Local Variation in Recombination Rates of the Honey Bee (*Apis mellifera*) Genome among Samples from Six Disparate Populations

Timothy DeLory[1, 2], Karen Funderburk[1,3], Katelyn Miller[1], Wendy Zuluaga Smith[1], Samantha McPherson[1,4], Christian W. Pirk[5], Cecilia Costa[6], Érica Weinstein Teixeira[7], Bjørn Dahle[8,9], Olav Rueppell[1*]

[1] Department of Biology, University of North Carolina at Greensboro, NC, USA

[2] Current address: Department of Biology, Utah State University, 5305 Old Main Hill, Logan, UT, USA

[3] Current address: Applied Mathematics for the Life & Social Sciences, College of Liberal Arts and Sciences, Arizona State University, Tempe, AZ, USA

[4] Current address: Current address: NCSU Department of Entomology & Plant Pathology, Campus Box 7613, 100 Derieux Place, Raleigh, NC, USA

[5] Social Insects Research Group, Department of Zoology & Entomology, University of Pretoria, South Africa

[6] Consiglio per la Ricerca in Agricolturae l'Analisi dell'Economia Agraria, Via Po, 14 - 00198 Rome, Italy

[7] Honey Bee Health Specialized Laboratory, Biological Institute, São Paulo State Agribusiness Technology Agency, Av. Prof. Manoel César Ribeiro, 1920, Pindamonhangaba, São Paulo 12411-010, Brazil

[8] Norwegian Beekeepers Association, Kløfta, Norway

[9] Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway

*Corresponding author: 312 Eberhart Bldg, 321 McIver Street, Greensboro NC 27403, USA. Phone: (+1) 336-2562591, Email: olav_rueppell@uncg.edu, ORCID: 0000-0001-5370-4229

Abstract

Meiotic recombination is an essential component of eukaryotic sexual reproduction but its frequency varies within and between genomes. Although it is well-established that honey bees have a high recombination rate with about 20 cM/Mbp, the proximate and ultimate causes of this exceptional rate are poorly understood. Here, we describe six linkage maps of the Western Honey Bee *Apis mellifera* that were produced with consistent methodology from samples from distinct parts of the species' near global distribution. We compared the genome-wide rates and distribution of meiotic crossovers among the six maps and found considerable differences. Overall similarity of local recombination rates among our samples was unrelated to geographic or phylogenetic distance of the populations that our samples were derived from. However, the limited sampling constrains the interpretation of our results because it is unclear how representative these samples are. In contrast to previous studies, we found only in two datasets a significant relation between local recombination rate and GC content. Focusing on regions of particularly increased or decreased recombination in specific maps, we identified several enriched gene ontologies in these regions and speculate about their local adaptive relevance. These data are contributing to an increasing comparative effort to gain an understanding of the intra-specific variability of recombination rates and their evolutionary role in honey bees and other social insects.

Keywords

Genome evolution, genetic diversity, meiotic recombination, intra-specific variability, crossover

Introduction

Meiotic recombination is inextricably linked to sexual reproduction, stabilizing chromosome pairing during meiosis (Fledel-Alon et al. 2009) and contributing to the shuffling of genetic material across generations (Hunter et al. 2016). Recombination events can either transfer only a short genetic sequence during noncrossover events (=gene conversion), or recombine long chromosomal segments when the intermediate double Holliday junction is resolved into a crossover event (Kohl and Sekelsky 2013). The rate of recombination varies considerably among species, partly as a function of physical genome size and chromosome number (De Villena and Sapienza 2001; Lynch 2006), but also due to other factors that remain less well explained (Lenormand et al. 2016; Wilfert et al. 2007). This variation in recombination rate indicates that recombination may evolve due to natural selection (Dumont et al. 2011). Selection for recombination is driven by potential fitness benefits, including the generation of new combinations of beneficial alleles in one individual or purging of deleterious mutations (Hartfield and Keightley 2012).

In a concerted effort to better understand why and how meiotic recombination rate evolves (Stapley et al. 2017), studies of intraspecific variation in recombination are increasingly emerging despite the efforts required. These studies either focus on population differences (Van Oers et al. 2014), inter-individual variation (Johnston et al. 2016), or variation within individuals over time (Langberg et al. 2018; Stevison et al. 2017) and among different genome regions (Myers et al. 2005). In a few cases variation in recombination rates can be directly linked to fitness (Cvetković and Tucić 1986; Kong et al. 2004; Pearson et al. 1970), but no overall consensus on the adaptive significance of recombination rate variation exists (Ritz et al. 2017).

Several social insects in the order Hymenoptera have genome-wide recombination rates that greatly exceed those of other insects, plants, and mammals (Beye et al. 2006; Jones et al. 2019; Liu et al. 2015; Wilfert et al. 2007). These excessive rates can only partly be explained by changes in chromosome number (Ross et al. 2015c). Two ant species (Sirviö et al. 2006), a wasp (Sirviö et al. 2011), and four honey bee species (Beye et al. 2006; Liu et al. 2015; Rueppell et al. 2016; Shi et al. 2013) all show high rates of recombination. Potential explanations of these high recombination rates can be fundamentally categorized into the following hypotheses. The "immune function hypothesis" and "worker diversification hypothesis" (Liu et al. 2015; Jones et al. 2019), both similarly based on the idea that recombination in parallel with multiple mating increases within-colony genetic diversity (Sirviö et al. 2006), benefitting disease resistance and division of labor (Mattila and Seeley 2007; Tarpy and Seeley 2006). However, the increase of genetic variance through high intra-chromosomal recombination is very small for quantitative traits (Rueppell et al. 2012). Accordingly, evidence for increased local recombination rates near immune genes is lacking (Liu et al. 2015), even though disease resistance can be increased by genetic variability (Conlon et al. 2018; Tarpy and Seeley 2006; Wallberg et al. 2014). The "social innovation hypothesis" focuses on the evolutionary history of social insects that is characterized by relatively long generation times, small effective population sizes, and strong selection for evolutionary innovation and independence of caste-specific genes (Kent and Zayed 2013). This idea is supported by the increased local recombination near worker-biased genes that are under positive selection (Harpur et al. 2014; Kent et al. 2012; Liu et al. 2015), although a similar result has also been found in *Megachile rotundata*, a solitary bee (Jones et al. 2019).

The high genome-wide recombination rate has been confirmed multiple times in the Western Honey Bee (*Apis mellifera* L.) in contrast to most other social insects that only have

data from a single linkage map. An early report of an average recombination rate of approximately 20 cM/Mb (Hunt and Page 1995) has been verified in later studies with different genetic markers (Beye et al. 2006; Ross et al. 2015a), and an even higher rate was reported when large numbers of markers are used (Liu et al. 2015). Local recombination rates are positively correlated with GC-content in the overall AT-rich *Apis mellifera* genome in all studies (Bessoltane et al. 2012; Beye et al. 2006; Liu et al. 2015; Mougel et al. 2014; Ross et al. 2015b; Wallberg et al. 2015). Local recombination rates have also been associated with specific sequence motifs (Bessoltane et al. 2012; Mougel et al. 2014), nucleotide diversity, DNA methylation (Wallberg et al. 2015), and gene expression patterns (Kent et al. 2012; Liu et al. 2015) in an effort to better understand the proximate and ultimate causes of the high recombination rate of *A. mellifera*. However, further empirical data are needed to evaluate intra-specific variation and evolution of genome-wide recombination in honey bees.

Although multiple linkage maps exist, few comparative studies have been attempted in *A. mellifera*. As predicted for a species with high recombination rates (Ubeda and Wilkins 2011), the intraspecific correlation of local recombination rates between different linkage maps is modest (Ross et al. 2015a). While local recombination rates across maps were best explained by local G/C content, the variability among maps was more related to microsatellite abundance and specific tri- and tetra-nucleotides (Ross et al. 2015b). However, specific chromosome identity and methodological differences in genotyping technology, marker density, and map construction also influenced the relation of local recombination rates to each other (Ross et al. 2015a). Thus, the methodological differences among different maps limited the conclusions that could be drawn. This prompted us to undertake a systematic comparison of genomic recombination rates in samples derived from six selected *A. mellifera* populations from different parts of the species

distribution. The populations were specifically chosen to test whether local recombination rates may be increased as a result of hybridization or by selection due to local adaptation. Linkage maps from these six sample sets were generated from SNP genotypes determined by sequencing. Correlates of local recombination rates across the corresponding six datasets were investigated to understand how general the association of dinucleotide content with variation in local recombination rate is (Bessoltane et al. 2012; Mougel et al. 2014; Ross et al. 2015b). Additionally, the gene content of genome regions with particularly increased recombination rates in specific datasets was analyzed and the gene content of these regions was interpreted in the context of the particular population histories.

Methods

Sample sets included in this study came from select populations that were very distinct despite the potential for human-mediated transport of honey bees. 1) North Carolina (NC), from the North American domestic population of mixed European origin, reflecting ancestral introductions and the current bee transport and breeding practices (Harpur et al. 2012), 2) Africanized honey bees from Brazil (BR) near the original hybridization zone between European honey bees (mainly *A. mellifera mellifera* and *A. mellifera ligustica*) and the African subspecies *A. mellifera scutellata* from which much of the Americas were colonized (Schneider et al. 2004), 3) native *A. mellifera ligustica*, considered isolated from other subspecies in Italy (IT) (Meixner et al. 2010, 2014), 4) a phenotypically distinguishable wild population of *A. mellifera scutellata* from Pretoria, South Africa (SA) (Ruttner 1988), (5) *A.m. mellifera* from the northern distribution limit of honey bees and isolated from current apiary locations in Norway (NO) (Pinto et al. 2014), and (6) an artificially selected population of Russian honey bees (RU)

imported into the U.S. in 1997 and kept under closely controlled inbreeding since (Bourgeois and Rinderer 2009). Drones were collected from a single colony, as pupae when possible, to ensure that they were the progeny of a single queen. Samples were shipped to North Carolina either frozen (RU), or preserved in acetone (NO), ethanol (IT, SA), or RNAlater (BR) before storage at -80 °C with the samples from North Carolina (NC) until DNA extraction.

Preserved samples were thawed and surplus solvent dried off the samples with Kimwipes™ (Fisher Scientific), while freshly frozen samples were directly mixed with CTAB extraction buffer (Hunt and Page 1995) and ground until no intact tissue remained visible. Subsequent DNA extraction followed standard procedures for genomic DNA isolation using the Epicentre MasterPure™ kits according to the manufacturer's recommendations. The precipitated DNA was resuspended in molecular grade water (G Biosciences), quantified by Nanodrop™ spectrophotometry (Thermo Fisher), and the concentration was adjusted to 100 ng/µl. Successfully extracted samples with sufficient DNA quality and quantity (BR = 35, IT = 91, SA = 70, NO = 97, RU = 91) were shipped to the Genomics Core Lab at Texas A&M University Corpus Christi (genomics.tamucc.edu/) for double-digest RAD-tag (ddRAD) sequencing. Digests were performed with MspI and EcoRI and size selection targeted 200-400bp. Each sample was run in two 192-plexed lanes of an Illumina HiSeq 2500 for 100bp paired-end sequencing (Peterson et al. 2012). A set of 187 drones from North Carolina was whole-genome resequenced at the UNC Chapel-Hill High Throughput Sequencing Facility (www.med.unc.edu/genomics). These samples were sheared with an E220 focused-ultrasonicator (Covaris Inc., Woburn, MA) and quantified before sequencing library construction using the high throughput kit of Kapa Biosystems® (Woburn, MA) following manufacturer's recommendations. Polyethylene glycol and magnetic beads were used for all clean-up and size

selection to 250 – 450 bp. Each library was tagged with dual adapter (2D) indexed adaptors. The samples were separated into two pools that were each sequenced in one lane of an Illumina™ HiSeq 2000 (San Diego, CA) in a 100-bp single-end run and two lanes in a 100-bp paired-end run for adequate coverage: The average read numbers did not vary significantly among datasets ($F_{(4,375)} = 1.6$, $p = 0.169$; mean ± S.D.: BR = 4036730 ± 891222, IT = 6976152 ± 1053596, SA = 14039781 ± 1889196, NO = 10756488 ± 1965725, RU = 8537843 ± 1856207).

Resulting ddRAD reads were combined in individual BAM files. Reads were aligned with Bowtie2 (Langmead and Salzberg 2012) to the Amel_4.5 assembly of the *A. mellifera* genome (Elsik et al. 2014). Aligned reads were processed with cSTACKS (Catchen et al. 2013) to identify SNPs and assign haplotypes at each locus, using a minimum read depth of three and default settings. Mapped loci were ordered according to their genome position and exported as specific csv files, while any markers that did not map to a chromosome were deleted. Further markers were excluded based on the expectation that only two genotypes were expected to be present in each mapping dataset. Occasional genotype calls that deviated from the two main genotypes at a locus were coded as missing data. Markers were also excluded due to an uneven allele distribution ($(G1 - G2)/G1 > .6$), which may indicate non-homologous alignment of the reads to duplicate genome regions. Furthermore, markers were excluded when the number of missing data points exceeded the number of individuals with the less common allele. Individuals with >90% overall missing data were excluded because they were not sufficiently informative. Similarly, the de-multiplexed resequencing data of the NC dataset was aligned to the reference genome with BWA (Li and Durbin 2009) and single nucleotide polymorphisms (SNPs) were identified with SAMtools utilities (Li et al. 2009) with minimal initial filtering. Subsequent quality filtering excluded markers when minor allele frequency was <20%, minor allele counts

<20, or a third allele was detected in >5% of individuals. Markers from this dataset were randomly subsampled to obtain a marker number that was comparable to the ddRAD datasets from the other datasets.

In the ordered marker file, marker genotypes were first assigned an arbitrary phase that was subsequently adjusted based on neighboring markers: The phase of each marker that exhibited >50% recombination to the previous marker was flipped. The resulting genotype matrix was further searched for discordances between the genetic linkage map and physical genome positions of markers by manually assessing the number of recombination events among all markers in close physical proximity. During visual inspection of the data, inconsistencies were recognized as an excess of crossovers between markers in close physical proximity (map expansion) that could be resolved by either rearranging local marker order or excluding markers and recalculating the map. Such markers may have been misaligned due to the evolutionary divergence of the studied populations from the reference genome. The finalized data matrix for each population (Supplemental File 1) was used to count crossover events that were divided by the number of individuals to estimate approximate recombination rates. In our estimation of the overall recombination rates, one crossover event was added for any chromosome in which no crossover was detected because a minimum of one crossover is required for proper chromosome segregation (Fledel-Alon et al. 2009). Double-crossovers are less likely to occur due to interference (Hillers 2004) and consequently the excessive number of double-crossovers in the IT population was removed by replacing the causative singular SNP genotypes with missing data. Local recombination rates (crossovers / Mbp) were estimated in 100 kbp windows as weighted average of the rates of recombination in each marker interval that fell within the respective window (Ross et al. 2015a).

The local recombination rates were compared among the six datasets, assessing overall and chromosome-specific similarity. The average recombination rate for each 100 kbp interval was calculated and the relative deviation of each individual value from that average quantified. For each dataset, the intervals that differed most strongly from the average (highest and lowest 5%) were selected as regions of high or low endemic recombination, respectively, and formatted into two separate BED files with all genes contained in these regions. The BED files were uploaded to HymenopteraMine version 1.3 for GO enrichment analyses based on RefSeq and OGS v.3.2 genes (Elsik et al. 2015). A FDR threshold of 0.1 was applied, based on Benjamini-Hochberg correction for multiple testing (Benjamini and Hochberg 1995).

In addition, the dinucleotide content between each pair of adjacent markers for each dataset was determined and correlated to local recombination rate, excluding two small intervals that exhibited unrealistically high (>1000 cM / Mbp) recombination rates. A Benjamini-Hochberg FDR threshold of 0.05 was applied to identify significant correlations.

Results and Discussion

The genotyping by sequencing of individual offspring from our six samples yielded numerous SNP markers (BR = 8035, IT = 5772, SA = 8035, NO = 5291, RU = 8054, NC = 931350). Conservative filtering reduced these markers to comparable sets of the most reliable markers (BR = 989, IT = 841, SA = 800, NO = 948, RU = 972, NC = 903) for all subsequent analyses. These markers were distributed relatively evenly across all chromosomes and about half of the genome fell within 100kb of a marker (BR = 51%, IT = 44%, SA = 44%, NO = 49%, RU = 49%, NC = 52%). These marker numbers equal or exceed that of some previous linkage maps of *Apis mellifera* (Solignac et al. 2007; Ross et al. 2015a,b) but are orders of magnitude below recent

ultra-high density maps (Liu et al. 2014; Wallberg et al. 2015, Kawakami et al. 2019). Therefore, our analysis is relatively coarse, which limits the exact localization of the crossover position and prevents detection of gene conversion events, an alternative important form of recombination (Comeron et al. 2012; Kohl and Sekelsky 2013). However, the resolution of linkage maps is also determined by the number of total crossovers that can be observed, which in turn is a function of sample size. The honey bee has the advantage of exhibiting a relatively high numbers of crossovers per meiosis and our sample size per dataset, ranging from 35 to 187, permitting us to record and analyze 42303 individual crossover events.

The overall numbers of crossover events per individual ranged from 66.7 (NC), 67.5 (SA), and 72.4 (NO) to 81.0 (BR) and 90.3 (RU), and to 123.5 (IT). Crossover detection in our study was unrelated to marker number (Pearson's R = -0.14, n = 6, p = 0.797), although a positive relation between marker density and recombination rate estimate was previously found at higher marker densities (Liu et al. 2015). Our crossover counts corresponded to genome-wide recombination rate estimates between 28 and 51 cM/Mb. This result suggested a much higher recombination rate in the sample from the old-world IT population compared to the other five datasets and previous estimates (Liu et al. 2015; Ross et al. 2015a; Solignac et al. 2007; Wallberg et al. 2015). Such an exceptional value is unlikely and we favor an alternative explanation: The results could be due to poorer genotyping quality as a result of DNA degradation in the IT samples, leading to inflated recombination rate estimates. This explanation was supported by an exceptional number of double crossovers, which are unlikely due to crossover interference (Hillers 2004; Solignac et al. 2007). Based on our data, it is impossible to determine whether the IT anomaly reflects a population- or individual-specific up-regulation of recombination (by gene conversion) or compromised sample quality. However, even after

removal of these double crossovers, the IT data resulted in an average number of crossovers per individual of 89.8, which was on the high end of estimates together with RU, the other dataset from Eurasia. Including the corrected IT value, the genome-wide recombination rate estimates from our data ranged from 28 to 38 cM/Mb. These values are slightly higher than most previous estimates with comparable marker densities (Beye et al. 2006; Ross et al. 2015b) but similar to the estimates derived from ultra-high marker densities (Liu et al. 2015; Wallberg et al. 2015). These increased values could be due to genotyping errors. In contrast to previous studies with medium marker densities (Ross et al. 2015b; Solignac et al. 2007), we did not remove double crossovers in general to avoid excluding crossover events with short track length (Liu et al. 2015), which also contributes to our higher recombination rate estimates. However, we were willing to risk an overestimation of the overall recombination rates because the inclusion of potential double crossovers increased the likelihood of identifying differences in local recombination rates among datasets, our main objective of this study. Another factor that influences recombination rate estimates could be intra-specific variation in physical genome size, which is commonly neglected in intra-specific comparisons but may be significant (Huang et al. 2014).

Recombination rates were variable among datasets and within each recombination map among genome regions (Figures 1 and S1), which corresponds to previous studies of honey bees (Ross et al. 2015a; Ross et al. 2015b) and other species (Comeron et al. 2012; Coop et al. 2008; Dumont et al. 2011). Even though chromosome-specific crossover rates were significantly correlated in half of the comparisons (all $n = 16$; $R_{BR_NO} = 0.73$, $p = 0.001$; $R_{BR_SA} = 0.70$, $p = 0.002$; $R_{IT_SA} = 0.62$, $p = 0.011$; $R_{BR_RU} = 0.61$, $p = 0.013$; $R_{BR_IT} = 0.52$, $p = 0.038$; $R_{NO_SA} = 0.52$, $p = 0.040$; $R_{IT_NO} = 0.52$, $p = 0.040$), considerable differences among the six maps in each

chromosome existed (Table 1). As expected, the correlations of crossover rates among our maps is much higher than previously reported correlations among maps that were generated with different methods and analyzed at a smaller scale (Ross et al. 2015a). Scale is critical when analyzing the conservation of recombination patterns (Smukowski and Noor 2011), although perhaps less so in honey bees (Ross et al. 2015a).

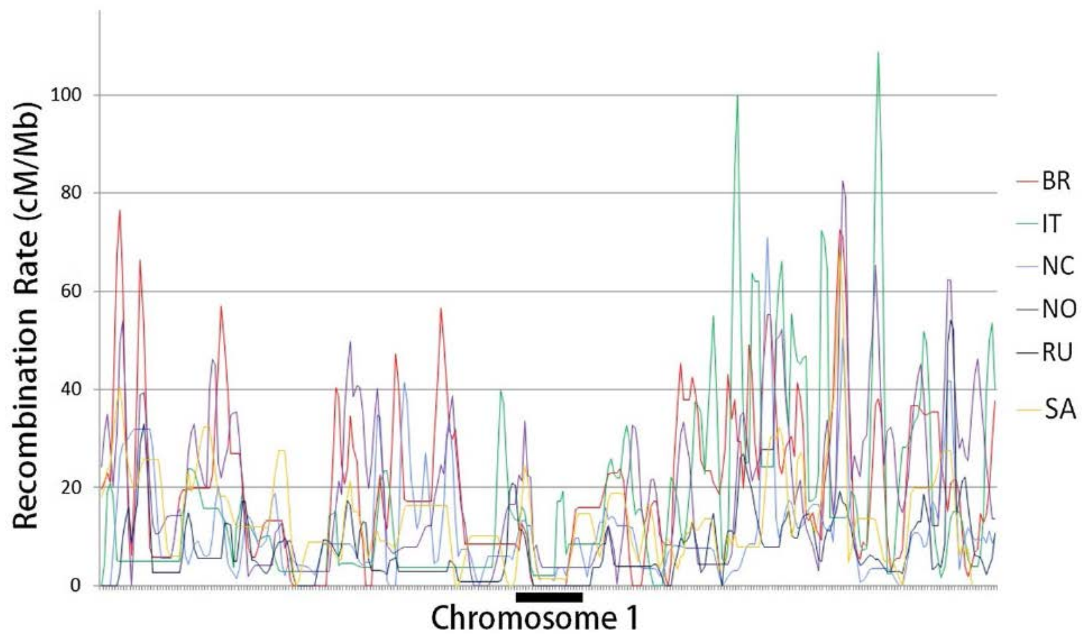

Figure 1: Recombination rate distribution across chromosome 1 in the six investigated populations. Recombination rates calculated for 100 kbp-windows (smoothed over a 300kbp sliding window) varied considerably across six different mapping populations and along the chromosomal location. For each mapping population, drones from a single queen were genotyped via RAD-tag sequencing. The centromere, as mapped by *AvaI* and low GC-regions in (Wallberg et al. 2019), is indicated by a black horizontal bar below the x-axis. Data for chromosome 1 is shown here, and figures for all chromosomes (#1-16) can be found in Supplemental File S2.

Table 1: Chromosome-specific crossover rates (per Mb) in six honey bee populations

Chr.	Brazil (BR)	Italy (IT)	North Carolina (NC)	Norway (NO)	Russians (RU)	South Africa (SA)	Variance
1	0.385	0.564	0.258	0.350	0.449	0.328	0.011
2	0.507	0.652	0.317	0.391	0.559	0.319	0.019
3	0.417	0.581	0.304	0.348	0.363	0.393	0.009
4	0.523	0.594	0.440	0.399	0.500	0.384	0.006
5	0.448	0.599	0.264	0.330	0.407	0.283	0.016
6	0.377	0.684	0.371	0.388	0.323	0.313	0.019
7	0.283	0.396	0.294	0.204	0.411	0.294	0.006
8	0.230	0.496	0.354	0.234	0.257	0.187	0.013
9	0.306	0.683	0.276	0.296	0.491	0.330	0.026
10	0.339	0.670	0.395	0.333	0.315	0.374	0.018
11	0.246	0.439	0.159	0.236	0.392	0.214	0.012
12	0.387	0.440	0.209	0.343	0.483	0.290	0.010
13	0.403	0.606	0.339	0.338	0.420	0.315	0.012
14	0.332	0.502	0.332	0.457	0.391	0.256	0.008
15	0.225	0.456	0.202	0.216	0.355	0.207	0.011
16	0.357	0.503	0.388	0.339	0.401	0.291	0.005

Correlations at the 100kb scale among our six datasets were all highly significant ($p < 0.001$) with Pearson's correlation coefficients ranging from 0.08 to 0.36. Contrary to our expectation, IT was least similar to all others, and overall similarity of recombination rates did

not reflect geographic or phylogenetic proximity of samples (Figure 2). The IT and RU samples both had high overall recombination rates and clustered outside of the other datasets. The BR and NO dataset were most correlated, which contrasts with their geographical distance. However, BR and NO contain genetic material from the M lineage of *A. mellifera* (Wallberg et al. 2014) because subspecies *A. mellifera mellifera* from northern Europe existed in Brazil prior to Africanization (Schneider et al. 2004). The dataset with the next closest pattern of local recombination is NC, representing US domestic stock and thus the C group (Wallberg et al. 2014). Thus, our data provides no evidence for phylogenetic inertia in local recombination rates of *A. mellifera*. This finding is consistent with the hypothesis that high recombination rates may quickly convert sequence motifs that may induce meiotic recombination and thus lead to evolutionary lability of local recombination rates (Ubeda and Wilkins 2011), although our results might also be explained by individual variation (see below). Additionally, the lack of a phylogenetic signal suggests that the conservative filtering of markers excluded most misalignment errors that might have arisen due to phylogenetic distance of our samples to the reference genome.

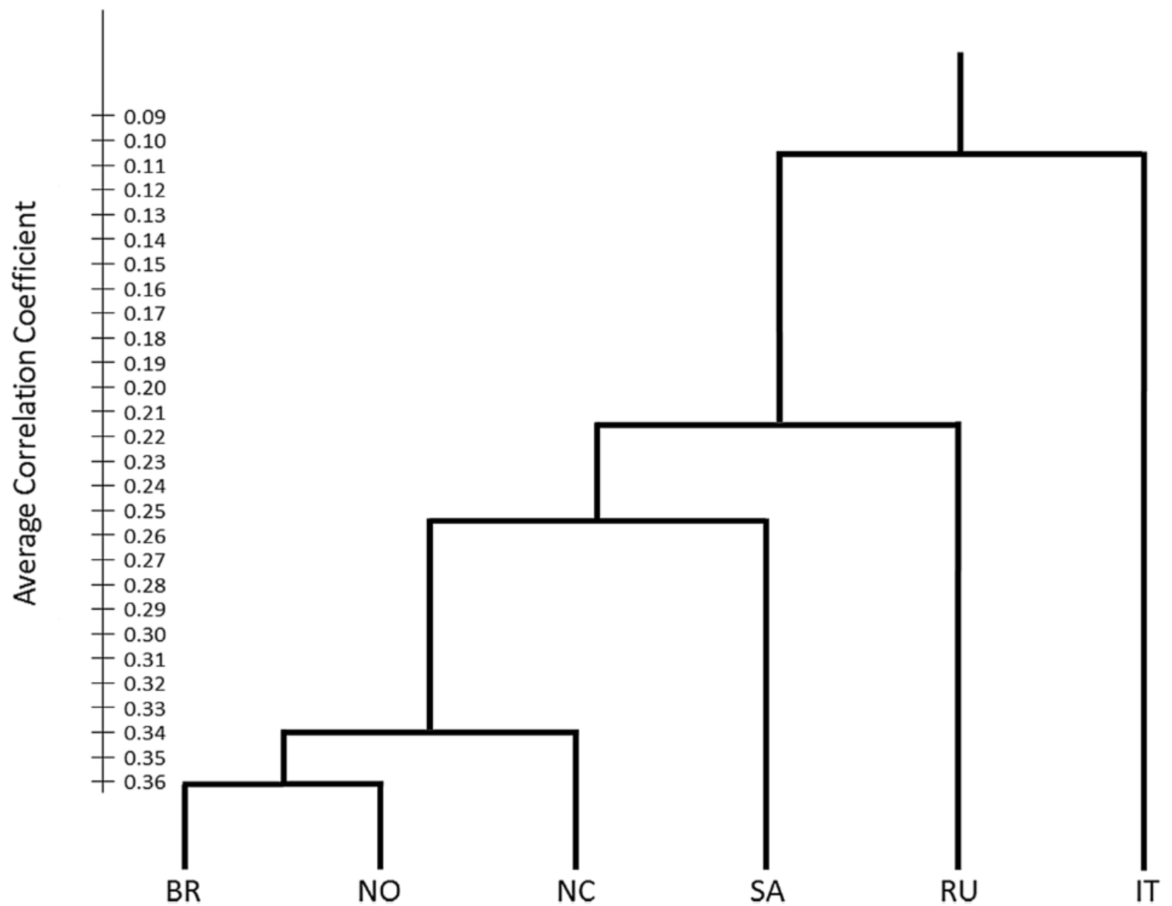

Figure 2: Cluster analysis of mapping populations based on local recombination rates. The recombination similarities did not show an obvious overall relation to evolutionary distance or climate experienced by the sampled populations. Unweighted average Pearson's correlation coefficients of recombination rates in 100kb-windows across all chromosomes were used for the hierarchical clustering.

Each of the six datasets in our study represents only the meiotic products of a single queen and therefore it is impossible to disentangle individual from population-level variation. In the absence of feasible solutions for genotyping thousands of individuals across the genome, this problem represents a severe limitation of the direct measurement of recombination rates by

analyzing individual families and favor estimates based on linkage disequilibrium (Jones et al. 2019). Within-population estimates of individual-level variation in global recombination rates are rare (Ritz et al. 2017) but heritability studies (Johnston et al. 2016; Kawakami et al. 2019) and artificial selection experiments (Charlesworth and Charlesworth 1985) indicate that such variation exists. Estimates of within-population variation in specific regions vary between 1.1- and 2-fold changes across a number of taxa (Ritz et al. 2017). The variation that we selectively analyzed in our dataset exceeded these values, indicating that we could have identified, at least partially, inter-population differences, which are larger than variation within-populations in *Apis mellifera* (Kawakami et al. 2019). Our singular measures cannot guarantee that any particular genome area identified in a particular dataset is representative of the population of origin of that particular queen. However, the analogous assumption that a singular linkage map is representative of its species is commonly made (Beye et al. 2006; Wilfert et al. 2007), and we minimize the risk of stochastic results due to genotyping errors or isolated rare genotypes that are not representative of the population by integrating our analyses to identify patterns across the genome.

Despite our consistent methodology, the datasets exhibited pronounced variation in the detected number of crossovers per individual. Our study cannot distinguish between the "social innovation", the "immune function", and the "worker diversification" hypothesis. However, population differences would be predicted by the former and not necessarily by the latter two. Furthermore, we did not confirm our prediction of high recombination rates on the fringes of the species distribution, which combine small population sizes and presumably strong environmental selection. While recombination was frequent in the RU samples, potentially due to small population size during recent breeding efforts (Bourgeois and Rinderer 2009), the NO samples

exhibited an intermediate recombination rate. The values from the two datasets that were collected from populations that may be deemed most ancestral were on both ends of the spectrum: The IT dataset, representing the C-lineage (Wallberg et al. 2014), had a very high rate of recombination, while recombination was very low in the SA samples, representing the A-lineage. Similarly, the two datasets from the Americas (BR and NC) were not concordant with each other, suggesting that recent hybridization does not invariably decrease overall recombination rates (Williams et al. 1995).

The relation between local dinucleotide content and recombination rate was variable among datasets (Table 2). Significant relations between dinucleotide content and local recombination rates were only found in the NC and the BR datasets. The significant correlations followed previous findings that GC-rich dinucleotides were generally positively associated with recombination rate, while the content of AT-rich dinucleotides showed negative correlations (Beye et al. 2006; Liu et al. 2015; Ross et al. 2015b). Our correlations were consistent among the two datasets and were also reflected in the other datasets, despite the lack of statistical significance of those analyses. This lack of significant correlations may have arisen due to lower power to detect these patterns when compared to studies with higher marker density (Liu et al. 2015, Wallberg et al. 2014), and it was unexpected because the general association between GC-content and local recombination rate (Mugal et al. 2015) results automatically in some positive and negative correlations of different dinucleotides with recombination rate (Ross et al. 2015b). The two new world datasets exhibit the expected correlations in contrast to the four remaining datasets from the old world. However, previous old world maps had also shown the expected correlations (Solignac et al. 2007), similar to new world crosses (Ross et al. 2015b). The reference genome Amel_4.5 (Elsik et al. 2014) has a larger phylogenetic distance to the samples

from the old world than from the new world, which may contribute to this difference. Particularly in severely bottlenecked populations derived from the distribution limit of *A. mellifera*, as it might be the case for our Russian and Norway samples, the high recombination rate combined with strong selection and drift may have altered local base composition, which is corroborated by a high sequence diversity in regions of high recombination in *A. mellifera* (Wallberg et al. 2015).

Table 2: Correlations between local (100kbp) dinucleotide content and crossover rates (FDR<0.05 in bold).

Dinucleotide	**Brazil**	**Italy**	**North Carolina**	**Norway**	**Russians**	**South Africa**
AA	**-0.11**	0.00	-0.04	-0.03	0.00	-0.03
AC	0.00	0.00	-0.01	-0.01	-0.08	0.03
AT	**-0.10**	0.00	**-0.10**	0.00	0.00	-0.04
AG	**0.10**	0.00	**0.11**	0.02	0.01	0.02
CA	-0.02	-0.02	**-0.12**	-0.02	-0.05	-0.01
CC	0.06	0.02	**0.09**	0.00	0.03	0.03
CG	**0.10**	0.01	0.06	0.00	0.02	0.05
CT	**0.10**	0.00	**0.10**	-0.04	0.00	0.03
GA	**0.07**	0.00	0.06	0.01	-0.01	0.02
GC	**0.11**	0.00	**0.08**	0.00	0.02	0.06
GG	**0.15**	0.02	**0.10**	0.00	-0.01	0.04
GT	**0.09**	-0.05	**0.10**	0.02	-0.01	0.01
TA	**-0.09**	0.00	**-0.08**	0.00	-0.01	-0.04
TC	**0.11**	0.01	0.05	-0.01	0.04	0.02
TG	**-0.07**	-0.10	0.03	-0.01	-0.08	-0.04
TT	**-0.10**	-0.01	**-0.07**	0.02	0.02	-0.03

Genome regions that had either exceptionally high or low crossover rates in one of the datasets compared to the others contained between 544 and 813 genes (BR_{High} = 544, BR_{Low} = 586; IT_{High} = 550, IT_{Low} = 598; NC_{High} = 656, NC_{Low} = 603; NO_{High} = 574, NO_{Low} = 609; RU_{High} = 595, RU_{Low} = 707; SA_{High} = 813, SA_{Low} = 562). Such regions of specifically increased or decreased recombination may just represent stochastic individual variation or arise through selectively neutral processes, but they may also indicate microevolution of local recombination in response specific alleles in these genome regions being selected due to local adaptation (Hill and Robertson 1966; Otto and Barton 2001). Such selection has not been empirically demonstrated but would occur presumably by cis-regulation (Hunter et al. 2016). Genome regions with elevated recombination rates might reflect regions under selection for decoupling allelic variants, while genome regions with suppressed recombination may be selected for maintaining allelic combinations. We selected a genome-wide, general analysis without investigating individual intervals because we cannot exclude that individual variation obscures some population differences and that some of the variation measured reflects stochastic variation. This problem should be particularly prominent in BR due to the smaller sample size, but no significantly enriched GO terms were found in BR, suggesting that our high-level analysis approach was effective at avoiding stochastic artefacts.

Across all 6 datasets, 33 significantly enriched GO terms were identified. Many terms were partially overlapping and the 33 terms consequently represented few distinct categories. The majority of these terms (20) were related to molecular functions (Table 3) and the remaining (13) were related to biological function (Table 4). All 13 enriched biological functions were found in genome regions where crossover rates were exceptionally high in the NO samples. The remaining 20 instances of enriched GO terms for molecular functions were identified from NO

(3 terms in comparatively high recombination regions and 7 in comparatively low regions), NC (9 terms in high regions), and RU (one term in high region). We did not find any enriched GO terms within regions of relatively high or low recombination in BR, IT or SA.

Table 3 Enriched molecular functions in regions with particularly high or low recombination in a specific population (FDR<.05).

Popu lation	**Rec. Rate**	**Biological Function**	**GO ID**	**FDR**
NC	High	olfactory receptor activity	GO:0004984	< 0.0001
NC	High	odorant binding	GO:0005549	0.0003
NC	High	transmembrane receptor activity	GO:0099600	0.005
NC	High	transmembrane signaling receptor activity	GO:0004888	0.005
NC	High	signaling receptor activity	GO:0038023	0.006
NC	High	receptor activity	GO:0004872	0.008
NC	High	molecular transducer activity	GO:0060089	0.008
NC	High	alcohol dehydrogenase (NAD) activity	GO:0004022	0.01
NC	High	signal transducer activity	GO:0004871	0.01
NO	Low	iron ion binding	GO:0005506	< 0.0001
NO	Low	monooxygenase activity	GO:0004497	< 0.0001
NO	Low	heme binding	GO:0020037	< 0.0001
NO	Low	tetrapyrrole binding	GO:0046906	< 0.0001
NO	Low	oxidoreductase activity, acting on paired donors	GO:0016705	< 0.0001
NO	Low	oxidoreductase activity	GO:0016491	0.0003
NO	Low	transition metal ion binding	GO:0046914	0.02
NO	High	inositol monophosphate 1-phosphatase activity	GO:0008934	0.02
NO	High	inositol phosphate phosphatase activity	GO:0052745	0.02
NO	High	inositol monophosphate phosphatase activity	GO:0052834	0.02
RU	High	peptidase activity	GO:0008233	0.03

Table 4 Enriched biological functions in regions with particularly high recombination in the NO population (FDR<.05).*

GO Term	GO ID	FDR	GO Term	GO ID	FDR
neurotransmitter secretion	GO:0007269	0.01	phosphatidyl-inositol phosphorylation	GO:0046854	0.02
signal release	GO:0023061	0.01	chemical synaptic transmission	GO:0007268	0.02
presynaptic process involved in chemical synaptic transmission	GO:0099531	0.01	anterograde trans-synaptic signaling	GO:0098916	0.02
signal release from synapse	GO:0099643	0.01	synaptic signaling	GO:0099536	0.02
regulation of neurotransmitter levels	GO:0001505	0.02	trans-synaptic signaling	GO:0099537	0.02
neurotransmitter transport	GO:0006836	0.02	lipid modification	GO:0030258	0.03
lipid phosphorylation	GO:0046834	0.02			

* No significant GO enrichment was found in low recombination regions of the NO population or in any other populations

All GO terms identified in the NC dataset were associated with olfactory signal transduction, except for "alcohol dehydrogenase activity". Chemosensation plays an important role in the adaptive success of insects (Leal 2013) and for plant-dependent honey bees in particular (Robertson and Wanner 2006). We speculate that the adaptation of honey bees of European descent to the North American continent might have increased selection for certain olfactory pathways, and odorant receptors are among positively selected genes across different

populations in *A. mellifera* (Wallberg et al. 2014). Olfactory genes have diversified in *A. mellifera* (Kapheim et al. 2015), and crossover events within gene families accelerate the rate at which novel alleles can be combined (Reich et al. 2002). Due to the many potential physiological roles of alcohol dehydrogenases (Kavanagh et al. 2008), their overrepresentation in highly recombining genome regions of the NC data is harder to interpret but could be linked to selection for detoxifying novel secondary plant metabolites or pheromone synthesis (Hasegawa et al. 2009). "Peptidase activity", the sole GO term found significantly enriched in genome regions of relatively high recombination in the RU dataset, is another very general GO term that could relate to selection for diverse functions, such as digestion, peptide hormone signaling, protein metabolism, and immune functions (Tundo et al. 2017).

The NO dataset had the highest number of GO terms. We speculate that this result could be due to stronger selection related to climate adaptation at the northern distribution limit of *A. mellifera* (Amdam et al. 2005). One cluster of GO terms in the regions of decreased recombination in the NO samples was related to iron-based oxidoreductase activity. The reduced recombination around oxidoreductase genes could indicate stabilizing selection for certain allelic combinations in energy metabolism and specifically adaptive thermogenesis (Lowell and Spiegelman 2000). However, oxidoreductases have numerous other functions, including detoxification and other oxidative transformations (Feyereisen 1999) that may be adaptive in an environment of extremely long winters that honey bees spend in their hives persisting on stored food (Southwick and Heldmaier 1987). The other cluster of significant GO terms for the NO samples was discovered in regions of increased recombination and related to the enzymatic function of inositol phosphatase, which was mirrored by numerous biological function GO terms related to neurotransmission. Inositol phosphatase, and signal transduction in general, may be

positively selected due to behavioral adaptations, as evidenced by their accelerated evolution in eusocial lineages (Kapheim et al. 2015). Furthermore, the inositol monophosphatase gene family plays key roles in lipid metabolism and may be related to diapause (Kocher et al. 2013). The related term of lipid phosphorylation has been correlated with anti-parasite defenses in social insects (Alleman et al. 2018) but it is unclear whether this can be related to disease pressure.

In conclusion, our comparison of recombination patterns among six datasets from across the world yielded more similarities than a previous study (Ross et al. 2015a), which made the analysis of local recombination rates comparisons more meaningful despite the fact that each dataset essentially only represents one data point per population for each genome region. It is unclear how much of the variation is population-specific and how much is individual variation and our understanding of this variation is only beginning to emerge (Kawakami et al. 2019). It is also important to note that the spatial resolution achieved by our marker density, although an improvement to previous comparative efforts (Ross et al. 2015a), is still insufficient to provide accurate estimates for each 100kb interval in the honey bee genome. Elevated recombination rates are generally associated with genetic diversity, tissue-specific expression, and evolutionary innovation in honey bees (Kent et al. 2012; Liu et al. 2015; Wallberg et al. 2015) and the results presented here complement these findings. Concordantly, we also find that locally increased recombination occurs in genome regions with genes that might play important roles in adaptive evolution. Despite the limited sampling, this study significantly increases the empirical data base for intra-specific studies of recombination in honey bees. However, to understand this important model for the evolution of high recombination rates in social insects, further population-level studies are needed.

Acknowledgements

We would like to thank all instructors and fellow students of the Math-Bio REU-site at UNCG for their support and collegial working atmosphere. We appreciate the opportunity to sample "Russian" colonies from Steven Coy (then President of the Russian Honey Bee Breeders) in Perkinston, Mississippi. We would like to thank Sharon Furiness and the Genomics Core Lab at TAMU Corpus Christi for performing the ddRAD library preparation, SE –sequencing, and SNP extraction. Esmaeil Amiri assisted in the data management. We would also like to thank three anonymous reviewers whose comments on a previous version of this manuscript helped us to improve the presentation of our study. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R15GM102753. Further financial support was provided by the US National Science Foundation (DMS #1359187) and UNCG.

References

Alleman A, Feldmeyer B, Foitzik S (2018) Comparative analyses of co-evolving host-parasite associations reveal unique gene expression patterns underlying slavemaker raiding and host defensive phenotypes. Scientific Reports 8:1951

Amdam GV, Norberg K, Omholt SW, Kryger P, Lourenco AP, Bitondi MMG, Simoes ZLP (2005) Higher vitellogenin concentrations in honey bee workers may be an adaptation to life in temperate climates. Insectes Sociaux 52:316-319

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 57:289-300

Bessoltane N, Toffano-Nioche C, Solignac M, Mougel F (2012) Fine scale analysis of crossover and non-crossover and detection of recombination sequence motifs in the honeybee (*Apis mellifera*). PLoS One 7:e36229

Beye M et al. (2006) Exceptionally high levels of recombination across the honey bee genome. Genome Res 16:1339-1344

Bourgeois AL, Rinderer TE (2009) Genetic characterization of Russian Honey Bee stock selected for improved resistance to *Varroa destructor*. J Econ Entomol 102:1233-1238

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. Mol Ecol 22:3124-3140

Charlesworth B, Charlesworth D (1985) Genetic variation in recombination. I. Responses to selection and preliminary genetic analysis. Heredity 54:71-83

Comeron JM, Ratnappan R, Bailin S (2012) The many landscapes of recombination in *Drosophila melanogaster*. PLoS Genet 8:e1002905

Conlon B, Frey E, Rosenkranz P, Locke B, Moritz R, Routtu J (2018) The role of epistatic interactions underpinning resistance to parasitic *Varroa* mites in haploid honeybee drones. Journal of Evolutionary Biology 31:801-809

Coop G, Wen XQ, Ober C, Pritchard JK, Przeworski M (2008) High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. Science 319:1395-1398

Cvetković D, Tucić N (1986) Female recombination rates and fitness in *Drosophila melanogaster*. Journal of Zoological Systematics and Evolutionary Research 24:198-207

De Villena FP-M, Sapienza C (2001) Recombination is proportional to the number of chromosome arms in mammals. Mammalian Genome 12:318-322

Dumont BL, White MA, Steffy B, Wiltshire T, Payseur BA (2011) Extensive recombination rate variation in the house mouse species complex inferred from genetic linkage maps. Genome Res 21:114-125

Elsik C et al. (2014) Finding the missing honey bee genes: lessons learned from a genome upgrade. BMC Genomics 15:86

Elsik CG, Tayal A, Diesh CM, Unni DR, Emery ML, Nguyen HN, Hagen DE (2015) Hymenoptera Genome Database: integrating genome annotations in HymenopteraMine. Nucleic Acids Res 44:D793-D800

Feyereisen R (1999) Insect P450 enzymes. Annu Rev Entomol 44:507-533

Fledel-Alon A, Wilson DJ, Broman K, Wen X, Ober C, Coop G, Przeworski M (2009) Broad-scale recombination patterns underlying proper disjunction in humans. PLoS Genet 5:e1000658

Harpur BA, Kent CF, Molodtsova D, Lebon JM, Alqarni AS, Owayss AA, Zayed A (2014) Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. Proceedings of the National Academy of Sciences 111:2614-2619

Harpur BA, Minaei S, Kent CF, Zayed A (2012) Management increases genetic diversity of honey bees via admixture. Mol Ecol 21:4414-4421

Hartfield M, Keightley PD (2012) Current hypotheses for the evolution of sex and recombination. Integrative Zoology 7:192-209

Hasegawa M et al. (2009) Differential gene expression in the mandibular glands of queen and worker honeybees, *Apis mellifera* L.: implications for caste-selective aldehyde and fatty acid metabolism. Insect Biochem Molec 39:661-667

Hill WG, Robertson A (1966) The effect of linkage on limits to artificial selection. Genet Res 8:269-294

Hillers KJ (2004) Crossover interference. Current Biology 14:R1036-R1037

Huang W et al. (2014) Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. Genome Res 24:1193-1208

Hunt GJ, Page RE (1995) Linkage map of the honey bee, *Apis mellifera*, based on RAPD markers. Genetics 139:1371-1382

Hunter CM, Huang W, Mackay TF, Singh ND (2016) The genetic architecture of natural variation in recombination rate in *Drosophila melanogaster.* PLoS Genet 12:e1005951

Johnston SE, Bérénos C, Slate J, Pemberton JM (2016) Conserved genetic architecture underlying individual recombination rate variation in a wild population of Soay sheep (*Ovis aries*). Genetics 203:583-598

Jones JC, Wallberg A, Christmas MJ, Kapheim KM, Webster MT (2019) Extreme differences in recombination rate between the genomes of a solitary and a social bee. Mol Biol Evol 36:2277-2291

Kapheim KM et al. (2015) Genomic signatures of evolutionary transitions from solitary to group living. Science 348:1139-1143

Kavanagh K, Jörnvall H, Persson B, Oppermann U (2008) Medium-and short-chain dehydrogenase/reductase gene and protein families. Cellular and Molecular Life Sciences 65:3895

Kawakami T, Wallberg A, Olsson A, Wintermantel D, de Miranda JR, Allsopp M, Rundlöf M, Webster MT (2019) Substantial heritable variation in recombination rate on multiple scales in honeybees and bumblebees. Genetics 212: 1101-1119

Kent CF, Minaei S, Harpur BA, Zayed A (2012) Recombination is associated with the evolution of genome structure and worker behavior in honey bees Proc Natl Acad Sci U S A 109:18012-18017

Kent CF, Zayed A (2013) Evolution of recombination and genome structure in eusocial insects. Commun Integr Biol 6:e22919

Kocher SD et al. (2013) The draft genome of a socially polymorphic halictid bee, *Lasioglossum albipes.* Genome Biology 14:R142

Kohl KP, Sekelsky J (2013) Meiotic and mitotic recombination in meiosis Genetics 194:327-334

Kong A et al. (2004) Recombination rate and reproductive success in humans. Nature Genetics 36:1203

Langberg K, Phillips M, Rueppell O (2018) Testing the effect of paraquat exposure on genomic recombination rates in queens of the western honey bee, *Apis mellifera.* Genetica 146:171-178

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357-359

Leal WS (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. Annu Rev Entomol 58:373-391

Lenormand T, Engelstädter J, Johnston SE, Wijnker E, Haag CR (2016) Evolutionary mysteries in meiosis. Phil Trans R Soc B 371:20160001

Lespinet O, Wolf YI, Koonin EV, Aravind L (2002) The role of lineage-specific gene family expansion in the evolution of eukaryotes. Genome Res 12:1048-1059

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754-1760

Li H et al. (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079

Liu H, Zhang X, Huang J, Chen J-Q, Tian D, Hurst L, Yang S (2015) Causes and consequences of crossing-over evidenced via a high-resolution recombinational landscape of the honey bee. Genome Biology 16:15

Lowell BB, Spiegelman BM (2000) Towards a molecular understanding of adaptive thermogenesis. Nature 404:652

Lynch M (2006) The origins of eukaryotic gene structure. Mol Biol Evol 23:450-468

Mattila HR, Seeley TD (2007) Genetic diversity in honey bee colonies enhances productivity and fitness. Science 317:362-364

Meixner MD, Costa C, Kryger P, Hatjina F, Bouga M, Ivanova E, Büchler R (2010) Conserving diversity and vitality for honey bee breeding. Journal of Apicultural Research 49:85-92

Meixner MD, Büchler R, Costa C, Francis RM, Hatjina F, Kryger P, Uzunov A, Carreck NL (2014) Honey bee genotypes and the environment. Journal of Apicultural Research 53:183-187

Mougel F, Poursat M-A, Beaume N, Vautrin D, Solignac M (2014) High-resolution linkage map for two honeybee chromosomes: the hotspot quest. Molecular Genetics and Genomics 289:11-24

Mugal CF, Weber CC, Ellegren H (2015) GC-biased gene conversion links the recombination landscape and demography to genomic base composition: GC-biased gene conversion drives genomic base composition across a wide range of species. Bioessays 37:1317-1326

Myers S, Bottolo L, Freeman C, McVean G, Donnelly P (2005) A fine-scale map of recombination rates and hotspots across the human genome. Science 310:321-324

Otto SP, Barton NH (2001) Selection for recombination in small populations. Evolution 55:1921-1931

Pearson P, Ellis J, Evans H (1970) A gross reduction in chiasma formation during meiotic prophase and a defective DNA repair mechanism associated with a case of human male infertility. Cytogenetic and Genome Research 9:460-467

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species PLoS One 7:e37135

Pinto MA et al. (2014) Genetic integrity of the Dark European honey bee (*Apis mellifera mellifera*) from protected populations: a genome-wide assessment using SNPs and mtDNA sequence data. Journal of Apicultural Research 53:269-278

Reich DE et al. (2002) Human genome sequence variation and the influence of gene history, mutation and recombination. Nature Genetics 32:135

Ritz KR, Noor MA, Singh ND (2017) Variation in recombination rate: adaptive or not? Trends in Genetics 33:364-374

Robertson HM, Wanner KW (2006) The chemoreceptor superfamily in the honey bee, *Apis mellifera*: Expansion of the odorant, but not gustatory, receptor family. Genome Res 16:1395-1403

Ross C, DeFelice D, Hunt G, Ihle K, Rueppell O (2015a) A comparison of multiple genome-wide recombination maps in *Apis mellifera*. In: Collaborative Mathematics and Statistics Research. Springer International Publishing, pp 91-98

Ross CR, DeFelice DS, Hunt GJ, Ihle KE, Amdam GV, Rueppell O (2015b) Genomic correlates of recombination rate and its variability across eight recombination maps in the western honey bee (*Apis mellifera* L.). BMC Genomics 16:107

Ross L, Blackmon H, Lorite P, Gokhman V, Hardy N (2015c) Recombination, chromosome number and eusociality in the Hymenoptera. Journal of Evolutionary Biology 28:105-116

Rueppell O et al. (2016) A new metazoan recombination rate record and consistently high recombination rates in the honey bee genus *Apis* accompanied by frequent inversions but not translocations. Genome Biology and Evolution 8:3653-3660

Rueppell O, Meier S, Deutsch R (2012) Multiple mating but not recombination causes quantitative increase in offspring genetic diversity for varying genetic architectures PLoS One 7:e47220

Ruttner F (1988) Biogeography and Taxonomy of Honeybees. Springer, Berlin

Schneider SS, DeGrandi-Hoffman G, Smith DR (2004) The African honey bee: Factors contributing to a successful biological invasion. Annu Rev Entomol 49:351–376

Shi YY, Sun LX, Huang ZY, Wu XB, Zhu YQ, Zheng HJ, Zeng ZJ (2013) A SNP based high-density linkage map of *Apis cerana* reveals a high recombination rate similar to *Apis mellifera.* PLoS One 8:e76459

Sirviö A, Gadau J, Rueppell O, Lamatsch D, Boomsma JJ, Pamilo P, Page RE (2006) High recombination frequency creates genotypic diversity in colonies of the leaf-cutting ant *Acromyrmex echinatior.* Journal of Evolutionary Biology 19:1475-1485

Sirviö A, Johnston JS, Wenseleers T, Pamilo P (2011) A high recombination rate in eusocial Hymenoptera: evidence from the common wasp *Vespula vulgaris.* BMC Genetics 12:95

Smukowski CS, Noor MAF (2011) Recombination rate variation in closely related species. Heredity 107:496-508

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

Southwick EE, Heldmaier G (1987) Temperature control in honey bee colonies. Bioscience 37:395-399

Stapley J, Feulner PG, Johnston SE, Santure AW, Smadja CM (2017) Variation in recombination frequency and distribution across eukaryotes: patterns and processes. Phil Trans R Soc B 372:20160455

Stevison LS, Sefick S, Rushton C, Graze RM (2017) Recombination rate plasticity: revealing mechanisms by design. Phil Trans R Soc B 372:20160459

Tarpy DR, Seeley TD (2006) Lower disease infections in honeybee (*Apis mellifera*) colonies headed by polyandrous vs monandrous queens. Naturwissenschaften 93:195-199

Tundo GR et al. (2017) Multiple functions of insulin-degrading enzyme: a metabolic crosslight? Critical Reviews in Biochemistry and Molecular Biology 52:554-582

Ubeda F, Wilkins JF (2011) The Red Queen theory of recombination hotspots Journal of Evolutionary Biology 24:541-553

Van Oers K et al. (2014) Replicated high-density genetic maps of two great tit populations reveal fine-scale genomic departures from sex-equal recombination rates. Heredity 112:307

Wallberg A et al. (2019) A hybrid de novo genome assembly of the honeybee, *Apis mellifera*, with chromosome-length scaffolds. BMC Genomics 20:275

Wallberg A, Glémin S, Webster MT (2015) Extreme recombination frequencies shape genome variation and evolution in the Honeybee, *Apis mellifera.* PLoS Genet 11:e1005189

Wallberg A et al. (2014) A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honeybee *Apis mellifera.* Nature Genetics 46:1081-1090

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

Williams CG, Goodman MM, Stuber CW (1995) Comparative recombination distances among *Zea mays* L. inbreds, wide crosses and interspecific hybrids. Genetics 141:1573-1581