Supplementary Methods Information

**DWV variant detection**

cDNA was generated from 1 g RNA per sample using qScript XLT reverse transcriptase (Quantabio, Beverly, USA). Samples were then analyzed in triplicate by qPCR with DWVQ\_F1 and R1 primers gene [1,2] and HRM (54 – 95 °C) using MeltDoctor HRM Master Mix and a QuantStudio 7 platform (Applied Biosystems/Thermo Fisher Scientific, USA).

Selected samples were Sanger sequenced directly following qPCR and HRM. Samples were diluted 1:5 with H2O and combined with 0.2 M DWVQ\_F1 primer. Sequencing was provided by the Massey Genome Service (Massey University, NZ).

HRM analyses revealed 45 individuals with more than a single HRM peak, indicating infection with multiple strains. For 43 of these individuals, we re-amplified the *RdRp* fragment for next-generation sequencing using the NEB-Next Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA). For each PCR product sample, we attached dual-indexed primers (NEBNext Multiplex Oligos Kit) to identify each sample and the sequencing adapter needed for the subsequent Illumina sequencing protocol. The amplification round was performed in a volume of 25 μL using 7.5 μL of every sample, 2.5 μL i5 primers and 2.5 μL i7 primers (E7600S, New England Biolabs), and 12.5 μL ofNEB PRC Master Mix. The mixture was denatured at 98˚C for 30 s followed by 8 cycles of 10 s at 98˚C and 75 s at 65˚C and a final cycle of 5 min at 65˚C. PCR products were then purified using Sera-Mag SpeedBeads (GE Healthcare) and resuspended in a volume of 15 μL of 0.1X TE buffer. We determined the DNA concentration in every pool using a Qubit High Sensitivity Assay (Qubit1) to prepare 4 nM of every pool. Pools were then combined and submitted to UCR Institute for Integrative Genome Biology Genomics Core for paired-end sequencing (2x300bp) on the Illumina MiSeq.

Following demultiplexing, MiSeq reads were processed through the linux cluster of the UCR High-Performance Computing Center. Using QIIME [3], forward and reverse primers were trimmed from the reads. Using the dada2 pipeline [4], reads were filtered, paired ends were merged, and sequence tables were constructed.

1. Martin SJ, Highfield AC, Brettell L, Villalobos EM, Budge GE, Powell M, Nikaido S, Schroeder DC. 2012 Global honey bee viral landscape altered by a parasitic mite. *Science (80-. ).* **336**, 1304–1306.

2. Highfield AC, El Nagar A, Mackinder LCM, Noël LMLJ, Hall MJ, Martin SJ, Schroeder DC. 2009 Deformed wing virus implicated in overwintering honeybee colony losses. *Appl. Environ. Microbiol.* **75**, 7212–7220.

3. Caporaso JG *et al.* 2010 correspondence QIIME allows analysis of high- throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nat. Methods* **7**, 335–336. (doi:10.1038/nmeth0510-335)

4. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016 DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583. (doi:10.1038/nmeth.3869)

5. Yue C, Genersch E. 2005 RT-PCR analysis of Deformed Wing Virus in honeybees (Apis mellifera) and mites (Varroa destructor). *J. Gen. Virol.* **86**, 3419–3424.

Supplemental Figures

Figure S1: Neighbor joining tree of 100bp fragment of DWV *RdRp*. Included are the 23 variants identified here, as well as the variants from Hawaii sequenced in Martin et al. (2012).

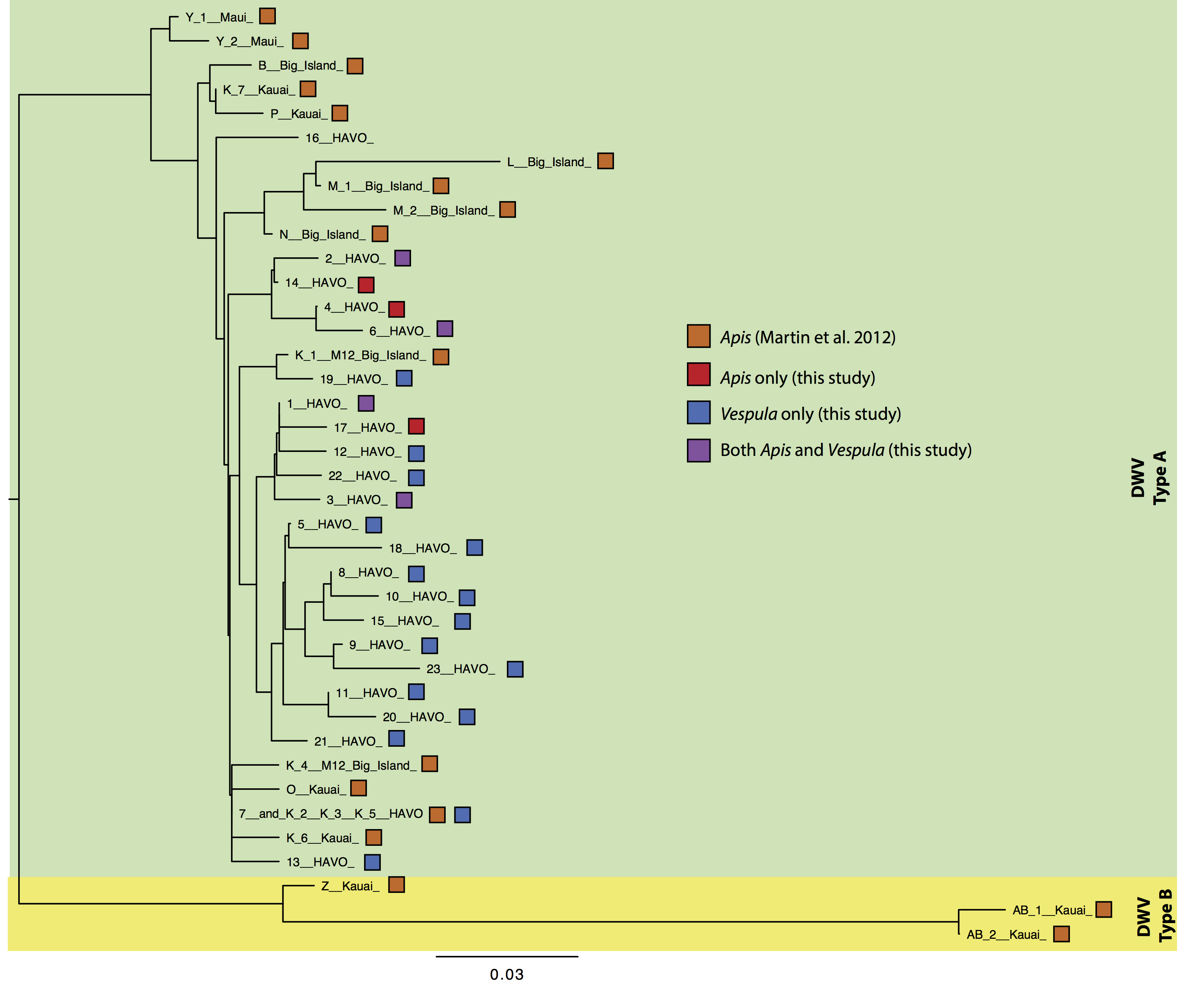


Figure S2. Haplotype map for sequences obtained before and after the arrival of *Varroa* (both host species included). Note that one clade on the left side of the figure is virtually extinguished after varroa arrival.

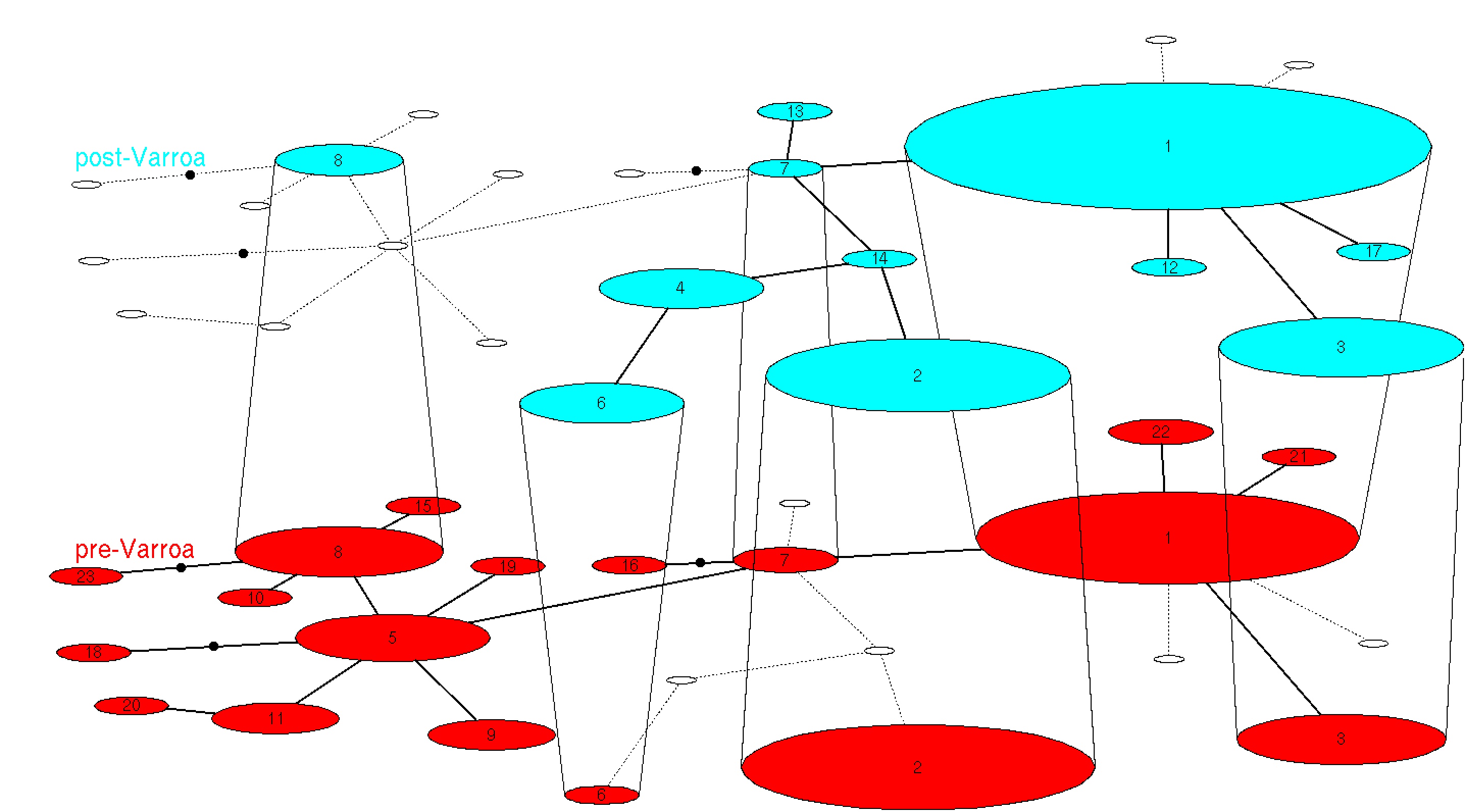


Figure S3. The number of individuals positive for each of 23 detected variants, by sample type. Variants are ordered by frequency in post-*Varroa* *Apis.*

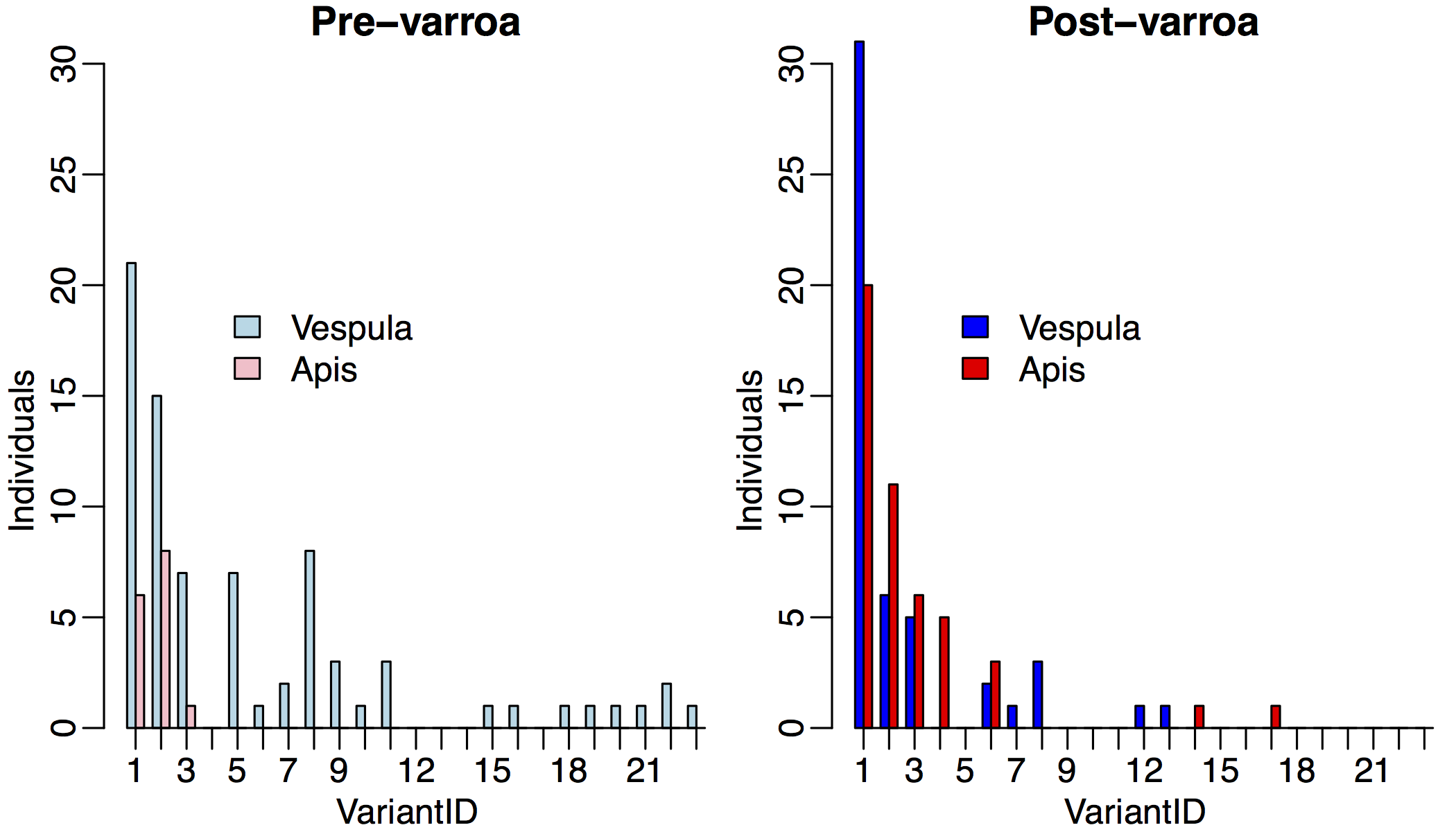


Figure S4 Negative strand rt-PCR for DWV. Tagged PCR was performed using tag and B23 primers on cDNA made with tag-F15 and B23 primers, as in ref [5]. 1: pool of n=12 *Apis* samples from California. 2: Pool of n=20 *Apis* post-*Varroa* samples from Hawaii Volcanoes National Park. 3: First pool of n=20 *Vespula* samples from Hawaii Volcanoes National Park; 4: Second pool of n=20 *Vespula* samples from Hawaii Volcanoes National Park. 5: Pool of n=14 *Apis* pre-*Varroa* samples; NTC: no template control. PCR products from samples 1-3 were sequenced to confirm DWV sequence identity.

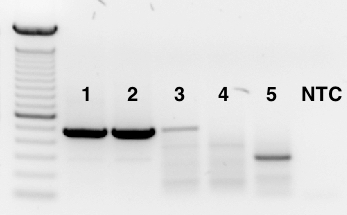


Table S1. Primers used in this study (excel sheet)

Table S2. Alternative analyses for pathogen community compositional changes before and after *Varroa* arrival in *Vespula* samples. We used several different methods to account for samples without detectable levels of pathogens (negatives) in our permutational MANOVA analyses. In Method 1, negative samples were coded as half of the minimum value of the pathogen in that species, representing the lowest detectable limit of that assay. Euclidean distances were used in adonis() following log transformation of the data. In Method 2, negatives were coded as zero, preventing log transformation. We then used Bray Curtis differences in the permutational MANOVA. In Method 3, we coded negatives as in Method 1, without log transformation, using Bray-Curtis differences in the permutational MANOVA.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Vespula*** |  | | **df** | **SS** | **Mean Sqs** | **F.model** | **R2** | ***P*** |
| Method 1 | age | 1 | 104.91 | 104.91 | 4.98 | 0.08 | 0.001 |
| colony | 18 | 854.22 | 47.46 | 2.26 | 0.62 | 0.002 |
| residuals | 20 | 421.33 | 21.07 |  | 0.31 |  |
| total | 39 | 1380.46 |  |  | 1 |  |
| Method 2 | age | 1 | 0.66 | 0.66 | 2.89 | 0.06 | 0.024 |
| colony | 18 | 5.94 | 0.33 | 1.45 | 0.53 | 0.031 |
| residuals | 20 | 4.56 | 0.23 |  | 0.41 |  |
| total | 39 | 11.16 |  |  | 1 |  |
| Method 3 | age | 1 | 0.66 | 0.66 | 2.89 | 0.06 | 0.024 |
| colony | 18 | 5.94 | 0.33 | 1.44 | 0.53 | 0.029 |
| residuals | 20 | 4.56 | 0.23 |  | 0.41 |  |
| total | 39 | 11.15 |  |  | 1 |  |