haplotypes were estimated as the squares of pairwise sequence differences. In an initial analysis, all 40 sites containing at least one *B*-like haplotype were clustered arbitrarily into ten regional groups (these groups occupied 100-300 km diameter areas). In a second analysis, only the 14 sites for which three or more sequences were available were used. In this case, goups of sites were clustered on the basis of patterns of regional differentiation previously detected at 14 presumed neutral nuclear loci (these groups typically occupied 100-600 km diameter areas; Ross et al. 2007). Parallel AMOVA analyses using the neutral nuclear data from the same 14 sites were conducted to provide a direct

comparison with the *Gp-9* results from the second analysis. All alleles at each neutral locus were assumed to be equally related to one another (i.e., an infinite alleles model of mutation was assumed). Statistical significance of genetic differentiation among sites or clusters of sites was determined by permuting haplotypes across individuals (20,000 replicates) for *Gp-9* or bootstrapping over loci (10,000 replicates) for the neutral markers.

Isolation-by-distance analyses were conducted for the *S. invicta* sequences to learn whether differentiation between sites in their *Gp-9* haplotype composition increases in parallel with their geographic separation. Only the 40 sites from which one or more *B*-like alleles were sampled were considered. The program GENEPOP (Raymond and Rousset 1995) was used to examine the relationship of Nei's net number of nucleotide differences (D_A; Nei and Li 1979) with the natural logarithm of geographic distances between sites (see Slatkin 1993; Rousset 1997). Significance of isolation-by-distance relationships was determined by means of Mantel tests based on 10,000 data permutations coupled with estimation of Spearman rank correlation coefficients.

RESULTS

Variation at Gp-9

The complete data set consisted of 185 full-length sequences (149 newly generated), of which 136 were from the focal species, *S. invicta*. For the entire data set, a consistent A+T bias in base composition was found, which is more substantial in the non-coding regions (0.816) than in the coding regions (0.574). No evidence of nonrandom codon usage was found over the entire set of sequences (scaled χ^2=0.104 using Yates correction;

effective number of codons=57.07 out of a maximum of 61; codon bias index=0.348). Also, there is no evidence for intragenic recombination having occurred (DDS and LRT: all potential recombination events below 95th percentile significance level [DDS: 84%; LRT: 46%]). This latter result parallels the lack of evidence for recombination at *Gp-9* reported by Krieger and Ross (2005) for a smaller data set from a more diverse set of *Solenopsis* species.

A total of 164 unique *Gp-9* haplotypes were identified, of which 121 were recovered from *S. invicta* specimens. This is a large increase over the six haplotypes previously described from the relatively few *S. invicta* sampled throughout the introduced USA range and from a single locality in Argentina (Krieger and Ross 2002, 2005). The great majority of *Gp-9* sequences in the complete fire ant data set (91%) were represented as singletons. Among the 13 haplotypes recovered from more than one specimen, three are shared between *S. invicta* and ants identified as belonging to the closely related species *S. quinquecuspis* or *S. megergates* (see Figure 4.4).

Most nucleotide sites (1976 of the 2321 in the aligned sequences, 85%) are invariant across all *Gp-9* haplotypes from the various *Solenopsis* species. Among the variable sites, most (66%) occur in the non-coding regions. However, because these regions encompass 80% of the total sequence length, proportionately more variable sites occur in the coding regions (25%) than in the non-coding regions (12%) across the study species. Of the 345 total variable sites, 134 are parsimony-informative.

Considering only *Gp-9* in *S. invicta*, the great majority of unique haplotypes are very similar to one another at the nucleotide sequence level; indeed, most haplotypes within the *B*-like and *b*-like allele classes differ by fewer than a dozen point substitutions (Figure

4.2). The two most divergent *B*-like alleles differ at only 36 of their nucleotides (1.5%), while the two most divergent *b*-like alleles differ at half that number. Considering both classes combined, an additional peak at 15-20 substitutions attributable to differences between *B*-like and *b*-like alleles becomes apparent (Figure. 4.2).

Summary diversity statistics for *Gp-9* in *S. invicta* are presented in Table 1. The mean number of nucleotide substitutions between pairs of haplotypes (*d*) varies from about three for the coding regions in *B*-like alleles to almost seven in the non-coding regions when all alleles are combined. In parallel with the observed patterns of site variation across sequences from all of the study species, the mean proportions of nucleotide substitutions (*p*) between *S. invicta* coding-region sequences consistently exceed those between non-coding sequences. For the *b*-like alleles and for all alleles combined, an approximately 3-fold excess of coding-region substitutions exists. Coding-region *p* estimates for the *B*-like and *b*-like classes considered separately are only about half the estimate for all alleles combined.

Within the coding regions, over 40% of codons and 35% of amino acids are variable across all *S. invicta* alleles (Table 1). Somewhat lower diversity for just the *B*-like alleles is indicated by these metrics, while substantially lower variable proportions are evident for the *b*-like alleles. The similar values for the two metrics within each allele class suggest that the great majority of coding-region nucleotide substitutions have yielded amino acid replacements; indeed, around 80% of all variable codon positions are also variable at the amino acid level in all three allele sets. The prevalence of nonsynonymous substitutions in the coding regions of *Gp-9* from *S. invicta* is reflected also in the similarities of the codon diversity and amino acid diversity estimates for all three sets of

alleles (Table 1). Considering all alleles together, about 90% of codon positions exhibit identical codon and amino acid diversities, indicating that every observed nucleotide substitution at these locations caused an amino acid replacement.

The numbers of different amino acids occurring at each codon position are depicted for the *B*-like and *b*-like allele classes from *S. invicta* in Figure 4.3. No consistent "hotspots" of residue variation are apparent across both allele types. Only seven positions harbor three or more residues within an allele class; two of these (positions 39 and 117) are implicated by the formal selection analyses as being under positive selection (see below).

An important finding of our survey is the existence in *S. invicta* of considerable additional amino acid variation across the three codon positions regarded as bearing residues diagnostic of the *B*-like and *b*-like allele classes in the socially polymorphic clade. Previously, all alleles of the *b*-like class have been found to encode Gly^{42}, Ile^{95}, and Ile^{139} residues, whereas virtually all alleles of the *B*-like class have been found to encode Ser^{42}, Met^{95}, and Val^{139} residues (Krieger and Ross 2002, 2005). Four sequences from native *S. invicta* encode amino acids at these three positions that do not conform to the previous patterns of association within each allele class, in that they feature some combination of *b*-like residues at one or two positions and *B*-like residues at the remaining position(s) (see Table 2). These sequences are discussed further below with respect to the form of social organization of the source colonies.

Aside from nucleotide and amino acid substitutions, a single previously unknown structural change in *Gp-9* also was detected. A sequence from an *S. invicta* colony in Santiago del Estero, Argentina, carries a unique point mutation in exon 5 that

transformed the stop codon (TAA) into a glutamine-encoding (CAA) codon, thereby extending the C-terminal tail of the resulting protein by 22 amino acids.

Association of polygyny with b-*like* Gp-9 *Variants in* S. invicta

All 29 native *S. invicta* colonies from which alleles were recovered that encode all three diagnostic *b*-like residues (Gly^{42}, Ile^{95}, Ile^{139}) were shown by microsatellite analysis to contain multiple offspring matrilines. Thus, we confirm the results of previous surveys, based on far fewer sequences or on allele-specific PCR assays (Kreiger and Ross 2002; Mescher et al. 2003), that possession of such alleles by a colony's workers invariably is linked to the expression of polygyne social organization.

Four additional *S. invicta* colonies yielded novel *Gp-9* variants that encode some combination of *b*-like and *B*-like residues at these three positions (Table 2). The social organization of three of these colonies was inferred by microsatellite analysis to be monogyny, whereas the fourth colony was inferred to be polygyne. Only codon 95 was completely predictive of the social organization of these colonies. Thus, we hypothesize that only this residue is obligately associated with colony social form in the socially polymorphic South American fire ants.

Phylogenetic Relationships of Gp-9 *Alleles*

Analyses of the complete data set revealed no evidence of compositional heterogeneity among species that might influence the phylogenetic analyses (χ^2=21.464, df=522, *P*=1.00; SH test comparing ME trees based on logDet and ML distances: Δ-lnL=7.543, *P*=0.397). All phylogenetic hypotheses recovered in this study by the differing methods are compatible with one another. However, the SH test identified the

poorly resolved MP tree as significantly worse than the BI tree (Δ-lnL=149.233, *P*=0.015), whereas the ME and BI trees did not differ significantly (Δ-lnL=36.944, *P*=0.376). The poor overall node resolution obtained with the MP analysis is consistent with the relatively few parsimony-informative sites across the 164 unique sequences. The relatively high standard deviations of split frequencies in the BI analyses (~0.085) also suggest limited information content of our data set, which results in an inability of the search algorithms to resolve some parts of the phylogeny with confidence (e.g., Ronquist et al. 2005).

The phylogenetic hypothesis produced by BI is shown as Figure 4.4. Three major *Gp-9* allele clades of unresolved relationship to one another are apparent. Clade I contains sequences from the North American fire ants *S. amblychila*, *S. aurea*, and *S. xyloni*, as well as from three other *Solenopsis* species also considered to be distant relatives of *S. invicta* (Pitts et al. 2005). Clade II contains sequences from an assortment of South American fire ant species more or less closely related to *S. invicta*, as well as two *S. invicta* sequences. Clade III contains two relatively well supported lineages. One includes three *S. invicta* sequences as well as sequences from two quite distant South American fire ant relatives of *S. invicta* (*S. altipunctata* and *S. saevissima*). The second, large lineage (Clade IIIa) contains only *Gp-9* alleles of the socially polymorphic South American fire ants (*S. invicta* and its close relatives *S. richteri*, *S. megergates*, *S. quinquecuspis*, *S. macdonaghi*, and *S.* species "X"). The relationships of the major *Gp-9* lineages depicted in Figure 4.4 reflect closely the relationships inferred by Krieger and Ross (2005) from a much smaller data set. Importantly, the previously detected paraphyly and polyphyly of *S. invicta Gp-9* alleles with respect to those of the other

socially polymorphic species (Krieger and Ross 2002, 2005) are extended to even more distantly related fire ant species in our analysis.

Within the lineage comprising exclusively alleles from the socially polymorphic species (Clade IIIa), a clade composed almost entirely of sequences from *S. richteri* forms a sister lineage to the remaining sequences. This evidence for a basal position of *S. richteri* alleles within the group of alleles of all socially polymorphic species again is consistent with earlier conclusions based on far smaller data sets (Krieger and Ross 2002, 2005).

Significantly, the *b*-like allele clade is recovered once again in our study, confirming a monophyletic origin of *Gp-9* variants associated with polygyny in the South American fire ants. As perhaps expected given our expanded sampling effort, support for such a clade is lower than in previous studies and, moreover, varies depending on how the clade is defined. If defined only by possession of the Ile^{139} residue (the first *b*-like substitution to appear in the lineage), the clade is supported by a posterior probability value <0.5. If defined by possession of all three residues previously considered diagnostic for the *b*-like allele class, the posterior probability value increases to 0.5 or greater (the Val^{139} residue from sampled colony SC665 apparently represents a reversal). Given the evidence that only residue Ile^{95} can be considered to be invariably associated with polygyny (Table 2), the *b*-like clade could be defined relatively more narrowly by possession of the codon for this residue (this would exclude from the group the allele from colony LP719 encoding Ile^{139} but not Ile^{95}).

Finally, our data verify that the *b* alleles of *S. invicta* form a relatively recently derived monophyletic group within the *b*-like clade. This confirms earlier evidence that

the radical, charge-changing Glu151Lys substitution characterizing these alleles occurred only once, presumably in an ancestral *S. invicta* population from northeastern Argentina or southeastern Brazil (where the alleles currently predominate; see Figure 4.1).

Selection on Gp-9

All of the selection analyses we undertook yielded evidence of selection of some form at various codon positions in *Gp-9* and on various branches of the allele phylogeny. Among the site-specific methods, the SLAC counting method identified three positively selected codon positions, while the Suzuki-Gojobori counting method identified only a single, separate position (Table 3). The former method also detected negative selection on position 32, while both methods detected such selection on position 99. The Bayesian random-effects method detected ten positively selected positions with very high confidence (posterior probability>95%) and another four with less confidence (posterior probability 90-95%) (Table 3). The observed tendency of the counting methods to produce more conservative results than the random-effects method has been reported previously (Kosakovsky Pond and Frost 2005a).

Considering branch-specific selection, the counting method of Zhang et al. (1997) revealed evidence of positive selection along three branches of the *Gp-9* phylogeny, two of which are within the *b*-like allele clade. The stem lineage of a broadly defined *b*-like clade including the allele from colony LP719 was not identified as experiencing selection (dN/dS=1/0), whereas the two successive descendant branches at the bases of more narrowly defined *b*-like clades were inferred to be under positive selection (dN/dS=3/0 [P=0.004] and dN/dS=2/0 [P=0.026], respectively). The third branch under positive

selection represents a relatively basal, well supported lineage within Clade IIIa consisting of one haplotype each from *S. invicta* and *S. macdonaghi* (see Figure 4.4) (dN/dS=4/1 [P=0.003]).

When we applied the branch-site-specific random-effects method to the *b*-like clade (broadly defined), it yielded significantly lower log-likelihoods for the null hypothesis of no selection than for the hypothesis of positive selection both on the stem lineage and over the internal branches of the clade. Despite this, dN/dS does not differ significantly from one on the stem (dN/dS=1.05, P=0.75) but considerably exceeds one over the interior branches (dN/dS=17.5, P<0.00001). The branch-site-specific Bayes empirical Bayes approach identified with confidence two positions under positive selection on these interior branches (39 and 117), both of which also were identified by two of the site-specific methods applied over the whole tree (Table 3).

When we applied the branch-site-specific analysis to just the clade of *b* alleles (those *b*-like alleles bearing a charge-changing amino acid substitution at position 151), no significant signature of selection was detected on its stem lineage nor on its internal branches, despite dN/dS=2.57 and ~200, respectively.

Phylogeography of Gp-9 *Variants in* S. invicta

Our sequence data confirm the occurrence of *b*-like alleles at all five sampling sites in the south-central portion of the native range previously reported to contain such alleles following allele-specific PCR analyses (Mescher et al. 2003). In addition, we discovered alleles of this class at several other sites well outside of their previously known area of occurrence (Corumba, Coxim, La Paz, and Suncho Corral; see Figure 4.1). These results

suggest that while polygyny may be concentrated in northern Argentina and southeastern Brazil, it is likely to occur at some frequency through much of the native range of *S. invicta*.

The initial AMOVA analysis of *B*-like alleles revealed that no detectable variation occurs among the ten arbitrarily clustered groups of 40 sites once the 30% of total variation found to reside among sites within groups is taken into account (significance of among-site differentiation; $P<0.001$). Similar results were obtained in a second analysis of 14 sites with at least three sequences that were clustered according to patterns of regional differentiation at several neutral genes; no variation occurs among the regional groups, but 21% of the total variation occurs among sites within groups ($P=0.001$ for among-site differentiation). As a comparison with the results of this second *Gp-9* analysis, 14% of the total variation at 14 neutral nuclear genes occurs among the regional groups, while 10% resides among sites within groups ($P<0.001$ for differentiation at both levels). Thus, *Gp-9* in the monogyne form of native *S. invicta* appears to exhibit somewhat stronger differentiation at very local scales than neutral markers of the nuclear genome but much weaker differentiation at broader geographic scales covering hundreds of kilometers. This lack of higher-level structure for *Gp-9* is evident as well from the fact that closely related *B*-like alleles frequently occur in widely separated local populations (data not shown).

No significant pattern of isolation-by-distance was detected using Nei's D_A values for the *B*-like alleles at 40 sites (one-tailed $P=0.087$). Thus, differentiation in *Gp-9* composition does not increase in parallel with geographic separation of sites, consistent with the absence of detectable structure at a regional scale. This finding again contrasts

with the strong isolation-by-distance patterns detected using neutral markers in native *S. invicta* populations (Ross et al. 2007).

DISCUSSION

The objective of this study was to survey naturally occurring molecular variation at *Gp-9*, a candidate gene of major effect on the expression of fire ant colony social organization, in *Solenopsis invicta*. Patterns of observed sequence variation at this gene were examined with the following specific goals: i) to identify the major mechanistic factors generating variation at the gene, ii) to reconstruct the evolutionary relationships of variant haplotypes from *S. invicta* and related fire ant species, iii) to examine the historical role of selection in shaping the variation, iv) to learn whether any single candidate amino acid residue in GP-9 protein is completely predictive of social organization, and v) to examine the geographic distribution of the observed variation. The motivation of the work was to help bridge the gap between functional biochemical information and molecular population genetic data in order to construct a cogent evolutionary narrative of the genetic underpinnings of a major social adaptation (see Gotzek and Ross 2007).

Our survey of samples collected over much of the native range of *S. invicta* succeeded in uncovering quite extensive *Gp-9* haplotype variation, with over 120 different sequences recovered. This represents a large increase over the six haplotypes previously documented from this species (Krieger and Ross 2002, 2005). Intragenic recombination was found to play little or no role in generating the observed sequence

variation at *Gp-9*, a conclusion reached also by Krieger and Ross (2005) for sequences obtained from a wide diversity of *Solenopsis* species.

The exon/intron structure of all observed *Gp-9* variants of *S. invicta* and our other study species is identical to that reported previously (see Figure 4.3; Krieger and Ross 2002, 2005), with one remarkable exception. A single *S. invicta* sequence from north-central Argentina contained a nonsynonymous nucleotide substitution in codon 154 that transformed it from a stop codon to a glutamine-encoding codon, thereby extending the C-terminal tail of the GP-9 protein by 22 amino acids. Two of these supernumerary residues are basic (K^{159}, H^{175}) while none is acidic, so that the mutant protein is likely to have a charge change mirroring or exceeding that of the proteins encoded by the distantly related *b* alleles of *S. invicta* (which have a K^{151} replacement). The charge change in the C-terminus of the *b*-encoded proteins is associated with recessive deleterious (lethal) effects not found in other *b*-like alleles (Hallar et al. 2007); these may stem from changes in the ligand binding/unloading properties or in the ability of the protein to form biologically active dimers, judging from the fact that the C-termini of odorant-binding proteins seem to be involved in these functions (Krieger 2005). Demonstration of similar deleterious effects of the elongated mutant protein could pave the way for functional experiments intended to clarify some of the basic biochemical features of GP-9 protein.

Only relatively modest nucleotide sequence variation was observed among the large number of *Gp-9* haplotypes recovered from *S. invicta*. Most pairs of haplotypes within the *B*-like or *b*-like allele classes differ by fewer than a dozen point substitutions across the 2300 bp sequence alignment, and the most divergent haplotypes in the entire data set differ at just 38 (1.6%) of their sites. This variation appears to be distributed unevenly

between the coding and non-coding regions of the gene, with the mean proportion of nucleotide substitutions (*p*) elevated as much as 3-fold in the coding regions. A similar trend is reflected across all of the species examined, where one-fourth of the coding-region nucleotide sites are variable but only half that proportion of non-coding sites are variable. Moreover, about 90% of the coding-region nucleotide variation in *S. invicta* results in amino acid replacements.

The resulting picture of *Gp-9* sequence evolution in *S. invicta* is that relatively few point mutations have accumulated across the gene during its history in the lineage, but a high proportion of these mutations occurred in the coding regions. Most such nucleotide substitutions led to amino acid replacements, so that over one-third of codon positions now have variable residues across all *S. invicta* haplotypes (e.g., Figure 4.3). This pattern is consistent with a general lack of negative selection acting to constrain amino acid replacements over much of the length of the protein, as inferred also for other insect odorant-binding proteins from their low amino acid sequence identities (the primary structure of these proteins evidently can be highly variable as long as the tertiary structure is maintained; Nagnan-Le Meillour and Jacquin-Joly 2003, Vogt 2005). This pattern might also be construed as consistent with the historical action of positive, diversifying selection on *Gp-9*.

More direct evidence of some role for positive selection on *Gp-9* in fire ants comes from the formal selection analyses we conducted. Several different codon positions were identified by various site-specific methods as having significantly elevated rates of nonsynonymous over synonymous substitutions across the species, and two consistently identified positions (39 and 117) were implicated as well by a branch-site-specific

method applied to the *b*-like clade. Position 117 also was identified by Krieger and Ross (2005) as subject to positive selection based on a more diverse set of *Solenopsis Gp-9* sequences. Neither of these two consistently identified positions appears to be in the binding cavity or C-terminal tail of GP-9 protein based on earlier structure prediction analyses (Krieger and Ross 2005). Thus, we cannot speculate about what physiological or other traits, if any, may have been affected by the amino acid replacements at these locations.

Perhaps more compelling evidence for positive selection on *Gp-9* comes from the branch-specific and branch-site-specific methods. These analyses consistently revealed elevated rates of amino acid replacement throughout the *b*-like clade of alleles associated with polygyny in the socially polymorphic species. In the early evolution of this clade, five nonsynonymous substitutions are inferred to have occurred on a background of zero synonymous substitutions, whereas the inferred ratio of the two types of substitutions over the entire history of the clade exceeds 17. Similar results were obtained from the earlier selection analyses of Krieger and Ross (2002, 2005). Thus, our study adds to the evidence that selection has played some positive creative role in the molecular evolution of *Gp-9* in fire ants, presumably in concert with the origin and elaboration of polygyne social behavior.

Two of the site-specific methods also yielded evidence for a single position being under negative selection, codon 99. This position is predicted to occur in the binding cavity of GP-9 (Krieger and Ross 2005), and so may represent a rare example where any variation in the encoded residue negatively impacts the binding capability and, hence, biochemical function of the protein.

Among the coding-region diversity uncovered in our survey was substantial variation across the three codon positions earlier believed to be diagnostic between *B*-like and *b*-like alleles (42, 95, 139) and, thus, considered wholly predictive of colony social organization (Krieger and Ross 2002). Rather than encoding distinctive suites of residues at these three positions, the four newly detected variants encode various combinations of *B*-like and *b*-like residues. These variants are important for several reasons. Based on our determination of the social organization of the four source colonies, we tentatively conclude that only the residue at position 95 is predictive of social behavior. By extension, studies aimed at resolving the role of *Gp-9* in mediating the expression of social behavior now should focus on the the biochemical, physiological and behavioral consequences of the polymorphism at this single site. The discovery of these new variants also has implications for how the *b*-like allele clade is defined; we propose that membership in the clade now be restricted to include only descendants of the variant in which the Ile^{95} residue first arose (this would exclude the sequence from colony LP719 in our study). Finally, the existence of the novel sequences from colonies Pi21 and O40 indicates that the *b*-like Gly^{42} residue has arisen independently on at least two occasions, given that the novel sequences fall far outside the *b*-like clade. An analogous finding concerns residue Ile^{95}, which apparently has arisen not only at the base of the *b*-like clade but also in a rather divergent lineage in *S.* species "X" (recovered from a colony of unknown social organization; Krieger and Ross 2005).

Our proposal that the Met95Ile amino acid replacement is completely associated with polygyne behavior in the socially polymorphic species, and potentially causally involved in its expression, contradicts the conclusion of Krieger (2005) and Krieger and Ross

(2005) that Val139Ile was the crucial replacement. This previous conclusion was based primarily on the inference from protein structure modeling that residue 139 forms part of the ligand-binding cavity. However, residue 95 also is predicted to lie in the binding cavity based on the same analysis. No study to date has identified position 139 as being under selection, whereas the Bayesian random-effects test implemented in this study implicated positive selection on position 95. (Position 42 also was identified by this method as subject to such selection. Given the lack of association of the *b*-like Gly^{42} residue with polygyny in the socially polymorphic species, as well as the occurrence of codons for this residue in sequences from fire ants outside of the socially polymorphic clade, this residue may influence some aspect of protein function unrelated to regulation of colony queen number.)

A previous survey of *Gp-9* polymorphism in *S. invicta* based on allele-specific PCR indicated that the *b*-like allele clade (and, by extension, polygyny) is restricted to the south-central portion of this species' range (Mescher et al. 2003). Our far more extensive sampling has extended the known area over which these alleles occur considerably northward. Nonetheless, in view of the fact that the ranges of the other socially polymorphic species harboring *b*-like alleles are restricted to eastern Argentina, Uruguay, and southeastern Brazil (Pitts et al. 2007), an origin for the *b*-like alleles of the South American fire ants in this area seems probable.

We found no indication of significant higher-level (regional) structure in the geographic distribution of *B*-like variants of *Gp-9* in *S. invicta*, although local populations are highly differentiated from one another. This pattern stands in contrast to the striking regional differentiation observed for numerous other, presumably neutral, nuclear loci.

One possibility for the difference is that, by chance, the distribution of *Gp-9* variation does not closely track that of the remaining nuclear genome because of the probabilistic nature of the time to reciprocal monophyly of allele lineages (e.g., Rosenberg 2003). Another possibility is that sporadic interspecific hybridization in different areas followed by introgression of heterospecific *Gp-9* alleles has broken down any geographic pattern that may have developed due to restricted inter-regional gene flow. Several lines of evidence support the plausibility of this latter scenario. First, the *B*-like alleles of *S. invicta* are extensively paraphyletic or polyphyletic with respect to the alleles of several other fire ants, including some species regarded as quite distant relatives (Pitts et al. 2005). Second, the *S. invicta* sequences that are polyphyletic with respect to these distant relatives often were obtained from colonies located within or close to the range of the other species (e.g., the closely related *S. invicta* and *S. saevissima* alleles in Clade II and the closely related *S. invicta*, *S. saevissima*, and *S. altipunctata* alleles in the sister lineage of Clade IIIA; see Figure 4.4). Finally, parallel patterns of minimal regional differentiation coupled with interspecific sequence paraphyly and polyphyly have been observed for the mtDNA of *S. invicta* (Shoemaker et al. 2006, Ross et al. 2007), with the latter features almost certainly the result of introgression.

This scenario posits that *Gp-9* (and the mtDNA) flows more freely between fire ant species (and, perhaps, among local *S. invicta* populations) than the bulk of the nuclear genome, perhaps because of a lack of selection against these introgressing genetic elements (or, in the case of the mtDNA, because of selection for introgression associated with the presence of the common fire ant endosymbiont *Wolbachia*; Ahrens and Shoemaker 2005). One important implication of this scenario, if true, is that polygyny

may not have arisen in the common ancestor of the socially polymorphic fire ants, with the *b*-like lineage persisting through multiple speciation events, as implied by Krieger and Ross (2002, 2005). Rather, this allele class and the alternate form of social behavior with which it is associated conceivably arose more recently then spread among species of the socially polymorphic clade through hybridization.

This study has yielded information of use in bridging functional and population genetic approaches to understanding the genetic basis of an important social trait. It is now time to move away from purely associative studies linking variation at *Gp-9* to variation in social organization in *S. invicta* and to take the next steps toward understanding the functional consequences of *Gp-9* variation. Only through an integrative approach investigating the biochemical pathways in which the gene product functions, the phenotypic effects of molecular variation at *Gp-9* and other pathway genes, and the potential involvement of other genes in linkage disequilibrium with *Gp-9* can substantial progress toward understanding the evolution of this key social adaptation be maintained.

ACKNOWLEDGMENTS

We thank Michael Strand for comments on an earlier version of this manuscript.

REFERENCES

Ahrens M, Shoemaker DD. 2005. Evolutionary history of *Wolbachia* infections in the fire ant *Solenopsis invicta*. BMC Evol Biol 5: 35-.

Akaike H. 1974. A new look at the statistical model identification. IEEE Trans Automatic Control 19: 716-723.

Schwarz G. 1978. Estimating the dimension of a model. Annals Stat 6: 461-464.

Bourke AFG, Franks NR. 1995. *Social evolution in ants*. Princeton: Princeton University Press.

Dupanloup I, Schneider S, Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11: 2571-2581.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479-491.

Gillespie JH. 1991. *The Causes of Molecular Evolution*. Oxford: Oxford University Press.

Golding GB, Dean AM. 1998. The structural basis of molecular adaptation. Mol Biol Evol 15: 355-369.

Hallar BL, Krieger MJB, Ross KG. 2006. Potential cause of lethality of an allele implicated in social evolution in fire ants. Genetica (in press).

Hedrick PW. 2000. *Genetics of Populations*. 2nd edition. Sudbury: Jones and Bartlett Publishers.

Ho JWK, Adams CE, Lew JB, Matthews TJ, Ng CC, Shahabi-Sirjani A, Tan LH, Zhao Y, Easteal S, Wilson SR, Jermiin LS. 2006. SeqVis: Visualization of

compositional heterogeneity in large alignments of nucleotides. Bioinformatics 22: 2162 - 2163

Horst R, Damberger F, Luginbuhl P, Guntert P, Peng G, Nikonova L, Leal WS, Wuthrich K. 2001. NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. Proc Natl Acad Sci USA 98: 14374-14379.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754-755.

Huelsenbeck JP, Dyer KA. 2004. Bayesian estimation of positively selected sites. J Mol Evol 58: 661-672.

Imyanitov EN, Buslov KG, Suspitsin EN, Kuligina Esh, Belogubova EV, Grigoriev M, Togo AV, Hanson KP. 2002. Improved reliability of allele-specific PCR. BioTechniques 33: 484-490.

Jermiin LS, Simon YWH, Ababneh F, Robinson J, Larkum AWD. 2004. The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. Syst Biol 53: 638-643.

Keller L, Ross K G. 1998. Selfish genes: a green beard in the red fire ant. Nature 394: 573-575.

Keller L, Ross KG. 1999. Major gene effects on phenotype and fitness: The relative roles of *Pgm-3* and *Gp-9* in introduced populations of the fire ant *Solenopsis invicta*. J Evol Biol 12: 672-680.

Kosakovsky Pond SL, Frost SWD. 2005a. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol 22: 1208-1222.

Kosakovsky Pond SL, Frost SWD. 2005b. DATAMONKEY: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21: 2531-2533.

Krieger MJB. 2005. To *b* or not to *b*: a pheromone–binding protein regulates colony social organization in fire ants. BioEssays 27: 91-99.

Krieger MJB, Keller L. 1997. Polymorphism at dinucleotide microsatellite loci in fire ant *Solenopsis invicta* populations. Molecular Ecology 6: 997-999.

Krieger MJB, Ross KG. 2002. Identification of a major gene regulating complex social behavior. Science 295: 328-332

Krieger MJB, Ross KG. 2005. Molecular evolutionary analyses of the odorant-binding protein gene *Gp-9* in fire ants and other *Solenopsis* species. Mol Biol Evol 22: 2090-2103

Kumar S, Tamura K, Nei M. 2004. MEGA 3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. Briefings Bioinformatics 5: 150-163.

Larget B. 2005. Introduction to Markov Chain Monte Carlo methods in molecular evolution. In Nielsen R. (ed.) *Statistical Methods in Molecular Evolution*. Springer, New York. Pp: 45-62.

Lartigue A, Gruez A, Spinelli S, Revière S, Brossut R**,** Tegoni M, Cambillau C. 2003. The crystal structure of a cockroach pheromone-binding protein suggests a new ligand binding and release mechanism. J Biol Chem 278: 30213-30218.

LeeD, Damberger FF, Peng GH, Horst G, Guntert P. 2002. NMR structure of the unliganded *Bombyx mori* pheromone-binding protein at physiological pH. Febs Letters 531: 314-318.

Lockhart PJ, Steel MA, Hendy MD, Penny D. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol Biol Evol 11:605-612.

Mescher MC, Ross KG, Shoemaker DD, Keller L, Krieger MJB. 2003. Distribution of the two social forms of the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae) in the native South American range. Annals of the Entomological Society of America 96: 810–817.

Mescher MC, Ross KG, Shoemaker DDS, Keller L, Krieger MJB. 2003. Distribution of the two social forms of the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae) in the native South American range. Ann Entomol Soc Am 96: 810-817.

Messier W, Stewart C-B. 1997. Episodic adaptive evolution of primate lysozymes. Nature 385: 151-154.

Milne I, Wright F, Rowe G, Marshal DF, Husmeier D, McGuire G. 2004. TOPALi: software for automatic identification of recombinant sequences within DNA multiple alignments. Bioinformatics 20: 1806-1807.

Nachman MW. 2005. The genetic basis of adaptation: lessons from concealing coloration in pocket mice. Genetica 123: 125-136.

Nagnan-Le Meillour P, Jacquin-Joly E. 2003. Biochemistry and diversity of insect odorant-binding proteins. In Blomquist GJ, Vogt RG. (eds.) Insect pheromone biochemistry and molecular biology: the biosynthesis and detection of pheromones and plant volatiles. London: Elsevier Academic. Pp: 509-537.

Nei M, Li HW. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76: 5269-5273.

Nielsen R, Yang Z. 1998. Likelihood models for detecting postively selected amino acid sites and applications to the HIV-1 envelope genes. Genetics 148: 929-936.

Nixon KC. 1999. The Parsimony Ratchet, a new method for rapid parsimony analysis. Cladistics 15: 407-414.

Phillips P C. 2005. Testing hypotheses regarding the genetics of adaptation. Genetica 123: 15-24.

Pitts JP, McHugh JV, Ross KG. 2005. Cladistic analysis of the fire ants of the *Solenopsis saevissima* species-group (Hymenoptera: Formicidae). Zool Scr 34: 493-505.

Pitts JP, McHugh JV, Ross KG. 2007. Revision of the fire ants of the *Solenopsis saevissima* species-group (Hymenoptera: Formicidae). Zootaxa (in press).

Posada D. 2002. Evaluation of methods for detecting recombination from DNA sequences: empirical data. Mol Biol Evol 19: 708-717.

Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817-818.

Posada D, Crandall KA. 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc Natl Acad Sci USA 98: 13757-13762.

Posada D, Buckley TR. 2004. Model selection and model averaging in phylogenetics: advantages of Aikake information criterion and Bayesian approaches over likelihood ratio tests. Syst Biol 53: 793-808.

Raymond M, Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Heredity 86: 248-249.

Ronquist F, Huelsenbeck JP, van der Mark P. 2005. MrBayes 3.1 Manual.

Rooney AP, Zhang J. 1999. Rapid evolution of a primate sperm proteinL relaxation of functional constraint or positive Darwinian selection? Mol Biol Evol 16: 706-710.

Rosenberg NA. 2003. The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. Evolution 57: 1465–1477.

Ross KG. 1997. Multilocus evolution in fire ants: effects of selection, gene flow, and recombination. Genetics 145: 961-974.

Ross KG, Keller L. 1998. Genetic control of social organization in an ant. Proc Natl Acad Sci USA 95: 14232-14237.

Ross KG, Vargo EL, Keller L, Trager JC. 1993. Effect of a founder event on variation in the genetic sex-determining system of the fire ant *Solenopsis invicta*. Genetics 135: 843-854.

Ross KG, Krieger MJB, Shoemaker DD. 2003. Alternative genetic foundations for a key social polymorphism in fire ants. Genetics 165: 1853-1867.

Ross KG, Shoemaker DD. 2005. Species delimination in native South American fire ants. Mol Ecol 14: 3419-3438.

Ross KG, Krieger MJB, KellerL , Shoemaker DD. 2007. Genetic variation and structure in native populations of the fire ant *Solenopsis invicta*: evolutionary and demographic implications. Biol J Linn Soc (*in press*).

Ross KG, Krieger MJB, Shoemaker DD, Vargo EL, Keller L. 1997. Hierarchical analysis of genetic structure in native fire ant populations: results from three classes of molecular markers. Genetics 147: 643-655.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics 145: 1219-1228.

Sandler BH, Nikinova L, Leal WS, Clardy J. 2000. Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. Chem Biol 7: 143-151.

Scheffler K, Seoighe C. 2005. A Bayesian model comparison approach to inferring positive selection. Mol Biol Evol 22: 2531-2540.

Schneider S, Roessli D, Excoffier L. 2000. *Arlequin: A software for population genetics data analysis*, Version 2.000. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva.

Shimodaira H, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol 16: 1114-1116.

Shoemaker DD, Ross KG. 1996. Effects of social organization on gene flow in the fire ant *Solenopsis invicta*. Nature 383: 613-616.

Shoemaker DD, Ahrens ME, Ross KG. 2006. Molecular phylogeny of fire ants of the *Solenopsis saevissima* species-group based on mtDNA sequences. Mol Phyl Evol 38: 200-215.

Shoemaker DD, DeHeer CJ, Krieger MJB, Ross KG. 2006. Population genetics of the invasive fire ant *Solenopsis invicta* in the United States. Ann Entomol Soc Am 99: 1213-1233.

Sikes DS, Lewis PO. 2001. PAUPRat v1: PAUP* implementation of the parsimony ratchet. Distributed by the authors Department Ecology and Evolutionary Biology, University of Connecticut, Storrs, USA.

Slatkin M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47: 264-279.

Tarrio R, Rodriguez-Trelles F, Ayala FJ. 2001. Shared nucleotide composition biases among species and their impact on phylogenetic reconstructions of the Drosophilidae. Mol Biol Evol 18: 1464-1473

Vasemägi A, Primmer CR. 2005. Challenges for identifying functionally important genetic variation:the promise of combining complementary research strategies. Mol Ecol 14: 3623-3642.

Vogt R G. 2005. Molecular basis of pheromone detection in insects. In Gilbert LI, Iatro K, Gill S (eds.) *Endocrinology vol. 3. Comprehensive Insect Physiology, Biochemistry, Pharmacology and Molecular Biology*. London: Elsevier. Pp: 753-804

Wiuf C, Christensen T, Hein J. 2001. A simulation study of the reliability of recombination detection methods. Mol Biol Evol 18: 1929-1939.

Xia X, Xie Z. 2001. DAMBE: Software package for data analysis in molecular biology and evolution. Journal of Heredity 92: 371-373.

Xu P, Zwiebel L, Smith D. 2003. Identification of a distinct family of genes encoding atypical odorant-binding proteins in the malaria vector mosquito *Anopheles gambiae*. Insect Mol Biol 12: 549-560.

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Appl BioSci 13: 555-556.

Yang, Z, Nielsen R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specified lineages. Mol Biol Evol 19: 908-917.

Yang Z, Nielsen R, Goldman N, Pedersen A-MK. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155: 431-449.

Yang Z, Wong WSW, Nielsen R. 2005. Bayes empirical bayes inference of amino acid sites under positive selection. Mol Biol Evol 22: 1107-1118.

Zhang J, Kumar S, Nei M. 1997. Small-sample tests of episodic adaptive evolution: a case study of primate lysozymes. Mol Biol Evol 14: 1335-1338.

Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol 22: 2472-2479.

Zhou J, Huang W, Zhang G, Pickett J, Field L,. 2004. “Plus-C” odorant-binding protein genes in two *Drosophila* species and the malaria mosquito *Anopheles gambiae*. Gene 327: 117-129.

TABLE LEGENDS

Table 4.1.—DIVERSITY STATISTICS FOR *GP-9* IN *S. INVICTA*.

The uncorrected mean number of nucleotide substitutions between unique haplotypes (*d*) and uncorrected mean proportion of nucleotide substitutions between unique haplotypes (*p*) are presented separately for the coding and non-coding regions of the gene.

TABLE 4.2.—AMINO ACID VARIATION AT THREE GP-9 CODON POSITIONS.

These codon positions are previously considered to contain residues diagnostic of the polygyny-associated *b*-like alleles (these *b*-like residues are underlined) in four specimens of *S. invicta*. Designation of the social organization of the colony of origin of each specimen is based on micosatellite analyses

Table 4.3.—IDENTIFICATION OF SITE-SPECIFIC POSITIVE SELECTION ON *GP-9* BY DIFFERENT METHODS.

Dashes indicate an absence for evidence of positive selection using a particular method.

Table 4.1.

	d		*p*					
	coding regions	non-coding regions	coding regions	non-coding regions	prop. variable codons	prop. variable amino acids	codon diversity	amino acid diversity
B-like alleles	3.05	6.20	0.0067	0.0036	0.3766	0.3052	0.0271	0.0211
b-like alleles	3.32	3.62	0.0072	0.0021	0.1242	0.0980	0.0253	0.0221
all *Gp-9* alleles	5.41	6.66	0.0118	0.0039	0.4351	0.3571	0.0362	0.0308

TABLE 4.2.

Specimen	Colony social organization	Codon position		
		42	95	139
Pi21	monogyne	Gly	Met	Val
O40	monogyne	Gly	Met	Val
LP719	monogyne	Ser	Met	Ile
SC665	polygyne	Gly	Ile	Val

Table 4.3.

Codon position	Method: Site-specific — SLAC counting[a] *(dN-dS, P)*	Site-specific — Suzuki-Gojobori counting *(dN-dS, P)*	Site-specific — Bayesian random-effects *(dN/dS*, post. prob.*)*	Branch-site-specific — Bayes empirical Bayes random-effects[b] *(dN/dS*, post. prob.*)*
3	—	19.00, 0.001	—	—
6	1.75, 0.22	—	—	—
39	2.49, 0.13	—	79, 1.00	17.5, 0.957
42	—	—	79, 0.986	—
45	—	—	79, 0.953	—
48	—	—	79, 0.973	—
61	—	—	79, 0.935	—
75	—	—	79, 0.995	—
78	—	—	79, 0.998	—
95	—	—	79, 0.993	—
117	2.50, 0.13	—	79, 1.00	17.5, 0.995
119	—	—	79, 0.900	—
120	—	—	79, 0.973	—
134	—	—	79, 0.998	—
145	—	—	79, 0.907	—
152	—	—	79, 0.908	—

[a] A nominal α-level of 0.25 is used to indicate statistical significance of selection in the SLAC analysis (Kosakovsky Pond and Frost 2005a).

[b] Only branches of the *b*-like clade were examined for positive selection using this method.

FIGURE LEGENDS

Figure 4.1.—LOCATIONS OF SAMPLING SITES FOR *S. INVICTA* IN ITS NATIVE RANGE.

The number of colonies sampled at a site (=number of sequences obtained) is indicated in parentheses in the key to the localities if this number is greater than one. Sites at which *b*-like alleles were found are highlighted with black squares; those at which *b* alleles were found are underlined as well. The native range is shown in gray shading.

Figure 4.2.—DISTRIBUTIONS OF NUMBERS OF NUCLEOTIDE SUBSTITUTIONS BETWEEN PAIRS OF UNIQUE *GP-9* HAPLOTYPES FROM *S. INVICTA*.

Figure 4.3.—NUMBERS OF DIFFERENT AMINO ACIDS OCCURRING AT EACH CODON POSITION OF *GP-9* IN THE *B*-LIKE AND *B*-LIKE ALLELE CLASSES IN *S. INVICTA*.

Exon boundaries are indicated by dotted lines, and signal peptide codons are highlighted with gray shading. Positions of the three amino acid residues previously considered to be diagnostic for *B*-like and *b*-like alleles are indicated by arrows.

Figure 4.4.—SUMMARY PHYLOGENETIC HYPOTHESIS OF *GP-9* HAPLOTYPE RELATIONSHIPS BASED ON FOUR INDEPENDENT MCMC RUNS USING BAYESIAN INFERENCE (BI).

Haplotypes of species other than *S. invicta* are indicated by specific abbreviations; *S. invicta* haplotypes are indicated by circles. Posterior clade probabilities >50% are indicated as percentages below the branches.

Figure 4.1

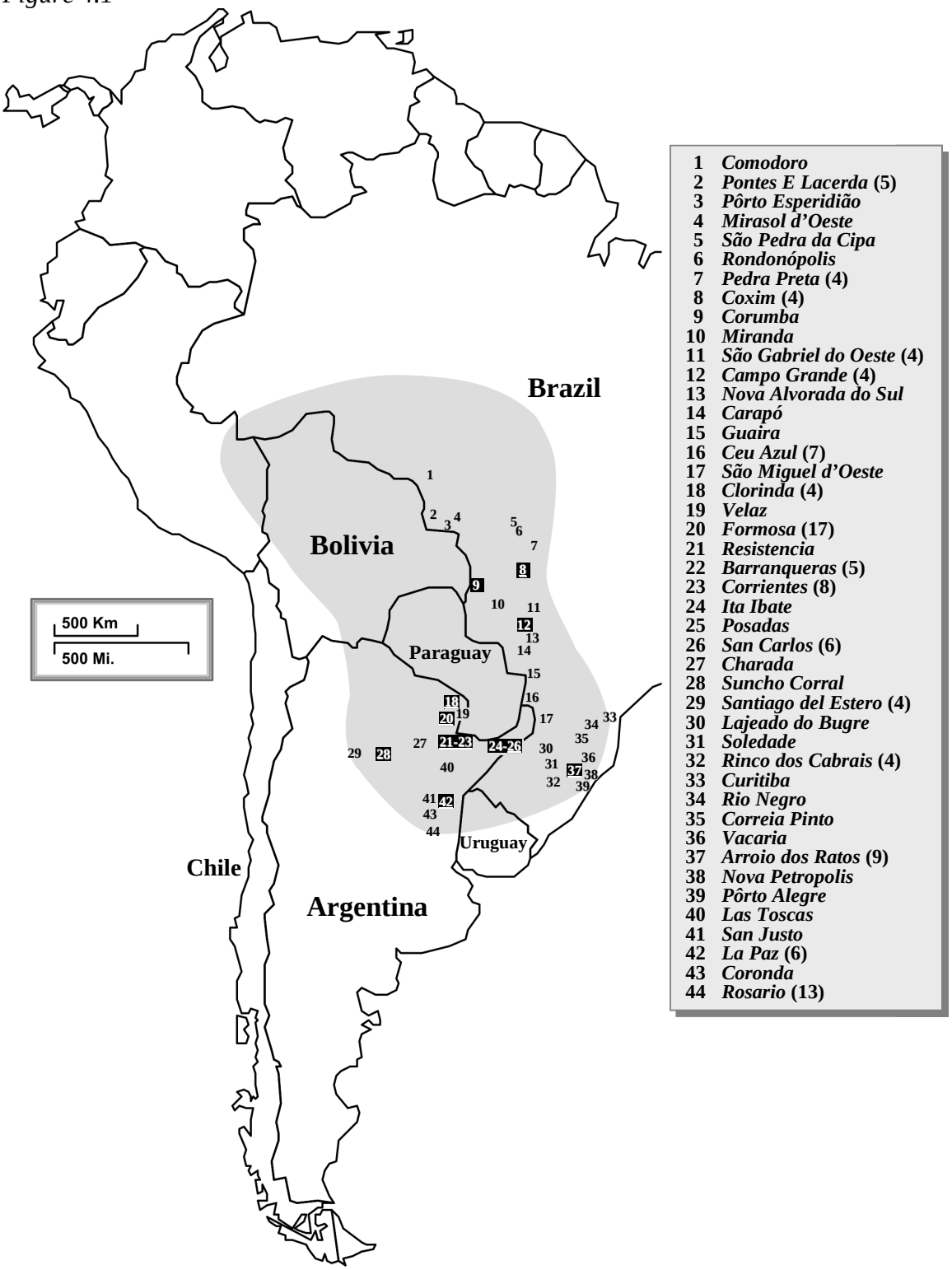

Brazil
Bolivia
Paraguay
Uruguay
Chile
Argentina
500 Km
500 Mi.
1 Comodoro
2 Pontes E Lacerda (5)
3 Pôrto Esperidião
4 Mirasol d'Oeste
5 São Pedra da Cipa
6 Rondonópolis
7 Pedra Preta (4)
8 Coxim (4)
9 Corumba
10 Miranda
11 São Gabriel do Oeste (4)
12 Campo Grande (4)
13 Nova Alvorada do Sul
14 Carapó
15 Guaira
16 Ceu Azul (7)
17 São Miguel d'Oeste
18 Clorinda (4)
19 Velaz
20 Formosa (17)
21 Resistencia
22 Barranqueras (5)
23 Corrientes (8)
24 Ita Ibate
25 Posadas
26 San Carlos (6)
27 Charada
28 Suncho Corral
29 Santiago del Estero (4)
30 Lajeado do Bugre
31 Soledade
32 Rinco dos Cabrais (4)
33 Curitiba
34 Rio Negro
35 Correia Pinto
36 Vacaria
37 Arroio dos Ratos (9)
38 Nova Petropolis
39 Pôrto Alegre
40 Las Toscas
41 San Justo
42 La Paz (6)
43 Coronda
44 Rosario (13)

Figure 4.2

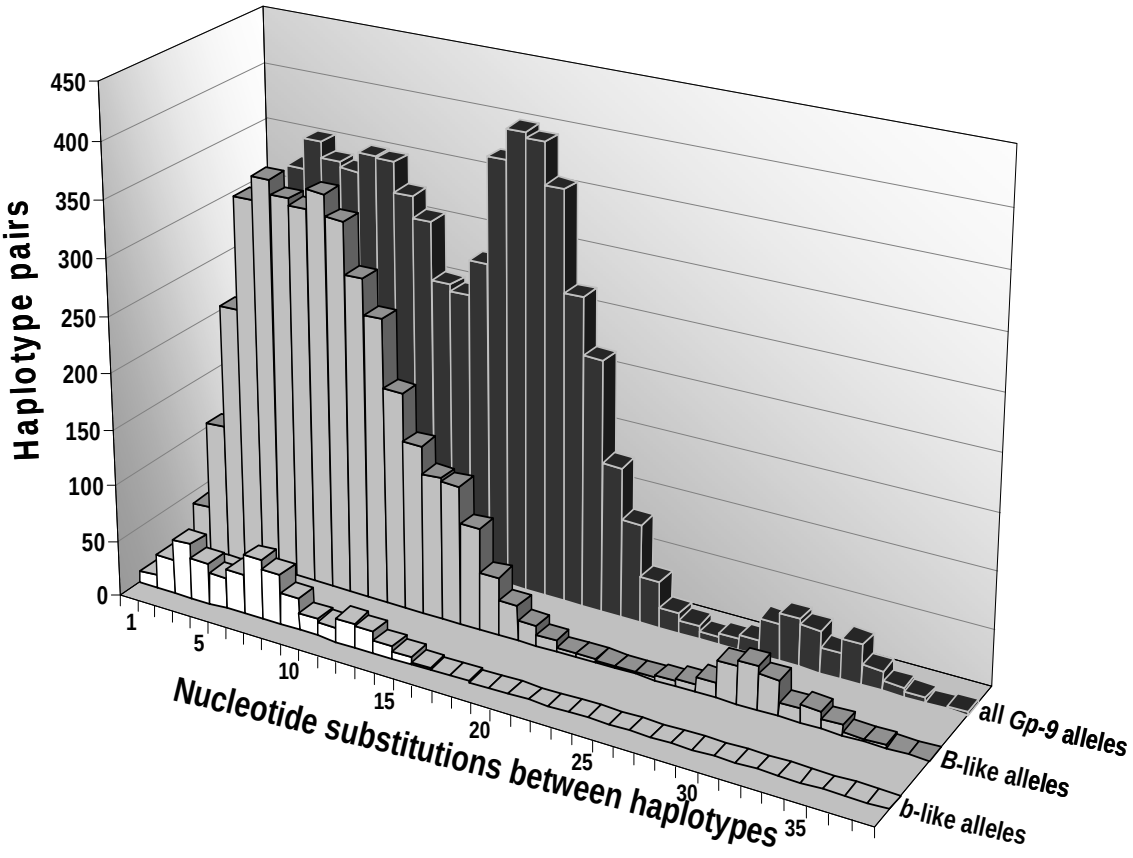

Haplotype pairs
450
400
350
300
250
200
150
100
50
0
1
5
10
15
20
25
30
35
Nucleotide substitutions between haplotypes
all Gp-9 alleles
B-like alleles
b-like alleles

Figure 4.3

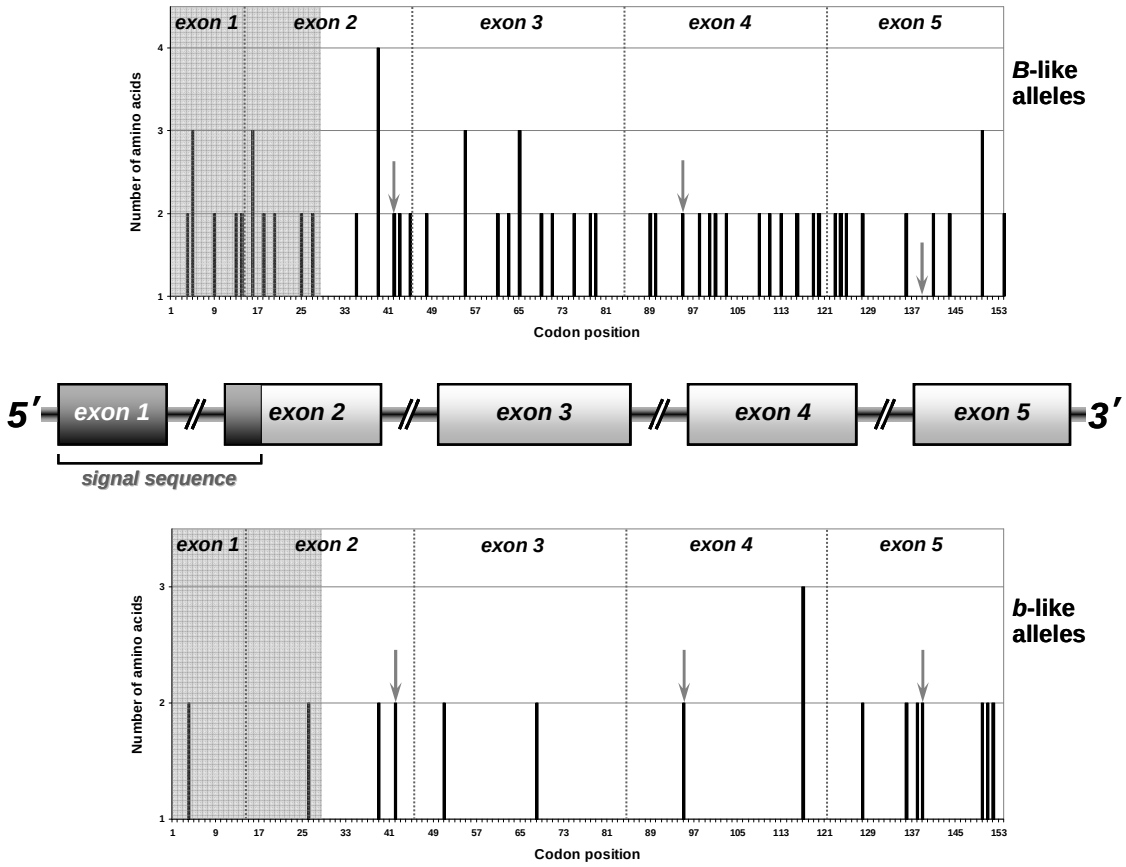
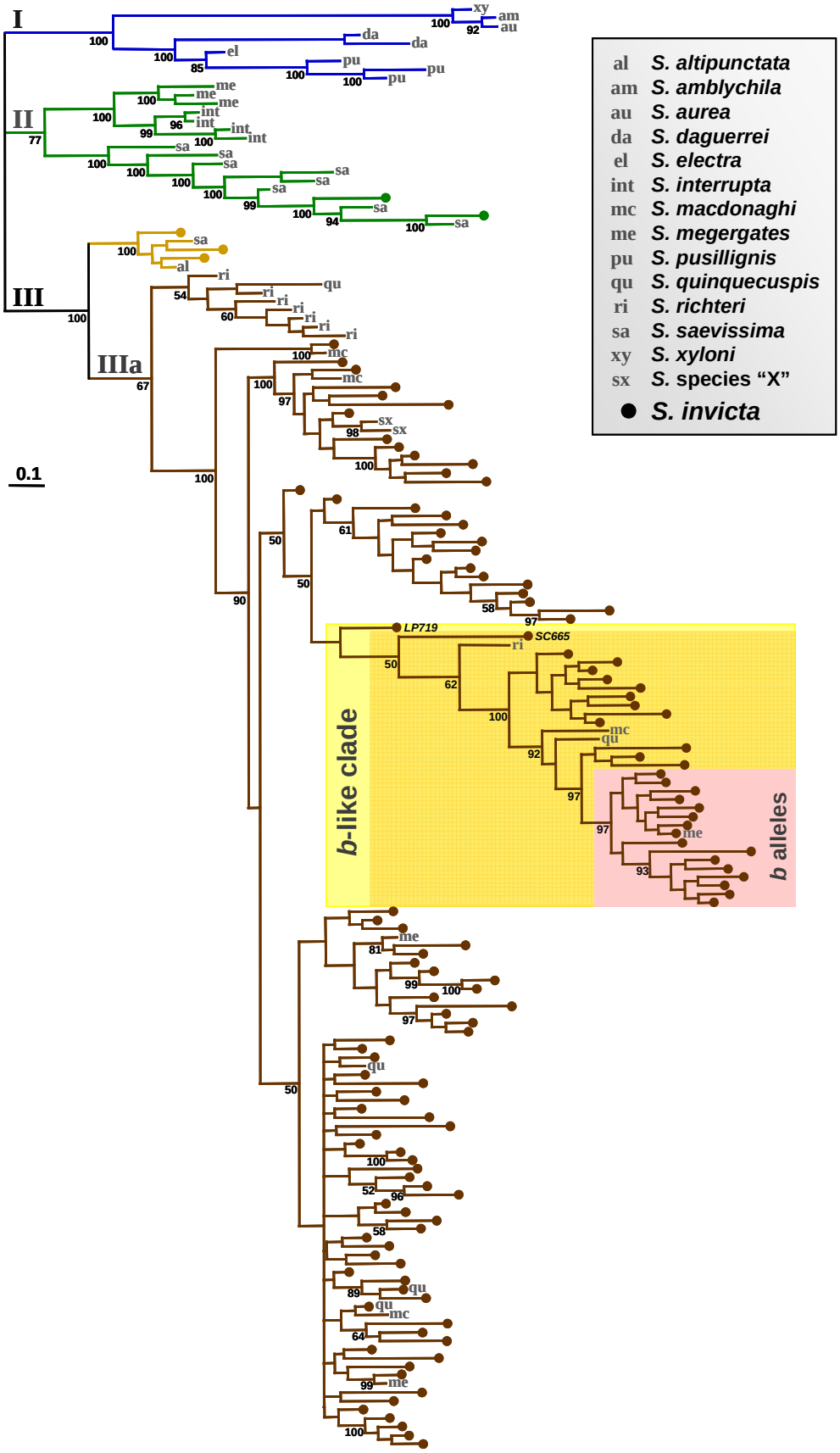

exon 1
exon 2
exon 3
exon 4
exon 5
B-like alleles
Number of amino acids
Codon position
5′
3′
signal sequence
b-like alleles

Figure 4,4


I
II
III
IIIa
0.1
al S. altipunctata
am S. amblychila
au S. aurea
da S. daguerrei
el S. electra
int S. interrupta
mc S. macdonaghi
me S. megergates
pu S. pusillignis
qu S. quinquecuspis
ri S. richteri
sa S. saevissima
xy S. xyloni
sx S. species "X"
S. invicta
LP719
SC665
b-like clade
b alleles

CHAPTER 5

CONCLUSIONS

The research presented here seeks to begin to bridge the gap between molecular pattern and process in our understanding of the association between *Gp-9* and polygyny in the socially polymorphic South American fire ants. I have sought to show how a relatively small number of the *b* allele carrying workers are sufficient to elicit polygyne behavior and that queen effects do not influence the acceptance of supernumerary queens. This complements and supports previous research and taken together these findings represent an important first step into understanding the ties between individual genotypes and colony level expression of social organization.

I have also attempted to show that single nucleotide substitutions at the *Gp-9* locus are not sufficiently associated with polygyne behavior, but that several such changes must occur for the expression of the polygyne phenotype. This raises interesting questions about the validity of the presumed causal role of the *b*-like alleles at *Gp-9* in inducing polygyne behavior in South American fire ants.

Finally, I discuss the state of our current understanding of *Gp-9* and its role in regulating social behavior, resulting in a simple, testable model, and suggestions for future avenues of research. Whether *Gp-9* remains the prime candidate gene for controlling social organization in fire ants remains to be seen. But even if it does not, it will deserve continued attention by students of social behavior.