

Original Article

Genomic signatures underlying the oogenesis of the ectoparasitic mite *Varroa destructor* on its new host *Apis mellifera*

Huoqing Zheng^{a,1}, Shuai Wang^{a,1}, Yuqi Wu^{a,1}, Shengmei Zou^{b,c}, Vincent Dietemann^{d,e}, Peter Neumann^{d,f}, Yanping Chen^g, Hongmei Li-Byarlay^{h,i}, Christian Pirk^j, Jay Evans^g, Fuliang Hu^{a,*,2}, Ye Feng^{b,c,*,2}

^a College of Animal Sciences, Zhejiang University, Hangzhou, China

^b Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China

^c Institute for Translational Medicine, Zhejiang University School of Medicine, Hangzhou, China

^d Swiss Bee Research Center, Agroscope, Bern, Switzerland

^e Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

^f Institute of Bee Health, Vetsuisse Faculty, University of Bern, Bern, Switzerland

^g USDA-ARS Bee Research Laboratory, Beltsville, MD, USA

^h Agricultural Research and Development Program, Central State University, Wilberforce, OH 45384, USA

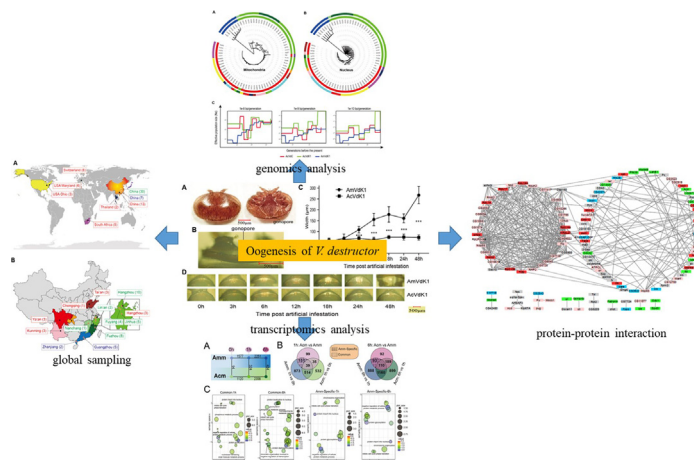
ⁱ Department of Agricultural and Life Science, Central State University, Wilberforce, OH 45384, USA

^j Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa

HIGHLIGHTS

- Genes related to the oogenesis of *V. destructor* on their new host species were studied.
- *A. cerana* and *A. mellifera* K1 mites exhibited a very close genetic relationship at the genome level.
- A total of 121 genes with nonsynonymous high- F_{ST} SNPs were found between the two types of mites.
- The transcriptomes of the two types of mites differentiated as early as 1 h post-infestation.
- Nine genes carrying nonsynonymous high- F_{ST} SNPs were associated with oogenesis on the new host.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 16 January 2022

Revised 21 April 2022

Accepted 22 April 2022

Available online 4 May 2022

ABSTRACT

Introduction: Host shift of parasites may have devastating effects on the novel hosts. One remarkable example is that of the ectoparasitic mite *Varroa destructor*, which has shifted its host from Eastern honey bees (*Apis cerana*) to Western honey bees (*Apis mellifera*) and posed a global threat to apiculture.

Abbreviations: DEG, Differentially Expressed Gene; GO, Gene Ontology; SNP, Single Nucleotide Polymorphism.

Peer review under responsibility of Cairo University.

* Corresponding authors.

¹ These authors contributed equally.

² Lead contact.

<https://doi.org/10.1016/j.jare.2022.04.014>

2090-1232/© 2022 The Authors. Published by Elsevier B.V. on behalf of Cairo University.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords:
Honey bee
Varroa destructor
Host shift
Oogenesis
Transcriptome

Objectives: To identify the genetic factors underlying the reproduction of host-shifted *V. destructor* on the new host.

Methods: Genome sequencing was conducted to construct the phylogeny of the host-shifted and non-shifted mites and to screen for genomic signatures that differentiated them. Artificial infestation experiment was conducted to compare the reproductive difference between the mites, and transcriptome sequencing was conducted to find differentially expressed genes (DEGs) during the reproduction process. **Results:** The host-shifted and non-shifted *V. destructor* mites constituted two genetically distinct lineages, with 15,362 high- F_{ST} SNPs identified between them. Oogenesis was upregulated in host-shifted mites on the new host *A. mellifera* relative to non-shifted mites. The transcriptomes of the host-shifted and non-shifted mites differed significantly as early as 1h post-infestation. The DEGs were associated with nine genes carrying nonsynonymous high- F_{ST} SNPs, including *mGluR2-like*, *Lamb2-like* and *Vitellogenin 6-like*, which were also differentially expressed, and *elF4G*, *CG5800*, *Dap160* and *Sas10*, which were located in the center of the networks regulating the DEGs based on protein-protein interaction analysis.

Conclusions: The annotated functions of these genes were all associated with oogenesis. These genes appear to be the key genetic determinants of the oogenesis of host-shifted mites on the new host. Further study of these candidate genes will help elucidate the key mechanism underlying the success of host shifts of *V. destructor*.

© 2022 The Authors. Published by Elsevier B.V. on behalf of Cairo University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Most parasites have a specific host range. However, the host range can be expanded when a host shift occurs (i.e., when a parasite infests a new host). Such shifts can lead to the emergence of devastating infectious diseases [1]. One species that has undergone such a shift is *Varroa destructor*, an ectoparasitic mite of honey bees. Originally a parasite of Eastern honey bees, *Apis cerana*, *V. destructor* shifted to Western honey bees, *Apis mellifera*, in the middle of the 20th century [2] after the latter was introduced to Asia for pollination and honey production. Benefiting from the international honey bee trade, the mite then attained a global distribution within a few decades, which has had disastrous consequences for global apiculture with European *A. mellifera* [2] and has, with few exceptions, eradicated wild and feral honey bee populations in Europe and North America [3,4]. Infestations by this parasite lead to host colony death due to its feeding on immature hosts [5] and the fact that it serves as a vector of viruses [6]. To mitigate the effects of this host shift and limit chances for further occurrences, identifying the mechanisms that permit this parasite to use alternative hosts requires investigation.

The mechanisms underlying successful host shifts are those that allow contact with an alternative host and enable successful infestations [7]. The former resulted from human transport of *A. mellifera* into the native range of *A. cerana* in Asia [8]. This novel sympatry has fostered the interspecific robbing of food stores [9,10], which has enabled mites infesting *A. cerana* to invade *A. mellifera* colonies [11]. The traits that permit *V. destructor* to successfully infest new hosts remain unclear [12]. A key mechanism required for successful infestation is the ability to exploit the new host for reproduction [12], and the recognition of host cues and the triggering of mite oogenesis are particularly important. Although there are approximately a dozen haplotypes of *V. destructor*, only two have successfully shifted to *A. mellifera*, the Korean haplotype (K1) and the Japanese/Thailand haplotype (J1) [8], which suggests that the host shift of *V. destructor* mites is a phenotype determined by specific genetics. A previous study revealed that some genes involved in oogenesis and reproduction are overexpressed in another species of this mite, *Varroa jacobsonii*, that was able to reproduce on the new host *A. mellifera* in Papua New Guinea [13]; however, nothing is known about the molecular mechanisms underlying the oogenesis of the *V. destructor* K1 haplotype which determines its ability to exploit its novel host *A. mellifera*.

Recently, a native population of the *V. destructor* K1 haplotype that is genetically highly similar but differs in its reproductive abil-

ities from the host-shifted lineage was found on the original host *A. cerana* in Eastern China [14,15]. This native mite population can reproduce only on the male brood of the original host, whereas the invasive lineage can reproduce on both the worker and male brood of both the original and new host and shows a lower host specificity [15]. Comparison of native nonshifted and invasive host-shifted mite populations provides an ideal approach for identifying the genetic factors underlying reproduction in *V. destructor*.

To facilitate their identification, we conducted a joint analysis of genomic, transcriptomic, epigenomic and reproductive data. We first performed whole-genome sequencing (WGS) to identify single nucleotide polymorphisms (SNPs) segregating the shifted and non-shifted lineages. Experimental infestations were then performed to obtain individuals differing in reproductive status for transcriptomics and epigenomics. Bioinformatics analyses combining genomic and transcriptomic data identified nine candidate genes that may have mutated to facilitate the exploitation of the new host *A. mellifera* by host-shifted mites.

Materials and methods

Sampling for genome sequencing

At 17 locations, *V. destructor* mite families (N = 1 to 10 according to mite availability, each family composed of a foundress, her son and one to three daughters) were collected from singly infested capped worker or drone brood cells of *A. mellifera* and capped drone brood cells of *A. cerana* in one to three colonies per apiary [16]. Details on the sampling locations, origins of honey bee species and haplotypes are shown in Fig. 1 and Table S1 at DOI <https://doi.org/10.17605/OSF.IO/ZS948>.

DNA extraction and mtDNA genotyping

Due to their small size [2], all the individuals composing each mite family were pooled for DNA extraction. Total DNA was extracted using the QIAamp Fast DNA Tissue Kit (QIAGEN, Hilden, Germany). The segment of the mitochondrial *cox1* gene was amplified using the primers reported by Wang et al. 2018 (Cox1_821_F: GGAGTAGGTACAGGTTGAACGG; Cox1_821_R: ACAACCCAGCAATAATAGCAA) [17]. PCR products were sequenced and the haplotype of each mite family was determined by comparison with the 458-bp sequence of the *cox1* gene reported by Anderson and Trueman (2000) [18].

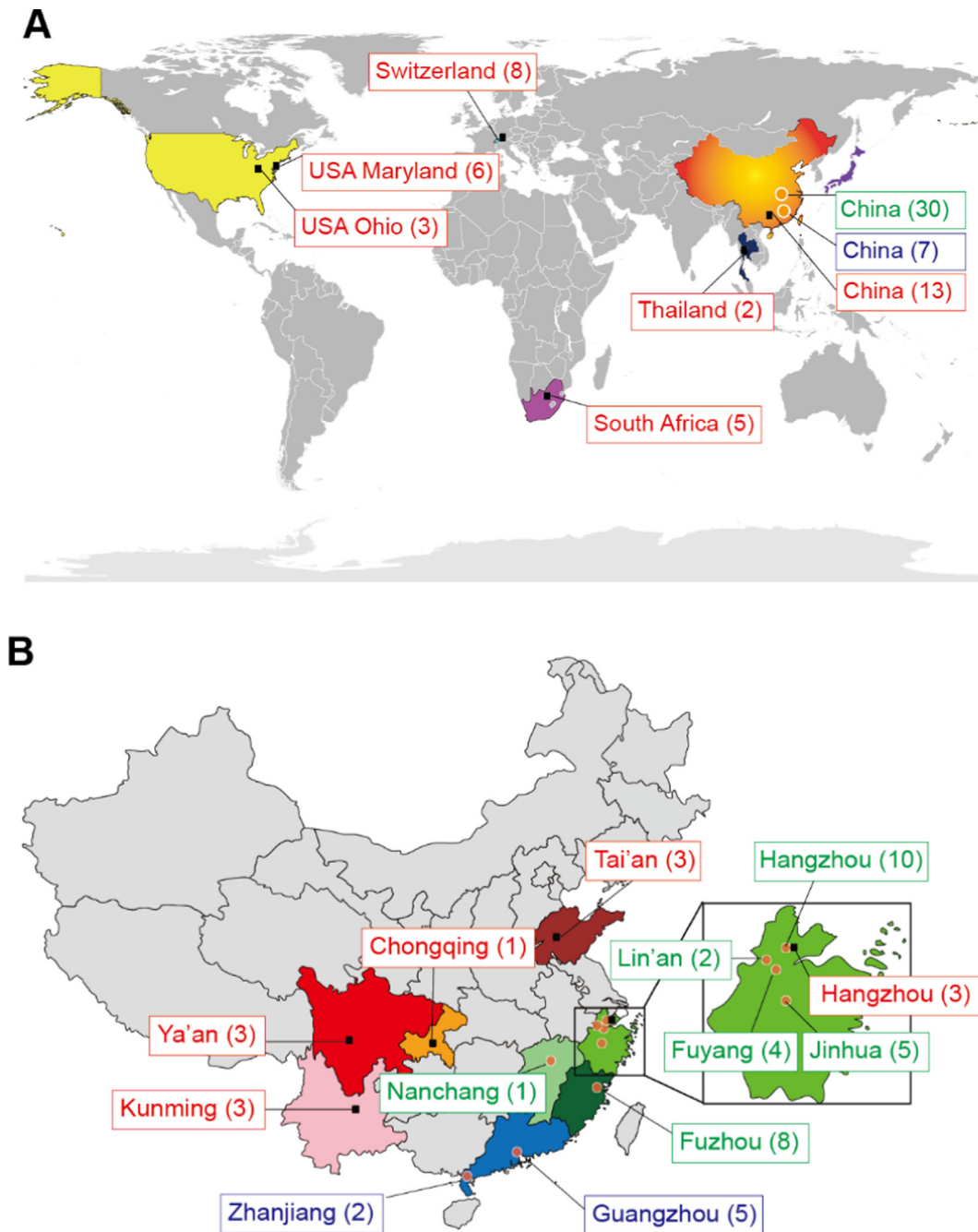


Fig. 1. Geographic origin of the *Varroa destructor* mite samples: (A) Global sampling (N in brackets), (B) Sampling in China from both *Apis mellifera* (black squares) and *Apis cerana* (orange circles). Countries and provinces are shown in different colors, which are used in the following graphs to indicate the geographic origin of the samples. Text color indicates the host species and *V. destructor* haplotypes (Red: K1 haplotype *V. destructor* mites infesting *A. mellifera*; Green: K1 haplotype *V. destructor* mites infesting *A. cerana*; Blue: C1 haplotype *V. destructor* mites infesting *A. cerana*).

Sequencing, quality control and variant calling

For each sample, we constructed a paired-end library with high quality DNA and sequenced the DNA on the Illumina HiSeq SE50 sequencing platform following the manufacturer's protocol. At least 10 Gb of 2×150 bp sequencing data were generated for each sample. Low-quality reads, including those with $> 50\%$ low-quality bases ($Q < 20$) or containing $> 10\%$ Ns, were discarded. Duplicated reads were further removed using Picard (<https://broadinstitute.github.io/picard/>). Clean reads were then mapped to the reference genome of *V. destructor* (GenBank Accession No.: BEIS00000000.1) [19] using the BWA aligner [20]. UnifiedGenotyper in the software Genome Analysis Tool Kit (GATK, [21]) was used to identify SNPs

and small indels. Low quality variants were then filtered with GATK Variant Filtration. Structural variation was detected using BreakDancer [22]; copy number variation was detected using CNVnator [23].

Diversity analysis and phylogenetic reconstruction

The pairwise genetic distances were measured on the basis of the number of shared alleles between each sample divided by the number of total SNPs; the obtained distance matrix was then used to construct a phylogenetic tree with the neighbor-joining method in PHYLIP [24]. Pairwise F_{ST} between groups were calculated using VcfTools [25]. Nucleotide diversity (π) was calculated

using DnaSP v6 [26]. ADMIXTURE was used to estimate the genetic ancestry of each sample, specifying a range of 2–5 hypothetical ancestral populations [27].

To tease apart the effects on the genetic structure from two sources of variation, namely, isolation by distance vs. host adaptation, Mantel and partial Mantel tests were performed using the *vegdist* and *mantel* packages in R, and significance was determined via 10,000 permutations. The genetic distance matrix was the same as that used for constructing the neighbor-joining tree. The geographic distance was calculated as the linear distance between the sampling sites and was estimated using Baidu Earth (<http://map.baidu.com>). The host distance matrix was built using the host phenotype as a discrete trait; pairs collected on the same honey bee host species were assigned a distance of 0, and those collected on different host species were assigned a distance of 1. SMC++ was used to estimate the historical effective population size [28]. Techer et al (2021) reported that the mutation rate in *V. destructor* was approximately 8×10^{-10} per bp per generation [29]. However, to increase confidence of our estimate, we calculated the population size with three possible mutation rates (1×10^{-8} , 1×10^{-9} , 1×10^{-10} per bp per generation).

Infestation experiments

Adult female *V. destructor* mites were collected from worker or drone brood cells of *A. mellifera* and drone brood cells of *A. cerana* colonies in the experimental apiary in Hangzhou, China, where both honey bee species were kept [15]. After being kept on *A. mellifera* workers to mimic the mite's mobile phase for two days, the two lineages of mites were manually transferred to freshly sealed (<6h post sealing) *A. mellifera* drone brood cells to record the fertility of each mite one day prior to the expected adult emergence (i.e., 13 days after infestation) [16].

Next, we carried out another three infestation experiments for the following purposes: 1) to measure the gravidity status index (i.e., gonopore size) of female mites to monitor their reproductive status morphologically; 2) to carry out transcriptomic sequencing; and 3) to perform DNA methylation sequencing. For 1), 10 *A. cerana* mites and 10 *A. mellifera* mites were sampled before being experimentally transferred to host larvae (0 h) and 3 h, 6 h, 12 h, 18 h, 24 h and 48 h after experimental infestation. Gravid mites have a swollen opisthosoma due to the presence of a matured egg which leads to the distending of intersegmental membranes [30]. As a result, gonopore size as the distance from the distal end of the anal sclerite (cribrum) to the dorsal shield increases. This distance was thus measured as a morphological index or the degree of gravidity and hence of reproductive activity. The experiment was replicated thrice in drone brood cells of three *A. mellifera* colonies. Data of the three replicates were pooled because of the high homogeneity across replicates. For 2) and 3), 15 *A. cerana* mites and 15 *A. mellifera* mites were sampled at 0 h, 1 h and 6 h post-infestation; the 15 mites were split into three groups as biological replicates.

RNA extraction, sequencing and data analysis

Total DNA and RNA were extracted from each mite with a DNA/RNA Extraction Kit (Tiangen, Beijing, China) following the manufacturer's protocol. To avoid the use of mites that recently drifted to colonies of the other host species in our analysis, mite lineages were confirmed a posteriori by mtDNA amplification and microsatellite DNA analysis [15].

RNA extracted from five individual mites in each treatment was pooled for transcriptome sequencing. The NEBNext Ultra TM RNA Library Prep Kit (New England Biolabs, Beijing, China) was used to construct the paired-end 150 bp library. RNA-seq was performed with an Illumina HiSeq 1500 system (Novogene, Tianjin,

China). At least 6 Gb of 2×150 bp sequencing data were generated for each sample.

The filter of low-quality reads was the same as that for genomic sequencing above. TopHat2 [31] was used to align the RNA-seq reads against the *V. destructor* genome (GenBank Accession No.: BEIS00000000.1). Gene expression was evaluated with RSEM [32]. Differentially expressed genes (DEGs) between individual samples were identified using DESeq2 [33]. The expression profile of these DEGs was clustered and a heatmap was produced using the *heatmap* package in R software.

DNA methylation sequencing and data analysis

The total DNA extraction and genotyping of individual mites followed the same procedures as those for genomic sequencing. The sequencing library was constructed by an EZ DNA Methylation Gold Kit (Zymo Research, Irvine, USA). High-throughput sequencing was performed with an Illumina HiSeq 1500 system (Novogene, Tianjin, China). At least 12 Gb of 2×150 bp sequencing data were generated for each sample.

Bismark software was used to map methylated cytosine (5mC) sequence reads against the *V. destructor* genome (GenBank Accession No.: BEIS00000000.1) and to determine cytosine methylation states [34,35]. The R library DSS and package methylKit were used to detect differentially methylated regions [36].

Gene Ontology (GO) annotation and enrichment analysis.

To perform GO enrichment analysis, we first search for GO terms based on their orthologs in the NCBI nr database. A BLAST (<https://blast.ncbi.nlm.nih.gov/>) protein search was conducted, using an e-value of 10^{-5} as the threshold; the highest hit with the GO term in the database was retrieved. The enrichment analysis was performed using the R package phyper; GO terms with q-value < 0.05 were considered significantly enriched. The EnrichmentMap plugin in Cytoscape software was used to visualize the enriched GO terms [37,38].

Protein–protein interactions

The protein–protein interactions (PPIs) were analyzed by the STRING database (<https://string-db.org/>). The fruit fly (*Drosophila melanogaster*) was chosen as the reference. The results were presented in visualization by Cytoscape [38].

Results

Genetic relationship between the host-shifted and nonshifted K1 haplotype . destructor

To profile the genomic structure of the K1 haplotype of *V. destructor* and to resolve the genetic relationship between the host-shifted and nonshifted mite lineages, 67 samples were processed for WGS. The samples included 37 *A. mellifera* K1 host-shifted mites from Switzerland, South Africa, the USA, Thailand, and five provinces in China, and 30 *A. cerana* K1 endemic, non-shifted mites from five provinces in China (Fig. 1A-B; Table S1 at DOI <https://doi.org/10.17605/OSF.IO/ZS948>). WGS data identified 173 SNPs in the mitochondrial genome, which generated a phylogenetic tree in which *A. mellifera* K1 mites could not be separated from *A. cerana* K1 mites, with five samples collected from *A. cerana* colonies being grouped in the *A. mellifera* K1 cluster (Fig. 2A).

At the nucleus level, 1,223,774 biallelic SNPs were discovered. In the nuclear SNP-based phylogenetic tree, the *A. mellifera* K1 mite samples were clearly separated from the *A. cerana* K1 mite samples

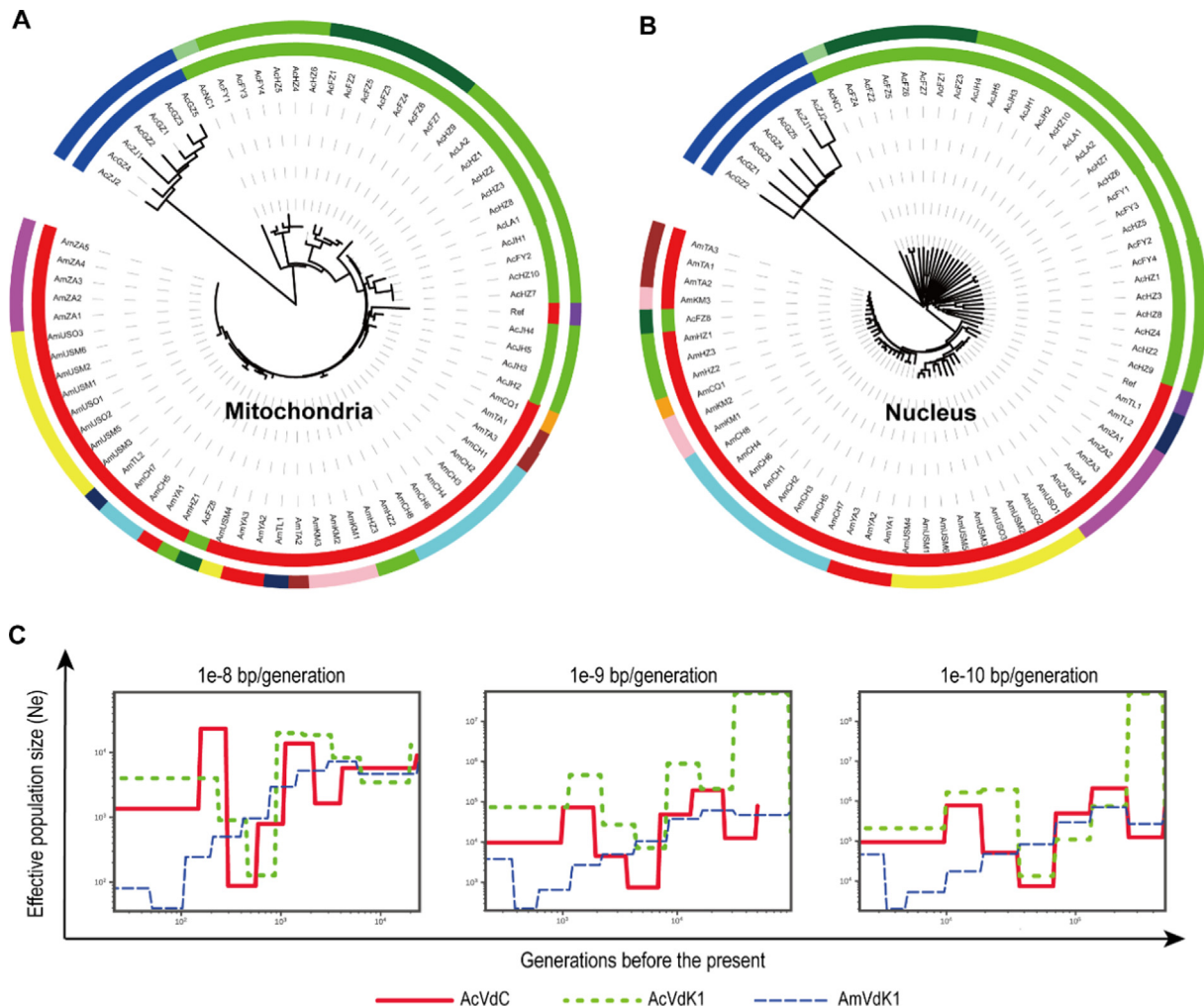


Fig. 2. Genomics analysis of *A. cerana* and *A. mellifera* mites. (A–B) Neighbor-joining tree constructed from SNPs on mitochondrial (A) and nuclear DNA (B). The inner circles indicate haplotypes: *A. mellifera* K1 (in red), *A. cerana* K1 (in green), and *A. cerana* C1 haplotype (in blue; as outgroup). The outer circles indicate the geographic origin of the samples, with country- and province- specific colors corresponding to Fig. 1A and 1B. (C) Inferred historical population size derived from different mutation rates. AcVdC: *V. destructor* C1 mite infesting *A. cerana*; AcVdK1: *V. destructor* K1 mite infesting *A. cerana*; AmVdK1: *V. destructor* K1 mite infesting *A. mellifera*. Detailed sample information is provided in Table S1.

(Fig. 2B). The only exception was a mite collected from an *A. cerana* colony in Fujian Province, China, which was classified in the *A. mellifera* K1 cluster and was thus excluded from further analysis. The outgroup C1 haplotype *V. destructor* infesting *A. cerana* samples was placed at the root of the phylogenetic tree. The nucleotide diversity of *A. cerana* C1, *A. cerana* K1 and *A. mellifera* K1 mites was 0.0003, 0.0004, and 0.0001, respectively. For *A. mellifera* K1 mites, the samples were clustered according to their geographic origin, with the Swiss samples closer to the Chinese samples (China-Switzerland clade) and the South African samples closer to the American samples (USA-South Africa clade) (Fig. 2B).

We next estimated the effective population size of *A. cerana* K1 and *A. mellifera* K1 mites over time. The exact curve plots derived from the three mutation rates selected varied, but all showed that the sampled *A. mellifera* K1 mites experienced a severe genetic bottleneck and were descendants of a recent common ancestral lineage (Fig. 2C).

We further tested whether patterns of population structure between *A. cerana* K1 and *A. mellifera* K1 mites were the product of isolation by geographic distance or of host adaptation. The genetic distance was not related to geographic distance (Mantel test, $R = -0.08$, $P = 0.87$). In contrast, there was a significant correlation between genetic distance and host range, even while con-

trolling for geographic distance (partial Mantel test, $R = 0.35$, $P < 0.001$). Thus, the genetic structure mainly reflected the host differences between the two mite lineages.

Genomic signatures of reproduction

The close genetic proximity between nonshifted *A. cerana* and host-shifted *A. mellifera* K1 mites (Fig. 2B) provided us with an opportunity to investigate the genomic signatures underlying differences in reproductive ability. We calculated F_{ST} -statistics (also known as the fixation index, F_{ST}) to measure the differentiation in variant (including structural variation and SNP) frequency among populations. Based on the low F_{ST} values, no structural variation events, including copy number variation, inversions and translocations, were identified between *A. cerana* and *A. mellifera* K1 mites. At the SNP level, 15,326 (1.24%) had F_{ST} values > 0.9 (high- F_{ST} SNPs). Among these SNPs, only 356 (2.32%) were situated in the coding regions (Fig. 3A), with synonymous and nonsynonymous SNPs at a ratio of 2:1. In addition, 227 of these SNPs (1.5%) were located within the promoter regions of 223 genes (Table S2 at DOI <https://doi.org/10.17605/OSF.IO/ZS948>). Notably, the percentage of promoter SNPs and the ratio of nonsynonymous/synonymous SNPs increased when $F_{ST} > 0.9$ (Fig. 3 B, C).

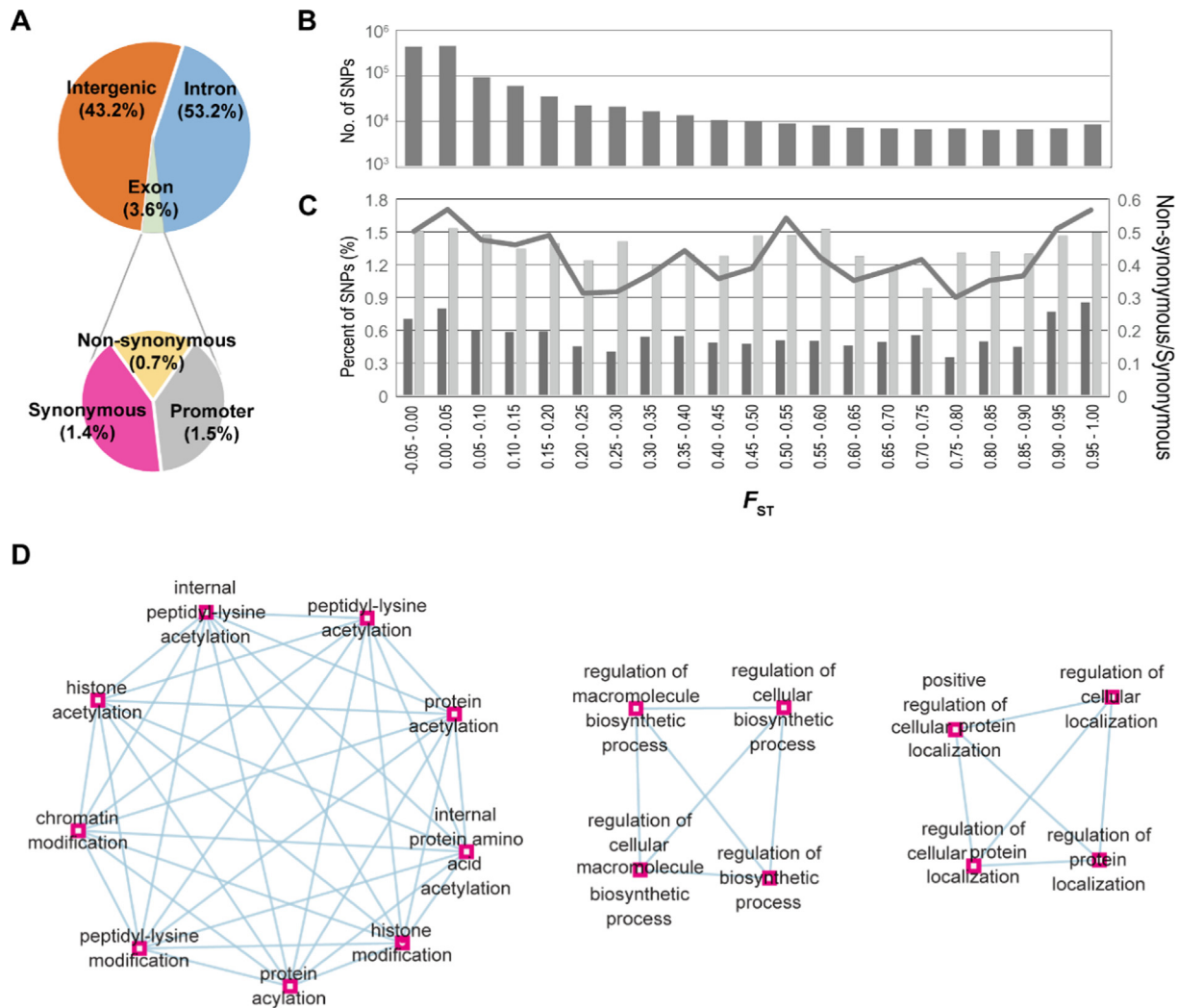


Fig. 3. Annotations of the SNPs with high F_{ST} values between *Apis cerana* and *Apis mellifera* K1 mites. (A) Partitioning of the SNPs according to their coding potential. (B) Number of SNPs divided by F_{ST} intervals. (C) The percent of nonsynonymous SNPs (black bars) and the SNPs located within the promoter regions (gray bars) at different F_{ST} intervals. Additionally, the rate of nonsynonymous SNPs to synonymous SNPs is shown (gray line). The x-axis below Panel C shows the F_{ST} intervals for both Panels B and C. (D) Enrichment maps of the enriched Gene Ontology (GO) biological process terms in the 121 genes carrying high- F_{ST} nonsynonymous SNPs. The nodes represent GO terms; the lines indicate the gene overlap between the connected GO terms.

The 125 nonsynonymous high- F_{ST} SNPs were distributed among 121 genes (Table S3 at DOI <https://doi.org/10.17605/OSF.IO/ZS948>). Of the 38 Gene Ontology (GO) biological processes enriched for the 121 genes, 20 were related to protein localization, protein modification and protein metabolism (Table S4 at DOI <https://doi.org/10.17605/OSF.IO/ZS948>; Fig. 3D).

Differentially expressed genes in reproduction initiation

To compare the reproductive ability of *A. cerana* and *A. mellifera* K1 mites in the novel host, K1 *V. destructor* mites collected from *A. cerana* (N = 59) and *A. mellifera* (N = 63) colonies from Hangzhou were individually transferred into newly sealed *A. mellifera* drone brood cells. Their reproductive output showed distinct patterns: $83.5\% \pm 9.1\%$ of *A. mellifera* K1 mites produced at least one offspring, whereas none of the *A. cerana* K1 mites did. We also opened *A. mellifera* drone brood cells (N = 30) experimentally infested with each mite type at seven time points (0, 3, 6, 12, 18, 24, and 48 h) and measured the gonopore size as a morphological index of reproductive activity. Consistent with their sterility and despite feeding activity to sustain themselves, this index did not change significantly for *A. cerana* K1 mites, whereas this index for *A. mellifera*

K1 mites steadily increased over time. The significant difference on gonopore size between these mites can be observed as early as 6 h after artificial infestation (Fig. 4).

To investigate the genes associated with the onset of oogenesis in *A. mellifera* K1 mites, *A. cerana* and *A. mellifera* K1 mites collected at 0 h, 1 h, and 6 h after experimental transfer to *A. mellifera* drone brood cells were used for transcriptomic analysis; 0 h and 1 h represented the time points before and shortly after contact with the host, respectively, before *A. mellifera* K1 mites showed visible signs of vitellogenesis with increased gonopore size at 6 h post-infestation (Fig. 4C).

We found that at 0 h, the *A. mellifera* and *A. cerana* K1 mites displayed a very similar transcriptome profile, while starting from 1 h, significantly different transcriptome profiles occurred between *A. mellifera* and *A. cerana* K1 mites (Fig. 5A; Table S5 at DOI <https://doi.org/10.17605/OSF.IO/ZS948>). At 1 h, 1,577 and 1,120 DEGs were identified in the *A. mellifera* and the *A. cerana* K1 mites, respectively, compared with samples at 0 h. At 6 h, 2,251 and 2,358 DEGs were identified in the *A. mellifera* and the *A. cerana* K1 mites, respectively, compared with samples at 1 h. A total of 324 and 484 DEGs were identified between the *A. mellifera* and the *A. cerana* K1 mites at 1 h and 6 h, respectively.

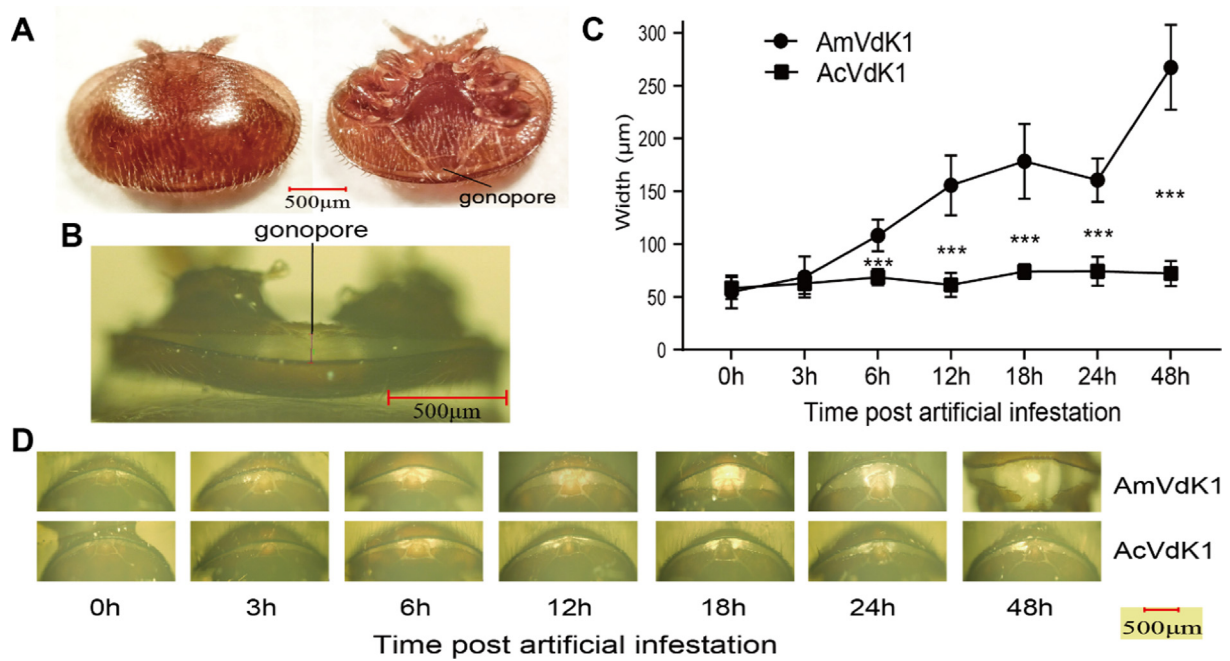


Fig. 4. Morphological determination of differences in the initiation of reproduction between ectoparasitic *A. cerana* and *A. mellifera* K1 mites at the early stage of *Apis mellifera* drone brood infestation. (A) Dorsal and ventral views of an *A. mellifera* K1 female mite, indicating the position of the gonopore. (B) Ventral view of an *A. mellifera* K1 female mite displaying gonopore size as the distance from the distal end of the anal sclerite (cribrum) to the dorsal shield. (C-D) Measurement and view of gonopore size of *A. cerana* and *A. mellifera* K1 mites infesting *A. mellifera* drone brood at seven time points post-infestation. AcVdK1: *V. destructor* K1 mite infesting *A. cerana*; AmVdK1: *V. destructor* K1 mite infesting *A. mellifera*.

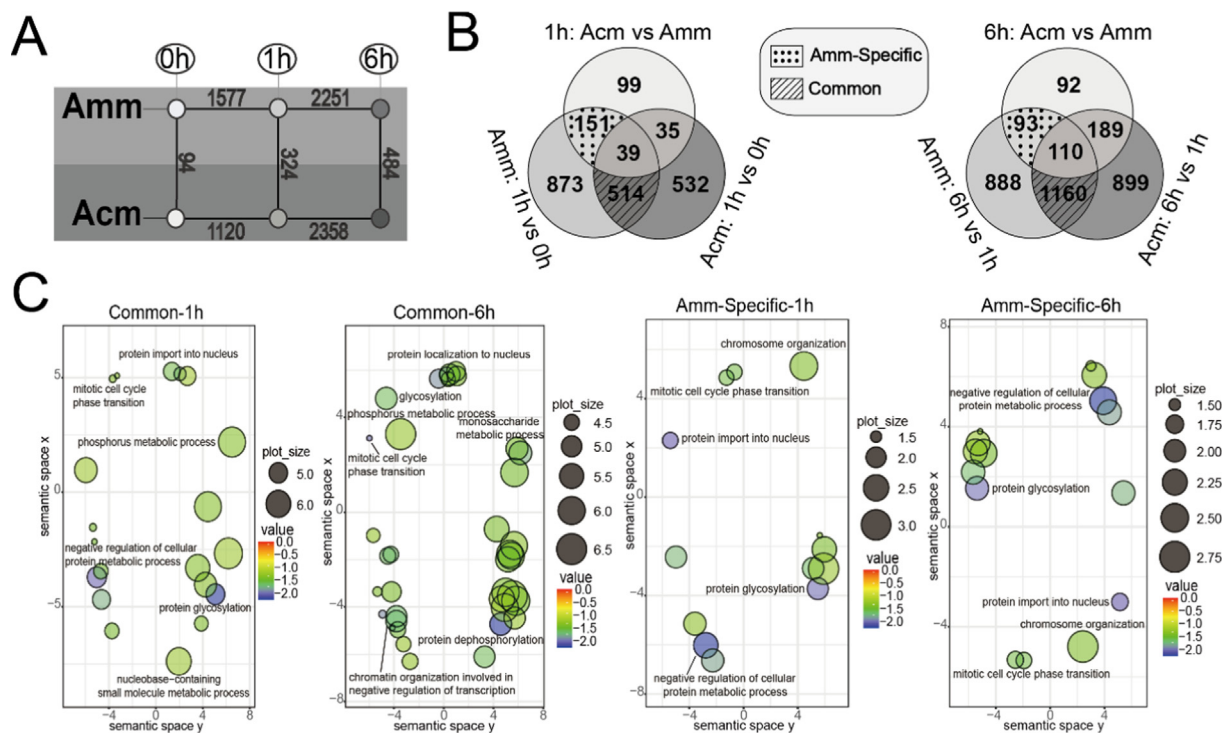


Fig. 5. Comparison of transcriptomic profiles in *A. cerana* and *A. mellifera* K1 mites. Acm and Amm indicate *A. cerana* and *A. mellifera* K1 mites, respectively. (A) Schematic showing the number of differentially expressed genes (DEGs) based on comparison of each time point and mite type. The time post-infestation is shown at the top of the figure. (B) Venn diagram showing the number of overlapping DEGs among the three pairs of comparisons (within and between species and times) at 1 h and 6 h post-infestation. The “common DEGs” (i.e., overlapping DEGs between time points within species, excluding those differentially expressed between species), are indicated by the lined shading; the “Amm-specific DEGs” (i.e., the DEGs specific to *A. mellifera* K1 mites) are indicated by dotted shading. (C) Semantic space analysis of the significantly enriched GO terms for biological processes over-represented among the “common DEGs” and “Amm-specific DEGs”, respectively. Bubble color indicated for the significance of how GO terms were enriched and size indicates the frequency of the GO term found in the GOA database. Displayed categories have been selected from a broader set to eliminate redundancy and prepared for visualization using the REVIGO tool available at <https://revigo.irb.hr/>; see Table S6 for an exhaustive listing.

At both 1 h and 6 h, approximately half the DEGs in *A. cerana* mites were shared with those in *A. mellifera* mites. The DEGs shared by the two mite types between each time point (0 h vs. 1 h and 1 h vs 6 h), excluding the genes differentially expressed between the two types of mites, were designated as “common DEGs” (Fig. 5B). The biological process GO terms enriched by the “common DEGs” were highly similar at 1 h and 6 h, mostly in chromosome organization, protein transportation and protein modification (Fig. 5C; Table S6 at DOI <https://doi.org/10.17605/OSF.IO/ZS948>). The DEG sets that were specific to *A. mellifera* K1 mites and differentially expressed between the two types of mites were designated as “Amm-specific DEGs” (Fig. 5B). Only one gene of the “Amm-specific DEGs” overlapped between 1 h and 6 h, and the enriched biological process GO terms were very similar at the two time points (Fig. 5C; Table S7 at DOI <https://doi.org/10.17605/OSF.IO/ZS948>). In these overlapping GO terms, “ribonucleoside monophosphate biosynthetic process” and “regulation of MAP kinase activity” were specifically enriched among “Amm-specific DEGs” compared with the GO terms enriched among the “common DEGs”.

To test whether the above transcriptomic differences stemmed from differences in DNA methylation levels, we compared the genome-wide DNA methylation variation between *A. mellifera* and *A. cerana* K1 mites at three time points. However, differential methylation was not observed in any DNA segment between the two mite groups. Therefore, we speculated that genomic variations rather than epigenetic phenomena was responsible for the transcriptomic changes in *A. mellifera* K1 mites.

Links between the genome and transcriptome

None of the 223 genes with promoters carrying high F_{ST} SNPs (Table S2 at DOI <https://doi.org/10.17605/OSF.IO/ZS948>) appeared in the Amm-specific DEGs. Next, we focused on the links between the Amm-specific DEGs and the 121 genes carrying nonsynonymous high- F_{ST} SNPs. Four genes appeared in the two gene sets at 1 h, *mGluR2-like* (LOC111252394; encoding metabotropic glutamate receptor 2-like isoform X6), *LOC111251149* (uncharacterized protein), *Lamb2-like* (LOC111245919; encoding laminin subunit beta-2-like isoform X5) and *LOC111253996* (uncharacterized protein), and one gene at 6 h, *Vitellogenin 6-like* (LOC111249976). Protein-protein interaction analysis between the three gene sets revealed four genes carrying nonsynonymous high- F_{ST} SNPs at the core putative regulatory networks (Fig. 6), *CG5800* (LOC111243975; encoding probable ATP-dependent RNA helicase DDX10 isoform X4), *Sas10* (LOC111248714; encoding silencing protein 10-like), *Dap160* (LOC111247635; encoding intersectin-1-like isoform X4) and *eIF4G* (LOC111252325; encoding eukaryotic translation initiation factor 4). In *A. mellifera* K1 mites, the DEGs they connected were all upregulated compared with *A. cerana* K1 mites at 1 h or 6 h.

Discussion

The aim of the present study was to identify the genomic, transcriptomic, and epigenomic mechanisms underlying the oogenesis of the *V. destructor* haplotype K1 in a new host species, *A. mellifera*. The main findings of the study were the following: i) *A. cerana* and *A. mellifera* K1 *V. destructor* mites were closely related at the genome level with 125 nonsynonymous high- F_{ST} SNPs distributed among the 121 genes, ii) the transcriptomes of the shifted and non-shifted mites were significantly differentiated as early as 1 h post-infestation of the new host *A. mellifera*, and iii) nine genes carrying nonsynonymous high- F_{ST} SNPs were associated with oogenesis on the new host.

The genetic relationship between *A. cerana* and *A. mellifera* K1 mites

To examine the genetic relationships between *A. cerana* and *A. mellifera* K1 mites, we performed, for the first time, a genomic comparison between host-shifted and nonshifted populations of the same mite haplotype. Compared with the outgroup (*A. cerana* C1 mites), the *A. cerana* and *A. mellifera* K1 mites exhibited a very close genetic relationship. Although indistinguishable based on the mitochondrial DNA, the two clades could be distinguished by nuclear DNA. Investigation of the K1 clade revealed much higher genetic diversity of the *A. cerana* mites than that of the *A. mellifera* samples, despite the former being restricted to southeastern China (Fig. 1B), which clarifies the findings of previous work on smaller genetic regions [15]. Inferred population analysis revealed that *A. mellifera* K1 mites experienced a genetic bottleneck, which is consistent with the host shift history of invasive mites [2], and indicates a recent common ancestral lineage worldwide. Because of its close genetic relationship with *V. destructor* K1 mites infesting *A. cerana* (Fig. 2A, B), this ancestral lineage appears to be a subclade of these native mites [8,15].

A successful host shift requires exposure to a novel host and a preadaptation of the parasite until a self-sustaining parasite population is established [7]. In the areas, where *A. mellifera* and *A. cerana* are kept in close sympatry (i.e., same apiaries), drifting of *V. destructor* between the two honey bee species can occur [11].

Additionally, one mite collected from an *A. cerana* colony clustered genetically with mites from *A. mellifera*. However, even if the drifting of mites occasionally occurs (1 out of the 37 detected in this study, [11,15]), *Varroa* spp. lineages are rarely able to reproduce on *A. mellifera* [18,39,40]. This suggests that *Varroa* spp. require novel mutations to overcome reproductive barriers, which occur at a low frequency. To putatively identify such mutations, we systematically analyzed the results of WGS, cross-fostering infestation experiments, and transcriptomic and epigenomic analyses.

Molecular mechanism underlying oogenesis on the new host

The *A. mellifera* K1 mites sampled from different continents have the same reproductive phenotype. They can reproduce on *A. mellifera* drone and worker brood [2,15]. The genetic determinants of this reproductive phenotype must thus have occurred within the ancestral population of the invasive lineage of *V. destructor* K1 mites. The high- F_{ST} SNPs identified in our study represent the genetic features that must have emerged in this particular lineage. These high- F_{ST} SNPs are more likely to contribute to phenotypic differences between these two kinds of mite than other SNPs.

Morphological changes linked to the initiation of reproductive activity were observed at 6 h post-artificial infestation in the *A. mellifera* K1 mites, and both the *A. cerana* and the *A. mellifera* K1 mites underwent significant changes in their transcriptome profiles as early as 1 h post-artificial infestation. At both 1 h and 6 h post-infestation, *A. cerana* and *A. mellifera* K1 mites exhibited a high percentage overlap of DEGs and enriched GO terms. Although similar transcriptomic responses can result from a variety of responses shared by the two types of mites, such as feeding, this also suggests that there is a sequence of cues required to lead to successful reproduction. The *A. cerana* K1 mites appear to recognize the early cues triggering oogenesis in *A. mellifera* drones but appear unable to complete this process. The DEGs specific to *A. mellifera* K1 mites differed at 1 h and 6 h post-artificial infestation, yet they were enriched in the same set of GO terms. The biological process GO terms “regulation of MAP kinase activity” and “ribonucleoside monophosphate biosynthetic process” were specifically enriched by “Amm-specific DEG” but not by the “common DEGs”, suggesting that these GO terms are key to the successful oogenesis of the host-shifted *V. destructor* K1 lineage. The mitogen-activated

Conclusions

The *V. destructor* K1 mites infesting *A. cerana* in southeastern China are highly similar in genetics with those infesting *A. mellifera*, thus providing good opportunities to investigate the genomic signatures underlying host shifts of the mites. Despite their close genetic relationship at the genome level, the transcriptomes of host-shifted and nonshifted *V. destructor* mites were significantly differentiated as early as 1 h post-infestation of the new host *A. mellifera*. Joint analysis of genomic, transcriptomic, and reproductive data of the two closely related lineages suggests that mite reproduction is affected by nine candidate genes, seven of which were related to upregulation of oogenesis, which may explain the striking differences in host specificity between *A. mellifera* and *A. cerana* K1 mites and have contributed to their ability to reproduce on the foreign host. Further study of the functions and regulation of these nine candidate genes will help elucidate the molecular basis of the host shift of *V. destructor*.

Funding

This work was supported by the National Natural Science Foundation of China (grant numbers 32072798 and 31672498), Science and Technology Department of Zhejiang Province, China (grant number 2021C02068-8); China Agriculture Research System of MOF and MARA (grant number CARS-44); and the United State Department of Agriculture National Institute of Food and Agriculture Award (grant number NI181445XXXXG007).

Availability of data and materials

All short-read sequence data of this study have been deposited in the NCBI Short Read Archive (genome sequencing: SRR11497693–SRR11497771; transcriptome sequencing: SRR11536640–SRR11536651; whole genome bisulfite sequencing: SRR11509808–SRR11509819), under the BioProjects PRJNA623528, PRJNA624005 and PRJNA623755. All the supplementary information files are included at OSF <https://doi.org/10.17605/OSF.IO/ZS948>.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. *Nature* 2008;451(7181):990–3.
- Rosenkranz P, Aumeier P, Ziegelmann B. Biology and control of *Varroa destructor*. *J Invertebr Pathol* 2010;103:S96–S119.
- Locke B. Natural *Varroa* mite-surviving *Apis mellifera* honeybee populations. *Apidologie* 2016;47(3):467–82.
- Jaffé R, Dietemann V, Allsopp MH, Costa C, Crewe RM, Dall'olio R, et al. Estimating the density of honeybee colonies across their natural range to fill the gap in pollinator decline censuses. *Conserv Biol* 2010;24(2):583–93.
- Ramsey SD, Ochoa R, Bauchan G, Gulbranson C, Mowery JD, Cohen A, et al. *Varroa destructor* feeds primarily on honey bee fat body tissue and not hemolymph. *Proc Natl Acad Sci USA* 2019;116(5):1792–801.
- Berthoud H, Imdorf A, Haueter M, Radloff S, Neumann P. Virus infections and winter losses of honey bee colonies (*Apis mellifera*). *J Apic Res* 2010;49(1):60–5.
- Woolhouse MEJ, Haydon DT, Antia R. Emerging pathogens: the epidemiology and evolution of species jumps. *Trends Ecol Evol* 2005;20(5):238–44.
- Traynor KS, Mondet F, de Miranda JR, Techer M, Kowalik V, Oddie MAY, et al. *Varroa destructor*: A complex parasite, crippling honey bees worldwide. *Trends Parasitol* 2020;36(7):592–606.
- Peck DT, Seeley TD, Nieh JC. Mite bombs or robber lures? The roles of drifting and robbing in *Varroa destructor* transmission from collapsing honey bee

- colonies to their neighbors. *PLoS ONE* 2019;14(6):e0218392. doi: <https://doi.org/10.1371/journal.pone.0218392>.
- Zheng H, Cao L, Huang S, Neumann P, Hu F. Current status of the beekeeping industry in China. pp. 129–158. In: Chantawannakul P, Williams G, Neumann P, editors. *Asian Beekeeping in the 21st Century*. Singapore: Springer Singapore; 2018. p. 129–58. doi: https://doi.org/10.1007/978-981-10-8222-1_6.
 - Dietemann V, Beaurepaire A, Page P, Yañez O, Buawangpong N, Chantawannakul P, et al. Population genetics of ectoparasitic mites *Varroa* spp. Eastern and Western honey bees *Parasitology* 2019;146(11):1429–39.
 - Agosta SJ, Janz N, Brooks DR. How specialists can be generalists: resolving the “parasite paradox” and implications for emerging infectious disease. *Zoologia* 2010;27(2):151–62.
 - Andino GK, Gribskov M, Anderson DL, Evans JD, Hunt GJ. Differential gene expression in *Varroa jacobsoni* mites following a host shift to European honey bees (*Apis mellifera*). *BMC Genomics* 2016;17(1):926.
 - Wang S, Lin Z, Chen G, Page P, Hu F, Niu Q, et al. Reproduction of ectoparasitic mites in a coevolved system: *Varroa* spp.—Eastern honey bees, *Apis cerana*. *Ecol Evol* 2020;10(24):14359–71.
 - Lin Z, Wang S, Neumann P, Chen G, Page P, Li Li, et al. Population genetics and host specificity of *Varroa destructor* mites infesting eastern and western honeybees. *J Pest Sci* 2021;94(4):1487–504.
 - Dietemann V, Nazzi F, Martin SJ, Anderson DL, Locke B, Delaplane KS, et al. Standard methods for *Varroa* research. *J Apic Res* 2013;52(1):1–54.
 - Wang S, Lin Z, Dietemann V, Neumann P, Wu Y, Hu F, et al. Ectoparasitic mites *Varroa underwoodi* (acarina: varroidae) in eastern honeybees, but not in western honeybees. *J Econ Entomol* 2019;112(1):25–32.
 - Anderson DL, Trueman J. *Varroa jacobsoni* (Acari: Varroidae) is more than one species. *Exp Appl Acarol* 2000;24(3):165–89.
 - Techer MA, Rane RV, Grau ML, Roberts JMK, Sullivan ST, Liachko I, et al. Divergent evolutionary trajectories following speciation in two ectoparasitic honey bee mites. *Commun Biol* 2019;2(1). doi: <https://doi.org/10.1038/s42003-019-0606-0>.
 - Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25(14):1754–60.
 - McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, et al. The genome analysis toolkit: a mapreduce framework for analyzing next-generation dna sequencing data. *Genome Res* 2010;20(9):1297–303.
 - Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods* 2009;6(9):677–81.
 - Abyzov A, Urban AE, Snyder M, Gerstein M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res* 2011;21(6):974–84.
 - Felsenstein J. Phylip-phylogeny inference package (version 3.2). *Cladistics* 1989;5:164–6.
 - Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics* 2011;27(15):2156–8.
 - Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol* 2017;34(12):3299–302.
 - Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 2009;19(9):1655–64.
 - Terhorst J, Kamm JA, Song YS. Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat Genet* 2017;49(2):303–9.
 - Techer MA, Roberts JMK, Cartwright RA, Mikheyev AS. The first steps toward a global pandemic: Reconstructing the demographic history of parasite host switches in its native range. *Mol Ecol* 2022;31(5):1358–74.
 - Akimov IA, Yastrebtsov AV. Reproductive system of *Varroa jacobsoni*. I. Female reproductive system and oogenesis. *Vestn zool* 1984;6:61–8.
 - Kim D, Perteza G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013;14(4):R36.
 - Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform* 2011;12:323.
 - Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15(12):550.
 - Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 2011;27(11):1571–2.
 - Li-Byarlay H, Li Y, Stroud H, Feng S, Newman TC, Kaneda M, et al. RNA interference knockdown of *DNA methyl-transferase 3* affects gene alternative splicing in the honey bee. *Proc Natl Acad Sci USA* 2013;110(31):12750–5.
 - Wu H, Xu T, Feng H, Chen L, Li B, Yao B, et al. Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates. *Nucleic Acids Res* 2015;43(21):e141–e141.
 - Merico D, Isserlin R, Stueker O, Emili A, Bader GD, Ravasi T. Enrichment map: A network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE* 2010;5(11):e13984. doi: <https://doi.org/10.1371/journal.pone.0013984>.
 - Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13(11):2498–504.
 - Beaurepaire AL, Truong TA, Fajardo AC, Dinh TQ, Cervancia C, Moritz RFA. Host specificity in the honeybee parasitic mite, *Varroa* spp. in *Apis mellifera* and *Apis cerana*. *PLoS One* 2015;10(8):e0135103.

- [40] Roberts JM, Anderson DL, Tay WT. Multiple host shifts by the emerging honeybee parasite, *Varroa jacobsoni*. *Mol Ecol* 2015;24(10):2379–91.
- [41] Su Y-Q, Denegre JM, Wigglesworth K, Pendola FL, O'Brien MJ, Eppig JJ. Oocyte-dependent activation of mitogen-activated protein kinase (ERK1/2) in cumulus cells is required for the maturation of the mouse oocyte–cumulus cell complex. *Dev Biol* 2003;263(1):126–38.
- [42] Gosden RG. Oogenesis as a foundation for embryogenesis. *Mol Cell Endocrinol* 2002;186(2):149–53.
- [43] Melchiorri D, Cappuccio I, Ciceroni C, Spinsanti P, Mosillo P, Sarichelou I, et al. Metabotropic glutamate receptors in stem/progenitor cells. *Neuropharmacology* 2007;53(4):473–80.
- [44] Ferraguti F, Shigemoto R. Metabotropic glutamate receptors. *Cell Tissue Res* 2006;326(2):483–504.
- [45] Brockmann K, Dechent P, Bönnemann C, Schreiber G, Frahm J, Hanefeld F. Quantitative proton MRS of cerebral metabolites in laminin $\alpha 2$ chain deficiency. *Brain Dev* 2007;29(6):357–64.
- [46] Tufail M, Takeda M. Molecular characteristics of insect vitellogenins. *J Insect Physiol* 2008;54(12):1447–58.
- [47] Sappington TW, Raikhel AS. Molecular characteristics of insect vitellogenins and vitellogenin receptors. *Insect BiochemMol Biol* 1998;28(5):277–300.
- [48] McAfee A, Chan QWT, Evans J, Foster LJ. A *Varroa destructor* protein atlas reveals molecular underpinnings of developmental transitions and sexual differentiation. *Mol Cell Proteomics* 2017;16(12):2125–37.
- [49] Linder P. Dead-box proteins: a family affair—active and passive players in RNP-remodeling. *Nucleic Acids Res* 2006;34(15):4168–80.
- [50] Chen Y-J, Wang H-J, Jauh G-Y, Mittelsten Scheid O. Dual role of a SAS10/C1D family protein in ribosomal RNA gene expression and processing is essential for reproduction in *Arabidopsis thaliana*. *PLoS Genet* 2016;12(10):e1006408. doi: <https://doi.org/10.1371/journal.pgen.1006408>.
- [51] Koh T-W, Verstreken P, Bellen HJ. Dap160/Intersectin Acts as a Stabilizing Scaffold Required for Synaptic Development and Vesicle Endocytosis. *Neuron* 2004;43(2):193–205.
- [52] Snigirevskaia E, Raikhel AS. Receptor-mediated endocytosis of yolk proteins in insect oocytes. In: Raikhel AS, Sappington TW. (eds.) *Progress in vitellogenesis. Reproductive biology of invertebrates*. USA: Science Publishers. pp. 199–228.
- [53] Prévôt D, Darlix JL, Ohlmann T. Conducting the initiation of protein synthesis: the role of eIF4G. *Biol cell* 2003;95(3–4):141–56.