The evolution, diversity and host associations of rhabdoviruses

Ben Longdon[1*], Gemma GR Murray[1], William J Palmer[1], Jonathan P Day[1], Darren J Parker[2], John J Welch[1], Darren J Obbard[3] and Francis M Jiggins[1].

[1]Department of Genetics
University of Cambridge
Cambridge
CB2 3EH
UK

[2] School of Biology
University of St. Andrews
St. Andrews
KY19 9ST
UK

[3] Institute of Evolutionary Biology, and Centre for Immunity Infection and Evolution
University of Edinburgh
Edinburgh
EH9 3JT
UK

*corresponding author
email: b.longdon@gen.cam.ac.uk
phone: +441223333945

Abstract

The rhabdoviruses are a diverse family of RNA viruses that includes important pathogens of humans, animals and plants. We have discovered the sequences of 32 new rhabdoviruses through a combination of our own RNA sequencing of insects and searching public sequence databases. Combining this with previously known sequences we reconstructed the phylogeny of 195 rhabdoviruses producing the most in depth analysis of the family to date. In most cases we know nothing about the biology of the viruses beyond the host they were isolated from, but our dataset provides a powerful way to phylogenetically predict which are vector-borne pathogens and which are specific to vertebrates or arthropods. This allowed us to identify 76 new likely vector-borne vertebrate pathogens among viruses isolated from vertebrates or biting insects. By reconstructing ancestral states, we found that switches between major groups of hosts have occurred rarely during rhabdovirus evolution, with single transitions giving rise to clades of plant pathogens, vertebrate-specific pathogens, and arthropod-borne pathogens of vertebrates. There are also two large clades of viruses that infect insects, including the sigma viruses, which are vertically transmitted. There are also few transitions between aquatic and terrestrial ecosystems. Our data suggest that throughout their evolution rhabdoviruses have occasionally made a long distance host jump, before spreading through related hosts in the same environment.

Keywords

Virus, Host shift, Arthropod, Insect, Rhabdoviridae, Mononegavirales

53 **Introduction**
54
55 RNA viruses are an abundant and diverse group of pathogens. In the past, viruses were
56 typically isolated from hosts displaying symptoms of infection, before being
57 characterized morphologically and then sequenced following PCR [1, 2]. PCR-based
58 sequencing of novel RNA viruses is problematic as there is no single conserved region of
59 the genome of viruses from a single family, let alone all RNA viruses. High throughput
60 next generation sequencing technology has revolutionized virus discovery, allowing
61 rapid detection and sequencing of divergent virus sequences simply by sequencing total
62 RNA from infected individuals [1, 2]
63
64 One particularly diverse family of RNA viruses is the *Rhabdoviridae*. Rhabdoviruses are
65 negative-sense single-stranded RNA viruses in the order Mononegavirales [3]. They
66 infect an extremely broad range of hosts and have been discovered in plants, fish,
67 mammals, reptiles and a broad range of insects and other arthropods [4]. The family
68 includes important pathogens of humans and livestock. Perhaps the most well known is
69 Rabies virus, which can infect a diverse array of mammals and causes a fatal infection; it
70 kills 59,000 humans per year with an estimated economic cost of US$8.6 billion [5].
71 Other rhabdoviruses such as Vesicular Stomatitis Virus and Bovine Ephemeral Fever
72 Virus are important pathogens of domesticated animals, whilst others are pathogens of
73 crops [3].
74
75 Arthropods play a key role in transmission of many rhabdoviruses. Many viruses found
76 in vertebrates have also been detected in arthropods, including sandflies, mosquitoes,
77 ticks and midges [6]. The rhabdoviruses that infect plants are also often transmitted by
78 arthropods [7]. Even the rhabdoviruses that infect fish have the potential to be vectored
79 by ectoparasitic copepod sea-lice [8]. Rhabdoviruses replicate upon infection of insects
80 (insects are not just mechanical vectors), which may explain why they are insects are
81 common rhabdovirus vectors [7].
82
83 Other rhabdoviruses are insect-specific. In particular, the sigma viruses are a clade of
84 vertically transmitted viruses that infect dipterans and are well-studied in *Drosophila*
85 [9-11]. Recently, a number of rhabdoviruses have been found to be associated with a
86 wide array of insect and other arthropod species, suggesting they may be common
87 arthropod pathogens [12, 13]. Furthermore, a number of arthropod genomes contain
88 integrated Endogenous Viral Elements (EVEs) with similarity to free-living
89 rhabdoviruses, suggesting that these species have been infected with rhabdoviruses
90 [14-17].
91
92 Here we aimed to uncover the diversity of the rhabdoviruses, and examine how they
93 have switched between different host taxa during their evolutionary history. Insects
94 infected with rhabdoviruses commonly become paralysed on exposure to CO_2 [18-20].
95 We exploited this fact to screen field collections of flies from several continents for novel
96 rhabdoviruses that were then sequenced using RNA-sequencing (RNA-seq). Additionally
97 we searched for rhabdovirus-like sequences in publicly available RNA-seq data. We
98 identified 34 novel rhabdovirus-like sequences from a wide array of invertebrates and
99 plants, and combined them with recently discovered viruses to produce the most

100 comprehensive phylogeny of the rhabdoviruses to date. For many of the viruses we do
101 not know their true host range, so we used the phylogeny to identify a large number of
102 new likely vector-borne pathogens and to reconstruct the evolutionary history of this
103 diverse group of viruses.
104
105
106 **Methods**
107
108 *Discovery of new rhabdoviruses by RNA sequencing*
109
110 Diptera species (flies, mostly Drosophilidae) were collected in the field from Spain, USA,
111 Kenya, France, Ghana and the UK (Data S1:
112 http://dx.doi.org/10.6084/m9.figshare.1425432). Infection with Rhabdoviruses can
113 cause *Drosophila* and other insects to become paralysed after exposure to CO_2 [18-20],
114 so we enriched our sample for infected individuals by exposing them to CO_2 at 12°C for
115 15mins, only retaining individuals that showed symptoms of paralysis 30mins later. We
116 extracted RNA from 79 individual insects (details in Data S1
117 http://dx.doi.org/10.6084/m9.figshare.1425432) using Trizol reagent (Invitrogen) and
118 combined the extracts into two pools. RNA was then rRNA depleted with the Ribo-Zero
119 Gold kit (epicenter, USA) and used to construct Truseq total RNA libraries (Illumina).
120 Libraries were constructed and sequenced by BGI (Hong Kong) on an Illumina Hi-Seq
121 2500 (one lane, 100bp paired end reads, generating ~175 million reads). Sequences
122 were quality trimmed with Trimmomatic (v3); Illumina adapters were clipped, bases
123 were removed from the beginning and end of reads if quality dropped below a
124 threshold, sequences were trimmed if the average quality within a window fell below a
125 threshold and reads less than 20 base pairs in length were removed. We *de novo*
126 assembled the RNA-seq reads with Trinity (release 2013-02-25) using default settings
127 and jaccard clip option for high gene density. The assembly was then blasted (tblastn)
128 with rhabdovirus coding sequences as the query to identify rhabdovirus-like sequences.
129 Contigs with hits were then reciprocally blasted against Genbank cDNA and RefSeq
130 databases and only retained if they hit a virus-like sequence. Raw read data were
131 deposited in the NCBI Sequence Read Archive (SRP057824). Putative viral sequences
132 have been submitted to Genbank (accessions in Tables S1 and S2).
133
134 As the RNA-seq was performed on pooled samples, we assigned rhabdovirus sequences
135 to individual insects by PCR. cDNA was produced using Promega GoScript Reverse
136 Transcriptase and random-hexamer primers, and PCR performed using primers
137 designed using the rhabdovirus sequences. Infected host species were identified by
138 sequencing the mitochondrial gene *COI*. We were unable to identify the host species of
139 the virus from a *Drosophila affinis* sub-group species (sequences appear similar to both
140 *Drosophila affinis* and the closely related Drosophila *athabasca*), despite using other
141 mitochondrial and nuclear genes to try and identify the species with certainty. We
142 confirmed all sequences were only present in RNA using PCR, and so are likely free
143 living viruses rather than being integrated into the insect genome (i.e. endogenous virus
144 elements or EVEs [16]).
145

146 We identified sigma virus sequences in RNA-seq data from *Drosophila montana* [21]. We
147 amplified the virus from an infected fly line by RT-PCR and carried out additional Sanger
148 sequencing with primers designed using the RNA-seq assembly. Additional sequences
149 from an RNA-seq analysis of pools of wild caught *Drosophila*: DImmSV from *Drosophila*
150 *immigrans* (collection and sequencing described [22]), DTriSV from a pool of *Drosophila*
151 *tristis* and SDefSV from *Scaptodrosophila deflexa* (both Darren Obbard, unpublished
152 data), accessions in tables S1 and S2.
153
154 *Discovery of rhabdoviruses in public sequence databases*
155
156 Rhabdovirus L gene sequences were used to search (tblastn) against expressed
157 sequence tag (EST) and transcriptome shotgun assembly (TSA) databases (NCBI). All
158 hits were reciprocally blasted against Genbank cDNA and RefSeq databases and only
159 retained if they hit a virus-like sequence. We used two approaches to examine whether
160 sequences were present as RNA but not DNA. First, where assemblies of whole-genome
161 shotgun sequences were available, sequences were blasted to check whether they were
162 integrated into the host genome. Second, for the virus sequences in the butterfly *Pararge*
163 *aegeria* and the medfly *Ceratitis capitata* we were able to obtain infected samples to
164 confirm the sequences are only present in RNA by performing PCR on both genomic
165 DNA and cDNA (samples kindly provided by Casper Breuker/Melanie Gibbs, and Philip
166 Leftwich respectively)
167
168 *Phylogenetic analysis*
169
170 All available rhabdovirus-like sequences were downloaded from Genbank (accessions in
171 Data S2: http://dx.doi.org/10.6084/m9.figshare.1425419). Amino acid sequences for
172 the L gene (encoding the RNA Dependent RNA Polymerase or RDRP) were used to infer
173 the phylogeny (L gene sequences: http://dx.doi.org/10.6084/m9.figshare.1425067), as
174 they contain conserved domains that can be aligned across this diverse group of viruses.
175 Sequences were aligned with MAFFT [23] under default settings and then poorly aligned
176 and divergent sites were removed with either TrimAl (v1.3 strict settings, implemented
177 on Phylemon v2.0 server, alignment: http://dx.doi.org/10.6084/m9.figshare.1425069)
178 [24] or Gblocks (v0.91b selecting smaller final blocks, allowing gap positions and less
179 strict flanking positions to produce a less stringent selection, alignment:
180 http://dx.doi.org/10.6084/m9.figshare.1425068) [25]. These resulted in alignments of
181 1492 And 829 amino acids respectively.
182
183 Phylogenetic trees were inferred using Maximum Likelihood in PhyML (v3.0) [26] using
184 the LG substitution model [27], with a gamma distribution of rate variation with four
185 categories and using a sub-tree pruning and regrafting topology searching algorithm.
186 Branch support was estimated using Approximate Likelihood-Ratio Tests (aLRT) that is
187 reported to outperform bootstrap methods [28]. Figures were created using FIGTREE
188 (v. 1.4) [29].
189
190 *Reconstruction of host associations*
191

192 Viruses were categorised as having one of four types of host association: arthropod-
193 specific, vertebrate-specific, arthropod-vectored plant, or arthropod-vectored
194 vertebrate. However, the host association of some viruses are uncertain when they have
195 been isolated from vertebrates, biting-arthropods or plant-sap-feeding arthropods. Due
196 to limited sampling it was not clear whether viruses isolated from vertebrates were
197 vertebrate specific or arthropod-vectored vertebrate viruses; or whether viruses
198 isolated from biting-arthropods were arthropod specific viruses or arthropod-vectored
199 vertebrate viruses; or if viruses isolated from plant-sap-feeding arthropods were
200 arthropod-specific or arthropod-vectored plant viruses. We omitted three viruses that
201 were isolated from hosts outside of these four categories from our analyses.
202
203 We simultaneously estimated both the current and ancestral host associations, and the
204 phylogeny of the viruses, using a Bayesian analysis, implemented in BEAST v1.8 [30, 31].
205 Since accurate branch lengths are essential for this analysis, we used a subset of the
206 sites and strains used in the Maximum Likelihood analysis. We retained 189 taxa (all
207 rhabdoviruses excluding the divergent fish-infecting novirhabdovirus clade and the
208 virus from *Hydra*, as well as the viruses from *Lolium perenne and Conwentzia*
209 *psociformis*, which had a large number of missing sites). Sequences were trimmed to a
210 conserved region of 414 amino acids where data was recorded for most of these viruses
211 (the Gblocks alignment trimmed further by eye:
212 http://dx.doi.org/10.6084/m9.figshare.1425431). We used the host-association
213 categories described above, which included ambiguous states. To model amino acid
214 evolution we used an LG substitution model with gamma distributed rate variation
215 across sites [27] and an uncorrelated lognormal relaxed clock model of rate variation
216 among lineages [32]. To model the evolution of the host associations we used an
217 asymmetric transition rate matrix (allowing transitions to and from a host association to
218 take place at different rates) and a strict clock model. We used a constant population
219 size coalescent prior for the relative node ages and the BEAUti v1.8 default priors for all
220 other parameters [30] (BEAUti xml http://dx.doi.org/10.6084/m9.figshare.1431922).
221 Convergence was assessed using Tracer v1.6 [33], and a burn-in (of 30%) removed prior
222 to the construction of a consensus tree, which included a description of ancestral host
223 associations. High effective sample sizes were achieved for all parameters (>200).
224
225
226 The maximum clade credibility tree estimated
227 (http://dx.doi.org/10.6084/m9.figshare.1425436) for the host association
228 reconstruction was very similar to the independently estimated maximum likelihood
229 phylogeny (http://dx.doi.org/10.6084/m9.figshare.1425083), which made no
230 assumptions about the appropriateness or otherwise of applying a clock model. The
231 minor topological differences may be expected in reconstructions that differ in their
232 assumptions about evolutionary rates. In Figure 2 we have transferred the ancestral
233 state reconstruction from the BEAST tree to the maximum likelihood tree.
234
235
236 **Results**
237
238 *Novel rhabdoviruses from RNA-seq*

239
240 To search for new rhabdoviruses we collected a variety of different species of flies,
241 screened them for CO_2 sensitivity and sequenced total RNA of these flies by RNA-seq. We
242 identified rhabdovirus-like sequences from a *de-novo* assembly by BLAST, and used PCR
243 to identify which samples these sequences came from.
244
245 This approach resulted in eleven rhabdovirus-like sequences from nine (possibly ten)
246 species of fly. Seven of these viruses were previously unknown and four had been
247 reported previously from shorter sequences (Tables S1 and S2). Rhabdoviruses known
248 from other species of *Drosophila* typically have genomes of ~12.5Kb [11, 34], and six of
249 our sequences were approximately this size, suggesting they are near-complete
250 genomes. None of the viruses discovered in our RNA-seq data appeared to be integrated
251 into the host genome (see Methods for details).
252
253 To investigate the putative gene content of the viruses, we predicted genes based on
254 open reading frames (ORFs). For the viruses with apparently complete genomes (Figure
255 1), we found that those from *Drosophila ananassae, Drosophila affinis, Drosophila*
256 *immigrans* and *Drosophila sturtvanti* contained the five core genes found across all
257 rhabdoviruses, with an additional gene between the P and M genes. This is the location
258 of the X gene found in sigma viruses, and in three of the four viruses it showed sequence
259 homology to the X gene of other sigma viruses. The virus from *Drosophila busckii* did not
260 contain an additional ORF between the P and M genes, but instead contained an ORF
261 between the G and L gene. Using the gene content and the phylogeny described below,
262 we have classified our newly discovered viruses as either sigma viruses or other
263 rhabdoviruses and named them after the host species they were isolated from (Figure
264 1) [35]. We also found one other novel mononegavirales-like sequence from *Drosophila*
265 *unispina* that groups with a recently discovered clade of arthropod associated viruses
266 (Nyamivirus clade [12], see Table S5 and the full phylogeny:
267 http://dx.doi.org/10.6084/m9.figshare.1425083), confirming our approach can detect a
268 wide range of divergent viruses.
269
270 [Figure 1 here]
271
272 *New Rhabdoviruses from public databases*
273
274 We identified a further 26 novel rhabdovirus-like sequences by searching public
275 databases of assembled RNA-seq data with BLAST. These included 19 viruses from
276 arthropods (Fleas, Crustacea, Lepidoptera, Diptera), one from a Cnidarian (*Hydra*) and 5
277 from plants (Table S3). Of these viruses, 19 had sufficient amounts of coding sequence
278 (>1000bp) to include in the phylogenetic analysis (Table S3), whilst the remainder were
279 too short (Table S4).
280
281 Four viruses from databases had near-complete genomes. These were from the moth
282 *Triodia sylvina*, the house fly *Musca domestica* (99% nucleotide identity to Wuhan house
283 fly virus 2 [12]), the butterfly *Pararge aegeria* and the medfly *Ceratitis capitata*, all of
284 which contain the five core rhabdovirus genes. The sequence from *C. capitata* had an
285 additional gene predicted between the P and M genes with sequence similarity to the X

286 gene in sigma viruses. There were several unusual sequences. Firstly, in the virus from
287 *P. aegeria* there appear to be two full-length glycoprotein genes between the M and L
288 gene (we confirmed the stop codon between the two genes was not an error by Sanger
289 sequencing). Secondly, the *Agave tequilana* transcriptome contained an L gene ORF on a
290 contig that was the length of a typical rhabdovirus genome but did not appear to contain
291 typical gene content, suggesting it has very atypical genome organization, has been
292 misassembled or is integrated into its host plant genome [36]. Finally, the virus from
293 *Hydra magnipapillata* contained six predicted genes, but the L gene (RDRP) was
294 unusually long. Some of the viruses we detected may well be EVEs inserted into the host
295 genome and subsequently expressed [17]. For example, this is likely the case for the
296 sequence from the silkworm *Bombyx mori* that we also found in the silkworm genome,
297 and the L gene sequence from *Spodoptera exigua* that contains stop codons. Assuming
298 viruses integrated into the host genome once infected those hosts, this does not affect
299 our conclusions below about the host range of these viruses [14-16]. We also found
300 nine other novel mononegavirale-like sequences that group with recently discovered
301 clades of insect viruses [12] (see Table S5 and
302 http://dx.doi.org/10.6084/m9.figshare.1425083).
303
304 *Rhabdovirus Phylogeny*
305
306 To reconstruct the evolution of the *Rhabdoviridae* we have produced the most complete
307 phylogeny of the group to date (Figure 2). We used an alignment of the relatively
308 conserved L gene (RNA Dependant RNA Polymerase) from our newly discovered viruses
309 with sequences of known rhabdoviruses to give an alignment of 195 rhabdoviruses (and
310 26 other mononegavirales as an outgroup). We reconstructed the phylogeny using
311 different sequence alignments and methodologies, and these all gave qualitatively
312 similar results with the same major clades being reconstructed (Gblocks:
313 http://dx.doi.org/10.6084/m9.figshare.1425083, TrimAl:
314 http://dx.doi.org/10.6084/m9.figshare.1425082 and BEAST:
315 http://dx.doi.org/10.6084/m9.figshare.1425436). The branching order between the
316 clades in the dimarhabdovirus supergroup was generally poorly supported and differed
317 between the methods. Eight sequences that we discovered were not included in this
318 analysis as they were considered too short, but their closest BLAST hits are listed in
319 Table S4.
320
321 We recovered all of the major clades described previously (Figure 2), and found that the
322 majority of known rhabdoviruses belong to the dimarhabdovirus clade (Figure 2b). The
323 RNA-seq viruses from *Drosophila* fall into either the sigma virus clade (Figure 2b) or the
324 arthropod clade sister to the cyto- and nucleo- rhabdoviruses (Figure 2a). The viruses
325 from sequence databases are diverse, coming from almost all of the major clades with
326 the exception of the lyssaviruses.
327
328 [Figure 2 here]
329
330 *Changes in host species*
331

332 With a few exceptions, rhabdoviruses are either arthropod-vectored pathogens of plants
333 or vertebrates, or are vertebrate- or arthropod- specific. In many cases the only
334 information about a virus is the host from which it was isolated. Therefore, it is not clear
335 whether viruses isolated from vertebrates are vertebrate-specific or arthropod-
336 vectored, or whether viruses isolated from biting arthropods (e.g. mosquitoes, sandflies,
337 ticks, midges and sea lice) are arthropod specific or also infect vertebrates. Likewise, it
338 is not clear whether viruses isolated from sap-sucking insects (all Hemiptera: aphids,
339 leafhoppers, scale insect and mealybugs) are arthropod-specific or arthropod-vectored
340 plant viruses. However, in the absence of these data, we used the phylogenetic
341 relationships of the viruses to predict both the ancestral and present host associations
342 (http://dx.doi.org/10.6084/m9.figshare.1425436).
343
344 This approach identified a large number of viruses that are likely to be new arthropod-
345 vectored vertebrate pathogens (Figure 2b). 87 of 92 viruses with ambiguous host
346 associations were assigned a host association with strong posterior support (>0.95). Of
347 the 52 viruses found in biting arthropods, 45 were predicted to be arthropod-vectored
348 vertebrate viruses, and 6 to be arthropod-specific. Of the 33 viruses found in
349 vertebrates, 31 were predicted to be arthropod-vectored vertebrate viruses, with none
350 being vertebrate-specific. Of the 7 viruses found in plant-sap-feeding arthropods (Figure
351 2a), 3 were predicted to be plant-associated and 2 arthropod-associated.
352
353 We were also able to infer the ancestral host association of 182 of 188 of the internal
354 nodes on the phylogenetic tree with strong support (>0.95). In addition to the switches
355 of host-type that occur deep in the tree, there are a small number of changes on the
356 terminal branches of the phylogeny. These could either represent errors in the host
357 assignment (e.g. cross-species contamination), or recent host shifts.
358
359
360 A striking pattern that emerged from our reconstructions of host associations is that
361 switches between major groups of hosts have occurred rarely during the evolution of
362 the rhabdoviruses, excluding a few rare transitions on terminal branches (Figure 2).
363 There has been one clear switch to become insect-vectored plant pathogens (cyto- and
364 nucleo- rhabdoviruses). A single virus isolated from the hop plant *Humulus lupulus* sits
365 in the dimarhabdovirus clade, but this may be because the plant was contaminated with
366 insect matter, as the same RNA-seq dataset contains *COI* sequences with high similarity
367 to thrips. A single transition to being vertebrate–specific has occurred in the
368 lyssaviruses clade [3]. There has also been a single transition to vertebrate viruses that
369 are vectored by arthropods in the dimarhabodovirus clade.
370
371 There are two main clades of arthropod-specific viruses. The first clade is a sister group
372 to the large plant virus clade. This novel group of largely insect-associated viruses are
373 associated with a broad range of insects, including flies, butterflies, moths, ants, thrips,
374 bedbugs, fleas, mosquitoes, water striders and leafhoppers. The mode of transmission
375 and biology of these viruses is yet to be examined. The second clade of insect associated
376 viruses is the sigma virus clade [10, 11, 18, 34]. These are derived from vector-borne
377 dimarhabdoviruses that have lost their vertebrate host and become vertically
378 transmitted pathogens of insects [10]. They are common in Drosophilidae, and our

379 results suggest that they may be widespread throughout the Diptera, with occurrences
380 in the Tephritid fruit fly *Ceratitis capitata*, the stable fly *Muscina stabulans,* several
381 divergent viruses in the housefly *Musca domestica* and louse flies removed from the skin
382 of bats. All of the sigma viruses characterised to date have been vertically transmitted
383 [10], but some of the recently described virus may be transmitted horizontally – it has
384 been speculated that the viruses from louse flies may infect bats [37] and Shayang Fly
385 Virus 2 has been reported in two fly species [12] (although contamination could also
386 explain this result). For the first time we have found sigma-like viruses outside of the
387 Diptera, with two Lepidoptera associated viruses and a virus from an aphid/parasitoid
388 wasp. Drosophila sigma virus genomes are characterised by an additional X gene
389 between the P and M genes [34]. Interestingly the two louse fly viruses with complete
390 genomes, Wuhan insect virus 7 from an aphid/parasitoid and Pararge aegeria
391 rhabdovirus do not have an X gene. Overall, our results suggest sigma-like viruses may
392 be common in a wide array of insect species.
393
394 The rhabdoviruses cluster on the phylogeny not only by the hosts they infect, but also by
395 whether they are found in terrestrial or aquatic environments. There has been one shift
396 from terrestrial to aquatic hosts during the evolution of the basal novirhabdoviruses,
397 which have a wide host range in fish. There have been other terrestrial to aquatic shifts
398 in the dimarhabdoviruses: in the clade of fish and cetacean viruses and the clade of
399 viruses isolated from sea-lice. The sea-lice viruses may be crustacean-specific as the two
400 viruses from *Lepeophtheirus salmonis* do not seem to infect the fish they parasitise and
401 are present in all developmental stages of the lice suggesting they may be transmitted
402 vertically [38].
403
404
405 **Discussion**
406
407 Viruses are ubiquitous in nature and recent developments in high-throughput
408 sequencing technology have led to the discovery and sequencing of a large number of
409 novel viruses in arthropods [12, 13]. Here we have identified 43 novel virus-like
410 sequences, from our own RNA-seq data and public sequence repositories. Of these, 32
411 were rhabdoviruses, and 26 of these were isolated from arthropods. Using these
412 sequences we have produced the most extensive phylogeny of the *Rhabdoviridae* to
413 date, including a total of 195 virus sequences.
414
415 In most cases we know nothing about the biology of the viruses beyond the host they
416 were isolated from, but our analysis provides a powerful way to predict which are
417 vector-borne pathogens and which are specific to vertebrates or arthropods. We have
418 identified a large number of new likely vector-borne pathogens – of 85 rhabdoviruses
419 isolated from vertebrates or biting insects we predict that 76 are arthropod-borne
420 viruses of vertebrates (arboviruses). Along with the known arboviruses, this suggests
421 the majority of known rhabdoviruses are arboviruses, and all of these fall in a single
422 clade known as the dimarhabdoviruses. In addition to the arboviruses, we also
423 identified two clades of likely insect-specific viruses associated with a wide range of
424 species, suggesting rhabdoviruses may be common arthropod pathogens.
425

426 We found that shifts between distantly related hosts are rare in the rhabdoviruses,
427 which is perhaps unsurprising as both rhabdoviruses of vertebrates (rabies virus in
428 bats) and invertebrates (sigma viruses in Drosophildae) show a declining ability to
429 infect hosts more distantly related to their natural host [39-41]. It is thought that sigma
430 viruses may sometimes jump into distantly related but highly susceptible species [40,
431 42, 43], but our results suggest that this rarely happens between major groups such as
432 vertebrates and arthropods. It is nonetheless surprising that arthropod-specific viruses
433 have arisen rarely, as one might naively assume that there would be fewer constraints
434 on vector-borne viruses losing one of their hosts. Within the major clades, closely
435 related viruses often infect closely related hosts (Figure 2). For example, within the
436 dimarhabdoviruses viruses isolated from mosquitoes, ticks, *Drosophila*, Muscid flies,
437 Lepidoptera and sea-lice all tend to cluster together (Figure 2B). However, it is also
438 clear that the virus phylogeny does not mirror the host phylogeny, suggesting that
439 following major transitions between distantly related host taxa, viruses preferentially
440 shift between more closely related species.
441
442 There has been a near four-fold increase in the number of rhabdovirus sequences in the
443 last five years. In part this may be due to the falling cost of sequencing transcriptomes
444 [44], and initiatives to sequence large numbers of insect and other arthropods [45]. The
445 use of high-throughput sequencing technologies should reduce the likelihood of
446 sampling biases associated with PCR based discovery where people look for similar
447 viruses in related hosts. This suggests that the pattern of viruses forming clades based
448 on the host taxa they infect is likely to be robust. However, these efforts are
449 disproportionately targeted at arthropods, and it is possible that there may be a great
450 undiscovered diversity of viruses in other organisms.
451
452 Rhabdoviruses infect a diverse assortment of host species, including a large number of
453 arthropod species. Our limited search has unearthed a large number of novel
454 rhabdovirus genomes, suggesting we are only just beginning to uncover the diversity of
455 these viruses.
456
457 **Data accessibility**
458
459 **All data has been made available in public repositories:**
460
461 NCBI Sequence Read Archive Data: SRP057824
462 Data S1, sample information: http://dx.doi.org/10.6084/m9.figshare.1425432
463 Data S2, virus ID, Genbank accessions and host information:
464 http://dx.doi.org/10.6084/m9.figshare.1425419
465 L gene sequences fasta: http://dx.doi.org/10.6084/m9.figshare.1425067
466 TrimAl alignment fasta: http://dx.doi.org/10.6084/m9.figshare.1425069
467 Gblocks alignment fasta: http://dx.doi.org/10.6084/m9.figshare.1425068
468 Phylogenetic tree Gblocks alignment: http://dx.doi.org/10.6084/m9.figshare.1425083
469 Phylogenetic tree TrimAl alignment: http://dx.doi.org/10.6084/m9.figshare.1425082
470 BEAST alignment fasta: http://dx.doi.org/10.6084/m9.figshare.1425431
471 BEAUti xml file: http://dx.doi.org/10.6084/m9.figshare.1431922
472 Bayesian analysis tree: http://dx.doi.org/10.6084/m9.figshare.1425436
473

Supplementary materials: Tables S1-5 List of newly discovered viruses.

Competing interests

We have no competing interests.

Acknowledgments

Many thanks to Mike Ritchie for providing the DMonSV infected fly line; Casper Breuker and Melanie Gibbs for PAegRV samples and Philip Leftwich for CCapSV samples. Thanks to everyone who provided fly collections.

Contributions

BL and FMJ conceived and designed the study. BL and JD carried out molecular work. BL, WJP, DJP and DJO carried out bioinformatic analysis. BL, GGRM and JJW carried out phylogenetic analysis. BL and FMJ wrote the manuscript with comments from all other authors. All authors gave final approval for publication.

Funding

BL and FMJ are supported by a NERC grant (NE/L004232/1), a European Research Council grant (281668, DrosophilaInfection), a Junior Research Fellowship from Christ's College, Cambridge (BL). GGRM is supported by an MRC studentship. The metagenomic sequencing of viruses from *D. immigrans, D. tristis* and *S. deflexa* was supported by a Wellcome Trust fellowship (WT085064) to DJO.

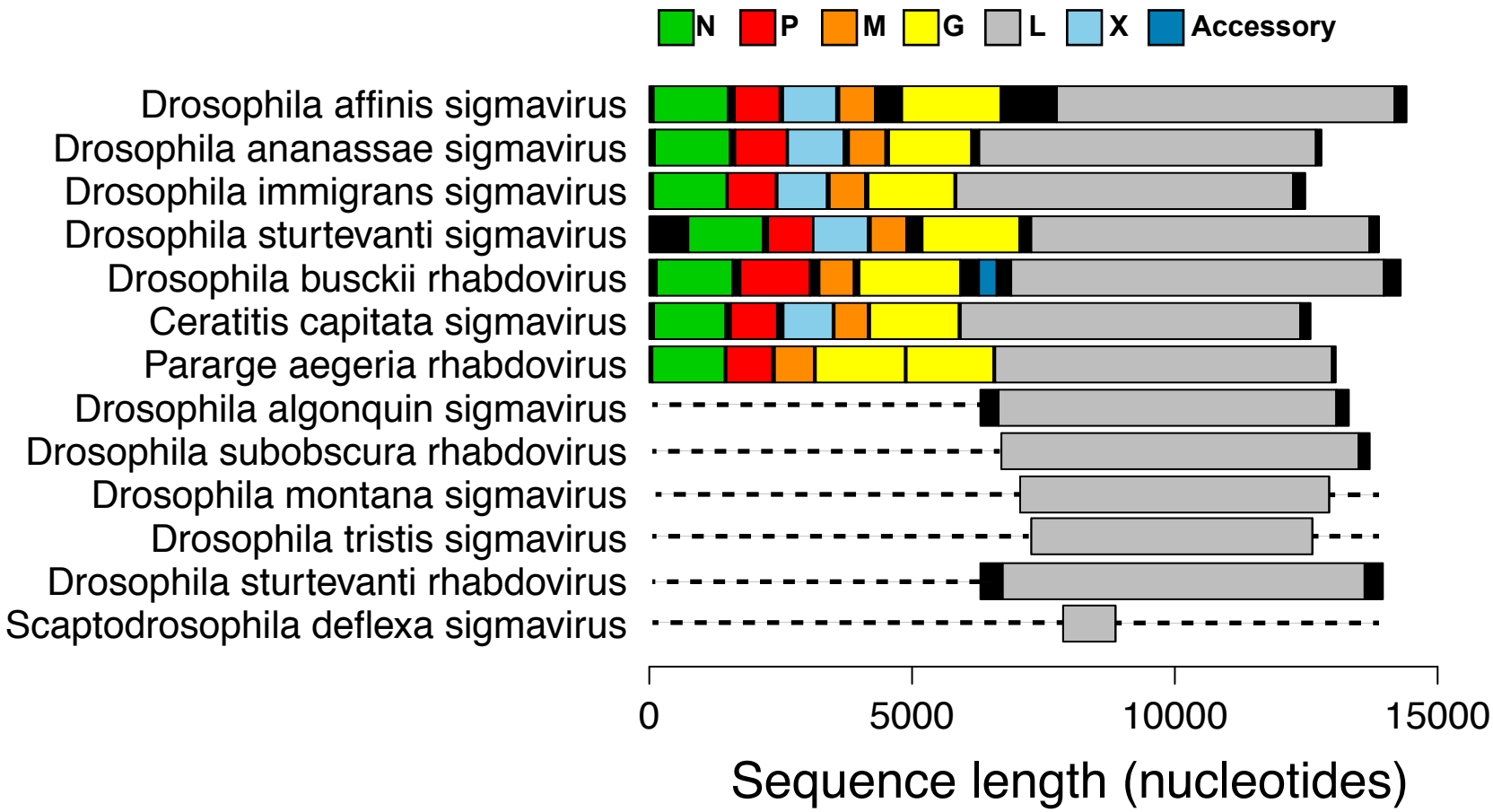
References

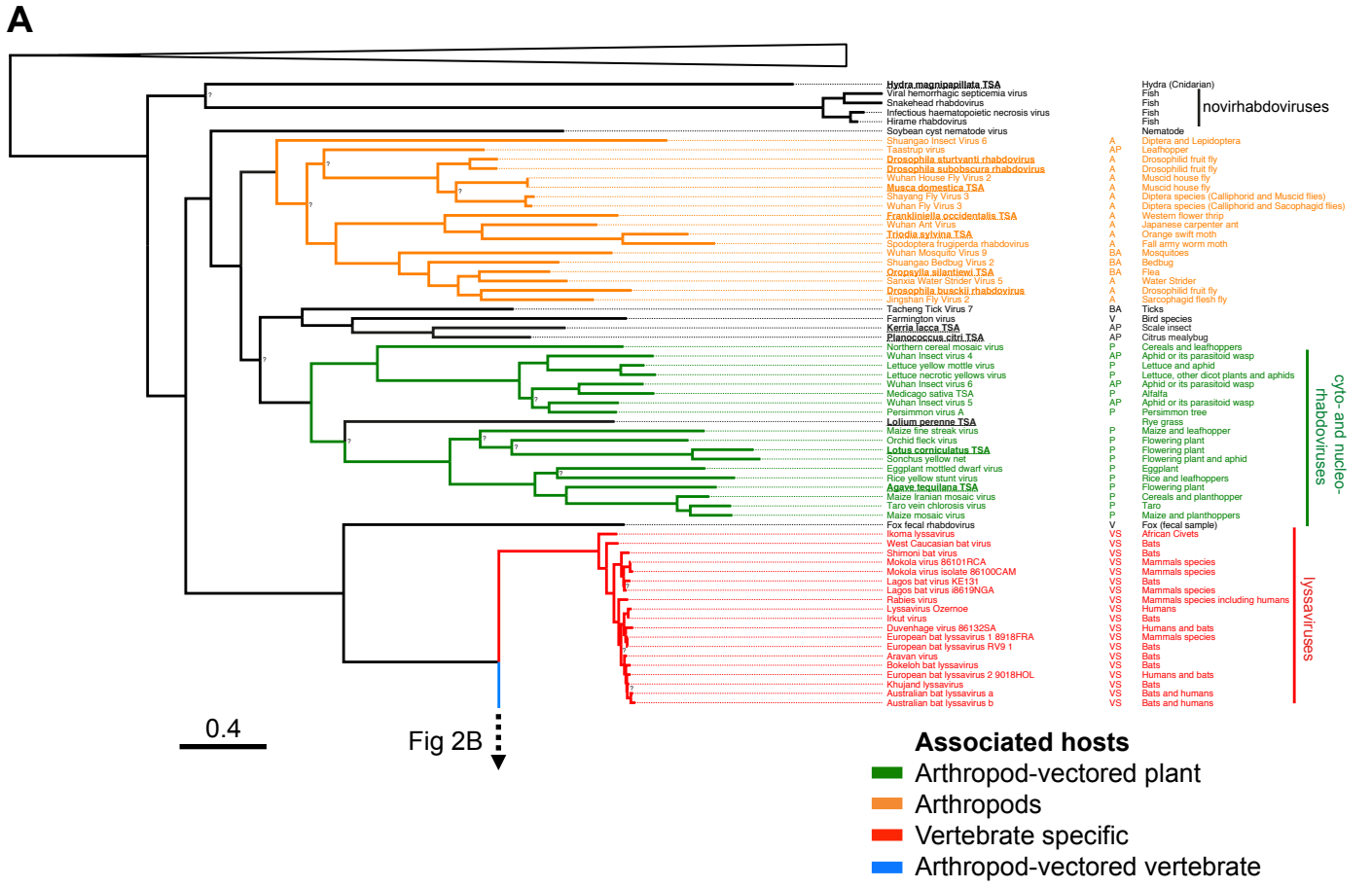
1. Lipkin W.I., Anthony S.J. 2015 Virus hunting. *Virology* **479-480C**, 194-199. (doi:10.1016/j.virol.2015.02.006).
2. Liu S., Vijayendran D., Bonning B.C. 2011 Next generation sequencing technologies for insect virus discovery. *Viruses* **3**(10), 1849-1869. (doi:10.3390/v3101849).
3. Dietzgen R.G., Kuzmin I.V. 2012 *Rhabdoviruses: Molecular Taxonomy, Evolution, Genomics, Ecology, Host-Vector Interactions, Cytopathology and Control*. Norfolk, UK, Caister Academic Press.
4. Bourhy H., Cowley J.A., Larrous F., Holmes E.C., Walker P.J. 2005 Phylogenetic relationships among rhabdoviruses inferred using the L polymerase gene. *Journal of General Virology* **86**, 2849-2858.
5. Hampson K., Coudeville L., Lembo T., Sambo M., Kieffer A., Attlan M., Barrat J., Blanton J.D., Briggs D.J., Cleaveland S., et al. 2015 Estimating the global burden of endemic canine rabies. *PLoS Negl Trop Dis* **9**(4), e0003709. (doi:10.1371/journal.pntd.0003709).
6. Walker P.J., Blasdell K.R., Joubert D.A. 2012 Ephemeroviruses: Athropod-borne Rhabdoviruses of Ruminants, with Large, Complex Genomes. In *Rhabdoviruses: Molecular Taxonomy, Evolution, Genomics, Ecology, Host-Vector Interactions, Cytopathology and Control* (eds. Dietzgen R.G., Kuzmin I.V.), pp. 59-88. Norfolk, UK, Caister Academic Press.

523 7. Hogenhout S.A., Redinbaugh M.G., Ammar E.D. 2003 Plant and animal
524 rhabdovirus host range: a bug's view. *Trends in Microbiology* **11**(6), 264-271. (doi:Doi
525 10.1016/S0966-842x(03)00120-3).
526 8. Ahne W. 1985 Argulus foliaceus L. and Piscicola geometra L. as mechanical
527 vectors of spring viraemia of carp virus (SVCV). *Journal of Fish Diseases* **8**(2), 241-242.
528 9. Longdon B., Jiggins F.M. 2012 Vertically transmitted viral endosymbionts of
529 insects: do sigma viruses walk alone? *Proc Biol Sci* **279**(1744), 3889-3898.
530 (doi:10.1098/rspb.2012.1208).
531 10. Longdon B., Wilfert L., Obbard D.J., Jiggins F.M. 2011 Rhabdoviruses in two
532 species of Drosophila: vertical transmission and a recent sweep. *Genetics* **188**(1), 141-
533 150. (doi:10.1534/genetics.111.127696).
534 11. Longdon B., Wilfert L., Osei-Poku J., Cagney H., Obbard D.J., Jiggins F.M. 2011
535 Host switching by a vertically-transmitted rhabdovirus in Drosophila. *Biology Letters*
536 **7**(5), 747-750. (doi:10.1098/rsbl.2011.0160).
537 12. Li C.X., Shi M., Tian J.H., Lin X.D., Kang Y.J., Chen L.J., Qin X.C., Xu J., Holmes E.C.,
538 Zhang Y.Z. 2015 Unprecedented genomic diversity of RNA viruses in arthropods reveals
539 the ancestry of negative-sense RNA viruses. *eLife* **4**. (doi:10.7554/eLife.05378).
540 13. Walker P.J., Firth C., Widen S.G., Blasdell K.R., Guzman H., Wood T.G., Paradkar
541 P.N., Holmes E.C., Tesh R.B., Vasilakis N. 2015 Evolution of genome size and complexity
542 in the rhabdoviridae. *PLoS Pathog* **11**(2), e1004664.
543 (doi:10.1371/journal.ppat.1004664).
544 14. Ballinger M.J., Bruenn J.A., Taylor D.J. 2012 Phylogeny, integration and
545 expression of sigma virus-like genes in Drosophila. *Mol Phylogenet Evol* **65**(1), 251-258.
546 (doi:10.1016/j.ympev.2012.06.008).
547 15. Fort P., Albertini A., Van-Hua A., Berthomieu A., Roche S., Delsuc F., Pasteur N.,
548 Capy P., Gaudin Y., Weill M. 2011 Fossil Rhabdoviral Sequences Integrated into
549 Arthropod Genomes: Ontogeny, Evolution, and Potential Functionality. *Mol Biol Evol*.
550 (doi:10.1093/molbev/msr226).
551 16. Katzourakis A., Gifford R.J. 2010 Endogenous Viral Elements in Animal Genomes.
552 *Plos Genet* **6**(11), e1001191. (doi:10.1371/journal.pgen.1001191).
553 17. Aiewsakun P., Katzourakis A. 2015 Endogenous viruses: Connecting recent and
554 ancient viral evolution. *Virology* **479-480C**, 26-37. (doi:10.1016/j.virol.2015.02.011).
555 18. Longdon B., Wilfert L., Jiggins F.M. 2012 *The Sigma Viruses of Drosophila*, Caister
556 Academic Press.
557 19. Rosen L. 1980 Carbon-dioxide sensitivity in mosquitos infected with sigma,
558 vesicular stomatitis, and other rhabdoviruses. *Science* **207**(4434), 989-991.
559 20. Shroyer D.A., Rosen L. 1983 Extrachromosomal-inheritance of carbon-dioxide
560 sensitivity in the mosquito culex-quinquefasciatus. *Genetics* **104**(4), 649-659.
561 21. Parker D.J., Vesala L., Ritchie M.G., Laiho A., Hoikkala A., Kankare M. 2015 How
562 consistent are the transcriptome changes associated with cold acclimation in two
563 species of the Drosophila virilis group? *Heredity (Edinb)*. (doi:10.1038/hdy.2015.6).
564 22. van Mierlo J.T., Overheul G.J., Obadia B., van Cleef K.W., Webster C.L., Saleh M.C.,
565 Obbard D.J., van Rij R.P. 2014 Novel Drosophila viruses encode host-specific suppressors
566 of RNAi. *PLoS Pathog* **10**(7), e1004256. (doi:10.1371/journal.ppat.1004256).
567 23. Katoh K., Standley D.M. 2013 MAFFT multiple sequence alignment software
568 version 7: improvements in performance and usability. *Mol Biol Evol* **30**(4), 772-780.
569 (doi:10.1093/molbev/mst010).
570 24. Capella-Gutierrez S., Silla-Martinez J.M., Gabaldon T. 2009 trimAl: a tool for
571 automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*
572 **25**(15), 1972-1973. (doi:10.1093/bioinformatics/btp348).
573 25. Talavera G., Castresana J. 2007 Improvement of phylogenies after removing
574 divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*
575 **56**(4), 564-577. (doi:Doi 10.1080/10635150701472164).

576 26. Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., Gascuel O. 2010
577 New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing
578 the Performance of PhyML 3.0. *Syst Biol* **59**(3), 307-321. (doi:Doi
579 10.1093/Sysbio/Syq010).
580 27. Le S.Q., Gascuel O. 2008 An improved general amino acid replacement matrix.
581 *Molecular Biology and Evolution* **25**(7), 1307-1320. (doi:Doi 10.1093/Molbev/Msn067).
582 28. Anisimova M., Gascuel O. 2006 Approximate likelihood-ratio test for branches: A
583 fast, accurate, and powerful alternative. *Syst Biol* **55**(4), 539-552.
584 (doi:10.1080/10635150600755453).
585 29. Rambaut A. 2011 FigTree. (v1.3 ed.
586 30. Drummond A.J., Suchard M.A., Xie D., Rambaut A. 2012 Bayesian phylogenetics
587 with BEAUti and the BEAST 1.7. *Mol Biol Evol* **29**(8), 1969-1973.
588 (doi:10.1093/molbev/mss075).
589 31. Weinert L.A., Welch J.J., Suchard M.A., Lemey P., Rambaut A., Fitzgerald J.R. 2012
590 Molecular dating of human-to-bovid host jumps by Staphylococcus aureus reveals an
591 association with the spread of domestication. *Biol Lett* **8**(5), 829-832.
592 (doi:10.1098/rsbl.2012.0290).
593 32. Drummond A.J., Ho S.Y., Phillips M.J., Rambaut A. 2006 Relaxed phylogenetics
594 and dating with confidence. *PLoS Biol* **4**(5), e88. (doi:10.1371/journal.pbio.0040088).
595 33. Rambaut A., Drummond A.J. 2007. *Tracer v16, Available from*
596 http://beastbioedacuk/Tracer
597 34. Longdon B., Obbard D.J., Jiggins F.M. 2010 Sigma viruses from three species of
598 Drosophila form a major new clade in the rhabdovirus phylogeny. *Proceedings of the*
599 *Royal Society B* **277**, 35-44. (doi:10.1098/rspb.2009.1472).
600 35. Longdon B., Walker P.J. 2011 ICTV sigmavirus species and genus proposal. (
601 36. Chiba S., Kondo H., Tani A., Saisho D., Sakamoto W., Kanematsu S., Suzuki N. 2011
602 Widespread Endogenization of Genome Sequences of Non-Retroviral RNA Viruses into
603 Plant Genomes. *Plos Pathogens* **7**(7). (doi:Artn E1002146
604 Doi 10.1371/Journal.Ppat.1002146).
605 37. Aznar-Lopez C., Vazquez-Moron S., Marston D.A., Juste J., Ibanez C., Berciano J.M.,
606 Salsamendi E., Aihartza J., Banyard A.C., McElhinney L., et al. 2013 Detection of
607 rhabdovirus viral RNA in oropharyngeal swabs and ectoparasites of Spanish bats. *J Gen*
608 *Virol* **94**(Pt 1), 69-75. (doi:10.1099/vir.0.046490-0).
609 38. Okland A.L., Nylund A., Overgard A.C., Blindheim S., Watanabe K., Grotmol S.,
610 Arnesen C.E., Plarre H. 2014 Genomic characterization and phylogenetic position of two
611 new species in Rhabdoviridae infecting the parasitic copepod, salmon louse
612 (Lepeophtheirus salmonis). *Plos One* **9**(11), e112517.
613 (doi:10.1371/journal.pone.0112517).
614 39. Faria N.R., Suchard M.A., Rambaut A., Streicker D.G., Lemey P. 2013
615 Simultaneously reconstructing viral cross-species transmission history and identifying
616 the underlying constraints. *Philos Trans R Soc Lond B Biol Sci* **368**(1614), 20120196.
617 (doi:10.1098/rstb.2012.0196).
618 40. Longdon B., Hadfield J.D., Webster C.L., Obbard D.J., Jiggins F.M. 2011 Host
619 phylogeny determines viral persistence and replication in novel hosts. *PLoS Pathogens*
620 **7**((9)), e1002260. (doi:10.1371/journal.ppat.1002260).
621 41. Streicker D.G., Turmelle A.S., Vonhof M.J., Kuzmin I.V., McCracken G.F., Rupprecht
622 C.E. 2010 Host Phylogeny Constrains Cross-Species Emergence and Establishment of
623 Rabies Virus in Bats. *Science* **329**(5992), 676-679. (doi:10.1126/science.1188836).
624 42. Longdon B., Brockhurst M.A., Russell C.A., Welch J.J., Jiggins F.M. 2014 The
625 Evolution and Genetics of Virus Host Shifts. *PLoS Pathog* **10**(11), e1004395.
626 (doi:10.1371/journal.ppat.1004395).
627 43. Longdon B., Hadfield J.D., Day J.P., Smith S.C., McGonigle J.E., Cogni R., Cao C.,
628 Jiggins F.M. 2015 The Causes and Consequences of Changes in Virulence following

629 Pathogen Host Shifts. *PLoS Pathog* **11**(3), e1004728.
630 (doi:10.1371/journal.ppat.1004728).
631 44. Wang Z., Gerstein M., Snyder M. 2009 RNA-Seq: a revolutionary tool for
632 transcriptomics. *Nat Rev Genet* **10**(1), 57-63. (doi:10.1038/nrg2484).
633 45. Misof B., Liu S., Meusemann K., Peters R.S., Donath A., Mayer C., Frandsen P.B.,
634 Ware J., Flouri T., Beutel R.G., et al. 2014 Phylogenomics resolves the timing and pattern
635 of insect evolution. *Science* **346**(6210), 763-767. (doi:10.1126/science.1257570).
636 46. Walker P.J., Dietzgen R.G., Joubert D.A., Blasdell K.R. 2011 Rhabdovirus accessory
637 genes. *Virus Res* **162**(1-2), 110-125. (doi:10.1016/j.virusres.2011.09.004).
638
639
640 **Figure legends**
641
642 **Figure 1.** Genome organization of newly discovered viruses. Predicted ORFs are shown
643 in colour, non-coding regions are shown in black. ORFs were designated as the first start
644 codon following the transcription termination sequence (7 U's) of the previous ORF to
645 the first stop codon. Dotted lines represent parts of the genome not sequenced. These
646 viruses were either from our own RNA-seq data, or were first found in in public
647 databases and key features verified by PCR and Sanger sequencing. Rhabdovirus
648 genomes are typically ~12-13kb long and contain five core genes 3'-N-P-M-G-L-5' [3].
649 However, a number of groups of rhabdoviruses contain additional accessory genes [46].
650 Online version in colour.
651
652 **Figure 2.** Maximum likelihood phylogeny of the *Rhabdoviridae.* **A** shows the basal fish-
653 infecting novirhabdoviruses, an unassigned group of arthropod associated viruses, the
654 plant infecting cyto- and nucleo- rhabdoviruses, as well as the vertebrate specific
655 lyssaviruses. **B** shows the dimarhabdovirus supergroup, which is predominantly
656 composed of arthropod-vectored vertebrate viruses, along with the arthropod specific
657 sigma virus clade. Branches are coloured based on the Bayesian host association
658 reconstruction analysis. Black represents taxa omitted from host-state reconstruction or
659 associations with <0.95 support. The tree was inferred from L gene sequences using the
660 Gblocks alignment. The columns of text are the virus name, the host category used for
661 reconstructions, and known hosts (from left to right). Codes for the host categories are:
662 vs= vertebrate-specific, vv= arthropod-vectored vertebrate, a= arthropod specific, ba =
663 biting-arthropod (ambiguous state), v = vertebrate (ambiguous state) and ap=plant-sap-
664 feeding-arthropod (ambiguous state). Names in bold and underlined are viruses
665 discovered in this study. The tree is rooted with the Chuvirus clade (root collapsed) as
666 identified in [12]. Nodes labelled with question marks represent nodes with aLRT
667 (approximate likelihood ratio test) values less than 0.75. Scale bar shows number of
668 amino-acid substitutions per site. Online version in colour.
669


N
P
M
G
L
X
Accessory
Drosophila affinis sigmavirus
Drosophila ananassae sigmavirus
Drosophila immigrans sigmavirus
Drosophila sturtevanti sigmavirus
Drosophila busckii rhabdovirus
Ceratitis capitata sigmavirus
Pararge aegeria rhabdovirus
Drosophila algonquin sigmavirus
Drosophila subobscura rhabdovirus
Drosophila montana sigmavirus
Drosophila tristis sigmavirus
Drosophila sturtevanti rhabdovirus
Scaptodrosophila deflexa sigmavirus
0
5000
10000
15000
Sequence length (nucleotides)


A
Hydra magnipapillata TSA
Viral hemorrhagic septicemia virus
Snakehead rhabdovirus
Infectious haematopoietic necrosis virus
Hirame rhabdovirus
Soybean cyst nematode virus
Shuangao Insect Virus 6
Taastrup virus
Drosophila sturtvanti rhabdovirus
Drosophila subobscura rhabdovirus
Wuhan House Fly Virus 2
Musca domestica TSA
Shuyang Fly Virus 3
Wuhan Fly Virus 3
Frankliniella occidentalis TSA
Wuhan Ant Virus
Triodia sylvina TSA
Spodoptera frugiperda rhabdovirus
Wuhan Mosquito Virus 9
Shuangao Bedbug Virus 2
Oropsylla silantiewi TSA
Sanxia Water Strider Virus 5
Drosophila lugaskii rhabdovirus
Jingshan Fly Virus 2
Tacheng Tick Virus 7
Farmington virus
Kerria lacca TSA
Planococcus citri TSA
Northern cereal mosaic virus
Wuhan Insect virus 4
Lettuce yellow mottle virus
Lettuce necrotic yellows virus
Wuhan Insect virus 6
Medicago sativa TSA
Wuhan Insect virus 5
Persimmon virus A
Lolium perenne TSA
Maize fine streak virus
Orchid fleck virus
Lotus corniculatus TSA
Sonchus yellow net
Eggplant mottled dwarf virus
Rice yellow stunt virus
Agave tequilana TSA
Maize Iranian mosaic virus
Taro vein chlorosis virus
Maize mosaic virus
Fox fecal rhabdovirus
Ikoma lyssavirus
West Caucasian bat virus
Shimoni bat virus
Mokola virus 86101RCA
Mokola virus isolate 86100CAM
Lagos bat virus KE131
Lagos bat virus 8619NGA
Rabies virus
Lyssavirus Ozernoe
Irkut virus
Duvenhage virus 86132SA
European bat lyssavirus 1 8918FRA
European bat lyssavirus RV9 1
Aravan virus
Bokeloh bat lyssavirus
European bat lyssavirus 2 9018HOL
Khujand lyssavirus
Australian bat lyssavirus a
Australian bat lyssavirus b
Hydra (Cnidarian)
Fish
Fish
Fish
Fish
Nematode
A Diptera and Lepidoptera
AP Leafhopper
A Drosophilid fruit fly
A Drosophilid fruit fly
A Muscid house fly
A Muscid house fly
A Diptera species (Calliphorid and Muscid flies)
A Diptera species (Calliphorid and Sacophagid flies)
A Western flower thrip
A Japanese carpenter ant
A Orange swift moth
A Fall army worm moth
BA Mosquitoes
BA Bedbug
BA Flea
A Water Strider
A Drosophilid fruit fly
A Sarcophagid flesh fly
BA Ticks
V Bird species
AP Scale insect
AP Citrus mealybug
P Cereals and leafhoppers
AP Aphid or its parasitoid wasp
P Lettuce and aphid
P Lettuce, other dicot plants and aphids
AP Aphid or its parasitoid wasp
P Alfalfa
AP Aphid or its parasitoid wasp
P Persimmon tree
Rye grass
P Maize and leafhopper
P Flowering plant
P Flowering plant
P Flowering plant and aphid
P Eggplant
P Rice and leafhoppers
P Flowering plant
P Cereals and planthopper
P Taro
P Maize and planthoppers
V Fox (fecal sample)
VS African Civets
VS Bats
VS Bats
VS Mammals species
VS Mammals species
VS Bats
VS Mammals species
VS Mammals species including humans
VS Humans
VS Bats
VS Humans and bats
VS Mammals species
VS Bats
VS Bats
VS Humans and bats
VS Bats
VS Bats and humans
VS Bats and humans
novirhabdoviruses
cyto- and nucleo-rhabdoviruses
lyssaviruses
0.4
Fig 2B
Associated hosts
Arthropod-vectored plant
Arthropods
Vertebrate specific
Arthropod-vectored vertebrate

B

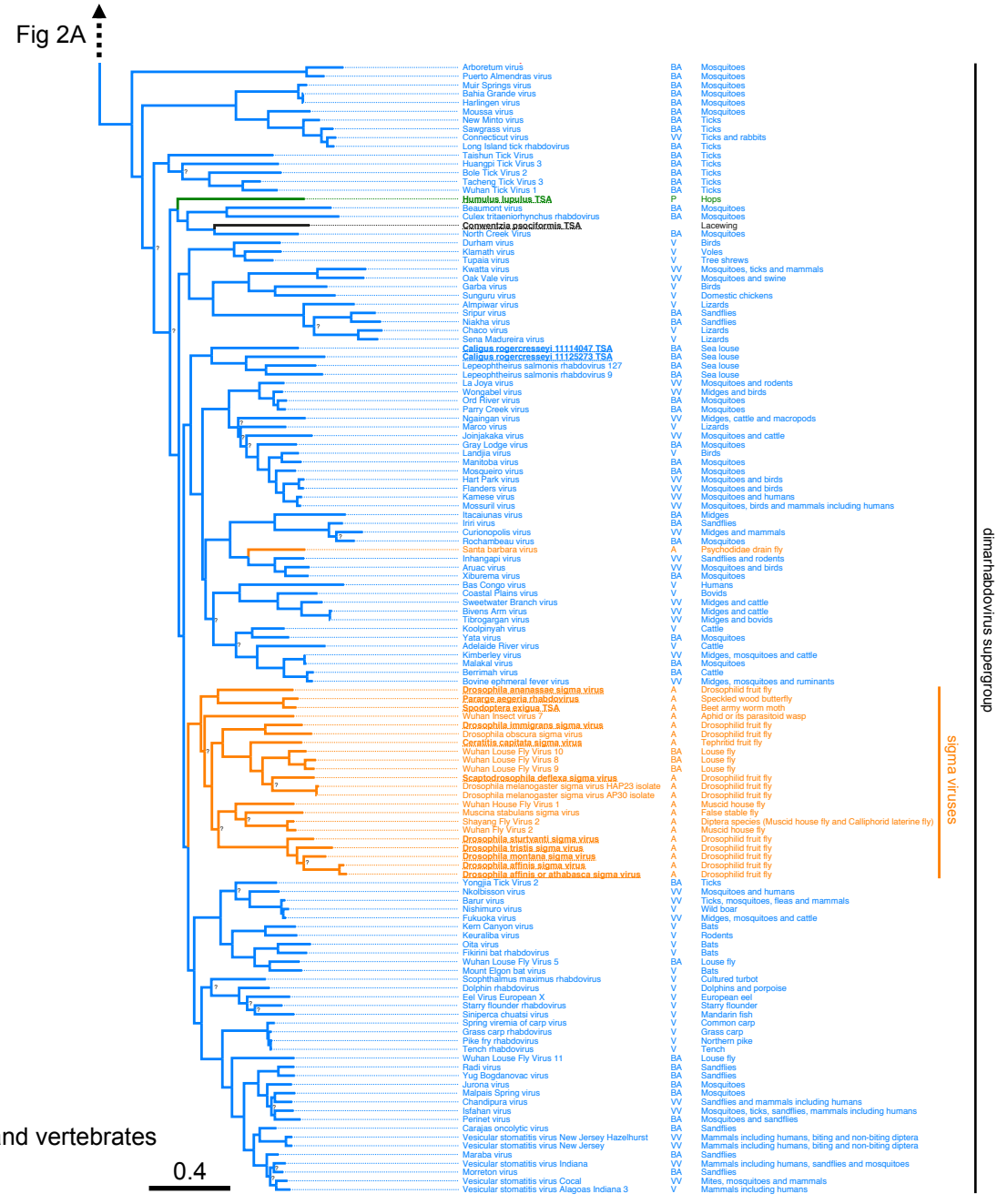


Fig 2A
Arboretum virus BA Mosquitoes
Puerto Almendras virus BA Mosquitoes
Muir Springs virus BA Mosquitoes
Bahia Grande virus BA Mosquitoes
Harlingen virus BA Mosquitoes
Moussa virus BA Mosquitoes
New Minto virus BA Ticks
Sawgrass virus BA Ticks
Connecticut virus VV Ticks and rabbits
Long Island tick rhabdovirus BA Ticks
Taishun Tick Virus BA Ticks
Huangpi Tick Virus 3 BA Ticks
Bole Tick Virus 2 BA Ticks
Tacheng Tick Virus 3 BA Ticks
Wuhan Tick Virus 1 BA Ticks
Humulus lupulus TSA P Hops
Beaumont virus BA Mosquitoes
Culex tritaeniorhynchus rhabdovirus BA Mosquitoes
Conwentzia psociformia TSA Lacewing
North Creek Virus BA Mosquitoes
Durham virus V Birds
Klamath virus V Voles
Tupaia virus V Tree shrews
Kwatta virus VV Mosquitoes, ticks and mammals
Oak Vale virus VV Mosquitoes and swine
Garba virus V Birds
Sunguru virus V Domestic chickens
Almpiwar virus V Lizards
Sripur virus BA Sandflies
Niakha virus BA Sandflies
Chaco virus V Lizards
Sena Madureira virus V Lizards
Caligus rogercresseyi 11114047 TSA BA Sea louse
Caligus rogercresseyi 11125273 TSA BA Sea louse
Lepeophtheirus salmonis rhabdovirus 127 BA Sea louse
Lepeophtheirus salmonis rhabdovirus 9 BA Sea louse
La Joya virus VV Mosquitoes and rodents
Wongabel virus VV Midges and birds
Ord River virus BA Mosquitoes
Parry Creek virus BA Mosquitoes
Ngaingan virus VV Midges, cattle and macropods
Marco virus V Lizards
Joinjakaka virus VV Mosquitoes and cattle
Gray Lodge virus BA Mosquitoes
Landjia virus V Birds
Manitoba virus BA Mosquitoes
Mosqueiro virus BA Mosquitoes
Hart Park virus VV Mosquitoes and birds
Flanders virus VV Mosquitoes and birds
Kamese virus VV Mosquitoes and humans
Mossuril virus VV Mosquitoes, birds and mammals including humans
Itacaiunas virus BA Midges
Iriri virus BA Sandflies
Curionopolis virus VV Midges and mammals
Rochambeau virus BA Mosquitoes
Santa barbara virus A Psychodidae drain fly
Inhangapi virus VV Sandflies and rodents
Aruac virus VV Mosquitoes and birds
Xiburema virus BA Mosquitoes
Bas Congo virus V Humans
Coastal Plains virus V Bovids
Sweetwater Branch virus VV Midges and cattle
Bivens Arm virus VV Midges and cattle
Tibrogargan virus VV Midges and bovids
Koolpinyah virus V Cattle
Yata virus BA Mosquitoes
Adelaide River virus V Cattle
Kimberley virus VV Midges, mosquitoes and cattle
Malakal virus BA Mosquitoes
Berrimah virus BA Cattle
Bovine ephmeral fever virus VV Midges, mosquitoes and ruminants
Drosophila ananassae sigma virus A Drosophilid fruit fly
Parargе aegeria rhabdovirus A Speckled wood butterfly
Spodoptera exigua TSA A Beet army worm moth
Wuhan insect virus 7 A Aphid or its parasitoid wasp
Drosophila immigrans sigma virus A Drosophilid fruit fly
Drosophila obscura sigma virus A Drosophilid fruit fly
Ceratitis capitata sigma virus A Tephritid fruit fly
Wuhan Louse Fly Virus 10 BA Louse fly
Wuhan Louse Fly Virus 8 BA Louse fly
Wuhan Louse Fly Virus 9 BA Louse fly
Scaptodrosophila deflexa sigma virus A Drosophilid fruit fly
Drosophila melanogaster sigma virus HAP23 isolate A Drosophilid fruit fly
Drosophila melanogaster sigma virus AP30 isolate A Drosophilid fruit fly
Wuhan House Fly Virus 1 A Muscid house fly
Muscina stabulans sigma virus A False stable fly
Shayang Fly Virus 2 A Diptera species (Muscid house fly and Calliphorid laterine fly)
Wuhan Fly Virus 2 A Muscid house fly
Drosophila sturtvanti sigma virus A Drosophilid fruit fly
Drosophila tristis sigma virus A Drosophilid fruit fly
Drosophila montana sigma virus A Drosophilid fruit fly
Drosophila affinis sigma virus A Drosophilid fruit fly
Drosophila affinis or athabasca sigma virus A Drosophilid fruit fly
Yongjia Tick Virus 2 BA Ticks
Nkolbisson virus VV Mosquitoes and humans
Barur virus VV Ticks, mosquitoes, fleas and mammals
Nishimuro virus V Wild boar
Fukuoka virus VV Midges, mosquitoes and cattle
Kern Canyon virus V Bats
Keuraliba virus V Rodents
Oita virus V Bats
Fikirini bat rhabdovirus V Bats
Wuhan Louse Fly Virus 5 BA Louse fly
Mount Elgon bat virus V Bats
Scophthalmus maximus rhabdovirus V Cultured turbot
Dolphin rhabdovirus V Dolphins and porpoise
Eel Virus European X V European eel
Starry flounder rhabdovirus V Starry flounder
Siniperca chuatsi virus V Mandarin fish
Spring viremia of carp virus V Common carp
Grass carp rhabdovirus V Grass carp
Pike fry rhabdovirus V Northern pike
Tench rhabdovirus V Tench
Wuhan Louse Fly Virus 11 BA Louse fly
Radi virus BA Sandflies
Yug Bogdanovac virus BA Sandflies
Jurona virus BA Mosquitoes
Malpais Spring virus BA Mosquitoes
Chandipura virus VV Sandflies and mammals including humans
Isfahan virus VV Mosquitoes, ticks, sandflies, mammals including humans
Perinet virus BA Mosquitoes and sandflies
Carajas oncolytic virus BA Sandflies
Vesicular stomatitis virus New Jersey Hazelhurst VV Mammals including humans, biting and non-biting diptera
Vesicular stomatitis virus New Jersey VV Mammals including humans, biting and non-biting diptera
Maraba virus BA Sandflies
Vesicular stomatitis virus Indiana VV Mammals including humans, sandflies and mosquitoes
Morreton virus BA Sandflies
Vesicular stomatitis virus Cocal VV Mites, mosquitoes and mammals
Vesicular stomatitis virus Alagoas Indiana 3 V Mammals including humans
dimarhabdovirus supergroup
sigma viruses
0.4
Associated hosts
Plants
Arthropods
Vertebrate specific
Arthropod-vectors and vertebrates