Title

Spatial scale influences taxon conservation in the eukaryotic virome of a mosquito

Running title

Virome conservation depends on spatial scale

Authors

Patricia Gil[a], Antoni Exbrayat[a], Etienne Loire[a], Ignace Rakotoarivony[a], Florian Charriat[a], Thierry Baldet[a†], Michel Boisseau[a], Albane Marie[b], Benoît Frances[b], Gregory L'Ambert[b], Mohamed Bessat[3], Yehia Otify[3], Maria Goffredo[d], Giuseppe Mancini[d], Núria Busquets-Martí[e], Lotty Birnberg[e], Sandra Talavera[e], Carlos Aranda[e,f], Emna Ayari[g], Selma Mejri[g], Soufien Sghaier[g], Amal Bennouna[h], Hicham El Rhaffouli[i], Thomas Balenghien[a,h], Ghita Chlyeh[j], Ouafaa Fassi Fihri[h], Julie Reveillaud[a‡], Yannick Simonin[a, k], Marc Eloit[l, m, n] and Serafin Gutierrez[a #]

[a]*ASTRE, CIRAD, INRAE, University of Montpellier, Montpellier, France.*

[b]*EID Mediterranée, Montpellier, France*

[c]*Department of Parasitology, Faculty of Veterinary Medicine, Alexandria University, Alexandria, Egypt.*

[d]*Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise 'G. Caporale'; Teramo, Italy.*

[e]*IRTA, Centre de Recerca en Sanitat Animal (CReSA, IRTA-UAB), Campus de la Universitat Autònoma de Barcelona, Bellaterra, Spain.*

[f]*Servei de Control de Mosquits del Consell Comarcal del Baix Llobregat, Barcelona, Spain.*

[g]*Institut de la Recherche Vétérinaire de Tunisie, Tunis, Tunisia.*

[h]*Department of animal pathology and public health. Hassan II Agronomy & Veterinary Institute, Rabat, Morocc*

[i]*Veterinary Division, FAR Military Health Service, Meknes, Morocco*

[j]*Département de Production, Protection et Biotechnologies Végétales, Unité de Zoologie, Institute of Agronomy and Veterinary Medicine Hassan II, Rabat, Morocco*

[k]*Pathogenesis and Control of Chronic Infections, University of Montpellier, INSERM, EFS, Montpellier, France.*

[l] *Institut Pasteur, Pathogen Discovery Laboratory, Paris, France.*

[m]*Institut Pasteur, The OIE Collaborating Centre for Detection and Identification in Humans of Emerging Animal Pathogens, Paris, France.*

[n]*École nationale vétérinaire d'Alfort, Maisons-Alfort, France.*

Address correspondence to Serafin Gutierrez, serafin.gutierrez@cirad.fr

[†]Present address: CIRAD, UMR ASTRE, dP OHOI, CYROI, 97491 Sainte-Clotilde, Reunion Island, France

[‡]Present address: MIVEGEC, University of Montpellier, INRAE, CNRS, IRD, Montpellier, France

Word count for the abstract/text = 220/5040

Abstract

Our knowledge on the diversity of eukaryotic viruses has recently undergone a massive expansion. This diversity could influence host physiology through yet unknown phenomena of potential interest to the fields of health and food production. However, the assembly processes in eukaryotic viromes of terrestrial animals remain elusive. This situation hinders hypothesis-driven tests of virome influence on host physiology. Here, we explore virome assembly at different spatial scales in the eukaryotic virome of the mosquito *Culex pipiens*. This mosquito is a vector of human pathogens worldwide. Several new aspects of virome assembly were unveiled through a sampling involving five countries in Africa and Europe and large sample sizes. A group of viruses was conserved in *C. pipiens* populations in all countries. This core group represented a relatively large and diverse fraction of the virome. However, core viruses were not shared by all host individuals in a given country, and their infection rates fluctuated between countries and years. Moreover, the distribution of co-infections in individual mosquitoes suggested random co-occurrence of certain core viruses. We also observed differences in the virome depending on geography, with viromes tending to cluster depending on the continent. Thus, our results unveil that taxon conservation in a eukaryotic virome changes with spatial scale. Thus, predictions on virome assembly seem possible at a large geographical scale in *C. pipiens*.

IMPORTANCE The study of the eukaryotic virome of mosquitoes is an emerging research field. Beyond its fundamental interest, this field could lead to the development of control tools against the transmission of mosquito-borne human pathogens. However, we yet know little on the assembly patterns in the eukaryotic viromes of mosquitoes, as well as of terrestrial animals in general. This situation hampers the design of hypothesis-driven studies on the influence of the virome on pathogen transmission. Here, we have analyzed virome assembly in the mosquito vector *Culex pipiens* within and between countries in Africa and Europe. Our results show that integrating contrasted spatial scales allows to identify deterministic patterns in virome assembly. Such patterns can guide future studies of virome influence on mosquito physiology.

60 ## Introduction

61 An impressive expansion in the known diversity of viruses has recently taken place thanks to
62 metagenomics (1). This expansion has been followed by major improvements in the emerging field of
63 community ecology of viruses. For example, viral communities or viromes have been shown to play a
64 major role in the regulation of the carbon cycle in the oceans (2), or in bacteriome dynamics (3, 4).
65 However, current knowledge is largely biased towards viromes of unicellular hosts. Our
66 understanding of the ecology of the eukaryotic virome in terrestrial animals largely lags behind (5, 6).
67 For example, we still know little about the potential deterministic patterns behind virome assembly
68 in those organisms (6-10). This situation severely hampers, among others, studies on the influence of
69 the eukaryotic virome on the physiology of those hosts.

70 The eukaryotic virome of mosquitoes has received considerable attention due to its potential
71 influence on the transmission of human pathogens with a huge impact on health worldwide (11). As
72 for most eukaryotic viromes, most studies on mosquito viromes have focused on describing
73 taxonomic diversity (*e.g.*, (9, 12-15)). Those studies have discovered an impressive number of new
74 viruses. These viruses are mostly RNA viruses, the dominant group in terrestrial eukaryotes (16), and
75 are thought to be mosquito commensals that do not infect vertebrates (17). Moreover, they are
76 often supposed to be vertically transmitted, although other transmission modes have been observed
77 (18-21). Interestingly, a few mosquito-specific viruses have been shown to influence the infection of
78 mosquitoes by human pathogens under laboratory conditions (21-26).

79 The improvement in the known diversity of mosquito viruses contrasts with the paucity of studies on
80 virome assembly patterns. For example, little is known on the potential influence of spatial scales on
81 virome assembly (27, 28), a main question in community ecology (29, 30). The study of virome
82 assembly across scales allows to infer whether specific viruses are associated to a host no matter the
83 spatial scale, that is, whether a core virome exists. The identification of core taxa is often a main
84 initial step in the study of microbial communities since core members are likely to have a key

85 influence on host physiology (31, 32). To date, a fistful of studies has explored the potential existence
86 of a core virome in different mosquitoes (9, 10, 33-39) . Overall, certain viruses have been found
87 repeatedly associated to a given mosquito species, suggesting that a cove virome may take place.
88 However, current studies analyze a single spatial scale (9, 10, 33-35, 38, 39), or different but limited
89 geographical ranges of the mosquito host (28). Moreover, sample sizes in those studies are usually
90 low despite the usual large population sizes of mosquitoes (10, 34-36). This situation hampers to
91 robustly elucidate whether a core virome is conserved across spatial scales.

92 Here, we analyze the influence of spatial scale on the eukaryotic virome of the common house
93 mosquito *Culex pipiens* Linnaeus, 1758. This mosquito is the main vector of important human
94 pathogens, like West Nile virus. Moreover, *C. pipiens* has nearly a worldwide distribution, probably
95 through introductions from Europe and Africa (40). We have analyzed the virome of *C. pipiens* both
96 between and within countries in Europe and Africa (Fig. 1). Our mosquito collection allowed to
97 analyze between hundreds and thousands of individuals per country, and from the main habitats of
98 *C. pipiens* in the study area. Using a blend of meta-transcriptomics (9, 12, 41) and PCR analyses, we
99 revealed that taxon conservation decreased with spatial scale. That is, a group of viruses was shared
100 between countries. However, a limited taxon conservation in the virome took place between
101 mosquitoes within a country, along with random co-infection patterns. Moreover, differences in
102 virome assembly were observed between European and African countries. Geographic structuration
103 did not seem to originate by strict isolation of virus populations but rather from differences in
104 relative abundancy within viromes.

105

106 ## Results

107 **High taxon conservation around the Mediterranean Basin**

108 We used meta-transcriptomics (12) to characterize the eukaryotic virome of *C. pipiens* in five
109 countries in Africa and Europe (Fig. 1). Two samples were generated per country, each comprising

110 711 females on average (Table S1). Samples were derived from different habitats to maximize the
111 exploration of virus diversity (Table S1). Meta-transcriptomics analysis of the ten samples provided
112 715 million reads (Table S1). Among them, 41 million reads (5.7%) mapped to virus-like sequences
113 associated with 324 viral operational taxonomic units (vOTUs) (36). The number of virus-like reads
114 significantly differed between samples (Table S1; Wilcoxon signed rank exact test, $p = 0.001$). We did
115 not detect a significant effect of raw reads or mosquito number on the yield of virus-like reads (Table
116 S1; Spearman's rank correlation, $p > 0.076$). Rarefaction curves showed that sampling effort, in terms
117 of viral reads, was overall adequate (Fig. S1).

118 After quality filtering, the dataset comprised 201 vOTUs, distributed over 44 clusters (an arbitrary
119 taxon similar to a Family, see Material and Methods), and 99.98% of the virus-like reads. The dataset
120 was dominated by vOTUs related to viruses detected in arthropods (167 vOTUs, 99.90% of the virus-
121 like reads). We were only interested in vOTUs potentially infecting *C. pipiens*. We thus discarded the
122 vOTUs related to non-arthropod hosts. None of the non-arthropod vOTUs was found in more than six
123 samples.

124 The final dataset consisted of 167 vOTUs and 36 clusters, mainly RNA viruses. A large proportion of
125 the vOTUs had identities at the amino-acid level below 80% with their best hits in GenBank (Fig. S2).
126 This situation suggests the presence of numerous new viruses, as usually observed in studies on the
127 mosquito virome (10, 27, 28). Selected sequences were submitted to GenBank (BioProject
128 PRJNA806751). Sequence submission was done with the Gsub tool, a new tool allowing easy
129 annotation and submission of virus contigs (Supplementary Annex 1). Most names of new virus
130 species were provided by children participating to the 2021 Science Days in Montpellier (France). A
131 detailed description of this diversity is beyond the scope of this study and was not further explored
132 here.

133 Cluster conservation was relatively large between libraries and countries (Fig. 2A). Ten clusters were
134 present in all the libraries (27% of the clusters; hereafter referred to as "highly-conserved clusters";

135 Fig. 2A). Moreover, almost half of the clusters were present in all countries (18 clusters, Fig. 2A).
136 Highly-conserved clusters tended to rank higher in terms of read abundance but some of them had
137 reproducibly low frequencies in most libraries (*i.e.* < 0.1 %; *e.g.,* see clusters Mononega_unclass and
138 Picorna_unclass in Fig. 2B). More generally, read distribution among clusters was rather
139 heterogeneous, with a few clusters providing most of the reads both globally and in each library (Fig.
140 2B).

141 We also observed a relatively high taxon conservation at the vOTU level (Fig. 3, Table S2). Seven out
142 of the ten highly-conserved clusters included a vOTU that was present in all libraries (Fig. 3, Table S2).
143 The three remaining clusters (Partiti, Mononega_unclass and Nege) had a vOTU present in at least
144 eight libraries (Fig. 3, Table S2).

145 **Highly-conserved taxa are found associated to *Culex pipiens* in other regions**

146 We analyzed if the highly-conserved taxa had been found associated to *C. pipiens* in other countries.
147 Our dataset was compared to four studies from China, Morocco, Serbia and Sweden (10, 38, 48, 49).
148 Despite important differences in methodology between studies (Table S3), we observed a relatively
149 good conservation of highly-conserved clusters (Fig. 4). All clusters were detected in Morocco and
150 Sweden. The study from China detected seven out of ten clusters despite of the fact that it only
151 provides vOTUs that shared more than a 90 % identity at the amino-acid level with their best hit.
152 Finally, the study from Serbia detected six clusters despite providing only those vOTUs for which
153 complete RdRp genes or genomes had been obtained. Taxon conservation was also observed at the
154 vOTU level. All highly-conserved clusters, except for the Partiti cluster, had vOTUs shared in other
155 countries (Fig. 4).

156 **Taxon conservation between countries does not imply conservation between individual**
157 **mosquitoes in a country**

158 We explored whether taxon conservation also took place within each country. We estimated the
159 infection rates of five vOTUs (Fig. S3), including two highly-conserved vOTUs, in each country. None

160 of the vOTUs reached an infection rate above 10% in any of the countries (Fig. 5A). We observed
161 significant differences in infection rates between vOTUs (Kruskal-Wallis, $p = 0.014$). More precisely,
162 the infection rates of the highly-conserved Culex.Iflavi_like.virus.4__96 were significantly higher than
163 those of Culex.Bunyavirus.2__98 and Culex.mosquito.virus.4__97 (Conover-Iman test, $p = 0.010$ and
164 0.002, respectively). Moreover, significant differences in infection rates were observed between
165 countries for a given virus (Fig. 5A). We did not detect a significant effect of the country on the
166 infection rates of all viruses (Kruskal-Wallis, $p = 0.37$).

167 We then explored whether infection rates could change among years. To do so, 748 females
168 collected in France in 2018 were screened for the five vOTUs. Infection rates in 2018 were often
169 significantly higher than in 2015 (Fig. 5B). For example, Alphamesonivirus.1__97 and
170 Culex.Iflavi_like.virus.4__96 showed infection rates close or above 50% in 2018. To further validate
171 our results, we estimated the infection rate of Alphamesonivirus.1__97 in over 1 300 adult females
172 of *C. pipiens* collected in another country and year (Morocco, 2016) (48). The infection rate lied
173 below 10% in the Moroccan samples, like in the other five countries in 2015 (Fig. 5B). Overall, our
174 results showed that the infection rates of highly-conserved vOTUs could be relatively low, and
175 differed between countries and years.

176 **Frequent random co-infection in individual mosquitoes**

177 We explored the potential non-random co-occurrence of different viruses within individual
178 mosquitoes. We thus screened 184 individual females for the five vOTUs. Those females had been
179 collected in France in 2018, the year with the highest infection rates for most vOTUs (Fig. 5). This
180 analysis showed that 54% of the females were infected by at least two different vOTUs (Fig. 6A).
181 Different combinations of vOTUs were observed among individuals (Fig. 6B). A probabilistic approach
182 was used to infer co-occurrence patterns in vOTU pairs (58). Of the ten possible pairs, one pair was
183 not included in the analysis because its expected co-occurrence was below 1% (the pair
184 Culex.negev.like_virus.1__99/Culex.mosquito.virus.1__97). Only one pair out of the nine pairs tested

185 showed an association deviating from random (positive association between
186 Culex.Iflavi_like.virus.4__96 and Culex.Bunyavirus.2__96, $p = 0.004$; other pairs: $p > 0.061$).

187 **Geography influences virome assembly**

188 We explored potential factors behind the virome heterogeneity observed between samples (Fig. 2A).
189 The habitat anthropization, the continent, and the number of mosquitoes or sampling sites per
190 library did not show a significant influence on taxon richness or the Simpson and Shannon indexes
191 (Kruskall Wallis test, $p > 0.33$; Fig. S4). Jaccard dissimilarities tended not to reach 0.5 but Bray-Curtis
192 dissimilarities were often higher than 0.5 (Fig. S5). These results suggest that the assemblies often
193 shared taxa although with different relative frequencies.

194 Next, we compared virome structure between libraries using non-parametric multidimensional
195 scaling (NMDS) and Bray-Curtis dissimilarities, as well as hierarchical clustering (Fig. 7). Viromes
196 tended to aggregate according to continent and country (Fig. 7). Clustering based on continent but
197 not on habitat anthropization was further confirmed with a PERMANOVA (continent: R2 = 0.445, p =
198 0.01; anthropization level: R2 = 0.062, $p = 0.35$).

199 **Virus phylogeny does not suggest strict isolation of viromes from different continents**

200 The observed differences in virome assembly could be explained, among others, by a limited
201 connectivity between the viromes of *C. pipiens* in Africa and Europe. We explored this question using
202 a phylogenetic approach. The existence of the Mediterranean Sea for over five million years should
203 have allowed for isolation, and thus divergent evolution, of virus populations located in Africa and
204 Europe. Sequences of three taxonomically-distant viruses (Culex Bunyavirus 2, Alphamesonivirus 1
205 and Culex Iflavi-like virus 4) from each country, together with sequences from other countries in the
206 GenBank database, were used to generate phylogenetic trees (Fig. 8). Sequence clustering based on
207 geography was only clear for Culex Bunyavirus 2.

208

209 ## Discussion

210 Our metagenomics analysis showed a relatively high taxon conservation in the eukaryotic virome of
211 *C. pipiens* over a large geographical area. Conserved taxa were phylogenetically distant and
212 represented a large fraction of the virome in terms of read abundancy. Moreover, taxon
213 conservation beyond the study region was supported by a literature analysis. This degree of taxon
214 conservation is striking given the experimental design. Here, sampling involved mosquito populations
215 separated by an important physical barrier, the Mediterranean Sea. Furthermore, the populations
216 from each country were analyzed separately according to habitat anthropization, a factor that
217 impacts virus diversity (59). This sampling design should favor differences in diversity between
218 samples.

219 To our knowledge, only one study has performed an in-depth comparison of the virome of a
220 mosquito species over a large geographical scale (28). Shi and colleagues compared the viromes from
221 field populations of *Aedes aegypti* from North America, Africa, Asia and Oceania. They also observed
222 numerous and diverse taxa conserved over several continents. Interestingly, the core virome in *A.*
223 *aegypti* does not fully overlap that observed in *C. pipiens* here (Fig. S6). This observation represents
224 the first evidence of specificities in the core virome of different mosquito species when compared
225 over large geographical ranges. Hence, our results suggest not only that a core virome may be a
226 common feature of the Culicidae, but also that core viromes could differ among mosquito species.

227 Two putative processes, non-mutually exclusive, could explain the observed taxon conservation in
228 the virome of *C. pipiens*. First, the conserved virus taxa may have spread along with *C. pipiens* during
229 geographical range expansions of the latter (40). However, the observed low similarity in the virome
230 among *C. pipiens* individuals should hamper virome conservation during migration. Such
231 conservation would have required the participation in migrations of a large number of *C. pipiens*
232 individuals (each with different viruses), or specific individuals carrying all core viruses. The second
233 process allowing taxon conservation involves the spread of the highly-conserved viruses by host

234 species other than *C. pipiens*. For example, several mosquito species have colonized parts of the *C.*
235 *pipiens* distribution (60). This second process requires that conserved viruses can be horizontally
236 transmitted between mosquito species, a transmission mode that is rarely invoked or demonstrated
237 in mosquito-specific viruses. In line with the second process, Alphamesonivirus 1, the best studied of
238 the highly-conserved viruses, has a host range including mosquito species beyond the *Culex* genus
239 and a large geographical range, although its transmission routes remain elusive (61). Currently, we
240 cannot disentangle the relative weight of the proposed processes in taxon conservation in the *C.*
241 *pipiens* virome. This situation is due to the paucity of data on the host range, transmission modes
242 and evolutionary history for most of the highly-conserved viruses (19).

243 Contrary to the observations at the between-country scale, virome conservation was limited at the
244 within-country scale. Our results thus suggest a new phenomenon: the level of taxon conservation in
245 a mosquito virome can depend on the spatial scale, with conservation decreasing along with scale. At
246 this point, we cannot provide a simple mechanistic explanation for this phenomenon. Whether this
247 phenomenon takes place in the eukaryotic virome of other mosquitoes or terrestrial animals is yet
248 unclear due to the limited number of studies robustly exploring different scales (6, 8, 62). A previous
249 study on the mosquito virome, geographically restricted to a Caribbean island, suggests that other
250 situations are possible (27). Shi and colleagues showed infection rates above 90% for three viruses
251 screened in 72 *A. aegypti* adults, and similar results were found in 24 *Culex quinquefasciatus* adults.

252 The co-infection rates in individual mosquitoes were analyzed to explore potential non-random co-
253 infection of certain viruses. Contrary to previous studies (27, 37, 63), a large number of individuals
254 was tested, thus improving the analysis robustness. Mosquitoes were often infected by more than
255 one virus, as previously observed in *C. pipiens* and other mosquito species (27, 37, 63). However, co-
256 infection by most virus couples did not deviate from random co-occurrence. Thus, the highly-
257 conserved viruses included in our analysis do not seem to be always simultaneously transmitted. This
258 observation is not expected in a core virome dominated by vertically-transmitted viruses, a
259 widespread view of the mosquito virome (19, 34).

260 Our prevalence data from individual mosquitoes also provided insights into an open question in
261 virome ecology, that of the potential similarity in assembly patterns between the virome and the
262 bacteriome (35, 64). Here, the low prevalence of certain highly-conserved viruses contrasts the high
263 prevalence of core taxa in the bacteriome of mosquitoes, including *C. pipiens* (65, 66).

264 Our analyses detected an influence of geography on virome assembly in *C. pipiens*, with viromes
265 clustering depending on their African or European origin. Jaccard and Bray-Curtis indexes suggested
266 that clustering due to geography was mainly due to differences in relative abundancy between
267 vOTUs, although specificities in taxonomic diversity were also observed. Moreover, the virus
268 phylogenies obtained did not suggest a strict isolation of virus populations in the study area (Fig. 8).
269 Our results thus suggest that viruses are shared and exchanged in the study area but, that, the
270 resulting viromes tend to differ in read abundancy due to unknown factors.

271 The observed patterns in virome assembly require further support due to limits in our experimental
272 design. First, our list of highly-conserved taxa may not be exhaustive. Our meta-transcriptomics
273 approach is considered to be among the most sensitive for virus detection (67), and the number of
274 mosquitoes was higher than in other studies (10, 27). Nevertheless, we cannot exclude that highly-
275 conserved viruses have not been detected due to, among others, low viral loads or temporal
276 dynamics of certain viruses. Secondly, certain highly-conserved taxa may not be fully conserved if the
277 spatial or temporal extent of the sampling is increased. For example, the number of libraries number
278 used in our metagenomics analysis is in the same range as those found in most studies on the
279 mosquito virome (10, 34-36, 38) but yet limited. Furthermore, our study only analyzed adult females.
280 Whether assembly patterns in the mosquito virome are conserved along the life cycle or between
281 sexes are questions that remain poorly studied (28, 34). Thirdly, the analyses of infection rates
282 involved a limited set of viruses. We chose to include the two highly-conserved viruses with read
283 counts and contig lengths warranting a robust detection. However, other highly-conserved viruses
284 may always have high infection rates in *C. pipiens* populations. Future work on a larger panel of
285 highly-conserved viruses could thus unveil the coexistence of contrasted infection rates. Another

286 point that requires further study is the potential influence of habitat anthropization on the mosquito
287 virome. Here, the limited number of libraries did not allow for a robust analysis.

288 Uncovering deterministic patterns in the eukaryotic viromes of terrestrial animals is considered to be
289 challenging (6, 62). This study shows that integrating different spatial scales allows to identify
290 deterministic patterns in the *C. pipiens* virome. Our data provide predictive power and can represent
291 a base for hypothesis-driven studies on, among others, the potential role of the *C. pipiens* virome in
292 pathogen transmission. For example, the highly-conserved viruses identified here represent
293 candidates for studies on their influence on pathogen transmission. Moreover, the virome
294 differences observed among individuals or regions could participate to the complex epidemiological
295 patterns associated to mosquito-borne arboviruses.

296

297 ## Material and Methods

298 **Mosquito sampling**

299 Mosquitoes were captured in five countries (Egypt, France, Italy, Spain and Tunisia; Fig. 1). The genetics
300 of the *C. pipiens* complex have been shown to be relatively homogeneous in the region (42). In each
301 country, mosquitoes were sampled in several sites (Fig. S7 and Table S4). Sites were representative of
302 habitats of *C. pipiens* in the study region, including marshlands, riverine woodlands, crop fields or
303 urbanized areas. Sites were grouped into two categories based on habitat anthropization (Table S4).
304 The two categories were broadly defined to take into account differences in the diversity and
305 availability of habitats between countries. A “high” anthropization level was broadly defined as a peri-
306 urban site, located less than 500 m from the perimeter of a densely populated area. Similarly, a “low”
307 level was largely defined as any agricultural or natural site with a low population density.

308 Several mosquito collections were carried out in each site between April and October in 2015. After
309 overnight trapping, mainly using CO2-baited CDC-type traps, mosquitoes were identified to the species

310 level on a chill table using morphological determination keys. Non-engorged females of *C. pipiens sensu*
311 *lato* were distributed in pools of around 30 individuals (Table S5), and stored at -80 °C until further
312 processing. An additional sampling was carried out in France in 2018. The latter sampling followed the
313 same protocol as in 2015 but for pool size (pools of six females or individual females, Table S5).

314 **Library preparation and sequencing**

315 Samples were analyzed using meta-transcriptomics (12). Ten libraries (one library per country and
316 anthropization level) were prepared for Illumina sequencing as previously described (43). Briefly,
317 mosquito pools were homogenized in ice-cold 1X PBS (Tissue Lyser, Qiagen). An aliquot of 150 µl was
318 used for total RNA isolation with the Nucleospin RNA virus kit (Macherey Nagel). Ten RNA samples
319 were generated by pooling aliquots of RNA extractions from all the pools from the same country and
320 anthropization level. The ten samples were retro-transcribed to cDNA (SuperScript IV reverse
321 transcriptase, Invitrogen) with random hexamers. Then, cDNA was amplified using multiple
322 displacement amplification with phi29 polymerase and random hexamers. Libraries were prepared
323 with the TruSeq DNA Nano Library Prep kit (Illumina). Libraries were sequenced at an expected depth
324 of 80 million reads on a HiSeq2000 (Illumina) in a 150-bp single-read format. Library preparations and
325 sequencing were outsourced to DNAVision (Charleroi, Belgium).

326 **Generation of viral operational taxonomic units**

327 We used an updated version of the SnakeVir bioinformatics pipeline to identify potential viral
328 sequences from reads (44). Briefly, library adapters and low-quality sequences in reads were removed
329 with Cutadapt 1.6 (Martin, 2011). Reads from rRNA sequences from diptera and bacteria were then
330 removed through mapping with BWA 0.7.15 (45) against SILVA rRNA bases. The remaining reads were
331 *de-novo* assembled with Megahit version 1.1.2 (46). The resulting contigs were used in a second *de-*
332 *novo* assembly with CAP3 (Huang, 1999). A contig was considered a virus-like sequence if a viral
333 sequence was found as the first best hit by Diamond (Buchfink et al., 2015) after comparison with the
334 NCBI nr database (release 234, 10/2019; e-value cutoff = 10^{-3}). Then, virus-like contigs were screened

335 for sequences potentially derived from endogenous viral elements (EVEs) using BLASTn against NCBI
336 nt (release 249, 6/2022; e-value cutoff = 10^{-3}). Contigs with a non-viral best hit with a coverage above
337 25% were considered as potential EVEs and removed.

338 The number of reads per virus-like contig and library was quantified through mapping with BWA 0.7.15
339 (45). Duplicate reads had been previously removed with the markdup tool in Samtools (47). Contigs
340 sharing the same best hit were grouped into a viral operational taxonomic unit (vOTU) as previously
341 described (27, 44). Names of vOTUs were generated through combining the best-hit name and the
342 mean percent identity at the amino-acid level of the contigs with the best hit. The percent identity was
343 placed at the end of the best-hit name after a double underscore. The vOTUs with percent identities
344 above 90 % can be considered to include the best-hit species. Spaces and hyphens in vOTU names were
345 replaced by dots and underscores respectively to facilitate analysis in R software.

346 The taxonomy of vOTUs was generated with SnakeVir (44). To avoid incomplete virus taxonomies, the
347 vOTUs were grouped into an arbitrary taxonomic level, here called "cluster", similar to the Family rank
348 (44). The cluster was defined as the family of the best hit whenever available in GenBank databases.
349 For those best hits without a full taxonomic classification, the most likely family was defined from the
350 associated literature and BLAST searches. Clusters were named using the family name without the
351 "viridae" suffix. If analyses failed to indicate a likely family for a given vOTU, the cluster was generated
352 as the name of the most probable Order of the vOTU and replacing the "virales" suffix with "_unclass".

353 **Validation of viral operational taxonomic units**

354 Two arbitrary detection thresholds were applied to the list of vOTUs as previously described (44). First,
355 a vOTUs was considered non-detected in a given library when the number of reads was below ten.
356 Consequently, the read number was set to zero in the library. After this step, vOTUs with less than 100
357 reads in total were removed. We next validated that taxon conservation was not due to contamination
358 during sample processing and sequencing. To do so, the vOTUs in the ten libraries of this study were
359 compared to those found in six libraries sequenced in parallel and processed in the same laboratory

360 for another project. These libraries derived from adult females of the mosquito *Aedes vexans* from
361 Senegal and have been previously described (44). This comparison found 25 vOTUs in common. Twelve
362 vOTUs were not considered as potential contaminants because they had a mean read abundance at
363 least two logs higher in the *C. pipiens* libraries. None of the remaining vOTUs was part of the most
364 prevalent vOTUs in the *C. pipiens* libraries (Table S2), or was present in all libraries.

365 **Analysis of published viromes of *Culex pipiens***

366 A search for articles with robust characterizations of the virome of *C. pipiens* was done in PubMed
367 (October, 2021). The criteria defined to select datasets included: (i) the minimum number of mosquito
368 individuals was set to 100, (ii) the minimal sequencing depth was set to 1 million reads per library, and
369 (iii) the virome diversity should be dominated by viruses from arthropods. Four studies were selected
370 for the analysis (10, 38, 48, 49). All studies analyzed populations in a single country (China, Morocco,
371 Serbia or Sweden; Table S3). There were important methodological differences among studies (Table
372 S3). For example, the studies widely differed in mosquito number, library preparation or sequencing
373 depth. Moreover, the studies had mosquito numbers, sequencing depths or a yield in virus-like reads
374 largely below the ones in this study. Hence, we limited the comparison to the presence of the highly-
375 conserved taxons found in our dataset.

376 **PCR detection of selected virus taxa**

377 We developed RT-qPCRs using SYBR green chemistry for the detection of several vOTUs. First, we
378 selected vOTUs in four highly-conserved clusters - the Mesoni, Ifla, Nege and Reo clusters – for the
379 analysis. We also selected vOTUs in two additional clusters present in nine libraries, the Chu and
380 Bunya_unclass clusters. Primer design was done with LightCycler Probe design Software 2.0 (Roche)
381 and based on the contigs of each vOTU. Up to three primer pairs were tested per vOTU. The PCR
382 output was not satisfactory in terms of specificity and sensibility for vOTUs in the Reo (not shown),
383 and vOTUs of this cluster were not included in the analysis. Moreover, the PCR design for the most
384 prevalent vOTU in the Nege was not satisfactory (Dezidougou.virus__98; not shown) and was

385 discarded from the analysis. A satisfactory PCR design was obtained for Culex.negev_like.virus.1__99,
386 the second most prevalent vOTU in the Nege. Culex.negev_like.virus.1__99 was thus included in the
387 analysis. The final set of five vOTUs and their prevalence among libraries are presented in Figure S2.
388 The primer pairs can be found in Table S6.

389 RT-qPCR conditions were as follows. First, total RNA was reverse transcribed using the RevertAid First
390 Strain cDNA kit (ThermoFisher) and random primers following manufacturer's instructions. Then, qPCR
391 amplifications were carried out with the Brilliant III-Sybrgreen QPCR Master Mix (Agilent) in a total
392 volume of 20 µl consisting of 2 µl of cDNA template, 0.3 µM of the forward and reverse primers, 10 µl
393 of Master Mix 2X and RNase-free H_2O. Thermal cycling was 95°C for 10 min, 40 cycles of 95°C for 20 s,
394 60°C for 20 s (55° for Culex.Bunyavirus.2__96), and a melting curve analysis involving 95°C for 10 s,
395 65°C for 60 s and 97°C for 1 s. Amplicons were Sanger sequenced (Genewiz, Germany) whenever an
396 unclear melting curve was observed.

397 **Phylogenetic analysis**

398 First, vOTUs were selected for phylogenetic analysis only if detected in at least two European
399 countries and the two African countries. Moreover, for a given vOTU, contigs from the different
400 countries had to encompass the same genomic region with at least 10X read coverage over 500 bp.
401 This analysis selected three vOTUs: Alphamesonivirus.1__97, Culex.Iflavi_like.virus.4__96 and
402 Culex.Bunyavirus.2__96. The contigs of those vOTUs had mean identities at the nucleotide level
403 above 90% with their best hits. The contigs in the vOTUs were thus considered as belonging to the
404 same species as their best hits.

405 The contigs selected for Alphamesonivirus.1__97 and Culex.Iflavi_like.virus.4__96 represented
406 nearly-full genomes. The contigs of Culex.Bunyavirus.2__96 represented a nearly-complete L
407 segment. In addition to those contigs, homologous sequences from different countries found in
408 GenBank were also included. The search for homologous sequences unveiled that sequences from
409 Culex picorna-like virus 1 and Aedes Ifla-like virus belong to the same species as Culex Iflavi-like virus

410 4. Similarly, sequences from Culex Bunya-like virus and Bunyanwera environmental sample belong to
411 the same species as Culex Bunyavirus 2.

412 Alignments were generated with Mega X using the ClustalW algorithm for each group of sequences
413 (50). All ambiguously aligned regions were removed using TrimAL (51). The alignments had 16000,
414 8800 and 7000 nucleotide positions for the sequences from Alphamesonivirus 1, Culex Iflavi-like virus
415 4 and Culex Bunyavirus 2 respectively. For each alignment, IQ-Tree (v.1.6.1) (52) and its module
416 *modelfinder* was used to determine the best-fit model of nucleotide substitution. The TN+F+R3
417 model (Tamura and Nei substitution model with empirical base frequency and Gamma distributions
418 for three categories) was selected for each alignment. Maximum likelihood phylogenetic trees were
419 then produced with 1000 bootstrapping replications. Trees were visualized and annotated with the R
420 package *ggtree* (53).

421 **Statistical analyses**

422 Statistical analyses are described along the text and were conducted using R version 3.5.1. The
423 associated figures were generated with the *ggplot2* package. Statistical analyses on virome data
424 were done, unless specifically stated, at the cluster level to avoid potential biases in vOTU
425 identification. The Jaccard and Bray-Curtis dissimilarities were estimated using the *vegan* package
426 (54) from data normalized using Total Sum Normalization (55). Rarefaction curves were generated
427 with the *vegan* package (54). Maximum-likelihood estimates of infection rates were obtained with
428 the *binGroup* package (56). A probabilistic analysis of virus co-occurrence was done with the *coocur*
429 package (57).

430 **Data availability**

431 Raw reads and selected contigs were deposited in GenBank associated to BioProject PRJNA806751.

432

433

Acknowledgements

We thank the personnel of the EID Méditerranée (agencies of Arles, Canet-en-Roussillon, Frejorgues and Montcalm) and Lou Carlier for their involvement in mosquito collection. We also thank Emmanuel Albina for his contribution in the initial phases of the project. This study was partially funded by EU grant FP7-613996 Vmerge and a grant from the Montpellier University of Excellence program (MUSE, ArboSud project). Moreover, SG was recipient of funding from the Direction Générale de l'Alimentation from the French Ministry of Agriculture and Food (DGAl grant agreement: SPA17 n◦0079-E). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

References

1. Kuhn JH, Wolf YI, Krupovic M, Zhang YZ, Maes P, Dolja VV, Koonin EV. 2019. Classify viruses - the gain is worth the pain. Nature 566:318-320.
2. Brum JR, Hurwitz BL, Schofield O, Ducklow HW, Sullivan MB. 2016. Seasonal time bombs: dominant temperate viruses affect Southern Ocean microbial dynamics. The ISME Journal 10:437-449.
3. Manrique P, Dills M, Young MJ. 2017. The Human Gut Phage Community and Its Implications for Health and Disease. Viruses 9.
4. Huang D, Yu P, Ye M, Schwarz C, Jiang X, Alvarez PJJ. 2021. Enhanced mutualistic symbiosis between soil phages and bacteria with elevated chromium-induced environmental stress. Microbiome 9:150.
5. Stulberg E, Fravel D, Proctor LM, Murray DM, LoTempio J, Chrisey L, Garland J, Goodwin K, Graber J, Harris MC, Jackson S, Mishkind M, Porterfield DM, Records A. 2016. An assessment of US microbiome research. Nat Microbiol 1:15015.
6. Anthony SJ, Islam A, Johnson C, Navarrete-Macias I, Liang E, Jain K, Hitchens PL, Che X, Soloyvov A, Hicks AL, Ojeda-Flores R, Zambrana-Torrelio C, Ulrich W, Rostal MK, Petrosov A, Garcia J, Haider N, Wolfe N, Goldstein T, Morse SS, Rahman M, Epstein JH, Mazet JK, Daszak P, Lipkin WI. 2015. Non-random patterns in viral diversity. Nat Commun 6:8147.
7. Bergner LM, Orton RJ, da Silva Filipe A, Shaw AE, Becker DJ, Tello C, Biek R, Streicker DG. 2019. Using noninvasive metagenomics to characterize viral communities from wildlife. Mol Ecol Resour 19:128-143.
8. Arze CA, Springer S, Dudas G, Patel S, Bhattacharyya A, Swaminathan H, Brugnara C, Delagrave S, Ong T, Kahvejian A, Echelard Y, Weinstein EG, Hajjar RJ, Andersen KG, Yozwiak NL. 2021. Global genome analysis reveals a vast and dynamic anellovirus landscape within the human virome. Cell Host & Microbe 29:1305-1315.e6.
9. Shi M, Neville P, Nicholson J, Eden JS, Imrie A, Holmes EC. 2017. High-Resolution Metatranscriptomics Reveals the Ecological Dynamics of Mosquito-Associated RNA Viruses in Western Australia. J Virol 91.
10. Pettersson JH, Shi M, Eden JS, Holmes EC, Hesson JC. 2019. Meta-Transcriptomic Comparison of the RNA Viromes of the Mosquito Vectors Culex pipiens and Culex torrentium in Northern Europe. Viruses 11.

11. de Almeida JP, Aguiar ER, Armache JN, Olmo RP, Marques JT. 2021. The virome of vector mosquitoes. Curr Opin Virol 49:7-12.
12. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ. 2016. Redefining the invertebrate RNA virosphere. Nature 540:539-543.
13. Atoni E, Wang Y, Karungu S, Waruhiu C, Zohaib A, Obanda V, Agwanda B, Mutua M, Xia H, Yuan Z. 2018. Metagenomic Virome Analysis of Culex Mosquitoes from Kenya and China. Viruses 10.
14. Fauver JR, Grubaugh ND, Krajacich BJ, Weger-Lucarelli J, Lakin SM, Fakoli LS, 3rd, Bolay FK, Diclaro JW, 2nd, Dabire KR, Foy BD, Brackney DE, Ebel GD, Stenglein MD. 2016. West African Anopheles gambiae mosquitoes harbor a taxonomically diverse virome including new insect-specific flaviviruses, mononegaviruses, and totiviruses. Virology 498:288-299.
15. Sadeghi M, Altan E, Deng X, Barker CM, Fang Y, Coffey LL, Delwart E. 2018. Virome of >12 thousand Culex mosquitoes from throughout California. Virology 523:74-88.
16. Wolf YI, Kazlauskas D, Iranzo J, Lucia-Sanz A, Kuhn JH, Krupovic M, Dolja VV, Koonin EV. 2018. Origins and Evolution of the Global RNA Virome. MBio 9.
17. Elrefaey AM, Abdelnabi R, Rosales Rosas AL, Wang L, Basu S, Delang L. 2020. Understanding the Mechanisms Underlying Host Restriction of Insect-Specific Viruses. Viruses 12.
18. Hall RA, Bielefeldt-Ohmann H, McLean BJ, O'Brien CA, Colmant AM, Piyasena TB, Harrison JJ, Newton ND, Barnard RT, Prow NA, Deerain JM, Mah MG, Hobson-Peters J. 2016. Commensal Viruses of Mosquitoes: Host Restriction, Transmission, and Interaction with Arboviral Pathogens. Evol Bioinform Online 12:35-44.
19. Altinli M, Schnettler E, Sicard M. 2021. Symbiotic Interactions Between Mosquitoes and Mosquito Viruses. Frontiers in cellular and infection microbiology 11:694020-694020.
20. Gaye A, Diagne MM, Ndiaye EH, Dior Ndione MH, Faye M, Talla C, Fall G, Ba Y, Diallo D, Dia I, Handschumacher P, Faye O, Sall AA, Diallo M. 2020. Vector competence of anthropophilic mosquitoes for a new mesonivirus in Senegal. Emerg Microbes Infect 9:496-504.
21. Ye G, Wang Y, Liu X, Dong Q, Cai Q, Yuan Z, Xia H. 2020. Transmission competence of a new mesonivirus, Yichang virus, in mosquitoes and its interference with representative flaviviruses. PLoS Negl Trop Dis 14:e0008920.
22. Baidaliuk A, Miot EF, Lequime S, Moltini-Conclois I, Delaigue F, Dabo S, Dickson LB, Aubry F, Merkling SH, Cao-Lormeau VM, Lambrechts L. 2019. Cell-Fusing Agent Virus Reduces Arbovirus Dissemination in Aedes aegypti Mosquitoes In Vivo. J Virol 93.
23. White AV, Fan M, Mazzara JM, Roper RL, Richards SL. 2021. Mosquito-infecting virus Espirito Santo virus inhibits replication and spread of dengue virus. J Med Virol 93:3362-3373.
24. Goenaga S, Kenney JL, Duggal NK, Delorey M, Ebel GD, Zhang B, Levis SC, Enria DA, Brault AC. 2015. Potential for Co-Infection of a Mosquito-Specific Flavivirus, Nhumirim Virus, to Block West Nile Virus Transmission in Mosquitoes. Viruses 7:5801-12.
25. Hall-Mendelin S, McLean BJ, Bielefeldt-Ohmann H, Hobson-Peters J, Hall RA, van den Hurk AF. 2016. The insect-specific Palm Creek virus modulates West Nile virus infection in and transmission by Australian mosquitoes. Parasit Vectors 9:414.
26. Bolling BG, Olea-Popelka FJ, Eisen L, Moore CG, Blair CD. 2012. Transmission dynamics of an insect-specific flavivirus in a naturally infected Culex pipiens laboratory colony and effects of co-infection on vector competence for West Nile virus. Virology 427:90-7.
27. Shi C, Beller L, Deboutte W, Yinda KC, Delang L, Vega-Rúa A, Failloux A-B, Matthijnssens J. 2019. Stable distinct core eukaryotic viromes in different mosquito species from Guadeloupe, using single mosquito viral metagenomics. Microbiome 7:121.
28. Shi C, Zhao L, Atoni E, Zeng W, Hu X, Matthijnssens J, Yuan Z, Xia H. 2020. Stability of the Virome in Lab- and Field-Collected Aedes albopictus Mosquitoes across Different Developmental Stages and Possible Core Viruses in the Publicly Available Virome Data of Aedes Mosquitoes. mSystems 5.

525 29. Agrawal AA, Ackerly DD, Adler F, Arnold AE, Cáceres C, Doak DF, Post E, Hudson PJ, Maron J,
526 Mooney KA, Power M, Schemske D, Stachowicz J, Strauss S, Turner MG, Werner E. 2007.
527 Filling key gaps in population and community ecology. Frontiers in Ecology and the
528 Environment 5:145-152.
529 30. Logue JB, Mouquet N, Peter H, Hillebrand H. 2011. Empirical approaches to
530 metacommunities: a review and comparison with theory. Trends in Ecology & Evolution
531 26:482-491.
532 31. Berg G, Rybakova D, Fischer D, Cernava T, Vergès M-CC, Charles T, Chen X, Cocolin L, Eversole
533 K, Corral GH, Kazou M, Kinkel L, Lange L, Lima N, Loy A, Macklin JA, Maguin E, Mauchline T,
534 McClure R, Mitter B, Ryan M, Sarand I, Smidt H, Schelkle B, Roume H, Kiran GS, Selvin J,
535 Souza RSCd, van Overbeek L, Singh BK, Wagner M, Walsh A, Sessitsch A, Schloter M. 2020.
536 Microbiome definition re-visited: old concepts and new challenges. Microbiome 8:103.
537 32. Shade A, Handelsman J. 2012. Beyond the Venn diagram: the hunt for a core microbiome.
538 Environ Microbiol 14:4-12.
539 33. Konstantinidis K, Dovrolis N, Kouvela A, Kassela K, Rosa Freitas MG, Nearchou A, de Courcy
540 Williams M, Veletza S, Karakasiliotis I. 2022. Defining virus-carrier networks that shape the
541 composition of the mosquito core virome of a local ecosystem. Virus Evol 8:veac036.
542 34. Coatsworth H, Bozic J, Carrillo J, Buckner EA, Rivers AR, Dinglasan RR, Mathias DK. 2022.
543 Intrinsic variation in the vertically transmitted core virome of the mosquito Aedes aegypti.
544 Molecular Ecology 31:2545-2561.
545 35. Thongsripong P, Chandler JA, Kittayapong P, Wilcox BA, Kapan DD, Bennett SN. 2021.
546 Metagenomic shotgun sequencing reveals host species as an important driver of virome
547 composition in mosquitoes. Sci Rep 11:8448.
548 36. Shi C, Beller L, Deboutte W, Yinda KC, Delang L, Vega-Rúa A, Failloux AB, Matthijnssens J.
549 2019. Stable distinct core eukaryotic viromes in different mosquito species from Guadeloupe,
550 using single mosquito viral metagenomics. Microbiome 7:121.
551 37. Parry R, James ME, Asgari S. 2021. Uncovering the Worldwide Diversity and Evolution of the
552 Virome of the Mosquitoes Aedes aegypti and Aedes albopictus. Microorganisms 9.
553 38. He X, Yin Q, Zhou L, Meng L, Hu W, Li F, Li Y, Han K, Zhang S, Fu S, Zhang X, Wang J, Xu S,
554 Zhang Y, He Y, Dong M, Shen X, Zhang Z, Nie K, Liang G, Ma X, Wang H. 2021. Metagenomic
555 sequencing reveals viral abundance and diversity in mosquitoes from the Shaanxi-Gansu-
556 Ningxia region, China. PLoS Negl Trop Dis 15:e0009381.
557 39. Feng Y, Gou QY, Yang WH, Wu WC, Wang J, Holmes EC, Liang G, Shi M. 2022. A time-series
558 meta-transcriptomic analysis reveals the seasonal, host, and gender structure of mosquito
559 viromes. Virus Evol 8:veac006.
560 40. Fonseca DM, Keyghobadi N, Malcolm CA, Mehmet C, Schaffner F, Mogi M, Fleischer RC,
561 Wilkerson RC. 2004. Emerging Vectors in the Culex pipiens Complex. Science 303:1535-1538.
562 41. Shi M, Zhang Y-Z, Holmes EC. 2018. Meta-transcriptomics and the evolutionary biology of
563 RNA viruses. Virus research 243:83-90.
564 42. Aardema ML, vonHoldt BM, Fritz ML, Davis SR. 2020. Global evaluation of taxonomic
565 relationships and admixture within the Culex pipiens complex of mosquitoes. Parasites &
566 vectors 13:8-8.
567 43. Cheval J, Sauvage V, Frangeul L, Dacheux L, Guigon G, Dumey N, Pariente K, Rousseaux C,
568 Dorange F, Berthet N, Brisse S, Moszer I, Bourhy H, Manuguerra CJ, Lecuit M, Burguiere A,
569 Caro V, Eloit M. 2011. Evaluation of high-throughput sequencing for identifying known and
570 unknown viruses in biological samples. Journal of clinical microbiology 49:3268-3275.
571 44. Gil P, Dupuy V, Koual R, Exbrayat A, Loire E, Fall AG, Gimonneau G, Biteye B, Talla Seck M,
572 Rakotoarivony I, Marie A, Frances B, Lambert G, Reveillaud J, Balenghien T, Garros C, Albina
573 E, Eloit M, Gutierrez S. 2021. A library preparation optimized for metagenomics of RNA
574 viruses. Mol Ecol Resour doi:10.1111/1755-0998.13378.
575 45. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler
576 transform. Bioinformatics 25:1754-1760.

577 46. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution
578 for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics
579 31:1674-1676.
580 47. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.
581 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-9.
582 48. Bennouna A, Gil P, El Rhaffouli H, Exbrayat A, Loire E, Balenghien T, Chlyeh G, Gutierrez S,
583 Fihri OF. 2019. Identification of Eilat virus and prevalence of infection among Culex pipiens L.
584 populations, Morocco, 2016. Virology 530:85-88.
585 49. Stanojevic M, Li K, Stamenkovic G, Ilic B, Paunovic M, Pesic B, Maslovara ID, Siljic M, Cirkovic
586 V, Zhang Y. 2020. Depicting the RNA Virome of Hematophagous Arthropods from Belgrade,
587 Serbia. Viruses 12.
588 50. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics
589 Analysis across Computing Platforms. Molecular Biology and Evolution 35:1547-1549.
590 51. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated
591 alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972-3.
592 52. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R.
593 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the
594 Genomic Era. Molecular Biology and Evolution 37:1530-1534.
595 53. Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. 2017. ggtree: an r package for visualization and
596 annotation of phylogenetic trees with their covariates and other associated data. Methods in
597 Ecology and Evolution 8:28-36.
598 54. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin P, O'Hara RB,
599 Simpson G, Solymos P, Stevens H, Szoecs E, Wagner H. 2019. vegan: Community Ecology
600 Package. R package version 2.5-6., https://CRAN.R-project.org/package=vegan.
601 55. McKnight DT, Huerlimann R, Bower DS, Schwarzkopf L, Alford RA, Zenger KR. 2019. Methods
602 for normalizing microbiome data: An ecological perspective. Methods in Ecology and
603 Evolution 10:389-400.
604 56. Biggerstaff BJ. 2008. Confidence Intervals for the Difference of Two Proportions Estimated
605 from Pooled Samples. Journal of Agricultural, Biological, and Environmental Statistics 13:478-
606 496.
607 57. Griffith DM, Veech JA, Marsh CJ. 2016. cooccur: Probabilistic Species Co-Occurrence Analysis
608 in R. Journal of Statistical Software, Code Snippets 69:1 - 17.
609 58. Veech JA. 2013. A probabilistic model for analysing species co-occurrence. Global Ecology
610 and Biogeography 22:252-260.
611 59. Junglen S, Kurth A, Kuehl H, Quan PL, Ellerbrok H, Pauli G, Nitsche A, Nunn C, Rich SM, Lipkin
612 WI, Briese T, Leendertz FH. 2009. Examining landscape factors influencing relative
613 distribution of mosquito genera and frequency of virus infection. Ecohealth 6:239-49.
614 60. Medlock JM, Hansford KM, Schaffner F, Versteirt V, Hendrickx G, Zeller H, Bortel WV. 2012. A
615 Review of the Invasive Mosquitoes in Europe: Ecology, Public Health Risks, and Control
616 Options. Vector-Borne and Zoonotic Diseases 12:435-447.
617 61. Vasilakis N, Guzman H, Firth C, Forrester NL, Widen SG, Wood TG, Rossi SL, Ghedin E, Popov
618 V, Blasdell KR, Walker PJ, Tesh RB. 2014. Mesoniviruses are mosquito-specific viruses with
619 extensive geographic distribution and host range. Virol J 11:97.
620 62. Bergner LM, Orton RJ, Benavides JA, Becker DJ, Tello C, Biek R, Streicker DG. 2020.
621 Demographic and environmental drivers of metagenomic viral diversity in vampire bats. Mol
622 Ecol 29:26-39.
623 63. Batson J, Dudas G, Haas-Stapleton E, Kistler AL, Li LM, Logan P, Ratnasiri K, Retallack H. 2021.
624 Single mosquito metatranscriptomics identifies vectors, emerging pathogens and reservoirs
625 in one assay. Elife 10.
626 64. Brum Jennifer R, Ignacio-Espinoza JC, Roux S, Doulcier G, Acinas Silvia G, Alberti A, Chaffron S,
627 Cruaud C, de Vargas C, Gasol Josep M, Gorsky G, Gregory Ann C, Guidi L, Hingamp P, Iudicone
628 D, Not F, Ogata H, Pesant S, Poulos Bonnie T, Schwenck Sarah M, Speich S, Dimier C, Kandels-

Lewis S, Picheral M, Searson S, null n, Bork P, Bowler C, Sunagawa S, Wincker P, Karsenti E, Sullivan Matthew B. 2015. Patterns and ecological drivers of ocean viral communities. Science 348:1261498.

65. Guégan M, Zouache K, Démichel C, Minard G, Tran Van V, Potier P, Mavingui P, Valiente Moro C. 2018. The mosquito holobiont: fresh insight into mosquito-microbiota interactions. Microbiome 6:49-49.

66. Muturi EJ, Kim C-H, Bara J, Bach EM, Siddappaji MH. 2016. Culex pipiens and Culex restuans mosquitoes harbor distinct microbiota dominated by few bacterial taxa. Parasites & vectors 9:18-18.

67. Zhang YZ, Shi M, Holmes EC. 2018. Using Metagenomics to Characterize an Expanding Virosphere. Cell 172:1168-1172.

Figures

Figure 1. Overview of mosquito sampling sites in the Mediterranean Basin. Detailed information on the sites can be found in Fig. S7 and Table S4.

Figure 2. Structure of the eukaryotic virome of *Culex pipiens* in the Mediterranean Basin. (A) Heatmap showing the relative abundance of clusters among libraries. Clusters are ranked on the y axis first by number of libraries with presence and, then, by total read number. The number of libraries in which a given cluster was found is shown on the right of the cluster name. This number is underlined in blue if the cluster was detected in all countries. The darker the tile color, the closer the relative frequency of a given cluster to one. Libraries are shown on the x axis with the European countries on the right and the African ones on the left to facilitate comparison. The terms "high" and "low" on the library names stand for levels of habitat anthropization. (B) Boxplots with relative abundance of clusters. Clusters are ranked as in the A panel. Dots indicate frequency in a given library and the dot color stands for the library as shown in the legend.

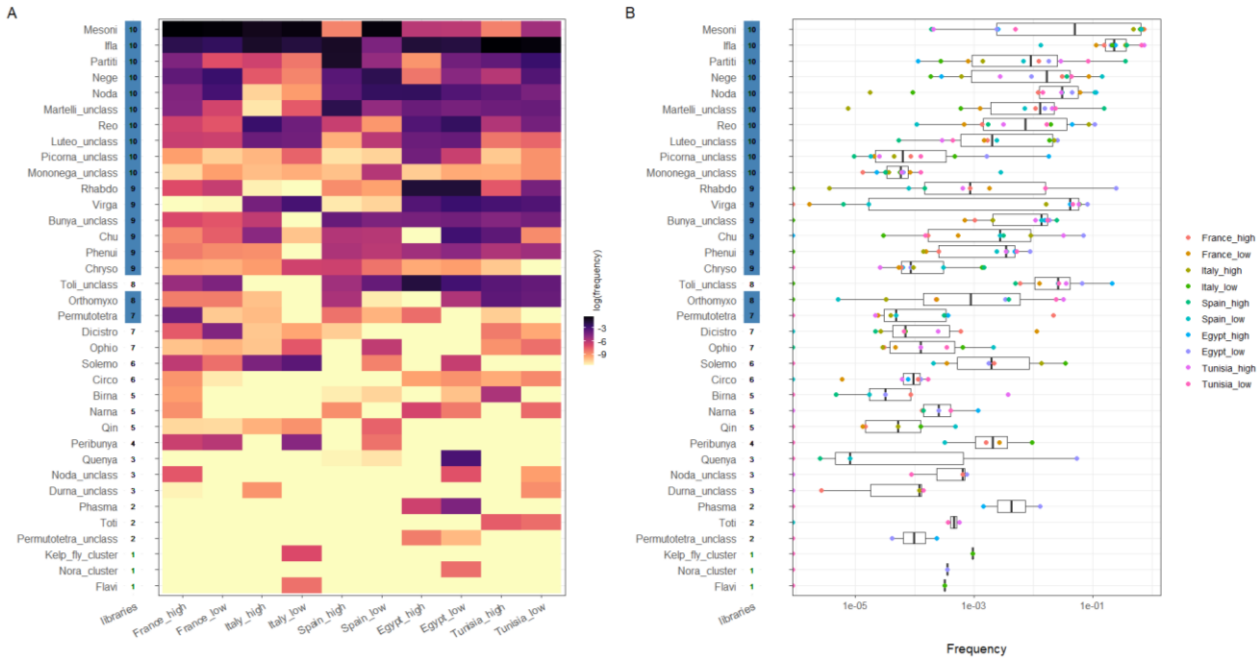


Figure 3. Prevalence of viral operational taxonomic units (vOTUs) from highly-prevalent clusters in the eukaryotic virome of *Culex pipiens* among libraries. The names of vOTUs and the associated clusters appear on the right and left of the heatmap, respectively. Colored tiles indicate detection in a given library. Tile color stands for cluster to facilitate visualization.

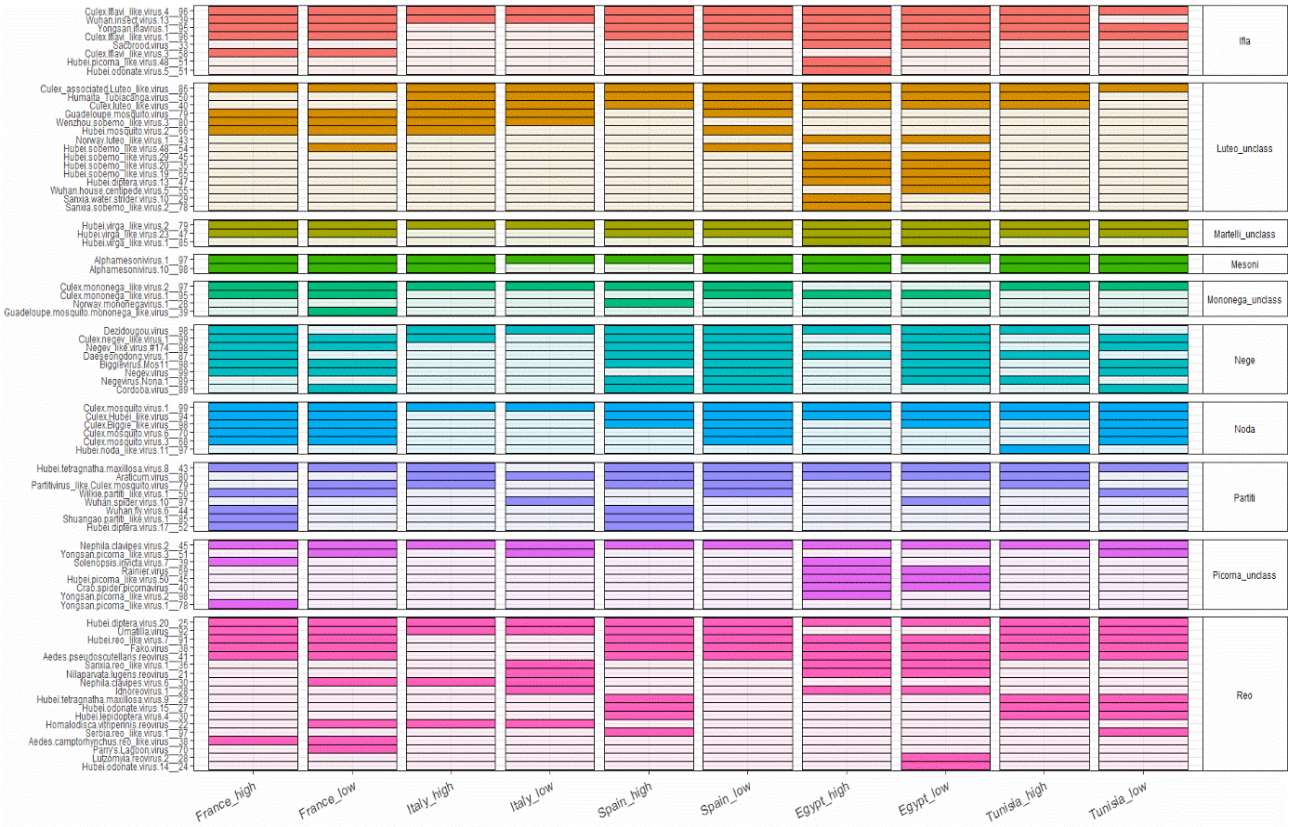


Figure 4. Taxon conservation in the virome of *Culex pipiens* between our study and studies from other countries. Grey tiles indicate highly-conserved clusters shared in our dataset and in a specific country (either China, Morocco, Serbia or Sweden). Numbers within tiles indicate the number of viral operational taxonomic units (vOTUs) of a given cluster also detected in our dataset. The vOTU number appears in bold if the vOTU with the highest prevalence in our dataset was also found in a given country (see Fig. 3 for those vOTUs).

	China	Morocco	Serbia	Sweden
Reo	0	0	1	1
Picorna_unclass	1	0	1	0
Partiti	0	0	0	0
Noda	1	**2**	0	1
Nege	1	4	0	1
Mononega_unclass	**2**	**2**	**1**	**2**
Mesoni	0	**1**	0	**1**
Martelli_unclass	**1**	**1**	**1**	1
Luteo_unclass	**4**	0	**2**	**2**
Ifla	**3**	**3**	**1**	0

Figure 5. Infection rates of five viral operational taxonomic units (vOTUs) in *Culex pipiens* females. The five vOTUs were Culex.Bunyavirus.2__98 (Bunya_unclass), Culex.mosquito.virus.4__97 (Chu), Culex.Iflavi_like.virus.4__96 (Iflavi), Alphamesonivirus.1__97 (Mesoni) and Culex.negev_like.virus.1__99 (Nege). Infection rates are grouped in panels depending on the cluster of the vOTU. Dots indicate the maximum likelihood estimates of the infection rates (infected females per 100 females) and vertical segments stand for the 95% confidence intervals. (A) Infection rates of the five vOTUs in the five countries sampled in 2015. EG: Egypt, ES: Spain, FR: France, IT: Italy, TN: Tunisia. (B) Infection rates of the five virus taxa in France in 2015 and 2018. The infection rate of the vOTU in the Mesoni cluster in Morocco in 2016 is also included to facilitate comparison (marked with a triangle).

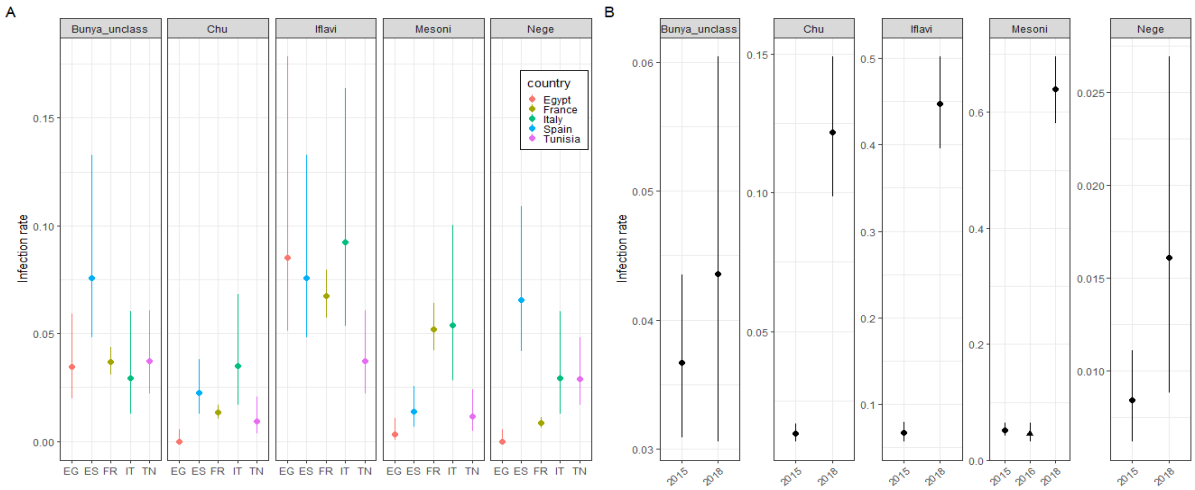
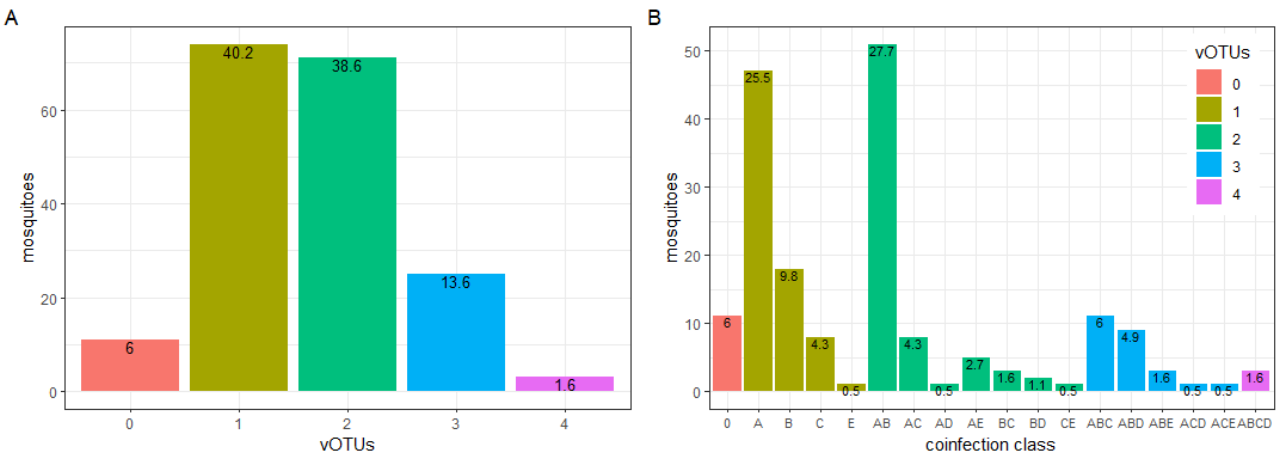
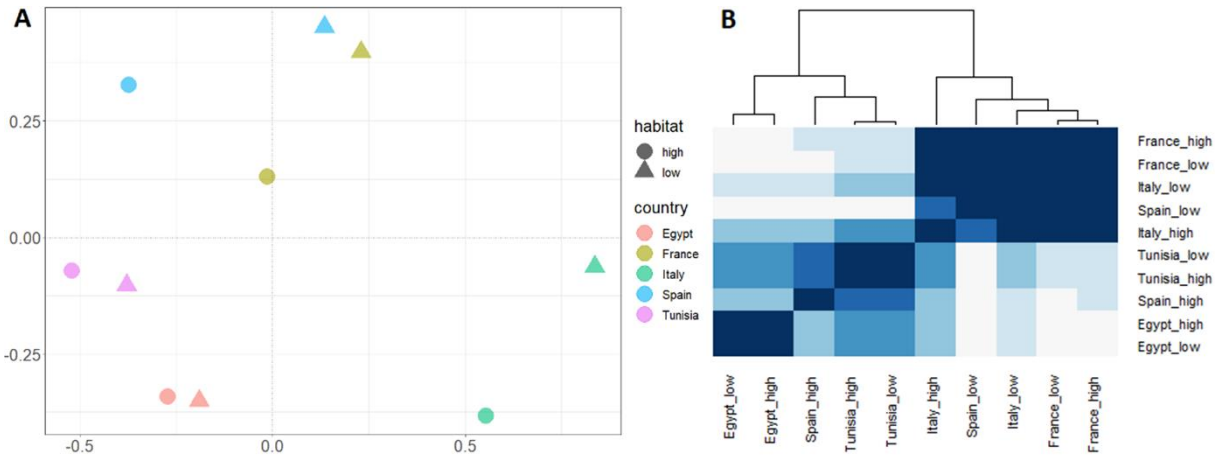


Figure 6. Distribution of viral operational taxonomic units (vOTUs) in individual *Culex pipiens* females. The five vOTUs were Culex.Bunyavirus.2__98 (Bunya_unclass), Culex.mosquito.virus.4__97 (Chu), Culex.Iflavi_like.virus.4__96 (Iflavi), Alphamesonivirus.1__97 (Mesoni) and Culex.negev_like.virus.1__99 (Nege). (A) Distribution of the number of vOTUs in individual mosquitoes. Bar color stands for the number of vOTUs per mosquito. The relative frequencies of each coinfection group are shown on top of the bars (in percentages). (B) Distribution of combinations of specific vOTUs (*i.e.* coinfection classes) found in individual *C. pipiens* females. Relative frequencies of each coinfection class are shown on top of the bars (in percentages). Bars are colored following the number of vOTUs per mosquito, as in panel A. Each capital letter stands for a given vOTU as follows. A: Alphamesonivirus.1__97 (Mesoni), B: Culex.Iflavi_like.virus.4__96 (Ifla), C: Culex.Bunyavirus.2__98 (Bunya_unclass), D: Culex.negev.like_virus.1__99 (Nege), and E: Culex.mosquito.virus.4__97 (Chu).



692 **Figure 7**. Influence of country and anthropization level (habitat) on virome structure. (A) Non-metric
693 multidimensional scaling plot on Bray-Curtis dissimilarities of the viromes in the ten libraries. (B)
694 Heatmap showing Pearson correlation coefficients between the viromes in the ten libraries. A
695 dendrogram based on hierarchical clustering is shown on the top. Tile color represents Pearson
696 coefficients, with color becoming darker along with the Pearson coefficient.



697

698

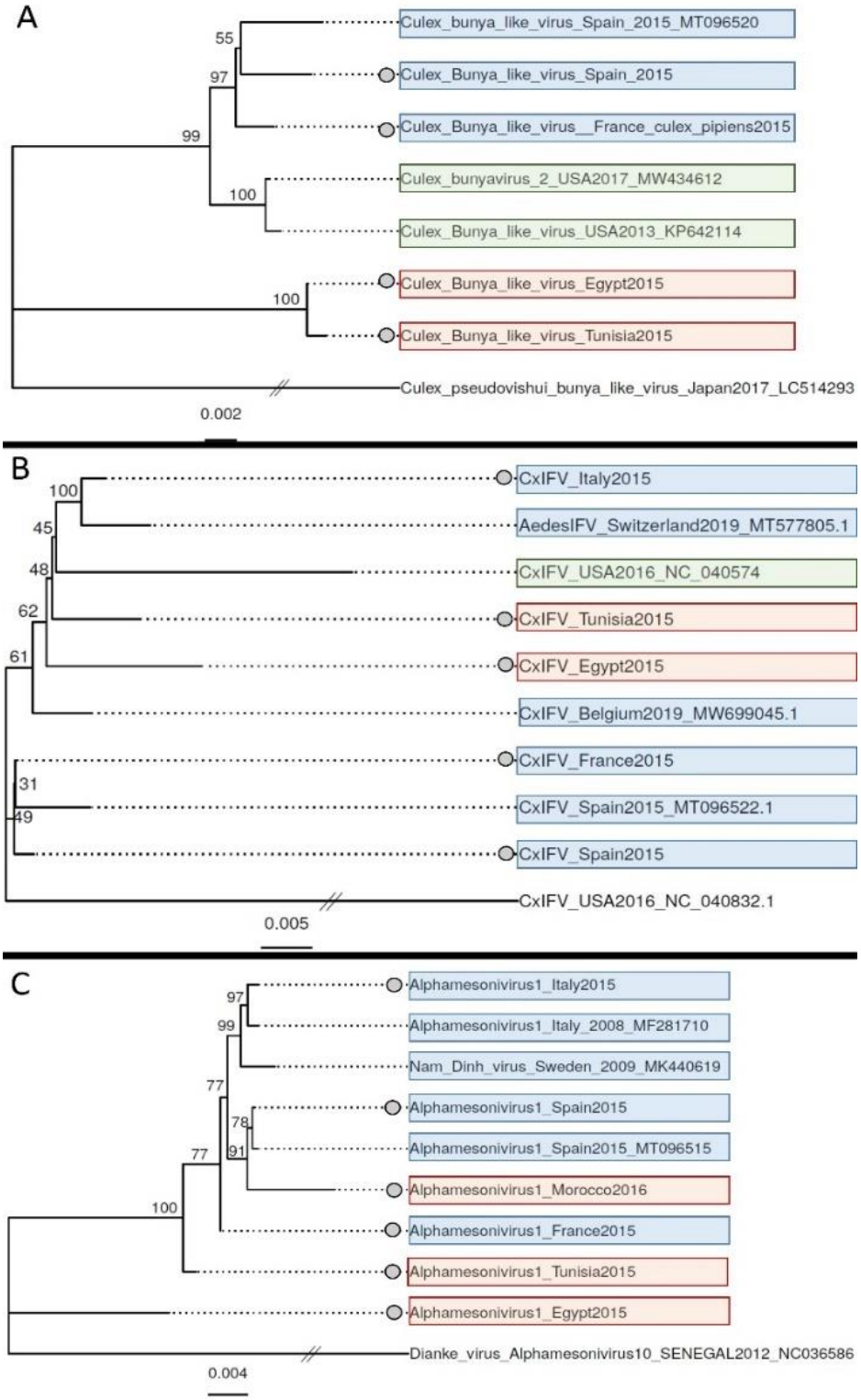



Figure 8. Maximum-likelihood phylogenetic trees of three viral operational taxonomic units: Culex.Bunyavirus.2__98 (A), Culex.Iflavi_like.virus.4__96 (CxIFV) (B), and Alphamesonivirus.1__97 (C). Sequences from this study are marked with a grey dot on the left of the sequence name. A divergent species from the same family was used as outgroup to root each tree (lowest branch in all the trees). The scale bar represents the number of nucleotide substitutions per site. Background color stands for geographical origin (Europe: blue, Africa: red, America: green). GenBank accession numbers of sequences from other studies are provided on the branch names.