Original article

The honey bee genome– what has it been good for?

Amy L. Toth[1], Amro Zayed[2]

[1]Department of Ecology, Evolution, and Organismal Biology and Department of Entomology, Iowa State University, 251 Bessey Hall, Ames, IA 50014, USA
[2]Department of Biology, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada

Abstract – In 2006, the full complement of DNA sequence information (or 'genome') of the Western honey bee, *Apis mellifera* , was published. This important resource was one of the most important advances in the history of honey bee research, with seemingly limitless applications to unlocking the secrets of honey bee biology and social life and for improving health, breeding and management. Honey bee genomics has seen immense growth in the past one and a half decades. In this article, we reflect on what the genome has added to our understanding of fundamental aspects of honey bee biology, including evolutionary origins, behaviour and health/disease. We conclude that while the genome has fuelled growth in many areas of honey bee research, it is only one part of an emerging systems-based, multi-omics approach. Moving forward, we posit that honey bee research will benefit most from an even fuller integration of genomics with classical approaches in evolution, ethology, physiology and microbiology.

Apis mellifera / genomics / transcriptomics / genetics / sequencing

1. INTRODUCTION

The Western honey bee *Apis mellifera* is without a doubt one of the most important insects to humanity, due to its global distribution, economic benefits and long history of interaction with humans. In 2006, *Apis mellifera* joined the ranks of the genome-enabled, being only the third insect (after the model *Drosophila melanogaster* and the malaria vector mosquito *Anopheles gambiae*) to have its genome sequenced (Robinson 2002). This advance also marked the first eusocial insect genome sequence, and this new influx of data signified an unprecedented promise for behavioural genetics, with hopes that the genome would reveal novel and dramatic differences from the genomes of non-social insects. On the cover of *Nature* , with the initial release of the genome, appeared the proclamation 'A blueprint for sociality' (Honey Bee Genome Sequencing Consortium 2006), illustrating the immense optimism surrounding the potential for the honey bee genome to elucidate the molecular secrets of complex social life. Today, most geneticists would agree that a genome is not a blueprint for building an organism; rather it is a dynamic set of instructions, subject to interpretation by the environment in which it occurs (Bell and Robinson 2011). Genomes are only the first step in a multi-level regulatory process that takes an organism from genotype to phenotype. Upon celebrating 50 years of *Apidologie* , and 14 years of furious research surrounding honey bee genomics, we believe we have reached a reasonable moment to retrospectively ask of the honey bee genome 'What *has it* been good for?'

The human genome (Venter et al., 2001, International Human Genome Sequencing

Consortium 2001), sequenced 5 years earlier than the genome of the honey bee, was also hailed to behold the secrets to mysteries of human traits, history and health. It is undeniable that human genomic information has led to major advances in our understanding of human diversity, evolutionary history, our relationship to other species, susceptibility to common diseases, identification of variants associated with rare diseases and identification of variants influencing inter-individual variation in phenotypic traits (Fine 2019). However, some have suggested that the human genome has failed to live up to its expectations (Daiger 2005), leaving more questions than answers. Numerous genotypic linkage and genome-wide association (GWAS) studies have consistently uncovered larger-than-expected amounts of 'missing heritability' (Maher 2008). That is, full genome sequence information not only fails to explain the majority of phenotypic variation, but also fails to explain much *hereditary* variation in a phenotype. Other factors, such as rare variants, gene by environment interactions and epigenetic inheritance, may be part of the mystery (Eichler et al. 2008, Manolio et al. 2009, Zuk et al. 2012). Nonetheless, this serves as a lesson that too big of a fixation on the DNA sequence may be putting too much faith in the explanatory power of the genome.

As doctoral students in 2006, we personally experienced the feverish atmosphere and great expectations arising from the publication of 'our' bee's 230 million As, Ts, Gs and Cs! The excitement was palpable; many bee researchers had a sense that everything we knew about the honey bee was about to change (see 2006 vintage T-shirt as evidence of this excitement, Figure 1). We had high hopes that bee health and management would be improved, great discoveries were going to be made, and longstanding hypotheses were going to be tested and perhaps even overthrown. Fourteen years and many, many studies later, it is a reasonable time to reflect on bee research progress since the honey bee genome's release. Today, honey bee genomics is an expanding and flourishing field of study, as exemplified by a large number of published papers related to the honey bee genome (Figure 2), along with numerous recent reviews of insights from genomics into honey bee biology, health and disease (Zayed and Robinson 2012, Grozinger and Robinson 2015,

Dolezal and Toth 2014, Grozinger and Zayed 2020). Thus, our goal in this article is not to review these topics exhaustively, but rather to present a few illustrative examples and ask—has the genome lived up to all of the hype? What have we learned from the honey bee genome and what have we not learned? If we imagine a world without the honey bee genome, what major insights into honey bee biology would be lost? What difference has this resource made to beekeeping and bee management and our understanding of bee health? What still remains to be learned from the genome that has not yet been addressed, of what new questions still remained to be cracked? What are some secrets of honey bee biology that the honey bee genome cannot or will never tell us? Below, we seek to assess these types of questions, providing a series of examples to illustrate some of the major successes and a few unresolved issues (no need to call them 'failures') from research based on the honey bee genome.

2. CASE STUDIES: WHAT HAPPENED WHEN HONEY BEE GENOMICS WAS APPLIED TO FUNDAMENTAL QUESTIONS IN BEE BIOLOGY AND HEALTH?

1) On the origin of *Apis:* honey bee lineages and evolutionary history

From a technological view point, the field of honey bee biogeography has come a long way since Ruttner's classic morphometric analyses of *Apis mellifera*'s native races in Africa, Europe and Asia (Ruttner 1988), although we are still searching for answers on where the honey bee came from. Based on principal component analysis of morphometric data, Ruttner hypothesized that different subspecies of *Apis mellifera* can be grouped into European, African or Asian lineages. Using maternally inherited mitochondrial markers, the first wave of population genetic analyses of honey bees confirmed the existence of genetically distinct subspecies that can be further grouped into distinct lineages in Europe (M or C), Africa (A) or Asia (Garnery et al. 1992, Garnery et al. 1993, Arias and Sheppard 1996, Sheppard and Meixner 2003). In the early 1990s, a large

Figure 1. The first author's 2006 vintage T-shirt (from the University of Illinois, Gene Robinson's research group) captures the excitement and great promise expected from the honey bee genome. **a** Front of the T-shirt reads 'The Honey Bee Genome Project'. **b** Back of the T-shirt reads 'Beenomics: For a Better World'.

number of microsatellite loci were sequenced (Estoup et al. 1993, Estoup et al. 1995), and this allowed researches to apply these loci to study the population genetics of native and managed honey bee populations (Garnery et al. 1998, Franck et al. 2000, De la Rua et al. 2001, Franck et al. 2001). The microsatellites were particularly helpful in sorting out which honey bee populations were pure and which were admixed.

The publication of the honey bee genome allowed bee researchers to directly query single nucleotide polymorphisms (Whitfield et al. 2006, Zayed and Whitfield 2008) and to carry out molecular evolution analyses on interesting genes
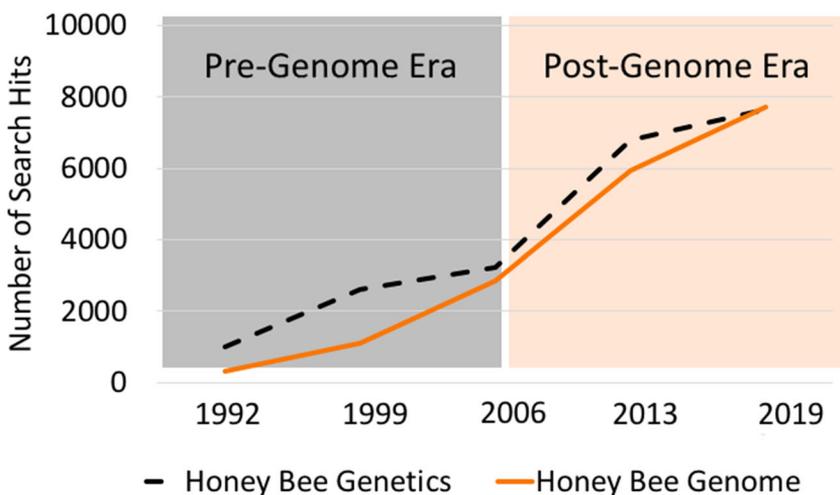

Figure 2. Research into both honey bee genetics and genomics has grown in the past three decades. The honey bee genome's publication in 2006 has fuelled a spike in these studies (as determined by a number of search hits on Google Scholar within a given publication year). In fact, studies that incorporate genomics have recently overtaken studies that mention only 'genetics', demonstrating that the honey bee genome has become an integral part of honey bee research.

(Hasselmann et al. 2008, Hasselmann et al. 2010, Kent et al. 2011, Harpur et al. 2013, Harpur and Zayed 2013). The earliest of these analyses supported an 'out of Africa' evolutionary origin for *Apis mellifera* (Whitfield et al. 2006); a surprising finding given hypothesized Asian or European origins. The 'out of Africa' story involved two separate expansions of honey bees out of Africa leading to the M lineage in western/northern Europe and the C and O lineages in Eastern Europe and the O lineage of Asia. However, a reanalysis of this dataset (Han et al. 2012) indicated that the out-of-Africa conclusion is very sensitive to inclusion of some admixed bees in Northern Africa; removal of these bees prevented the researchers from determining the ancestral *A. mellifera* population with certainty.

The single nucleotide polymorphism- (SNP) based studies were quickly followed by population genomic studies that sequenced the individual genomes of different honey bee species revealing a great degree of genetic diversity segregating within individuals and populations of native and managed honey bees (Harpur et al. 2014, Molodtsova et al. 2014, Wallberg et al. 2014, Chen et al. 2016, Kadri et al. 2016, Wallberg et al. 2016, Wragg et al. 2016, Wallberg et al. 2017). These studies discovered that a native honey bee of the middle east, *A. m. yementica* (lineage Y), is actually genetically distinct from both the African A lineage and the Asian O lineage (Harpur et al. 2014, Cridland et al. 2017) and uncovered a new Asian subspecies that paradoxically belongs to the M lineage—a lineage that was previously considered to only inhabit Europe (Chen et al. 2016). Despite the influx of all of these genomic resources and substantial progress in understanding *A. mellifera* diversity and evolutionary history, the evolutionary origin of *A. mellifera* still remains uncertain, with both 'out of Africa' and 'out of Asia' scenarios possible (Cridland et al. 2017). While we certainly have the tools to rapidly sequence and analyse bee genomes, the limiting factor in understanding the evolutionary origins of *A. mellifera* is actually the availability of a 'pure' honey bee samples especially from Africa and Asia (Dogantzis and Zayed 2019). This example illustrates how genomics cannot operate in a vacuum of biological understanding, i.e. genomic expertise can never replace the importance of well-trained biologists that are able to identify and sample the large number of native honey bee subspecies in the field.

2) Thelytoky: the strange case of worker reproduction

While honey bee workers still retain the ability to lay unfertilized eggs that develop into drones, workers of *A. m. capensis* in the cape of South Africa are able to lay unfertilized eggs that develop into females (Anderson 1963)—a phenomenon called thelytokous parthenogenesis. The genetics that underlie this ability of Cape honey bee workers to act as queens have remained a mystery for nearly 60 years. Here, the honey bee genome substantially energized research on the genetics of thelytoky, and we appear to have finally solved this mystery after several missed steps. Using genetic crosses and microsatellite loci anchored by the newly published bee genome, Lattorff and colleagues (Lattorff et al. 2007) first identified a single locus on chromosome 13 as the candidate loci for thelytoky. The authors hypothesized that one transcription factor, *Grainyhead*, within this genomic region was the candidate gene influencing thelytoky. Harnessing the power of the bee genome to carry out functional genomics and gene knockdowns, another transcription factor *Gemini* was suggested as a candidate gene for thelytoky (Jarosch et al. 2011) as alternative splicing of this gene, along with the presence or absence of a 9 bp insertion/deletion polymorphism, was associated worker clonal reproduction. However, genomics-enabled population studies overturned the 9 bp indel in *Gemini* as it was also found in other African honey bees that do not exhibit thelytoky (Chapman et al. 2015, Wallberg et al. 2016, Aumer et al. 2017).

Charting the genotype-phenotype map in honey bees has been substantially enhanced by the availability of a genome sequence because it allows powerful techniques such as the genome-wide association mapping of complex traits (Dogantzis and Zayed 2019, Kent et al. 2019; Grozinger and Zayed 2020). These approaches were recently applied to study the genetics of

thelytoky yielding a different candidate gene on chromosome 1, GB46427 (Aumer et al. 2017, Aumer et al. 2019); this uncharacterized gene contains a non-synonymous single nucleotide mutation that was strongly associated with worker clonal reduction in a single Cape honey bee colony. Yet again, population genomic data refuted this hypothesis, as the thelytoky-associated mutation in GB46427 in Aumer et al.'s study was also common in non-thelytokous populations (Christmas et al. 2019). In parallel, population genomic comparisons of Cape honey bees with *A. mellifera scutellata* of South Africa, which does not exhibit thelytoky, identified a very small number of mutations that show fixed difference or nearly fixed differences between *capensis* and *scutellata*—these loci may be responsible for thelytoky or other phenotypic differences between these two subspecies (Wallberg et al. 2016). One of these mutations is found on a chromosome 11 gene (GB45239) that likely plays a role in chromosome alignment during meiosis (Wallberg et al. 2016). The potentially last chapter of the thelytokous saga strongly supports GB45239 as the causal gene for clonal reproduction in the Cape honey bee. Using full genome sequencing of nearly 50 females from a backcross between *capensis* and *scutellata*, Yagound et al. (2020) found several markers in GB45239 that showed consistent co-segregation with thelytoky. The author consulted population genomic datasets and found that the thelytoky-associated alleles were present in all *A. m. capensis* genomes sequenced to date but were absent from all other honey bees in African, Asia and Europe. Finally, expression patterns of GB45239 were consistent with its putative function; it is expressed in ovaries and downregulated in clonal workers.

The thelytoky saga clearly illustrates how genomics, especially when used in an integrative manner (i.e. combination of quantitative genomics, population genomics and transcriptomics) can be used to solve major questions in the field. There is a cautionary tale here as well; analyses based on small datasets or studies that neglect population and evolutionary contexts may lead to biased or misleading results. The last chapter in the thelytoky saga has yet to be written; researchers have yet to show that manipulating GB45239 leads to changes in reproduction in honey bees. Excitingly, recent studies have demonstrated the feasibility of germ-line transformation in honey bees (e.g. Schulte et al. 2014, McAfee et al. 2019); thus, the reality of making genetically modified honey bee strains is not far away.

3) Hygienic behaviour, a key trait for breeding and selection

Like thelytoky, the genetic basis underlying hygienic behaviour has mystified bee biologists for about 60 years. Hygienic behaviour is an important social immune trait in honey bees. Worker bees are able to recognize and remove dead or dying brood from the colony, which leads to lower provenance of parasitic *Varroa* mites and diseases such as American foulbrood and chalkbrood (Spivak 1996; Spivak and Gilliam, 1998a; Spivak and Gilliam, 1998b). In 1965, Rothenbuhler carried out a set of genetic crosses between hygienic and non-hygienic colonies and, based on the number of phenotypic clusters observed in workers from these crosses, concluded that hygienic behaviour was under the control of two loci—one responsible for uncapping cells of dead pupae and another responsible for removal of the dead pupae (Rothenbuhler 1964). However, this conclusion was very sensitive to Rothenbuhler's discrete classification of colonies as hygienic or non-hygienic, when in fact, hygienic behaviour varies on a continuous scale (Moritz 1988). Indeed, Mortiz's reanalysis of Rothenbuhler's data suggested at least 3 loci, and likely many more, influencing this complex trait (Moritz 1988). Using quantitative trait loci mapping, Lapidge et al. (2002) highlighted 7 suggested loci that influence general hygienic (i.e. assayed by freeze killing brood and scoring the percentage of dead brood removed a day later); although without a reference genome to compare to, it was impossible to learn more about the actual genes underlying this trait without carrying out additional crosses and positional cloning—a very tedious process.

The publication of the honey bee genome invigorated research on the genetics and molecular biology of hygienic behaviour. Oxley et al. (2010) carried out a QTL study of general hygienic behaviour using 437 microsatellites. By anchoring

the microsatellites on the newly published genome, the researchers were able to determine the physical positions of these loci on chromosomes 2, 5, 9, 10 and 16 and highlight interesting candidate genes among the hundreds of genes residing among the mapped loci. Shortly after, Behrens et al. (2011) carried out a QTL study using approximately 200 microsatellites and mapped three candidate loci on chromosomes 4, 7 and 9 that influence hygienic behaviour as it specifically pertains to the removal of brood infected with *Varroa* mites—often referred to as *Varroa*-sensitive hygiene (VSH). Similar to Oxley et al.'s study, the implicated loci were several megabases large and contained hundreds of genes. A year later, Tsuruda et al. (2012) carried out another QTL study—this time employing 1340 newly developed SNP markers and discovered 2 loci on chromosome 9 and 1 that influence *Varroa*-sensitive hygiene. These regions were approximately 3 million bases in size and contained approximately 100 genes. While the QTL studies were a large step forward in terms of understanding the genetic architecture of hygienic behaviour, the QTL approach lacks the required resolution to pinpoint the specific genes and mutations responsible for causing differences in behaviour between individuals and colonies (Dogantzis and Zayed 2019, Kent et al. 2019; Grozinger and Zayed 2020). Moreover, the typical QTL designs used in honey bee experiments include a very small number of bees in the genetic crosses; this both serves to underestimate the actual number of loci influencing a trait and overestimates the effects of the discovered loci (Xu 2003)—a phenomenon called the Beavis effect.

Leveraging both the availability of a reference genome and cost-effective short-read sequencing technology, it is now possible to use genome-wide association studies (GWAS) to link genetic mutations with differences in traits between individuals, colonies and populations (reviewed by Grozinger and Zayed 2020). Spötter and colleagues (2012, 2016) used a 44,000 SNP assay to genotype 122 workers that exhibited the strongest abilities to detect and uncap *Varroa*-parasitied brood and compared them to 122 control workers that did not exhibit hygienic behaviour. The authors discovered 6 SNP markers on

chromosome 2, 3 (2×), 5, and 7 that were associated with the trait. While these SNPs were not found in genes, the authors used the bee's published genome and identified 4 candidate genes that were nearby (i.e. 65 to790 kB away). More recently, Harpur et al. (2019) sequenced the genomes of 125 drones from either a control population with average hygienic behaviour or 2 artificially selected populations with high levels of hygienic behaviour. Here, the authors used two types of analyses to link genetics with phenotype. They first searched for genomic regions that show signs of artificial selection in the hygienic populations and then asked if these regions predict hygienic behaviour in the control population. This integrative analysis highlighted approximately 73 genes associated with hygienic behaviour (Harpur et al. 2019). While this number of loci is certainly higher relative to other studies, it is in line with the number of loci that regulates several behaviours in *Drosophila*. These candidates have not been functionally validated as of yet but several of the loci identified by Harpur et al. were within known QTL regions (Harpur et al. 2019). We expect that the application of GWAS will further enhance our understanding of the genetics of some of the most complex and charismatic honey bee traits. These efforts can be further enhanced with a development of tools to disrupt gene function (e.g. RNAi or CRISPR-Cas9 knockouts) to directly confirm causal relationships between genes and phenotypes. Nonetheless, the extensive resources provided by the honey bee genome have catapulted forward our understanding of the genetic basis of hygienic behaviour, which can facilitate marker-assisted selection and bee breeding programs.

4) The iconic dance language of the bees

The famous dance language of the honey bee, decoded by Karl von Frisch in the mid-twentieth century, resulting in his receipt of the 1972 Nobel Prize in Medicine or Physiology, is arguably *the single most* quintessential honey bee behaviour (Couvillon 2012). This form of symbolic communication, which will be so familiar to readers of *Apidologie*, is expressed by foraging workers upon returning back to the hive from a profitable foraging or nesting site. It involves an intricate

series of movements performed in a figure 8-like pattern, punctuated by a frenetic 'waggle run' in the middle. Subtle variations in the tempo, duration and angle of the dance communicate specific information about the quality, distance and direction of profitable sites to recruit bees, the receivers of the dance information (von Frisch 1967). Recruits then leave the hive in search of these sites, allowing honey bees to code, decode, and share with each other information about specific locations in their environment. It remains one of the most complex and spectacular examples of animal communication outside of human language.

The dance language has fascinated generations of bee biologists, and thus, there has been a cottage industry of studies on different aspects of the dance (Dyer 2002), elucidating even finer and more fascinating aspects of the dance. These include the discovery of sound, vibrational (Michelson et al. 1986) and pheromonal components (Thom et al. 2007) to the dance, investigations into how bees measure distance and time (Esch et al. 2001), and the existence of species differences and regional dialects (Rinderer and Beaman 1995). Thus, this is perhaps one of the single best-studied forms of behaviour in all of ethology. What can genomics possibly have to add?

Pre-genomic studies of the honey bee dance based on traditional crossing-phenotyping between subspecies of bees with different dance dialects suggested a genetic basis for one aspect of the dance, the transition distance at which bees switch from round dancing to waggle dancing (Johnson et al. 2002, Rinderer and Beaman 1995). However, there have been no subsequent genome-enabled studies that have attempted to identify specific genes or mechanisms related to neither this behaviour nor dialect differences. This is an area that remains ripe for future study, especially given the fact that we now have genomes from multiple different Apis subspecies (Fuller et al. 2015) and species (Park et al. 2015). Such studies have the potential to elucidate heredity influences and evolutionary trajectories of the dance.

Transcriptomic studies have utilized honey bee genome-derived gene sequence information to facilitate the study of global gene expression patterns related to dancing honey bees, including across honey bee species. Sen Sarma et al. (2009) examined brain region-specific gene expression in dancing Apis mellifera, florea and dorsata, which differ greatly in aspects of their dance behaviour. This study revealed commonalities in gene expression patterns across dancing workers of the three species in the mushroom bodies (brain region associated with learning and memory), including some related to metabolic and energy production processes. There were also numerous differences in the gene expression between Apis species, suggesting that there may be a core set of genes across species expressed in 'the dancing brain', whereas others were unique to each species, perhaps contributing to species differences in the dance. The mushroom body showed the most distinct expression profile compared with other brain regions, suggesting that this brain region, involved in learning, memory and sensory integration, is important in the regulation of the dance. However, this study was largely correlative as it did not compare dancers and non-dancers and thus does not allow for clear identification of genes that actually regulate dance, although it does give some promising leads for candidate genes for this behaviour. In another transcriptomic study focusing on Apis mellifera, a subcomponent of the dance, distance estimation, was studied in trained bees flying through tunnels with different distance-simulated environments (Sen Sarma et al., 2010). The study revealed differences in the gene expression in the optic lobes and mushroom bodies in bees flying at different perceived distances. This study suggested some candidate genes and pathways (related to synaptic remodeling, transcription factors and protein metabolism) related to the distance estimation aspect of the dance. Another study reports differences in the expression of microRNAs between dancing and non-dancing foragers of Apis mellifera and also uncovered several novel microRNAs in dancers and foragers, and the target genes of these microRNAs were suggested to be related to kinase activity, neural function, and energy production (Li et al. 2012). Overall, these studies suggest that dynamic gene expression changes in the brain, especially in the mushroom body, may be involved in the production of the dance language.

However, these studies have been all been correlative, and thus, our understanding of how, or even *if* genes, regulate dancing is still very much in in its infancy.

Thus far, the neurogenomic and heritable genetic bases of the dance language remain rather elusive (Barron and Plath 2017). This may be an example from honey bee biology in which genomics has not yet not added much to our understanding or appreciation of this particular aspect of honey bee biology. Genomics is most powerful when it can be applied in a high-throughput manner, but unfortunately, 'measuring' the waggle dance requires careful and time-consuming observation by well-trained ethologists. Here, the limiting factor is 'phenotyping', not genotyping. The dance language remains a special challenge for honey bee genomicists; nonetheless; honey bee behavioural genomics has had its share of successes. Numerous honey bee behaviours that can be more readily phenotyped have been studied in detail on a genomic level including pollen foraging (Page Jr et al., 2012, Rueppell 2014), age polyethism (Zayed and Robinson 2012), learning and memory (Müller, 2012) and defensive behaviour (Avalos et al. 2020, Harpur et al. 2020). These bodies of work have resulted in profound leaps forward in our understanding of the molecular underpinnings of other aspects of honey bee behaviour.

5) Beyond the honey bee genome: epigenetics and DNA methylation

Going beyond the sequence of the genome, there has been a growing appreciation of the importance of epigenetics in social and health-related phenotypes, that is, alterations of chromatin structure via chemical modifications to DNA or histones (Szyf and Meaney, 2008). Epigenetics refers to chemical modifications to DNA that do not change the DNA sequence but can affect gene expression and thus phenotypic traits. Such modifications can be induced by the environment, including in response to stress, toxins and the social environment (Crews 2008). Most strikingly, epigenetic modifications can be stable over many cell divisions, so they have the potential to be passed from parents

to offspring resulting in epigenetic inheritance, although they can also be reversible (reviewed in Crews 2008). Thus, there has been increasing interest in the role of epigenetics in honey bees, in multiple contexts: as related to regulation of gene expression, caste formation, behavioural plasticity and imprinting, in which paternal and maternal genes may be differentially expressed (Queller 2003, Haig 2000).

DNA methylation—a major epigenetic mechanism—was thought to be absent in insects because the DNA methylation machinery was absent from the two fly genomes available at the time. As such, a major surprise from the honey bee genome was the presence of genes for DNA methylation and a functional DNA methylation system (Wang et al. 2006). Two decades and many genomes later, it turns out that the flies were exceptions, and most insects are now known to possess DNA methylation systems; although interestingly, DNA methylation has been lost multiple times within the insects (Glastad et al. 2019). This exciting discovery in honey bees provided fuel for a growing interest and appreciation of the role of epigenetics in honey bee social life. Subsequent studies addressed the possible role of DNA methylation in honey bee caste differentiation. DNA methylation patterns across the entire genome differ between queens and workers, with over 550 differentially methylated genes between the brains of adult members of these two castes (Lyko et al. 2010). Furthermore, the knockdown of the DNA methyl transferase gene *Dnmt3* during larval development causes a bias toward adult bees emerging with the queen phenotype (Kucharski et al. 2008). Inhibiting DNA methylation during larval development also led to caste differential gene expression and changes in alternative splicing (Li-Byarlay et al. 2013). Further studies suggest that DNA methylation may affect behavioural maturation in honey bee workers, as there are differences in global methylation patterns between nurses and foragers (Herb et al. 2012), and some of these changes may be dynamic and responsive to the social environment (Lockett et al. 2011).

DNA methylation may also play a role in genomic imprinting, that is, a situation in which the expression of alternate alleles of a gene in a

diploid organism is biased toward one parent's allele vs the other. Because mothers and fathers may have competing interests, this sets up the potential for intragenomic conflict (Haig 2000). One general prediction is that patrigenes should promote phenotypes that overuse maternal energy reserves, while matrigenes should 'fight back' towards conserving maternal resources. Queller (2003) extended this for haplodiploid social insects predicting various parent-of-origin effects on social traits. Cross-breeding studies between two strains of honey bees found evidence of a strong paternal effect on worker fertility (Oldroyd et al. 2014). On the molecular level, many genes in the honey bee genome show parent-of-origin effects on gene expression (Kocher et al. 2015). A subsequent study found in reproductive tissue expression tends to be patrigene biased, and this bias is stronger in workers which activated their ovaries in response to queenless conditions (Galbraith et al. 2016).

Despite these advances, honey bee epigenetics holds a dark secret—that is, we still lack a clear understanding of what DNA methylation *actually does* on a molecular level, in honey bee and in all insects (Glastad et al. 2019, Li-Byarlay 2016). Insect DNA methylation departs from its canonical patterns in other taxa in multiple ways: it does not appear to be associated with gene silencing as in other animals; it is found at much lower levels overall across the genome; it is found mainly in gene bodies; and there have been multiple losses in different groups of insects (Glastad et al. 2019), including in Hymenoptera (Standage et al. 2016). Recent studies have also cast doubt upon the dynamism of DNA methylation in social insects (Libbrecht et al. 2016). Using a more conservative approach to methylation calling from 'bisulfite resequenced' genome data, these authors cast doubt on previous results and go as far as suggesting that there is no caste-related variation in DNA methylation in honey bees.

Perhaps part of the challenge is that by focusing so much on DNA methylation, other mechanisms have been ignored; have we been barking up the wrong epigenetic tree? Other forms of epigenetic modifications have also shown promising connections to behavioural plasticity and caste differences in various social insect species, including histone

acetylation (Simola et al. 2016) and microRNAs. Beyond DNA methylation, there have begun to be additional studies on other types of epigenetic mechanisms in honey bees. The honey bee genome shows extensive evidence of various forms of histone acetylation; however, no differences were found between queen ovaries and young (pre-caste differentiated) larvae (Dickman et al. 2013). The treatment of adult worker bees with histone deacetylase inhibitors resulted in alterations in aversive memory formation (Lockett et al. 2014), suggesting a potential role for histone acetylation in learning and memory. However, there have been few studies in general of histone acetylation in honey bees; thus, we know relatively little about its importance in honey bee epigenetics. Another epigenetic regulator, non-coding microRNAs (miRNAs), can also alter gene expression; they are thought to function in gene silencing by inhibiting protein production via binding to complementary mRNA strands and blocking translation. Numerous (over 300) miRNAs have been found in the genomes of honey bees (Weaver et al. 2007, Ashby et al. 2016), and differential expressions of numerous miRNAs were found to be differentially expressed between queens, workers and/or drones (Ashby et al. 2016, Guo et al. 2016). In addition, some miRNAs are hypothesized to be active components of royal jelly (Guo et al. 2013, Guo et al. 2016). There are also differences in miRNA expression between honey bee worker behavioural castes (Behura and Whitfield, 2010, Liu et al. 2012) suggesting miRNA profiles change over development. With respect to the epigenetic realm, this is one of the most active fields of honey bee post-genomic research. However, at this point, there are more questions than answers with respect to the role of epigenetics in honey bee social life.

6) Holo-bee-onts: the honey bee hive as a genomic community

When we sequence the genome of a honey bee, in fact, we sequence the genome of an entire ecological community of interacting species. The honey bee can thus be viewed as a 'holobiont' (Schwartz et al. 2015) consisting of a diverse community of bacteria, viruses, mites and fungi.

Furthermore, the hive itself is a superorganism that is host to additional species of insects (e.g. wax moths, hive beetles), yeasts and other microbes. Upon the initial sequencing of the honey bee genome, researchers were focused on the organism itself, trying to assemble the genome of the isolated species and remove DNA 'contamination' from other species that may have been sequenced along with the honey bee's own DNA. However, biologists have been gaining a deeper appreciation of the highly integrated way in which genomes of interacting species are inextricably interconnected. For example, there are cases of mutualisms between species that are so tight (e.g. aphids and their bacterial symbionts *Buchnera*) that the genomes of the host and the symbiont are said to 'collaborate', whereby genes in one species' genome can functionally replace or complement steps in the biochemical pathways controlled by genes in the cooperating species' genome (Wilson and Duncan 2015). This is a beautiful illustration of the idea that a species' genome and evolutionary history are incomplete unless viewed in conjunction with the other species with which it commonly interacts. Thus, genomicists must sequence past the individual organism or even superorganism and appreciate the genomic context of ecological units consisting of communities of genes of multiple interacting species. The honey bee is no exception. Given this, what do we now know about the collective genome of the honey bee 'holobiont', or the 'holo-bee-ont', and what insights does such a view give us about honey bee biology?

One of the most important interacting species for *Apis mellifera* and other *Apis* species are mites in the genus *Varroa*. The genome of the common, worldwide apiary pest *Varroa destructor* has been sequenced (Cornman et al. 2010). This information is of great interest in developing genome-directed control strategies that may be *Varroa* mite-specific, such as RNA interference, or excitingly even accomplish RNAi by engineered honey bee gut bacteria (Hunter et al. 2010, Garbian et al. 2012, Leonard et al. 2020). In addition, *Varroa* mites have been associated with vectoring numerous honey bee viruses, many of which are pathogens of dire concern for bee

health (Wilfert et al. 2016). Genomic studies have also provided important insights into the virus community associated with honey bees (Brutscher et al. 2016). In addition to having complete genome sequences (Fung et al. 2018) for many common honey bee viruses (such as deformed wing virus, Israeli acute paralysis virus, sacbrood virus and acute bee paralysis virus (Govan et al. 2000)), high throughput genomic techniques have provided 'metagenomic' surveys of honey bee viruses (Cox-Foster et al. 2007). Having the full genome sequence of the honey bee available facilitated these studies, allowing for 'subtraction' of the host's genome information from metagenomic datasets, so that non-bee sequences could be easily identified and further studied. This approach has led to the discovery of many novel viruses (Galbraith et al. 2018, Remnant et al. 2017) that were not previously known in honey bees. Studying honey bee viruses has also provided a deeper understanding of the *Varroa* virus complex and how these pathogens may interact to subvert host immunity (Brutscher et al. 2015) and behavioural defenses (Geffre et al. 2020) since many viruses are likely vectored by *Varroa* mites.

Another burgeoning area of study is related to the honey bee bacterial microbiome (Moran 2015). The individual genomes of well-known honey bee pathogenic bacteria have been sequenced, such as *Paenibacillus larvae* (the cause of American foulbrood, Chan et al. 2011). Beyond pathogens, the gut-associated microbiome of adult worker honey bees has been a major focus, and metagenomic sequencing of the honey bee gut has added greatly to our understanding of this community. The honey bee gut bacterial community appears to consist of relatively few species compared to other insects, but despite being host to few species, there appears to be high strain and functional diversity (Engel et al. 2012). The honey bee gut microbiome consists predominantly of 9 bacterial species, representing 6–8 phylotypes, and each appears to occupy distinct niches within the bee's gut. Some of these appear to be novel, bee-specific bacteria, including an acetic acid bacterium found in the

midgut, *Bombella apis* (Yun et al. 2017), which may aid in the breakdown of pollen walls (Bonilla-Rosso and Engel, 2018). The main species in the ileum are *Snodgrassella alvi*, *Gilliamella apicola* and *Frischella perrara*, and in the rectum, *Lactobacillus* (firm 5 and 4) and *Bifidobacterium.* Major functions of these bacterial groups include carbohydrate metabolism, saccharide breakdown, fermentation and biofilm formation, and they may thus play an important role in proper honey bee nutrition (Lee et al. 2015, Kwong and Moran, 2016). Although the taxonomy of honey bee-associated bacteria is still an area of active study, there is some debate about whether the gut is host to more bacterial species than previously thought (Sabree et al. 2012, Mattila et al. 2012). Perturbations in the honey bee's gut microbiome are associated with disease states, as well as invasions by pathogenic bacteria (Anderson and Ricigliano 2017, Kwong and Moran, 2016). Targeted sequencing of bacterial genomes (Smith et al. 2019) and metagenomic screens are providing useful insights into the bacterial community and its role in health and disease (Raymann and Moran 2018).

Beyond the worker bee gut, the honey bee hive itself harbors many additional microbial communities in the wax comb, propolis envelope and food stores, especially pollen stores ('bee bread'). The bee bread has been the focus of much microbial work, the classic view establishing the presence of yeasts and lactic acid bacteria such as *Bacillus* (Gilliam 1979), thought to aid in the breakdown of stored pollen, aiding in its digestibility. However, genomic and metagenomic studies of bee bread have found a diverse array of microbes, not only including core lactic acid bacteria (Vásquez and Olofsson 2009), but also including non-core bacteria such as *Fructobacillus* and other *Lactobacillaceae* that may promote the growth of other honey bee-specific bacterial species (Rokop et al. 2015). It has also been suggested that these processes aid in preserving and storing pollen, rather than breaking it down as previously thought (Anderson et al. 2014). Interestingly, propolis (bee-collected plant resins) may also play an important role in maintaining healthy microbiomes within the hive and also in the bee gut (Simone-Finstrom et al. 2017), although the mechanism of action is not yet understood.

In conclusion, the genomic community of honey bees has revealed numerous new surprises with respect to the intricacy of interactions with microbes in bee life, health and disease. This is an area in which our understanding of bee biology has skyrocketed with integration of genomic techniques, albeit by applying them beyond 'the honey bee genome'.

3. CONCLUSION

It is hard to think of a field of honey bee biology that has not been impacted to some degree by the publication of the honey bee genome. Honey bee research is certainly a growing enterprise, although it is clear that the honey bee genome has benefited some fields over others. We quantified the pace of research in a large number of disciplines and found a substantive across-the-board increase in the productivity of bee research over time (Figure 3). Indeed, some of the 'hottest' bee research fields, such as epigenetics and metagenomics, were essentially non-existent until 2006. Other fields, such as bee quantitative and population genetics and sociobiology, clearly benefited from the publication of the bee genome and the array of tools it enabled; we can now rapidly measure gene expression and quantify genetic diversity at a previously unfathomed scale. Some fields have shown less growth, e.g. nutrition, and we suggest that these fields stand to benefit from even further integration of genome information (i.e. 'nutrigenomics' (e.g. Alaux et al. 2011)). While we are still far from understanding the bee's remarkable social biology in 'molecular terms', we have collectively taken large strides toward that goal and have already achieved important milestones as mentioned above. In hindsight, we think that the great fanfare and expectations arising from the honey bee genome itself were somewhat misplaced; the reference genome in itself is static and—dare we say—somewhat boring. We are ultimately interested in differences; differences in behaviour of individual workers over time, differences between castes,
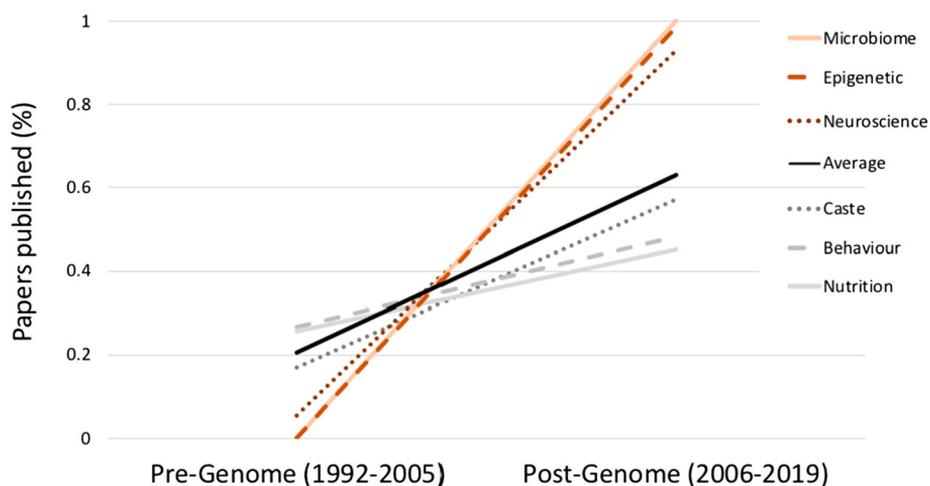
Figure 3. Growth of honey bee research areas in the post-genome era. We counted the number of papers in the Web of Knowledge database published between 1970 and 2019 with the keywords *Apis mellifera* and 'behaviour', 'breeding', 'caste', 'demography', 'division of labour', 'epigenetic', 'foraging', 'genetic', 'learning', 'memory', 'metabolism', 'microbiome', 'neuroscience', 'nutrition', 'parasite', 'pathogen', 'physiology', 'population' and 'toxicology.' We then compared how each term grew in number of publication from a period prior to (1992 to 2005) and after (2006 to 2019) the publication of the honey bee genome. All terms experienced a substantive increase over time; on average, there were approximately 3 times as many papers published in the 13 years *post-genome* relative to the 13 years *pre-genome.* Some terms/fields stood out among the general trends. 'Microbiome', 'epigenetic' and 'neuroscience' were never (former two) or rarely published on in the pre-genomic era relative to the post-genomic era. On the other hand, the following terms experienced the slowest growth in our analysis: 'behaviour', 'caste' and 'nutrition'.

differences in the behaviour and health between colonies within a yard and between subspecies inhabiting different environments, differences between honey bee species that show remarkable differences in biology and differences between the social honey bees and their solitary ancestors. These differences that we find so fascinating are all essentially orchestrated by differences in genetics or differences in the regulation of gene activity—something that a static reference genome does not directly provide. Thus, the genome has been a *single component* of a multi-omics toolbox; in reality, honey bee genomics is a systems-level enterprise that must include transcriptomics, proteomics, metabolomics and epigenomics. Many of the major research milestones discussed herein were made possible by the ability to probe the dynamic nature of genomes over multiple biological levels and time scales. Population genomics—a topic that has received

many recent reviews (e.g. Dogantzis and Zayed, 2019, Hasselmann et al. 2015)—has been particularly useful in uncovering potentially functional genetic mutations that affect phenotypic traits and colony fitness.

Finally, we would like to stress that bee genomics cannot operate in a vacuum. Genomics is a means to an end, but not the end itself. The field of apidology still needs talented ethologists, physiologists, geneticists, ecologists and evolutionary biologists. Indeed, we strongly believe that such integration, with genomics as an important component, is the key to new discoveries going forward.

AUTHOR CONTRIBUTIONS

FUNDING

OPEN ACCESS

Le génome de l'abeille: A quoi a-t-il servi?

Apis mellifera / génomique / transcriptomique / génétique / séquençage.

Das Genom der Honigbiene: Wozu wurde es verwendet?

Apis mellifera / Genomik / Transkriptomik / Genetik / Sequenzierung.

REFERENCES

Alaux C, Dantec C, Parrinello H, Le Conte Y (2011) Nutrigenomics in honey bees: digital gene expression analysis of pollen's nutritive effects on healthy and varroa-parasitized bees. BMC Genomics 12:496.

Anderson RH (1963) The laying worker in the Cape honeybee *Apis mellifera capensis*. J Apic Res 2(2):85-92.

Anderson KE, Ricigliano VA (2017) Honey bee gut dysbiosis: a novel context of disease ecology. Curr Opin Insect Sci 22:125-132.

Anderson KE, Carroll MJ, Sheehan TIM, Mott BM, Maes P, Corby-Harris V (2014) Hive-stored pollen of honey bees: many lines of evidence are consistent with pollen preservation, not nutrient conversion. Mol Ecol 23:5904-5917.

Arias MC, Sheppard WS (1996) Molecular phylogenetics of honey bee subspecies (*Apis mellifera* L.) inferred from mitochondrial DNA sequence. Mol Phylogenet Evol 5(3):557-566.

Ashby R, Forêt S, Searle I, Maleszka R (2016) MicroRNAs in honey bee caste determination. Sci Re 6:18794.

Aumer D, Allsopp MH, Lattorff HMG, Moritz RFA, Jarosch-Perlow A (2017) Thelytoky in Cape honeybees (*Apis mellifera capensis*) is controlled by a single recessive locus. Apidologie 48(3): 401-410.

Aumer D, Stolle E, Allsopp M, Mumoki F, Pirk, CWW et al. (2019) A single SNP turns a social honey bee (*Apis mellifera*) worker into a selfish parasite. Mol Biol Evol 36(3): 516-526.

Avalos A., Fang M, Pan H, Lluch AR, Lipka AE, et al. (2020) Genomic regions influencing aggressive behavior in honey bees are defined by colony allele frequencies. Proc Natl Acad Sci USA 117:17135-17141.

Barron AB, Plath, JA (2017) The evolution of honey bee dance communication: a mechanistic perspective. J Exper Biol 220(23):4339-4346.

Behrens D, Huang Q, Gessner C, Rosenkranz P, Frey E, et al. (2011) Three QTL in the honey bee *Apis mellifera* L. suppress reproduction of the parasitic mite *Varroa destructor*. Ecol Evol 1(4): 451-458.

Behura SK, Whitfield CW (2010) Correlated expression patterns of microRNA genes with age-dependent behavioural changes in honeybee. Insect Molec Biol 19(4):431-439.

Bell AM, Robinson, GE (2011) Behavior and the dynamic genome. Science 332(6034):1161-1162.

Bonilla-Rosso G, Engel P (2018) Functional roles and metabolic niches in the honey bee gut microbiota. Curr Opin Microbiol 43:69-76.

Brutscher LM, Daughenbaugh KF, Flenniken ML (2015) Antiviral defense mechanisms in honey bees. Curr Opin Insect Sci 10:71-82.

Brutscher LM, McMenamin AJ, Flenniken ML (2016) The buzz about honey bee viruses. PLoS Pathogens 12(8):e1005757.

Chan QW, Cornman RS, Birol I, Liao NY, Chan SK, Docking TR, Jackman SD, Taylor, GA, Jones SJ, de Graaf DC, Evans JD (2011) Updated genome assembly and annotation of *Paenibacillus larvae*, the agent of American foulbrood disease of honey bees. BMC Genomics 12(1):450.

Chapman NC, Beekman M, Allsopp MH, Rinderer TE, Lim J et al. (2015) Inheritance of thelytoky in the honey bee *Apis mellifera capensis*. Heredity 114: 584-592.

Chen C, Liu Z, Pan Q, Chen X, Wang H, et al. (2016) Genomic analyses reveal demographic history and temperate adaptation of the newly discovered honey bee subspecies *Apis mellifera sinisxinyuan n. ssp*. Mol Biol Evol msw017.

Christmas MJ, Smith NMA, Oldroyd BP, Webster MT (2019) Social parasitism in the honeybee (*Apis mellifera*) Is not controlled by a single SNP. Mol Biol Evol 36(8):1764-1767.

Cornman RS, Schatz MC, Johnston JS, Chen YP, Pettis J, Hunt G, Bourgeois L, Elsik C, Anderson D, Grozinger CM, Evans JD (2010) Genomic survey of the ectoparasitic mite *Varroa destructor*, a major pest of the honey bee *Apis mellifera*. BMC Genomics 11(1): 602.

Couvillon MJ (2012) The dance legacy of Karl von Frisch. Insectes Sociaux 59(3):297-306.

Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, Moran NA, Quan PL, Briese T, Hornig M, Geiser DM, Martinson V (2007) A metagenomic survey of microbes in honey bee colony collapse disorder. Science 318(5848):283-287.

Crews D (2008) Epigenetics and its implications for behavioral neuroendocrinology. Front Neuroendocrinol 29(3):344-357.

Cridland JM, Tsutsui ND, Ramirez SR (2017) The complex demographic history and evolutionary origin of the western honey bee, *Apis mellifera*. Genome Biol Evol 9(2):457-472.

Daiger SP (2005) Was the human genome project worth the effort?. Science 308(5720):362-364.

De la Rua P, Galian J, Serrano J, Moritz RFA (2001) Genetic structure and distinctness of *Apis mellifera* L. populations from the Canary Islands. Mol Ecol 10(7):1733-1742.

Dickman, MJ, Kucharski R, Maleszka R, Hurd PJ (2013) Extensive histone post-translational modification in honey bees. Insect Biochem Molec Biol 43(2):125-137.

Dogantzis KA, Zayed A (2019) Recent advances in population and quantitative genomics of honey bees. Curr Opin Insect Sci 31:93-98.

Dolezal AG, Toth AL (2014). Honey bee sociogenomics: a genome-scale perspective on bee social behavior and health. Apidologie 45(3):375-395.

Dyer FC (2002) The biology of the dance language. Ann Rev Entomol 47(1):917-949.

Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH Nadeau JH (2010) Missing heritability and strategies for finding the underlying causes of complex disease. Nature Rev Genet 11(6):446-450.

Esch HE, Zhang S, Srinivasan MV, Tautz J (2001) Honeybee dances communicate distances measured by optic flow. Nature 411(6837):581-583.

Estoup A, Solignac M, Harry M, Cornuet JM (1993) Characterization of (GT)N AND (CT)N microsatellites in two insect species - *Apis mellifera* and *Bombus terrestris*. Nucleic Acids Res 21(6):1427-1431.

Estoup A, Garnery L, Solignac M, Cornuet JM (1995) Microsatellite variation in honey bee (*Apis mellifera* L.) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. Genetics 140(2):679-695.

Fine R (2019) Lessons from the Human Genome Project. http://sitn.hms.harvard.edu/flash/2019/lessons-from-the-human-genome-project/

Franck P, Garnery L, Celebrano G, Solignac M, Cornuet JM (2000) Hybrid origins of honeybees from Italy (*Apis mellifera ligustica*) and Sicily (*A. m. sicula*). Mol Ecol 9(7):907-921.

Franck P, Garnery L, Loiseau, Oldroyd BP, Hepburn HR, et al. (2001) Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data. Heredity 86: 420-430.

Fuller ZL, Niño EL, Patch HM, Bedoya-Reina OC, Baumgarten T, Muli E, Mumoki F, Ratan A, McGraw J, Frazier M, Masiga D (2015) Genome-wide analysis of signatures of selection in populations of African honey bees (*Apis mellifera*) using new web-based tools. BMC Genomics 16(1):518.

Fung E, Hill K, Hogendoorn K, Glatz RV, Napier KR, Bellgard MI, Barrero RA (2018) De novo assembly of honey bee RNA viral genomes by tapping into the innate insect antiviral response pathway. J Invert Pathol 152:38-47.

Galbraith DA, Kocher SD, Glenn T, Albert I, Hunt GJ, Strassmann JE, Grozinger CM (2016) Testing the kinship theory of intragenomic conflict in honey bees (*Apis mellifera*). Proc Nat Acad Sci 113(4): 1020-1025.

Galbraith DA, Fuller ZL, Ray AM, Brockmann A, Frazier M, Gikungu MW, Martinez JFI, Kapheim KM, Kerby JT, Kocher SD, Losyev O (2018) Investigating the viral ecology of global bee communities with high-throughput metagenomics. Sci Reports 8(1): 1-11.

Garbian Y, Maori E, Kalev H, Shafir S, Sela I (2012) Bidirectional transfer of RNAi between honey bee and *Varroa destructor*: Varroa gene silencing reduces Varroa population. PLoS Pathogens 8(12).

Garnery L, Cornuet JM, Solignac M (1992) Evolutionary history of the honey bee *Apis mellifera* inferred from mitochondrial DNA analysis. Mol Ecol 1(3):145-154.

What has the honey bee genome been good for?

Garnery L, Solignac M, Celebrano G, Cornuet JM (1993) A simple test using restricted PCR amplified mitochondrial DNA to study the genetic structure of *Apis mellifera* L. Experientia 49(11):1016-1021.

Garnery L, Franck P, Baudry E, Vautrin D, Cornuet JM, et al. (1998) Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). II. Microsatellite loci. Genet Sel Evol 30:S49-S74.

Geffre AC, Gernat T, Harwood GP, Jones,BM, Gysi DM, Hamilton AR, Bonning BC, Toth AL, Robinson GE, Dolezal AG (2020) Honey bee virus causes context-dependent changes in host social behavior. Proc Nat Acad Sci 117(19):10406-10413.

Gilliam M (1979) Microbiology of pollen and bee bread: the genus *Bacillus*. Apidologie 10(3):269-274.

Glastad KM, Hunt BG, Goodisman MA (2019) Epigenetics in insects: genome regulation and the generation of phenotypic diversity. Ann Rev Entomol 64:185-203.

Govan VA, Leat N, Allsopp M, Davison S (2000) Analysis of the complete genome sequence of acute bee paralysis virus shows that it belongs to the novel group of insect-infecting RNA viruses. Virology 277(2): 457-463.

Grozinger CM, Robinson GE (2015) The power and promise of applying genomics to honey bee health. Curr Opin Insect Sci 10:124-132.

Grozinger CM, Zayed A (2020) Improving bee health through genomics. Nat Rev Genet **21**: 277–291.

Guo X, Su S, Skogerboe G, Dai S, Li W, Li S, Zhang S (2013) Recipe for a busy bee: microRNAs in honey bee caste determination. PloS One 8(12):e81661.

Haig D (2000) Genomic imprinting, sex-biased dispersal, and social behavior. Ann New York Acad Sci 907(1):149-163.

Han F, Wallberg A, Webster MT (2012) From where did the Western honeybee (*Apis mellifera*) originate? Ecol Evol 2(8):1949-1957.

Harpur BA, Zayed A (2013) Accelerated evolution of innate immunity proteins in social insects: adaptive evolution or relaxed constraint? Mol Biol Evol 30:1665-1674.

Harpur BA, Minaei S, Kent CF, Zayed A (2013) Admixture increases diversity in managed honey bees. Reply to De la Rúa et al. (2013). Mol Ecol 22: 3211-3215.

Harpur BA, Kent CF, Molodtsova D, Lebon JMD, Alqarni AS, et al. (2014) Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. Proc Natl Acad Sci USA 111(7):2614-2619.

Harpur BA, Guarna MM, Huxter E, Higo H, Moon KM, et al. (2019) Integrative genomics reveals the genetics and evolution of the honey bee's social immune system. Genome Biol Evol 11: 937-948.

Harpur BA, Kadri SM, Orsi, RO, Whitfield, CW, Zayed A (2020) Defense response in Brazilian honey bees (*Apis mellifera scutellata* × spp.) is underpinned by complex patterns of admixture. Genome Biol Evol 12:1367-1377.

Hasselmann M, Gempe T, Schiott M, Nunes-Silva CG, Otte M, et al. (2008) Evidence for the evolutionary nascence of a novel sex determination pathway in honeybees. Nature 454(7203): 519-522.

Hasselmann M, Lechner S, Schulte C, Beye M (2010) Origin of a function by tandem gene duplication limits the evolutionary capability of its sister copy. Proc Nat Acad Sci 107(30):13378-13383

Hasselmann M, Ferretti L, Zayed A (2015) Beyond fruit-flies: population genomics advances in non-*Drosophila* arthropods. *Brief Funct Genomics 14:424-431.*

Herb BR, Wolschin F, Hansen KD, Aryee MJ, Langmead B, Irizarry R, Feinberg AP (2012) Reversible switching between epigenetic states in honeybee behavioral subcastes. Nature Neuroscience 15(10):1371-1373.

Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. Nature 443(7114):931.

Hunter W, Ellis J, Hayes J, Westervelt D, Glick E, Williams M, Sela I, Maori E, Pettis J, Cox-Foster D, Paldi N (2010) Large-scale field application of RNAi technology reducing Israeli acute paralysis virus disease in honey bees (*Apis mellifera*, Hymenoptera: Apidae). PLoS Pathog 6(12): e1001160.

International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409:860-921.

Jarosch A, Stolle E, Crewe RM, Moritz RFA (2011) Alternative splicing of a single transcription factor drives selfish reproductive behavior in honeybee workers (*Apis mellifera*). Proc Natl Acad Sci USA 108(37):15282-15287.

Johnson RN, Oldroyd BP, Barron AB, Crozier RH (2002) Genetic control of the honey bee (*Apis mellifera*) dance language: segregating dance forms in a backcrossed colony. J Hered 93(3):170-173.

Kadri SM, Harpur BA, Orsi RO, Zayed A (2016) A variant reference data set for the Africanized honeybee, *Apis mellifera*. Sci Data 3:160097.

Kent CF, Issa A, Bunting AC, Zayed A (2011) Adaptive evolution of a key gene affecting queen and worker traits in the honey bee, *Apis mellifera*. Mol Ecol 20:5226-5235.

Kent CF, Tiwari T, Rose S, Patel H, Conflitti IM, et al. (2019) Studying the genetics of behavior in the genomics era. Encyclopedia of Animal Behavior (Second Edition). Elsevier: 223-233.

Kocher SD, Tsuruda JM, Gibson JD, Emore CM, Arechavaleta-Velasco ME, Queller DC, Westerman R (2015) A search for parent-of-origin effects on honey bee gene expression. G3: Genes Genomes Genet 5(8):1657-1662.

Kucharski R, Maleszka J, Foret S, Maleszka R (2008) Nutritional control of reproductive status in honeybees via DNA methylation. Science 319(5871):1827-1830.

Kwong WK, Moran NA (2016) Gut microbial communities of social bees. Nat Rev Microbiol 14(6):374-384.

Kwong, WK, Steele MI, Moran NA (2018). Genome sequences of *Apibacter* spp., gut symbionts of Asian honey bees. Genome Biol Evolution 10(4):1174-1179.

Lapidge KL, Oldroyd BP, Spivak M (2002) Seven suggestive quantitative trait loci influence hygienic behavior of honey bees. Naturwissenschaften 89(12):565-568.

Lattorff HM, Moritz RF, Crewe RM, Solignac M (2007) Control of reproductive dominance by the thelytoky gene in honeybees. Biol Letters 3(3):292-295.

Lee FJ, Rusch DB, Stewart FJ, Mattila HR, Newton IL (2015) Saccharide breakdown and fermentation by the honey bee gut microbiome. Environ Microbiol 17(3):796-815.

Leonard SP, Powell JE, Perutka J, Geng P, Heckmann LC, Horak RD, Davies BW, Ellington AD, Barrick JE, Moran NA (2020) Engineered symbionts activate honey bee immunity and limit pathogens. Science 367(6477):573-576.

Li, L, Liu F, Li W, Li Z, Pan J, Yan L, Zhang S, Huang ZY Su S (2012) Differences in microRNAs and their expressions between foraging and dancing honey bees, *Apis mellifera* L. J Insect Physiol 58(11):1438-1443.

Libbrecht R, Oxley PR, Keller L, Kronauer DJC (2016). Robust DNA methylation in the clonal raider ant brain. Curr Biol 26(3):391-395.

Li-Byarlay H (2016) The function of DNA methylation marks in social insects. Front Ecol Evol 4: 57.

Li-Byarlay H, Li Y, Stroud H, Feng S, Newman TC, Kaneda M, ... Jacobsen SE (2013). RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative splicing in the honey bee. Proc Nat Acad Sci 110(31):12750-12755.

Liu F, Peng W, Li Z, Li W, Li L, Pan J, Su S (2012) Next-generation small RNA sequencing for microRNAs profiling in *Apis mellifera*: comparison between nurses and foragers. Insect Molec Biol 21(3):297-303.

Lockett GA, Kucharski R, Maleszka R (2012) DNA methylation changes elicited by social stimuli in the brains of worker honey bees. Genes Brain Behav 11(2):235-242.

Lockett GA, Wilkes F, Helliwell P, Maleszka R (2014) Contrasting effects of histone deacetylase inhibitors on reward and aversive olfactory memories in the honey bee. Insects 5(2):377-398.

Lyko F, Foret S, Kucharski R, Wolf S, Falckenhayn C, Maleszka R (2010). The honey bee epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol 8(11): e1000506.

Maher B (2008) The case of the missing heritability. Nature 456(7218):18-22.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, Cho JH (2009). Finding the missing heritability of complex diseases. Nature 461(7265):747-753.

Mattila HR, Rios D, Walker-Sperling VE, Roeselers G, Newton IL (2012) Characterization of the active microbiotas associated with honey bees reveals healthier and broader communities when colonies are genetically diverse. PLoS ONE 7: e32962.

McAfee A, Pettis JS, Tarpy DR, Foster LJ (2019) Feminizer and doublesex knock-outs cause honey bees to switch sexes. PLoS Biol 17(5): e3000256.

Michelsen A, Kirchner WH, Lindauer M (1986) Sound and vibrational signals in the dance language of the honeybee, *Apis mellifera*. Behav Ecol Sociobiol 18(3):207-212.

Molodtsova D, Harpur BA, Kent CF, Seevananthan K, Zayed A (2014) Pleiotropy constrains the evolution of protein but not regulatory sequences in a transcription regulatory network influencing complex social behaviours. Front Genet 5:431.

Moran, NA (2015) Genomics of the honey bee microbiome. Curr Opin Insect Sci 10:22-28.

Moritz RFA (1988) A reevaluation of the two-locus model hygienic behavior in honey bees, *Apis mellifera* L. J. Hered. 79(4):257-62.

Müller U (2012) The molecular signalling processes underlying olfactory learning and memory formation in honeybees. Apidologie *43*(3):322-333.

Oldroyd BP, Allsopp MH, Roth KM, Remnant EJ, Drewell RA, Beekman M. (2014). A parent-of-origin effect on honeybee worker ovary size. Proc Roy Soc B: Biol Sci 281(1775):20132388.

Oxley PR, Spivak M, Oldroyd BP (2010) Six quantitative trait loci influence task thresholds for hygienic behaviour in honeybees (*Apis mellifera*). Mol Ecol 19(7):1452-1461.

Page Jr, RE, Rueppell O, Amdam GV (2012) Genetics of reproduction and regulation of honeybee (*Apis mellifera* L.) social behavior. Ann Rev Genet 46:97-119.

Park D, Jung JW, Choi BS, Jayakodi M, Lee J, Lim J, Yu Y, Choi YS, Lee ML, Park Y, Choi IY (2015) Uncovering the novel characteristics of Asian honey bee, *Apis cerana*, by whole genome sequencing. BMC Genomics 16(1):1.

Queller DC (2003) Theory of genomic imprinting conflict in social insects. BMC Evol Biol 3(1):15.

Raymann K, Moran NA (2018) The role of the gut microbiome in health and disease of adult honey bee workers. Curr Opin Insect Science 26:97-104.

Remnant EJ, Shi M, Buchmann G, Blacquière T, Holmes EC, Beekman M, Ashe A (2017) A diverse range of novel RNA viruses in geographically distinct honey bee populations. J Virol 91(16).

Rinderer TE, Beaman LD (1995) Genic control of honey bee dance language dialect. Theoret Appl Genet 91(5):727-732.

Robinson GE (2002) Sociogenomics takes flight. Science 297(5579):204-205.

Rokop ZP, Horton MA Newton ILG (2015) Interactions between cooccurring lactic acid bacteria in honey bee hives. Appl Environ Microbiol 81(20):7261-7270.

Rothenbuhler WC (1964) Behavior genetics of nest cleaning in honey bees. Iv. Responses of F1 and back-cross generations to disease-killed blood. Am Zool 4:111-123.

Rueppell O (2014) The architecture of the pollen hoarding syndrome in honey bees: implications for understanding social evolution, behavioral syndromes, and selective breeding. Apidologie 45(3):364-374.

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

Sabree ZL, Hansen AK, Moran NA (2012) Independent studies using deep sequencing resolve the same set of core bacterial species dominating gut communities of honey bees. PloS One 7(7):e41250.

Schulte C, Theilenberg E, Müller-Borg M, Gempe T, Beye M (2014). Highly efficient integration and expression of piggyBac-derived cassettes in the honeybee (*Apis mellifera*). Proc Nat Acad Sci 111(24):9003-9008.

Schwarz RS, Huang Q, Evans JD (2015) Hologenome theory and the honey bee pathosphere. Curr Opinion Insect Sci 10:1-7.

Sen Sarma M, Rodriguez-Zas SL, Hong F, Zhong S, Robinson GE (2009) Transcriptomic profiling of central nervous system regions in three species of honey bee during dance communication behaviour. PloS One 4(7):887

Sen Sarma M, Rodriguez-Zas SL, Gernat T, Nguyen T, Newman T, Robinson GE (2010) Distance-responsive genes found in dancing honey bees. Genes Brain Behav 9(7):825-830.

Sheppard W, Meixner M (2003) *Apis mellifera pomonella*, a new honey bee subspecies from Central Asia. Apidologie 34(4):367-375.

Simola DF, Graham RJ, Brady CM, Enzmann BL, Desplan C, Ray A, Berger SL (2016) Epigenetic (re) programming of caste-specific behavior in the ant *Camponotus floridanus*. Science 351(6268).

Simone-Finstrom M, Borba RS, Wilson M, Spivak M (2017) Propolis counteracts some threats to honey bee health. Insects 8(2):46.

Smith EA, Martin-Eberhardt SA, Miller DL, Parish AJ Newton IL (2019) Draft genome sequence of a *Bombella apis* strain isolated from honey bees. Microbiol Resour Announce 8(47).

Spivak M (1996) Honey bee hygienic behavior and defense against *Varroa jacobsoni*. Apidologie 27:245-260.

Spivak M, Gilliam M (1998a) Hygienic behaviour of honey bees and its application for control of brood diseases and Varroa - Part II. Studies on hygienic behaviour since the Rothenbuhler era. Bee World 79(4):169-186.

Spivak M, Gilliam M (1998b) Hygienic behaviour of honey bees and its application for control of brood diseases and Varroa Part I. Hygienic behaviour and resistance to American foulbrood. Bee World 79(3):124-134.

Spotter A, Gupta S, Nurnberg G, Peinsch N, Bienefeld K (2012) Development of a 44K SNP assay focussing on the analysis of a Varroa-specific defence behaviour in honye bees (*Apis mellifera carnica*). Mol Ecol Resour 12: 323-332.

Spotter A, Gupta P, Mayer M, Reinsch N, Bienefeld K (2016) Genome-wide association study of a Varroa-specific defense behavior in honeybees (*Apis mellifera*). J Hered 107(3):220-227.

Standage DS, Berens AJ, Glastad KM, Severin AJ, Brendel VP, Toth AL (2016) Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. Mol Ecol 25(8):1769-1784.

Szyf M, Meaney MJ (2008) Epigenetics, behaviour, and health. Allergy Asthma Clin Immunol 4(1):37.

Thom C, Gilley DC, Hooper J Esch HE (2007) The scent of the waggle dance. PLoS Biol 5(9).

Tsuruda JM, Harris JW, Bourgeois L, Danka RG, Hunt GJ (2012) High-resolution linkage analyses to identify genes that influence Varroa sensitive hygiene behavior in honey bees. PLoS ONE 7(11).

Vásquez A, Olofsson, TC (2009) The lactic acid bacteria involved in the production of bee pollen and bee bread. J Apicult Res 48(3):189-195.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD (2001) The sequence of the human genome. Science 291(5507):1304-1351.

Von Frisch K (1967) The Dance Language and Orientation of Bees. Belknap Press of Harvard University Press.

Wallberg A, Han F, Wellhagen G, Dahle B, Kawata M, et al. (2014) A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honeybee *Apis mellifera*. Nat Genet 46:1081-1088.

Wallberg A, Pirk CW, Allsopp MH, Webster MT (2016) Identification of multiple loci associated with social parasitism in honeybees. PLoS Genet 12(6):e1006097.

Wallberg A, Schöning C, Webster MT, Hasselmann M (2017) Two extended haplotype blocks are associated with adaptation to high altitude habitats in East African honey bees. PLoS Genet 13(5):e1006792.

Wang Y, Jorda M, Jones PL, Maleszka R, Ling X, Robertson HM, Robinson GE (2006) Functional CpG methylation system in a social insect. Science 314(5799):645-647.

Weaver DB, Anzola JM, Evans JD, Reid JG, Reese JT, Childs KL, Zdobnov EM, Samanta MP, Miller J, Elsik CG (2007) Computational and transcriptional evidence for microRNAs in the honey bee genome. Genome Biol 8(6):R97.

Whitfield CW, Behura SK, Berlocher SH, Clark AG, Johnston JS et al. (2006) Thrice out of Africa: ancient and

recent expansions of the honey bee, *Apis mellifera*. Science 314(5799):642-645.

Wilfert L, Long G, Leggett HC, Schmid-Hempel P, Butlin R, Martin SJM, Boots M (2016) Deformed wing virus is a recent global epidemic in honeybees driven by Varroa mites. Science *351* (6273):594-597.

Wilson AC, Duncan RP (2015) Signatures of host/symbiont genome coevolution in insect nutritional endosymbioses. Proc Nat Acad Sci 112(33):10255-10261.

Wragg D, Marti-Marimon M, Basso B, Bidanel J-P, Labarthe E, et al. (2016) Whole-genome resequencing of honeybee drones to detect genomic selection in a population managed for royal jelly. Sci Rep 6:27168.

Xu S (2003) Theoretical basis of the Beavis effect. Genet 165(4):2259-2268.

Yagound B, Dogantzis KA, Zayed A, Lim J, Broekhuyse P, Remnant EJ, Beekman M, Allsopp MH, Aamidor SE, Dim O, Buchmann G (2020) A single gene causes thelytokous parthenogenesis, the defining feature of the Cape honey bee *Apis mellifera capensis*. Curr Biol 30(12):2248-2259.

Yun JH, Lee JY, Hyun DW, Jung MJ, Bae JW (2017) *Bombella apis sp. nov.*, an acetic acid bacterium isolated from the midgut of a honey bee. Int J Systematic Evol Microbiol 67(7): 2184-2188.

Zayed A, Robinson GE (2012). Understanding the relationship between brain gene expression and social behavior: lessons from the honey bee. Ann Rev Genet 46:591-615.

Zayed A, Whitfield CW (2008) A genome-wide signature of positive selection in ancient and recent invasive expansions of the honey bee *Apis mellifera*. Proc Nat Acad Sci USA 105:3421-3426.

Zuk O, Hechter E, Sunyaev SR, Lander ES (2012) The mystery of missing heritability: Genetic interactions create phantom heritability. Proc Nat Acad Sci 109(4):1193-1198.