Divergent evolutionary trajectories following speciation in two ectoparasitic honey bee mites

Maeva A. Techer¹*, Rahul V. Rane²,³, Miguel L. Grau¹, John M.K. Roberts², Shawn T. Sullivan⁴, Ivan Liachko⁴, Anna K. Childers⁵, Jay D. Evans⁵ & Alexander S. Mikheyev¹,⁶*

Multispecies host-parasite evolution is common, but how parasites evolve after speciating remains poorly understood. Shared evolutionary history and physiology may propel species along similar evolutionary trajectories whereas pursuing different strategies can reduce competition. We test these scenarios in the economically important association between honey bees and ectoparasitic mites by sequencing the genomes of the sister mite species Varroa destructor and Varroa jacobsoni. These genomes were closely related, with 99.7% sequence identity. Among the 9,628 orthologous genes, 4.8% showed signs of positive selection in at least one species. Divergent selective trajectories were discovered in conserved chemosensory gene families (IGR, SNMP), and Halloween genes (CYP) involved in moulting and reproduction. However, there was little overlap in these gene sets and associated GO terms, indicating different selective regimes operating on each of the parasites. Based on our findings, we suggest that species-specific strategies may be needed to combat evolving parasite communities.
Interactions between hosts and parasites shape biodiversity by driving coevolutionary arms races and by regulating populations over ecological timescales. Parasitism is a remarkably successful strategy, occurring in nearly half of the recognized animal phyla\textsuperscript{1,2}. Parasites depend on their hosts which form their principal ecological niches and provide essential resources to survive\textsuperscript{3}. Commonly, multiple parasite generations elapse per host generation. Thus parasites are prone to evolve rapidly and may circumvent any gained host advantage\textsuperscript{4}. The similar selective regimes acting on phylogenetically unrelated parasites resulted in reproductive evolution such as genome size reduction, functional losses associated with reduced metabolism, and adaptive functional gains linked to parasite virulence\textsuperscript{5}. Most models of parasite coevolution are limited to interactions between single focal parasites and their hosts, in part because multispecies dynamics are much more complex\textsuperscript{6,7}. In nature, the reality is that parasite’s success and evolution will also depend on interactions with co-infecting conspecific and heterospecific parasites within its host niche\textsuperscript{8-10}. Yet, coevolutionary dynamics involving more than one parasitic species remain poorly understood\textsuperscript{11,12}.

Co-infecting parasites occupying the same host can exclude competitors via direct, or interference competition\textsuperscript{13,14}. One strategy to reduce interspecific competition can be spatial separation within regions of the host gut as adopted by cestodes that naturally co-infect sticklebacks\textsuperscript{15}. Within the same host, spatial segregation that allows parasite coexistence can ultimately lead to intrahost speciation, as found in gill parasites (Dactylogyrus) of freshwater fish (Cyprinidae)\textsuperscript{16}. Niches may also be partitioned temporally, which allows two fungal sibling species (Erysipe quercicola and E. alpithoides) to exploit the same host tree (Quercus robur) by reducing direct interference\textsuperscript{17}. Alternatively, parasites may shift their strategies to exploit different aspects of the host niche. For example, fire ants (Solenopsis spp.) are parasitized by a genus of decapitating phorid fly (Pseudacteon), which are highly host-specific and have the same life cycle, but specialize on attacking ants in different contexts (e.g., while foraging, vs. at the nest)\textsuperscript{18}. However, almost nothing is known about evolutionary trajectories of speciating parasites. On one hand, they share physiology and host-specific adaptations, which may predispose them to evolve along common lines, particularly in allopatry. On the other hand, selective regimes may be less predictable, or even disruptive, in the case of character displacement.

We investigated the evolution of speciating parasites by sequencing genomes of two economically important mite species that specialize on honey bees (Apis sp.); Varroa jacobsoni and Varroa destructor. The honey bee colony, which is a densely packed community of genetically similar individuals, can host many diseases and parasites. Varroa mites are obligatory ectoparasites which feed on fat bodies of developing larvae/pupae and adult honey bees (shown in V. destructor)\textsuperscript{19} and reproduce in the brood\textsuperscript{20}. Obviously, host-feeding by Varroa mites injures vital organs of the honey bee, which can also lead to abnormal development\textsuperscript{21} and to reduced lifespan\textsuperscript{22}. In addition, Varroa mites are known to vector several viruses\textsuperscript{23} and to transmit pathogenic bacteria\textsuperscript{24}. These combined negative effects on individual honey bee decrease the fitness of the colony as a superorganism. Without acaricide treatments, managed western honey bee colonies were reported to decline and die within two to 3 years\textsuperscript{20}. As its name indicates, V. destructor is considered as the main devastating force behind colony collapses and impact deeply western honey bee (A. mellifera) beekeeping\textsuperscript{25,26}.

The first record of Varroa mite parasitism dates back to 1904 in Java in where female mites were described from an Indo-Malayan Apis cerana honey bee colony (ex. Apis indica)\textsuperscript{27}. Not until the mid-20th century did Varroa draw attention again when they were found to infest western honey bee A. mellifera colonies introduced in far eastern Asia\textsuperscript{28}. Originally from Europe, Africa, and the Middle East, the western honey bee was allopatric to the rest of the Asian Apis species, and had no indigenous brood ectoparasites\textsuperscript{29}. The host-switched mites rapidly spread worldwide and caused dramatic losses of A. mellifera colonies. They were identified as Varroa jacobsoni until the 2000s when a key taxonomic revision of the genus described them as a cryptic sister species V. destructor\textsuperscript{30}. The taxonomic confusion existed and persisted for a while because both mites have similar life cycles and morphology\textsuperscript{30} and both originally parasitized the widespread eastern honey bee (Apis cerana). The Varroa cryptic species complex was resolved by a combination of morphometrics and mitochondrial sequencing, which identified at least four mite species: V. destructor, V. jacobsoni, V. rindereri, and V. underwoodi\textsuperscript{30}.

As western honey bees were brought into contact with the eastern honey bee in Asia and Oceania, both V. destructor repeatedly colonized this novel host, followed by V. jacobsoni in 2008 in Papua New Guinea\textsuperscript{28,31}. In contrast to the population collapses of A. mellifera, the host-parasite interaction with A. cerana is a balanced relationship in which a tolerance equilibrium is maintained by a range of parasite adaptations and host counter-adaptations\textsuperscript{32-37}. Specifically, the eastern honey bee has impressive host defense traits, including (i) self and collective grooming behaviors\textsuperscript{38}, (ii) hygienic behavior where mites are removed from infested cell\textsuperscript{39}, (iii) worker brood death that stops parasite reproduction\textsuperscript{34,35}, (iv) entombing of mites in drone brood, and (v) shorter worker development cycle limiting the mites’ offspring number\textsuperscript{33}. By contrast, most western honey bee populations lack these defensive mechanisms, and hives collapse quickly after Varroa infestation. By contrast, Varroa mites have evolved highly specialized strategies to find, select, feed, reproduce, and hide on their host\textsuperscript{40}. While many mite adaptations have already been discovered\textsuperscript{20,40}, the recent use of ‘omics approaches has quickly broadened our understanding on V. destructor hidden specialization to host\textsuperscript{37,41-43}. For example, transcriptomic and proteomic profiling of different V. destructor life stages boosted the identification of genomic regions related to chemosensory system\textsuperscript{41,43,44}, reproductive success\textsuperscript{42}, and linking oogenesis initiation or failure to parasitic reduced metabolism and host defense\textsuperscript{45,46}.

Here, we ask how both mite species evolved toward a tolerance equilibrium with A. cerana—did they follow similar or different evolutionary paths? On one hand, the two species naturally occur in parapatry, which is expected to result from divergent local adaptation, perhaps to the distribution of different A. cerana lineages\textsuperscript{47,48} or to the climate. In actuality, what drives the parapatric distribution is unclear and recent evolutionary history shows that massive host-associated range shifts are possible. For instance, V. destructor, which was historically absent from Southeast Asia, now occurs there on its novel host A. mellifera in sympathy with V. jacobsoni, which does not parasitize A. mellifera outside of Melanesia. Thus, it is possible that the current geographical distribution resulted from secondary contact following allopatric speciation. This view is further supported by the fact that both species have the genetic capacity to shift to A. mellifera as a novel host, suggesting a level of physiological similarity. We tested these alternative hypotheses (parallel vs. divergent evolution) by generating high-quality hybrid de novo genome assembly of both species and examining them for signatures of adaptive evolution. We also compared them to the recently published genome of a distantly related ectoparasitic mite Tropilaelaps mercedesae, to see whether the patterns appear universal at different timescales. We found here that although the two Varroa species share 99.7% sequence identity across the nuclear genome.
they have gone on largely dissimilar evolutionary trajectories after their split.

**Results**

**Two new reference genomes for *V. destructor* and *V. jacobsoni***.

We took advantage of the haplodiploid sex determination of both *V. destructor* and *V. jacobsoni* to generate data to reduce the ambiguity due to heterozygous sites for a high contiguous hybrid genome assembly. The sequencing strategy adopted in this study combined short reads (Illumina) for which input DNA can be obtained from a single mite individual, long reads of tens of kilobases (PacBio), which requires a large quantity of high-molecular-weight DNA and Hi-C to capture chromosome conformation and interactions. We generated 28.6 Gb (333 ± S.D. 56 bp) and 56.2 Gb (409 ± S.D. 61 bp) of de-duplicated, decontaminated, and merged Illumina reads for *V. jacobsoni* and *V. destructor*, corresponding to a coverage of 80x and 152x, respectively. PacBio sequencing yielded 20.3 Gb of *V. destructor* reads with an N50 of 13.5 kb (66× coverage), which was used to fill 84.4% of the gaps in the scaffolded HiC assembly. Final de novo assemblies were named Vdes_3.0 (GCA_002443255.1) and Vjacob_1.0 (GCA_002532875.1) and are available on NCBI.

Through the k-mer analysis, the genome size estimates ranged from 369.02 to 365.13 Mbp using k = 42 (Supplementary Fig. 1). Both Vdes_3.0 and Vjacob_1.0 final assembly size were close to the predicted estimates but *V. jacobsoni* genome size was slightly shorter than *V. destructor* with 365.59 < 368.94 Mbp, respectively (Table 1). Previous assemblies size and contiguity for *V. destructor* with Vdes_1.0 and Vdes_2.0 were improved as detailed in Table 1. According to the N50 statistics, more than 50% of the genome was in scaffolds of 58.5 Mbp or more for Vdes_3.0 while the Vdes_2.0 N50 was around 0.1 Mbp. When comparing to invertebrate assemblies, Vdes_3.0 genome is among the most highly contiguous with to date the best-reported scaffold N50 and best contig N50 for Acari and Mesostigmata as reported on the i5k database (i5k.github.io/arthropod_genomes_at_ncbi).

Using Hi-C data, seven major scaffolds were constructed for *V. destructor* ranging from 39.4 to 76.9 Mbp with additional of 1,418 nuclear minor scaffolds > 0.2 Mbp. The longest genomic scaffolds correspond to the seven haploid chromosomes (2n = 14) previously detected in *V. destructor* karyotype. Contrary to Vdes_3.0, which benefited from long reads and chromosome capture sequencing, *V. jacobsoni* Vjacob_1.0 genome was assembled from 8241 contigs into 4881 scaffolds with an N50 of 233,810 bp. Similar GC content composed 40.9% were found for both sister species Vdes_3.0 and Vjacob_1.0. The automated NCBI Eukaryote annotation pipeline predicted 10,247 protein-coding genes (out of 12,846 genes and pseudo-genes) for Vdes_3.0, and 10,739 protein-coding genes (out of 15,486 genes and pseudo-genes) for Vjacob_1.0 (based upon transcriptome data available on the database in October 2017). When single haploid male of each Varroa species reads were mapped onto the Vdes_3.0 reference genome, a total of 1,404,212 quality-filtered SNPs were detected. Relative to Vdes_3.0 genome size, sequence divergence with Vjacob_1.0 amounted to 0.38% across the nuclear genome sequence.

**Identification and confirmation of Varroa mite species and mtDNA lineages.** Mapping and pairwise whole mitochondrial alignment (16,505 bp) showed that Vdes_3.0 diverged of only 0.1% to the reference *V. destructor* mitogenome (NC004454.2). Despite presenting similar morphological phenotypes mainly distinguishable through female body size (Fig. 1), the *V. jacobsoni* mitochondrial genome diverged from Vdes_3.0 mtDNA at 5.1% of all nucleotides.

Given that COX1 is a standard marker used to identify the *Varroa* mite haplogroup, this region was extracted and realigned with 26 described sequences from native and introduced populations. Divergence within the complete COX1 sequences
of the *V. destructor* and *V. jacobsoni* reference genomes rose up to 5.9%. The partial COX1 sequence (426 bp) of Vdes_3.0 clustered with *V. destructor* sequences and was clearly distinct from Vjacob_1.0, which clustered with *V. jacobsoni* haplotypes (Fig. 2a, b). The concatenated COX1, ATP6, COX3, and CYTB sequence (2696 bp) from the Vdes_3.0 mitochondrial genome was identical to that of the invasive Korean K1-1/K1-2 sequence (Fig. 2c).

Orthologue genes between *Varroa* and other Parasitiformes. The arthropod genome database at NCBI presently contains only two other published representative Mesostigmata genomes: the brood parasitic honey bee mite *Tropilaelaps mercedesae* (GCA_002081605.1) and the non-parasitic Western predatory mite *Metaseiulus (=Typhlodromus or Galendromus) occidentalis* (GCA_00025335.1). Orthologue prediction was performed by comparing annotated *Varroa* genes to those of the two Mesostigmata species gene sets and those of two additional Acari: the parasitic tick *Ixodes scapularis* (GCA_000208615.1) and the free-living two-spotted mite *Tetranychus urticae* (GCA_000239435.1). The tick presents the largest set of genes (n = 20,486) and also possessing a haploid genome more than four times larger 1.76 Gb. Orthology prediction analysis using the six genomes clustered the genes by domain into 7502 super-orthologue groups consisting of orthologues and inparalogues. A total of 11,123 orthologue groups were predicted and 42.8% were shared by the five parasitiformes and the outgroup spider mite *Tetranychus* (Supplementary Data 1). The concatenated phylogenetic tree built on the genomic sets based on 4758 genes confirmed the close relationship between the three mite species compared to other Acari (Fig. 3). The honey bee ectoparasitic mites, *Varroa*, and *Tropilaelaps* shared 813 lineage-specific orthogroups (akin to gene families) of which 65.9% were Varroidae-lineage-specific. The two sibling *Varroa* species shared 9628 orthologous genes including 536, which were private to Varroidae.
Orthologous genes specific to the Parasitiformes ancestral branch (i.e., before *T. urticae* and other Acari species split) was enriched for molecular functions and biological processes involved in several transporters and channel activities (Supplementary Data 2). When considering the Mesostigmata ancestral branch (i.e., before *M. occidentalis* and the three honey bee parasites species split), GO terms related with structural constituents of the cuticle were the most represented. At the Varroidae ancestral branch level (i.e., before the two species split), orthologous genes showed GO terms mainly related to endonuclease activity.

Divergent patterns of selection among honey bee mites. A total of 234 and 225 genes were detected under positive selection for *V. destructor* and *V. jacobsoni* branches, respectively, though only 13 were shared among this list (2.8%) (Fisher’s exact test \( p = 0.003 \)) (Fig. 3). Additionally, 40 genes were detected under positive selection when looking at the Varroa ancestral branch, i.e., before the two species split. On these set combined, *V. destructor* and *V. jacobsoni* presented 81.0% and 80.4% species-specific genes with positive selection signals, respectively (Supplementary Fig. 2 and Supplementary Data 3). The remaining proportion of positively selected genes shared involved the other honey bee mite *T. mites* rather than with any other Acari. The genomic regions associated with genes under positive selection were found and distributed along all major chromosomal sites of *V. destructor*, and, assuming preservation of synteny, in *V. jacobsoni* as well (Fig. 4a).

Semantic distribution of the GO terms for the genes under positive selection in each *Varroa* species showed non-significant overlap \( (p = 0.385 > 0.050) \). The few terms that overlapped involved biological processes in essential functions such as the molting cycle—chitin-based, fatty acid, and lipid metabolism and responses to pH (Fig. 4b). More specifically, chitin is the main constituent of the mite cuticle hardening it against both bees and the environment. Its metabolism varies throughout the mite life cycle, especially during reproduction or moltng\(^2\). Fatty acid metabolism could be related to the feeding on the fat body on the host, as well as oogenesis. Although few overlaps exists when GO terms list were reduced using Resnik’s measure, the REVIGO scatter plot and treemap visually suggest functionally different routes to host adaptation (Fig. 4b, Supplementary Data 4). In *V. destructor*, GO terms were mainly associated with (1) regulation of membrane depolarization, (2) retinal cone cell development and myofibrilubt differentiation, (3) mRNAS cis splicing via spliceosome whereas GO terms in *V. jacobsoni* related to (1) protein processing involved in targeting to mitochondria, (2) vitamin K metabolic process, (3) response to pH and other processes detailed in (Supplementary Fig. 3).

Tandem duplication was detected in 51 genes for *V. destructor*, 35 in *V. jacobsoni* and 70 in *T. mites* (Fig. 3). Numbers of duplication events and gene clusters within a species varied from two to six units and generally involved adjacent chromosomal locations in *V. destructor* (Fig. 5a, Supplementary Data 5). In duplicated genes found only in *V. jacobsoni*, 23 had no orthologs in the *V. destructor* genome. However, most duplications are on the same scaffold with the exception of the ABC gene family where genes are present on two scaffolds. Manual checks of duplicated clusters with more than two genes validated all of the duplications except for the Histone 2A family where LOC1112553102 gene located on the minor scaffold BEIS0100444.1 was removed since it is 100% identical to LOC111259948 located on chromosome 1 and may be an artefact of mis-assembly.

Only five duplicated gene clusters were shared by all three honey bee parasites. These amounted to 9.8% and 14.3% of
detected tandem gene duplications for *V. destructor* and *V. jacobsoni*, respectively (Fig. 2, Supplementary Data 6). GO terms associated with these duplicated gene clusters related to chitin, amino sugars, drug and carbohydrate derivatives. Conversely, the two *Varroa* sister species shared 14 gene clusters undergoing tandem duplication, representing 27.5% and 40.0% of the total size for *V. destructor* and *V. jacobsoni*, respectively. When summarizing the list of duplicated shared genes by both *Varroa* species, GO terms were associated with biological processes in glycerophospholipid catabolism, body development, regulation of DNA-templated transcription/elongation and amino-sugar metabolism (Supplementary Data 7). Finally, 62.7% of the duplicated genes were exclusive to *V. destructor* whereas 45.7% were only detected in *V. jacobsoni*. Genes duplicated in *V. destructor* were associated with GO terms related to biological processes such as mechanical/external stimulus, regulation of endoribonuclease activity and phagocytosis (Fig. 5b, Supplementary Data 7-8). In contrast, genes duplicated in *V. jacobsoni* were involved in biological processes of the skeletal myofibril assembly, Golgi calcium ion transport, and striated muscle contraction (Fig. 5b).

**Fig. 3** Positive selection and gene duplication of orthologous genes in parasitic and free-living Acari. For the honey bee mite parasites, host ranges are shown as checkboxes and individual records had to meet the following conditions: (1) Mites had to be found reproducing or observed in bee colonies in independent surveys, and (2) the identity of cryptic *Tropilaelaps* and *Varroa* species was confirmed by molecular markers (e.g., COX1 barcoding). The number of orthologous genes in each lineage is circled at the nodes. The number of positively selected genes for each branch is shown in green, with the number of species-specific genes shown in purple. Gene duplications within honey bee parasitic mites are shown in boxes: white squares = shared genes, black squares = *Varroa*-specific genes by, red and blue squares = *V. destructor* and *V. jacobsoni* specific genes, respectively. The two *Varroa* mites have similar host ranges, through *V. jacobsoni* also occurs on *A. nigrocincta*, a close relative of *A. cerana*, possibly as a result of a recent secondary host shift. Despite sharing the same ancestral host, the *Varroa* sister species show different patterns of adaptive molecular evolution and exhibit different gene expansion patterns, suggesting different evolutionary trajectories.

**Chemosenory and cytochrome P450 genes in honey bee mites.**

**Lipid carrier proteins:** Based on the comparison of six amino acid sequences of OBPs previously described from *V. destructor* legs transcripts, we found seven homologs (with conserved domain IPR036728 or OBP-like) in *V. destructor* annotated genome (Supplementary Fig. 4). Each OBP gene found in *V. destructor* presented an ortholog in *V. jacobsoni* coding for an identical OBP sequence (Supplementary Data 9). An analogous pattern of strong conservation was found for eight NCP2s (IPR003172) or similar with a Ganglioside GM2 activator conserved domain (IPR028996) in both sibling Varroa mites (Supplementary Fig. 5). None of the genes encoding for OBP and NCP2 showed signatures of positive selection or tandem duplication in Varroa sister species.

**Membrane-bound receptors:** A total of four genes per Varroa genome were found to encode for GR proteins presenting either the 7m_7 or Trehalose receptor conserved domains (IPR013604 and IPR009318, respectively). All four GR were closely related to GRs proteins found either in the honey bee parasitic mite *T. mercedesae* or in the parasitic tick *I. scapularis* (Supplementary Fig. 6). As previously reported, IGR protein profiles are the largest in *V. destructor*. According to the IGR phylogeny, this is also the case for *V. jacobsoni* with 35 identified genes (mainly presenting IPR01828, IPR019594 or IPR001320 signatures) (Fig. 6). Among these, 19 proteins found in Varroa genomes had closest homolog in *T. mercedesae* indicating high conservation of the gene family within honey bee parasites. Interestingly, three genes encoding for IGR or IGR-like proteins out of the 37 identified in *V. destructor* had no orthologous in *V. jacobsoni*. The opposite situation was also true but only for one gene in *V. jacobsoni* (LOC1112727277). Additionally, divergent selective trajectories were found as a glutamate [NMDA] receptor subunit 1-like gene (LOC111270396) was under positive selection for *V. jacobsoni* while a metabotropic glutamate receptor 3-like gene (LOC111245193) show positive selection signal for *V. destructor*. Finally, despite a small number of SNMPs (with conserved domain IPR002159) detected in *V. jacobsoni*, we also found that one scavenger receptor class B member 1-like gene (LOC111267160) was under positive selection but none for *V. destructor* (Fig. S7).
**Fig. 4** Genes and pathways under positive selection in the *Varroa* sister species. 

**a** Red and blue bars (in 5 kb windows) represent locations of genes in *V. destructor* and *V. jacobsoni*, respectively. *V. jacobsoni* data are mapped to *V. destructor* scaffold positions for comparison. Black asterisks indicated genes under positive selection, shared by both species (n = 12) and detected in the *Varroa* ancestral lineage prior to the split of the two species (n = 40).

**b** Semantic space analysis of significantly enriched GO terms (BP, CC, and MF) over-represented in genes detected under positive selection for *Varroa* mites. Bubble color indicates the species for which GO terms were enriched (all p-value < 0.05) and size indicates the frequency of the GO term found in the GOA database. There was little overlap in analyses at both gene (2.8%) and functional levels (0.8%), suggesting different selective pressures on the two sister species since they split.
**Fig. 5** Duplicated genes in *V. destructor* are found throughout the genome and are involved in different biological pathways. **a** Chromosomal location and similarity of duplicated genes (arrows) for *V. destructor*. Different genes and chromosomal regions underwent duplication in the two species. **b** Cluster analysis of significantly enriched GO terms for biological processes over-represented among duplicating genes in Varroa mites. Not only were most of the GO terms species-specific, but they also comprised non-overlapping categories of biological processes. Neither gene duplication nor selection analysis suggests a substantial degree of parallel evolution in these mites.
Ecdysteroids and Halloween genes: All CYP enzymes sequences from *V. destructor* and *V. jacobsoni* annotation were extracted by filtering gene description list containing cytochrome p450 or exhibiting the associated conserved domain cl12078 or pfam00067. A total of 36 and 38 CYP (one isoform per gene) were aligned together with the seven Halloween genes reported from *D. melanogaster* and *A. aegypti* and other CYPs from Acari. Previously, only some members of this pathway, namely orthologues of *spook*, *shade*, and *disembodied* were found from a genomic sequence generated by 454 pyrosequencing. Homologs proteins of CYP302A1 and CYP314A1 (~32% and ~28% sequence identity to *disembodied* and *shade* of *D. melanogaster*, respectively) were found for both Varroa sibling species but surprisingly no homolog *spook* (CYP307A1) was found for *V. jacobsoni*, though there was one in *V. destructor* (Fig. 7). Phylogenetic relationships among Varroa mites, *T. urticae* and *M. occidentalis* and *D. melanogaster* CYP315A1 sequences (~36% identity with the fruit fly) suggest that both *V. destructor* and *V. jacobsoni* Varroa mites possess *shadow* homologs. Likely, *shroud* and *phantom* are also presents in both Varroa species but not *neverland*. In the CYP family, we detected three genes under positive selection only in *V. jacobsoni*, including *shadow-like* (LOC111271433) and *disembodied* (LOC111265704) (Fig. 7 and Supplementary Data 9).

Other genes possibly related to nutrition and detoxification: Some annotated genes detected under positive selection or undergoing tandem duplication in Varroa have been linked to play a role in external stress tolerance (e.g., in-hive temperature), nutrition, molting, reproduction, and metabolizing toxic xenobiotics possibly leading to acaricide resistance (detailed in Table 2). On *V. destructor* chromosome 4 (NW_019211457.1), two duplicated genes were also under positive selection and were involved in cuticular proteins pathway. These genes are involved in physical properties of the mite cuticle, which in addition to providing protection may constitute the first barrier to reduce the penetration of external chemical agents. On the other hand, two duplicate clusters were also detected undergoing positive selection and annotated as part of the ABC family (ATP-Binding Cassette sub-family A) and to aminopeptidase M1-B-like, in *V. jacobsoni*. Finally, the most remarkable tandem duplication occurred on Vdes_3.0 chromosome 1 (NW_019211454.1) with six duplicated genes coding for heat shock 70kDa protein. These were not detected for *V. jacobsoni* (Table 2). Both positive gene selection and duplication patterns in the *V. destructor* and *V. jacobsoni*
Fig. 7 CYP repertoire in Varroa mites is broader than expected with two functionally important Halloween genes shade and disembodied under positive selection in V. jacobsoni. The phylogenetic tree was constructed using 174 amino acid sequences from Acari V. destructor (Vdes = genome and Var = protein sequences), V. jacobsoni (Vjac), T. mercedesae (Tmer), M. occidentalis (Mocc), I. scapularis (Isca), and T. urticae (Turt). Additionally, sequences from the fruit fly Drosophila melanogaster (Dmel), mosquito Aedes aegypti (Aaeg) and the honey bee A. mellifera (Amel) and A. cerana (Acer) were downloaded from NCBI and aligned with MAFFT. Best fit model computed for the tree is LG + R5 using IQ-TREE. Bootstrap values were estimated using a SH-like aLRT with 1000 and bootstraps over 95% are shown by a purple circle. V. destructor was previously believed to have only disembodied (CYP302a1), shade (CYP314a1), and spook (CYP307a1) homologs of the seven Halloween genes (background in black).

Discussion

In this study, we developed high-quality de novo reference genomes and annotations for economically important honey bee parasites V. destructor and V. jacobsoni. Because Varroa mites are, to a large extent, responsible for the global honey bee crisis, and research efforts worldwide target these parasites, these genomic resources respond to a pressing need. Improved genome assembly and annotation with transcriptomic data from different life stages produced the most comprehensive gene catalogue to date, as reflected by our discovery of previously missing Halloween genes homologs. The supposed absence of such genes has led to speculation on control measures development or that host counter adaptation could in the future target this pathway. Our analysis also revealed that the two Varroa species share 99.7% sequence identity across the nuclear genome. Yet, after the two species have split, they have gone on largely dissimilar evolutionary trajectories, as evidenced by different patterns of positive selection and gene duplication. In particular, signals of positive selection were found in functionally important genes involved in chemosensing, molting, and reproduction. Thus, it seems likely that the two species underwent different patterns of adaptation while continuing to exploit the same host, the eastern honey bee (A. cerana). Differences between the two species most likely result from local adaptation to host geographic subspecies, the environment or their interaction. Though not enough data exist to distinguish between these alternatives, below we summarize what is already known and outline promising avenues for investigation.

Inferring what constitutes the ancestral ecological niche of the two mites is key to understanding what selective pressures they have faced and how they responded to them. V. destructor and V. jacobsoni are very similar morphologically, and until 20 years ago...
they were mistakenly considered the same species. Their host specificity overlaps as they both parasitize the eastern honey bee A. cerana throughout Asia. The only exception reported for the Indonesian A. nigrocincta infested by V. jacobsoni, though possibly as a result of a recent host switch from A. cerana and A. melifera, might have also been a parasite of the A. cerana lineage. This hypothesis is supported by the host range of V. underwoodi, another closely related mite that also parasitizes A. cerana in addition to A. nigrocincta and A. nulens. With no known varroa fossils, no information exists about speciation timing in Varroa mites. Based on the COXI gene, a recognized molecular clock for insects with a divergence rate of 2.3% per Myr, the divergence between V. destructor and V. jacobsoni of 5.9% suggests that the two species split 2–3 Myr ago. This is after the split between A. cerana and A. melifera, which occurred 6–9 Myr ago based on COXI and ND2 to 17–19 Myr using a combination of mitochondrial, ribosomal, and nuclear loci. Thus, host ranges and molecular estimates suggest that most Varroa evolution occurred on ancestors of A. cerana.

The major ecological difference between these mites is their parapatric distribution (Fig. 1). Whether parapatry resulted from their respective biogeographic history, adaptation to different subspecies of A. cerana, or is shaped by competition is not known. In fact, even the extent of the geographical overlap between the mite distributions is uncertain, since few surveys actually used molecular markers for identification and most have focused on associations with A. melifera. Yet, parapatry is clearly not due to the inability of the mites to complete their phoretic life stage outside the hive, as host-switched V. destructor successfully expanded its range following the range extension of A. melifera. While allopatry could result from adaptation to local hosts/conditions or competitive exclusion, both mites have been reported from the same Thai apiary on A. cerana. Unfortunately, there are not enough data on A. cerana infestation under natural conditions to determine the likelihood that the two Varroa species interact and compete. One survey in southern Thailand found 98% parasitism by V. jacobsoni under natural conditions, albeit at variable levels of infestation, suggesting that encounters between the two species are indeed likely. If the divergent evolutionary history of the two Varroa species resulted from competition, we predict that long-term coexistence within a hive of A. cerana should be unlikely, a hypothesis that we hope will be tested in the future in the zone of sympatry.

In addition, or as an alternative to interspecific competition, it is possible that the divergence between V. destructor and V. jacobsoni was driven by local adaptation to different A. cerana populations. For instance, V. destructor was found to infest the Northern, Himalayan and the Indo-Chinese A. cerana morphoclusters identified whereas V. jacobsoni was reported on the Indo-Chinese and the Indo-Malayan one. These A. cerana populations differ in their nesting behaviour, such as the number of combs or even in their body size. In China alone, A. cerana populations have been shown to be as genetically divergent for more than 300,000–500,000 years as recognized subspecies of A. melifera. In addition to genetic differences between the hosts, the climactic environments across Asia are highly variable and could result in local adaptation, though, as mentioned earlier, this did not prevent range expansions after switching hosts (Fig. 1). Consequently, the dissimilar selective pressures observed in the mites may have resulted from selection due to different host genotypes.

Understanding the dynamics of co-infestation by both mites will become important if the host-switched V. destructor and V.

---

Table 2 Genes undergoing positive selection or duplication, that have been implicated in tolerance to external stressors and stimuli

|                      | V. destructor (Vdes_3.0) | V. jacobsoni (Vjac_1.0) |
|----------------------|-------------------------|-------------------------|
|                      | Annotated gene under positive selection | Annotated gene undergoing duplication | Annotated gene under positive selection | Annotated gene undergoing duplication |
| Stress from in-hive temperature | Heat Shock 70 Kda protein-like | 6 | 1 |
| Nutrition            | Guanine metabolism      |               | 1 |
| Detoxification       | Glutathione-S-transferase | 80            | 1 |
|                      | Sodium channel related  | 125           | 1 |
|                      | Transmembrane proteins  | 3             | 2 |
|                      | G-protein coupled receptors | 1          | 4 |
|                      | Esterases | 126          | 2 |
|                      | GABA-gated ion channels | 127         | 1 |
|                      | ATP-binding cassette gene family | 128 | 2 |
| Molting and reproduction | Ecysteoides pathway with cytochrome P450 | 45 | 3 |
|                      | Cuticular protein       | 2             | 4 |
|                      | Nutrition and reproduction | Fatty acid metabolism | 3 |
|                      | Chemosensory receptors  |               | 1 |
|                      | Sensory neuron membrane protein |               | 1 |
|                      | Ionotopic glutamate receptor | 1           | 1 |

Research on Acari and V. destructor, in particular, has identified a large number of genes involved in stress and detoxification (possibly leading to acaricide resistance), nutrition, reproduction and chemosensing in interaction with the host. Intriguingly, many of these genes appear to be under ancestral positive selection, as V. jacobsoni has not been extensively targeted by acaricides on A. melifera, and most of the evolutionary history captured by positive selection analyses occurred in A. cerana, given that coevolution with A. melifera is a relatively recent evolutionary phenomenon, particularly for V. jacobsoni.
jacobsoni are ever found to co-occur on A. mellifera. Presently, host-switched V. jacobsoni is restricted to Papua New Guinea, where V. destructor is absent31. However, if its range expands, it may impose an additional parasitic burden on A. mellifera, but whether the effects of this extra parasite are cumulative or whether the parasites will simply replace each other locally is unknown. Our data suggest that if the parasites are able to tolerate each other within A. mellifera colonies, given their different suites of molecular adaptations, they may in fact impose a greater cost than either parasite alone. Clearly, this is another avenue where further experimental work is urgently needed.

Likewise, it is unknown whether chemical treatment options developed for V. destructor will also be effective against V. jacobsoni. Interestingly, both mitre sequences show signs of selection on a range of genes previously found to be involved in stress and pesticide resistance (Table 2). Since V. destructor has evolved resistance to acaricides, it cannot be ruled out that positive selection resulted from recent acaricide pressures rather than coevolution with their host. By contrast, acaricides are not yet widely used on V. jacobsoni, though the number of genes under selection is also significant, suggesting future monitoring. In the case of V. jacobsoni, and possibly generally, these genes have the potential to become co-opted for pesticide resistance. This has been the case with pesticide tolerance in the sheep blowfly (Lucilia cuprina), where sequencing of museum specimens has shown that resistance variants pre-dated pesticide use75. Similarly, the two-spotted mite (Tetranychus urticae) appears to have co-opted existing detoxification machinery in a selective response to miticides76. However, more detailed population-level work comparing switching and unswitched populations of mites will be necessary to test this hypothesis. Museum specimens or other historical material would be particularly useful in understanding evolutionary trends.

Despite the economic importance of brood parasitic mites in A. mellifera, it is clear that various pressing research needs still exist concerning their biology, cryptic diversity in the native range, possible geographic ranges of host-switched species, and how co-infections may proceed. These new reference genomes will allow researchers to pursue such investigations, develop better tools for Varroa control, and to better understand the invasion and evolution of Varroa on A. mellifera. In particular, we hope that these new resources will greatly benefit future studies to track the population genetic changes and demography of V. destructor’s worldwide spread, which has been difficult due to the low polymorphism of previously available markers70. More Varroa spp. are continuously testing new host78,79 and may remain undetected. Finally, the recent boom in -omics methods for the study of honey bee parasites and diseases underlines the need to improve our knowledge of genome structure and organization to investigate the genetic basis of newly reported acaricide resistance, virulence level or even new host tolerance. Such knowledge combined with progress in novel manipulation tools, such as RNAi80,81, and more experimental tools like gene drive82, may provide ways of specifically controlling mite pests without harming honey bees or other insects.

**Methods**

**Sequencing strategy and Varroa mite collection.** Haploid males (0.7 mm by 0.7 mm) are only found inside capped brood cells after initiation of reproduction while diploid females (~1 mm by 1.5 mm) can be found in the brood and on adult honey bees. All V. destructor samples were collected from A. mellifera colonies of Okinawa Institute of Science and Technology (OIST) experimental apiary in Japan (26.477 N 127.835 E) in August–September 2016. One mature male V. destructor was collected within a worker cell at the red-eye pupa developmental stage (day 12−16) using a brush and kept in absolute ethanol at −80 °C. Mature adult females were collected from adult honey bee workers. In order to obtain a large number of mites infecting adult workers, we modified the standard powdered sugar method35 for an entire managed colony as follows: one hive box containing three to four frames with honey bee workers and no queen was placed on top of a collecting box. Approximately 500 g of icing sugar was then placed on filter powder strainer; the sugar and honey bees were shaken in the tray for 2 min and then separated using a sieve filter, placing the sugared-workers directly within the hive. This process was repeated on adjacent colonies. Varroa mites were separated from their hosts and trapped in the icing sugar. Following a water rinse on a gauze mesh, a total of 1207 alive Varroa females were snap frozen at −80 °C for laboratory processing (inactive, sluggish or dead mites were discarded). Samples of V. jacobsoni were obtained from its first detection survey where it was found reproducing in the western honey bee31. One mature male (VJ856) collected from A. mellifera drone brood cell in Papua New Guinea (26.477 N 127.835 E) in August 2016 was used for a whole-genome library.

One sclerotized mature female V. destructor (VDOI-K1) collected from OIST apiary in the reproductive phase and one sclerotized mature female V. jacobsoni (VJ347-11) collected from A. cerana originating in Java (Cililin central Java, 13/06/1998 by D. Anderson) were processed for X-ray microtomography (micro-CT) using a Zeiss Xradia 510 Versa scanner. 3D surface models were reconstructed from micro-CT outputs using Amira software to visualize the slight morphological differences when focusing on idiosoma ventral and dorsal views and whole body in profile (Fig. 1a).

DNA extraction, genome sequencing, assembly, and annotation. Genomic DNA was extracted from each single haploid male V. jacobsoni and V. destructor (VDOI-K1) using Qiagen DNA Micro Kit. Short-insert (100–200 bp) paired-end reads were then sequenced using Illumina HiSeq 2500 at the OIST sequencing center in paired-end 250 cycle mode. Prior to assembly paired-end reads were de-duplicated using ParDRe83, and host and microbial contamination was filtered using DeConSeq (parameters: -i 97 -c 0.05) screening against genomes of Apis cerana, Apis mellifera, as well as bacterial genomes and genomes of bee diseases and parasites. The cleaned reads were then merged using PEAR (parameters: -min-overlap 10 –memory 80 G-threads 10 -n 200 -m 600 -p 0.0001)38 and assembled using Newbler (v. 2.6) (parameters: -large -mcpu 24 -m -ratio 390 δ -i 1000 -a 5000 -u -rtd -novs 10000 -s 150000 -o 95 -m 1000). The complete genome datasets were then filtered using DeConSeq (parameters: -i 97 -c 0.05) to remove known RNA-seq data assembled using Trinity87,88. For V. destructor further scaffolding was carried out using Hi-C, performed by Phase Genomics. Hi-C read pairs were aligned to the initial assembly using BWA-MEM with the -t option89, keeping only primary properly paired alignments (samtools with the -F 2316 filter)90. Scaffolding was then performed using Phase Genomics’ Proximo Hi-C scaffolding pipeline91, which builds on the LACHESIS method82.

To estimate genome characteristics, K-mers from cleaned reads were counted using jellyfish version 2.2.793 for the following values of K: 19, 25, 31, and 42 and coverage cutoff of 1000. Frequencies were then computed on HiSeq 2500 reads and mapped per individual haploid species. Contigs with K-mers histograms for k = 42 were analyzed with GenomeScope to estimate genome size, kcover, number of duplicates91. We used the bioinformatically derived genome sizes for genome coverage estimation.

Small batches of adult females V. destructor (1,207 mites total) were pooled and crushed in 1.5 mL Eppendorf tubes. The powder was then lysed using proteinase K and lysis buffer at 56°C for 20 min. Sample lysate was then transferred to a Maxwell Cartridge. The remaining step was fully performed on the Maxwell 16 automated purification system (Promega) with the LEV Blood DNA mode. Library preparation for PacBio RSII sequencing was performed using SMRTbell Template Prep Kit 1.0 (Pacific Biosciences) following the manufacturer’s recommendations. The manufacturer’s RNA-seq data, as well as extensive protein homology data stored in GenBank52.

**Identification of Varroa mite strains using mtDNA sequences.** Most previous taxonomic work relied on mitochondrial fragments. We used mitochondrial analysis to identify and compare the lineages of Varroa species used for genome sequencing to the previously described haplotypes. The whole mitogenome of V. destructor (accession NC_004454.234) was available from the NCBI database and used as a reference for read mapping. Raw Illumina reads for both V. destructor and V. jacobsoni were mapped onto the entire reference sequence (16,476 bp) using NextGenMap 0.5.097. Mapped reads were sorted and PCR duplicates were removed using Samtools 1.3.198. Maximum coverage of 500x was subsampled on all mapped reads using VariantBar 1.4.398. Subsequently, variant calling was
performed using FreeBayes v1.1.0-469, with the following parameters others than the default: minimum mapping quality of 20 and minimum base quality of 5. Consensus Vcdflib barcode filtering was used to remove non-uniformly expressed genes. We then aligned each family of protein sequences with other sequences from other Acari and insects using MAFFT version 7115 (parameters: E-INS-i, the BLOSUM62 matrix, and MAFFT-homolog option with E threshold = 0.1). We used IQ-TREE116 to find the best substitution model for each family alignment, with state frequency estimated by maximum likelihood, 1000 ultra-fast bootstraps, 1000 replicate for SH-aLRT branch test. All generated phylogenetic trees were visualized and exported using iTOL v4117.

We ran the same process for the cytochrome P450 enzyme (CYP) genes (including those from V. destructor and V. jacobsoni), known to be involved in the steroid molting hormone ecdysone biosynthesis pathway in Arthropods. Ecdysteroids derivatives are also important as they are involved in the vitellogenin production and oogenesis initiation in V. destructor after feeding on its host45,118. To assess which CYP genes and in particular the Halloween genes were present in both V. destructor and V. jacobsoni genomes, we used sequences from the fruit fly Drosophila melanogaster, the mosquito Aedes aegypti, both honey bee hosts A. mellifera and A. cerana, as well as others Acari (same name code as in a previous study on M. occidentalis119).

Statistics and reproducibility. GO terms enrichment was conducted with the R package GOstats using the hypergeometric test for association of categories and genes. The test parameters for each species and each ontology (BP, CC, and MF) using gene ID from NCBI were as follow: p-value cutoff = 0.05, not conditional and with detection of over-represented GO terms (testDirection = >). Regarding the test conducted for the significance in gene overlap, a Fisher’s exact two-tailed test with GeneOverlap R package and can be reproduced directly from the online supplementary data.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
All data generated or analysed during this study are included in this published article (and its supplementary information files). V. destructor (Vdes_3.0) and V. jacobsoni (Vjacob_1.0) assembled reference genomes and annotation are available on NCBI database respectively with the accession number GCF_002443255 and GCF_002328725. These genomes are also available for search and annotation at the i5k Workspace40 (https://i5k.nal.usda.gov/). Raw Illumina and PacBio reads are also available in the Sequence Read Archive (SRA) under the Bioprojects PRJDR66279 and PRJNA391052. Raw data of both Varroa mites micro-CT scans are available upon request.

Code availability
Online supplemental codes and map are available with an R markdown hosted from on github.com/MikheyevLab/varroa-denoovo-genomes410. Received: 20 January 2019; Accepted: 10 September 2019; Published online: 01 October 2019

References
1. de Meelis, T. & Renaud, F. Parasites within the new phylogeny of eukaryotes. Trends Parasitol. 18, 247–251 (2002).
2. Weinstein, S. R. & Kuris, A. M. Independent origins of parasitism in Animalia. Biol. Lett. 12, 20160324 (2016).
3. Poulin, R., Krasnov, B. R., Mouillot, D. & Thieüges, D. W. The comparative ecology and biogeography of parasites. Philos. Trans. R. Soc. Lond. B Biol. Sci. 366, 2379–2390 (2011).
4. Kochin, B. F., Bull, J. J. & Antia, R. Parasite evolution and life history theory. PLoS Biol. 8, e1000524 (2010).
5. Poulin, R. & Randhawa, H. S. Evolution of parasitism along convergent lines: from ecology to genomics. Parasitology 142(Suppl 1), S6–S15 (2015).
6. Rigaud, T., Perrot-Minnot, M.-J. & Brown, M. J. F. Parasite and host assemblages: enhancing the reality will improve our knowledge of parasite transmission and virulence. Proc. Biol. Sci. 277, 3693–3702 (2010).
7. Hempel, P. S. Evolutionary parasitology: the integrated study of infections, immunology, ecology, and genetics. (Oxford University Press, 2011).
8. Petney, T. N. & Andrews, R. H. Multiparasite communities in animals and humans: frequency, structure and pathogenecic significance. Int. J. Parasitol. 28, 377–393 (1998).
9. Vasmourin, E., Vourch, G., Gasqui, P. & Vayssier-Taussat, M. The importance of multiparasitism: examining the consequences of co-infections for human and animal health. Parasit. Vectors 8, 545 (2015).
Acknowledgements
We would like to acknowledge Yoann Portugal (Ecology and Evolution OIST) for helping with the fieldwork collection of V. destructor mites, Jo Tan (Ecology and Evolution OIST) for processing in laboratory Varroa males and Hiroki Goto (OIST DNA sequencing section – Onna, Okinawa) for leading and processing the PacBio RSII sequencing. We also thank Francisco Hita Garcia (Biodiversity and Biocomplexity OIST) for his support in preparation and scanning of specimens using the micro-CT. We thank Steven Aird (OIST technical editor) for carefully checking and helping to correct our manuscript. This work was supported by the Okinawa Institute of Science and Technology Graduate University, as well as grants from the Japan Society for the Promotion of Science (Kakenhi grants 16KK0175, 16H06209 and 18H02216) and the Australian Research Council (Future Fellowship Scheme FT160100178). Genome assembly was partially supported by Driscoll’s, Inc.

Author contributions
M.A.T. collected samples for sequencing, carried out analyses, interpreted the data, and wrote the manuscript. R.V.R. carried out analyses and wrote the manuscript. J.M.K.R. collected samples for sequencing. M.L.G. carried out the assembly and quality analyses. S.T.S. carried out the computational analyses for Hi-C genome assembly. I.L. carried out the experimental protocols for Hi-C genome assembly. A.K.C. conceived and designed experiments. J.D.E. conceived and designed experiments. A.S.M. conceived and designed experiments, carried out the assembly and analyses, interpreted the data, and wrote the manuscript.

Competing interests
I.L. and S.T.S. declare the following competing interests: I.L. and S.T.S. are employees and shareholders of Phase Genomics, a company commercializing proximity-ligation technology. M.A.T., R.V.R., J.M.K.R., M.L.G., A.K.C., J.D.E., and A.S.M. declare no financial or non-financial competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s42003-019-0606-0.

Correspondence and requests for materials should be addressed to M.A.T. or A.S.M.

Reprints and permission information is available at http://www.nature.com/reprints

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019