



# Improving bee health through genomics

Christina M. Crozinger<sup>1</sup> and Amro Zayed<sup>2</sup>

**Abstract** | Declines in bee populations across the world threaten food security and ecosystem function. It is currently not possible to routinely predict which specific stressors lead to declines in different populations or contexts, hindering efforts to improve bee health. Genomics has the potential to dramatically improve our ability to identify, monitor and predict the effects of stressors, as well as to mitigate their impacts through the use of marker-assisted selection, RNA interference and potentially gene editing. Here we discuss the most compelling recent applications of genomics to investigate the mechanisms underpinning bee population declines and to improve the health of both wild and managed bee populations.

## Allozymes

Structural but not necessarily functional enzyme variants that migrate at different rates during electrophoresis, which are the result of genetic changes in the enzyme-encoding gene.

## Restriction fragment length polymorphisms

Genetic variations that create or abolish a restriction enzyme recognition site, leading to variation in DNA fragment sizes following digestion by restriction endonucleases, as revealed by gel electrophoresis.

Bees provide indispensable pollination services, and their decline threatens the integrity of ecosystem function and food security<sup>1,2</sup>. Several factors have been associated with declines of populations of bees, such as habitat loss, pathogens and parasites, exposure to toxins, malnutrition and climate change. However, the exact causes of decline in a specific population or species often remain elusive, which hinders conservation and management efforts.

Concerns over pollinator population decline surfaced in the 1990s and intensified in the early 2000s. These concerns motivated research on the conservation biology of bees, which initially focused on gathering evidence on the extent and scope of bee population declines. Ecological studies were helpful in highlighting vulnerable taxa and habitats<sup>3</sup>, determining the timescale over which declines occur<sup>4</sup> and formulating general hypotheses for the causes of declines<sup>5</sup> and their effects on ecosystem function<sup>6,7</sup>. Using genetic markers such as allozymes, restriction fragment length polymorphisms and microsatellites, genetic research was focused on understanding how haplodiploidy and other life-history traits of bees influence their conservation biology (BOX 1).

In honeybees (genus *Apis*) and bumblebees (genus *Bombus*), genetic markers were also used to construct linkage maps and identify broad regions of the genome that influence phenotypic traits<sup>8,9</sup>. While these foundational studies were critical in shaping the field, they often lacked the power and genetic resolution to uncover the specific mechanisms causing pollinator population decline, were too expensive or inefficient to be applied to studies of different bee species in the field and could not be readily adapted to develop tools and solutions to improve bee health.

Advances in next-generation sequencing (NGS) technologies and tools have overcome many of these

limitations, and genomic approaches such as population genomics, transcriptomics and metagenomics can now provide the power, resolution and efficiency required to understand declines in bee populations at a mechanistic level<sup>10–12</sup>. These genomic approaches have been applied most extensively in managed western honeybees (*Apis mellifera*) because theirs was the first bee genome to be sequenced<sup>13</sup>, but, as the number of sequenced bee genomes increases, these approaches could soon be applied more broadly (FIG. 1). It is important to note that bees are complex animals that interact with diverse biotic and abiotic environmental factors, and these interactions can greatly influence their genomic, transcriptomic and metagenomic profiles. Nevertheless, the ability to generate and share these sequence profiles at an increasingly large scale now makes it possible to conduct studies in ecologically relevant conditions to identify the core genes, pathways and microbial communities associated with both stressed and healthy bees. The scale at which genomics is being applied to the understanding of bee population declines generates the potential to harness and directly apply omic tools to improve bee health (TABLE 1).

In this Review, we discuss several existing and emerging genomic approaches for studying bee health. By whole-genome sequencing of samples of bees, population genomics can be used to identify genomic variation associated with populations subject to decline or selection, to identify target subspecies or species for conservation or monitoring and to identify specific traits, which may subsequently be used for marker-assisted selection. In particular, we discuss genome-wide association studies (GWAS), which require large-scale sequencing but do not depend on crosses and thus can be used for populations and species not amenable to breeding. We next discuss how transcriptomics can improve our

<sup>1</sup>Department of Entomology, Center for Pollinator Research, Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, PA, USA.

<sup>2</sup>Department of Biology, York University, Toronto, ON, Canada.

✉e-mail: cmgroat@psu.edu; zayed@yorku.ca

<https://doi.org/10.1038/s41576-020-0216-1>

**Box 1 | Bee traits and conservation**

Bees are a diverse taxon, with more than 20,000 described species<sup>136</sup> (FIG. 1). These species differ dramatically in their life history traits, with the vast majority of species exhibiting a solitary life cycle, where the female creates a nest and collects pollen to provision her eggs and the offspring develop without further interactions with the mother. At the other extreme are eusocial bees, such as bumblebees and honeybees, which form social groups with cooperative brood care, overlapping generations and reproductive division of labour, including caste differentiation between reproductive queens and non-reproductive workers.

Bees are haplodiploid; under typical circumstances, females are diploid and males are haploid. Sex is determined by genotype at a hypervariable autosomal gene called complementary sex determiner (*csd*) in honeybees<sup>137</sup>. Homozygosity at this locus leads to the production of inviable or effectively sterile diploid males instead of females<sup>138</sup>. This mode of sex determination reduces the breeding effective population size of bees<sup>139</sup> and imposes a severe genetic load in small isolated populations, which can result in rapid extirpation via the 'diploid male extinction vortex'<sup>140</sup>.

In addition to haplodiploidy, nest building and central place foraging<sup>14,141</sup>, sociality<sup>8</sup> and diet specialization<sup>142</sup> have all been hypothesized to affect important population genetic parameters that influence population viability. Eusociality has been shown to limit effective population size in insects<sup>143</sup>, but there is a need for larger studies that control for phylogeny. A recent meta-analysis examined the role of life history traits on patterns of genetic differentiation — and, by extension, rates of gene flow — in bees<sup>141</sup>. It showed that body size was the strongest predictor of population genetic structure in bees; smaller bees (which cannot fly as far and thus have limited gene flow) exhibited higher levels of genetic differentiation relative to larger bees. Sociality had a weaker but significant effect on population structure, whereby eusocial bees had lower levels of population structure (that is, higher rates of dispersal) relative to solitary bees<sup>141</sup>. It is not clear whether the reduced effective population size caused by eusociality is offset by the ability of eusocial bees to disperse over greater distances. In the same analysis, diet specialization had no influence on the population structure of bees<sup>141</sup>. Thus, overall, these studies suggest that small, solitary bees are likely the most sensitive to local habitat change (including patterns of pesticide use<sup>23</sup>) and climate change.

**Microsatellites**

Stretches of repetitive DNA with a core motif (typically two or three bases) that is repeated many times. Microsatellites have high mutation rates, leading to many alleles segregating within populations. Because of their polymorphic nature, they are naturally suited for DNA fingerprinting applications.

**Haplodiploidy**

A genetic system in which females are diploid and males are haploid.

**Population genetics**

Analysis of genetic diversity at a genome scale in populations to estimate population genetic parameters or to link genotype with phenotype.

**Transcriptomics**

Analysis of transcripts within specific tissues using microarrays or RNA sequencing.

understanding of molecular, physiological and behavioural responses of bees to diverse stressors, allowing the development of approaches which may increase resilience of these bees. Studies have demonstrated that there are conserved pathways which respond to certain stressors under a wide variety of conditions, and we propose that these transcriptional responses can serve as diagnostic biomarkers for rapid screening and assessment of bee populations. We then discuss how metagenomics has been used to characterize populations of pathogenic and beneficial microbial species within bees or their nests and to identify the species of flowering plants from which they are collecting resources. Finally, we discuss the potential and challenges associated with implementing tools to manipulate gene function in bees, including RNA interference (RNAi) and gene editing.

**Population genomics**

Sequencing the genomes of bees has allowed us to identify populations that are at risk or in decline and associate declining populations with specific current or historical conditions. Population genomics has also led to the development of tools to secure the genetic integrity and enhance the resilience of bee populations (FIG. 2).

**Population and comparative genomic studies reveal the context of bee population declines.** Early research on the population genetics of bees helped identify species that are at risk because of life history traits that intrinsically increase the probability of extinction and reduce

the ability of pollinators to cope with biotic and abiotic stressors<sup>2,14</sup> (BOX 1). Recent population genetic studies have also enhanced our understanding of how landscape context influences dispersal or population density in social species. While simply evaluating the abundance of solitary bees within a landscape can provide information on overall population size, this approach is intractable for social species in which an individual colony can produce hundreds or (in the case of honeybees) thousands of individual workers. Manual assessment of colony density is extremely challenging because nests are often inconspicuous and bees can travel hundreds or thousands of metres from their nest to forage<sup>15</sup>. By using genetic markers, researchers can identify nestmates among collected samples, allowing them to calculate the number of colonies contributing to the population of bees collected at a given site<sup>16</sup>. By comparing samples across sites, researchers can evaluate foraging distances or how landscape quality can influence colony density<sup>16</sup>. Similarly, comparison of samples across time enables researchers to evaluate how landscape context affects colony growth and survival and the production of the following season's reproductives<sup>17</sup>. These studies demonstrate that landscapes that provide abundant floral resources, particularly during critical periods of colony establishment and growth in the spring, can positively influence colony survival, growth and reproduction<sup>17,18</sup>, factors that likely benefit solitary species as well.

Population genetic studies typically use a small number (tens to hundreds) of genetically anonymous markers for inference; as these markers are not associated with particular genes, their ability to address mechanistic questions and study adaptive evolutionary processes is limited. By contrast, population genomics interrogates patterns of genetic diversity across the genome and can therefore identify populations that are in decline, determine the time frame of these declines and begin to ascertain the causes (FIG. 2a). A recent study of the at-risk yellow-banded bumblebee (*Bombus terricola*) illustrates the power of population genomics for bee conservation<sup>19</sup>. Researchers generated a reference genome for *B. terricola* and then conducted a population genomic study in which genomes from declining and stable bee populations were sequenced and compared. The resulting dense single-nucleotide polymorphism (SNP) dataset allowed estimations of the date and severity of *B. terricola*'s decline and provided evidence of substantial inbreeding in the species. Unlike traditional markers used in conservation genetics (such as microsatellites), NGS data allow researchers to directly query levels of variation and selection acting on specific genes; for example, tests of very recent positive selection can implicate genes that may have played a role in ameliorating the effects of certain stressors. In *B. terricola*, genes showing signatures of recent selective sweeps were often enriched for innate immune system function, which lends some support to the hypothesis that spillover of pathogens from commercial bee populations contributes to the decline of native bumblebee populations<sup>5</sup>.

Genome sequences from additional bee species will certainly improve our ability to understand why some bee lineages may be particularly susceptible to some

**Metagenomics**

Analysis of DNA or RNA from communities of organisms using high-throughput sequencing approaches. Metagenomics does not require isolation of specific species or strains from collected samples before sequencing and thus can be used to identify all species or variants within a sample using bioinformatics.

**Marker-assisted selection**

Artificial selection programmes using predictive genetic markers to select individuals for breeding

**Genome-wide association studies**

(GWAS). Studies that investigate the association between genotypes across the genome and their influence on phenotypic traits in natural populations.

**Effective population size**

The size of an 'ideal' population (a random mating population of constant size with Poisson variation in family sizes) that would have the same genetic parameters as the actual population under study.

**RNA interference**

(RNAi). The application of double-stranded RNA molecules to reduce or silence the expression of target genes.

**Single-nucleotide polymorphism**

(SNP). A point mutation in a DNA sequence that introduces variation between individuals of a group or species.

**Positive selection**

An evolutionary force that increases the frequency of beneficial mutations within populations.

**Selective sweeps**

Processes by which strong positive selection on a mutation results in reduced genetic diversity at nearby linked loci.

**Admixed**

Pertaining to genomes that have been generated by hybridization of typically distinct genomes.

**Extinction via hybridization**

The loss of naturally distinct evolutionary lineages as a result of hybridization with other — typically managed — populations.

**stressors.** For example, comparative functional genomic approaches supported by several community-led databases (such as BeeBase and HymenopteraMine) have been used to identify the molecular mechanisms responsible for variation in sensitivity or resilience to agrochemicals in different bee species. Studies in honeybees (*A. mellifera*), bumblebees (*Bombus terrestris*), mason bees (*Osmia cornifrons*) and leafcutter bees (*Megachile rotundata*) identified the genes responsible for detoxification of neonicotinoid pesticides and demonstrated species-specific differences in the cytochrome P450 genes involved in these processes, thereby revealing potential mechanisms underlying differences in sensitivity to these pesticides<sup>20–22</sup>. This approach can be used more broadly to develop or deploy pesticides that are better able to target pest species rather than beneficial insect species or non-target species; however, given the number and diversity of non-target species, efforts to reduce insecticide use and exposure while maintaining agricultural output are likely to be more effective than 'targeted' pesticides<sup>23</sup>.

**Tools for bee biosecurity.** The movement of bees, unintentionally through global trade routes, intentionally through commercial beekeeping practices or through habitat alterations caused by land use regimes or climate change, can introduce new bee species, subspecies, pathogens and parasites to a region. Indeed, commercial beekeeping and trade facilitated the global spread of *Varroa destructor* mites, the most deadly parasite of honeybees<sup>24</sup>. Vectoring of deformed wing virus (DWV) by *Varroa* seems to have selected for specific viral strains that have subsequently spread<sup>25,26</sup>. In South America, the introduction of exotic European bumblebee species for commercial pollination services led to the introduction of a novel parasite, the protozoan *Apicystis bombi*, into native bumblebee populations<sup>27</sup>. Identifying recently introduced bee, pathogen or parasite species or strains, their distribution and factors influencing this distribution are greatly facilitated, and in some cases only possible, through the use of genomic approaches that can specifically detect and distinguish among genetic variants.

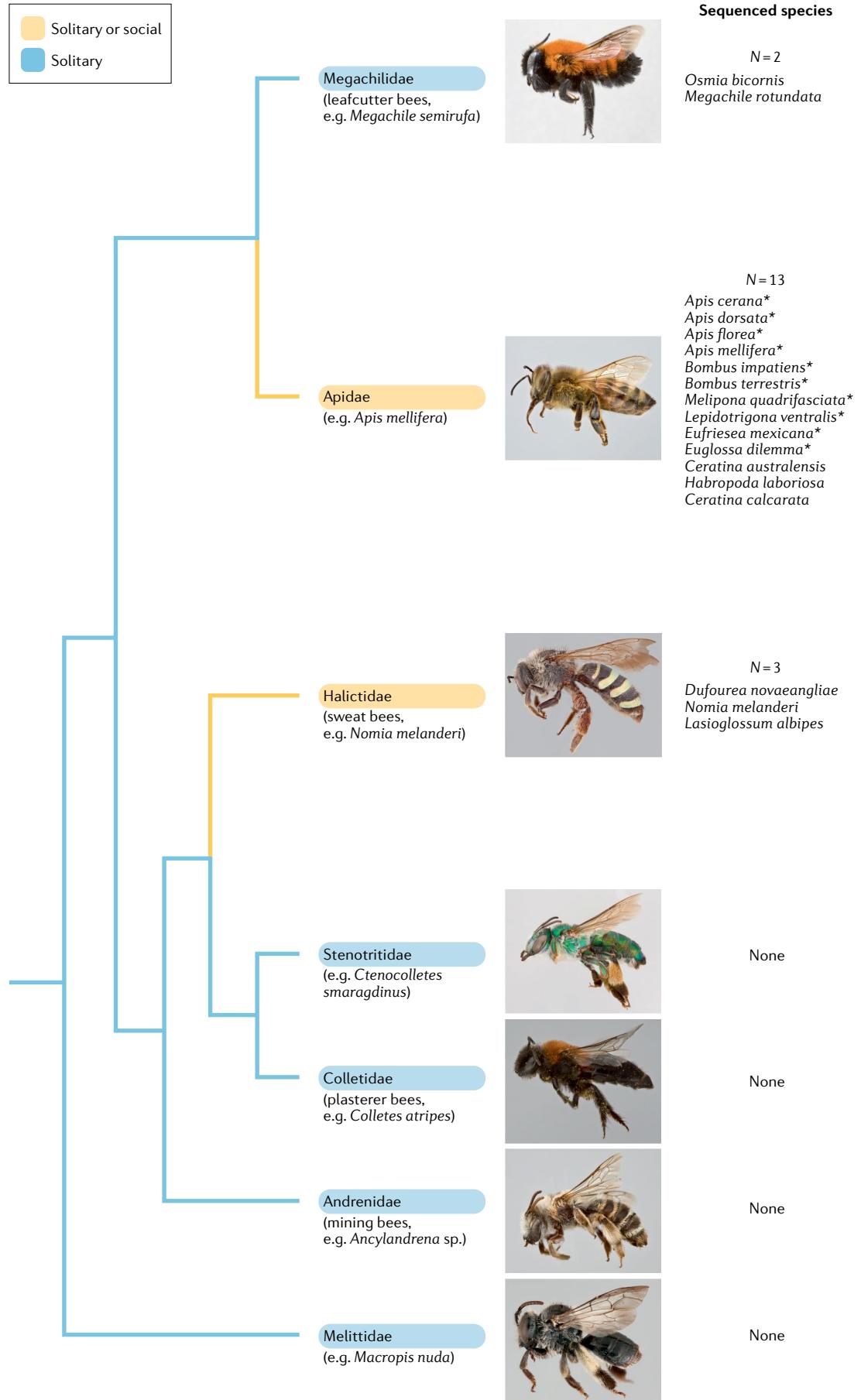
For example, population genomic studies led to the development of tools to differentiate between different strains of honeybees, some of which are targets for conservation, whereas others can become invasive pests (FIG. 2b). The honeybee has at least 27 geographically restricted and locally adapted subspecies in its native range in Africa, Asia and Europe<sup>10</sup>. While managed honeybees are often derived from European honeybee subspecies<sup>28</sup>, the introduction of African honeybees to South America in 1956 resulted in the spread of Africanized 'killer' bees across the continent and into North America<sup>29</sup>. The genomes of Africanized bees are highly admixed (approximately 75% African descent and approximately 25% European descent<sup>30</sup>) and these bees are highly aggressive, which renders them undesirable for many beekeepers. Identifying Africanized bees has historically been problematic as methods based solely on mitochondrial DNA have no power to detect invasion by Africanized drones, while morphometric analyses

are unable to assign hybrid individuals with high precision<sup>31</sup>. Whole-genome sequencing has revealed large genetic differences between African and European honeybees<sup>32</sup>, which have been exploited to develop an ancestry-informative SNP panel that can identify Africanized honeybees and their hybrids with high accuracy<sup>33,34</sup>. The deployment of these SNP tests is expected to enhance bee biosecurity and limit secondary introductions of Africanized killer bees.

Ancestry-informative SNP panels have also been developed to protect native honeybee subspecies from admixture with commercial stocks (FIG. 2b). Native subspecies in the Old World are at constant risk of extinction via hybridization<sup>35</sup> with nearby commercial colonies, which are often highly admixed<sup>28</sup>. In Europe, there are specific concerns over the loss of the genetic integrity of several native subspecies, including *A. mellifera mellifera* and *A. mellifera iberiensis*. Population genomic studies have demonstrated that a small number of informative SNPs can be used to differentiate native *A. mellifera mellifera* and *A. mellifera iberiensis* from introduced commercial honeybees and other native subspecies found in Europe<sup>36,37</sup>. These tools will be valuable in establishing conservation breeding programmes for these threatened subspecies.

**Selection of stocks with resilient traits**

Many traits that influence bee health show considerable heritability<sup>38,39</sup>. For example, several social and innate immune traits in honeybees have a narrow sense heritability in the range of 25–50%<sup>38,40,41</sup>. In the bumblebee *B. terrestris*, distinct loci are known to control a total of 10–15% of the phenotypic variance for several immune traits<sup>42,43</sup>. While many non-genomic factors can influence the ability of bees to cope with both biotic and abiotic stressors, it should still be possible to select for populations of bees that have increased resilience to common stressors. However, even in honeybees, which have a long history of management and tools for instrumental insemination<sup>44</sup>, selective breeding programmes have often been small in scale<sup>45,46</sup>, and their efficacy is threatened by 'genetic dilution' because queen honeybees naturally mate with a large number of drones from geographically distant (that is, likely unselected) colonies<sup>47</sup>. In theory, high-throughput assessment of a large number of colonies, facilitated by genetic markers, can allow the development of large-scale breeding programmes that could lead to substantial and persistent improvements in colony traits. However, identifying markers and mutations associated with specific traits has been challenging in bees. Before the introduction of genomics, researchers used quantitative trait loci (QTLs) to study the genetics of social and innate immunity in honeybees and bumblebees<sup>43,48–50</sup>. This approach involved carrying out crosses between stocks with distinct traits. The resulting F1 hybrids were then backcrossed either to the parental stocks or to other hybrids, and the offspring of these crosses where phenotyped and genotyped at several hundred neutral genetic markers, such as microsatellites. Statistical associations between genotypes and phenotypes provided some clues about the number and location of QTLs that influence phenotypic traits,



◀ Fig. 1 | Evolutionary relationships of bees depicting the number of published genomes per family. There are more than 20,000 described bee species<sup>136</sup>, and they are classified into seven families<sup>144</sup>. These include the Melittidae, a group of small ground-nesting bees; the ‘mining bees’ of the Andrenidae; the ‘plasterer bees’ of the Colletidae and their sister clade, the Stenotritidae; and the ‘leafcutter bees’ of the Megachilidae. Species in the aforementioned families tend to be solitary, although some species nest communally. The ‘sweat bees’ of the Halictidae and the Apidae both show a remarkable diversity of social organization, including species with solitary, semisocial, subsocial or eusocial behaviour. One subfamily within the Apidae, the Apinae, includes the corbiculate bees (honeybees, bumblebees, orchid bees and stingless bees). The corbiculate bees include species that are important for commercial pollination and/or that serve as model organisms for studying social behaviour and have therefore received substantial research and sequencing attention (10 of the 13 sequenced Apidae genomes are from corbiculate bees, denoted by asterisks). It will be important to sequence species from the understudied families to better understand the evolutionary history of bees, and because these taxa both contribute to pollination and are threatened by similar stressors as the corbiculate bees. All photographs of bees are courtesy of L. Packer, York University, Toronto, Canada (from the [Packer lab Bee Galleries](#)).

**Narrow sense heritability**  
The proportion of phenotypic variance attributed to additive genetic variance.

**Quantitative trait loci (QTLs)**  
Genomic loci that contain variants influencing a quantitative trait. Quantitative traits exhibit continuous variation within populations, such as height in humans or the amount of pollen collected by bee colonies.

**Haplotype blocks**  
Stretches of DNA characterized by high levels of linkage disequilibrium.

**Thelytokous parthenogenesis**  
A form of asexual reproduction found in honeybees from the Cape region of South Africa. Worker bees have the ability to lay unfertilized diploid eggs that develop into daughter workers.

**Adverse outcomes pathway**  
A conceptual framework that links a molecular phenotype (such as a change in gene expression or level of activity of a receptor) to a phenotypic change at another level of biological organization (such as physiology or behaviour) that is associated with a particular end point of interest (such as changes in survival, population demography or size). Adverse outcomes pathways are developed and refined using empirical data and are commonly used in ecotoxicological research.

but the QTLs were often large (thousands of kilobases) and encompassed hundreds to thousands of genes. These large QTLs were not informative for breeding resilient populations as the few microsatellite markers denoting QTLs lost linkage with causal mutations and the phenotypic trait after a small number of generations of selective breeding.

Population genomic approaches, including GWAS, have tremendous potential for identifying loci that influence population resilience and health in bees (FIG. 2c). Such studies typically assay millions of genetic markers in natural or seminatural populations where linkage disequilibrium is low, resulting in higher mapping resolution relative to QTL approaches<sup>51,52</sup>. Thus far, several studies have identified important loci by sequencing the genomes of distinct populations of bees that differ in phenotypes as a result of natural selection<sup>53,54</sup> or artificial selection<sup>55,56</sup>. Sequencing the genomes of highland and lowland honeybees revealed two haplotype blocks that are likely involved in adaptation of bees to high altitudes<sup>53</sup>. Sequencing of several genomes of *A. mellifera scutellata* and *A. mellifera capensis* identified loci with high genetic differentiation between the two subspecies, which may underlie the ability of *A. mellifera capensis* workers to lay diploid eggs via thelytokous parthenogenesis<sup>54</sup> — an economically damaging phenomenon for African beekeepers. The genetics of honeybee aggression has been analysed by comparing the genomes of aggressive and gentle Africanized honeybees in South America<sup>57</sup>. In addition to comparing natural populations of bees with distinct phenotypes, comparisons between selected and unselected populations have been used to identify genetic variants associated with economically valuable traits in honeybees, such as royal jelly production and population growth rate of parasitic *Varroa* mites<sup>55,58</sup>. Similarly, the genetics underlying a social immune trait called ‘hygienic behaviour’ has been studied by sequencing the genomes of selected and unselected queens and identifying loci that display large differences in allele frequencies between the two populations and covary with phenotype in the unselected population<sup>56</sup>. Compared with previous QTL studies, which implicated several large stretches of DNA (approximately 12 megabases in total) that influenced hygienic behaviour, this genomic

approach was able to highlight 73 genes (approximately 0.9 megabases in total) associated with this trait.

Diminishing costs of sequencing now allow GWAS in unselected honeybee populations (FIG. 2c). This process involves correlating patterns of genetic polymorphism across the genome to phenotypic differences<sup>59</sup>. GWAS benefits from the high recombination rates that occur in honeybee genomes because the resulting rapid decay of linkage disequilibrium means that the genomic regions showing a statistical association with traits of interest are small (tens of kilobases or smaller); however, adequately powered experiments are needed to achieve such results<sup>60</sup>. The possibility of discovering causal variants underlying bee traits affecting pathogen resistance and overall health will be very useful for breeding programmes powered by marker-assisted selection<sup>61</sup>. Importantly, because the GWAS approach does not require experimental crosses, it can be easily extended to study wild bee species.

## Transcriptomics

Transcriptomic studies using NGS have been instrumental in characterizing the molecular and physiological mechanisms by which bees respond to diverse stressors, in addition to providing possible tools for rapid diagnosis of bee health (FIG. 3).

**Revealing the effects of biotic and abiotic stressors on bee health.** Several studies have used transcriptomic approaches to understand how biotic and abiotic stressors affect bee populations. Most studies have been conducted on honeybees, using RNA sequencing analyses to evaluate changes in genome-wide gene expression patterns associated with parasitization<sup>62</sup>, viral infections<sup>63</sup>, pesticide exposure<sup>64</sup>, nutritional stress<sup>65</sup> and exposure to multiple stressors<sup>66</sup> (FIG. 3). Transcriptomic analyses have also been conducted to understand the pathways involved in the response to temperatures associated with overwintering and diapause in diverse bee species<sup>67–69</sup>. These studies have revealed that these stressors have complex and interconnected effects on core physiological pathways in pollinators; this information can help us understand, and possibly mitigate, the effects of multiple stressors on pollinator health.

In some cases, genes are associated with an important life history trait, and thus gene expression changes can be used to infer long-term changes in an organism’s behaviour, physiology or lifespan, similarly to the adverse outcomes pathway approach used in ecotoxicology<sup>70</sup>. For example, in honeybees, workers that are exposed to diverse stressors exhibit accelerated behavioural maturation<sup>71</sup>, which is associated with shortened lifespans of individual workers and imbalanced colony demography that can lead to colony collapse<sup>71</sup>. In *A. mellifera*, accelerated behavioural maturation is mediated by increased levels of juvenile hormone and reduced expression of the vitellogenin gene (*Vg*). Vitellogenin is involved in nutritional storage, immune function and oxidative stress resilience, and also negatively regulates levels of juvenile hormone<sup>72</sup>. Thus, levels of *Vg* RNA can serve as a biomarker of ‘biological age’ of the workers in a colony, and increased biological age can be an indication

Table 1 | An omics toolkit for understanding and improving bee health

Goals	Applications and interventions	Outlook	Examples
<b>Population genomics</b>			
Identify at-risk populations	Focus efforts for conservation on key populations	In use	SNPs used to quantify and protect the genetic integrity of the 'dark' European honeybee <sup>37,145</sup>
Characterize population densities and foraging range	Improve habitat structure to support bee populations	In use	Microsatellites used to estimate habitat features that support healthy bumblebee populations <sup>17</sup>
Identify markers that distinguish among strains or species	Manage and mitigate spread of invasive or exotic strains or species	In use	Development of a SNP chip for detection of invasive Africanized bees <sup>33</sup>
Identify genes underlying resilience	Select for more resilient stock	Emerging	Identification of candidate genes associated with social immunity in honeybees <sup>36</sup>
	Increase resilience through genetic manipulation (RNAi and CRISPR–Cas9-based genome editing) of genes associated with resilience	Future	NA
<b>Transcriptomics</b>			
Characterize responses to diverse stressors	Reveal unexpected associations between stressors and physiological or behavioural traits	In use	Elucidation of common transcriptional responses to pollen diets and pesticide exposure <sup>76</sup>
	Develop novel approaches to increase resilience by leveraging knowledge of physiological and behavioural impacts of stressors	In use	Honey constituents upregulate detoxification and immunity genes in honeybees <sup>76</sup>
	Develop diagnostic markers for diverse stressors	Emerging	Identification of common and unique transcriptional responses to different pathogens and parasites in honeybees <sup>86</sup> Identification of biomarkers of honeybee colony collapse <sup>62,74</sup>
	Increase resilience through genetic manipulation (RNAi and CRISPR–Cas9-based genome editing) of genes differentially regulated in response to stressors	Future	RNAi targeting of Varroa gene reduces Varroa population <sup>127</sup> (laboratory-based study) RNAi silencing of the honeybee Naked Cuticle gene ( <i>nkd</i> ; also known as LOC724997) improves host immune function and reduces <i>Nosema ceranae</i> infections <sup>128</sup> (laboratory-based study)
<b>Metagenomics</b>			
Characterize foraging preferences	Increase floral diversity to support bee health	In use	Identification of preferred floral species for honeybees in spring in US Midwestern agroecosystems <sup>118</sup>
Characterize pathogenic microbial communities	Manage and mitigate disease spread	Emerging	Bee-infecting viruses identified by metagenomics are broadly distributed in US bee populations <sup>93,95</sup>
Characterize gut microbiota	Manage gut microbiota (including developing probiotics) to increase resilience	Emerging	Treatment with gut microorganisms increased resilience of honeybees to <i>Nosema</i> infection <sup>111</sup> (laboratory-based study)

NA, not applicable; RNAi, RNA interference; SNP, single-nucleotide polymorphism.

of chronic stress<sup>73</sup>. Furthermore, reduced levels of Vg RNA are associated with increased risk of colony death, in both the winter and the summer<sup>62,74</sup>. Thus, Vg RNA levels can be used to monitor the potential sublethal, long-term impacts of diverse stressors, individually or in combination, in individual bees and as a measure of exposure to stress under ecological conditions, such as in samples collected from the field<sup>75</sup> (see Biomarkers for rapid diagnostics of stressors).

Transcriptomic studies can also reveal unexpected effects of stressors and potential methods for mitigating these effects. For example, pesticide exposure alters

expression of genes that also change in response to pollen consumption in honeybees<sup>76</sup>. This observation led to studies that demonstrated that feeding bees pollen-based diets increased their resilience to pesticides. Pollen contains several macronutrients (such as protein and lipids) and micronutrients<sup>77</sup>, but feeding the bees protein alone did not confer resilience<sup>76</sup>; thus, other components of the pollen are also critical for bee health. Similarly, constituents of honey upregulated expression of genes involved in detoxification, and feeding honeybees these components increased pesticide detoxification<sup>78</sup>. Additionally, exposure to neonicotinoid pesticides was

found to reduce expression of a gene involved in antiviral response; this observation led to studies demonstrating that exposure to neonicotinoids (but not other types of pesticide) increased viral titres in honeybees<sup>79</sup>.

**Biomarkers for rapid diagnostics of stressors.** Whereas the general stressors that affect bee health are known, the specific stressors acting on any given population will obviously vary over space and time. Beekeepers and conservation practitioners lack appropriate diagnostic tools that are common for other managed animal species, which has hindered our abilities to fully understand the factors that govern bee health and to mitigate declines in populations of managed and native bees. Several platforms have been developed to screen sampled bees for the presence of RNA or DNA from pathogens or parasites<sup>80</sup>, and there are manual approaches for identifying and quantifying some of the larger ectoparasites (such as *Varroa* mites). However, by use of transcriptional biomarkers, sampled bees can be screened for markers of responses to a broad range of biotic and abiotic stressors, in a high-throughput manner, thus allowing identification or diagnosis of more diverse stressors than specific pathogens and parasites (FIG. 3). Biomarkers are commonly used in human health care<sup>81</sup> and livestock management<sup>82</sup> for diagnosing disease, and a growing body of research suggests that expression profiling can be used to identify stressor-specific biomarkers in bees, because they have distinct genetic pathways for dealing with pathogenic<sup>83</sup>, nutritional<sup>84</sup> and xenobiotic<sup>85</sup> challenges that are associated with specific and thus diagnostic changes in gene expression. Although gene expression is typically sensitive to the environment, a large meta-analysis of transcriptomic studies conducted in *A. mellifera* honeybees under different conditions discovered both common and pathogen-specific changes in gene expression associated with infections<sup>86</sup>. For example, *Atg2* (also known as LOC726497), *Metap2* (also known as LOC551771) and *Dnr1* (also known as LOC412897), among other genes, were differentially regulated only in the presence of a microsporidian parasite, *Nosema*, whereas *Iap2* (also known as LOC413374), *Rel* (also known as LOC552247), *Tube* (also known as LOC725368) and *Def2*, among others, were differentially regulated only in the presence of *Varroa* and viral infections; thus, expression of these genes can potentially serve as a biomarker for infection by *Nosema* or *Varroa*, respectively. Fewer studies have been conducted

on nutritional<sup>87,88</sup> and xenotoxic<sup>76,89</sup> challenges, but the current datasets also suggest that specific stressors often lead to unique changes in the expression of some genes that can be used for diagnostic purposes. In the bumblebee *B. terrestris*, a proof-of-concept study supported the application of biomarkers in diagnosing nutritional stress<sup>90</sup>, while studies in honeybees have also demonstrated that biomarkers of nutritional stress (Vg RNA levels) differ across landscapes that differ in floral resources<sup>75</sup>.

To develop a biomarker-based diagnostic platform, researchers need to conduct a systematic effort to expose bees to field-realistic levels of a large number of relevant stressors. Honeybees are particularly amenable

to experimental manipulation and are thus suited for identifying stressor-specific biomarkers. Work on this front has begun with the BeeCSI project, which strives to use expression profiling to identify biomarkers for a diversity of stressors in managed honeybees in North America. Through such studies, it will also be possible to differentiate between gene expression differences that vary as a result of exposure to a stressor versus those that enhance resilience to a stressor. The biomarkers discovered in honeybees can be adopted for use in other species provided they are validated either via experiments on a few experimentally tractable bees (such as bumblebees or leafcutter bees) or via studies of natural populations known to be exposed to specific stressor regimes.

## Metagenomics

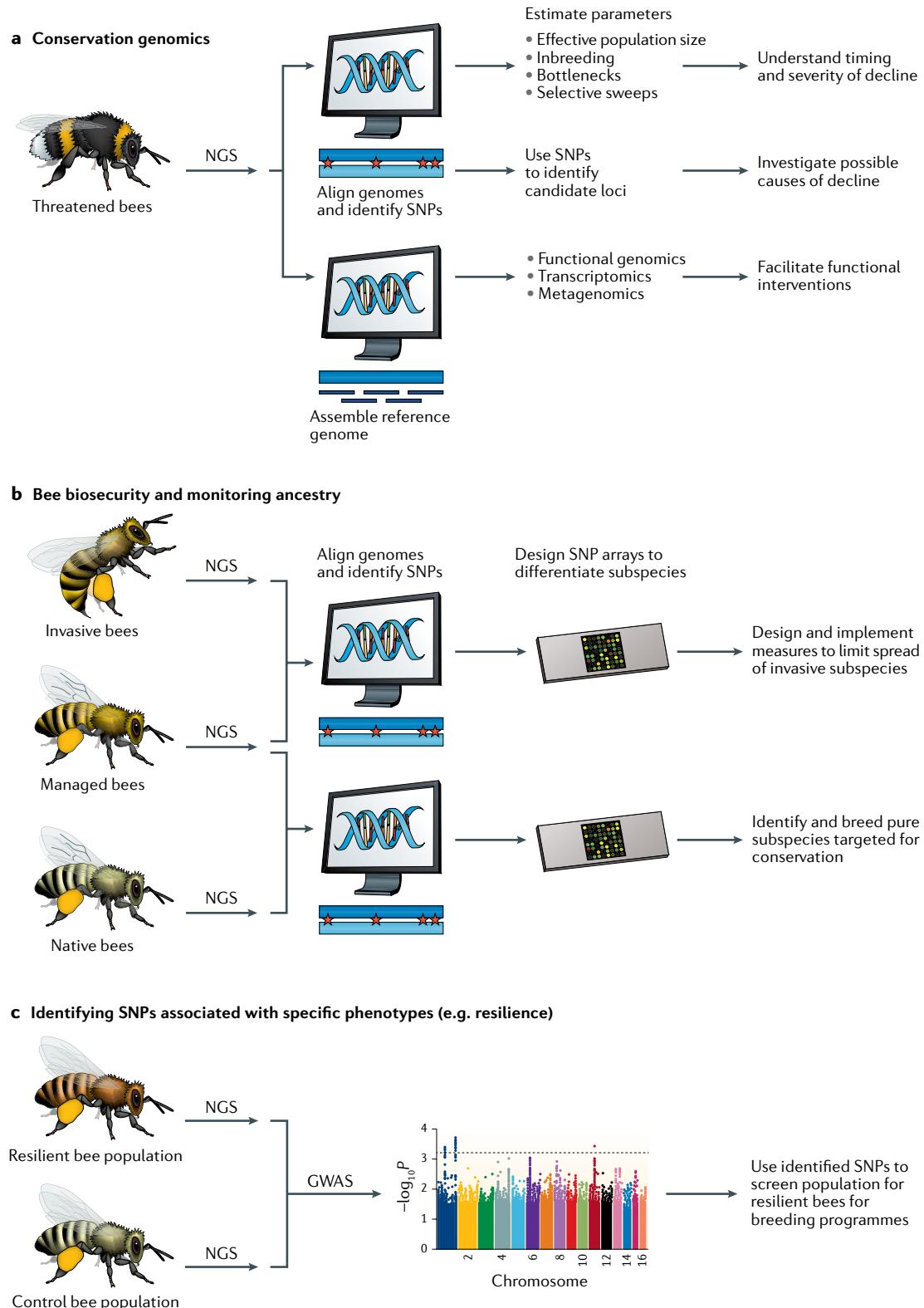
Bees interact with complex communities of plants and microorganisms. They are hosts to a broad array of both beneficial and pathogenic microorganisms. Moreover, bees collect nectar and pollen from a diversity of flowering plant species and bring these resources back to their nests (where they can be further processed by microorganisms) as the sole source of nutrition for the developing and adult bees in the colony. Metagenomic approaches, including metabarcoding, allow researchers to rapidly and efficiently identify the microbial and plant communities associated with healthy and declining bee populations (FIG. 4).

**Characterizing pathogenic and beneficial microbial communities.** Microorganisms are important contributors to bee health, serving as pathogens<sup>11</sup>, beneficial components of the gut microbiota<sup>91</sup> or as additions to improve the quality or increase the longevity of food stores<sup>92</sup>. Metagenomics bypasses the need to isolate and individually culture each species, thereby allowing efficient identification and evaluation of microbial communities (FIG. 4a,b). Researchers can readily determine which microorganisms are present in the microbial community, where they are distributed (within an organism, across populations or across species) and what influences their presence and abundance. It is important to note, however, that although several studies have documented the association of a particular microbial species with a bee host, in many cases additional research is needed to understand the fitness consequences of these associations.

Metagenomic studies can be used to identify pathogens and parasites infecting bee populations<sup>93</sup> (FIG. 4a), and then molecular approaches (such as PCR and quantitative PCR) can be used for streamlined screens for known pathogens and parasites<sup>80,95</sup>. Thus far, most studies have been retrospective, and have used previously collected samples to document the historical distributions of pathogens and parasites and how these are associated with changes in bee populations<sup>26,27</sup>. However, with ongoing, consistent collections and rapid processing of samples, it is possible to identify emerging diseases or changes in disease distributions. The US National Honey Bee Pest and Disease Survey, funded by the US Department of Agriculture's Animal

### Metabarcoding

A metagenomic method in which specific genomic regions (rather than whole genomes) are amplified and sequenced. These regions are selected to show high interspecific variation and low intraspecific variation, allowing identification of the different species within the sample. Metabarcoding is often used to minimize sequencing costs or to investigate fairly well-characterized communities.



and Plant Health Inspection Service, collects samples from managed honeybee colonies across the USA and evaluates these for a panel of common bee viruses. With this dataset, it has been demonstrated that the prevalence of a strain of DWV (DWV-B or *Varroa destructor* virus 1) has increased dramatically in recent years<sup>94</sup>, and metagenomic approaches have been used to evaluate

the distribution of a suite of viruses recently identified in bees<sup>95</sup>. The maintenance of well-curated and accessible collections of bees or omic data collected from bee samples<sup>96</sup>, across broad spatiotemporal scales, is invaluable for studies of changes in the distribution and prevalence of species and strains of pathogens, parasites and bees.

◀ Fig. 2 | Population genomic approaches to study and improve bee health. Population genomic approaches can identify genomic variation within and between populations or species of bees, allowing biologists to understand the context and possible causes of species decline and to design specific interventions to improve health outcomes (TABLE 1). **a** | Applied to at-risk species, population genomics powered by next-generation sequencing (NGS) can uncover and quantify variation at a large number of single-nucleotide polymorphisms (SNPs). The resulting analyses can help inform the severity and duration of bottlenecks, assess genetic threats to population viability such as inbreeding depression and troubleshoot the potential causes of population crashes. Moreover, the availability of genomic sequences for at-risk species allows biologists to undertake functional genomic, transcriptomic and metagenomics studies that can lead to more directed intervention to circumvent declines (FIGS 3, 4 and 5; TABLE 1). **b** | Whole-genome sequencing of populations leads to the identification of SNPs that differ between species, which can facilitate the development of rapid diagnostic platforms that can be applied to improve bee biosecurity by limiting the spread of invasive and economically damaging strains of bees, or to preserve the genetic purity of native locally adapted bee populations that are at risk of extinction via hybridization with other populations. **c** | Population genomics can also identify SNPs that differ with specific traits (such as resilience), which in turn can be used for high-throughput breeding assisted by marker selection. GWAS, genome-wide association study.

An interesting outcome of metagenomic studies of virus populations in bees has been the identification of several viruses shared among populations of co-foraging bee species and between bees and plants<sup>11</sup>. Initial PCR-based screens for known honeybee-infecting viruses in other bee species found that these viruses can be shared by pollen collected from flowers<sup>97</sup>; subsequent metagenomic studies demonstrated that multiple types of virus are shared among bee species<sup>93,98</sup>. Thus, although pollinator plantings create important nutritional resources for bees, they may also unintentionally exacerbate disease transfer if they create common foraging hubs in areas with limited natural forage<sup>99</sup>. Further study is needed to determine whether plant viruses found in metagenome sequences generated from bee tissue are derived from contaminating plant material or whether bees are serving as hosts; however, there is evidence that these viruses can replicate in bees<sup>100</sup>. As bees forage broadly, they can ‘bioaccumulate’ plant viruses in a landscape, allowing efficient detection of these viruses: indeed, researchers in Australia identified 28 plant viruses from the metagenomes of 9 pooled *A. mellifera* samples in 2013 and 2014. One of these viruses — the exotic cucumber green mottle mosaic virus — was detected only in plants by Australian biosecurity programmes at dates later than dates on which bees were sampled<sup>101</sup>. For microorganisms with known pathogenic effects (such as DWV), sequences generated by NGS have allowed researchers to evaluate population genomics of viral strains, including recombination events, across individual hosts and treatment groups<sup>102,103</sup>.

In the honeybee, removal of the gut microbiota through antibiotic use negatively impacts immune function, insulin signalling and metabolism<sup>104,105</sup>. To understand how the microbiota influences bee health, it is necessary to characterize this microbial community through metagenomics (FIG. 4b). In these studies, researchers either sequence genetic material from the entire community or use a metabarcoding approach, in which amplicons from specific regions of the genome (such as the 16S ribosomal RNA gene) are sequenced<sup>106</sup>. Although much remains to be discovered about the

function of the gut microbiota in honeybees, it is clear that changes in the composition of the gut microbiota — caused by extrinsic stressors such as exposure to pesticides and antibiotics — can make honeybees less resilient to pathogen infections<sup>107,108</sup>. Moreover, some rare members of the honeybee’s gut microbiota, such as *Serratia marcescens*, can themselves become pathogenic when the gut microbiota is disturbed by stressors<sup>109</sup>. The mechanisms underlying the protective effects of the bee’s gut microbiota are only beginning to be studied<sup>105,110</sup>. Once aspects of the gut microbiota that promote bee health are well understood, it may be possible to directly improve bee health by the application of probiotics<sup>111</sup>. It may also be possible to use shifts in the composition of the gut microbiota as a diagnostic tool for identifying stressors affecting bee populations.

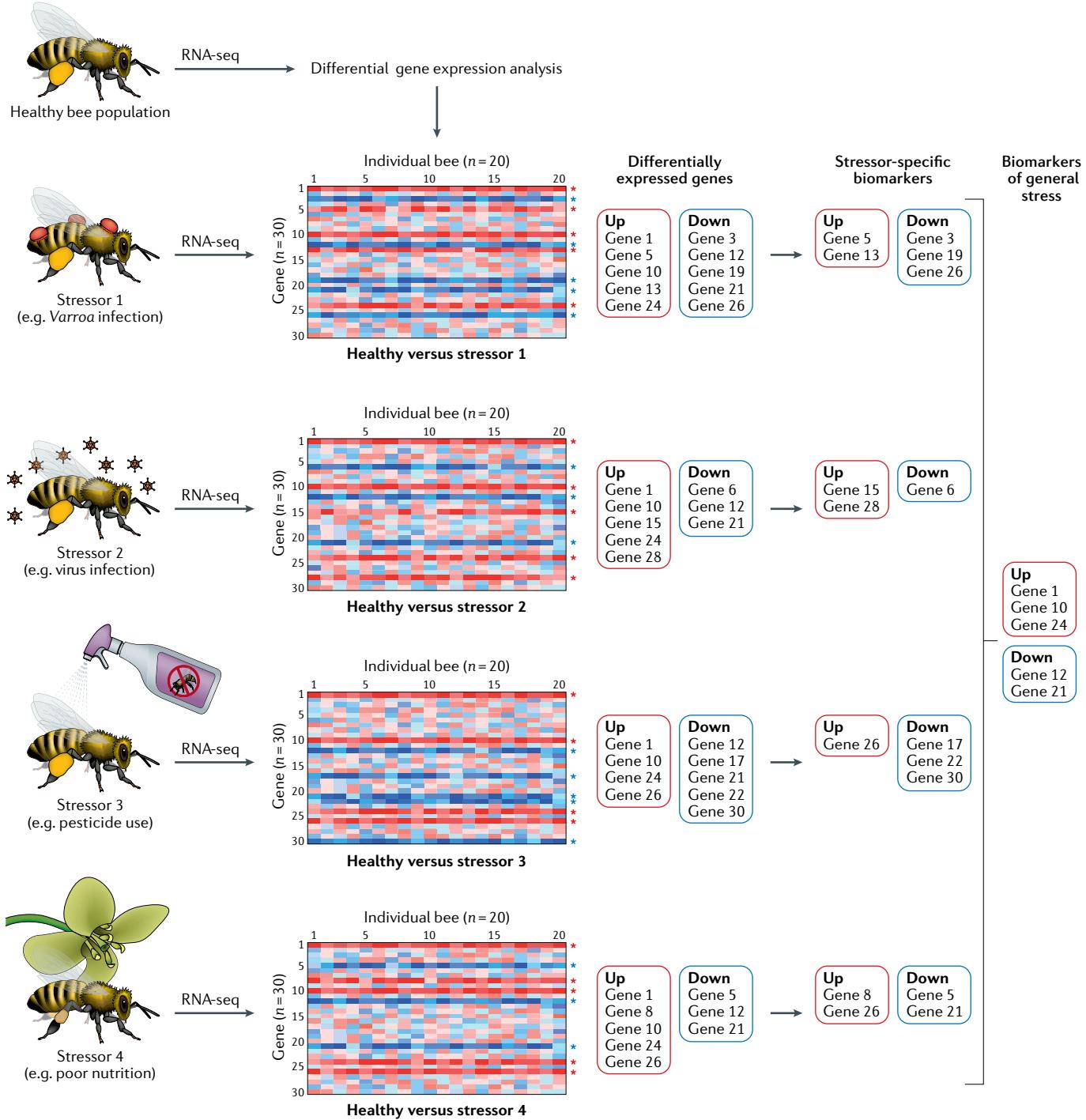
**Revealing the nutritional needs of bees.** Plant–pollinator interactions represent an ancient mutualism, where pollinators obtain nutritional resources from plants and plants use pollinators to transfer pollen and reproduce<sup>112</sup>. For bees, the nectar collected from flowering plants is their primary source of carbohydrates and pollen is their primary source of proteins and lipids<sup>77</sup>. Plant species differ dramatically in the nutritional quality and quantity of the nectar and pollen that they produce, and bees forage preferentially on different plant species to meet their nutritional needs<sup>113</sup>. Thus, understanding which plant species are preferred by a particular bee species is critical for understanding bee health. However, it can be challenging to identify the preferred plant species in a given environment because the vast majority of bee species are polylectic (that is, they forage on a broad variety of plant species<sup>114</sup>) and can forage over large distances (several kilometres from their nest site in the case of honeybees<sup>115,116</sup>). It is possible to extract the DNA from the pollen collected by bees and use metabarcoding to identify the original plant sources of the pollen<sup>117–119</sup> (FIG. 4c). This is an extremely versatile approach, as pollen may be collected from the bodies of foraging bees<sup>120</sup>, including from the corbicula loads of honeybees and bumblebees<sup>117,118</sup> or bees in museums<sup>121</sup>, or pollen can be sampled from the nest<sup>122</sup>, including from nectar and honey stores that contain small quantities of pollen grains<sup>123</sup>. Unlike traditional microscopy-based palynology, metabarcoding is amenable to high-throughput, parallel-sample processing and thus can be used for large sample sets. However, fully leveraging the power of metabarcoding approaches requires a comprehensive database of plant species sequences, ideally curated to provide information on plant distributions and phenologies, and these databases are well developed only in certain regions and for specific plant groups<sup>124</sup>. Moreover, current metabarcoding approaches rely on a small set of biomarkers, which limits identification, discrimination and quantification of plant species<sup>125</sup>. Decreasing sequencing costs are allowing researchers to explore alternative sequencing approaches (such as genome skimming, to obtain near-complete sequences from mitochondria, plastids and ribosomal RNA), which will greatly facilitate pollen identification through genomic approaches<sup>126</sup>. In conjunction with laboratory studies that experimentally

#### Corbicula loads

Pollen loads that are carried on the modified pollen basket (that is, the corbicula) of some bees of the subfamily Apinae.

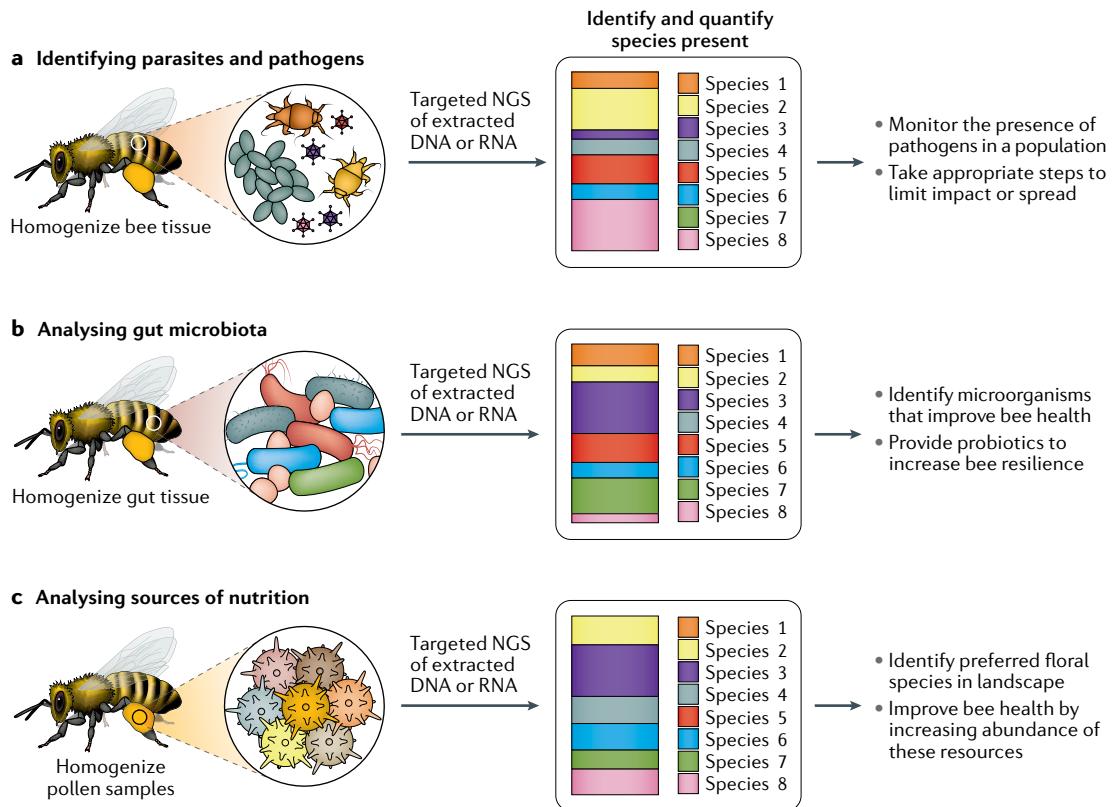
#### Palynology

The taxonomic analysis of pollen grains.



**Fig. 3 | Transcriptomic approaches to study, diagnose and improve bee health.** Exposure to a stressor (such as a parasite, pathogen, pesticide or poor nutrition) triggers a specific transcriptional response in a bee. We demonstrate how these transcriptional responses, revealed by RNA sequencing (RNA-seq), can be used to identify general and stressor-specific biomarkers that can be applied to monitor and manage bee health. In this hypothetical example, we profile gene expression in either whole bodies or specific tissues of healthy control bees and bees exposed to Varroa mites (stressor 1), a virus (stressor 2), a pesticide (stressor 3) or poor nutrition (stressor 4). The heatmaps represent expression of 30 genes in the stressed experimental groups relative to the unmanipulated controls. Statistical analyses are used to identify genes that are differentially upregulated

(red asterisks) or downregulated (blue asterisks) in bees exposed to different stressors relative to controls. These changes in gene expression can be used to better understand the molecular and physiological mechanisms through which a stressor undermines bee health, which can then lead to new approaches for increasing resilience. A meta-analysis of these gene lists is then used to identify gene expression patterns that are diagnostic of specific or general stressors. For example, downregulation of gene 6 indicates virus exposure, while upregulation of gene 8 indicates poor nutrition. Upregulation of gene 24 indicates that bees have been exposed to a stressor that influences their health relative to control bees. This knowledge can be applied in routine biomonitoring of bee health or to investigate specific cases of bee death.



**Fig. 4 | Metagenomic approaches to study and improve bee health.** Whole-genome sequencing (metagenomics) or targeted gene sequencing (metabarcoding) can be used to identify and quantify the complex communities of organisms with which bees and other pollinators interact. **a** | Metagenomic approaches can help identify and monitor the pathogens and parasites associated with diseases, thereby allowing beekeepers and others to take steps to mitigate the spread and impacts of these diseases. **b** | Metagenomic approaches have identified the microbial communities present in the bee digestive tract or stored food products in the nest. Several studies have highlighted the protective effects of the bee gut microbiota<sup>104,105,111</sup>, which can potentially be developed into probiotic dietary supplements to improve bee health. **c** | Metagenomic approaches can be used to identify the plants that bees preferentially forage on in urban, agricultural and natural landscapes at different times of year. This information can be used to improve floral enhancement and restoration schemes, resulting in improved nutrition for bees. NGS, next-generation sequencing.

manipulate diet and examine its consequences on bee health, these tools will allow researchers to identify the plants that provide key nutritional resources for diverse bee species throughout the growing season, and to improve recommendations for pollinator forage and habitat restoration and management.

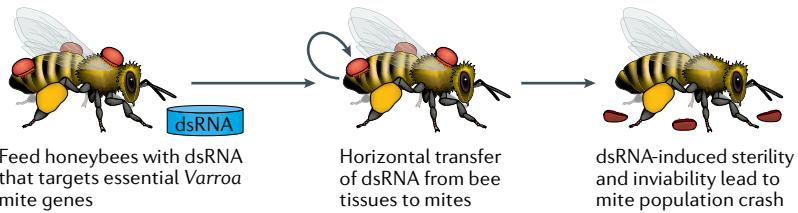
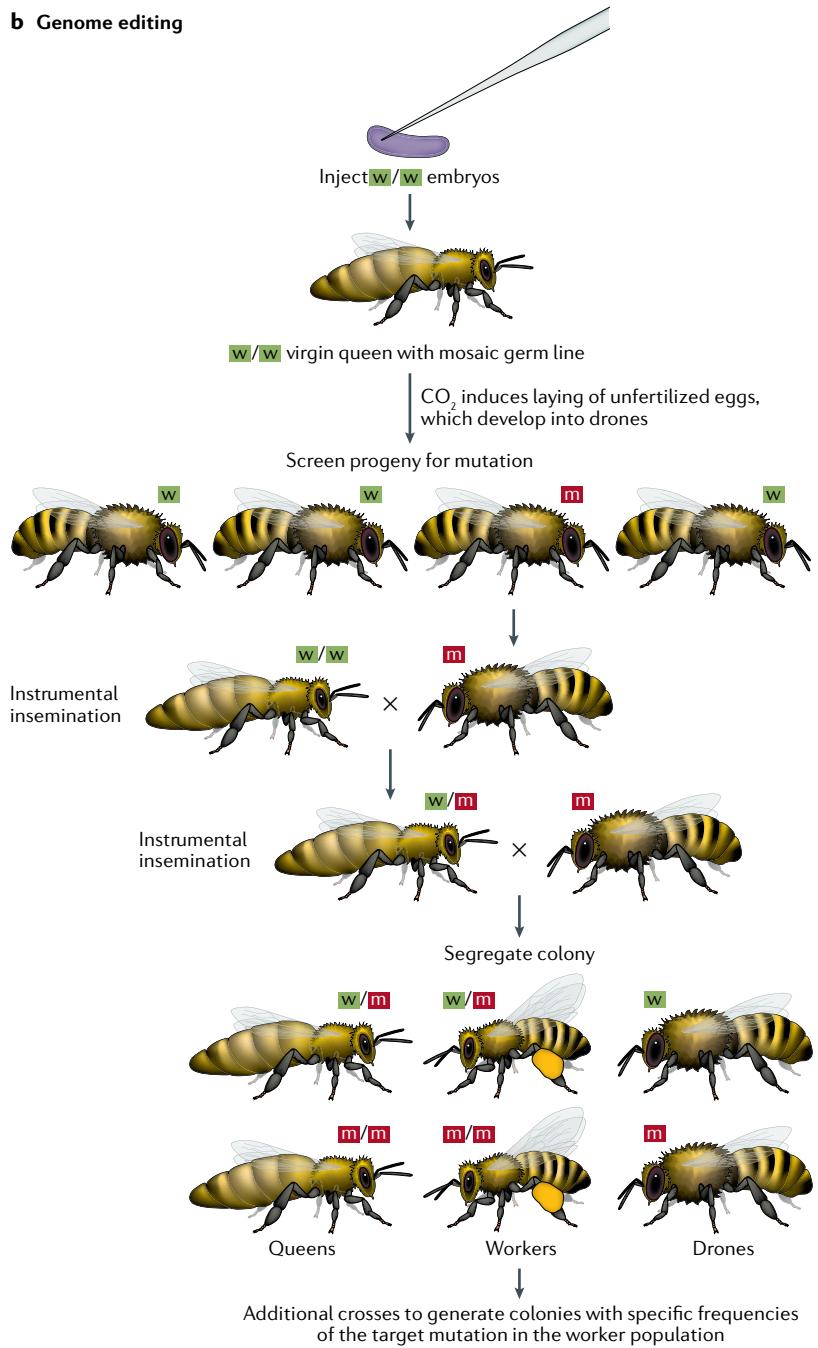
### Manipulating bee genomes

Once genes regulating key traits, either of the bee or of its parasites and pathogens, have been identified, tools that modify the function of these genes can be used to improve the health of bees and/or reduce the levels of their associated pathogens and parasites. RNAi approaches, in which double-stranded RNA with sequences that match target genes in the parasite, host or pathogen, have been used successfully to reduce levels of *Varroa* mites<sup>127</sup>, *Nosema* microsporidia<sup>128</sup> and viral populations<sup>129</sup> in honeybees in controlled studies (FIG. 5a). However, there are substantial challenges and limitations with use of RNAi as a therapeutic treatment in the field, including the presence of multiple species and strains of pathogens and parasites, difficulties with introducing double-stranded RNA at levels high enough

to be effective, low efficiency, high cost and off-target effects, which may serve to compromise the health of bees or other species<sup>129</sup>.

Stable changes in expression of bee genes can be obtained through genome editing (FIG. 5b). With the development of CRISPR-Cas9 systems, genome editing has become feasible in a wide array of species, including insects<sup>130</sup>. The gene encoding the major royal jelly protein, which is important for the production of brood food in adult worker honeybees, has been successfully edited in both queen and male honeybees as a proof of principle to demonstrate the effectiveness of this approach for mutating genes that are not required for development<sup>131</sup>. Editing of the honeybee *Dsx* and *Fem* genes showed that *Dsx* controls the size of reproductive organs regardless of nutrition, whereas *Fem* modulates the size of reproductive organs in response to nutrition, a key process underlying caste differentiation in honeybees<sup>132</sup>. In both of these studies, the researchers injected very young embryos (less than 3 hours after oviposition), thereby creating mosaic individuals. In the first study, one of the edited queens produced mutant male offspring, indicating that genome editing

occurred in the developing gametes as well as in somatic tissue. In the second study, more than 70% of the injected embryos possessed only the edited allele, suggesting a high degree of efficiency. An alternative approach, which has been used successfully in mosquitoes, is to

**a RNAi****b Genome editing**

target CRISPR–Cas9 to developed oocytes of reproductive females, thereby targeting the germ line directly<sup>133</sup>. Such a system would be technically easier, as it would eliminate the need for specialized equipment for embryo injection and the modified larvae can develop under normal conditions within the colony, bypassing the need for *in vitro* rearing. Regardless of the approach used, once a mutant is created, additional crosses and rearing are needed to obtain a diploid mutant strain (FIG. 5b). One obvious complication for the application of genome editing to improve honeybee health is that colony fitness is largely influenced by within-colony genetic diversity<sup>134</sup>; creating a colony of bees carrying homozygous or heterozygous CRISPR mutations will inadvertently lead to reduced levels of genetic diversity and likely reduced fitness.

Genome editing technology is clearly feasible in bee species and can be valuable for functional evaluation of candidate genes. However, rigorous debate about the ethics and risk of such modifications is needed if this approach were ever to be considered for genetically modifying free-living populations. The case of the Africanized honeybee provides a poignant example of what can go wrong when experimenting with bee genetics: in this case, honeybees from Africa were imported to South America in an effort to improve the performance of managed honeybees in this region, but the behavioural characteristics of African bees (defensiveness, migratory behaviour, rapid reproduction) that made them well adapted to African environments also made them difficult and dangerous to manage, and allowed them to spread throughout South America and Central America and into North America<sup>29</sup>. It is important to recognize that genes function within large networks, and thus modifying gene function can have

**Fig. 5 | Methods for manipulating gene function in bees.**

**a |** RNA interference (RNAi) can be used to knockdown specific gene targets in either the host bee genome or in the genome of the parasite or pathogen. Double-stranded RNA (dsRNA) specific to the targeted gene (or genes) is fed to the bee and can be horizontally transferred from the bee to the pathogen. The dsRNA stimulates the RNAi pathway, which results in reduced levels of the target RNA. This can either impair viability or fertility of the parasite or pathogen (as in the case of genes targeting Varroa<sup>127</sup>) or improve the immune function of the bee (if the target is a host gene that inhibits immune function<sup>128</sup>). The RNAi response is typically temporary, necessitating frequent dosing with dsRNA. **b |** CRISPR–Cas9-based gene editing results in permanent editing of the target gene in the genome. Typically, wild-type (green 'w') embryos are injected with the required gene editing constructs and are then reared as queens with a mosaic germ line. These queens are induced to lay unfertilized eggs, and the resulting drones are screened for the CRISPR–Cas9-induced mutation (red 'm'). By instrumental insemination, wild-type queens can be mated with mutant drones, leading to an F1 heterozygous queen. Repeated selection and breeding is necessary to generate a line of bees in which both copies of a gene are mutated, which can be challenging in a system in which rearing and breeding are not routine. Furthermore, only genes that are not required for development or normal function can be mutated by this approach.

**Pleiotropic**

Pertaining to genetic loci that affect two or more phenotypic traits.

pleiotropic effects that may unexpectedly undermine bee health or may lead to undesirable traits that make management difficult. Furthermore, although it may be possible to increase the resilience of a particular bee stock to a specific stressor, the native bee community — which often surpasses honeybees in terms of pollination efficiency — will continue to be sensitive to these stressors, and thus these genetic modification approaches would not support the larger ecological community. Given the complex interconnectedness of plant–pollinator communities<sup>112</sup>, resilience and food security require the maintenance of diverse communities of bee species.

**Conclusions**

Genomics is revolutionizing the field of pollinator conservation. Population genomics can be used to identify at-risk or resilient populations, thereby informing and enhancing conservation efforts. Transcriptomic studies reveal the multitude of stressors that undermine bee health, as well as the intricate mechanisms by which bees combat these stressors. Metagenomics provides a high-throughput strategy to characterize the diverse and complex communities of microorganisms that interact with bees and to explore bees' foraging preferences and dietary requirements. Combined, these approaches can

generate an omics toolkit that can be used to improve bee health through selection for resilient traits, rapid and comprehensive diagnostic approaches, development of dietary probiotics or targeted floral restoration schemes, and potential genetic modifications to increase bee resilience. However, it is very clear that landscape and environmental conditions, including weather, landscape configuration, availability of floral resources and pesticide use, interact with genotype and greatly influence the health and survival of managed and wild bee populations<sup>135</sup>. Moreover, with more than 20,000 described species of bees<sup>136</sup>, it is possible to conduct detailed studies on only a handful of species, and many of the genomic approaches that can serve to mitigate the effects of stressors are feasible only in managed bees. Thus, while genomics can be used to enhance the intrinsic resilience of bee populations, it is vital that we continue to work towards reducing the diverse extrinsic stressors that affect bee health. We believe that while omic tools can be effectively leveraged to support the health of bees, they should be embedded within holistic programmes that consider the health of the overall ecological community, including other animal and plant species.

Published online 25 February 2020

1. Steffan-Dewenter, I., Potts, S. G. & Packer, L. Pollinator diversity and crop pollination services are at risk. *Trends Ecol. Evol.* **20**, 651–652 (2005).
2. Klein, A.-M. et al. Importance of pollinators in changing landscapes for world crops. *Proc. R. Soc. Lond. B Biol. Sci.* **274**, 303–313 (2007).
3. Williams, P., Colla, S. & Xie, Z. Bumblebee vulnerability: common correlates of winners and losers across three continents. *Conserv. Biol.* **23**, 931–940 (2009).
4. Grixti, J. C., Wong, L. T., Cameron, S. A. & Favret, C. Decline of bumble bees (*Bombus*) in the North American Midwest. *Biol. Conserv.* **142**, 75–84 (2009).
5. Colla, S. R., Otterstatter, M. C., Geiger, R. J. & Thomson, J. D. Plight of the bumblebee: pathogen spillover from commercial to wild populations. *Biol. Conserv.* **129**, 461–467 (2006).
6. Allen-Wardell, G. et al. The potential consequences of pollinator declines on the conservation of biodiversity and stability of food crop yields. *Conserv. Biol.* **12**, 8–17 (1998).
7. Gallai, N., Salles, J., Settele, J. & Vaissière, B. E. Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. *Ecol. Econ.* **68**, 810–821 (2009).
8. Hunt, G. J. & Page, R. E. J. Linkage map of the honey bee, *Apis mellifera*, based on RAPD markers. *Genetics* **139**, 1371–1382 (1995).
9. Hunt, G. J., Page, R. E. J., Fondrk, M. K. & Dullum, C. J. Major quantitative trait loci affecting honey bee foraging behavior. *Genetics* **141**, 1537–1545 (1995).
10. Dogantzi, K. A. & Zayed, A. Recent advances in population and quantitative genomics of honey bees. *Curr. Opin. Insect Sci.* **31**, 93–98 (2019).
11. Grozinger, C. M. & Flenniken, M. L. Bee viruses: ecology, pathogenicity, and impacts. *Annu. Rev. Entomol.* **64**, 205–226 (2019).
12. Lozier, J. D. & Zayed, A. Bee conservation in the age of genomics. *Conserv. Genet.* **18**, 713–729 (2017).
13. Weinstock, G. M. et al. Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* **443**, 931–949 (2006).
14. Zayed, A. Bee genetics and conservation. *Apidologie* **40**, 237–262 (2009).
15. Laczner, A. R. & Colla, S. R. A systematic review of the nesting and overwintering habitat of bumble bees globally. *J. Insect Conserv.* <https://doi.org/10.1007/s10841-019-00173-7> (2019).
16. Darvill, B., Knight, M. E. & Goulson, D. Use of genetic markers to quantify bumblebee foraging range and nest density. *Oikos* **107**, 471–478 (2004).
17. Carvell, C. et al. Bumblebee family lineage survival is enhanced in high-quality landscapes. *Nature* **543**, 547–549 (2017). **This study provides an excellent example of how population genetics can be combined with landscape data to understand the habitat requirements of at-risk pollinators.**
18. Jha, S. & Kremen, C. Urban land use limits regional bumble bee gene flow. *Mol. Ecol.* **22**, 2483–2495 (2013).
19. Kent, C. F. et al. Conservation genomics of the declining North American bumblebee *Bombus terricola* reveals inbreeding and selection on immune genes. *Front. Genet.* **9**, 316 (2018).
20. Beadle, K. et al. Genomic insights into neonicotinoid sensitivity in the solitary bee *Osmia bicoloris*. *PLoS Genet.* **15**, e1007903 (2019).
21. Manjón, C. et al. Unravelling the molecular determinants of bee sensitivity to neonicotinoid insecticides. *Curr. Biol.* **28**, 1137–1143 (2018).
22. Hayward, A. et al. The leafcutter bee, *Megachile rotundata*, is more sensitive to N-cyanoamidine neonicotinoid and butenolide insecticides than other managed bees. *Nat. Ecol. Evol.* **3**, 1521–1524 (2019).
23. Sponsler, D. B. et al. Pesticides and pollinators: a socioecological synthesis. *Sci. Total. Env.* **662**, 1012–1027 (2019).
24. Nazzi, F. & Le Conte, Y. Ecology of Varroa destructor, the major ectoparasite of the western honey bee, *Apis mellifera*. *Annu. Rev. Entomol.* **61**, 417–432 (2016).
25. Martin, S. J. & Brettell, L. E. Deformed wing virus in honeybees and other insects. *Annu. Rev. Virol.* <https://doi.org/10.1146/annurev-virology-092818-015700> (2019).
26. Wilfert, L. et al. Deformed wing virus is a recent global epidemic in honeybees driven by Varroa mites. *Sci.* **351**, 594–597 (2016).
27. Arbetman, M. O., Meeus, I., Morales, C. L., Aizen, M. A. & Smagghe, G. Alien parasite hitchhikes to Patagonia on invasive bumblebee. *Biol. Invasion* **15**, 489–494 (2012).
28. Harpur, B. A., Minaei, S., Kent, C. F. & Zayed, A. Admixture increases diversity in managed honey bees. Reply to De la Rúa et al. (2013). *Mol. Ecol.* **22**, 3211–3215 (2013).
29. Schneider, S. S., DeGrandi-Hoffman, G. & Smith, D. R. The African honey bee: factors contributing to a successful biological invasion. *Annu. Rev. Entomol.* **49**, 351–376 (2004).
30. Zayed, A. & Whitfield, C. W. A genome-wide signature of positive selection in ancient and recent invasive expansions of the honey bee *Apis mellifera*. *Proc. Natl. Acad. Sci. USA* **105**, 3421–3426 (2008).
31. Francoy, T. M., Gonçalves, L. S. & De Jong, D. Rapid morphological changes in populations of hybrids between Africanized and European honey bees. *Genet. Mol. Res.* **11**, 3349–3356 (2012).
32. Harpur, B. A. et al. Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proc. Natl. Acad. Sci. USA* **111**, 2614–2619 (2014).
33. Chapman, N. C. et al. A SNP test to identify Africanized honeybees via proportion of 'African' ancestry. *Mol. Ecol. Resour.* **15**, 1346–1355 (2015).
34. Harpur, B. A. et al. Assessing patterns of admixture and ancestry in Canadian honey bees. *Insect Soc.* **62**, 479–489 (2015).
35. Requier, F. et al. The conservation of native honey bees is crucial. *Trends Ecol. Evol.* <https://doi.org/10.1016/j.tree.2019.04.008> (2019).
36. Muñoz, I. et al. SNPs selected by information content outperform randomly selected microsatellite loci for delineating genetic identification and introgression in the endangered dark European honeybee (*Apis mellifera mellifera*). *Mol. Ecol. Resour.* **17**, 783–795 (2017).
37. Henrikus, D. et al. Developing reduced SNP assays from whole-genome sequence data to estimate introgression in an organism with complex genetic patterns, the Iberian honeybee (*Apis mellifera iberiensis*). *Evolut. Appl.* **11**, 1270–1282 (2018).
38. Harbo, J. R. & Harris, J. W. Heritability in honey bees (Hymenoptera: Apidae) of characteristics associated with resistance to *Varroa jacobsoni* (Mesostigmata: Varroidae). *J. Econ. Entomol.* **92**, 261–265 (1999).
39. Koffler, S., Kleinert, A. d. M. P. & Jaffé, R. Quantitative conservation genetics of wild and managed bees. *Conserv. Genet.* **18**, 689–700 (2017).
40. Stanimirovic, Z., Stevanovic, J., Aleksić, N. & Velibor, S. Heritability of grooming behaviour in grey honey bees (*Apis mellifera Carnica*). *Acta Vet.* **60**, 313–323 (2009).
41. Decanini, L. I., Collins, A. M. & Evans, J. D. Variation and heritability in immune gene expression by diseased honeybees. *J. Hered.* **98**, 195–201 (2007).
42. Wilfert, L., Gadau, J., Baer, B. & Schmid-Hempel, P. Natural variation in the genetic architecture of a host-parasite interaction in the bumblebee *Bombus terrestris*. *Mol. Ecol.* **16**, 1327–1339 (2007).
43. Wilfert, L., Gadau, J. & Schmid-Hempel, P. The genetic architecture of immune defense and reproduction in male *Bombus terrestris* bumblebees. *Evolution* **61**, 804–815 (2007).

44. Nino, E. L. & Cameron Jasper, W. Improving the future of honey bee breeding programs by employing recent scientific advances. *Curr. Opin. Insect Sci.* **10**, 163–169 (2015).
45. Ibrahim, A. G., Reuter, M. & Spivak, M. Field trials of honey bee colonies bred for mechanisms of resistance against *Varroa destructor*. *Apidologie* **38**, 67–76 (2006).
46. Spivak, M. & Gilliam, M. 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**, 124–134 (1998).
47. Baudry, E. et al. Relatedness among honeybees (*Apis mellifera*) of a drone congregation. *Proc. Biol. Sci.* **265**, 2009–2014 (1998).
48. Behrens, D. et al. Three QTL in the honey bee *Apis mellifera* L. suppress reproduction of the parasitic mite *Varroa destructor*. *Ecol. Evol.* **1**, 451–458 (2011).
49. Behrens, D. & Moritz, R. F. QTL-mapping of individual resistance against American foulbrood in haploid honeybee drone larvae (*Apis mellifera*). *Apidologie* **45**, 409–417 (2014).
50. Guzman-Novoa, E., Hunt, G. J., Uribe, J. L., Smith, C. & Arechavala-Velasco, M. E. Confirmation of QTL effects and evidence of genetic dominance of honeybee defensive behavior: results of colony and individual behavioral assays. *Behav. Genet.* **32**, 95–102 (2002).
51. McCarthy, M. I. et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat. Rev. Genet.* **9**, 356–369 (2008).
52. Korte, A. & Farlow, A. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* **9**, 29 (2013).
53. Wallberg, A., Schöning, C., Webster, M. T. & Hasselmann, M. Two extended haplotype blocks are associated with adaptation to high altitude habitats in East African honey bees. *PLoS Genet.* **13**, e1006792 (2017).
54. Wallberg, A., Pirk, C. W., Allsopp, M. H. & Webster, M. T. Identification of multiple loci associated with social parasitism in honeybees. *PLoS Genet.* **12**, e1006097 (2016).
55. Wragg, D. et al. Whole-genome resequencing of honeybee drones to detect genomic selection in a population managed for royal jelly. *Sci. Rep.* **6**, 27168 (2016).
56. Harpur, B. A. et al. Integrative genomics reveals the genetics and evolution of the honey bee's social immune system. *Genome Biol. Evol.* **11**, 937–948 (2019). **In this study, the combined use of artificially selected lines and population genomics allows the researchers to unravel the complex genetics of social immunity in honeybees.**
57. Avalos, A. et al. A soft selective sweep during rapid evolution of gentle behaviour in an Africanized honeybee. *Nat. Commun.* **8**, 1550 (2017).
58. Broeckx, B. J. G. et al. Honey bee predisposition of resistance to ubiquitous mite infestations. *Sci. Rep.* **9**, 7794 (2019).
59. Kruglyak, L. The road to genome-wide association studies. *Nat. Rev. Genet.* **8**, 314–318 (2008).
60. Wu, Y., Zheng, Z., Visscher, P. M. & Yang, J. Quantifying the mapping precision of genome-wide association studies using whole-genome sequencing data. *Genome Biol.* **18**, 86 (2017).
61. Guarna, M. M. et al. Peptide biomarkers used for the selective breeding of a complex polygenic trait in honey bees. *Sci. Rep.* **7**, 8381 (2017). **This study provides a clear demonstration of the use of marker-assisted selection in honeybees.**
62. Zanni, V., Galbraith, D. A., Annoscia, D., Grozinger, C. M. & Nazzi, F. Transcriptional signatures of parasitization and markers of colony decline in *Varroa*-infested honey bees (*Apis mellifera*). *Insect Biochem. Mol. Biol.* **87**, 1–13 (2017).
63. Brutscher, L. M., Daughenbaugh, K. F. & Fleniken, M. L. Virus and dsRNA-triggered transcriptional responses reveal key components of honey bee antiviral defense. *Sci. Rep.* **7**, 6448 (2017).
64. Shi, T. F., Wang, Y. F., Liu, F., Qi, L. & Yu, L. S. Sublethal effects of the neonicotinoid insecticide thiamethoxam on the transcriptome of the honey bees (Hymenoptera: Apidae). *J. Econ. Entomol.* **110**, 2283–2289 (2017).
65. Corby-Harris, V., Jones, B. M., Walton, A., Schwan, M. R. & Anderson, K. E. Transcriptional markers of sub-optimal nutrition in developing *Apis mellifera* nurse workers. *BMC Genom.* **15**, 134 (2014).
66. Azzouz-Olden, F., Hunt, A. & DeGrandi-Hoffman, G. Transcriptional response of honey bee (*Apis mellifera*) to differential nutritional status and *Nosema* infection. *BMC Genom.* **19**, 628 (2018).
67. Amsalem, E., Galbraith, D. A., Chaani, J., Teal, P. E. & Grozinger, C. M. Conservation and modification of genetic and physiological toolkits underpinning diapause in bumble bee queens. *Mol. Ecol.* **24**, 5596–5615 (2015).
68. Torson, A. S. et al. Physiological responses to fluctuating temperatures are characterized by distinct transcriptional profiles in a solitary bee. *J. Exp. Biol.* **220**, 3372–3380 (2017).
69. Durant, D. R., Berens, A. J., Toth, A. L. & Rehan, S. M. Transcriptional profiling of overwintering gene expression in the small carpenter bee, *Ceratina calcarata*. *Apidologie* **47**, 572–582 (2016).
70. Ankley, G. T. et al. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Env. Toxicol. Chem.* **29**, 730–741 (2010).
71. Barron, A. B. Death of the bee hive: understanding the failure of an insect society. *Curr. Opin. Insect Sci.* **10**, 45–50 (2015).
72. Amdam, G. V., Fennern, E. & Havukainen, H. in *Honey Bee Neurobiology and Behavior* (eds Galizia, C., Eisenhardt, D., & Guirfa, M.) 17–29 (Springer, 2012).
73. Alaux, C. et al. Measuring biological age to assess colony demographics in honeybees. *PLoS One* **13**, e0209192 (2018). **In this study, molecular markers are used to measure the 'biological age' of individual honeybees and evaluate the impacts of different stressors and management regimes.**
74. Dainat, B., Evans, J. D., Chen, Y. P., Gauthier, L. & Neumann, P. Predictive markers of honey bee colony collapse. *PLoS One* **7**, e32151 (2012).
75. Ricigliano, V. A. et al. Honey bee colony performance and health are enhanced by apiary proximity to US Conservation Reserve Program (CRP) lands. *Sci. Rep.* **9**, 4894 (2019).
76. Schmehl, D. R., Teal, P. E., Frazier, J. L. & Grozinger, C. M. Genomic analysis of the interaction between pesticide exposure and nutrition in honey bees (*Apis mellifera*). *J. Insect Physiol.* **71**, 177–190 (2014).
77. Vaudo, A. D., Tooker, J. F., Grozinger, C. M. & Patch, H. M. Bee nutrition and floral resource restoration. *Curr. Opin. Insect Sci.* **10**, 133–141 (2015).
78. Mao, W., Schuler, M. A. & Berenbaum, M. R. Honey constituents up-regulate detoxification and immunity genes in the western honey bee *Apis mellifera*. *Proc. Natl Acad. Sci. USA* **110**, 8842–8846 (2013).
79. Di Prisco, G. et al. Neonicotinoid clothianidin adversely affects insect immunity and promotes replication of a viral pathogen in honey bees. *Proc. Natl Acad. Sci. USA* **110**, 18466–18471 (2013).
80. De Smet, L. et al. BeeDoctor, a versatile MLPA-based diagnostic tool for screening bee viruses. *PLoS One* **7**, e47953 (2012).
81. Massuti, B., Sanchez, J. M., Hernando-Trancho, F., Karachaliou, N. & Rosell, R. Are we ready to use biomarkers for staging, prognosis and treatment selection in early-stage non-small-cell lung cancer? *Transl Lung Cancer Res.* **2**, 208–221 (2013).
82. Adnane, M., Kelly, P., Chapwanya, A., Meade, K. G. & O'Farrelly, C. Improved detection of biomarkers in cervico-vaginal mucus (CVM) from postpartum cattle. *BMC Vet. Res.* **14**, 297 (2018).
83. Brutscher, L. M., Daughenbaugh, K. F. & Fleniken, M. L. Antiviral defense mechanisms in honey bees. *Curr. Opin. Insect Sci.* **10**, 71–82 (2015).
84. Kunieda, T. et al. Carbohydrate metabolism genes and pathways in insects: insights from the honey bee genome. *Insect Mol. Biol.* **15**, 563–576 (2006).
85. Claudiano, C. et al. A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee. *Insect Mol. Biol.* **15**, 615–636 (2006).
86. Doublet, V. et al. Unity in defence: honeybee workers exhibit conserved molecular responses to diverse pathogens. *BMC Genom.* **18**, 207 (2017). **This article describes a meta-analysis of several transcriptomic studies that identified genes that responded either to several or to only one of the pathogens or parasites tested, indicating that host transcriptional responses can be used to diagnose the particular pathogen or parasite that is infecting a bee.**
87. Wheeler, M. M. & Robinson, G. E. Diet-dependent gene expression in honey bees: honey vs. sucrose or high fructose corn syrup. *Sci. Rep.* **4**, 5726 (2014).
88. Alaux, C., Dantec, C., Parrinello, H. & Le Conte, Y. Nutrigenomics in honey bees: digital gene expression analysis of pollen's nutritive effects on healthy and varroa-parasitized bees. *BMC Genom.* **12**, 12–496 (2011).
89. Wu, M. C., Chang, Y. W., Lu, K. H. & Yang, E. C. Gene expression changes in honey bees induced by sublethal imidacloprid exposure during the larval stage. *Insect Biochem. Mol. Biol.* **88**, 12–20 (2017).
90. Wang, L. et al. Metabolomics-based biomarker discovery for bee health monitoring: a proof of concept study concerning nutritional stress in *Bombus terrestris*. *Sci. Rep.* **9**, 11423 (2019).
91. Raymann, K. & Moran, N. A. The role of the gut microbiome in health and disease of adult honey bee workers. *Curr. Opin. Insect Sci.* **26**, 97–104 (2018).
92. Dharampal, P. S., Carlson, C., Currie, C. R. & Steffan, S. A. Pollen-borne microbes shape bee fitness. *Proc. Biol. Sci.* **286**, 20182894 (2019).
93. Galbraith, D. A. et al. Investigating the viral ecology of global bee communities with high-throughput metagenomics. *Sci. Rep.* **8**, 8879 (2018). **In this study, metagenomics is used to characterize the viruses infecting 12 bee species in populations collected from nine countries.**
94. Ryabov, E. V. et al. Recent spread of *Varroa destructor* virus-1, a honey bee pathogen, in the United States. *Sci. Rep.* **7**, 17447 (2017).
95. Ray, A. M. et al. Distribution of recently identified bee-infecting viruses in managed honey bee (*Apis mellifera*) populations in the United States. *Apidologie* (in the press).
96. Corman, R. S. Relative abundance of deformed wing virus, *Varroa destructor* virus 1, and their recombinants in honey bees (*Apis mellifera*) assessed by kmer analysis of public RNA-Seq data. *J. Invertebr. Pathol.* **149**, 44–50 (2017).
97. Singh, R. et al. RNA viruses in hymenopteran pollinators: evidence of inter-taxa virus transmission via pollen and potential impact on non-*Apis* hymenopteran species. *PLoS One* **5**, e14357 (2010).
98. Schoonaerde, K., Smagghe, G., Francis, F. & de Graaf, D. C. Study of the metatranscriptome of eight social and solitary wild bee species reveals novel viruses and bee parasites. *Front. Microbiol.* **9**, 177 (2018).
99. Piot, N. et al. Establishment of wildflower fields in poor quality landscapes enhances micro-parasite prevalence in wild bumble bees. *Oecologia* **189**, 149–158 (2019).
100. Li, J. L. et al. Systemic spread and propagation of a plant-pathogenic virus in European honeybees, *Apis mellifera*. *mBio* **5**, e00898-13 (2014).
101. Roberts, J. M. K., Ireland, K. B., Tay, W. T. & Paini, D. Honey bee-assisted surveillance for early plant virus detection. *Ann. Appl. Biol.* **173**, 285–293 (2018).
102. Fung, E. et al. De novo assembly of honey bee RNA viral genomes by tapping into the innate insect antiviral response pathway. *J. Invertebr. Pathol.* **152**, 38–47 (2018).
103. Ryabov, E. V. et al. A virulent strain of deformed wing virus (DWV) of honeybees (*Apis mellifera*) prevails after *Varroa destructor*-mediated, or *in vitro*, transmission. *PLoS Pathog.* **10**, e1004230 (2014).
104. Zheng, H., Powell, J. E., Steele, M. I., Dietrich, C. & Moran, N. A. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proc. Natl Acad. Sci. USA* **114**, 4775–4780 (2017).
105. Kwong, W. K., Mancenido, A. L. & Moran, N. A. Immune system stimulation by the native gut microbiota of honey bees. *R. Soc. Open Sci.* **4**, 170003 (2017).
106. Moran, N. A. Genomics of the honey bee microbiome. *Curr. Opin. Insect Sci.* **10**, 22–28 (2015).
107. Motta, E. V. S., Raymann, K. & Moran, N. A. Glyphosate perturbs the gut microbiota of honey bees. *Proc. Natl Acad. Sci. USA* **115**, 10305–10310 (2018).
108. Raymann, K., Shaffer, Z. & Moran, N. A. Antibiotic exposure perturbs the gut microbiota and elevates mortality in honeybees. *PLoS Biol.* **15**, e2001861 (2017).
109. Raymann, K., Coon, K. L., Shaffer, Z., Salisbury, S. & Moran, N. A. Pathogenicity of *Serratia marcescens* strains in honey bees. *mBio* **9**, e01649-18 (2018).
110. Steele, M. I., Kwong, W. K., Whiteley, M. & Moran, N. A. Diversification of type VI secretion system toxins reveals ancient antagonism among bee gut microbes. *mBio* **8**, e01630-17 (2017).
111. El Khoury, S. et al. deleterious interaction between honeybees (*Apis mellifera*) and its microsporidian intracellular parasite *Nosema ceranae* was mitigated

- by administrating either endogenous or allochthonous gut microbiota strains. *Front. Ecol. Evol.* <https://doi.org/10.3389/fevo.2018.00058> (2018).
112. Bascompte, J. & Jordano, P. Networks: the architecture of plant-animal mutualistic biodiversity. *Annu. Rev. Ecol. Evol. Syst.* **38**, 567–593 (2007).
  113. Vaudo, A. D., Patch, H. M., Mortensen, D. A., Tooker, J. F. & Grozinger, C. M. Macronutrient ratios in pollen shape bumble bee (*Bombus impatiens*) foraging strategies and floral preferences. *Proc. Natl Acad. Sci. USA* **113**, E4035–E4042 (2016).
  114. Cane, J. H. & Snipes, S. in *Plant-Pollinator Interactions: From Specializations to Generalization* (eds N. M. Wasser & J. Ollerton) 99–122 (University of Chicago Press, 2006).
  115. Couvillon, M. J., Schurk, R. & Ratnieks, F. L. Waggle dance distances are integrative indicators of seasonal foraging challenges. *PLoS One* **9**, e93495 (2014).
  116. Greenleaf, S. S., Williams, N. M., Winfree, R. & Kremen, C. Bee foraging ranges and their relationship to body size. *Oecologia* **155**, 589–596 (2007).
  117. Richardson, R. T. et al. Rank-based characterization of pollen assemblages collected by honey bees using a multi-locus metabarcoding approach. *Appl. Plant Sci.* <https://doi.org/10.3732/apps.1500043> (2015).
  118. Richardson, R. T. et al. Quantitative multi-locus metabarcoding and waggle dance interpretation reveal honey bee spring foraging patterns in Midwest agroecosystems. *Mol. Ecol.* **28**, 686–697 (2018). **This promising study shows the power of metagenomics and barcoding in providing quantitative data on the diversity and abundance of pollen collected by honeybees.**
  119. Bell, K. L. et al. Pollen DNA barcoding: current applications and future prospects. *Genome* **59**, 629–640 (2016).
  120. Pornon, A., Andalo, C., Burrus, M. & Escaravage, N. DNA metabarcoding data unveils invisible pollination networks. *Sci. Rep.* **7**, 16828 (2017).
  121. Gous, A., Swanevelder, D. Z. H., Eardley, C. D. & Willows-Munro, S. Plant-pollinator interactions over time: pollen metabarcoding from bees in a historic collection. *Evol. Appl.* **12**, 187–197 (2019).
  122. Sickel, W. et al. Increased efficiency in identifying mixed pollen samples by meta-barcoding with a dual-indexing approach. *BMC Ecol.* **15**, 20 (2015).
  123. Hawkins, J. et al. Using DNA metabarcoding to identify the floral composition of honey: a new tool for investigating honey bee foraging preferences. *PLoS One* **10**, e0134735 (2015).
  124. Bengtsson-Palme, J. et al. Metaxa2 Database Builder: enabling taxonomic identification from metagenomic or metabarcoding data using any genetic marker. *Bioinforma.* **34**, 4027–4033 (2018).
  125. Bell, K. L. et al. Quantitative and qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Mol. Ecol.* **28**, 431–455 (2019).
  126. Hollingsworth, P. M., Li, D. Z., van der Bank, M. & Twyford, A. D. Telling plant species apart with DNA: from barcodes to genomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* <https://doi.org/10.1098/rstb.2015.0338> (2016).
  127. Garbian, Y., Maori, E., Kalev, H., Shafir, S. & Sela, I. Bidirectional transfer of RNAi between honey bee and varroa destructor: varroa gene silencing reduces varroa population. *PLoS Pathog.* **8**, e1003035 (2012).
  128. Li, W. et al. Silencing the honey bee (*Apis mellifera*) Naked Cuticle gene (*nkd*) improves host immune function and reduces nosema ceranae infections. *Appl. Env. Microbiol.* **82**, 6779–6787 (2016). **This study demonstrates the utility of RNAi in controlling pathogen and parasite populations in bees.**
  129. Yang, D. et al. Diverse factors affecting efficiency of RNAi in honey bee viruses. *Front. Genet.* **9**, 384 (2018).
  130. Taning, C. N. T., Van Eynde, B., Yu, N., Ma, S. & Smagghe, G. CRISPR/Cas9 in insects: Applications, best practices and biosafety concerns. *J. Insect Physiol.* **98**, 245–257 (2017).
  131. Kohno, H., Suenami, S., Takeuchi, H., Sasaki, T. & Kubo, T. Production of knockout mutants by CRISPR/Cas9 in the European honeybee, *Apis mellifera* L. *Zool. Sci.* **33**, 505–512 (2016).
  132. Roth, A. et al. A genetic switch for worker nutrition-mediated traits in honeybees. *PLoS Biol.* **17**, e3000171 (2019). **This study demonstrates the use of gene editing with CRISPR–Cas9 to characterize the function of candidate genes.**
  133. Chaverra-Rodriguez, D. et al. Targeted delivery of CRISPR-Cas9 ribonucleoprotein into arthropod ovaries for heritable germline gene editing. *Nat. Commun.* **9**, 3008 (2018).
  134. Mattila, H. R. & Seeley, T. D. Genetic diversity in honey bee colonies enhances productivity and fitness. *Science* **317**, 362–364 (2007).
  135. Meixner, M. D. et al. Honey bee genotypes and the environment. *J. Apicul. Res.* **53**, 183–187 (2014). **This study summarizes a series of honeybee genotype–environment studies conducted in Europe.**
  136. Danforth, B. N., Minckley, R. L., Neff, J. L. & Fawcett, F. *The Solitary Bees: Biology, Evolution and Conservation*. (Princeton University Press, 2019).
  137. Beye, M., Hasselmann, M., Fondrk, M. K., Page, R. E. & Omholt, S. W. The gene csd is the primary signal for sexual development in the honey bee and encodes a new SR-type protein. *Cell* **114**, 419–429 (2003).
  138. Harpur, B. A., Sobhani, M. & Zayed, A. A review of the consequences of complementary sex determination and diploid male production on mating failures in the Hymenoptera. *Entomol. Exp. Appl.* **146**, 156–164 (2013).
  139. Zayed, A. Effective population size in Hymenoptera with complementary sex determination. *Heredity* **93**, 627–630 (2004).
  140. Zayed, A. & Packer, L. Complementary sex determination substantially increases extinction proneness of haplodiploid populations. *Proc. Natl Acad. Sci. USA* **102**, 10742–10746 (2005).
  141. Lopez-Uribe, M. M., Jha, S. & Soro, A. A trait-based approach to predict population genetic structure in bees. *Mol. Ecol.* **28**, 1919–1929 (2019).
  142. Packer, L. et al. Conservation genetics of potentially endangered mutualisms: reduced levels of genetic variation in specialist versus generalist bees. *Conserv. Biol.* **19**, 195–202 (2005).
  143. Romiguier, J. et al. Population genomics of eusocial insects: the costs of a vertebrate-like effective population size. *J. Evol. Biol.* **27**, 593–603 (2014).
  144. Murray, E. A., Bossert, S. & Danforth, B. N. Pollinivory and the diversification dynamics of bees. *Biol. Lett.* <https://doi.org/10.1098/rsbl.2018.0530> (2018).
  145. Pinto, M. A. et al. Genetic integrity of the dark European honey bee (*Apis mellifera mellifera*) from protected populations: a genome-wide assessment using SNPs and mtDNA sequence data. *J. Apicult. Res.* **53**, 269–278 (2014).

**Acknowledgements**

Research in the laboratory of C.M.G. is supported by funding from the US National Science Foundation and the US Department of Agriculture. Research in the laboratory of A.Z. is supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada, Large Scale Applied Research Projects (BeeOMICs and BeeCSI) from Genome Canada and a York University Research Chair in Genomics.

**Author contributions**

The authors contributed equally to all aspects of the article.

**Competing interests**

The authors declare no competing interests.

**Publisher's note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Reviewer information**

*Nature Reviews Genetics* thanks D. de Graaf, K. Raymann and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**RELATED LINKS**

BeeBase: <http://hymenopteragenome.org/beebase/>

HymenopteraMine: <http://hymenopteragenome.org/>

Packer lab Bee Galleries: <https://www.yorku.ca/bugsrus/resources/resources>

The BeeCSI project: <http://beecsi.ca>

US National Honey Bee Pest and Disease Survey: <https://www.aphis.usda.gov/aphis/ourfocus/planthealth/plant-pest-and-disease-programs/honeybees/survey>

Reproduced with permission of copyright owner. Further reproduction  
prohibited without permission.