

# Sequence Recombination and Conservation of *Varroa destructor Virus-1* and *Deformed Wing Virus* in Field Collected Honey Bees (*Apis mellifera*)

Hui Wang<sup>1\*</sup>, Jiazheng Xie<sup>2</sup>, Tim G. Shreeve<sup>3</sup>, Jinmin Ma<sup>2</sup>, Denise W. Pallett<sup>1</sup>, Linda A. King<sup>3</sup>, Robert D. Possee<sup>1</sup>

**1** Centre for Ecology and Hydrology, Natural Environmental Research Council, Wallingford, Oxfordshire, United Kingdom, **2** Beijing Genome Institute, Yantian District, Shenzhen, China, **3** Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, United Kingdom

## Abstract

We sequenced small (s) RNAs from field collected honeybees (*Apis mellifera*) and bumblebees (*Bombus pascuorum*) using the Illumina technology. The sRNA reads were assembled and resulting contigs were used to search for virus homologues in GenBank. Matches with *Varroa destructor virus-1* (VDV1) and Deformed wing virus (DWV) genomic sequences were obtained for *A. mellifera* but not *B. pascuorum*. Further analyses suggested that the prevalent virus population was composed of VDV-1 and a chimera of 5'-DWV-VDV1-DWV-3'. The recombination junctions in the chimera genomes were confirmed by using RT-PCR, cDNA cloning and Sanger sequencing. We then focused on conserved short fragments (CSF, size > 25 nt) in the virus genomes by using GenBank sequences and the deep sequencing data obtained in this study. The majority of CSF sites confirmed conservation at both between-species (GenBank sequences) and within-population (dataset of this study) levels. However, conserved nucleotide positions in the GenBank sequences might be variable at the within-population level. High mutation rates (Pi>10%) were observed at a number of sites using the deep sequencing data, suggesting that sequence conservation might not always be maintained at the population level. Virus-host interactions and strategies for developing RNAi treatments against VDV1/DWV infections are discussed.

**Citation:** Wang H, Xie J, Shreeve TG, Ma J, Pallett DW, et al. (2013) Sequence Recombination and Conservation of *Varroa destructor Virus-1* and *Deformed Wing Virus* in Field Collected Honey Bees (*Apis mellifera*). PLoS ONE 8(9): e74508. doi:10.1371/journal.pone.0074508

**Editor:** Robert Belshaw, Plymouth University, United Kingdom

**Received:** May 16, 2013; **Accepted:** August 2, 2013; **Published:** September 18, 2013

**Copyright:** © 2013 Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Funding was provided by Natural Environment Research Council (United Kingdom, grant no. NE/I000593/1), Oxford Brookes University and Ministry of Science and Technology of the Peoples Republic of China (China, grant no. 2011DFA33220). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

\* E-mail: huw@ceh.ac.uk

## Introduction

Virus infections are closely associated with colony collapse in the European honeybee *Apis mellifera* (reviewed by [1,2]). Recent pandemics of colony collapse around the world have also been associated with invasion of honeybee populations by the mite *Varroa destructor*. Colonisation of hives by *V. destructor* in areas previously mite free has resulted in the selection of certain Deformed wing virus (DWV) strains [3-6]. The emergence of these selected strains appears to disturb the balance between the virus and honeybee populations, thus initiating collapse of honeybee colonies [3-7]. In areas where *V. destructor* is established, virus recombination occurs and chimeric viruses between DWV and the *Varroa destructor virus-1* (VDV1) [8-10] have been reported [10,11].

Both DWV [1,5,6,11-42] and VDV-1 [8-10] belong to the Iflaviridae (genus *Iflavirus*) [8,43], a picorna-like family of insect

viruses. There is also a Kakugo virus (KV, genus *Iflavirus*) [44-46], which is very closely related to DWV genetically but is associated with aggressive behaviour in guard bees. These viruses have positive-sense, single-stranded RNA (+ssRNA) genomes of approximately 10 Kb, containing a single open reading frame flanked by 5' and 3' untranslated regions (UTR). The two viruses are phylogenetically related and form a DWV/VDV-1 clade with conserved amino acid motifs and a 3'-UTR [8]. As expected for picorna-like viruses, DWV populations display high mutation rates and quasi-species characteristics (reviewed by [25]). Chimeric viruses between DWV and VDV-1 have been reported in Israel [10] and the UK [11], suggesting that genome recombination between the two viruses may not be rare. In both cases, the chimera viruses co-persisted with one of the parents, either VDV-1 [10] or DWV [11]. In the latter case, two chimera viruses distinguishable by recombination patterns were detected [11]. The apparent unpredictable

recombination pattern has added to the complexity of evolutionary trends in the DWV/VDV-1 group. Most importantly, rapid viral evolution represents an elevated challenge for developing treatments and/or controlling measures against the virus infections.

One of the new strategies to control viral diseases is RNA interfering (RNAi) technology (reviewed in [1]). RNAi (also known as post-transcriptional gene silencing) is an ancient intracellular mechanism shared by both prokaryotes and eukaryotes and can mediate a specific anti-virus function (e.g., reviewed by [47,48]). In brief, host RNase III-like Dicer enzymes target double-stranded (ds) RNAs, such as viral RNA replication intermediates. A cascade of host enzyme actions is initiated to produce primary and secondary viral derived small interfering RNAs (vsiRNA) that specifically target the viral ssRNAs based on sequence complementary homology [49]. Based on their specific anti-virus function, vsiRNA sequencing has become a new technique for detecting unknown virus prevalence and discovering novel viruses, particularly from plant and insect materials [50-52]. High throughput next generation sequencing (NGS) technologies makes sequencing based approaches very efficient for virus survey and discovery [53-56]. Based on vsiRNAs' virus inhibitory function, RNAi technology has also been developed to control virus infections [1,57] including within insects [58]. In honeybees, RNAi has been successfully employed against Israeli acute paralysis virus (IAPV) infections in both laboratory and large-scale trials [59,60]. Anti-IAPV immunity was introduced by feeding double-stranded viral RNA fragments to induce vsiRNA production. Feeding double-stranded DWV RNA fragments to honeybees has also been shown to reduce virus load, wing deformity and increase longevity when compared with mock treated controls [26].

In addition to honeybees, bumblebees (e.g. *Bombus pascuorum*) are important wild pollinators [61] and cross-species virus infections have been reported for honeybees and wild pollinators including bumblebees, reflecting a scenario that viruses may be transmitted from wild pollinators to hive bees [61,62]. To investigate potential virus infections, we sequenced small RNA (sRNA) from extracts from *B. pascuorum* and *A. mellifera* collected from the field in Oxfordshire, UK. Small RNA deep sequencing was deployed to detect virus prevalence followed by validations using RT-PCR and Sanger sequencing. We also conducted a within-population analysis of single nucleotide polymorphism (SNP) using the deep sequencing dataset and analysed conserved short fragments (CSF) in the virus genomes to identify possible targets for RNAi development against infections of the VDV-1/DWV group of viruses that associate with collapse of honeybee colonies worldwide.

## Results

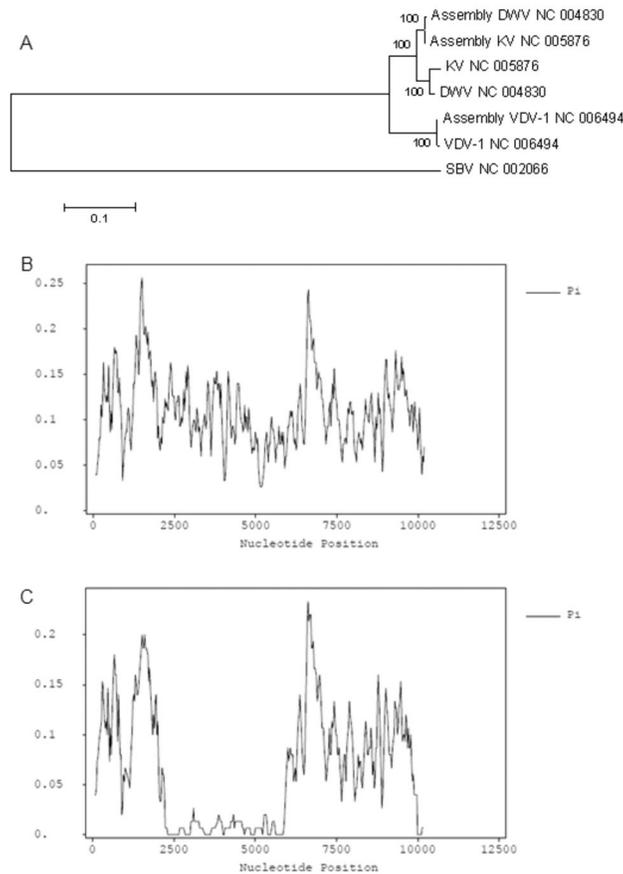
From the pooled honeybee and bumblebee sRNA extracts, approximately 95 million (M) high quality sRNA reads were obtained in total. There were 1.08 M unique reads ranging between 18-44 nt with the dominating species being 22-nt long Dicer products (Figure S1). After removal of ncRNAs (Table S1

and Figure S2), 3.77 M of total reads remained in the library containing 0.93 M unique reads (Figure S1). These unique reads were assembled firstly by the SOAPdenovo program and the resulting contigs were subjected to BLAST searches against virus reference genomes in the NCBI. Matches with three viruses (Text S1) were obtained with high scores. These were the Deformed wing virus (DWV), *Varroa destructor virus-1* (VDV1) and *Kakugo virus* (KV). A later result showed that these viral signals were only associated with the honeybee extracts. Viral genomes (NCBI, DWV, NC\_004830; VDV-1, NC\_006494; KV, NC\_005876) were used as the references to enable MAQ assembly of the virus genome. The assembled consensus genomes (NCBI accession numbers: VDV-1\_Ox, KC786222; DWV\_Ox, KC786223; KVlike\_Ox, KC786224) displayed close but distinguishable phylogenetic relationships (Figure 1), suggesting mixed infections in the sample. When sRNA reads were mapped against the reference sequences, 31,192 unique reads could be perfectly matched to either the VDV-1, WDV or KV genomes (Figure 2, Table S2 and Figure S3). These vsiRNAs were also dominated by 22-nt long species (Figure S1), distributed throughout the genomes with hotspots, and originated from both plus and minus RNA strands (Figure 2).

The vsiRNAs could be mapped through-out the whole VDV-1\_Ox genome (Figure 2) but represented a gap against the DWV\_Ox (2168-5853 nt) and KVlike\_Ox (2164-5865 nt) genomes (Figure 2). This gap was located between the VP2 encoding region and helicase encoding region, and did not appear to be due to sequence similarities among the VDV-1, DWV and KV genomes (Figure 2). Sequence conservations among the three viruses resulted in a few vsiRNA islands being shared by these viruses (Figure 2). SNP analysis using the virus reference genomes (Figure 1) and the MAQ assembled genomes confirmed this conclusion (Figure 1). The absence of DWV/KV sequences in the middle fractions suggested that the VP2-Helicase fragment of DWV\_Ox/KVlike\_Ox was replaced by the sequences of VDV-1\_Ox.

To confirm that the viral population was a mixture of a VDV-1 virus and a 5' DWV-VDV1-DWV3' chimera, we amplified the putative recombination junction regions from the total RNA extract by using RT-PCR. The honeybee RNA extracts produced single band products at both nt 1874-3001 (VDV-1) and nt 5386-6252 (VDV-1) regions, whereas RNA extracts of the bumblebee were negative (Figure S4), suggesting that the local infection might be limited in honeybees although DWV and VDV-1 infections have been reported in bumblebees [28,32,62,69]. Negative bumblebee samples also demonstrated that there was no viral contamination during sample processing. The RT-PCR products were cloned and then sequenced using the Sanger method. Phylogenetic trees of sequences in these two fragments clearly showed that there were mixed infections (Figure 3). SNP analyses using the Sanger sequences confirmed the sRNA assembly and mapping result (Figure 1, Figure 2) that sequences in the nt 2100-5800 nt region (Figure 2) contained much less polymorphism than the 5'- and 3'-fragments (Figure 3, Text S2, Text S3), the presence of chimera virus and that the sRNA

Figure 1



**Figure 1. Comparison of assembled virus genomes.** (A) A Neighbour Joint tree constructed with 1000 bootstraps. NCBI reference sequences were labelled with virus names followed by the GenBank accession numbers. The Maq assembled sequences were labelled accordingly to the reference sequences used. The *Sacbrood virus* (SBV) reference genome was used as an outer sequence. (B) SNP profile using the 3 viral reference sequences. (C) SNP profile using the 3 Maq assembled virus genomes.

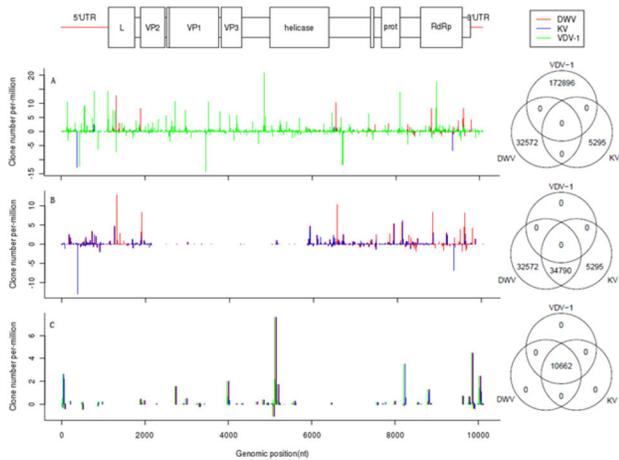
doi: 10.1371/journal.pone.0074508.g001

deep sequencing result was genuine rather than an artefact due to unexpected sequencing bias.

Deep sequencing technology provides new opportunities to analyse mutation profiles at each nucleotide position of a virus genome. Here we focused on the conserved positions. Reads were mapped against the reference sequences with 2-nt mismatches allowed. A mutation profile of the infecting virus population was obtained (Data not shown). To further evaluate the observation, we downloaded all *Flavivirus* sequences from the GenBank (by Dec 2012, Figure S5). These GenBank reference sequences were aligned and conserved regions were screened by using sliding windows (Window size 25-nt) in Program DnaSP [67]. Thirty conserved short fragments (CSF) were observed (Figure 4 and Table S3) from the sequences of the VDV-1/DWV/KV clade (Figure S5). The largest CSF was 61-nt long and located at the 3'-end of the virus genome (Figure 4). Using the sRNA library, mutation rates ( $\Pi$ ) at each nucleotide positions were calculated and they were not

correlated to sequencing coverage in these CSF (Figure S6), indicating that the SNPs observed in the deep sequencing dataset were not largely due to sequencing error which could be expected at a rate of 0.02% with the Illumina Solexa technology. Although these sites appeared conserved at the between-species level, significant SNP (e.g.,  $\Pi > 10\%$ ) was observed for some of the positions at the within-population level (Figure 4, Figure S6, Table S3). However, the majorities of the CSF sites did display low (e.g.  $\Pi < 5\%$ ) or nil SNP (Figure 4). The observed high mutation rate (Figure 4) is indicative that nucleotide constraint may not be initiated at the within-population level. There were eight CSFs having a consistent 20-nt long region (Figure 4) in which the observed mutation rates were lower than 5%.

Figure 2



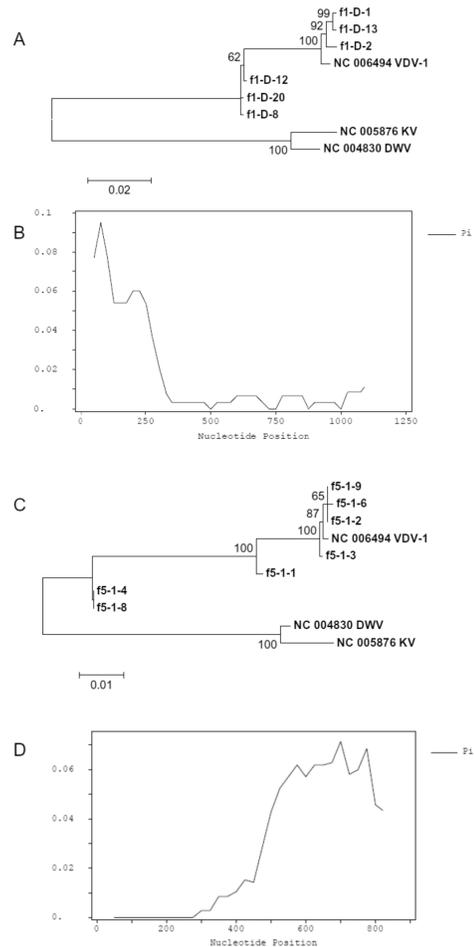
**Figure 2. Virus derived small interfering (vsiRNA) mapping profiles.** (A) Unique (non-redundant) vsiRNAs mapped to either VDV-1 or DWV or KV genomes. Shared reads were excluded. (B) Positions of vsiRNAs mapped to DWV and KV genomes. Shared reads with VDV-1 were excluded. (C) Shared vsiRNAs mapped to all of the three virus genomes. Note that the Y-axis represents both plus and minus strands and the scale is different among the three panels.

doi: 10.1371/journal.pone.0074508.g002

**Discussion**

RNAi is an ancient anti-viral mechanism. The vsiRNAs generated by host RNase III enzymes are integrated with host Argonout (AGO) enzymes and lead the RISC (RNA-Induced Silencing Complex) to target viral single-stranded RNAs based on complementary sequence homology. In turn, viruses have evolved suppressor proteins against host RNAi (reviewed by [70]). On the other hand, due to error prone RNA dependent RNA polymerase, RNA viruses have high mutation rates and their populations display quasi-species characteristics. It is well known that animal viruses may mutate to escape host antibody derived immunities (reviewed by [71]). Regarding host RNAi immunity (reviewed by [47,48]), viruses may also evolve to change their genome composition to reduce host RNAi response during infections [72-74]. Single nucleotide mutations were not considered as a feasible evolutionary strategy for viruses to combat the natural host RNAi because RNAi targets the whole of the virus genome simultaneously (e.g. Figure 2), making it impossible for the virus to deploy sufficient site mutations to escape. Human introduced anti-virus RNAi immunity, however, often targets a small region in the virus genome, allowing the virus to accumulate sufficient mutations in the targeted region [75,76]. It therefore becomes necessary to develop strategies to minimise the chances of virus escape by mutation. A logical approach is to target conserved viral sequences in RNAi development. This is because conserved sequences are likely to have important functions constrained

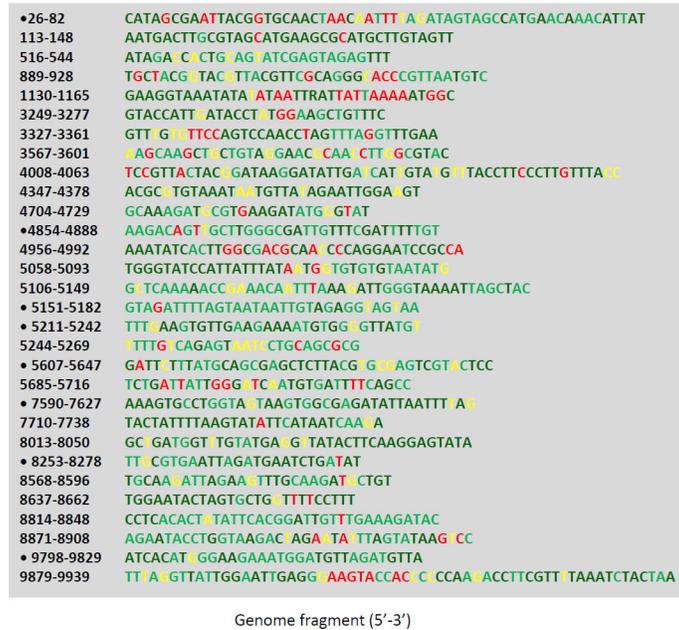
Fig. 3



**Figure 3. Sanger sequencing validation of chimera virus.** (A) Unrooted Neighbour Joint tree of the 5' recombination junction using the sequences of RT-PCR products amplified by primers F1R1. The tree was constructed with 1000 bootstraps and scores higher than 50% were displayed. Reference sequences are labelled with the NCBI accession numbers followed by virus names. The Sanger sequences obtained in this work (NCBI Accession No: KC691296-KC691301) are labelled as F1 followed by in-house sequence IDs. (B) SNP profile of the 5' recombination junction using the Sanger sequences, showing a decrease of polymorphism after the recombination junction. (C) Unrooted Neighbour Joint tree of the 3' recombination junction using the sequences of RT-PCR products amplified by primers F5R5. The tree was constructed with 1000 bootstraps and scores higher than 50% are displayed. Reference sequences are labelled with the NCBI accession numbers followed by virus names. The Sanger sequences obtained in this work (GenBank Accession No: KC691302-KC691308) are labelled as F5 followed by in-house sequence IDs. (D) SNP profile of the 3' recombination junction using the Sanger sequences, showing that nucleotide SNPs increased after the recombination junction.

doi: 10.1371/journal.pone.0074508.g003

Fig. 4



**Figure 4. Conserved short fragments (CSF) of the virus population.** Between-species conserved regions (consecutively larger than 25 nt, without any SNP) were identified by using sequence alignment of the VDV-1/DWV/KV clade (Figure S5). Within-population SNP was calculated by using the sRNA deep sequencing dataset. Each CSF was labelled using DWV genome positions. Colour coded positions represented within-population mutation rate (Pi) as Pi = 0 (dark green), Pi < 5% (light green), 5% ≤ Pi < 10% (yellow) and Pi ≥ 10% (red). The mutation rate at each nucleotide position was reported in Table S3. Black dots label CSFs with 20 consecutive low Pi sites at the within-population level.

doi: 10.1371/journal.pone.0074508.g004

by strong negative evolutionary selection. We have identified eight CSF that were relatively constrained at both within-species and between-species levels (Figure 4). These are likely to be functionally important motifs for virus survival and it would be unlikely that the virus could generate escape mutations if they were targeted. To obtain a comprehensive CSF profile of the VDV-1/DWV/KV clade viruses, spatially distinct virus populations need to be investigated.

Flaviviruses have diverse population structures that can shift dynamically according to selection pressure and environmental changes (e.g., [6]). Between-species genome recombination and chimera virus infections have been reported for the DWV/VDV-1 group [10,11]. By comparing sequences deposited in public databases, the research community builds up a capacity to determine conserved regions of virus genomes. Although codon constraint [77] and RNA folding structure [78] have been well considered, knowledge about function and maintenance of conserved nucleotide positions in virus genomes is still limited. It may be logical to assume that such sequence conservation is maintained at the very basic level, i.e., within-population level. By deep sequencing the sRNA population in natural conditions, here we provide clear evidence that observed between-species sequence conservations were not always maintained at the within-population level (Figure 4). Errors generated during library construction and sequencing cannot be completely ruled

out. However, it is unlikely to be wholly due to the possible PCR and Solexa sequencing errors that both occur at about 10<sup>-5</sup> per nucleotide [79,80], which is incomparable with the observed mutation rates (e.g. Pi >10%) shown in Figure 4. This surprising observation on high mutation rates at some CSF sites (Figure 4) introduces a new task for RNAi development: to prevent viral escape via population structure shift, i.e., RNAi needs to be designed to target true CSFs rather than those only displayed at higher (between-isolate/between-species) levels. Highly polymorphic sites at the within-population level should be avoided in RNAi design. Otherwise, existing polymorphism in the virus population can mediate rapid escape from designed RNAi [75,81].

Ample evidence shows that VDV-1/DWV/KV clade viruses contribute to the collapse of honeybee colonies in the northern hemisphere (reviewed by [1,2]). It is necessary to develop treatment and control measures against infections by these viruses. Following the trials against the *Israeli Acute Paralysis* (IAPV) in *A. mellifera* [59,60], immunisation by feeding with double-stranded viral RNA has been shown to be an effective approach to reduce DWV load and disease symptoms in honeybees [26]. Desai et al. [26] targeted a 700 nt long fragment (8565-9355 nt, AY292384) at the DWV 3'-end fraction, which contained two CSFs described in Figure 4. Although an effective siRNA profile was not provided to clarify

the mechanism(s), RNAi was the primary candidate that mediates the ds-RNA induced immunity [26]. The vsiRNA map shown in Figure 2 supports this speculation, i.e., the vsiRNA against VDV-1/DWV/KV originated from both plus and minus strands and was mainly Dicer products of 21-23 nt long, similarly as characterised for the IAPV RNAi [59,60]. New knowledge on vsiRNA and CSF profiles provided in this study will help to develop more efficient RNAi protocols to combat the virus infections.

DWV and VDV-1 can infect bumblebees and other pollinators [28,32,62,69]. Due to the relatively small sample size used in this work, lack of detection of DWV/VDV-1 in *B. pascuorum* cannot be used to conclude that the local infection may be limited to *A. mellifera*. A full picture of the virus population would only be revealed when all major host and vector species are tested for virus. This work in *A. mellifera* demonstrates that virus evolution events like genome recombination are ongoing as well as nucleotide conservation. Virus evolution may lead to emergence of unexpected diseases whereas conservation provides possible targets for developing treatments and control measures.

## Materials and Methods

### Ethics Statement

No permits were required for the described study, which complied with all relevant regulations. The field site is a public place and the study did not involve endangered or protected species.

### Sampling, RNA extraction, and small RNA sequencing

Foraging wild female worker honeybee (*Apis mellifera*) and bumblebee (*Bombus pascuorum*) individuals were net caught in Sept 2010 whilst foraging on a single large patch of *Succisa pratensis* in Bernwood Forest, Oxfordshire, UK. The samples were stored at -80°C until use. The thoraxes of 10 individual *A. mellifera* and *B. pascuorum* were pooled, respectively, and extracted for total RNA (>200 nt) and small RNA (<200 nt) fractions using the mirVana miRNA isolation kit (Cat AM1560, Ambion, Life Technologies Ltd, Paisley, UK) following the manufacturer's protocols. Concentrations of the sRNA extracts were measured by using the Nanodrop 1000 UV-Vis spectrophotometer (Nanodrop Products, Wilmington, USA). The two sRNA extracts were quantitatively (1:1) pooled and despatched for small RNA sequencing (at BGI, Hong Kong, China). The deep sequencing dataset has been deposited in the NIH Short Read Archive: accession number SRA069239

### Bioinformatic analyses

Raw sequence reads were subjected to a standard Illumina Solexa quality control pipeline, which also removed the adaptor sequences. The high quality sRNA reads were filtered against the non-coding RNA database (<http://biobases.ibch.poznan.pl/ncRNA/>). After removing the ncRNAs, the sRNA library was assembled using the SOAPdenovo2 program (<http://soap.genomics.org.cn/soapdenovo.html>) [63]. Resulting contigs were used to search against the nt/nr database for

virus hits using the BLAST program [64]. The sRNA library was then mapped to the NCBI reference sequences of the VDV-1, DWV and KV genomes. After identification of viruses present in the sample, the sRNA library was also used to assemble the virus genomes with the NCBI virus reference sequences (DWV, NC\_004830; VDV-1, NC\_006494; KV, NC\_005876) with the Maq program (Maq 0.6.6, <http://maq.sourceforge.net/>) [65]. The assembled virus genomes (NCBI accession number: :VDV-1\_Ox, KC786222; DWV\_Ox, KC786223; KVlike\_Ox, KC786224) were aligned with the reference genomes and the other NCBI virus genome sequences to identify conserved regions using ClusterX [66]. Single nucleotide polymorphism (SNP) profiles were identified by using the DnaSP program [67].

### RT-PCR, Cloning and Sanger sequencing

Consensus sequences shared among the assembled genomes were used to design experiments to amplify recombination junction regions. DNA primers were made as F1: 5'-tggagtagatggactagtaatgatgt-3', R1: 5'-aggaacataacctacaattaaccta-3' (synthesized by Eurofins MWG/Operon, Ebersberg, Germany) for amplifying the 5' recombination junction located at VDV-1 genome positions of nt 1874-3001. A second set of primers was made as F5: 5'-tgactaacggctgaatgaatca-3' and R5: 5'-ttcatttctcctaagcgctgatt-3' for covering the 3' recombination site at the VDV-1 genome position of nt 5386-6252. The total RNA extracts of honeybee and bumblebee were used as templates for RT-PCR reactions by using the Qiagen one step RT-PCR kit (Cat 210210, Qiagen, Crawley, UK). The reactions were performed on an Omn-E Thermal Cycler (Hybaid Ltd, Cambridge, UK) with a program of 50 °C 30 min, 95°C 2 min; 30 cycles of 95°C 1 min, 55°C 1 min, 72°C 1 min; and 72°C 10 min. The DNA products were examined using 1% Agarose gel electrophoresis, then cloned into the pGEM-T vector plasmid (Cat A1360, Promega, Southampton, UK) and sequenced from both directions using the Big Dye Terminator kit (Applied Biosystems by Life Technologies Corporation, Carlsbad, USA) with the T7 and SP6 primers. Resulting sequences (NCBI Accession No: KC691296-KC691301) were used to construct phylogenetic trees using the Mega5 program [68].

## Supporting Information

**Figure S1. Reads length distribution.**  
(PDF)

**Figure S2. Percentage of reads matched to non-coding RNAs.**  
(PDF)

**Figure S3. Count of reads aligned to DWV, VDV-1 and KV.**  
(PDF)

**Figure S4. RT-PCR amplification of viral fragments.**  
(PDF)

**Figure S5. Neighbour Joint consensus tree of Iflavirus genomes.**

(PDF)

**Figure S6. Site mutations in conserved short fragments (CSF) among DWV, VDV-1, and KV.**

(PDF)

**Table S1. Non coding RNA classification.**

(PDF)

**Table S2. Number and percentage of reads mapped to the three reference sequences.**

(PDF)

**Table S3. Mutation rate (%) of nucleotide positions in conserved short fragments.**

(XLSX)

**Text S1. Contigs matched to genome sequences of DWV, VDV1 and KV.**

(TXT)

**Text S2. CLUSTAL X (1.8) multiple sequence alignment of the 5'-end of the recombination junction (F1\_6Seq).**

(TXT)

**Text S3. CLUSTAL X (1.8) multiple sequence alignment of the 3'-end of the recombination junction (F5\_7Seq).**

(TXT)

**Acknowledgements**

We thank Andrew J. Lack for discussions and helping with sampling.

**Author Contributions**

Conceived and designed the experiments: HW TGS LAK RDP. Performed the experiments: HW TGS DWP. Analyzed the data: HW JX JM. Contributed reagents/materials/analysis tools: HW JM TGS. Wrote the manuscript: HW TGS LAK RDP.

**References**

1. Genersch E, Aubert M (2010) Emerging and re-emerging viruses of the honey bee (*Apis mellifera* L.). *Vet Res* 41: 54. doi:10.1051/vetres/2010027. PubMed: 20423694.
2. Evans JD, Schwarz RS (2011) Bees brought to their knees: microbes affecting honey bee health. *Trends Microbiol* 19: 614-620. doi:10.1016/j.tim.2011.09.003. PubMed: 22032828.
3. Yang X, Cox-Foster D (2007) Effects of parasitization by *Varroa destructor* on survivorship and physiological traits of *Apis mellifera* in correlation with viral incidence and microbial challenge. *Parasitology* 134: 405-412. doi:10.1017/S0031182006000710. PubMed: 17078903.
4. Yang X, Cox-Foster DL (2005) Impact of an ectoparasite on the immunity and pathology of an invertebrate: evidence for host immunosuppression and viral amplification. *Proc Natl Acad Sci U S A* 102: 7470-7475. doi:10.1073/pnas.0501860102. PubMed: 15897457.
5. Nazzi F, Brown SP, Annoscia D, Del Piccolo F, Di Prisco G et al. (2012) Synergistic parasite-pathogen interactions mediated by host immunity can drive the collapse of honeybee colonies. *PLOS Pathog* 8: e1002735. PubMed: 22719246.
6. Martin SJ, Highfield AC, Brettell L, Villalobos EM, Budge GE et al. (2012) Global honey bee viral landscape altered by a parasitic mite. *Science* 336: 1304-1306. doi:10.1126/science.1220941. PubMed: 22679096.
7. Schroeder DC, Martin SJ (2012) Deformed wing virus: The main suspect in unexplained honeybee deaths worldwide. *Virulence* 3: 589-591. doi:10.4161/viru.22219. PubMed: 23154287.
8. Ongus JR, Peters D, Bonmatin JM, Bengsch E, Vlaskovic JM et al. (2004) Complete sequence of a picorna-like virus of the genus *Iflavirus* replicating in the mite *Varroa destructor*. *J Gen Virol* 85: 3747-3755. doi:10.1099/vir.0.80470-0. PubMed: 15557248.
9. Ongus JR, Roode EC, Pleij CW, Vlaskovic JM, van Oers MM (2006) The 5' non-translated region of *Varroa destructor* virus 1 (genus *Iflavirus*): structure prediction and IRES activity in *Lymantria dispar* cells. *J Gen Virol* 87: 3397-3407. doi:10.1099/vir.0.82122-0. PubMed: 17030876.
10. Zioni N, Soroker V, Chejanovsky N (2011) Replication of *Varroa destructor* virus 1 (VDV-1) and a *Varroa destructor* virus 1-deformed wing virus recombinant (VDV-1-DWV) in the head of the honey bee. *Virology* 417: 106-112. doi:10.1016/j.virol.2011.05.009. PubMed: 21652054.
11. Moore J, Jironkin A, Chandler D, Burroughs N, Evans DJ et al. (2011) Recombinants between Deformed wing virus and *Varroa destructor* virus-1 may prevail in *Varroa destructor*-infested honeybee colonies. *J Gen Virol* 92: 156-161. doi:10.1099/vir.0.025965-0. PubMed: 20926636.
12. Ai H, Yan X, Han R (2012) Occurrence and prevalence of seven bee viruses in *Apis mellifera* and *Apis cerana* apiaries in China. *J Invertebr Pathol* 109: 160-164. doi:10.1016/j.jip.2011.10.006. PubMed: 22062807.
13. Antúnez K, Anido M, Garrido-Bailón E, Botías C, Zunino P et al. (2012) Low prevalence of honeybee viruses in Spain during 2006 and 2007. *Res Vet Sci* 93: 1441-1445. doi:10.1016/j.rvsc.2012.03.006. PubMed: 22513127.
14. Bacandritsos N, Granato A, Budge G, Papanastasiou I, Roinioti E et al. (2010) Sudden deaths and colony population decline in Greek honey bee colonies. *J Invertebr Pathol* 105: 335-340. doi:10.1016/j.jip.2010.08.004. PubMed: 20804765.
15. Barriga GP, Cifuentes-Muñoz N, Rivera PA, Gutierrez M, Shmaryahu A et al. (2012) First detection and complete genome sequence of Deformed wing virus in Chilean honeybees. *Virus Genes* 45: 606-609. doi:10.1007/s11262-012-0791-0. PubMed: 22836560.
16. Boncristiani HF Jr., Di Prisco G, Pettis JS, Hamilton M, Chen YP (2009) Molecular approaches to the analysis of deformed wing virus replication and pathogenesis in the honey bee, *Apis mellifera*. *Virology* 393: 221. doi:10.1016/j.virol.2009.06.021. PubMed: 20003360.
17. Choe SE, Nguyen LT, Noh JH, Koh HB, Jean YH et al. (2012) Prevalence and distribution of six bee viruses in Korean *Apis cerana* populations. *J Invertebr Pathol* 109: 330-333. doi:10.1016/j.jip.2012.01.003. PubMed: 22273697.
18. Chen Y, Zhao Y, Hammond J, Hsu HT, Evans J et al. (2004) Multiple virus infections in the honey bee and genome divergence of honey bee viruses. *J Invertebr Pathol* 87: 84-93. doi:10.1016/j.jip.2004.07.005. PubMed: 15579317.
19. Chen YP, Higgins JA, Feldlaufer MF (2005) Quantitative real-time reverse transcription-PCR analysis of deformed wing virus infection in the honeybee (*Apis mellifera* L.). *Appl Environ Microbiol* 71: 436-441. doi:10.1128/AEM.71.1.436-441.2005. PubMed: 15640219.
20. Chen YP, Pettis JS, Collins A, Feldlaufer MF (2006) Prevalence and transmission of honeybee viruses. *Appl Environ Microbiol* 72: 606-611. doi:10.1128/AEM.72.1.606-611.2006. PubMed: 16391097.
21. Costa C, Tanner G, Lodesani M, Maistrello L, Neumann P (2011) Negative correlation between *Nosema ceranae* spore loads and deformed wing virus infection levels in adult honey bee workers. *J Invertebr Pathol* 108: 224-225. doi:10.1016/j.jip.2011.08.012. PubMed: 21939664.
22. Dainat B, Evans JD, Chen YP, Gauthier L, Neumann P (2012) Dead or alive: deformed wing virus and *Varroa destructor* reduce the life span of winter honeybees. *Appl Environ Microbiol* 78: 981-987. doi:10.1128/AEM.06537-11. PubMed: 22179240.
23. Dainat B, Evans JD, Chen YP, Neumann P (2011) Sampling and RNA quality for diagnosis of honey bee viruses using quantitative PCR. *J*

- Virol Methods 174: 150-152. doi:10.1016/j.jviromet.2011.03.029. PubMed: 21473885.
24. Dainat B, Neumann P (2012) Clinical signs of deformed wing virus infection are predictive markers for honey bee colony losses. *J Invertebr Pathol* 112: 278-280. PubMed: 23270875.
  25. de Miranda JR, Genersch E (2010) Deformed wing virus. *J Invertebr Pathol* 103 Suppl 1: S48-S61. doi:10.1016/j.jip.2009.06.012. PubMed: 19909976.
  26. Desai SD, Eu YJ, Whyard S, Currie RW (2012) Reduction in deformed wing virus infection in larval and adult honey bees (*Apis mellifera* L.) by double-stranded RNA ingestion. *Insect Mol Biol* 21: 446-455. doi: 10.1111/j.1365-2583.2012.01150.x. PubMed: 22690671.
  27. Francis RM, Nielsen SL, Kryger P (2013) Patterns of viral infection in honey bee queens. *J Gen Virol* 94: 668-676. doi:10.1099/vir.0.047019-0. PubMed: 23223622.
  28. Genersch E, Yue C, Fries I, de Miranda JR (2006) Detection of Deformed wing virus, a honey bee viral pathogen, in bumble bees (*Bombus terrestris* and *Bombus pascuorum*) with wing deformities. *J Invertebr Pathol* 91: 61-63. doi:10.1016/j.jip.2005.10.002. PubMed: 16300785.
  29. Gisder S, Aumeier P, Genersch E (2009) Deformed wing virus: replication and viral load in mites (*Varroa destructor*). *J Gen Virol* 90: 463-467. doi:10.1099/vir.0.005579-0. PubMed: 19141457.
  30. Highfield AC, El Nagar A, Mackinder LC, Noël LM, Hall MJ et al. (2009) Deformed wing virus implicated in overwintering honeybee colony losses. *Appl Environ Microbiol* 75: 7212-7220. doi:10.1128/AEM.02227-09. PubMed: 19783750.
  31. Kojima Y, Toki T, Morimoto T, Yoshiyama M, Kimura K et al. (2011) Infestation of Japanese native honey bees by tracheal mite and virus from non-native European honey bees in Japan. *Microb Ecol* 62: 895-906. doi:10.1007/s00248-011-9947-z. PubMed: 21960435.
  32. Li J, Peng W, Wu J, Strange JP, Boncristiani H et al. (2011) Cross-species infection of deformed wing virus poses a new threat to pollinator conservation. *J Econ Entomol* 104: 732-739. doi:10.1603/EC10355. PubMed: 21735887.
  33. Li J, Qin H, Wu J, Sadd BM, Wang X et al. (2012) The prevalence of parasites and pathogens in Asian honeybees *Apis cerana* in China. *PLOS ONE* 7: e47955. doi:10.1371/journal.pone.0047955. PubMed: 23144838.
  34. Möckel N, Gisder S, Genersch E (2011) Horizontal transmission of deformed wing virus: pathological consequences in adult bees (*Apis mellifera*) depend on the transmission route. *J Gen Virol* 92: 370-377. doi:10.1099/vir.0.025940-0. PubMed: 20965988.
  35. Prisco GD, Zhang X, Pennacchio F, Caprio E, Li J et al. (2011) Dynamics of persistent and acute deformed wing virus infections in honey bees, *Apis mellifera*. *Viruses* 3: 2425-2441. doi:10.3390/v3122425. PubMed: 22355447.
  36. Rortais A, Tentcheva D, Papachristoforou A, Gauthier L, Arnold G et al. (2006) Deformed wing virus is not related to honey bees' aggressiveness. *Virol J* 3: 61. doi:10.1186/1743-422X-3-61. PubMed: 16942620.
  37. Sanpa S, Chantawannakul P (2009) Survey of six bee viruses using RT-PCR in Northern Thailand. *J Invertebr Pathol* 100: 116-119. doi: 10.1016/j.jip.2008.11.010. PubMed: 19105966.
  38. Santillán-Galicia MT, Carzaniga R, Ball BV, Alderson PG (2008) Immunolocalization of deformed wing virus particles within the mite *Varroa destructor*. *J Gen Virol* 89: 1685-1689. doi:10.1099/vir.0.83223-0. PubMed: 18559939.
  39. Terio V, Martella V, Camero M, Decaro N, Testini G et al. (2008) Detection of a honeybee flavivirus with intermediate characteristics between kakugo virus and deformed wing virus. *New Microbiol* 31: 439-444. PubMed: 19123297.
  40. Welch A, Drummond F, Tewari S, Averill A, Burand JP (2009) Presence and prevalence of viruses in local and migratory honeybees (*Apis mellifera*) in Massachusetts. *Appl Environ Microbiol* 75: 7862-7865. doi:10.1128/AEM.01319-09. PubMed: 19854916.
  41. Williams GR, Rogers RE, Kalkstein AL, Taylor BA, Shuttler D et al. (2009) Deformed wing virus in western honey bees (*Apis mellifera*) from Atlantic Canada and the first description of an overtly-infected emerging queen. *J Invertebr Pathol* 101: 77-79. doi:10.1016/j.jip.2009.01.004. PubMed: 19373971.
  42. Zhang X, He SY, Evans JD, Pettis JS, Yin GF et al. (2012) New evidence that deformed wing virus and black queen cell virus are multi-host pathogens. *J Invertebr Pathol* 109: 156-159. doi:10.1016/j.jip.2011.09.010. PubMed: 22001629.
  43. Lanzi G, de Miranda JR, Boniotti MB, Cameron CE, Lavazza A et al. (2006) Molecular and biological characterization of deformed wing virus of honeybees (*Apis mellifera* L.). *J Virol* 80: 4998-5009. doi:10.1128/JVI.80.10.4998-5009.2006. PubMed: 16641291.
  44. Fujiyuki T, Matsuzaka E, Nakaoka T, Takeuchi H, Wakamoto A et al. (2009) Distribution of Kakugo virus and its effects on the gene expression profile in the brain of the worker honeybee *Apis mellifera* L. *J Virol* 83: 11560-11568.
  45. Fujiyuki T, Ohka S, Takeuchi H, Ono M, Nomoto A et al. (2006) Prevalence and phylogeny of Kakugo virus, a novel insect picorna-like virus that infects the honeybee (*Apis mellifera* L.), under various colony conditions. *J Virol* 80: 11528-11538. doi:10.1128/JVI.00754-06. PubMed: 16971448.
  46. Fujiyuki T, Takeuchi H, Ono M, Ohka S, Sasaki T et al. (2004) Novel insect picorna-like virus identified in the brains of aggressive worker honeybees. *J Virol* 78: 1093-1100. doi:10.1128/JVI.78.3.1093-1100.2004. PubMed: 14722264.
  47. Ding SW (2010) RNA-based antiviral immunity. *Nat Rev Immunol* 10: 632-644. doi:10.1038/nri2824. PubMed: 20706278.
  48. Ding SW, Lu R (2011) Virus-derived siRNAs and piRNAs in immunity and pathogenesis. *Curr Opin Virol* 1: 533-544. doi:10.1016/j.coviro.2011.10.028. PubMed: 22180767.
  49. Aliyari R, Ding SW (2009) RNA-based viral immunity initiated by the Dicer family of host immune receptors. *Immunol Rev* 227: 176-188. doi: 10.1111/j.1600-065X.2008.00722.x. PubMed: 19120484.
  50. Kreuze JF, Perez A, Untiveros M, Quispe D, Fuentes S et al. (2009) Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. *Virology* 388: 1-7. doi:10.1016/j.virol.2009.03.024. PubMed: 19394993.
  51. Wu Q, Luo Y, Lu R, Lau N, Lai EC et al. (2010) Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. *Proc Natl Acad Sci U S A* 107: 1606-1611. doi:10.1073/pnas.0911353107. PubMed: 20080648.
  52. Pallett DW, Ho T, Cooper I, Wang H (2010) Detection of Cereal yellow dwarf virus using small interfering RNAs and enhanced infection rate with Cocksfoot streak virus in wild cocksfoot grass (*Dactylis glomerata*). *J Virol Methods* 168: 223-227. doi:10.1016/j.jviromet.2010.06.003. PubMed: 20558208.
  53. Delwart EL (2007) Viral metagenomics. *Rev Med Virol* 17: 115-131. doi:10.1002/rmv.532. PubMed: 17295196.
  54. Tang P, Chiu C (2010) Metagenomics for the discovery of novel human viruses. *Future Microbiol* 5: 177-189. doi:10.2217/fmb.09.120. PubMed: 20143943.
  55. Mokili JL, Rohwer F, Dutilh BE (2012) Metagenomics and future perspectives in virus discovery. *Curr Opin Virol* 2: 63-77. doi:10.1016/j.coviro.2011.12.004. PubMed: 22440968.
  56. Radford AD, Chapman D, Dixon L, Chantrey J, Darby AC et al. (2012) Application of next-generation sequencing technologies in virology. *J Gen Virol* 93: 1853-1868. doi:10.1099/vir.0.043182-0. PubMed: 22647373.
  57. Haasnoot J, Westerhout EM, Berkhout B (2007) RNA interference against viruses: strike and counterstrike. *Nat Biotechnol* 25: 1435-1443. doi:10.1038/nbt1369. PubMed: 18066040.
  58. Huvenne H, Smagghe G (2010) Mechanisms of dsRNA uptake in insects and potential of RNAi for pest control: a review. *J Insect Physiol* 56: 227-235. doi:10.1016/j.jinsphys.2009.10.004. PubMed: 19837076.
  59. Hunter W, Ellis J, Vanengelsdorp D, Hayes J, Westervelt D et al. (2010) Large-scale field application of RNAi technology reducing Israeli acute paralysis virus disease in honey bees (*Apis mellifera*, Hymenoptera: Apidae). *PLOS Pathog* 6: e1001160.
  60. Maori E, Paldi N, Shafir S, Kalev H, Tsur E et al. (2009) IAPV, a bee-affecting virus associated with Colony Collapse Disorder can be silenced by dsRNA ingestion. *Insect Mol Biol* 18: 55-60. doi:10.1111/j.1365-2583.2009.00847.x. PubMed: 19196347.
  61. Evison SE, Roberts KE, Laurenson L, Pietravalle S, Hui J et al. (2012) Pervasiveness of parasites in pollinators. *PLOS ONE* 7: e30641. doi: 10.1371/journal.pone.0030641. PubMed: 22347356.
  62. Meeus I, Smagghe G, Siede R, Jans K, de Graaf DC (2010) Multiplex RT-PCR with broad-range primers and an exogenous internal amplification control for the detection of honeybee viruses in bumblebees. *J Invertebr Pathol* 105: 200-203. doi:10.1016/j.jip.2010.06.012. PubMed: 20600092.
  63. Luo R, Liu B, Xie Y, Li Z, Huang W et al. (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* 1: 18. doi:10.1186/2047-217X-1-18. PubMed: 23587118.
  64. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403-410. doi:10.1016/S0022-2836(05)80360-2. PubMed: 2231712.
  65. Baker A, Schroeder D (2008) Occurrence and genetic analysis of picorna-like viruses infecting worker bees of *Apis mellifera* L. populations in Devon, South West England. *J Invertebr Pathol* 98: 239-242. doi:10.1016/j.jip.2008.02.010. PubMed: 18359042.

66. Higgins DG, Sharp PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* 73: 237-244. doi:10.1016/0378-1119(88)90330-7. PubMed: 3243435.
67. Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496-2497. doi:10.1093/bioinformatics/btg359. PubMed: 14668244.
68. Tamura K, Peterson D, Peterson N, Stecher G, Nei M et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731-2739. doi:10.1093/molbev/msr121. PubMed: 21546353.
69. Singh R, Levitt AL, Rajotte EG, Holmes EC, Ostiguy N et al. (2010) RNA viruses in hymenopteran pollinators: evidence of inter-Taxa virus transmission via pollen and potential impact on non-Apis hymenopteran species. *PLOS ONE* 5: e14357. doi:10.1371/journal.pone.0014357. PubMed: 21203504.
70. Bivalkar-Mehla S, Vakharia J, Mehla R, Abreha M, Kanwar JR et al. (2011) Viral RNA silencing suppressors (RSS): novel strategy of viruses to ablate the host RNA interference (RNAi) defense system. *Virus Res* 155: 1-9. doi:10.1016/j.virusres.2010.10.003. PubMed: 20951748.
71. Corti D, Lanzavecchia A (2013) Broadly Neutralizing Antiviral Antibodies. *Annu Rev Immunol* 31: 705-742. doi:10.1146/annurev-immunol-032712-095916. PubMed: 23330954.
72. Ho T, Wang H, Pallett D, Dalmay T (2007) Evidence for targeting common siRNA hotspots and GC preference by plant Dicer-like proteins. *FEBS Lett* 581: 3267-3272. doi:10.1016/j.febslet.2007.06.022. PubMed: 17597620.
73. Ho T, Rusholme Pilcher RL, Edwards ML, Cooper I, Dalmay T et al. (2008) Evidence for GC preference by monocot Dicer-like proteins. *Biochem Biophys Res Commun* 368: 433-437. doi:10.1016/j.bbrc.2008.01.110. PubMed: 18243131.
74. Ho T, Wang L, Huang L, Li Z, Pallett DW et al. (2010) Nucleotide bias of DCL and AGO in plant anti-virus gene silencing. *Proteins Cell* 1: 847-858. doi:10.1007/s13238-010-0100-4. PubMed: 21203927.
75. Martínez F, Lafforgue G, Morelli MJ, González-Candelas F, Chua NH et al. (2012) Ultradeep sequencing analysis of population dynamics of virus escape mutants in RNAi-mediated resistant plants. *Mol Biol Evol* 29: 3297-3307. doi:10.1093/molbev/mss135. PubMed: 22593223.
76. Berkhout B, Sanders RW (2011) Molecular strategies to design an escape-proof antiviral therapy. *Antiviral Res* 92: 7-14. doi:10.1016/j.antiviral.2011.04.002. PubMed: 21513746.
77. Davis BK (1999) Evolution of the genetic code. *Prog Biophys Mol Biol* 72: 157-243. doi:10.1016/S0079-6107(99)00006-1. PubMed: 10511799.
78. Seetin MG, Mathews DH (2012) RNA structure prediction: an overview of methods. *Methods Mol Biol* 905: 99-122. PubMed: 22736001.
79. Zhou YH, Zhang XP, Ebright RH (1991) Random mutagenesis of gene-sized DNA molecules by use of PCR with Taq DNA polymerase. *Nucleic Acids Res* 19: 6052. doi:10.1093/nar/19.21.6052. PubMed: 1658751.
80. Dohm JC, Lottaz C, Borodina T, Himmelbauer H (2008) Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res* 36: e105. doi:10.1093/nar/gkn425. PubMed: 18660515.
81. van Rij RP, Andino R (2006) The silent treatment: RNAi as a defense against virus infection in mammals. *Trends Biotechnol* 24: 186-193. doi:10.1016/j.tibtech.2006.02.006. PubMed: 16503061.