The first steps toward a global pandemic: Reconstructing the demographic history of parasite host switches in its native range

Maeva A. Techer[1]*, John M. K. Roberts[2], Reed A Cartwright[3,4], and Alexander S. Mikheyev[1,5]*

1 Okinawa Institute of Science and Technology, 1919-1 Tancha Onna-son, 904-0495 Okinawa, Japan

2 Commonwealth Scientific & Industrial Research Organisation, Canberra, ACT, 2601, Australia

3 The Biodesign Institute, Arizona State University, Tempe, Arizona, USA

4 School of Life Sciences, Arizona State University, Tempe, Arizona, USA

5 Australian National University, Canberra, ACT, 2600, Australia

Corresponding Authors *

Maeva A. Techer maeva.techer@oist.jp, Alexander S. Mikheyev sasha@homologo.us

Maeva A. Techer 0000-0001-5417-5103

John M. K. Roberts 0000-0001-9739-5595

Reed A. Cartwright 0000-0002-0837-9380

Alexander S. Mikheyev 0000-0003-4369-1019

Introductory paragraph

Host switching allows parasites to expand their niches. However, successful switching may require suites of adaptations and may decrease performance on the old host. As a result, reductions in gene flow accompany many host switches, driving speciation. Because host switches tend to be rapid, it is difficult to study them in real time and their demographic parameters remain poorly understood. As a result, fundamental factors that control subsequent parasite evolution, such as the size of the switching population or the extent of immigration from the original host, remain largely unknown. To shed light on the host switching process, we explored how host switches occur in independent host shifts by two ectoparasitic honey bee mites (*Varroa destructor* and *V. jacobsoni*). Both switched to the western honey bee (*Apis mellifera*) after it was brought into contact with their ancestral host (*Apis cerana*), ~70 and ~12 years ago, respectively. *Varroa destructor* subsequently caused worldwide collapses of honey bee populations. Using whole-genome sequencing on 63 mites collected in their native ranges from both the ancestral and novel hosts, we were able to reconstruct the known temporal dynamics of the switch. We further found multiple previously undiscovered mitochondrial lineages on the novel host, along with genetic equivalent of tens of individuals that were involved in the initial host switch. Despite being greatly reduced, some gene flow remains between mites adapted to different hosts. Our findings suggest that while reproductive isolation may facilitate fixation of traits beneficial for exploitation of the new host, ongoing genetic exchange may allow genetic amelioration of inbreeding effects.

Keywords

Host switch, population genomics, demographic history, founder size, sympatry.

Arms races between parasites and their hosts drive evolutionary innovation. Novel parasites can decimate host populations or drive them extinct unless counter-adaptations evolve. Similarly, parasite evolution accompanies the very act of host switching, since it requires adaptations to novel host physiology to persist and to spread. Because parasite adaptations tend to be host-specific, host switches are often associated with host-associated genetic differentiation and eventual speciation. However, only the endpoint of this process is typically observed, as host switches tend to occur rapidly, and the original host is often unknown. As a result, many unanswered questions remain about how parasites acquire new hosts. For instance, if host switches are accompanied by a bottleneck due to reduced gene flow from the ancestral host, how does the parasite have sufficient genetic diversity to adapt? Does gene flow cease completely, or does it continue at a low level, potentially providing additional genetic material for adaptations?

One of the major limiting factors for host switching is the geographic separation between parasites and potential hosts. Globalization has eased these barriers, sometimes giving rise to pandemics (Hatcher, Dick, and Dunn 2012) As a result, host switches are easier to observe and to study in something approaching real-time. One of the most dramatic and economically important switches involved the two ectoparasitic mite species, *Varroa destructor* and *Varroa jacobsoni*, which acquired the western honey bee (*Apis mellifera*) as a new host, ~70 and ~12 years ago, respectively (D. L. Anderson and Trueman 2000; Roberts, Anderson, and Tay 2015). *V. destructor,* in particular, spread worldwide, causing extensive honey bee population collapses, whereas *V. jacobsoni* has so far remained in Oceania (Traynor et al. 2020; Noël, Le Conte, and Mondet 2020). Both mites were originally found on the sister species, *Apis cerana*, and came into contact with *A. mellifera,* which was brought in for purposes of beekeeping (Crane 1968; Chantawannakul et al. 2016). These two host switches occurred in parallel and relatively recently, allowing the reconstruction of how the host switches took place, using genomic tools.

Both switches have been investigated using microsatellite markers and relatively short mitochondrial markers, which revealed that in both species populations on the new host were strongly differentiated and genetically depauperate (Roberts, Anderson, and Tay 2015; Solignac et al. 2005; Warrit, Smith, and Lekprayoon 2006; Navajas et al. 2010; Beaurepaire et al. 2015; Dietemann et al. 2019). While *V. destructor*, in particular, was described as "quasi-clonal" (Solignac et al. 2005), it successfully spread worldwide and has shown a remarkable ability to parasitize genetically diverse *A. mellifera* populations, as well as to evolve resistance to human counter-measures, such as pesticides (González-Cabrera et al. 2016). How does a bottlenecked species achieve such a level of success? Increasing evidence from population genomic analysis of fungal pathogens suggests that the success of many pathogens appears to rely on maintaining some level of adaptive diversity despite the presence of bottlenecks during host switches (Fry 2016; McMullan et al. 2018), but similar work is lacking in animal systems.

To answer this question and to gain broader insight into how host switches happen, we sequenced nuclear and mitochondrial genomes from sympatric populations of the two mites across Asia and Oceania, collected on both novel and introduced hosts. This allowed us far greater power to examine how the host switch took place with much greater precision than was possible previously. We found strikingly parallel dynamics at play in both host switches, which were characterized by a surprisingly large effective population size at the time of the switch and ongoing gene flow with cryptic population genetic processes that may have helped *Varroa* succeed.

Results

Cryptic diversity in mitogenomes suggests multiple foundresses at the origins of the host switch

We examined the genome-wide variation and divergence among 31 *V. destructor* and 32 *V. jacobsoni* females from their original and novel hosts (Table 1). *Varroa* species identity was confirmed by extracting and aligning mitogenomes together with known reference sequences of the mtDNA *COX1* 458-bp standard marker (Table S1 and Figure S1). A sole mismatch between taxonomic identification and mtDNA barcoding was detected in a *V. jacobsoni* specimen (THV003_1), which was collected in Thailand on *A. mellifera,* though its reproductive status was not confirmed.

Using the set of 2,091 SNPs detected across the mitogenome, we reconstructed the phylogenetic relationship among *Varroa* lineages for each species (Figure 1). While exhibiting more genomic variability than previously reported, the mitogenome sequences networks were consistent with the standard classification of *Varroa* geographic lineages. We found that both sister species on 3.6% of mitochondrial consensus sequences. Relative to these levels, divergence within *V. destructor* was very low (0.5% with 85 SNPs) but moderate in *V. jacobsoni* with up to 1.3% (216 SNPs).

We identified 18 distinct mitogenome sequences (thereafter named mitotypes) for *V. destructor* that clustered into seven lineages (Figure 1A and Table S2). Throughout our Asian sampling sites, all *V. destructor* mitotypes on *A. mellifera* host exclusively belonged to the host switched Korean K1 lineage, which was also found in one *A. cerana* sample (Figure 1B). We found three additional derived K1 mitotypes in China that were not previously reported (Table S1). We confirmed that the SNPs differentiating these K1 mitotypes were not artefacts by manually checking raw reads mapping (cut-off 200x). Interestingly, the two publicly available mitochondrial genomes of the K1 haplotype (France: NC_004454.2 and Japan: AP0195523.1) appear slightly

divergent from the dominant K1 mitotype. Though we found the other reported host switch Japanese J1 lineage in the source population from Japan, we did not detect it on *A. mellifera*. In addition, a re-analysis of 22 *COX1* sequences available on NCBI, combined with our new data also indicates diverse mitochondrial backgrounds of host-switched bees, both in *V. destructor* and *V. jacobsoni* (Figure S1).

Considering a similar sample size as for *V. destructor*, we found a much higher genomic variability in the mtDNA in *V. jacobsoni* samples. Our analysis detected 27 unique mitotypes that belonged to seven lineages (Figure 1C and Table S2). Unexpectedly, we found that the misidentified mite on *A. mellifera* in Thailand (THV003_1) clustered with a new lineage VJ North Thailand 4 (Table 1). This lineage was closely related to the North Thailand 1-3 *COX1* (Warrit, Smith, and Lekprayoon 2006; Dietemann et al. 2019) reference sequences (Figure S1) on and Laos 1 mitotypes (Figure 1C). All other *V. jacobsoni* mites collected on *A. mellifera* (*N* = 10), unequivocally belonged to the reported host switched lineage Papua New Guinea 1. The median joining network showed that six mitotypes found in the A. *mellifera* population were derived (up to 8 SNPs difference) or identical to native *V. jacobsoni* mitotypes. Two founders PNG 1 mitotypes detected during the host switch event (2008) were retrieved several generations later in *A. mellifera* colonies (2014) (Table S1).

Host-associated genomic differentiation following host shifts

Our variant call analysis revealed that 2,728,471 biallelic SNPs in whole-genome sequences of *Varroa* mites, following a quality and coverage bias filtering. However, multiple results pointed toward a loss a genome-wide loss of variability associated with the transition from the original host *A. cerana* to the novel one *A. mellifera*. When looking at polymorphic sites only, a 2.8 fold reduction in SNPs numbers between was observed in *V. destructor* ($N_{VdesAcer}$ = 1,219,990 vs $N_{VdesAmel}$ = 429,420 SNPs) and 3.6 fold in *V. jacobsoni* ($N_{VdesAcer}$ = 1,546,366 vs $N_{VdesAmel}$ = 630,534 SNPs). Reduction of genetic diversity with the founding of new populations was also

observed with the average genomic nucleotide diversity π in both *V. destructor* ($\pi_{VdesAcer}$ = 0.0020 vs $\pi_{VdesAmel}$ = 0.0005) and *V. jacobsoni* ($\pi_{VjacAcer}$ = 0.0014 vs $\pi_{VjacAmel}$ = 0.0008). Genome scans using 20kb sliding windows showed that π values in the new host was overall lower than in the original host but that some regions exhibited higher nucleotide diversity (Figure S2). Analysis of the genomic heterozygosity revealed extremely high levels of inbreeding coefficient in mite genomes with the lowest value F = 0.686 in *V. destructor* and F = 0.721 in *V. jacobsoni* (Table S3). However, when comparing *A. mellifera* and *A. cerana* host populations, we found no significant difference in F estimates for *V. destructor* (two-tailed t-test *p-value* = 0.263) and *V. jacobsoni* (two-tailed t-test *p-value* = 0.297). Such levels in both species were previously reported using microsatellites (Roberts, Anderson, and Tay 2015) and more likely resulting from incestuous *Varroa* mating.

PCA further illustrated the loss of genetic diversity in host-switched mites coupled with genome-wide differentiation (Figure 2AC). F_{ST} and D_{xy} pairwise estimates indicated strong differentiation between nuclear genomes from divergent lineages (Table S4). Within each species, high levels of genetic differentiation all along all seven chromosomes were detected between host-adapted lineages (Figure S2). Interestingly, when comparing the host-switched *V. destructor* specimens to native-host samples from the same mitotype (K1) there was far less genetic differentiation (F_{ST} = 0.004) compared with any other native-range population (F_{ST} > 0.6) (Table S4). Thus, host-switched populations still carry many of the genes present in the population of origin, a scenario consistent with rapid reproductive isolation post-switch. Similar patterns were obtained genome-wide for the larger sliding windows of 20kb of F_{ST} and D_{xy} estimates show that *A. cerana* and *A. mellifera* K1 mites diverged massively (Supplemental online markdown).

We found similar patterns of population structure and divergence among hosts and mtDNA lineages in *V. jacobsoni*, while host switch was more recent. In our study, we did not find clonal genomes in either *V. jacobsoni* or *V. destructor* mites collected from *A. mellifera* populations.

Ancestry of host switched populations and evidence of gene flow

NGSadmix analysis conducted on 1,276,602 unlinked biallelic SNPs confirmed the population structure observed between *A. mellifera* and *A. cerana* mites and helped to identify the source of host populations. At K = 2, population structure quickly emerged among *V. destructor* mites. J1 and K1 *A. cerana* mites were assigned unequivocally to the same genetic cluster as all K1 *A. mellifera* mites (Figure 2B). Therefore, genome-wide sequencing confirmed the contribution of K1 and related J1 mite sources in the founding of novel host populations. In *V. jacobsoni*, admixture estimates suggested that two large genetic clusters exist: Northern/Asian *V. jacobsoni* lineages differing from the South/Oceanian (Figure 2D). At K = 3, unreported spillover events were evidenced with the mite (THV003_1) found in *A. mellifera* in Thailand that was genetically related to mites in Viet Nam (Laos 1). Here, population genomics also supports the origin of *V. jacobsoni* host switch in Papua New Guinea to be exclusively from *A. cerana* PNG1 and Java 1. Unlike in *V. destructor*, signal of recent migration was exposed by one individual (PGV956_4) in *A. mellifera* Papua New Guinea population.

Recent migration and spillover events were confirmed by analyzing mite's body metagenomic reads. Our predictions that honey bee hosts mtDNA reads could be retrieved from the whole crushed parasitic mite was validated. The proof of concept for this method was first tested and demonstrated with mite families (used for *de novo* mutation measurements) that were actively feeding at the time of collection (Table S5). For population genomics mites, up to 0.18% of all reads from each mite library (n = 115,581 reads) matched honey bee host mitochondrial reference genomes. Interestingly, the NGSadmix *V. jacobsoni* outlier (PGV956_4) was actually "caught" in migration with more *A. cerana* reads (*N* = 1,622) than of the collected host *A. mellifera* (*N* = 69). While admixture estimates did not spot such outliers in *V. destructor*, the diet molecular analysis revealed reverse cases in Viet Nam (VNV475_1) and in Japan (JPV025_1). *A. mellifera* and *A. cerana* honey bees present enough important morphological and behavior differences that

experts field collectors would have assuredly not missed. Such convergent observations more likely indicated that these mites migrated between hosts shortly before collection.

Low *de novo* mutation rate from pedigree-based measurements

We examined whether the rapid parallel differentiation observed between *Varroa* host populations was driven by high frequency of new mutations by estimating the *de novo* mutation rate, following a previously published approach (Ramu et al. 2013). As the *Varroa* reproductive sequence cycle is well known (one haploid son, followed by up to four diploid daughters), we sequenced six diploid mothers and their respective sons at a coverage between 18x - 33x. Before filtering, the number of detected mutations was 3,082 but there was extreme variation between samples in the number of mutations called. Notably mother-son pairs with less coverage contained several folds more mutation calls than mother-son pairs with higher coverage. Samples with higher mutation calls also showed stronger evidence of clustering of mutation calls, which typically indicates false-positive mutation calls. After filtering, two mutation calls remained, but showed low coverage and were discarded as unreliable.

We conducted a set of simulations to deal with coverage bias issues in the data set. Out of 6,000 simulated mutations, 3,464 could be detected by our pipeline, indicating that our recall rate was approximately 58%. The total haploid length of the 7 chromosomes considered in the mutation scan was 362,833,908 base pairs. Given that we found no mutations among 6 mother-son pairs, we estimate that the mutation rate in *Varroa* is less than 8×10^{-10} per bp per generation ($\frac{1}{7 \times 362833908 \times 0.5773}$) or less than 0.28 mutations per son per generation.

Estimating of host switch demographic history and founding population size

We aimed to further parameterize the demography of host switches detected in mitochondrial data using coalescent-based modeling. Demographic inference using fastsimcoal2 was conducted using an isolation with migration scenario (Figure S3). For *V. destructor*, 32 haploid

230 mite genomes were sequenced from *A. mellifera* host and 22 haploid mite genomes were sequenced from *A. cerana* host. For *V. jacobsoni*, 18 haploid genomes of *A. mellifera* mites and 6 haploid genomes of *A. cerana* mites from Papua New Guineas were projected (excluding the older sample from 1997). Given that we were only able to estimate the upper limit of *de novo* mutation rate, we explored whether lower mutation rate improve the observed maximum
235 likelihood distribution of 100 replicates (Figure S4). We examined rates as low as 1.0×10^{-11}, at the lower limit of rates in arthropods (Fazalova and Nevado 2020). We also tested whether including an inbreeding coefficient would help improving the model likelihood by implement 70% of inbreeding (minimum detected in both species). However, lower mutation rates did not improve model fit for both *V. destructor* (Kruskal-Wallis chi-squared = 497.61, df = 494, *p-value* = 0.446)
240 and *V. jacobsoni* (Kruskal-Wallis chi-squared = 499, df = 498, *p-value* = 0.479) (Figure S4). The inbreeding coefficient did also not significantly improve the likelihood with equal mutation rate (available in online supplemental markdown). In addition, 2D-joint SFS visual inspection did not reveal striking differences, and the parameter estimates overall were qualitatively similar over the range of mutation and inbreeding parameters we examined.

245

Hence, demographic parameters and their bootstraps values were estimated using $\mu = 8 \times 10^{-10}$ per bp per generation and no inbreeding value. Mean values and 95% confidence intervals for each species are graphically summarized in Figure 3**.** The older host switch with *V. destructor* was estimated to have occurred 887 generations before sampling (~88 years ago considering 10
250 generations / year (Cornuet et al. 2006)) with a founder effect lasting for 177 generations. Estimates suggested that 156 [CI95 = 146-165] haploid genomes of *V. destructor* contributed to the founding event in Asia. From that point on, the novel mite population quickly growing on *A. mellifera* diverged from the sympatric *A. cerana* source populations. While the best scenario assumed migration, our results suggest gene flow with low but continuous migrating rates
255 between hosts. Despite being a more recent event, estimated around 183 generations ago (~18 years ago), demographic parameters for *V. jacobsoni* also supported a large founding size

reaching 138 haploid mites. The effective population size of *V. jacobsoni* mites on *A. mellifera* was lower than the source population, contrary to the demographic history estimated for *V. destructor*. Finally, bidirectional migration rates suggested more important ongoing gene flow among host populations.

Discussion

Virtually every living species has at least one species of parasite, making parasitism perhaps the most successful mode of life. Most studies examine parasitism from either a macroevolutionary perspective, for instance, through host-parasite co-phylogenies, or at the population level, by examining patterns of differentiation between parasites and hosts (Poulin et al. 2011; Penczykowski, Laine, and Koskella 2016). Theoretical modelling links the two scales, suggesting that specialization coupled with trade-offs in performance on alternative hosts should lead to speciation (Sexton et al. 2017). Yet, empirical observations of this process remain scarce, particularly of host switch demography, which includes key parameters such as population sizes during switches that are necessary to ensure sufficient evolutionary potential for parasites (McDonald and Linde 2002). For this reason, pandemic spread of *Varroa* mites among honey bees has long been puzzling, given their “quasi-clonal” structure in the invasive portion of their range (Solignac et al. 2005). Using high-resolution mitochondrial and nuclear genomic data from sympatric populations, we show that in both *V. destructor* and *V. jacobsoni,* (a) genetic bottlenecks were far less severe than previously estimated and (b) while gene flow was greatly reduced post-host switch, consistent with models of incipient speciation that may accompany acquisition of novel hosts, gene flow has not completely stopped. Our data highlight the importance of genetic diversity during initial stages of parasitic host switches.

While loss of genetic diversity is common in parasites, it often occurs in the parasite’s introduced geographic range (Auger-Rozenberg et al. 2012; Demastes et al. 2019). Consequently, parasites

may possess sufficient genetic diversity to parasitize their hosts. For example, the fungal parasite responsible for ash dieback in Europe arrived from Asia, and while bottlenecked to only two haplotypes, it nonetheless maintains adaptive diversity in key host interaction genes (Fry 2016). To the best of our knowledge, our study is unique in quantifying demographics within the native range of the parasite, where it has opportunities for additional gene flow from genetically diverse sympatric populations. At least for *Varroa,* the initial host switch requires an unexpected amount of genetic diversity at the point of the switch, but rapidly leads to reproductive isolation between sympatric populations on novel and original hosts. The amount of differentiation appears to increase over time, being greatest in *V. destructor*, which switched ~50 years earlier than *V. jacobsoni*. *Varroa* mites coevolve quickly with their new hosts (Beaurepaire et al. 2019) and such sympatric isolation could ultimately lead to speciation.

Post host-switch speciation can occur in the presence of gene flow, if specialization is adaptively favored and selection acts on many genomic regions (Pinho and Hey 2010; Feder, Egan, and Nosil 2012). By promoting genomic heterogeneity and introducing beneficial mutations (Tomasini and Peischl 2020), gene flow can ameliorate lost diversity due to selection or inbreeding (Lenormand 2002; Whiteley et al. 2015), both of which would impact host-switched *Varroa* populations. Alternatively, *de novo* mutation rates could contribute to population genetic variation, but our direct estimation of mutation rate did not show exceptionally high mutation rates ($< 8 \times 10^{-10}$ per bp per generation), compared to other arthropods (H. Liu et al. 2017; Fazalova and Nevado 2020). Consequently, immigration may be the primary means of introducing new alleles into populations on the novel host during early stages of the host switch.

While host switch events by specialized parasites were once assumed to be rare, this assumption is increasingly challenged (Rigaud, Perrot-Minnot, and Brown 2010; Nylin et al. 2018). Ecological fitting theory suggests that shifts could readily occur in species with a pre-existing ability to use novel hosts (Araujo et al. 2015). Host switches in *Varroa* appear to fit this model. While the

310 western honey bee has been introduced throughout the *Varroa* native range, mitochondrial data, which are geographically informative (Figure 1), suggest that switches occurred only in Korea, Japan, Philippines (*V. destructor*), Papua New Guinea and possibly Thailand (*V. jacobsoni*) (Roberts, Anderson, and Tay 2015; Navajas et al. 2010; Beaurepaire et al. 2015; Dietemann et al. 2019). *A. cerana* subspecies are strongly differentiated geographically (Radloff et al. 2010; K.
315 Tan et al. 2016; Chen et al. 2018; Yancan et al. 2019; Ilyasov et al. 2019; Ji et al. 2020), and *Varroa* mitotypes mirror host biogeography and subspecies distribution (Warrit, Smith, and Lekprayoon 2006; Beaurepaire et al. 2015; Rueppell et al. 2011; Dietemann et al. 2019). This suggests that *Varroa* populations may vary in traits, such as host specificity, as a result of their previous coevolutionary interaction with local *A. cerana* subspecies. This may have allowed some
320 populations to switch, while others were unable, in spite of available novel hosts (W. Li et al. 2019). However, the dynamic nature of the interaction does not preclude additional switches in the future.

One potential limitation of our data set is its sparse geographic and temporal sampling, given the
325 size of the native range. While small sample size is enough to retrace ancestral event, large sample sizes can increase the confidence in model selection and parameter estimates for recent demographic events (Robinson et al. 2014). Yet, past scenario and demographic estimates were consistent with mite biology and reflected direct observations made during their invasion of *A. mellifera*. Furthermore, while our data set captured most of the described *Varroa* mitochondrial
330 lineages and even discovered new ones (D. L. Anderson and Trueman 2000; Solignac et al. 2005; Dietemann et al. 2019; Zhou et al. 2004), some reported mitotypes were missing. In the future, incorporating larger sample size from modern source populations could help better estimate each lineage genetic contribution to host switch success. In addition it would allow to estimate how host switch may have affected native diversity as for the case of J1 lineage
335 displacement (de Guzman, Rinderer, and Anthony Stelzer 1999; Ogihara et al. 2020). Despite these limitations, our analysis correctly reconstructed known aspects of *Varroa* host switch

demography, such as the times of the switches. However, these times need to be regarded as approximate, since generation times may fluctuate due to a variety of factors, such as seasonality or brood availability. We relied on a previous approximation of 10 generations per year (Cornuet et al. 2006). As a result, we caution against over-relying on the apparent precision of the numerical estimates (Table 2). Nonetheless, data from both mitogenome sequencing and coalescent analysis all point to a relatively large founding population size of the *Varroa* mites (Figures 1 and 3).

Our findings highlight the dynamic and ongoing nature of host switches in the native range, and the need to better understand native mite populations. Mitochondrial DNA has already detected the likely presence of several host switches, such as those in Philippines in 2015 (Beaurepaire et al. 2015) and the presence of new haplogroups in Eastern Europe (Gajic et al. 2013), indicating that populations of *Varroa* are continuously testing *A. mellifera* as a new host, and may have spread without our awareness. Additionally, while this work provides insight into the initial host switch, the demographics of the subsequent worldwide pandemic remain largely unknown, for instance the critical population size necessary to establish a regional infestation. Future research on global demographic parameters such as genetic diversity and gene flow is crucial (Evans and Cook 2018). Since *V. jacobsoni* presents striking parallels to *V. destructor*, this knowledge could be applied to forecast and prevent its spread.

Materials and Methods

Mite sampling on original and novel honey bee hosts

We sequenced adult females throughout the *Varroa* native ranges where original and novel hosts occur in sympatry (*V. destructor N* = 31 and *V. jacobsoni N* = 32; Table 1). For *V. destructor,* specimens were collected between 1996 and 2003 as part of the taxonomic revision and diversity survey carried by Anderson and Trueman (D. L. Anderson and Trueman 2000) and Navajas *et* al. (Navajas et al. 2010) and included 14 samples from *A. cerana* and 17 from *A. mellifera* collected

from Japan, South Korea, China, Thailand, Viet Nam, Myanmar and Nepal. For *V. jacobsoni*, female specimens were obtained from Anderson and Trueman (D. L. Anderson and Trueman 2000) (1994-2004) and Roberts *et al*. (Roberts, Anderson, and Tay 2015) (2008-2015) surveys of host shifts reports in Papua New Guinea. A total of 19 adult female *V. jacobsoni* mites were collected from *A. cerana* across different areas in India, Thailand, Viet Nam, Malaysia, Indonesia and Papua New Guinea. In contrast to *V. destructor*, *V. jacobsoni* infestation on *A. mellifera* is still restricted to Papua New Guinea from which 10 samples were collected. Additionally, we included three *V. jacobsoni* mites collected from an alternative original host *A. nigrocincta* found in Indonesia. *A. nigrocincta* is not believed to be the source of the host-switched mites, so these samples were included for comparison purposes.

Sampling complete details regarding host, location and year are available in Table S1**.** Exact geographical coordinates were not always available for samples collected before 2008, and were approximated by the locality provided, or published survey maps (D. L. Anderson and Trueman 2000; Navajas et al. 2010; Zhou et al. 2004; Denis L. Anderson and Fuchs 1998). All individuals were mature sclerotized females collected from single colonies and were preserved in individual Eppendorf tubes. Each collection tube was stored with 70% ethanol and kept at -20°C at the CSIRO in Canberra, Australia.

DNA extraction and whole genome resequencing

Mites were surface sterilized by cleaning them in absolute ethanol using a sterile brush to remove any external debris, and then gently shaken in a 2.0 mL Eppendorf filled with absolute ethanol. Each mite was then dried for 10 sec on a sterile paper towel before being placed in a 1.5 mL Eppendorf tube in liquid nitrogen. Genomic DNA was extracted from each mite by crushing the whole body, using a sterile pestle to obtain a fine powder and processed with a QIAamp DNA Micro Kit (© Qiagen) following the manufacturer's instructions. Final elution volume was 15 µL.

390 Total dsDNA was measured using a Qubit™ 4 Fluorometer with an Invitrogen dsDNA HS Assay Kit.

For population genomic samples, short-inserts of 150-bp paired-end libraries were prepared for each individual using a Nextera XT DNA Library Preparation Kit (Illumina ®). Size-selection and
395 cleanup were accomplished using CA-magnetic beads (Dynabeads® MyOne Carboxylic Acid, Invitrogen), and 11-11.5% PEG 6000 (Sigma-Aldrich © LLC). Library quality and size were assessed using a Bioanalyzer High Sensitivity DNA kit (Agilent). Libraries were run on HiSeq 4000 and NovaSeq6000 in 150 bp x 2 paired-end mode (Illumina ®) at the OIST Sequencing Center. Biosamples and DRA accession as well as sequencing coverage for each sample are
400 provided in Table S1.

Host DNA contents in mite whole-body metagenomics

We cross-checked honey bee host identities reported during sampling with host read identity to detect potentially migrating mites. Mites feed on honey bees during the phoretic phase (Ramsey
405 et al. 2019) and maintain a consistent feeding regimen by consuming ~1 µL of host fluid (digested fat-body and haemolymph) per day (Posada-Florez et al. 2019). Therefore, we assumed that host DNA would be retrieved from crushed whole-mite tissues. Mitochondrial DNA was targeted, as it is more abundant than nuclear DNA. We mapped raw fastq reads on honey bee host mitochondrial reference genomes using [NC_001566.1] for *A. mellifera ligustica* (Crozier and
410 Crozier 1993) and [NC_014295] for *A. cerana* (H.-W. Tan et al. 2011). The number of reads mapped to either one of these honey bee host reference genomes was counted and compared to sampled host identities.

Data filtering, mapping and genotype calling

Commands used for each analysis step are available on our Snakemake script (Köster and
415 Rahmann 2012) available on https://github.com/MaevaTecher/varroa-host-jump. Briefly, we assessed demultiplexed fastq read quality using FastQC (Andrews and Others 2010). We then

mapped reads to the *V. destructor* reference genome on NCBI [GCF_002443255.1] (Techer et al. 2019) separately from the complete mitogenome [NC_004454.2] (Navajas et al. 2002) using the soft-clipping and very sensitive mode of NextGenMap v0.5.0 (Sedlazeck, Rescheneder, and von
420 Haeseler 2013) (following a comparison with Bowtie2 v2.6 (Langmead and Salzberg 2012)). Reads were sorted and duplicates were removed using SAMtools (H. Li et al. 2009), and subsampled to a maximum coverage of 200 using VariantBam (Wala et al. 2016) to speed up processing. Mapping rates and reads depths were computed from the generated BAM files.

425 We generated three data sets for the mapped reads depending on analysis requirements, following Yamasaki *et al.* (Yamasaki et al. 2020). First, we obtained a "*SNP-only dataset*" containing only variant sites using FreeBayes v1.1.0 (Garrison and Marth 2012) with the following parameters: minimum mapping = 10, minimum base quality = 5, use of the four best SNP alleles, and no populations prior. In order to correct for coverage bias and sequencing artefacts in
430 problematic regions, we estimated twice the mean reads depth along the seven main chromosomes. Subsequently, variants were filtered using using VCFtools v0.1.12b (Danecek et al. 2011). The "*SNP-only dataset*" resulted in 2,728,471 SNPs. Second, we computed an "*all-sites dataset*" by calling both variants and invariants sites using BCFtools v1.9 mpileup (H. Li 2011). We removed indels and sites at over twice the computed mean depth, as well as sites with any
435 missing data. We excluded sites that were not placed on the seven chromosomes. After filtering, the "*all-sites dataset*" resulted in 2,130,335 SNPs and 120,279,163 monomorphic sites. Third, we obtained a "*mtDNA SNP-only dataset*" by calling polymorphic sites only on the mitochondrial genome NC_004454.2. We used FreeBayes v1.1.0 with strict quality parameters. The dataset resulted in 2,091 SNPs. These were validated in the *COX1* region by Sanger sequencing (see
440 next section).

Comparative mitogenomes analysis and lineages identification

We used the "*mtDNA SNP-only dataset*" for mitochondrial variation genomic analysis by converting and generating individual nucleotide sequences using the option *vcf2fasta* in vcflib
445 scripts (github.com/vcflib). Nucleotide sequences were screened and aligned (16,476 bp) using Geneious Prime® 2019.2.3 together with the reference mitochondrial sequences NC_004454.2 (*V. destructor* on *A. mellifera*, France (Navajas et al. 2002)) and also, to avoid problems with reference bias, from AP019523.1 (*V. destructor* on *A. mellifera*, Japan (Harada et al. 2020)). Mitotype diversity for each *Varroa* species and distribution in hosts was assessed using the
450 median joining network method from SplitsTree5 (Huson and Bryant 2006). Divergence levels between *Varroa* species was estimated by comparing the whole-mitochondrial consensus sequences with Geneious Prime®.

Standard methods exists to determine *Varroa* species and lineage (also named haplogroup)
455 identity by using the following mitochondrial markers: 1) a 458-bp fragment from the *COX1* gene (D. L. Anderson and Trueman 2000; Dietemann et al. 2013) and 2) a concatenated 2,696-bp fragment including *COX1*, *ATP6*, *COX3*, and *CYTB* (Navajas et al. 2010). Thus, we determined the lineage (e.g. Korean K1, Japan J1, *etc.*) of each mitotype by extracting the *COX1* barcoding region of interest and aligning it together with 60 unique *COX1* reference sequences from NCBI
460 using ClustalW and manual check. A neighbor-joining tree was computed on the fasta alignment using IQ-tree (Nguyen et al. 2015) and exported with iTOL (Letunic and Bork 2019). Following *Varroa* revised nomenclature (Traynor et al. 2020), a sequence was considered from a known lineage only if 100% identical to reference *COX1* haplotype. The same process was used for sub-lineages (e.g. Korean K1-1/K1-2, Korean K1-3,...) identification by with the *COX1-ATP6-COX3-*
465 *CYTB* concatenated region sequences which were aligned with 22 available reference concatenated sequences for *V. destructor* only (Navajas et al. 2010; Gajic et al. 2013; Elbeaino et al. 2016).

To ensure that filtering of the "*mtDNA SNP-only dataset*" did not remove existing and known
470 variants, we also sequenced the *COX1* gene using Sanger sequencing using primers 10kbCOIF1 and 6,5KbCOIR (Navajas et al. 2010). PCR reactions were carried out in 25µL containing 5 µL of 5X Phusion® HF buffer; 0.5 µL of dNTP mix (10mM); 0.25 µL of Phusion® High-Fidelity DNA Polymerase (NEB); 1.25 µL of each oligo primer 10kbCOIF1and 6,5KbCOIR (10mM); 1 µL of template DNA (0.5 ng/µL) and Milli-Q water. Samples were denatured at 98°C for 30 sec, and
475 then PCR was performed for 35 cycles of 10 sec denaturation at 98°C, 15 sec of annealing at 59°C and 15 sec of extension at 72°C with a 5 min final elongation at 72°C. DNA amplification success was visualized by loading 3 µL of PCR product with 3 µL of loading dye on 1% agarose gel (110V for 20 min). PCR products were then cleaned-up using Dynabeads® MyOne Carboxylic Acid, CA-beads (Invitrogen) and 19% Polyethylene glycol PEG. Directly, fragments were
480 sequenced in the two directions using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher) in a capillary sequencer Applied Biosystems 3730xl DNA Analyzer (Thermo Fisher). FASTA sequences generated by mitogenome mapping and Sanger sequencing were then aligned and checked for differences.

485 **Genome-wide variation and analysis of divergence**

We examined levels of genetic diversity and differentiation between host-adapted population of each *Varroa* species. To investigate differentiation among populations, we first measured the absolute divergence D_{xy} which requires both variants and invariants sites (Cruickshank and Hahn 2014). We also investigated levels of standing genetic variation (π). For both analyses genome
490 scans were computed using a sliding window method on the "*all-sites dataset*" (50kb window, sliding every 20kb and containing a minimum of 100 SNPs). This method was implemented using the *parseVCF.py* and *popgenWindow.py* python scripts (github.com/simonhmartin/genomics_general) (Martin, Davey, and Jiggins 2015).

495 Using the biallelic "*SNP-only dataset*", we also estimated the genome-wide differentiation by calculating Weir and Cockerham's F_{ST} per site using VCFtools. Subsequently, we assessed the genetic structure within species using a principal component analysis (PCA) performed with the R package vcfR (Knaus and Grünwald 2017). To further investigate the population structure and related ancestry, we conducted an incremental K step analysis using NGSadmix (Skotte,
500 Korneliussen, and Albrechtsen 2013). We reduce the effect of linkage disequilibrium between SNPs by conducting a pruning using PLINK v1.90b3 (Purcell et al. 2007) on the "S*NP-only dataset*" with the following parameters: --indep-pairwise 20 10 0.5. After removing linked SNPs, the pruned dataset containing 1,276,602 SNPs was converted into BEAGLE format and directly used in NGSadmix. A total of 10 replicates per K, from 2 to 20, were run and the model with the
505 highest likelihood was selected for plotting each K. The number of genetic clusters among and within species was determined following guidelines in (Meirmans 2015; Lawson, van Dorp, and Falush 2018) and making biological sense.

Mutation rate estimation from mother-son pairs

510 Coalescent analysis requires an estimation of mutation rate, yet none is available for closely related species. To do this we collected and sequenced six *V. destructor* diploid mothers and their haploid sons for this purpose. These samples were collected from individual capped cells containing *A. mellifera* red-eye drone pupae. Each family was selected under the following conditions: 1) each cell was only infected by a single mother-foundress mite, 2) a *Varroa* family at
515 this stage should contained three to four offspring, including one haploid son and two to three diploid daughters, 3) the son was at the deutonymph or adult phase to avoid misidentification with protonymph sisters. We collected three families and preserved in absolute ethanol in February 2018 (#1, #2, and #7), and three more in May 2018 (#15, #17, and #19) from the same colony at the OIST Ecology and Evolution experimental apiary.

520

Libraries were prepared using an NEBNext® Ultra™ II FS DNA Library Prep Kit (New England Biolabs, Inc) with a fragmentation size incubation time of 15 min at 37°C, and three PCR cycles. Libraries were cleaned up using CA-magnetic beads (Dynabeads® MyOne Carboxylic Acid, Invitrogen), and 17% PEG 6000 (Sigma-Aldrich © LLC). Libraries were pooled and sequenced on two lanes of HiSeq 4000 (Illumina ®) at the OIST sequencing center.

Mutation rate was estimated from GATK variant calls (McKenna et al. 2010) using DeNovoGear's dng-call algorithm (Ramu et al. 2013) which can model *Varroa*'s haplodiploidy sex-determination system. Single-nucleotide mutations were called on the 7 largest contigs in the *Varroa* genome, avoiding sites that were within 100 bp of an indel. Mutation calls were de-duplicated, retaining only sites that were biallelic and not part of a mutation cluster (with 100 bp of another call or part of a long-run of calls from the same sample). After deduplication, the remaining calls were filtered such that P (denovo| data) was high (DNP >= 0.75) and the fit to the model was good (LLS >= -3). DNP (de-novo probability) and LLS (log-likelihood scaled) are both per-site statistics generated by dng-call. The denominator for the mutation rate analysis was estimated by simulating mutations at 1,000 locations in each son and calculating what fraction of these simulated mutations was recovered by our pipeline. A VCF was generated for these locations, and mutations were simulated by changing the son's haplotype to one of the other three bases. ALT and AD fields were updated as needed.

Demographic analyses of host switch using SFS

We inferred the demographic history of both *V. destructor* and *V. jacobsoni* mites with the coalescent simulator fastsimcoal v2.6 using the site frequency spectrum (SFS) (Excoffier et al. 2013). Demographic inferences were computed independently for each species and samples were subsetted following genome-wide analysis results while considering geographical sympatry and continuity. For *V. destructor* (*N* = 27), we excluded Nepal and the Japanese mites (continental island) whereas for *V. jacobsoni* (*N* = 12), we included only mites from Papua New

Guinea (host switch region). To reduce the effect of selection which can biais the SFS (71), we kept SNPs and invariants from the "*all-sites dataset*" that were at least 50 kb away from any
550 annotated genes regions (VCFtools exclude-bed option). The 2D-joint folded-SFS was computed for each species/host population on the filtered 12,594,802 sites including 224,568 SNPs, using the *vcf2sfs* R scripts (github.com/shenglin-liu/vcf2sfs) (S. Liu et al. 2018).

We considered a scenario consistent with the known history of *Varroa* jumps (Figure S3). We
555 incorporated the following demographic parameters (Table 2): the estimated effective population size for *Varroa* mites parasitizing *A. cerana* N_{VAC1} (in haploid genomes) to be stable before, during, and after the host switch. On the other hand, the mite population size on the new host *A. mellifera* expanded from N_{BOTAM} or not with a growth rate G_{AM} after the host switch to reach modern population size N_{VAM0} (in haploid genomes). The host switch founder event occurred at a
560 time T_{JUMP} (in generations) and ended at a time T_{BOTEND} (in generations). Finally, in a case of bidirectional migrations due to sympatry we estimated M_{AMtoAC}/M_{ACtoAM}, to be the proportion of haploid genomes migrating from one population to another.

We ran 100 replicates using the observed SFS as follow: a minimum of 20 loops (--minnumloops
565 20) and a maximum of 150 loops (e.g., ECM cycles, --numloops 100) were performed to estimate the parameters, with one million coalescent simulations per loop (--numsims 1,000,000), and a 0.001 minimum relative difference in parameter values estimated by the maximum composite likelihood (--maxlhood 0.001). The replicate with the highest likelihood set of estimated parameters was retained for model comparison.

570

To ensure that the upper limit genome-wide mutation rate $\mu = 8.0 \times 10^{-10}$ previously used was adequate to estimate parameters from the observed SFS, we tested different μ levels. For this we ran scenario 4 in the same conditions as for scenario choice but μ ranging from 8.0×10^{-10} to 1.0×10^{-11}. Additionally, fastsimcoal2.6 offers the possibility to input inbreeding coefficient. As *Varroa*

575 populations showed high inbreeding, we also ran scenario 4 with (F = 0.7, preliminary tests also ran with 0.8 and 0.9) and without inbreeding coefficient. For all these conditions combined, we ran 100 replicates for each of the 10 combinations per species.

Finally, the overall best estimate parameters set was used to generate 100 pseudo-observed
580 SFS (for a similar number of polymorphic SNPs) for parametric bootstraps. We repeated the parameters estimation for each pseudo-observed SFS and kept the best run. The top 100 runs estimated parameters values were used to calculate 95% confidence intervals. The mean and 95 percentile confidence intervals were computed using the 'boot' R package (Canty and Ripley 2019). Finally, the fit of the best expected SFS was visualized against the observed using
585 SFStools R scripts (github.com/marqueda/SFS-scripts).

References

Anderson, D. L., and J. W. Trueman. 2000. "Varroa Jacobsoni (Acari: Varroidae) Is More than One Species." *Experimental & Applied Acarology* 24 (3): 165–89.

590 Anderson, Denis L., and Stefan Fuchs. 1998. "Two Genetically Distinct Populations of Varroa Jacobsoni with Contrasting Reproductive Abilities on Apis Mellifera." *Journal of Apicultural Research* 37 (2): 69–78.

Andrews, Simon, and Others. 2010. "FastQC: A Quality Control Tool for High Throughput Sequence Data."

595 Araujo, Sabrina B. L., Mariana Pires Braga, Daniel R. Brooks, Salvatore J. Agosta, Eric P. Hoberg, Francisco W. von Hartenthal, and Walter A. Boeger. 2015. "Understanding Host-Switching by Ecological Fitting." *PloS One* 10 (10): e0139225.

Auger-Rozenberg, M-A, T. Boivin, E. Magnoux, C. Courtin, A. Roques, and C. Kerdelhué. 2012. "Inferences on Population History of a Seed Chalcid Wasp: Invasion Success despite a
600 Severe Founder Effect from an Unexpected Source Population." *Molecular Ecology* 21 (24): 6086–6103.

Beaurepaire, Alexis L., Arrigo Moro, Fanny Mondet, Yves Le Conte, Peter Neumann, and Barbara Locke. 2019. "Population Genetics of Ectoparasitic Mites Suggest Arms Race with Honeybee Hosts." *Scientific Reports* 9 (1): 11355.

605 Beaurepaire, Alexis L., Tuan A. Truong, Alejandro C. Fajardo, Tam Q. Dinh, Cleofas Cervancia, and Robin F. A. Moritz. 2015. "Host Specificity in the Honeybee Parasitic Mite, Varroa Spp. in Apis Mellifera and Apis Cerana." *PloS One* 10 (8): e0135103.

Canty, Angelo, and Brian Ripley. 2019. “Package ‘Boot.’” Version. ftp://ftp.math.ethz.ch/sfs/R-CRAN/web/packages/boot/boot.pdf.

Chantawannakul, Panuwan, Lilia I. de Guzman, Jilian Li, and Geoffrey R. Williams. 2016. “Parasites, Pathogens, and Pests of Honeybees in Asia.” *Apidologie* 47 (3): 301–24.

Chen, Chao, Huihua Wang, Zhiguang Liu, Xiao Chen, Jiao Tang, Fanming Meng, and Wei Shi. 2018. “Population Genomics Provide Insights into the Evolution and Adaptation of the Eastern Honey Bee (Apis Cerana).” *Molecular Biology and Evolution* 35 (9): 2260–71.

Cornuet, J. M., M. A. Beaumont, A. Estoup, and M. Solignac. 2006. “Inference on Microsatellite Mutation Processes in the Invasive Mite, Varroa Destructor, Using Reversible Jump Markov Chain Monte Carlo.” *Theoretical Population Biology* 69 (2): 129–44.

Crane, Eva. 1968. “Beekeeping Round the World.” *Bee World* 49 (3): 113–14.

Crozier, R. H., and Y. C. Crozier. 1993. “The Mitochondrial Genome of the Honeybee Apis Mellifera: Complete Sequence and Genome Organization.” *Genetics* 133 (1): 97–117.

Cruickshank, Tami E., and Matthew W. Hahn. 2014. “Reanalysis Suggests That Genomic Islands of Speciation Are Due to Reduced Diversity, Not Reduced Gene Flow.” *Molecular Ecology* 23 (13): 3133–57.

Danecek, Petr, Adam Auton, Goncalo Abecasis, Cornelis A. Albers, Eric Banks, Mark A. DePristo, Robert E. Handsaker, et al. 2011. “The Variant Call Format and VCFtools.” *Bioinformatics* 27 (15): 2156–58.

Demastes, James W., David J. Hafner, Mark S. Hafner, Jessica E. Light, and Theresa A. Spradling. 2019. “Loss of Genetic Diversity, Recovery and Allele Surfing in a Colonizing Parasite, Geomydoecus Aurei.” *Molecular Ecology* 28 (4): 703–20.

Dietemann, Vincent, Alexis Beaurepaire, Paul Page, Orlando Yañez, Ninat Buawangpong, Panuwan Chantawannakul, and Peter Neumann. 2019. “Population Genetics of Ectoparasitic Mites Varroa Spp. in Eastern and Western Honey Bees.” *Parasitology* 146 (11): 1429–39.

Dietemann, Vincent, Francesco Nazzi, Stephen J. Martin, Denis L. Anderson, Barbara Locke, Keith S. Delaplane, Quentin Wauquiez, et al. 2013. “Standard Methods for Varroa Research.” *Journal of Apicultural Research* 52 (1): 1–54.

Elbeaino, Toufic, Nouraldin Daher-Hjaij, Faiz Ismaeil, Jamal Mando, Bassem Solaiman Khaled, and Raied Abou Kubaa. 2016. “Occurrence of Deformed Wing Virus, Chronic Bee Paralysis Virus and MtDNA Variants in Haplotype K of Varroa Destructor Mites in Syrian Apiaries.” *Experimental & Applied Acarology* 69 (1): 11–19.

Evans, Jay D., and Steven C. Cook. 2018. “Genetics and Physiology of Varroa Mites.” *Current Opinion in Insect Science* 26 (April): 130–35.

Excoffier, Laurent, Isabelle Dupanloup, Emilia Huerta-Sánchez, Vitor C. Sousa, and Matthieu Foll. 2013. “Robust Demographic Inference from Genomic and SNP Data.” *PLoS Genetics* 9 (10): e1003905.

Fazalova, Varvara, and Bruno Nevado. 2020. “Low Spontaneous Mutation Rate and Pleistocene Radiation of Pea Aphids.” *Molecular Biology and Evolution* 37 (7): 2045–51.

Feder, Jeffrey L., Scott P. Egan, and Patrik Nosil. 2012. “The Genomics of Speciation-with-Gene-Flow.” *Trends in Genetics: TIG* 28 (7): 342–50.

Fry, William E. 2016. “Phytophthora Infestans: New Tools (and Old Ones) Lead to New Understanding and Precision Management.” *Annual Review of Phytopathology* 54 (1): 529–47.

Gajic, Bojan, Zeljko Radulovic, Jevrosima Stevanovic, Zoran Kulisic, Milos Vucicevic, Predrag Simeunovic, and Zoran Stanimirovic. 2013. “Variability of the Honey Bee Mite Varroa Destructor in Serbia, Based on MtDNA Analysis.” *Experimental & Applied Acarology* 61 (1): 97–105.

Garrison, Erik, and Gabor Marth. 2012. “Haplotype-Based Variant Detection from Short-Read Sequencing.” *ArXiv [q-Bio.GN]*. arXiv. http://arxiv.org/abs/1207.3907.

González-Cabrera, Joel, Sonia Rodríguez-Vargas, T. G. Emyr Davies, Linda M. Field, Daniel Schmehl, James D. Ellis, Klemens Krieger, and Martin S. Williamson. 2016. “Novel Mutations in the Voltage-Gated Sodium Channel of Pyrethroid-Resistant Varroa Destructor Populations from the Southeastern USA.” *PloS One* 11 (5): e0155332.

Guzman, Lilia I. de, Thomas E. Rinderer, and J. Anthony Stelzer. 1999. “Occurrence of Two Genotypes of Varroa Jacobsoni Oud. in North America.” *Apidologie* 30 (1): 31–36.

Harada, Reona, Masato Yoshioka, Hisashi Okuyama, Manabu Kato, Stephen J. Martin, and Jun-Ichi Takahashi. 2020. “Complete Mitochondrial DNA Sequence of the Parasitic Honey Bee Mite Varroa Destructor (Mesostigmata: Varroidae).” *Mitochondrial DNA. Part B, Resources* 5 (1): 635–36.

Hatcher, Melanie J., Jaimie T. A. Dick, and Alison M. Dunn. 2012. “Disease Emergence and Invasions.” *Functional Ecology* 26 (6): 1275–87.

Huson, Daniel H., and David Bryant. 2006. “Application of Phylogenetic Networks in Evolutionary Studies.” *Molecular Biology and Evolution* 23 (2): 254–67.

Ilyasov, Rustem A., Han Gi Youn, Myeong-Lyeol Lee, Kil Won Kim, Maxim Y. Proshchalykin, Arkady S. Lelej, Jun-Ichi Takahashi, and Hyung Wook Kwon. 2019. “Phylogenetic Relationships of Russian Far-East Apis Cerana with Other North Asian Populations.” *Journal of Apicultural Science* 63 (2): 289–314.

Ji, Yongkun, Xingan Li, Ting Ji, Junbo Tang, Lifei Qiu, Jiahui Hu, Jiangxing Dong, et al. 2020. “Gene Reuse Facilitates Rapid Radiation and Independent Adaptation to Diverse Habitats in the Asian Honeybee.” *Science Advances* 6 (51): eabd3590.

Knaus, Brian J., and Niklaus J. Grünwald. 2017. “Vcfr: A Package to Manipulate and Visualize Variant Call Format Data in R.” *Molecular Ecology Resources* 17 (1): 44–53.

Köster, Johannes, and Sven Rahmann. 2012. “Snakemake--a Scalable Bioinformatics Workflow Engine.” *Bioinformatics* 28 (19): 2520–22.

Langmead, Ben, and Steven L. Salzberg. 2012. “Fast Gapped-Read Alignment with Bowtie 2.” *Nature Methods* 9 (4): 357–59.

Lawson, Daniel J., Lucy van Dorp, and Daniel Falush. 2018. “A Tutorial on How Not to Over-Interpret STRUCTURE and ADMIXTURE Bar Plots.” *Nature Communications* 9 (1): 3258.

Lenormand, Thomas. 2002. “Gene Flow and the Limits to Natural Selection.” *Trends in Ecology & Evolution* 17 (4): 183–89.

Letunic, Ivica, and Peer Bork. 2019. “Interactive Tree Of Life (ITOL) v4: Recent Updates and New Developments.” *Nucleic Acids Research* 47 (W1): W256–59.

Li, Heng. 2011. “A Statistical Framework for SNP Calling, Mutation Discovery, Association Mapping and Population Genetical Parameter Estimation from Sequencing Data.” *Bioinformatics* 27 (21): 2987–93.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. “The Sequence Alignment/Map Format and SAMtools.” *Bioinformatics* 25 (16): 2078–79.

Li, Wenfeng, Cheng Wang, Zachary Y. Huang, Yanping Chen, and Richou Han. 2019. “Reproduction of Distinct Varroa Destructor Genotypes on Honey Bee Worker Brood.” *Insects* 10 (11): 372.

Liu, Haoxuan, Yanxiao Jia, Xiaoguang Sun, Dacheng Tian, Laurence D. Hurst, and Sihai Yang. 2017. “Direct Determination of the Mutation Rate in the Bumblebee Reveals Evidence for Weak Recombination-Associated Mutation and an Approximate Rate Constancy in Insects.” *Molecular Biology and Evolution* 34 (1): 119–30.

Liu, Shenglin, Anne-Laure Ferchaud, Peter Grønkjaer, Rasmus Nygaard, and Michael M. Hansen. 2018. “Genomic Parallelism and Lack Thereof in Contrasting Systems of Three-Spined Sticklebacks.” *Molecular Ecology* 27 (23): 4725–43.

Martin, Simon H., John W. Davey, and Chris D. Jiggins. 2015. “Evaluating the Use of ABBA-BABA Statistics to Locate Introgressed Loci.” *Molecular Biology and Evolution* 32 (1): 244–57.

McDonald, Bruce A., and Celeste Linde. 2002. “Pathogen Population Genetics, Evolutionary Potential, and Durable Resistance.” *Annual Review of Phytopathology* 40 (February): 349–79.

McKenna, Aaron, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella, et al. 2010. “The Genome Analysis Toolkit: A MapReduce Framework for Analyzing next-Generation DNA Sequencing Data.” *Genome Research* 20 (9): 1297–1303.

McMullan, Mark, Maryam Rafiqi, Gemy Kaithakottil, Bernardo J. Clavijo, Lorelei Bilham, Elizabeth Orton, Lawrence Percival-Alwyn, et al. 2018. “The Ash Dieback Invasion of Europe Was Founded by Two Genetically Divergent Individuals.” *Nature Ecology & Evolution* 2 (6): 1000–1008.

Meirmans, Patrick G. 2015. “Seven Common Mistakes in Population Genetics and How to Avoid Them.” *Molecular Ecology* 24 (13): 3223–31.

Navajas, Maria, Denis L. Anderson, Lilia I. de Guzman, Zachary Y. Huang, Jeremy Clement, Ting Zhou, and Yves Le Conte. 2010. “New Asian Types of Varroa Destructor: A Potential New Threat for World Apiculture.” *Apidologie* 41 (2): 181–93.

Navajas, Maria, Y. Le Conte, M. Solignac, S. Cros-Arteil, and J-M Cornuet. 2002. “The Complete Sequence of the Mitochondrial Genome of the Honeybee Ectoparasite Mite Varroa Destructor (Acari: Mesostigmata).” *Molecular Biology and Evolution* 19 (12): 2313–17.

Nguyen, Lam-Tung, Heiko A. Schmidt, Arndt von Haeseler, and Bui Quang Minh. 2015. “IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies.” *Molecular Biology and Evolution* 32 (1): 268–74.

Noël, Amélie, Yves Le Conte, and Fanny Mondet. 2020. “Varroa Destructor: How Does It Harm Apis Mellifera Honey Bees and What Can Be Done about It?” *Emerging Topics in Life Sciences* 4 (1): 45–57.

Nylin, Sören, Salvatore Agosta, Staffan Bensch, Walter A. Boeger, Mariana P. Braga, Daniel R. Brooks, Matthew L. Forister, et al. 2018. “Embracing Colonizations: A New Paradigm for Species Association Dynamics.” *Trends in Ecology & Evolution* 33 (1): 4–14.

Ogihara, Mari H., Mikio Yoshiyama, Nobuo Morimoto, and Kiyoshi Kimura. 2020. “Dominant Honeybee Colony Infestation by Varroa Destructor (Acari: Varroidae) K Haplotype in Japan.” *Applied Entomology and Zoology* 55 (2): 189–97.

Penczykowski, Rachel M., Anna-Liisa Laine, and Britt Koskella. 2016. “Understanding the Ecology and Evolution of Host-Parasite Interactions across Scales.” *Evolutionary Applications* 9 (1): 37–52.

Pinho, Catarina, and Jody Hey. 2010. “Divergence with Gene Flow: Models and Data.” *Annual Review of Ecology, Evolution, and Systematics* 41 (1): 215–30.

Posada-Florez, Francisco, Daniel E. Sonenshine, Noble I. Egekwu, Clifford Rice, Robert Lupitskyy, and Steven C. Cook. 2019. “Insights into the Metabolism and Behaviour of Varroa Destructor Mites from Analysis of Their Waste Excretions.” *Parasitology* 146 (4): 527–32.

Poulin, Robert, Boris R. Krasnov, David Mouillot, and David W. Thieltges. 2011. “The Comparative Ecology and Biogeography of Parasites.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 366 (1576): 2379–90.

Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel A. R. Ferreira, David Bender, Julian Maller, et al. 2007. “PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses.” *American Journal of Human Genetics* 81 (3): 559–75.

Radloff, Sarah E., Colleen Hepburn, H. Randall Hepburn, Stefan Fuchs, Soesilawati Hadisoesilo, Ken Tan, Michael S. Engel, and Viktor Kuznetsov. 2010. “Population Structure and Classification of Apis Cerana.” *Apidologie* 41 (6): 589–601.

Ramsey, Samuel D., Ronald Ochoa, Gary Bauchan, Connor Gulbronson, Joseph D. Mowery, Allen Cohen, David Lim, et al. 2019. “Varroa Destructor Feeds Primarily on Honey Bee Fat

Body Tissue and Not Hemolymph." *Proceedings of the National Academy of Sciences of the United States of America* 116 (5): 1792–1801.

Ramu, Avinash, Michiel J. Noordam, Rachel S. Schwartz, Arthur Wuster, Matthew E. Hurles, Reed A. Cartwright, and Donald F. Conrad. 2013. "DeNovoGear: De Novo Indel and Point Mutation Discovery and Phasing." *Nature Methods* 10 (10): 985–87.

Rigaud, Thierry, Marie-Jeanne Perrot-Minnot, and Mark J. F. Brown. 2010. "Parasite and Host Assemblages: Embracing the Reality Will Improve Our Knowledge of Parasite Transmission and Virulence." *Proceedings. Biological Sciences / The Royal Society* 277 (1701): 3693–3702.

Roberts, J. M. K., D. L. Anderson, and W. T. Tay. 2015. "Multiple Host Shifts by the Emerging Honeybee Parasite, Varroa Jacobsoni." *Molecular Ecology* 24 (10): 2379–91.

Robinson, John D., Alec J. Coffman, Michael J. Hickerson, and Ryan N. Gutenkunst. 2014. "Sampling Strategies for Frequency Spectrum-Based Population Genomic Inference." *BMC Evolutionary Biology* 14 (December): 254.

Rueppell, O., A. M. Hayes, N. Warrit, and D. R. Smith. 2011. "Population Structure of Apis Cerana in Thailand Reflects Biogeography and Current Gene Flow Rather than Varroa Mite Association." *Insectes Sociaux* 58 (4): 445–52.

Sedlazeck, Fritz J., Philipp Rescheneder, and Arndt von Haeseler. 2013. "NextGenMap: Fast and Accurate Read Mapping in Highly Polymorphic Genomes." *Bioinformatics* 29 (21): 2790–91.

Sexton, Jason P., Jorge Montiel, Jackie E. Shay, Molly R. Stephens, and Rachel A. Slatyer. 2017. "Evolution of Ecological Niche Breadth." *Annual Review of Ecology, Evolution, and Systematics* 48 (1): 183–206.

Skotte, Line, Thorfinn Sand Korneliussen, and Anders Albrechtsen. 2013. "Estimating Individual Admixture Proportions from next Generation Sequencing Data." *Genetics* 195 (3): 693–702.

Solignac, Michel, Jean-Marie Cornuet, Dominique Vautrin, Yves Le Conte, Denis Anderson, Jay Evans, Sandrine Cros-Arteil, and Maria Navajas. 2005. "The Invasive Korea and Japan Types of Varroa Destructor, Ectoparasitic Mites of the Western Honeybee (Apis Mellifera), Are Two Partly Isolated Clones." *Proceedings. Biological Sciences / The Royal Society* 272 (1561): 411–19.

Tan, Hong-Wei, Guo-Hua Liu, Xia Dong, Rui-Qing Lin, Hui-Qun Song, Si-Yang Huang, Zi-Guo Yuan, Guang-Hui Zhao, and Xing-Quan Zhu. 2011. "The Complete Mitochondrial Genome of the Asiatic Cavity-Nesting Honeybee Apis Cerana (Hymenoptera: Apidae)." *PloS One* 6 (8): e23008.

Tan, Ken, Yufeng Qu, Zhengwei Wang, Zhiwei Liu, and Michael S. Engel. 2016. "Haplotype Diversity and Genetic Similarity among Populations of the Eastern Honey Bee from Himalaya-Southwest China and Nepal (Hymenoptera: Apidae)." *Apidologie* 47 (2): 197–205.

Techer, Maeva A., Rahul V. Rane, Miguel L. Grau, John M. K. Roberts, Shawn T. Sullivan, Ivan Liachko, Anna K. Childers, Jay D. Evans, and Alexander S. Mikheyev. 2019. "Divergent Evolutionary Trajectories Following Speciation in Two Ectoparasitic Honey Bee Mites." *Communications Biology* 2 (1): 357.

Tomasini, Matteo, and Stephan Peischl. 2020. “When Does Gene Flow Facilitate Evolutionary Rescue?” *Evolution; International Journal of Organic Evolution* 74 (8): 1640–53.

Traynor, Kirsten S., Fanny Mondet, Joachim R. de Miranda, Maeva Techer, Vienna Kowallik, Melissa A. Y. Oddie, Panuwan Chantawannakul, and Alison McAfee. 2020. “Varroa Destructor: A Complex Parasite, Crippling Honey Bees Worldwide.” *Trends in Parasitology* 36 (7): 592–606.

Wala, Jeremiah, Cheng-Zhong Zhang, Matthew Meyerson, and Rameen Beroukhim. 2016. “VariantBam: Filtering and Profiling of next-Generational Sequencing Data Using Region-Specific Rules.” *Bioinformatics* 32 (13): 2029–31.

Warrit, Natapot, Deborah Roan Smith, and Chariya Lekprayoon. 2006. “Genetic Subpopulations of Varroa Mites and Their Apis Cerana Hosts in Thailand.” *Apidologie* 37 (1): 19–30.

Whiteley, Andrew R., Sarah W. Fitzpatrick, W. Chris Funk, and David A. Tallmon. 2015. “Genetic Rescue to the Rescue.” *Trends in Ecology & Evolution* 30 (1): 42–49.

Yamasaki, Yo Y., Ryo Kakioka, Hiroshi Takahashi, Atsushi Toyoda, Atsushi J. Nagano, Yoshiyasu Machida, Peter R. Møller, and Jun Kitano. 2020. “Genome-Wide Patterns of Divergence and Introgression after Secondary Contact between Pungitius Sticklebacks.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 375 (1806): 20190548.

Yancan, Li, Chao Tianle, Fan Yunhan, Lou Delong, and Wang Guizhi. 2019. “Population Genomics and Morphological Features Underlying the Adaptive Evolution of the Eastern Honey Bee (Apis Cerana).” *BMC Genomics* 20 (1): 869.

Zhou, Ting, Denis L. Anderson, Zachary Y. Huang, Shuangxiu Huang, Jun Yao, Tan Ken, and Qingwen Zhang. 2004. “Identification of Varroa Mites (Acari: Varroidae) Infesting Apis Cerana and Apis Mellifera in China.” *Apidologie* 35 (6): 645–54.

Acknowledgments

MAT’s research was supported by a postdoctoral fellowship from the Japan Society for Promotion of Science (JSPS) (P19723), Kakenhi Grant-in-Aid, (19F19723) and the Okinawa Institute of Science and Technology (OIST). ASM was supported by a Future Fellowship from the Australian Research Council (FT160100178) and a Kakenhi Grant-in-Aid for Scientific Research from the JSPS (18H02216). JMKR was supported by the Australian Centre for International Agricultural Research (ACIAR) and the Australian Department of Agriculture, Water and Environment (DAWE). RAC was supported by National Institutes of Health (R01-HG007178). We are grateful to Jo Si Lay Tan and Lijun Qiu for their advice and guidance in developing the wet lab workflow for *Varroa* mite sequencing. We also would like to thank the OIST Sequencing Center for

assisting us in the sample sequencing. We wish to thank Steven D. Aird, technical editor, for reviewing and improving our manuscript.

Author Contributions

MAT and ASM designed research, analyzed population genetics data and wrote the manuscript. MAT processed the samples in the wet lab until library preparation and ran the demographic inferences. JMKR collected and provided samples from the mite CSIRO collection, and helped in the data interpretation. RAC analyzed data, wrote the manuscript, provided reproducible online resources and data for the estimation of mutation rate.

Competing interests

The authors declare that they have no competing interests.

Data accessibility

Raw FASTQ files for all individuals sequenced were submitted to the DNA Databank of Japan and uploaded to GenBank under the bioproject PRJDB9195 with the run accessions DRR209082-DRR209125 and DRR212369-DRR212380. The pipeline developed using Snakemake, Rmarkdown for data exploration, demographic inferences input files and FASTA sequences alignment (mtDNA) are made available for reproducibility on https://github.com/MaevaTecher/varroa-host-jump. Variant calling files, genome indexing files and input lists are readily available at DRYAD.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Table 1: *Varroa* specimens used for population genomics were collected from their native range across 11 countries, both on their original or new honey bee host. Newly reported mtDNA lineages are indicated in bold

Species	Host	Country	Date	N	mtDNA lineages as defined by standard *COX1*
V. destructor	Original host	Japan	1994	2	VD Japan J1
		South Korea	1996	1	VD Korea K1
		China	2001-2002	6	VD China C1
		Thailand	2003	1	VD Viet Nam V1
		Viet Nam	1996	2	VD Viet Nam V1
		Myanmar	2003	1	VD Viet Nam V1
		Nepal	2000	1	**VD Nepal 2**
V. destructor	New host	Japan	1996	1	VD Korea K1
		South Korea	1996	3	VD Korea K1
		China	2001	5	VD Korea K1
		Viet Nam	1996, 1998	6	VD Korea K1
		Myanmar	2003	2	VD Korea K1
V. jacobsoni	Original host	India	2002	2	VJ Bengaluru 1
		Thailand	2003	1	VJ Laos L1
		Viet Nam	2003-2004	5	VJ Laos L1
		Malaysia	1995	1	VJ Malaysia 2
		Indonesia	1998, 2002	6	VJ Java 1
		Papua New Guinea	1997, 2008, 2015	4	VJ PNG 1
V. jacobsoni	Original host	Indonesia	1996	3	VJ Java 1
V. jacobsoni	New host	Thailand	1994	1	**VJ North Thailand 4**
		Papua New Guinea	2008, 2014	9	VJ PNG 1

Table 2: Parameter estimates for the most likely demographic scenarios with isolation with migration and population expansion for each species jump. As stated in fastsimcoal2 manual, these upper-bound priors indicated with an asterisk were not bounded by the analysis.

	Priors		*V. destructor*		*V. jacobsoni*	
	Distribution	Range	Mean	CI 95%	Mean	CI 95%
N_{VAC1} (haploid genomes) At present, mite effective population size on original host *A. cerana*	uniform	[2; 20,000*]	17,166	[16,598; 17,731]	17,376	[16589; 18144]
N_{VAM0} (haploid genomes) At present, mite effective population size on novel host *A. mellifera*	uniform	[2; 20,000*]	29,659	[28,722; 30,666]	5,220	[4,787; 5,683]
N_{BOTAM} (haploid genomes) Past mite effective founding population size during host switch	uniform	[2; 1,000*]	156	[146; 165]	138	[126; 151]
T_{JUMP} (generation) Time of new founding population split by host switching	uniform and bounded	[1; 1,000]	887	[866; 906]	183	[179; 187]
T_{BOTEND} (generation) End of the founder effect period	uniform and bounded	[1; 1,000]	710	[693; 726]	170	[165; 175]
G_{AM} (haploid genomes per generation) Growth rate of the expanding newly founded mite population (from present to past)	uniform	[-1; 0*]	-0.0014	[-0.0015; -0.0013]	-0.4636	[-0.4895; -0.4399]
M_{ACtoAM} (haploid genomes per generation)	log uniform	[$1x10^{-6}$; 100*]	0.0003	[0.0003; 0.0003]	0.0161	[0.0154; 0.0168]

Migration rate of parasite spillover (original toward new host)						
M_{AMtoAC} (haploid genomes per generation) Migration rate of parasite spillback (new toward new original host)	log uniform	[$1x10^{-6}$; 100*]	0.0001	[0.0001; 0.0001]	0.0295	[0.0283; 0.0309]
Mutation rate (fixed-parameter)	-	8.0×10^{-10}				
Inbreeding (fixed-parameter)	-	none				

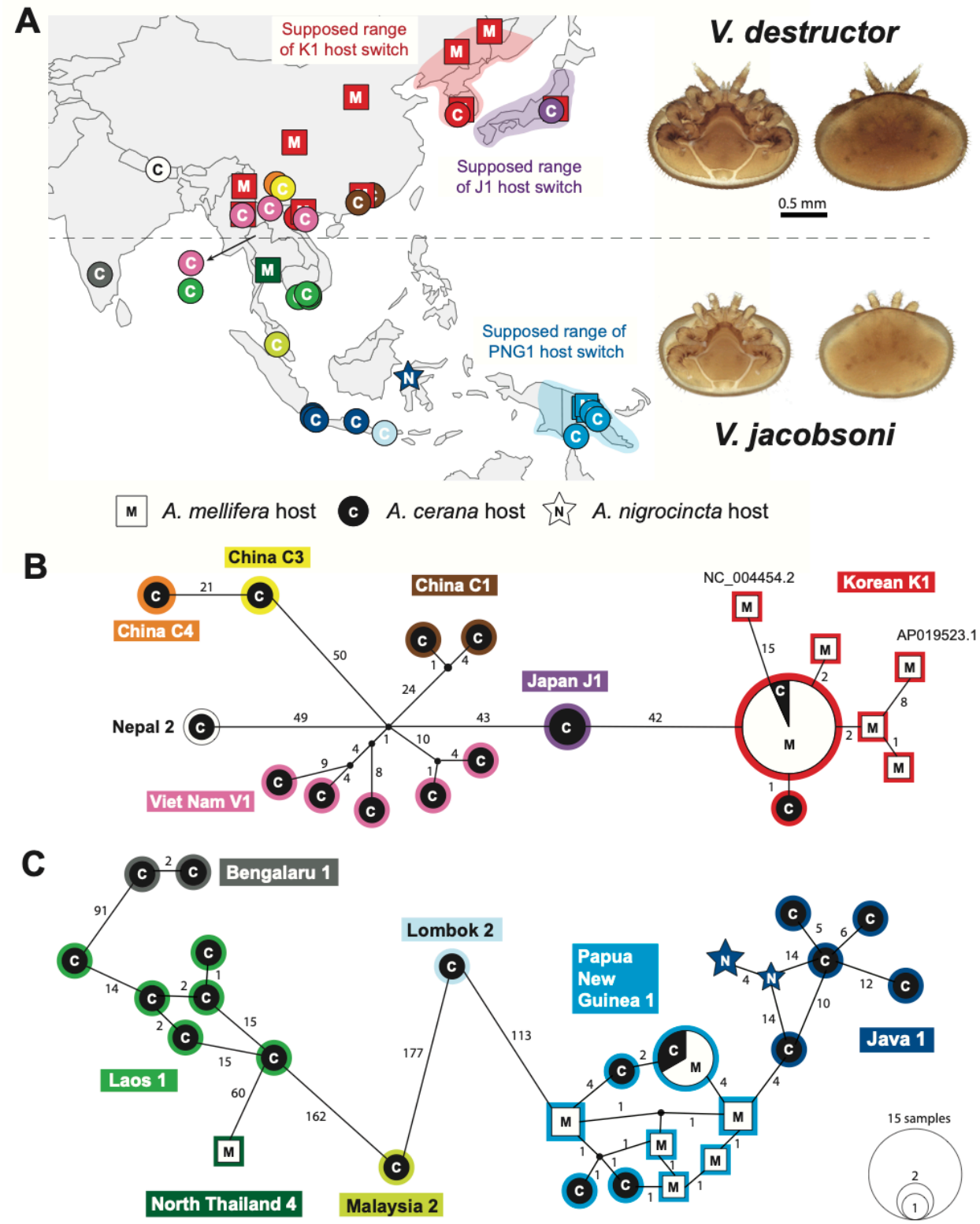

A
Supposed range of K1 host switch
Supposed range of J1 host switch
Supposed range of PNG1 host switch
V. destructor
0.5 mm
V. jacobsoni
A. mellifera host
A. cerana host
A. nigrocincta host
B
China C3
China C4
China C1
Japan J1
Korean K1
NC_004454.2
AP019523.1
Nepal 2
Viet Nam V1
C
Bengalaru 1
Lombok 2
Papua New Guinea 1
Java 1
Laos 1
North Thailand 4
Malaysia 2
15 samples

Figure 1: Mitogenomic phylogeographic networks uncover cryptic diversity in host-switched lineages. Both mite species exhibit geographic structure on their original host, with lineages classified using their *COX1* sequences (A). The color of each circle/square indicates a unique mitogenome sequence and size is proportionate to the sample size (B). Previous studies based on *COX1* found that *V. destructor* host-switched lineages largely belong to the so-called 'Korean K1' haplotype. However, the mitogenome-wide analysis revealed four to six mitotypes on the new host, suggesting previously unreported switches by mites from the same geographic region (B). For the most recent jump in 2008, out the seven lineages identified for *V. jacobsoni* only the known PNG1 succeeded in crossing the species barrier (Roberts, Anderson, and Tay 2015) (C). Yet, looking at the whole mitochondrial genome, it becomes evident that the host switch involved a number of independent female founders from closely related populations, or even the same population. Thus, mitogenomic data suggest a relatively diverse population of host-switched mites in both species.

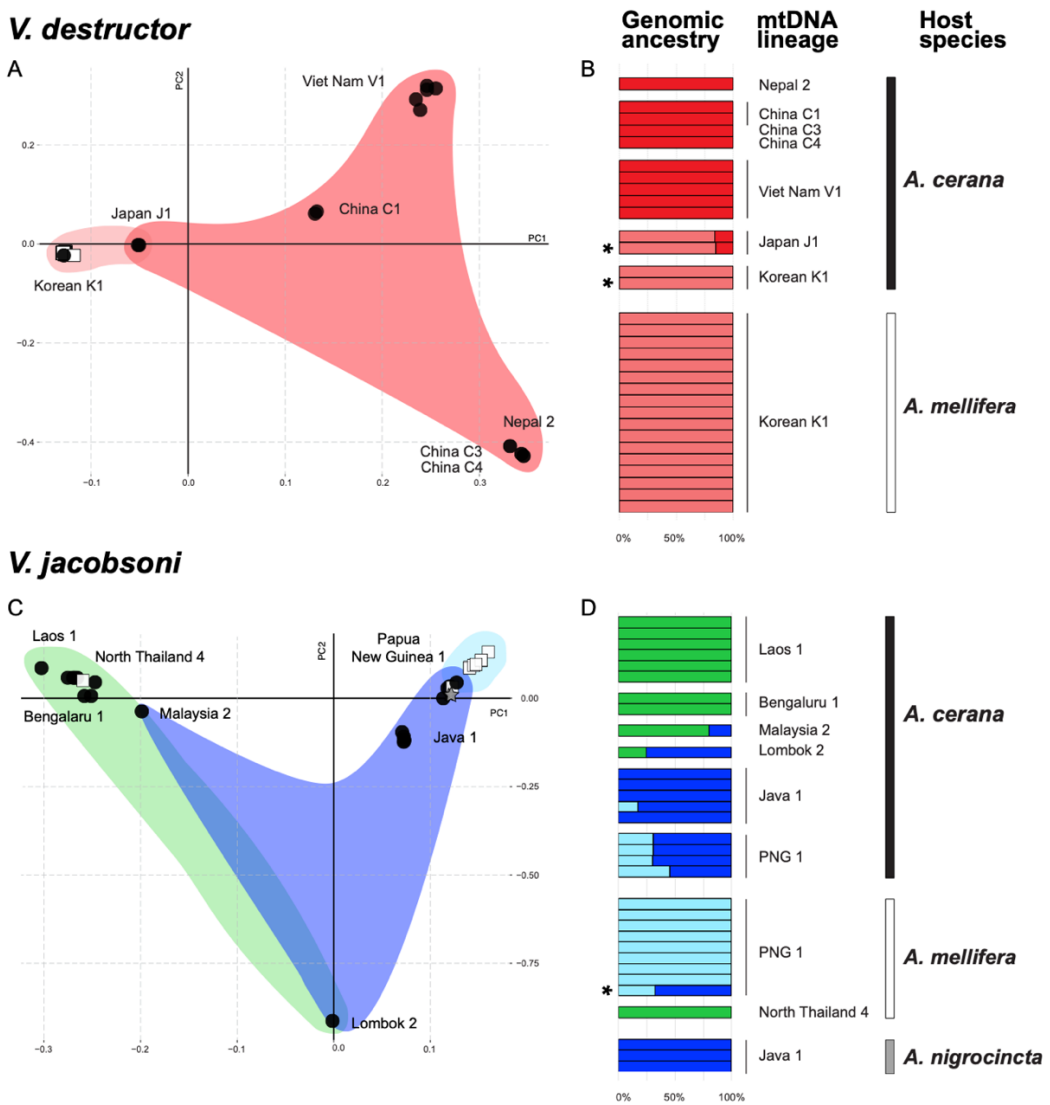


Figure 2: Loss of genetic diversity and rapid genomic differentiation occurred in spite of large founding size and migration between sympatric honey bee hosts. *V. destructor* populations on its original host *A. cerana* (black circle) show geographic structure across their native range (A). Envelopes around the PCA points are color-coded according to nuclear genetic ancestry (B). Host-switched populations (white square) are genetically homogenous throughout Asia and are most similar to the populations with the same mitotype (K1) (B). Despite a more recent host switch (~2008), a similar pattern is observed for *V. jacobsoni* in South-Asia and Oceania, where the PNG1 lineage underwent a host switch (C). However, other switches appear possible elsewhere in *V. jacobsoni*'s range, such as the specimen collected from *A. mellifera* in North Thailand (green cluster) (D). In both species, there is evidence of mites drifting between hosts, as indicated by genetic analysis of their gut contents, indicated by asterisks. The overall picture for both species is similar, where reproductive differentiation after host switches is coupled with additional opportunities for gene flow via drifting, and potential for additional switches happening elsewhere in the range.

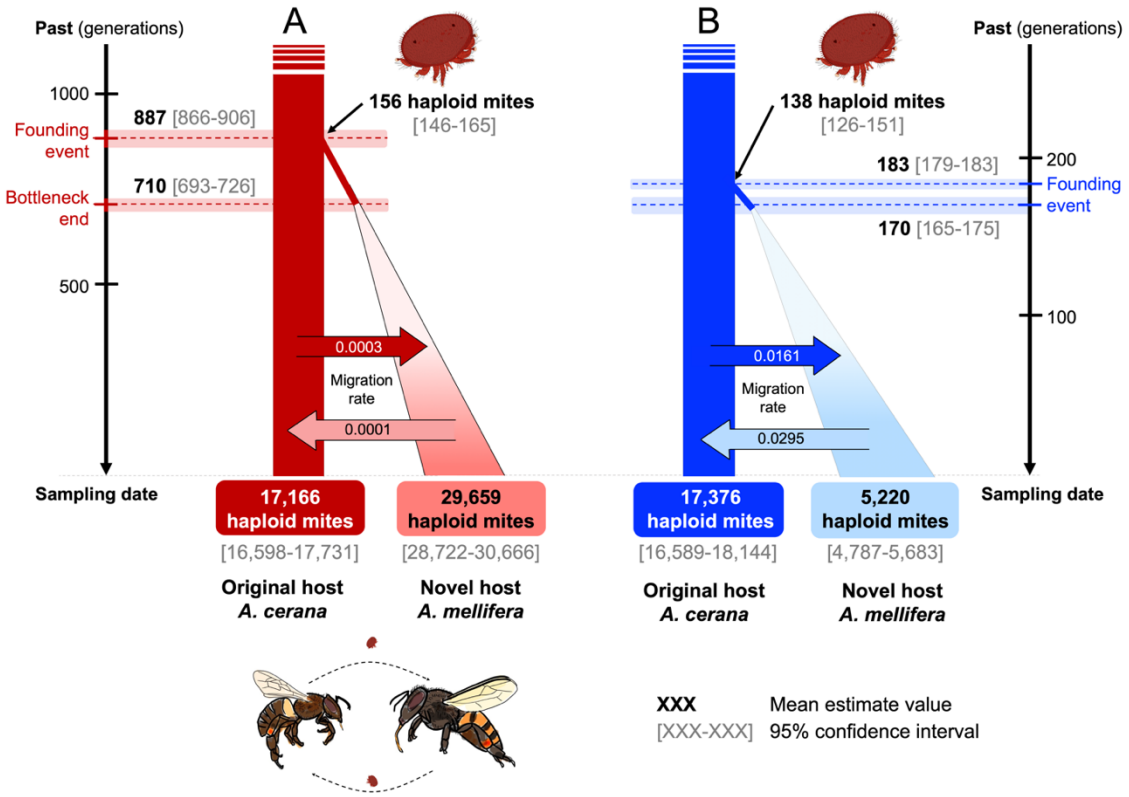


Figure 3: Multiple jumps to *A. mellifera* and isolation with migration likely allowed *V. destructor* and *V. jacobsoni* to successfully switch hosts and persist. Graphical illustration of the inferred scenario with mean parameter estimates in bold and associated 95% confidence interval in [light gray] with isolation with bidirectional migration (A) for *V. destructor* and (B) for *V. jacobsoni*. Estimated population sizes are given in haploid genome size. Migration rates were kept as default values in haploid genome per generation. The coalescent model reconstructed timings of expansion that are consistent with known observations. Yet it suggests that the host switch was not instantaneous and may have involved a period of progressive adaptation before the mites spread beyond their source populations.

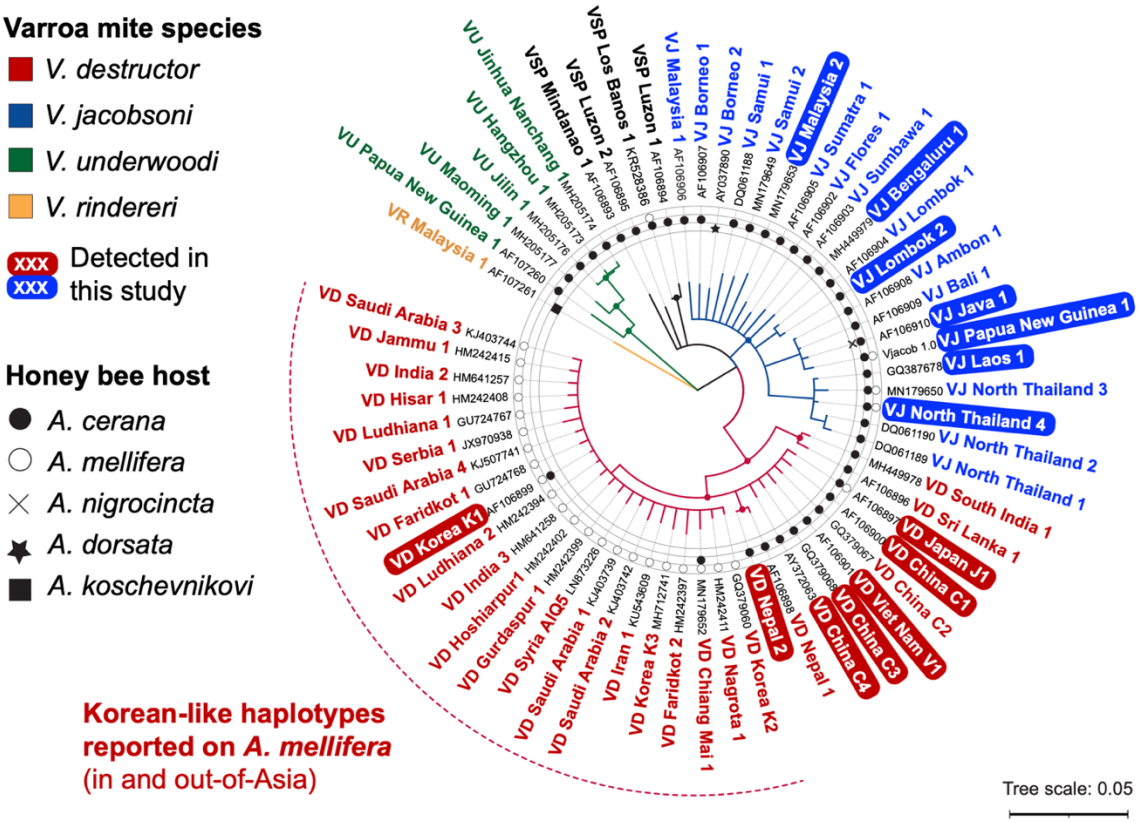
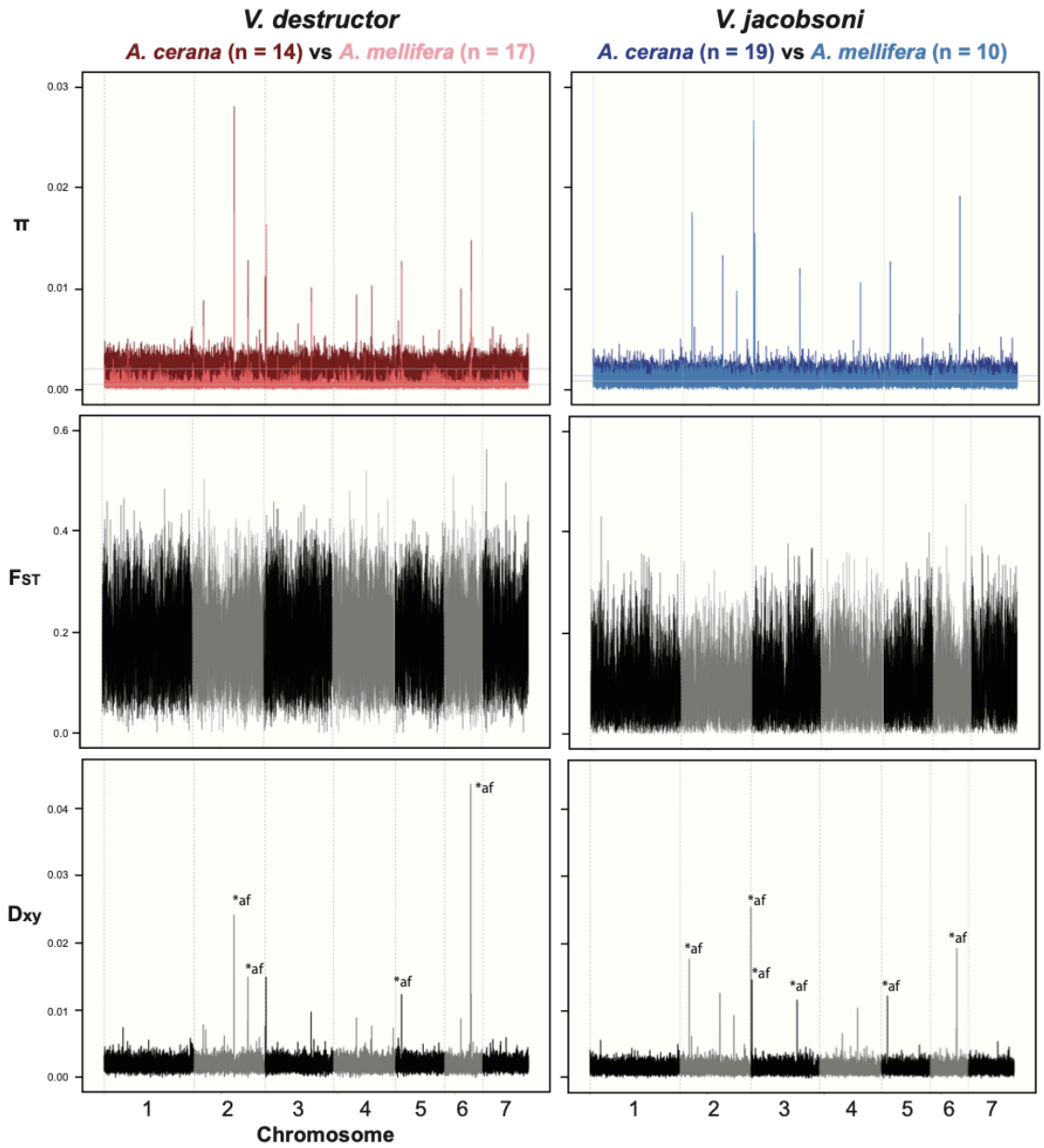


Figure S1: Sequenced *V. destructor* and *V. jacobsoni* mites originated from diverse maternal genetic backgrounds (lineages) shown by mtDNA *COX1* barcoding. A neighbor-joining tree was built using the Tamura-Nei genetic distance model with 1,000 bootstraps (> 80% are shown as filled circles in the base of each branch). *V. rindereri* was considered as an outgroup and the host range indicated (D. L. Anderson and Trueman 2000). *V. destructor* mites collected throughout Asia on their original host belong to seven distinct lineages (see details in Table S2). By contrast, all the host-switched mites across that range belong to the Korean K1 haplotype [AF106899], which is also found worldwide (Traynor et al. 2020). *V. jacobsoni* mites collected throughout Southeast Asia and Oceania, we identified seven lineages including two newly described: Lombok 2 and North Thailand 4. All mites parasitizing *A. mellifera* in Papua New Guinea were infested only by the *V. jacobsoni* Papua New Guinea 1 haplotype. The novel haplotype, North Thailand 4, was found on one *A. mellifera* colony in the northern limit of *V. jacobsoni* native range.



940 **Figure S2: Parallel reduction of genetic diversity and genomic divergence between hosts in *V. destructor* and *V. jacobsoni*.** First row: Nucleotide diversity in 20kb window along all seven chromosomes with mites parasitizing the same host pooled as one population for each *Varroa* species. Genetic divergence estimated by pairwise F_{ST} (second row) and absolute divergence using D_{xy} estimate (third row) in 20kb window,
945 show a strong host differentiation all along the genome with few outliers. All sites with D_{xy} value > 0.01 were investigated manually by examining individual genotypes, which showed an excess in heterozygosity levels. These regions were also poorly conserved within species and had BLAST matches elsewhere in the genome, suggesting the presence of sequencing artefacts rather than biological signal (labelled *af on the plots).

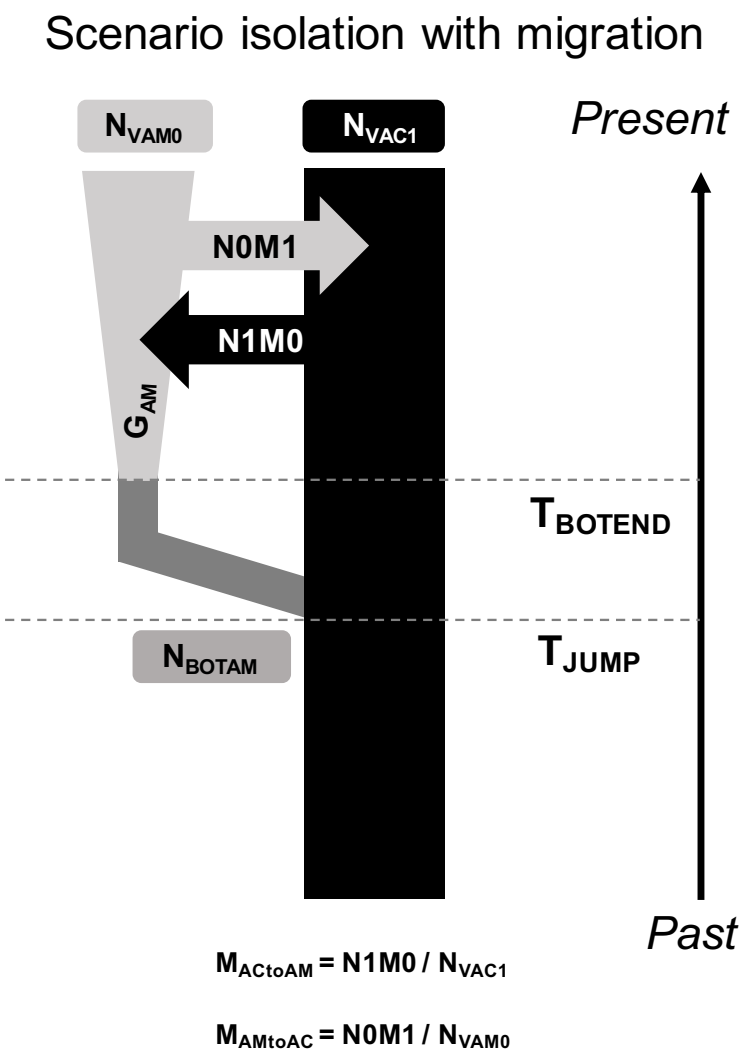


Figure S3 Graphical representation of *Varroa* host switch demographic scenario designed following population genomics analysis. The model was run independently for each *Varroa* species but contained the same parameters. N_{VAM0} is the present effective population size on the novel host (*A. mellifera*) whereas N_{VAC1} is the present population size on the original host (*A. cerana*). At a time in the past, T_{JUMP}, a founding size of N_{BOTAM} switched durably host from *A. cerana* to *A. mellifera*. The founder effect associated with host switch ended at T_{BOTEND} and population on novel host expanded at a growth rate G_{AM}. As both hosts are in artificial sympatry with evidence of gene flow, migration rates were estimated for spillover (M_{ACtoAM}) and spillback (M_{AMtoAC}). See Table 2 for parameters prior details.

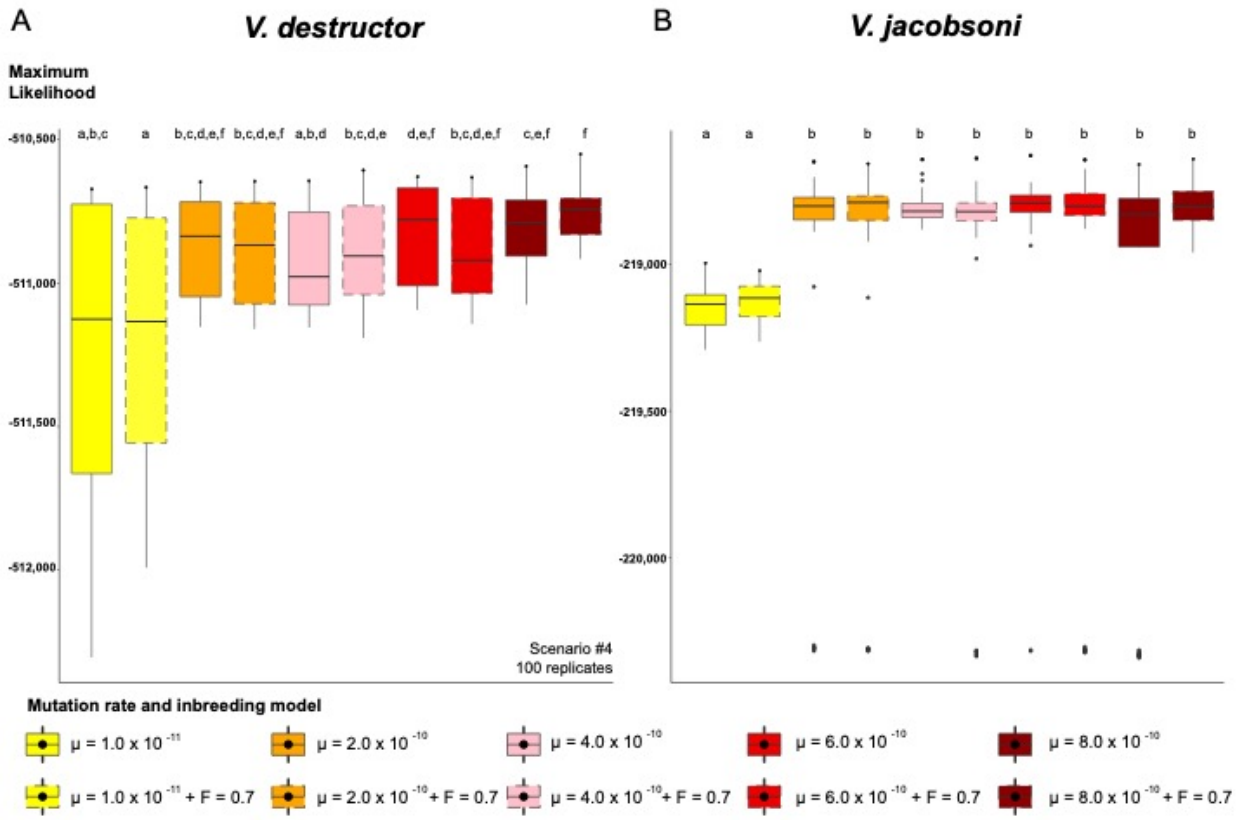
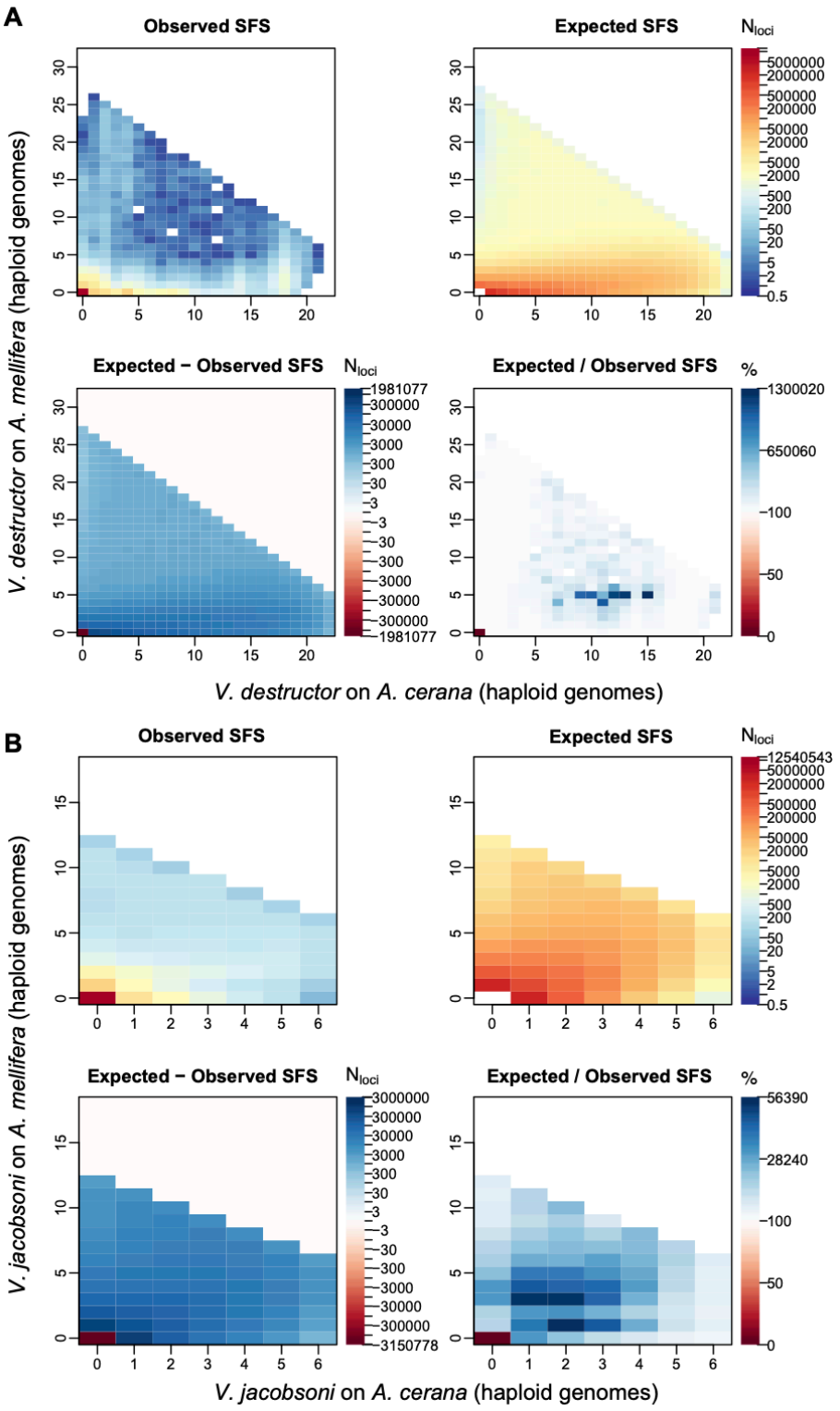


Figure S4 Distribution of the maximum likelihood obtained from the best pseudo-observed SFS generated by demographic inferences across different mutation rates and inbreeding levels in *Varroa* species. Replicates generated with lower mutation rates than the directly estimated 8.0 x 10^{-10} were not associated with a better fit to the observed SFS from our genomic data for *V. destructor* (A) and *V. jacobsoni* (B). No significant differences were detected in the likelihood variance among different mutation rates. However, pairwise Wilcoxon test significance (letters) after Bonferroni corrections showed that the lower mutation rate differentiated from other rates, especially in *V. jacobsoni*.



975 **Figure S5: Observed vs expected joint 2D-SFS from the best replicate of the isolation with migration scenario for *V. destructor* and *V. jacobsoni*.**

Table S1: Detailed sampling information for each of the *Varroa* mite sequenced for population genomics as well as for estimation of the mutation rate. Sampling details include *Varroa* species, host identity, country and locality, date and geographical coordinates. Sequencing details include the DRA accession on NCBI, the total number of reads, mapped reads, mapping ratio and mean read depth. Mitochondrial lineages and mitotype identity are also included for each specimen.

Table S2: Repertoire of standard mtDNA markers used to determine mite species and lineages of sequenced specimens. Names given to new mitotype were chosen to follow the revised nomenclature and compared to *V. destructor* and *V. jacobsoni* reference sequences (source and sequence accession indicated).

Table S3: Observed (O_{HOM}) and expected (E_{HOM}) homozygosity and inbreeding coefficient (F) values computed for each sequenced mite on N_{SITES} for the population genomics analysis.

Table S4: Pairwise Weir and Cockerham F_{ST} values computed between host-adapted *Varroa* species and between host-adapted *Varroa* mtDNA lineages. The color shades indicate relatively high genetic differentiation (in red), moderate (in yellow) to a low one (in green).

Table S5: Number of reads from whole-body *Varroa* sample mapped against *A. mellifera* and *A. cerana* mitochondrial reference genomes. Host metagenomics identity was assigned by the higher number of reads and compared to the field collection information. The proof of concept was confirmed with Mom and Son samples collected while feeding on their *A. mellifera* host.