THE ROLES OF GENETIC DIVERSITY AND SINV-2 VIRAL INFECTION IN FITNESS OF THE INVASIVE FIRE ANT *SOLENOPSIS INVICTA*

by

Mark A. Fisher

(Under the Direction of Kenneth G. Ross)

ABSTRACT

Variation in fitness is an underlying tenet of adaptive evolution, and as such it is an essential goal of evolutionary biology to understand the mechanisms and forces that generate it. We explored two mechanisms that could potentially generate variation in fitness in the invasive fire ant, *Solenopsis invicta*: within-individual genetic diversity and viral infection. In order to explore the former, we conducted a heterozygosity-fitness association (HFA) study across twelve fitness-related traits in *S. invicta* queens and incipient colonies. Analyses revealed that HFAs are uncommon in our study population, with queen multilocus heterozygosity a significant predictor of just a small subset of the fitness traits that we measured. The same analyses also highlighted the influence exerted by stress, life-history tradeoffs, and variation in life-cycle stages assayed on HFAs in our system. An analytical pipeline, HeFPipe, was constructed to facilitate and streamline future HFA studies.

In order to explore the effects of viral infection on fitness of *S. invicta* queens and incipient colonies, we looked for differences across the same fitness-related traits leveraged in the HFA study between queens infected by various permutations of three

positive, single-stranded RNA viruses, SINV-1, -2, and -3, and uninfected queens. We demonstrated that at least one virus infecting fire ants, SINV-2, has significant, negative effects on several traits associated with early queen and colony growth and survival. Our evidence for clear fitness costs of SINV-2 on *S. invicta* queens and colonies runs contrary to previous studies suggesting this virus is largely asymptomatic. Our study highlights how rigorous, quantitative measurements of fitness-related traits may reveal significant host effects of viruses that otherwise would go undetected. Such quantitative studies will become important from medical, agricultural, economic, and conservation perspectives as globalization continues to bring viruses into contact with an increasing diversity of hosts.

INDEX WORDS: fire ants, heterozygosity, fitness, HFA, HFC, inbreeding, SINV-2, *Solenopsis invicta*, virus, HeFPipe, population genetics software

THE ROLES OF GENETIC DIVERSITY AND SINV-2 VIRAL INFECTION IN FITNESS OF THE INVASIVE FIRE ANT *SOLENOPSIS INVICTA*

by

Mark A. Fisher

B.S., Yale University, 2007

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

2013

THE ROLES OF GENETIC DIVERSITY AND SINV-2 VIRAL INFECTION IN FITNESS OF THE INVASIVE FIRE ANT *SOLENOPSIS INVICTA*

by

MARK AARON FISHER

Major Professor: Kenneth Ross

Committee: Daniel Promislow
DeWayne Shoemaker
David Hall
Kelly Dyer

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2013

DEDICATION

To my nestmates in Genetics. Without you, I wouldn't have endured this fitness crucible. May we continue cooperating and growing together in the future.

To all of my friends at Athens Fitness and MMA and Global MMA. Growing beside you as a martial artist in many ways reflected growing as a scientist. It was a true honor. And it certainly helped to have a hobby where violence sometimes was the answer.

To Cass for many years of happy pleometrosis.

To Philip, Mark, Melissa, and Jenna. You made Athens for me.

To the runners. You made Gainesville.

To my family. It was you who planted the seeds of science in my mind. How could someone be so lucky?

And to Atticus.

ACKNOWLEDGEMENTS

I thank my adviser, Dr. Kenneth G. Ross, for making me feel like a collaborator from day one; it's not every day that one finds an adviser who does field work and bench work side-by-side with his students. I'm happy to call him a friend at the end of all of this. There are so many things that I will take with me from his mentoring style. I am grateful for everything he's given.

I thank Dr. DeWayne Shoemaker for the sense of community he brings to his laboratory, for teaching me many new laboratory techniques, and for his impeccable taste in food. The Gainesville lab became second home, and he went a long way in making me feel like it. I thank Dr. Daniel Promislow for taking me on as a teaching mentee, and for his commitment to and passion for science. I thank Dr. David Hall for bearing the brunt of my most major molecular genetics growing pains. I'm still sorry for breaking John Wares's cold centrifuge while rotating in his lab and for attempting to cast a gel without using the rig. The rigor that he brings to everything he does is something to aspire to. I thank Dr. Kelly Dyer for bringing a sense of balance to committee meetings. Not once was her feedback anything but eminently useful, constructive, and relevant.

I thank Dr. Jessica Kissinger and Dr. John Wares for opening their laboratories to my nomadic work style. I thank Dr. Chin-Cheng (Scotty) Yang, Dr. Fabio Manfredini, and Dr. Lucinda Lawson for all of the extremely useful technical and intellectual insight; you became like siblings to me. I thank Eileen Carroll for her technical expertise; I can't even imagine what genotyping would have been like without her help. I thank Ariel Bromlow, Andrew Shahan, Haley Martinez, Brandi Jo Petronio, Stephanie Tiv, and Tori Staples for their help in the laboratories. I thank Patrice David, Juanita Olano-Marin, Bill

Amos, Jon Slate, and especially Jo Chapman for the timely and insightful responses to all of my HFC-related questions.

I thank the staff in the genetics department, especially Janice, Erika, and Susan. You got to know my face quite well over the years, huh? Thank you for saving me from disaster time and time again!

I thank my friends in the Genetics department, especially Emily, Jen, Sarah, Marly, Cass, Louisa, Jodie, Eileen, Scott, John, Tina, Monica, Kerin, Katie, Christine, Liza, Paul, Christina, and Jenna. I will never forget your wise counsel and comfort.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION

Variation in fitness is an underlying tenet of adaptive evolution, and as such it is an essential goal of evolutionary biology to understand the mechanisms and forces that generate it. Of particular interest is the role of genetic diversity in the survival and reproductive success of individuals. In order to characterize this role, it is important to measure the genetic variability of individuals, usually with some estimator of multilocus heterozygosity, and to determine the strength of an association between this measure and relative fitness. Studies that investigate such relationships are called heterozygosity-fitness correlation (HFC) studies or, perhaps more appropriately, heterozygosity-fitness association (HFA) studies (e.g.,1–7).

Another mechanism that generates variation in fitness is viral infection. Viral infection is ubiquitous, and viruses or some type of selfish genetic element infect every cellular organism (8, 9). However, the effects—especially the subtle and cryptic ones—of viruses on host ecology, life history, and fitness are poorly understood for all but the best-characterized viruses (10).

This review provides context for both historical and contemporary empirical HFA studies, establishes the current theoretical framework of HFAs, explores the suitability of the invasive fire ant *Solenopsis invicta* as a system in which to study HFAs, highlights the diversity of statistical and analytical tools that are currently available to distinguish

among various theories, thereby justifying the need for an analytical pipeline to provide ecumenicism across these tools, provides context for both historical and contemporary empirical insect virus research, discusses the status of SINV-1,-2, and -3 as candidates for biological control in *S. invicta*, and finally emphasizes the need for careful, quantitative analyses of these viral infections on queen and colony phenotypes.

Heterozygosity-Fitness Associations

HFA theory and practice

Although most heterozygosity-fitness association studies have reported a weak, positive effect, effect sizes to date generally are relatively small (11) and there is substantial variation in results among studies. Furthermore, measures of multilocus heterozygosity determined by the small panels of marker loci typically employed (11) are notoriously poor predictors of genome-wide heterozygosity (12, 13), raising doubts as to the meaning of the variable results obtained thus far.

Several potential explanations have been advanced for HFAs: i) the surveyed marker loci directly affect the fitness trait(s) of interest, a possibility only if these loci are non-neutral, ii) the marker loci are in linkage disequilibrium (LD) with loci directly influencing the fitness trait(s), and iii) individuals with lower heterozygosity suffer costs of inbreeding, or individuals with higher heterozygosity suffer costs of outbreeding, relative to the remainder of the population. The two latter explanations both rely on associations between coding loci and marker loci to explain HFAs, but the third ("general effect" hypothesis) differs from the second ("local effect" hypothesis) in that it typically invokes some form of inbreeding *sensu lato* (nonrandom mating, small population size,

bottlenecks, or admixture, which can render non-admixed individuals inbred relative to admixed ones) to generate extensive intra-genomic associations (14).

HFAs attributable to the general effects of inbreeding depression make intuitive sense in terms of the genetic architecture of fitness traits, which are believed largely to be the products of many genes of small effect distributed throughout the genome (15, 16). By causing an association between homozygous genotypes at different loci — thereby creating identity disequilibrium (ID) — inbreeding can drastically affect the phenotype of polygenic fitness-associated traits by altering the genotypes of several loci at once; such traits are otherwise considered to be resistant to rapid change because of the large and diffuse mutational targets they represent. Thus, it is this third, general effect explanation for HFAs that currently is favored by many theoreticians and empiricists (14, 17, 18). Nonetheless, an association between population demographic history and effect size of observed HFAs, which theory suggests should exist if inbreeding is the predominant cause of positive HFAs (14), generally is absent (11), lending some support to the two alternate hypotheses not invoking general effects. This theoretical framework applies as well to negative HFAs, in which heterozygosity is negatively associated with fitness-trait values, although the distinction between general effects versus direct or local effects usually is couched instead in terms of outbreeding depression versus single-locus genotype-trait associations (SLAs), respectively (e.g.,5, 19–22).

Incomplete knowledge of the extent of departures from panmixia in many study systems, coupled with uncertainty about the functional roles and selective consequences of inbreeding and outbreeding depression in generating HFAs, likely contribute along with small marker panel size to the variability in results and consequent lack of consensus

regarding the importance of direct, local, or general effects (11). No doubt also contributing are substantial variation in the types and quality of fitness-associated traits measured, the sample sizes employed, and the potential for alleles under directional selection to generate transient HFAs (23, 24).

Solenopsis invicta as a candidate for HFA

Solenopsis invicta is an invasive fire ant introduced to the U.S. from Argentina in the early 1930s (25). It has since spread rapidly throughout the southeastern and parts of the southwestern U.S., where it has emerged as a major agricultural and urban pest (26). Several considerations make *S. invicta* a promising subject for an HFA study, including potential increases in additive genetic variance in fitness and in LD between marker and fitness loci, both induced by the founder event, availability of a large complement of genetic markers with which to measure heterozygosity, and a firm understanding of traits that are intimately associated with fitness, especially in the early life of a colony.

The effective founding population of North American *S. invicta* is estimated to have comprised 9 to 20 mated queens (27), and only 10 to 20 generations have passed since the founding event (27, 28). Empirical observations (29) and theoretical work (30) suggest that both additive genetic variance and fitness variance can increase after such a genetic bottleneck, and subsequent spread of rare or novel genotypes and the fitness phenotypes they influence in the expanding population may enhance HFAs. A bottleneck also increases ID and LD among loci (14), a fact that can be leveraged to detect the effects of fitness loci using associated marker loci when sufficiently few generations have passed for such disequilibrium to decay. The potential for detecting HFAs in our study system also is enhanced by the availability of a large number (>100) of polymorphic

markers. This large panel size enables us to accurately assess genome-wide heterozygosity, a necessity made salient by two recent publications that used large panels (~70 loci) to recant evidence for a general effect detected by smaller marker panels in earlier studies of the same systems (31, 32). Finally, knowledge of *S. invicta* natural history implicates several measurable morphological, physiological, and life history traits intimately tied to fitness during the crucial colony-founding phase of the life cycle, including queen size, queen weight, queen resistance to viral infection, and colony developmental rate.

On the other hand, some factors may disfavor detection of HFA in invasive *S. invicta*. These ants engage in massive aerial mating/dispersal swarms that often are synchronized over vast areas (28, 33, 34), so there is little reason to expect the existence of any consanguineous matings that result in ID and, consequently, HFAs. Indeed, earlier empirical studies have failed to detect evidence for significant levels of inbreeding in invasive U.S. populations (35–37), the situation expected for species such as *S. invicta* with single-locus complementary sex-determination systems that impose severe fitness penalties for inbreeding (38). HFA signals may also be weak in *S. invicta* because males are haploid (typically) and females are diploid. Recessive deleterious alleles — presumably responsible for positive HFAs due to inbreeding depression in a portion of the population — can be relatively readily removed via selection on hemizygous males (38). But this purging scenario assumes that reproductives of the two sexes generally express genes at similar levels, a scenario not well supported for *S. invicta* males and queens (39).

The types of relationships expected between individual heterozygosity and fitness in the invasive U.S. population of *S. invicta* thus are unclear. More broadly, little empirical work intended to clarify such relationships has been conducted either in social Hymenoptera (but see Vitikainen et al. 2011; Kureck et al. 2012) or in invasive populations (but see Wetzel et al. 2012).

Tools available for conducting HFA studies

Modern HFC studies almost exclusively employ microsatellite markers rather than allozymes, and the shift to larger marker panels containing potentially neutral loci has brought with it a wider availability of statistical tests that help researchers explore the nature of heterozygosity-fitness correlations in their study systems. These tests make it possible to determine i) whether the MLH of the marker panel is reflective of genome-wide MLH, for instance with the heterozygosity-heterozygosity correlation (HHC) test (12, 44), ii) whether there is identity disequilibrium (ID) among the markers and consequently inbreeding *sensu lato* in the study system, for instance by testing for second-order heterozygosity disequilibrium, or g2 (14, 45), and iii) whether there is evidence for single-marker effects on the trait(s) of interest, for instance by using *GEPHAST* (46) or the F-ratio test described by David (14, 17). There is furthermore the question of how best to weigh rare alleles or influential loci when calculating multilocus heterozygosity estimators (11, 44, 47, 48). The software now available to conduct these tests and calculate these estimators as well as run the regressions and correlations that are the core of HFC analyses require input files of different formats, and there is currently no software that provides a cohesive framework across them or ecumenicism among them.

Fitness effects of insect viruses

Viruses are the most abundant biological entities on the planet (49), and they infect every type of cellular organism (8, 9). Numerous viruses have been discovered and well-characterized, especially highly virulent or disease-causing viruses infecting plants and animals (e.g.,10, 50–55). Despite this knowledge for many viruses of medical, agricultural, or economical importance—often focused on molecular and biochemical mechanisms of infection and replication (10, 50, 52, 54, 55)—very little is known for the vast majority of viruses. The case is no different for insect viruses, which represent a large portion of virus diversity and include both DNA and RNA viruses (10). Many examples exist in the insect virus literature that highlight the imbalance between mechanistic and ecological knowledge of insect viruses (10, 56).

In contrast to this backdrop of early virus research, recent studies have revealed a diversity of other insect viruses with less devastating effects on their hosts. Indeed, viruses have been implicated in phenotypic effects as various as alteration of morphology (57), behavioral modification (58), and manipulation of sex ratio (59), as well as persisting with no overt signs of disease but resulting in mortality under certain conditions (55). In light of these findings, it seems likely that a number of viruses that have been described as asymptomatic because of the lack of any overt or gross phenotypic effects may in fact exert significant but cryptic effects on their hosts, although the possibility remains that some viruses may persist with no apparent costs to fitness whatsoever (60, 61).

Viruses infecting *Solenopsis invicta*

Such cryptic phenotypic effects of infection may be true for one or more viruses infecting the fire ant *Solenopsis invicta*. Recent studies have identified three positive, single-stranded RNA viruses, SINV-1, SINV-2, and SINV-3 (62–64). SINV-3 infections appear to be systemic, infecting all tissues (63), and infected colonies exhibit large midden piles of dead ants, brood mortality, loss of queen physogastry (distension of the abdominal intersegmental membranes), and colony collapse (63), with occasional rebound to normal brood production (63). However, no observable symptoms were detected in the field in colonies infected by either SINV-1 or SINV-2 (62, 65), which also infect all castes and developmental stages in *S. invicta* (62, 66) but bear more localized infections, which occur predominantly in the alimentary canal (specifically the midgut) (66, 67). Furthermore, only minor brood die-off was detected in small sample sizes of SINV-1- or SINV-2-infected colonies reared in laboratory conditions (62, 67).

SINV-1 and SINV-2 infections thus have been described as largely asymptomatic. However, given the clear precedent in the virus literature for cryptic or late-onset fitness costs to persistent viral infections (10, 53, 55, 68, 69), a more rigorous study measuring the effects of infection of these three viruses on a variety of queen and colony phenotypes is much needed. Rigorous characterization of the phenotypic effects of infection by these three viruses is of course of additional interest in *S. invicta* because of its status as an invasive pest.

This review establishes the need and validity of an HFA study of *Solenopsis invicta*, justifies the usefulness of an analytical pipeline to provide ecumenicism across the tools available to conduct HFAs, and emphasizes the need for careful, quantitative analyses of

viral infections on queen and colony phenotypes of invasive *S. invicta*. Accordingly, Chapter 2 describes a heterozygosity-fitness correlation conducted in a population of monogyne *S. invicta* queens from northeast Georgia, Chapter 3 describes the analytical pipeline, *HeFPipe*, which provides a framework for conducting microsatellite-based HFAs in a single sitting, as well as ecumenicism across several distinct file formats, and Chapter 4 describes a study that attempts to identify associations between viral infection by SINV-1,-2-, and -3 and several queen and colony phenotypic traits.

REFERENCES

1. **Mitton JB, Grant MC**. 1984. Associations among protein heterozygosity, growth rate, and developmental homeostasis. Annual Review of Ecology and Systematics **15**:479–499.

2. **Pogson GH**. 1991. Expression of overdominance for specific activity at the Phosphoglucomutase-2 locus in the pacific oyster, Crassostrea gigas. Genetics **128**:133–141.

3. **Pogson GH, Zouros E**. 1994. Allozyme and RFLP heterozygosities as correlates of growth rate in the scallop Placopecten magellanicus: A test of the associative overdominance hypothesis. Genetics **137**:221–231.

4. **Hedgecock D, McGoldrick DJ, Manahan DT, Vavra J, Appelmans N, Bayne BL**. 1996. Quantitative and molecular genetic analyses of heterosis in bivalve molluscs. Journal of Experimental Marine Biology and Ecology **203**:49–59.

5. **Olano-Marin J, Mueller JC, Kempenaers B**. 2011. Heterozygosity and survival in blue tits (Cyanistes caeruleus): contrasting effects of presumably functional and neutral loci. Molecular Ecology **20**:4028–4041.

6. **Forstmeier W, Schielzeth H, Mueller JC, Ellegren H, Kempenaers B**. 2012. Heterozygosity-fitness correlations in zebra finches: microsatellite markers can be better than their reputation. Molecular Ecology **21**:3237–3249.

7. **Ruiz-López MJ, Gañan N, Godoy JA, Del Olmo A, Garde J, Espeso G, Vargas A, Martinez F, Roldán ERS, Gomendio M, Olmo ANADEL**. 2012. Heterozygosity-fitness correlations and inbreeding depression in two critically endangered mammals. Conservation Biology **26**:1121–1129.

8. **Koonin E V, Senkevich TG, Dolja V V**. 2006. The ancient Virus World and evolution of cells. Biology Direct **1**:29.

9. **Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA**. 2005. Virus Taxonomy: VIIIth Report of the International Committee on Taxonomy of Viruses, 8th ed. Elsevier Inc., Oxford, UK.

10. **Asgari S, Johnson K**. 2010. Insect VirologyInsect Virology. Caister Academic Press, Norfolk, UK.

11. **Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC**. 2009. A quantitative review of heterozygosity-fitness correlations in animal populations. Molecular Ecology **18**:2746–2765.

12. **Balloux F, Amos W, Coulson T**. 2004. Does heterozygosity estimate inbreeding in real populations? Molecular Ecology **13**:3021–3031.

13. **Slate J, David P, Dodds KG, Veenvliet BA, Glass BC, Broad TE, McEwan JC**. 2004. Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. Heredity **93**:255–265.

14. **Szulkin M, Bierne N, David P**. 2010. Heterozygosity-fitness correlations: a time for reappraisal. Evolution **64**:1–16.

15. **Houle D, Morikawa B, Lynch M**. 1996. Comparing Mutational Variabilities. Genetics **143**:1467–1483.

16. **Orr HA, Coyne JA**. 1992. The Genetics of Adaptation : A Reassessment. The American Naturalist **140**:725–742.

17. **David P**. 1997. Modeling the genetic basis of heterosis: tests of alternative hypotheses. Evolution **51**:1049–1057.

18. **Chapman JR, Sheldon BC**. 2011. Heterozygosity is unrelated to adult fitness measures in a large, noninbred population of great tits (Parus major). Journal of Evolutionary Biology **24**:1715–1726.

19. **Coulson TN, Pemberton JM, Albon SD, Beaumont M, Marshall TC, Slate J, Guinness FE, Clutton-Brock TH**. 1998. Microsatellites reveal heterosis in red deer. Proceedings of the Royal Society of London Series B-Biological Sciences **265**:489–495.

20. **Marshall TC, Spalton JA**. 2000. Simultaneous inbreeding and outbreeding depression in reintroduced Arabian oryx. Animal Conservation **3**:241–248.

21. **Jourdan-Pineau H, Folly J, Crochet P-A, David P**. 2012. Testing the influence of family structure and outbreeding depression on heterozygosity-fitness correlations in small populations. Evolution **66**:3624–3631.

22. **Laine VN, Herczeg G, Shikano T, Primmer CR**. 2012. Heterozygosity-behaviour correlations in nine-spined stickleback (Pungitius pungitius) populations: contrasting effects at random and functional loci. Molecular Ecology **21**:4872–4884.

23. **Mueller JC, Hermisson J, Olano-Marin J, Hansson B, Kempenaers B**. 2010. Linking genetic mechanisms of heterozygosity-fitness correlations to footprints of selection at single loci. Evolutionary Ecology **25**:1–11.

24. **Deng H-W, Fu Y-X**. 1998. Conditions for positive and negative correlations between fitness and heterozygosity in equilibrium populations. Genetics **148**:1333–40.

25. **Caldera EJ, Ross KG, DeHeer CJ, Shoemaker DD**. 2008. Putative native source of the invasive fire ant Solenopsis invicta in the USA. Biological Invasions **10**:1457–1479.

26. **IUCN TWCU**. 2000. One-hundred of the World's Worst Invasive Alien Species: A Selection from the Global Invasive Species Database.

http://www.iucn.org/places/medoffice/invasive_species/docs/invasive_species_booklet.pdf.

27. **Ross KG, Shoemaker DD**. 2008. Estimation of the number of founders of an invasive pest insect population: the fire ant Solenopsis invicta in the USA. Proceedings of the Royal Society of London Series B-Biological Sciences **275**:2231–2240.

28. **Tschinkel WR**. 2006. The Fire Ants, 1st ed. The Belknap Press of Harvard University Press, Cambridge, MA.

29. **Bryant EH, McCommas SA, Combs LM**. 1986. The effect of an experimental bottleneck upon quantitative genetic variation in the housefly. Genetics **114**:1191–1211.

30. **Goodnight CJ**. 1988. Epistasis and the Effect of Founder Events on the Additive Genetic Variance. Evolution **42**:441–454.

31. **Hoffman JI, Hanson N, Forcada J, Trathan PN, Amos W**. 2010. Getting long in the tooth: a strong positive correlation between canine size and heterozygosity in Antarctic fur seals Arctocephalus gazella. The Journal of Heredity **101**:527–538.

32. **Slate J, Pemberton JM**. 2002. Comparing molecular measures for detecting inbreeding depression. Journal of Evolutionary Biology **15**:20–31.

33. **Morrill W**. 1974. Production and flight of alate red imported fire ants. Environmental Entomology **3**:265–271.

34. **Markin, Markin GP, Dillier JH, Hill SO, Blum MS, Hermann HR**. 1971. Nuptial flight and flight ranges of the imported fire ant, Solenopsis invicta. Journal of the Georgia Entomological Society **6**:145–156.

35. **Shoemaker DD, Deheer CJ, Krieger MJB, Ross KG**. 2006. Population genetics of the invasive fire ant Solenopsis invicta (Hymenoptera : Formicidae) in the United States. Annals of the Entomological Society of America **99**:1213–1233.

36. **Ross KG, Shoemaker DD**. 1997. Nuclear and mitochondrial genetic structure in two social forms of the fire ant Solenopsis invicta: insights into transitions to an alternate social organization. Heredity **78**:590–602.

37. **Ross KG, Shoemaker DD, Krieger MJ, DeHeer CJ, Keller L**. 1999. Assessing genetic structure with multiple classes of molecular markers: a case study involving the introduced fire ant Solenopsis invicta. Molecular Biology and Evolution **16**:525–543.

38. **Werren JH**. 1993. The Evolution of Inbreeding in Haplodiploid Organisms, p. 42–59. *In* Thornhill, NW (ed.), The Natural History of Inbreeding and Outbreeding, 1st ed. The University of Chicago Press, Chicago.

39. **Ometto L, Shoemaker D, Ross KG, Keller L**. 2011. Evolution of gene expression in fire ants: the effects of developmental stage, caste, and species. Molecular Biology and Evolution **28**:1381–1392.

40. **Haag-Liautard C, Vitikainen E, Keller L, Sundström L, Sundstrom L**. 2009. Fitness and the level of homozygosity in a social insect. Journal of Evolutionary Biology **22**:134–142.

41. **Vitikainen E, Haag-Liautard C, Sundström L**. 2011. Inbreeding and reproductive investment in the ant formica exsecta. Evolution **65**:2026–2037.

42. **Kureck IM, Jongepier E, Nicolai B, Foitzik S**. 2012. No inbreeding depression but increased sexual investment in highly inbred ant colonies. Molecular Ecology **21**:5613–5623.

43. **Wetzel DP, Stewart IRK, Westneat DF**. 2012. Heterozygosity predicts clutch and egg size but not plasticity in a house sparrow population with no evidence of inbreeding. Molecular Ecology **21**:406–420.

44. **Alho JS, Välimäki K, Merilä J, Valimaki K, Merila J**. 2010. Rhh: an R extension for estimating multilocus heterozygosity and heterozygosity-heterozygosity correlation. Molecular Ecology Resources **10**:720–722.

45. **David P, Pujol B, Viard F, Castella V, Goudet J**. 2007. Reliable selfing rate estimates from imperfect population genetic data. Molecular Ecology **16**:2474–2487.

46. **Amos W, Acevedo-Whitehouse K**. 2009. A new test for genotype-fitness associations reveals a single microsatellite allele that strongly predicts the nature

of tuberculosis infections in wild boar. Molecular Ecology Resources **9**:1102–1111.

47. **Amos W, Wilmer JW, Fullard K, Burg TM, Croxall JP, Bloch D, Coulson T**. 2001. The influence of parental relatedness on reproductive success. Proceedings of the Royal Society of London Series B-Biological Sciences **268**:2021–2027.

48. **Aparicio JM, Ortego J, Cordero PJ**. 2006. What should we weigh to estimate heterozygosity, alleles or loci? Molecular Ecology **15**:4659–4665.

49. **Breitbart M, Rohwer F**. 2005. Here a virus, there a virus, everywhere the same virus? Trends in Microbiology **13**:278–284.

50. **Huger AM**. 1966. A virus disease of the Indian rhinoceros beetle, Oryctes rhinoceros(Linnaeus), caused by a new type of insect virus, Rhabdionvirus oryctes gen. n., sp. n. Journal of Invertebrate Pathology **8**:38–51.

51. **Klatzmann D, Gluckman JC**. 1987. The pathophysiology of HIV infection: A complex pathway of host-virus interaction. Annales de l'Institut Pasteur-Virologie **138**:21–29.

52. **Tijssen P, Bergoin M**. 1995. Densonucleosis viruses constitute an increasingly diversified subfamily among the parvoviruses. Seminars in Virology **6**:347–355.

53. **Pelletier I, Duncan G, Pavio N, Colbère-Garapin F**. 1998. Molecular mechanisms of poliovirus persistence: key role of capsid determinants during the establishment phase. Cellular and Molecular Life Sciences **54**:1385–1402.

54. **Jousset FX, Baquerizo E, Bergoin M**. 2000. A new densovirus isolated from the mosquito Culex pipiens (Diptera: culicidae). Virus Research **67**:11–16.

55. **Chen YP, Siede R**. 2007. Honey bee viruses. Advances in Virus Research **70**:33–80.

56. **Steinhaus EA**. 1956. Microbial control-the emergence of an idea. A Journal of Agricultural Science **26**:107–160.

57. **De Miranda JR, Genersch E**. 2010. Deformed wing virus. Journal of Invertebrate Pathology **103**:S48–561.

58. **Fujiyuki T, Takeuchi H, Ono M, Ohka S, Sasaki T, Nomoto A, Kubo T**. 2004. Novel Insect Picorna-Like Virus Identified in the Brains of Aggressive Worker Honeybees. Journal of Virology **78**:1093–1100.

59. **Nakanishi K, Hoshino M, Nakai M, Kunimi Y**. 2008. Novel RNA sequences associated with late male killing in Homona magnanima. Proceedings of the Royal Society B-Biological Sciences **275**:1249–1254.

60. **Habayeb MS, Ekengren SK, Hultmark D**. 2006. Nora virus, a persistent virus in Drosophila, defines a new picorna-like virus family. The Journal of General Virology **87**:3045–51.

61. **Habayeb MS, Cantera R, Casanova G, Ekström J-O, Albright S, Hultmark D**. 2009. The Drosophila Nora virus is an enteric virus, transmitted via feces. Journal of Invertebrate Pathology **101**:29–33.

62. **Valles SM, Strong CA, Dang PM, Hunter WB, Pereira RM, Oi DH, Shapiro AM, Williams DF**. 2004. A picorna-like virus from the red imported fire ant, Solenopsis invicta: initial discovery, genome sequence, and characterization. Virology **328**:151–157.

63. **Valles SM, Hashimoto Y**. 2009. Isolation and characterization of Solenopsis invicta virus 3, a new positive-strand RNA virus infecting the red imported fire ant, Solenopsis invicta. Virology **388**:354–361.

64. **Valles SM, Strong CA, Hashimoto Y**. 2007. A new positive-strand RNA virus with unique genome characteristics from the red imported fire ant, Solenopsis invicta. Virology **365**:457–463.

65. **Valles SM, Hashimoto Y**. 2008. Detection and quantitation of Solenopsis invicta virus-2 genomic and intermediary replicating viral RNA in fire ant workers and larvae. Journal of Invertebrate Pathology **98**:243–245.

66. **Hashimoto Y, Valles SM**. 2008. Infection characteristics of Solenopsis invicta virus 2 in the red imported fire ant, Solenopsis invicta. Journal of Invertebrate Pathology **99**:136–140.

67. **Valles SM**. 2012. Positive-Strand RNA Viruses Infecting the Red Imported Fire Ant, Solenopsis invicta. Psyche: A Journal of Entomology **2012**:1–14.

68. **Julien J, Leparc-Goffart I, Lina B, Fuchs F, Foray S, Janatova I, Aymard M, Kopecka H**. 1999. Postpolio syndrome: poliovirus persistence is involved in the pathogenesis. Journal of Neurology **246**:472–476.

69. **Klingel K, Hohenadl C, Canu A, Albrecht M, Seemann M, Mall G, Kandolf R**. 1992. Ongoing enterovirus-induced myocarditis is associated with persistent heart muscle infection: quantitative analysis of virus replication, tissue damage, and inflammation. Proceedings of the National Academy of Sciences of the United States of America **89**:314–318.

CHAPTER 2

HETEROZYGOSITY-FITNESS ASSOCIATIONS IN AN INVASIVE POPULATION OF THE FIRE ANT *SOLENOPSIS INVICTA*[1]

[1] Fisher, Mark A., Shoemaker, D.D., Ross, K.G. To be submitted to *Evolution.*

ABSTRACT

A main goal of evolutionary biology is to clarify the forces that maintain genetic diversity. Heterozygosity-fitness association (HFA) studies, which characterize the relationship between individual multilocus heterozygosity and fitness, have commonly been conducted to address this issue. Empirical difficulties and variation in population characteristics have contributed to a lack of uniformity in HFA study outcomes, with the resulting explanatory theory correspondingly complex and contentious. We conducted an HFA study on several hundred colony-founding *Solenopsis invicta* queens by genotyping them at 105 polymorphic loci and measuring 12 fitness-associated traits. We discovered relatively few significant HFAs despite the presence of many population and study features favoring their detection (including a low level of inbreeding). Both "positive" and "negative" HFAs were among these, and in one case virus-infected (stressed) and uninfected queens differed in the direction of the association for the same trait. Neither inbreeding depression nor single-locus genotype-trait associations (SLAs) appear to explain the observed positive HFAs, while neither outbreeding depression nor SLAs fully explain the negative HFAs. We suggest that the pattern of HFAs we found instead may arise from life-history tradeoffs and the effects of stress on inter-related traits, emphasizing these as important factors to consider in future HFA studies.

INTRODUCTION

An essential goal of evolutionary biology is to determine the role of genetic diversity both in the vigor of populations facing new selection pressures and in the survival and reproductive success of individuals within such populations. In the latter case, it is important to measure the genetic variability of individuals, usually with some estimator of multilocus heterozygosity, and to determine the strength of an association between this measure and relative fitness. Studies that investigate such relationships are called heterozygosity-fitness correlation (HFC) studies or, perhaps more appropriately, heterozygosity-fitness association (HFA) studies (e.g., 1–7). Although most such studies have reported a weak, positive effect (8), there is substantial variation among studies in their results. However, measures of multilocus heterozygosity determined by the small panels of marker loci typically employed (8) are notoriously poor predictors of genome-wide heterozygosity (9, 10), raising doubts as to the meaning of the variable results obtained thus far. New studies employing large numbers of markers thus are required to more accurately resolve relationships between individual heterozygosity and fitness in wild populations.

Several potential explanations have been advanced for those associations between multilocus heterozygosity and survival or reproductive success observed to date: i) the surveyed marker loci directly affect the fitness trait(s) of interest, a possibility only if these loci are non-neutral (are located within coding sequence or regulatory elements), ii) the marker loci are in linkage disequilibrium (LD) with loci directly influencing the fitness trait(s), and iii) individuals with lower heterozygosity suffer costs of inbreeding, or individuals with higher heterozygosity suffer costs of outbreeding, relative to the remainder of the population. While the latter two explanations both rely on associations between coding and marker loci to explain HFAs, the third

("general effect" hypothesis) differs from the second ("local effect" hypothesis) in that it typically invokes some form of inbreeding arising from nonrandom mating, small population size, bottlenecks, or admixture (which can render non-admixed individuals inbred relative to admixed ones) to generate extensive intra-genomic associations (11).

HFAs attributable to the general effects of inbreeding depression make intuitive sense in terms of the genetic architecture of fitness traits, which are believed largely to be the products of many genes of small effect distributed throughout the genome (12, 13). By causing an association between homozygous genotypes at different loci — thereby creating identity disequilibrium (ID) — inbreeding can drastically affect the phenotype of polygenic traits, like those commonly associated with fitness, by altering the genotypes of several loci at once; such traits are otherwise considered to be resistant to rapid change because of the large and diffuse mutational targets they represent. Thus, it is this third, general effect explanation for HFAs that currently is favored by many theoreticians and empiricists (11, 14, 15). Nonetheless, an association between population demographic history and effect size of observed positive HFAs, which theory suggests should exist if inbreeding is their predominant cause (11), generally is absent (8), lending some support to the two alternate hypotheses not invoking general effects. This theoretical framework applies as well to negative HFAs, in which heterozygosity is negatively associated with fitness-trait values, although the distinction between general effects versus direct/local effects usually is couched instead in terms of outbreeding depression versus single-locus genotype-trait associations (SLAs) (e.g., 5, 16–19).

Incomplete knowledge of the extent of departures from panmixia in many study systems, coupled with uncertainty about the functional roles and selective consequences of inbreeding and outbreeding depression in generating HFAs, likely contribute along with small marker panel size

to the variability in results and consequent lack of consensus regarding the importance of direct, local, or general effects (8). No doubt also contributing are substantial variation in the types and quality of fitness-associated traits measured and in the sample sizes employed, as well as the potential for alleles under directional selection to generate transient HFAs (20, 21).

Solenopsis invicta is an invasive fire ant introduced to the U.S.A. from Argentina in the early 1930s (22). It has since spread rapidly throughout the southeastern and parts of the southwestern U.S.A., where it has emerged as a major agricultural and urban pest (23). Several considerations make *S. invicta* a promising subject for an HFA study, including potential increases in additive genetic variance in fitness and in LD between marker and fitness loci, both induced by the founder event, availability of a large complement of genetic markers with which to measure heterozygosity, and a firm understanding of traits that are intimately associated with fitness, especially in the early life of a colony.

The effective founding population of North American *S. invicta* is estimated to have comprised 9 to 20 mated queens (24), and only 10 to 20 generations have passed since the founding event (24, 25). Empirical observations (26) and theoretical work (27) suggest that both additive genetic variance and fitness variance can increase after such a genetic bottleneck, and subsequent spread of rare or novel genotypes and the fitness phenotypes they influence in the expanding population may enhance HFAs. A bottleneck also increases ID and LD (11), a fact that can be leveraged to detect the effects of fitness loci using associated marker loci when sufficiently few generations have passed for such disequilibrium to decay. The potential for detecting HFAs in our study system also is enhanced by the availability of a large number (> 100) of polymorphic markers. This large panel size enables us to more accurately assess genome-wide heterozygosity, a necessity made salient by two recent publications that used large panels (~

70 loci) to recant evidence for a general effect detected by smaller marker panels in earlier studies of the same systems (28, 29). Finally, rich background knowledge of *S. invicta* natural history implicates several measurable individual and colony-level traits intimately tied to fitness during the crucial colony-founding phase of the life cycle.

As in most wild populations, some factors also may obscure detection of HFA in invasive *S. invicta*. For instance, fire ants engage in massive aerial mating/dispersal swarms that often are synchronized over vast areas (25, 30, 31), so there is little reason to expect the existence of any consanguineous matings that result in ID and, consequently, positive HFAs. Also, HFA signals may be weak in *S. invicta* because recessive deleterious alleles — presumably responsible for positive HFAs due to inbreeding depression in a portion of the population — may be relatively readily removed via selection on hemizygous males (32). However, this purging scenario, which relies on the haploidy of males common to the order Hymenoptera, also assumes that reproductives of the two sexes generally express genes at similar levels, a pattern not well supported for *S. invicta* males and queens (33).

The types of relationships expected between individual heterozygosity and fitness in the invasive U.S. population of *S. invicta* thus are unclear. More broadly, little empirical work intended to clarify such relationships has been conducted either in social Hymenoptera (but see 34, 35) or in invasive populations (but see 36). We therefore undertook a study that remedies these deficiencies by combining genotyping at a large number of markers with careful measurement of clear fitness-related traits in a large sample of colony-founding queens of invasive *S. invicta*.

MATERIALS AND METHODS

Samples

We collected 652 newly-mated *S. invicta* queens from a single site in Oglethorpe County, Georgia, U.S. on 12 June 2008 as they landed from their mating flights. These queens were all 2-6 weeks of age (post-eclosion), judging from the time of year and lack of previous late-spring flights in the area. The Oglethorpe County site was selected because it was previously identified as having predominantly monogyne (single-queen) colonies. We collected only once at a single site to avoid any possibility of population genetic structure that can result from sampling multiple sites or time points and can weaken the power of HFA analyses (37; but see 38, 39 for evidence that spatial genetic structure is weak in the U.S.A.). On the same day that they were collected, a subset of 400 queens was transferred individually to small, stoppered, plastic tubes held in our rearing room (14:10 h light:dark cycle, 28-30°C, 40-70% RH); these tubes served as nesting sites for the queens to initiate their colonies (see 40, 41). Queens were allowed to lay eggs and rear their first cohort of brood without being fed, as colony-founding queens of the monogyne form fast during this claustral founding period (25). At the time that their first adult worker offspring eclosed (emerged as adults from the pupal stage, marking the end of the claustral period), a subset of 300 of the 400 founding queens was collected along with all of their offspring and stored in a -80°C freezer. The remaining 100 queens and their offspring were transferred to larger rearing units, and the colonies were allowed to develop for another six weeks before being terminated and frozen. These post-claustral colonies 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), supplemented with frozen crickets provided on a twice-weekly

basis. Workers in post-claustral colonies in the wild forage in order to support the continued growth and survival of the colony.

Fitness-associated traits

We selected fitness-related traits for measurement in incipient colonies based on extensive previous studies of fire ant natural history (25). These studies show that rapid growth is essential to newly founded monogyne *S. invicta* colonies because the primary sources of early mortality in U.S. populations predominantly afflict relatively slowly growing colonies. These sources are competition with other fire ant colonies (other incipient colonies in newly colonized areas) and climate-related stress during winter (42, 43). Higher worker numbers enable a young colony to compete successfully for an exclusive foraging territory and to excavate its nest to a sufficient depth in the soil to prevent freezing in winter (43, 44). Indeed, a positive correlation between colony size at some critical early period and likelihood of survival may be a general feature of many hymenopteran societies in temperate areas (45, 46).

We measured two sets of fitness traits, the first set comprising individual traits of queens and the second comprising traits of the colonies they founded, which, in highly eusocial insects such as fire ants, represent the extended phenotype of the queen. Individual traits included queen size, initial queen weight, proportional queen weight loss during the claustral period, whether the queen was infected with any of three different viruses, queen survivorship through the claustral period, and queen survivorship during the six weeks following the claustral period (which can also be considered a colony-level trait). Colony traits included the number of days between a queen's mating flight and eclosion of her first adult worker (i.e., the length of the claustral period), colony weight at the end of the claustral period, colony population (number of 4^{th}-instar larval, pupal, and adult workers) at the end of the claustral period, colony weight six weeks after

the end of the claustral period, colony population six weeks after the end of the claustral period, and proportional colony growth between the end of the claustral period and six weeks later (based on colony weight). Because colony population and weight were highly correlated at both time points at which they were measured (Spearman's rank correlation coefficient $r = 0.869$ and 0.712, respectively, both $P < 0.0001$), we report only analyses using weight. Where convenient, the twelve fitness traits employed are henceforth referenced using the abbreviations in Table 1.

Queen size measurements were obtained by removing the left antenna, left hindleg, and head of each queen, placing these on a microscope slide, and measuring the length of the antennal scape, length of the metathoracic femur, and width of the head (between the lateral-most points of the compound eyes) to the nearest 0.001 mm using the *Leica Application Suite* software (v.3.8.0) on a Leica DFC295 microscope. In holometabolous insects such as ants, the size of such external adult structures is set in the pupal stage in response to genotype or conditions during larval development (e.g., feeding rate) and does not change during adult ontogeny. We conducted principal components analysis (PCA) on the covariance matrix for these three measurements using the princomp() function in *R* (v.2.15.1) (47). Because the first principal component, which loaded heavily for all three measures, explains more than 50% of the total variance in the measurement data, we used it as a proxy for size in all subsequent analyses. Queen size is regarded as a potential fitness-related trait because greater size favors success in competition among queens during pleometrotic (cooperative) colony founding in monogyne *S. invicta* (48) and is associated with enhanced queen survival in the harvester ant *Pogonomyrmex occidentalis* (49).

Queen weight was measured to the nearest 0.1 mg on the day of the mating flight and at the end of the claustral period. Weight is a highly labile trait in young fire ant queens, reflecting

such factors as the amount of fat reserves sequestered prior to the mating flight and the rate at which such reserves are depleted during the claustral period (50). In general, queens with greater reserves are expected to have greater success in bringing their incipient colonies through the claustral phase (25).

Viral infection of queens was determined by assaying for the presence of three viruses (SINV-1, SINV-2, SINV-3) known to infect and cause mortality at varying rates in *S. invicta* in the U.S.A. (51–53). RNA was extracted from the anterior gaster segments of each queen (see below) using an AllPrep DNA/RNA Mini Kit (Qiagen); the manufacturer's instructions were followed, except the final product was eluted into 10 μL EB buffer. The RNA was then subjected to a two-step RT-PCR protocol (Fisher et al. unpubl. ms.), with each queen extract assayed for the three viruses and the transcript of the nuclear gene *Gp-9* — the last as a control for successful viral RNA extraction and cDNA synthesis.

During the claustral period, colonies were checked twice daily for the presence of adults, and the time (to the nearest complete day) until the first adult worker eclosed was recorded. Colony weight upon eclosion of the first adult worker and six weeks thereafter were measured to the nearest 0.1 mg. Proportional colony growth was measured as number-fold increase in colony weight (excluding the queen) between the two time points. Queen (colony) survival was assessed as a binary variable at two separate intervals: i) through the claustral phase, during which queen survival is independent of worker foraging, and ii) through a six-week period immediately following the end of the claustral phase, when colonies become well-established and workers actively forage in support of the queen and colony.

Protein, DNA, and RNA extraction

Prior to extractions, the body of each queen was divided into three parts: the head, mesosoma (thorax plus propodeum), and gaster (abdomen minus propodeum). The gaster was further subdivided into anterior, middle, and posterior sections. Protein, DNA, and RNA for allozyme, microsatellite, and viral assays, respectively, were extracted from the different parts of each queen. Soluble proteins were extracted separately from the thorax and middle gaster sections (54). DNA was extracted from the tissue remaining after these protein extractions using a DNeasy Blood and Tissue Kit (Qiagen) and following the manufacturer's instructions. DNA and RNA were simultaneously extracted from the anterior gaster section using AllPrep DNA/RNA Mini Kits (Qiagen) as described above. The posterior gaster section (including the spermatheca) was archived for future study.

Genotyping

Candidate microsatellites were developed as in Ascunce et al. (55) by searching for microsatellite motifs either in an *S. invicta* expressed sequence tag (EST) database available in GenBank (56) or in a normalized cDNA library sequenced using 454 pyrosequencing technology. Primers were designed using *Primer3* (57); primer sequences are listed along with cycling profile, core repeat unit, and other marker-specific information in Appendix A.1. PCR was performed on extracted DNA as in Ascunce et al. (55) with the following modifications: the total reaction volume was 10.4 μL, and each reaction contained 0.05-0.7 μL of a 10 mM solution of each primer, 6.4 μL of 2X *Taq*-Pro COMPLETE (2.0 mM $MgCl_2$, Denville Scientific), 1.0 μL of total genomic DNA, and H_2O to volume. Genotype analysis of the fluorescence intensity spectra was conducted using *GeneMarker* (v.2.2.0, SoftGenetics). After removing from the original pool of 652 those queens that were triploid or that came from polygyne (multiple-queen) colonies (see

below), genotypes of 593 queens were surveyed at 99 polymorphic microsatellite loci, 45 of which were developed previously (55, 58) and 54 of which were developed for this study. No null alleles were detected empirically at any of these loci in a survey of 96 haploid *S. invicta* males from various locations in the U.S.A. (data not shown).

Soluble proteins extracted from the 652 collected queens were used for genotyping at six polymorphic allozyme loci (Appendix A.1) as well as the protein-coding gene *Gp-9* by means of specific histochemical or non-specific protein staining in 14% horizontal starch gels (54, 59). No evidence exists to suggest the presence of null alleles (pronounced heterozygote deficiencies for diploid females, dropouts for haploid males) at any of the six allozyme loci (e.g., 38, 39, 60).

Linkage disequilibrium (LD)

We tested for composite LD between each pair of marker loci (5356 comparisons) using *GenePop* (v.4.0.10). Correction for multiple testing was made using the Benjamini-Hochberg procedure in *R*, with an expected false discovery rate (FDR) of 5%.

Individual heterozygosity

We calculated individual multilocus heterozygosity (MLH), the proportion of heterozygous loci among those successfully genotyped in a queen (8), using a custom script written in *Python* (v.2.7.3); we focus on this metric in presenting the results of the HFA and other analyses because of its simplicity and straightforward interpretation (8). We further assessed individual heterozygosity by calculating homozygosity by loci (HL) (61), internal relatedness (IR) (62), and standardized multilocus heterozygosity (SH) (63) using the *Rhh* package (v.1.0.2) in *R* (64). These metrics were employed in the HFA analyses to assess consistency with the results obtained using the MLH metric. IR values were used as well to directly estimate the inbreeding coefficient, *F* (62).

Identity disequilibrium (ID) and departure from Hardy-Weinberg Equilibrium (HWE)

We tested whether multilocus heterozygosity as measured by the marker panel was reflective of genome-wide heterozygosity using the heterozygosity-heterozygosity correlation (HHC) test (9). This test calculates the correlation coefficient between heterozygosity measured as HL for two subsets of randomly chosen markers from the entire panel; this process is repeated a large number of times (250 in our case) to obtain a distribution of correlation coefficients that is tested for significant departure from a distribution centered on zero. By quantifying how strongly heterozygosity in randomized marker subsets covaries, the HHC test also assesses genomic ID and, consequently, level of inbreeding (11). We further explored the issue of whether our marker panel heterozygosity reflects genome-wide heterozygosity, as well as quantified ID, by estimating the second-order heterozygosity disequilibrium (g2) metric, which assesses the excess of double-heterozygotes and double-homozygotes relative to expectations of independent assortment using all pairs of markers in the panel (65). The g2 metric, unlike the HHC test, can be used to directly derive the expected correlation between multilocus heterozygosity (SH) and the inbreeding coefficient, *F* (11). Values of g2, along with their associated *P*-values, were calculated based on 1000 iterations using the software *RMES* (65).

To learn whether marker genotype distributions departed significantly from HWE expectations, we performed exact tests on each locus using *GenePop*. We determined the direction of deviation (heterozygote excess or deficiency) by calculating single-locus estimates of Wright's fixation index F_{IS} (66) using *GenePop*. Global (multilocus) estimates of F_{IS} were calculated using the software *GDA* (v.1.1) (67).

Simulation analyses

We used the allele frequencies in our original dataset to randomly assign multilocus genotypes to 3000 simulated individuals at the entire panel of markers using a custom *Python* script. We then conducted inbreeding analyses on the simulated dataset to generate the expected results for a randomly mating (HWE) population with allele frequencies identical to those in our study population.

Heterozygosity-Fitness Associations (HFAs)

We evaluated the ability of individual multilocus heterozygosity to explain fitness-associated trait values by employing it as a predictor in generalized linear models (GLMs) with (i) an identity link function and Gaussian error distribution for the response variables *QnSize*, *QnWt*, $QnWtLs_{CL}$, $ColWt_{CL}$, and $ColGrowth_{CL\text{-}6}$, (ii) a logit link function and binomial error distribution for the response variables *QnSINV-1, QnSINV-2, QnSINV-3*, $QnSurv_{CL}$, and $QnSurv_{CL\text{-}6}$, or (iii) a log link function and Poisson error distribution for the response variables *ClaustPer* and $ColGrowth_{CL\text{-}6}$. Corrections for multiple testing were made using the Benjamini-Hochberg procedure with a 5% FDR.

We ran additional GLM regressions on four data subsets to learn whether queens with more extreme heterozygosity values exert a strong influence on the regression results. These subsets of the complete dataset excluded queens with MLH values (i) in the lowest quartile of the observed distribution, (ii) in the 5% lowest values of the simulated dataset, (iii) in the highest quartile of the observed distribution, or (iv) in the 5% highest values of the simulated dataset.

Heterozygosity-fitness associations were further examined by testing for significant differences between mean trait values in "low-heterozygosity" and "high-heterozygosity" queens using Mann-Whitney *U* tests for continuous traits and Fisher's exact tests for binary traits.

Membership in the two groups was assigned either i) based on the appearance of a clear break in the distribution of individual MLH values, with queens having values ≤ 0.330 placed in the former category and the remaining queens placed in the latter category, or ii) by placing queens in the lowest quartile of MLH values (MLH ≤ 0.406) in the low-heterozygosity category and queens in the highest quartile of MLH values (MLH ≥ 0.476) in the high-heterozygosity category.

In an effort to rule out any systematic bias that poorly-scored samples or loci might exert on the regression results, we also ran all of the GLM analyses separately on data subsets including i) only individuals whose genotypes were scored successfully for at least 75, 80, 85, or 95 loci, and ii) only the 75 loci that were scored successfully in at least 490 queens (a threshold based on inspection of the distribution of numbers of queens scored at each locus).

Several studies have found that HFAs are observed more readily in individuals challenged by some stressor (e.g., 68, 69). We accordingly also ran regressions for predictors that were significant for the complete dataset on the data subset comprising only queens infected with SINV-2 (the only surveyed virus to show strong associations with other trait measures [Fisher et al. unpubl. ms.]), under the expectation that such individuals were stressed by this pathogen.

Single-Locus Associations with Traits (SLAs)

To determine whether heterozygosity at individual loci is significantly associated with trait values in our study, we compared a regression model invoking MLH as the single predictor of each fitness trait to a multiple regression model invoking the heterozygosities of each marker locus as separate predictors (11). In each multiple regression model, heterozygosity of each locus was reported as 1 or 0, with missing genotypes replaced by the average heterozygosity at that locus (11). We used an *F*-ratio test to determine whether the second model explained

significantly more of the variance in trait values than the first, in which case heterozygosity at one or more individual loci is implicated in influencing the trait values. Corrections for multiple testing were made using the Benjamini-Hochberg procedure with a 5% FDR.

We next scanned each locus for associations of any genotype(s) with each fitness trait using the genotype-by-phenotype association test implemented in the *GEPHAST* program (70). The *GEPHAST* algorithm classifies genotypes at a locus into "high risk" and "low risk" categories according to whether or not they are found predominantly in individuals with above-average trait values. The score of a *t*-test comparing the trait value mean between "high risk" and "low risk" genotypes is then compared to *t*-scores for 100,000 datasets in which the genotypes are randomly assigned the trait values from the original dataset. The percentile represented by the *t*-score of the observed dataset within the larger distribution of randomization *t*-scores is taken as the *P*-value. These *P*-values were adjusted using the Benjamini-Hochberg procedure with a 5% FDR.

Effects on trait values of single-locus genotypes, or of single-locus genotypes pooled on the basis of shared alleles, were examined further for eight loci implicated by *GEPHAST*. Significance of trait value differences was assessed by means of Mann-Whitney *U* tests for the loci affecting *QnSize* and $QnWtLs_{CL}$ and by Fisher's exact tests for the loci affecting $QnSurv_{CL}$ and $QnSurv_{CL\text{-}6}$. For the trait *ClaustPer*, values were pooled into two categories, short claustral periods (< 23 days) and long claustral periods (≥ 23 days), and genotype-associated differences were assessed by means of Fisher's exact tests.

Excluded Samples

We excluded from all analyses queens bearing the *b* allele of the gene *Gp-9* because such queens, which originate exclusively from colonies of the polygyne social form (featuring multiple

queens per colony), generally exhibit reproductive strategies and physiologies that differ strongly from those of monogyne queens (41, 71). We further excluded from all HFA analyses queens whose microsatellite genotypes indicated they were triploids (products of matings between a diploid male and queen), because these individuals appear generally to have low fitness compared to diploid queens (72). A low proportion of *S. invicta* queens in the U.S.A. mate with a male bearing an allele at the sex-determining locus that matches one of the queen's alleles (73). Because such queens produce (diploid) male offspring during claustral founding, which invariably leads to colony failure (40), we excluded such queens from all HFA analyses other than those involving queen size, initial queen weight, and viral presence. Finally, we excluded from all analyses related to viral presence samples for which RT-PCR of *Gp-9* failed (successful *Gp-9* amplification was used as a control for successful viral RNA extraction and RT-PCR). After application of these filtering criteria, the most inclusive analyses featured sample sizes of 593 queens.

RESULTS

General results

Specific information for each locus included in the final marker panel is reported in Appendix A.1. For both microsatellites and allozymes, this includes the number of individuals scored at each locus and several measures of marker diversity, while for the microsatellites only, primer sequences, cycling profiles, core repeat units, repeat size, and scaffold in the current *S. invicta* genome assembly on which each locus is located also are reported. The microsatellites featured a mean of 3.88 alleles observed per locus (range 2, 13; mean effective number of alleles = 2.16, range 1.02, 5.67), with mean single-locus H_{obs} = 0.452 (range 0.015, 0.839) and H_{exp} = 0.468 (range 0.015, 0.824) for this marker class. All allozyme loci had only two alleles (mean

effective number of alleles = 1.42, range 1.19, 1.95), with mean H_{obs} = 0.267 (range 0.155, 0.476) and H_{exp} = 0.276 (range 0.160, 0.489). Allele frequencies for each locus are reported in Appendix A.2.

Results of the tests for composite linkage disequilibrium (LD) between each pair of loci are shown in Appendix A.3. Among the 5356 comparisons, 325 (6.06%) comprise significantly non-random allelic associations, with only 24 (0.45%) remaining significant after correction for multiple comparisons (Benjamini-Hochberg procedure, 5% FDR). Among these latter cases are one triplet (*Bertha/C147/Heartland*) and seven pairs of loci (*Anna/Starr*, *Baez/Coyote*, *Blackbird/Nettie_Moore*, *Blue_Jay/Seastones*, *C485/Sway*, *Jokerman/Sol-20*, *Maggie_Mae/Sun_King*) that are located on the same genomic scaffolds (Appendix A.1), indicating that physical linkage is the likely cause of their LD. The great majority of markers used in this study thus appear to be independent of one another with respect to their allelic composition.

Individual multilocus heterozygosity estimates correlated strongly and significantly between all four metrics (MLH, HL, IR, SH), with the weakest correlation between HL and SH (Pearson product moment correlation coefficient r = -0.952, $P < 0.0001$). Individual heterozygosity values for the complete sample of 593 queens ranged as follows: MLH (0.211, 0.574), HL (0.337, 0.736), IR (-0.233, 0.492), and SH (0.522, 1.33).

Associations between traits

Associations between fitness-related traits are reported in Table 2.2 (parametric statistics) and Appendix A.4 (non-parametric statistics). Many of the traits are significantly associated with one another, as expected if most are at least roughly predictive of incipient colony success. For instance, queens that were heavier on the day of their mating flight tended to be larger in terms of

overall body size and also tended to lose proportionately less weight during the claustral phase, require a shorter claustral period, be more likely to survive the claustral period and the six-week period thereafter, and have larger colonies both at the end of the claustral period and six weeks later compared to lighter queens. Focusing on colony attributes, colonies with short claustral periods tended to be relatively large at the end of this period as well as six weeks later.

Inbreeding, identity disequilibrium (ID), and departure from Hardy-Weinberg Equilibrium (HWE)

The level of inbreeding for each queen was estimated as internal relatedness (IR) from her multilocus genotype. Shown in Figure 2.1 are distributions of IR values for the whole dataset, for the subset of queens infected with SINV-2 virus (the only commonly occurring virus in our samples; Fisher et al. unpubl. ms.), and for individuals in a simulated panmictic population. Distributions for queens in both of the real datasets are shifted toward higher values than those obtained for the simulated individuals (mean IR = 0.038, 0.043, and 0.002, respectively), suggesting the occurrence of some level of inbreeding in our study population. No significant difference in IR was found between SINV-2-infected and uninfected queens within the real dataset (Mann-Whitney *U* test, $W = 127{,}964$, $P = 0.520$), indicating that relatively inbred queens (and/or the colonies from which they originated) are not more susceptible to infection with this pathogen. Closer examination of the IR distributions shows a clear enrichment of queens with high IR values (highly homozygous individuals) in both of the datasets for real queens compared to the simulated data (Fig. 2.2). If we consider the 55 queens in the whole dataset with IR values exceeding the 95th percentile of simulated values (0.165-0.492) to be relatively inbred, these IR values correspond to offspring inbreeding coefficients in haplodiploids that would result from

scenarios ranging from cousin matings ($F = 0.187$) to three generations of sibling mating ($F = 0.5$).

Application of the heterozygosity-heterozygosity correlation (HHC) test revealed the existence of significant identity disequilibrium (ID), as predicted under some level of inbreeding, in both the whole dataset and the subset including only queens infected with SINV-2. Mean HHC was 0.168 (95% confidence interval [CI_{95}] = 0.098, 0.229) and 0.254 (CI_{95} = 0.120, 0.373), respectively, for these datasets. As expected, mean HHC was very small for the simulated dataset (-0.003), and the CI_{95} (-0.030, 0.024) encompassed zero but did not overlap those derived from the datasets for real queens.

We detected a small but significantly positive value for second-order heterozygosity disequilibrium in our whole dataset (g2 = 0.0097, $P < 0.0001$), consistent again with some level of inbreeding in our study population. To ensure that some inability of the program *RMES* to handle our large dataset was not unduly affecting these results (P. David, pers. comm.), we also calculated g2 using smaller panels of markers consisting of 10-90 randomly chosen loci. Variation both in g2 values and their associated *P*-values decreased as marker-panel size increased, with g2 stabilizing at 40 loci at values very close to those obtained for the whole panel (data not shown). A similar g2 value was obtained when only queens infected with SINV-2 were included in an analysis using all 105 loci (g2 = 0.0126, $P < 0.0001$). Values of g2 calculated for all markers were used to derive the expected correlation coefficients between standardized heterozygosity (SH) and the inbreeding coefficient (*F*) (11); this correlation coefficient is r = -0.781 and -0.892, respectively, for the two sets of queens, meaning that our estimators of individual multilocus heterozygosity are likely to be reliable guides to the actual level of inbreeding. The g2 value calculated for the simulated dataset using all loci was several orders of

magnitude smaller than that for the real data and did not differ significantly from zero (g2 = 7.81×10^{-5}, $P = 0.426$); the expected correlation coefficient between SH and F consequently is very small ($r = -0.080$), meaning that, as expected, values of multilocus heterozygosity in such a hypothetical population provide no information on levels of individual inbreeding.

Nineteen of the 105 loci (all microsatellites) displayed genotype frequencies that depart significantly from frequencies expected under HWE (Appendix A.5), several-fold more than the number expected to depart by chance in a panmictic population. Even after correction for multiple comparisons (Benjamini-Hochberg procedure, 5% FDR), nine of these loci remain significantly out of HWE. Estimates of F_{IS} indicate that heterozygote deficiencies are responsible for the departures at all but one of the original 19 loci and at all of the nine loci retained after correction. Importantly, calculation of multilocus F_{IS} also revealed a significant global deficiency of heterozygotes relative to HWE expectation ($F_{IS} = 0.\ 0347$, $CI_{95} = 0.0226, 0.0475$). We note that the presence of null alleles is not likely to have confounded any of our analyses of inbreeding, given that such alleles are unknown for any of the markers in our panel.

Heterozygosity-Fitness Associations (HFAs)

Results of regression models investigating the effects of individual queen multilocus heterozygosity on each fitness trait are shown in Table 2.3. After correction for multiple comparisons, MLH was a significant predictor of only one trait, $QnSurv_{CL}$ (queen survival until first worker eclosion). The regression coefficient in this case is negative, indicating that more highly heterozygous queens were less likely to survive the claustral period. Our correction for multiple comparisons likely is overly conservative because correlated traits were treated as response variables in independent models and because the correction was applied across all of the

essentially redundant models using the various heterozygosity measures (MLH, HL, IR, SH). Thus, we highlight also the two MLH regressions that are significant before correction, which feature *ClaustPer* and $ColGrowth_{CL\text{-}6}$ as response variables (Table 2.3). These regressions suggest that increased heterozygosity is associated with relatively short claustral periods as well as relatively slow post-claustral growth. These same three traits also were identified as potentially being affected by heterozygosity using the three other heterozygosity measures (HL, IR, SH) as predictors (Appendices A.6-A.8), with *ClaustPer* as well as $QnSurv_{CL}$ significant after correction in one or more of these models.

The potential influence of queens with extreme levels of heterozygosity on HFAs was examined by removing from the original dataset queens with the highest or lowest MLH values. Fitness traits implicated as being predicted by MLH in various of these regressions include the three identified for the whole dataset as well as *QnSize* and *QnSINV-1* (Appendix A.9). Although the significance of none of these models is retained after correction for multiple comparisons, the two traits with the strongest evidence for HFAs in the complete dataset ($QnSurv_{CL}$, *ClaustPer*) continue to show some evidence for these even after removal of either highly heterozygous or highly homozygous individuals.

Comparison of trait values between "low-heterozygosity" and "high-heterozygosity" queens defined as having MLH values below or above 0.330, respectively, revealed significant differences between the two groups in $ColGrowth_{CL\text{-}6}$ (Appendix A.10), one of the three response variables implicated as potentially being influenced by heterozygosity in the general regression models. When the low-heterozygosity group was defined as including queens in the lowest quartile of the MLH distribution and the high-heterozygosity group as including queens in the highest quartile, the response variable most strongly implicated by regression analysis, $QnSurv_{CL}$,

differed significantly between the groups, but $ColGrowth_{CL\text{-}6}$ did not (Appendix A.11). Both of these differences were in the direction expected based on the regression results, although neither remained statistically significant after correction for multiple comparisons.

We re-ran all regression analyses separately on individuals that were successfully scored for at least 75, 80, 85, or 95 loci, as well as on all samples using just the 75 loci that were successfully scored in at least 490 queens. The regression results from these smaller subsets were similar to those from the whole dataset (results not shown). In the one instance where the sign of the regression coefficient changed for one of the three implicated fitness-associated traits, it was in the dataset with the smallest sample size (containing only individuals scored at 95 or more loci) and the regression coefficient was small and not significantly different from zero.

We also re-ran regression analyses on data subsets including only queens infected with the SINV-2 virus. The results largely recapitulated the HFA patterns from the complete dataset, with one exception: MLH was marginally positively associated with $ColGrowth_{CL\text{-}6}$ (regression coefficient $b = 32.38$, $P = 0.0741$), meaning that colonies founded by infected queens tended to grow relatively slowly in the post-claustral phase if the queen had low heterozygosity, the opposite of the pattern weakly supported for the whole dataset. The expectation that queens uninfected with SINV-2 exhibited the pattern found for the whole data set is confirmed; MLH was significantly negatively associated with $ColGrowth_{CL\text{-}6}$ for this class (regression coefficient $b = -12.23$, $P = 0.0289$), meaning that colonies founded by uninfected queens tended to grow relatively rapidly in the post-claustral phase if the queen had low heterozygosity. Thus, queens of different infection status exhibit contrasting HFAs for post-claustral colony growth, presumably illustrating an unusual effect of stress on the manifestation of HFAs.

Finally, we re-ran all regression analyses after excluding all eight loci implicated in SLAs (see below). All previously significant models for the complete dataset were retained as such with the exception of the one featuring $ColGrowth_{CL\text{-}6}$ as the response variable; the new model for this trait produced a marginally-significant (uncorrected) *P*-value of 0.0587. Consideration of just SINV-2-infected or uninfected queens in such analyses led to essentially similar (but often weaker) associations as obtained when all loci were included. Most important, MLH was retained as a significant negative predictor of $ColGrowth_{CL\text{-}6}$ in uninfected queens and a marginally significant positive predictor of $ColGrowth_{CL\text{-}6}$ in infected queens (uncorrected $P = 0.0715$).

Single-Locus Associations with Traits (SLAs)

Significant single-locus heterozygosity effects were detected for four fitness traits using the *F*-ratio test after correction for multiple comparisons (*QnSINV-1,* $ColGrowth_{CL\text{-}6}$, *ClaustPer*, $QnSurv_{CL}$). While the *F*-ratio test does not identify the specific locus (or loci) responsible for genotype-by-trait associations, the latter two implicated traits also were found to be affected by specific loci in the *GEPHAST* analyses, although in neither case do the loci involved exhibit overdominance (see below). The *F*-ratio test generally is regarded as conservative in detecting SLAs (11); indeed, to our knowledge this is the first report of evidence for SLAs based on this approach.

We detected eight significant associations between genotypes at specific marker loci and fitness trait values using *GEPHAST* after correction for multiple comparisons: the loci *Beatles* and *Acoh-1* with the trait *QnSize*, locus *St_Stephen* with trait $QnWtLs_{CL}$, loci *Imagine* and *Workingman_Blues* with trait *ClaustPer*, locus *Baez* with trait $QnSurv_{CL}$, and loci *Seastones* and *St_Stephen* with trait $QnSurv_{CL\text{-}6}$. Further analysis revealed the following specific genotypic effects (Appendix A.12): i) queens bearing the common homozygous *Beatles* genotype were

larger than queens of the alternate, heterozygous genotype (Mann-Whitney *U* test, $P = 0.0034$); ii) queens bearing at least one *Acoh-1*^{82} allele were larger than queens homozygous for the alternate allele segregating at that locus (Mann-Whitney *U* test, $P = 0.0457$), iii) queens of the only homozygous class at the locus *St_Stephen* lost proportionately less weight during the claustral period than did heterozygous queens, a marginally non-significant trend (Mann-Whitney *U* test, $P = 0.0697$); iv) queens homozygous at *Imagine* founded colonies with shorter claustral periods than did queens with the alternate, heterozygous genotype (Fisher's exact test, $P = 0.0107$); v) queens bearing at least one *Workingman_Blues*^{256} allele founded colonies with shorter claustral periods than did queens lacking the allele (Fisher's exact test, $P < 0.0001$); and vi) queens bearing at least one *Baez*^{157} allele had a higher probability of surviving the claustral period than did queens homozygous for the alternate allele segregating at that locus (Fisher's exact test, $P < 0.0001$). No single genotypes or groups of genotypes pooled on the basis of shared alleles could be identified as being significantly associated with $QnSurv_{CL\text{-}6}$ for either locus identified by *GEPHAST*.

DISCUSSION

Studies that relate individual multi-locus heterozygosity to fitness are important for elucidating both the basic evolutionary mechanisms of adaptation and the forces that maintain population genetic variation, and such studies have a long history and growing presence in the literature (e.g., 1–3, 5, 7, 15). Overall, there is immense variation among such studies in terms of whether heterozygosity-fitness associations (HFAs) are detected and, if so, what the direction and magnitude of the effects are. This variation has fostered the development of a number of competing theories to explain HFAs and their role in the maintenance of genetic diversity (1, 11,

74, 75). Moreover, variation in study outcomes has generated increasing criticism of the field and its empirical methods, including claims of bias arising from the use of too few (or, perhaps as in the case of allozymes, unrepresentative) markers to track genome-wide heterozygosity, small sample sizes, inadequate knowledge of the demographics and/or breeding biology of study populations, and inadequate knowledge of traits that ostensibly are closely tied to fitness (8, 11, 76, 77). In the present study, we attempted to overcome some of these difficulties by employing a large panel of diverse markers, large sample sizes, and modern analytical approaches to examine the potential effects of individual heterozygosity on fitness-related traits in *Solenopsis invicta*, a species with a well-known natural history. The work additionally contributes to the HFA literature because invasive or eusocial insects seldom have been the subjects of such studies.

Fitness traits in colony-founding *S. invicta* queens

Our monitoring of colony-founding fire ant queens and their incipient colonies revealed strong associations between many of the traits assumed to be important in early colony survival and growth. For example, heavier queens tended to be relatively larger, lose proportionally less weight during the claustral phase, require a shorter claustral period before their first daughter workers eclosed, be more likely to survive the claustral period, and produce larger colonies compared to lighter queens. These results support the hypothesis that this suite of traits contributes in a cohesive manner to enhanced early growth and survival of colonies, the mechanisms of which seem clear. Specifically, traits that favor queens successfully rearing adult workers before starving and that allow young colonies to become sufficiently populous to rapidly secure exclusive foraging territories and construct a nest of sufficient depth to avoid winter freezing are paramount. Although various studies have suggested the likely importance of several

of these traits for early colony fitness (25, 40, 41, 78), our work provides comprehensive empirical evidence consistent with their adaptive importance.

Inbreeding and identity disequilibrium (ID)

Our analyses of the genotypes of many hundred queens at more than 100 loci revealed the unexpected presence of significant inbreeding in a small portion of our study population. Several lines of evidence support this conclusion. The first is the distribution of internal relatedness (IR) values for queens, which is enriched for high values (i.e., more homozygous individuals) compared to the distribution of values from a simulated population in Hardy-Weinberg equilibrium (HWE). Also, both the heterozygosity-heterozygosity correlation (HHC) test and calculation of the g2 metric suggested the existence of significant ID. Importantly, this ID occurs in conjunction with minimal linkage disequilibrium (LD), the combination of which constitutes a hallmark signature of inbreeding (11). Finally, we observed an excess of homozygotes both at a substantial number of individual loci and globally across our marker panel compared to HWE expectations.

The occurrence of inbreeding, at any level, in a monogyne (single queen per colony) population of *S. invicta* is surprising for two reasons: first, knowledge of the breeding biology of this social form of *S. invicta* suggests that consanguineous matings seldom are possible, and second, previous population genetic studies of invasive populations of the ant failed to reveal any evidence for inbreeding. Fire ant mating flights, triggered by locally heavy rains throughout warm periods of the year, often occur simultaneously over very large regions, and sexuals mate on the wing, often at altitudes above 100 m (25, 31, 79); these facts, combined with the enormous density of colonies throughout the invasive range (80), appear to make it highly unlikely that related queens and males would ever encounter one another during a mating flight. The

credibility of this scenario suggesting minimal opportunities for consanguineous matings has been bolstered by the general concordance of nuclear genotype distributions to HWE expectations (including low and insignificant F_{IS} values) reported previously for monogyne populations throughout the native and U.S. ranges (e.g., Ross et al. 1997, 1999; Shoemaker et al. 2006). The most recent and complete of such studies used a 16-marker dataset (microsatellites and allozymes) obtained from over 500 monogyne colonies located throughout the southern U.S.A. (39). We re-analyzed those data to calculate global F_{IS} in the manner described herein to directly compare values between the two studies. The point estimate for the Shoemaker et al. (2006) study is very similar to that for the current study (F_{IS} = 0.0379 and 0.0347, respectively); however, the former estimate does not differ significantly from zero whereas the latter does (CI_{95} = -0.0172, 0.0899 and 0.0226, 0.0475, respectively). Given the very large sample sizes in both studies, this difference in precision of the estimates is almost certainly due to the much larger size of the marker panel in the present study. This result parallels similar findings of the importance of using large numbers of markers to obtain robust estimates of various population genetic parameters relevant to within- and between-population diversity (e.g., 82, 83) and genome-wide heterozygosity (9).

Finally, our conclusion of a low level of inbreeding in monogyne *S. invicta* raises the question of how any matings between relatives could occur in this social form, given the evident strictures of its mating system. One possibility is that some lineages are pre-disposed to disperse only very short distances during mating flights, such that local clusters of related colonies would emerge, increasing the probability of mating between related queens and males once these took flight. However, the expected signature of local (microgeographic) genetic structure expected under such a scenario has not been detected in native or invasive monogyne populations (38, 39,

81). In a variant of this scenario, queens might disperse from a localized aerial aggregation after mating, behavior unlikely to result in microgeographic genetic structure. A final, and perhaps most likely, possibility is that queens and males from the same nest occasionally mate on the surface or within their nest prior to the queens departing on flights; such "intranidal" mating behavior is considered to be possible, and perhaps even common, in the polygyne (multiple queens per colony) social form of this species (71).

Heterozygosity-fitness associations (HFAs)

Analyses of potential heterozygosity-fitness associations revealed that these are uncommon in our study population, with queen multilocus heterozygosity a significant predictor of just a small subset of the fitness traits that we measured: survival of colony-founding queens until eclosion of their first workers ($QnSurv_{CL}$) is strongly and consistently implicated, while the length of the claustral period (*ClaustPer*) is implicated as well in some analyses. A third trait, proportional colony growth for the six-week period immediately following the claustral period ($ColGrowth_{CL\text{-}6}$), was weakly supported as being influenced by heterozygosity in the complete dataset. For the first two HFAs, increased heterozygosity corresponds with a decreased probability of queen survival and a decreased duration of the claustral period. For the third, the situation is complicated by viral infection status of queens — higher heterozygosity weakly corresponds with faster post-claustral growth in queens infected with the SINV-2 virus, but it corresponds with slower post-claustral growth in uninfected queens.

Assuming that shorter claustral periods and faster post-claustral growth reflect higher fitness (higher probability of survival) for a queen and her colony, we detected evidence for both positive (*ClaustPer*, $ColGrowth_{CL\text{-}6}$ in infected queens) and negative ($QnSurv_{CL}$, $ColGrowth_{CL\text{-}6}$ in uninfected queens) HFAs in our study population. That is, the former associations feature

fitness trait values that tend to increase with higher heterozygosity, whereas the latter feature values that decrease with higher heterozygosity. Several points can be made regarding this diversity of positive and negative HFAs. Perhaps most important, the paucity of positive HFAs seems surprising in the context of a study population that evidently features a subset of individuals that are inbred. Indeed, conditions in invasive *S. invicta* arguably are ideal for detection of positive HFAs driven by general effects manifested as inbreeding depression, especially when studied using many samples, a large marker panel, and a diverse set of fitness-related traits. The evident failure of general effects to generate widespread positive HFAs is made more noteworthy in light of our finding that some evidence for a positive HFA involving *ClaustPer* remains even after the most highly homozygous (inbred) individuals are removed from the analyses; this suggests that, even for this trait, general effects alone are not sufficient to explain the positive HFA.

The occurrence of negative HFAs in our study system is, at first glance, equally puzzling. Negative HFAs usually are attributed either to single-locus association(s) (SLAs) (e.g., 5) or to outbreeding depression (e.g., 17). Although we detected evidence for several SLAs in our system, many involving the same traits implicated in HFAs (see below), removal of the SLA loci from the analyses did not change the main HFA results, suggesting that these single-locus effects are not driving the negative HFAs we observed. Outbreeding depression also is unlikely to be an important mechanism generating negative HFAs in our system. Outbreeding depression, which stems from the admixture of divergent conspecific lineages, presumably produces negative HFAs by breaking up co-adapted gene complexes or complementary alleles in highly admixed and consequently highly heterozygous individuals, thereby reducing their fitness (84–86). Monogyne *S. invicta* populations appear to be relatively uniform genetically across the invasive U.S. range,

perhaps as a result of common descent from a single recent founder population (cf. 39), making unlikely the existence of lineages that are divergent with respect to genes underlying fundamental fitness traits. Further mitigating against any role for outbreeding depression is the fact that the two traits exhibiting negative HFAs ($QnSurv_{CL}$ and $ColGrowth_{CL\text{-}6}$ in uninfected queens) continue to show some evidence of these HFAs even after removal of the most highly heterozygous individuals from the analysis.

A third, novel explanation for the occurrence of a negative HFA, in combination with one or more positive HFAs such as we observed, is that these result from life-history or physiological tradeoffs. That is, energetic and/or nutritional constraints placed on organisms by finite resources lead to enhanced expression of one fitness-related trait at the expense of reduced expression of another (87–89). Thus, if higher heterozygosity favors higher-fitness values of a first trait, it necessarily will also be associated with reduced values of a second. In such cases, an ensuing optimal balance in trait values that depends in part on environment likely exists (88). Just such a tradeoff might be expected between the duration of the claustral period and queen survival during this period, where a reduction in the former (e.g., by means of queens provisioning eggs and/or larvae more intensively), would necessarily result in increased mortality risk for the nutrient-limited foundress. The existence of such a tradeoff would be fully consistent with the pattern of a positive HFA for *ClaustPer* and negative HFA for $QnSurv_{CL}$ that we detected. In this more nuanced view of relationships between genomic heterozygosity and fitness, a focus on positive and negative HFAs for single traits that are components of multi-dimensional suites of fitness phenotypes may be overly simplistic; indeed, use of the terms “positive” and “negative” with respect to HFAs seems ill-advised when based, as it often is, on incomplete knowledge and untested assumptions regarding single elements of complex fitness phenotypes.

The evidence in our study population for the joint presence of both a positive and negative HFA for a single trait ($ColGrowth_{CL\text{-}6}$), the sign of which depends on the viral infection status of queens, also is noteworthy. Stress previously has been shown to affect the occurrence and strength of HFAs in various organisms (e.g., 68, 69), and infection with SINV-2 virus appears to constitute a stressor in foundress *S. invicta* queens (Fisher et al., unpubl. ms.). Remarkably, viral infection abolishes or even reverses the significant relationship between queen multilocus heterozygosity and post-claustral growth of the colony she founds such that higher heterozygosity no longer is associated with reduced growth upon infection. Although the mechanism underlying this pattern is not apparent, it may be related to the complex interdependencies of the various traits influencing incipient colony fitness, several of which are dramatically altered by viral infection (Fisher et al., unpubl. ms.). Regardless of the cause, this example illustrates the potential difficulties facing any HFA study in which individuals are pooled for analysis without regard to potential stress factors such as pathogenic challenge, individual age, or environmental conditions. We note that the most important pathogens of monogyne *S. invicta* in the U.S.A. were surveyed in our study (90) and that the latter two categories of stressors were controlled by virtue of i) our collection strategy, which ensured an even-age cohort of queens and ii) a common rearing environment.

Single-Locus associations with traits (SLAs) and linkage disequilibrium (LD)

We identified seven microsatellite loci and one allozyme locus evidently involved in significant SLAs with five fitness-related traits, including all of the traits for which HFAs also were detected. Because the microsatellites are derived from EST or cDNA libraries (i.e., they occur within or near coding regions) and the allozymes comprise protein products of single genes, variation at any of these loci potentially could be directly responsible for the

corresponding variation in the fitness-related traits. Although single genes with major effects on important fitness-related traits generally may be expected to be uncommon (13), several recent examples of such genes have been reported in wild populations (e.g., 70, 91, 92). Alternatively, the SLA loci we detected could be markers that are in strong LD with other coding loci or regulatory elements exerting major effects on the relevant traits, an appealing explanation for an invasive population that has gone through a recent bottleneck that would, at least transiently, generate enhanced genome-wide LD. However, our analyses indicate surprisingly low levels of pairwise LD in the study population, much of which apparently arises from physical linkage, suggesting that "local effect" associations between major-effect and marker loci probably are not manifested largely as a result of high LD.

The minimal linkage disequilibrium we detected in invasive *S. invicta* may seem surprising given that the U.S. founder population experienced a bottleneck featuring no more than a few dozen haploid genomes and that only 10-20 generations have passed since that bottleneck (24). The likely reason for rapid decay of the LD presumed to be present in the early colonizers is the documented rapid population expansion and vast subsequent effective population sizes established by this invasive pest insect (e.g., 93, 94), resulting in an enormous field for genetic recombination. Given the large number and correspondingly small size of scaffolds in the *S. invicta* genome assembly (10,543 scaffolds with an N50 of 720,578bp [95]), those marked by SLA loci represent promising targets for further investigation and, ultimately, identification of candidate fitness-trait genes, an endeavor facilitated by ongoing, improved assembly and annotation of the *S. invicta* genome (96).

Finally, we note that our discovery of several SLAs is not to be viewed as evidence for the local and/or direct effect hypotheses for HFAs for two reasons. First, we found no evidence

for overdominance at any locus involved in an SLA, although directional selection on particular alleles, as may occur at some of our SLA loci, potentially can generate transient HFAs (20, 21). Second, removal of all loci implicated in SLAs did not substantially affect our HFA regression results.

Concluding remarks

The sweep of HFA studies over the past few decades has spanned from relatively simple explanations of influential focal study outcomes that invoked overdominance at single loci (1, 74, 97) to progressively more complex explanations of diverse outcomes that feature the general effects of inbreeding depression, sometimes in combination with direct or local effects of single loci (5, 11, 19, 75). As the complicated dynamics of HFAs, and the network of evolutionary forces potentially at play, have become increasingly apparent, empirical studies face the challenge of ever more sophisticated study designs necessary to address specific issues. Future studies not only must heed the now well-known imperative to deploy large sample sizes and marker panels (8, 77) but also should consider the effects of stress, life-history tradeoffs, SLAs, and variation in life-cycle stages assayed (e.g., 69) in producing novel study outcomes such as the presence of both "positive" and "negative" HFAs (5, 17, 19). In our study, the apparent effects of stress and tradeoffs, in particular, were only possible to detect because diverse, readily measured fitness traits were investigated, made possible by detailed prior understanding of the study organism's natural history. Based on our results, we advocate exclusion of SLA-implicated loci in subsets of HFA analyses, or at the very least, separate treatment of loci known to have some functional role (e.g., 5, 19, 20), in order to help clarify single-gene effects in generating HFAs. We also endorse investigation of different life-cycle stages because of the possibility that heterozygosity changes due to selection at an early stage affect the form of HFAs at later stages;

as an example, the negative HFA for queen survival during the claustral period in our study evidently truncates the range of heterozygosity available for expression of HFAs in the post-claustral phase. Finally, comparative HFA studies that take advantage of different population structures or genome architectures may yield important data for formulating a more comprehensive theory of the role of HFAs in maintaining individual and population genetic diversity (8, 98–100). In our case, this can be accomplished by investigating HFAs in native South American and recently introduced Pacific Rim *S. invicta* populations; because of the very different population histories (55, 101), each may be unique in terms of such features as additive genetic variance and linkage disequilibrium, which may facilitate formulation of specific predictions about the extent and patterns of HFA. (8)

ACKNOWLEDGMENTS

We thank Eileen Carroll and Stephanie Tiv for help in the laboratory, Jaxk Reeves and Kim Love-Meyers for statistical consultation, and Jo Chapman, Patrice David, Bill Amos, and Juanita Olano-Marin for their timely replies to technical HFA questions. This work was supported by grants from the U.S. National Science Foundation to KGR and from the Agriculture and Food Research Initiative of the U.S. Department of Agriculture to DDS.

REFERENCES

1. **Mitton JB, Grant MC**. 1984. Associations among protein heterozygosity, growth rate, and developmental homeostasis. Annual Review of Ecology and Systematics **15**:479–499.

2. **Pogson GH**. 1991. Expression of overdominance for specific activity at the Phosphoglucomutase-2 locus in the pacific oyster, *Crassostrea gigas*. Genetics **128**:133–141.

3. **Pogson GH, Zouros E**. 1994. Allozyme and RFLP heterozygosities as correlates of growth rate in the scallop *Placopecten magellanicus*: A test of the associative overdominance hypothesis. Genetics **137**:221–231.

4. **Hedgecock D, McGoldrick DJ, Manahan DT, Vavra J, Appelmans N, Bayne BL**. 1996. Quantitative and molecular genetic analyses of heterosis in bivalve molluscs. Journal of Experimental Marine Biology and Ecology **203**:49–59.

5. **Olano-Marin J, Mueller JC, Kempenaers B**. 2011. Heterozygosity and survival in blue tits (*Cyanistes caeruleus*): contrasting effects of presumably functional and neutral loci. Molecular Ecology **20**:4028–4041.

6. **Forstmeier W, Schielzeth H, Mueller JC, Ellegren H, Kempenaers B**. 2012. Heterozygosity-fitness correlations in zebra finches: microsatellite markers can be better than their reputation. Molecular Ecology **21**:3237–3249.

7. **Ruiz-López MJ, Gañan N, Godoy JA, Del Olmo A, Garde J, Espeso G, Vargas A, Martinez F, Roldán ERS, Gomendio M**. 2012. Heterozygosity-fitness correlations and inbreeding depression in two critically endangered mammals. Conservation Biology **26**:1121–1129.

8. **Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC**. 2009. A quantitative review of heterozygosity-fitness correlations in animal populations. Molecular Ecology **18**:2746–2765.

9. **Balloux F, Amos W, Coulson T**. 2004. Does heterozygosity estimate inbreeding in real populations? Molecular Ecology **13**:3021–3031.

10. **Slate J, David P, Dodds KG, Veenvliet BA, Glass BC, Broad TE, McEwan JC**. 2004. Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. Heredity **93**:255–265.

11. **Szulkin M, Bierne N, David P**. 2010. Heterozygosity-fitness correlations: a time for reappraisal. Evolution **64**:1–16.

12. **Houle D, Morikawa B, Lynch M**. 1996. Comparing mutational variabilities. Genetics **143**:1467–1483.

13. **Orr HA, Coyne JA**. 1992. The genetics of adaptation: a reassessment. The American Naturalist **140**:725–742.

14. **David P**. 1997. Modeling the genetic basis of heterosis: tests of alternative hypotheses. Evolution **51**:1049–1057.

15. **Chapman JR, Sheldon BC**. 2011. Heterozygosity is unrelated to adult fitness measures in a large, noninbred population of great tits (*Parus major*). Journal of Evolutionary Biology **24**:1715–1726.

16. **Coulson TN, Pemberton JM, Albon SD, Beaumont M, Marshall TC, Slate J, Guinness FE, Clutton-Brock TH**. 1998. Microsatellites reveal heterosis in red deer. Proceedings of the Royal Society of London Series B-Biological Sciences **265**:489–495.

17. **Marshall TC, Spalton JA**. 2000. Simultaneous inbreeding and outbreeding depression in reintroduced Arabian oryx. Animal Conservation **3**:241–248.

18. **Jourdan-Pineau H, Folly J, Crochet P-A, David P**. 2012. Testing the influence of family structure and outbreeding depression on heterozygosity-fitness correlations in small populations. Evolution **66**:3624–3631.

19. **Laine VN, Herczeg G, Shikano T, Primmer CR**. 2012. Heterozygosity-behaviour correlations in nine-spined stickleback (*Pungitius pungitius*) populations: contrasting effects at random and functional loci. Molecular Ecology **21**:4872–4884.

20. **Mueller JC, Hermisson J, Olano-Marin J, Hansson B, Kempenaers B**. 2010. Linking genetic mechanisms of heterozygosity-fitness correlations to footprints of selection at single loci. Evolutionary Ecology **25**:1–11.

21. **Deng H-W, Fu Y-X**. 1998. Conditions for positive and negative correlations between fitness and heterozygosity in equilibrium populations. Genetics **148**:1333–40.

22. **Caldera EJ, Ross KG, DeHeer CJ, Shoemaker DD**. 2008. Putative native source of the invasive fire ant *Solenopsis invicta* in the USA. Biological Invasions **10**:1457–1479.

23. **IUCN**. 2000. One-hundred of the world's worst invasive alien species: a selection from the Global Invasive Species Database. http://www.iucn.org/places/medoffice/invasive_species/docs/invasive_species_booklet.pdf .

24. **Ross KG, Shoemaker DD**. 2008. Estimation of the number of founders of an invasive pest insect population: the fire ant *Solenopsis invicta* in the USA. Proceedings of the Royal Society of London Series B-Biological Sciences **275**:2231–2240.

25. **Tschinkel WR**. 2006. The Fire Ants, 1st ed. The Belknap Press of Harvard University Press, Cambridge, MA.

26. **Bryant EH, McCommas SA, Combs LM**. 1986. The effect of an experimental bottleneck upon quantitative genetic variation in the housefly. Genetics **114**:1191–1211.

27. **Goodnight CJ**. 1988. Epistasis and the effect of founder events on the additive genetic variance. Evolution **42**:441–454.

28. **Hoffman JI, Hanson N, Forcada J, Trathan PN, Amos W**. 2010. Getting long in the tooth: a strong positive correlation between canine size and heterozygosity in Antarctic fur seals *Arctocephalus gazella*. The Journal of Heredity **101**:527–538.

29. **Slate J, Pemberton JM**. 2002. Comparing molecular measures for detecting inbreeding depression. Journal of Evolutionary Biology **15**:20–31.

30. **Morrill W**. 1974. Production and flight of alate Red Imported Fire Ants. Environmental Entomology **3**:265–271.

31. **Markin, Markin GP, Dillier JH, Hill SO, Blum MS, Hermann HR**. 1971. Nuptial flight and flight ranges of the imported fire ant, *Solenopsis invicta*. Journal of the Georgia Entomological Society **6**:145–156.

32. **Werren JH**. 1993. The evolution of inbreeding in haplodiploid organisms, p. 42–59. *In* Thornhill, NW (ed.), The Natural History of Inbreeding and Outbreeding, 1st ed. The University of Chicago Press, Chicago.

33. **Ometto L, Shoemaker D, Ross KG, Keller L**. 2011. Evolution of gene expression in fire ants: the effects of developmental stage, caste, and species. Molecular Biology and Evolution **28**:1381–1392.

34. **Vitikainen E, Haag-Liautard C, Sundström L**. 2011. Inbreeding and reproductive investment in the ant *Formica exsecta*. Evolution **65**:2026–2037.

35. **Kureck IM, Jongepier E, Nicolai B, Foitzik S**. 2012. No inbreeding depression but increased sexual investment in highly inbred ant colonies. Molecular Ecology **21**:5613–5623.

36. **Wetzel DP, Stewart IRK, Westneat DF**. 2012. Heterozygosity predicts clutch and egg size but not plasticity in a house sparrow population with no evidence of inbreeding. Molecular Ecology **21**:406–420.

37. **Slate J, Pemberton J**. 2006. Does reduced heterozygosity depress sperm quality in wild rabbits (*Oryctolagus cuniculus*)? Current Biology **16**:790–791.

38. **Ross KG, Shoemaker DD, Krieger MJ, DeHeer CJ, Keller L**. 1999. Assessing genetic structure with multiple classes of molecular markers: a case study involving the introduced fire ant *Solenopsis invicta*. Molecular Biology and Evolution **16**:525–543.

39. **Shoemaker DD, Deheer CJ, Krieger MJB, Ross KG**. 2006. Population genetics of the invasive fire ant *Solenopsis invicta* (Hymenoptera : Formicidae) in the United States. Annals of the Entomological Society of America **99**:1213–1233.

40. **Ross KG, Fletcher DJC**. 1986. Diploid male production-a significant colony mortality factor in the fire ant *Solenopsis invicta* (Hymenoptera:Formicidae). Behavioral Ecology and Sociobiology **19**:283–291.

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

42. **Wilson EO**. 1971. The Insect Societies. Harvard University Press, Cambridge, MA.

43. **Markin G, Dillier JH, Collins H**. 1973. Growth and development of colonies of the red imported fire ant, *Solenopsis invicta*. Annals of the Entomological Society of America **66**:803–808.

44. **Tschinkel WR, Howard DF**. 1983. Colony founding by pleometrosis in the fire ant, *Solenopsis invicta*. Behavioral Ecology and Sociobiology **12**:103–113.

45. **Oster GF, Wilson EO**. 1978. Caste and ecology in the social insects. Monographs in Population Biology no. 12. Princeton University Press, Princeton, NJ.

46. **Brian M**. 1983. Social insects: ecology and behavioural biology. Chapman and Hall, New York, NY.

47. **R Development Core Team**. 2011. R: A Language and Environment for Statistical Computing. U.S. Patent 2.13.2. R Foundation for Statistical Computing, Vienna, Austria.

48. **Bernasconi G, Keller L**. 1999. Effect of queen phenotype and social environment on early queen mortality in incipient colonies of the fire ant, *Solenopsis invicta*. Animal Behaviour **57**:371–377.

49. **Wiernasz DC, Cole BJ**. 2003. Queen size mediates queen survival and colony fitness in harvester ants. Evolution **57**:2179–2183.

50. **Tschinkel WR**. 1993. Resource-allocation, brood production and cannibalism during colony founding in the fire ant, *Solenopsis invicta*. Behavioral Ecology and Sociobiology **33**:209–223.

51. **Valles SM, Strong CA, Dang PM, Hunter WB, Pereira RM, Oi DH, Shapiro AM, Williams DF**. 2004. A picorna-like virus from the red imported fire ant, *Solenopsis invicta*: initial discovery, genome sequence, and characterization. Virology **328**:151–157.

52. **Hashimoto Y, Valles SM**. 2008. Infection characteristics of Solenopsis invicta virus 2 in the red imported fire ant, *Solenopsis invicta*. Journal of Invertebrate Pathology **99**:136–140.

53. **Valles SM, Hashimoto Y**. 2009. Isolation and characterization of Solenopsis invicta virus 3, a new positive-strand RNA virus infecting the red imported fire ant, *Solenopsis invicta*. Virology **388**:354–361.

54. **Shoemaker DD, Costa J, Ross KG**. 1992. Estimates of heterozygosity in two social insects using a large number of electrophoretic markers. Heredity **69**:573–582.

55. **Ascunce MS, Yang C, Oakey J, Calcaterra L, Wu W, Shih C, Goudet J, Ross KG, Shoemaker D**. 2011. Global invasion history of the fire ant. Science **331**:1066–1068.

56. **Wang J, Jemielity S, Uva P, Wurm Y**. 2007. An annotated cDNA library and microarray for large-scale gene-expression studies in the ant Solenopsis invicta. Genome biology **8**.

57. **Rozen S, Skaletsky HJ**. 2000. Primer3 on the WWW for general users and for biologist programmers. Bioinformatics Methods and Protocols: Methods in Molecular Biology 365–386.

58. **Ascunce MS, Bouwma AM, Shoemaker D**. 2009. Characterization of 24 microsatellite markers in 11 species of fire ants in the genus *Solenopsis* (Hymenoptera: Formicidae). Molecular Ecology Resources **9**:1475–1479.

59. **Deheer CJ, Goodisman MAD, Ross KG**. 1999. Queen dispersal strategies in the multiple-queen form of the fire ant *Solenopsis invicta*. American Naturalist **153**:660–675.

60. **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.

61. **Aparicio JM, Ortego J, Cordero PJ**. 2006. What should we weigh to estimate heterozygosity, alleles or loci? Molecular Ecology **15**:4659–4665.

62. **Amos W, Wilmer JW, Fullard K, Burg TM, Croxall JP, Bloch D, Coulson T**. 2001. The influence of parental relatedness on reproductive success. Proceedings of the Royal Society of London Series B-Biological Sciences **268**:2021–2027.

63. **Coltman DW, Pilkington JG, Smith JA, Pemberton JM**. 1999. Parasite-mediated selection against inbred soay sheep in a free-living, island population. Evolution **53**:1259–1267.

64. **Alho JS, Välimäki K, Merilä J, Valimaki K, Merila J**. 2010. Rhh: an R extension for estimating multilocus heterozygosity and heterozygosity-heterozygosity correlation. Molecular Ecology Resources **10**:720–722.

65. **David P, Pujol B, Viard F, Castella V, Goudet J**. 2007. Reliable selfing rate estimates from imperfect population genetic data. Molecular Ecology **16**:2474–2487.

66. **Weir BSS, Cockerham CC**. 1984. Estimating F-statistics for population structure. Evolution **38**:1358–1370.

67. **Lewis PO, Zaykin D**. 2001. Genetic Data Analysis: computer program for the analysis of allelic data, Version 1.0. Available at http://lewis.eeb.uconn.edu/lewishome/software.html.

68. **Lesbarrères D, Primmer CR, Laurila A, Merilä J**. 2005. Environmental and population dependency of genetic variability-fitness correlations in *Rana temporaria*. Molecular Ecology **14**:311–323.

69. **Pujolar JM, Maes GE, Vancoillie C, Volckaert FAM**. 2006. Environmental stress and life-stage dependence on the detection of heterozygosity-fitness correlations in the European eel, *Anguilla anguilla*. Genome **49**:1428–1437.

70. **Amos W, Acevedo-Whitehouse K**. 2009. A new test for genotype-fitness associations reveals a single microsatellite allele that strongly predicts the nature of tuberculosis infections in wild boar. Molecular Ecology Resources **9**:1102–1111.

71. **Ross KG, Keller L**. 1995. Ecology and evolution of social organization: insights from fire ants and other highly eusocial insects. Annual Review of Ecology and Systematics **26**:631–656.

72. **Krieger MJB, Ross KG, Chang CWY, Keller L**. 1999. Frequency and origin of triploidy in the fire ant *Solenopsis invicta*. Heredity **82**:142–150.

73. **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.

74. **Smouse PE**. 1986. The fitness consequences of multiple-locus heterozygosity under the multiplicative overdominance and inbreeding depression models. Evolution **40**:946–957.

75. **David P**. 1998. Heterozygosity-fitness correlations: new perspectives on old problems. Heredity **80**:531–537.

76. **Britten HB**. 1996. Meta-analyses of the association between multilocus heterozygosity and fitness. Evolution **50**:2158–2164.

77. **Coltman DW, Slate J**. 2003. Microsatellite measures of inbreeding: a meta-analysis. Evolution **57**:971–983.

78. **Bernasconi G, Keller L**. 1998. Phenotype and individual investment in cooperative foundress associations of the fire ant, *Solenopsis invicta*. Behavioral Ecology **9**:478–485.

79. **Fritz GN, Fritz AH, Vander Meer RK**. 2011. Sampling high-altitude and stratified mating flights of Red Imported Fire Ant. Entomological Society of America **48**:508–512.

80. **Porter SD, Fowler HG, Mackay WP**. 1992. Fire ant mound densities in the United States and Brazil (Hymenoptera, Formicidae). Journal of Economic Entomology **85**:1155–1161.

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

82. **Nei M, Roychoudhury AK**. 1974. Sampling variances of heterozygosity and genetic distance. Genetics **76**:379–390.

83. **Nybom H**. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Molecular Ecology **13**:1143–1155.

84. **Brncic D**. 1954. Heterosis and the integration of the genotype in geographic populations of *Drosophila pseudoobscura*. Genetics **39**:77–88.

85. **Wallace B**. 1995. Inter-population hybrids in *Drosophila melanogaster*. Evolution **9**:302–316.

86. **Lynch M**. 1991. The genetic interpretation of inbreeding depression and outbreeding depression. Evolution **45**:622–629.

87. **Cole LC**. 1954. The population consequences of life history phenomena. The Quarterly Review of Biology **29**:103–137.

88. **Stearns SC**. 1977. The evolution of life history traits: a critique of the theory and a review of the data. Annual Review of Ecology and Systematics **8**:145–171.

89. **Pianka ER, Parker WS**. 1975. Age-specific reproductive tactics. The American Naturalist **109**:453–464.

90. **Valles SM, Oi DH, Porter SD**. 2010. Seasonal variation and the co-occurrence of four pathogens and a group of parasites among monogyne and polygyne fire ant colonies. Biological Control **54**:342–348.

91. **Queitsch C, Sangster T a, Lindquist S**. 2002. Hsp90 as a capacitor of phenotypic variation. Nature **417**:618–624.

92. **Shapiro MD, Kronenberg Z, Li C, Domyan ET, Pan H, Campbell M, Tan H, Huff CD, Hu H, Vickrey AI, Nielsen SCA, Stringham SA, Hu H, Willerslev E, Gilbert MTP, Yandell M, Zhang G, Wang J**. 2013. Genomic diversity and evolution of the head crest in the rock pigeon. Science **339**:1063–1067.

93. **Porter SD, Williams DF, Patterson RS**. 1997. Intercontinental differences in the abundance of *Solenopsis* fire ants (Hymenoptera: Formicidae): escape from natural enemies? Environmental Entomology **26**:373–384.

94. **Buhs JB**. 2004. The Fire Ant Wars: Nature, Science, and Public Policy in Twentieth-Century America. The University of Chicago Press, Chicago.

95. **Wurm Y, Wang J, Riba-Grognuz O, Corona M, Nygaard S, Hunt BG, Ingram KK, Falquet L, Nipitwattanaphon M, Gotzek D, et al.** 2011. The genome of the fire ant *Solenopsis invicta*. Proceedings of the National Academy of Sciences of the United States of America **108**:5679–5684.

96. **Wurm Y, Uva P, Ricci F, Wang J, Jemielity S, Iseli C, Falquet L, Keller L**. 2009. Fourmidable: a database for ant genomics. BMC Genomics **10**:5.

97. **Zouros E, Foltz DW**. 1987. The use of allelic isozyme variation for the study of heterosis. Isozymes-Current Topics in Biological and Medical Research **13**:1–59.

98. **Grueber CE, Wallis GP, Jamieson IG**. 2008. Heterozygosity-fitness correlations and their relevance to studies on inbreeding depression in threatened species. Molecular Ecology **17**:3978–3984.

99. **Luquet E, Léna J-P, David P, Prunier J, Joly P, Lengagne T, Perrin N, Plénet S**. 2013. Within- and among-population impact of genetic erosion on adult fitness-related traits in the European tree frog *Hyla arborea*. Heredity **110**:347–54.

100. **González-Varo JP, Aparicio A, Lavergne S, Arroyo J, Albaladejo RG**. 2012. Contrasting heterozygosity-fitness correlations between populations of a self-compatible shrub in a fragmented landscape. Genetica **140**:31–38.

101. **Ross K, Krieger M, Keller L, Shoemaker DD**. 2007. Genetic variation and structure in native populations of the fire ant *Solenopsis invicta*: evolutionary and demographic implications. Biological Journal of the Linnean Society **92**:541–560.

Table 2.1. Abbreviations and descriptions for fitness-related traits studied.

Abbreviation	Trait description
QnSize	Queen size: PC1 from a principal components analysis of queen antennal scape length, metathoracic femur length, and head width.
QnWt	Queen weight on the day of her mating flight.
$QnWtLs_{CL}$	Queen weight loss during the claustral period: proportional weight loss of queen between the day of her mating flight and eclosion of her first worker.
QnSINV-1	Queen infection with SINV-1 virus.
QnSINV-2	Queen infection with SINV-2 virus.
QnSINV-3	Queen infection with SINV-3 virus.
ClaustPer	Claustral period: number of days between a queen's mating flight and eclosion of her first worker.
$QnSurv_{CL}$	Queen survival until the end of the claustral period (eclosion of her first worker).
$ColWt_{CL}$	Colony weight at the end of the claustral period (time of first worker eclosion).
$ColWt_{CL+6}$	Colony weight six weeks after the end of the claustral period (six weeks after first worker eclosion).
$ColGrowth_{CL-6}$	Colony growth between the end of the claustral period and six weeks thereafter: proportional increase in colony weight between the day of first worker eclosion and six weeks thereafter.
$QnSurv_{CL-6}$	Queen survival between the end of the claustral period (eclosion of her first worker) and six weeks thereafter.

Table 2.2. Effect sizes (above diagonal) and *P*-values (below diagonal) for associations between fitness-related traits (excluding viral infection) analyzed using parametric statistical tests. Comparisons between traits that are effectively continuous (*QnSize*, *QnWt*, $QnWtLs_{CL}$, *ClaustPer*, $ColWt_{CL}$, $ColWt_{CL+6}$, and $ColGrowth_{CL-6}$) were made using Pearson product moment correlation, with the correlation coefficients (*r*) as effect sizes. Comparisons between binary traits ($QnSurv_{CL}$ and $QnSurv_{CL-6}$) and continuous traits were made using the *t*-test, with effect sizes obtained by converting the *t*-scores to Pearson *r* values (77). *P*-values ≤ 0.05 are shown in bold, while *P*-values significant after adjustment for multiple comparisons (Benjamini-Hochberg procedure, 5% FDR) are bold and underlined. NA = comparison not applicable. Associations between viral infection status and the other traits are described elsewhere (Fisher et al. unpubl. ms.).

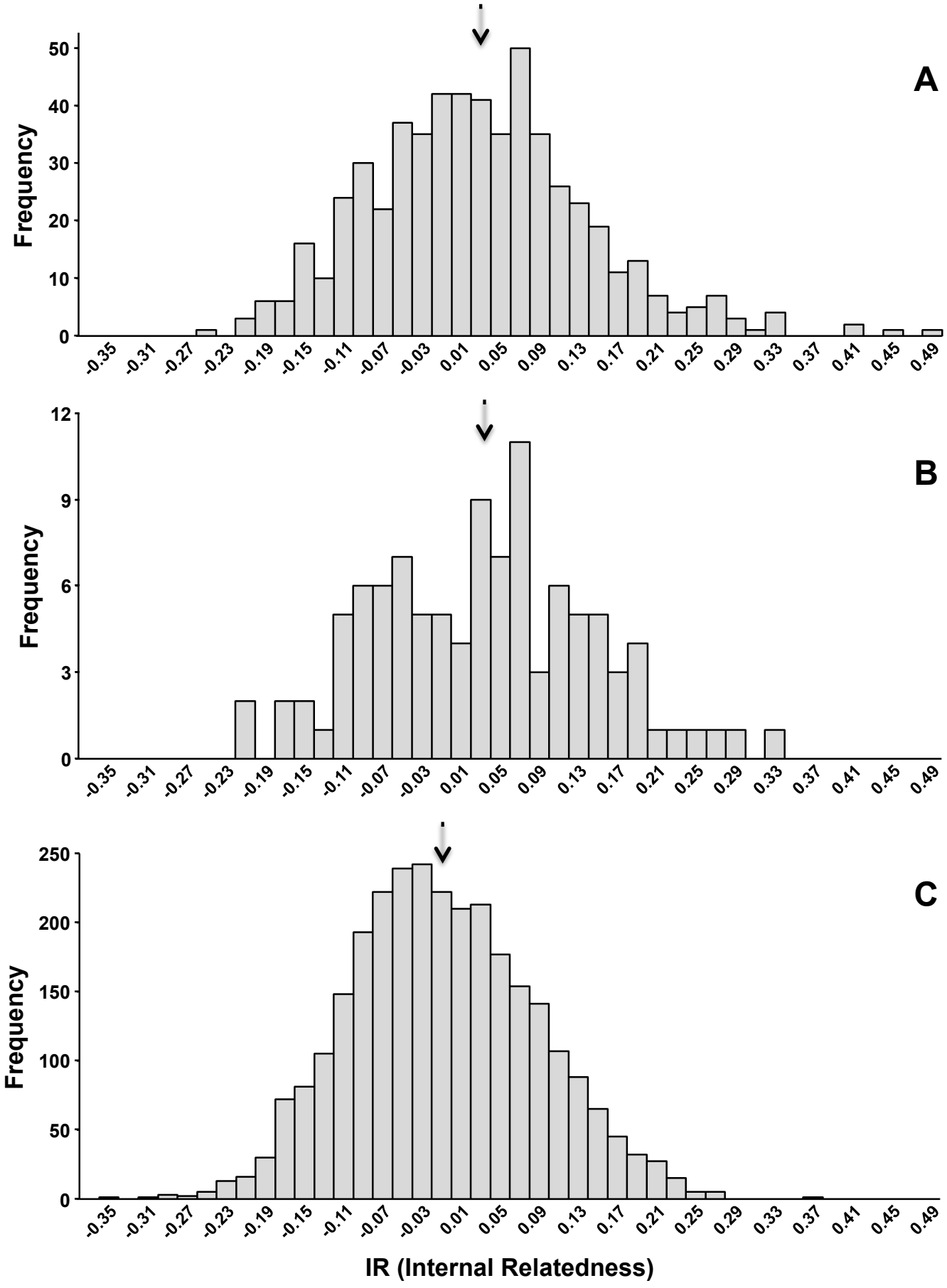
	QnSize	*QnWt*	$QnWtLs_{CL}$	*ClaustPer*	$QnSurv_{CL}$	$ColWt_{CL}$	$ColWt_{CL+6}$	$ColGrowth_{CL-6}$	$QnSurv_{CL-6}$
QnSize	—	0.1310	-0.0028	0.0573	0.2134	-0.0778	-0.0239	-0.0029	-0.0198
QnWt	**0.0081**	—	-0.1491	-0.2010	0.4993	0.3692	0.2255	0.0843	0.5599
$QnWtLs_{CL}$	0.9679	**0.0222**	—	0.2132	NA	-0.7143	-0.1993	0.0215	-0.5361
ClaustPer	0.4229	**0.0023**	**0.0012**	—	NA	-0.5213	-0.3742	-0.2516	-0.7799
$QnSurv_{CL}$	0.1907	**0.0006**	NA	NA	—	NA	NA	NA	NA
$ColWt_{CL}$	0.2689	**<0.0001**	**<0.0001**	**<0.0001**	NA	—	0.4225	0.0533	0.7648
$ColWt_{CL+6}$	0.8456	**0.0457**	0.0783	**0.0007**	NA	**0.0001**	—	0.8793	NA
$ColGrowth_{CL-6}$	0.9810	0.4600	0.8505	**0.0253**	NA	0.6407	**<0.0001**	—	NA
$QnSurv_{CL-6}$	0.9475	**0.0355**	0.1206	**0.0256**	NA	**0.0118**	NA	NA	—

Signs of effect sizes reflect directions of associations; for the binary traits $QnSurv_{CL}$ and $QnSurv_{CL-6}$, positive effect sizes indicate that surviving queens had higher associated trait values than non-survivors.

Table 2.3. Regression models for analysis of HFAs (using MLH metric), with relevant model information reported for each response variable (fitness-related trait). *P*-values ≤ 0.05 are shown in bold, while *P*-values significant after adjustment for multiple comparisons (Benjamini-Hochberg procedure, 5% FDR) are bold and underlined. Effect sizes are reported directly as Pearson product moments (continuous traits) or by converting from the Mann-Whitney *U* test *Z*-score (binary traits) (77).

Response variable/trait	Regression coefficient	Standard error	Error distribution of GLM model	Type of score	Score	*P*-value	Sample size	Effect size	R^2
QnSize	0.020	0.045	Gaussian	*t*	0.43	0.666	454	0.01	0.0001
QnWt	-0.251	8.080	Gaussian	*t*	-0.03	0.975	591	-0.01	0.0001
$QnWtLs_{CL}$	-0.042	0.095	Gaussian	*t*	-0.44	0.663	235	-0.03	0.0008
QnSINV-1	-4.130	3.156	Binomial	*Z*	-1.31	0.191	562	-0.06	0.0038
QnSINV-2	-1.117	1.895	Binomial	*Z*	-0.59	0.555	562	-0.04	0.0019
QnSINV-3	5.239	7.090	Binomial	*Z*	0.74	0.460	562	0.03	0.0008
ClaustPer	-2.119	0.910	Poisson	*Z*	-2.33	**0.020**	228	-0.12	0.0133
$QnSurv_{CL}$	-11.363	3.693	Binomial	*Z*	-3.08	**<u>0.002</u>**	282	-0.13	0.0169
$ColWt_{CL}$	40.610	31.130	Gaussian	*t*	1.31	0.193	235	0.09	0.0073
$ColWt_{CL+6}$	-141.200	521.500	Gaussian	*t*	-0.27	0.787	79	-0.03	0.0010
$ColGrowth_{CL-6}$	-1.817	0.822	Poisson	*Z*	-2.21	**0.027**	79	-0.20	0.0415
$QnSurv_{CL-6}$	-6.478	6.306	Binomial	*Z*	-1.03	0.304	92	-0.05	0.0026

Figure 2.1. Distributions of IR values for (A) queens in the whole dataset, (B) queens infected with the SINV-2 virus, and (C) individuals in a simulated dataset. Values on the x-axis indicate lower bounds of alternate bins. Arrows show the mean for each distribution.


A
Frequency
50
40
30
20
10
0
-0.35 -0.31 -0.27 -0.23 -0.19 -0.15 -0.11 -0.07 -0.03 0.01 0.05 0.09 0.13 0.17 0.21 0.25 0.29 0.33 0.37 0.41 0.45 0.49
B
Frequency
12
9
6
3
0
-0.35 -0.31 -0.27 -0.23 -0.19 -0.15 -0.11 -0.07 -0.03 0.01 0.05 0.09 0.13 0.17 0.21 0.25 0.29 0.33 0.37 0.41 0.45 0.49
C
Frequency
250
200
150
100
50
0
-0.35 -0.31 -0.27 -0.23 -0.19 -0.15 -0.11 -0.07 -0.03 0.01 0.05 0.09 0.13 0.17 0.21 0.25 0.29 0.33 0.37 0.41 0.45 0.49
IR (Internal Relatedness)

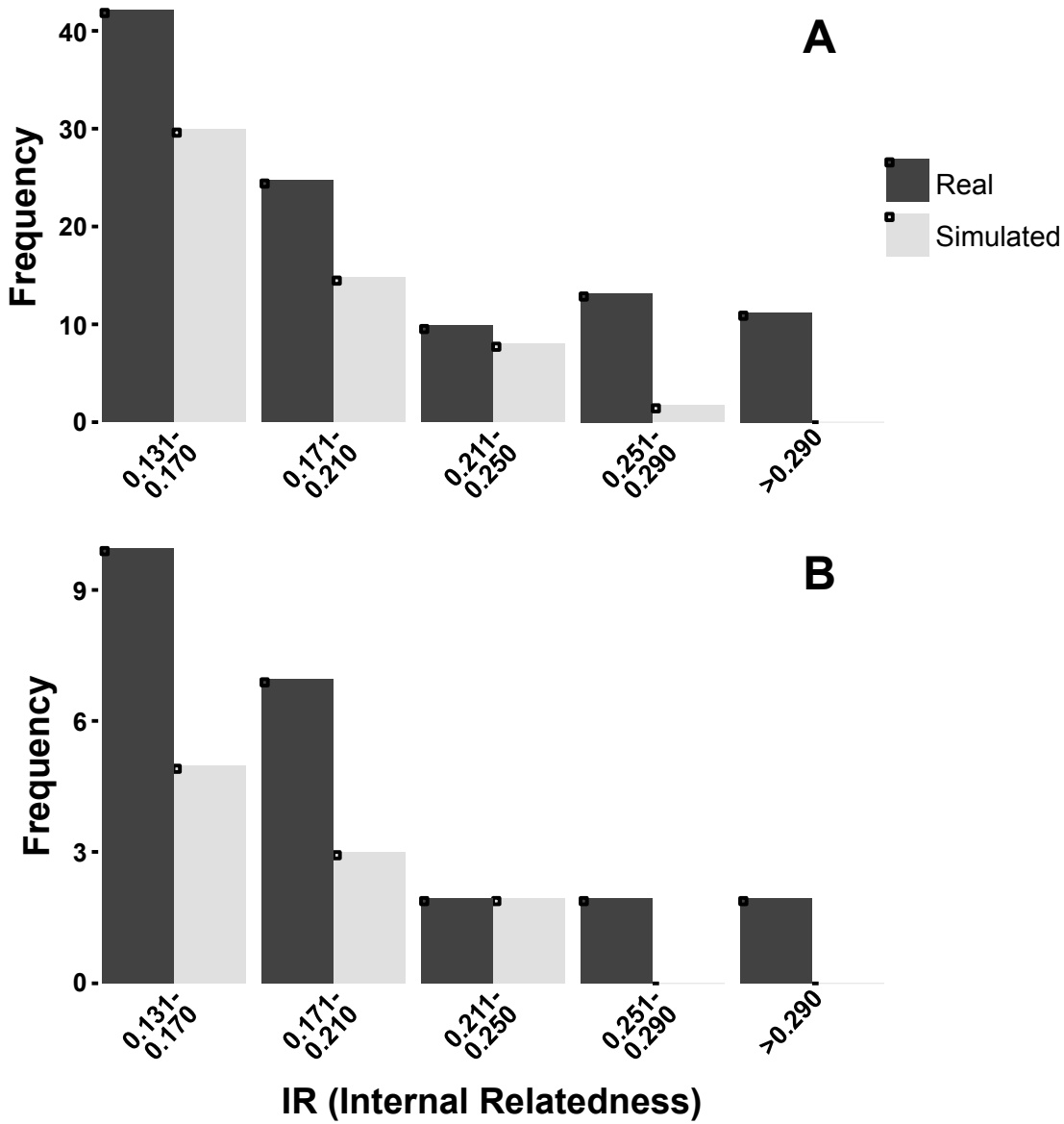


Figure 2.2. Distributions of IR values in high-homozygosity bins for (A) queens in the whole dataset and (B) queens infected with the SINV-2 virus, compared with individuals in a simulated dataset. Counts of simulated individuals are scaled according to the number of queens in the group with which they are compared.

CHAPTER 3

HEFPIPE: A COMPLETE ANALYTICAL PIPELINE FOR HETEROZYGOSITY-FITNESS CORRELATION STUDIES[1]

[1] Fisher, Mark A. Submitted to *Molecular Ecology Resources*, 5/29/2013

ABSTRACT

As the body of heterozygosity-fitness correlation (HFC) research grows, more and increasingly-complicated tests have become an integral part of a typical HFC analysis. Currently, no software is available to undertake conversion between the file formats required to conduct all of these tests and to conduct the main regression analyses at the core of all HFCs. *HeFPipe* is a script written in *Python* that accomplishes both of these tasks for studies based on microsatellite data. *HeFPipe* takes input in the form of allele reports from the genotype-calling software, *GeneMarker*, and reconfigures the data into *GENEPOP*, *Rhh*, *RMES*, and *GEPHAST* formats. The script is also equipped to re-format the output from *GENEPOP* on the Web (option 5) and *Rhh* into csv spreadsheets that can be incorporated into downstream analyses. *HeFPipe* accommodates user-provided lists of samples and markers to be included in or excluded from analyses. *HeFPipe* is equipped to create generalized linear models (GLMs) from both the main dataset and subsets of the data. Finally, *HeFPipe* allows users to explore single-marker effects and conduct correlation analyses. The script, a comprehensive manual, a link to a series of video tutorials, and an example dataset are available from GitHub (http://github.com/Atticus29/*HeFPipe*_repos).

RESOURCE ARTICLE: COMPUTER PROGRAM

Identifying associations between heterozygosity and fitness is a lynchpin of studies intending to clarify the role that genetic diversity plays in the survival and reproductive success of individuals. Heterozygosity-fitness correlations (HFCs) have a long history in the population genetics literature (1–5), and they were originally conducted by genotyping samples at a modest panel of allozyme loci and regressing a trait(s) associated with fitness (e.g., survival, growth rate) on multi-locus heterozygosity (MLH) as measured by the panel (3). Modern HFC studies almost exclusively employ microsatellite markers rather than allozymes, and an emphasis has been placed on the use of large numbers of markers (6), although few studies have yet to fulfill this recommendation (5). The shift to larger marker panels containing potentially neutral loci has brought with it a wider availability of statistical tests that help researchers explore the nature of heterozygosity-fitness correlations in their study systems. These tests make it possible to determine i) whether the MLH of the marker panel is reflective of genome-wide MLH (6, 7), ii) whether there is identity disequilibrium (ID) among the markers and consequently inbreeding *sensu lato* in the study system (8, 9), and iii) whether there is evidence for single-marker effects on the trait(s) of interest (9–11) (Figure 3.1). The software now available to conduct these tests as well as run the regressions and correlations that are the core of HFC analyses require input files of different formats, and there is currently no software that provides ecumenicism across these formats.

HeFPipe (short for **He**terozygosity-**F**itness **Pipe**line) is a script written in *Python* that conducts analyses typically performed in heterozygosity-fitness correlation (HFC) studies. It also tests for evidence of single-marker effects on a trait(s). More specifically, *HeFPipe* takes input in the form of allele reports in the “Marker Table” style from the microsatellite genotype-calling

software, *GeneMarker*, and reconfigures the data into *GENEPOP* (12), *Rhh* (7), *RMES* (8), and *GEPHAST* (10) formats (Figure 3.1). The script is also equipped to re-format the *output* from *GENEPOP on the Web* (option 5) and *Rhh* into comma-separated values (csv) formatted spreadsheets and incorporate them into downstream analyses. The *HeFPipe* script accommodates user-provided lists of markers to be included in or excluded from analyses, a list of samples to exclude from analyses, and a spreadsheet containing trait values on which to perform the HFCs and search for single-marker effects. These input files allow the user to refine and repeat each analysis with ease. With regards to the analyses that require regression—HFCs and one of the tests for single-marker effects—*HeFPipe* is equipped to run generalized linear models (GLMs) using the *Python* package *PypeR* (13), a package that enables the statistics software R (14) to be used in the context of a *Python* script. By using GLMs, the user is able to assign a link function and error distribution appropriate for the response variable in a particular model, which can be used to relax some of the assumptions of general linear regression. The script is also equipped to conduct the regression analyses on subsets of the dataset, which might be desirable in various scenarios, such as where HFCs are predicted to appear in stressed individuals (e.g., (15, 16)). Single-marker effects are explored using methodologies described in (11) and using *GEPHAST* (10). Correlations (both Pearson and Spearman) among the traits provided are reported in several different formats (as text, spreadsheets, and images); significance tests are conducted on these correlations, and the *P*-values (both adjusted and unadjusted for multiple comparisons) are also reported in the various formats.

HeFPipe is designed to be used from the command line terminal and will run on any Mac OSX computer that has *Python* v 2.7.3 and *R* v 2.15.1 (or compatible versions) installed. The script, a users' manual, a link to a video tutorial, and an example dataset are available at GitHub

(https://github.com/Atticus29/*HeFPipe*_repos). Software dependencies, including packages in both *R* and *Python* required for the pipeline (e.g., *PypeR*), are listed in the manual, as are brief instructions for their installation.

ACKNOWLEDGEMENTS

I thank Xiao-Qin Xia for help with *PypeR* and Emily Bewick and Kerin Bentley for providing additional datasets on which *HeFPipe* was tested. Kenneth Ross and DeWayne Shoemaker reviewed an early version of the manuscript. This work was funded in part by the Georgia Agricultural Experiment Stations, University of Georgia.

REFERENCES

1. **Mitton JB, Grant MC**. 1984. Associations among protein heterozygosity, growth rate, and developmental homeostasis. Annual Review of Ecology and Systematics **15**:479–499.

2. **Pogson GH**. 1991. Expression of overdominance for specific activity at the Phosphoglucomutase-2 locus in the pacific oyster, *Crassostrea gigas*. Genetics **128**:133–141.

3. **Britten HB**. 1996. Meta-analyses of the association between multilocus heterozygosity and fitness. Evolution **50**:2158–2164.

4. **Coltman DW, Slate J**. 2003. Microsatellite measures of inbreeding: a meta-analysis. Evolution **57**:971–983.

5. **Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC**. 2009. A quantitative review of heterozygosity-fitness correlations in animal populations. Molecular Ecology **18**:2746–2765.

6. **Balloux F, Amos W, Coulson T**. 2004. Does heterozygosity estimate inbreeding in real populations? Molecular Ecology **13**:3021–3031.

7. **Alho JS, Välimäki K, Merilä J, Valimaki K, Merila J**. 2010. Rhh: an R extension for estimating multilocus heterozygosity and heterozygosity-heterozygosity correlation. Molecular Ecology Resources **10**:720–722.

8. **David P, Pujol B, Viard F, Castella V, Goudet J**. 2007. Reliable selfing rate estimates from imperfect population genetic data. Molecular Ecology **16**:2474–2487.

9. **Szulkin M, Bierne N, David P**. 2010. Heterozygosity-fitness correlations: a time for reappraisal. Evolution **64**:1–16.

10. **Amos W, Acevedo-Whitehouse K**. 2009. A new test for genotype-fitness associations reveals a single microsatellite allele that strongly predicts the nature of tuberculosis infections in wild boar. Molecular Ecology Resources **9**:1102–1111.

11. **David P**. 1997. Modeling the genetic basis of heterosis: tests of alternative hypotheses. Evolution **51**:1049–1057.

12. **Raymond M, Rousset F**. 1995. GENEPOP (Version 1.2): Population Genetics Sofware for Exact Tests and Ecumenicism. Heredity **86**:248–249.

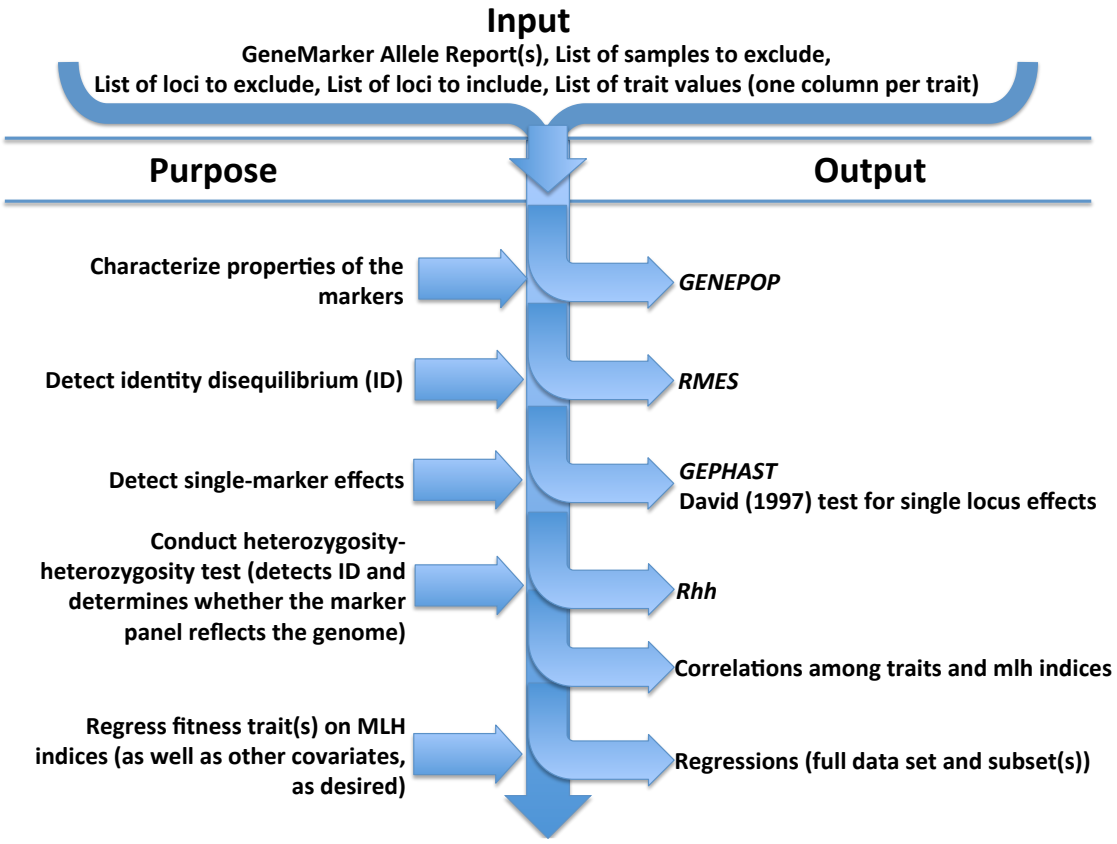
13. **Xia X-Q, McClelland M, Wang Y**. 2010. PypeR, A Python Package for Using R in Python. Journal of Statistical Software **35**.

14. **R Development Core Team**. 2011. R: A Language and Environment for Statistical Computing. U.S. Patent 2.13.2. R Foundation for Statistical Computing, Vienna, Austria.

15. **Schmeller DS, Schregel J, Veith M**. 2007. The importance of heterozygosity in a frog's life. Naturwissenschaften **94**:360–366.

16. **Pujolar JM, Maes GE, Vancoillie C, Volckaert FAM**. 2006. Environmental stress and life-stage dependence on the detection of heterozygosity-fitness correlations in the European eel, *Anguilla anguilla*. Genome **49**:1428–1437.

Figure 3.1. Simplified flowchart depicting the *HeFPipe* pipeline. The input files listed in the top section are used at various relevant points throughout the pipeline. The chronological flow of the pipeline is depicted by the direction of the arrow in the middle of the figure, and the output of the pipeline is depicted on the right, while a brief description of the relevance of each output item to an HFC analysis is described on the left. The items listed under "Output" are files generated by *HeFPipe* that are either useable by external programs (*GENEPOP*, *RMES*, *GEPHAST*, *Rhh*) or are themselves core results of HFC analyses (correlations, regressions). Other, less-essential output files that are also products of *HeFPipe* are not described in this figure but are discussed in the *HeFPipe* manual and tutorial videos along with instructions for how to use the output from the external programs listed in this figure as input in subsequent steps of the *HeFPipe* pipeline.


Input
GeneMarker Allele Report(s), List of samples to exclude,
List of loci to exclude, List of loci to include, List of trait values (one column per trait)
Purpose
Output
Characterize properties of the markers
GENEPOP
Detect identity disequilibrium (ID)
RMES
Detect single-marker effects
GEPHAST
David (1997) test for single locus effects
Conduct heterozygosity-heterozygosity test (detects ID and determines whether the marker panel reflects the genome)
Rhh
Correlations among traits and mlh indices
Regress fitness trait(s) on MLH indices (as well as other covariates, as desired)
Regressions (full data set and subset(s))

CHAPTER 4

FITNESS EFFECTS OF INFECTION BY THREE RNA VIRUSES ON FOUNDRESS QUEENS AND INCIPIENT COLONIES OF THE INVASIVE FIRE ANT *SOLENOPSIS INVICTA*[1]

[1] Fisher, Mark A., Ross, K.G., Shoemaker, D.D. To be submitted to *Journal of Virology.*

ABSTRACT

We investigated the effects of SINV-1, SINV-2, and SINV-3 viral infections on colony-founding *Solenopsis invicta* queens and their incipient colonies for nine fitness-associated traits. We detected significant differences between SINV-2-infected queens and uninfected queens for several of these traits, including initial queen weight, queen weight loss during the claustral period, duration of the claustral period, early colony weight, and early colony growth. These differences likely result in substantial, evolutionarily meaningful reductions in fitness in the SINV-2-infected cohort. No significant differences between queens infected with SINV-1 or SINV-3 and uninfected queens were evident for any of the traits surveyed. This in-depth study highlights the importance of examining a wide array of fitness-associated traits in detail when investigating the phenotypic effects of viral infections in insects, some of which may be subtle or cryptic. Finally, our results suggest that additional studies of SINV-2 infection dynamics are warranted to explore its potential application as a biopesticide of *S. invicta* in its invasive range.

INTRODUCTION

Viruses are the most abundant biological entities on the planet (1), and they infect every type of cellular organism (2, 3). Numerous viruses have been discovered and characterized, especially highly virulent or disease-causing viruses infecting plants and animals (e.g.,4–10). Despite detailed knowledge for many viruses of medical, agricultural, or economical importance, very little is known about the vast majority of viruses. Further, even among many well-described viruses, knowledge often is focused on the molecular and biochemical mechanisms of infection and replication (4–8) rather than on the phenotypic and fitness effects on their hosts. This disparity between knowledge of mechanistic and host-effect dynamics of viruses is especially evident for many that infect insects, which as a group harbor a diversity of DNA and RNA viruses (4). One example of such disparities is represented by *infectious flacherie virus* (IFV), which nearly collapsed the European silkworm industry in the 1860s (11). IFV has been well-characterized over the decades in terms of genome structure, replication mechanisms, and transmission dynamics (4), but its effects on the host were not well understood for much of this period, during which they often were confused with the effects of several other infectious and non-infectious diseases that influence an array of silkworm fitness phenotypes (11).

While information regarding phenotypic effects on hosts currently is unavailable for the large majority of viruses, numerous studies in insects conducted to date have revealed a variety of phenotypic effects of viral infection. Examples of known effects include increased mortality rates, alteration of morphology (12), behavioral modification (13), and manipulation of sex ratio (14), as well as asymptomatic persistence until the host becomes stressed (8). A number of viruses also have been described strictly as asymptomatic (e.g., the *Drosophila* Nora virus; 15, 16). However, it is not clear in these cases whether the viruses are truly asymptomatic or simply

described as such due to the lack of study beyond simple testing for gross phenotypic effects on the hosts.

The observed diversity of viral effects on insect hosts may apply to three recently-identified, positive, single-stranded RNA viruses, SINV-1, SINV-2, and SINV-3, that infect the invasive fire ant *Solenopsis invicta* (17–19). Indeed, the putative effects of these three infections on fire ants appear to range from severe effects (high mortality and complete colony collapse) to apparently little or no effect (asymptomatic). Specifically, previous studies showed that SINV-3-infected colonies exhibit large midden piles of dead ants, some brood mortality, loss of queen physogastry (distension of the abdomen [gaster] due to extensive ovarian development), and colony collapse (18), with only occasional rebound to normal brood production (18). Infections by this virus appear to be systemic, infecting all tissues, castes, and developmental stages of the ants (18). Conversely, no observable symptoms were observed in wild *S. invicta* colonies infected by either SINV-1 or SINV-2 (17, 20), which also infect all castes and developmental stages but evidently are not systemic (17, 21). Furthermore, only minor brood mortality was detected in a limited sample of SINV-1- or SINV-2-infected colonies reared under laboratory conditions (17, 22). However, these earlier studies were based largely on anecdotal evidence. Thus, a more rigorous study measuring the effects of infection of these three viruses on a variety of queen and colony fitness-associated phenotypes is much needed, especially given the possibility for cryptic or late-onset phenotypic effects of persistent viral infections (e.g.,4, 8, 9, 23, 24). For the present study, we measured nine fitness-associated traits related to colony-founding queen survival and early colony growth to explore the consequences of SINV-1, SINV-2, and SINV-3 infection on their fire ant hosts.

This study is of additional interest because of the status of *S. invicta* as an invasive pest in the U.S.A., China, Australia, and elsewhere on the Pacific Rim (25, 26). The damage caused by and effort exerted to manage this invasive pest in the U.S.A. alone costs approximately $6 billion annually (27). Biological control has been recommended as the most promising long-term management strategy for fire ants (28), and several parasites and pathogens, including viruses, have been suggested as agents of sustained population suppression (28). Clearly, basic information on the host effects of presumptive biological control agents and biopesticides is desirable as these are developed as management tools.

MATERIALS AND METHODS

Samples

We collected 652 newly-mated *S. invicta* queens from a single site in Oglethorpe County, Georgia, U.S.A. in June 2008 as they landed from their mating flights. This site was selected because it was previously identified as having predominantly monogyne (single egg-laying queen per colony) colonies. We focused on the monogyne social form primarily because the life history and founding strategies of queens of this form have been studied in greater detail than for queens of the alternate, polygyne social form (multiple egg-laying queens per colony) (29). Furthermore, the relatively few studies that have focused on polygyne colony reproduction suggest that the founding strategies of queens of the two social forms differ dramatically (e.g., 30). All queens were weighed individually upon return to the laboratory; 252 of these queens were selected at random and subsequently stored in a -80 °C freezer pending DNA/RNA analyses. The remaining 400 queens were placed individually in small plastic, screw-top tubes overlaid with a thin layer of castone (dental plaster), which helps maintain high humidity, prevents ants from escaping, and

serves as a nesting site for queens to initiate their colonies in isolation (see 30, 31 for methodological details). All tubes were maintained in a rearing room under the following conditions: 14:10 h light: dark cycle, 28-30 °C, 40-70% RH. Queens were allowed to lay eggs and rear their first cohort of brood without being fed, as colony-founding queens of the monogyne form typically fast during this claustral founding period (29). All tubes were monitored daily for queen mortality and emergence of adult workers. When the first adult worker emerged (marking the end of the claustral period), the mother queen and her entire brood were weighed separately. At this point, 300 randomly selected founding colonies were terminated: these founding queens and all of their offspring were stored in a -80 °C freezer. For the remaining 100 colonies, each queen and her offspring were transferred to larger rearing units. These post-claustral colonies 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), supplemented with frozen crickets provided on a twice-weekly basis (32). Colonies were inspected daily for queen mortality, and dead queens along with their workers and brood were placed immediately in a -80 °C freezer. Surviving queens and their colonies were weighed separately six weeks after the end of the claustral period; after another ten weeks each colony was terminated, with each queen and her offspring stored at -80 °C.

Fitness-Associated Traits

We selected fitness-associated traits for measurement in foundress queens and incipient colonies based on extensive previous studies of fire ant natural history (29). These studies show that rapid growth is essential to newly founded monogyne *S. invicta* colonies because the primary sources of early mortality in North American populations most likely afflict relatively slowly growing colonies. These sources are competition with other fire ant colonies (including other

incipient colonies in newly colonized areas) and climate-related stress during winter (33, 34). Higher worker numbers enable a young colony to compete successfully for an exclusive foraging territory and to excavate the nest to a sufficient soil depth to prevent freezing in winter (34, 35).

We measured two sets of fitness-related traits, the first comprising individual traits of queens and the second comprising traits of the colonies they founded, which, in highly eusocial insects such as fire ants, represent the extended phenotype of the queen. Individual traits included queen size (*QnSize*), initial queen weight (*QnWt*), proportional queen weight loss during the claustral period ($QnWtLs_{CL}$), queen survivorship through the claustral period ($QnSurv_{CL}$), and queen survivorship during the six weeks following the claustral period ($QnSurv_{CL\text{-}6}$, which can also be considered a colony-level trait) (see Table 4.1 for further description of studied traits). Colony traits included the number of days between a queen's mating flight and eclosion of her first adult worker (i.e., the length of the claustral period, *ClaustPer*), colony weight at the end of the claustral period ($ColWt_{CL}$), colony population (number of 4^{th}-instar larval, pupal, and adult workers) at the end of the claustral period, colony weight six weeks after the end of the claustral period ($ColWt_{CL+6}$), colony population six weeks after the end of the claustral period, and proportional colony growth between the end of the claustral period and six weeks later (based on colony weight, $ColGrowth_{CL\text{-}6}$). Because colony weight and population were highly correlated at both time points at which they were measured (Spearman correlation coefficient $r = 0.869$ and 0.712, respectively, both $P < 0.0001$), we report only analyses using weight. Trait measurements were obtained as described in (Fisher, Shoemaker, and Ross, submitted for publication). SINV-1, SINV-2, or SINV-3 infections also were treated as individual traits (*SINV-1, SINV-2,* and *SINV-3*) for the generalized linear model (GLM) analyses described below. Where convenient, all fitness traits are henceforth referenced using the abbreviations above (Table 4.1).

Viral Infections

Molecular assays were used to survey each queen for the presence of three RNA viruses known to infect *S. invicta*: SINV-1, SINV-2, and SINV-3 (17–19). RNA was extracted from the anterior gaster segments of each queen using the AllPrep DNA/RNA Mini Kit (Qiagen) following the manufacturer's instructions (RNA was eluted in 10 μL EB buffer at the final step). Total RNA was subjected to a two-step RT-PCR protocol similar to (36), with the following modifications: a 0.5 μL aliquot of each RNA sample was added to a solution containing 0.37 μL Oligo(dT) primer (Superscript III First strand synthesis kit, Invitrogen), 0.35 μL Annealing Buffer (Superscript III kit), and 1.75 μL H_2O. The solution was heated at 65 °C for 5 min, placed in an ice bath for 1 min, and then centrifuged. Next, 3.75 μL of 2X first strand reaction mix (Superscript III kit), 0.25 μL Superscript III enzyme mix (Superscript III kit), and 0.5 μL H_2O were added to each reaction, heated to 50 °C for 50 min followed by 85 °C for 5 min, and chilled in an ice bath for 2 min. The subsequent PCR reaction contained 1 μL of the cDNA resulting from the previous steps, 6.1 μL H_2O, 7.5 μL of 2X Taq-Pro COMPLETE (2.0 mM $MgCl_2$, Denville Scientific), and 0.2 μL of each of the forward and reverse primers (50 mM) that amplify portions of transcripts of either SINV-1, SINV-2, SINV-3, or the nuclear gene *Gp-9* (see Appendix B1 for primer sequences) — the last as a control for successful RNA extraction and cDNA synthesis. The PCR cycling profile was: 94 °C for 2 min, 35 cycles of 94 °C for 15 s, 60 °C for 15 s, and 68 °C for 30 s, followed by 68 °C for 5 min, and 4 °C indefinitely.

Excluded Samples

We excluded a given queen from further analysis if any of the following four conditions was met: 1) RT-PCR of *Gp-9* transcript was not successful, 2) queen possessed the *b* allele of the

gene *Gp-9*, 3) queen was triploid, or 4) queen produced diploid males. Because RT-PCR of *Gp-9* transcript served as a control for successful RNA extraction/cDNA synthesis, its failure invalidated the viral molecular assay of a queen. Queen *Gp-9* genotype was determined via horizontal starch gel electrophoresis using thorax tissues as described in DeHeer et al. (37); queens bearing a $Gp\text{-}9^{b}$ allele were excluded because they originate exclusively from colonies of the polygyne social form (38). We determined ploidy of each queen by utilizing a previously generated genotype dataset for 99 microsatellite loci for all queens used in the current study (described in Fisher, Shoemaker, and Ross, submitted for publication); the low proportion of queens determined to be triploid (products of matings between a diploid male and queen) were excluded because they appear to have low fitness compared to diploid queens (39). Finally, previous studies have shown that a low proportion of *S. invicta* queens in the U.S.A. mate with a male bearing an allele at the sex-determination locus that matches one of the queen's alleles (40). We excluded such queens from all analyses because they produce diploid male offspring during claustral founding, which invariably leads to colony failure (31). After applying these filtering criteria, 563 queens were retained for the analyses below.

Statistical Tests of Fitness Trait/Viral Infection Associations

Associations of fitness traits with viral infection status initially were examined by comparing trait value distributions between two groups, virus-infected queens and uninfected queens (i.e., queens free of any viral infections), using Mann-Whitney *U* tests for continuous traits and Fisher's exact tests for binary traits. Queens infected by more than one virus were excluded from these analyses. Corrections for multiple testing were made using the Benjamini-Hochberg procedure (5% false discovery rate [FDR]) in *R* (v. 2.15.1) (41). These tests could not

be conducted for traits $ColWt_{CL+6}$, $ColGrowth_{CL-6}$, and $QnSurv_{CL-6}$ in SINV-1-infected queens because of insufficient samples of infected queens.

We next compared mean trait values between another two groups, queens singly infected with either SINV-1 or SINV-2 and queens that were co-infected with both SINV-1 and SINV-2, using Mann-Whitney *U* tests. Corrections for multiple testing again were made using the Benjamini-Hochberg procedure (5% FDR). These tests could not be conducted for traits $ColWt_{CL+6}$ and $ColGrowth_{CL-6}$, as well as for trait $QnSurv_{CL-6}$ in queens singly infected with SINV-1, because of insufficient samples of infected queens.

We next assessed whether viral infections of queens affect fitness-associated trait values by employing the viruses as predictors (i.e., covariates) in generalized linear models (GLMs) with the following error distribution for each response variable: (i) an identity link function and Gaussian error distribution for the response variables *QnSize*, *QnWt*, $QnWtLs_{CL}$, $ColWt_{CL}$, and $ColWt_{CL+6}$, (ii) a logit link function and binomial error distribution for the response variables *SINV-1, SINV-2, SINV-3* (where appropriate), $QnSurv_{CL}$, and $QnSurv_{CL-6}$, or (iii) a log link function and Poisson error distribution for the response variables *ClaustPer* and $ColGrowth_{CL-6}$. Corrections for multiple testing were made using the Benjamini-Hochberg procedure (5% FDR). Models could not be run for traits $ColWt_{CL+6}$ and $ColGrowth_{CL-6}$ in SINV-1-infected queens due to a lack of appropriate samples.

RESULTS

Viral Prevalence

Our molecular assays revealed that all three viruses were present in queens from our study population. We found that twelve of 563 queens (2.1%) were singly infected with SINV-1, 85

(15.1%) were singly infected with SINV-2, and five (0.9%) were singly infected with SINV-3. We also found that 17 queens (2.9%) were co-infected with SINV-1 and SINV-2, a frequency of co-infection significantly greater than expected by chance (log-linear analysis; $P < 0.001$). One queen (0.2%) was co-infected with SINV-2 and SINV-3, and one queen was infected with all three viruses, the latter also a highly improbable event (log-linear analysis; $P < 0.001$). None of the sampled queens was co-infected with SINV-1 and SINV-3. Prevalence estimates based on just the subset of queens frozen immediately upon collection in the field are very similar to those for the whole data set (data not shown), suggesting that transmission, contamination, or selective mortality in the laboratory did not strongly effect our prevalence estimates.

Fitness Trait/Viral Infection Associations

Mean trait values for singly-infected queens and uninfected queens are reported in Table 4.2 along with results of the statistical tests for trait differences between the groups. Values for five traits (*QnWt*, $QnWtLs_{CL}$, *ClaustPer*, $ColWt_{CL}$, and $ColGrowth_{Cl\text{-}6}$) differ significantly between SINV-2-infected and uninfected queens after correction for multiple comparisons. On average, SINV-2-infected queens weigh 0.7 mg (4.5%) less than uninfected queens and proportionally lose 10.0% more of their body weight during the claustral phase, which, on average, lasts 1.6 days longer (see Table 4.2, Figure 4.1). Also, colonies founded by SINV-2-infected queens weigh 28% less at the end of the claustral phase than those founded by uninfected queens, a difference that is compounded six weeks later (average colony growth between the time points is 4.12-fold lower for infected than uninfected queens) (Table 4.2, Figure 4.1).

There are no significant differences in the values of any studied traits between queens infected with SINV-1 or SINV-3 and uninfected queens based on the group comparisons.

However, these results should be interpreted cautiously because of the small number of queens singly infected with these two viruses.

Comparisons of trait values between singly-infected queens and SINV-1/SINV-2 co-infected queens revealed that traits *QnWt* and $ColWt_{CL}$ differ significantly between SINV-1-infected and co-infected queens, while traits $ColWt_{CL}$ and $QnSurv_{CL\text{-}6}$ differ significantly between SINV-2-infected and co-infected queens (Appendix B2). Co-infected queens in all four of these cases have lower trait values than singly-infected queens.

Results of our analyses employing each of the viruses as predictors (covariates) in generalized linear models (GLMs) are presented in Table 4.3. *SINV-2* is a significant predictor of *QnWt*, $QnWtLs_{CL}$, *ClaustPer*, $ColWt_{CL}$, $ColWt_{CL+6}$, and $ColGrowth_{CL\text{-}6}$, as well as of co-infection with SINV-1, paralleling the previous results. In addition, *SINV-1* is a significant predictor of $QnWtLs_{CL}$ and $ColWt_{CL}$, as well as of co-infection with SINV-2, the first two results being novel to this analysis. *SINV-3* is not a significant predictor of any trait.

DISCUSSION

Viruses are a taxonomically diverse group that infect almost every known cellular organism (2, 3). The phenotypic and fitness effects that viruses exert on their hosts also are diverse, and range from severe effects resulting in host mortality or reduced host lifespan (e.g., 7,8,10) to mild or no apparent effects (e.g., 15). However, our current understanding of viral effects on their hosts is limited largely to research on a small subset of well-studied viruses, and little is known about the vast majority that remain. Adding to this deficiency is the fact that many studies are based solely on anecdotal evidence or on laboratory work of limited scope, which may allow detection of extreme viral effects but likely precludes detection of weaker (e.g., mild fitness

reduction) or more specific (e.g., limited to a single life stage or sex) effects of potential importance. Thus, detailed studies, such as those involving thorough examination of numerous fitness-related traits in infected and uninfected hosts, are needed to fully understand the effects of viral infections on their hosts.

To fully characterize the effects of viral infection in wild populations, it is first necessary to understand traits that are closely associated with fitness. Detailed studies of the invasive fire ant *Solenopsis invicta* have revealed a suite of traits correlated with the survival and success of colony-founding queens and their incipient colonies. Specifically, larger queens tend to be heavier, to lose proportionally less weight during the claustral phase, to require a shorter claustral period before first worker eclosion, and to produce larger colonies than smaller queens (Fisher, Shoemaker, and Ross, submitted for publication). These traits contribute in a cohesive manner to enhance the early growth and survival of colonies; rapid growth is essential for continued colony survival because colonies must become sufficiently populous to secure exclusive foraging territories and construct a nest of sufficient depth in the soil to avoid winter freezing (34, 35).

In this study, we measured trait-values for the suite of traits associated with early colony fitness in a single-age cohort of colony-founding monogyne *S. invicta* queens. We assayed these queens for the presence of the three viruses SINV-1, SINV-2, and SINV-3 in order to examine their potential fitness effects on their fire ant hosts at a critical phase of colony ontogeny. The estimated frequencies of each virus reflect viral prevalence in the field when the queens were collected in June 2008 and any dynamics occurring under laboratory rearing conditions in the subset of queens reared there for up to four months; however, the similarity between our estimates for the whole data set and the subset including only queens frozen immediately upon collection suggests that such dynamics were unimportant. Observed prevalences of SINV-1 and

SINV-2 differ slightly from estimates previously reported for these viruses in north-central Florida (28), whereas the infection prevalence of SINV-3 falls within the previously reported range for the time period of sampling (28). This variation in viral prevalence is not surprising, given that SINV-1 prevalence and temperature are correlated (28) and that quite pronounced climatic differences exist between north-central Florida and north Georgia, where our samples were collected.

Analyses of our data reveal significant differences between SINV-2-infected and uninfected queens for several traits: SINV-2-infected queens tend to be lighter, lose proportionally more weight during the claustral period, endure longer claustral periods, and produce colonies that both weigh less at the end of the claustral period and grow more slowly thereafter in comparison to uninfected queens. These infection-related changes associated with reduced early growth and survival strongly implicate SINV-2 infection in reducing queen/early colony fitness (Fisher, Shoemaker, and Ross, submitted for publication; 29, 30).

We did not observe any significant differences in measured trait values between queens singly-infected with either SINV-1 or SINV-3 and uninfected queens. In the GLM analyses, which included queens co-infected with any permutation of the three viruses, SINV-1 was a significant predictor of $QnWtLs_{CL}$, $ColWt_{CL}$, and *SINV-2*. Consideration of these results together suggests that it is only in the presence of SINV-2 that SINV-1 alters fitness-related trait expression. This in turn suggests that an increased probability of SINV-1 infection is likely yet another fitness burden borne by SINV-2-infected queens (cf.,28), a hypothesis supported by the higher than expected numbers of queens co-infected with the two viruses. Although our results suggest that infection with SINV-1 alone does not strongly negatively affect the fitness of young fire ant queens or their incipient colonies, we cannot rule out the possibility of more subtle

impacts or of effects manifested in different life-history stages of both queens and/or the members of their colonies. This will require additional studies of the appropriate material.

Our finding that SINV-3-infected queens did not differ significantly from uninfected queens for any of the study traits is somewhat surprising, given the symptoms of substantial worker and brood mortality, loss of queen physogastry, and colony collapse noted for the virus in a previous study (18). One explanation may be the low power of our analyses caused by the low prevalence of the virus in our study population. Another possibility is that symptoms of SINV-3 infection are limited to mature (several years old) *S. invicta* colonies and have little effect on incipient colony survival and growth, a possibility that may seem doubtful given the conspicuous and severe symptoms suggested by the previous work (18). Future studies that include larger numbers of SINV-3-infected queens and explore the effects of variables such as developmental stage, season, and location will be required to clarify the phenotypic effects of SINV-3 infection on its fire ant host.

In summary, we demonstrate that at least one virus infecting fire ants, SINV-2, has profound negative effects on several traits associated with foundress queen and early colony survival and growth. Our evidence for clear fitness costs of SINV-2 infection on *S. invicta* queens and colonies runs contrary to previous studies suggesting this virus is largely asymptomatic (18). This discrepancy is likely due to the fact that the previous studies were based largely on anecdotal evidence. Our study highlights how rigorous, quantitative measurements of fitness-related traits can reveal significant viral host effects that might otherwise go undetected. Future studies that leverage large sample sizes and survey a wide array of fitness-related traits likely will be required to distinguish viruses with subtle fitness effects from those bearing no fitness costs (e.g., 15, 16). Such distinctions become increasingly important from medical, agricultural, and

conservation perspectives as globalization continues to bring viruses into contact with an increasing diversity of hosts. Finally, our identification of significant fitness costs of SINV-2 infection in *S. invicta* is of special interest because of the status of this ant as a high-impact pest in the U.S.A. (25, 27) and the potential application of viruses as biopesticides in the invasive range (22, 42). Careful evaluation of the host effects can help inform decisions about any role that SINV-1, SINV-2, and SINV-3 may play in management of invasive populations of *S. invicta*.

REFERENCES

1. **Breitbart M, Rohwer F**. 2005. Here a virus, there a virus, everywhere the same virus? Trends in Microbiology **13**:278–284.
2. **Koonin E V, Senkevich TG, Dolja V V**. 2006. The ancient Virus World and evolution of cells. Biology Direct **1**:29.
3. **Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA**. 2005. Virus Taxonomy: VIIIth Report of the International Committee on Taxonomy of Viruses, 8th ed. Elsevier Inc., Oxford, UK.
4. **Asgari S, Johnson K**. 2010. Insect Virology. Caister Academic Press, Norfolk, UK.
5. **Tijssen P, Bergoin M**. 1995. Densonucleosis viruses constitute an increasingly diversified subfamily among the parvoviruses. Seminars in Virology **6**:347–355.
6. **Jousset FX, Baquerizo E, Bergoin M**. 2000. A new densovirus isolated from the mosquito *Culex pipiens* (Diptera: culicidae). Virus Research **67**:11–16.
7. **Huger AM**. 1966. A virus disease of the Indian rhinoceros beetle, *Oryctes rhinoceros* (Linnaeus), caused by a new type of insect virus, Rhabdionvirus oryctes gen. n., sp. n. Journal of Invertebrate Pathology **8**:38–51.
8. **Chen YP, Siede R**. 2007. Honey bee viruses. Advances in Virus Research **70**:33–80.
9. **Pelletier I, Duncan G, Pavio N, Colbère-Garapin F**. 1998. Molecular mechanisms of poliovirus persistence: key role of capsid determinants during the establishment phase. Cellular and Molecular Life Sciences **54**:1385–1402.

10. **Klatzmann D, Gluckman JC**. 1987. The pathophysiology of HIV infection: A complex pathway of host-virus interaction. Annales de l'Institut Pasteur-Virologie **138**:21–29.

11. **Steinhaus EA**. 1956. Microbial control-the emergence of an idea. A Journal of Agricultural Science **26**:107–160.

12. **De Miranda JR, Genersch E**. 2010. Deformed wing virus. Journal of Invertebrate Pathology **103**:S48–561.

13. **Fujiyuki T, Takeuchi H, Ono M, Ohka S, Sasaki T, Nomoto A, Kubo T**. 2004. Novel insect picorna-like virus identified in the brains of aggressive worker honeybees. Journal of Virology **78**:1093–1100.

14. **Nakanishi K, Hoshino M, Nakai M, Kunimi Y**. 2008. Novel RNA sequences associated with late male killing in *Homona magnanima*. Proceedings of the Royal Society B-Biological Sciences **275**:1249–1254.

15. **Habayeb MS, Cantera R, Casanova G, Ekström J-O, Albright S, Hultmark D**. 2009. The *Drosophila* Nora virus is an enteric virus, transmitted via feces. Journal of Invertebrate Pathology **101**:29–33.

16. **Habayeb MS, Ekengren SK, Hultmark D**. 2006. Nora virus, a persistent virus in *Drosophila*, defines a new picorna-like virus family. The Journal of General Virology **87**:3045–3051.

17. **Valles SM, Strong CA, Dang PM, Hunter WB, Pereira RM, Oi DH, Shapiro AM, Williams DF**. 2004. A picorna-like virus from the red imported fire ant, *Solenopsis invicta*: initial discovery, genome sequence, and characterization. Virology **328**:151–157.

18. **Valles SM, Hashimoto Y**. 2009. Isolation and characterization of Solenopsis invicta virus 3, a new positive-strand RNA virus infecting the red imported fire ant, *Solenopsis invicta*. Virology **388**:354–361.

19. **Valles SM, Strong CA, Hashimoto Y**. 2007. A new positive-strand RNA virus with unique genome characteristics from the red imported fire ant, *Solenopsis invicta*. Virology **365**:457–463.

20. **Valles SM, Hashimoto Y**. 2008. Detection and quantitation of *Solenopsis invicta* virus-2 genomic and intermediary replicating viral RNA in fire ant workers and larvae. Journal of Invertebrate Pathology **98**:243–245.

21. **Hashimoto Y, Valles SM**. 2008. Infection characteristics of Solenopsis invicta virus 2 in the red imported fire ant, *Solenopsis invicta*. Journal of Invertebrate Pathology **99**:136–140.

22. **Valles SM**. 2012. Positive-Strand RNA Viruses Infecting the Red Imported Fire Ant, *Solenopsis invicta*. Psyche: A Journal of Entomology **2012**:1–14.

23. **Julien J, Leparc-Goffart I, Lina B, Fuchs F, Foray S, Janatova I, Aymard M, Kopecka H**. 1999. Postpolio syndrome: poliovirus persistence is involved in the pathogenesis. Journal of Neurology **246**:472–476.

24. **Klingel K, Hohenadl C, Canu A, Albrecht M, Seemann M, Mall G, Kandolf R**. 1992. Ongoing enterovirus-induced myocarditis is associated with persistent heart muscle infection: quantitative analysis of virus replication, tissue damage, and inflammation. Proceedings of the National Academy of Sciences of the United States of America **89**:314–318.

25. **IUCN**. 2000. One-hundred of the world's worst invasive alien species: a selection from the Global Invasive Species Database. http://www.iucn.org/places/medoffice/invasive_species/docs/invasive_species_booklet.pdf.

26. **Ascunce MS, Yang C, Oakey J, Calcaterra L, Wu W, Shih C, Goudet J, Ross KG, Shoemaker D**. 2011. Global invasion history of the fire ant. Science **331**:1066–1068.

27. **Pereira RM**. 2003. Areawide suppression of fire ant populations in pastures: project update. Journal of Agricultural and Urban Entomology **20**:123–130.

28. **Valles SM, Oi DH, Porter SD**. 2010. Seasonal variation and the co-occurrence of four pathogens and a group of parasites among monogyne and polygyne fire ant colonies. Biological Control **54**:342–348.

29. **Tschinkel WR**. 2006. The Fire Ants, 1st ed. The Belknap Press of Harvard University Press, Cambridge, MA.

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

31. **Ross KG, Fletcher DJC**. 1986. Diploid male production-a significant colony mortality factor in the fire ant *Solenopsis invicta* (Hymenoptera:Formicidae). Behavioral Ecology and Sociobiology **19**:283–291.

32. **Eliyahu D, Ross KG, Haight KL, Keller L, Liebig J**. 2011. Venom alkaloid and cuticular hydrocarbon profiles are associated with social organization, queen fertility status, and queen genotype in the fire ant *Solenopsis invicta*. Journal of Chemical Ecology **37**:1242–1254.

33. **Wilson EO**. 1971. The Insect Societies. Harvard University Press, Cambridge, MA.

34. **Markin G, Dillier JH, Collins H**. 1973. Growth and development of colonies of the red imported fire ant, *Solenopsis invicta*. Annals of the Entomological Society of America **66**:803–808.

35. **Tschinkel WR, Howard DF**. 1983. Colony founding by pleometrosis in the fire ant, *Solenopsis invicta*. Behavioral Ecology and Sociobiology **12**:103–113.

36. **Valles SM, Varone L, Ramírez L, Briano J**. 2009. Multiplex detection of *Solenopsis invicta* viruses-1, -2, and -3. Journal of Virological Methods **162**:276–279.

37. **Deheer CJ, Goodisman MAD, Ross KG**. 1999. Queen dispersal strategies in the multiple-queen form of the fire ant *Solenopsis invicta*. American Naturalist **153**:660–675.

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

39. **Krieger MJB, Ross KG, Chang CWY, Keller L**. 1999. Frequency and origin of triploidy in the fire ant *Solenopsis invicta*. Heredity **82**:142–150.

40. **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.

41. **R Development Core Team**. 2011. R: A Language and Environment for Statistical Computing. U.S. Patent 2.13.2. R Foundation for Statistical Computing, Vienna, Austria.

42. **Tufts DM, Spencer K, Hunter WB, Bextine B**. 2011. Delivery system using sodium alginate virus loaded pellets to Red Imported Fire Ants (*Solenopsis invicta*, Hymenoptera : Formicidae). Florida Entomologist **94**:237–241.

Table 4.1. Trait abbreviations and descriptions.

Abbreviation	Trait description
QnSize	Queen size: PC1 from a principle components analysis of queen metathoracic femur length, antennal scape length, and head width
QnWt	Queen weight on the day of her nuptial flight
$QnWtLs_{CL}$	Queen weight loss during the claustral period: proportional weight loss of queen between the day of her nuptial flight and eclosion of her first worker
ClaustPer	Claustral period: number of days between a queen's mating flight and eclosion of her first worker
$QnSurv_{CL}$	Queen survival until the end of the claustral period (eclosion of her first worker)
$ColWt_{CL}$	Colony weight at the end of the claustral period (time of first worker eclosion)
$ColWt_{CL+6}$	Colony weight six weeks after the end of the claustral period (six weeks after eclosion of the first worker)
$ColGrowth_{CL-6}$	Colony growth between the end of the claustral period and six weeks thereafter: number-fold increase in colony weight between the day of first worker eclosion and six weeks after that day
$QnSurv_{CL-6}$	Queen survival between the end of the claustral period (time of first worker eclosion) and six weeks thereafter

Table 4.2. Results of statistical tests for trait-value differences between singly-infected and uninfected queens.

Virus[a]	Trait	Test	*W*-score	*P*-value[b]	Sample size	Mean / Frequency[c]		Difference[d]
						Infected	Uninfected	
SINV-1	*QnSize*	Mann-Whitney *U*	1821	0.296	345	0.02	0.00	0.02
SINV-1	*QnWt*	Mann-Whitney *U*	3048	0.378	454	15.36	15.04	0.32
SINV-1	$QnWtLs_{CL}$	Mann-Whitney *U*	264	0.926	185	0.49	0.50	-0.01
SINV-1	*ClaustPer*	Mann-Whitney *U*	318	0.618	185	21.33	21.25	0.09
SINV-1	$QnSurv_{CL}$	Fisher's Exact	NA	1.000	260	0.75	0.75	0.00
SINV-1	$ColWt_{CL}$	Mann-Whitney *U*	427	0.096	185	11.80	10.20	1.60
SINV-2	*QnSize*	Mann-Whitney *U*	11218	0.431	407	-0.01	0.00	-0.01
SINV-2	*QnWt*	Mann-Whitney *U*	13053	**<0.001**	527	14.37	15.04	-0.66
SINV-2	$QnWtLs_{CL}$	Mann-Whitney *U*	2950	**0.001**	205	0.60	0.50	0.10
SINV-2	*ClaustPer*	Mann-Whitney *U*	2818	**0.001**	204	22.86	21.25	1.62
SINV-2	$QnSurv_{CL}$	Fisher's Exact	NA	**0.020**	302	0.57	0.75	-0.18
SINV-2	$ColWt_{CL}$	Mann-Whitney *U*	1009	**<0.001**	205	7.38	10.20	-2.82
SINV-2	$ColWt_{CL+6}$	Mann-Whitney *U*	49	**0.013**	65	18.84	58.65	-39.81
SINV-2	$ColGrowth_{CL-6}$	Mann-Whitney *U*	40	**0.007**	65	1.70	5.82	-4.12
SINV-2	$QnSurv_{CL-6}$	Fisher's Exact	NA	1.000	71	1.00	0.92	0.08
SINV-3	*QnSize*	Mann-Whitney *U*	587	0.666	340	-0.01	0.00	-0.01

SINV-3	*QnWt*	Mann-Whitney *U*	887	0.449	447	15.06	15.04	0.02
SINV-3	$QnWtLs_{CL}$	Mann-Whitney *U*	351.5	0.910	186	0.50	0.50	0.00
SINV-3	*ClaustPer*	Mann-Whitney *U*	424.5	0.557	186	1.50	1.25	0.25
SINV-3	$QnSurv_{CL}$	Fisher's Exact	NA	1.000	261	0.80	0.75	0.05
SINV-3	$ColWt_{CL}$	Mann-Whitney *U*	513.5	0.162	186	11.55	10.20	1.35
SINV-3	$ColWt_{CL+6}$	Mann-Whitney *U*	143	0.533	64	63.78	58.65	5.13
SINV-3	$ColGrowth_{CL-6}$	Mann-Whitney *U*	119	0.989	64	5.55	5.82	-0.28
SINV-3	$QnSurv_{CL-6}$	Fisher's Exact	NA	1.000	69	1.00	0.92	0.08

[a] Queens infected with a given virus were compared to queens uninfected with any of the three viruses.

[b] *P*-values ≤0.05 are shown in bold, while *P*-values significant after correction for multiple comparisons (Benjamini-Hochberg procedure, 5% FDR) are bold and underlined.

[c] Means are reported for continuous traits, while frequencies are reported for the binary traits $QnSurv_{CL}$ and $QnSurv_{CL-6}$. *QnSize* is reported as the PC1 score from principal components analysis (PCA), while *QnWt*, $ColWt_{CL}$, and $ColWt_{CL+6}$ are reported as weight in mg. $QnWtLs_{CL}$ is the proportional weight loss of queens during the claustral period. $ColGrowth_{CL-6}$ is the proportional weight gain of colonies during the six-week period immediately following the end of the claustral period.

[d] Mean / frequency of infected queens minus mean / frequency of uninfected queens.

Table 4.3. Results of generalized linear models for fitness-associated traits and viral presence using *SINV-1*, *SINV-2*, and *SINV-3* infection as predictors (covariates).

Response variable/trait	Predictor	Regression coefficient	Standard error	Type of score	Score	*P*-value[a]	Sample size	Effect size[b]	R^2
QnSize	*SINV-1*	0.02	0.01	*t*	1.48	0.140	433	0.06	<0.01
QnSize	*SINV-2*	-0.01	0.01	*t*	-1.35	0.177	433	-0.05	<0.01
QnSize	*SINV-3*	-0.01	0.02	*t*	-0.33	0.743	433	-0.01	<0.01
QnWt	*SINV-1*	0.10	0.21	*t*	0.47	0.641	563	0.04	<0.01
QnWt	*SINV-2*	-0.72	0.12	*t*	-5.98	**<0.001**	563	-0.25	0.06
QnWt	*SINV-3*	-0.01	0.41	*t*	-0.03	0.973	563	-0.01	<0.01
$QnWtLs_{CL}$	*SINV-1*	0.06	0.03	*t*	2.42	**0.016**	219	0.31	0.09
$QnWtLs_{CL}$	*SINV-2*	0.11	0.02	*t*	7.02	**<0.001**	219	0.49	0.24
$QnWtLs_{CL}$	*SINV-3*	-0.01	0.03	*t*	-0.41	0.683	219	-0.01	<0.01
ClaustPer	*SINV-1*	-0.49	0.50	*Z*	-0.98	0.329	212	-0.07	<0.01
ClaustPer	*SINV-2*	0.80	0.14	*Z*	5.61	**<0.001**	212	0.39	0.15
ClaustPer	*SINV-3*	0.18	0.41	*Z*	0.44	0.663	212	0.03	<0.01
$QnSurv_{CL}$	*SINV-1*	1.05	0.82	*Z*	1.28	0.200	319	0.07	0.01
$QnSurv_{CL}$	*SINV-2*	-0.75	0.33	*Z*	-2.29	**0.022**	319	-0.13	0.02
$QnSurv_{CL}$	*SINV-3*	0.51	1.11	*Z*	0.46	0.648	319	0.03	<0.01
$ColWt_{CL}$	*SINV-1*	-2.71	0.80	*t*	-3.37	**0.001**	219	-0.36	0.13
$ColWt_{CL}$	*SINV-2*	-3.48	0.49	*t*	-7.17	**<0.001**	219	-0.51	0.26
$ColWt_{CL}$	*SINV-3*	0.39	1.05	*t*	0.37	0.709	219	0.01	<0.01
$ColWt_{CL+6}$	*SINV-2*	-40.04	11.19	*t*	-3.58	**0.001**	69	-0.41	0.17

$ColWt_{CL+6}$	*SINV-3*	4.89	12.41	*t*	0.39	0.695	69	0.07	0.01
$ColGrowth_{CL-6}$	*SINV-2*	-1.23	0.35	*Z*	-3.54	**<u>≤0.001</u>**	69	-0.43	0.18
$ColGrowth_{CL-6}$	*SINV-3*	-0.05	0.22	*Z*	-0.22	0.827	69	-0.03	<0.01
$QnSurv_{CL-6}$	*SINV-1*	-39.13	8780.61	*Z*	<-0.01	0.996	77	<-0.01	<0.01
$QnSurv_{CL-6}$	*SINV-2*	-1.45	0.94	*Z*	-1.55	0.122	77	-0.18	0.03
$QnSurv_{CL-6}$	*SINV-3*	17.08	5377.01	*Z*	<0.01	0.997	77	<-0.01	<0.01
SINV-1	*SINV-2*	2.04	0.39	*Z*	5.22	**<u>≤0.001</u>**	563	0.22	0.05
SINV-1	*SINV-3*	0.93	1.17	*Z*	0.80	0.427	563	0.03	<0.01
SINV-2	*SINV-1*	2.04	0.39	*Z*	5.22	**<u>≤0.001</u>**	563	0.22	0.05
SINV-2	*SINV-3*	0.38	0.90	*Z*	0.42	0.674	563	0.02	<0.01
SINV-3	*SINV-1*	0.93	1.17	*Z*	0.80	0.427	563	0.03	<0.01
SINV-3	*SINV-2*	0.38	0.90	*Z*	0.42	0.674	563	0.02	<0.01

[a] *P*-values ≤0.05 are shown in bold, while *P*-values significant after correction for multiple comparisons (Benjamini-Hochberg procedure, 5% FDR) are bold and underlined.

[b] Effect sizes are reported directly as Pearson product moments (continuous traits) or by conversion from the regression *Z*-score (binary traits) (Coltman & Slate 2003). Signs of the effect sizes converted from *Z*-scores were recovered from the signs of the regression coefficients.

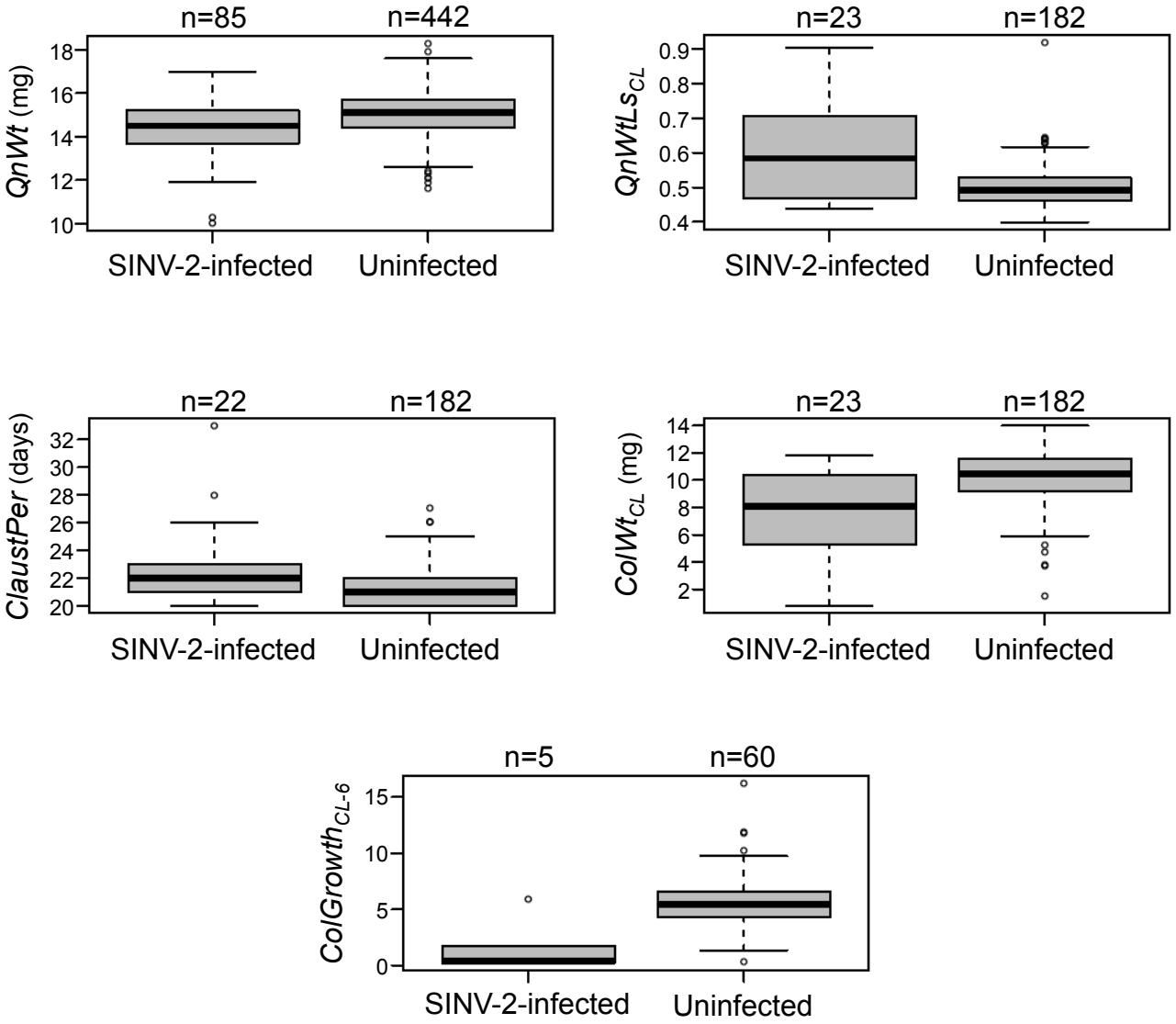


Figure 4.1. Distributions of values for fitness traits exhibiting significant differences between SINV-2-infected queens and those uninfected with any virus. Boxplots depict the first quartile as the bottom of the box, third quartile as the top, and median as a horizontal line. Boxplot whiskers encompass values below the first and above the third quartile by an amount as much as 1.5 times the difference in the third and first quartile values; open circles indicate outliers beyond these limits. See Table 4.1 for definitions of trait abbreviations.

CHAPTER 5

CONCLUSIONS

Both within-individual genetic diversity and viral infection can have effects on individual fitness ranging from substantial to inconsequential. Characterizing the magnitude of these effects in an organism can inform the role that they play in adaptive evolution, which can further our understanding both of the importance of these forces in a particular system and of evolutionary theory in general. In this dissertation, I conducted a heterozygosity-fitness association study on an invasive population of the fire ant *Solenopsis invicta* (Chapter 2), because this system was poised to answer important questions surrounding HFA theory. I also developed an analytical pipeline, HeFPipe (Chapter 3), which in addition to being useful in the HFA study in *S. invicta* aims to facilitate and streamline future HFA analyses. Finally, I explored the phenotypic effects of viral infection by SIN-1, -2, and -3 on newly-mated *S. invicta* queens and their incipient colonies (Chapter 4).

In the heterozygosity-fitness association study conducted across 12 fitness-related traits in *S. invicta* queens and incipient colonies (Chapter 2), analyses revealed that HFAs are uncommon in our study population, with queen multilocus heterozygosity a significant predictor of just a small subset of the fitness traits that we measured: survival of colony-founding queens until eclosion of their first workers ($QnSurv_{CL}$) is strongly and consistently implicated, while the length of the claustral period (*ClaustPer*) is implicated as well in some analyses. A third trait, proportional colony growth for the six-week

period immediately following the claustral period ($ColGrowth_{CL\text{-}6}$), was weakly supported as being influenced by heterozygosity in the complete dataset. For the first two HFAs, increased heterozygosity corresponds with a decreased probability of queen survival and a decreased duration of the claustral period. For the third, the situation is complicated by viral infection status of queens — higher heterozygosity weakly corresponds with faster post-claustral growth in queens infected with the SINV-2 virus, but it corresponds with slower post-claustral growth in uninfected queens. Assuming that shorter claustral periods and faster post-claustral growth reflect higher fitness (higher probability of survival) for a queen and her colony, we thus detected evidence for both positive (*ClaustPer*, $ColGrowth_{CL\text{-}6}$ in infected queens) and negative ($QnSurv_{CL}$, $ColGrowth_{CL\text{-}6}$ in uninfected queens) HFAs in our study population.

The sweep of HFA studies over the past few decades has spanned from relatively simple explanations of influential focal study outcomes that invoked overdominance at single loci (1–3) to progressively more complex explanations of diverse outcomes that feature the general effects of inbreeding depression, sometimes in combination with direct or local effects of single loci (4–7). As the complicated dynamics of HFAs, and the network of evolutionary forces potentially at play, have become increasingly apparent, empirical studies face the challenge of ever more sophisticated study designs necessary to address specific issues. Future studies not only must heed the now well-known imperative to deploy large sample sizes and marker panels (8, 9) but also should consider the effects of stress, life-history tradeoffs, SLAs, and variation in life-cycle stages assayed (e.g., 10) in novel study outcomes such as the presence of both “positive” and “negative” HFAs (6, 7, 11). Based on our results, we advocate exclusion of SLA-

implicated loci in subsets of HFA analyses, or at the very least, separate treatment of loci known to have some functional role (e.g., 6, 7, 12), in order to help clarify single-gene effects in generating HFAs. We also endorse investigation of different life-cycle stages because of the possibility that heterozygosity changes due to selection at an early stage affect the form of HFAs at later stages (10); as an example, the negative HFA for queen survival during the claustral period in our study evidently truncates the range of heterozygosity available for expression of HFAs in the post-claustral phase. Finally, comparative HFA studies that take advantage of different population structures or genome architectures may yield important data for formulating a more comprehensive theory of the role of HFAs in maintaining individual and population genetic diversity (9, 13–15). In our case, this can be accomplished by investigating HFAs in native South American and recently introduced Pacific Rim *S. invicta* populations; because of the very different population histories (16, 17), each may be unique in terms of such features as additive genetic variance and linkage disequilibrium, which may facilitate formulation of specific predictions about the extent and patterns of HFA.

Our findings in Chapter 2 make salient the increasing complexity of HFA analyses and in particular emphasize the importance of considering life-history tradeoffs, variation in life-cycle stages, SLAs, and stress. In light of this, I constructed HeFPipe to accommodate such complexity (Chapter 3). Specifically, HeFPipe enables users to conduct HFA analyses of several traits, including interactions among those traits, which of course would be necessary to detect life-history tradeoffs among traits and to identify variation in the strength of HFAs at different life stages. HeFPipe is also equipped with two separate algorithms that can identify SLAs, *GEPHAST* (18) and David's (19) F-ratio

test. Finally, HeFPipe allows users to divide their datasets into subsets, which would be desirable to investigate the role of stress (say, viral infection) on the HFAs of a system. All of these functionalities can be accomplished in one run of the pipeline.

Chapter 4 involves the investigation of just such a stressor in an invasive population of *S. invicta*, namely viral infection. Rather than investigating the role of viral infection on HFAs (covered in Chapter 2), Chapter 4 directly investigated the role of three, positive, single-stranded RNA viruses on queen and incipient colony fitness-related traits. We demonstrated that at least one virus infecting fire ants, SINV-2, has profound negative effects on several traits associated with early queen and colony growth and survival. Our evidence for clear fitness costs of SINV-2 infection on *S. invicta* queens and colonies runs contrary to previous studies suggesting this virus is largely asymptomatic (20). This discrepancy is likely due to the fact that the previous studies were based largely on anecdotal evidence. Our study highlights how rigorous, quantitative measurements of fitness-related traits can reveal significant viral host effects that might otherwise go undetected. Future studies that leverage large sample sizes and survey a wide array of fitness-related traits likely will be required to distinguish viruses with subtle fitness effects from those bearing no fitness costs (e.g., 21, 22). Such distinctions become increasingly important from medical, agricultural, and conservation perspectives as globalization continues to bring viruses into contact with an increasing diversity of hosts. Finally, our identification of significant fitness costs of SINV-2 infection in *S. invicta* is of special interest because of the status of this ant as a high-impact pest in the U.S.A. (23, 24) and the potential application of viruses as biopesticides in the invasive range (25, 26). Careful evaluation of the host effects can help inform

decisions about any role that SINV-1, SINV-2, and SINV-3 may play in management of invasive populations of *S. invicta*.

Thus, both multilocus heterozygosity and viral infection are associated with fitness traits of queens and their incipient colonies in an invasive population of *S. invicta*, but neither in the manner in which we initially expected. These associations bear directly on the fitness dynamics of the invasive *S. invicta* population, but also have implications for other systems. In particular, each study highlighted several factors that made such associations detectable in our population, and we advocate the consideration of these factors in future studies.

REFERENCES

1. **Mitton JB, Grant MC**. 1984. Associations among protein heterozygosity, growth rate, and developmental homeostasis. Annual Review of Ecology and Systematics **15**:479–499.

2. **Zouros E, Foltz DW**. 1987. The use of allelic isozyme variation for the study of heterosis. Isozymes-Current Topics in Biological and Medical Research **13**:1–59.

3. **Smouse PE**. 1986. The fitness consequences of multiple-locus heterozygosity under the multiplicative overdominance and inbreeding depression models. Evolution **40**:946–957.

4. **David P**. 1998. Heterozygosity-fitness correlations: new perspectives on old problems. Heredity **80**:531–537.

5. **Szulkin M, Bierne N, David P**. 2010. Heterozygosity-fitness correlations: a time for reappraisal. Evolution **64**:1–16.

6. **Olano-Marin J, Mueller JC, Kempenaers B**. 2011. Heterozygosity and survival in blue tits (*Cyanistes caeruleus*): contrasting effects of presumably functional and neutral loci. Molecular Ecology **20**:4028–4041.

7. **Laine VN, Herczeg G, Shikano T, Primmer CR**. 2012. Heterozygosity-behaviour correlations in nine-spined stickleback (*Pungitius pungitius*) populations: contrasting effects at random and functional loci. Molecular Ecology **21**:4872–4884.

8. **Coltman DW, Slate J**. 2003. Microsatellite measures of inbreeding: a meta-analysis. Evolution **57**:971–983.

9. **Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC**. 2009. A quantitative review of heterozygosity-fitness correlations in animal populations. Molecular Ecology **18**:2746–2765.

10. **Pujolar JM, Maes GE, Vancoillie C, Volckaert FAM**. 2006. Environmental stress and life-stage dependence on the detection of heterozygosity-fitness correlations in the European eel, *Anguilla anguilla*. Genome **49**:1428–1437.

11. **Marshall TC, Spalton JA**. 2000. Simultaneous inbreeding and outbreeding depression in reintroduced Arabian oryx. Animal Conservation **3**:241–248.

12. **Mueller JC, Hermisson J, Olano-Marin J, Hansson B, Kempenaers B**. 2010. Linking genetic mechanisms of heterozygosity-fitness correlations to footprints of selection at single loci. Evolutionary Ecology **25**:1–11.

13. **Grueber CE, Wallis GP, Jamieson IG**. 2008. Heterozygosity-fitness correlations and their relevance to studies on inbreeding depression in threatened species. Molecular Ecology **17**:3978–3984.

14. **Luquet E, Léna J-P, David P, Prunier J, Joly P, Lengagne T, Perrin N, Plénet S**. 2013. Within- and among-population impact of genetic erosion on adult fitness-related traits in the European tree frog *Hyla arborea*. Heredity **110**:347–54.

15. **González-Varo JP, Aparicio A, Lavergne S, Arroyo J, Albaladejo RG**. 2012. Contrasting heterozygosity-fitness correlations between populations of a self-compatible shrub in a fragmented landscape. Genetica **140**:31–38.

16. **Ross K, Krieger M, Keller L, Shoemaker DD**. 2007. Genetic variation and structure in native populations of the fire ant *Solenopsis invicta*: evolutionary and demographic implications. Biological Journal of the Linnean Society **92**:541–560.

17. **Ascunce MS, Yang C, Oakey J, Calcaterra L, Wu W, Shih C, Goudet J, Ross KG, Shoemaker D**. 2011. Global invasion history of the fire ant. Science **331**:1066–1068.

18. **Amos W, Acevedo-Whitehouse K**. 2009. A new test for genotype-fitness associations reveals a single microsatellite allele that strongly predicts the nature of tuberculosis infections in wild boar. Molecular Ecology Resources **9**:1102–1111.

19. **David P**. 1997. Modeling the genetic basis of heterosis: tests of alternative hypotheses. Evolution **51**:1049–1057.

20. **Valles SM, Hashimoto Y**. 2009. Isolation and characterization of Solenopsis invicta virus 3, a new positive-strand RNA virus infecting the red imported fire ant, *Solenopsis invicta*. Virology **388**:354–361.

21. **Habayeb MS, Cantera R, Casanova G, Ekström J-O, Albright S, Hultmark D**. 2009. The *Drosophila* Nora virus is an enteric virus, transmitted via feces. Journal of Invertebrate Pathology **101**:29–33.

22. **Habayeb MS, Ekengren SK, Hultmark D**. 2006. Nora virus, a persistent virus in *Drosophila*, defines a new picorna-like virus family. The Journal of General Virology **87**:3045–51.

23. **IUCN**. 2000. One-hundred of the world's worst invasive alien species: a selection from the Global Invasive Species Database. http://www.iucn.org/places/medoffice/invasive_species/docs/invasive_species_booklet.pdf.

24. **Pereira RM**. 2003. Areawide suppression of fire ant populations in pastures: project update. Journal of Agricultural and Urban Entomology **20**:123–130.

25. **Valles SM**. 2012. Positive-Strand RNA Viruses Infecting the Red Imported Fire Ant, *Solenopsis invicta*. Psyche: A Journal of Entomology **2012**:1–14.

26. **Tufts DM, Spencer K, Hunter WB, Bextine B**. 2011. Delivery system using sodium alginate virus loaded pellets to Red Imported Fire Ants (*Solenopsis invicta*, Hymenoptera : Formicidae). Florida Entomologist **94**:237–241.

27. **Valles SM, Oi DH, Porter SD**. 2010. Seasonal variation and the co-occurrence of four pathogens and a group of parasites among monogyne and polygyne fire ant colonies. Biological Control **54**:342–348.

APPENDICES

Appendix A.1. Information for each locus included in the marker panel. Diversity statistics are reported separately for each marker and each marker class (microsatellites and allozymes). NA = information not applicable to allozyme markers. Information for markers developed initially but ultimately not included in the final panel is available upon request from M.A.F.

Marker type	Locus	Forward primer sequence (5'-3')[1]	Reverse primer sequence (5'-3')[1]	PCR cycling profile[2]	Core repeat	Size of repeat	Scaffold in current assembly[3]	Number of individuals scored	Observed number of alleles	Effective number of alleles[4]	H_{obs}[5]	H_{exp}[5]	Reference for marker development
Microsatellites													
	Anna	NED-TACGAATTTCAAAGGCAAACGAG	GCTTCTCTCACGCGAATCGCCAAC	Msat v4	$(CGAT)_6$	4	5431	529	3	2.382	0.5615	0.5807	Present study
	Baez	NED-AGTAGATTTCAACGCATGACACC	GCTTCTGAATGTATGCGGGAGGGATT	Msat v4	$(CGT)_6$	3	9348	529	2	1.929	0.4688	0.4822	Present study
	Baron_Harkonnen	NED-TTAACCGACGAATGATCAAACA	GCTTCTCTTAAATGCCGATACAAAGAGTC	Msat v1	$(AT)_9$	2	4510	484	2	1.118	0.0944	0.1055	Present study
	Beatles	VIC-TTTGGATTATCGACATACATACATT	GCTTCTGCGCTCTTTCAGTGAATTTT	Msat v4	$(ACAT)_6$	4	2940	548	2	1.016	0.0153	0.0152	Present study
	Bertha	6FAM-CGTTTAACTTTAACTTTCCCGAAC	GCATATTTCACGACGATCCTTAAT	Msat v4	$(CT)_{12}$	2	1203	540	5	2.378	0.6023	0.5801	Ascunce et al. 2011
	Black_Crow	NED-GAGGTGCGTTAGCTTCGTAGAAT	GCTTCTGCCATTCTCTGCAAATGTTAAA	Msat v4	$(AG)_8$	2	7705	527	5	3.485	0.6753	0.7138	Present study
	Blackbird	FAM-CGTAAATAGAGAGACGCCGAGTT	GCTTCTGTTTCTCTCTCGCCGGATTA	Msat v4	$(AAC)_6$	3	4103	558	2	1.250	0.1876	0.2000	Present study
	Blue_Jay	VIC-CTAAATTCACTGGGGACTGCTAA	GCTTCTAAGGATTCTTTTGCCATCTTTTG	Msat v4	$(AAG)_6$	3	10511	553	3	1.688	0.4102	0.4081	Present study
	C1	VIC-AAGCAATAACGATTTGGGGAGG	CAGATTTTCACAAGTAGCGGGTTC	Msat v1	$(CA)_7$	2	2923	512	6	2.080	0.5122	0.5199	Ascunce et al. 2009
	C147	FAM-GAATTTCGTCGCAGCAGGG	ACGACCGTCCTCTCGCTTTG	Msat v1	$(GA)_3AA(GA)_{11}$	2	1203	559	4	2.108	0.5169	0.5262	Ascunce et al. 2009
	C204	NED-GTTCGCCAAGGGAAATGAGG	AGTCAAACCGTCGCAGTATTGC	Msat v1	$(TC)_5A(CT)_5ATCG(CT)_4$	2	6724	541	7	1.979	0.4709	0.4953	Ascunce et al. 2009
	C21	VIC-CCCGCCACATAAGAACCTAA	GCAGGGAATGCTTTTTACGA	Msat v1	$(GA)_5$	2	2694	553	2	1.507	0.3258	0.3367	Ascunce et al. 2009
	C216_PigTail	NED-GGTTCGCTTTCTTGTCCCCG	GCTTCTTCGCTTTCATCGTGTCCGTG	Msat v4	$(AG)_3TCTT(AG)_{14}$	2	2648	549	4	3.693	0.7036	0.7299	Ascunce et al. 2009
	C234	FAM-GCGCAGTAGCATTCCGATAG	TGGGAGAGAAGGAAGTGAACA	Msat v1	$(GTC)_7(CGG)_6$	3	6564	544	8	4.200	0.7784	0.7627	Ascunce et al. 2009
	C259	PET-TCTCGGAGACAGACGATGAA	GCTAAGGCAGCATTACTCTCGT	Msat v1	$(CA)_{21}$	2	3126	525	3	1.960	0.5131	0.4903	Ascunce et al. 2009
	C278_PT	VIC-CTCACTCACTCACTCACGCTTT	GCTTCTCGGGATTTGTTGGGATAGAG	Msat v1	$(CA)_{27}(CT)_3$	2	5901	543	2	1.430	0.2802	0.3009	Ascunce et al. 2009
	C294	FAM-CGTCCTCGTTCATATTTTCAG	GGGAGAGGCTTCGAGGATAG	Msat v4	$(CT)_5AT(CT)_9$	2	3404	552	7	3.027	0.5837	0.6703	Ascunce et al. 2009
	C316	VIC-TACTTGCGGCGGAGTTTAAT	TTATAGTGTTGCTCGCGTGTC	Msat v1	$(TA)_3(CA)_8CC(CA)_9$	2	6788	532	8	2.109	0.5187	0.5264	Ascunce et al. 2009
	C334	VIC-TTTCTTTCTTTCTGTCTCTTTCTCG	ACGGAAGGCACGAATGAAGTC	Msat v1	$(TC)_{14}$	2	6514	539	2	1.278	0.1981	0.2179	Ascunce et al. 2009
	C367	CCCGAACAATGAGACATCCT	PET-TTAAATCTTAGCGCCGACGA	Msat v1	$(GA)_7$	2	5901	539	5	1.381	0.2777	0.2762	Ascunce et al. 2009
	C368_PigTail	FAM-TCCCTCCTCTTAAGCCTATCC	GCTTCTACGATTCGATGCTCGCTTAT	Msat v1	$(GT)_{12}(GC)_7(GT)_{13}$	2	8991	530	6	1.554	0.3531	0.3567	Ascunce et al. 2009
	C485	ATAGCGGGAATTGCAGGTCA	NED-TCGGAAGGACTGAAGGAGTG	Msat v1	$(CA)_{19}$	2	6843	558	7	4.195	0.7556	0.7624	Ascunce et al. 2009
	C487_PigTail	FAM-CCACGAGATGGAAGATGAGTTTACG	GCTTCTCCTTTTGTCTGGCGAGGGAC	Msat v1	$(CT)_4A(CT)_4$	2	—	452	4	1.808	0.4253	0.4476	Ascunce et al. 2009
	C536	PET-GAACGTGACGCGGAATAAA	GAAACGGTAGGCCAAGTCC	Msat v1	$(GA)_6$	2	6790	527	12	5.668	0.8393	0.8244	Ascunce et al. 2009
	Cactus_tree	FAM-AAGTACGATGACGACGATGAAAA	GCTTCTTTCGGACGATGATATATTTTGAGA	Msat v1	$(GAT)_8$	3	3190	510	2	1.944	0.4481	0.4862	Present study
	Carey	NED-TTGAGATCATCGAGATTGTAGCA	GCTTCTCACTTGCGTTACTACACGTCACT	Msat v4	$(AG)_8$	2	9376	542	2	1.814	0.4826	0.4491	Present study
	Carnival	PET-GCCCGCTACTTATAGAATCGAAA	GCTTCTACCTTTGTTTATCGGTGGGATT	Msat v4	$(AG)_8$	2	805	535	3	1.870	0.4813	0.4657	Present study
	Cool_water	PET-TTAGGTAAGAGAGAGACGCGAAGT	GCTTCTCTTTCTCTTCGATGTTCTTGCAT	Msat v1	$(AG)_8$	2	2127	543	2	1.389	0.2596	0.2802	Present study
	Coyote	FAM-CTAAAGCGATAACACCAATCGTT	GCTTCTTCATTTGTTTGTATCAACACGGTA	Msat v4	$(AG)_8$	2	9348	519	2	1.530	0.3700	0.3469	Present study
	Darwi_Odrade	PET-GACGTCCATTGAACCTCAATTT	GCTTCTGTTTGTATCAACTTATGCCACTAAA	Msat v4	$(AT)_7$	2	6126	319	12	4.754	0.6885	0.7910	Present study
	Daytripper	NED-GCGCGATACAGAGAAAGATAGAA	GCTTCTCAGAGTTTTGGGATCAGTGGTAA	Msat v4	$(AAT)_7$	3	4077	555	2	1.667	0.3943	0.4004	Present study
	Deadpool	PET-AATGCGAAGAACCATCAAGATT	GCTTCTCTCGAATCACTCGATCTCTTCA	Msat v1	$(AT)_8$	2	1122	396	2	1.177	0.1325	0.1504	Present study
	Diamond_Joe	CCAGACAGCGAAGATAAACGTA	GCTTCTCGGGCTAGTTGTGTATCCCTAC	Msat v4	$(AG)_6$	2	—	522	2	1.379	0.2646	0.2754	Present study
	Emma_Frost	VIC-GCGTAACGATACAACATCAACG	GCTTCTGCCGTATGCCCGTGTAAA	Msat v4	$(AT)_6$	2	4103	513	2	1.604	0.3483	0.3769	Present study
	Glass_onion_v2	FAM-ACTAGTCGAAGGTATCGGGAGAT	GCTTCTGCAACGCTGTATTCTCTCACTCT	Msat v1	$(AG)_8$	2	2940	548	3	1.572	0.3659	0.3642	Present study
	Handy_dandy	NED-CTCGCAATCTCACCTCGTAATC	GCTTCTGCGAGGAGCTTCATTGAATTAT	Msat v4	$(AG)_6$	2	2551	505	2	1.725	0.3846	0.4208	Present study
	Harvest	NED-CGTGATTTTACGATTTCATCAGC	GCTTCTCTGGTCCGGATGGGTTTT	Msat v4	$(ACC)_8$	3	779	549	3	2.510	0.6183	0.6022	Present study
	Heartland	VIC-TTTCCTGAATTCACCGAGACTAA	GCTTCTAATTAATGGATTACAGACTCGGTTC	Msat v4	$(AG)_7$	2	1203	513	3	2.082	0.5082	0.5202	Present study
	i-109	NED-TACGTCCATACCGGGTTCTTTAT	GTAAGGATAACGGGCTGTTGTTT	Msat v4	$(GT)_5GA(GT)_4AT(GT)_{13}$	2	6340	549	4	3.494	0.7094	0.7145	Ascunce et al. 2011
	i-113	6FAM-ACACCGCCGACTACACGATAA	CGGTCAAAGATGAAATAACAACGA	Msat v4	$(CT)_{12}$	2	4863	528	6	3.136	0.6654	0.6818	Ascunce et al. 2011
	i-114	NED-GTCGAAACGATAGGACGGAAAT	GTTCTCACCGTGGCGAAG	Msat v1	$(AG)_{10}$	2	5712	491	8	2.750	0.6118	0.6371	Ascunce et al. 2011
	i-120_PigTail	PET-ATTCTCGCCACGTTCGATT	GCTTCTACGAAAGCGTTTCGATATTAGC	Msat v4	$(AG)_{13}$	2	1571	478	8	3.275	0.6304	0.6954	Ascunce et al. 2011
	i-126	NED-AGAAGCCCTCGACCGACTAAC	GGATTCGATCGGCAATAACAAC	Msat v1	$(ACT)_6AGT(ACT)_{12}$	3	1466	515	8	2.989	0.6531	0.6662	Ascunce et al. 2011
	i-127	GCCTCTTCTTACGCTTTCTATTCG	6FAM-ACAACTGTACCTTTGCACACGACT	Msat v4	$(AAC)_9$	3	1533	509	4	2.060	0.5020	0.5150	Ascunce et al. 2011
	i-132	NED-TATAACGTCCGTAGGCGATAACAG	GCCTGACTAAGAACTTGTGTGAGC	Msat v4	$(AC)_{10}$	2	335	491	3	2.754	0.5874	0.6376	Ascunce et al. 2011
	i-135	VIC-CAGTTCGACGTAGTGGCGTTTAT	ACTACGACGACGACGTTCAGC	Msat v4	$(GT)_7$	2	1722	545	3	1.684	0.3920	0.4066	Ascunce et al. 2011
	i129	PET-GCGAGTGTGAAACACAACGTG	GCGCGAGAGAGGAAATTATCACTA	Msat v4	$(ACGC)_7$	4	2974	548	3	1.455	0.2983	0.3131	Ascunce et al. 2011
	Imagine	NED-GTTTATTACGGCAAACGATGCAC	GCTTCTCTCCGCGGCACCCTTACT	Msat v4	$(ACG)_6$	3	—	554	2	1.057	0.0549	0.0535	Present study
	Jackstraw_PigTail	PET-AAGCTGGAAATGAGAAAAGGAA	GCTTCTTGTTCTAAGAGAGTCTTCTCGGAGT	Msat v4	$(GA)_{12}$	2	335	547	7	2.535	0.6000	0.6061	Ascunce et al. 2011
	Jam_session	VIC-CAATTAAAGCGAGCGGTAAAATC	GCTTCTTGGCTTGGAGTTTGTAAATTTGT	Msat v4	$(CGT)_{10}$	3	2797	521	3	1.691	0.4315	0.4090	Present study
	Jericho	VIC-TGTGAAAGTTTGTTTCCATTTCA	GCTTCTGAGGATAGAGAACCGATGGTCTT	Msat v1	$(CT)_7$	2	1235	444	5	3.031	0.6787	0.6709	Present study
	Jerry-Garcia	NED-GTTGTCGACCACCCGTTG	AAACTTGTTCCGGCTTCTCAAT	Msat v1	$(CGG)_9$	3	8068	537	5	2.383	0.5558	0.5809	Ascunce et al. 2011
	Jokerman	NED-AGAACTATTAAAGATGACGAGTTCTGA	GCTTCTAAAATATACAAGTTCCGTGATGACAA	Msat v4	$(AT)_6$	2	6514	409	5	3.157	0.6580	0.6842	Present study
	Jude	FAM-AAAATCCTGCACACGACAA	GCTTCTAATTCCACGATCAATCGAATAC	Msat v1	$(ACG)_6$	3	5747	545	2	1.620	0.3346	0.3829	Present study
	Julia	PET-CGAGGTACATATGGCACTAACG	GCTTCTTTCCTCGGAGAGGGTTGC	Msat v4	$(ACG)_6$	3	2797	524	2	1.548	0.3635	0.3545	Present study

	Kitty_Pryde	PET-GCACGGATACTTTCGATGTGTA	GCTTCTCACAAACAGTCACCTCAGAAGC	Msat v4	$(AT)_6$	2	6050	530	3	1.676	0.3917	0.4037	Present study
	Maggie_Mae	FAM-GCTGCCGCTGAATAACGAT	GCTTCTACGTCGGGGAGTAGGTTACG	Msat v4	$(ACG)_6$	3	7090	521	3	1.215	0.1710	0.1771	Present study
	Majesty	NED-CACGTATCACGTAGCGCACATA	GCTTCTGACGCTACGCAACGCAAC	Msat v1	$(ACG)_6$	3	6735	483	2	1.789	0.4249	0.4415	Present study
	Miles_teg	PET-CATCCTTTGTACATCATTCTAACAG	GCTTCTTCTTTGAAGTTAATCAGTATGTTACTGTG	Msat v1	$(AT)_6$	2	6002	420	2	1.798	0.3550	0.4444	Present study
	Million_miles	NED-GGAGTGGAAAATATGACGGAAG	GCTTCTGGTCGTCTGCTTTATTCGTTATG	Msat v1	$(AT)_6$	2	805	456	3	2.096	0.5228	0.5236	Present study
	Monkey	NED-AACCTCCGAGTATCTCTGTCCA	GCTTCTCAGCTCGATAGTAGCGGTAGAAG	Msat v1	$(ACT)_8$	3	715	539	3	1.279	0.2171	0.2185	Present study
	Mozambique	PET-AAAACACGTTAAAATCAATTTCGTG	GCTTCTTAAATAGAACGTTCGTGGCAAA	Msat v1	$(AT)_7$	2	874	471	3	2.127	0.5055	0.5305	Present study
	Neil_Young	VIC-GTAACACCTAAAGCCGAAAGCTC	GCTTCTTTAAGTGCCTCAAAATGTGTTGT	Msat v4	$(AAT)_8$	3	8068	544	3	2.360	0.5307	0.5768	Present study
	Nettie_Moore	FAM-GCAACGCGTAAATAATTGATGA	GCTTCTCTTATCGTCAATCTTGCTGCAC	Msat v4	$(AT)_7$	2	4103	489	5	3.499	0.6829	0.7150	Present study
	Pam	PET-CAGATTCTCACGATGAACCTTTG	GCTTCTTTTAGGTCATTGTTGAGGCTACAT	Msat v4	$(ATC)_{11}$	3	5118	554	4	2.326	0.5777	0.5707	Present study
	Percy	NED-TGAAACGCTCGAATTACACTCA	GCTTCTCCCATCACCCTTCTCTATTCAT	Msat v4	$(AT)_7$	2	10535	418	2	1.900	0.5064	0.4744	Present study
	Piggies	FAM-CAAAAGGCGAAGAAAATAATTGA	GCTTCTTTACTCGGATATTGTGGTTTGCT	Msat v1	$(ATT)_6$	3	1426	518	2	1.964	0.4518	0.4914	Present study
	Psylocke	VIC-TGTATAAGCGCTGCTACTGGTT	GCTTCTTTAGCAATCCTTCCTTCACAAA	Msat v1	$(AT)_7$	2	6792	472	4	3.092	0.6200	0.6773	Present study
	Queen_Jane	PET-GTGTACAACCGATCAAAACGAT	GCTTCTAAACGCCAAAAGAATGAGTCAC	Msat v4	$(AT)_7$	2	5577	499	2	1.952	0.3452	0.4882	Present study
	Ramble	VIC-TTCAACGTGTACGCTATTTACAA	GCTTCTATACGCATCTGCTAGTGCTGAG	Msat v4	$(AG)_7$	2	468	511	2	1.268	0.2156	0.2116	Present study
	Red_ant	6FAM-TCGTGTCGTCGTCATCATCA	CAGCGGCGGCGACTAAAC	Msat v4	$(CGT)_9$	3	5901	556	5	1.561	0.3604	0.3599	Ascunce et al. 2011
	Sam_Stonev2	VIC-TGTGGTACACGTGTTTACAAAGG	GCTTCTTGCGATGTAACAGAAAGAGAGAG	Msat v4	$(CT)_8$	2	2742	386	3	2.855	0.6278	0.6507	Present study
	Seastones	NED-CAAAGCAGTTGCTCGCATAATAA	ATCGTCCCTGCCCTAAACACT	Msat v4	$(CTT)_6$	3	10511	536	2	1.414	0.2868	0.2933	Ascunce et al. 2011
	Sinv-25	PET-GTGGGATCGGGTGAAGTG	ACGCGTTTCAACAGTAGCG	Msat v1	$(GA)_{10}$	2	2162	547	3	1.410	0.3027	0.2910	Ascunce et al. 2011
	Sinv12	6FAM-ACGCCAGCGAAGTGACTAAT	CAGCAATTTGCCGAGATGTA	Msat v1	$(TA)_{12}$	2	1104	553	5	1.049	0.0361	0.0466	Ascunce et al. 2011
	Siona	VIC-AGTAAGCGAACGGCTCTCAAC	GCTTCTCTGATCGCGATTGCAGGAT	Msat v4	$(AT)_6$	2	779	501	2	1.771	0.4181	0.4359	Present study
	Sol-11	NED-ACTGGAGCCTCCGAGACC	CACTCCGGAAGAGTAACTTTGC	Msat v4	$(TC)_{15}$	2	10455	534	5	3.947	0.7539	0.7474	Ascunce et al. 2011
	Sol-18	VIC-CGTTTGCGTTGTTTTGGC	ACCGTCTCCCTCTTTTCTG	Msat v1	$(TC)_{11}$	2	7919	538	3	1.649	0.4035	0.3941	Ascunce et al. 2011
	Sol-20	6FAM-GACTTCCCTACTTTGTCTCTCTCC	AGCATGAAAAATCGGGAGC	Msat v1	$(TC)_{13}$	2	6514	551	13	3.012	0.6571	0.6687	Ascunce et al. 2011
	Sol-42f	VIC-ATGCCGGTTTTTATTGGTAGG	GGAATTTCCGTCGCCATT	Msat v1	$(TC)_{26}$	2	3712	550	10	3.766	0.7424	0.7352	Ascunce et al. 2011
	Sol-49	6FAM-GTCATCGGTGACCCACAAG	GTATCTCGACACGAAACTCGG	Msat v1	$(TC)_{19}$	2	6143	516	5	4.200	0.7490	0.7627	Ascunce et al. 2011
	Sol-6	PET-TTTACAGCGAATGGAACACG	CATTAAGTGCAATTATGCTCGC	Msat v1	$(TC)_4TT(TC)_8$	2	6243	551	4	2.017	0.5000	0.5047	Ascunce et al. 2011
	Sol-J1	NED-AAAGTGCGTTCCATTATCTGAC	GAAAAAGTGTTTGCCCACG	Msat v4	$(CA)_5TA(CA)_{13}$	2	1968	503	6	2.495	0.5735	0.5998	Ascunce et al. 2011
	Sol-M2	PET-CAATTACGTCTCGGTTATCGACTC	GGGTTGATCGCAAAACGATTG	Msat v1	$(TG)_{17}$	2	1565	530	5	2.827	0.6008	0.6468	Ascunce et al. 2011
	Sol-M3	PET-ATTGTACGAGCGGGAGAAGCAC	TCGAGATCGACCAGGTAGGAACAC	Msat v1	$(AG)_9$	2	3126	532	3	2.369	0.5933	0.5784	Ascunce et al. 2011
	St_Stephen	PET-CTGCCTGCAAATTTGAACGATT	ATTATTAGAGCGGCCTCCCATT	Msat v4	$(AAT)_9$	3	5901	512	3	1.066	0.0598	0.0622	Ascunce et al. 2011
	Starr	FAM-GAACCCCAGTCACGGTCTC	GCTTCTTATGTGTATGTCTCGCTCCGAAT	Msat v4	$(AC)_7$	2	5431	555	2	1.242	0.1887	0.1951	Present study
	Sun_King	PET-AAGTTTCCATTGGTCGGTGA	GCTTCTCTGCCGCTGAATAACGATAAC	Msat v4	$(CGT)_6$	3	7090	417	2	1.194	0.1538	0.1629	Present study
	Sunrise	NED-AACAGAATTGGGTATTGAATGT	TTCCTGATTGTCTTCACTAATCTCA	Msat v4	$(AAG)_8$	3	1122	559	3	2.058	0.4991	0.5146	Ascunce et al. 2011
	Sway	NED-AGCTAACCTCACAGCATCAAA	GCTTCTTGTATGCCGGTAAATTCTTGA	Msat v4	$(AT)_8$	2	6843	530	3	2.655	0.6230	0.6240	Present study
	Taxman	VIC-TTGATACAGCGCAATGTACCTCT	GCTTCTCGTTGTCTTCAGCGTGTGA	Msat v4	$(CGT)_8$	3	9607	535	4	2.591	0.5918	0.6146	Present study
	Tombstone_blues	FAM-CGCATACATTCGCACACATAC	GCTTCTAGATCAATTTCTAAGGCAGGTTT	Msat v4	$(CT)_6$	2	2974	523	2	1.600	0.3560	0.3754	Present study
	Trouble	PET-CGTCGGTCCATAAAGTCGTT	GCTTCTGTGACAGAGCTGTCGTCCTC	Msat v4	$(CT)_7$	2	319	464	2	1.995	0.4920	0.4993	Present study
	Tweedy	PET-TTGATCATCATCTTTATTATTGTTGG	GCTTCTCGTTACTTCGATTGTCCCTAAAAC	Msat v1	$(AGT)_9$	3	3776	475	3	1.955	0.4525	0.4889	Present study
	Walrus	VIC-GTTAACCCACGATCCGGTATT	GCTTCTAGTCAATCGTAATCCTCCTGAGC	Msat v4	$(CTT)_6$	3	1858	522	3	1.615	0.3909	0.3812	Present study
	Weight	FAM-GCAACTCTTGCAGGGTTTGTAT	GCTTCTGAGAGAGAGAATGTCACGAAAGC	Msat v1	$(CT)_7$	2	10535	506	2	1.243	0.1784	0.1959	Present study
	Wigwam	FAM-CGTGGCATGATACCTAAGCTAGT	GCTTCTCGGTCACATGTAGGTAGAGCAG	Msat v1	$(GT)_6$	2	2797	387	2	1.390	0.2676	0.2812	Present study
	Workingman_Blues	PET-CACACATCGACCATGAAGAGTT	GCTTCTGGAGCAATGGATAAAGGACTTG	Msat v4	$(GT)_7$	2	2797	520	3	2.487	0.4758	0.5986	Present study
	Yellow_submarine	FAM-CGATACTTTCGGAAACAAGAGG	GCTTCTAATTATTGAGAAGCTCGTAAAACGTA	Msat v4	$(CTT)_6$	3	4103	534	2	1.613	0.3569	0.3802	Present study
	Mean								3.879	2.164	0.4520	0.4677	
Allozymes													
	Aat-2	NA	NA	NA	NA	NA	NA	555	2	1.298	0.2190	0.2299	Shoemaker et al. 1992
	Acoh-1	NA	NA	NA	NA	NA	NA	249	2	1.204	0.1634	0.1697	Shoemaker et al. 1992
	Acoh-5	NA	NA	NA	NA	NA	NA	531	2	1.536	0.3546	0.3492	Shoemaker et al. 1992
	Acy1	NA	NA	NA	NA	NA	NA	549	2	1.190	0.1553	0.1596	Shoemaker et al. 1992
	G3pdh-1	NA	NA	NA	NA	NA	NA	529	2	1.954	0.4761	0.4887	Shoemaker et al. 1992
	Pgm-1	NA	NA	NA	NA	NA	NA	546	2	1.344	0.2354	0.2563	Shoemaker et al. 1992
	Mean								2.000	1.421	0.2673	0.2756	

[1]Microsatellite primers are reported with the fluorescent label at the 5' end of either the forward or reverse primer.

[2]See Ascunce et al. (2009).

[3]Number of the scaffold with best match in a BLASTn query including both forward and reverse primer sequences against current *Solenopsis invicta* genome assembly (see Wurm et al. 2009). Dash indicates that the primers were not detected on any scaffold.

[4]Effective number of alleles was calculated using the equation $n_e=1/\Sigma(p_i^2)$ (Frankham et al. 2010).

[5]H_{obs} and H_{exp} were calculated using *GenePop* (v.4.0.10).

Appendix A.2. Allele frequencies for each locus in the marker panel.

Locus	Allele	Frequency
Aat-2	*100*	0.8679
	144	0.1321
Acoh-1	*82*	0.0923
	100	0.9077
Acoh-5	*93*	0.2247
	100	0.7753
Acy1	*93*	0.0882
	100	0.9118
Anna	*69*	0.1307
	73	0.5446
	77	0.3248
Baez	*151*	0.4035
	157	0.5965
Baron_Harkonnen	*310*	0.9443
	311	0.0557
Beatles	*116*	0.9924
	124	0.0076
Bertha	*200*	0.1712
	206	0.5730
	212	0.2490
	214	0.0019
	225	0.0049
Black_Crow	*161*	0.2644
	165	0.1610
	167	0.1700
	169	0.4026
	171	0.0020
Blackbird	*97*	0.8867
	100	0.1133
Blue_Jay	*93*	0.1774
	101	0.0774
	104	0.7453
C1	*273*	0.0673
	274	0.0051
	275	0.6429
	277	0.2480
	284	0.0357
	288	0.0010
C147	*67*	0.5366
	69	0.0319
	71	0.0009
	73	0.4306
C204	*163*	0.2718
	167	0.0600
	171	0.6538
	173	0.0010

	194	0.0010
	202	0.0010
	204	0.0116
C21	*81*	0.7845
	83	0.2155
C216_PigTail	*203*	0.2872
	205	0.1927
	209	0.1698
	211	0.3502
C234	*95*	0.0019
	109	0.0010
	118	0.1885
	124	0.1683
	133	0.0712
	136	0.1962
	139	0.3615
	142	0.0115
C259	*184*	0.5743
	186	0.4247
	188	0.0010
C278_PT	*112*	0.8161
	114	0.1839
C294	*96*	0.0123
	100	0.2770
	106	0.4573
	108	0.0446
	110	0.0009
	112	0.2068
	114	0.0009
C316	*205*	0.1703
	207	0.6427
	211	0.0010
	213	0.0010
	232	0.0010
	234	0.1781
	238	0.0010
	250	0.0049
C334	*189*	0.8757
	191	0.1243
C367	*228*	0.0233
	230	0.8398
	232	0.1350
	233	0.0010
	236	0.0010
C368_PigTail	*188*	0.7923
	190	0.0039
	194	0.1073
	196	0.0335

	209	0.0620
	212	0.0010
C485	*101*	0.3593
	114	0.1144
	116	0.2345
	120	0.0019
	124	0.0009
	132	0.1435
	134	0.1426
	137	0.0028
C487_PigTail	*320*	0.7236
	322	0.1239
	326	0.0367
	328	0.1158
C536	*87*	0.1752
	91	0.0040
	93	0.2327
	95	0.1020
	99	0.1168
	101	0.0129
	103	0.1149
	105	0.0030
	107	0.2327
	111	0.0010
	113	0.0040
	117	0.0010
Cactus_tree	*140*	0.5843
	143	0.4157
Carey	*103*	0.3401
	105	0.6599
Carnival	*129*	0.6801
	132	0.0581
	139	0.2618
Cool_water	*144*	0.1683
	146	0.8317
Coyote	*163*	0.7774
	165	0.2226
Darwi_Odrade	*333*	0.1781
	335	0.0016
	339	0.1340
	341	0.0997
	343	0.0016
	344	0.0033
	345	0.0082
	347	0.3611
	350	0.1193
	352	0.0082
	354	0.0768

	356	0.0082
Daytripper	*100*	0.7241
	103	0.2759
Deadpool	*439*	0.9182
	441	0.0818
Diamond_Joe	*297*	0.1643
	298	0.8357
Emma_Frost	*308*	0.2520
	311	0.7480
G3pdh-1	*40*	0.4225
	100	0.5775
Glass_onion_v2	*161*	0.7677
	163	0.2170
	165	0.0153
Handy_dandy	*255*	0.7002
	257	0.2998
Harvest	*151*	0.3705
	154	0.4924
	157	0.1371
Heartland	*245*	0.5368
	246	0.0256
	250	0.4376
i-109	*145*	0.2634
	147	0.3998
	149	0.1536
	154	0.1832
i-113	*159*	0.0432
	161	0.0010
	163	0.3762
	165	0.1562
	167	0.0373
	169	0.3861
i-114	*297*	0.0011
	301	0.5579
	303	0.1421
	305	0.0011
	307	0.1516
	309	0.0800
	313	0.0168
	317	0.0495
i-120_PigTail	*314*	0.2863
	318	0.4555
	320	0.0803
	324	0.0076
	326	0.0575
	328	0.0640
	330	0.0477
	332	0.0011

i-126	*206*	0.0010
	209	0.2318
	212	0.0010
	218	0.4109
	221	0.3340
	224	0.0182
	230	0.0010
	242	0.0020
i-127	*217*	0.0051
	220	0.0784
	223	0.6314
	226	0.2851
i-132	*240*	0.4044
	242	0.1922
	244	0.4034
i-135	*91*	0.7195
	93	0.2767
	99	0.0038
i129	*146*	0.1279
	150	0.0563
	154	0.8158
Imagine	*109*	0.9726
	116	0.0274
Jackstraw_PigTail	*115*	0.0010
	117	0.1891
	119	0.5470
	121	0.0010
	127	0.0202
	129	0.2409
	133	0.0010
Jam_session	*238*	0.1338
	244	0.7475
	253	0.1187
Jericho	*158*	0.4229
	162	0.0048
	176	0.3578
	178	0.1157
	180	0.0988
Jerry-Garcia	*213*	0.1162
	216	0.0020
	219	0.5361
	223	0.0010
	224	0.3447
Jokerman	*241*	0.0801
	243	0.1886
	245	0.4858
	247	0.1886
	251	0.0568

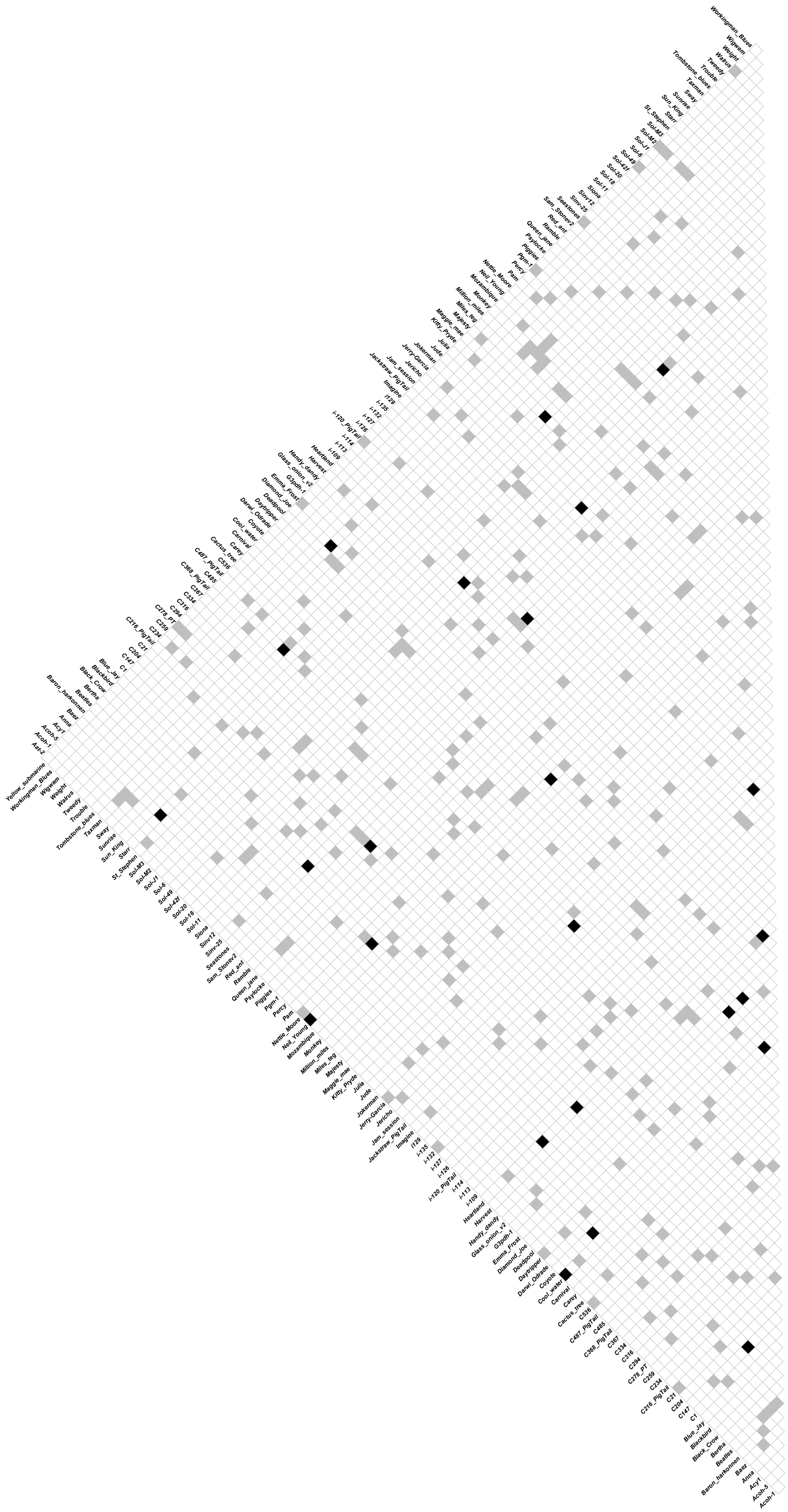
Jude	*68*	0.7428
	71	0.2572
Julia	*116*	0.7695
	119	0.2305
Kitty_Pryde	*167*	0.7299
	169	0.2534
	171	0.0167
Maggie_Mae	*81*	0.9026
	84	0.0040
	88	0.0934
Majesty	*91*	0.3276
	94	0.6724
Miles_teg	*331*	0.3317
	333	0.6683
Million_miles	*376*	0.2654
	378	0.6287
	380	0.1059
Monkey	*107*	0.0735
	109	0.8791
	118	0.0474
Mozambique	*246*	0.6162
	248	0.2851
	254	0.0987
Neil_Young	*145*	0.0880
	148	0.4665
	151	0.4455
Nettie_Moore	*258*	0.0549
	260	0.1203
	261	0.2036
	264	0.4420
	270	0.1793
Pam	*82*	0.1834
	85	0.1049
	88	0.6134
	91	0.0983
Percy	*244*	0.3845
	246	0.6155
Pgm-1	*96*	0.1505
	100	0.8495
Piggies	*160*	0.5671
	163	0.4329
Psylocke	*325*	0.0511
	327	0.3733
	329	0.3767
	331	0.1989
Queen_Jane	*255*	0.5785
	257	0.4215
Ramble	*246*	0.1199

	248	0.8801
Red_ant	*79*	0.0019
	90	0.0019
	92	0.2288
	95	0.7665
	98	0.0009
Sam_Stonev2	*159*	0.4134
	163	0.2330
	167	0.3537
Seastones	*84*	0.8221
	91	0.1779
Sinv-25	*155*	0.0010
	157	0.1750
	161	0.8241
Sinv12	*120*	0.9763
	121	0.0038
	122	0.0019
	126	0.0170
	132	0.0009
Siona	*311*	0.6792
	313	0.3208
Sol-11	*141*	0.3538
	143	0.0819
	145	0.1228
	149	0.2865
	153	0.1550
Sol-18	*230*	0.7374
	232	0.2490
	234	0.0136
Sol-20	*103*	0.1302
	105	0.5295
	107	0.1055
	109	0.0143
	113	0.0010
	115	0.0010
	121	0.0038
	122	0.0010
	123	0.1245
	125	0.0019
	129	0.0019
	131	0.0827
	133	0.0029
Sol-42f	*112*	0.0019
	114	0.4019
	116	0.2267
	118	0.1324
	122	0.0019
	126	0.0029

	128	0.1800
	130	0.0029
	138	0.0029
	140	0.0467
Sol-49	*137*	0.1446
	144	0.1507
	156	0.3717
	158	0.1446
	162	0.1884
Sol-6	*107*	0.1290
	109	0.0019
	111	0.6594
	113	0.2097
Sol-J1	*248*	0.0010
	253	0.0041
	254	0.5372
	256	0.1591
	258	0.2944
	260	0.0041
Sol-M2	*276*	0.3205
	278	0.4813
	280	0.0828
	296	0.1144
	297	0.0010
Sol-M3	*204*	0.4676
	206	0.0902
	208	0.4422
Starr	*85*	0.8898
	88	0.1102
St_Stephen	*254*	0.0051
	257	0.9681
	260	0.0267
Sun_King	*150*	0.9109
	156	0.0891
Sunrise	*79*	0.5777
	82	0.0337
	85	0.3886
Sway	*82*	0.5020
	86	0.2673
	90	0.2307
Taxman	*188*	0.3662
	191	0.4775
	194	0.1543
	198	0.0020
Tombstone_blues	*242*	0.7485
	245	0.2515
Trouble	*247*	0.5240
	249	0.4760

Tweedy	*222*	0.0044
	227	0.4075
	231	0.5881
Walrus	*165*	0.2248
	168	0.7535
	171	0.0218
Weight	*310*	0.8903
	312	0.1097
Wigwam	*240*	0.1689
	243	0.8311
Workingman_Blues	*254*	0.1590
	256	0.5282
	257	0.3129
Yellow_submarine	*273*	0.7456
	276	0.2544

Appendix A.3. Results of pairwise composite linkage disequilibrium (LD) tests. Light grey cells reflect tests significant at the $P = 0.05$ level, while black cells reflect tests that remain significant after correction for multiple comparisons (Benjamini-Hochberg procedure, false discovery rate [FDR] of 5%).



Appendix A.4. Effect sizes (above diagonal) and *P*-values (below diagonal) for associations between fitness-related traits (excluding viral infection) analyzed using non-parametric statistical tests. Comparisons between traits that are effectively continuous (*QnSize*, *QnWt*, $QnWtLs_{CL}$, *ClaustPer*, $ColWt_{CL}$, $ColWt_{CL+6}$, and $ColGrowth_{CL-6}$) were made using Spearman's rank correlation, with the correlation coefficients (*r*) as effect sizes. Comparisons between binary traits ($QnSurv_{CL}$ and $QnSurv_{CL-6}$) and continuous traits were made using the Mann-Whitney *U* test, with effect sizes obtained by converting the *Z*-scores to Pearson *r* values (Coltman and Slate 2003). *P*-values ≤ 0.05 are shown in bold, while *P*-values significant after adjustment for multiple comparisons (Benjamini-Hochberg procedure, 5% FDR) are bold and underlined. NA = comparison not applicable.

	QnSize	*QnWt*	$QnWtLs_{CL}$	*ClaustPer*	$QnSurv_{CL}$	*ColWtCL*	$ColWt_{CL+6}$	$ColGrowth_{CL-6}$	$QnSurv_{CL-6}$
QnSize	—	0.1112	-0.0394	0.0498	0.1315	-0.0338	0.0883	0.0953	-0.0162
QnWt	**0.0247**	—	-0.2066	-0.2093	0.4447	0.5244	0.2683	0.1095	0.0947
$QnWtLs_{CL}$	0.5756	**<u>0.0014</u>**	—	-0.0141	NA	-0.5526	-0.1351	0.0310	-0.1475
ClaustPer	0.4863	**<u>0.0015</u>**	0.8318	—	NA	-0.4163	-0.3329	-0.2025	-0.3408
$QnSurv_{CL}$	0.2415	**<u><0.0001</u>**	NA	NA	—	NA	NA	NA	NA
ColWtCL	0.6311	**<u><0.0001</u>**	**<u><0.0001</u>**	**<u><0.0001</u>**	NA	—	0.2785	-0.0368	0.3762
$ColWt_{CL+6}$	0.4705	**0.0168**	0.2351	**0.0027**	NA	**0.0130**	—	0.8978	NA
$ColGrowth_{CL-6}$	0.4352	0.3368	0.7863	0.0735	NA	0.7474	**<u><0.0001</u>**	—	NA
$QnSurv_{CL-6}$	0.7532	**0.0268**	**0.0217**	**<u>0.0009</u>**	NA	**<u>0.0003</u>**	NA	NA	—

Signs of effect sizes reflect directions of associations; for the binary traits $QnSurv_{CL}$ and $QnSurv_{CL-6}$, positive effect sizes indicate that surviving queens had higher associated trait values than non-survivors.

Appendix A.5. Results of Fisher's exact tests for Hardy-Weinberg Equilibrium (HWE) and values of F_{IS} estimated for each marker locus. *P*-values and their standard errors are listed for the exact tests. F_{IS} values were calculated using two algorithms (WC=Weir and Cockerham 1984; RH=Robertson and Hill 1984). *P*-values ≤ 0.05 are shown in bold, while *P*-values significant after adjustment for multiple comparisons (Benjamini-Hochberg procedure, 5% FDR) are bold and underlined.

Locus	*P*-value	Standard error	F_{IS} (WC)	F_{IS} (RH)
Aat-2	0.2126	0.0048	0.0533	0.0533
Acoh-1	0.4469	0.0034	0.0439	0.0440
Acoh-5	1.0000	0.0000	-0.0039	-0.0039
Acy1	0.4093	0.0041	0.0304	0.0304
Anna	0.2981	0.0128	0.0292	0.0451
Baez	0.4742	0.0094	0.0342	0.0343
Baron_Harkonnen	0.0768	0.0023	0.0876	0.0877
Beatles	1.0000	0.0000	-0.0055	-0.0055
Bertha	0.4797	0.0237	-0.0329	-0.0139
Black_Crow	0.3020	0.0225	0.0646	0.0446
Blackbird	0.0639	0.0025	0.0806	0.0806
Blue_Jay	0.3152	0.0101	0.0009	-0.0099
C1	0.0803	0.0108	0.0439	0.0418
C147	**0.0224**	0.0046	0.0267	0.0588
C204	0.9355	0.0096	0.0360	0.0026
C21	0.5219	0.0064	0.0262	0.0262
C216_PigTail	0.4983	0.0175	0.0424	0.0345
C234	0.3856	0.0230	-0.0240	0.1257
C259	0.6064	0.0204	-0.0351	-0.0175
C278_PT	**0.0322**	0.0052	0.0759	0.0367
C294	**<u><0.0001</u>**	0.0000	0.1600	0.1538
C316	0.7944	0.0191	0.0075	0.0005
C334	**0.0065**	0.0008	0.1233	0.1234
C367	0.4428	0.0169	-0.0183	0.0081
C368_PigTail	0.8387	0.0121	0.0204	0.0045
C485	0.2016	0.0259	0.0050	0.0010
C487_PigTail	0.2059	0.0100	0.0488	0.0478
C536	0.4379	0.0325	-0.0126	-0.0042
Cactus_tree	0.1537	0.0078	0.0645	0.0646
Carey	0.0534	0.0043	-0.0822	-0.0823
Carnival	0.1574	0.0086	-0.0186	0.0170
Cool_water	**0.0404**	0.0029	0.0926	0.0927
Coyote	0.1356	0.0061	-0.0694	-0.0695
Darwi_Odrade	**0.0119**	0.0052	0.1230	0.0888

Daytripper	0.4581	0.0089	0.0328	0.0328
Deadpool	**0.0276**	0.0012	0.1218	0.1219
Diamond_Joe	0.7548	0.0039	0.0125	0.0125
Emma_Frost	0.1045	0.0052	0.0755	0.0756
G3pdh-1	0.5885	0.0086	0.0262	0.0262
Glass_onion_v2	0.5973	0.0094	0.0036	-0.0087
Handy_dandy	0.0817	0.0050	0.0798	0.0799
Harvest	0.7773	0.0106	-0.0360	-0.0326
Heartland	0.5818	0.0115	0.0243	0.0300
i-109	0.7183	0.0139	0.0067	0.0036
i-113	0.1494	0.0153	0.0295	0.0248
i-114	0.4926	0.0203	0.0212	0.0008
i-120_PigTail	0.3111	0.0204	0.0826	0.0463
i-126	**0.0102**	0.0051	0.0340	0.0212
i-127	0.0740	0.0109	0.0207	0.0265
i-132	**0.0460**	0.0062	0.0736	0.0749
i-135	0.7710	0.0082	0.0328	0.0154
i129	0.1383	0.0061	0.0446	0.0123
Imagine	1.0000	0.0000	-0.0288	-0.0289
Jackstraw_PigTail	0.2302	0.0227	0.0089	0.0012
Jam_session	0.3240	0.0086	-0.0624	-0.0551
Jericho	0.0989	0.0098	-0.0057	0.1050
Jerry-Garcia	**0.0015**	0.0008	0.0363	-0.0031
Jokerman	0.0623	0.0068	0.0440	0.0408
Jude	**0.0037**	0.0008	0.1317	0.1319
Julia	0.5332	0.0070	-0.0330	-0.0331
Kitty_Pryde	0.4237	0.0109	0.0543	0.0223
Maggie_Mae	0.6933	0.0071	0.0195	0.0104
Majesty	0.4526	0.0078	0.0351	0.0351
Miles_teg	**<0.0001**	0.0000	0.2133	0.2136
Million_miles	0.8560	0.0058	0.0172	0.0175
Monkey	0.2498	0.0073	0.0047	-0.0048
Mozambique	0.4596	0.0113	0.0507	0.0304
Neil_Young	0.0545	0.0066	0.0916	0.0729
Nettie_Moore	0.5851	0.0171	0.0427	0.0450
Pam	0.7479	0.0105	-0.0271	-0.0279
Percy	**0.0343**	0.0032	-0.1025	-0.1027
Pgm-1	0.1018	0.0038	0.0749	0.0749
Piggies	0.0986	0.0066	0.0756	0.0757
Psylocke	**0.0334**	0.0043	0.1074	0.0838
Queen_Jane	**<0.0001**	0.0000	0.3185	0.3189
Ramble	0.8325	0.0019	0.0120	0.0121
Red_ant	0.1448	0.0118	-0.0150	-0.0049
Sam_Stonev2	0.2295	0.0117	0.0341	0.0443
Seastones	0.2971	0.0064	0.0438	0.0439

Sinv-25	0.9014	0.0048	-0.0152	-0.0073
Sinv12	**<0.0001**	0.0000	0.2172	0.4967
Siona	0.1815	0.0074	0.0610	0.0610
Sol-11	0.8635	0.0094	-0.0087	-0.0037
Sol-18	0.1644	0.0080	-0.0058	0.0594
Sol-20	0.1366	0.0267	0.0251	0.0019
Sol-42f	0.3088	0.0272	-0.0072	0.0054
Sol-49	0.8790	0.0094	0.0259	0.0248
Sol-6	0.8308	0.0110	-0.0075	0.0026
Sol-J1	**0.0001**	0.0001	0.0562	0.2272
Sol-M2	0.3313	0.0211	0.0688	0.0359
Sol-M3	0.0747	0.0069	-0.0282	-0.0030
St_Stephen	0.3897	0.0068	0.0385	0.0259
Starr	0.3877	0.0047	0.0386	0.0386
Sun_King	0.5117	0.0027	0.0374	0.0374
Sunrise	0.4083	0.0126	0.0337	0.0052
Sway	0.7291	0.0118	0.0035	0.0053
Taxman	0.3837	0.0212	0.0488	0.0374
Tombstone_blues	0.3441	0.0082	0.0436	0.0436
Trouble	0.5695	0.0090	0.0284	0.0285
Tweedy	**0.0002**	0.0001	0.1112	0.3027
Walrus	0.3329	0.0099	-0.0482	-0.0387
Weight	0.0859	0.0032	0.0800	0.0801
Wigwam	0.2709	0.0055	0.0618	0.0618
Workingman_Blues	**<0.0001**	0.0000	0.2049	0.1712
Yellow_submarine	0.1147	0.0056	0.0709	0.0710

Appendix A.6. Regression models for analysis of HFAs (using HL metric), with relevant model information reported for each response variable (fitness-related trait). See Table 2.3 for additional information.

Response variable/trait	Regression coefficient	Standard error	Error distribution of GLM model	Type of score	Score	*P*-value	Sample size	Effect size	R^2
QnSize	-0.010	0.040	Gaussian	*t*	-0.25	0.800	454	-0.01	0.0002
QnWt	1.000	7.251	Gaussian	*t*	0.14	0.890	591	0.01	0.0001
$QnWtLs_{CL}$	0.043	0.085	Gaussian	*t*	0.50	0.615	235	0.03	0.0011
QnSINV-1	3.515	2.830	Binomial	*Z*	1.24	0.214	562	0.06	0.0035
QnSINV-2	0.658	1.704	Binomial	*Z*	0.39	0.700	562	0.03	0.0011
QnSINV-3	-2.454	6.208	Binomial	*Z*	-0.40	0.693	562	-0.01	0.0002
ClaustPer	2.450	0.809	Poisson	*Z*	3.03	**0.002**	228	0.15	0.0224
$QnSurv_{CL}$	9.555	3.243	Binomial	*Z*	2.95	**0.003**	282	0.13	0.0160
$ColWt_{CL}$	-37.400	27.810	Gaussian	*t*	-1.35	0.180	235	-0.09	0.0077
$ColWt_{CL+6}$	316.600	487.600	Gaussian	*t*	0.65	0.518	79	0.07	0.0054
$ColGrowth_{CL-6}$	2.054	0.759	Poisson	*Z*	2.71	**0.007**	79	0.25	0.0620
$QnSurv_{CL-6}$	6.514	5.991	Binomial	*Z*	1.09	0.277	92	0.06	0.0038

Appendix A.7. Regression models for analysis of HFAs (using IR metric), with relevant model information reported for each response variable (fitness-related trait). See Table 2.3 for additional information.

Response variable/trait	Regression coefficient	Standard error	Error distribution of GLM model	Type of score	Score	*P*-value	Sample size	Effect size	R^2
QnSize	-0.001	0.023	Gaussian	*t*	-0.03	0.977	454	-0.02	0.0005
QnWt	1.727	4.135	Gaussian	*t*	0.42	0.676	591	0.00	0.0000
$QnWtLs_{CL}$	0.025	0.048	Gaussian	*t*	0.52	0.602	235	0.03	0.0012
QnSINV-1	1.875	1.614	Binomial	*Z*	1.16	0.245	562	0.05	0.0025
QnSINV-2	0.426	0.973	Binomial	*Z*	0.44	0.661	562	0.03	0.0007
QnSINV-3	-3.651	3.747	Binomial	*Z*	-0.97	0.330	562	-0.04	0.0014
ClaustPer	1.075	0.454	Poisson	*Z*	2.37	**0.018**	228	0.12	0.0137
$QnSurv_{CL}$	5.310	1.844	Binomial	*Z*	2.88	**0.004**	282	0.12	0.0145
$ColWt_{CL}$	-16.460	15.620	Gaussian	*t*	-1.05	0.293	235	-0.07	0.0047
$ColWt_{CL+6}$	155.700	257.300	Gaussian	*t*	0.61	0.547	79	0.07	0.0047
$ColGrowth_{CL-6}$	1.031	0.397	Poisson	*Z*	2.60	**0.009**	79	0.24	0.0572
$QnSurv_{CL-6}$	4.306	3.354	Binomial	*Z*	1.28	0.199	92	0.07	0.0046

Appendix A.8. Regression models for analysis of HFAs (using SH metric), with relevant model information reported for each response variable (fitness-related trait). See Table 2.3 for additional information.

Response variable/trait	Regression coefficient	Standard error	Error distribution of GLM model	Type of score	Score	*P*-value	Sample size	Effect size	R^2
QnSize	0.005	0.020	Gaussian	*t*	0.24	0.812	454	0.02	0.0020
QnWt	-0.278	3.631	Gaussian	*t*	-0.08	0.939	591	-0.01	0.0000
$QnWtLs_{CL}$	-0.026	0.043	Gaussian	*t*	-0.62	0.537	235	-0.04	0.0020
QnSINV-1	-1.890	1.427	Binomial	*Z*	-1.33	0.185	562	-0.06	0.0030
QnSINV-2	-0.706	0.849	Binomial	*Z*	-0.83	0.406	562	-0.05	0.0010
QnSINV-3	1.640	3.105	Binomial	*Z*	0.53	0.597	562	0.02	0.0000
ClaustPer	-0.963	0.407	Poisson	*Z*	-2.37	**0.018**	228	-0.12	0.0140
$QnSurv_{CL}$	-4.958	1.634	Binomial	*Z*	-3.04	**0.002**	282	-0.13	0.0330
$ColWt_{CL}$	20.860	13.900	Gaussian	*t*	1.50	0.135	235	0.10	0.0100
$ColWt_{CL+6}$	-68.670	231.790	Gaussian	*t*	-0.30	0.768	79	-0.03	0.0010
$ColGrowth_{CL-6}$	-0.811	0.364	Poisson	*Z*	-2.23	**0.026**	79	-0.21	0.0420
$QnSurv_{CL-6}$	-3.248	2.883	Binomial	*Z*	-1.13	0.260	92	-0.05	0.0140

Appendix A.9. Regression models for analysis of HFAs (using MLH metric) for data sets excluding low-heterozygosity or high-heterozygosity queens, with relevant model information reported for each response variable (fitness-related trait). Only traits for which MLH was found to be a significant predictor in any of these models are reported. See Table 2.3 for additional information.

Data set	Response variable/trait	Regression coefficient	Standard error	Error distribution of GLM model	Type of score	Score	*P*-value	Sample size	Effect size	R^2
Lowest quartile observed MLH removed	$QnSurv_{CL}$	-13.915	5.122	Binomial	*Z*	-2.72	**0.007**	214	-0.19	0.0345
Lowest 5% simulated MLH removed	*QnSINV-1*	-8.163	3.874	Binomial	*Z*	-2.11	**0.035**	541	-0.09	0.0082
Lowest 5% simulated MLH removed	*ClaustPer*	-3.114	1.131	Poisson	*Z*	-2.75	**0.006**	218	-0.19	0.0348
Lowest 5% simulated MLH removed	$QnSurv_{CL}$	-10.740	3.847	Binomial	*Z*	-2.79	**0.005**	271	-0.17	0.0288
Highest quartile observed MLH removed	*QnSize*	0.229	0.113	Gaussian	*t*	2.03	**0.044**	252	0.13	0.0161
Highest quartile observed MLH removed	$ColGrowth_{CL\text{-}6}$	-16.634	8.116	Poisson	*t*	-2.05	**0.045**	62	-0.26	0.0654
Highest 5% simulated MLH removed	*ClaustPer*	-2.356	1.034	Poisson	*Z*	-2.28	**0.023**	206	-0.16	0.0252
Highest 5% simulated MLH removed	$QnSurv_{CL}$	-11.619	4.983	Binomial	*Z*	-2.33	**0.020**	251	-0.15	0.0217

Appendix A.10. Results of analysis of HFAs (using MLH metric) based on classification of queens into low-heterozygosity or high-heterozygosity categories. Membership in the two groups was assigned based on a visible break in the distribution of individual MLH values (see text). The *P*-value shown in bold represents the sole test for which the difference between the groups is significant at ≤ 0.05 before correction for multiple comparisons; the test was not significant after correction (Benjamini-Hochberg procedure, 5% FDR). NA=not applicable.

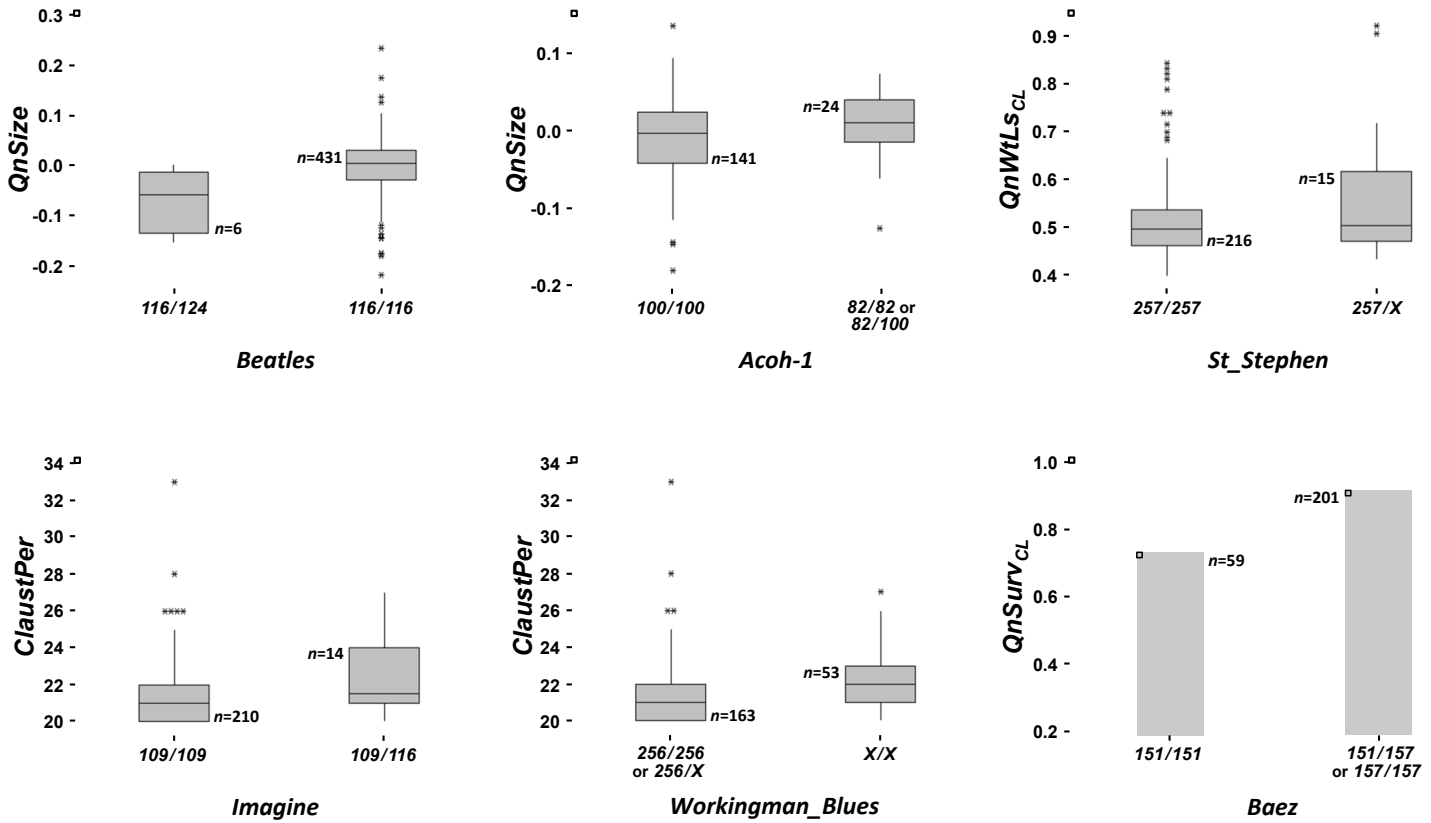
Trait	Test	*W*-score	*P*-value	Direction of association
QnSize	Mann-Whitney *U*	3258	0.4017	high-het < low-het
QnWt	Mann-Whitney *U*	6835	0.7062	high-het < low-het
$QnWtLs_{CL}$	Mann-Whitney *U*	1570	0.9273	high-het > low-het
QnSINV-1	Fisher's exact	NA	0.6268	high-het > low-het
QnSINV-2	Fisher's exact	NA	0.7831	high-het > low-het
QnSINV-3	Fisher's exact	NA	1.0000	high-het > low-het
ClaustPer	Mann-Whitney *U*	1750	0.2768	high-het < low-het
$QnSurv_{CL}$	Fisher's exact	NA	0.2297	high-het < low-het
$ColWt_{CL}$	Mann-Whitney *U*	1242	0.2162	high-het > low-het
$ColWt_{CL+6}$	Mann-Whitney *U*	237	0.3045	high-het < low-het

$ColGrowth_{CL\text{-}6}$	Mann-Whitney *U*	304	**0.0200**	high-het < low-het
$QnSurv_{CL\text{-}6}$	Fisher's exact	NA	0.5449	high-het > low-het

Appendix A.11. Results of analysis of HFAs (using MLH metric) based on classification of queens into low-heterozygosity or high-heterozygosity categories. Queens with MLH values in the lowest quartile of the MLH distribution were placed in the low-heterozygosity category, while queens with values in the highest quartile were placed in the high-heterozygosity category. The *P*-value shown in bold represents the sole test for which the difference between the groups is significant at ≤ 0.05 before correction for multiple comparisons; the test was not significant after correction (Benjamini-Hochberg procedure, 5% FDR). NA = not applicable.

Trait	Test	*W*-Score	*P*-value	Direction of association
QnSize	Mann-Whitney *U*	5602	0.4913	high-het > low-het
QnWt	Mann-Whitney *U*	10814	0.9896	high-het > low-het
$QnWtLs_{CL}$	Mann-Whitney *U*	1663	0.7938	high-het < low-het
QnSINV-1	Fisher's exact	NA	0.4424	high-het < low-het
QnSINV-2	Fisher's exact	NA	0.8812	high-het < low-het
QnSINV-3	Fisher's exact	NA	0.3637	high-het > low-het
ClaustPer	Mann-Whitney *U*	1606	0.5298	high-het < low-het
$QnSurv_{CL}$	Fisher's exact	NA	**0.0200**	high-het < low-het
$ColWt_{CL}$	Mann-Whitney *U*	1318.5	0.0909	high-het > low-het
$ColWt_{CL+6}$	Mann-Whitney *U*	168	0.9640	high-het > low-het

$ColGrowth_{CL\text{-}6}$	Mann-Whitney *U*	191	0.5366	high-het < low-het
$QnSurv_{CL\text{-}6}$	Fisher's exact	NA	0.3327	high-het < low-het



Appendix A.12. Differences in fitness-trait values between genotypes or pooled genotypes at single loci found to exhibit statistically significant (or marginally non-significant in the case of locus *St_Stephen* and trait $QnWtLs_{CL}$,) single locus genotype-trait associations using *GEPHAST* (SLAs). Distributions for continuous traits are shown with boxplots, whereas those for the binary trait $QnSurv_{CL}$ are shown with a bar graph. Locus names are indicated beneath the plots, with genotypes labeled on the x-axis (*X* denotes any allele other than those explicitly identified). Boxplots depict the first quartile of values as the bottom of the box, third quartile as the top, and median as a horizontal line. Whiskers encompass values that extend below the first and above the third quartile by amounts as much as 1.5 times the difference in the third and first quartile values; outliers beyond these limits are indicated by asterisks.

Appendix B.1. Primer sequences used for RT-PCR

Primer name	Sequence (5'-3')	Reference
P62 (SINV-1 forward)	GGAAGTCATTACGTGGTCGAAAACG	(1)
P63 (SINV-1 reverse)	CGTCCTGTATGAAAACCGGTCTTTACCACAGAAATCTTA	(1)
P64 (SINV-2 forward)	ATTTGTTTTGGCCACGGTCAACA	(2)
P65 (SINV-2 reverse)	GATGATACAAAGCATTAGCGTAGGTAAACG	(2)
P705 (SINV-3 forward)	CTGCTGGTATGATGGCAACAGATCCTTCTGT	(3)
P707 (SINV-3 reverse)	AAGGAGTTTGTGTATTAGTTGCAATGCCAGAATCT	(3)
Gp-9 33 (*Gp-9* control forward)	CATTCAAAGTACAGTAGAATAACTGCC	(4)
Gp-9 490 (*Gp-9* control reverse)	GTATGCCAGCTGTTTTTAATTGC	(4)

Appendix B.2. Results of Mann-Whitney *U* tests for trait-value differences between singly infected and co-infected queens.

				Sample size		Mean / Frequency[a]		
Single infection	Trait	*W*-score	*P*-value[b]	Singly infected	Co-infected[c]	Singly infected	Co-infected	Difference[d]
SINV-1	*QnSize*	62	0.602	9	12	0.02	0.01	0.01
SINV-1	*QnWt*	164	**0.006**	12	17	15.36	14.22	1.13
SINV-1	$QnWtLs_{CL}$	3	0.167	3	6	0.49	0.72	-0.23
SINV-1	*ClaustPer*	3	0.346	3	1	21.33	20.00	1.33
SINV-1	$QnSurv_{CL}$	13	0.778	4	7	0.75	0.86	-0.11
SINV-1	$ColWt_{CL}$	18	**0.024**	3	6	11.80	2.58	9.22
SINV-2	*QnSize*	372	0.489	71	12	-0.01	0.01	-0.02
SINV-2	*QnWt*	791	0.544	85	17	14.37	14.22	0.15
SINV-2	$QnWtLs_{CL}$	36	0.080	23	6	0.60	0.72	-0.12
SINV-2	*ClaustPer*	21	0.168	22	1	22.86	20.00	2.86
SINV-2	$QnSurv_{CL}$	114	0.149	46	7	0.57	0.86	-0.29
SINV-2	$ColWt_{CL}$	117	**0.011**	23	6	7.38	2.58	4.79
SINV-2	$QnSurv_{CL\text{-}6}$	12	**0.015**	6	2	1.00	0.00	1.00

[a] Means are reported for continuous traits, while frequencies are reported for the binary traits $QnSurv_{CL}$ and $QnSurv_{CL\text{-}6}$. *QnSize* is reported as the PC1 score from principal components analysis (PCA), while *QnWt* and $ColWt_{CL}$ are reported as weight in mg. $QnWtLs_{CL}$ is the proportional weight loss of queens during the claustral period. $ColGrowth_{CL\text{-}6}$ is the proportional weight gain of colonies during the six-week period immediately following the end of the claustral period.

[b] *P*-values significant after correction for multiple comparisons are shown in bold and underlined.

[c] Excludes a queen triply infected with SINV-1, SINV-2, and SINV-3.

[d] Mean / frequency of singly infected queens minus mean / frequency of co-infected queens.