

# GENES WITH SOCIAL EFFECTS ARE EXPECTED TO HARBOR MORE SEQUENCE VARIATION WITHIN AND BETWEEN SPECIES

Timothy A. Linksvayer<sup>1,2,3</sup> and Michael J. Wade<sup>4,5</sup>

<sup>1</sup>*School of Life Sciences, Arizona State University, P.O. Box 874501, Tempe, Arizona 85287-4501*

<sup>2</sup>*E-mail: tlinksvayer@gmail.com*

<sup>4</sup>*Department of Biology, Indiana University, Bloomington, Indiana 47405*

<sup>5</sup>*E-mail: mjwade@indiana.edu*

Received October 31, 2008

Accepted February 6, 2009

The equilibrium sequence diversity of genes within a population and the rate of sequence divergence between populations or species depends on a variety of factors, including expression pattern, mutation rate, nature of selection, random drift, and mating system. Here, we extend population genetic theory developed for maternal-effect genes to predict the equilibrium polymorphism within species and sequence divergence among species for genes with social effects on fitness. We show how the fitness effects of genes, mating system, and genetic system affect predicted gene polymorphism. We find that, because genes with indirect social effects on fitness effectively experience weaker selection, they are expected to harbor higher levels of polymorphism relative to genes with direct fitness effects. The relative increase in polymorphism is proportional to the inverse of the genetic relatedness between individuals expressing the gene and their social partners that experience the fitness effects of the gene. We find a similar pattern of more rapid divergence between populations or species for genes with indirect social effects relative to genes with direct effects. We focus our discussion on the social insects, organisms with diverse indirect genetic effects, mating and genetic systems, and we suggest specific examples for testing our predictions with emerging sociogenomic tools.

**KEY WORDS:** Indirect genetic effects, kin selection, mutation–selection balance, relaxed selective constraint, social evolution, sociogenomics.

The equilibrium sequence diversity of a gene within a population is affected by many features of its population biology, including the rate of mutation, the nature of selection (e.g., balancing vs. directional), its pattern of expression (i.e., which individuals express the gene and its fitness effects), the effective population size, the mating system, and the genetic structure. For example, consider the different effects of recessive and dominant gene expression on the equilibrium polymorphism at mutation–selection balance.

<sup>3</sup>Present address: Centre for Social Evolution, Department of Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark

A deleterious recessive allele will achieve a higher equilibrium frequency than a deleterious dominant allele with the same selection coefficient, because, when rare, the majority of copies of the recessive are unexpressed and hidden away from natural selection in heterozygotes. Mating system interacts with expression because a deleterious recessive will reach a higher frequency in a randomly mating population than it will in an inbreeding population whereas a deleterious dominant is unaffected by inbreeding because its expression remains constant across mating systems. Similarly, a diploid gene expressed only by one sex, say females, will have twice the equilibrium heterozygosity as a gene expressed in both sexes, all else being equal, because of the

number of unexpressed copies in males (Wright 1960; Whitlock and Wade 1995; Wade 1998, 2001; Barker et al. 2005; Demuth and Wade 2007; Wade et al. 2008). More generally, all factors that decrease the association between genotype and phenotype decrease the proportion of gene copies that are *seen* by selection and, thus, result in higher equilibrium polymorphism within populations at mutation–selection balance. When the phenotypic effects of a gene expressed in one individual are manifested in a different individual, as for example, when maternal genes affect offspring phenotype (i.e., maternal genetic effects), the association between genotype and phenotype is decreased and within-population polymorphism levels are increased (Wade 1998, 2001). In each of these cases, the higher level of polymorphism is the result of the relaxed selection, often called *relaxed selective constraint* (Demuth and Wade 2007; J. D. Van Dyken and M. J. Wade, unpubl. ms.) that attends the weaker association between genotype and phenotype.

Higher levels of polymorphism from relaxed selective constraint lead to higher levels of equilibrium additive genetic variance for fitness (Whitlock and Wade 1995; Wade 1998). When genetic variance within populations is converted to variance among populations by the combined action of random genetic drift and local selection, it results in higher rates of evolution measured as the number of fixed, nonsynonymous substitutions between two taxa (Demuth and Wade 2007; Cruickshank and Wade 2008; J. D. Van Dyken and M. J. Wade, unpubl. ms.). Divergence between species is often interpreted as evidence of adaptive evolution by diversifying natural selection (i.e., local selection), although relatively rapid divergence can also result from relaxed selective constraint (i.e., reduced effectiveness of purifying selection followed by divergence by random genetic drift). For example, maternally expressed genes can accumulate fixed differences between taxa several times more rapidly than a homologue with direct effects expressed in both sexes (Demuth and Wade 2007; Cruickshank and Wade 2008).

Here, we extend population genetic theory developed to predict polymorphism at maternal-effect genes to genes with other types of indirect effects arising from social interactions, such as kin selected effects. These indirect genetic effects, also known as associative effects, occur when the phenotype of an individual is affected by genes expressed in its social partners (Cheverud and Moore 1994; Moore et al. 1997; Wolf et al. 1998; Bijma et al. 2007). We are specifically interested in determining the “relative” levels of polymorphism at genes with direct effects, maternal effects, and sib effects, in haplo-diploid (HD) and diplo-diploid (DD) populations with both single and multiple mating. Thus, we are interested in how the patterns of gene expression, genetic system, and mating system affect expected sequence diversities. Higher levels of polymorphism within species for some genes allow these genes to respond more quickly to selection under

changing conditions or to diversify more rapidly among populations when genetically subdivided.

Although our results apply to all social organisms in which phenotypes are influenced by the genotypes of interacting social partners, including model organisms usually considered to be solitary, for example, *Drosophila* (Wolf 2003) and *Arabidopsis* (Mitic and Wolf 2007), we focus our discussion on the social insects for three reasons: (1) Social interactions have special significance in social insect colonies; as described below, a variety of different indirect effects are theoretically important in the expression, evolution, and development of social insect phenotypes (Linksvayer and Wade 2005), and indeed recent empirical studies explicitly designed to estimate the relative contributions of direct and indirect genetic effects to variation for social insect phenotypes demonstrate the importance of indirect effects arising from social interactions (Linksvayer 2006, 2007; Wang et al. 2008; Linksvayer et al. 2009). (2) Social insects have diverse genetic and mating systems, so that our discussion focused on social insects applies broadly to organisms with divergent genetic and mating systems. (3) The emerging field of sociogenomics is rapidly growing (e.g., Robinson et al. 2005; Sinha et al. 2006; Barchuk et al. 2007; Toth et al. 2007), especially with the availability of the honey bee genome (Honey Bee Genome Sequencing Consortium 2006), and theory explicitly incorporating the complexities of patterns of gene expression, mating systems, and genetic systems in social organisms are required to make specific evolutionary genetic predictions to be tested with this new empirical resource.

In the social insects, phenotypes are influenced by a variety of different genetic and environmental pathways. The genetic pathways include genes with direct effects, expressed during the development of males and females, genes with maternal effects expressed only in queens or paternal effects expressed only in males, as well as genes with sib effects, expressed in adult workers or in sibling larvae (Linksvayer and Wade 2005; Linksvayer 2006, 2007). As we show here, much of the phenotypic variation among the social insects may be the result of relaxed selective constraint and its interaction with diversifying, adaptive selection. Furthermore, when worker phenotypes have indirect fitness effects on queen and male phenotypes, then genes affecting worker phenotypes and division of colony labor should harbor higher levels of polymorphism. This may be an important factor contributing to the well-known diversity of species and social phenotypes in these taxa.

## *Maternal Genetic Effects Theory*

Using a single locus population genetic model, Wade (1998) argued that, because selection on a maternal-effect gene is half as effective as selection on a direct-effect gene expressed within populations, nonsynonymous or replacement site diversity should be

two times higher, at equilibrium, for a maternal-effect gene than for a direct-effect gene. This prediction assumes that the maternal- and direct-effect loci being compared experience similar strengths of selection when expressed, and furthermore that there have not been recent selective sweeps at the two loci or closely linked loci distorting the expected equilibrium. Among populations or lineages, for purifying selection, a greater rate of replacement site divergence is predicted for a maternal-effect gene than for a direct-effect gene because slightly deleterious mutations for maternal-effect genes are more likely to fix (Demuth and Wade 2007). Once again, this prediction assumes similar selection coefficients between the loci being compared as well as similar effective population sizes between the lineages being compared. When two genes with different expression patterns are compared for their *relative* levels of polymorphism within the same species or their *relative* levels of diversification across the same set of taxa, they necessarily have similar effective population sizes. Selection must be *more than twice as strong* on a maternal-effect gene in order for its diversification by positive selection to be more rapid than the rate of diversification by positive selection of a direct-effect gene.

In quantitative support of this theoretical prediction, Barker et al. (2005) compared sequence diversity of the maternal-effect gene, *bicoid*, with that of its zygotically expressed, tandem duplicate, *zerknüllt* (*zen*) within *Drosophila melanogaster* and *D. simulans*. They found that sequence diversity in the coding regions of *bicoid* is at least twice that of *zen* in both species as predicted. In contrast, the ratios of neutral, silent-site sequence diversity (third position base pairs and introns) were approximately one. This result is consistent with both loci experiencing similar strengths of purifying selection, but selection at the maternal effect *bicoid* being half as effective as selection at the direct effect *zen*, as predicted. For genes acting in early *Drosophila* development, Cruickshank and Wade (2008) found a twofold elevation in replacement site diversity for maternal-effect genes as a class relative to zygotic genes, once again consistent with these highly conserved genes experiencing on average similar strengths of purifying selection when expressed.

Following Kimura (1962), Demuth and Wade (2007, see their Fig. 1) quantify the rate of fixation of a maternal-effect gene relative to that of a direct-effect gene as a function of  $s$ , the average selection coefficient, assuming no dominance. For positively selected genes (where  $s > 0$ ), the probability of fixation of a maternal-effect gene is always less than that of a direct-effect gene, with the relative rate asymptoting at a value of 0.50 as  $s$  exceeds +0.01. That is, positive selection acting on a maternal-effect gene is only half as effective as similar selection on a direct-effect gene. Conversely, for purifying selection acting against deleterious alleles, where the average  $s$  is negative (i.e.,  $s < 0$ ), the relative probability of fixation for a maternal-effect gene exceeds

that of a direct-effect gene. For example, for  $s \sim -0.01$ , the probability of fixation is more than four times higher. Among taxa in the cyclorrhaphan flies (Stauber et al. 1999, 2002), the coding region of *bicoid* is evolving 1.67 times faster than *zen* (Demuth and Wade 2007), consistent with a predominance of purifying selection. However, among species within the genus *Drosophila*, where it is known that *bicoid* is under positive selection, it is evolving at a rate of only 0.68 that of *zen*, close to the limiting ratio of 0.50. Together these empirical tests demonstrate the utility of population genetic theory for predicting patterns of sequence variation within and between populations.

Wade (2001) extended population genetic models for maternal-effect genes from diploid autosomal genes to X-linked and haplo-diploid genes. He showed that the effect of selection on maternal-effect genes is greater in haplo-diploids than in diploids and that it is not affected by inbreeding or mate numbers. However, these theoretical predictions have not yet been tested either with estimates of standing sequence diversity within populations or with estimates of the rates of phylogenetic divergence among taxa.

## Social Genetics Theory

Sequence diversity of sib-effect genes in social insect populations is affected by the mating system and the number of colony queens (i.e., any factor that affects nestmate relatedness). In contrast, polymorphism of maternal-effect genes is insensitive to numbers of mates and the colony founding system does not affect those maternal genes whose expression impacts only their own progeny (Wade 1998, 2001). For genes with direct effects expressed in both sexes, equilibrium replacement site diversity should be lower given the same average intensity of selection, assuming predominantly purifying selection. For comparisons within species among genes with different expression patterns, we assume that the strength of selection,  $s$  and effective population size,  $N_e$  are the same (or essentially so) for all genes, regardless of expression pattern. That is, we assume that the genes experience approximately the same average distribution of  $s$  values for the period of comparison, and we ask what are the expected differences in sequence diversity? For the *Hox* genes and other highly conserved gene families experiencing strong purifying selection, this assumption is probably adequate. However, for positive selection on genes involved in the origin and refinement of adaptive novelty, this assumption is less likely to hold. Nevertheless, because the ratio of fixation probabilities for genes with different patterns of expression asymptotes for positive selection with values of  $s > 0.01$  (Demuth and Wade 2007; see Fig. 3 below), some of our predictions are less sensitive to variations in  $s$  than others. Below, we will relax the assumption of equal distributions of  $s$  and examine its effect on the relative rates of fixation for the different kinds

**Table 1.** Fitness of offspring genotypes in families of a single-mating diplo-diploid population. Frequency (top of each cell) and fitness (bottom of each cell) of offspring depends on their own genotype and the genotypes of family members. Shaded boxes are offspring genotypes not produced.  $W_{DD\ monandry} = 1 + 2p(s_d + s_m + s_s)$ .  $\Delta p_{DD\ monandry} = pq(s_d + 1/2s_m + 1/2s_s) / W_{DD\ monandry}$ .

Mother	Father	A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub> A <sub>2</sub>	A <sub>2</sub> A <sub>2</sub>
A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub> A <sub>1</sub>	1 1 + 2s <sub>d</sub> + 2s <sub>m</sub> + 2s <sub>s</sub>		
A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub> A <sub>2</sub>	1/2 1 + 2s <sub>d</sub> + 2s <sub>m</sub> + 3/2s <sub>s</sub>	1/2 1 + s <sub>d</sub> + 2s <sub>m</sub> + 3/2s <sub>s</sub>	
A <sub>1</sub> A <sub>1</sub>	A <sub>2</sub> A <sub>2</sub>		1 1 + s <sub>d</sub> + 2s <sub>m</sub>	
A <sub>1</sub> A <sub>2</sub>	A <sub>1</sub> A <sub>1</sub>	1/2 1 + 2s <sub>d</sub> + s <sub>m</sub> + 3/2s <sub>s</sub>	1/2 1 + s <sub>d</sub> + s <sub>m</sub> + 3/2s <sub>s</sub>	
A <sub>1</sub> A <sub>2</sub>	A <sub>1</sub> A <sub>2</sub>	1/4 1 + 2s <sub>d</sub> + s <sub>m</sub> + s <sub>s</sub>	1/2 1 + s <sub>d</sub> + s <sub>m</sub> + s <sub>s</sub>	1/4 1 + s <sub>m</sub> + s <sub>s</sub>
A <sub>1</sub> A <sub>2</sub>	A <sub>2</sub> A <sub>2</sub>		1/2 1 + s <sub>d</sub> + s <sub>m</sub> + 1/2s <sub>s</sub>	1/2 1 + s <sub>m</sub> + 1/2s <sub>s</sub>
A <sub>2</sub> A <sub>2</sub>	A <sub>1</sub> A <sub>1</sub>		1 1 + s <sub>d</sub> + s <sub>s</sub>	
A <sub>2</sub> A <sub>2</sub>	A <sub>1</sub> A <sub>2</sub>		1/2 1 + s <sub>d</sub> + 1/2s <sub>s</sub>	1/2 1 + 1/2s <sub>s</sub>
A <sub>2</sub> A <sub>2</sub>	A <sub>2</sub> A <sub>2</sub>			1 1

of gene expression patterns. As for the maternal effects theory described above, we also assume that the loci being compared are at equilibrium under mutation–selection balance and have not experienced recent selective sweeps disrupting this equilibrium.

For predictions of the *relative* levels of sequence diversity for genes within the same species, the assumption of equal  $N_e$  is straightforward as long as neither gene is closely linked to some other gene experiencing strong selection. However, among taxa, we can expect some variation in  $N_e$  and this may affect our predictions for the *relative* probabilities of fixation and rates of sequence evolution among the same set of taxa.

### The Model

#### DIPLO-DIPLOIDS

Consider a diploid, autosomal locus, A, with two alternate alleles, A<sub>1</sub> and A<sub>2</sub> in frequencies  $p$  and  $q$  ( $= 1 - p$ ), respectively. If there are more than two alleles, we make the standard assumption that the frequency of A<sub>2</sub> includes all alleles other than A<sub>1</sub>. Let A<sub>1</sub> be expressed in individuals, their mothers, and their siblings, with the following additive effects on individual fitness; direct effect,  $s_d$ , maternal effect,  $s_m$ , and sib effect,  $s_s$ . (Dominance deviations from additive expectation affect the  $\Delta p$  expressions below, but not their *ratios* which is our concern here.) Tables 1 and 2 summarize how the fitness of individuals depends on their own genotype and that of their family members, for families with a singly mated (monandrous) or multiply mated (polyandrous) mother. With these fitness effects, the change in allele frequency

of A<sub>1</sub> in the next generation after selection for a monandrous population is given by

$$\Delta p_{DD\ monandry} = pq(s_d + 1/2s_m + 1/2s_s) / W, \tag{1}$$

where the mean fitness,  $W$  equals  $[1 + 2p(s_d + s_m + s_s)]$ . When  $s_m$  and  $s_s$  equal zero, equation (1) is the standard population genetic equation for the change in allele frequency for an allele with an additive direct effect,  $s_d$ . For a polyandrous population, the coefficients of  $s_d$  and  $s_m$  remain constant, but that of  $s_s$  is halved:

$$\Delta p_{DD\ polyandry} = pq(s_d + 1/2s_m + 1/4s_s) / W. \tag{2}$$

The change in mating system from monandry to polyandry does not affect the contribution of either the direct effect or the maternal effect to the rate of evolution, but it does affect the coefficient of the sib effect, which decreases from  $1/2$  to  $1/4$ . The coefficients of  $s_d$ ,  $s_m$ , and  $s_s$  are the genetic regressions between the individual expressing the genes and the individual expressing its phenotypic (i.e., fitness) effects. For  $s_m$ , the coefficient is  $1/2$ , corresponding to the genetic relatedness between mother and offspring, and, for  $s_s$ , the coefficient is  $1/2$  and  $1/4$ , the average relatedness between full-siblings and half-siblings, respectively, as determined by the mating system (see Fig. 1). For haplo-diploids, these genetic regressions change and, below, we describe how this affects the sequence diversity maintained at mutation–selection balance within populations, and how this contributes to sequence divergence between lineages.

**Table 2.** Fitness of offspring genotypes in families of a multiple-mating diplo-diploid population.  $W_{DD\text{polyandry}}=1+2p(s_d+s_m+s_s)$ .  $\Delta p_{DD\text{polyandry}}=pq(s_d+1/2s_m+1/4s_s)/W_{DD\text{polyandry}}$ .

Mother	A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub> A <sub>2</sub>	A <sub>2</sub> A <sub>2</sub>
A <sub>1</sub> A <sub>1</sub>	$p$ $1+2s_d+2s_m+(1+p)s_s$	$1-p$ $1+s_d+2s_m+(1+p)s_s$	
A <sub>1</sub> A <sub>2</sub>	$p/2$ $1+2s_d+s_m+(p+1/2)s_s$	$1/2$ $1+s_d+s_m+(p+1/2)s_s$	$(1-p)/2$ $1+s_m+(p+1/2)s_s$
A <sub>2</sub> A <sub>2</sub>		$p$ $1+s_d+ps_s$	$1-p$ $1+ps_s$

**HAPLO-DIPLOIDS**

HD populations have an added complication because males and females must be modeled separately (Tables 3–6). One of the important considerations is how to assign fitnesses to males relative to females. Comparing Tables 3 and 4, note that we have assigned a fitness of  $(1 + 2s_d)$  to A<sub>1</sub>A<sub>1</sub> females and a fitness of  $(1 + s_d)$  to A<sub>1</sub> males. This means that our derivation assumes no dosage compensation of the direct effects of alleles. An alternative, not explored here, is to assign haploid A<sub>1</sub> males a fitness of  $(1 + 2s_d)$ . The equations for sex-specific changes in allele frequency per generation due to selection in a multiple-mating HD population are:

$$\Delta p_{HD\text{polyandry}}^{\sigma} = pq(s_d + s_m + 1/2s_s) / W_{HD\text{polyandry}}^{\sigma} \quad (3)$$

$$\Delta p_{HD\text{polyandry}}^{\varphi} = pq(s_d + 1/2s_m + 1/4s_s) / W_{HD\text{polyandry}}^{\varphi} \quad (4)$$

The mean fitness of haplo-diploid females,  $W_{HD\text{polyandry}}^{\varphi}$ , is  $[1 + 2p(s_d + s_m + s_s)]$ , identical to that of diplo-diploid females. With no dosage compensation, the mean fitness of haplo-diploid males,  $W_{HD\text{polyandry}}^{\sigma}$  equals  $[1 + p(s_d + 2s_m + 2s_s)]$ .

The coefficient of  $s_m$ , the maternal effect on offspring fitness, is 1 for males and  $1/2$  for females, representing the maternal-son and maternal-daughter genetic regressions, respectively, which do not change with maternal mate number. In contrast, the coefficient of  $s_s$ , the sib-social effect on daughter fitness, does change from  $1/2$  (full-siblings) with monandry to  $1/4$  (maternal half-siblings) with polyandry.

The total change in allele frequency in the population is the weighted average of that in each sex. The weightings equal the proportion of alleles found in males and females, that is,  $1/3$  and  $2/3$ , respectively. Assuming weak selection, mean fitness is approximately equal in both sexes, and, we have

$$\Delta p_{HD\text{polyandry}} = pq(s_d + 2/3s_m + 1/3s_s) / W_{HD\text{polyandry}} \quad (5)$$

In a single-mating HD population, male and female change in allele frequency is:

$$\Delta p_{HD\text{monandry}}^{\sigma} = pq(s_d + s_m + 1/2s_s) / W_{HD\text{monandry}}^{\sigma} \quad (6)$$

$$\Delta p_{HD\text{monandry}}^{\varphi} = pq(s_d + 1/2s_m + 3/4s_s) / W_{HD\text{monandry}}^{\varphi} \quad (7)$$

**Table 3.** Fitness of daughter genotypes in families of a single-mating haplo-diploid population.  $W_{HD\text{monandry}}=1+2p(s_d+s_m+s_s)$ .  $\Delta p_{HD\text{monandry}}=pq(s_d+1/2s_m+3/4s_s)/W_{HD\text{monandry}}$ .

Mother	Father	A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub> A <sub>2</sub>	A <sub>2</sub> A <sub>2</sub>
A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub>	1 $1+2s_d+2s_m+2s_s$		
A <sub>1</sub> A <sub>1</sub>	A <sub>2</sub>		1 $1+s_d+2s_m+s_s$	
A <sub>1</sub> A <sub>2</sub>	A <sub>1</sub>	$1/2$ $1+2s_d+s_m+3/2s_s$	$1/2$ $1+s_d+s_m+3/2s_s$	
A <sub>1</sub> A <sub>2</sub>	A <sub>2</sub>		$1/2$ $1+s_d+s_m+1/2s_s$	$1/2$ $1+s_m+1/2s_s$
A <sub>2</sub> A <sub>2</sub>	A <sub>1</sub>		1 $1+s_d+s_s$	
A <sub>2</sub> A <sub>2</sub>	A <sub>2</sub>			1 1

**Table 4.** Fitness of daughter genotypes in families of a multiple-mating haplo-diploid population.  $W_{HD\text{polyandry}} = 1 + 2p(s_d + s_m + s_s)$ .  $\Delta P_{HD\text{polyandry}} = p q (s_d + 1/2 s_m + 1/4 s_s) / W_{HD\text{polyandry}}$ .

Mother	A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub> A <sub>2</sub>	A <sub>2</sub> A <sub>2</sub>
A <sub>1</sub> A <sub>1</sub>	$p$ $1 + 2s_d + 2s_m + (1+p)s_s$	$Q$ $1 + s_d + 2s_m + (1+p)s_s$	
A <sub>1</sub> A <sub>2</sub>	$p/2$ $1 + 2s_d + s_m + (1+2p)/2 s_s$	$1/2$ $1 + s_d + s_m + (1+2p)/2 s_s$	$q/2$ $1 + s_m + (1+2p)/2 s_s$
A <sub>2</sub> A <sub>2</sub>		$P$ $1 + s_d + p s_s$	$Q$ $1 + p s_s$

Combining the weighted changes and assuming approximately equal fitness in males and females, we find that total change in allele frequency is

$$\Delta P_{HD\text{ monadry}} = pq(s_d + 2/3s_m + 2/3s_s) / W_{HD\text{ monadry}} \quad (8)$$

Comparing the coefficients of  $s_m$  and  $s_s$  in equations (5) and (8), we note that the coefficient of  $s_m$  is 2/3 in both cases, equal to the coefficient of  $s_s$  with monadry but twice as large as is the value of 1/3 with polyandry. Thus, a change in the mating system from monadry to polyandry will affect the relative rate of evolution of maternal to sib-effect genes.

**EXPECTED POLYMORPHISM WITHIN POPULATIONS**

Following Whitlock and Wade (1995), we compute the expected allele frequencies at mutation–selection balance (i.e., assuming purifying selection) given  $\mu$ , the mutation rate per generation (Table 7). As shown by Wade (1998, 2001), maternal effect loci are predicted to be more polymorphic (twice as much for DD populations) than direct effect loci at mutation–selection balance. The expected polymorphism for maternal-effect genes does not

depend on mating system. In contrast, for sib-effect loci, the predicted level of polymorphism maintained at mutation–selection balance within populations does depend on mating system. The most sequence variation is predicted to be maintained at sib-effect genes in multiple-mating populations. In DD, this is four times the amount of polymorphism relative to direct-effect genes and twice the polymorphism relative to maternal-effect genes. Generally, the predicted amount of polymorphism maintained within populations depends on the relatedness between the individual expressing the gene and the individual in which the phenotypic effects are manifested. For direct-effect genes, these individuals are one and the same and relatedness is one. The average relatedness between mothers and offspring is 1/2 in DD and 2/3 in HD, and the average relatedness between siblings depends on mating system (e.g., for DD, it is 1/2 with monadry and 1/4 with polyandry).

**SEQUENCE VARIATION BETWEEN LINEAGES**

Here we consider rates of sequence divergence between loci with different patterns of expression, estimated as probabilities of fixation. The rate of substitution of new allelic variants is governed by the mutation rate,  $\mu$  and the probability of fixation,  $P$ . For alleles with additive effects, Kimura (1962) showed that the probability of fixation,  $P_{d,DD}$  of an allele at DD locus with a direct effect on fitness of  $s$  is

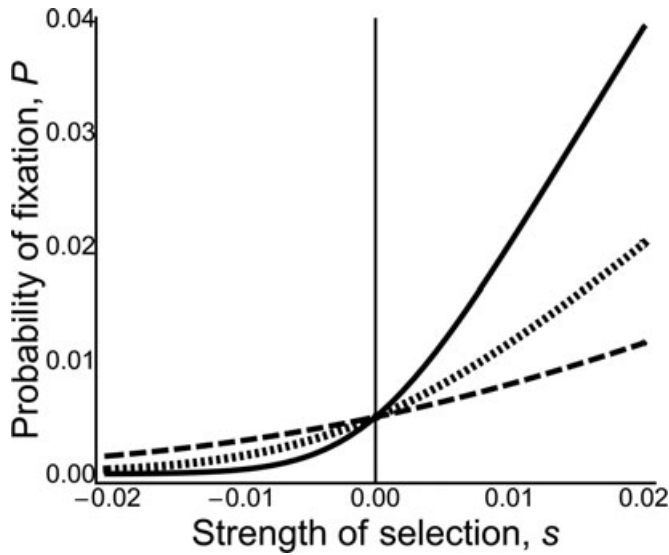
$$P_{d,DD} = \frac{1 - e^{-4N_e s p}}{1 - e^{-4N_e s}} \quad (9)$$

**Table 5.** Fitness of son genotypes in families of a single-mating haplo-diploid population.  $W_{HD\text{monadry}} = 1 + p(s_d + 2s_m + 2s_s)$ .  $\Delta P_{HD\text{monadry}} = p q (s_d + s_m + 1/2 s_s) / W_{HD\text{monadry}}$ .

Mother	Father	A <sub>1</sub>	A <sub>2</sub>
A <sub>1</sub> A <sub>1</sub>	A <sub>1</sub>	1 $1 + s_d + 2s_m + 2s_s$	
A <sub>1</sub> A <sub>1</sub>	A <sub>2</sub>	1 $1 + s_d + 2s_m + s_s$	
A <sub>1</sub> A <sub>2</sub>	A <sub>1</sub>	1/2 $1 + s_d + s_m + 3/2 s_s$	1/2 $1 + s_m + 3/2 s_s$
A <sub>1</sub> A <sub>2</sub>	A <sub>2</sub>	1/2 $1 + s_d + s_m + 1/2 s_s$	1/2 $1 + s_m + 1/2 s_s$
A <sub>2</sub> A <sub>2</sub>	A <sub>1</sub>		1 $1 + s_s$
A <sub>2</sub> A <sub>2</sub>	A <sub>2</sub>		1 1

**Table 6.** Fitness of son genotypes in families of a multiple-mating haplo-diploid population.  $W_{HD\text{polyandry}} = 1 + p(s_d + s_m + 2s_s)$ .  $\Delta P_{HD\text{polyandry}} = p q (s_d + s_m + 1/2 s_s) / W_{HD\text{polyandry}}$ .

Mother	A <sub>1</sub>	A <sub>2</sub>
A <sub>1</sub> A <sub>1</sub>	1 $1 + s_d + 2s_m + (1+p)s_s$	
A <sub>1</sub> A <sub>2</sub>	1/2 $1 + s_d + s_m + (1+2p)/2 s_s$	1/2 $1 + s_m + (1+2p)/2 s_s$
A <sub>2</sub> A <sub>2</sub>		1 $1 + p s_s$



**Figure 1.** Probability of fixation of new mutants in a diplo-diploid population with different patterns of expression for weak negative to weak positive selection. Alleles with direct effects (solid line); alleles with maternal effects or sib effects in single-mating full-sibling family (dotted line); and alleles with sib effects in a multiple-mating half-sibling family (dashed line). Alleles with indirect effects are less likely to fix when beneficial and less likely to be lost when harmful.  $N_e = 100$ ,  $p = 1/(2N_e)$ .

Substituting  $(s/2)$ , the selection coefficient for a maternal effect relative to a direct-effect gene (i.e., half the strength of selection), Demuth and Wade (2007), following Kimura (1962), showed that the probability of fixation of a maternal effects allele in a DD population is

$$P_{m,DD} = \frac{1 - e^{-2N_e s p}}{1 - e^{-2N_e s}} \quad (10)$$

We can generalize the results of Demuth and Wade (2007) by substituting the appropriate regression coefficients for the different genetic systems and mating systems as derived above. We find that the probability of fixation of alleles with direct, maternal, or sib effects in a DD population is

$$P_{DD} = \frac{1 - e^{-4rN_e s p}}{1 - e^{-4rN_e s}}, \quad (11)$$

where  $r$ , the weighted average of the genetic regression between the individual expressing the allele and the individuals receiving its fitness effects, takes a value of 1 for a direct effect allele,  $1/2$  for a maternal-effect allele in a DD population or a sib-effect allele in a DD monandrous population, and  $1/4$  for a sib-effect allele in a DD polyandrous population. These fixation probabilities for direct, maternal, and sib effects in DD populations are shown in Figure 1.

We can so determine the “relative” rate of divergence for diploid loci with different expression patterns using the ratio of

fixation probabilities. For example, the relative rate of fixation for maternal effect alleles relative to direct effect alleles is

$$\frac{P_{m,DD}}{P_{d,DD}} = \frac{(1 - e^{-4N_e s}) (1 - e^{-2N_e s p})}{(1 - e^{-2N_e s}) (1 - e^{-4N_e s p})}, \quad (12)$$

which simplifies to  $\frac{e^{-2N_e s} + 1}{e^{-2N_e s p} + 1}$  (Demuth and Wade 2007). Assuming that the loci have similar mutation rates, this equation describes the relative rates of divergence between maternal and direct-effect genes. We can relax the assumption of equal selection between differently expressed loci by assuming that selection on one locus is  $k$  times as strong as selection at the second locus, changing the ratio of fixation probabilities in equation (12) to

$$\frac{P_{1,DD}}{P_{2,DD}} = \frac{(1 - e^{-4r_1 k N_e s}) (1 - e^{-4r_2 N_e s p})}{(1 - e^{-4r_2 N_e s}) (1 - e^{-4r_1 k N_e s p})}. \quad (13)$$

Here,  $P_{1,DD}$  and  $P_{2,DD}$  are the fixation probabilities for loci 1 and 2 with different patterns of expression,  $r_1$  and  $r_2$  are the relatedness coefficients for loci 1 and 2 that depend on expression pattern, and  $k$  is strength of selection on locus 1 (when expressed) relative to locus 2.

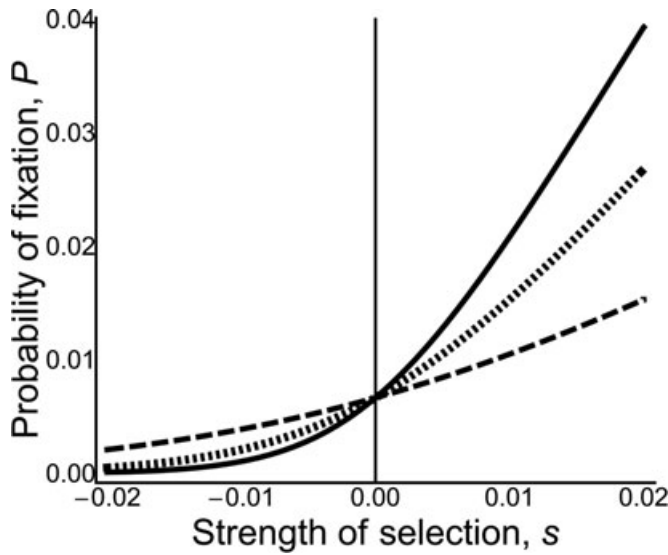
For haplo-diploid species, we need the change in the mean allele frequency by weak selection from equations (5) and (8) as well as the variance of the allele frequency change by random genetic drift to derive the probability of fixation using the diffusion methods of Kimura (1962). Assuming  $W \sim 1$  for weak selection, the average change in frequency,  $M_d$ , by selection for a gene with direct effect,  $s_d$ , is  $pqs_d$  and the variance owing to drift,  $V$ , is  $2pq/3N$  (assuming equal numbers of males and females). The exponent in the equation for fixation probability is given by the ratio  $M_d/(V/2)$  or  $2M_d/V$ . This gives us  $3s_d N$  for a gene with direct effects on male and female fitness. The resulting equation for fixation probability of a direct effect allele in a HD population is

$$P_{d,HD} = \frac{1 - e^{-3N_e s p}}{1 - e^{-3N_e s}} \quad (14)$$

For a gene with only a maternal effect, the ratio changes because the mean change by weak selection,  $M_m$ , is  $2pqs_m/3$  whereas the variance caused by random drift remains the same,

**Table 7.** Expected allele frequency at mutation–selection balance depending on pattern of expression (direct-, maternal-, or sib-effect loci), mating system (monandry or polyandry), and genetic system (haplo-diploid or diplo-diploid).

	Monandry DD	Polyandry DD	Monandry HD	Polyandry HD
Direct effects	$1 \mu/s_d$	$1 \mu/s_d$	$1 \mu/s_d$	$1 \mu/s_d$
Maternal effects	$2 \mu/s_m$	$2 \mu/s_m$	$3/2 \mu/s_m$	$3/2 \mu/s_m$
Sib effects	$2 \mu/s_s$	$4 \mu/s_s$	$3/2 \mu/s_s$	$3 \mu/s_s$



**Figure 2.** Probability of fixation of new mutants in a haplo-diploid population with different patterns of expression. Alleles with direct effects (solid line); alleles with maternal effects or sib effects in single-mating full-sibling family (dotted line); and alleles with sib effects in a multiple-mating half-sibling family (dashed line).  $N_e = 100$ ,  $p = 2/(3N_e)$ .

independent of how a gene is expressed. The ratio,  $2M_m/V$ , equals  $2s_mN$

$$P_{m,DD} = \frac{1 - e^{-2N_e s p}}{1 - e^{-2N_e s}} \tag{15}$$

For a gene with only a sib effect,  $s_s$ , the change in mean,  $M_s$ , varies with the mating system, from  $2pqs_s/3$  with monandry to  $pqs_s/3$  with polyandry. Again, the variance owing to drift remains the same, so our exponent is  $2s_sN$  for monandry but only  $1s_sN$  for polyandry. As we did for a DD population (eq. 11), we can generalize these results to find the probability of fixation of alleles with direct, maternal, or sib effects in a HD population:

$$P_{HD} = \frac{1 - e^{-3rN_e s p}}{1 - e^{-3rN_e s}} \tag{16}$$

Where  $r$  equals 1 for a direct effect allele,  $2/3$  for a maternal effect allele or sib-effect allele with monandry, and  $1/3$  for a sib-effect allele with polyandry (see Figure 2).

Similarly for the DD case above (eq. 14), the relative fixation rates of two loci,  $P_{1,HD}$  and  $P_{2,HD}$  with different patterns of expression is

$$\frac{P_{1,HD}}{P_{2,HD}} = \frac{(1 - e^{-3r_1kN_e s}) (1 - e^{-3r_2N_e s p})}{(1 - e^{-3r_2N_e s}) (1 - e^{-3r_1kN_e s p})} \tag{17}$$

Where,  $r_1$  and  $r_2$  are the relatedness coefficients for loci 1 and 2 that depend on expression pattern, and  $k$  is strength of selection on locus 1 (when expressed) relative to locus 2.

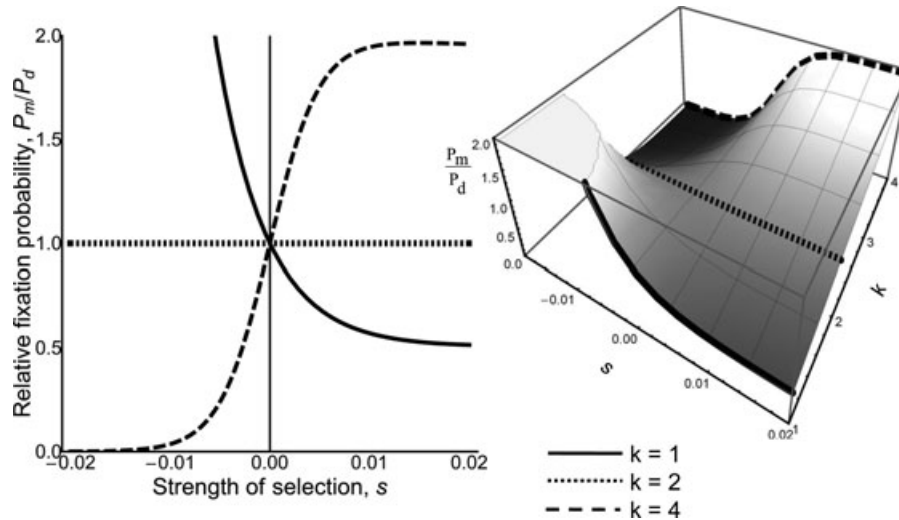
Assuming that new mutations at diploid loci initially occur at a frequency  $p = 1/2N$ , and  $N$  equals  $N_e$ , we show in Figure 3 how the relative fixation probabilities, and assuming equal mutation rates the relative rates of divergence, depend on pattern of expression and strength of selection. Notice that, changing a gene’s expression pattern via  $r$  is equivalent to changing the relative strength of selection experienced by a gene through  $k$ . As a result, two genes can experience the same effective strength of selection, and thus the same fixation probabilities, with different, but balancing patterns of expression and relative strength of selection. For example, in a diploid population, a maternal-effect gene experiencing *twice* the strength of selection as a direct-effect gene will have the same fixation probability as the direct-effect gene (i.e., the  $r$  of  $1/2$  will balance out the  $k$  of 2 for the maternal-effect gene relative to the  $r = k = 1$  for the direct-effect gene; see Figs. 3 and 4).

### Discussion

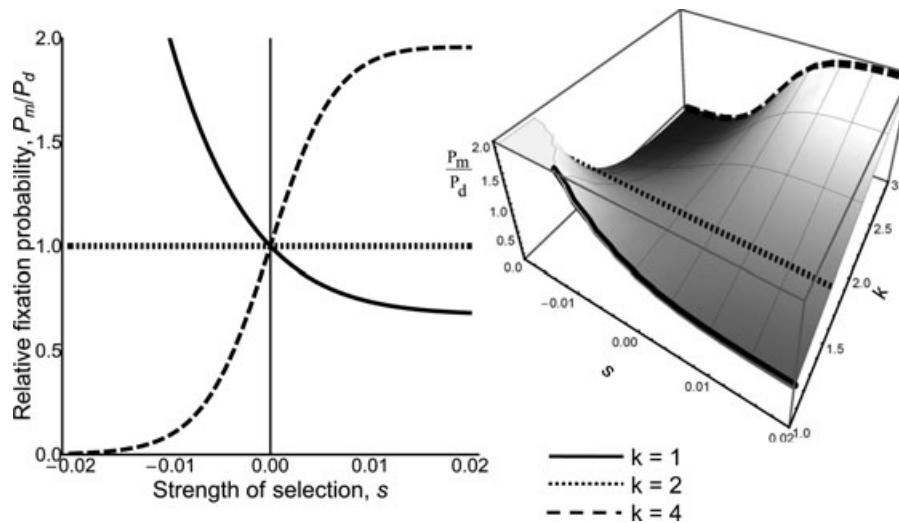
We show that all else being equal, selection acts less effectively on loci with indirect social effects on fitness than on loci with direct effects on fitness. As a result, at mutation–selection balance, indirect effect loci are expected to harbor higher levels of sequence polymorphism within populations. Furthermore, indirect effect loci are expected to diverge more rapidly between lineages than direct effect loci because slightly deleterious indirect effect alleles are more likely to drift to fixation and contribute to fixed differences than similarly deleterious direct effect alleles.

Generally, the relatedness between the individual expressing the gene and the individual manifesting its fitness effects determines the rate of change in allele frequency per generation, and as a result, the expected sequence polymorphism. Thus, decreasing relatedness between social partners is equivalent to weakening the average strength of selection experienced by genes with indirect fitness effects, and as a result increasing the standing genetic variation. This makes intuitive sense. As relatedness decreases, so does the association between genotype and phenotype. Direct-effect genes are expressed in the same individual in which the phenotypic effects are manifested and, barring factors that cause gene copies to be unexpressed (e.g., dominance, penetrance), there is a one-to-one correspondence between genotype and phenotype. Differently put, the genetic relatedness between an individual and itself is one. In contrast, maternal- and sib-effect genes are expressed in mothers and sibs, respectively, but have phenotypic effects on other individuals. The relatedness between mothers and offspring or between siblings is less than one. For example, the genetic relatedness between full-siblings and half-siblings in a randomly mating DD population is  $1/2$  and  $1/4$ , respectively. This results in selection half to one-fourth as weak and two to four times the amount of variation at mutation–selection balance for





**Figure 3.** Relative fixation probability of maternal effect (or sib effect in a monandrous population) to direct effect loci in DD with weak negative to weak positive selection. The relative strength of selection experienced by the indirect effect locus,  $k = 1$  (solid line),  $k = 2$  (dotted line), and  $k = 4$  (dashed line). The full three-dimensional graph of relative fixation probabilities is shown to the right. When  $k = 1$ , the direct and indirect effect loci experience the same strength of selection yet the direct effect locus is much more likely to be lost when it is deleterious ( $s < 0$ ) and becomes twice as likely to be fixed when advantageous ( $s > 0$ ) because the indirect effect locus experiences relaxed selective constraint. When  $k = 2$ , the increased selection experienced by the indirect effect locus is exactly balanced out by its relaxed selective constraint, as measured by the regression coefficient  $r$ , relative to the indirect effect locus, and the relative fixation probability of direct to indirect effects is equal (i.e.,  $2 \times 1/2 = 1$ ). When  $k = 4$ , the indirect effect locus effectively experiences selection twice as strong as the direct effect locus and the indirect effect locus becomes twice as likely to be fixed when advantageous and much more likely to be lost when deleterious.



**Figure 4.** Relative fixation probability of maternal effect (or sib effect in a monandrous population) to direct effect loci in HD with weak negative to weak positive selection. The relative strength of selection experienced by the indirect effect locus,  $k = 1$  (solid line),  $k = 1.5$  (dotted line), and  $k = 3$  (dashed line). The full three-dimensional graph of relative fixation probabilities is shown to the right. In contrast to the DD case, in HD, the relaxed selective constraint experienced by the indirect effect locus is exactly balanced by increased strength of selection when  $k = 3/2$  (i.e.,  $3/2 \times 2/3 = 1$ ). And when  $k = 3$ , the indirect effect locus effectively experiences selection twice as strong as the direct effect locus (i.e.,  $3 \times 2/3 = 2$ ). The relatively flatter surface for HD relative to DD indicates that maternal versus direct expression has less of an effect on the probability of fixation in HD.

sib-effects genes relative to direct-effect genes (Table 7; Figs. 1 and 2). The fact that loci with indirect effects experience less effective selection than loci with direct effects is also captured by kin selection theory. In the well-known Hamilton's Rule,  $rb - c > 0$ , the indirect fitness benefit to a social partner,  $b$ , is decremented by the relatedness between social partners,  $r$ , whereas the direct fitness cost to the actor,  $c$ , has a coefficient of one (Hamilton 1964). These two fitness effects are associated with a single gene in kin selection theory, whereas we have explored the relative rates of evolution and relative levels of polymorphism of pairs of genes with different types of direct or indirect fitness effects.

Mating system does not affect the relatedness between mothers and their offspring, so the expected sequence polymorphism for maternal-effect genes does not depend on mating system, all things being equal (Wade 1998). Mating system does affect the relatedness between siblings. In HD and DD populations with single mating, the expected sequence diversities for maternal and sib effects are the same, but for polyandrous DD and HD populations, sib-effect loci experience less selective constraint and hence accumulate more polymorphism at mutation–selection balance. Other factors that influence the relatedness between social partners, such as queen number and worker egg-laying, will also influence the predicted sequence diversity of caste- or worker-specific genes because such effects change the genetic regression between the genotype where the gene is expressed and the phenotype of the receiving individual.

As summarized above, what is important in terms of our theory is whether a gene has direct or indirect effects on *fitness*. In social insects, queens and males reproduce and workers typically do not (although in many HD social insects, workers can lay unfertilized male-destined eggs in certain situations). As a result, genes affecting most worker traits have mainly indirect fitness effects, dependent on the number and quality of new queens and males produced by colonies. Thus, genes expressed exclusively by workers are likely to have indirect sib effects on fitness; genes expressed by male and female sexuals are likely to have direct fitness effects; and queen genes that affect the quality of her sexual offspring have maternal fitness effects. (Genes that affect egg number and are sex-limited in their expression [i.e., expressed only by queens but also transmitted by males] also experience relaxed selection independent of mating system.)

The weaker purifying selection (i.e., relaxed constraint) on genes with indirect fitness effects potentially enables wider exploration of phenotypic and genotypic space both within and among populations. As a result, social insect populations may harbor more genetic and phenotypic variation within and between populations for worker traits relative to queen and male traits. However, worker and sexual traits are also likely to experience different selective regimes. For example, the traits of males and queens may often experience strong stabilizing selection whereas in some

cases phenotypic variation for worker traits within a colony may be favorable. Generally, there is a tension between the reduced strength of selection that decreases the probability advantageous alleles spread, and relaxed selective constraint that enables the accumulation of genetic variation within population and facilitates rapid divergence between populations. The former applies to alleles that experience mainly positive selection whereas the latter applies to alleles that are mildly deleterious under most environmental conditions, experiencing mainly purifying selection.

All else equal, polyandrous populations should harbor more variation than monandrous populations for sib-effect genes expressed by workers. The discovery of genetic components to worker polymorphism in several highly polyandrous species (Hughes et al. 2003; Rheindt et al. 2005; Jaffe et al. 2007) lends tentative support to this prediction. Indeed alleles for caste determination may generally be mildly deleterious (see review by Anderson et al. 2008), but as explained above, genetically based variation in worker phenotypes may be adaptive in some cases, and not mainly governed by mutation–selection balance.

Given the diversity of genetic (DD or HD) and mating system found in social insects, it is relatively easy to select suitable study systems for testing our predictions. Mating frequency theoretically affects population sex ratio, and studies of sex-ratio evolution have identified several social insect species with populations differing primarily in mating frequency (Bourke and Franks 1995). Comparisons can also be made between lineages that differ in mating frequency, for example, between species in *Pogonomyrmex* harvester ants (Rheindt et al. 2004; Holbrook et al. 2007); between genera and species in fungus-growing ants (Villessen et al. 1999); or between species of stingless bees (Paxton et al. 1999). However, effective population size is likely a complicating factor in these comparisons. As such, it may often be easier to make comparisons of sequence variation within populations or sequence divergence between lineages using comparisons between loci as in equations (11–12), where both loci, although differing in expression pattern, are subject to the same variations in  $N_e$ .

Currently, an incisive test of our theory is limited by the absence of loci that can be defined as having indirect fitness effects. As explained above, genes expressed in adults of the worker caste are likely to have indirect fitness effects, whereas genes expressed in adult sexual females and males, or all larvae are likely to have direct fitness effects. Genes affecting traits involved in sib and maternal care, including provisioning and defense of brood, have sib and maternal effects, respectively (Linksvayer and Wade 2005). The simple predictions discussed here for relative polymorphism and sequence divergence between genes with different expression patterns assume each gene has exclusively direct, maternal, or sib fitness effects. However, in reality, genes may often have multiple pleiotropic effects and care must be taken to identify all the various fitness effects of genes of interest. Expected polymorphism

and sequence divergence depends on the total strength of selection experienced by a gene (see eqs. 1–8). When genes have pleiotropic effects, coefficients for the various direct and indirect selection components must be added (assuming each effect experiences similar strengths of selection), and then expected polymorphism and divergence calculated. A good starting point for testing our predictions is focusing on the simplest case of genes with single types of direct or indirect fitness effects. The rapid growth of sociogenomics promises to provide a wealth of expression data in the near future and these data are the starting point for identifying differences such as sex- and caste-specific patterns of gene expression. Indeed a large number of genes with caste-specific patterns of expression have already been found in termites, ants, and especially honey bees (Miura et al. 1999; Evans and Wheeler 2001; Wheeler et al. 2006; Whitfield et al. 2006; Barchuk et al. 2007).

A notable set of candidate genes for testing our predictions are the major royal jelly protein genes (*mrjp1–9*) found in honey bees. These genes are thought to have arisen via tandem duplication events from the *yellow* genes found in *Drosophila* and other insects (Drapeau et al. 2006). Some are expressed primarily in the hypopharyngeal glands of nurse workers, the site of synthesis of royal jelly proteins that are fed by nurses to larvae and influence whether the larvae develop into queens or workers. These genes thus have indirect sib effects on fitness. Other major royal jelly protein genes are expressed in different body parts of developing queen, worker, and male larvae (summarized in Drapeau et al. 2006, Table 3). These can be considered to be direct-effect genes because they are expressed in both sexes as well as in workers. Comparisons of sequence diversity within populations and divergence between *Apis* species for the different major royal jelly protein loci that differ in patterns of expression may provide a suitable test of our predictions.

Socially parasitic species in which the worker caste is rarely or never expressed also provide a very promising test case. Because genes affecting worker traits are rarely/never expressed, these genes should experience drastically reduced selective constraint and should harbor even more sequence variation than sib-effect genes in normal social insect populations. Furthermore, the closest relative of socially parasitic species is often the host (Emery's Rule, Buschinger 1990; Sumner et al. 2004), offering a unique opportunity for paired comparison of sequence variation for direct and indirect-effect genes in parasite and host.

#### ACKNOWLEDGMENTS

We thank two anonymous reviewers and the editor for helpful comments on the manuscript as well as D. van Dyken, Y. Brandvain, and T. Cruickshank. TAL was funded by a National Science Foundation postdoctoral fellowship in biological informatics; MJW was supported by National Institutes of Health 1R01GM084238-01.

#### LITERATURE CITED

- Anderson, K. E., T. A. Linksvayer, and C. R. Smith. 2008. The causes and consequences of genetic caste determination in ants (Hymenoptera: Formicidae). *Myrmecol. News* 11:119–132.
- Barchuk, A. R., A. S. Cristino, R. Kucharski, L. F. Costa, Z. L. P. Simoes, and R. Maleszka. 2007. Molecular determinants of caste differentiation in the highly eusocial honeybee *Apis mellifera*. *BMC Dev. Biol.* 7:70.
- Barker, M. S., J. P. Demuth, and M. J. Wade. 2005. Maternal expression relaxes constraint on innovation of the anterior determinant, bicoid. *PLoS Genet.* 1:527–530.
- Bijma, P., W. A. Muir, and J. A. M. Van Arendonk. 2007. Multilevel selection I: quantitative genetics of inheritance and response to selection. *Genetics* 175:277–288.
- Bourke, A. F. G., and N. R. Franks. 1995. *Social evolution in ants*. Princeton Univ. Press, Princeton, NJ.
- Buschinger, A. 1990. Sympatric speciation and radiative evolution of socially parasitic ants: heretic hypotheses and their factual background. *J. Zoolog. Syst. Evol. Res.* 28:241–260.
- Cheverud, J. M., and A. J. Moore. 1994. Quantitative genetics and the role of the environment provided by relatives in behavioral evolution. Pp. 67–100 in C. R. B. Boake, ed. *Quantitative genetic studies of behavioral evolution*. Univ. of Chicago Press, Chicago, IL.
- Cruickshank, T., and M. J. Wade. 2008. Microevolutionary support for a developmental hourglass: gene expression patterns shape sequence variation and divergence in *Drosophila*. *Evol. Dev.* 10:583–590.
- Demuth, J. P., and M. J. Wade. 2007. Maternal expression increases the rate of bicoid evolution by relaxing selective constraint. *Genetica* 129:37–43.
- Drapeau, M. D., S. Albert, R. Kucharski, C. Prusko, and R. Maleszka. 2006. Evolution of the Yellow/Major Royal Jelly Protein family and the emergence of social behavior in honey bees. *Genome Res.* 16:1385–1394.
- Evans, J. D., and D. E. Wheeler. 2001. Gene expression and the evolution of insect polyphenisms. *Bioessays* 23:62–68.
- Hamilton, W. D. 1964. The genetical evolution of social behaviour. I, II. *Journal of Theoretical Biology* 7:1–16, 17–52.
- Holbrook, C. T., C. P. Strehl, R. A. Johnson, and J. Gadau. 2007. Low queen mating frequency in the seed-harvester ant *Pogonomyrmex (Ephebomyrmex) pima*: implications for the evolution of polyandry. *Behav. Ecol. Sociobiol.* 62:229–236.
- Honey Bee Genome Sequencing Consortium. 2006. Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443:931–949.
- Hughes, W. O. H., S. Sumner, S. Van Borm, and J. J. Boomsma. 2003. Worker caste polymorphism has a genetic basis in *Acromyrmex* leaf-cutting ants. *Proc. Natl. Acad. Sci. USA* 100:9394–9397.
- Jaffe, R., D. J. C. Kronauer, F. B. Kraus, J. J. Boomsma, and R. F. A. Moritz. 2007. Worker caste determination in the army ant *Eciton burchellii*. *Biol. Lett.* 3:513–516.
- Kimura, M. 1962. On the probability of fixation of mutant genes in a population. *Genetics* 47:713–719.
- Linksvayer, T. A. 2006. Direct, maternal, and subsocial genetic effects on individual and colony traits in an ant. *Evolution* 60:2552–2561.
- . 2007. Ant species differences determined by epistasis between brood and worker genomes. *PLoS ONE* 2:e994.
- Linksvayer, T. A., and M. J. Wade. 2005. The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: maternal effects, sib-social effects, and heterochrony. *Q. Rev. Biol.* 80:317–336.
- Linksvayer, T. A., M. K. Fondrk, and R. E. Page, Jr. 2009. Honey bee social regulatory networks are shaped by colony-level selection. *Am. Nat.* 173:E99–E107.
- Miura, T., A. Kamikouchi, M. Sawata, H. Takeuchi, S. Natori, T. Kubo, and T. Matsumoto. 1999. Soldier caste-specific gene expression in the

- mandibular glands of *Hodotermopsis japonica* (Isoptera : Termopsidae). *Proc. Natl. Acad. Sci. USA* 96:13874–13879.
- Moore, A. J., E. D. Brodie, III, and J. B. Wolf. 1997. Interacting phenotypes and the evolutionary process .1. Direct and indirect genetic effects of social interactions. *Evolution* 51:1352–1362.
- Mutic, J. J., and J. B. Wolf. 2007. Indirect genetic effects from ecological interactions in *Arabidopsis thaliana*. *Mol. Ecol.* 16:2371–2381.
- Paxton, R. J., N. Weissschuh, W. Engels, K. Hartfelder, J. Javier, and G. Quezada-Euan. 1999. Not only single mating in stingless bees. *Naturwissenschaften* 86:143–146.
- Rheindt, F. E., J. Gadau, C. P. Strehl, and B. Holldobler. 2004. Extremely high mating frequency in the Florida harvester ant (*Pogonomyrmex badius*). *Behav. Ecol. Sociobiol.* 56:472–481.
- Rheindt, F. E., C. P. Strehl, and J. Gadau. 2005. A genetic component in the determination of worker polymorphism in the Florida harvester ant *Pogonomyrmex badius*. *Insectes Soc.* 52:163–168.
- Robinson, G. E., C. M. Grozinger, and C. W. Whitfield. 2005. Sociogenomics: social life in molecular terms. *Nat. Rev. Genet.* 6:257–270.
- Sinha, S., X. Ling, C. W. Whitfield, C. X. Zhai, and G. E. Robinson. 2006. Genome scan for cis-regulatory DNA motifs associated with social behavior in honey bees. *Proc. Natl. Acad. Sci. USA* 103:16352–16357.
- Stauber, M., H. Jackle, and U. Schmidt-Ott. 1999. The anterior determinant bicoid of *Drosophila* is a derived Hox class 3 gene. *Proc. Natl. Acad. Sci. USA* 96:3786–3789.
- Stauber, M., A. Prell, and U. Schmidt-Ott. 2002. A single Hox3 gene with composite bicoid and zerknullt expression characteristics in non-cyclorrhaphan flies. *Proc. Natl. Acad. Sci. USA* 99:274–279.
- Sumner, S., D. K. Aanen, J. Delabie, and J. J. Boomsma. 2004. The evolution of social parasitism in *Acromyrmex* leaf-cutting ants: a test of Emery's rule. *Insectes Soc.* 51:37–42.
- Toth, A. L., K. Varala, T. C. Newman, F. E. Miguez, S. K. Hutchison, D. A. Willoughby, J. F. Simons, M. Egholm, J. H. Hunt, M. E. Hudson, and G. E. Robinson. 2007. Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. *Science* 318:441–444.
- Villesen, P., P. J. Gertsch, J. Frydenberg, U. G. Mueller, and J. J. Boomsma. 1999. Evolutionary transition from single to multiple mating in fungus-growing ants. *Mol. Ecol.* 8:1819–1825.
- Wade, M. J. 1998. The evolutionary genetics of maternal effects. Pp. 5–21 in T. A. Mousseau, and C. W. Fox, eds. *Maternal effects as adaptations*. Oxford Univ. Press, New York.
- . 2001. Maternal effect genes and the evolution of sociality in haplodiploid organisms. *Evolution* 55:453–458.
- Wade, M. J., N. K. Priest, and T. Cruickshank. 2008. A theoretical overview of maternal genetic effects: evolutionary predictions and empirical tests using sequence data within and across mammalian taxa. *Maternal effects in mammals*. Univ. of Chicago Press, Chicago, IL.
- Wang, J., K. G. Ross, and L. Keller. 2008. Genome-wide expression patterns and the genetic architecture of a fundamental social trait. *PLoS Genetics* 47:e1000127. Available at <http://www.plosgenetics.org/article/info:doi%2F10.1371%2Fjournal.pgen.1000127>.
- Wheeler, D. E., N. Buck, and J. D. Evans. 2006. Expression of insulin pathway genes during the period of caste determination in the honey bee, *Apis mellifera*. *Insect Mol. Biol.* 15:597–602.
- Whitfield, C. W., Y. Ben-Shahar, C. Brillet, I. Leoncini, D. Crauser, Y. LeConte, S. Rodriguez-Zas, and G. E. Robinson. 2006. Genomic dissection of behavioral maturation in the honey bee. *Proc. Natl. Acad. Sci. USA* 103:16068–16075.
- Whitlock, M. C., and M. J. Wade. 1995. Speciation: founder events and their effects on X-linked and autosomal genes. *Am. Nat.* 145:676–685.
- Wolf, J. B. 2003. Genetic architecture and evolutionary constraint when the environment contains genes. *Proc. Natl. Acad. Sci. USA* 100:4655–4660.
- Wolf, J. B., E. D. Brodie, III, J. M. Cheverud, A. J. Moore, and M. J. Wade. 1998. Evolutionary consequences of indirect genetic effects. *Trends Ecol. Evol.* 13:64–69.
- Wright, S. 1960. On the number of self-incompatibility alleles maintained in equilibrium by a given mutation rate in a population of given size: a reexamination. *Biometrics* 16:61–85.

Associate Editor: R. Snook