






The Virome of Healthy Honey Bee Colonies: Ubiquitous Occurrence of Known and New Viruses in Bee Populations

 Dominika Kadlečková,^a Ruth Tachezy,^a  Tomáš Erban,^b Ward Deboutte,^c Jaroslav Nunvár,^a Martina Saláková,^a  Jelle Matthijnsens^d

^aDepartment of Genetics and Microbiology, Faculty of Science BIOCEV, Charles University, Prumyslova, Vestec, Czech Republic

^bCrop Research Institute, Prague, Czech Republic

^cMax Planck Institute for Immunobiology and Epigenetics, Freiburg, Germany

^dKU Leuven—Rega Institute for Medical Research, Laboratory of Viral Metagenomics, Leuven, Belgium

ABSTRACT Honey bees are globally important pollinators threatened by many different pathogens, including viruses. We investigated the virome of honey bees collected at the end of the beekeeping season (August/September) in Czechia, a Central European country. Samples were examined in biological replicates to assess the homogeneity, stability, and composition of the virome inside a single hive. By choice of healthy workers from colonies, where *Varroa destructor* was under control, we could identify ubiquitous bee viruses. Deformed wing virus (DWV) was highly prevalent, even though the bees were healthy, without any noticeable disease signs. The overall virome composition (consisting of honey bee-, plant-, and bacterium-infecting viruses) was driven primarily by the hive and its location. However, honey bee-specific viruses showed an uneven distribution within the same hive. In addition, our results point to an unusual cooccurrence between two rhabdoviruses and reveal the presence of five distinct lineages of Lake Sinai viruses (LSVs) clustering with other LSV strains described globally. Comparison of our results with the virome of Australian honey bees, the last truly *Varroa*- and DWV-free population, showed a strong difference with respect to DWV and a set of diverse members of the *Picornavirales*, of which the latter were absent in our samples. We hypothesize that the occurrence of DWV introduced by *Varroa* strongly affects the virome structure despite the mite being under control.

IMPORTANCE The Western honey bee, *Apis mellifera*, is a vital part of our ecosystem as well as cultural heritage. Annual colony losses endanger beekeeping. In this study, we examined healthy bees from the heart of Central Europe, where honey bee colonies have been commonly affected by varroosis over 5 decades. Our virome analysis showed the presence of ubiquitous viruses in colonies where the mite *Varroa destructor* was under control and no honey bee disease signs were observed. Compared to previous studies, an important part of our study was the analysis of multiple replicates from individual hives. Our overall results indicate that the virome structure (including bee-infecting viruses, plant-infecting viruses, and bacteriophages) is stable within hives; however, the bee-infecting viruses varied largely within interhive replicates, suggesting variation of honey bee viruses within individual bees. Of interest was the striking difference between the viromes of our 39 pools and 9 pools of honey bee viromes previously analyzed in Australia. It could be suggested that *Varroa* not only affects DWV spread in bee colonies but also affects diverse members of the *Picornavirales*, which were strongly decreased in Czech bees compared to the *Varroa*- and DWV-naive Australian bees.

KEYWORDS *Apis mellifera*, *Picornavirales*, rhabdovirus, Lake Sinai virus, metagenomics, viruses

Editor Seth Bordenstein, Vanderbilt University

Copyright © 2022 Kadlečková et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Jelle Matthijnsens, jelle.matthijnsens@kuleuven.be.

The authors declare no conflict of interest.

Received 24 January 2022

Accepted 18 April 2022

Published 9 May 2022

The European honey bee, *Apis mellifera* Linnaeus 1758, is used for the production of honey, propolis, beeswax, venom, pollen, and royal jelly (1). However, the most crucial beneficial feature of honey bees lies in pollination in both agricultural (2, 3) and natural (4) habitats. Annual colony losses jeopardize these benefits provided by honey bees (5). In temperate zones of Europe, the main colony losses occur over winter and are thus referred to as wintering losses (6). Another problem, not yet fully understood, is colony losses in the United States, also known as colony collapse disorder (CCD) (7). This phenomenon is probably due to a combination of several factors, mainly *Varroa destructor* (8), other viral pathogens, and their interaction (7, 9). Conservation of honey bees is difficult, especially in countries where the density of managed colonies is very high, and this is precisely the case in Czechia, in the heart of Central Europe (10).

The global spread of *V. destructor* has had a severe effect on the transmission and virulence of certain honey bee viruses such as deformed wing virus (DWV); DWV variant B (DWV-B), also described as *Varroa destructor virus 1* (11, 12); and viruses belonging to the acute-Kashmir-Israeli complex (13–15). However, without high mite infestations, DWV infections are often benign or asymptomatic. Important from this point of view is the presence of diverse viruses in Australia, where honey bees are free of both *V. destructor* and DWV, although some viruses from the *Picornavirales* order can be found (16). Interestingly, the interaction between viruses and the mite can affect virus strain distribution, as described previously for DWV (17–20). The DWV-A/B strain ratio is affected by the level of mite infestation in a colony (21). Another parasite, *Nosema ceranae*, has been shown to aggravate black queen cell virus (BQCV) infection (22). However, the synergistic effect of *N. ceranae* in combination with different viruses such as DWV was negated (23). Thus, for other known or newly identified viruses, similar or unexpected interactions may exist. Various nonviral pathogens may play an important role in the prevalence and severity of diseases.

Until lately, honey bee virus research focused mainly on 23 described viral species, as reviewed in 2015 (24). In the last few years, more viruses that demonstrably or presumably infect honey bees were discovered due to the increased use of next-generation sequencing (NGS) technologies (16, 25, 26). Moreover, it was recently shown that the honey bee gut virome contains many bacteriophages (27, 28). Previous knowledge was limited to phages from pathogens such as *Paenibacillus larvae* (27). In contrast to the best-characterized bee-infecting viruses, which belong to the *Picornavirales* or other positive-sense single-stranded RNA (+ssRNA) virus groups (e.g., *Flaviridae* and *Dicistroviridae*), some novel viruses belonging to viral families like the *Rhabdoviridae* or *Orthomyxoviridae* have recently been reported in honey bees (29, 30) and the parasite *V. destructor* (30). Most infections with these novel viruses are not yet known to manifest symptomatically but could impact colony health through fitness costs, even though subtle, for the host and/or through interactions with the host and other pathogens/parasites associated with honey bees. Furthermore, the spread of viral infections from honey bees to wild pollinators is also of great concern (31, 32).

In this study, we explored the diversity and composition of the virome in honey bees from healthy colonies from beekeepers breeding various honey bee genetic lines in Czechia. To see the robustness of the virome analyses, we analyzed three biological repeats from each hive. To our knowledge, this is the first such analysis performed on honey bees. We focused on the virome composition (common versus new viruses, plant viruses, and bacteriophages). In addition, we compared our results with those for nine Australian viromes from bees with no exposure to *V. destructor* or DWV.

RESULTS

Composition and similarity of virome samples. NGS of 39 samples (3 replicates of 9 pooled bees from 13 colonies) yielded a total of 398,231,288 reads, with an average of 10 million reads (range, 1,920,148 to 30,170,502; median, 11,650,042) per sample containing 9 bees. The reads were classified as follows: (i) 46.66% eukaryotic, (ii) 30.98% viral,

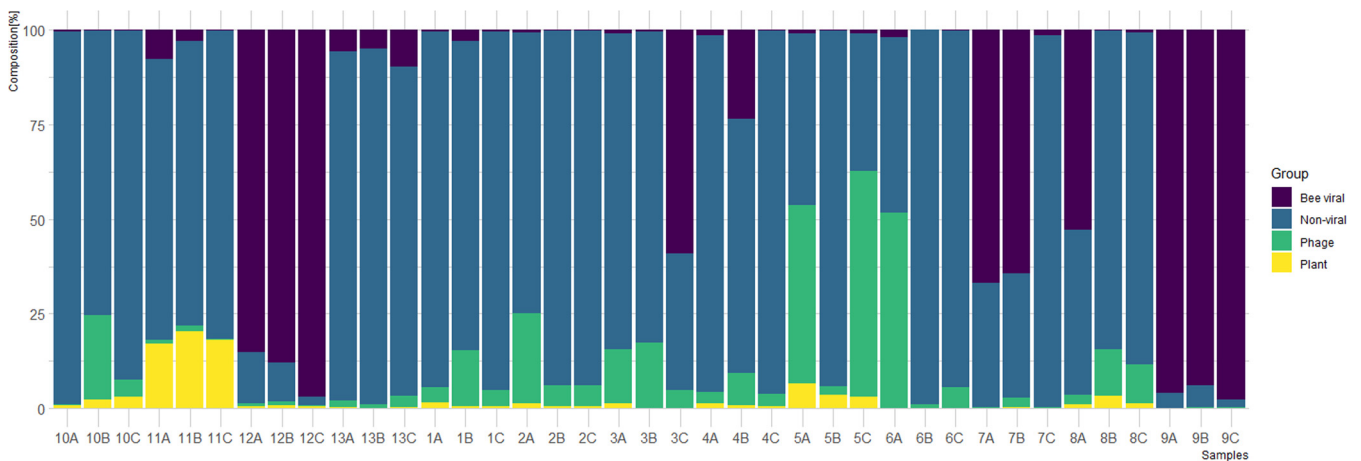


FIG 1 Composition of sequencing reads in all 39 analyzed samples. Each hive was analyzed in 3 independent replicates consisting of nine individual bees. The percentages of reads of different taxonomic affiliations (plant viruses, bacteriophages, bee viruses, and nonviral) are denoted in different colors.

(iii) 13.14% bacterial, and (iv) 9.21% not identified as homologous to any reference sequence.

First, we analyzed the composition of sequencing reads in each of the 39 samples by determining the proportion of reads originating from nonviral sequences (e.g., honey bee genome and bacterial microbiome), known bee viruses (33), bacteriophages, and plant viruses (Fig. 1). Although some replicates seemed to be rather consistent, notable heterogeneity was detected among several other samples (e.g., samples 3, 5, 6, 7, and 8). Next, we analyzed the taxonomic composition of the bee virome with respect to the (i) abundance and (ii) diversity of viral families (see Fig. S1 in the supplemental material).

To visualize the virome similarity of the 39 samples, a heatmap was constructed from the relative abundances of all viral sequences (Fig. 2; Table S1). For almost all samples, replicates originating from the same hive clustered together and thus exhibited similar total viromes. Samples obtained from different sites (hives and apiaries) at the same location also exhibited related viromes. The Adonis test confirmed the highly significant association of virome composition with hive and location ($P < 0.0001$ by an Adonis-Bray test; R^2 , 0.62964 and 0.45389). k-means clustering had for k 13 (representing hives) an adjusted mutual information score (taking on values from 1 for identical to around 0 for random) of 0.14, suggesting the existence of a pattern in clustering (Fig. S1).

Bee viruses. Furthermore, we focused our analysis on eukaryotic viruses that were demonstrated or predicted to infect honey bees (33). Altogether, the analyzed samples revealed the presence of one DNA and nine RNA viruses. Besides the well-known viruses belonging to the *Dicistroviridae* and *Iflaviridae*, we found viruses belonging to the families *Rhabdoviridae* and *Orthomyxoviridae* as well as several variants of Lake Sinai virus (LSV) and the DNA virus *Apis mellifera* filamentous virus (AmFV) (Fig. 3).

Overall distribution of bee viruses. A heatmap was constructed based on the relative abundances of all detected bee-infecting viruses (Fig. 3). The most commonly present viruses were Deformed wing virus variant A (DWV-A) and variant B (DWV-B), Black queen cell virus (BQCV), Aphid lethal paralysis virus (ALPV), AmFV, and Sacbrood virus (SBV). In sharp contrast to the heatmap constructed from all viral sequences (Fig. 2), clustering between bee viruses of most replicate samples was no longer discernible (Fig. 3). We presume that the lack of geographic clustering can be attributed to the absence of bacteriophages and plant viruses (see below) in this analysis. For bee-infecting viruses, the adjusted mutual information score for k 13 (hive) was low (-0.006), further confirming that clustering between samples is absent.

High differences in abundance among the replicate samples were observed for all known bee-infecting viruses, even though each sample consisted of nine pooled bees. This implies that the pooling of nine bees per hive is not sufficient to compensate for the variability in the occurrence of viruses in individual bees. Importantly, in some

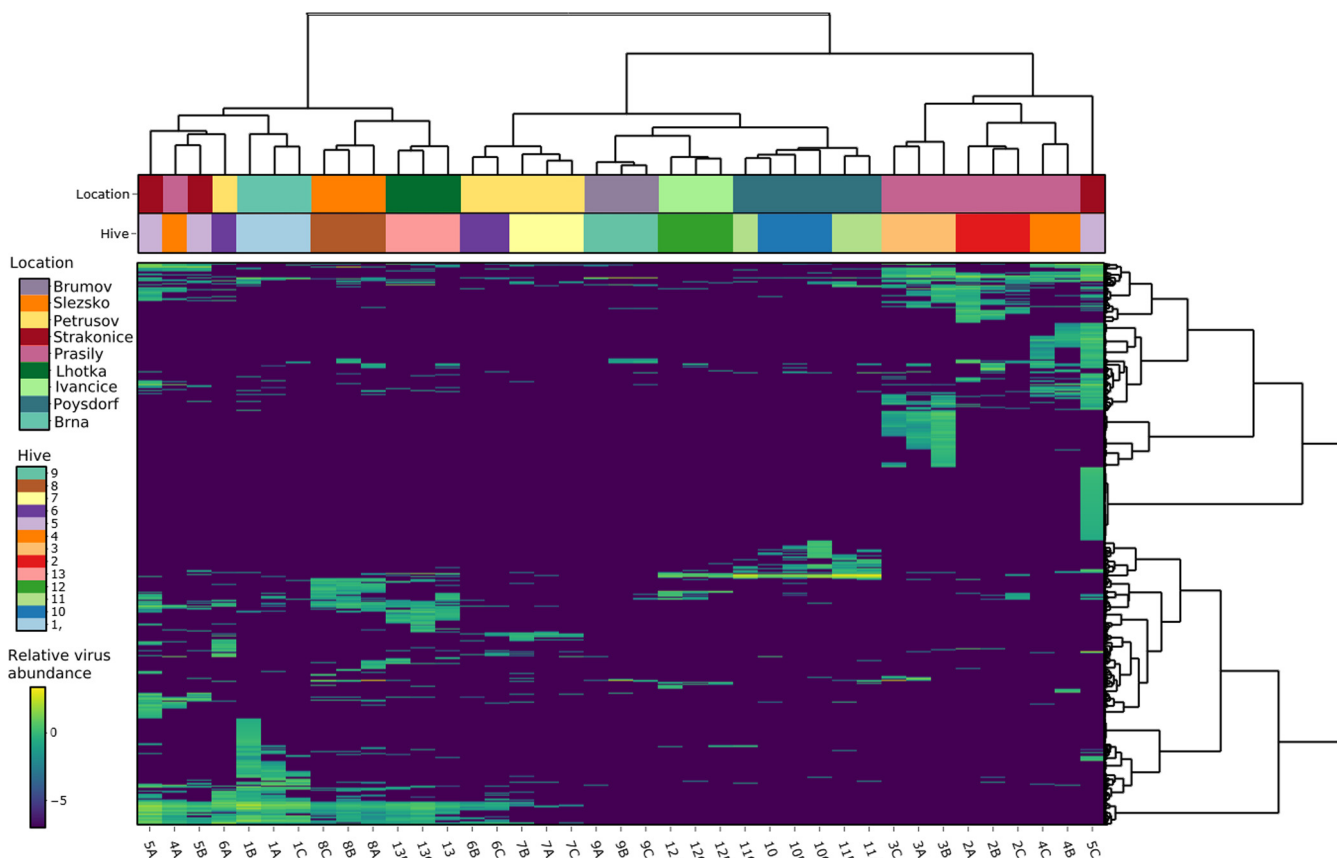


FIG 2 Heatmap constructed from all classified viral sequences in all 39 samples. The viral sequences are taxonomically assigned to the family or species level. Relative abundances (viruses per 1 million sequencing reads) are shown on a log₁₀ scale. Samples (columns) and contigs (rows) are clustered by Ward’s minimum variance method; both columns and rows are seriated by optimal leaf ordering.

samples, a single virus accounted for over 50% of the total sequencing reads (58.8% LSV-A reads in pool 3C and 59.4% acute bee paralysis virus [ABPV] reads in pool 12C) while being present in negligible quantities (<1% reads) in each of the two remaining replicate samples. This indicates the (sporadic) presence of individual bees with very high viral loads compared to those in other bees within the same honey bee colony.

Virome comparison with the *Varroa*- and DWV-naive honey bees (Australia).

Since we used healthy asymptomatic bees where varroosis was under control, we compared our data with those reported previously by Roberts et al. (16). The heatmap in Fig. 4 shows that the virome of *Varroa*- and DWV-naive Australian honey bees is different from that of the Czech samples. All the Australian viromes clustered together and were separated from the Czech viromes. The Australian viromes included several abundant and diverse viruses belonging to the *Picornavirales* (e.g., Perth bee virus, Darwin bee virus, or Robinvale bee virus). In contrast, we did not detect any of the diverse *Picornavirales* (Fig. 4). Conversely, DWV-A/B and ABPV found in Czech honey bees were absent in the nine Australian viromes. Finally, several viruses (BQCV, variants of LSV, SBV, and ALPV) were present in both Australian and Czech viromes.

Prompted by these results, we decided to investigate if this difference was also discernible in other studies of non-Australian viromes. We used public NGS data from Belgium (34); Israel (30); South Africa, The Netherlands, and Tonga (29); and the United States, Central America, Europe, Kenya, India, and New Zealand (26). This allowed us to compare the bee populations where *V. destructor* is present (most of the world) with *Varroa*-naive populations (16). The results showed that the difference observed between the Czech and Australian bee viromes can be generalized to other parts of the world: the above-mentioned diverse *Picornavirales* were absent from the honey

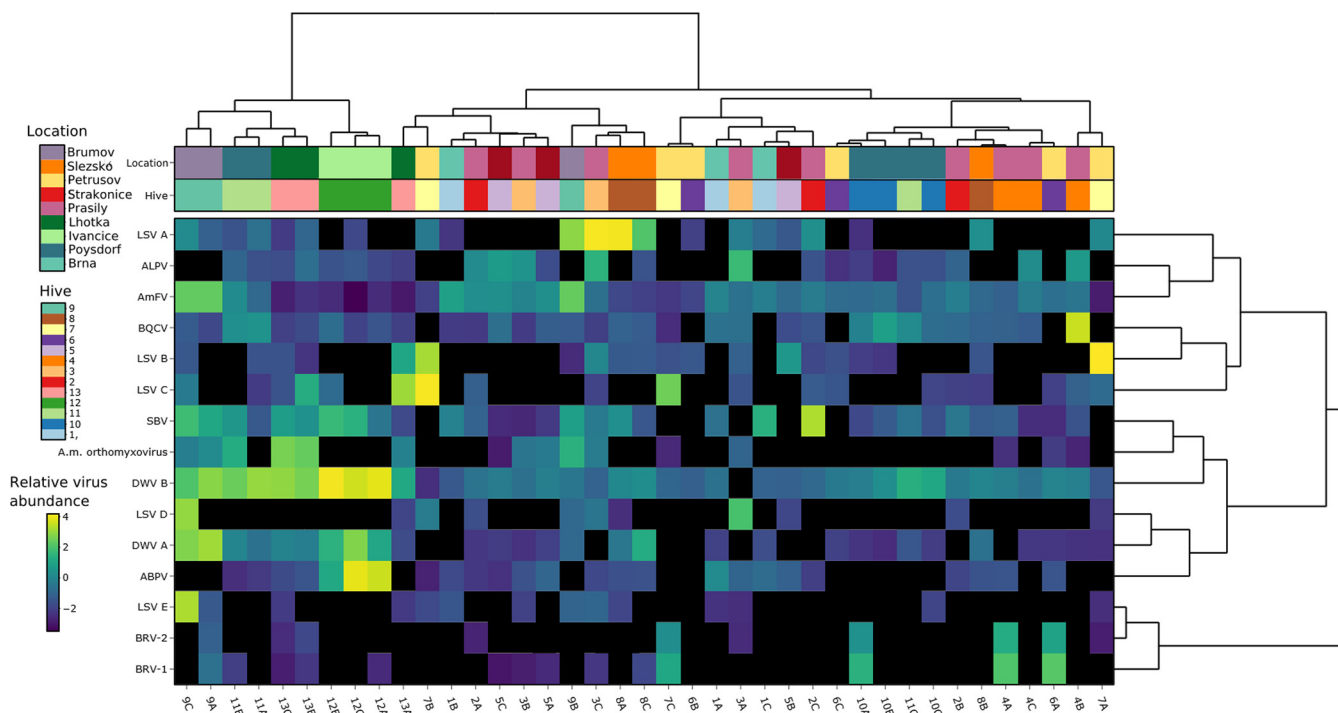


FIG 3 Diversity of viruses infecting honey bees in healthy bee colonies from Czechia. Relative abundances (viruses per 1 million sequencing reads, calculated from reference genome coverage [see Materials and Methods]) are shown on a \log_{10} scale. Samples (columns) and viruses (rows) are clustered by Ward’s minimum variance method algorithm and seriated by optimal leaf ordering. ABPV, acute bee paralysis virus; ALPV, aphid lethal paralysis virus; AmFV, *Apis mellifera* filamentous virus; BRV, bee rhabdovirus; BQCV, black queen cell virus; DWV, deformed wing virus; SBV, Sacbrood virus; LSV, Lake Sinai virus.

bee viromes in all other geographic regions except Australia (the trace amount can be attributed to read misalignment), whereas DWV-A and DWV-B were globally abundant. Australia clustered aside from other interleaved samples (for the principal-coordinate analysis [PCoA] and heatmap, see Fig. S1; for raw data, see Table S3).

Emerging viruses in samples from Czechia. The heatmap of bee virus abundances (Fig. 3) revealed a conspicuous pattern of cooccurrence of two recently discovered rhabdoviruses, bee rhabdovirus 1 (BRV-1) and BRV-2, which possibly infect both the honey bee and the mite *V. destructor* (29). Among our samples, BRV-1 and BRV-2 were always present together (pools 4A, 6A, 7C, 9A, and 10A), whereas both identified rhabdoviruses were absent from the remaining replicates (Fig. 3). The positivity of one out of three replicates probably implicates a low prevalence of BRV-positive bees within hives. In addition, BRV-1 always showed a higher abundance than BRV-2 in individual samples (Fig. 3). These viruses, albeit related, are phylogenetically distinct and display very limited sequence similarity (Fig. S1).

Furthermore, we reanalyzed NGS data from BRV-1-positive samples reported in three previous studies (26, 29, 35) by differentially mapping the sequencing reads to BRV-1 and BRV-2 reference genomes. Despite differences in multiple sample characteristics (i.e., bee pooling, nucleic acid isolation, and library preparation), we detected BRV-2 in all BRV-1-positive samples (Table 1). As in the Czech samples, BRV-2 was always present at a lower abundance. We consider this to be an indicator of an unusual relationship between the two rhabdoviruses, which has been left unnoticed previously (see Discussion).

Surprisingly, our analysis also revealed a large diversity of LSVs. The LSV genomes identified in the Czech samples were distributed in five positions in the global LSV phylogenetic tree (Fig. 5). Five variants were thus designated *de novo* as LSV-A to LSV-E for the purpose of this study. The interlineage nucleotide identity among Czech genomic sequences ranged from 72% to 79%, and the intralinear identity ranged from 86% to 99%. As shown in Fig. 5, among the closest relatives of LSV variants A, B, C, and E were sequences originating from different continents, suggesting that these variants

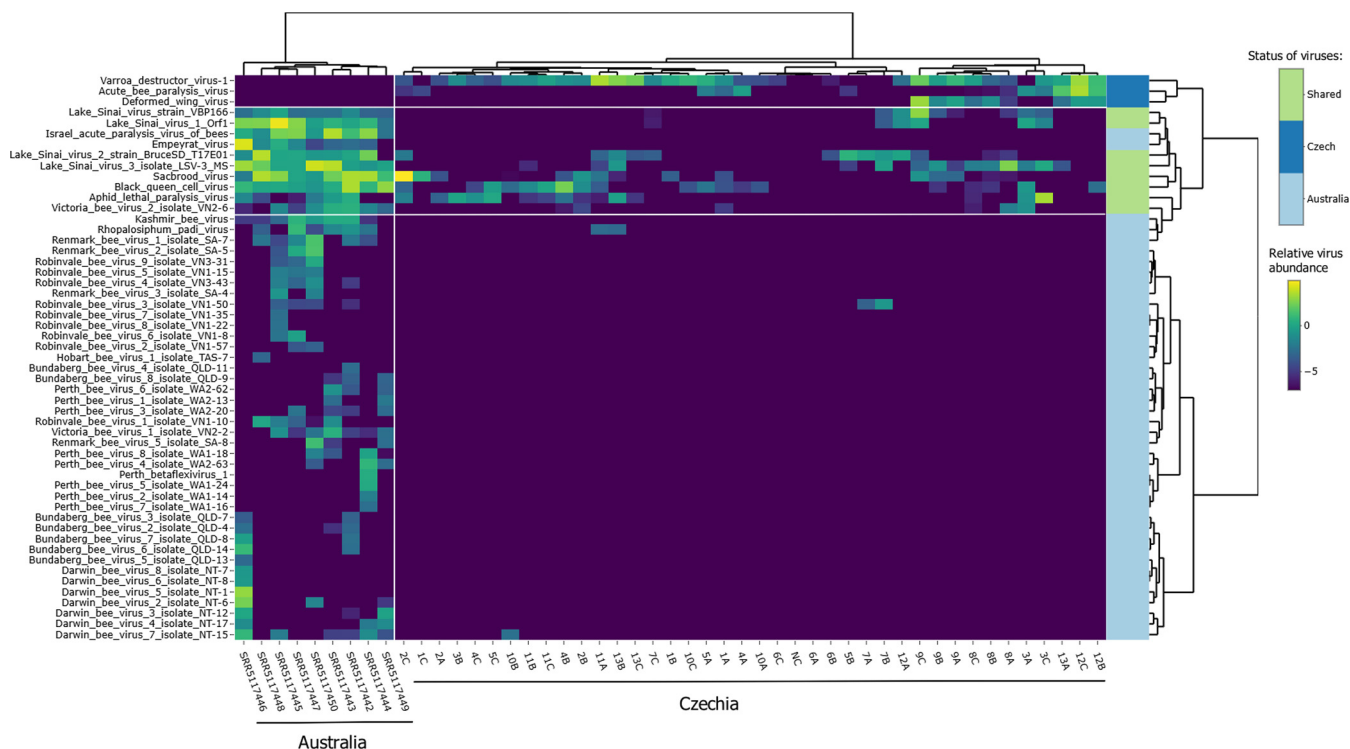


FIG 4 Heatmap of bee-infecting viruses in a comparison between the data from our study and Australian bees (SRA accession numbers [SRR5117442](#) to [SRR5117450](#)). Relative abundances are shown on a log₁₀ scale. Samples (columns) and viruses (rows) are clustered by Ward’s minimum variance method algorithm and seriated by optimal leaf ordering. Row colors show if the virus is present in Australia, Czechia, or both countries. Czech, viruses found only in Czechia; Australia, viruses found only in Australia; shared, viruses present in both regions. White lines separate the heatmap into several parts, Australian/Czech samples and viruses present/absent in the given regions.

belonged to LSV lineages with an intercontinental or even a global distribution. The RNA-dependent RNA polymerase (RdRp) phylogeny places LSV-D among a dozen European and one Iranian LSV genotypes; however, the resolution of the RdRp fragment phylogeny is insufficient, as indicated by the level of bootstrap support (Fig. S2).

In general, the distribution of LSV variants in Czechia was variable in both between- and within-hive comparisons. Interestingly, two or more variants of LSV were detected in individual hives (LSV-B and LSV-C in hives 7 and 12 and LSV-A, LSV-D, and LSV-E in hive 3).

Bacteriophages and clustering. The results showed the high diversity and abundance of bacteriophages (Fig. S1). Therefore, we analyzed the bee phageome in detail.

TABLE 1 Cooccurrence of bee rhabdoviruses 1 and 2 in NGS samples here and in three other studies where BRV-1 was identified^a

Sample ID	No. of BRV-1 reads	No. of BRV-2 reads	BRV-1/BRV-2 ratio	Country	SRA accession no.	No. of bees/sample	NA for library construction
SRR3927497	9,509	765	12.4	Israel	SRR3927497	30	Total RNA
DWV	41,335	1,024	40.4	USA	SRR6033679	10	Virus-enriched (encapsulated) DNA + RNA
NE_AWD_1442	34,115	737	46.3	The Netherlands	SRR5109823	5	Total RNA (rRNA depleted)
SA_RI_49	83,757	6,791	12.3	South Africa	SRR5109831	5	Total RNA (rRNA depleted)
T_V9	104,233	181	575.9	Tonga	SRR5109822	5 (thoraces only)	Total RNA (rRNA depleted)
T_V10	38,738	12,120	3.2	Tonga	SRR5109821	5 (thoraces only)	Total RNA (rRNA depleted)
T_T12	449,293	4,501	99.8	Tonga	SRR5109828	5 (thoraces only)	Total RNA (rRNA depleted)
T_T23	331,169	113	2,930.7	Tonga	SRR5109834	5 (thoraces only)	Total RNA (rRNA depleted)
4A	44,492	6,741	6.6	Czechia	SRS11094606	9	Virus-enriched (encapsulated) DNA + RNA
6A	83,517	4,475	18.7	Czechia	SRS11094614	9	Virus-enriched (encapsulated) DNA + RNA
7C	7,029	1,072	6.6	Czechia	SRS11094620	9	Virus-enriched (encapsulated) DNA + RNA
9A	127	40	3.2	Czechia	SRS11094624	9	Virus-enriched (encapsulated) DNA + RNA
10A	16,770	1,493	11.2	Czechia	SRS11094627	9	Virus-enriched (encapsulated) DNA + RNA

^aSequencing reads were mapped to a hybrid reference sequence consisting of combined BRV-1 and BRV-2 genomes (BRV-1, GenBank accession number [MH267692](#); BRV-2, GenBank accession number [KY354234](#)) to prevent the interference of multiple mapped reads. The BRV-1/BRV-2 ratio was calculated from read counts.

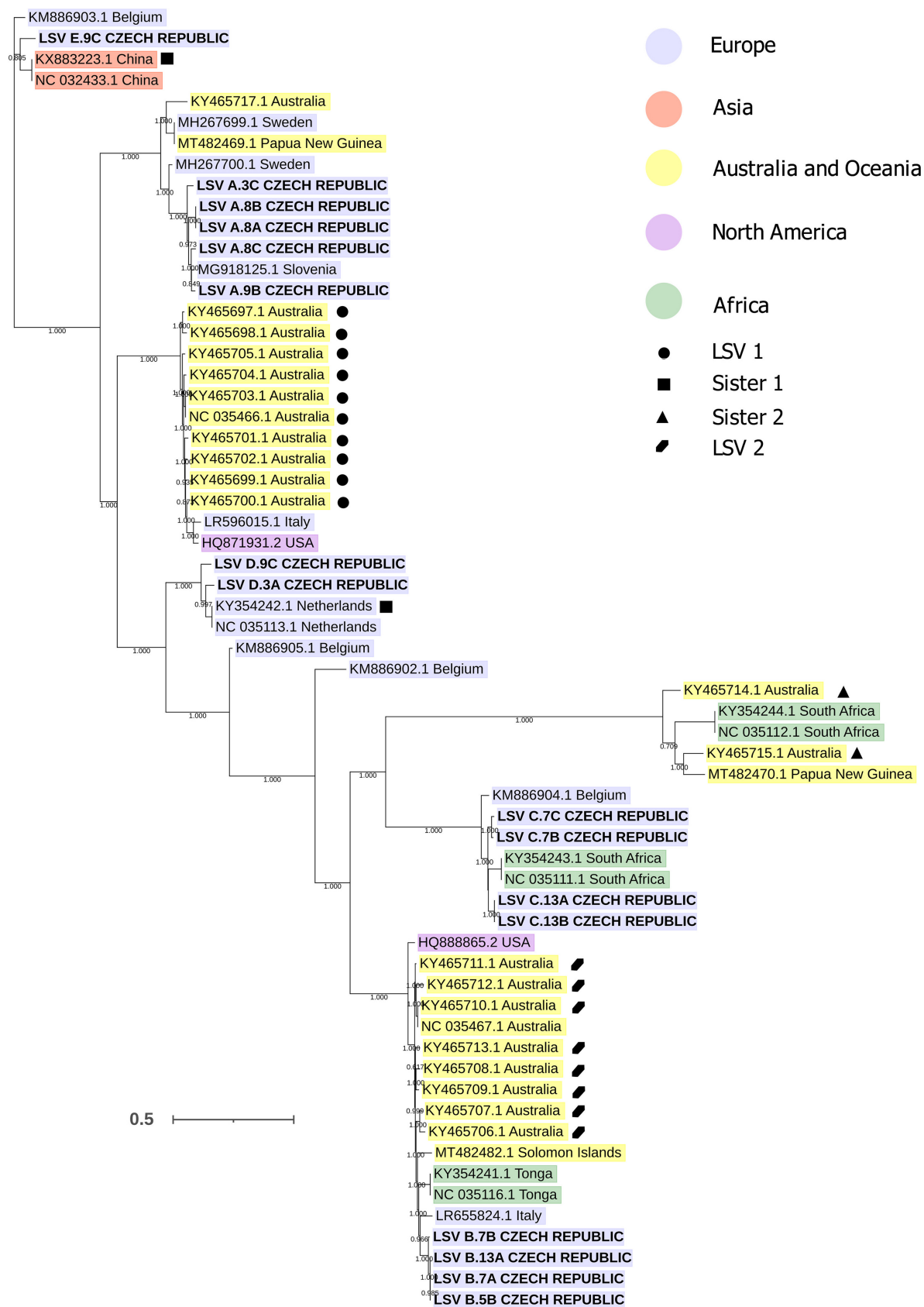


FIG 5 Phylogenetic tree of full LSV genomic sequences. LSV sequences are highlighted in color according to the continent of origin. Symbols mark isolates assigned to LSV lineages by Cornman (72). Sequences obtained in this study are in boldface capital letters. Bootstrap values are shown.

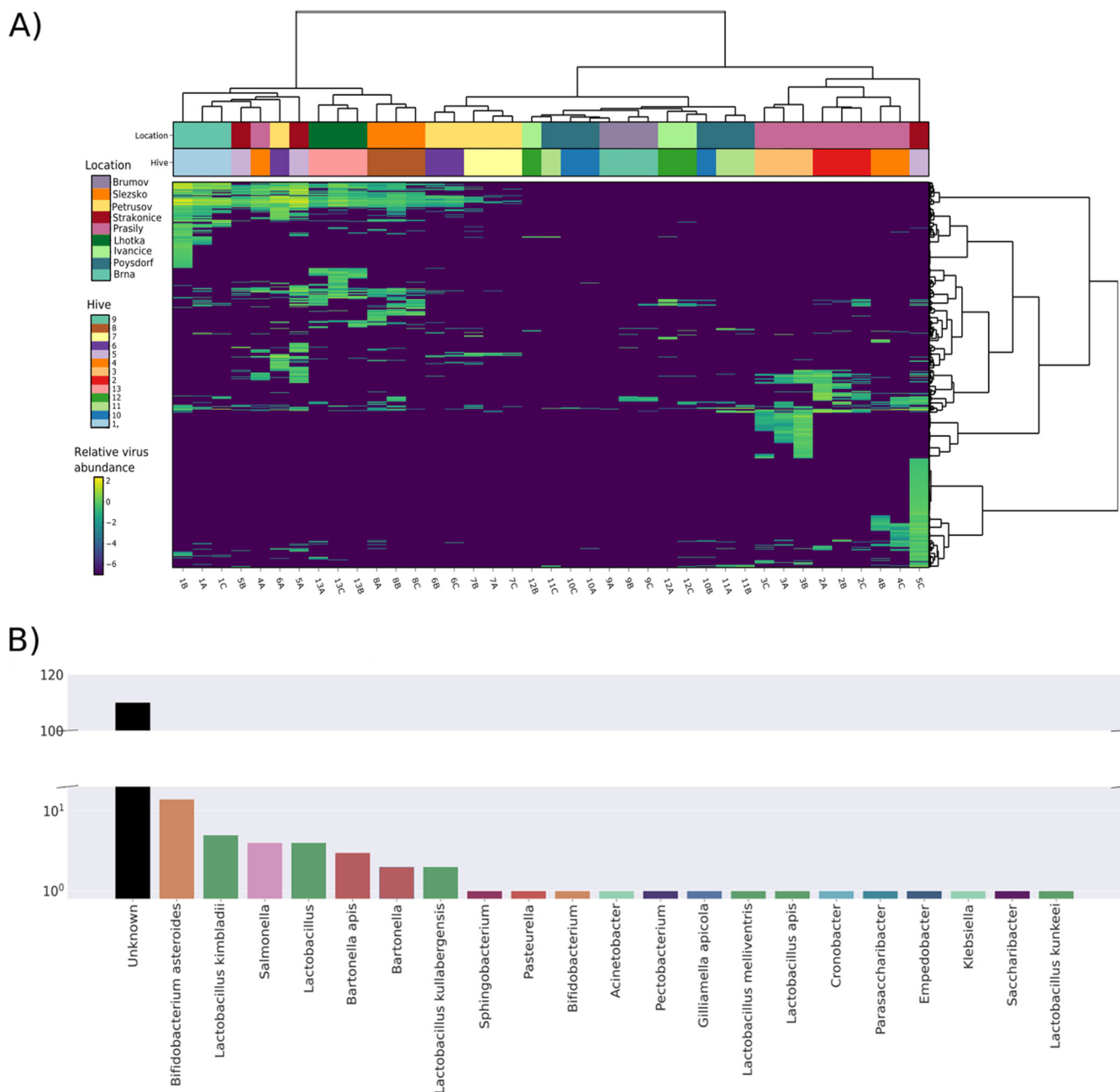


FIG 6 Host calling and clustering of prokaryotic viruses. (A) Heatmap of all contigs classified as bacteriophages. Relative abundances (viruses per 1 million sequencing reads) are shown on a log₁₀ scale. Samples (columns) and contigs (rows) are clustered by Ward's minimum variance method; both columns and rows are seriated by optimal leaf ordering. (B) Count of predicted hosts for 158 bacteriophage contigs identified by VirSorter2. Contigs with no prediction are in the "Unknown" category.

A heatmap was created by mapping reads to the set of phages identified by VirSorter2 (Table S1). Clustering of the phageome by the geographic origins of the samples was weaker than that in the heatmap of all viral sequences (Fig. 2), but still, replicate samples from seven hives clustered together. In five other hives, two out of three replicate samples clustered together; the replicate samples from hive 13 did not cluster (Fig. 6A). The Adonis test showed a highly significant correlation between the phageome and hive/location ($P < 0.0001$ by an Adonis-Bray test; R^2 , 0.47355/0.34024).

Since the plant viruses originating from pollen were abundant (Fig. S1), we further explored their clustering. The plant viruses clustered almost perfectly (Fig. S1) ($P < 0.0001$ by an Adonis-Bray test; R^2 , 0.70959 for hive and 0.53102 for location).

Interestingly, k-means clustering had higher adjusted mutual information scores for both bacteriophages and plant viruses than for all viral sequences (0.22 and 0.23, respectively, against 0.14) (Fig. S1).

To classify the phage genomes, which were predicted to be more than 50% complete (by CheckV), we used vConTACT2, which clustered the sequences with phages in the RefSeq database by their encoded protein profile. The resulting network had 398 individual viral clusters (roughly equivalent to genus-level assignment). Visualization of the resulting sequence similarity network (Fig. S3) shows the distribution of putative phage contigs through the network. Out of 158 individual phage contigs, 71 were unambiguously clustered. These formed 22 clusters, 15 of which were composed entirely of putative phage contigs from this study; 22 viral clusters (representing 26 putative bacteriophage contigs) were clustered with at least one reference sequence. Thanks to the reference sequences, the clusters could be tentatively classified as belonging to the family *Myoviridae* but also as belonging to the *Siphoviridae*, *Podoviridae*, and *Microviridae*. A single cluster contained strains from more than one viral family.

Host calling for each of the 158 detected phage contigs was performed through matches with CRISPR spacers identified using MinCED (36) and an additional analysis with CrisprOpenDB (37). Over 200 spacers matched the detected viral contigs, yet due to duplicate assignments (one contig matching spacers from multiple strains of one bacterial host species), only 22 (14%) of the phage contigs could be assigned to a host. The most common phage hosts were *Lactobacillus* species, *Bifidobacterium*, *Bartonella*, and *Salmonella* (Fig. 6B).

DISCUSSION

This study is the first to analyze the honey bee virome in Czechia, a country located in the heart of Central Europe. In addition, to our knowledge, this is the first work where the honey bee virome was examined in biological replicate samples within hives/colonies. Virome variation in sample replicates raises the question of to what extent single bees can affect the virome structure of an entire colony. However, the open question that remains is how NGS-based virome analysis of pools of bees can be affected by a single bee with a distinct virome. It is relevant to what we have attempted: to examine the homogeneity of virus infection within hives. A major factor that could affect the virome is that honey bee colonies were from a Central European country with one of the highest colony densities worldwide (10). Thus, despite the fact that *V. destructor* occurrence was low in all investigated colonies, and no signs of varroosis were observed, a virus(es) introduced by the mite (17, 38) was expected to be found. Furthermore, we compared our viromes with previously described viromes of Australian honey bees, which have never been exposed to varroosis (DWV and *Varroa* mite) (16, 39). When it came to bee-infecting viruses, our sets of viruses were diametrically different from those in Australia. Notably, one substantial difference was the lack of diverse members of the *Picornavirales* in our data set, which could be explained by *Varroa*-DWV interaction pressure indicated previously (17–22). Further investigation will be necessary because of methodological differences in this study compared to the study of Roberts et al. (16). For future comparability with other studies, it is also of importance to note that our samples were collected at the end of the beekeeping season in August/September because seasonal variations in virus occurrence have been previously described (40), whereas the Australian samples were collected between a longer period (August 2013 to April 2015, but the majority were collected in August).

Traditional and new bee viruses. The following viruses were detected in Czech samples: BQCV, AmFV, DWV-A and -B, SBV, ALPV, ABPV, BRV-1 and -2, *Apis mellifera* orthomyxovirus, and variants of LSV. The prevailing honey bee virus, in both abundance and prevalence, was DWV-B. Importantly, this result is in accordance with its recent global spread (20). Notably, the prevalence of DWV-B observed in asymptomatic (healthy) hives/colonies is in agreement with the results of a study by Norton et al., who showed DWV-B persistence in colonies with low *V. destructor* levels and those

treated with miticides (21). Finally, the prevalence of DWV-A was low in our virome samples, consistent with the fact that varroosis was under control in our colonies (21).

We determined the viromes in 39 pools, each consisting of nine individual honey bees. Importantly, we show that none of the structures of our Czech virome resembled those of the Australian viromes (16). Moreover, unique viruses were identified in each data set. These results may correspond to the fact that the virome structure is affected by the presence/absence of *V. destructor* since Australian honey bees are *Varroa*-free (16, 39), while the Czech viromes originated from regions where *V. destructor* has been widely distributed since the late 1970s (41, 42). Although we worked with “healthy” honey bee colonies, where *V. destructor* was under control, the effect of the mite is still to be considered. In Czech samples, the diverse set of viruses belonging to the *Picornavirales* detected in Australian honey bees was absent. Another important observation is the complete absence of DWV in the Australian viromes (16) and the contrary wide DWV presence in our samples. This shift from diverse *Picornavirales* members to a primarily DWV-dominated virome could signify that the presence of *V. destructor* and its interaction with DWV change the virome of honey bees. We suggest that the absence of diverse *Picornavirales* in other samples from honey bee populations around the world, even though varroosis was under control in some, adds support to this assumption.

The two rhabdoviruses (BRV-1 and -2) exhibited a conspicuous pattern of distribution among our samples: (i) both viruses were always present simultaneously; (ii) in each sample, BRV-1 was more abundant than BRV-2; and (iii) BRV-1 and -2 were always present in only one out of three replicate nine-bee samples per hive. Reanalysis of previously reported BRV-positive NGS samples (26, 29, 35) independently validated both the first and second phenomena (see above). So far unnoticed, this relationship may indicate a type of (inter)dependence between these two viruses. It is possible that an interaction takes place, either a nondirect one (e.g., through the immune system) or a direct one (e.g., a defect in protein synthesis) (43). This phenomenon can also be caused by a common transmission route or another unknown mechanism. Either way, the precise nature of the relationship between BRV-1 and BRV-2 should be confirmed by a single-bee analysis.

Since the first description of LSVs by Runckel et al. (44), various variants of this group of new viruses have been discovered in different regions around the world, including Australia (39, 45–48). Such diversity, in combination with a global distribution, could indicate a long coexistence of LSV with honey bees (39). According to current knowledge, LSV infections occur asymptotically, with no described adverse effect, which might have hindered their discovery in the pre-omics era. It was suggested that, based on the specific antibody detection LSV was described as bee virus X or Y (39, 49, 50). Among 39 Czech samples, we detected five distinct variants (designated LSV-A to -E) that were distributed in 17/39 samples (read count of ≥ 1 per million reads). Several cases of the simultaneous presence of several variants in one colony/hive were observed. A single-bee analysis is needed to determine the precise prevalence of LSV variants and the frequency of their cooccurrence as the simultaneous presence of several LSV variants in a single bee has been reported previously (47). The phylogenetic analysis of LSV sequences suggests that the LSV variants identified in the Czech samples belong to lineages that are globally distributed. Given the variable distribution and extraordinary global diversity of LSV, additional LSV variants are likely to be found in future studies. However, the available LSV sequences, both partial and complete, suffer from a heavy geographic sampling bias, implying that the true diversity might be even (much) larger than currently observed.

Virome structure stability. When all viral sequences present in our samples were analyzed, most samples clustered in accordance with their site of origin. Clustering was also observed for samples originating from different hives from the same apiary and for hives from different apiaries located in the same municipality. It appears that this clustering is driven mainly by plant viruses and to a lesser extent by bacteriophages

(see Fig. S1 in the supplemental material). In this study, bacteriophage sequences were one of the most diverse and abundant virome components in our samples. However, bacteriophages still represent a novel topic in honey bee research (27, 28) and are often excluded from analyses. In the bacteriophage heatmap, replicate samples still clustered based on hive and location, like the whole virome (plant viruses plus bacteriophages and other viral sequences). This is in accordance with previously reported findings that the phageome should be relatively stable in the location for the given year (28). The plant viruses in the total bee virome display the strongest geographic dependence (Fig. S1). Unfortunately, to our knowledge, no studies examining the stability and geographic differences of the pollen virome have been carried out yet.

The robust overall clustering pattern was lost when only viruses known to infect honey bees were analyzed. The uneven distribution of bee-infecting viruses among replicate nine-bee samples from the same hive raises serious questions for the NGS study design. NGS studies are typically carried out on pools of a small number (~ 1 to 5 [e.g., see reference 34]) or large numbers (>30 [e.g., see reference 35]) of individual bees.

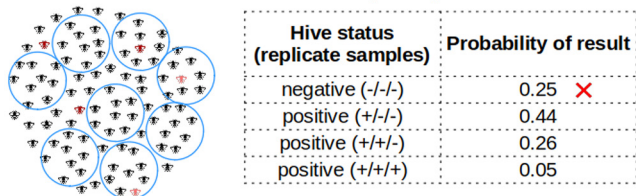
In our case, the pooling of nine bees per sample was not enough to balance the uneven representation of bee viruses among individuals. The recurring observations that only one nine-bee pool out of three yielded positivity for some viruses (e.g., BRV-1/2 [see “Traditional and new bee viruses,” above] indicate that the actual prevalence of infected worker bees per hive is low. In addition, we suspect that individual bees with relatively high viral loads are present in hives, and when they are randomly included in a pool for NGS virome analysis, this virus may dominate the virome of this pool. To accurately determine the abundance and prevalence of bee viruses in a hive, NGS of a large number of libraries, each prepared from an individual bee, appears to be the methodologically correct solution, which unfortunately is labor-intensive, expensive, and, thus, often beyond feasibility. Another possibility is to pool a large number of bees into a single master NGS library per hive (Fig. 7) and carry out NGS at a sequencing depth that is the same as or higher than the one in this study. The actual number of bees per master NGS library pool, our results suggest, should optimally be 50 individuals or more. Such an approach is expected to (i) detect the diversity of the viruses present, including those of low prevalence, and (ii) determine the actual genotypes of the viruses present in an affordable fashion (even though it can be challenging to disentangle closely related genomes). In this case, however, the viral prevalence in hives would remain to be assessed by analyzing similar quantities of individual bees, preferably by a rapid method like quantitative PCR (qPCR) (Fig. 7).

Bacteriophages. The phages identified in the Czech viromes belonged to several families, with *Myoviridae*, *Podoviridae*, *Microviridae*, and *Siphoviridae* being the most common. Out of the total of 158 (at least 50% complete) phage genomes, 26 could be directly clustered with the reference sequences from the database (approximately genus-level assignment), while the majority of contigs could not be assigned to a known genus or family (132), suggesting that multiple novel phage genera are present in honey bees. Host calling through CRISPR spacers predicted bacterial hosts for around 15% of bacteriophages. The predicted hosts (mainly *Lactobacillus* species, *Bifidobacterium*, *Bartonella*, and *Salmonella*) are those residing in the honey bee gut (51, 52). Even though this assignment level is higher than, e.g., that in the human gut, it did not reach the levels described previously for honey bees (27), with the majority of contigs being assigned to their hosts. This difference may be due to our approach for assigning a host to individual phage sequences instead of a viral cluster or to our stringent setting when comparing CRISPR spacers with phage sequences.

Conclusion. In conclusion, we identified an important aspect of the total bee virome: the bee-infecting viruses vary widely among individual bees, while the complete virome, which is composed predominantly of bacteriophages and plant viruses, is largely stable and geographically dependent. As the samples originate from the heart of Central Europe, our virome could be considered representative for the region. We revealed the absence of diverse *Picornavirales* in Czech and other global non-Australian honey bees, probably resulting from the cooccurrence/interaction of *V. destructor* and



This study's design:
3 x 9-bee samples per hive



Proposed design:
1 x 50-bee sample per hive

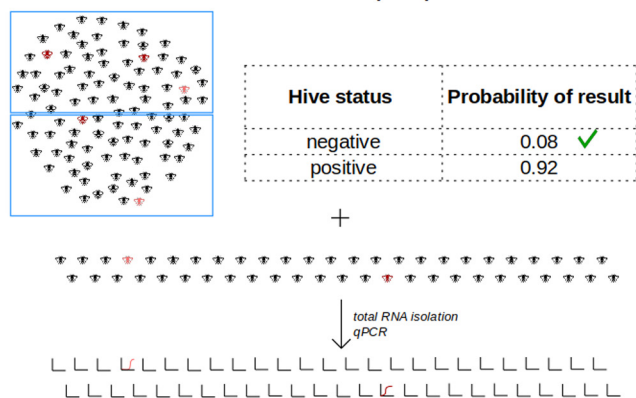


FIG 7 Effect of the NGS experimental design on hive status determination (positive/negative) for low-prevalence bee viruses (5%; 1/20 bees infected). + or – shows positivity or negativity, respectively, for that replicate sample (9-bee sample).

DWV. We report the wide presence of LSVs among Czech bees and their unexpected diversity, consisting of five globally represented variants. We provide the first description of a tentative close relationship between two related honey bee-infecting rhabdoviruses. Finally, we provide a preliminary characterization of bacteriophages present in the Czech honey bee virome samples.

MATERIALS AND METHODS

Sample origin. The worker bees were collected from 8 August to 29 September 2018 with owner permission from 13 representative bee colonies of nine beekeepers, some of whom provided more subspecies/subtype information (see Fig. S1 and Table S1 in the supplemental material). The honey bee workers were shaken off the brood combs into plastic bags, which were immediately placed into a box with dry ice for transport. The samples were then stored at –80°C until analyses. An overview of sampling sites (13 hives, 11 apiaries, and 9 locations) is shown in Fig. S1. All apiaries enrolled in this study were healthy; i.e., they did not exhibit symptoms of common pathogen infections or *V. destructor* infestation. All hives (except hive 9) were previously treated against *V. destructor* by either organic acids or amitraz. Three subspecies/subtypes of honey bee as provided by the beekeeper were included: Buckfast honey bee (*A. m. buckfast*), a hybrid of dark honey bee (*A. m. mellifera* Linnaeus 1758), and Carniolan honey bee (*A. m. carnica* Pollman 1879). Some bees originated from different apiaries from the same location or from different hives of the same apiary (metadata are available in Table S1).

Nucleic acid extraction and sequencing library preparation. Bees were processed according to the NetoVir protocol (28, 53). Three randomly chosen bees from each hive were homogenized in a tube with 2.8-mm ceramic beads (zirconium oxide) (Bertin Technologies, Montigny-le-Bretonneux, France) and 1 mL of sterile phosphate-buffered saline (PBS) using a Minilys homogenizer (Bertin Technologies, Montigny-le-Bretonneux, France) at 3,000 rpm for 5 min. After centrifugation at 17,000 × g for 3 min,

the supernatant was filtered through a 0.8- μ m filter (polyethersulfone [PES]) (Sartorius, Göttingen, Germany) and centrifuged again at 17,000 $\times g$ for 1 min. The supernatant from three homogenates, each consisting of three individuals, gave a pool of nine bees after mixing in one tube. Overall, three distinct replicates of the pooled nine individuals from a hive, denoted replicates A, B, and C, were used in further analyses.

For nuclease treatment, 260 μ L of each pooled sample was used and treated with 4 μ L of Benzonase nuclease and 2 μ L of micrococcal nuclease (New England BioLabs, Ipswich, MA, USA). The total nucleic acids were extracted using the QIAamp viral RNA minikit (Qiagen, Hilden, Germany) according to the manufacturer's instructions, without using carrier RNA. Extracted nucleic acids were reverse transcribed and amplified with the WTA2 kit using 17 amplification cycles (Sigma-Aldrich, St. Louis, MO, USA). The concentration of samples was measured by the Qubit dsDNA (double-stranded DNA) HS (high-sensitivity) assay kit (Thermo Fisher Scientific, Waltham, MA, USA) on the Qubit 2.0 fluorometer. Three nanograms of isolated DNA was processed with the Nextera XT library preparation kit (Illumina, San Diego, CA, USA). The quality of the DNA libraries and size distribution were evaluated using a high-sensitivity DNA assay on a Bioanalyzer 2100 instrument (Agilent Technologies, CA, USA), and the concentration was measured on the Qubit 2.0 fluorometer. The libraries were sent on dry ice to the KU Leuven Nucleomics Core (VIB), Leuven, Belgium, for analysis. Sequencing was performed on the HiSeq2000 platform (Illumina, CA, USA) for 2 \times 150-bp paired-end cycles.

Bioinformatic analysis. (i) Sequencing data processing and assembly. Sequencing quality was assessed using FastQC (Babraham Bioinformatics, Cambridge, UK) both before and after trimming. Adapters and low-quality bases were removed using Trimmomatic (54) with settings of 4, 20; leading of 19; trailing of 15; and minlen of 50. The assembly of trimmed reads was done with SPAdes (55) with the metagenomic option and using the following k-mers: 21, 33, 55, and 77. Contigs larger than 500 nucleotides (nt) and with identities of >95% and coverage of at least 80% of the length of the shortest scaffold were merged with ClusterGenomes (<https://bitbucket.org/MAVERICLab/docker-clustergenomes>). DIAMOND (56) with the settings -sensitive and -c 1 was used to compare sequences against the nonredundant protein database (NCBI) downloaded on 30 September 2018 and subsequently annotated via Kronatools (57). Individual reads were mapped against the nonredundant contigs with BWA-MEM (58), and BamM was used to determine coverage (tpmean) (<https://github.com/Ecogenomics/BamM>).

(ii) Targeted analysis of bee-infecting viruses. To obtain precise mapping information about the viruses that infect honey bees, we performed an additional analysis targeted on individual bee viruses. Reference genomic sequences of all currently recognized viral species known to infect honey bees (33) were retrieved from the GenBank database. Sequencing reads were mapped to these reference sequences under conditions of a maximum of 20% mismatches and a maximum of 20% gaps using the Geneious 6.0.3 Read Mapper (59). Consensus nucleotide sequences (majority rule) were called for viruses with complete or nearly complete coverage of reference sequences; the terminal and low-coverage regions of consensus sequences were visually inspected and manually curated. Virus abundance values (virus reads per 1 million sequencing reads) were determined from the sequencing read coverage of the actual viral sequences present in the samples from Czechia.

Comparison with other studies. For comparisons, we addressed data from previous honey bee virome studies (16, 26, 29, 30, 34). FastQ files were retrieved with the prefetch and fasterq-dump tools available in the SRA toolkit (NCBI). Reads were mapped against viruses known to infect bees (33) with BWA-MEM (58), and coverage was extracted with BamM using the tpmean method. Only RNA viruses (DNA viruses were not investigated in the Australian samples) with sum tpmean values over all samples of ≥ 20 were included for downstream analysis. Furthermore, we specifically screened our data for novel viruses from the order *Picornavirales*.

Phylogenetic analysis. For phylogenetic analysis, complete and partial (>500-bp; RdRp) LSV sequences (as of 6 January 2021) were retrieved from the NCBI database and combined with the LSV sequences obtained in this study. Sequence alignment was done with MAFFT with -localpair; -maxiterate 1000 (60). Alignments were trimmed with trimAL -automated1 (61), and the best model was determined by ModelTest-NG (62). Phylogenetic trees were created with PhyML (63) with the model best suited for alignment.

Bacteriophage identification. Bacteriophages were identified from all nonredundant contigs (>500 bp) with VirSorter2 (64), ignoring the groups Nucleocytoplasmic large DNA viruses and *Lavidaviridae*. Presumed phage contigs were checked with CheckV (65) to determine their "completeness." Putative phage sequences that were at least 50% complete were then classified with VConTACT2 (66) with the BLASTP mode using the Prokaryotic Viral RefSeq 88 database MCL for protein clustering and ClusterONE for genome clustering. Host calling was performed using CRISPR spacers retrieved from a set of 304 genomes of bacterial species described to reside in the honey bee gut (NCBI and JGI IMG/M) (Table S2). Bacterial genomic sequences were processed with MinCED (36), and the predicted spacers were pulled and analyzed by BLAST against phage contigs with the stringent settings -ungapped and -perc_identity 100. A complementary host-calling approach was performed by utilizing a CrisprOpenDB (37) search against all complete bacterial genomes to identify bacterial hosts outside the common spectrum (only level 1 predictions). All predicted hosts from CrisprOpenDB were bee-infecting bacteria less frequently mentioned in the literature and therefore were not included in our database (MinCED). The results of the two predictions were merged.

Statistical analysis and visualization. Statistical analysis was done in R with the Adonis test (permutations, 10,000) implemented in the vegan package (67). Heatmaps were created using the heatmaply package (68), and trees were visualized with iTOL (69). The VConTACT2 network was visualized in Python with the graph-tool library (70). k-means clustering was done in Python with sklearn (71). For each data set, data were first scaled by a standard scaler, and clusters were predicted with k-means

(13 clusters). Labels of these predicted clusters were compared with labels of real clusters (1 to 13). Gained scores are available in Fig. S1.

Data availability. The sequencing reads were deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number [PRJNA781422](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA781422). The assembled viral genomic sequences were deposited in GenBank (accession numbers [OL803813](https://www.ncbi.nlm.nih.gov/nuclseq/OL803813) to [OL803870](https://www.ncbi.nlm.nih.gov/nuclseq/OL803870)).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 0.6 MB.

FIG S2, PDF file, 0.3 MB.

FIG S3, PDF file, 0.2 MB.

TABLE S1, XLSX file, 1.6 MB.

TABLE S2, XLSX file, 0.01 MB.

TABLE S3, XLSX file, 0.05 MB.

ACKNOWLEDGMENTS

We thank M. Simonovsky for valuable help with sample collection. We thank beekeepers for allowing us to collect bee samples. We also thank N. Vaclavikova for technical support.

This study was supported by CELSA [Metagenomic Analysis of Honey Bee (*Apis mellifera*) in the Czech Republic, 2017 to 2019], grant number QK1910018 of the Ministry of Agriculture of the Czech Republic (<http://www.eagri.cz>), and the MICOBION project funded from EU H2020 (number 810224).

REFERENCES

- Schmidt JO. 1997. Bee products, p 15–26. In Mizrahi A, Lensky Y (ed), *Bee products: properties, applications, and apitherapy*. Springer, Boston, MA.
- Klein A-M, Vaissière BE, Cane JH, Steffan-Dewenter I, Cunningham SA, Kremen C, Tscharntke T. 2007. Importance of pollinators in changing landscapes for world crops. *Proc Biol Sci* 274:303–313. <https://doi.org/10.1098/rspb.2006.3721>.
- Calderone NW. 2012. Insect pollinated crops, insect pollinators and US agriculture: trend analysis of aggregate data for the period 1992–2009. *PLoS One* 7:e37235. <https://doi.org/10.1371/journal.pone.0037235>.
- Hung K-LJ, Kingston JM, Albrecht M, Holway DA, Kohn JR. 2018. The worldwide importance of honey bees as pollinators in natural habitats. *Proc Biol Sci* 285:20172140. <https://doi.org/10.1098/rspb.2017.2140>.
- Neumann P, Carreck NL. 2010. Honey bee colony losses. *J Apic Res* 49: 1–6. <https://doi.org/10.3896/IBRA.1.49.1.01>.
- Gray A, Adjlane N, Arab A, Ballis A, Brusbardis V, Charrière J-D, Chlebo R, Coffey MF, Cornelissen B, da Costa CA, Dahle B, Danihlík J, Dražić MM, Evans G, Fedoriak M, Forsythe I, Gajda A, de Graaf DC, Gregorc A, Ilieva I, Johannesen J, Kauko L, Kristiansen P, Martikkala M, Martín-Hernández R, Medina-Flores CA, Mutinelli F, Patalano S, Raudmets A, Martin GS, Soroker V, Stevanovic J, Uzunov A, Vejsnaes F, Williams A, Zammit-Mangion M, Brodschneider R. 2020. Honey bee colony winter loss rates for 35 countries participating in the COLOSS survey for winter 2018–2019, and the effects of a new queen on the risk of colony winter loss. *J Apic Res* 59: 744–751. <https://doi.org/10.1080/00218839.2020.1797272>.
- vanEngelsdorp D, Evans JD, Saegerman C, Mullin C, Haubruge E, Nguyen BK, Frazier M, Frazier J, Cox-Foster D, Chen Y, Underwood R, Tarpy DR, Pettis JS. 2009. Colony collapse disorder: a descriptive study. *PLoS One* 4: e6481. <https://doi.org/10.1371/journal.pone.0006481>.
- Anderson DL, Trueman J. 2000. Varroa jacobsoni (Acari: Varroidae) is more than one species. *Exp Appl Acarol* 24:165–189. <https://doi.org/10.1023/A:1006456720416>.
- Cornman RS, Tarpy DR, Chen Y, Jeffreys L, Lopez D, Pettis JS, vanEngelsdorp D, Evans JD. 2012. Pathogen webs in collapsing honey bee colonies. *PLoS One* 7:e43562. <https://doi.org/10.1371/journal.pone.0043562>.
- De la Rúa P, Jaffé R, Dall'Olio R, Muñoz I, Serrano J. 2009. Biodiversity, conservation and current threats to European honeybees. *Apidologie* 40: 263–284. <https://doi.org/10.1051/apido/2009027>.
- Möckel N, Gisdler 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. <https://doi.org/10.1099/vir.0.025940-0>.
- Wilfert L, Long G, Leggett HC, Schmid-Hempel P, Butlin R, Martin SJM, Boots M. 2016. Deformed wing virus is a recent global epidemic in honeybees driven by Varroa mites. *Science* 351:594–597. <https://doi.org/10.1126/science.aac9976>.
- Brødsgaard CJ, Ritter W, Hansen H, Brødsgaard HF. 2000. Interactions among Varroa jacobsoni mites, acute paralysis virus, and Paenibacillus larvae larvae and their influence on mortality of larval honeybees in vitro. *Apidologie* 31:543–554. <https://doi.org/10.1051/apido:2000145>.
- Martin SJ. 2001. The role of Varroa and viral pathogens in the collapse of honeybee colonies: a modelling approach. *J Appl Ecol* 38:1082–1093. <https://doi.org/10.1046/j.1365-2664.2001.00662.x>.
- Francis RM, Nielsen SL, Kryger P. 2013. Varroa-virus interaction in collapsing honey bee colonies. *PLoS One* 8:e57540. <https://doi.org/10.1371/journal.pone.0057540>.
- Roberts JMK, Anderson DL, Durr PA. 2018. Metagenomic analysis of Varroa-free Australian honey bees (*Apis mellifera*) shows a diverse Picornavirales virome. *J Gen Virol* 99:818–826. <https://doi.org/10.1099/jgv.0.001073>.
- 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* 336:1304–1306. <https://doi.org/10.1126/science.1220941>.
- Ryabov EV, Wood GR, Fannon JM, Moore JD, Bull JC, Chandler D, Mead A, Burroughs N, Evans DJ. 2014. A virulent strain of deformed wing virus (DWV) of honeybees (*Apis mellifera*) prevails after Varroa destructor-mediated, or in vitro, transmission. *PLoS Pathog* 10:e1004230. <https://doi.org/10.1371/journal.ppat.1004230>.
- Posada-Florez F, Childers AK, Heerman MC, Egekwu NI, Cook SC, Chen Y, Evans JD, Ryabov EV. 2019. Deformed wing virus type A, a major honey bee pathogen, is vectored by the mite Varroa destructor in a non-propagative manner. *Sci Rep* 9:12445. <https://doi.org/10.1038/s41598-019-47447-3>.
- Norton AM, Remnant EJ, Buchmann G, Beekman M. 2020. Accumulation and competition amongst deformed wing virus genotypes in naïve Australian honeybees provides insight into the increasing global prevalence of genotype B. *Front Microbiol* 11:620. <https://doi.org/10.3389/fmicb.2020.00620>.
- Norton AM, Remnant EJ, Tom J, Buchmann G, Blacquiere T, Beekman M. 2021. Adaptation to vector-based transmission in a honeybee virus. *J Anim Ecol* 90:2254–2267. <https://doi.org/10.1111/1365-2656.13493>.
- Doublet V, Labarussias M, de Miranda JR, Moritz RFA, Paxton RJ. 2015. Bees under stress: sublethal doses of a neonicotinoid pesticide and

- pathogens interact to elevate honey bee mortality across the life cycle. *Environ Microbiol* 17:969–983. <https://doi.org/10.1111/1462-2920.12426>.
23. Martin SJ, Hardy J, Villalobos E, Martín-Hernández R, Nikaido S, Higes M. 2013. Do the honeybee pathogens *Nosema ceranae* and deformed wing virus act synergistically? *Environ Microbiol Rep* 5:506–510. <https://doi.org/10.1111/1758-2229.12052>.
 24. McMenamin AJ, Genersch E. 2015. Honey bee colony losses and associated viruses. *Curr Opin Insect Sci* 8:121–129. <https://doi.org/10.1016/j.cois.2015.01.015>.
 25. McMenamin AJ, Flenniken ML. 2018. Recently identified bee viruses and their impact on bee pollinators. *Curr Opin Insect Sci* 26:120–129. <https://doi.org/10.1016/j.cois.2018.02.009>.
 26. Galbraith DA, Fuller ZL, Ray AM, Brockmann A, Frazier M, Gikungu MW, Martinez JFI, Kapheim KM, Kerby JT, Kocher SD, Losyev O, Muli E, Patch HM, Rosa C, Sakamoto JM, Stanley S, Vaudo AD, Grozinger CM. 2018. Investigating the viral ecology of global bee communities with high-throughput metagenomics. *Sci Rep* 8:8879. <https://doi.org/10.1038/s41598-018-27164-z>.
 27. Bonilla-Rosso G, Steiner T, Wichmann F, Bexkens E, Engel P. 2020. Honey bees harbor a diverse gut virome engaging in nested strain-level interactions with the microbiota. *Proc Natl Acad Sci U S A* 117:7355–7362. <https://doi.org/10.1073/pnas.2000228117>.
 28. Deboutte W, Beller L, Yinda CK, Maes P, de Graaf DC, Matthijnsens J. 2020. Honey-bee-associated prokaryotic viral communities reveal wide viral diversity and a profound metabolic coding potential. *Proc Natl Acad Sci U S A* 117:10511–10519. <https://doi.org/10.1073/pnas.1921859117>.
 29. Remnant EJ, Shi M, Buchmann G, Blacquièrre T, Holmes EC, Beekman M, Ashe A. 2017. A diverse range of novel RNA viruses in geographically distinct honey bee populations. *J Virol* 91:e00158-17. <https://doi.org/10.1128/JVI.00158-17>.
 30. Levin S, Sela N, Erez T, Nestel D, Pettis J, Neumann P, Chejanovsky N. 2019. New viruses from the ectoparasite mite *Varroa destructor* infesting *Apis mellifera* and *Apis cerana*. *Viruses* 11:94. <https://doi.org/10.3390/v11020094>.
 31. Martin SJ, Brettell LE. 2019. Deformed wing virus in honeybees and other insects. *Annu Rev Virol* 6:49–69. <https://doi.org/10.1146/annurev-virology-092818-015700>.
 32. Alger SA, Burnham PA, Noncristiani HF, Brody AK. 2019. RNA virus spillover from managed honeybees (*Apis mellifera*) to wild bumblebees (*Bombus* spp.). *PLoS One* 14:e0217822. <https://doi.org/10.1371/journal.pone.0217822>.
 33. Beaupaire A, Piot N, Doublet V, Antunez K, Campbell E, Chantawannakul P, Chejanovsky N, Gajda A, Heerman M, Panziera D, Smagghe G, Yañez O, de Miranda JR, Dalmon A. 2020. Diversity and global distribution of viruses of the Western honey bee, *Apis mellifera*. *Insects* 11:239. <https://doi.org/10.3390/insects11040239>.
 34. Deboutte W, Beller L, Yinda CK, Shi C, Smets L, Vanmechelen B, Conceição-Neto N, Dallmeier K, Maes P, de Graaf DC, Matthijnsens J. 2020. Hymenoptera associated eukaryotic virome lacks host specificity. *bioRxiv* <https://doi.org/10.1101/2020.09.15.298042>.
 35. Levin S, Sela N, Chejanovsky N. 2016. Two novel viruses associated with the *Apis mellifera* pathogenic mite *Varroa destructor*. *Sci Rep* 6:37710. <https://doi.org/10.1038/srep37710>.
 36. Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpidis NC, Hugenholtz P. 2007. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics* 8:209. <https://doi.org/10.1186/1471-2105-8-209>.
 37. Pourcel C, Touchon M, Villeriot N, Vernadet J-P, Couvin D, Toffano-Nioche C, Vergnaud G. 2020. CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and cas genes from complete genome sequences, and tools to download and query lists of repeats and spacers. *Nucleic Acids Res* 48:D535–D544. <https://doi.org/10.1093/nar/gkz915>.
 38. Bowen-Walker PL, Martin SJ, Gunn A. 1999. The transmission of deformed wing virus between honeybees (*Apis mellifera* L.) by the ectoparasitic mite *Varroa jacobsoni* Oud. *J Invertebr Pathol* 73:101–106. <https://doi.org/10.1006/jipa.1998.4807>.
 39. Roberts JMK, Anderson DL, Durr PA. 2017. Absence of deformed wing virus and *Varroa destructor* in Australia provides unique perspectives on honeybee viral landscapes and colony losses. *Sci Rep* 7:6925. <https://doi.org/10.1038/s41598-017-07290-w>.
 40. Gauthier L, Tentcheva D, Tournaire M, Dainat B, Cousserans F, Colin ME, Bergoin M. 2007. Viral load estimation in asymptomatic honey bee colonies using the quantitative RT-PCR technique. *Apidologie* 38:426–435. <https://doi.org/10.1051/apido:2007026>.
 41. Thompson HM, Brown MA, Ball RF, Bew MH. 2002. First report of *Varroa destructor* resistance to pyrethroids in the UK. *Apidologie* 33:357–366. <https://doi.org/10.1051/apido:2002027>.
 42. Roth MA, Wilson JM, Tignor KR, Gross AD. 2020. Biology and management of *Varroa destructor* (Mesostigmata: Varroidae) in *Apis mellifera* (Hymenoptera: Apidae) colonies. *J Integr Pest Manag* 11:1. <https://doi.org/10.1093/jipm/pmz036>.
 43. DaPalma T, Doonan BP, Trager NM, Kasman LM. 2010. A systematic approach to virus-virus interactions. *Virus Res* 149:1–9. <https://doi.org/10.1016/j.virusres.2010.01.002>.
 44. Runckel C, Flenniken ML, Engel JC, Ruby JG, Ganem D, Andino R, DeRisi JL. 2011. Temporal analysis of the honey bee microbiome reveals four novel viruses and seasonal prevalence of known viruses, *Nosema*, and *Criethidia*. *PLoS One* 6:e20656. <https://doi.org/10.1371/journal.pone.0020656>.
 45. Ravooet J, Maharramov J, Meeus I, De Smet L, Wenseleers T, Smagghe G, de Graaf DC. 2013. Comprehensive bee pathogen screening in Belgium reveals *Criethidia mellificae* as a new contributory factor to winter mortality. *PLoS One* 8:e72443. <https://doi.org/10.1371/journal.pone.0072443>.
 46. Granberg F, Vicente-Rubiano M, Rubio-Guerri C, Karlsson OE, Kukielka D, Belák S, Sánchez-Vizcaino JM. 2013. Metagenomic detection of viral pathogens in Spanish honeybees: co-infection by aphid lethal paralysis, Israel acute paralysis and Lake Sinai viruses. *PLoS One* 8:e57459. <https://doi.org/10.1371/journal.pone.0057459>.
 47. Ravooet J, De Smet L, Wenseleers T, de Graaf DC. 2015. Genome sequence heterogeneity of Lake Sinai virus found in honey bees and Orf1/RdRp-based polymorphisms in a single host. *Virus Res* 201:67–72. <https://doi.org/10.1016/j.virusres.2015.02.019>.
 48. Daughenbaugh KF, Martin M, Brutscher LM, Cavigli I, Garcia E, Lavin M, Flenniken ML. 2015. Honey bee infecting Lake Sinai viruses. *Viruses* 7:3285–3309. <https://doi.org/10.3390/v7062772>.
 49. Bailey L, Ball B, Carpenter J, Woods R. 1980. Small virus-like particles in honey bees associated with chronic paralysis virus and with a previously undescribed disease. *J Gen Virol* 46:149–155. <https://doi.org/10.1099/0022-1317-46-1-149>.
 50. de Miranda JR, Bailey L, Ball BV, Blanchard P, Budge GE, Chejanovsky N, Chen Y-P, Gauthier L, Genersch E, de Graaf DC, Ribière M, Ryabov E, De Smet L, van der Steen JJM. 2013. Standard methods for virus research in *Apis mellifera*. *J Apic Res* 52:1–56. <https://doi.org/10.3896/IBRA.1.52.4.22>.
 51. Ellegaard KM, Engel P. 2019. Genomic diversity landscape of the honey bee gut microbiota. *Nat Commun* 10:446. <https://doi.org/10.1038/s41467-019-08303-0>.
 52. Anjum SI, Aldakheel F, Shah AH, Khan S, Ullah A, Hussain R, Khan H, Ansari MJ, Mahmoud AH, Mohammed OB. 2021. Honey bee gut an unexpected niche of human pathogen. *J King Saud Univ Sci* 33:101247. <https://doi.org/10.1016/j.jksus.2020.101247>.
 53. Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L, Deboutte W, Yinda CK, Lavigne R, Maes P, Ranst MV, Heylen E, Matthijnsens J. 2015. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep* 5:16532. <https://doi.org/10.1038/srep16532>.
 54. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
 55. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
 56. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>.
 57. Ondov BD, Bergman NH, Phillippy AM. 2011. Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics* 12:385. <https://doi.org/10.1186/1471-2105-12-385>.
 58. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
 59. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.

60. Katoh K. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* 33:511–518. <https://doi.org/10.1093/nar/gki198>.
61. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
62. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. 2020. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol* 37:291–294. <https://doi.org/10.1093/molbev/msz189>.
63. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321. <https://doi.org/10.1093/sysbio/syq010>.
64. Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, Delmont TO, Pratama AA, Gazitúa MC, Vik D, Sullivan MB, Roux S. 2021. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* 9:37. <https://doi.org/10.1186/s40168-020-00990-y>.
65. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat Biotechnol* 39:578–585. <https://doi.org/10.1038/s41587-020-00774-7>.
66. Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, Brister JR, Kropinski AM, Krupovic M, Lavigne R, Turner D, Sullivan MB. 2019. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat Biotechnol* 37:632–639. <https://doi.org/10.1038/s41587-019-0100-8>.
67. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlenn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2019. vegan: community ecology package (2.5-6). <https://CRAN.R-project.org/package=vegan>. Accessed 20 July 2020.
68. Galili T, O'Callaghan A, Sidi J, Sievert C. 2018. heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* 34:1600–1602. <https://doi.org/10.1093/bioinformatics/btx657>.
69. Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47:W256–W259. <https://doi.org/10.1093/nar/gkz239>.
70. Peixoto TP. 2017. The graph-tool python library. figshare.
71. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É. 2011. Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830.
72. Cornman RS. 2019. Relative abundance and molecular evolution of Lake Sinai virus (Sinaivirus) clades. *PeerJ* 7:e6305. <https://doi.org/10.7717/peerj.6305>.

Discovery and characterization of novel DNA viruses in *Apis mellifera*: expanding the honey bee virome through metagenomic analysis

Dominika Kadlečková,¹ Martina Saláková,¹ Tomáš Erban,² Ruth Tachezy¹

AUTHOR AFFILIATIONS See affiliation list on p. 17.

ABSTRACT To date, many viruses have been discovered to infect honey bees. In this study, we used high-throughput sequencing to expand the known virome of the honey bee, *Apis mellifera*, by identifying several novel DNA viruses. While the majority of previously identified bee viruses are RNA, our study reveals nine new genomes from the *Parvoviridae* family, tentatively named Bee densovirus 1 to 9. In addition, we characterized a large DNA virus, *Apis mellifera* filamentous-like virus (AmFLV), which shares limited protein identities with the known *Apis mellifera* filamentous virus. The complete sequence of AmFLV, obtained by a combination of laboratory techniques and bioinformatics, spans 152,678 bp. Linear dsDNA genome encodes for 112 proteins, of which 49 are annotated. Another large virus we discovered is *Apis mellifera* nudivirus, which belongs to a group of *Alphanudivirus*. The virus has a length of 129,467 bp and a circular dsDNA genome, and has 106 protein encoding genes. The virus contains most of the core genes of the family *Nudiviridae*. This research demonstrates the effectiveness of viral binning in identifying viruses in honey bee virology, showcasing its initial application in this field.

IMPORTANCE Honey bees contribute significantly to food security by providing pollination services. Understanding the virome of honey bees is crucial for the health and conservation of bee populations and also for the stability of the ecosystems and economies for which they are indispensable. This study unveils previously unknown DNA viruses in the honey bee virome, expanding our knowledge of potential threats to bee health. The use of the viral binning approach we employed in this study offers a promising method to uncovering and understanding the vast viral diversity in these essential pollinators.

KEYWORDS metagenomics, honey bee viruses, virome, virus discovery, honey bee

The honey bee (*Apis mellifera* Linnaeus, 1758) is an indispensable global pollinator, playing a pivotal role in global agriculture and ecosystem dynamics (1, 2) and is a model organism for biological research on eusocial insects (3, 4). Therefore, it is alarming that the high annual losses of honey bee colonies continue around the world. The overwintering losses in some countries exceed 30%, but in some regions, losses exceeding 50% have been reported (5, 6). Many different factors have been described that could account for these colony losses and most likely a combination of several of them is the cause (6, 7). One of the many factors is viral infections (8), which may interact with diverse biotic and abiotic stressors (9). Although the focus is often on pesticides, the main threat is *Varroa destructor* which triggers viral infections in honey bees (10, 11). It is important to consider that viruses are a common and abundant part of the bee colony, similar to the more thoroughly studied bacterial microbiome.

Editor Morgan G. I. Langille, Dalhousie University, Halifax, Canada

Address correspondence to Tomáš Erban, arachnid@centrum.cz; erban@vurv.cz, or Ruth Tachezy, ruth.tachezy@natur.cuni.cz.

The authors declare no conflict of interest.

See the funding table on p. 17.

Received 19 January 2024

Accepted 9 February 2024

Published 5 March 2024

Copyright © 2024 Kadlečková et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

All viruses found in certain spaces are called viromes. In bees, the virome consists of bee-infecting viruses, viruses infecting other eukaryotes living in/on bees (viruses of parasites), bacteriophages, and transient viruses present in pollinator resources such as pollen (12). Several new bee viruses have been discovered using next-generation sequencing (NGS) techniques (13–15). Thanks to NGS, approximately 72 viruses have been described in honey bees (16). Most of viruses found in honey bees belong to the +ssRNA group, such as the *Iflaviridae* and *Dicistroviridae*. So far, only a few DNA viruses have been identified: the large dsDNA virus *Apis mellifera* filamentous virus (AmFV), discovered in 1978 (17) but fully sequenced in 2015 (17, 18), and the recently discovered small single-stranded DNA viruses belonging to the families *Genomoviridae* and *Microviridae* (19).

In our recent project focusing on the honey bee virome, we found sequences of bee DNA viruses belonging to the *Parvoviridae* family, a large DNA virus related to the AmFV virus, and another to *Nudiviridae*. Parvoviruses, especially those belonging to the subfamily *Densovirinae*, are well-known small nonenveloped viruses that infect insects (20). The genome of parvoviruses consists of linear ssDNA, 4–6 kb long, which contains two major expression cassettes with open reading frames (ORFs) that code non-structural proteins and structural capsid proteins (21, 22). The known members of this family are highly pathogenic for their insect hosts (20). The other sequences we have identified are similar to a large dsDNA AmFV with a genome of over 450 kb. This virus has not yet been classified and some of its proteins show identities with proteins of the family *Baculoviridae* (18). Individuals with high viremia have milky hemolymph due to cellular degradation caused by the presence of virions and show signs of weakness in crawling bees at the entrance. Although the virus is widespread in colonies in different parts of the world (23, 24), clinical symptoms are detected only sporadically (17, 25). *Nudiviridae* are also a group of viruses that infect insects and crustaceans; they are enveloped dsDNA viruses with a large genome of 90–230 kb (26).

In this study, we, for the first time, describe bee parvoviruses, *Apis mellifera* filamentous-like virus and *Apis mellifera* nudivirus (AmNV). Laboratory and bioinformatic approaches were combined to complete the genomes of these new honey bee viruses. Viral binning and the creation of vMAGs (viral metagenome-assembled genomes) were used for the first time, to our knowledge, for the genome assembly of honey bee viruses.

MATERIALS AND METHODS

De novo sequencing

Honey bee collection, sample processing, and libraries preparation

Bees were collected from 18 hives in five different locations in Czechia [Lisnice, Libečov, Práslav, Brdý/Neřežín, and Prague-Ružyně in the Crop Research Institute (abbreviated as VURV)]. In addition, two of the five VURV colonies used for analyses were moved from a different site out of the flying range to a demarcated place in VURV two weeks before sampling. Immediately after sampling by shaking from a brood comb frame into a plastic bag, they were placed in a polystyrene box on dry ice. Then, the bees were divided to sterile centrifugal tubes and stored at -80°C until further use. Except for two colonies, which had obvious signs of varroosis, i.e. crippled wings, all hives were denoted to be “healthy,” that is, the colonies had rapid build-up, showed no damaged capping, and no signs of overt diseases or *Varroa* infestation were observed. The list and specifications of the samples are available on GitHub (<https://github.com/kadlck/NAZV19>).

Fifty randomly selected bees from each hive were pooled in one sample. As we suggested in (24) this number of bees in the sample should allow for capturing the full diversity of eukaryotic viruses in the colonies. We processed the bees as described in detail before (24). Briefly, the homogenization was done in four 5mL tubes with ceramic beads (Bertin technologies, Montigny-le-bretonneux, Ile-de-France, France), after centrifugation and filtration the bees were pooled into two separate aliquots of 25

bee pools, and after extraction of encapsulated nucleic acids with QIAamp Viral RNA Kit (Qiagen, Hilden, Germany), the two 25 bee pools were combined back into one sample of 50 bees. The reverse transcription of RNA and amplification of cDNA/DNA was done with WTA2 kit (Sigma-Aldrich, St. Louis, Missouri, United States). Libraries were prepared with Nextera XT (Illumina, San Diego, California, USA). The sequencing was performed in several runs on NextSeq 500 using Mid Output Kits (Illumina, San Diego, California, USA) for 2×150 bp paired-end cycles, with minimum of 10M reads per sample. The 1xPBS was used as a negative control and has been processed through all the steps (from homogenization to sequencing) to exclude possible contamination of samples and reagents during processing (see Fig. 1 for overview).

Bioinformatics—data analysis

The quality of sequences was checked with FastQC v0.11.9 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>). Trimming was performed with Trimmomatic v0.39.10 (27) using adapters/primers from both amplification steps with ILLUMINACLIP, HEADCROP:19, LEADING:15, TRAILING:15, SLIDINGWINDOW:4:20, and MINLEN:50. After trimming, the quality of the reads was checked again. The reads were assembled with SPAdes v3.15.3 (28) --meta using the following k-mers: k 21, 33, 55, 77, and the gained contigs were classified with Diamond v2.0.11 (29) blastx against the non-redundant (NR) database ((National Center for Biotechnology Information (NCBI) downloaded on 9 August 2021). The classification for each sample was displayed using Kronatools v2.8.1 (30). Contigs larger than 500 bp from all samples were clustered based on pairwise ANI (average nucleotide identity) with 95% identity over 85% length using scripts and instructions provided by CheckV v1.0.1 (31). Classification was done again using Diamond v2.0.11 (29) blastx, but with --sensitive and --c 1 settings. The reads were

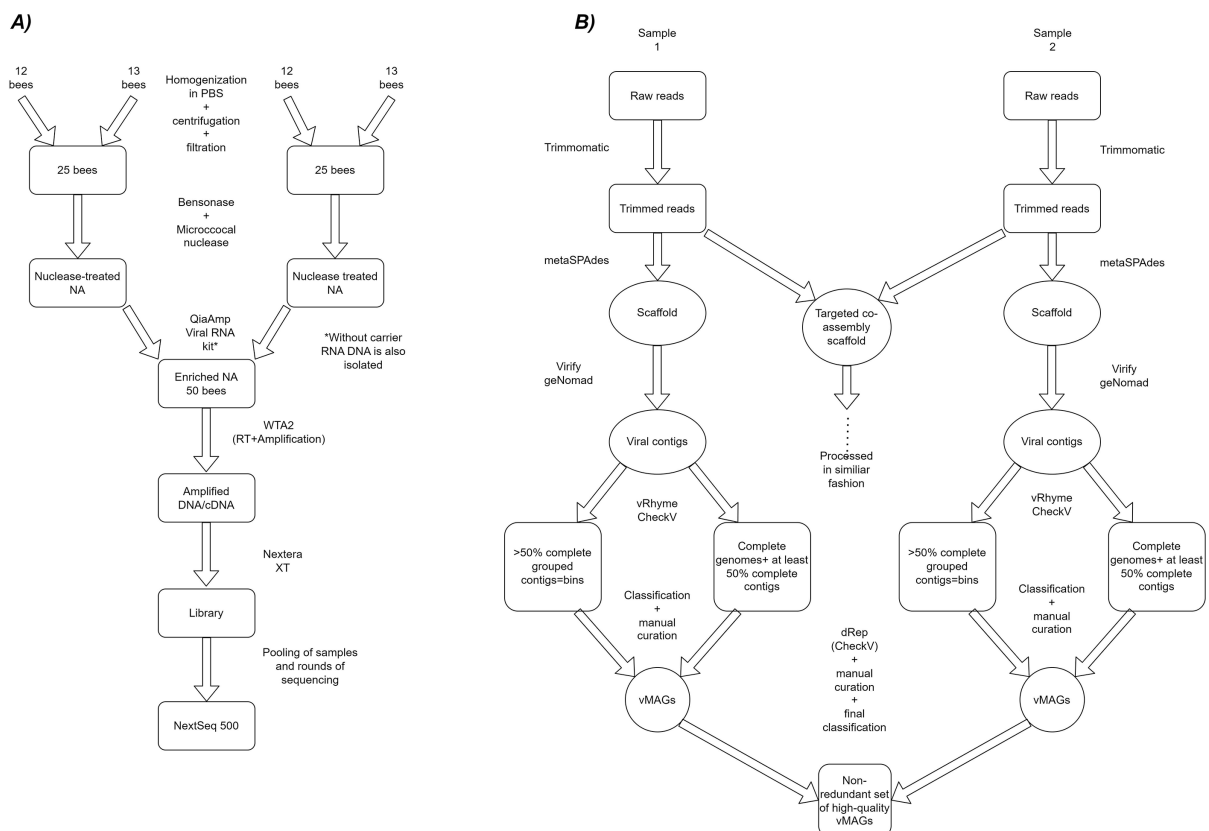


FIG 1 Overview of the workflow of sample preparation (A) and bioinformatical procedure (B) used in this study.

mapped back to the contigs using bwa-mem2 v2.2.1 (32) and the mapping of reads was extracted using CoverM v0.6.1 (<https://github.com/wwood/CoverM>).

We considered removing the host reads from the samples prior to analysis but they made up on average ~36% of all reads. When we compared the results of assembly of dehosted reads with those non-dehosted we found that the statistic of the scaffold was in very slight favour of non-dehosting option on most of the samples. Therefore, we continued with reads containing the host reads (Table 1 and full statistics on [GitHub](#)).

Determination of the complete genome of *Apis mellifera* filamentous-like virus

Bioinformatics—the creation of MAGs

Since we found contigs belonging to large viruses, we tried binning and creating vMAGs, see Fig. 1. Several different steps and softwares were tried on test samples, and the best-performing software was used further. First, we predicted viral contigs from the scaffold (>1,500 bp) using Virify v0.4.0 (33) and geNomad v1.2.0 (<https://github.com/apcamargo/genomad>). Trimmed reads were mapped back to the viral contigs and then were binned using vRhyme v1.1.0 (34). The viral contigs were also checked with CheckV v1.0.1 (31), whose output was used for dRep v3.4.0 (35), removing bins with less than 50% completeness. The complete genomes (ITR/DTR/circular genomes) identified by vRhyme (34) and CheckV v1.0.1 (31) were extracted and added to the bins. Bins were checked for mismatched classifications [Diamond (29) against IMG/VR v4 (36) and NR, geNomad classification], and split if necessary. They were further split on the classification which we were able to designate with certainty, considering the percent of identity and CheckV statistics (contamination/warnings). Furthermore, the high-quality vMAGs/genomes from all samples were gathered, and CheckV v1.0.1 (31) with dRep v3.4.0 (35) was run again. This resulted in a non-redundant set of vMAGs that were at least 50% complete. They were classified against IMG/VR v4 (36) and the NR protein database (NCBI). All suspicious vMAGs (based on CheckV statistics and warnings; or with uncertain or disputable classification), with a focus on eukaryotic viruses, were checked again over the whole length of the sequence against the NR in BLAST at NCBI and manually curated. Most of these cases with difficult assignment/splitting were phages.

TABLE 1 The most important sequencing statistics of the samples^a.

Sample	Reads	% of remaining reads	% of assembled	% of mapped to host	% of viral reads	N50	N50/dehosted	Longest contig	Longest contig/dehosted
Lisnice11	41,440,100	74.99	95.88	57.95	25.31	2,126	1,972	44,207	44,207
Lisnice24	30,596,366	62.40	94.37	66.67	25.38	2,214	2,011	28,235	28,235
Lisnice333	22,914,070	78.46	99.16	19.03	88.86	1,675	1,047	8,256	8,255
Brdy1	55,271,986	73.82	92.85	40.76	35.67	2,021	2,125	103,615	43,060
Brdy2	32,128,740	71.95	92.56	45.01	35.10	2,429	2,279	49,872	44,479
Brdy3	30,099,480	49.37	89.81	28.58	54.30	2,839	2,483	54,474	32,004
Libechov11	39,834,524	69.30	92.84	37.24	28.49	2,218	2,186	64,021	63,918
Libechov14	18,261,622	70.76	96.92	90.57	6.63	2,908	2,845	31,936	31,936
Libechov6	18,149,762	74.32	94.54	44.83	28.76	2,126	2,115	59,551	47,373
Prasily1	13,634,908	82.14	99.59	6.61	99.00	1,325	1,406	5,956	5,956
Prasily2	15,098,038	81.19	99.61	8.39	95.61	3,632	3,720	30,652	30,652
Prasily3	45,293,424	78.05	98.48	49.74	43.63	1,984	1,967	42,461	42,461
VURV1	26,401,994	76.93	98.18	32.37	60.10	1,965	1,942	47,953	47,953
VURV5	33,526,926	75.66	97.59	33.31	68.74	2,085	1,879	18,154	15,974
VURV4	44,055,956	84.95	99.57	12.82	94.37	2,349	2,279	72,955	72,955
VURV7	27,161,514	84.68	99.71	15.45	87.60	1,703	1,669	7,761	7,742
VURV_H	18,125,694	78.71	97.11	29.24	66.73	1,707	1,652	22,873	27,430
VURV_D	11,303,362	87.06	99.88	11.30	99.26	4,987	4,846	5,359	5,358

^aNumber of remaining reads after trimming, number of those which mapped back at scaffold, to the host genome and to identified viral sequences. See full Table on [GitHub](https://github.com/kadlick/NAZV19) (<https://github.com/kadlick/NAZV19>). Statistics were done on contigs >1,000 bp.

Completing vMAG with additional sequencing

For the processing of AmFLV), the sequences of the three largest contigs were extended by amplicon sequencing. Primers were designed at both ends of the three contigs (available on GitHub, see Data Availability for link), and PCR was performed with primer combinations using Phusion polymerase (Thermo Fisher, Waltham, Massachusetts, USA) according to manufacturer's protocol with three different extension times (2.5, 5, and 10 min). The resulting amplicons were purified using MSB SPIN PCRAPACE (Invitex Molecular, Berlin, Berlin, Germany). The concentrations were measured using Qubit dsDNA HS Assay kit (Thermo Fisher, Waltham, Massachusetts, USA). The library was prepared using Nextera XT kit (see NetoVIR protocol) with 5-min tagmentation, sequenced, and analyzed as described above.

Determination of the complete nudivirus genome

First, we identified a few relatively short contigs which could be attributed to nudiviruses, the longest contig of length 16,645 bp was obtained from one sample (Lisnice11). To enhanced detection of DNA, the sample was sequenced again without WTA2 preamplification which includes reverse transcription. One sample contained more than 50,000 reads attributable to *Nudiviridae* contigs (Lisnice24). All obtained reads from three samples were co-assembled with setting of SPAdes as described above. Then, we mapped individual samples on the scaffold, used vRhyme to get bins, extended them from the assembly graph with binSPReader, pre-release (37). The extended nudivirus bin was checked with Virsorter2 v2.2.4 (38) and DRAM v1.4.6 (39) to ensure that all contigs we have gained had a majority of *Nudiviridae* alignments. Then, we tried to get cleaner reads to resolve some regions. We mapped the reads of these three samples on the bin of four contigs undoubtedly belonging to *Nudiviridae* and extracted mapped reads (with a mate in case of one read from pair mapping). We repeated the mapping but on scaffold (>1 kbp) gained in co-assembly and extracted unmapped reads. The reads were from the three samples, mapped on nudivirus and unmapped on anything else (scaffold from co-assembly of three samples, greater than 1,000 bp). With these reads, we tried assembly, the best was obtained with SPAdes v3.15.5, --metaviral -k 21, 33, 55, and 77 settings which resulted in three contigs, 76,482, 24,683, and 22,709 bp (= 123,874 bp).

Therefore, we design primers similarly as for AmFLV (available at GitHub, see Data Availability for link), on the ends of contigs aiming outward. We performed PCR with different combinations of primers with Phusion polymerase and using different elongation times. The obtained fragments were purified and the library using Nextera XT (tagmentation based on lengths of amplicons: 2, 3, and 7 min) was prepared and sequenced on MiSeq with reads 2×250 bp.

Comparison with published data sets

After finding possible new viruses in our *de novo* data, we analyzed sequences obtained by NGS from studies which used protocol for sample processing that allows detection of both RNA and DNA viruses (40). FastQ files obtained in the study by Deboutte et al. (41) were pulled from the Sequence Read Archive (SRA) archive using prefetch and fastq-dump available from NCBI, and all non-redundant scaffolds from all the samples were made available on GitHub. Additionally, we used data from our previous study (24). The non-redundant scaffolds from both studies were classified with Diamond blastx against NR with --sensitive and -c 1.

The viral set we created for further analyses consisted of novel viral genomes detected in the current study combined with viruses found in the previous studies (24, 41) as described above. The viral set contains all the vMAGs/genomes. Reads from the studies were mapped on them using bwa-mem2 v2.2.1 (32), and the coverage was extracted using CoverM v0.6.1 (<https://github.com/wwood/CoverM>).

Phylogeny, visualization, annotations, and data processing

We gathered NS1 proteins of *Densovirinae* available in the database (NCBI, 06.16.2022), and the data set was filtered several times for partial and misnamed proteins. NS1-superfamily region was identified using Batch CD-Search (42) and extracted in Python. For AmFLV, the sequences of DNA polymerases were randomly selected from each group of large viruses (<https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>). For the described AmNV, the core genes of 13 nudiviruses were extracted (DNA polymerase B, DNA helicase, integrase, p47, lef-4, lef-8, lef-9, vlf-1, p74, pif-1, pif-2, pif-3, pif-4/19 kda, 38k, vp39, vp91, ac81); for outgroup, we used *Baculovirus* with the same set of genes but without integrase (see <https://github.com/kadlck/NAZV19> for accessions) as described in reference (26).

Alignment was done with mafft v7.520 --auto (43), then it was trimmed with trimAl v1.4.1 --gappyout (44). The best model was determined with ModelTest-NG v0.1.7 (45), and the tree was built using Phyml v3.3.20220408 (46) using the model best suited for our data set. Visualization was done using iTOL (47).

Annotations of densovirus, *Apis mellifera* filamentous-like virus, and *Apis mellifera* nudivirus were performed using BLAST in NCBI (48) and DRAM-v v1.4.6 (39). Schemes of the genomes were generated using the Python library DNA Features Viewer v3.1.2 (49) and pyCirclize v0.5.0 (<https://github.com/moshi4/pyCirclize>), and other libraries were used to process data (Pandas, Numpy, Biopython). The terminal repeats were predicted by RNA-fold as specified in reference (50). The genomes of AmFLV and AmNV were polished using Pilon v1.24 (51).

RESULTS

Sequencing statistics

The sequencing statistics are listed in Table 1. We gained more than 10M of reads per sample. Also, a large number of reads mapped back to our curated vMAGs, showing efficiency of our protocol.

Parvoviridae

NGS of 18 analysed samples resulted in 77,423 reads which belong to the *Parvoviridae* family. This presents approximately 0.02% of all reads and 0.03% of those determined to be viral reads. These reads were present in 7 out of 18 samples. In the analyzed sequences from three studies from which we gained contigs, we found 16 unique contigs corresponding to the subfamily *Densovirinae*: nine were complete and seven were incomplete genomes. The complete genomes were named Bee densovirus 1 to 9. The lengths of the complete genomes ranged from 3.6 to 6 kbp. All complete genomes contained two major ORFs that code structural and non-structural proteins with *Parvoviridae*-specific motifs and other short ORFs encoding additional hypothetical proteins surrounded by non-coding regions at both ends (Table 2; Fig. 2). The number of proteins and their predicted function based on similarity are shown in Table 3.

Terminal repeats (TRs) were present in five of the nine complete genomes (Bee densovirus 1, 2, 3, 6, 9). Detailed information and predicted structure of TRs are shown in Table 2; Fig. 2 in the file with additional densovirus figures at <https://github.com/kadlck/NAZV19>. The TRs were completely identical for Bee densovirus 2 and 6 or had 1–2 bp mismatches for the other Bee densovirus. The rest of the complete genomes lack terminal repetition; however, the ORFs are flanked by non-coding regions in length ranging from 14 bp (Bee densovirus 4) to 414 bp (Bee densovirus 7) at the 5' end of the genomes and from 104 bp (Bee densovirus 4) to 229 bp (Bee densovirus 5) at the 3' ends.

Incomplete genomes (3.7–4.5 kbp) contain neither non-coding regions nor complete ORFs. They were therefore deemed incomplete and are available on GitHub (see Materials and Methods for link).

The bee densoviral genomes were highly variable; they differed in length and terminal repetitions, and predicted ORFs. The sequence similarity of the genomes ranged

TABLE 2 Detailed information about the TRs and ends of the densoviruses^a

Virus	TR length	TR type	TR mismatches (bp)	5' Flank (before TR)	3' Flank (after TR)
Bee densovirus 1	232	Homotelomeric	1	38	38
Bee densovirus 2	155	Homotelomeric	0	14	0
Bee densovirus 3	76	Homotelomeric	2	1	0
Bee densovirus 4	NP	NP	NP	14	104
Bee densovirus 5	NP	NP	NP	323	229
Bee densovirus 6	77	Homotelomeric	0	0	0
Bee densovirus 7	NP	NP	NP	414	223
Bee densovirus 8	NP	NP	NP	251	126
Bee densovirus 9	169	Homotelomeric	1	66	0

^aNP, not present.

from 24% to 44% across the whole genome (Table 2 in file with additional densoviruses figures at <https://github.com/kadlck/NAZV19>). This variability was confirmed by phylogenetic analysis using sequences retrieved from NCBI (Fig. 3 in file with additional densoviruses figures at <https://github.com/kadlck/NAZV19>). The genomes were distributed through the phylogenetic tree based on NS1-superfamily region-conserved sequences of densoviruses retrieved from NCBI. Bee densovirus 4 was closest to a member of *Miniambidensovirus*, *Acheta domestica* mini ambidensovirus (37.0% similarity over the whole sequence), and Bee densovirus 7 to *Scindoambidensovirus* but also to members from the family *Densoviridae* (previously subfamily *Ambidensovirus*) that lack recent classification. Bee densovirus 8 was closest to sequences belonging to *Atrato Denso*-like viruses and Broome densovirus (31.2%, 30.3%, 31.2%). The group which included Bee densoviruses 1 and 5 had the highest similarity to *Ambidensovirus* sp., 40.6% for Bee densovirus 1 and 45.6% for Bee densovirus 5. Bee densovirus 3 was closest to *Tarsiger cyanurus* ambidensovirus with a similarity of 36.6% over the whole sequence. Bat-associated densovirus had a 50.6% similarity to Bee densovirus 7. Bee densovirus 6 had the highest similarity with 75.7% to *Periparus ater* ambidensovirus. For Bee densovirus 2, the similarity to *Phylloscopus inornatus* ambidensovirus was 59.9%. And finally, Bee densovirus 9 had a similarity of 46.7% to the *Ambidensovirus* sp.

Apis mellifera filamentous-like virus (AmFLV)

From *de novo* sequencing, we identified one contig 103,615 bp long with low similarity to AmFV polymerase (30.8%). The retrieved sequence was extended as a vMAG with five more contigs (77,565 bp, 43,940 bp, 8,006 bp, 5,756 bp, and 4,443 bp), and CheckV predicted the vMAG as 92.9% complete. A ~10 kb amplicon was obtained using PCR with specific primers designed at the end of the large contigs (see GitHub, <https://github.com/kadlck/NAZV19>, for the scheme). After sequencing the amplicon, assembly, and polishing, a linear genome with TRs was obtained. Out of 39,950,966 reads, 448,034 reads (1.1% of the sample where contigs were discovered, region Brdy, specifically Brdy1) mapped to the complete genome. The complete genome length, provisionally named AmF-like virus, was 152,678 bp long with a GC content of 49.8%. A scheme of the genome is shown in Fig. 3A. The sequence of AmFLV was flanked by inverted homologous TRs 77 bp long, forming a Y-shape at both ends (Fig. 3B). A total of 112 ORFs were identified; the putative ORFs were distributed on both strands, and 49 (43.8%) were identified in protein databases (Viral, Peptidase, Pfam, Cazy, Vogdb, KEEG).

The predicted proteins mostly showed protein similarity to AmFV (similarity from 16.8% to 51.5%), but also other large viruses (see GitHub, <https://github.com/kadlck/NAZV19>, for all hits): protein numbered AmFLV_89 to trimeric dUTPase of *Vombatid gammaherpesvirus 1* (55.1%), protein AmFLV_76 to orf66 gene product of *Helicoverpa zea* nudivirus 2 (58.4%), protein AmFLV_2 to inhibitor of apoptosis (iap) of *Trichoplusia ni* single nucleopolyhedrovirus (43.0%), AmFLV_34 to ribonucleotide reductase/HP APL35_gp114 of AmFV (50.4%), and protein AmFLV_74 to putative endonuclease of

TABLE 3 BLASTp results for complete densoviruses genome proteins^a

ORF	Length (bp)	Length (aa)	Motif	Hit	Query coverage	E value	Percent identity	Accession
Bee densovirus 1 (4,865 bp) GC: 44.8%								
HP3	489	163	No hit					
SP	2,253	751		Structural protein (Tarsiger cyanurus ambidensovirus)	76%	2×10^{-96}	38.6%	QTE04081.1
HP2	651	217	No hit					
NS1	1,554	518	NS1 motif	Non-structural ORFs (<i>Ambidensovirus</i> sp.)	88%	3×10^{-68}	31.6%	AWV66973.1
HP1	723	241	No hit					
Bee densovirus 2 (5,601 bp) GC:39.4%								
NS1	1,635	545	NS1 motif	Non-structural protein 1 (<i>Phylloscopus inornatus</i> 99% ambidensovirus)	0		75.5%	QVW56790.1
SP-B	1,689	563	Denso_VP4	Structural protein VP1 (<i>Phylloscopus inornatus</i> ambidensovirus)	83%	0	70.8%	QVW56791.1
SP-A	960	320	Nterminal region of VP1 coat protein	Structural protein VP1 (<i>Phylloscopus inornatus</i> ambidensovirus)	81%	1×10^{-53}	42.3%	QVW56792.1
NS2	819	273		Non-structural protein NS-2 (<i>Culex pipiens</i> densovirus)	100%	1×10^{-97}	54.0%	YP_002887626.1
HP1	438	146	No hit					
HP2	414	138	No hit					
Bee densovirus 3 (3,632 bp) GC: 40.2%								
NS1	1,020	340	NS1 motif	TPA: MAG TPA: Rep 40 protein helicase (<i>Parvoviridae</i> sp.)	96%	3×10^{-175}	69.1%	DAN51445.1
SP	1,611	537	Nterminal region of VP1 coat protein	TPA: MAG TPA: capsid protein (<i>Parvoviridae</i> sp.)	94%	4×10^{-142}	45.2%	DAN51446.1
NS	486	162		TPA: MAG TPA: Rep 40 protein helicase (<i>Parvoviridae</i> sp.)	54%	4×10^{-17}	51.1%	DAN51445.1
Bee densovirus 4 (5,160 bp) GC:41.8%								
SP	2,823	941	Nterminal region of VP1 coat protein	VP (uncultured densovirus)	60%	4×10^{-33}	24.9%	QOD39535.1
NS1	1,533	511	NS1 motif	Putative non-structural protein (<i>Phylloscopus schwarzi parvoviridae</i> sp.)	82%	1×10^{-40}	26.8%	QTE04075.1
HP2	1,023	341	No hit					
HP4	423	141	No hit					
HP3	498	166	No hit					
HP1	423	141	No hit					
Bee densovirus 5 (3,910 bp) GC: 40.2%								
SP	1,728	576		Structural protein (<i>Phylloscopus ambidensovirus</i>)	95%	6×10^{-165}	48.7%	QTE03896.1
NS1	1,542	514	NS1 motif	Non-structural protein (Tarsiger cyanurus ambidensovirus)	85%	4×10^{-83}	36.0%	QTE04079.1
HP	669	223	No hit					
Bee densovirus 6 (5,058 bp) GC: 42.0%								
SP-A	1,821	607	Denso_VP4	Structural protein VP2 (Tarsiger cyanurus <i>parvoviridae</i> sp.)	99%	0	93.7%	QVW56816.1
NS1	1,599	533	NS1 motif	Non-structural protein (bat-associated densovirus)	100%	0	91.9%	QOR29557.1
SP-B	741	247		Structural protein (Tarsiger cyanurus <i>parvoviridae</i> sp.)	100%	3×10^{-164}	91.1%	QVW56817.1
HP2	417	139	No hit					

(Continued on next page)

TABLE 3 BLASTp results for complete densoviruses genome proteins^a (Continued)

ORF	Length (bp)	Length (aa)	Motif	Hit	Query coverage	E value	Percent identity	Accession
HP1	576	192	No hit					
NS2	789	263		Non-structural protein NS-2 (<i>Blattella germanica</i> densovirus 1)	95%	6×10^{-117}	65.7%	NP_874382.1
NS	693	231		Non-structural protein (<i>Nandayus nenday parvoviridae</i> sp.)	100%	6×10^{-155}	90.4%	QTE03739.1
Bee densovirus 7 (4,841 bp) GC: 38.3%								
NS2	855	285		Non-structural protein 2 (<i>Grus japonensis parvoviridae</i> sp.)	74%	2×10^{-141}	94.3%	QTE03769.1
NS1	1,680	560	NS1 motif	Non-structural protein (bat-associated densovirus)	95%	0	57.5%	QOR29553.1
SP-B	765	255		Putative structural protein (<i>Grus japonensis parvoviridae</i> sp.)	96%	8×10^{-45}	41.7%	QTZ83145.1
SP-A	1,722	574	Denso_VP4	VP1 (bat-associated densovirus)	99%	1×10^{-172}	47.3%	QOR29554.1
Bee densovirus 8 (6,023 bp) GC: 41.5%								
HP1	708	263	No hit					
NS1	1,293	431	NS1 motif	Putative non-structural protein (<i>Ambidensovirus</i> sp.)	63%	6×10^{-27}	29.6%	UGV24202.1
HP2	777	259	No hit					
HP3	543	181	No hit					
SP	2,862	954		Capsid protein (<i>Emberiza spodocephala ambidensovirus</i>)	24%	3×10^{-6}	25.3%	QTE04116.1
Bee densovirus 9 (5,202 bp) GC: 40.4%								
HP1	597	199	No hit					
NS1	1,509	503	NS1 motif	Non-structural protein NS1 (<i>Densovirinae</i> sp.)	94%	5×10^{-179}	52.1%	QJ153739.1
NS2	798	266		Non-structural protein NS-2 (<i>Culex pipiens</i> densovirus)	92%	9×10^{-52}	40.0%	YP_002887626.1
HP2	513	171	No hit					
SP	2,553	851	Denso_VP4	Viral polypeptide VP1 (<i>Diatraea saccharalis</i> densovirus)	72%	6×10^{-157}	45.1%	NP_046815.1

^aHP, hypothetical protein; NS, non-structural protein; SP, structural protein.

Emiliania huxleyi virus 86 (42.9%) and AmFLV_87 to DNA polymerase of AmFV (28.4%). Like other large viruses, this virus encodes its own DNA polymerase and proteins like dUTPase and metalloproteinase that can affect host cell metabolism. Additionally, this virus codes proteins like inhibitors of apoptosis or *per os* infectivity factors. On average, the identity of proteins with known function was higher than that of hypothetical proteins, and even higher for proteins that affect cell metabolism. However, even with several significant alignments (mostly to AmFV), most of the proteins are hypothetical or have no significant alignments detected.

Some characteristics of the new virus are similar to AmFV, like the presence of the *Baculoviridae*-related regions (pif-1/2/3), which are important for cell entry and are essential for *per os* infection. The similarities were relatively high (41.7%, 51.5%, and 41.7%) and we found pif 1–3 which form the conserved *per os* infectivity complex in *Baculoviridae* (52) or the presence of the kinesin motor domain which could be one of the components responsible for affecting cytoskeletal dynamics by viral infection. It may be important that the virus has significant similarity to one hypothetical protein of AmFV (APL35_gp042, 42.4%), which could encode integrase/recombinase closest to the phage integrase family (Pfam, Vogdb). The identified AmFLV virus was also detected in our previous study (24) and in the study by Deboutte et al. (41) (Table 4).

The phylogenetic analysis was done with a representative of large viruses (from NCBI, random selection from RefSeq DNA polymerases sequences in each family) and shows

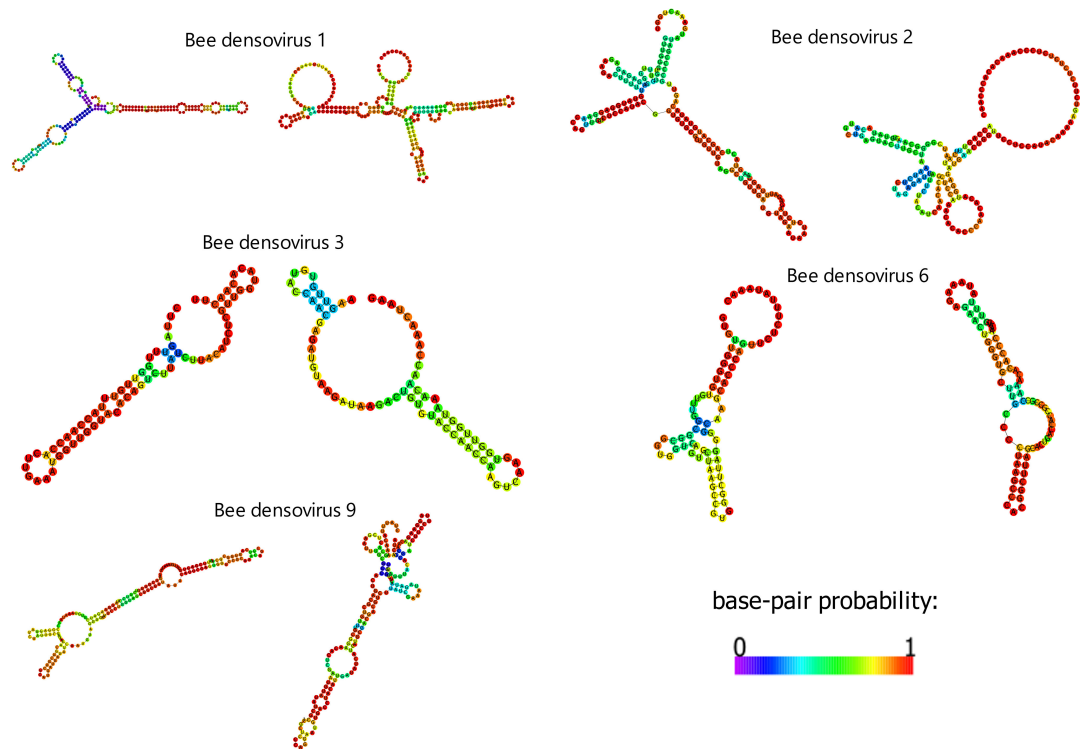


FIG 2 Predicted structures of terminal repeat from described densovirus (RNA-fold program) as in reference (50). We can see familiar structures like Y-shape or I-shape. The color scale of pair-base probability is shown.

that the AmFV, based on sequences of polymerases, is the closest representative, but still clearly distinct (Fig. 3C).

The vMAG was made up mainly of three large contigs (103, 77, and 44 kbp), but only the 103 and 44 kbp contigs were mapped to the AmFLV. All three contigs have very similar mapping patterns (Table 4) across the samples where the contigs were present. The sequence of the third 77 kbp contig (made available on GitHub, <https://github.com/kadlck/NAZV19>) probably belongs to another large virus that infects honey bees. The inclusion of the 77 kbp contig in vMAG isn't unexpected. The probability that the 77 kbp contig doesn't contain a sequence of AmFLV is supported by the finding of a higher number of reads mapping to the shorter 77 kbp fragment in comparison to the 103 kbp contig and only two similarities to AmFV detected in the 77 kbp contig in comparison to dozens in the 103 and 44 kbp contigs (Table 4). Additionally, when we include the 77 kbp contig into vMAG, the two genes of ribonucleoside-diphosphate reductase were identified.

Apis mellifera nudivirus (AmNV)

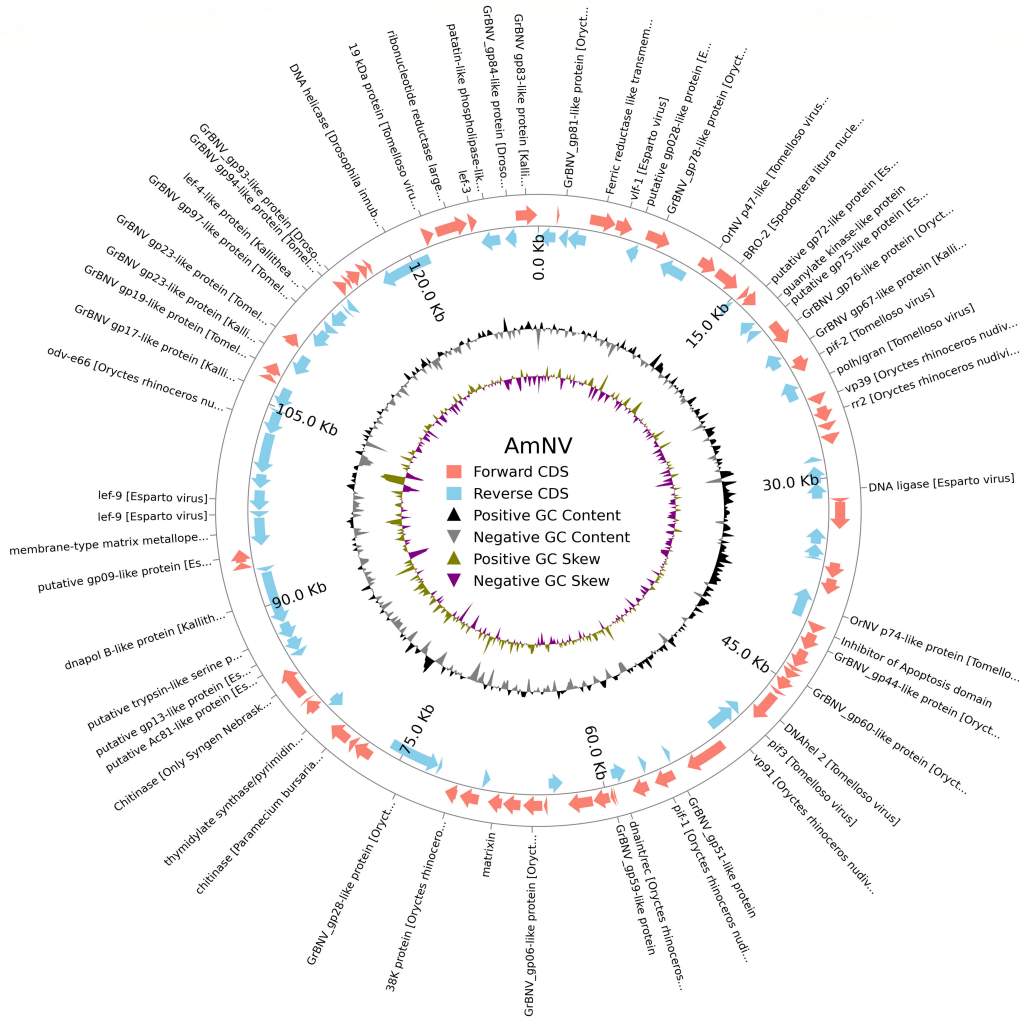
At the beginning, we had 12 contigs which we reduced to three contigs by co-assembly (see Materials and Methods), 76,482, 24,683, and 22,709 bp (= 123,874 bp) with high reliability and continuity. They all had a number of hits to *Nudiviridae*. With primers that we designed at the ends of these contigs, and directed outward utilizing PCR, we were able to obtain three amplicons (<500 bp, <900 bp, <10 kbp) in reasonable combinations which make the virus circular and complete (see GitHub, <https://github.com/kadlck/NAZV19>, for the scheme). These amplicons were sequenced and the genome of the new nudivirus found in honey bees was completed. We gained a circular genome of 129,467 bp, coding 106 proteins on both strands of the viral genome. The virus had 40.3% GC content, was named *Apis mellifera nudivirus*, and taxonomically belongs to *Alphanudivirus* (Fig. 4). On the virus, mapped 219,147 reads out of 31,076,599 (0.7%)

TABLE 4 Mapping of different samples from analyzed studies on three contigs belonging into one vMAG and to the whole AmFLV sequence completed by analyses of NGS available sequences and PCR^a

Current study	Brdy1	Brdy2	Brdy3	Libechov11	Libechov6
AmFLV	448,034	213,635	20,568	2,113	1,851
NODE_1_length_103615_cov_192.460671	300,031	141,132	13,644	968	1,060
NODE_2_length_77565_cov_239.767590	552,407	252,903	23,319	2,608	2,195
NODE_4_length_43940_cov_193.303262	126,097	58,462	4,910	533	516
Kadlečková et al. (24)	BeeCelsaA2	BeeCelsaA3	BeeCelsaA4	BeeCelsaA5	BeeCelsaA6
AmFLV	1,606	7,663	715	2,214	45,358
NODE_1_length_103615_cov_192.460671	866	4,721	409	1,345	26,226
NODE_2_length_77565_cov_239.767590	792	5,252	542	1,362	22,194
NODE_4_length_43940_cov_193.303262	390	2,158	209	635	12,685
Deboutte et. al. (41)	SRR10418323	SRR10418338	SRR10418344	SRR10418371	SRR10418385
AmFLV	1,072	2,481	534	770	13,731
NODE_1_length_103615_cov_192.460671	286	1,649	325	175	8,797
NODE_2_length_77565_cov_239.767590	464	2,755	594	523	17,569
NODE_4_length_43940_cov_193.303262	192	726	134	175	3,861

^aThe number of reads mapping is shown.

A)



B)

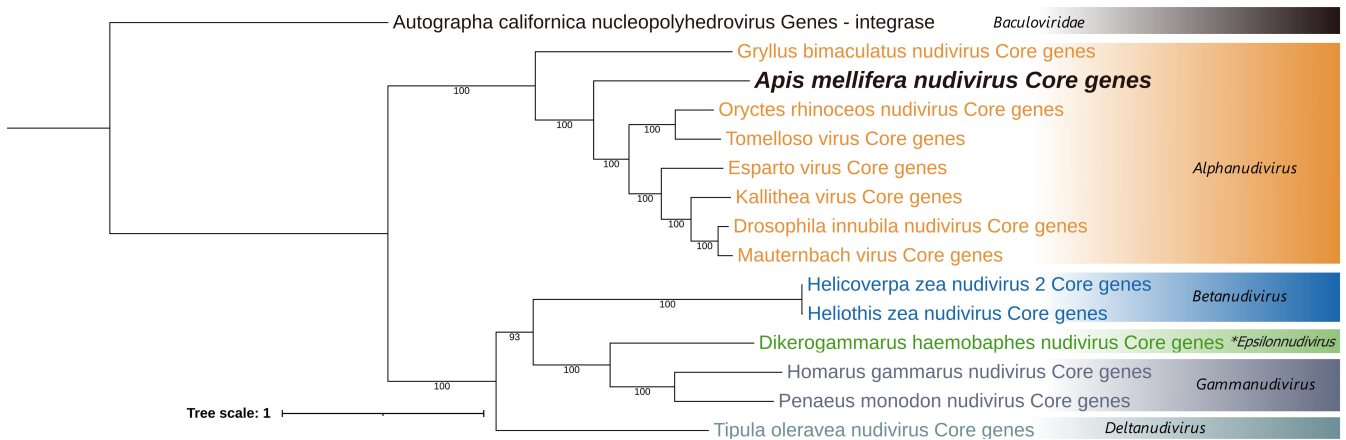


FIG 4 Information about new honey bee nudivirus. (A) Scheme of the genome. (B) Taxonomy of core genes * proposed genus (53).

Lisnice 11 from sample sequenced using the full NetovIR protocol (RNA/DNA preparation), 110,763 out of 13,409,876 (0.8%) Lisnice11 with only DNA sequenced (no WTA2 preamplification), and 67,247 out of 19,091,873 (0.4%) Lisnice24. The virus was also detected in other samples, but the number of mapping reads was below 50,000 reads

and was also detected in studies which analyzed viromes of Czech (24) and Belgian bees (41).

Out of 106 proteins, 79 had some significant alignments (74.5%), mainly to four viruses: *Oryctes rhinoceros nudivirus*, Esparto virus, *Kallithea virus*, and Tomelloso virus which are all *Alphanudivirus*. The other significant alignments were AmNV_10 to BRO-2 of *Spodoptera litura nucleopolyhedrovirus II* (11.8%), AmNV_39 to inhibitor of apoptosis 3 of *Choristoneura rosaceana entomopoxvirus "L"* (27.8%), AmNV_64 to hypothetical protein APL35_gp193 of *Apis mellifera filamentous virus* (47.1%), AmNV_66 to chitinase of *Paramecium bursaria Chlorella virus NYs1* (23.2%), another AmNV_70 to chitinase of *Only-Syngen Nebraska virus 5* (26.6%), and AmNV_78 to membrane-type matrix metalloproteinase-1 of *Anopheles minimus iridovirus* (27.2%). All annotations are available on GitHub (see Materials and Methods for link), and a schematic of the genome is shown in Fig. 4.

The AmNV contains significant alignments to 17 core genes of the *Nudiviridae* family: AmNV_74 to DNA polymerase to *Kallithea virus* (42.9%), AmNV_49 integrase/recombinase of *Oryctes rhinoceros nudivirus* (51.9%), AmNV_41 to helicase 2 of Tomelloso virus (40.2%), AmNV_92 lef-4 (subunit DdRp) *Kallithea virus* (45.6%), AmFV_44 to lef-8 (catalytic subunit DdRp) of *Kallithea virus* (58.9%), AmNV_80 lef-9 (subunit DdRp) Esparto virus (51.9%), AmNV_46 to pif-1 *Oryctes rhinoceros nudivirus* (49.3%), AmNV_17 to pif-2 Tomelloso virus (68.3%), AmNV_42 to pif-3 of Tomelloso virus (58.2%), AmNV_5 to vlf-1 of Esparto virus (41.9%), AmNV_19 vp39 *Oryctes rhinoceros nudivirus* (53.7%), AmNV_33 to p74 of Tomelloso virus (50.6%), AmNV_43 vp91 of *Oryctes rhinoceros nudivirus* (36.1%), AmNV_61 to 38K *Oryctes rhinoceros nudivirus* (47.7%), AmNV_101 19 kDa protein Tomelloso virus (65.2%), and AmNV_71 to Ac81 protein of Esparto virus (49.5%). These core genes had, on average, greater similarity than the other significant alignments. We did not find any similarity for 4 out of 29 core genes (lef-5, p6.9, fen1, and 11K-like protein).

The virus contains a complete core of *per os* infectivity complex (pif-1/2/3) and all genes we expect to see in a member of *Nudiviridae* (core genes), mainly DNA polymerase B, integrase/recombinase, DNA-directed RNA polymerase, and DNA helicase 2. Apart from *Nudiviridae* proteins, AmNV has other significant alignments, which might be specific for honey bees (like alignment to APL35_gp193 of AmFV).

DISCUSSION

Historically, honey bees have been thought to be primarily associated with a plethora of RNA viruses, belonging to the order *Picornavirales*, together with the most common members of *Iflaviridae* and *Dicistroviridae*. In stark contrast, only a handful of DNA viruses have been documented in this managed pollinator (18, 19). Our analysis in this study of 18 NGS samples, each representing a pool of 50 bees, together with sequences from two other studies (24, 41), has significantly expanded this DNA virome landscape that may be present in different honey bee populations. In our previous article, we suggested to use a number of pooled bees per sample, because this approach should allow to detect the diversity of honey bee viruses in the colony, and also low prevalence viruses (i.e., those present in only part of bees in a colony). Even if the viruses are in low abundance in the pooled sample, they can be used for *de novo* genome assembly (24). *de novo* assembly and the subsequent generation of vMAGs from the sequences performed in this study revealed several new DNA viruses. The benefit of analyzing larger numbers of bees can be documented by the detection of low prevalence viruses and proves to be beneficial in reflecting the diversity of DNA viruses in honey bees.

Most of the new DNA viruses could be classified as members of the *Densovirinae* subfamily. Densoviruses have a classical structure, two coding ORF cassettes and untranslated regions at the ends, with five out of nine having terminal repeats while others lack them. Not all known members of *Densovirinae* have terminal repetition at their genome ends (50). Our constructed phylogenetic tree was based on the conserved domain of NS1 even though the most described genera of *Densovirinae*

clustered together, group previously classified as *Ambidensovirus*, and encompassing all the new -ambidensoviruses, were distributed through the tree with low bootstrap support. Members of the group are known to have even less than 30% similarity within the genus (22). The latest revision of the *Parvoviridae* taxonomy split the family into seven more lineages, some of which having greater similarity to *Iteradensovirus* than to other members of the group previously classified as *Ambidensovirus* (22).

The impact of *Densovirinae* on arthropods varies, ranging from overt pathologies (20, 54) to mutualistic relationships (55). Most members of the subfamily *Densovirinae* cause lethal infection of their hosts. The first symptoms are often anorexia and lethargy, followed by flaccidity, progressive paralysis, slow melanization, and tumor development (20). With a well-defined and distinguishable set of symptoms and a high viral load in a sick or dead animal, these viruses were relatively easy to identify before the era of NGS. Our bees that were selected for the analysis showed no overt signs of infection, such as deformities. To determine whether the newly described viruses are truly asymptomatic will require further study. The large number of found densoviruses with a low degree of nucleic acid similarities (23.9% to 44.4%) in two European countries (Czech Republic, Belgium) may indicate the presence of these and similar viruses also in other countries. The diversity of *Densovirinae* seems to be steadily increasing and new genera are being identified within the viruses previously known as ambidensoviruses. This high variability may not be limited to honey bee viruses like the *Densovirinae*. For instance, a recent study identified nine complete and six incomplete new species of chaphamaparvoviruses in six chickens (56).

After the initial analysis of the obtained sequences, we found a long fragment with low identities to AmFV. With viral binning, we found possible other assembly fragments of the new DNA virus. The virus AmFLV was completed using PCR that connected two contigs in one linear viral genome with TRs and the length over a 152 kbp. The terminal repeats form Y-structures at the ends of the genome. It contains a wide range of ORFs encoding a range of proteins from polymerase to those that affect the host cytoskeleton. Notably, the newly described AmFLV contains all three pifs (1, 2, 3) which have been documented as the core of the *per os* infectivity complex in baculoviruses (52). Apart from that, of interest might be the homology of the viral sequence with the phage integrase family and probable integrase/recombinase. This protein is also present in AmFV but denoted as hypothetical. It is possible, that under certain conditions, the virus is able to utilize this protein for integration into the host genome. The protein has been described in some large viruses, for example, in *Nudiviridae*. Therefore, it could be involved in establishing of latency and integration into the host genome (26). Overall, AmFLV encodes 112 ORFs, but the function of most of them remains undescribed (hypothetical), even for those that we were able to find similarity, mainly to AmFV.

The remaining long contig has a very similar mapping pattern across all samples where the contigs were present. The contig is approximately 77 kbp in size. However, we didn't succeed in gaining its complete genome. It is also possible that this large fragment represents part of the genome of another large honey bee virus. Similar large DNA viruses with the same mapping patterns could easily be sorted as part of the vMAG of interest; therefore, the generation of vMAGs is extremely helpful in gaining genomic sequences but should be confirmed by a wet laboratory approach for some uses.

We completed the genome of nudivirus of honey bees using a combination of vMAGs generation, co-assembly, and amplicon sequencing. We gained a circular genome 129,467 bp long that contains the core genes of the *Nudiviridae*. In particular, lef-4, lef-5, lef-8, lef-9, and p47 are important for transcription, whereas pif-1-3, pif-4/19 kDa, and p74 are necessary for infection. There are several proteins important for viral morphogenesis, 38K, vp91, vlf-1, ac81, and vp39. The core genes also include proteins necessary for replication/recombination and repair like DNAPol-B, helicase, and integrase. There are other important genes involved in nucleotide metabolism and some with unknown function (26). It is not surprising that the virus lacks 4 of the 29 suggested core genes (26) since, until present, only few nudiviruses have been described and the list of core

genes is still changing. The virus encompasses a total of 108 ORFs, with significant alignments of predicted proteins to *Nudiviridae*, more precisely to *Alphanudivirus* (36.1% to 68.3%). The virus contains some proteins with reliable alignments to other viruses, like one alignment to AmFV, nucleopolyhedrovirus, or entomopoxvirus, but the similarities are low (11.8% to 47.1%). Phylogenetic analysis also revealed that the virus belongs into *Alphanudivirus*. The pathologies of *Nudiviridae* vary from virus to virus and between host's life stages. In insects, they can cause lethargy, weakness, malformations, stunted growth, reduced longevity, or fertility. They can also cause changes in the viscosity and color of the hemolymph (opalescent), or they can concentrate in the abdomen and form a "waxy plug" (26). However, further studies are needed to find out if a specific pathology linked to AmNV exists. Generally, the infection with these viruses is less symptomatic in comparison to *Baculoviridae*, and they seem to be restricted to certain cell types (26).

With the exception of two colonies showing varroosis symptoms, all other colonies used as source colonies for sampling were considered healthy, i.e., they built up quickly, had no damaged capping and no signs of overt disease, or *Varroa* infestation were observed. However, all the bees we selected for the virome analysis were free of malformation and showed no signs of overt pathology. However, it is not possible to completely exclude the possibility that individual bees may have signs of pathology that were not obvious and visible when the individual bees were selected prior to virome analysis. Further studies with different experimental design and sampling are needed to clarify this issue.

Even though the binning of sequencing data is regularly performed for bacterial and eukaryotic species, the binning and generation of MAGs is still a novel method for viruses. Only in recent years, new software that is designed for binning of viral sequences has been released. The tools like Coconet (57), vRhyme (34), and vamb (58) perform better for creating viral MAGs since they provide a larger number of cleaner bins with fewer misclassified contigs in comparison to other programs that are better suited for creating bacterial and eukaryotic MAGs (34, 57, 58). This was confirmed by testing on a limited number of samples as we tried MetaWRAP (59) and all the viral binning software mentioned above. For our data set, vRhyme (34) performed best, but it would be beneficial to have a bin refinement tool such as MetaWRAP (59) with a combination of different binning software for viruses. The vMAGs generated were instrumental in piecing together a significant portion of a new large viral genomes, which otherwise might have been overlooked since its similarities to large viruses affect the statistics (e.g., in CheckV) and are deemed very incomplete. The generation of vMAGs then allows for keeping maximum information while still filtering out very incomplete and poor fragments. It allows us to "connect" the contigs and treat them as fragmented genomes.

In conclusion, our study underscores the richness of the honey bee DNA virome, which was previously overshadowed by their RNA counterparts. The plethora of densoviruses identified, coupled with the discovery of the AmFLV with its predominantly hypothetical proteins, and first nudivirus in honey bee (AmNV), paves the way for deeper investigations into the ecological and pathological implications of these viruses in bee populations.

ACKNOWLEDGMENTS

We thank Milan Šimonovský for collecting samples and other beekeepers for allowing us access to their apiaries, and Martin Markovič for helping us with correcting the article citations. And we also thank the core facility OMICS-Genomics in BIOCEV run by Štěpánka Hrdá for the rapid processing of our samples.

This research was supported by the Ministry of Agriculture of the Czech Republic, Project No. QK1910018 (NAZV) and MZE-RO0423, by the National Institute of Virology and Bacteriology Project (EXCELES Program, ID Project No. LX22NPO5103), funded by the European Union, Next Generation EU, by Charles University Grant No. SVV 260679, and by the European Union, EU4Health Program (EU4H), Project No. 101102733 DURABLE. However, the views and opinions expressed are those of the authors and do not

necessarily reflect those of the European Union or the European Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

AUTHOR AFFILIATIONS

¹Department of Genetics and Microbiology, Faculty of Science BIOCEV, Charles University, Vestec, Průmyslová, Czechia

²Crop Research Institute, Drnovská, Prague, Czechia

AUTHOR ORCIDs

Dominika Kadlečková  <http://orcid.org/0000-0002-4312-7373>

Martina Saláková  <http://orcid.org/0000-0003-0827-1211>

Tomáš Erban  <http://orcid.org/0000-0003-1730-779X>

Ruth Tachezy  <http://orcid.org/0000-0001-7689-9727>

FUNDING

Funder	Grant(s)	Author(s)
Ministerstvo Zemědělství (Ministry of Agriculture)	QK1910018; MZE-RO0423	Tomáš Erban
Ministerstvo Školství, Mládeže a Tělovýchovy (MŠMT)	LX22NPO5103	Ruth Tachezy Martina Saláková Dominika Kadlečková
Charles University	SVV 260679	Dominika Kadlečková
EU4H Project No.101102733 (DURABLE)	101102733	Ruth Tachezy Martina Saláková Dominika Kadlečková

AUTHOR CONTRIBUTIONS

Dominika Kadlečková, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft | Martina Saláková, Conceptualization, Methodology, Supervision, Writing – review and editing | Tomáš Erban, Conceptualization, Funding acquisition, Project administration, Resources, Writing – review and editing | Ruth Tachezy, Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review and editing

DATA AVAILABILITY

The raw data of the current study were deposited in SRA under BioProject id [PRJNA1008242](https://doi.org/10.1093/bioinformatics/btad001) and the described densovirus under accessions nos. [OR553295-OR553303](https://doi.org/10.1093/bioinformatics/btad001), AmFLV under [OR553294](https://doi.org/10.1093/bioinformatics/btad001) and AmNV under [OR596894](https://doi.org/10.1093/bioinformatics/btad001). Other files (incomplete *Parvoviridae*; third contig of the predicted vMAG of AmFLV; primers used in the completion of AmFLV and nudivirus; annotation tables for AmFLV, third contig and nudivirus; selection of proteins for nudivirus taxonomy) are available on GitHub (<https://github.com/kadlck/NAZV19>) together with some of the code (jupyter-notebooks).

REFERENCES





- Hung K-LJ, Kingston JM, Albrecht M, Holway DA, Kohn JR. 2018. The worldwide importance of honey bees as pollinators in natural habitats. *Proc Biol Sci* 285:20172140. <https://doi.org/10.1098/rspb.2017.2140>
- Calderone NW. 2012. Insect pollinated crops, insect pollinators and US agriculture: trend analysis of aggregate data for the period 1992–2009. *PLoS One* 7:e37235. <https://doi.org/10.1371/journal.pone.0037235>
- Webster MT. 2019. *Apis mellifera*. *Trends Genet* 35:880–881. <https://doi.org/10.1016/j.tig.2019.08.003>
- Wang X, Zhang X, Zhang Z, Lang H, Zheng H. 2018. Honey bee as a model organism to study gut microbiota and diseases. *Drug Discov Today Dis Models* 28:35–42. <https://doi.org/10.1016/j.ddmod.2019.08.010>

5. Gray A, Adjlane N, Arab A, Ballis A, Brusbardis V, Bugeja Douglas A, Cadahía L, Charrière J-D, Chlebo R, Coffey MF, et al. 2023. Honey bee colony loss rates in 37 countries using the COLOSS survey for winter 2019–2020: the combined effects of operation size, migration and queen replacement. *J Apic Res* 62:204–210. <https://doi.org/10.1080/00218839.2022.2113329>
6. Insolia L, Molinari R, Rogers SR, Williams GR, Chiaromonte F, Calovi M. 2022. Honey bee colony loss linked to parasites, pesticides and extreme weather across the United States. *Sci Rep* 12. <https://doi.org/10.1038/s41598-022-24946-4>
7. Hristov P, Shumkova R, Palova N, Neov B. 2020. Factors associated with honey bee colony losses: a mini-review. *Vet Sci* 7:166. <https://doi.org/10.3390/vetsci7040166>
8. Grozinger CM, Fleniken ML. 2019. Bee viruses: ecology, pathogenicity, and impacts. *Annu Rev Entomol* 64:205–226. <https://doi.org/10.1146/annurev-ento-011118-111942>
9. Ullah A, Tlak Gajger I, Majoros A, Dar SA, Khan S, KalimullahA, Nasir Khabir M, Hussain R, Khan HU, Hameed M, Anjum SI. 2021. Viral impacts on honey bee populations: a review. *Saudi J Biol Sci* 28:523–530. <https://doi.org/10.1016/j.sjbs.2020.10.037>
10. 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* 336:1304–1306. <https://doi.org/10.1126/science.1220941>
11. Traynor KS, Mondet F, de Miranda JR, Techer M, Kowallik V, Oddie MAY, Chantawannakul P, McAfee A. 2020. *Varroa destructor*: a complex parasite, crippling honey bees worldwide. *Trends Parasitol* 36:592–606. <https://doi.org/10.1016/j.pt.2020.04.004>
12. Fetters AM, Ashman T-L. 2023. The pollen virome: a review of pollen-associated viruses and consequences for plants and their interactions with pollinators. *Am J Bot* 110:e16144. <https://doi.org/10.1002/ajb2.16144>
13. Remnant EJ, Shi M, Buchmann G, Blacquière T, Holmes EC, Beekman M, Ashe A. 2017. A diverse range of novel RNA viruses in geographically distinct honey bee populations. *J Virol* 91:e00158–17. <https://doi.org/10.1128/JVI.00158-17>
14. Roberts JMK, Anderson DL, Durr PA. 2018. Metagenomic analysis of *Varroa*-free Australian honey bees (*Apis mellifera*) shows a diverse *Picornavirales* virome. *J Gen Virol* 99:818–826. <https://doi.org/10.1099/jgv.0.001073>
15. Galbraith DA, Fuller ZL, Ray AM, Brockmann A, Frazier M, Gikungu MW, Martinez JFI, Kapheim KM, Kerby JT, Kocher SD, Losyev O, Muli E, Patch HM, Rosa C, Sakamoto JM, Stanley S, Vaudo AD, Grozinger CM. 2018. Investigating the viral ecology of global bee communities with high-throughput metagenomics. *Sci Rep* 8:8879. <https://doi.org/10.1038/s41598-018-27164-z>
16. Beaufort A, Piot N, Doublet V, Antunez K, Campbell E, Chantawannakul P, Chejanovsky N, Gajda A, Heerman M, Panziera D, Smaghe G, Yañez O, de Miranda JR, Dalmon A. 2020. Diversity and global distribution of viruses of the western honey bee, *Apis mellifera*. *Insects* 11:239. <https://doi.org/10.3390/insects11040239>
17. Clark TB. 1978. A filamentous virus of the honey bee. *J Invertebr Pathol* 32:332–340. [https://doi.org/10.1016/0022-2011\(78\)90197-0](https://doi.org/10.1016/0022-2011(78)90197-0)
18. Gauthier L, Cornman S, Hartmann U, Cousserans F, Evans J, de Miranda J, Neumann P. 2015. The *Apis mellifera* filamentous virus genome. *Viruses* 7:3798–3815. <https://doi.org/10.3390/v7072798>
19. Kraberger S, Cook CN, Schmidlin K, Fontenele RS, Bautista J, Smith B, Varsani A. 2019. Diverse single-stranded DNA viruses associated with honey bees (*Apis mellifera*). *Infect Genet Evol* 71:179–188. <https://doi.org/10.1016/j.meegid.2019.03.024>
20. Fédière G. 2000. Epidemiology and pathology of *Densovirinae*. *Contrib Microbiol* 4:1–11. <https://doi.org/10.1159/000060332>
21. Cotmore SF, Agbandje-McKenna M, Chiorini JA, Mukha DV, Pintel DJ, Qiu J, Soderlund-Venermo M, Tattersall P, Tijssen P, Gatherer D, Davison AJ. 2014. The family *Parvoviridae*. *Arch Virol* 159:1239–1247. <https://doi.org/10.1007/s00705-013-1914-1>
22. Péntzes JJ, Söderlund-Venermo M, Canuti M, Eis-Hübinger AM, Hughes J, Cotmore SF, Harrach B. 2020. Reorganizing the family *Parvoviridae*: a revised taxonomy independent of the canonical approach based on host association. *Arch Virol* 165:2133–2146. <https://doi.org/10.1007/s00705-020-04632-4>
23. Hartmann U, Forsgren E, Charrière J-D, Neumann P, Gauthier L. 2015. Dynamics of *Apis mellifera* filamentous virus (AmFV) infections in honey bees and relationships with other parasites. *Viruses* 7:2654–2667. <https://doi.org/10.3390/v7052654>
24. Kadlečková D, Tachezy R, Erban T, Deboutte W, Nunvář J, Saláková M, Matthijnsens J. 2022. The virome of healthy honey bee colonies: ubiquitous occurrence of known and new viruses in bee populations. *mSystems* 7:e0007222. <https://doi.org/10.1128/mSystems.00072-22>
25. Bailey L. 1982. Viruses of honeybees. *Bee World* 63:165–173. <https://doi.org/10.1080/0005772X.1982.11097891>
26. Petersen JM, Bézier A, Drezén J-M, van Oers MM. 2022. The naked truth: an updated review on nudiviruses and their relationship to bracoviruses and baculoviruses. *J Invertebr Pathol* 189:107718. <https://doi.org/10.1016/j.jip.2022.107718>
27. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
28. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834. <https://doi.org/10.1101/gr.213959.116>
29. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>
30. Ondov BD, Bergman NH, Phillippy AM. 2011. Interactive metagenomic visualization in a web browser. *BMC Bioinformatics* 12:385. <https://doi.org/10.1186/1471-2105-12-385>
31. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpidis NC. 2021. Checkv assesses the quality and completeness of metagenome-assembled viral genomes. *Nat Biotechnol* 39:578–585. <https://doi.org/10.1038/s41587-020-00774-7>
32. Vasimuddin M, Misra S, Li H, Aluru S. 2019. “Efficient architecture-aware acceleration of BWA-MEM for multicore systems” 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS); Rio de Janeiro, Brazil, p 314–324. <https://doi.org/10.1109/IPDPS.2019.00041>
33. Rangel-Pineros G, Almeida A, Beracochea M, Sakharova E, Marz M, Reyes Muñoz A, Hölzer M, Finn RD. 2023. VIRify: an integrated detection, annotation and taxonomic classification pipeline using virus-specific protein profile hidden Markov models. *PLoS Comput Biol* 19:e1011422. <https://doi.org/10.1371/journal.pcbi.1011422>
34. Kieft K, Adams A, Salamzade R, Kalan L, Anantharaman K. 2022. vRhyme enables binning of viral genomes from metagenomes. *Nucleic Acids Res* 50:14. <https://doi.org/10.1093/nar/gkac341>
35. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* 11:2864–2868. <https://doi.org/10.1038/ismej.2017.126>
36. Roux S, Páez-Espino D, Chen I-M, Palaniappan K, Ratner A, Chu K, Reddy TBK, Nayfach S, Schulz F, Call L, Neches RY, Woyke T, Ivanova NN, Eloe-Fadrosh EA, Kyrpidis NC. 2021. IMG/VR V3: an integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses. *Nucleic Acids Res* 49:D764–D775. <https://doi.org/10.1093/nar/gkaa946>
37. Tolstoganov I, Kamenev Y, Kruglikov R, Ochkalova S, Korobeynikov A. 2022. BinSPreader: refine binning results for fuller MAG reconstruction. *iScience* 25:104770. <https://doi.org/10.1016/j.isci.2022.104770>
38. Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, Delmont TO, Pratama AA, Gazitúa MC, Vik D, Sullivan MB, Roux S. 2021. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* 9:37. <https://doi.org/10.1186/s40168-020-00990-y>
39. Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL, Solden LM, Liu P, Narrowe AB, Rodríguez-Ramos J, Bolduc B, Gazitúa MC, Daly RA, Smith GJ, Vik DR, Pope PB, Sullivan MB, Roux S, Wrighton KC. 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res* 48:8883–8900. <https://doi.org/10.1093/nar/gkaa621>
40. Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L, Deboutte W, Yinda CK, Lavigne R, Maes P, Van Ranst M, Heylen E, Matthijnsens J. 2015. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep* 5:16532. <https://doi.org/10.1038/srep16532>

41. Deboutte W, Beller L, Yinda CK, Maes P, de Graaf DC, Matthijssens J. 2020. Honey-bee-associated prokaryotic viral communities reveal wide viral diversity and a profound metabolic coding potential. *Proc Natl Acad Sci U S A* 117:10511–10519. <https://doi.org/10.1073/pnas.1921859117>
42. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang D, Zheng C, Lanczycki CJ, Marchler-Bauer A. 2020. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* 48:D265–D268. <https://doi.org/10.1093/nar/gkz991>
43. Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* 33:511–518. <https://doi.org/10.1093/nar/gki198>
44. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>
45. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T. 2020. Modeltest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol* 37:291–294. <https://doi.org/10.1093/molbev/msz189>
46. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321. <https://doi.org/10.1093/sysbio/syq010>
47. Letunic I, Bork P. 2019. Interactive tree of life (iTOL) V4: recent updates and new developments. *Nucleic Acids Res* 47:W256–W259. <https://doi.org/10.1093/nar/gkz239>
48. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>
49. Zulkower V, Rosser S. 2020. DNA features viewer: a sequence annotation formatting and plotting library for python. *Bioinformatics* 36:4350–4352. <https://doi.org/10.1093/bioinformatics/btaa213>
50. Laugel M, Lecomte E, Ayuso E, Adjali O, Mével M, Penaud-Budloo M. 2023. The diversity of *Parvovirus* telomeres. In Fonseca-Alves CE (ed), *Recent advances in canine medicine*, IntechOpen. London.
51. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>
52. Wang X, Shang Y, Chen C, Liu S, Chang M, Zhang N, Hu H, Zhang F, Zhang T, Wang Z, Liu X, Lin Z, Deng F, Wang H, Zou Z, Vlak JM, Wang M, Hu Z. 2019. Baculovirus per os infectivity factor complex: components and assembly. *J Virol* 93:e02053-18. <https://doi.org/10.1128/JVI.02053-18>
53. Allain TW, Stentiford GD, Bass D, Behringer DC, Bojko J. 2020. A novel nudivirus infecting the invasive demon shrimp *Dikerogammarus haemobaphes* (Amphipoda). *Sci Rep* 10:14816. <https://doi.org/10.1038/s41598-020-71776-3>
54. Tijssen P, Péntzes JJ, Yu Q, Pham HT, Bergoin M. 2016. Diversity of small, single-stranded DNA viruses of invertebrates and their chaotic evolutionary past. *J Invertebr Pathol* 140:83–96. <https://doi.org/10.1016/j.jip.2016.09.005>
55. Xu P, Liu Y, Graham RI, Wilson K, Wu K. 2014. Densovirus is a mutualistic symbiont of a global crop pest (*Helicoverpa armigera*) and protects against a *Baculovirus* and Bt biopesticide. *PLoS Pathog* 10:e1004490. <https://doi.org/10.1371/journal.ppat.1004490>
56. Sarker S, Talukder S, Anwar A, Van TTH, Petrovski S. 2022. Unravelling bile viromes of free-range laying chickens clinically diagnosed with spotty liver disease: emergence of many novel chaphamaparvoviruses into multiple lineages. *Viruses* 14:2543. <https://doi.org/10.3390/v14112543>
57. Liu J, Lin R, Wu G, Liu R, Luo Z, Fan X. 2022. CoCoNet: coupled contrastive learning network with multi-level feature ensemble for multi-modality image fusion. *arXiv:2211.10960*. <https://doi.org/10.48550/arXiv.2211.10960>
58. Nissen JN, Johansen J, Allesøe RL, Sønderby CK, Armenteros JJA, Grønbech CH, Jensen LJ, Nielsen HB, Petersen TN, Winther O, Rasmussen S. 2021. Improved metagenome binning and assembly using deep variational autoencoders. *Nat Biotechnol* 39:555–560. <https://doi.org/10.1038/s41587-020-00777-4>
59. Uritskiy GV, DiRuggiero J, Taylor J. 2018. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6:158. <https://doi.org/10.1186/s40168-018-0541-1>

RESEARCH ARTICLE

Varroa destructor parasitism and Deformed wing virus infection in honey bees are linked to peroxisome-induced pathways

Tomas Erban¹  | Dominika Kadleckova²  | Bruno Sopko¹  | Karel Harant³  |
 Pavel Talacko³  | Martin Markovic¹  | Martina Salakova²  | Klara Kadlikova¹  |
 Ruth Tachezy²  | Jan Tachezy⁴ 

¹Proteomics and Metabolomics Laboratory, Crop Research Institute, Prague 6-Ruzyne, Czechia

²Department of Genetics and Microbiology, Faculty of Science BIOCEV, Charles University, Vestec, Czechia

³Proteomics Core Facility, Faculty of Science BIOCEV, Charles University, Vestec, Czechia

⁴Department of Parasitology, Faculty of Science BIOCEV, Charles University, Vestec, Czechia

Correspondence

Tomas Erban, Proteomics and Metabolomics Laboratory, Crop Research Institute, Drnovska 507/73, Prague 6-Ruzyne, CZ-16106, Czechia. Email: erban@vurv.cz
arachnid@centrum.cz

Funding information

Ministry of Agriculture of the Czech Republic, Grant/Award Numbers: QK1910018, RO0423; Charles University, Grant/Award Number: SVV260679; European Regional Development Fund (ERDF) project CePaVIP, Grant/Award Number: CZ.02.1.01/0.0/0.0/16_019/0000759

Abstract

The ectoparasitic mite *Varroa destructor* transmits and triggers viral infections that have deleterious effects on honey bee colonies worldwide. We performed a manipulative experiment in which worker bees collected at emergence were exposed to *Varroa* for 72 h, and their proteomes were compared with those of untreated control bees. Label-free quantitative proteomics identified 77 differentially expressed *A. mellifera* proteins (DEPs). In addition, viral proteins were identified by orthogonal analysis, and most importantly, Deformed wing virus (DWV) was found at high levels/intensity in *Varroa*-exposed bees. Pathway enrichment analysis suggested that the main pathways affected included peroxisomal metabolism, cyto-/exoskeleton reorganization, and cuticular proteins. Detailed examination of individual DEPs revealed that additional changes in DEPs were associated with peroxisomal function. In addition, the proteome data support the importance of TGF- β signaling in *Varroa*-DWV interaction and the involvement of the mTORC1 and Hippo pathways. These results suggest that the effect of DWV on bees associated with *Varroa* feeding results in aberrant autophagy. In particular, autophagy is selectively modulated by peroxisomes, to which the observed proteome changes strongly corresponded. This study complements previous research with different study designs and suggests the importance of the peroxisome, which plays a key role in viral infections.

KEYWORDS

Apis mellifera, autophagy, DWV, host-pathogen interaction, lipid metabolism

Abbreviations: 2D-E-MS/MS, 2D electrophoresis tandem mass spectrometry; AFV, Apis flavivirus; A. mellifera, Apis mellifera; AmFV, Apis mellifera filamentous virus; BlastP, protein-protein BLAST; BMLV, Bee macula-like virus; BQCV, Black queen cell virus; CCD, Conserved Domains Database; CoA, coenzyme A; CRI, Crop Research Institute; DEPs, differentially expressed A. mellifera proteins; DOI, digital object identifier; DWV, Deformed wing virus; ER, endoplasmic reticulum; FDR, false discovery rate; IMD, immune deficiency; JAK/STAT, Janus kinases (JAKs) / signal transducer and activator of transcription proteins (STATs); KBV, Kashmir bee virus; KEGG, Kyoto Encyclopedia of Genes and Genomes; LFQ, label-free quantitative; MassIVE, Mass Spectrometry Interactive Virtual Environment; MMTS, methanethiosulfonate; MOB1, MOB kinase activator-like 1; MS, mass spectrometry; MS/MS, tandem mass spectrometry; mTOR, mammalian target of rapamycin; MZE, Ministry of Agriculture of the Czech Republic; nanoLC, nanoliquid chromatography; NAZV, National Agency for Agricultural Research of the Ministry of Agriculture of the Czech Republic; NCBI, National Center for Biotechnology Information; NetoVIR, Novel Enrichment Technique of Viromes; NF- κ B, nuclear factor kappa-light-chain-enhancer of activated B cells; PBS, phosphate-buffered saline; PCA, principal component analysis; PPO, prophenoloxidase; Prx5, peroxiredoxin-5; Romo1, reactive oxygen species modulator 1; qPCR, quantitative polymerase chain reaction; RefSeq, Reference Sequence; RNA, ribonucleic acid; ROS, reactive oxygen species; SBV, Sacbrood virus; SDC, sodium deoxycholate; TCEP, Tris(2-carboxyethyl) phosphine hydrochloride; TLR4, Toll-like receptor 4; Tmed7, transmembrane emp24 domain-containing protein 7; TGF- β , transforming growth factor β ; V. destructor, Varroa destructor.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Proteomics* published by Wiley-VCH GmbH

1 | INTRODUCTION

The invasive ectoparasitic mite *Varroa destructor* (Anderson & Trueman, 2000) is currently distributed worldwide [1, 2], including in Australia, where the first outbreak was recorded in 2022 [3, 4]. The mite has become a major threat to honey bees since it shifted and broadened in the 1950s from the eastern honey bee *Apis cerana* Fabricius, 1793, to the western honey bee *Apis mellifera* Linnaeus, 1758. The infestation of honey bee colonies by *Varroa* is connected to virus transmission and an increase in viral load [5, 6]. *Varroa* infestation is closely associated with Deformed wing virus (DWV), which is a cause of bee deformation and is considered the most common sign of varroosis worldwide [2]. Moreover, DWV infection is driven by the spread of *Varroa* mite populations throughout the world [7]. Notably, a virome analysis of bee populations in central Europe and the subsequent comparison of this population with naive Australian bee populations, which were free of both *Varroa* and DWV at the time of the study, indicated that DWV profoundly affects the composition of the virome in the new host [8, 9]. Thus, the *Varroa*–DWV interaction is profound and affects the survival and function of bee colonies. Understanding the mechanisms of interaction among mites, viruses, and honey bees is important for the implementation of strategies to prevent the weakening and death of bee colonies.

The negative effect of mite parasitism is connected to the weakening of the host, as the parasite consumes nutrients, resulting in lower protein and carbohydrate levels in bees [10]. Using a proteomic approach, *Varroa* bodies were shown to carry numerous proteins acquired from host honey bees [11, 12], and viral proteins have also been identified in these mites [11]. Additionally, cell components such as those from the fat body [13] together with hemolymph components have been identified in mites [11]. 2D electrophoresis–tandem mass spectrometry (2D-E-MS/MS) proteomic analysis of 1-day-old bee hemolymph revealed the effect of *Varroa* on diverse physiological processes, such as energy metabolism, detoxification, the oxidative stress response, and olfaction [14].

The effect of *Varroa* parasitism and DWV infection on the immune response of bees is controversial. An increase in the DWV load mediated by *Varroa* was associated with the suppression of immunity-related markers (e.g., antimicrobial peptides and enzymes) due to *Varroa* parasite in capped comb cells [15, 16]. Studies on *Varroa*-parasitized larvae have indicated that the negative effect of DWV on humoral and cellular immune responses is mediated through NF- κ B signaling, as indicated by the observation of dorsal-1A, Amel-LRR, and apidaecin expression [16, 17]. A quantitative proteomic analysis of emerging bees suggested that *Varroa* infection disrupts host autophagy via modulation of TGF- β signaling. In addition, the synergistic effects of *Varroa* and DWV modulate NF- κ B and JAK/STAT signaling, both of which are associated with autophagy [18]. However, in other studies, only a weak association between *Varroa* infection and immunosuppression was observed [19–21]. Most recently, a significant upregulation of the immune response of bees to *Varroa* infection was observed in 10-day-old bees from *Varroa*-infested colonies compared with those from control colonies [22]. Specifically, the antimicrobial peptides apidaecin,

Significance statement

Varroa destructor is the greatest threat to honey bees worldwide. The disease caused by mites is called varroosis and is associated with the appearance of crippled bees and leads to weakened colonies. It is known that *Varroa* parasitism triggers DWV infection, while in the absence of a mite, the virus can reach a persistent state. To date, different studies have produced different results regarding the effect of *Varroa*–DWV interaction on honey bees. This variation is caused by differences in developmental stages and study designs. We designed a manipulative experiment to study the *Varroa*–DWV interaction. This study is the first to report the effect of *Varroa* on the proteome of 3-day-old worker bees exposed to high *Varroa* loads after emergence. Using high-throughput proteomic analysis, we identified significant changes in *Varroa*-treated and control bees, suggesting the importance of the role of peroxisomes. Until now, the role of peroxisomes in *Varroa*–DWV interactions has not been emphasized, although peroxisomes play a crucial role in viral infections. In addition, we complemented the previously suggested role of the TGF- β signaling pathway in *Varroa* parasitism.

abaecin, defensin, and hymenoptaecin were strongly upregulated in *Varroa*-parasitized bees. Conversely, the expression of the NF- κ B-like *Relish* gene, which induces a humoral response (in the IMD pathway), was downregulated [22]. The observed differences in the impact of *Varroa* mites on bees reported in various studies most likely resulted from differences in experimental designs and the use of different developmental stages and castes of bees [23, 24].

In this study, we exposed previously unparasitized emerged worker bees for 3 days (72 h) to *Varroa* mites collected from a collapsing diseased bee colony with typical symptoms of varroosis. Using label-free quantitative proteomics, we observed proteomic changes in honey bees after exposure to the mites, and from the same data, we retrieved information about the viruses present in the experimental bees. We analyzed the key pathways involved and compared the changes with those found in previous studies that investigated *Varroa*–virus exposure.

2 | MATERIALS AND METHODS

2.1 | Approach and samples

All the biological samples were collected in mid-July, when *Varroa* infestations are generally very infrequent and collapsing colonies are exceptional. The experimental scheme, including an overview of the analyses, is shown in Figure 1.

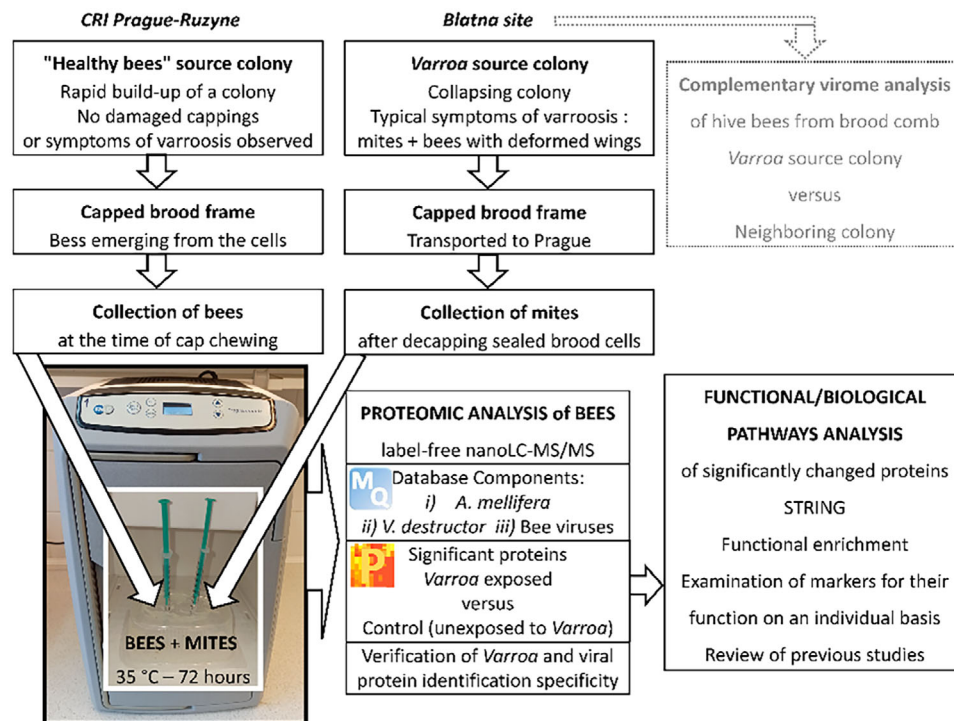


FIGURE 1 The overall approach used in this study.

2.1.1 | Unparasitized emerging *Apis mellifera* worker bees

The bees used for the manipulative experiment originated from a colony at the Crop Research Institute (CRI), Prague. The worker bees were collected from a brood frame at the time of emergence when they were chewing the cell cap [18, 25, 26]. The experimental bees had no obvious defects, were vital, and had no mites found on them or in their cells when they were collected. In addition, the colony had a rapid buildup with no damaged cappings, and no obvious symptoms of bee diseases were observed in the colony throughout the season.

2.1.2 | *Varroa destructor*

The mites came from a collapsing colony with typical symptoms of varroosis, located at another site about 100 km away from CRI in the South Bohemian municipality of Blatna, Czechia. On July 12, a frame with a sealed brood was taken from the diseased colony and transported to the laboratory in Prague, where the female mites were collected after the comb caps were disturbed by circular motion using a toothpick.

2.1.3 | The manipulative experiment

The emergent bees were placed in plastic chambers with mesh lids on the bottom. The chambers were equipped with two 1-mL syringes placed at the top of the chamber to provide a feed consisting of a 50% (w/v) sugar solution prepared by dissolving granulated beet sugar. The

experiment was started by placing 50 mites on 11 worker bees with a brush in a chamber through the opening of the syringe. The same number of bees were placed in a different chamber without mites. Both chambers were incubated in a portable ICT-P mini-incubator (FALC Instruments, Treviglio, Italy) of 18 L in volume (internal dimensions 275 × 185 × 335 mm) with circulating air, a precision of $\pm 0.2^\circ\text{C}$ and a stability of $\pm 0.2^\circ\text{C}$. The exposure lasted for 72 h at 35°C . The availability of the sugar feed supply in the syringes was continuously checked. At the end of the exposure period, the bees were collected, and mites were removed with a brush. The bees were placed in 0.5 mL microvials and immediately frozen on dry ice ($\sim -70^\circ\text{C}$). Samples were then stored in a deep freezer at -80°C until analysis. At the end of the experiment, the number of dead and live mites were counted.

2.1.4 | Use of bees for virome analysis

At the Blatna site, additional samples were collected from the *Varroa* source colony and a neighboring colony. The samples included bees that were shaken from the brood frame into a plastic bag and then frozen on dry ice. Later, the samples were stored at -80°C and used for analysis of the bee virome, which provided complementary data to the manipulative experiment.

2.2 | Proteomic analysis of bees

Ten control and ten *Varroa*-exposed bees from the manipulative experiment were subjected to label-free quantitative (LFQ) mass

spectrometry (MS) analysis. Each bee was homogenized in 2.5 mL of homogenization buffer consisting of 100 mM triethylammonium bicarbonate buffer (TEAB; Cat No. 90360; Sigma–Aldrich) and 2% sodium deoxycholate (SDC; Cat No. 30970, BioXtra, Sigma–Aldrich) in a 5-mL glass Potter–Elvehjem tissue grinder (Kartell Labware division, Noviglio, Italy) with a Teflon pestle operated by a drill. The homogenization was completed on ice in three cycles, with each cycle involving 1 min of homogenization followed by incubation on ice for 10 min. The homogenate was transferred to a 15-mL centrifugation tube and centrifuged for 10 min at $3000 \times g$ at 4°C in an MR 23i centrifuge (Jouan Industries, France). The supernatant was collected, divided into aliquots, and stored at –80°C until analysis. Further sample processing and MS analysis were performed as previously described [27]. Briefly, the protein concentration was determined using a BCA protein assay kit (Cat No. 23225, Thermo Fisher Scientific, MA, USA). Cysteines were reduced with Tris(2-carboxyethyl) phosphine hydrochloride (TCEP) and blocked with methyl methanethiosulfonate (MMTS). The samples were digested with porcine trypsin. Then, the samples were acidified with trifluoroacetic acid to a final concentration of 1%. SDC was removed by extraction to ethylacetate and subsequently to hexane. Peptides were desalted using a Michrom C18 column. The dried peptides were resuspended in 25 μ L of water containing 2% acetonitrile and 0.1% trifluoroacetic acid. The samples were injected into the LC in the order they were numbered, with all control samples analyzed first, followed by the *Varroa*-exposed samples. Peptides were separated by nanoliquid chromatography (nanoLC) on a Dionex Ultimate 3000 system (Thermo Fisher Scientific) and analyzed on an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific).

2.3 | Proteomic data evaluation

The data were evaluated with MaxQuant version 2.2.0.0 using LFQ algorithms [28, 29] and the Andromeda search engine [30]. The key criteria used in the data analysis were a false discovery rate (FDR) of 0.01 for proteins and peptides, a minimum length of seven amino acids, a fixed modification (methylthio), and variable modifications of N-terminal protein acetylation and methionine oxidation. The data were locally evaluated against selected databases of nonredundant protein sequences that were downloaded from NCBI on February 13, 2023. Sequences related to the honey bee host and the mite ectoparasite were downloaded from the Reference Sequence protein database (RefSeq, [31]) and consisted of 23,251 sequences from *A. mellifera* (txid7460) and 30,221 sequences from *V. destructor* (txid109461). In addition, 6785 viral sequences restricted to *A. mellifera* (excluding bacteriophages) were downloaded from the NCBI database. The local database is available along with the raw data and result files in MasSIVE (see the Data Availability section). The data were processed in Perseus version 2.0.7.0 [32]. Results representing standard contaminants, reverse sequences (decoys), and sequences identified only with modified peptides were discarded. The dataset containing proteins with at least five valid LFQ values in at least one experimental group was further analyzed. The dataset was log₂ transformed. Missing LFQ

values were replaced from the normal distribution (width 0.3; downshift 1.8) and histograms before and after the data imputation were examined. Heatmap with hierarchical clustering was used to check the uniformity of the sample proteomes between the two subgroups. Significant differences were determined by a permutation-based two-sided *t*-test with an error-corrected *p*-value (FDR = 0.05; S0 = 0.1). The number of randomizations in the analysis was 1000. The results were visualized in a volcano plot, and the significant proteins were exported in tabular form. In addition, the significantly differentially expressed proteins (DEPs) were normalized by z-scores across rows, and hierarchical clustering using average Euclidian distance was performed in OriginPro 2024 (OriginLab, Northampton, MA, USA).

To identify functional protein network interactions, we performed an analysis in STRING v11.5 [33, 34]. The STRING network analysis was performed with the identified 77 *A. mellifera* DEPs, for which a list of the curated gene names (including alternative) was found. The analysis was performed at a confidence of 0.3. Only those terms that were significantly enriched with a Benjamini–Hochberg FDR-corrected *p*-value of less than 0.05 were included. Furthermore, the proteins were analyzed individually to correct and complement automated protein annotations. The representative GenBank accession numbers of the significant proteins were provided with a hyperlink to the NCBI protein database, facilitating direct individual examination via the Conserved Domains Database (CCD) [35]. In addition, we performed individual analyses of the DEPs via the Kyoto Encyclopedia of Genes and Genomes (KEGG) [36], InterPro [37], and UniProt. The information retrieved from these databases was added to the proteins in tabular form and provided with hyperlinks, and the key caption was reflected in the visualization of the STRING analysis to the gene symbols. Finally, all the observed proteome changes, together with reference studies, were used to create a schematic of relevant pathways using BioRender (<https://biorender.com>).

We analyzed the presence of unique peptides retrieved from a MaxQuant peptide table in the samples and checked whether the peptides in the individual samples were identified by MS/MS or only based on mass matching without MS/MS spectra. To increase the reliability of the identified viruses, we performed a qualitative analysis of the viruses via peptide searches based on MaxQuant intensity values with the disabled match-between-runs option [38].

2.4 | Virome analysis of colonies from the mite source site

Virome analysis was performed on randomly selected bees from the brood comb of the colony that was the source of *Varroa* mites and a neighboring colony at the Blatna site. The modified NetoVIR protocol was utilized as described previously [39]. Briefly, two bees per colony were homogenized in 1 mL of 1x phosphate-buffered saline (PBS) in tubes with 2.8-mm ceramic (zirconium oxide) beads (cat. no. P000911-LYSKO-A, Bertin Technologies, Montigny-le-bretonneux, Ile-de-France, France). Homogenization was performed for 1 min at 3000 rpm with a MINILYSIS tissue homogenizer (Bertin Technologies,

Montigny-le-bretonneux, Ile-de-France, France). The samples were treated with benzonase nuclease and micrococcal nuclease (New England Biolabs, Ipswich, MA, USA), and nucleic acids were subsequently isolated with a QIAamp Viral RNA Mini Kit (QIAGEN, Hilden, Germany). Reverse transcription and the first amplification procedure were performed via a modified WTA2 protocol (MERCK, Rahway, NJ, USA). Libraries were prepared via a modified Nextera XT protocol (Illumina, San Diego, CA, USA). The samples were subsequently sequenced at the KU Leuven Nucleomics Core (VIB), Leuven, Belgium, on the HiSeq2000 platform (Illumina, CA, USA) for 2x cycles, during which 150-bp paired-end reads were obtained. The reads were evaluated with FastQC [40] and subsequently trimmed with Trimmomatic [41]. After another evaluation with FastQC, the reads were assembled with SPAdes [42] with the settings `-meta` and `-k 21, 33, 55, 77`. The contigs obtained were blasted with Diamond [43] against a nonredundant protein database (NCBI, downloaded on September 8, 2022). The reads were subsequently mapped to contigs with `bwa-mem2` [44], and CoverM [45] was used to identify the reads mapped to individual contigs. The hits and number of reads that had been mapped were uploaded into KronaTools [46].

3 | RESULTS

3.1 | Survival of bees and mites in the manipulative experiment

All 11 control and 11 *Varroa*-exposed bees survived 72 h of exposure to 50 *Varroa* mites. Nineteen dead mites were found at the end of the experiment, while 31 out of 50 mites remained viable throughout the experiment. Ten bees from each group were then subjected to proteomic analysis.

3.2 | Overall proteomic analysis results

LFQ proteomics revealed 26,343 peptides (Table S1) in all the analyzed samples, which were assigned to 2663 proteins (Table S2A). Using the LFQ algorithm, we found six groups of viral proteins (Table S2B; Section 3.3). In addition, we found several proteins identified by *V. destructor* sequence information, of which those identified only by *V. destructor* sequence information were filtered out from further data analysis. For quantitative analysis of honey bee DEPs, we further filtered the data to a threshold of at least five valid LFQ values in an experimental group, resulting in a dataset of 2421 proteins. The histograms before (Figure S1) and after (Figure S2) the imputation of missing data (width 0.3; downshift 1.8) were both symmetric/unimodal. Exploratory analysis of the proteomes using hierarchical clustering on columns (Figure S3) revealed sample homogeneity in the two experimental groups.

3.3 | Identification of viral proteins

Initial analysis using LFQ algorithms revealed that four protein groups matched polyproteins of DWV origin. Importantly, DWV was identified by a total of 16 unique peptides exclusively expressed by *Varroa*-exposed bees. The high number of MS/MS counts (149 for unique peptides) indicates a high level of DWV-B polyprotein in the treated bees. The polyprotein is a large precursor protein that includes the structural proteins VP1, VP2, VP3 and possibly V4 [47] at the N-terminal domain and nonstructural regions (RNA helicase, a chymotrypsin-like 3C protease, and an RNA-dependent RNA polymerase) at the C-terminal domain [48]. Detailed inspection of the identified peptides revealed that all the peptides mapped to the amino acid sequence of the DWV-B polyprotein (665 identical or closely related sequences at NCBI), of which APP91308 was used as a representative sequence. The peptides covered both the structural and nonstructural domains of the protein. In addition, we identified a single unique peptide with a single amino acid substitution D/N at position 871 that matched the VP1 domain of the DWV-A protein (AMK01489) and a unique peptide with a substitution A/T at position 916 (the VP3 domain) corresponding to the polyprotein AUI41303 (Figure 2, Table S3). Because only a single peptide was found for each protein, the interpretation of these findings is inconclusive.

As quantitative LFQ methods may lead to false negatives, we verified the distribution of viral peptides in control and *Varroa*-treated bees via qualitative analysis based on peptide intensities and mapping to the reference sequence. This analysis confirmed the absence of DWV in the control samples, while all *Varroa*-treated bees were DWV positive (Figure 2, Table S3).

In addition to DWV, LFQ-based analysis revealed the presence of Apis flavivirus (AFV) and Apis mellifera filamentous virus (AmFV), which was confirmed by qualitative analysis. Moreover, the latter analysis suggested the presence of two additional viral species, Kashmir bee virus (KBV) and Bee macula-like virus (BMLV) (Figure 2). Polyprotein of AFV origin (YP_009388303) [49] was detected by two unique peptides found only in four *Varroa*-treated bees, while no AFV infection was observed in the controls. In three control bees, we found two unique peptides that matched the protein of AmFV origin with the C-terminal domain of the putative metalloproteinase (AKY03074). The presence of KBV was suggested by the identification of two overlapping unique peptides that matched a domain of the 3C-protease of the nonstructural KBV protein [50]. These peptides were detected in two control bees and six *Varroa*-treated bees (Figure 2). Finally, BMLV infection was suggested only for a single *Varroa*-treated bee (v5). In this sample, three unique peptides matched the BMLV coat protein (UYM19103). Importantly, all these proteins were identified with very low MS/MS count values (3–8; Figure 2), suggesting a low load of the corresponding viruses in comparison to DWV. Collectively, our data indicate significant *Varroa*-dependent DWV infection in *Varroa*-treated bees. Persistent infections with AFV and BMLV might also be associated with *Varroa*, while infections with KBV and AmFV in our

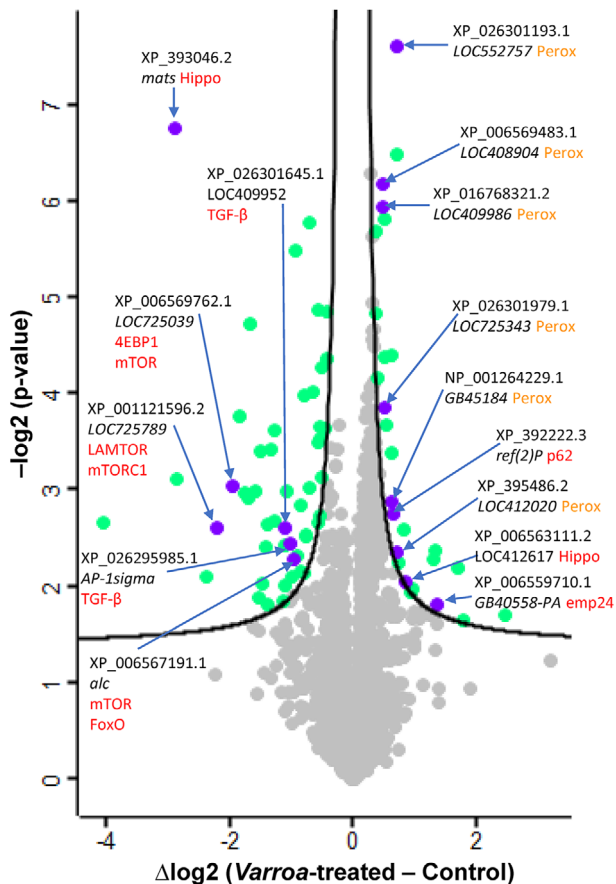


FIGURE 4 Volcano plot identifying 77 significantly (FDR = 0.05; $S_0 = 0.1$; permutations = 1,000) differentially expressed proteins (DEPs) between *Varroa*-treated and control samples of bees. Selected DEPs are labeled in the volcano plot.

category using the Benjamini–Hochberg procedure (FDR) are shown in Table S6.

The most significantly enriched term was the term peroxisomes in the following three categories: cellular component (GO:0005777; 7/91 – observed/background gene count; q -value = 5.2×10^{-4}), KEGG pathways (ame04146; 6/59 – observed/background; q -value = 9.64×10^{-5}), and compartments (GOCC:0005777; 7/93 – observed/background; q -value = 8.9×10^{-4}). In addition, the peroxisome term was combined with other enriched KEGGs: fatty acid metabolism (ame01212; 4/50 – observed/background; q -value = 6.1×10^{-3}), biosynthesis of unsaturated fatty acids (ame01040; 4/27 – observed/background; q -value = 9.9×10^{-4}), β -alanine metabolism (ame00410; 3/21 – observed/background; q -value = 7×10^{-3}), and alpha-linolenic acid metabolism (ame00592; 2/11 – observed/background; q -value = 4.7×10^{-2}). Functional enrichment analysis revealed that structural molecule activity (GO:0005198; 12/355 – observed/background; q -value = 3.6×10^{-4}), was the only significantly enriched term in the category Molecular function. The next important enriched terms were those connected to the cuticle in the Annotated UniProt keywords (KW-0193; 4/42 – observed/background; q -value = 1.05×10^{-2}) and Local

STRING network cluster CL19319; 5/92 – observed/background; q -value = 3.71×10^{-2}). Many counts in the network were found in the enriched term “cellular anatomical entity” in two categories: cellular component (GO:0110165; 62/8868 – observed/background; q -value = 1.3×10^{-3}) and compartments (GOCC:0110165; 58/8023 – observed/background; q -value = 4.2×10^{-3}). The results of hyper-linking the individual analysis of the significant proteins in the bioinformatic resources (KEGGs, CDs, InterPro) are shown in Table S7. The key information obtained is reflected in the STRING interaction network (Figure 6) together with the gene symbols.

The key group of interacting proteins in the STRING network (Figure 6) was found to be six peroxisomal proteins. This was supported by the heatmap (Figure 5), which shows the coexpression of the peroxisomal proteins, as five of them clustered immediately next to each other and the sixth clustered nearby. Furthermore, the three nodes directly related to the peroxisomal proteins identified via STRING, that is, proteasomal proteins, proteins involved in steroid degradation and p24, were also upregulated. Notably, p62 was found to be coexpressed in the heatmap, although the direct interaction between p62 and peroxisomal proteins was not detected via STRING. STRING analysis also revealed other groups of interacting DEPs, and the detailed analysis collectively suggested that bee exposure to *Varroa* and consequently DWV infection has an impact on peroxisomal function and cyto-/exoskeleton reorganization and supports the importance of TGF- β signaling and the involvement of the mTORC1 and Hippo pathways in *Varroa*-DWV interactions.

4 | DISCUSSION

4.1 | Original system for investigating the specific developmental stage of bees and *Varroa*/DWV exposure status

In this study, we used a novel, well-defined biological system to investigate the effect of *Varroa*/DWV on the bee proteome. In our system, we used 3-day-old worker bees and particular *Varroa* loads, including the viruses, most notably high levels of DWV in *Varroa*-treated bees. We increased the uniformity of the experimental bees for manipulative experiments by collecting them at the time of emergence [18, 25], ensuring the absence of *Varroa* mites in the cell and on the bees. Thus, since their emergence after the completion of metamorphosis, the experimental bees have had no prior contact with colony members and could not have been contaminated by sharing the feeding route with other bees or the comb, including food stores, as can occur when newly emerged bees are collected collectively from brood combs. In addition, the bees were collected from a single colony with no signs of disease to increase the sample uniformity, but this reduces the biological variability that can be useful in other studies. Importantly, we recorded the number of bees and mites used at the beginning of the experiment and the number that survived. In particular, similar data are required for chronic toxicity tests on bees [51]. Somewhat unexpectedly, all the *Varroa*-exposed bees survived the 72-h treatment, as did

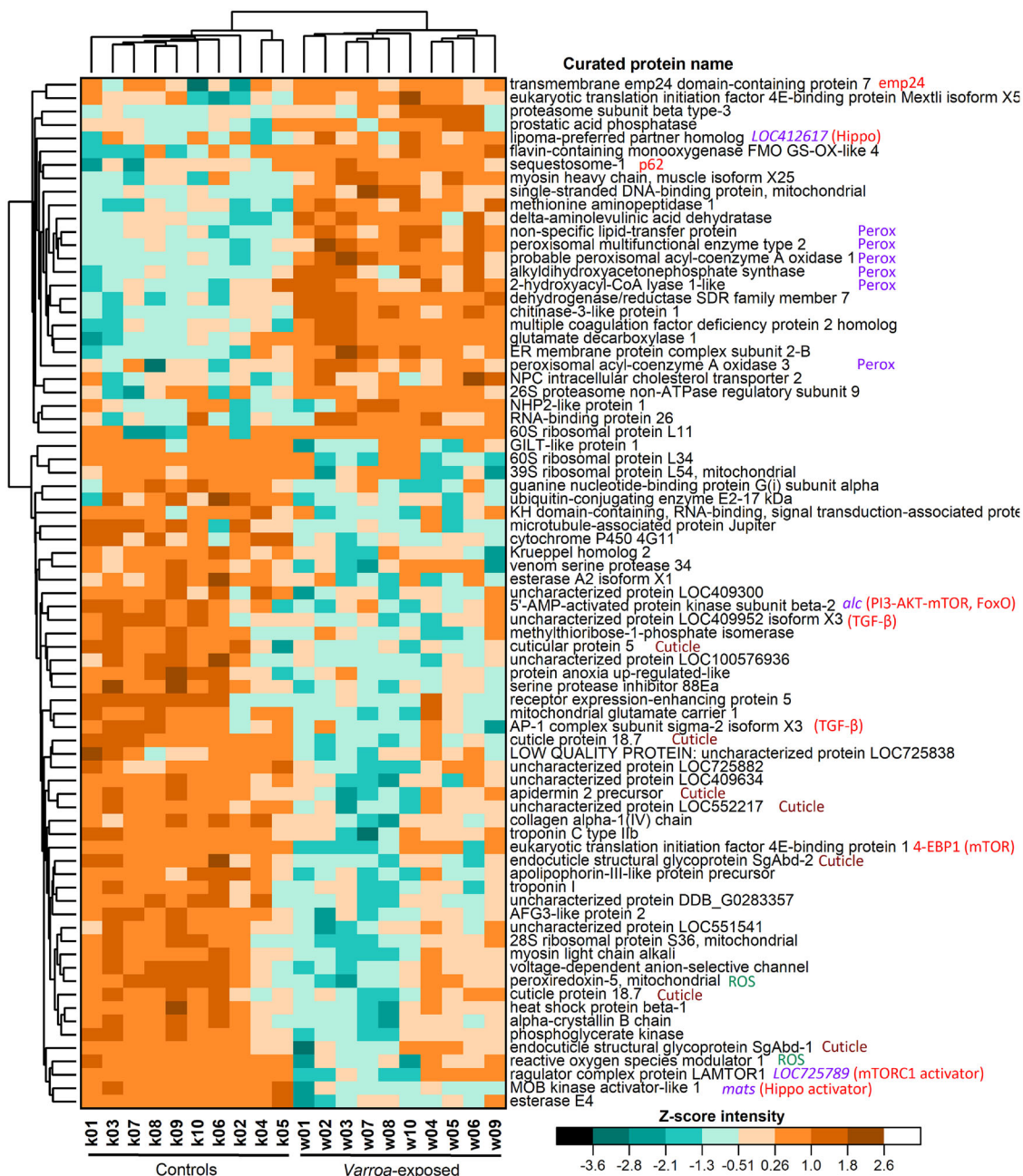


FIGURE 5 Heatmap visualization of the protein expression profiles of all 77 identified DEPs. LFQ values were z-score normalized prior to hierarchical clustering using Euclidean distance in Origin software. In addition to the curated protein names, specific marks are provided for selected proteins.

all the controls. However, there was a 38% loss of live mites during the exposure, which could be due to a decrease in their fitness during the experiment and the previous manipulation, that is, mite collection and placement in the experimental cage. Thus, this study involved *Varroa* parasitization of 11 bees by 31 fully functional mites that were alive at the end of the experiment out of 50 initial mites.

Overall, the results of this study complement those of other studies that have produced similar or different results and interpretations most likely due to differences in study designs and/or analyses of developmental stages, castes, and ages of the bees [23, 24]. For instance,

proteomic studies have analyzed the effect of *Varroa* parasitization at the time of emergence [18], in 10-day-old bees [22], in the hemolymph of 1-day-old bees [14], and in worker and drone pupae [52].

4.2 | Virus identification using proteomics

The *Varroa* mite is known to transmit DWV and cause the transition from latent to acute symptomatic DWV infection [2, 53]. In our experiment, we expected a transmitted DWV infection in *Varroa*-exposed

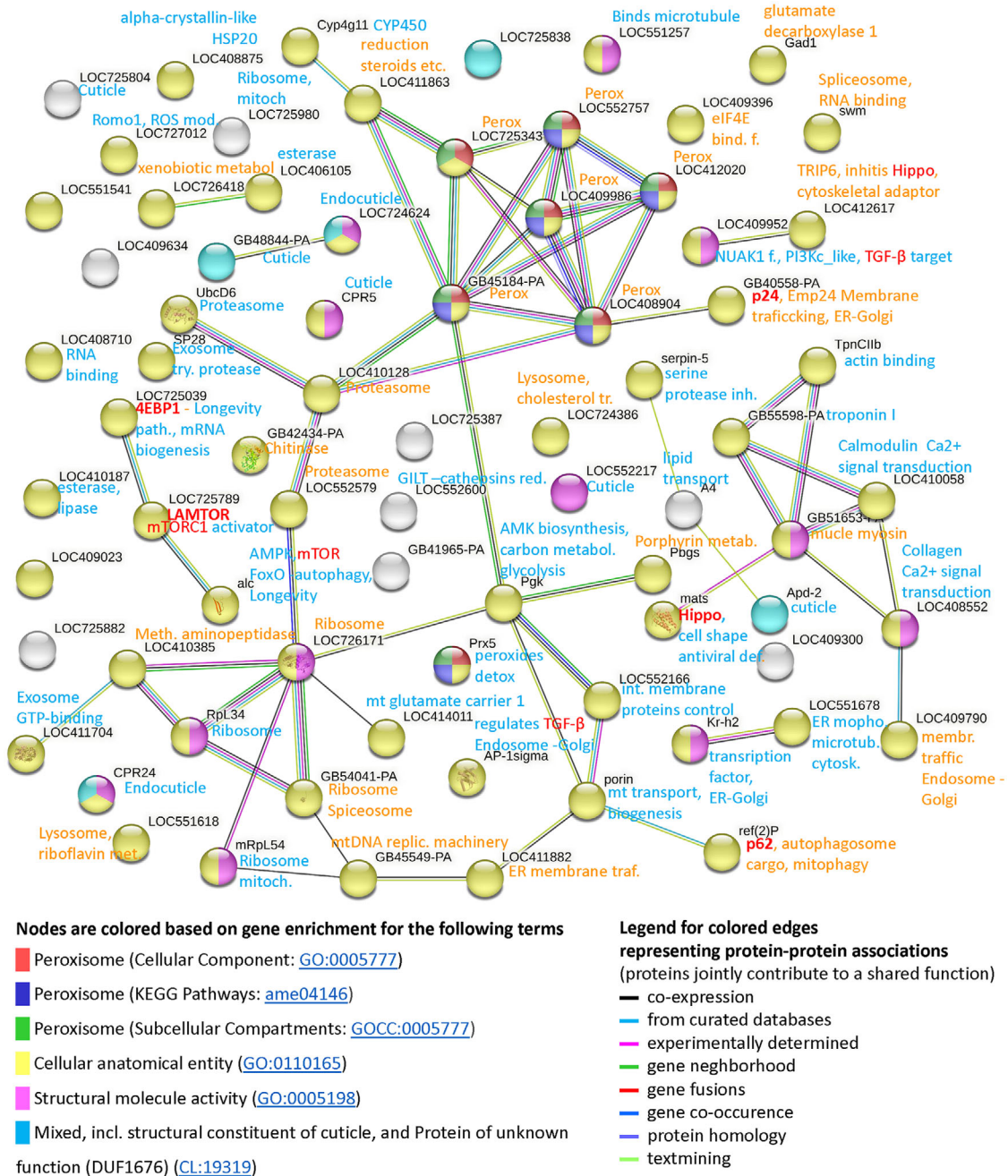


FIGURE 6 Functional protein association network of the significant DEPs determined via STRING. The description added to the gene symbols is derived from the data in Table S7, where annotations from bioinformatic resources (i.e., KEGGs, CDs, InterPro) are provided; orange indicates upregulated DEPs, while blue indicates downregulated DEPs. The red text highlights proteins of specific importance (see Discussion).

bees because the colony from which the mites came showed typical symptoms of varroosis. In addition, we found that the virome of honey bees in the *Varroa*-source colony contained a considerable portion of DWV-B, which is the main genotype associated with *Varroa* occurrence in Europe [54]. Indeed, our proteomic analysis confirmed the presence of DWV-B at high intensity in all 10 *Varroa*-exposed bees, while the virus was not detected in the control bees. For pathogen identification, we exploited the ability of proteomics to identify viral proteins instead of RNA-based qPCR. This approach allowed reliable identification of the viral proteins and host proteins from the same

data. Because we endeavored to avoid false-positive identifications of pathogen proteins in honey bees [55], we carefully verified proteins that were identified from the virus alongside honey bee and mite identification. An important issue to address was the occurrence of low LFQ values for the positives for DWV in 6/10 control bees, which could indicate latent DWV infection. However, detailed qualitative analysis revealed that these identifications were based exclusively on the MaxQuant matching-between-runs algorithm, while no MS/MS spectra were recorded for the controls. Thus, DWV in the controls could be considered false positives. In addition, a virus-focused analysis

indicated the presence of persistent viral infection of AmFV, AFV, BMLV, and KBV in some bees. The distribution of viral positives in the bee samples, particularly AmFV and KBV in the controls, suggested that these could be viral legacies that survived metamorphosis in some bees at the time of emergence [15].

4.3 | Pathways affected in bees

The most prominent group of proteins affected by *Varroa*-DWV exposure was a group of peroxisomal proteins, all of which were collectively upregulated. A greater abundance of peroxisomal proteins suggested the upregulation of fatty acid β -oxidation with acyl-CoA oxidase (*LOC412020*; *LOC552757*) and multifunctional enzyme type 2 (*LOC409986*), alpha-oxidation (*GB45184*), the biosynthesis of ether phospholipids (alkyldihydroxyacetonephosphate synthase; *LOC725343*), and lipid transfer (sterol carrier protein-2; *LOC408904*). In general, peroxisomes play key roles in diverse cellular processes, including fatty acid/lipid and amino acid metabolism and the maintenance of reactive oxygen species (ROS) [56–58]. Peroxisomes are functionally and in some cases physically associated with other organelles, including the endoplasmic reticulum (ER), lysosomes, autophagosomes, mitochondria, and lipid droplets [57, 59, 60]. However, peroxisomes are less known for functioning as cell signaling organelles, and ROS generated by peroxisomes can activate intracellular signaling cascades [58, 61]. The observed DEPs suggested that these peroxisome-related functions, whether affected by peroxisomes or those affecting peroxisomes, were modulated by pathogen exposure, as depicted in Figure 7.

Increased peroxisomal metabolism is associated with increased production of ROS that should be eliminated; however, the downregulation of peroxiredoxin-5 (*Prx5*) observed by us and others [62] suggests the opposite. This downregulation may be part of the response to inflammation [62]; in particular, a trade-off has been observed between the protective role of *Prx5* in immune and antioxidant functions [63]. Furthermore, one of the critical regulators of ROS is mitochondrial reactive oxygen species modulator 1 (*Romo1*; *LOC727012*), whose downregulation suggests an attempt to decrease cellular ROS levels [64] and reduce NF- κ B activation caused by DWV, as viral infections can increase oxidative stress through *Romo1* associated with NF- κ B [65]. Taken together, these results suggest that the downregulation of both *Prx5* and *Romo1* is associated with the immune response to DWV through the modulation of ROS levels, which act as active signaling molecules.

Upregulated transmembrane emp24 domain-containing protein 7 (*Tmed7*; *GB40558-PA*) is associated with Toll-like receptor 4 (TLR4)-mediated viral infection [66], as it is required for TLR4 trafficking from the ER through the Golgi to the cell surface [67]. Notably, emp24/p24 localizes to the ER, Golgi, COP vesicles, and peroxisomes [68]. Upregulated sequestosome-1 (*p62/SQSM1*; *ref(2)P*) is an autophagy receptor that connects autophagosomes to peroxisomes via ubiquitinated peroxisomal PEX5 [59]. Furthermore, p62 was found to be required for the autophagic clearance of ubiquitinated proteins and for the delivery of

ubiquitinated cargo to the proteasome, and its levels correlated with autophagic degradation [69]. Incidentally, according to our STRING analysis, proteasomal proteins (*LOC410128*; *LOC552579*) that are connected to peroxisomal proteins were upregulated.

In *Varroa*-exposed bees, we found substantial downregulation of the Regulator complex protein LAMTOR1 (*Lamtor1/p18*; *LOC725789*), which is involved in amino acid sensing and activation of mTORC1 [70, 71]. This finding suggested an increase in autophagy, particularly pexophagy [58, 72]. In addition, suppression of mTORC1 activity has been linked to alterations in cholesterol transport [73], the hijacking (i.e., upregulation of *NPC2*; *LOC724386*) of which is essential for infection and the life cycle [74]. Because of the link to the nutrient-sensing pathway, the collective changes in the mTORC1 pathway and peroxisomal β -oxidation [75] may explain the increase in the sphingolipid metabolic rate observed by Kunc et al. [22] in 10 day old parasitized bees.

A downregulated *Lamtor1/p18* was a key marker in our study. According to the STRING analysis, it interacts with eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1; *LOC725039*), which was downregulated to a similar extent. Importantly, 4E-BP1 harbors numerous phosphorylation sites, and this phosphorylation is regulated by the mTOR pathway [76], suggesting that this effect is important for our interpretation (Figure 7). mTORC1 stimulates translation via phosphorylation of 4E-BP1 [77, 78]. 4E-BP1 binds eIF4E, through which some viruses modulate protein synthesis; in particular, picornaviruses (which include DWV) induce the dephosphorylation of 4E-BP, leading to cap-dependent inhibition of protein synthesis [79]. According to our data, 4E-BP1 was downregulated and may have been affected by dephosphorylation triggered by viral activity [80–82]. Importantly, the mTORC1/4E-BP1 axis is related to TGF- β -induced pathways and has been identified as a key pathway involved in wound healing and repair [83, 84]. Injury-induced mTOR activation in epithelial cells is likely conserved across species, that is, insects and vertebrates [85]. Furthermore, *Lamtor1/p18* was linked in STRING to a regulatory subunit (*PRKAB2*; *alc*) of AMP-activated protein kinase (AMPK), which directly regulates cell growth and macroautophagy and microautophagy via mTORC1. This function has also been linked to the aforementioned p62 [86].

We found that some DEPs were directly related to TGF- β , a signaling pathway that regulates cell proliferation, differentiation, and apoptosis and is important for inflammation, various immune responses against microbes, and wound healing [87–90]. TGF- β has been suggested to play an important role in *Varroa*-parasitized bees at the time of emergence [18]. It has been shown that there is crosstalk between the TGF- β , Hippo, and mTORC1 pathways [91–93]. Thus, the downregulation of MOB kinase activator-like 1 (*MOB1*; *mats*), a key signaling adaptor of the Hippo pathway, which is essential for organ growth control and tissue homeostasis, is important in this context [94]. Since *MOB1* was previously shown to be upregulated in *Varroa*-DWV variant emerging bees [18], we suggest that differentially regulated *MOB1* may be associated with differentially altered innate immune defenses [95, 96] in two different developmental stages (3-day-old bees vs. emerging bees). Notably, modulation of the Hippo pathway interferes

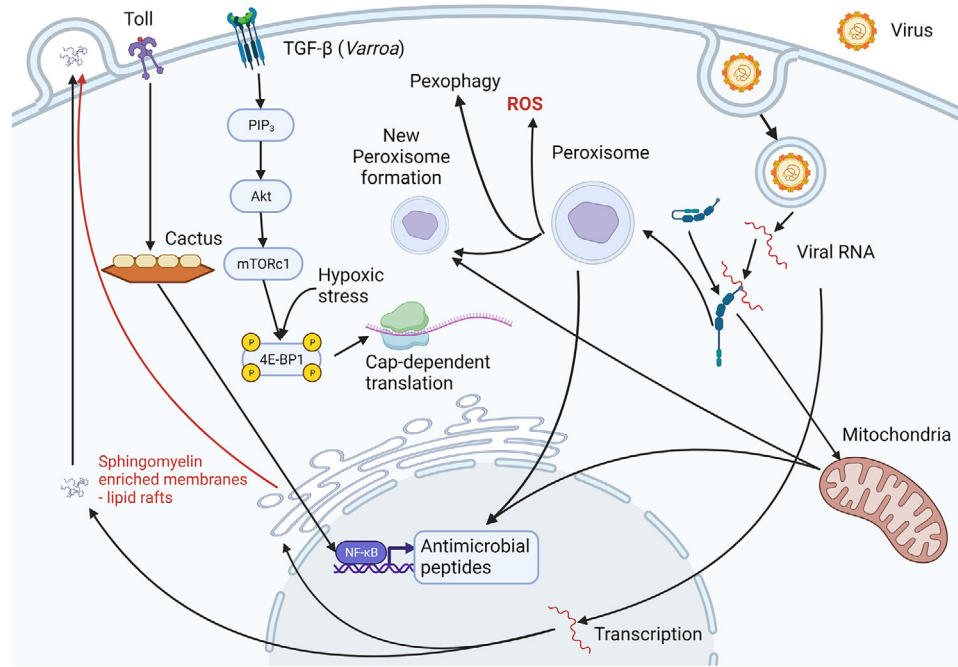


FIGURE 7 The overall scheme of host pathways affected by *Varroa*/DWV exposure. The scheme summarizes the results of this study and accounts for the previous studies that have been conducted. The effect of picornavirus multiplication on both peroxisomes and mitochondria is suggested, as is the cell antiviral response via the Toll signaling pathway. Furthermore, peroxisome signaling is associated with ROS production and pexophagy. In addition, the results indicated that the TGF- β pathway is affected by *Varroa* feeding on the host and plays a role in DWV infection. An important factor that is likely to affect viral replication is impaired autophagy, which is associated with stress (mainly due to *Varroa* parasitization) that affects mTORC1.

with antiviral defenses [97, 98], which are likely differentially regulated in the two different developmental stages.

The impact of *Varroa* and/or DWV on cyto-/exoskeleton reorganization has been previously identified, and the relevant markers in our study included changes in cuticle and cytoskeletal proteins, including muscle proteins [14, 18, 52, 99]. In this context, through direct interaction via STRING, we found MOB1 in the Hippo pathway [100]. Furthermore, the Lipoma-preferred partner TRIP6 [101; LOC412617], which is involved in the Hippo pathway, affects actin cytoskeletal reorganization, cell adhesion and migration [102]. The function of TRIP6 is linked to another protein (NUAK1; LOC409952) that is a TGF- β target. The downregulation of the microtubule-associated protein Jupiter (LOC551257) plays an important role in microtubule stability in *Drosophila* [103, 104]. Overall, changes related to cytoskeletal remodeling may be linked primarily to the DWV life cycle since viruses need to hijack cytoplasmic membrane trafficking machines for their own replication, assembly and release from a cell [105–109]. However, cytoskeleton remodeling is also due to *Varroa* syndrome affecting some of these pathways, and the interplay among these pathways leads to synergistic effects with DWV during parasite sucking of a host [18].

The number of downregulated proteins was related to the structural constituents of the cuticle (LOC724624, CPR24, CPR5, GB48844, LOC725804, LOC552217, *Apd-2*, and *Cyp4g11*). Like in our study, *Varroa*-parasitized purple-eye pupae exhibited a downregulated array of cuticular proteins [52]. Additionally, parasitization of honey bees by *Vairimorpha* (*Nosema ceranae* (5 days postinfection) downregulated

seven cuticle genes [110], some of which were the same as those in our study. Changes in exoskeletal proteins due to infection with DWV might be linked to weakening of the external barrier of a host to facilitate transmission [110]; however, transmission of this virus has been connected mainly to the ectoparasite *Varroa*, which creates wounds [111] and delays healing in a host [18]. Interestingly, differences in the expression of cuticle genes have been implicated in different hygienic behaviors of honey bees [112] and resistance to DWV infection [99].

An important change affecting immunity is the downregulation of the serine protease inhibitor 88Ea (*serpin-5*; *serpin-5*), which has been shown to regulate prophenoloxidase (PPO) activation and antimicrobial peptide pathway activation. *Serpin-5* has been shown to function as a secreted negative modulator of Toll signaling [113, 114]. It also acts as a component of the extracellular surveillance system in epithelial cells, and the mechanism has been linked to the Hippo pathway [115].

Overall, our observations summarized in Figure 7 are in agreement with the following previous results on the effect of *Varroa*–DWV on bees: (1) the levels of lipid metabolism-related proteins increase in parasitized worker pupae [52]; (2) the levels of proteins related to the metabolism of lipids and branched amino acids increase in the hemolymph of *Varroa*-parasitized bees [14]; and (3) the sphingolipid metabolism rate increases in *Varroa*-parasitized 10-day-old bees [22]. Some of the contrasting differences in expression levels nicely illustrate the influence of the experimental design. For instance, the opposite regulatory effects on *Lamtor1/p18* and *MOB1* illustrate the influence of the experimental design and the age of the bees analyzed

on the results; that is, 3-day-old with 3 days of *Varroa*-DWV exposure (this study) vs. 0-day-old (emerging) [18].

5 | CONCLUSION

This study investigated the effect of *Varroa*-DWV exposure on honey bees in a specific manipulative experiment. High-throughput proteomics revealed certain characteristics that allowed the identification of markers and pathways affected by *Varroa*-DWV interaction in 3-day-old bees postemergence. We found that a key set of affected proteins was associated with increased peroxisomal metabolism, and the other proteomic changes, such as ROS homeostasis/signaling, were also associated with peroxisomes. These results are consistent with the finding that peroxisomes play an important role in viral infections. In the future, it will be necessary to determine to what extent peroxisomes act as antiviral agents or whether they also exert proviral functions in honey bees parasitized by *Varroa*. These results also support the importance of TGF- β signaling in the *Varroa*-DWV interaction and reveal the link between the mTORC1 and Hippo pathways. The ability of *Varroa* to promote DWV infection likely occurs through the subversion of host autophagy caused by the modulation of apoptosis and proliferation associated with the suppression of wound healing and repair by *Varroa* and the reprogramming of cellular functions by DWV. Cytoskeletal remodeling and changes in cuticular proteins seem to be concomitant processes associated with viral infection and *Varroa* feeding on the host. The limitation of this study is that whole bees were analyzed, and in the future, it will be necessary to localize the changes to specific cell types and tissues.

AUTHOR CONTRIBUTIONS

Tomas Erban wrote the main manuscript, designed the study, and evaluated the data in detail. Klara Kadlikova processed the samples. Karel Harant and Pavel Talacko performed the nano-LC-MS/MS analysis. Jan Tachezy and Karel Harant contributed to the data evaluation. Dominika Kadleckova, Martina Salakova, and Ruth Tachezy performed and evaluated the virome analysis. Bruno Sopko created the scheme of affected pathways. Martin Markovic verified the references, edited the draft version and checked the formality. Jan Tachezy, Ruth Tachezy, Dominika Kadleckova, and Martina Salakova contributed to the writing. All the authors have read and approved the manuscript.

ACKNOWLEDGMENTS

The authors are grateful to the beekeepers for allowing us to collect the mite and bee samples. This research was supported by project no. QK1910018 (NAZV), Institutional Support for R&D no. MZE-RO0423 of the Ministry of Agriculture of the Czech Republic, the Charles University grant number SVV 260679 and ERDF project CePaViP no. CZ.02.1.01/0.0/0.0/16_019/0000759.

CONFLICT OF INTEREST STATEMENT

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

DATA AVAILABILITY STATEMENT

The raw data from the nanoLC-MS/MS runs, peak list files, search database, search engine parameters and zipped combined_txt from MaxQuant can be accessed via the complete MassIVE MSV000094006 submission (doi:10.25345/C5H12VJ8N) or ProteomeXchange (PXD049168). The reads from the virome analysis were deposited into the NCBI database (PRJNA943204, SRR23800994-5).

ORCID

Tomas Erban  <https://orcid.org/0000-0003-1730-779X>
 Dominika Kadleckova  <https://orcid.org/0000-0002-4312-7373>
 Bruno Sopko  <https://orcid.org/0000-0002-5580-1871>
 Karel Harant  <https://orcid.org/0000-0002-9654-5392>
 Pavel Talacko  <https://orcid.org/0000-0002-0943-4564>
 Martin Markovic  <https://orcid.org/0000-0003-3435-1215>
 Martina Salakova  <https://orcid.org/0000-0003-0827-1211>
 Klara Kadlikova  <https://orcid.org/0000-0002-9321-9380>
 Ruth Tachezy  <https://orcid.org/0000-0001-7689-9727>
 Jan Tachezy  <https://orcid.org/0000-0001-6976-8446>

REFERENCES

- Anderson, D. L., & Trueman, J. W. H. (2000). *Varroa jacobsoni* (Acari: Varroidea) is more than one species. *Experimental and Applied Acarology*, 24(3), 165–189. <https://doi.org/10.1023/A:1006456720416>
- Traynor, K. S., Mondet, F., de Miranda, J. R., Techer, M., Kowalik, V., Oddie, M. A. Y., Chantawannakul, P., & McAfee, A. (2020). *Varroa destructor*: A complex parasite, crippling honey bees worldwide. *Trends in Parasitology*, 36(7), 592–606. <https://doi.org/10.1016/j.jpt.2020.04.004>
- Australian Government (2022). *Varroa* mite (*Varroa destructor*): pest situation. National pest & disease outbreaks. (Accessed date: 9 March 2023). <https://www.outbreak.gov.au/current-responses-to-outbreaks/varroa-mite>
- Rooth, M. (2018). *Varroa* mite detected at Port of Melbourne on a ship from United States. *ABC Rural*, 29 June 2018. <https://www.abc.net.au/news/rural/2018-06-29/varroa-mite-detected-in-melbourne/9923972>
- Martin, S. J. (2001). The role of *Varroa* and viral pathogens in the collapse of honeybee colonies: a modelling approach. *Journal of Applied Ecology*, 38(5), 1082–1093. <https://doi.org/10.1046/j.1365-2664.2001.00662.x>
- Rosenkranz, P., Aumeier, P., & Ziegelmann, B. (2010). Biology and control of *Varroa destructor*. *Journal of Invertebrate Pathology*, 103(Supplement 1), S96–S119. <https://doi.org/10.1016/j.jip.2009.07.016>
- Wilfert, L., Long, G., Leggett, H. C., Schmid-Hempel, P., Butlin, R., Martin, S. J. M., & Boots, M. (2016). Deformed wing virus is a recent global epidemic in honeybees driven by *Varroa* mites. *Science*, 351(6273), 594–597. <https://doi.org/10.1126/science.aac9976>
- Roberts, J. M. K., Anderson, D. L., & Durr, P. A. (2018). Metagenomic analysis of *Varroa*-free Australian honey bees (*Apis mellifera*) shows a diverse Picornavirales virome. *Journal of General Virology*, 99(6), 818–826. <https://doi.org/10.1099/jgv.0.001073>
- Kadleckova, D., Tachezy, R., Erban, T. S., Deboutte, W., Nunvar, J., Salakova, M., & Matthijssens, J. (2022). The virome of healthy honey bee colonies: ubiquitous occurrence of known and new viruses in bee populations. *mSystems*, 7(3), e0007222. <https://doi.org/10.1128/mSystems.00072-22>
- Bowen-Walker, P. L., & Gunn, A. (2001). The effect of the ectoparasitic mite, *Varroa destructor* on adult worker honeybee (*Apis mellifera*) emergence weights, water, protein, carbohydrate, and lipid levels.

- Entomologia Experimentalis et Applicata*, 101(3), 207–217. <https://doi.org/10.1046/j.1570-7458.2001.00905.x>
11. Erban, T., Harant, K., Hubalek, M., Vitamvas, P., Kamler, M., Poltronieri, P., Tyl, J., Markovic, M., & Titera, D. (2015). In-depth proteomic analysis of *Varroa destructor*: Detection of DWV-complex, ABPV, VdMLV and honeybee proteins in the mite. *Scientific Reports*, 5(1), 13907. <https://doi.org/10.1038/srep13907>
 12. McAfee, A., Chan, Q. W. T., Evans, J., & Foster, L. J. (2017). A *Varroa destructor* protein atlas reveals molecular underpinnings of developmental transitions and sexual differentiation. *Molecular & Cellular Proteomics*, 16(12), 2125–2137. <https://doi.org/10.1074/mcp.RA117.000104>
 13. Ramsey, S. D., Ochoa, R., Bauchan, G., Gulbranson, C., Mowery, J. D., Cohen, A., Lim, D., Joklik, J., Cicero, J. M., Ellis, J. D., Hawthorne, D., & Vanengelsdorp, D. (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. <https://doi.org/10.1073/pnas.1818371116>
 14. Slowinska, M., Nynca, J., Bak, B., Wilde, J., Siuda, M., & Ciereszko, A. (2019). 2D-DIGE proteomic analysis reveals changes in haemolymph proteome of 1-day-old honey bee (*Apis mellifera*) workers in response to infection with *Varroa destructor* mites. *Apidologie*, 50(5), 632–656. <https://doi.org/10.1007/s13592-019-00674-z>
 15. Yang, X., & Cox-Foster, D. L. (2005). Impact of an ectoparasite on the immunity and pathology of an invertebrate: Evidence for host immunosuppression and viral amplification. *Proceedings of the National Academy of Sciences of the United States of America*, 102(21), 7470–7475. <https://doi.org/10.1073/pnas.0501860102>
 16. Di Prisco, G., Annoscia, D., Margiotta, M., Ferrara, R., Varricchio, P., Zanni, V., Caprio, E., Nazzi, F., & Pennacchio, F. (2016). A mutualistic symbiosis between a parasitic mite and a pathogenic virus undermines honey bee immunity and health. *Proceedings of the National Academy of Sciences of the United States of America*, 113(12), 3203–3208. <https://doi.org/10.1073/pnas.1523515113>
 17. Nazzi, F., Brown, S. P., Annoscia, D., Del Piccolo, F., Di Prisco, G., Varricchio, P., Della Vedova, G., Cattonaro, F., Caprio, E., & Pennacchio, F. (2012). Synergistic parasite-pathogen interactions mediated by host immunity can drive the collapse of honeybee colonies. *PLOS Pathogens*, 8(6), e1002735. <https://doi.org/10.1371/journal.ppat.1002735>
 18. Erban, T., Sopko, B., Kadlikova, K., Talacko, P., & Harant, K. (2019). *Varroa destructor* parasitism has a greater effect on proteome changes than the deformed wing virus and activates TGF- β signaling pathways. *Scientific Reports*, 9(1), 9400. <https://doi.org/10.1038/s41598-019-45764-1>
 19. Kuster, R. D., Boncristiani, H. F., & Rueppell, O. (2014). Immunogene and viral transcript dynamics during parasitic *Varroa destructor* mite infection of developing honey bee (*Apis mellifera*) pupae. *Journal of Experimental Biology*, 217(10), 1710–1718. <https://doi.org/10.1242/jeb.097766>
 20. Navajas, M., Migeon, A., Alaux, C., Martin-Magniette, M. L., Robinson, G. E., Evans, J. D., Cros-Arteil, S., Crauser, D., & Le Conte, Y. (2008). Differential gene expression of the honey bee *Apis mellifera* associated with *Varroa destructor* infection. *BMC Genomics*, 9(1), 301. <https://doi.org/10.1186/1471-2164-9-301>
 21. Johnson, R. M., Evans, J. D., Robinson, G. E., & Berenbaum, M. R. (2009). Changes in transcript abundance relating to colony collapse disorder in honey bees (*Apis mellifera*). *Proceedings of the National Academy of Sciences of the United States of America*, 106(35), 14790–14795. <https://doi.org/10.1073/pnas.0906970106>
 22. Kunc, M., Dobes, P., Ward, R., Lee, S., Cegan, R., Dostalkova, S., Holusova, K., Hurychova, J., Elias, S., Pindakova, E., Cukanova, E., Prodelalova, J., Petrivalsky, M., Danihlik, J., Havlik, J., Hobza, R., Kavanagh, K., & Hyrs, P. (2023). Omics-based analysis of honey bee (*Apis mellifera*) response to *Varroa* sp. parasitisation and associated factors reveals changes impairing winter bee generation. *Insect Biochemistry and Molecular Biology*, 152, 103877. <https://doi.org/10.1016/j.ibmb.2022.103877>
 23. Ryabov, E. V., Wood, G. R., Fannon, J. M., Moore, J. D., Bull, J. C., Chandler, D., Mead, A., Burroughs, N., & Evans, D. J. (2014). A virulent strain of deformed wing virus (DWV) of honeybees (*Apis mellifera*) prevails after *Varroa destructor*-mediated, or *in vitro*, transmission. *PLOS Pathogens*, 10(6), e1004230. <https://doi.org/10.1371/journal.ppat.1004230>
 24. Doublet, V., Poeschl, Y., Gogol-Doring, A., Alaux, C., Annoscia, D., Aurori, C., Barribeau, S. M., Bedoya-Reina, O. C., Brown, M. J. F., Bull, J. C., Flenniken, M. L., Galbraith, D. A., Genersch, E., Gisder, S., Grosse, I., Holt, H. L., Hultmark, D., Lattorff, H. M. G., Le Conte, Y., ... Grozinger, C. M. (2017). Unity in defence: Honeybee workers exhibit conserved molecular responses to diverse pathogens. *BMC Genomics*, 18(1), 207. <https://doi.org/10.1186/s12864-017-3597-6>
 25. Erban, T., Harant, K., Kamler, M., Markovic, M., & Titera, D. (2016). Detailed proteome mapping of newly emerged honeybee worker hemolymph and comparison with the red-eye pupal stage. *Apidologie*, 47(6), 805–817. <https://doi.org/10.1007/s13592-016-0437-7>
 26. Erban, T., Parizkova, K., Sopko, B., Talacko, P., Markovic, M., Jarosova, J., & Votypka, J. (2023). Imidacloprid increases the prevalence of the intestinal parasite *Lotmaria passim* in honey bee workers. *Science of The Total Environment*, 905, 166973. <https://doi.org/10.1016/j.scitotenv.2023.166973>
 27. Erban, T., Sopko, B., Talacko, P., Harant, K., Kadlikova, K., Halesova, T., Riddellova, K., & Pekas, A. (2019). Chronic exposure of bumblebees to neonicotinoid imidacloprid suppresses the entire mevalonate pathway and fatty acid synthesis. *Journal of Proteomics*, 196, 69–80. <https://doi.org/10.1016/j.jprot.2018.12.022>
 28. Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, 26(12), 1367–1372. <https://doi.org/10.1038/nbt.1511>
 29. Cox, J., Hein, M. Y., Lubner, C. A., Paron, I., Nagaraj, N., & Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & Cellular Proteomics*, 13(9), 2513–2526. <https://doi.org/10.1074/mcp.M113.031591>
 30. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., & Mann, M. (2011). Andromeda: A peptide search engine integrated into the MaxQuant environment. *Journal of Proteome Research*, 10(4), 1794–1805. <https://doi.org/10.1021/pr101065j>
 31. O'Leary, N. A., Wright, M. W., Brister, J. R., Ciuffo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., ... Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, 44(D1), D733–D745. <https://doi.org/10.1093/nar/gkv1189>
 32. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., & Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, 13(9), 731–740. <https://doi.org/10.1038/nmeth.3901>
 33. Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & von Mering, C. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47(D1), D607–D613. <https://doi.org/10.1093/nar/gky1131>
 34. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2023). The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest.

- Nucleic Acids Research*, 51(D1), D638–D646. <https://doi.org/10.1093/nar/gkac1000>
35. Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J., & Marchler-Bauer, A. (2020). CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Research*, 48(D1), D265–D268. <https://doi.org/10.1093/nar/gkz991>
 36. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., & Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research*, 51(D1), D587–D592. <https://doi.org/10.1093/nar/gkac963>
 37. Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., Bileschi, M. L., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunic, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Orengo, C. A., Pandurangan, A. P., Rivoire, C., ... Bateman, A. (2023). InterPro in 2022. *Nucleic Acids Research*, 51(D1), D418–D427. <https://doi.org/10.1093/nar/gkac993>
 38. Lim, M. Y., Paulo, J. A., & Gygi, S. P. (2019). Evaluating false transfer rates from the match-between-runs algorithm with a two-proteome model. *Journal of Proteome Research*, 18(11), 4020–4026. <https://doi.org/10.1021/acs.jproteome.9b00492>
 39. Conceicao-Neto, N., Yinda, K. C., Van Ranst, M., & Matthijssens, J. (2018). *NetoVIR: modular approach to customize sample preparation procedures for viral metagenomics*. In: A. Moya, & V. P. Brocal (Eds.), *The human virome* (pp. 85–95). New York, NY: Humana Press. https://doi.org/10.1007/978-1-4939-8682-8_7
 40. Babraham Bioinformatics. (2019). FastQC. Cambridge: Babraham Institute. (Accessed date: 9 March 2023). <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
 41. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
 42. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Pribelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, 19(5), 455–477. <https://doi.org/10.1089/cmb.2012.0021>
 43. Buchfink, B., Reuter, K., & Drost, H. G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature Methods*, 18(4), 366–368. <https://doi.org/10.1038/s41592-021-01101-x>
 44. Vasimuddin, M., Misra, S., Li, H., & Aluru, S. (2019). Efficient architecture-aware acceleration of BWA-MEM for multicore systems. In: L. Alvarez, L. Arantes, E. Arima, T. Benson, J. L. Bez, S. Bhalachandra, ... A. Zlateski (Eds.), 2019 IEEE 33rd International Parallel and Distributed Processing Symposium (IPDPS 2019) (pp. 314–324). <https://doi.org/10.1109/ipdps.2019.00041>
 45. Woodcroft, B. J. (2021). CoverM. GitHub. (Accessed date: 9 March 2023). <https://github.com/wwood/CoverM>
 46. Ondov, B. D., Bergman, N. H., & Phillippy, A. M. (2011). Interactive metagenomic visualization in a web browser. *BMC Bioinformatics*, 12(1), 385. <https://doi.org/10.1186/1471-2105-12-385>
 47. Skubnik, K., Novacek, J., Fuzik, T., Pridal, A., Paxton, R. J., & Plevka, P. (2017). Structure of deformed wing virus, a major honey bee pathogen. *Proceedings of the National Academy of Sciences of the United States of America*, 114(12), 3210–3215. <https://doi.org/10.1073/pnas.1615695114>
 48. Lanzi, G., de Miranda, J. R., Boniotti, M. B., Cameron, C. E., Lavazza, A., Capucci, L., Camazine, S. M., & Rossi, C. (2006). Molecular and biological characterization of deformed wing virus of honeybees (*Apis mellifera* L.). *Journal of Virology*, 80(10), 4998–5009. <https://doi.org/10.1128/JVI.80.10.4998-5009.2006>
 49. Remnant, E. J., Shi, M., Buchmann, G., Blacquiere, T., Holmes, E. C., Beekman, M., & Ashe, A. (2017). A diverse range of novel RNA viruses in geographically distinct honey bee populations. *Journal of Virology*, 91(16), e00158–00117. <https://doi.org/10.1128/jvi.00158-17>
 50. Reddy, K. E., Yoo, M.-S., Kim, Y.-H., Kim, N.-H., Jung, H.-N., Thao, L. T. B., Ramya, M., Doan, H. T. T., Nguyen, L. T. K., Jung, S.-C., & Kang, S.-W. (2014). Analysis of the RdRp, intergenic and structural polyprotein regions, and the complete genome sequence of Kashmir bee virus from infected honeybees (*Apis mellifera*) in Korea. *Virus Genes*, 49(1), 137–144. <https://doi.org/10.1007/s11262-014-1074-8>
 51. Organisation for Economic Co-operation and Development (OECD). (2017). Test No. 245: Honey bee (*Apis mellifera* L.), chronic oral toxicity test (10-day feeding). In: OECD (Ed.), *OECD guidelines for the testing of chemicals, section 2: Effects on biotic systems*. Paris: OECD Publishing. https://www.oecd-ilibrary.org/environment/test-no-245-honey-bee-apis-mellifera-l-chronic-oral-toxicity-test-10-day-feeding_9789264284081-en
 52. Surlis, C., Carolan, J. C., Coffey, M., & Kavanagh, K. (2018). Quantitative proteomics reveals divergent responses in *Apis mellifera* worker and drone pupae to parasitization by *Varroa destructor*. *Journal of Insect Physiology*, 107, 291–301. <https://doi.org/10.1016/j.jinsphys.2017.12.004>
 53. de Miranda, J. R., & Genersch, E. (2010). Deformed wing virus. *Journal of Invertebrate Pathology*, 103(Supplement 1), S48–S61. <https://doi.org/10.1016/j.jip.2009.06.012>
 54. Paxton, R. J., Schafer, M. O., Nazzi, F., Zanni, V., Annoscia, D., Marroni, F., Bigot, D., Laws-Quinn, E. R., Panziera, D., Jenkins, C., & Shafiey, H. (2022). Epidemiology of a major honey bee pathogen, deformed wing virus: Potential worldwide replacement of genotype A by genotype B. *International Journal for Parasitology: Parasites and Wildlife*, 18, 157–171. <https://doi.org/10.1016/j.ijppaw.2022.04.013>
 55. Foster, L. J. (2011). Interpretation of data underlying the link between colony collapse disorder (CCD) and an invertebrate iridescent virus. *Molecular & Cellular Proteomics*, 10(3), M110006387. <https://doi.org/10.1074/mcp.M110.006387>
 56. Morita, M., & Imanaka, T. (2019). The function of the peroxisome. In: T. Imanaka, & N. Shimozawa (Eds.), *Peroxisomes: biogenesis, function, and role in human disease* (pp. 59–104). Singapore: Springer. https://doi.org/10.1007/978-981-15-1169-1_4
 57. Pridie, C., Ueda, K., & Simmonds, A. J. (2020). Rosy beginnings: Studying peroxisomes in *Drosophila*. *Frontiers in Cell and Developmental Biology*, 8, 835. <https://doi.org/10.3389/fcell.2020.00835>
 58. Tripathi, D. N., & Walker, C. L. (2016). The peroxisome as a cell signaling organelle. *Current Opinion in Cell Biology*, 39, 109–112. <https://doi.org/10.1016/j.ceb.2016.02.017>
 59. Sargsyan, Y., & Thoms, S. (2020). Staying in healthy contact: How peroxisomes interact with other cell organelles. *Trends in Molecular Medicine*, 26(2), 201–214. <https://doi.org/10.1016/j.molmed.2019.09.012>
 60. Schrader, M., Kamoshita, M., & Islinger, M. (2020). Organelle interplay—peroxisome interactions in health and disease. *Journal of Inherited Metabolic Disease*, 43(1), 71–89. <https://doi.org/10.1002/jimd.12083>
 61. Ivashchenko, O., Van Veldhoven, P. P., Brees, C., Ho, Y. E.-S., Terlecky, S. R., & Fransen, M. (2011). Intraperoxisomal redox balance in mammalian cells: Oxidative stress and interorganellar cross-talk. *Molecular Biology of the Cell*, 22(9), 1440–1451. <https://doi.org/10.1091/mbc.E10-11-0919>
 62. Knoops, B., Argyropoulou, V., Becker, S., Ferte, L., & Kuznetsova, O. (2016). Multiple roles of peroxiredoxins in inflammation. *Molecules and Cells*, 39(1), 60–64. <https://doi.org/10.14348/molcells.2016.2341>
 63. Radyuk, S. N., Michalak, K., Klichko, V. I., Benes, J., & Orr, W. C. (2010). Peroxiredoxin 5 modulates immune response in *Drosophila*. *Biochimica et Biophysica Acta (BBA)—General Subjects*, 1800(11), 1153–1163. <https://doi.org/10.1016/j.bbagen.2010.06.010>

64. Shin, J. A., Chung, J. S., Cho, S.-H., Kim, H. J., & Yoo, Y. D. (2013). Romo1 expression contributes to oxidative stress-induced death of lung epithelial cells. *Biochemical and Biophysical Research Communications*, 439(2), 315–320. <https://doi.org/10.1016/j.bbrc.2013.07.012>
65. Amini, M. A., Karimi, J., Talebi, S. S., & Piri, H. (2022). The association of COVID-19 and reactive oxygen species modulator 1 (ROMO1) with oxidative stress. *Chonnam Medical Journal*, 58(1), 1–5. <https://doi.org/10.4068/cmj.2022.58.1.1>
66. Olejnik, J., Hume, A. J., & Muhlberger, E. (2018). Toll-like receptor 4 in acute viral infection: Too much of a good thing. *PLOS Pathogens*, 14(12), e1007390. <https://doi.org/10.1371/journal.ppat.1007390>
67. Liaunardy-Jopeace, A., Bryant, C. E., & Gay, N. J. (2014). The COP II adaptor protein TMED7 is required to initiate and mediate the delivery of TLR4 to the plasma membrane. *Science Signaling*, 7(336), ra70. <https://doi.org/10.1126/scisignal.2005275>
68. Saraya, R., Veenhuis, M., & van der Klei, I. J. (2010). Peroxisomes as dynamic organelles: Peroxisome abundance in yeast. *The FEBS Journal*, 277(16), 3279–3288. <https://doi.org/10.1111/j.1742-4658.2010.07740.x>
69. Lippai, M., & Low, P. (2014). The role of the selective adaptor p62 and ubiquitin-like proteins in autophagy. *BioMed Research International*, 2014, 832704. <https://doi.org/10.1155/2014/832704>
70. Bar-Peled, L., Schweitzer, L. D., Zoncu, R., & Sabatini, D. M. (2012). Ragulator is a GEF for the Rag GTPases that signal amino acid levels to mTORC1. *Cell*, 150(6), 1196–1208. <https://doi.org/10.1016/j.cell.2012.07.032>
71. Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A. L., Nada, S., & Sabatini, D. M. (2010). Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell*, 141(2), 290–303. <https://doi.org/10.1016/j.cell.2010.02.024>
72. Eun, S. Y., Lee, J. N., Nam, I.-K., Liu, Z.-q., So, H.-S., Choe, S.-K., & Park, R. K. (2018). PEX5 regulates autophagy via the mTORC1-TFEB axis during starvation. *Experimental & Molecular Medicine*, 50(4), 1–12. <https://doi.org/10.1038/s12276-017-0007-8>
73. Eid, W., Dauner, K., Courtney, K. C., Gagnon, A., Parks, R. J., Sorisky, A., & Zha, X. (2017). mTORC1 activates SREBP-2 by suppressing cholesterol trafficking to lysosomes in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, 114(30), 7999–8004. <https://doi.org/10.1073/pnas.1705304114>
74. Glitscher, M., & Hildt, E. (2021). Endosomal cholesterol in viral infections—a common denominator? *Frontiers in Physiology*, 12, 750544. <https://doi.org/10.3389/fphys.2021.750544>
75. Li, N. A., Hua, B., Chen, Q., Teng, F., Ruan, M., Zhu, M., Zhang, L. I., Huo, Y., Liu, H., Zhuang, M., Shen, H., & Zhu, H. (2022). A sphingolipid-mTORC1 nutrient-sensing pathway regulates animal development by an intestinal peroxisome relocation-based gut-brain crosstalk. *Cell Reports*, 40(4), 111140. <https://doi.org/10.1016/j.celrep.2022.111140>
76. Wang, X., Beugnet, A., Murakami, M., Yamanaka, S., & Proud, C. G. (2005). Distinct signaling events downstream of mTOR cooperate to mediate the effects of amino acids and insulin on initiation factor 4E-binding proteins. *Molecular and Cellular Biology*, 25(7), 2558–2572. <https://doi.org/10.1128/MCB.25.7.2558-2572.2005>
77. Thedieck, K., & Hall, M. N. (2010). Translational control by amino acids and energy. In: R. A. Bradshaw, & E. A. Dennis (Eds.), *Handbook of cell signaling*, Vol. 3, 2nd edn. (pp. 2285–2293). Amsterdam: Elsevier/Academic Press. <https://doi.org/10.1016/B978-0-12-374145-5.00274-6>
78. Gingras, A.-C., Gygi, S. P., Raught, B., Polakiewicz, R. D., Abraham, R. T., Hoekstra, M. F., Aebersold, R., & Sonenberg, N. (1999). Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes & Development*, 13(11), 1422–1437. <https://doi.org/10.1101/gad.13.11.1422>
79. Karaki, S., Andrieu, C., Ziouziou, H., & Rocchi, P. (2015). The eukaryotic translation initiation factor 4E (eIF4E) as a therapeutic target for cancer. *Advances in Protein Chemistry and Structural Biology*, 101, 1–26. <https://doi.org/10.1016/bs.apcsb.2015.09.001>
80. Perez-Gil, G., Landa-Cardena, A., Coutino, R., Garcia-Roman, R., Sampieri, C. L., Mora, S. I., & Montero, H. (2015). 4EBP1 is dephosphorylated by respiratory syncytial virus infection. *Intervirology*, 58(4), 205–208. <https://doi.org/10.1159/000435774>
81. Connor, J. H., & Lyles, D. S. (2002). Vesicular stomatitis virus infection alters the eIF4F translation initiation complex and causes dephosphorylation of the eIF4E binding protein 4E-BP1. *Journal of Virology*, 76(20), 10177–10187. <https://doi.org/10.1128/jvi.76.20.10177-10187.2002>
82. Stern-Ginossar, N., Thompson, S. R., Mathews, M. B., & Mohr, I. (2019). Translational control in virus-infected cells. *Cold Spring Harbor Perspectives in Biology*, 11(3), a033001. <https://doi.org/10.1101/cshperspect.a033001>
83. Woodcock, H. V., Eley, J. D., Guillotin, D., Plate, M., Nanthakumar, C. B., Martufi, M., Peace, S., Joberty, G., Poekkel, D., Good, R. B., Taylor, A. R., Zinn, N., Redding, M., Forty, E. J., Hynds, R. E., Swanton, C., Karsdal, M., Maher, T. M., Fisher, A., ... Chambers, R. C. (2019). The mTORC1/4E-BP1 axis represents a critical signaling node during fibrogenesis. *Nature Communications*, 10(1), 6. <https://doi.org/10.1038/s41467-018-07858-8>
84. Kisseleva, T., & Brenner, D. A. (2008). Mechanisms of fibrogenesis. *Experimental Biology and Medicine*, 233(2), 109–122. <https://doi.org/10.3181/0707-MR-190>
85. Wang, J., Cui, B., Chen, Z., & Ding, X. (2022). The regulation of skin homeostasis, repair and the pathogenesis of skin diseases by spatiotemporal activation of epidermal mTOR signaling. *Frontiers in Cell and Developmental Biology*, 10, 950973. <https://doi.org/10.3389/fcell.2022.950973>
86. Mihaylova, M. M., & Shaw, R. J. (2011). The AMPK signalling pathway coordinates cell growth, autophagy and metabolism. *Nature Cell Biology*, 13(9), 1016–1023. <https://doi.org/10.1038/ncb2329>
87. Chaudhury, A., & Howe, P. H. (2009). The tale of transforming growth factor-beta (TGFβ) signaling: A soigne enigma. *IUBMB Life*, 61(10), 929–939. <https://doi.org/10.1002/iub.239>
88. Ramirez, H., Patel, S. B., & Pastar, I. (2014). The role of TGFβ signaling in wound epithelialization. *Advances in Wound Care*, 3(7), 482–491. <https://doi.org/10.1089/wound.2013.0466>
89. Sanjabi, S., Oh, S. A., & Li, M. O. (2017). Regulation of the immune response by TGF-β: From conception to autoimmunity and infection. *Cold Spring Harbor Perspectives in Biology*, 9(6), a022236. <https://doi.org/10.1101/cshperspect.a022236>
90. Upadhyay, A., Moss-Taylor, L., Kim, M.-J., Ghosh, A. C., & O'Connor, M. B. (2017). TGF-β family signaling in drosophila. *Cold Spring Harbor Perspectives in Biology*, 9(9), a022152. <https://doi.org/10.1101/cshperspect.a022152>
91. Lamouille, S., & Derynck, R. (2007). Cell size and invasion in TGF-β-induced epithelial to mesenchymal transition is regulated by activation of the mTOR pathway. *Journal of Cell Biology*, 178(3), 437–451. <https://doi.org/10.1083/jcb.200611146>
92. Honda, D., Okumura, M., & Chihara, T. (2023). Crosstalk between the mTOR and Hippo pathways. *Development, Growth & Differentiation*, 65(6), 337–347. <https://doi.org/10.1111/dgd.12867>
93. Zhang, S., Liang, S., Wu, D., Guo, H., Ma, K., & Liu, L. (2021). LncRNA coordinates Hippo and mTORC1 pathway activation in cancer. *Cell Death & Disease*, 12(9), 822. <https://doi.org/10.1038/s41419-021-04112-w>
94. Kulaberoglu, Y., Lin, K., Holder, M., Gai, Z., Gomez, M., Assefa Shifa, B., Mavis, M., Hoa, L., Sharif, A. A. D., Lujan, C., Smith, E. S. J., Bjedov, I., Tapon, N., Wu, G., & Hergovich, A. (2017). Stable MOB1 interaction with Hippo/MST is not essential for development and tissue growth

- control. *Nature Communications*, 8(1), 695. <https://doi.org/10.1038/s41467-017-00795-y>
95. Liu, B. O., Zheng, Y., Yin, F., Yu, J., Silverman, N., & Pan, D. (2016). Toll receptor-mediated Hippo signaling controls innate immunity in *Drosophila*. *Cell*, 164(3), 406–419. <https://doi.org/10.1016/j.cell.2015.12.029>
 96. Huang, Y., Ma, F.-t., & Ren, Q. (2020). Function of the MOB kinase activator-like 1 in the innate immune defense of the oriental river prawn (*Macrobrachium nipponense*). *Fish & Shellfish Immunology*, 102, 440–448.
 97. Munoz-Wolf, N., & Lavelle, E. C. (2017). Hippo interferes with antiviral defences. *Nature Cell Biology*, 19(4), 267–269. <https://doi.org/10.1038/ncb3502>
 98. Wang, Z., Lu, W., Zhang, Y., Zou, F., Jin, Z., & Zhao, T. (2020). The Hippo pathway and viral infections. *Frontiers in Microbiology*, 10, 3033. <https://doi.org/10.3389/fmicb.2019.03033>
 99. Weaver, D. B., Cantarel, B. L., Elsik, C. G., Boncristiani, D. L., & Evans, J. D. (2021). Multi-tiered analyses of honey bees that resist or succumb to parasitic mites and viruses. *BMC Genomics*, 22(1), 720. <https://doi.org/10.1186/s12864-021-08032-z>
 100. Yu, F.-X., & Guan, K.-L. (2013). The Hippo pathway: Regulators and regulations. *Genes & Development*, 27(4), 355–371. <https://doi.org/10.1101/gad.210773.112>
 101. Dutta, S., Mana-Capelli, S., Paramasivam, M., Dasgupta, I., Cirka, H., Billiar, K., & Mccollum, D. (2018). TRIP6 inhibits Hippo signaling in response to tension at adherens junctions. *EMBO Reports*, 19(2), 337–350. <https://doi.org/10.15252/embr.201744777>
 102. Lin, V. T. G., & Lin, F.-T. (2011). TRIP6: An adaptor protein that regulates cell motility, antiapoptotic signaling and transcriptional activity. *Cellular Signalling*, 23(11), 1691–1697. <https://doi.org/10.1016/j.cellsig.2011.06.004>
 103. Karpova, N., Bobinnec, Y., Fouix, S., Huitorel, P., & Debec, A. (2006). Jupiter, a new *Drosophila* protein associated with microtubules. *Cell Motility and the Cytoskeleton*, 63(5), 301–312. <https://doi.org/10.1002/cm.20124>
 104. Martinez, D., Zhu, M., Guidry, J. J., Majeste, N., Mao, H., Yanofsky, S. T., Tian, X., & Wu, C. (2021). Mask, the *Drosophila* ankyrin repeat and KH domain-containing protein, affects microtubule stability. *Journal of Cell Science*, 134(20), jcs258512. <https://doi.org/10.1242/jcs.258512>
 105. Walsh, D., & Naghavi, M. H. (2019). Exploitation of cytoskeletal networks during early viral infection. *Trends in Microbiology*, 27(1), 39–50. <https://doi.org/10.1016/j.tim.2018.06.008>
 106. Marsh, M. (2005). *Membrane trafficking in viral replication*. Berlin, Heidelberg: Springer. <https://doi.org/10.1007/b138037>
 107. Hernandez-Gonzalez, M., Laroque, G., & Way, M. (2021). Viral use and subversion of membrane organization and trafficking. *Journal of Cell Science*, 134(5), jcs252676. <https://doi.org/10.1242/jcs.25267>
 108. Taylor, M. P., Koyuncu, O. O., & Enquist, L. W. (2011). Subversion of the actin cytoskeleton during viral infection. *Nature Reviews Microbiology*, 9(6), 427–439. <https://doi.org/10.1038/nrmicro2574>
 109. Seo, D., & Gammon, D. B. (2022). Manipulation of host microtubule networks by viral microtubule-associated proteins. *Viruses*, 14(5), 979. <https://doi.org/10.3390/v14050979>
 110. Badaoui, B., Fougeroux, A., Petit, F., Anselmo, A., Gorni, C., Cucurachi, M., Cersini, A., Granato, A., Cardeti, G., Formato, G., Mutinelli, F., Giuffra, E., Williams, J. L., & Botti, S. (2017). RNA-sequence analysis of gene expression from honeybees (*Apis mellifera*) infected with *Nosema ceranae*. *PLOS ONE*, 12(3), e0173438. <https://doi.org/10.1371/journal.pone.0173438>
 111. Kanbar, G., & Engels, W. (2004). Visualisation by vital staining with trypan blue of wounds punctured by *Varroa destructor* mites in pupae of the honey bee (*Apis mellifera*). *Apidologie*, 35(1), 25–29. <https://doi.org/10.1051/apido:2003057>
 112. Parker, R., Guarna, M. M., Melathopoulos, A. P., Moon, K.-M., White, R., Huxter, E., Pernal, S. F., & Foster, L. J. (2012). Correlation of proteome-wide changes with social immunity behaviors provides insight into resistance to the parasitic mite, *Varroa destructor*, in the honey bee (*Apis mellifera*). *Genome Biology*, 13(9), R81. <https://doi.org/10.1186/gb-2012-13-9-r81>
 113. Li, J., Ma, L., Lin, Z., Zou, Z., & Lu, Z. (2016). Serpin-5 regulates prophenoloxidase activation and antimicrobial peptide pathways in the silkworm, *Bombyx mori*. *Insect Biochemistry and Molecular Biology*, 73, 27–37. <https://doi.org/10.1016/j.ibmb.2016.04.003>
 114. An, C., & Kanost, M. R. (2010). *Manduca sexta* serpin-5 regulates prophenoloxidase activation and the Toll signaling pathway by inhibiting hemolymph proteinase HP6. *Insect Biochemistry and Molecular Biology*, 40(9), 683–689. <https://doi.org/10.1016/j.ibmb.2010.07.001>
 115. Katsukawa, M., Ohsawa, S., Zhang, L., Yan, Y., & Igaki, T. (2018). Serpin facilitates tumor-suppressive cell competition by blocking Toll-mediated Yki activation in *Drosophila*. *Current Biology*, 28(11), 1756–1767.e6.e1756. <https://doi.org/10.1016/j.cub.2018.04.022>

SUPPORTING INFORMATION

Additional supporting information may be found online <https://doi.org/10.1002/pmhc.202300312> in the Supporting Information section at the end of the article.

How to cite this article: Erban, T., Kadleckova, D., Sopko, B., Harant, K., Talacko, P., Markovic, M., Salakova, M., Kadlikova, K., Tachezy, R., & Tachezy, J. (2024). *Varroa destructor* parasitism and Deformed wing virus infection in honey bees are linked to peroxisome-induced pathways. *Proteomics*, e2300312. <https://doi.org/10.1002/pmhc.202300312>