MEZNAR, EMILY RACHEL, M.S. Genomic Synteny and Comparison of Recombination between *Apis mellifera* (the European Honey Bee) and *Apis florea (*the Red Dwarf Honey Bee). (2009)
Directed by Dr. Olav Rueppell. 75pp.

In all sexually reproducing organisms meiotic recombination increases genetic diversity among offspring and creates new genomes through chromosomal reconfiguration. The rate at which recombination occurs varies among and within species. Recombination rates are exceptionally high in social insect species and the European honey bee (*Apis mellifera)* has the highest recombination rate known in multi-cellular eukaryotes. To explain this, three major, non-exclusive hypotheses have been proposed. High recombination may be the result of strong selection during domestication of the honey bee (1). Alternatively, it may benefit division of labor (2) or disease resistance (3) by increasing the genetic diversity among colony members. Therefore, I compared the genomic recombination rate in *Apis mellifera* to that of *A. florea* (the red dwarf honey bee). This species is undomesticated, has a relatively low degree of pathogen pressure and a complex division of labor. I screened 684 microsatellite markers from the *A. mellifera* genome for polymorphism in *A. florea;* 37 polymorphic markers were identified and genotyped in a mapping population of 96 *A. florea* drones. Pairwise recombinational distances were calculated using MapMaker3.0. I examined genomic synteny over two chromosomes, finding marker order conserved in both. Overall recombination distances are comparable between the species, though I found one interval of significantly higher recombinational distances in *A. florea* and no intervals of significantly lower distances. I conclude that the high recombination rate in *A. mellifera*

is not a result of domestication and provide further support for the link between complex division of labor and high recombination rate.

GENOMIC SYNTENY AND COMPARISON OF RECOMBINATION BETWEEN *APIS MELLIFERA* (THE EUROPEAN HONEY BEE) AND *APIS FLOREA* (THE RED DWARF HONEY BEE).

by

Emily Rachel Meznar

A Thesis Submitted to
the Faculty of The Graduate School
at The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Greensboro
2009

Approved by

_____
Committee Chair

APPROVAL PAGE

This thesis has been approved by the following committee of the Faculty of The Graduate School at the University of the North Carolina at Greensboro.

Committee Chair_____

Committee Members_____

_____

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Evolution and Genetic Variability

Genetic variability is the fundamental raw material of evolution and selection for advantageous mutations is the consequential mechanism of evolution. Mutations that produce genetic variability can be induced by environmental or chemical factors and they also appear spontaneously in the genome. In any case, mutations occur randomly and beneficial mutations are believed to be relatively rare (Peck 1993). Selection, initially formalized by Charles Darwin in the mid-nineteenth century, results in changes in the frequency of alleles in the next generation. Alleles that are beneficial have a selective advantage and are passed on in a proportion that relates to that advantage. Thus, natural selection constantly decreases genetic variability by favoring certain alleles over others (Cockburn 1991).

This poses an interesting problem: genetic variability is necessary for selection to occur and the process of selection works to reduce genetic variability. Yet, observable genetic variation is ubiquitous today. Therefore, some mechanism must exist to counteract the homogenizing effects of selection (Muller 1932; Felsenstein 1974). The homogenizing effects of selection can be noted especially in asexually reproducing populations. Asexual reproduction requires only one parent who transmits a copy of its

entire genome to each of its offspring. This method has several possible advantages, the most significant of which is the guaranteed perpetuation of successful genotypes. However, the only genetic variability introduced in asexually reproducing organisms is that caused by mutations, the vast majority of which are deleterious (Muller 1932). In fact, asexual reproduction subjects a population to an ever-increasing mutational load (Muller 1964) where all the mutations accumulated in the parental generation must be passed to the offspring.

To prevent deleterious mutations from being passed from parent to offspring the parent must pass only the portion of its DNA that does not contain the mutation to its offspring. Sexual reproduction provides a means by which each parent contributes a complimentary half to their offspring's genome, with the other portion of the genome being contributed by a sexual partner. Sexual reproduction is more costly than asexual reproduction and the risks involved in mixing genes with another individual are significant (reviewed in Michod and Levin 1988; and Otto and Lenormand 2002). The process of mating is costly in time and energy (spent both in courting a mate and mating) and exposes the organism to increased risk of predation and sexually transmitted diseases. The production of males is another cost of sexual reproduction (Maynard Smith 1978). For a sexually reproducing couple to be as successful as an asexually reproducing individual the couple must produce twice as many offspring. Sexual reproduction also does not guarantee that the offspring's genotype will be as successful as either of its parents. Yet sexual reproduction is the predominant mode of reproduction throughout the animal kingdom.

Presumably, sexual reproduction has both long term benefits for populations and short term benefits for individuals (Muller 1932; Van Valen 1973; Felsenstein 1974; Ghiselin 1974). Three main hypotheses propose benefits of sexual reproduction: 1) In the long term, sexual reproduction provides a means to prevent deleterious mutations from accumulating in future generations because each parent contributes only half of the genetic material in the offspring. Therefore, even highly mutated parental genomes can be combined in such a way as to produce offspring with low mutational loads (Muller 1932). 2) In the long term, sexual reproduction can accelerate the rate of evolution by recombining beneficial mutations. Favorable mutations which arise in different lineages can ultimately be combined in one individual (Fisher 1930; Felsenstein 1974). 3) In the short term, sex increases genetic variability among offspring. Increased variability can reduce competition if each individual is more suited to a particular niche (Ghiselin 1974) and also may provide increased pathogen resistance in a population (Van Valen 1973). Increased genetic diversity also benefits populations in the long term by providing more raw material for evolution (Smith 1989; Cockburn 1991). Genetic diversity may be especially important to organisms living in highly structured societies, such as the social insects (Sirvio et al. 2006; Wilfert, Gadau, and Schmid-Hempel 2007)

The ability to select one's sexual partner introduces other benefits to sexual reproduction (Hadany and Beker 2007). In the long term, sexual selection may reduce mutational load (Siller 2001) and give organisms an increased ability to adapt in a changing environment (Lorch et al. 2003). In the short term, unfit females benefit dramatically from combining genes with a highly fit male and highly fit females may

benefit as well. Highly differential mating success among males may allow the male offspring of a highly fit female such a great advantage in the second generation that it compensates for the cost of producing male offspring in the first generation (Hadany and Beker 2007).

Meiotic Recombination

The advantages of sexual reproduction (excluding those introduced solely by sexual selection) are enhanced through the process of meiotic recombination (Felsenstein 1974; Cockburn 1991; Otto and Lenormand 2002). The simple combination of two parental genomes during sexual reproduction provides a significant increase in genetic diversity. Recombination, the exchange of genetic material between homologous chromosomes, exponentially increases the number of potential offspring genotypes. The benefits of recombination extend beyond increasing genetic diversity. Without recombination, each parent contributes information from only one member of a homologous pair of chromosomes to each of its offspring. Recombination has the potential to combine information from both members of the homologous pair (in the same way separating information contained on a single chromosome) in the production of gametes. This can be beneficial when two advantageous mutations are brought together, or when advantageous and deleterious mutations are separated. This process allows a generation of offspring to contain genetic information from all four of their grandparents. Without recombination only one maternal grandparent and one paternal grandparent would be represented in each homologous chromosome pair of the second generation. Recombination allows each homologous pair to contain information from all four

grandparents (whose chromosomes, in turn, contain information from all four of their grandparents). Mutations in all parental genomes can therefore continue being transmitted in the population, accelerating the rate of evolution, and genomes that have deleterious mutations can contribute to the next generation without transmitting those deleterious mutations (Muller 1932). Therefore, recombination increases the rate at which evolution can occur as well as the number of offspring genotypes. This is vital to both the short and long term success of a population (Cockburn 1991).

Recombination has adaptive advantages (Muller 1932; Van Valen 1973; Felsenstein 1974; Barton 1995). However, crossing over between homologous pairs is also necessary for correct chromosomal segregation in meiosis (Kleckner 1996; Zetka et al. 1999). Recombination occurs during prophase I of meiosis. During early prophase I homologous chromosomes pair up, align loosely and crossing-over occurs. Crossing-over is the reciprocal physical exchange of segments between homologous pairs of chromosomes. The site where crossing over occurs is called a chiasma. Without chiasma formation, chromosomes will not segregate properly during the following phases of meiosis (Zetka, Kawasaki et al. 1999). The reciprocal nature of crossing-over usually prevents the loss or addition of genetic material to recombinant chromosomes. However, on the rare occasion that an unequal cross-over does occur, the resulting gene duplication can provide important raw material for future evolution (Koonin and Galperin 2003).

Recombination Rate on Genomic Maps

Measuring the frequency and location of recombination is an extremely useful tool in genetic research. Recombination rates are extremely variable (as detailed below).

Both physical distances and recombination frequencies between markers are necessary to quantify variability in recombination rate in any given organism. This requires the construction of a physical map of the genome and a genetic linkage map of the genome but these processes are not always entirely independent of one another (e.g., Honeybee Genome Sequencing Consortium 2006).

Physical maps are constructed by aligning contiguous stretches of chromosomal DNA. This map represents actual physical distances, measured in base pairs, between certain sequences of DNA. Physical maps provide an anchor for markers used in genetic linkage maps, as well as serve to resolve ambiguities about closely linked markers (Gibson and Muse 2001). Short stretches of DNA (contigs) are sequenced and these contigs are aligned by matching overlapping sequences. Overlapping sequences are primarily identified by matching restriction fragment length profiles of different clones or by using a common probe to identify clones that are likely to be contiguous. Established genetic linkage maps can also be used to assemble contigs (e.g., Solignac et al. 2007b).

To construct a linkage map two individuals are genotyped at multiple loci across the genome and their offspring are genotyped at the same loci. Whenever an offspring displays a genotype from the maternal (or paternal) grandmother and grandfather on the same chromosome, recombination has occurred in the parental genome. Recombination events between markers can be scored in large, single families or on pedigrees that combine information from multiple families.

In the past, genotyping was done by using classical mutants such as enzyme polymorphisms and visible mutations as genetic markers. However, this method left

many species “un-mappable,” due to low numbers of these classical mutants (Solignac et al. 2004). DNA marker technology has opened the door to map many more genomes. One particularly useful marker class that will be used in my study consists of short tandem DNA repeats called microsatellite DNA markers (Hoy 2003; Solignac et al. 2004). Microsatellites are short sequences of non-coding DNA that have a di-, tri-, or tetra-nucleotide repeating motif. The number of repeats is usually variable, due to errors in DNA replication. Microsatellite alleles are detectable by designing PCR primers which amplify only a specific locus in a given genome and then by analyzing the PCR product sizes from different individuals (Hoy 2003).

When using DNA markers to construct a linkage map, the number of markers that can be used to discriminate between genotypes is theoretically unlimited. The first step in constructing a linkage map with a large number of markers is to sort the markers into linkage groups by their linked or correlated inheritance pattern (Stam 1993). Marker pairs with low recombination frequencies will be assigned to the same linkage group. A computerized search of all marker pairs will establish a number of linkage groups. The higher the stringency of the linkage threshold (i.e., LOD score) the more linkage groups will be generated. Ideally, the number of linkage groups should be the same as the haploid number of chromosomes in the organism. When beginning a linkage map, the threshold recombination frequency should be relatively high, in order to prevent groups of markers on different chromosomes being assigned to the same linkage group (Stam 1993).

The most computationally difficult aspect of developing a linkage map is assigning the correct order to markers within the same linkage group. There are several different approaches software packages take to accomplish this task. With the introduction of each new marker to be ordered, the number of possible orders increases rapidly (for n markers there are (½)n! possible orders). It is computationally expensive and virtually impossible to consider all the possible marker orders for a set of greater than ten markers (Stam 1993). Ten markers cannot provide enough information for a complete linkage map. Therefore, software packages use various "shortcut" algorithms to provide an acceptable marker order in a reasonable amount of time.

Once the markers have been ordered based on pairwise recombination frequencies among them, these recombination frequencies can be translated into map distances (measured in centiMorgans). There are many methods for making this translation, but the two most widely accepted formulas were developed by Haldane and Kosambi. Assuming no interference between recombination events, Haldane developed the mapping function $r = ½(1-e^{-2x})$, where r is recombination frequency and x is map distance (Haldane 1919). However, interference between recombination events has been observed in many organisms (for instance, Solignac et al. 2007b). Positive interference is the suppression of recombination events in the neighborhood of a given one. Negative interference refers to the opposite: a given recombination event encouraging further recombination in the same area (Stam 1993). Assuming independence of recombination events (no interference) results in higher numbers of double recombinants than actually observed in most organisms (Felsenstein 1979). This led to the conclusion that positive

interference is common and widespread. The relationship between recombination frequency and map distance is dependent on the strength of interference, which is variable, but Kosambi's mapping function (which takes positive interference into account) seems to work well for many organisms (Felsenstein 1979). His mapping function is as follows: $r = \frac{1}{2} \tanh(2x)$ (Kosambi 1944). This results in a shorter map distance for a given recombination frequency than Haldane's method and has been shown to be the correct model in the genus *Apis* (Solignac et al. 2007b)

Variability in Recombination Rate

For structural reasons one recombination event is required per homologous pair of chromosomes in meiosis. Therefore, recombination might be interpreted as simply a byproduct of chromosomal pairing in meiosis. This idea is supported by the fact that the number of chiasmata formed between homologous chromosomes is highly correlated across a wide range of taxa. Most organisms have about 1.6 recombination events per chromosome pair per meiosis (Baker et al. 1976).

However, the frequency at which recombination occurs is variable within and between species which suggests that recombination is not simply a structural necessity for meiosis but evolves in response to selection. Comparing linkage and physical maps helps quantify this variation. For species in which genetic and physical maps can be compared, it has become clear that the physical equivalent of 1cM varies between and within genomes. For instance, in human automsomal euchromatin, 1cM represents approximately 0.91Mb (Kong et al. 2002), while in the honey bee 1cM corresponds to approximately 0.05Mb (Solignac et al. 2007a).

Comparing recombination rates systematically across kingdoms provides further evidence for both adaptive and structural causes of variation in recombination rate (Wilfert, Gadau, and Schmid-Hempel 2007). Overall, the highest recombination rates are found in the fungi and protozoa, both of which are characterized by small physical genomes. For instance, a yeast-like fungus (*Cryptococcus neoformans)* demonstrates a recombination rate of 75cM/Mb (Marra et al. 2004). Other fungi have rates ranging from 25.7cM/Mb to 37.4cM/Mb (Kullman, Tamm, and Kullman 2005). All other taxa show considerably lower recombination rates. Humans, for example, have a recombination rate of 1.1 cM/Mb while the recombination rate in mice has been found to be 0.4cM/Mb (Dietrich et al. 1996; Kong et al. 2002). Plants and most insects also fall in the lowest range of recombination rates. *Pinus pinaster* (maritime pine) has a recombination rate of 0.07cM/Mb (Chagne et al. 2002) and the parasitic wasp, *Nasonia vitripennis*, has a recombination rate of 2.5cM/Mb (Gadau, Page, and Werren 1999). Though the social insects of the order Hymenoptera (discussed in detail below) are closely related to insects such as *N. vitripennis* they fall into a third range of recombination rates, lower than the fungi and protozoa but higher than all other taxa (Wilfert, Gadau, and Schmid-Hempel 2007).

This distribution of recombination rates suggests both adaptive processes and structural features influence recombination rate. The correlation between small genome size and high recombination rates in the fungi and protozoa suggests a structural control of recombination rate, while the high recombination rates of the social insects (a group

that shares behavioral, but not physical characteristics) suggests adaptive control of recombination rate (Sirvio et al. 2006; Wilfert, Gadau, and Schmid-Hempel 2007).

Comparisons of recombination rate between related species can reveal structural influences on recombination rate. This has been demonstrated in *Drosophila* sister species. The frequency and distribution of recombination among *D. mauritiana, simulans* and *melanogaster* were shown to be significantly different (True, Mercer, and Laurie 1996). Even between homologous portions of the genome significant differences were found in recombination rate with *D. mauritiana* having (on average) a recombinational size 1.8 times greater than that of *D. melanogaster* and *D. simulans* consistently demonstrating an intermediate size (True, Mercer, and Laurie 1996). *D. mauritiana* showed little centromeric suppression of recombination rate, while this effect was significant in *D. melanogaster*. Structural variations may be responsible for these differences in recombination rate. Centric heterochromatin may be responsible for suppression of recombination rate around centromeres and has been known to evolve rapidly between species, making it a good candidate to explain the differences seen in the above study (John and Miklos 1979; John 1988; True, Mercer, and Laurie 1996).

Similar variation in recombination rate, although probably due to adaptive causes, has been observed within a mouse species. Dumont et al. (2009) used genotype data from over 10,000 SNP markers in approximately 2300 genetically admixed stock mice from 85 families to determine the average number of crossovers per meiosis in male and female mice. They found considerable variability in recombination rate, with female crossover averages ranging from 9.0-17.3 and male crossover averages ranging from 7.7-

14.7. Variation in recombination rates was heritable, but only approximately 45% of the variation could be explained by heredity (Dumont, Broman, and Payseur 2009). The high degree of relatedness between individuals in this study makes it unlikely that structural features explain the additional variability.

In plants, domesticated species have been shown to have higher recombination rates than their wild progenitors (Ross-Ibarra 2004), again suggesting recombination rate is influenced by adaptive processes. Recombination rate may be controlled by a theoretical "recombinational load" above which recombination is selected against (Barton 1995). Evidence for this idea is seen in plant species and the same evidence suggests that increased recombination in domesticated plants is in fact a product of selection and not of other causes, such as increased homozygosity (Ross-Ibarra 2004). Domestication has also been shown to increase chiasma frequency in mammals (Burt and Bell 1987). Increased recombination rate may be an adaptation to environments with intense selective pressure for novel combinations of traits (Van Valen 1973).

This diversity of recombination rates across taxa without ubiquitous structural genome correlates suggests that both structural features and adaptive processes affect recombination rate. Similarly, the variation of recombination rates within particular genomes also suggests both structural and adaptive control of recombination rate.

Within genomes structural features have been shown to exert some control over recombination frequency and chiasma location (Dumas and Britton-Davidian 2002). The distribution of chiasmata over chromosomes is known to be nonrandom (Kaback et al. 1992) and influenced by factors such as nucleotide content (Eyre-Walker 1993),

chiasma interference (Lawrie, Tease, and Hulten 1995; Gorlov and Gorlova 2001), and sex (Hawley, McKim, and Arbel 1993; Hassold, Sherman, and Hunt 2000). In human, mouse and rat genomes GC content and number of CA repeats are highly correlated with recombination and recombination is often suppressed near the centromere and elevated near the telomeres (Jensen-Seaman et al. 2004).

Structural features may also explain some of the variation in recombination rate between chromosomes within a genome. Chromosome form and size influence chiasma number as has been found in yeast and humans (Kaback et al. 1992; Kaback 1996). Human metacentric chromosomes seem to require two crossovers (Hassold et al. 2004) and at least one chiasma per chromosomal arm is required to ensure correct chromosomal segregation during meiosis (John 1990; Koehler et al. 1996; Hassold, Sherman, and Hunt 2000; Paliulis and Nicklas 2000).

Some variation in recombination rate may be caused by chromosome size. In many species there is a significant tendency for the smallest chromosomes to recombine more often (per Mb) than the largest ones. Among mammals, this tendency has been shown to be more significant in humans than in rodents (Jensen-Seaman et al. 2004). An extreme example is the chicken, where the recombination rate in the smallest chromosome is approximately ten times larger than in the largest chromosomes (International Chicken Genome Sequencing Consortium 2004). Notably, there is no relation between chromosome size and recombination rate in the honey bee (Beye et al. 2006; Solignac et al. 2007a).

Though many structural features play a role in determining the location and likelihood of chiasma formation, some of the observed variability in recombination rate suggests that adaptive processes are at work. For instance, distance from the telomere is the best indicator of recombination rate in the human genome, but in the rat and mouse genomes this variable explains much less of the variation in recombination rate than other sequence parameters (Jensen-Seaman et al. 2004). The overall recombination rate is significantly different between species with humans having about twice as much recombination per generation as the rat or mouse (Jensen-Seaman et al. 2004). In humans, higher recombination rates are associated with higher neutral mutation rates (Hardison et al. 2003), though a causal relationship has yet to be determined. In the honey bee recombination rate has also been shown to correlate positively with distance between genes, though not with gene density (Beye et al. 2006). This may be due to larger introns in genes in areas of low recombination. Intron length and recombination rate are negatively correlated in *Drosophila* and humans (Comeron and Kreitman 2000; Beye et al. 2006)

The evolution of sex chromosomes (reviewed in Bergero and Charlesworth 2009) also suggests that recombination rate responds to selection. Recombination between heterogametic sex chromosomes (X and Y in humans, for instance) is restricted to small regions of the chromosomes. This is probably an adaptive response to prevent recombination between sex-determining loci when such recombination could cause sterility or hermaphroditism (Charlesworth 2002).

In honey bees, sex is determined by the complimentary sex determination gene (Beye et al. 1999; Beye et al. 2003), not sex chromosomes. Recombination rate is inflated at the sex-determination locus (Beye et al. 1999). This is due to negative frequency dependent selection because individuals who are homozygous develop into infertile males with zero fitness. Since rare alleles at the sex-determination locus (*csd*) are selected for, a high recombination rate around this locus works to prevent deleterious mutations surrounding a rare *csd* allele from being incorporated into the genome. The association of high recombination rate with strong diversifying selection suggests an adaptive explanation for the increased recombination around the *csd* in honey bees (Beye et al. 1999).

Variation in recombination rate within genomes provides a strong argument for adaptive control of recombination. Structural correlates with recombination rate within a genome are limited to features such as GC content and distance from the telomere and these features fail to completely explain the wide range of recombination rates demonstrated within single genomes.

Comparisons in recombination rate between genomes with significant structural differences also provide evidence suggesting adaptive processes control recombination rate. The consistently high recombination rate in the social Hymenoptera provides a compelling argument for the role of natural selection (Gadau et al. 2000; Wilfert, Gadau, and Schmid-Hempel 2007). Recombination rates are consistently higher in the social vs. non-social hymenopterans (Wilfert, Gadau, and Schmid-Hempel 2007) and this high rate

has been maintained in multiple phylogenetic groups and appears to be unassociated with genome size or chromosome number (Crozier 1977).

Sociality and Recombination Rate

To date, recombination rate has been studied in four social insect species. The first studies in *Apis mellifera* (the European honeybee) suggested recombination rates about 19cM/Mb (Hunt and Page 1995) and current estimates based on more complete information are about 22cM/Mb (Beye et al. 2006; Solignac et al. 2007a). Another species of the same family (Apidae), the primitively eusocial bumble bee (*Bombus terrestris),* has a recombination rate of approximately 4.4cM/Mb (Gadau et al. 2001). Bumble bees have a lower degree of social complexity than honey bees with smaller nest sizes, an annual life cycle, and a lower degree of caste differentiation. These results suggest a correlation between recombination rate and sociality. *Acromyrmex echinatior* (a leaf cutter ant), is another member of the social Hymenoptera (with a colony structure well developed and similar to that of *A. mellifera*) who demonstrates a high recombination rate of 6.7cM/Mb (Sirvio et al. 2006). Though not as high as *A. mellifera*, this recombination frequency does support a correlation between sociality and recombination rate. Data have also been collected for another eusocial ant species, *Pogonomyrex rugosus,* and its recombination rate of 14cM/Mb is similar to that of the honey bee (Sirvio et al. 2006).

Three classes of hypotheses have been postulated to explain the high recombination rate in the social Hymenoptera (Sirvio et al. 2006):

(1) Hypotheses referring to genomic features such as haplo-diploid sex determination or the number and/or size of chromosomes in the genome. Haplo-diploidy is shared across all Hymenoptera and would not sufficiently explain the variability within the order because some non-social species have low recombination (Gadau et al. 2000). The social Hymenoptera have been shown to have higher chromosome numbers than their non-social counterparts (Crozier 1977; Wilfert, Gadau, and Schmid-Hempel 2007). Though this structural feature may conceivably play a role in difference in recombination rates between social and non-social Hymenoptera, it does not explain the variation seen within the social insects, because chromosome number and genome-wide recombination rates are not correlated (Gadau et al. 2001; Sirvio et al. 2006; Solignac et al. 2007a).

(2) Hypotheses based on the variance-reducing effects of recombination (Fig. 1). Recombination makes kinship more homogenous among colony members (Gadau et al. 2000; Sirvio et al. 2006), and this is potentially an important factor in the evolution of sociality by kin selection (Sirvio et al. 2006). High levels of relatedness are important in the evolution of eusociality (Hughes et al. 2008). However, once eusociality has been irreversibly established (by workers losing their reproductive capabilities) these high levels of relatedness seem to be less important, as is demonstrated by the appearance of multiple mating (Hughes et al. 2008). In modern eusocial societies the majority of variance between individuals in a colony is between patrilines and a direct consequence of the multiple mating of the queen (Crozier and Page 1985; Winston 1987). If high levels of relatedness were important to maintain eusociality, modification in mating

behavior of the queen presumably would produce a much larger variance-reducing effect than recombination does. The persistence of the multiple mating behavior of the queen suggests that variance between individuals does not have severe detrimental consequences for the colony (Hughes et al. 2008). Also, nepotism and kin recognition do not play a large role in honey bee swarms (Kryger and Moritz 1997), queen rearing (Breed, Welch, and Cruz 1994; Tarpy and Fletcher 1998) or worker rearing (Noonan and Kolmes 1989). Certainly, it is possible that the variance reducing effects of recombination mediate some of the nepotistic pressure on honey bee colonies (and other social insect societies). However, the lack of observable nepotistic behavior in honey bees suggests that the high recombination rate observed in *A. mellifera* is not primarily due to its variance reducing effects.

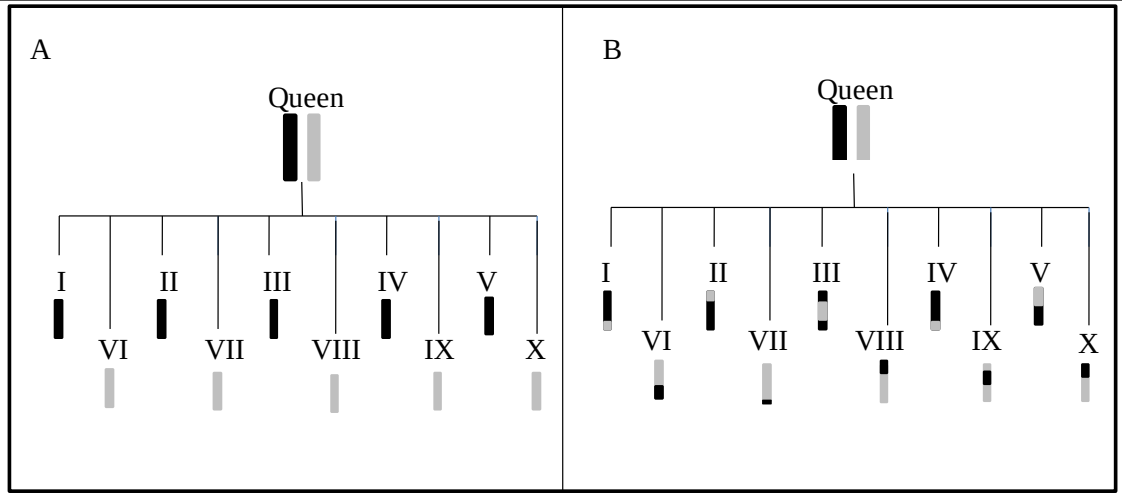


Figure 1: Variance reducing and diversity increasing effects of recombination. By looking at the theoretical inheritance of one chromosome we see the effects of recombination on the genetic similarity of haploid drones laid by a single queen. Without recombination (A) there would be two groups of genetically identical drones. Recombination (B) reduces the variance between offspring because all individuals share at least some genetic material and increases the diversity between offspring by creating novel multilocus genotypes

(3) Hypotheses dependent on the diversity-increasing effects of recombination (Fig. 1). The benefits of increased genetic diversity as a result of multiple mating have been well documented. Increased genetic diversity can reduce diploid male production (Page 1980; Tarpy and Page 2001), increase disease resistance (Brown and Schmid-Hempel 2003; Seeley and Tarpy 2007) and enhance the division of labor in social insect colonies (Mattila and Seeley 2007; Oldroyd and Fewell 2007). Recombination creates novel genomic combinations in the offspring and this increased genetic diversity has the potential to also benefit these colony-level traits. The high levels of recombination seen in the social insects (Wilfert, Gadau, and Schmid-Hempel 2007) suggest that the diversity-increasing effects of recombination play a key role in eusocial societies. That role is still unclear and comparative studies (such as this one) are needed to gather more information.

The domestication of *A. mellifera* (and subsequent use of the species in research) has been proposed as an explanation for its extraordinarily high recombination rate (Schmid-Hempel and Jokela 2002). Little is known about the effects of domestication on recombination rates in insects. The high recombination rates seen across the social Hymenoptera are not an artifact of domestication, but there is a possibility that domestication has influenced the extremely high rate in *A. mellifera.* Studies in non-domesticated honeybee species should shed light on this issue.

Honey Bee Biology and Phylogeny

There are at least 9 species within the genus *Apis* (Raffiudin and Crozier 2006) which allows for detailed comparative studies of the high recombination discovered in *A.*

mellifera. In addition to the construction of multiple linkage maps in *A. mellifera*, its genome has been completely sequenced, assembled and annotated (Honeybee Genome Sequencing Consortium 2006) and its biology is understood in detail, thanks to a long and rich tradition of using this species in research (Winston 1987).

Honey bee colonies contain three classes of individuals. Each colony has one female queen who is responsible for reproduction for the entire colony. At the beginning of her life, she mates with multiple drones (male bees) whose sole responsibility is to provide sperm to a queen outside of their own colony. All non-reproductive tasks of the colony are performed by workers, which are the alternative female caste to the queen. Upon emerging from a cell, a worker bee begins her life doing in-hive tasks such as cleaning cells or caring for brood. At approximately 20 days old a worker bee will begin to forage for pollen or nectar outside of the hive (Winston 1987). Genetic and environmental factors affect the age of first foraging and foraging preference of individual workers (Robinson 2002; Rueppell et al. 2004). The age of first foraging is a significant predictor of worker lifespan (Rueppell et al. 2008) and worker lifespan is also dependent on seasonal factors. In the summer, workers live for approximately three weeks, while in the winter they can live for up to 20 weeks (Winston 1987).

Like all Hymenoptera and some other insects, honey bees display haplo-diploid sex determination: males are haploid and arise from unfertilized eggs and females are diploid and thus arise from fertilized eggs. This not only shapes the genetic structure of colonies with important consequences for kin selection (Winston 1987; Oldroyd and Wongsiri 2006) but also makes honey bee drones ideal genetic mapping populations

(Page, Gadau, and Beye 2002; Solignac et al. 2007a) which will be discussed further below.

Individuals in the same colony are highly related because, under normal circumstances, they are all offspring of the resident queen. In genetic terms, drones are living representations of one of the many potential queen haplotypes, arising directly from her meiotically produced, unfertilized eggs. The haploid nature of the drone dictates that each of his offspring will have identical paternal genetic information. These offspring will also share 50% of the genetic information inherited from the queen, resulting in 75% total shared genetic material between workers of the same patriline. Such worker bees are "super-sisters" (Page and Laidlaw 1988) and worker bees with different drone fathers are half-sisters, sharing only 25% of their genetic information.

Colony organization, relatedness among individuals, and most other aspects of basic biology are shared by all *Apis* species. Though it is often difficult to distinguish between species, using the biological species concept (Mayr 1942) and additional genetic information (Raffiudin and Crozier 2006), at least 9 honey bee species have been determined that fall into three distinct groups (Table 1). Only two of these species (*A. mellifera* and *A. cerana)* have been domesticated**.** Despite their common biology, there are important differences between the *Apis* species, such as colony size, nest location, and dance behavior (Oldroyd and Wongsiri 2006).

Group	Subgenus	Species	Common Name
Dwarf honey bees	Microapis	*Apis florea* *Apis andreniformis*	Red dwarf honey bee Black dwarf honey bee
Giant honey bees	Megapis	*Apis dorsata* *Apis laboriosa*	Common giant honey bee Giant mountain honey bee
Cavity-nesting honey bees	Apis	*Apis cerana* *Apis koschevnikovi* *Apis nuluenis* *Apis nigrocincta* *Apis mellifera*	Eastern hive bee Red honey bee Mountain honey bee Sulawesian honey bee Western honey bee

Table 1: *Apis* species can be divided into three distinct taxonomic groups.

Using genetic data from three mitochondrial and one nuclear gene Raffiudin and Crozier (2006) established a three-group phylogeny for the *Apis* genus: dwarf, giant and cavity nesting bees. Molecular data shows the dwarf honey bees as basal to the other two groups, the giant honey bees and the cavity nesting bees (Raffiudin and Crozier 2006). This is in agreement with previous phylogenies established using morphological data (Alexander 1991; Engel and Schultz 1997) and molecular sequence analysis (Arias and Sheppard 2005). This study will focus on a comparison between *A. mellifera* and one of the undomesticated basal dwarf species, *A. florea.*

Though *A. florea* is significantly smaller than *A. mellifera* and *A. florea* queens are about three times the size of workers (Oldroyd and Wongsiri 2006), differences between the species extend beyond physical size. *A. florea* has a distinctive red colored abdomen and other differences between the species include nest location and construction, and waggle dance form. *A. florea* builds a single comb hive that surrounds a branch or a twig, while *A. mellifera* has a multiple comb nest, usually built in a cavity (Oldroyd and Wongsiri 2006; Raffiudin and Crozier 2006). *A. florea* uses a horizontal stage to perform its horizontal waggle dance to communicate locations of food (Koeniger

et al. 1982), while *A. mellifera* employs a vertical dance style unless its dance stage is experimentally manipulated (Lindauer 1961).

All groups of honey bees are highly susceptible to pathogen infestation (Seger and Hamilton 1988). Bacterial diseases infect bees on every continent and multiple viruses infect honey bees, including several types that cause paralysis (Graham et al. 1992). Honey bees are also susceptible to fungal and protozoan diseases, some of which can be exacerbated by cold or damp living conditions (Root and Root 1980). Honey bees are especially vulnerable to infestation by mites. *Varroa jacobsoni* and *Acarapis woodi* are two species of mites which cause serious problems for European bees in the United States (Root and Root 1980). Different groups of honey bees are more susceptible to certain mites, this is dependent on factors such as life history, nest location and perhaps chemical production (Root and Root 1980).

More pathogens of *A. mellifera* have been identified than of any other honey bee species. This may be due to the fact that *A. mellifera* has been more thoroughly studied than other *Apis* species. However, it is possible that due to factors such as nest location, migratory behavior and cleaning habits (Woyke, Wilde, and Reddy 2004) more pathogens exist that target *A. mellifera* than other bee species. Many pathogens affect all *Apis* species. However a review of parasites in the social insects (Schmid-Hempel 1998) identifies 3 viruses, 11 bacteria, 9 fungi, 5 protozoa, 1 nematode and 9 mites that are specific to *A. mellifera.* In contrast, only 1 protozoan and 1 or 2 mites have been identified as specific to *A. florea* (Schmid-Hempel 1998; Oldroyd and Wongsiri 2006).

A higher pathogen pressure in cavity nesting bees, such as *A. mellifera*, is also suggested by the occurrence of more sophisticated hygienic behavior in cavity nesting bees than open air nesting species. Cavity nesting bees (*A. cerana* and *A. mellifera*) are more likely to uncap and remove diseased brood from their cells. Open air nesting bees (*A. dorsata* and *A. laboriosa*) have been shown to leave diseased brood cells capped and cease using that part of their comb (Woyke, Wilde, and Reddy 2004). Woyke et. al. (2004) hypothesize that this is due to the more frequent migrations observed in open-air-nesting species. Here, I hypothesize that frequent migrations and nest abandonment (Oldroyd and Wongsiri 2006) and the open-air nest environment lower the pathogen pressure on open-air nesting species such as *A. florea*.

Linkage Mapping in A. mellifera

Much work has confirmed the exceptionally high recombination rate in *A. mellifera*, based on the construction of multiple genetic linkage maps (Table 2). The first microsatellite-based linkage map in *A. mellifera* established 24 linkage groups and estimated the total genome size to be 4061.2 cM, (intermediate between the previous estimates) with a recombination rate of approximately 22cM/Mb (Solignac et al. 2004). Solignac et. al (2007) published a more complete microsatellite based linkage map with 2,008 markers and the recombinational size and recombinational frequency of the genome remained constant, around 4000cM and 22cM/Mb respectively. Microsatellites are a better resource for anchoring a genetic map on a physical map than other genetic markers (such as AFLPs and RAPDs) and should therefore provide a more accurate estimate of recombination rate (Solignac et al. 2004). Some of these markers have

already been shown to amplify in other *Apis* species (Solignac et al. 2003). This information will allow for calculating the recombination rate in *A. florea*, which is the goal of my study.

Authors, year	Markers	Linkage Groups	Total Genome Size	Recombination Rate
Hunt and Page, 1995	365 RAPD	26	3450cM	19cM/Mb
Rueppell et. al, 2004	387 AFLP	35	4610cM	25cM/Mb
Rueppell et. al, 2004	396 AFLP	38	4527cM	24cM/Mb
Solignac et. al, 2004	541 micro-satellites	24	4061.2cM	22cM/Mb
Solignac et. al, 2007	2,008 micro-satellites	16	4,114.5cM	22cM/Mb

Table 2: Summary of linkage mapping in *A. mellifera*.

Hypotheses and Predictions

The comparison of global and local recombination rates between *A. florea* and *A. mellifera* will provide information about possible explanations for the high recombination rates in social insects and in particular in the honey bee. My main hypotheses are that pathogen pressure, division of labor and /or domestication could have selected for high recombination rate in *Apis mellifera*. Accordingly, I predict:

1) If the high recombination rate observed in *A. mellifera* is due primarily to pathogen pressure, a significantly lower recombination rate is expected in *A. florea* due to its lower pathogen pressure.

2) If the high recombination rate observed in *A. mellifera* is due primarily to domestication, a significantly lower recombination rate will be observed in *A. florea* because it has never been domesticated.

3) If the high recombination rate observed in *A. mellifera* is due primarily to division of labor comparable recombination rates will be observed in *A. florea* due to its equally complex division of labor.

By comparing marker order between the species this study I will also investigate evolutionary synteny and could provide support for the current *A. florea* genome assembly project.

CHAPTER II

MATERIALS AND METHODS

Collection

The mapping population of *Apis florea* that provided the biological material for this project has been identified and collected in 2007 by Dr. N. Koeniger in Thailand (in Mahasarakham, 300km North-East of Bangkok). About 220 drone pupae were directly collected from one comb of an *A. florea* colony and stored in 95% ethanol in individual tubes. After transport to UNCG, the samples were transferred to -80 °C until DNA extraction.

Drones were collected because they make an excellent mapping population. Being haploid, drones are a living representation of a particular haplotype found in the queen. All variation between drones is a consequence of independent assortment of chromosomes and/or recombination during meiosis. In contrast, variation between workers can be attributed to different patrilines within one colony. Workers can be used as a mapping population, but drones are preferable when available.

Extraction

DNA was extracted from the thoraxes of 120 drones using a Qiagen DNeasy Blood and Tissue Kit. The manufacturers' protocol was followed exactly using a

disposable polypropylene pestle for homogenization of the tissue. After obtaining the 200 µl eluate according to the kit protocol, the DNA concentration was determined using an average of two duplicate readings on a Nanodrop® instrument, and DNA dilutions were prepared to give a final concentration of 10 ng/µl in TE buffer.

Genotyping

To collect informative genetic markers for map construction I first screened primer pairs that had been characterized to amplify microsatellite loci in *A. mellifera* (Solignac et al. 2007a) for their amplification and variability in my specific *A. florea* mapping population. This screening process included two different PCR and electrophoresis protocols because I used primers that were either labeled for detection on a MegaBACE® or LiCor® DNA analyzer, depending on their availability in the lab. Eight individuals were used in each initial screen. Subsequently, every polymorphic marker was genotyped in each member of a randomly selected mapping population of 96 individuals using either the MegaBACE® or LiCor® protocol. The detailed protocols for the two systems are described in the following sections.

PCR Amplification

MegaBACE® Sequencer

Forty-six MegaBACE® primers previously known to amplify loci in the *A. mellifera* genome were screened for amplification in *A. florea.* For each screen, 1µl of extracted DNA (1 ng/µl) was be combined with 9.3µl ddH20, 1.5µl 10x PCR buffer, 1.5µl 2mM dNTP, 0.75µl 5mM forward primer, 0.75µl 5mM reverse primer and 0.2µl Taq polymerase enzyme. The standard PCR program used was a touch-down program.

The specific protocol used was 4min 45sec at 95°C for initial denaturing, followed by 5 cycles of 4min at annealing temperature (beginning at 68°C for the first cycle and decreasing by 2°C for the next four cycles), 1min at 70°C for elongation and 45sec at 95°C for denaturation. These 5 cycles were followed by 5 cycles of 2min annealing temperature (beginning at 58°C and decreasing by 2°C each cycle), 1min at 70°C for elongation and 45sec at 95°C for denaturation. This was followed by 25 cycles with 2min annealing at 50°C, 1min at 70°C for elongation and 45sec at 95°C for denaturation, with 4min at 70°C for final extension.

LiCor® Sequencer

I screened 460 primers (from sites throughout the *A. mellifera* genome) designated for the Licor® sequencer following a slightly different PCR protocol. 1µl of extracted DNA (1ng/ µl) was combined with 8.95µl ddH20, 1.5µl 10x PCR buffer, 1.5µl 2mM dNTP, 0.35µl 5mM forward primer, 0.75µl 5mM reverse primer, 0.75µl IRD700-labeled M13 oligonucleotide (LiCor®) and 0.2u Taq polymerase enzyme. The labeled universal M13 oligonucleotide incorporates into the PCR product at lower annealing temperatures and allows the product to be read by the automated sequencer. These reactions were run with the same touch-down PCR protocol as the samples for the MegaBACE® sequencer.

Amplification Detection

The PCR products for each sequencer were then subjected to agarose electrophoresis in 25cm gels consisting of 0.5% Synergel and 0.3% agarose dissolved in 0.5X TBE. Gels were run at 200 volts for 90 minutes and stained with ethidium bromide

for 30 minutes to determine if there was an amplification product and determine its approximate size. The primers that produced non-specific products, or no product at all, were generally eliminated from subsequent consideration.

Polymorphism Detection and Subsequent Genotyping

MegaBACE® Sequencer

Nineteen primers amplified during the PCR protocol were therefore screened on the MegaBACE® sequencer for polymorphisms, using eight individuals. Following PCR amplification the samples were multiplexed according to size and desalted using a Millipore desalting plate. After desalting, 2µl of each sample was combined with 7.75µl 0.1% Tween buffer and 0.25µl Et-Rox size standard. The MegaBACE® sequencer was run using the genotyping application and the sequencer protocol was followed through the "Inject Samples and Run" protocol (Amersham Biosciences Corp 2002).

Screening revealed two polymorphic markers which were amplified in each individual of the mapping population using the PCR protocol described above. The PCR products were then analyzed on the MegaBACE® sequencer following the procedure described for screening above. Genotypes were determined using FragmentProfiler® and scoring was done manually.

LiCor® Sequencer

One hundred forty-six primers amplified during the PCR process and were then screened for polymorphism on the LiCor® sequencer, using six individuals. 1µl of each PCR was multiplexed according to product size (avoiding any overlap between loci) and diluted with 4µl of 1x PCR buffer. 2.5 µl LiCor® loading buffer was added to each

multiplex. These samples were loaded into a polyacrylamide gel, along with a 50-350bp LiCor® size standard. The LiCor® sequencer was run at for 90-120 minutes at default settings for temperature, voltage, current and power (45°C, 1500 volts, 40 amps and 40 watts). Upon completion of a run the gels were scored by eye, using Gene Profiler® software.

Screening identified 18 polymorphic markers which were amplified in each member of the mapping population. The PCR products from these individuals were genotyped following the procedure for screening described above.

Based on the limited number of polymorphic loci I was able to genotype in the entire genome, I subsequently focused my efforts on loci found on chromosomes 3 and 12 in the *A. mellifera* genome. I screened all remaining 101 microsatellite loci that had been identified on chromosome 3 (Solignac et al. 2007a), resulting in seven additional polymorphic loci that were amplified and scored in the entire mapping population. These additional screens left two large (<50cM) linkage gaps (analysis procedures are described below). Sixteen new primer pairs were designed with Primer3 software around microsatellites that were directly identified from the *A. mellifera* genome sequence, specifically in the two poorly covered regions. These primers were screened for amplification and polymorphism following the protocol described above and one polymorphic marker was successfully scored in the *A. florea* population. Similarly, I also screened all additionally available 77 microsatellite primer pairs from chromosome 12 (Solignac et al. 2007a) resulting in 10 additional polymorphic loci that were then amplified and scored in the entire mapping population.

Analysis

Genome-wide recombination rate screen

Pairwise distances for linked loci from the initial genome-wide primer screen were calculated using JoinMap 4.0 (Stam 1993). Marker data was entered in both linkage phases and the calculations were done phase-unknown. Linkage was determined using the default program settings of a LOD score greater than 2.0 and map distances for linked markers were determined using the Kosambi function.

The recombinational distances of these intervals were tested for significant differences between *A. florea* and *A. mellifera.* Recombinational distances given by MapMaker3.0 (*A. florea)* or reported by Solignac (*A. mellifera)* were converted into recombinational frequencies using the Kosambi function and I calculated the number of individual crossovers in each interval based on these frequencies and the respective sample sizes. For each interval the number of recombinant and non-recombinant individuals was compared between species using 2x2 contingency tables. P-values were calculated using Fisher's Exact Test and the Bonferroni correction was used to account for multiple tests. To provide descriptive statistics, recombination rates for intervals in *A. florea* were calculated using the recombinational distances provided by the software and the physical distances known in *A. mellifera.*

Map Construction in Chromosomes 3 and 12

Genotypic data for each polymorphic loci was used for manual linkage mapping, using MapMaker3.0 (Lander et al. 1987; Lincoln, Daly, and Lander 1992). Marker data

was entered in both phases because the linkage phase was not known and default linkage criteria were relaxed to <50cM and LOD scores > 1. First, all marker linkages were investigated with the "near" command. Using this data, linkage groups were established and all possible orders of these groups were tested manually, using the "map," "sequence," "drop marker," and "try" commands.

All marker data used to construct the chromosomal maps resulted from at least two consensus genotyping runs. Markers that could not be genotyped successfully in more than one run were generally eliminated from consideration (see Discussion for details on individual markers).

Interval comparison

Each interval between two polymorphic, mappable microsatellites found in *A. florea* was compared to the corresponding interval in *A. mellifera.* Significant differences in map distance were identified using 2x2 contingency tables and the Fisher's Exact Test as outlined above.

A sign test was used to determine if there was a significant difference between the overall number of intervals larger in *A. mellifera* and the number of intervals larger in *A. florea.*

Additional Comparisons

To eliminate spurious results due to methodological differences between my *A. florea* map intervals and those of Solignac et. al (2007) for *A. mellifera,* the *A. florea* data was also compared to two additional *A. mellifera* mapping populations. The Social Insect Lab at UNCG had previously constructed maps of two *A. mellifera* crosses

(designated as P1 and P2 throughout the remainder of this thesis) using a low density of SNP and microsatellite markers with methods similar to those described here. The markers used for map construction in these populations were not identical to those used in *A. florea* so the chi-square test could not be used to test for significant differences.

To ensure that maps constructed with a low density of markers are not systematically larger than Solignac et. al's map I tested for correlation in recombination rates between P1, P2 and Solignac's population. To do so, I first calculated a sliding recombination rate window of 1Mb for the highly resolved Solignac map. The SNP markers used in constructing the P1 and P2 maps were not included in Solignac's map; however, the sliding window allowed for estimation of recombination rate across intervals that corresponded to the marker intervals in P1 and P2. I then calculated the Spearman correlation coefficient between recombination rates in P1 and P2, P1 and Solignac et. al, and P2 and Solignac et. over identical physical intervals using SPSS16.0. I also tested for significant differences in average map distances between P1, P2 and Solignac et. al's mapping population using a paired samples t-test.

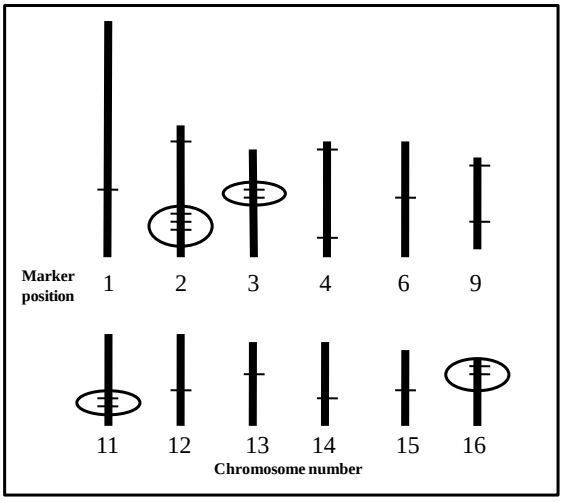
CHAPTER III

RESULTS

Genome-Wide Recombination Rate

Before screening for additional markers on chromosome 12 and chromosome 3 was completed 20 loci had been successfully amplified in 96 *A. florea* individuals. Of these 20 loci, there were five pairs that were linked in *A. mellifera* (Fig. 2). Four of these pairs were also linked in *A. florea* and one pair had a significantly different recombinational distance (Table 3, Fig. 3). On chromosome 2, the interval between AC033 and BI251 was found to be significantly different between *A. mellifera* and *A. florea.* In *A. florea* this interval is more than twice as long as in *A. mellifera.* One interval on chromosome 11 was linked at 20.4cM in *A. mellifera* and was not linked in *A. florea*. Without a linkage distance the number of crossovers in that interval could not be calculated and statistical significance could not be established.

Figure 2: Distribution of initial amplified markers in *A. florea*. 20 markers were distributed across 12 chromosomes. Markers linked at <50cM in *A. mellifera* are circled.



Interval	Chr	Physical distance in *A. mellifera* (bp)[a]	*A. florea* recombinational distance (cM)[b]	*A. florea* sample size[c]	*A. mellifera* recombinational distance (cM)[d]	*A. mellifera* sample size[e]	P-value[f]	*A. florea* rec. rate (cM/Mb)[g]	*A. mellifera* rec. rate (cM/Mb)[e]
A008-Ac033	2	916126	19.0	71	32.9	187	0.1113	20.8	35.9
AC033-BI251	2	78484	24.3	71	9.3	187	0.0062*	309.4	118.5
SV171-K0230	3	54291	3.1	68	7.1	92	0.4684	57.6	130.8
K1168-BI001	11	761632	unlinked	56	20.4	92	unlinked	unlinked	26.8
Ap217-Ap057	16	1274261	27.5	81	14	92	0.0846	21.6	11.0

Table 3: Comparison of linkage distances in *A. mellifera* and *A. florea*. Across the five intervals linked in *A. mellifera* 4 were also linked in *A. florea*. Intervals between microsatellite markers were compared for differences in recombination distances between *A. mellifera* and *A. florea*. [a]Calculated from data presented in Solignac, Mougel et al. 2007. [b]From JoinMap4.0. [c,e]Number of individuals that were scored for both markers of the interval. [d,e]From Solignac, Moguel et al. 2007. [f] Boneferroni corrected significance level: $P \leq 0.0125$. [g]Assumed conservation of physical size between genomes.

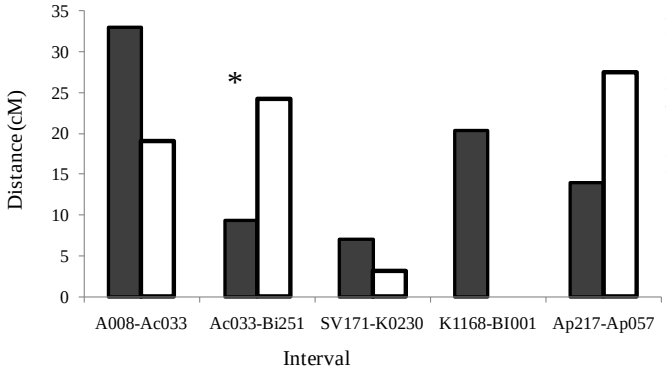


Figure 3: Genome wide comparison of recombinational distances. Empty bars represent *A. florea*. Significant difference in the interval between AP217 and AP057 (P=.0062). K1168 and BI001 were not linked in *A. florea*.

Chromosome Models

Chromosome 3

Of the 141 microsatellites found on chromosome 3 in *A. mellifera* 10 scoreable polymorphic loci were found in *A. florea.* The data from these 10 loci were entered in MapMaker3.0 in both phases and using the "near" command all linkages to each marker were identified (Table 4).

BI257 and AP036 failed to link to any other marker in chromosome 3 and were therefore excluded from further analyses. Using MapMaker3.0 all possible orders of the 8 remaining markers were compared. The null model (conserved *A. mellifera* order) and the most likely order are reported in Table 5. The likelihood scores between these two models are not significantly different. Therefore, the null model was used for further analyses.

Position in *A. mellifera* (cM)	Distance (cM) \ LOD	SV197	6230	AT160	SV171	K0320	BI257	AP036	OR3_27b	SLE	K0302
27.5	SV197		10.17								
30.7	6230	14.11									
96.8	AT160				3.46	2.64					
130.5	SV171			32.48		17.74			1.21		
137.6	K0320			35.07	4.89				1.44		
158.8	BI257										
178.5	AP036										
~200	OR3_27b				48.65	44.97					
250.3	SLE										2.21
270.8	K0302									40.24	

Table 4: Pairwise distances and LOD scores for all loci from Chromosome 3 in *A. mellifera* that are linked in *A. florea*. Pairwise linkage distances were calculated by MapMaker3.0. All linkages with a distance of <50cM and LOD >1.0 are reported here. Markers are listed according to order in the *A. mellifera* genome.

Table 5: Possible map orders for chromosome 3. Marker orders, interval distances and likelihood scores were determined by MapMaker3.0. Dotted line indicates inversion as compared to the null model. Note the large intervals where no linkage was established in Table 4. Differences in map distances between here and Table 4 are the consequence of ordering markers in a linkage group.

	Null Model		Most Likely Model	
	Marker Order	Pairwise Distance (cM)	Marker Order	Pairwise Distance (cM)
	SV197	14.1	6230	14.1
	6230	180.3	SV197	87.3
	AT160	32.9	AT160	32.8
	SV171	5.0	SV171	5.0
	K0320	47.6	K0320	47.6
	OR3_27b	59.0	OR3_27b	59.0
	SLE	40.2	SLE	40.2
	K0302	--	K0302	--
Total Distance (cM)	379.2		286.1	
Log likelihood	-179.87		-179.80	

Chromosome 12

Of the 101 microsatellites found on chromosome 12 in *A. mellifera* 12 polymorphic markers provided scoreable results in my *A. florea* mapping population. The data from these 12 loci were entered in MapMaker3.0 in both phases and using the "near" command all linkages to each marker were identified all linkage distances below 50cM with LOD scores above 1are reported in Table 6.

BI125 was dropped from further analyses, because it failed to link significantly to any marker. UN353B and UN237B are only loosely linked to other markers and additional marker orders were established excluding these markers (Table 7).

Position in *A. mellifera* (cM)	LOD / Distance (cM)	K1201B	6301	UN353B	UNDW17	K0305	K1257	C5828T	UN237B	BI125	AT136	UN246B	SV085
5.9	K1201B		21.09	1.41									
9.9	6301	3.38											
36.4	UN353B	42.94				1.64							
47.1	UNDW17					3.8							
56.5	K0305			35.63	28.85		1.02					1.19	
90.8	K1257					47.84		4.09	1.79				
133.5	C5828T						27.47		1.08				
139.4	UN237B						41.02	46.15					
139.4	BI125												
173.3	AT136											14.64	2.25
177.2	UN246B					46.41					10.02		3.71
212.4	SV085										38.83	31.22	

Table 6: Pairwise distances and LOD scores for all loci from Chromosome 12 in *A. mellifera* that are linked in *A. florea*. Linkage distances were calculated by MapMaker3.0. All linkages with a distance of <50cM and LOD >1.0 are reported here. Markers are listed according to order in the *A. mellifera* genome.

Table 7: Marker orders for chromosome 12. Dotted lines indicate where an inversion would have taken place.

	All Markers				Excluding UN237B			Excluding UN353B	
	Null Model	Most Likely	One Inversion	One Inversion	Null model	Most Likely	One inversion	Null model	Most likely
	K1201B	K0305	K0305	UN237B	K1201B	C5828T	C5828T	K1201B	SV085
	6301	UNDW17	UNDW17	C5828T	6301	K1257	K1257	6301	UN246B
	UN353B	UN353B	UN353B	K1257	UN353B	K0305	K0305	UNDW17	AT136
	UNDW17	K1201B	6301	K0305	UNDW17	UNDW17	UNDW17	K0305	6301
	K0305	6301	K1201B	UNDW17	K0305	UN353B	UN353B	K1257	K1201B
	K1257	AT136	K1257	UN353B	K1257	K1201B	6301	C5828T	UN237B
	C5828T	UN246B	C5828T	6301	C5828T	6301	K1201B	UN237B	C5828T
	UN237B	SV085	UN237B	K1201B	AT136	AT136	AT136	AT136	K1257
	AT136	UN237B	AT136	AT136	UN246B	UN246B	UN246B	UN246B	K0305
	UN246B	K1257	UN246B	UN246B	SV085	SV085	SV085	SV085	UNDW17
	SV085	C5828T	SV085	SV085					
Total Distance (cM)	342.6	326.5	377.7	332.0	297.4	280.0	286.0	346.3	301.6
Log likelihood	-224.96	-223.72	-225.77	-224.48	-202.03	-201.17	-201.54	-212.48	-210.98

Interval Comparisons

Chromosome 3

In chromosome 3 two markers (BI257 and AP036) were dropped from consideration in map construction. Pairwise distances between the remaining markers are shown for the most likely chromosome models in Table 5. No intervals were significantly different between the two species. The intervals between AT160 and SV171 and SV171 and K0302 were highly conserved (Fig 4A).

Interval	Physical distance in *A. mellifera* (bp)[a]	*A. florea* recombinational distance (cM) [b]	*A. florea* sample size[c]	*A. mellifera* recombinational distance (cM) [d]	*A. mellifera* sample size[e]	P-value[f]	*A. florea* rec. rate (cM/Mb) [g]	*A. mellifera* rec. rate (cM/Mb) [e]
SV197-6230	157726	14.1	80	3.2	92	0.0222	89.4	20.3
6230-At160	2381279	unlinked	82	66.1	92	unlinked	unlinked	27.8
At160-SV171	1391572	32.9	82	33.7	92	1	23.6	24.2
SV171-K0320	54291	5.0	75	7.1	92	1	92.1	130.8
K0320-OR3_27b	2675947	47.6	75	62.4	92	0.5293	17.8	23.3
OR3_27b-SLE	1317526	59	90	50.3	92	0.762	44.8	38.2
SLE-K0302	1366274	40.2	81	20.5	92	0.0555	29.4	15.0

Table 8: Interval comparisons in chromosome 3. Intervals between microsatellite markers were compared for differences in recombination distances between *A. mellifera* and *A. florea*. [a]Calculated from data presented in Solignac, Mougel et al. 2007. [b]From Table 5, null model. [c,e]Number of individuals that were scored for both markers of the interval. [d,e]From Solignac, Moguel et al. 2007. [f] Boneferroni corrected significance level $\leq$ 0.008333. [g]Assumed conservation of physical size between genomes.

Chromosome 12

In chromosome 12 one marker (BI125) failed to link significantly to any other markers and was dropped from consideration in map construction. Models which also dropped UN237B produced the best results (see discussion). Further analysis was done using the null model which excluded BI125 and UN237B (Table 7). Pairwise distances were conserved between all models which excluded these markers (Table 9).

No intervals in chromosome 12 were significantly different between *A. mellifera* and *A. florea*. The interval between 6301 and K1201B is highly conserved between the two species (Table 10, Fig. 4B).

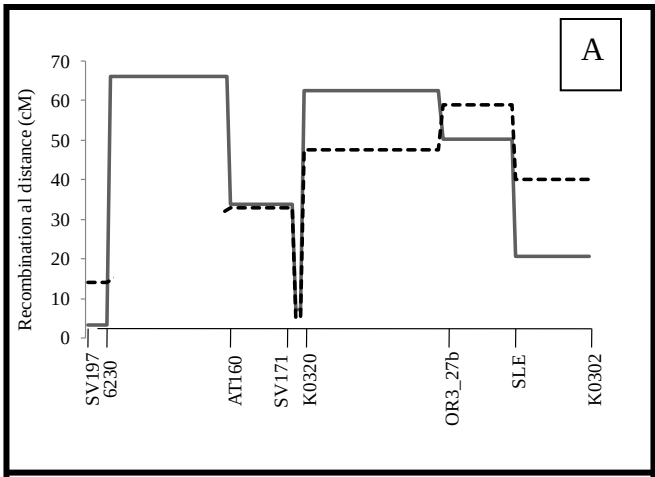
Table 9: Marker orders for chromosome 12 excluding UN237B. Pairwise distances and likelihood scores were calculated by MapMaker3.0. Some map inflation around K0305/UN353B (see table 6) in all models, but most distances are conserved between table 6 and all models here. Dotted lines indicate inversion from the null model.

Null Model		Most Likely		One inversion	
Marker Order	Pairwise Distance (cM)	Most Likely	Pairwise Distance (cM)	One inversion	Pairwise Distance (cM)
K1201B	3.4	C5828T	27.7	C5828T	27.7
6301	49.6	K1257	50.7	K1257	50.7
UN353B	27.8	K0305	29.0	K0305	29.0
UNDW17	29	UNDW17	28.3	UNDW17	27.8
K0305	50.7	UN353B	45.9	UN353B	49.6
K1257	27.5	K1201B	3.3	6301	3.4
C5828T	67.5	6301	53.2	K1201B	55.9
AT136	10.1	AT136	10.2	AT136	10.2
UN246B	31.7	UN246B	31.7	UN246B	31.7
SV085	--	SV085	--	SV085	--
Total Distance (cM)	297.4		280.0		286.0	
Log likelihood	-202.03		-201.17		-201.54	

Interval	Physical distance in *A. mellifera* (bp)[a]	*A. florea* recombinational distance (cM) [b]	*A. florea* sample size[c]	*A. mellifera* recombinational distance (cM) [d]	*A. mellifera* sample size[e]	P-value[f]	*A. florea* rec. rate (cM/Mb) [g]	*A. mellifera* rec. rate (cM/Mb) [e]
K1201B-6301	46861	3.4	89	4	92	1	72.6	85.4
6301-UN353B	1008680	49.6	53	26.5	92	0.0891	49.2	26.3
UN353B-UNDW17	335812	27.8	45	10.7	92	0.0461	82.8	31.9
UNDW17-K0305	647118	29	39	9.4	92	0.0284	44.8	14.5
K0305-K1257	1275450	50.7	39	34.3	92	0.3142	39.8	26.9
K1257-C5828T	1512708	27.7	53	37.8	92	0.4485	18.3	25.0
C5828T-AT136	1847933	67.5	53	40	92	0.2859	36.5	21.6
AT136-UN246B	146622	10.2	83	3.9	92	0.2326	69.6	26.6
UN246B-SV085	1015938	31.7	80	39.1	92	0.5083	31.2	38.5

Table 10: Interval comparisons in chromosome 12. Intervals between microsatellite markers were compared for differences in recombination distances between *A. mellifera* and *A. florea*. [a]Calculated from data presented in Solignac, Mougel et al. 2007. [b]From Table 5, null model. [c,e]Number of individuals that were scored for both markers of the interval. [d,e]From Solignac, Moguel et al. 2007. [f] Boneferroni corrected significance level ≤ 0.0056. [g]Assumed conservation of physical size between genomes.

In total, comparisons were made over 19 intervals in 6 chromosomes. In 8 intervals *A. mellifera* had a larger recombinational distance than *A. florea.* This is not a significant difference (sign test, P=0.6476).



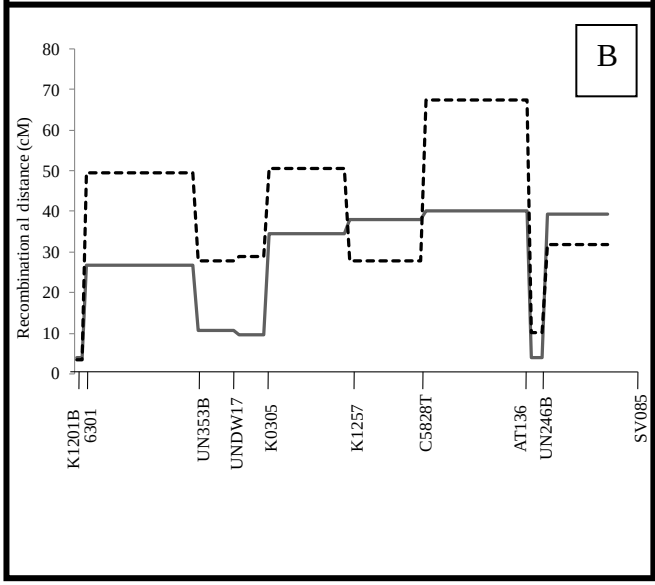


Figure 4: Comparisons of recombinational distances in chromosomes 3 and 12. Markers are placed along the X-axis according to physical position in the *A. mellifera* genome. The dashed line represents *A. florea*. The interval between 6230 and AT160 is unlinked in *A. florea.* **A:** Comparison of recombinational distances across chromosome 3. **B:** Comparison of recombinational distances across chromosome 12.

Additional Comparisons

Recombination rates in P1 and P2 were significantly correlated (R=0.390, *P*<0.000, Fig 5A). Recombination rates in P1 and P2 were both significantly correlated with recombination rates in Solignac et. al's population (P1: R=0.282, *P*<0.000; P2: R=0.269, *P*<0.000; Fig. 5B). The average recombination rate over intervals in P1 was not significantly different from the average recombination rate over intervals in

Solignac's map (paired t-test, *P*=.072). Intervals in P1 have an average recombination rate of 30.57cM/Mb compared to 28.32cM/Mb in Solignac's map. However, the average recombination rate in P2 is 31.79cM/Mb and this is a slight but significant difference from the average in Solignac et. al (paired t-test, t=0.2609, *P*=.01).

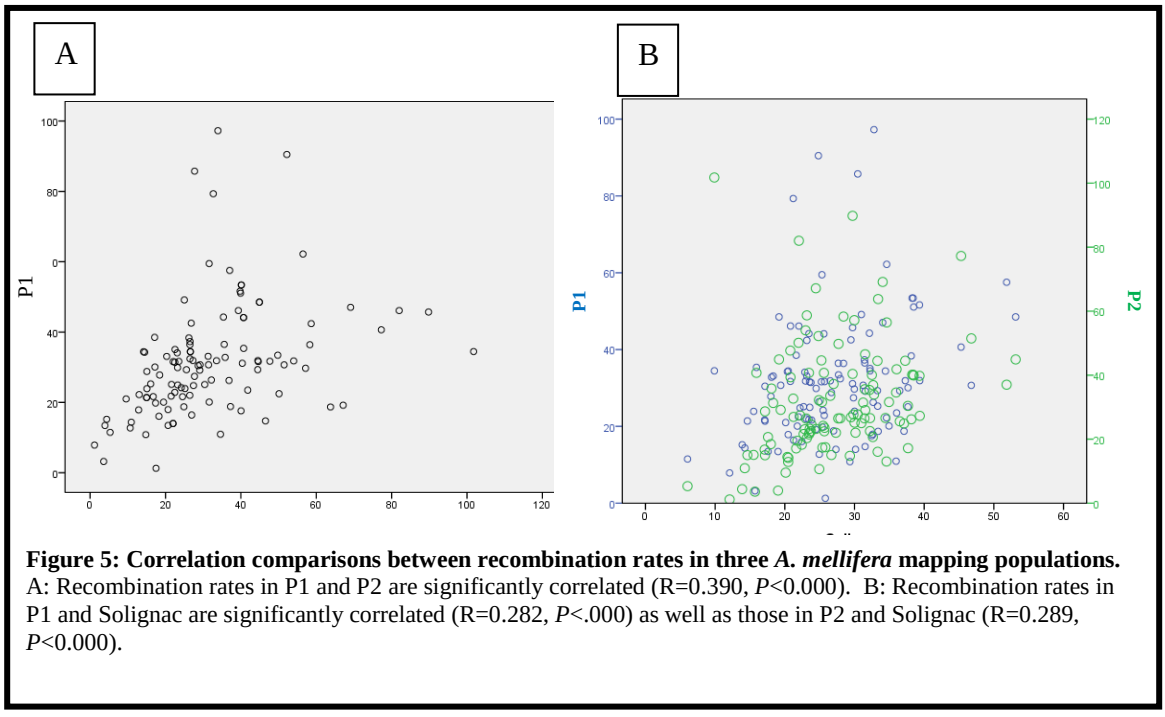


Figure 5: Correlation comparisons between recombination rates in three *A. mellifera* mapping populations. A: Recombination rates in P1 and P2 are significantly correlated (R=0.390, *P*<0.000). B: Recombination rates in P1 and Solignac are significantly correlated (R=0.282, *P*<.000) as well as those in P2 and Solignac (R=0.289, *P*<0.000).

Error Rate

The error rate was calculated for each marker that was double scored (Table 11). Seven markers had no mismatches between scorings and the highest error rate was for 3.2% for locus 6230. The overall error rate for the dataset was 0.8%.

Table 11: Genotyping error rate. Error was calculated for each marker that was double scored and for the complete dataset. Error for individual markers ranged from 0-3.2% and the overall dataset had an error rate of 0.8%.

Locus	Comparisons	Mismatches	Error
SV171	64	1	0.7%
SLE	83	0	0.0%
OR3_27b	61	2	1.6%
K0302	78	0	0.0%
SV197	63	0	0.0%
6230	62	4	3.2%
AT160	61	1	0.8%
K1201B	59	0	0.0%
K12577	67	1	0.7%
6301	48	0	0.0%
C5828T	71	0	0.0%
K0305	50	3	3.0%
SV085	72	0	0.0%
AT136	89	3	1.7%
K0320	56	1	0.9%
Total	984	16	0.8%

CHAPTER IV

DISCUSSION

I found marker order to be highly conserved between *A. florea* and *A. mellifera* but recombination rate variable. Recombination rates appear comparable between species, but contrary to all my predictions, there is a trend towards higher recombination rates in *A. florea*. This main conclusion is independent of chromosome model and genotyping methodology and dismisses domestication or pathogen pressure as a main cause for the high recombination rate of *A. mellifera* but the even higher rate of *A. florea* requires a new explanation.

Chromosome Models

Significant differences in recombinational distances can be observed regardless of marker order. However, the established chromosome models are valuable for the insights they give into genome evolution in *Apis.* Multiple criteria were used to determine the most likely marker order and markers were excluded from consideration only after careful examination. In chromosome three, omitting BI257 from map construction was justified because this locus was completely unlinked and it could not be replicated. This suggests that the bands observed in the first genotyping run were not the BI257 locus.

Accepting the null model of conserved marker order in chromosome three was also appropriate. The likelihood scores between the null model and the most likely

model are not significantly different. The possible inversion occurs between markers 6230 and AT160. In *A. mellifera* these markers are 66.1cM apart, too far to establish reliable linkage. Additional efforts to fill this gap by designing and amplifying new primers were unsuccessful. This gap remains in *A. florea* and resulted in two models of comparable likelihood. The lack of linkage between 6230 and AT160 allowed MapMaker3.0 to invert the first two markers with little consequence because 6230 and SV197 are linked strongly to each other, but neither marker is consequentially linked to AT160. My data is not sufficient to dismiss the null model based on likelihood scores. A significant difference may be evident with a higher map resolution.

Another indicator of the reliability of a marker order is the overall recombinational length of the linkage group. I would expect the length of the most probable chromosome model in *A. florea* to be most comparable to the length of the group in *A. mellifera* (270.8cM). The most likely model provided by MapMaker3.0 is approximately 15cM longer than the group in *A. mellifera*, whereas the null model is over 100cM longer than the *A. mellifera* group. This difference can be entirely attributed to the linkage gap between markers 6230 and AT160. However this distance estimate is unreliable because these markers are essentially unlinked. The distance between 6230 and AT160 in the null model and the distance between SV197 and AT160 in the most likely model are both too large to be considered as evidence of linkage. Therefore, the difference in length of the linkage group does not provide enough evidence to dismiss the null model.

In chromosome 12 excluding BI125 from map construction was justified. Not only was this locus completely unlinked, but I attempted to re-amplify and genotype this marker twice and was not able to replicate the results from the first trial. This suggests that the bands observed in the first genotyping run were not the BI125 locus.

Dropping UN237B from consideration in determining marker order, while including UN353B, was also justified. Though marker UN237B linked loosely with K1257 and C5828T, both linkages were greater than 40cM and had a LOD score of less than 2.0. UN353B was also linked loosely to only two markers (K0305 and K1201B) and therefore considered for exclusion. However, at least one of these linkages was with less than 40cM distance. The main reason for excluding UN237B and not UN353B was map expansion caused by these markers. Map expansion is another criterion used to exclude markers from a linkage group (Hunt and Page 1995). Including a marker between two other markers should not inflate the overall map distance because linkage distances should be largely additive. Including UN237B increased the interval between C5828T and AT136 by 65%. On the other hand, including UN353B decreased the interval between 6301 and UNDW17 by 4% (data not shown). Lastly, UN237B also could not be successfully re-amplified or genotyped while UN353B was replicated though not re-scored.

In the *A. mellifera* none of the markers in the chromosome 12 linkage group are linked at greater than 50cM. Excluding UN237B and BI125 does leave a gap of approximately 40cM between C5828T and AT136. The next largest gap is approximately 26cM between 6301 and UN353B.

The three models for chromosome 12 do not have significantly different LOD scores. The most likely model indicates an inversion event at two largest gaps in the *A. mellifera* group (noted above). Such an event would be extremely rare and is probably postulated by MapMaker3.0 due to the large distances between C5828T/AT136 and 6301/UN353B. K1201B and 6301 are tightly linked in both *A. mellifera* and *A. florea*. The double inversion model rotates this group and the single inversion model does not. Between these two models, the difference in distance in interval length on both sides of this group is negligible. The simplicity of the single inversion model makes it more likely than the double inversion model (Gibson and Muse 2001).

Between the single inversion model and the null model there is a difference in overall linkage group length. In *A. mellifera* the chromosome 12 linkage group has a recombinational length of 212.4cM. The length of the null model in *A. florea* is 297.4cM and the single inversion model is 286.0cM. The difference between the two *A. florea* models is found primarily in the interval around the proposed inversion site. In the null model C5828T is 66.7cM away from AT136. In the inverted model K1201B is 55.9cM from AT136. Both of these intervals are conventionally too large to establish reliable linkage, and the likelihood scores of the two models are not significantly different. Therefore, my data suggests that the null model is the correct marker order for chromosome 12 in *A. florea.*

Interval Comparisons

The results of the sign test suggest that overall recombinational distances are not significantly different between *A. mellifera* and *A. florea.* However, most intervals were

non-significantly larger in *A. florea* and I found one interval where *A. florea* demonstrates a significantly larger recombinational distance than *A. mellifera*. In chromosome 2, the interval between AC033 and BI251 is 161% larger in *A. florea*. This interval appears to be in a recombinational hotspot, with a recombination rate of over 100cM/Mb in *A. mellifera* and over 300cM/Mb in *A. florea*. Two intervals (between K1168 and BI001 on chromosome 2 and between 6230 and AT160 in chromosome 3) were unlinked in *A. florea* and linked in *A. mellifera*, also suggesting a larger recombinational distance in *A. florea*.

Genotyping errors typically inflate map distances (Stam 1993; Gibson and Muse 2001) and might be considered a possible explanation for the significantly higher recombinational distances in *A. florea*. However, the estimated genotyping error rate for this data set is low. Marker 6230 had the highest error rate for an individual marker at 3.2%. This marker surrounds a large linkage gap in chromosome 3. The interval between 6230 and AT160 was too large to establish linkage in both *A. mellifera* and *A. florea* and therefore the map-inflating effects of genotyping errors here are not reflected in the results because 6230 was treated as unlinked.

Intra-specific variation in recombination rate has been demonstrated in many species (Serre, Nadon, and Hudson 2005; Graffelman et al. 2007; Dumont, Broman, and Payseur 2009; Lowe, Riaz, and Walker 2009). Presumably, my results may be the consequence of especially high recombination rates in this particular *A. florea* population. Linkage mapping studies in *A. mellifera* have shown some population dependent variation in recombination. However the correlation between overall recombination rates

in different populations is strong (Hunt and Page 1995; Solignac et al. 2004; Beye et al. 2006; Solignac et al. 2007a). My intra-specific comparisons between P1, P2 and the map of Solignac et. al (2007) are consistent with conserved recombination rates. Assuming that recombination rate behaves similarly between *A. florea* populations it seems unlikely that my results are solely due to an atypical *A. florea* mapping population.

It is possible that some of the difference observed between my mapping population of *A. florea* and Solignac's mapping population of *A. mellifera* is a consequence of difference in map construction methods. The significant recombination rate differences between P2 and Solignac's cross suggest that methodology plays a role but that resulting rate differences are generally minor (<10%). P1 and P2 are significantly correlated and the recombination rates in P1 were not significantly different from Solignac et. al. In this study, intervals in *A. mellifera* ranged from 77% smaller than *A. florea* to 126% larger than *A. florea*. On average, intervals in *A. mellifera* were 11% smaller than the intervals in *A. florea* which is consistent with the comparison between P2 and Solignac's population. However, the significantly different interval between *A. florea* and Solignac's map is 161% larger in *A. florea*. This suggests that the significance of this interval cannot be explained by methodological differences in map construction alone.

Implications of Results

My data gives the first insight into the genomic recombination rate of *A. florea*. The results of my study are not substantially influenced by genotyping errors and it is unlikely that intra-specific variation in recombination rate plays a significant role in

comparisons between *A. florea* and *A. mellifera.* My maps are less resolved than the map of *A. mellifera*, but I have found that this likely does not affect the significance of my findings. Though limited to primarily two chromosomes, my results suggest a genome-wide recombination rate in *A. florea* that is at least as high of that in *A. mellifera.*

Thus, my study provides another line of evidence supporting the correlation between eusociality and high recombination rates. Social hymenoptera are known to have more chromosomes than their non-social counterparts (Crozier 1977; Sherman 1979), but this structural commonality is not enough to explain the high recombination rates in the social insects. Adaptive processes must influence recombination rate as well (Beye et al. 2006; Sirvio et al. 2006; Wilfert, Gadau, and Schmid-Hempel 2007).

My findings in *A. florea* provide additional support for this argument**.** As this study demonstrates, marker order is conserved between the *A. mellifera* and *A. florea,* and structurally their genomes are similar (Arias and Sheppard 2005; Raffiudin and Crozier 2006). Adaptive causes of variation in recombination rate between the species seem likely. This study tested the adaptive effects of domestication, pathogen pressure and division of labor on recombination rates in *Apis.*

High recombination rates in the undomesticated *A. florea* contradict the claim that the exceptionally high recombination rate seen in *A. mellifera* is an artifact of domestication. Domestication has been shown to increase recombination rates in some plants and animals (Burt and Bell 1987; Ross-Ibarra 2004), but the high recombination rate seen here in *A. florea* suggests that domestication has not had an increasing effect on recombination rate in *A. mellifera* because it is certain that *A. florea* was never

domesticated (Oldroyd and Wongsiri 2006). Two possible explanations for this are: 1) *A. mellifera* is not truly domesticated and 2) domestication effects on recombination rate may be taxon specific. Domestication implies breeder selection for particular traits. Particular traits (honey production, low aggression, etc.) have been selected for in the honey bee, but it is difficult to know if this selection has been effective (Crane 1999). Wild honey bee populations can be abundant and readily interbreed with commercial honey bees which could dilute any domestication effects. It seems more likely that domestication simply has different effects on recombination rate in plants and animals than it does in insects. Even in non-domesticated social insects recombination rates are similar to that of the honeybee (Sirvio et al. 2006; Wilfert, Gadau, and Schmid-Hempel 2007) and in other domesticated insects such as *Bombyx mori*, the silkworm, recombination rates are comparable with closely related, non-domesticated insects (Miao et al. 2005; Yamamoto et al. 2006; Wilfert, Gadau, and Schmid-Hempel 2007). This suggests that insect genomes may respond differently to artificial selection than plant and animal genomes. Perhaps this is due to a shared structural feature of insect genomes, such as compactness or low GC content compared to vertebrates (Berry 1985; Honeybee Genome Sequencing Consortium 2006).

My hypothesis that recombination rate is correlated with pathogen pressure in honey bees is not supported by these findings. I postulated that the open-air nesting *A. florea* should have a lower degree of pathogen pressure than the cavity-nesting *A. mellifera* (Root and Root 1980; Schmid-Hempel 1998; Woyke, Wilde, and Reddy 2004; Oldroyd and Wongsiri 2006) and I did not find lower recombination rates in *A. florea*.

This suggests that the adaptive benefit of high recombination rates in *Apis* is not due to pathogen pressure. My findings do not eliminate the possibility that pathogen pressure played a role in establishing the high recombination rate in *Apis* and other social insects. However, the results of this study do suggest that the current role of pathogen pressure in maintaining high recombination rates in *Apis* is limited.

The hypothesis that recombination rate is correlated with division of labor in social insects is most consistent with my findings. Complex division of labor is correlated with high recombination rates in *A. florea* and well as other insect species (Beye et al. 2006; Sirvio et al. 2006; Wilfert, Gadau, and Schmid-Hempel 2006). No scale has been established to quantitatively measure the complexity of social insect societies, but honey bees have extraordinarily complex societies compared to most other Hymenoptera. Factors such as colony size, reproductive behavior of workers, and division of colony tasks establish *A. mellifera* as one of the most developed eusocial insect species (Winston 1987; Sirvio et al. 2006; Wilfert, Gadau, and Schmid-Hempel 2007). *A. florea* has been less extensively studied. However, many characteristics are shared between the species and it is highly unlikely that *A. florea* has a more complex division of labor than *A. mellifera*. On the contrary, smaller colony sizes and the basal phylogenetic position of *A. florea* suggest that *A. florea* perhaps has a less complex social structure than *A. mellifera.* While the high recombination rate observed in *A. florea* is expected compared to other social Hymenoptera with lower degrees of social complexity and other non-social organisms, I did not expect to see a higher recombination rate in *A. florea* than in *A. mellifera* based on the division of labor hypothesis.

Possible additional explanations for the high recombination rate in *A. florea* include differences in mating behavior and effective population size. *A. florea* queens mate with an average of 8 males, whereas *A. mellifera* queens mate with an average of 13.8 males (Oldroyd et al. 1997). Genetic diversity is increased through multiple mating and the consequential introduction of multiple patrilines in the colony. Recombination also increases genetic diversity (Muller 1932; Felsenstein 1974; Michod and Levin 1988; Sirvio et al. 2006). Currently, it is unknown to what extent these processes can compensate for each other. However, both recombination and multiple mating (Sherman 1979; Schmid-Hempel and Crozier 1999) contribute to genetic diversity among colony offspring that is beneficial to the colony. Theoretically, a loss of genetic diversity due to reduction in the number of mating partners could be compensated for by an increase in recombination. This could explain the elevated recombination rate in specific intervals in *A. florea.* More studies on the correlation between multiple mating and recombination rate are needed to support this possible conclusion.

Another possible explanation for the high recombination rate in *A. florea* is the effect of population structure on recombination rate. Theoretical studies suggest that population structure influences recombination rate. In particular, it has been suggested that recombination does not easily spread widely through highly structured societies, such as those seen in social insects (Martin, Otto, and Lenormand 2006). This contradicts overall empirical findings (Beye et al. 2006; Sirvio et al. 2006; Wilfert, Gadau, and Schmid-Hempel 2006; Wilfert, Gadau, and Schmid-Hempel 2007), but differences in population structure may account specifically for differences in recombination rate

between *A. mellifera* and *A. florea.* One difference between species may be in effective population size (Wright 1931; Wright 1938). In social insects, effective population size is determined primarily by the number of colonies in a population, not the number of individuals (Kerr 1967; Kraus et al. 2005). Honey bees have small but variable population sizes and determining this size with genetic data is straightforward and reliable (Kraus et al. 2005), yet few practical studies of effective population size have been done in honey bees. My data shows recombination rate is variable, though correlated, between different populations of *A. mellifera* and that recombination rate is somewhat variable between species within *Apis.* The exceptionally high recombination rates in conjunction with the small population sizes in *Apis* make the honey bee an especially interesting model for the investigating the correlation between population size and recombination rate.

Future Studies

Here, I have determined that domestication and pathogen pressure are not responsible for the high recombination rate in *A. mellifera.* In order to determine what precise adaptive benefits recombination offers to honey bees (and social insects in general) more studies must be done. One avenue of research is further comparisons within *Apis.* These comparisons must be two–fold. We need to investigate recombination rate across *Apis* more generally and we also need a better understanding of the biology of the different *Apis* species, especially in regards to effective population sizes, pathogen pressure and the complexity of their division of labor.

Based on my findings here, to determine the cause of variation in recombination rate between species within *Apis* studies should be done which take mating behavior and population size into account. More comparisons between species with different mating behaviors is one approach. Another approach is to identify isolated populations with small effective population sizes and compare recombination rates of these populations to other populations of the same species.

Other avenues of research can be pursued using social insects outside of *Apis*. This approach establishes phylogentically independent contrasts and offers the opportunity to investigate more clearly defined differences in pathogen pressure and division of labor. Particularly, we should investigate social insects with a lower degree of division of labor and/or a higher pathogen pressure than the honey bee, such as soil-dwelling ants (Hèolldobler and Wilson 1990). If high recombination rates were found in such species, this would suggest that recombination rate is an adaptive benefit in the presence of extreme pathogen pressure and a lower recombination rate in these species would suggest that high levels of recombination are an adaptive benefit promoting division of labor.

Another particularly interesting species for future study is *Bombus terrestris* (the large earth bumble bee). As stated above, the bumble bee has a recombination rate of 4.4cm/Mb (Wilfert, Gadau, and Schmid-Hempel 2006). Many factors may play a role in establishing this low recombination rate in the bumble bee. Colonies sizes are small and only young queens generally survive the winter (reviewed in Velthuis and van Doorn 2006). These factors may reduce pathogen pressure in the bumble bee. Division of labor

is not as complex in *B. terrestris* as in *A. mellifera* (Velthuis and van Doorn 2006). Effective population size has not yet been studied in *B. terrestris* and would provide more information about the link between recombination rate and population size in social insects. *B. terrestris* is relatively unique among the Hymenoptera in that queens are singly-mated (Schmid-Hempel and Schmid-Hempel 2000). More studies are needed to determine which of these factors play a significant role in contributing to the low recombination rate observed in the bumble bee.

Another interesting approach to the question of the adaptive benefits of recombination is to study social insects outside of the Hymenoptera. Currently, it is not known whether high recombination rate is exclusive to Hymenoptran social insects, or if high recombination rates are correlated with sociality in other taxa. Degrees of sociality vary within other taxa, from aggregations of some butterfly larvae (first identified by Riley and Howard 1890) to complex eusocial societies seen in termites (Order Isoptera). Low recombination rates in social insects outside of the Hymenoptera would suggest a structural explanation for the high recombination rates within the Hymenoptera and high recombination rates correlated with eusociality outside of the Hymenoptera would suggest an adaptive benefit of recombination correlated with sociality.

CHAPTER V

CONCLUSION

This study contributes to the growing body of evidence establishing a link between sociality and recombination rate. Here, I report findings of possibly higher recombination rates in *A. florea* than in *A. mellifera*. This could make *A. florea* the species with the highest recombination rate among all metazoans. It suggests that the adaptive benefit of increased recombination rate in social insects may be more closely associated with division of labor than with pathogen pressure. However, without a clear scale by which to measure these features this conclusion remains speculative. My results provide evidence that the extraordinarily high recombination rate of *A. mellifera* is not the result of selection pressure exerted through the process of domestication. However, more studies within *Apis* and other social insect genera are needed for conclusive results on what is responsible.

My results support the general association between high recombination rate and sociality in insects, one of the strongest arguments for the hypothesis that meiotic recombination may be adaptive and not just a structural necessity to guarantee proper chromosome segregation.

REFERENCES

Alexander, B. A. 1991. Phylogenetic analysis of the genus *Apis* (Hymenoptera: Apidae). Ann. Entomol. Soc. Amer. **84**:137-149.

Amersham Biosciences Corp. 2002. MegaBACE: Instrument Operator's Guide. Amersham Biosciences AB, Uppsala.

Arias, M. C., and W. S. Sheppard. 2005. Phylogenetic relationships of honey bees (Hymenoptera:Apinae:Apini) inferred from nuclear and mitochondrial DNA sequence data. Mol Phylogenet Evol **37**:25-35.

Baker, B. S., A. T. Carpenter, M. S. Esposito, R. E. Esposito, and L. Sandler. 1976. The genetic control of meiosis. Annu Rev Genet **10**:53-134.

Barton, N. H. 1995. A general model for the evolution of recombination. Genet Res **65**:123-145.

Bergero, R., and D. Charlesworth. 2009. The evolution of restricted recombination in sex chromosomes. Trends Ecol Evol **24**:94-102.

Berry, S. J. 1985. Insect nucleic acids *in* G. A. Kerkut, and L. I. Gilbert, eds. Insect Physiology Biochemistry and Pharmacology. Oxford, Pergamon.

Beye, M., I. Gattermeier, M. Hasselmann, T. Gempe, M. Schioett, J. F. Baines, D. Schlipalius, F. Mougel, C. Emore, O. Rueppell, A. Sirvio, E. Guzman-Novoa, G. Hunt, M. Solignac, and R. E. Page, Jr. 2006. Exceptionally high levels of recombination across the honey bee genome. Genome Res **16**:1339-1344.

Beye, M., M. Hasselmann, M. K. Fondrk, R. E. Page, and S. W. Omholt. 2003. The gene csd is the primary signal for sexual development in the honeybee and encodes an SR-type protein. Cell **114**:419-429.

Beye, M., G. J. Hunt, R. E. Page, M. K. Fondrk, L. Grohmann, and R. F. Moritz. 1999. Unusually high recombination rate detected in the sex locus region of the honey bee (Apis mellifera). Genetics **153**:1701-1708.

Breed, M., C. Welch, and R. Cruz. 1994. Kin Discrimination Within Honey-Bee (Apis-mellifera) Colonies: An Analysis of the Evidence. Behavioural Processes **33**:25-39.

Brown, M. J., and P. Schmid-Hempel. 2003. The evolution of female multiple mating in social hymenoptera. Evolution **57**:2067-2081.

Burt, A., and G. Bell. 1987. Mammalian chiasma frequencies as a test of two theories of recombination. Nature **326**:803-805.

Chagne, D., C. Lalanne, D. Madur, S. Kumar, J. Frigerio, C. Krier, S. Decroocq, A. Savoure, M. Bou-Dagher-Kharrat, E. Bertocchi, J. Brach, and C. Plomion. 2002. A high density genetic map of maritime pine based on AFLPs. Annals of Forest Science **59**:627-636.

Charlesworth, D. 2002. Plant sex determination and sex chromosomes. Heredity **88**:94-101.

Cockburn, A. 1991. An Introduction to Evolutionary Ecology. Blackwell Scientific Publications, Oxford.

Comeron, J. M., and M. Kreitman. 2000. The correlation between intron length and recombination in *Drosophila*. Dynamic equilibrium between mutational and selective forces. Genetics **156**:1175-1190.

Crane, E. 1999. The world history of beekeeping and honey hunting. Routledge, New York.

Crozier, R. H. 1977. Evolutionary Genetics of the Hymenoptera. Ann. Rev. Entomol. **22**:263-288.

Crozier, R. H., and R. E. Page. 1985. On being the right size: male contributions and multiple mating in social Hymenoptera. Behavioral Ecology and Sociobiology **18**:105-115.

Dietrich, W. F., J. Miller, R. Steen, M. A. Merchant, D. Damron-Boles, Z. Husain, R. Dredge, M. J. Daly, K. A. Ingalls, and T. J. O'Connor. 1996. A comprehensive genetic map of the mouse genome. Nature **380**:149-152.

Dumas, D., and J. Britton-Davidian. 2002. Chromosomal rearrangements and evolution of recombination: comparison of chiasma distribution patterns in standard and robertsonian populations of the house mouse. Genetics **162**:1355-1366.

Dumont, B. L., K. W. Broman, and B. A. Payseur. 2009. Variation in Genomic Recombination Rates among Heterogeneous Stock Mice. Genetics **In Press**.

Engel, M. S., and T. R. Schultz. 1997. Phylogeny and behavior in honey bees (Hymenoptera, Apidae). Ann. Entomol. Soc. Amer. **90**:43-53.

Eyre-Walker, A. 1993. Recombination and mammalian genome evolution. Proc Biol Sci **252**:237-243.

Felsenstein, J. 1974. The evolutionary advantage of recombination. Genetics **78**:737-756
.
Felsenstein, J. 1979. A mathematically tractable family of genetic mapping functions with different amounts of interference. Genetics **91**:769-775.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford.

Gadau, J., C. U. Gerloff, N. Kruger, H. Chan, P. Schmid-Hempel, A. Wille, and R. E. Page. 2001. A linkage analysis of sex determination in Bombus terrestris (L.) (Hymenoptera: Apidae). Heredity **87**:234-242.

Gadau, J., R. E. Page, Jr., and J. H. Werren. 1999. Mapping of hybrid incompatibility loci in Nasonia. Genetics **153**:1731-1741.

Gadau, J., R. E. Page, Jr., J. H. Werren, and P. Schmid-Hempel. 2000. Genome organization and social evolution in Hymenoptera. Naturwissenschaften **87**:87-89
.
Ghiselin, M. T. 1974. The economy of nature and the evolution of sex. University of California Press, Berkeley.

Gibson, G., and S. V. Muse. 2001. A primer of genome science. Sinauer, Sunderland, Mass.

Gorlov, I. P., and O. Y. Gorlova. 2001. Cost-benefit analysis of recombination and its application for understanding of chiasma interference. J Theor Biol **213**:1-8.

Graffelman, J., D. J. Balding, A. Gonzalez-Neira, and J. Bertranpetit. 2007. Variation in estimated recombination rates across human populations. Hum Genet **122**:301-310.

Graham, J. M., J. T. Ambrose, L. L. Langstroth, and Dadant & Sons. 1992. The Hive and the honey bee : a new book on beekeeping which continues the tradition of "Langstroth on the hive and the honeybee". Dadant, Hamilton, Ill.

Hadany, L., and T. Beker. 2007. Sexual selection and the evolution of obligatory sex. BMC Evol Biol **7**:245.

Haldane, J. B. B. 1919. The combination of linkage values, and the calculation of distance between linked factors. J. Genet. **8**:299-309.

Hardison, R. C., K. M. Roskin, S. Yang, M. Diekhans, W. J. Kent, R. Weber, L. Elnitski, J. Li, M. O'Connor, D. Kolbe, S. Schwartz, T. S. Furey, S. Whelan, N. Goldman, A. Smit, W. Miller, F. Chiaromonte, and D. Haussler. 2003. Covariation in

frequencies of substitution, deletion, transposition, and recombination during eutherian evolution. Genome Res **13**:13-26.

Hassold, T., L. Judis, E. R. Chan, S. Schwartz, A. Seftel, and A. Lynn. 2004. Cytological studies of meiotic recombination in human males. Cytogenet Genome Res **107**:249-255.

Hassold, T., S. Sherman, and P. Hunt. 2000. Counting cross-overs: characterizing meiotic recombination in mammals. Hum Mol Genet **9**:2409-2419.

Hawley, R. S., K. S. McKim, and T. Arbel. 1993. Meiotic segregation in Drosophila melanogaster females: molecules, mechanisms, and myths. Annu Rev Genet **27**:281-317.

Hèolldobler, B., and E. O. Wilson. 1990. The ants. Belknap Press of Harvard University Press, Cambridge, Mass.

Honeybee Genome Sequencing Consortium. 2006. Insights into social insects from the genome of the honeybee *Apis mellifera*. Nature **443**:931-949.

Hoy, M. A. 2003. Insect Molecular Genetics: An Introduction to Principles and Applications. Academic Press, Amsterdam.

Hughes, W. O., B. P. Oldroyd, M. Beekman, and F. L. Ratnieks. 2008. Ancestral monogamy shows kin selection is key to the evolution of eusociality. Science **320**:1213-1216.

Hunt, G. J., and R. E. Page, Jr. 1995. Linkage map of the honey bee, Apis mellifera, based on RAPD markers. Genetics **139**:1371-1382.

International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature **432**:695-716.

Jensen-Seaman, M. I., T. S. Furey, B. A. Payseur, Y. Lu, K. M. Roskin, C. F. Chen, M. A. Thomas, D. Haussler, and H. J. Jacob. 2004. Comparative recombination rates in the rat, mouse, and human genomes. Genome Res **14**:528-538.

John, B. 1990. Meiosis. Pp. 29-102. Cambridge University Press, Cambridge.

John, B. 1988. The biology of heterochromatin. Pp. 1-147 *in* R. S. Verma, ed. Heterochromatin: Molecular and Structural Aspects. Cambridge University Press, Cambridge.

John, B., and G. L. Miklos. 1979. Functional aspects of satellite DNA and heterochromatin. Int Rev Cytol **58**:1-114.

Kaback, D. B. 1996. Chromosome-size dependent control of meiotic recombination in humans. Nat Genet **13**:20-21.

Kaback, D. B., V. Guacci, D. Barber, and J. W. Mahon. 1992. Chromosome size-dependent control of meiotic recombination. Science **256**:228-232.

Kerr, W. 1967. Multiple alleles and genetic load in bees. J. Apicult. Res. **6**:61-64.

Kleckner, N. 1996. Meiosis: how could it work? Proc Natl Acad Sci U S A **93**:8167-8174.

Koehler, K. E., R. S. Hawley, S. Sherman, and T. Hassold. 1996. Recombination and nondisjunction in humans and flies. Hum Mol Genet **5 Spec No**:1495-1504.

Koeniger, N., G. Koeniger, R. W. K. Punchihewa, and M. Fabritius. 1982. Observations and experiments on dance communication in *Apis florea* in Sri Lanka. J. Apicult. Res. **21**:45-52.

Kong, A., D. F. Gudbjartsson, J. Sainz, G. M. Jonsdottir, S. A. Gudjonsson, B. Richardsson, S. Sigurdardottir, J. Barnard, B. Hallbeck, G. Masson, A. Shlien, S. T. Palsson, M. L. Frigge, T. E. Thorgeirsson, J. R. Gulcher, and K. Stefansson. 2002. A high-resolution recombination map of the human genome. Nat Genet **31**:241-247.

Koonin, E. V., and M. Y. Galperin. 2003. Sequence-Evolution-Function: Computational Approaches in Comparative Genomics. Kluwer Academic Publishers.

Kosambi, D. D. 1944. The estimation of map distance from recombination values. Ann. Eugen. **12**:172-175.

Kraus, F. B., N. Koeniger, S. Tingek, and R. F. A. Moritz. 2005. Using drones for estimating colony number by microsatellite DNA analyses of haploid maies in *Apis*. Apidologie **36**:223-229.

Kryger, P., and R. F. A. Moritz. 1997. Lack of Kin Recognition in Swarming Honeybees (Apis mellifera). Behavioral Ecology and Sociobiology **40**:271-276.

Kullman, B., H. Tamm, and K. Kullman. 2005. Fungal Genome Size Database.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly, S. E. Lincoln, and L. A. Newberg. 1987. MAPMAKER: an interactive computer package for constructing

primary genetic linkage maps of experimental and natural populations. Genomics **1**:174-181.

Lawrie, N. M., C. Tease, and M. A. Hulten. 1995. Chiasma frequency, distribution and interference maps of mouse autosomes. Chromosoma **104**:308-314.

Lincoln, S. E., M. J. Daly, and E. S. Lander. 1992. Constructing genetic maps with MAPMAKER/EXP 3.0. Whitehead Institute Technical Report, Cambridge, M.A.

Lindauer, M. 1961. Communication among social bees. Harvard University Press, Cambridge.

Lorch, P., S. Proulx, L. Rowe, and T. Day. 2003. Condition-dependent sexual selection can accelerate adaptation. Evol Ecol Res **5**:867-881.

Lowe, K. M., S. Riaz, and M. A. Walker. 2009. Variation in recombination rates across the *Vitis* genome. Tree Genetics & Genomes **5**:71-80.

Marra, R. E., J. C. Huang, E. Fung, K. Nielsen, J. Heitman, R. Vilgalys, and T. G. Mitchell. 2004. A genetic linkage map of Cryptococcus neoformans variety neoformans serotype D (Filobasidiella neoformans). Genetics **167**:619-631.

Martin, G., S. P. Otto, and T. Lenormand. 2006. Selection for recombination in structured populations. Genetics **172**:593-609.

Mattila, H. R., and T. D. Seeley. 2007. Genetic diversity in honey bee colonies enhances productivity and fitness. Science **317**:362-364.

Maynard Smith, J. 1978. The evolution of sex. Cambridge University Press.

Mayr, E. 1942. Systematics and the origin of species. Columbia Univeristy Press, New York.

Miao, X. X., S. J. Xub, M. H. Li, M. W. Li, J. H. Huang, F. Y. Dai, S. W. Marino, D. R. Mills, P. Zeng, K. Mita, S. H. Jia, Y. Zhang, W. B. Liu, H. Xiang, Q. H. Guo, A. Y. Xu, X. Y. Kong, H. X. Lin, Y. Z. Shi, G. Lu, X. Zhang, W. Huang, Y. Yasukochi, T. Sugasaki, T. Shimada, J. Nagaraju, Z. H. Xiang, S. Y. Wang, M. R. Goldsmith, C. Lu, G. P. Zhao, and Y. P. Huang. 2005. Simple sequence repeat-based consensus linkage map of Bombyx mori. Proc Natl Acad Sci U S A **102**:16303-16308.

Michod, R. E., and B. R. Levin. 1988. The evolution of sex: An examination of current ideas. Sinauer Associates, Sunderland.

Muller, H. J. 1932. Some genetic aspects of sex. American Naturalist **66**:118-138.

Muller, H. J. 1964. The Relation of Recombination to Mutational Advance. Mutat Res **106**:2-9.

Noonan, K. C., and S. A. Kolmes. 1989. Kin recognition of worker brood by worker honey bees, *Apis mellifera* L. Journal of Insect Behavior **2**:473-485.

Oldroyd, B. P., M. J. Clifton, S. Wongsiri, T. E. Rinderer, H. A. Sylvester, and R. H. Crozier. 1997. Polyandry in the genus *Apis*, particularly *Apis andreniformis*. Behavioral Ecology and Sociobiology **40**:17-26.

Oldroyd, B. P., and J. H. Fewell. 2007. Genetic diversity promotes homeostasis in insect colonies. Trends Ecol Evol **22**:408-413.

Oldroyd, B. P., and S. Wongsiri. 2006. Asian honey bees : biology, conservation, and human interactions. Harvard University Press, Cambridge, Mass.

Otto, S. P., and T. Lenormand. 2002. Resolving the paradox of sex and recombination. Nat Rev Genet **3**:252-261.

Page, R. E., Jr. 1980. The evolution of multiple mating behavior by honey bee queens (Apis mellifera L.). Genetics **96**:263-273.

Page, R. E., Jr., J. Gadau, and M. Beye. 2002. The emergence of hymenopteran genetics. Genetics **160**:375-379.

Page, R. E., and H. H. Laidlaw. 1988. Full Sisters and Super Sisters - A terminological paradigm. Animal Behaviour **36**:944-945.

Paliulis, L. V., and R. B. Nicklas. 2000. The reduction of chromosome number in meiosis is determined by properties built into the chromosomes. J Cell Biol **150**:1223-1232.

Peck, J. R. 1993. Frequency-dependent selection, beneficial mutations, and the evolution of sex. Proc Biol Sci **254**:87-92.

Raffiudin, R., and R. H. Crozier. 2006. Phylogenetic analysis of honey bee behavioral evolution. Molecular Phylogenetics and Evolution.

Riley, C. V., and O. Howard. 1890. A social *Papilio* larva. Ins. Life **3**:32-33.

Robinson, G. 2002. Genomics and integrative analyses of division of labor in honeybee colonies. American Naturalist **160**:S160-S172.

Root, A. I., and E. R. Root. 1980. The ABC and XYZ of bee culture : an encyclopedia pertaining to scientific and practical culture of bees. A. I. Root Co., Medina, Ohio.

Ross-Ibarra, J. 2004. The evolution of recombination under domestication: a test of two hypotheses. Am Nat **163**:105-112.

Rueppell, O., R. Linford, P. Gardner, J. Coleman, and K. Fine. 2008. Aging and demographic plasticity in response to experimental age structures in honeybees (Apis mellifera L). Behav Ecol Sociobiol **62**:1621-1631.

Rueppell, O., T. Pankiw, D. I. Nielsen, M. K. Fondrk, M. Beye, and R. E. Page, Jr. 2004. The genetic architecture of the behavioral ontogeny of foraging in honeybee workers. Genetics **167**:1767-1779.

Schmid-Hempel, P. 1998. Parasites in social insects. Princeton University Press, Princeton, N.J.

Schmid-Hempel, P., and R. H. Crozier. 1999. Polyandry versus polygyny versus parasites. Philos Trans R Soc Lond Ser B-Biol Sci **354**:507-515.

Schmid-Hempel, P., and J. Jokela. 2002. Socially structured populations and evolution of recombination under antagonistic coevolution. Am Nat **160**:403-408.

Schmid-Hempel, R., and P. Schmid-Hempel. 2000. Female mating frequencies in *Bombus* spp. from central Europe. Insect. Soc. **47**:36-41.

Seeley, T. D., and D. R. Tarpy. 2007. Queen promiscuity lowers disease within honeybee colonies. Proc Biol Sci **274**:67-72.

Seger, J., and W. D. Hamilton. 1988. Parasites and Sex. Pp. 176-193 *in* R. E. Michod, and B. R. Levin, eds. The Evolution of Sex. Sinauer Associates, Inc, MA.

Serre, D., R. Nadon, and T. J. Hudson. 2005. Large-scale recombination rate patterns are conserved among human populations. Genome Res **15**:1547-1552.

Sherman, P. W. 1979. Insect chromosome-numbers and eusociality. American Naturalist **113**:925-935.

Siller, S. 2001. Sexual selection and the maintenance of sex. Nature **411**:689-692.

Sirvio, A., J. Gadau, O. Rueppell, D. Lamatsch, J. J. Boomsma, P. Pamilo, and R. E. Page, Jr. 2006. High recombination frequency creates genotypic diversity in colonies of the leaf-cutting ant Acromyrmex echinatior. J Evol Biol **19**:1475-1485.

Smith, J. M. 1989. Evolutionary Genetics. Oxford University Press, Oxford.

Solignac, M., F. Mougel, D. Vautrin, M. Monnerot, and J. M. Cornuet. 2007a. A third-generation microsatellite-based linkage map of the honey bee, Apis mellifera, and its comparison with the sequence-based physical map. Genome Biol **8**:R66.

Solignac, M., D. Vautrin, E. Baudry, F. Mougel, A. Loiseau, and J. M. Cornuet. 2004. A microsatellite-based linkage map of the honeybee, Apis mellifera L. Genetics **167**:253-262.

Solignac, M., D. Vautrin, A. Loiseau, F. Mougel, E. Baudry, A. Estoup, L. Garnery, M. Haberl, and J. Cornuet. 2003. Five hundred and fifty microsatellite markers for the study of the honeybee (Apis mellifera L.) genome. Molecular Ecology Notes **3**:307-311.

Solignac, M., L. Zhang, F. Mougel, B. Li, D. Vautrin, M. Monnerot, J. M. Cornuet, K. C. Worley, G. M. Weinstock, and R. A. Gibbs. 2007b. The genome of Apis mellifera: dialog between linkage mapping and sequence assembly. Genome Biol **8**:403.

Stam, P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. The Plant Jounal **3**:739-744.

Tarpy, D. R., and D. J. C. Fletcher. 1998. Effects of relatedness on queen competition within honey bee colonies. Anim Behav **55**:537-543.

Tarpy, D. R., and R. E. Page. 2001. The curious promiscuity of queen honey bees (*Apis mellifera):* evolutionary and behavioral mechanisms. Ann. Zool. Fenn. **38**:255-265.

True, J. R., J. M. Mercer, and C. C. Laurie. 1996. Differences in Crossover Frequency and Distribution Among Three Sibling Species of Drosophila. Genetics **142**:507-523.

Van Valen, L. 1973. A new evolutionary law. Evolutionary Theory **1**:1-30.

Velthuis, H., and A. van Doorn. 2006. A century of advances in bumblebee domestication and the economic and environmental aspects of its commercialization for pollination. Apidologie **37**:421-451.

Wilfert, L., J. Gadau, and P. Schmid-Hempel. 2006. A core linkage map of the bumblebee Bombus terrestris. Genome **49**:1215-1226.

Wilfert, L., J. Gadau, and P. Schmid-Hempel. 2007. Variation in genomic recombination rates among animal taxa and the case of social insects. Heredity **98**:189-197.

Winston, M. L. 1987. The biology of the honey bee. Harvard University Press, Cambridge, Mass.

Woyke, J., J. Wilde, and C. C. Reddy. 2004. Open-air-nesting honey bees *Apis dorsata and Apis laboriosa* differ from the cavity-nesting *Apis mellifera* and *Apis cerana* in brood hygiene behaviour. Journal of Invertebrate Pathology **86**:1-6.

Wright, S. 1938. Size of population and breeding structure in relation to evolution. Science **87**:430-431.

Wright, S. 1931. Evolution in Mendelian populations. Genetics **16**:97-159.

Yamamoto, K., J. Narukawa, K. Kadono-Okuda, J. Nohata, M. Sasanuma, Y. Suetsugu, Y. Banno, H. Fujii, M. R. Goldsmith, and K. Mita. 2006. Construction of a single nucleotide polymorphism linkage map for the silkworm, Bombyx mori, based on bacterial artificial chromosome end sequences. Genetics **173**:151-161.

Zetka, M. C., I. Kawasaki, S. Strome, and F. Muller. 1999. Synapsis and chiasma formation in Caenorhabditis elegans require HIM-3, a meiotic chromosome core component that functions in chromosome segregation. Genes Dev **13**:2258-2270.