# Comparing virome composition between honey bees (*Apis mellifera* L.) and small carpenter bees (*Ceratina calcarata* Robertson) in response to various agricultural landscapes

Zixiao Zhao[1]; M. Catherine Farrell[2]; Jayla Marvin[2] ; Sandra M. Rehan[3]; Hongmei Li-Byarlay[1,2 *]

[1]Agricultural Research Development Program, Central State University, Wilberforce, OH, USA

[2]Department of Agricultural and Life Sciences, Central State Univesity, Wilberforce, OH, USA

[3]Department of Biology, York University, Toronto, ON, Canada

**Author contact information**

ZZ: zxzhao5@gmail.com

https://orcid.org/0000-0002-3627-5503

MCF: mfarrell@centralstate.edu

https://orcid.org/0000-0002-0081-3285

JM: JaylaMarvin832@gmail.com

SMR: sandra.rehan@gmail.com

https://orcid.org/0000-0002-6441-5155

HL-B: Hli-byarlay@centralstate.edu

https://orcid.org/0000-0001-6579-5172

***Corresponding Author**

Hongmei Li-Byarlay

Hli-byarlay@centralstate.edu

## Abstract:

Managed pollinators, such as honeybees (*Apis mellifera* L.) and wild pollinators, such as small carpenter bees (*Ceratina calcarata* Robertson), are important for food and agriculture in North America. Viral prevalence as a pathogenic stress and agricultural landscapes such as organic and conventional farms as foraging resources for pollinators are important for their health. However, it is unclear if the virus is distributed differently between managed and wild pollinators when they share similar foraging area resources. We investigated the virome composition of *Apis* and *Ceratina* collected from conventional, organic farms, or roadside (as a control without agricultural activity) using RNA-Seq analysis. The result demonstrates that honeybees carry a wide range of insect and plant viruses. The black queen cell virus is the only pathogenic virus discovered from both species in one conventional farm. Three novel insect viruses (1 entomopoxvirus and 2 iflaviruses) were detected in small carpenter bees, 1 novel iflavirus of which was identified in honey bee samples. The biological and ecological significance of novel viruses are yet to be studied. The results of our project shed light on the status of the biotic stress of viruses on pollinators.

## Keywords:

Honey bees, agricultural landscape, transcriptomics, bee virus, iflavirus, entomopoxvirus

## Introduction:

Animal pollinators have critical ecological functions in maintaining the diversity of global ecosystems [1, 2]. About 87.5% of flowering plants rely on animal pollinators to support reproduction [3]. In agriculture, approximately 75% of crop species require pollinators, contributing about 30-35% of total crop production yield [4, 5]. Bees (Hymenoptera: Apidae) are considered the most abundant pollinators. Honey bee species (*Apis* spp.), with a global distribution due to domestication, contribute about 80% of total pollination and create about $700 million value annually in the US [6]. However, the ecological significance of non-*Apis* wild bee species shouldn't be neglected. It is estimated over 17,000 wild bee species worldwide, with diverse social structures from eusocial to solitary. Wild bees provide pollination services to many crops, especially when honey bees are not available [7, 8]. For last several decades, both wild and honey bees are suffering from colony declines [9-12], which not only causes economical losses to apiary industry, but more importantly, reduces pollination service and imposes potential risks for global food security.

The causes for pollinator colony declines are complicated. Factors include habitat loss and fragmentation, increasing use of chemical pesticides and other agrochemcials, introductions of pests and pathogens, and climate changes [11, 13]. Among those stressors, pathogenic viruses are considered the direct and the most significant harrzard to both wild and honey bee populations. Currently identified bee pathogenic viruses belong to positive-sense single-stranded RNA (+ss RNA) virus category. The majority of the pathogenic bee viruses are from two families under the order Picornavirales: Iflaviridae and Dicistroviridae. Iflaviridae has only one genus: *Iflavirus*, which contains Sacbrood virus (SBV), Deformed wing virus (DWV), Varroa destructor virus-1 (VDV-1), and Slow bee paralysis virus (SBPV). While the family

75 Dicistroviridae contains Black queen cell virus (BQCV), Kashmir bee virus (KBV), Acute
76 bee paralysis virus (ABPV), and Israeli acute paralysis virus (IAPV) [14]. Two extra
77 pathogenic viruses of bees: Chronic bee paralysis virus (CBPV) is an unclassified virus
78 with two genomic RNA segments [15], and Lake Sinai viruses (LSVs) are from a
79 separate order Nodamuvirales [14], with 8 identified linages [16]. The symptoms
80 caused by pathogenic viruses vary, including body deformities, color changed,
81 uncapped cell with larvae or pupae, paralysis adults, which eventually increase
82 mortality and the degradation of the colony. IAPV infections have been discovered
83 association with colony collapse disorder (CCD), which responsible for large-scale
84 honey bee colony loss in the US [17].
85 Despite the focus on honey bees, the pathogenic viruses can also infect wild bee
86 species [18-20]. For example, DWV have been reported infecting a wide range of
87 hymenopteran insects, including some social and solitary bees and several non-bee
88 wasp species [18, 20, 21]. Dolezal et al. [19] reported DWV as well as SBV were
89 widely detected in wild bees collected in Iowa, while IAPV, LSV and BQCV were also
90 detected in wild be species. Furthermore, DWV, BQCV and SBV can be deposited to
91 pollen and remain infective, which can contribute to inter-taxa transmission among
92 bees, wild bees, and wasps [20]. These studies suggested virus can be a potential
93 factor responsible for colony declines in wild bee species.
94 Agriculture are considered one of the major reason for colony declines. It have
95 converted natural ecosystems to farmland primary for food and fiber productions.
96 This process is accompanied by loss of biodiversity and use of agrochemicals causing
97 habitat loss, fragmentation and degradation. Agriculture can also alter host-pathogen
98 interactions at multiple levels. Study in the plants demonstrated the revalence of
99 plant viruses is negatively correlated with the biodiversity in agricultural landscape

100 [22]. However, whether agricultural activities affect viral prevalence in insects,
101 especially in wild and honey bees, is unknown.

102 The small carpenter bee, *Ceratina calcarata* Robertson, is a wild bee species
103 native to eastern part of North America. It is considered as indicator species and
104 pollinators for several native and agricultural plants [23]. Due to the feature of
105 subsociality, *C. calcarata* is used for studying pollinator ecology, behavior, evolution
106 and nutrition [24-27]. Little is known about the virology of this species. The impact of
107 viruses to *C. calcarata* populations and whether the viral prevalence is related to
108 agricultural activities are not clear.

109 Here we presented our metagenomic studies on viruses associated with western
110 honey bee (*A. mellifera* L.) and *C. calcarata* in response to different agricultural
111 intensification. Bees were collected from three landscape types representing levels of
112 agricultural intensification: conventional farming, organic farming and roadside with
113 no agricultural activities. The viral RNA was extracted from collected samples and
114 subjected to high-throughput sequencing. We analyzed known and novel viruses
115 associated with both pollinators, which are highly specific to location and pollinator
116 species. Our study provided new methodology and unique knowledge for
117 pollinator-virus interactions.

## 118 Results:

119 *Reconstruct virome from viral RNA-Seq reads*

120 The RNA-Seq of extracted viral RNA generates approximately 17 to 36 million
121 raw read pairs for each *A. mellifera* and *C. calcarata* sample. After quality trimming
122 and partially removing reads from respective host genome (Amel HAv3.1
123 PRJNA471592 and Ccalc.v3 PRJNA791561), 5 to 25 million pairs were kept for each
124 sample (**Supplementary file S1**). Trinity *de nove* assembled 366,289 cotings for *A.*

125 *mellifera* and 203,317 contigs for *C. calcarata* virome, which consists of 188,433 and
126 91,314 non-redundant clusters. The assembled contigs were subject to BLASTN
127 searches against non-redundant database (nr), which claimed 541 clusters and 1187
128 contigs from *A. mellifera* assembly poteintially belong to viruses. In contrast, 62
129 clusters and 84 contigs from *C. calcarata* assembly are potentially from viruses. A
130 large portion of assembled contigs of are from either host genes or the plants they
131 have visited.
132
133 *Viruses infecting bees and other insects detected in A. mellifera and C. calcarata*
134 *virome assemblies*
135 BLASTN search identified four known bee pathogenic viruses from *A. mellifera*
136 virome: BQCV, DWV, SBV, and LSV3 with fully assembled genome sequences. Two
137 other members of the LSV group: LSV1 and LSV2 were detected. But their sequences
138 were only partially assembled. Positive and negative strands were detected in all of
139 the above viruses except LSV2, from which only positive strand is detected. We also
140 detected *Apis mellifera* filamentous virus (AmFV) [28] and Hubei partiti-like virus 34
141 [29] in *A. mellifera*. The AmFV were partially assembled and the Hubei partiti-like
142 virus 34 demonstrated a fully assembled genome (**Table 1A**). In *C. calcarata*, BQCV
143 was the only detected bee pathogenic virus and appeared as partial-assembled as
144 negative-strand contig (**Table 1B**).
145
146 *Plant viruses discovered in A. mellifera and C. calcarata virome assemblies*
147 A variety of plant viruses were detected in virome assemblies of both species. In
148 *A. mellifera,* white clover cryptic virus 1 (WCCV1) and white clover cryptic virus 2
149 (WCCV2) from genus *Alphapartitivirus* (family: Partitiviridae, +ssRNA), as well as
150 white clover mosaic virus (WCMV) genus *Potexvirus* (family: Alphaflexiviridae,

151 +ssRNA) were fully assembled. Blackberry chlorotic ringspot virus (BCRV) from genus
152 *Ilavirus* (family: Bromoviridae, +ssRNA) and *Raphanus sativus* cryptic virus 1 (RsPV1)
153 from Partitiviridae (+ssRNA) have one genomic segment fully assembled, while other
154 segments were either partially assembled or completely missing. Partially assembled
155 plant viruses include 7 +ssRNA viruses: peanut stunt virus (PSV) and tobacco ringspot
156 virus (TRSV) from Bromoviridae, clover yellow mosaic virus (ClYMV) from
157 Alphaflexiviridae, red clover nepovirus A (RCNVA) , tobacco streak virus (TSV), and
158 tomato ringspot virus (ToRSV) from genus *Nepovirus* (family: Secoviridae), red clover
159 vein mosaic virus (RCVMV) from Betaflexiviridae, and one -ssRNA virus: *pinus flexilis*
160 virus 1 (PiFleV1) from Rhabdoviridae (**Table 1A**). The plant viruses in *C. calcarata*
161 virome are much less diversed. WCCV-2 is the only fully-assembled plant viruses,
162 while WCCV-1, BCRV, ToRSV, and WClMV were partially assembled (**Table 1B**).
163
164 *Novel viruses in A. mellifera and C. calcarata virome assemblies*
165 The genus *Iflavirus* (family: Iflaviridae) consists of +ssRNA viruses infecting
166 arthropods, including three bee pathogenic viruses SBV, DWV, and SBPV. Two novel
167 Iflavirus, referred to as unknown iflavirus 1 and unknown inflavirus 2, were identified
168 in *C. calcarata* virome assembly as positive strands. The phylogenetic analysis using
169 translated polyprotein amino acid sequence suggests both viruses are in the same
170 group with SBV. The unknown iflavirus-1 has the highest identity with Darwin bee
171 virus-2, an unclassified picornavirus infecting honey bee [30] (**Figure 1A**). This virus is
172 also detected in *A. mellifera* virome assembly as a negative strand, suggesting that it
173 can infect and replicate in honey bees.
174 In *C. calcarata*, we detected a novel poxvirus (double-strand DNA virus). Its
175 DNA-dependent RNA polymerase RPO132 is partially assembled. Phylogenetic
176 analysis showed this virus is closely related to the Yalta virus, an uncategorized

177 poxvirus from *Drosophila melanogaster* and *Diachasmimorpha* entomopoxvirus
178 (DIEPV), from parasitoid wasp *Diachasmimorpha longicaudata*. The three viruses
179 form a sister cluster to other entomopoxvirus (genus: *Alphaentomopoxvirus*,
180 *Betaentomopoxvirus*, and *Deltaentomopoxvirus*), while all the vertebrate poxvirus
181 (genus: *Avipoxvirus*, *Capripoxvirus*, and *Parapoxvirus*) are clustered into a separate
182 group. The result suggests this novel virus belongs to the subfamily
183 Entomopoxvirinae (EPV) that primarily infect insects (**Figure 1E**).
184 Three additional novel plant viruses were detected in the *A. mellifera* virome,
185 including one from the genus *Nepovirus* (unknown nepovirus 1, family: Secoviridae),
186 one from family Partitiviridae (unknown parititivirus 1, family: Partitiviridae), and one
187 of genus *Amalgavirus* (unknown Amalgavirus 1, family: Amalgaviridae). Nepovirus
188 has two positive-strand genomic RNAs [31]. For unknown nepovirus 1, the ORF of
189 polyprotein RNA2 was fully assembled. However, the correspondent RNA1 encoding
190 viral coat protein was not detected. The phylogeny of polyprotein sequence shows it
191 is most close related to Blueberry leaf mottle virus and Grapevine Bulgarian latent
192 virus, forming a separate cluster (**Figure 1B**). Similarly, only RNA1 encoding
193 RNA-dependent RNA polymerase (RDRP) were fully assembled for another potential
194 plant virus: unknown parititivirus 1. The phylogenetic analysis of RDRP showed it is
195 clustered with several uncategorized partitivirus viruses associated with plant:
196 *Polygonatum kingianum* cryptic virus 1, Panax cryptic virus 3, and Dichroa partitivirus
197 1. This cluster is sister group to plant-infecting genus *Alphapartitivirus* (**Figure 1D**).
198 The genome of unknown amalgavirus 1 was fully assembled with two complete ORF:
199 the ORF1+2 fusion protein responsible for genome replication, and a coat protein.
200 The phylogeny of fusion proteins showed the it is the most closed to Vicia cryptic
201 virus M (VCV-M) and Medicago sativa amalgavirus 1 (MsAV1) and have a clear
202 separation with *Zygosaccharomuces bailii* virus Z (ZbV-Z), a yeast-infecting virus

203 belongs to the genus *Zybavirus* (**Figure 1C**).

204 In addition, other viruses, including phages and a human virus: Merkel cell
205 polymavirus (Family: Polyomaviridae) were identified in both virome assemblies. The
206 viral genome consists of circular dsRNA, which resulted in two identical full-length
207 viral sequences were assembled into one contig in *C. calcarata* virome. In *A. mellifera*,
208 only partial sequence was assembled. How the Merkel cell polymavirus introduced
209 into virome and their effects in both species are unknown.

210

211 *Quantification of viruses in response to different agricultural landscape*

212 The distributions of the above detected viruses are highly site-specific. In *A.*
213 *mellifera*, among all identified insect viruses, five known bee viruses: BQCV, DWV,
214 LSVs, SBV, and AmFV were detected from all of the experimental sites regardless
215 of the type of agricultural use, with AmFV only marginally detected in conventional
216 farm C1. In contrast, the novel unknown ilfavirus 1 was only detected in samples
217 from organic farm O2 (**Figure 2A**). For five known bee viruses, the edgeR did not
218 show any significantly differences at virus level. While, a separate 2834-bp LSV3
219 segment, containing ORFs of two viral proteins, were significantly increased in
220 organic samples comparing to roadside samples (FDR: 0.04). The Hubei partiti-like
221 virus was also identified on all sites except conventional farm C1 (**Figure 2A**). Its
222 full-length sequence significantly increased in organic farms comparing to both
223 conventional and roadside samples (FDR: 0.006 to conventional, 0.003 to roadside).

224 The presence and viral levels of plant viruses are highly depended to the
225 location. For instance, TSV and PiFleV1 were only detected respectively on one
226 roadside site, RCNVA and TRSV were only found respectively on one conventional
227 and one organic site. RCVMV were identified across on three conventional sites, but
228 in relatively low level. WCIMV, WCCV2, and ToRSV were expressed on all

229 experimental sites (**Figure 2B**). The WCCV-1 is the only significant plant virus
230 detected by edgeR. The levels of both genome segments decreased significantly in
231 organic sites, comparing to both conventional (FDR, RNA1: 0.011, RNA2: 0.006) and
232 roadside (FDR, RNA1: 0.006, RNA2: 0.003). The viral levels of BRCV at road sites were
233 either very low or undetected, but the overall level is too low to carry out statistical
234 analysis in edgeR.

235 *C. calcarata* hosts much less viruses. The presence of insect viruses was high
236 site-specific. BQCV and unknown entomopoxvirus 1 were only found respectively in
237 one conventional site, where two novel ilfavirus were detected in roadside sites and
238 conventional farm C1. None of the viruses were detected from samples collected
239 from both organic sites, conventional farm C3, and roadside R2 (**Figure 2C**).
240 Regarding to the plant viruses, WCCV1 and WCCV2 have a similar pattern, being
241 detected only on conventional farms, while WCIMV and ToRSV were detected in
242 some roadside samples as well (**Figure 2D**). However, their levels were not claimed
243 significant. Similarly, none of the plant viruses were detected from samples collected
244 organic sites.

245

## 246 Discussion

247 The metagenomic approach based on viral RNA-Seq was adopted for identifying
248 and quantifying known viruses, as well as discovering novel viruses. The presented
249 study surveyed the abundance of viruses associated with two pollinator species. The
250 virome compositions are site-specific. However, the *A. mellifera* carries a wider range
251 of both insect and plant viruses than *C. calcarata*. The results also indicate that the *A.*
252 *mellifera* may transmit plant diseases more effectively than *C. calcarata*.

253

254 *The Honey bee viruses in A. mellifera and spillover to C. calcarata*

255 The levels of BQCV, DWV and SBV are high in *A. mellifera* virome at all
256 experimental sites, suggested the high prevalence of above viruses. In contrast, none
257 of the acute bee paralysis virus complex (ABPV, KBV and IAPV) were detected. This
258 result is consistent with other studies in the Northeast of the US [20, 32]. LSV is a
259 highly-variable multi-strain bee virus that do not have any significant symptom, but
260 likely cause week colonies. The LSV contigs assembled in this study are mostly closely
261 related to LSV1, LSV2 and LSV3. The LSV3 were detected from all experimental sites
262 and the only LSV with full-length genome sequence assembled. The LSV3 were
263 originally identified Slovenia [33] and has been detected in the Western US [34]. The
264 study confirmed LSV3 has been spread to the New England area.

265 Two known viruses associated with *A. mellifera* were also detected: AmFV and
266 Hubei partiti-like virus 34. The AmFV is a weak pathogen in *A. mellifera* and rarely
267 causes significant symptoms. Its infection has been reported related to BQCV and
268 DWV [28, 35]. We detected AmFV from all sites, while its viral loads were from
269 extremely low at conventional farm C1 to medium-high at conventional farm C3. The
270 Hubei partiti-like virus 34 was first discovered as an arthropod RNA virus [36] and
271 was reported associated with hymenopteran species, including *A. mellifera* and Asian
272 hornet *Vespa velutina* in East Asia and South Europe [29, 37]. The current study
273 reported the low to moderate infection of Hubei partiti-like virus 34 from all except
274 conventional farm C3 site. The potential impact of this virus to honey bee health is
275 still unknown.

276 The spillover of honey bee pathogen to wild bees has been considered as a
277 threaten to native pollinator populations [38, 39]. In genus *Ceratina*, DWV was found
278 spillover to two species *C. dupla* [20] and *C. smaragdula* [40]. The current study
279 showed given the high prevalence of both BQCV and DWV in *A. mellifera*, BQCV can

280 spilled over to *C. calcarata*, while no DWV spill over was observed. BQCV is only
281 detected from conventional farm C2 in relatively low prevalence and viral load,
282 indicating that this spillover to *C. calcarata* is not a common scenario.
283
284 *Potential viruses infecting A. mellifera and C. calcarata*
285 The primary hosts for iflaviruses were arthropods. We discovered two closely
286 related novel ifla-like viruses. Unknown iflavirus 1 were detected in both *A. mellifera*
287 and *C. calcarata*. The unknown iflavirus 2 were only detected in *C. calcarata*. The
288 prevalence of both viruses was low, however, the viral load of unknown ilfavirus 2
289 were slightly higher of *C. calcarata* samples collected from conventional farm C1 and
290 roadside site 1, comparing to unknown ilfavirus 1 of the samples.
291 The EPVs are dsDNA viruses infecting insects and predominantly insect
292 pathogens. The unknown entomopoxivirus 1 detected in *C. calcarata* are more
293 closely related to Yalta viruses and DIEPV than other entomopathogenic viruses.
294 Although the role of Yalta virus is still unknown, the DIEPV has been found replicate
295 in both parasitoid wasp *D. longicaudata* and its host fruit fly *Anastrepha suspensa*.
296 The replication of virus is highly virulent to the host fruit fly and no negative impact
297 to wasp host, suggesting a mutualistic symbiotic relationship of DIEPV to parastoid
298 wasp [41]. Although both *C. calcarata* and *D. longicaudata* are from order
299 Hymenoptera, no such parasitoid relationship was observed in *C. calcarata.* Whether
300 this novel entomopoxivirus primary infects *C. calcarata* or just an overspill event to *C.*
301 *calcarata* yet to be studied.
302
303 *Pollinators carrying plant viruses*
304 Many plant viruses can travel within or attach to pollen grains [42, 43], including
305 BRCV, TSV, PSV, WCCV1, WCCV2, TRSV, ToRSV, and WClMV from presented. However,

306 only TSV, TRSV, and ToRSV have been proven vertically and horizontally transmitted
307 by the pollens. The rest were only detected by RNA-Seq [43]. Pollinator such as *A.*
308 *mellifera* posing potential risks. However none of the detected viruses were on the
309 list of pollinator-transmitting virus, according to Fetters et al [43]. TRSV was reported
310 infect honey bee [44], however whether cross-kingdom infection assist transmitting
311 diseases is unknown. On the other hand, pollinators are a powerful tool for
312 monitoring plant diseases [45]. The plant viruses detected in *A. mellifera* resembles
313 to the results of other metagenomic studies of bees and pollen at family and order
314 levels [45, 46], while the species composition is determined by plant and prevalence
315 of plant viruses at the location.

316 Comparing to *A. mellifera*, plant viruses in *C. calcarata* are much less diverse.
317 Previous studies have shown pollens from clover and rose are major food recourses
318 [47], and white clover pollen significantly increases the fitness of *C. calcarata* [23]. In
319 consist with that, among four plant viruses detected in *C. calcarata* virome, WCCV1,
320 WCCV2, WClMV primarily infect white clover (*Trifolium repens*), and the host range
321 of ToRSV include rose and red clover (*T. pratense*). None of the viruses beyond *C.*
322 *calcarata's* food range were detected. This suggested *C. calcarata* is better for
323 monitoring white clover pathogens in a smaller area.
324

## 325 Material and method:

326 *Sample collection*

327 Both honey bee and small carpenter bee samples were collected from sites
328 representing three different types of land use: conventional farm, organic farm and
329 roadside. Honey bees and the small carpenter bees were collected by sweep nets
330 from May to July, 2018 in Strafford County, New Hampshire. We collected both

331 species within three conventional farms and three organic farms. The roadside
332 samples were collected in the urbanized area within the University of New
333 Hampshire campus. The detailed location and GPS Coordinates are described in
334 previous publication [48, 49]. Bee samples were flash-frozen in liquid nitrogen and
335 temporarily stored at -80 °C.
336
337 *RNA extraction, sequencing and data cleaning*
338 The viral RNA was extracted from *A. mellifera* and *C. calcarata* individually by
339 Quick-RNA viral kit (ZYMO RESEARCH, Irvine, CA). DNase I (ZYMO RESEARCH)
340 in-column digestions were applied to remove genomic and environmental DNA. For
341 mRNA library generation, 200ng of total RNA was treated with NEBNext Poly mRNA
342 Magnetic Isolation Module E7490L (New England Biolabs, USA) following the
343 manufacturer protocol. Subsequently, the isolated mRNA was fragmented for 10
344 minutes. cDNA was synthesized and amplified for 12 PCR cycles using NEBNext Ultra
345 II Directional (stranded) RNA Library Prep Kit E7760L for Illumina (New England
346 Biolabs, USA) with NEBNext Multilex oligos indexes kit E6440S following the
347 manufacturer directions. Distributions of the template length and adapter-dimer
348 contamination were assessed using an Agilent 2100 Bioanalyzer and High Sensitivity
349 DNA kit (Agilent Technologies, Inc). The concentration of cDNA libraries was
350 determined using Invitrogen Qubit dsDNA HS reagents and read on a Qubit
351 Fluorometer (Thermo Fisher), and the cDNA libraries were paired-end 150bp format
352 sequenced on a Novaseq SP system (Illumina, San Diego, CA). In total, 17 honey bee
353 samples and 16 *C. calcarata* samples were sequenced (**Supplementary File S1**). Raw
354 reads were subject to adapter and quality trimming by Trimmomatic [50]. To remove
355 endogenous transcripts from bee host, the reads were respectively aligned to honey
356 bee and small carpenter bee [25] genomes (including mitochondrial sequences)

357 respectively using Bowtie2 [51]. Read pairs concordantly aligned to bee genomes
358 were considered as endogenous transcripts and subject to removal.
359
360 *Virome assembly and viral identification*
361 To reconstruct the virome of *A. mellifera* or *C. calcarata,* we first merged all the
362 pre-cleaned reads from the same organism to increase the sequence coverage. Then
363 two *de novo* assembler: Trinity [52] and metaSpade [53] were used with default
364 setting. The results are comparable but the total runtime for Trinity is lower. We
365 decided to adopt Trinity for following proceeding analysis.
366 With the pre-cleaning process the assembled virome still contains host
367 sequences. To identify virus sequences, we use BLASTN to search assembled contigs
368 to default nr/nt database with E-value cutoff of 1e-7. Contigs with first three top hits
369 to viral sequence from database will be retained as viral contigs. To further virus
370 identification, we exam the blast results of contigs >500 nt. A contig will be assigned
371 as known virus if it aligned stringently (E-value = 0) to existing virus. If the alignment
372 is less stringent (0<E-value<1e-7), it will be assigned as novel virus. The "full-length"
373 or "partial" labels are based on the completeness and intactness of viral protein of
374 each viral contig. For potential novel viral contigs, ORF finder and BLASTP were used
375 to translate and predict viral protein. MEGA 11 [54] was used to construct
376 Maximum-likelihood phylogenetic tree based on the sequence of homologous
377 proteins of known and novel viruses.
378
379 *Quantification of virus and differential analysis between landscapes*
380 To quantify the abundance of each contigs, the clean reads were mapped to
381 virome assembly of respective species using Bowtie 2 [51]. Multiple alignment is
382 allowed to obtain maximum sensitivity. The read counts were obtained by

featureCount [55] and were normalized and converted to count per million reads (CPM) by R edgeR package [56]. We also use edgeR to conduct differential analysis between different landscape type. Prior to that, each contig, including endogenous and viral, was subject to the filter based on its expression level [57]. Only contigs with enough high levels were retained in statistical test. The contigs with FDR less than 0.05 were claim significantly difference between landscapes. The viral contigs, defined by viral identification step, were focused.


## Acknowledgements:

We are grateful for the help received from Stephanie Gardner, Katherine Odanaka, and Wyatt Shell with collecting bee samples. We thank the Genomic Shared Resource Shared Resource at The Ohio State University Comprehensive Cancer Center, Columbus, OH for generating the virus RNA-seq data reported in this publication. Research reported in this publication was supported by USDA Evans Allen Fund NI231445XXXXG004, NI241445XXXXG004 for HL-B and ZZ, USDA CBG GRANT13202344 for HL-B, USDA award 2020-67014-31557 for MCF, and the NSF Louis Stokes Alliances for Minority Participation (LSAMP) program Ohio Alliance scholarship for JM, the Ohio State University Comprehensive Cancer Center and the National Institutes of Health under grant number P30 CA016058.


## Author contributions:

HL-B and SMR designed the experiment. SMR conducted the field studies and provided the samples. MCF and JM prepared samples for RNA-seq. ZZ conducted data analysis and wrote the first draft of the manuscript. All authors reviewed the final manuscript.

## Figures and Tables:

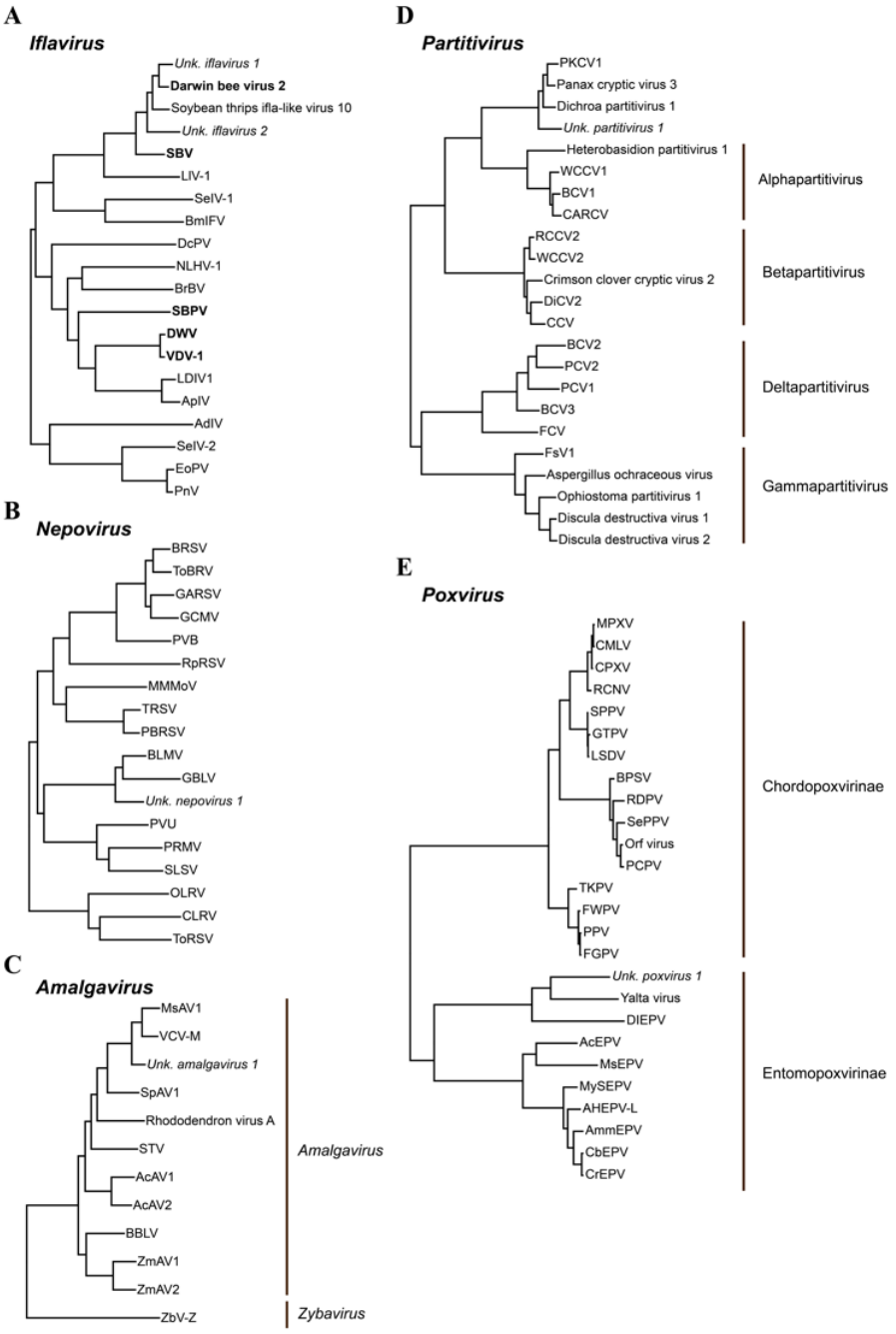


**Figure 1.** Phylogenetic trees of novel viruses discovered from A. mellifera and C. *calcarata* virome assemblies by Maximum-likelihood analysis of sequences of viral proteins: **(A)** polyprotein (iflavirus), **(B)** polyprotein from RNA2 (neprovirus), **(C)** RNA-dependent RNA polymerase (amalgavirus), **(D)** RNA-dependent RNA polymerase from RNA1 (partitivirus), and **(E)** RNA-dependent RNA polymerase submit RPO132 (poxvirus). The names in bold correspond to honeybee-associated viruses.

415 Abbreviations and protein accessions of known viruses are included in **Supplementary File S2**.

416

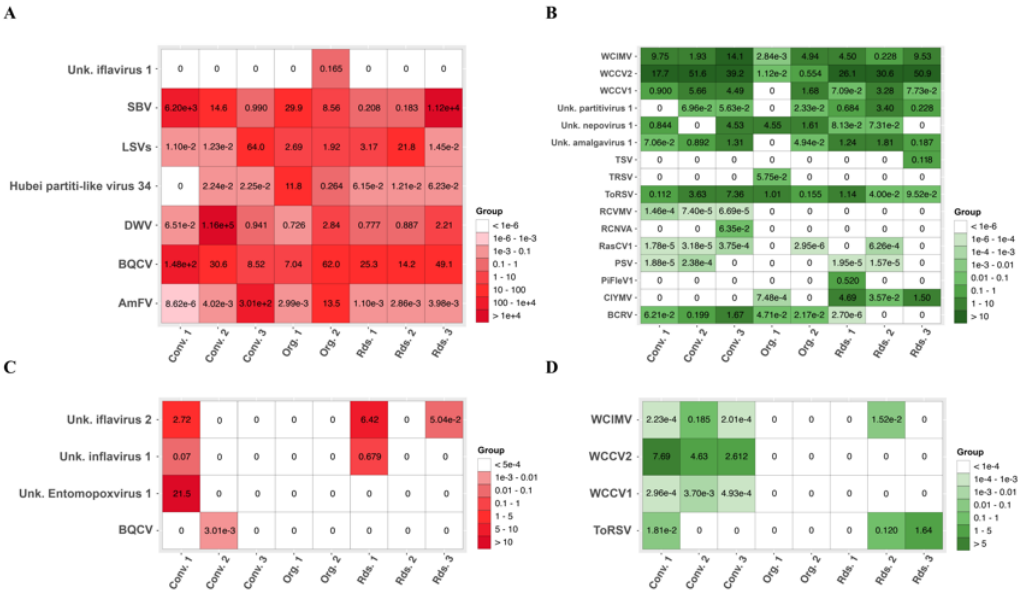


**Figure 2.** Heatmaps demonstrating the relative abundance of **(A)** insect viruses in *A. mellifera*, **(B)** plant viruses in *A. mellifera*, **(C)** insect viruses in C. *calcarata*, and **(D)** plant viruses in *C. calcarata* at each sampling site. Average normalized count-per-million reads (cpm) from RNA-Seq was calculated for each virus of each site and used in the heatmap. White color indicates the virus is undetected on that site. Potential insect and plant viruses were included.

424 **Table 1**. List of identified insect and plant viruses in **(A)** *A. mellifera* and **(B)** *C. calcarata* virome

425 assemblies, assembled from RNA-Seq data

426 **A**

| Virus name | Abbr. | Type | Family | Genus | Host | Assemble status | Strand |
|---|---|---|---|---|---|---|---|
| Black queen cell virus | BQCV | (+)ssRNA | Dicistroviridae | *Triatovirus* | Bees | Full | +/- |
| Deformed wing virus | DWV | (+)ssRNA | Iflaviridae | *Iflavirus* | Bees | Full | +/- |
| Sacbrood virus | SBV | (+)ssRNA | Iflaviridae | *Iflavirus* | Bees | Full | +/- |
| Lake Sinai virus 3 | LSV3 | (+)ssRNA | Sinhaliviridae | *Sinaivirus* | Bees | Full | +/- |
| Lake Sinai virus 1 | LSV1 | (+)ssRNA | Sinhaliviridae | *Sinaivirus* | Bees | Partial | +/- |
| Lake Sinai virus 2 | LSV2 | (+)ssRNA | Sinhaliviridae | *Sinaivirus* | Bees | Partial | + |
| Apis mellifera filamentous virus | AmFV | dsDNA | *unclassified* | *unclassified* | Bees | Partial | +/- |
| Hubei partiti-like virus 34 | - | RNA | *unclassified* | *unclassified* | Bees | Full | + |
| Blackberry chlorotic ringspot virus | BCRV | (+)ssRNA | Bromoviridae | *Ilarvirus* | Plants | RNA1: Full<br>RNA2: Partial<br>RNA3: Partial | +<br>-<br>+ |
| Tobacco streak virus | TSV | (+)ssRNA | Bromoviridae | *Ilarvirus* | Plants | RNA1: Partial<br>RNA2: Partial<br>RNA3: Partial | +/-<br>+/-<br>+ |
| Peanut stunt virus | PSV | (+)ssRNA | Bromoviridae | *Cucumovirus* | Plants | RNA1: Partial<br>RNA2: Partial<br>RNA3: Partial | +<br>+<br>- |
| Raphanus sativus cryptic virus 1 | RasCV1 | dsRNA | Partitiviridae | *unclassified* | Plants | RNA1: Full<br>*RNA2: Missing* | - |
| White clover cryptic virus 1 | WCCV1 | dsRNA | Partitiviridae | *Alphapartitivirus* | Plants | RNA1: Full<br>RNA2: Full | -<br>- |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| White clover cryptic virus 2 | WCCV2 | dsRNA | Partitiviridae | *Alphapartitivirus* | Plants | RNA1: Full<br>RNA2: Full | -<br>- |
| Red clover nepovirus A | RCNVA | (+)ssRNA | Secoviridae | *Nepovirus* | Plants | RNA1: Partial<br>RNA2: Partial | +<br>+/- |
| Tobacco ringspot virus | TRSV | (+)ssRNA | Secoviridae | *Nepovirus* | Plants | RNA1: Partial<br>RNA2: Partial | +/-<br>- |
| Tomato ringspot virus | ToRSV | (+)ssRNA | Secoviridae | *Nepovirus* | Plants | RNA1: Partial<br>RNA2: Partial | +/-<br>- |
| White clover mosaic virus | WClMV | (+)ssRNA | Alphaflexiviridae | *Potexvirus* | Plants | Full | +/- |
| Clover yellow mosaic virus | ClYMV | (+)ssRNA | Alphaflexiviridae | *Potexvirus* | Plants | Partial | +/- |
| Red clover vein mosaic virus | RCVMV | (+)ssRNA | Betaflexiviridae | *Carlavirus* | Plants | Partial | +/- |
| Pinus flexilis virus 1 | PiFleV1 | (-)ssRNA | Rhabdoviridae | *unclassified* | Plants | Partial | +/- |

**B**

| Virus name | Abbr. | Type | Family | Genus | Host | Assembly status | Strand |
|---|---|---|---|---|---|---|---|
| Black queen cell virus | BQCV | (+)ssRNA | Dicistroviridae | *Triatovirus* | Bees | Partial | - |
| Tomato ringspot virus | ToRSV | (+)ssRNA | Secoviridae | *Nepovirus* | Plants | RNA1: Partial<br>RNA2: Partial | +/-<br>+/- |
| White clover cryptic virus 1 | WCCV-1 | dsRNA | Partitiviridae | *Alphapartitivirus* | Plants | RNA1: Full<br>*RNA2: Missing* | - |
| White clover cryptic virus 2 | WCCV-2 | dsRNA | Partitiviridae | *Alphapartitivirus* | Plants | RNA1: Full<br>RNA2: Full | +<br>- |
| White clover mosaic virus | WClMV | (+)ssRNA | Alphaflexiviridae | *Potexvirus* | Plants | Partial | +/- |

431 **Table 2.** List of novel viruses detected in *A. mellifera* and *C. calcarata* virome assemblies assembled

432 from RNA-Seq data

433

| Virus | Potential host | Length (bp) | Viral protein ORF status | Detected in |
|---|---|---|---|---|
| Merkel cell polyomavirus | Vertebrate | 3330 | Partial | *A. mellifera* |
| | | 10809 | Full | *C. calcarata* |
| Unknown iflavirus 1 | Arthropods | 1574 | Partial | *A. mellifera* |
| | | 616 | Partial | *C. calcarata* |
| Unknown iflavirus 2 | Arthropods | 7343 | Full | *C. calcarata* |
| Unknown nepovirus 1 | Plants | 6501 | Full | *A. mellifera* |
| Unknow partitivirus 1 | Plants | 1895 | Full | *A. mellifera* |
| Unknown amalgavirus 1 | Plants | 3425 | Full | *A. mellifera* |
| Unknown entomopoxvirus 1 | Arthropods | 2543 | Partial | *C. calcarata* |

434

435

**Reference:**


1. Ollerton J. Pollinator diversity: distribution, ecological function, and conservation. Annual Review of Ecology, Evolution, and Systematics. 2017;48:353-76. doi: https://doi.org/10.1146/annurev-ecolsys-110316-022919.
2. Katumo DM, Liang H, Ochola AC, Lv M, Wang Q-F, Yang C-F. Pollinator diversity benefits natural and agricultural ecosystems, environmental health, and human welfare. Plant Diversity. 2022;44(5):429-35. doi: https://doi.org/10.1016/j.pld.2022.01.005.
3. Ollerton J, Winfree R, Tarrant S. How many flowering plants are pollinated by animals? Oikos. 2011;120(3):321-6. doi: https://doi.org/10.1111/j.1600-0706.2010.18644.x.
4. Khalifa SAM, Elshafiey EH, Shetaia AA, El-Wahed AAA, Algethami AF, Musharraf SG, et al. Overview of bee pollination and its economic value for crop production. Insects. 2021;12(8):688. PubMed PMID: doi:10.3390/insects12080688.
5. Klein A-M, Vaissière BE, Cane JH, Steffan-Dewenter I, Cunningham SA, Kremen C, et al. Importance of pollinators in changing landscapes for world crops. Proceedings of the Royal Society B: Biological Sciences. 2007;274(1608):303-13. doi: doi:10.1098/rspb.2006.3721.
6. Breeze TD, Bailey AP, Balcombe KG, Potts SG. Pollination services in the UK: How important are honeybees? Agriculture, Ecosystems & Environment. 2011;142(3):137-43. doi: https://doi.org/10.1016/j.agee.2011.03.020.
7. Winfree R, Williams NM, Dushoff J, Kremen C. Native bees provide insurance against ongoing honey bee losses. Ecology Letters. 2007;10(11):1105-13. doi: https://doi.org/10.1111/j.1461-0248.2007.01110.x.
8. Brittain C, Kremen C, Klein A-M. Biodiversity buffers pollination from changes in environmental conditions. Global Change Biology. 2013;19(2):540-7. doi: https://doi.org/10.1111/gcb.12043.
9. Goulson D, Nicholls E, Botías C, Rotheray EL. Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. Science. 2015;347(6229):1255957. doi: 10.1126/science.1255957.
10. Roulston TaH, Goodell K. The role of resources and risks in regulating wild bee populations. Annual Review of Entomology. 2010;56(1):293-312. doi: 10.1146/annurev-ento-120709-144802.
11. Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, Kunin WE. Global pollinator declines: trends, impacts and drivers. Trends in Ecology & Evolution. 2010;25(6):345-53. doi: 10.1016/j.tree.2010.01.007.
12. Drossart M, Gérard M. Beyond the decline of wild bees: optimizing conservation measures and bringing together the actors. Insects. 2020;11(9):649. PubMed PMID:

474 doi:10.3390/insects11090649.
475 13. Brown MJF, Dicks LV, Paxton RJ, Baldock KCR, Barron AB, Chauzat M-P, et al. A
476 horizon scan of future threats and opportunities for pollinators and pollination. PeerJ.
477 2016;4:e2249. doi: 10.7717/peerj.2249.
478 14. Walker PJ, Siddell SG, Lefkowitz EJ, Mushegian AR, Adriaenssens EM,
479 Alfenas-Zerbini P, et al. Recent changes to virus taxonomy ratified by the
480 International Committee on Taxonomy of Viruses (2022). Archives of Virology.
481 2022;167(11):2429-40. doi: 10.1007/s00705-022-05516-5.
482 15. Seitz K, Buczolich K, Dikunová A, Plevka P, Power K, Rümenapf T, et al. A
483 molecular clone of Chronic Bee Paralysis Virus (CBPV) causes mortality in honey bee
484 pupae (*Apis mellifera*). Scientific Reports. 2019;9(1):16274. doi:
485 10.1038/s41598-019-52822-1.
486 16. Chen N-C, Wang C-H, Yoshimura M, Yeh Y-Q, Guan H-H, Chuankhayan P, et al.
487 Structures of honeybee-infecting Lake Sinai virus reveal domain functions and capsid
488 assembly with dynamic motions. Nature Communications. 2023;14(1):545. doi:
489 10.1038/s41467-023-36235-3.
490 17. Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, Moran NA, et al. A
491 metagenomic survey of microbes in honey bee colony collapse disorder. Science.
492 2007;318(5848):283-7. doi: doi:10.1126/science.1146498.
493 18. Tehel A, Brown MJF, Paxton RJ. Impact of managed honey bee viruses on wild
494 bees. Current Opinion in Virology. 2016;19:16-22. doi:
495 https://doi.org/10.1016/j.coviro.2016.06.006.
496 19. Dolezal AG, Hendrix SD, Scavo NA, Carrillo-Tripp J, Harris MA, Wheelock MJ, et al.
497 Honey bee viruses in wild bees: viral prevalence, loads, and experimental Inoculation.
498 PLOS ONE. 2016;11(11):e0166190. doi: 10.1371/journal.pone.0166190.
499 20. Singh R, Levitt AL, Rajotte EG, Holmes EC, Ostiguy N, vanEngelsdorp D, et al. RNA
500 viruses in hymenopteran pollinators: evidence of inter-taxa virus transmission via
501 pollen and potential impact on non-apis hymenopteran species. PLOS ONE.
502 2010;5(12):e14357. doi: 10.1371/journal.pone.0014357.
503 21. Zhang X, He SY, Evans JD, Pettis JS, Yin GF, Chen YP. New evidence that deformed
504 wing virus and black queen cell virus are multi-host pathogens. Journal of
505 Invertebrate Pathology. 2012;109(1):156-9. doi:
506 https://doi.org/10.1016/j.jip.2011.09.010.
507 22. Susi H, Laine AL. Agricultural land use disrupts biodiversity mediation of virus
508 infections in wild plant populations. New Phytologist. 2021;230(6):2447-58.
509 23. Lawson SP, Kennedy KB, Rehan SM. Pollen composition significantly impacts the
510 development and survival of the native small carpenter bee, *Ceratina calcarata*.
511 Ecological Entomology. 2021;46(2):232-9. doi: https://doi.org/10.1111/een.12955.

24. Rehan SM, Richards MH. Nesting biology and subsociality in *Ceratina calcarata* (Hymenoptera: Apidae). The Canadian Entomologist. 2010;142(1):65-74. Epub 2012/04/02. doi: 10.4039/n09-056.
25. Rehan SM, Glastad KM, Lawson SP, Hunt BG. The genome and methylome of a subsocial small carpenter bee, *Ceratina calcarata*. Genome Biology and Evolution. 2016;8(5):1401-10. doi: 10.1093/gbe/evw079 %J Genome Biology and Evolution.
26. Lawson SP, Helmreich SL, Rehan SM. Effects of nutritional deprivation on development and behavior in the subsocial bee *Ceratina calcarata* (Hymenoptera: Xylocopinae). Journal of Experimental Biology. 2017;220(23):4456-62. doi: 10.1242/jeb.160531.
27. Shell WA, Rehan SM. The price of insurance: costs and benefits of worker production in a facultatively social bee. Behavioral Ecology. 2017;29(1):204-11. doi: 10.1093/beheco/arx146.
28. Hartmann U, Forsgren E, Charrière J-D, Neumann P, Gauthier L. Dynamics of *Apis mellifera* filamentous virus (AmFV) infections in honey bees and relationships with other parasites. Viruses [Internet]. 2015; 7(5):[2654-67 pp.].
29. Marzoli F, Forzan M, Bortolotti L, Pacini MI, Rodríguez-Flores MS, Felicioli A, et al. Next generation sequencing study on RNA viruses of *Vespa velutina* and *Apis mellifera* sharing the same foraging area. Transboundary and Emerging Diseases. 2021;68(4):2261-73. doi: https://doi.org/10.1111/tbed.13878.
30. Roberts JMK, Anderson DL, Durr PA. Metagenomic analysis of Varroa-free Australian honey bees (*Apis mellifera*) shows a diverse Picornavirales virome. Journal of General Virology. 2018;99(6):818-26. doi: https://doi.org/10.1099/jgv.0.001073.
31. Rowhani A, Daubert SD, Uyemoto JK, Al Rwahnih M, Fuchs M. American Nepoviruses. In: Meng B, Martelli GP, Golino DA, Fuchs M, editors. Grapevine Viruses: Molecular Biology, Diagnostics and Management. Cham: Springer International Publishing; 2017. p. 109-26.
32. Drummond FA, Lund J, Eitzer B. Honey bee health in Maine wild blueberry production. Insects. 2021;12(6). doi: 10.3390/insects12060523.
33. Šimenc L, Kuhar U, Jamnikar-Ciglenečki U, Toplak I. First complete genome of Lake Sinai virus lineage 3 and genetic diversity of Lake Sinai virus strains from honey bees and bumble bees. Journal of Economic Entomology. 2020;113(3):1055-61. doi: 10.1093/jee/toaa049.
34. Daughenbaugh KF, Martin M, Brutscher LM, Cavigli I, Garcia E, Lavin M, et al. Honey bee infecting Lake Sinai viruses. Viruses. 2015;7(6):3285-309. doi: 10.3390/v7062772.
35. Abou Kubaa R, Molinatto G, Solaiman Khaled B, Daher-Hjaij N, Heinoun K, Saponari M. First detection of black queen cell virus, *Varroa destructor* macula-like

550 virus, *Apis mellifera* filamentous virus and *Nosema ceranae* in Syrian honey bees *Apis*
551 *mellifera syriaca*. Bulletin of Insectology. 2018;71(2):217–24.
552 36. Shi M, Lin X-D, Tian J-H, Chen L-J, Chen X, Li C-X, et al. Redefining the
553 invertebrate RNA virosphere. Nature. 2016;540(7634):539-43. doi:
554 10.1038/nature20167.
555 37. Kwon M, Jung C, Kil E-J. Metagenomic analysis of viromes in honey bee colonies
556 (*Apis mellifera*; Hymenoptera: Apidae) after mass disappearance in Korea. Frontiers
557 in Cellular and Infection Microbiology. 2023;13. doi: 10.3389/fcimb.2023.1124596.
558 38. Nanetti A, Bortolotti L, Cilia G. Pathogens spillover from honey bees to other
559 arthropods. Pathogens [Internet]. 2021; 10(8).
560 39. Fleites-Ayil FA, Medina-Medina LA, Quezada Euán JJG, Stolle E, Theodorou P,
561 Tragust S, et al. Trouble in the tropics: Pathogen spillover is a threat for native
562 stingless bees. Biological Conservation. 2023;284:110150. doi:
563 https://doi.org/10.1016/j.biocon.2023.110150.
564 40. Santamaria J, Villalobos EM, Brettell LE, Nikaido S, Graham JR, Martin S.
565 Evidence of Varroa-mediated deformed wing virus spillover in Hawaii. Journal for
566 Parasitology: Parasites and Wildlife. 2018;151:126-30.
567 41. Coffman KA, Burke GR. Genomic analysis reveals an exogenous viral symbiont
568 with dual functionality in parasitoid wasps and their hosts. PLOS Pathogens.
569 2020;16(11):e1009069. doi: 10.1371/journal.ppat.1009069.
570 42. Card SD, Pearson MN, Clover GRG. Plant pathogens transmitted by pollen.
571 Australasian Plant Pathology. 2007;36(5):455-61. doi: 10.1071/AP07050.
572 43. Fetters AM, Ashman T-L. The pollen virome: A review of pollen-associated
573 viruses and consequences for plants and their interactions with pollinators. American
574 Journal of Botany. 2023;110(6):e16144. doi: https://doi.org/10.1002/ajb2.16144.
575 44. Li JL, Cornman RS, Evans Jay D, Pettis Jeffery S, Zhao Y, Murphy C, et al. Systemic
576 spread and propagation of a plant-pathogenic virus in European honeybees, *Apis*
577 *mellifera*. mBio. 2014;5(1):10.1128/mbio.00898-13. doi: 10.1128/mbio.00898-13.
578 45. Smadi M, Lee E, Phelan J, Wang A, Bilodeau GJ, Pernal SF, et al. Plant virus
579 diversity in bee and pollen samples from apple (*Malus domestica*) and sweet cherry
580 (*Prunus avium*) agroecosystems. Frontiers in Plant Science. 2024;15:1335281.
581 46. Fetters AM, Cantalupo PG, Wei N, Robles MTS, Stanley A, Stephens JD, et al. The
582 pollen virome of wild plants and its association with variation in floral traits and land
583 use. Nature Communications. 2022;13(1):523. doi: 10.1038/s41467-022-28143-9.
584 47. Lawson SP, Ciaccio KN, Rehan SM. Maternal manipulation of pollen provisions
585 affects worker production in a small carpenter bee. Behavioral Ecology and
586 Sociobiology. 2016;70(11):1891-900. doi: 10.1007/s00265-016-2194-z.
587 48. Tucker EM, Rehan SM. Wild Bee Community Assemblages Across Agricultural

588 Landscapes1. Journal of Agricultural and Urban Entomology. 2017;33(1):77-104.
589 49. Tucker EM, Rehan SM. Farming for bees: annual variation in pollinator
590 populations across agricultural landscapes. Agricultural and Forest Entomology.
591 2018;20(4):541-8.
592 50. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
593 sequence data. Bioinformatics. 2014. doi: 10.1093/bioinformatics/btu170.
594 51. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature
595 Methods. 2012;9(4):357-9. doi: 10.1038/nmeth.1923
596 http://www.nature.com/nmeth/journal/v9/n4/abs/nmeth.1923.html#supplementar
597 y-information.
598 52. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al.
599 Full-length transcriptome assembly from RNA-Seq data without a reference genome.
600 Nature Biotechnology. 2011;29(7):644-52.
601 53. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile
602 metagenomic assembler. Genome Research. 2017;27(5):824-34. doi:
603 10.1101/gr.213959.116.
604 54. Tamura K, Stecher G, Kumar S. MEGA11: Molecular evolutionary genetics
605 analysis version 11. Molecular Biology and Evolution. 2021;38(7):3022-7. doi:
606 10.1093/molbev/msab120.
607 55. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for
608 assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30.
609 doi: 10.1093/bioinformatics/btt656.
610 56. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for
611 differential expression analysis of digital gene expression data. Bioinformatics.
612 2010;26(1):139-40. doi: 10.1093/bioinformatics/btp616.
613 57. Chen Y, Lun AT, Smyth GK. From reads to genes to pathways: differential
614 expression analysis of RNA-Seq experiments using Rsubread and the edgeR
615 quasi-likelihood pipeline. F1000Research. 2016;5:1438. Epub 2016/08/11. doi:
616 10.12688/f1000research.8987.2. PubMed PMID: 27508061; PubMed Central PMCID:
617 PMCPMC4934518.
618