



26 **Significance statement**

27 Emerging disease epidemics often result from a pathogen establishing transmission in a novel host  
28 species. However, most cross-species transmissions fail to establish in the newly infected species  
29 for reasons that remain poorly understood. Examining cross-species inoculations involving rabies,  
30 a widespread viral zoonosis, we show that mismatches in virulence, which are predictable from  
31 host and viral factors, make sustained transmission in the novel host less likely. In particular,  
32 disease progression was accelerated and virus excretion decreased when the reservoir and novel  
33 host were physiologically or genetically more dissimilar. These mechanistic insights help to  
34 explain and predict host shift events and highlight meta-analyses of existing experimental  
35 inoculation data as a powerful and generalisable approach for understanding the dynamics of index  
36 infections in novel species.

## 37 **Introduction**

38 Cross-species transmission is an important source of emerging and endemic disease. Viruses such  
39 as West Nile virus, rabies virus and Lassa virus cause tens of thousands of human infections  
40 annually through transmission from animal reservoirs (Chancey et al., 2015; Hampson et al., 2015;  
41 Ogbu et al., 2007). Cross-species transmission is also the first step towards host shifts, where  
42 pathogens establish transmission cycles in novel hosts (Wolfe et al., 2007). While the broader-scale  
43 epidemiological and ecological factors driving cross-species transmission are beginning to be  
44 understood (reviewed in Lloyd-Smith et al., 2009; Plowright et al., 2014, 2017), we remain unable  
45 to anticipate whether cross-species transmission will cause ‘dead-end’ infections or lead to  
46 sustained transmission. Infection dynamics at the cross-species interface, specifically the  
47 probability of infection given exposure and the progression of index infections in novel hosts, are  
48 generally unobservable in nature. This is a crucial gap given that the outcomes of cross-species  
49 infections have profound implications for host shifts and disease emergence.

50 Cross-infection studies, in which viruses from a natural reservoir are experimentally inoculated  
51 into novel host species, provide a rare view into the dynamics of index infections. Since the dose,  
52 route, timing and origin of viral exposure are known, these factors can be controlled for to identify  
53 the ecological and evolutionary rules that govern the outcomes of cross-species transmission. We  
54 focus on *Rabies lyssavirus* (family *Rhabdoviridae*) as a model pathogen for understanding cross-  
55 species transmission (Fisher et al., 2018). Rabies virus is a zoonotic RNA virus, transmitted through  
56 an infectious bite, that infects all mammals and, untreated, has the highest case fatality ratio of any  
57 infectious disease (Hemachudha et al., 2002; Rupprecht et al., 2002). Rabies virus naturally infects  
58 multiple carnivore and bat species, which each perpetuate species-specific maintenance cycles  
59 (Mollentze et al., 2014). Although most cross-species transmission events do not lead to onward  
60 transmission, each maintenance cycle represents a rare past cross-species transmission event that  
61 established transmission in a novel host. Dead-end cross-species transmissions and historical host  
62 shifts are detectable in rabies virus phylogenies, and epidemiological surveillance reveals that  
63 nascent hosts shifts remain commonplace (Kuzmin et al., 2012; Mollentze et al., 2014; Streicker et  
64 al., 2010). As such, rabies virus exhibits extensive variation in the epidemiological outcomes of  
65 cross-species transmission. Here, we exploit cross-infection studies conducted over several  
66 decades, in which diverse mammalian species were inoculated with rabies viruses of bat and  
67 carnivore origin, to investigate the individual-level outcomes of index infections.

68 The potential for sustained onward transmission of rabies virus is likely to depend on the incubation  
69 period (from bite to the appearance of clinical signs) and the duration of clinical signs prior to death  
70 (here, the clinical period) of infected hosts. Longer incubation periods are associated with greater  
71 distribution of virus through the central nervous system (Fekadu et al., 1982), spread to a wider  
72 range of tissues (Davis et al., 2013), and higher virus titres in the salivary glands, all of which  
73 should facilitate onward transmission (Baer and Bales, 1967). Conversely, faster progression of  
74 infection has been associated with lower virus excretion, and in extreme cases animals die before  
75 the virus reaches the salivary glands, making transmission highly unlikely (Baer and Bales, 1967;  
76 Charlton et al., 1987; Davis et al., 2013; Fekadu et al., 1982). Further, the clinical period of rabies  
77 coincides with the period of greatest infectivity, when excretion of virus in the saliva often  
78 coincides with clinical signs such as aggression which promote transmission (Hanlon, 2013).  
79 Testing for shifts in incubation and clinical period durations, and in the amount of virus excreted,  
80 allows us to examine what constrains onward transmission of rabies in index hosts following cross-  
81 species transmission.

82 Based on previous work on rabies virus and in other host-pathogen systems, several mechanisms  
83 can be hypothesised to influence infection dynamics and thus the outcome of cross-species  
84 transmission:

- 85 1. Features of exposed host species (*host effects*) influencing the outcome of cross-species  
86 transmissions, irrespective of the infecting virus. For example, larger-bodied species may  
87 be more resistant to infection and thus require either higher infectious doses or a longer  
88 period of virus replication before symptoms become apparent. More generally,  
89 evolutionarily conserved similarities in host physiology mean that groups of related taxa  
90 might have similar susceptibility or clinical outcomes of infection (Longdon et al., 2011).
- 91 2. Features inherent to the virus lineage involved (*virus effects*), irrespective of the infected  
92 host, likely due to adaptation of individual lineages to reservoir host species. For example,  
93 key differences in disease expression between rabies viruses adapted to bats and those  
94 adapted to carnivores have been noted in humans, although it remains unclear whether this  
95 is a feature of the virus or due to differing routes of exposure (Begeman et al., 2018).  
96 Ultimately, the effect of a specific virus and its reservoir host cannot be disentangled.

97 3. *Host-virus interactions.* Both initial cross-species transmission and successful  
98 establishment occur most often between closely related hosts, often referred to as the  
99 phylogenetic distance effect (Gilbert and Webb, 2007; Longdon et al., 2015, 2011; Streicker  
100 et al., 2010). However, the mechanisms underlying this pattern remain obscure. Mammals  
101 also exhibit considerable variability in physiological features that are only moderately  
102 constrained by phylogenetic relatedness, such as body temperature (Clarke and Rothery,  
103 2008). This may create distantly related pairs of reservoir and novel host species which  
104 nevertheless share key physiological features affecting disease outcome (Longdon et al.,  
105 2014), a potential explanation for the occurrence of host shifts over wide phylogenetic  
106 scales (e.g. the bat to primate host shifts involving SARS-coronavirus and Ebola viruses).

107 Here, we test these hypotheses by conducting a meta-analysis of individual-level data from 514  
108 published experimental cross-species infections involving rabies virus. We show that features of  
109 the virus and of the inoculated host species interact with the initial conditions of exposure to  
110 influence the outcome of cross-species transmission in ways expected to affect the likelihood of  
111 onward transmission in the novel host species.

## 112 **Results**

113 Our meta-analysis of cross-species inoculation experiments yielded results from 20 mammal  
114 species (in the orders Carnivora, Chiroptera, Cetartiodactyla and Rodentia), inoculated with 39  
115 unique inocula from 7 reservoir species in the orders Carnivora and Chiroptera (Figure 1 & Table  
116 1). Three outcome measures were recorded – the duration of the incubation period (N=443) and  
117 clinical period (N=175) and the amount of virus excreted (N=278).

### 118 Incubation period

119 The time period between inoculation and the appearance of symptoms was highly variable, with a  
120 median duration of 15 days. While incubation periods ranged between 4 and 141 days, 95% lasted  
121  $\leq 28$  days (all estimates based on a non-parametric Kaplan-Meier fit to the censored event times).

122 We modelled incubation period duration using log-normal generalized linear mixed models  
123 (GLMMs), correcting for phylogenetic non-independence among inoculated species and among  
124 reservoir species, as well as for clustering within experiments (see Methods). As expected,

125 incubation periods were shortened by both increased dose and by inoculation routes which were  
126 closer to the brain, although the effect size estimate for inoculation distance included zero (Figure  
127 2A). More importantly, the duration of incubation periods was also influenced by features of the  
128 virus as well as its interaction with the inoculated host. Specifically, significant differences in  
129 incubation period duration were associated with reservoir type (bat vs. carnivore) and with body  
130 temperature differences between source and inoculated hosts (Figure 2A). Both effects depended  
131 on viral dose. At low doses, viruses from bat reservoirs were associated with shorter incubation  
132 periods compared to viruses from carnivores, though this effect diminished at higher doses (Figure  
133 2B). A similar effect was seen when the inoculated species was a known rabies reservoir (Figure  
134 2C) although the 95% highest posterior density interval (HPD) of this effect size included zero  
135 (Figure 2A). The difference in typical body temperature between the virus reservoir and the  
136 inoculated species had a more marked effect on incubation period durations (Figure 2A). The onset  
137 of symptoms was delayed when the virus was inoculated into species with a warmer body  
138 temperature than its reservoir (negative values in Figure 2D), although this delay reduced with  
139 higher doses. The opposite was also true – in hosts with lower body temperatures than the virus  
140 reservoir, the incubation period tended to be shorter (Figure 2D, Supplementary figure S 1). Models  
141 fitting effects for inoculated and reservoir species body temperature separately allowed us to  
142 explore this temperature effect further. Inoculated species with higher typical body temperatures  
143 tended to have longer incubation periods (Supplementary figure S 2). However, viruses from  
144 reservoirs with higher body temperatures were associated with shorter incubation periods,  
145 suggesting that these viruses had adapted to counteract any losses in efficiency caused by the body  
146 temperature of their reservoir host. Importantly, there was no correlation between phylogenetic  
147 distance and body temperature difference among species (Supplementary figure S 3), indicating  
148 that the observed temperature effects were not explainable by the level of taxonomic relatedness  
149 among species. All fixed effects combined explained 19.2% of the variation in incubation period  
150 durations (HPD: 2.8 – 40.2%).

### 151 Clinical period

152 Once symptoms appeared, the median time to death (the clinical period) was 3 days, ranging from  
153 < 1 day to 8 days. We modelled clinical period duration using log-normal GLMMs, correcting for  
154 phylogenetic clustering among inoculated species and for clustering within experiments. In  
155 contrast to results for the incubation period, phylogenetic distance between reservoir and inoculated

156 host species appeared to affect the duration of clinical periods. Cross-species inoculations between  
157 phylogenetically more distant species were associated with an increased sensitivity to high viral  
158 doses, resulting in shorter clinical periods (Figure 2E). However, the overall effect size estimate  
159 for this interaction included zero (Figure 2A). Bat-associated viruses appeared to have shorter  
160 clinical periods, and – as also observed for incubation periods – this effect depended on dose, but  
161 here it was poorly estimated, with the HPD again including zero (Figure 2A). All other fixed effects  
162 were small, and none could be clearly separated from zero, but combined the fixed effects  
163 explained 40.3% of the variation in clinical period duration (HPD: 12.9 – 64.0%).

#### 164 Virus titre in salivary glands

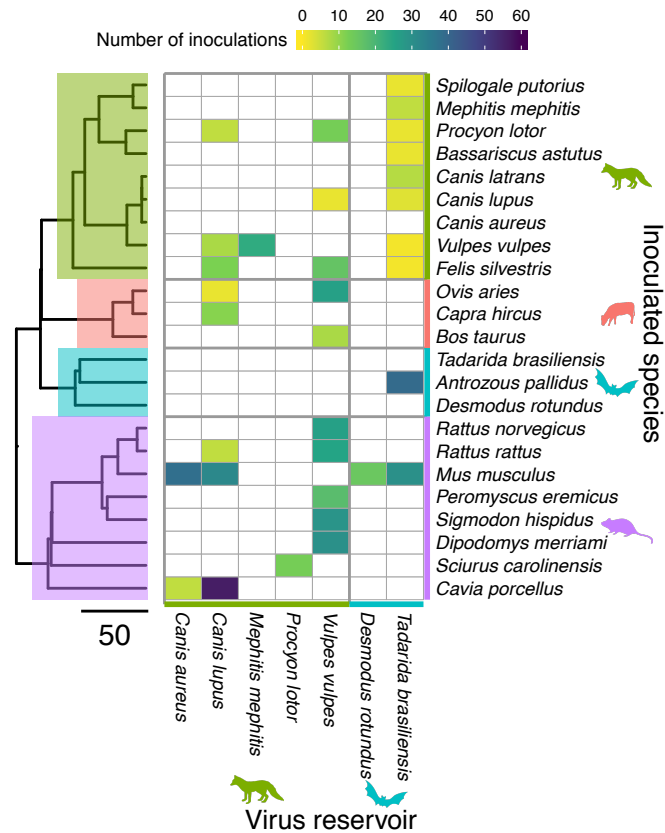
165 Onward transmission of rabies virus, which is mediated by an animal bite, requires presence of the  
166 virus in sufficiently high titres in the salivary glands. To test how the host-virus context of cross-  
167 species transmission affects the amount of virus excreted, we investigated the virus titre detected  
168 in salivary glands *post mortem* as a proxy. To simultaneously investigate potential explanations for  
169 the previously reported correlation between salivary gland virus titre and incubation period  
170 duration (Baer and Bales, 1967; Davis et al., 2013; Fekadu et al., 1982), we modelled the virus titre  
171 excreted jointly with incubation periods using a multi-response log-normal GLMM. When  
172 accounting only for clustering within experiments, salivary gland titres showed a moderate positive  
173 correlation with incubation period duration (Pearson correlation: 0.298, 95% HPD: 0.115 – 0.471;  
174 Figure 3A). Thus, consistent with previous work, animals which experienced longer incubation  
175 periods tended to have more virus in their salivary glands *post mortem*. Part of this correlation is  
176 accounted for by the inoculated species phylogeny (Figure 3B). The remaining residual correlation  
177 is explained by differences in dose (Figure 3C), with higher doses leading to decreased salivary  
178 gland titres (Figure 3D). The inoculation of species with a lower body temperature than the  
179 reservoir also tended to reduce the virus titre in the salivary glands, although the size of this effect  
180 could not be estimated precisely enough to make it distinct from zero (positive body temperature  
181 differences, Figure 3D & E). A clearer effect was observed for the interaction of reservoir status  
182 and dose: at low doses, known rabies reservoir species produced higher virus titres in the salivary  
183 glands than non-reservoirs (Figure 3G). At very high doses, however, we detected no difference in  
184 salivary gland titres, possibly because animals succumb too fast for any differences to develop.

185 Table 1: Summary of data available for each response variable

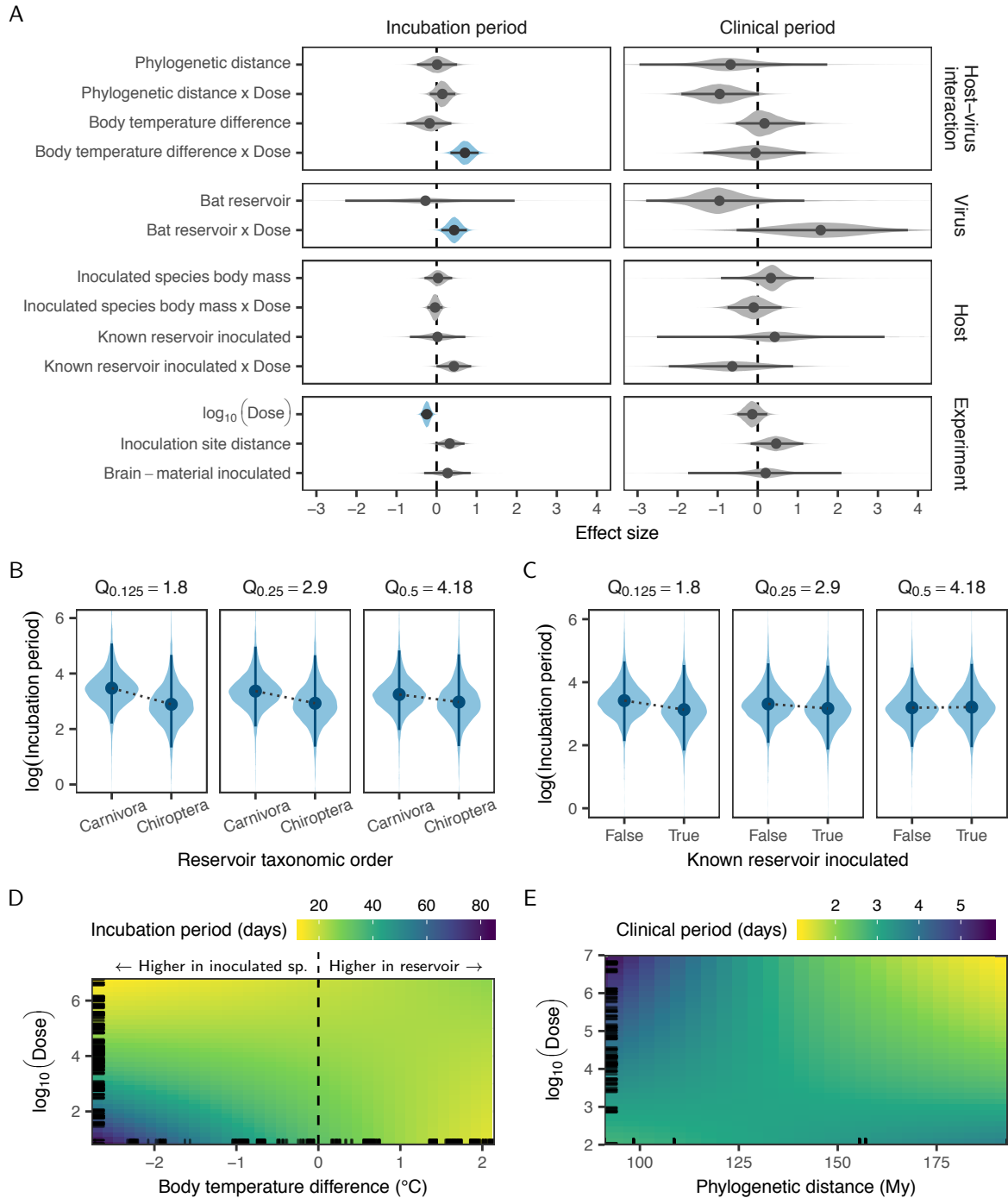
	Incubation period	Clinical period	Presence of virus in salivary glands*
Individual inoculations	443	175	514 (278)
Studies	19	14	23 (16)
Experiments	25	19	30 (20)
Inocula	35	17	39 (22)
Source taxonomic orders / species	2 / 7	2 / 4	2 / 7 (2 / 4)
Inoculated taxonomic orders / species	4 / 19	4 / 13	4 / 20 (4 / 18)
Source-inoculated species combinations	30	19	33 (25)

186 \*Joint fit, dataset includes inoculations providing data for incubation period only. Data for  
187 inoculations specifically providing information on the virus tires in salivary glands given in  
188 brackets





189  
 190 Figure 1: Source-inoculated species combinations for which data were included in this study, and  
 191 the number of animals inoculated. Phylogenetic relationships are demonstrated using a composite  
 192 time-scaled phylogeny, with Carnivora highlighted in green, Artiodactyla in red, Chiroptera in  
 193 blue, and Rodentia in purple. Branch-lengths are in millions of years. The shading of cells indicates  
 194 the number of inoculated animals for which data are available, with white cells indicating  
 195 combinations for which no data were available. Three reservoirs occurred as sources only,  
 196 receiving no inoculations which met our inclusion criteria (*C. aureus*, *T. brasiliensis* and *D.*  
 197 *rotundus*). These species were nevertheless retained on the y-axis to demonstrate phylogenetic  
 198 relationships.

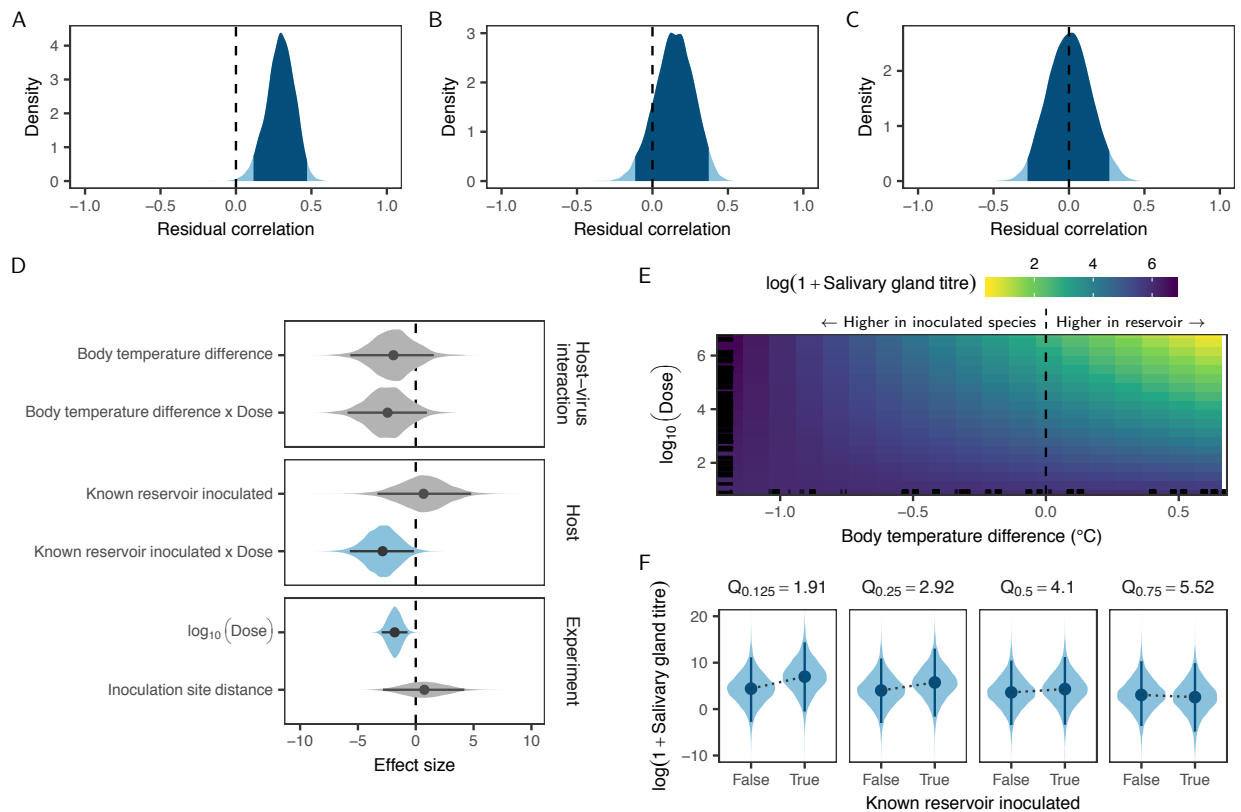


199

200 Figure 2: Factors affecting the progression of disease following cross-species inoculations. A:  
 201 Coefficient estimates from independent models fit to the duration of incubation and clinical  
 202 periods. In addition to the variables shown, both models contained corrections for non-  
 203 independence between observations from the same experiment and for inoculated species  
 204 phylogeny. The regression on incubation period durations additionally contained corrections for

205 reservoir phylogeny. Lines indicate the extent of the 95% highest posterior density (HPD), while  
 206 points indicate the posterior median. Shaded areas show the posterior distribution, with colour used  
 207 to indicate estimates whose 95% HPD excludes zero. B–D: Predicted incubation period durations  
 208 when varying dose along with reservoir taxonomic order (B), whether or not the inoculated species  
 209 was a known reservoir of rabies virus (C), or the difference in typical body temperatures between  
 210 the virus reservoir and the inoculated species (specifically reservoir temperature minus inoculated  
 211 species temperature, D). For each set of predictions, all other explanatory variables in the model  
 212 were held constant at their median observed value. In B & C, each sub-panel shows the predicted  
 213 effect at different quantiles (Q) of observed doses (in  $\log_{10}$  mouse LD50), indicated above the  
 214 panel. E: Predicted duration of clinical periods as a function of phylogenetic distance and dose.  
 215 Panels D-E show the posterior median of predictions.

216



217  
 218 Figure 3: Factors predicting the titre of virus in salivary glands following cross-species inoculation.  
 219 A: Correlation between virus titre and the duration of incubation periods, after accounting for  
 220 clustering within experiments. Shown is the estimated posterior distribution from a multi-response  
 221 regression fitted using the mcmcGLMM library in R, with the darker shaded area indicating the

222 extent of the 95% highest posterior density (HPD). B-C: The remaining correlation is reduced when  
223 also accounting for correlation of virus titre and incubation period duration within the inoculated  
224 species phylogeny (B), and further reduced to approximately 0 when additionally accounting for  
225 differences in inoculated dose (C). D: Coefficient estimates for the regression on log virus titres in  
226 the salivary glands, when accounting for clustering within experiments and the inoculated species  
227 phylogeny. Lines indicate the extent of the 95% HPD, while points show the posterior median and  
228 shaded areas the shape of the posterior distribution, with colour used to indicate estimates whose  
229 95% HPD excludes zero. E-F: Predicted salivary gland titres when varying dose and either the  
230 difference in body temperatures between the virus reservoir and the inoculated species (E) or  
231 whether or not the inoculated species is a known rabies virus reservoir (F), while keeping all other  
232 variables in the model constant. Panel E shows the posterior median of predicted virus titres. In F,  
233 predictions are shown at different quantiles of dose (in  $\log_{10}$  mouse LD50), as indicated above each  
234 sub-panel.

## 235 **Discussion**

236 The progression of viral infections within the index host following cross-species transmission is a  
237 crucial determinant of onward transmission, but is generally unobservable in nature. By analysing  
238 a unique dataset of experimental cross-species infections, we demonstrate that phylogenetic  
239 distance and specific physiological differences between the host species involved alter the  
240 progression of infections in ways that are expected to influence whether transmission in the novel  
241 host is sustained.

242 The association between incubation period duration, the extent of centrifugal spread to other  
243 tissues, and the amount of virus excreted suggests a direct mechanism linking longer incubation  
244 periods to onward transmission. Species with higher body temperatures than the reservoir host  
245 tended to have longer incubation periods, specifically at lower viral inoculation doses (Fig. 2D).  
246 Although one might expect body temperature to be a phylogenetically conserved trait, we found  
247 no correlation with phylogenetic distance (Supplementary figure S 3A), and others have shown  
248 that body temperatures are clustered primarily at higher taxonomic levels (Clarke and Rothery,  
249 2008). As a result, most bat species have lower body temperatures than most carnivore species, but  
250 the distributions of body temperatures across bat and carnivore species still overlap (Supplementary  
251 figure S 3B). The finding that incubation period duration is influenced by body temperature is

252 consistent with *in vitro* experiments showing that temperature can affect the infectivity of rabies  
253 virus, possibly by altering the rate of cell-to-cell spread (Morimoto et al., 1996). However, the  
254 specific mechanisms that could shorten incubation periods in a novel host environment which is  
255 colder than the host environment to which the virus is adapted, remain to be identified.

256 Crucially, we found clear evidence for virus adaptation to host body temperature (Supplementary  
257 figure S 2), consistent with infection progression being matched to each host species. Given this  
258 host adaptation and relatively poor correspondence between body temperature differences and  
259 phylogenetic distance, the observed temperature effect may help explain rabies virus host shifts  
260 across large phylogenetic distances. For example, despite host shifts from bats into carnivores  
261 being generally very rare, rabies virus has shifted repeatedly from big brown bats (*Eptesicus fuscus*,  
262 36 °C) to striped skunks, (*Mephitis mephitis*, 36.45 °C; Kuzmin et al., 2012; Tacutu et al., 2013).  
263 More generally, our results suggest that transmissions to species with warmer typical body  
264 temperatures than the current reservoir are more likely to become established, since this would be  
265 expected to result in longer incubation periods and higher virus excretion. This might explain  
266 observations suggesting sustained transmission of rabies virus lineages associated with common  
267 vampire bats (*Desmodus rotundus*, 35 °C) in sympatric frugivorous bats (*Artibeus lituratus*; 37.3  
268 °C; Kobayashi et al., 2007; Obregón-Morales et al., 2017; Tacutu et al., 2013).

269 The observation that low doses of viruses from bat reservoirs resulted in shorter incubation periods  
270 relative to those from carnivores (Figure 2A & B) suggests increased infectivity and/or faster  
271 within-host spread in bat-adapted rabies viruses. Since our data were limited to viruses from two  
272 bat reservoirs – with 82% of these associated with one species, *Tadarida brasiliensis* (Figure 1) –  
273 it remains unclear whether this is a general feature of bat-associated rabies viruses. But, similar  
274 results have been observed in humans, where both incubation and clinical periods were shorter  
275 when the virus originated from bats rather than carnivores (Begeman et al., 2018). This bat-  
276 associated effect may be the result of body temperature differences – at 35 and 36 °C, the two bat  
277 reservoirs included had cooler body temperatures than almost all inoculated species  
278 (Supplementary figure 3). Alternatively, since bats are considerably smaller than known carnivore  
279 reservoirs and likely transfer much smaller volumes of saliva during transmission, bat-associated  
280 rabies viruses may be adapted to transmit at lower doses. Although it may be expected that smaller  
281 animals require less virus to become infected and that body mass would modulate subsequent

282 disease progression, we found no evidence that the body mass of inoculated species or its  
283 interaction with dose affects the duration of incubation or clinical periods (Figure 2A).

284 Following the incubation period, the appearance of clinical signs of disease typically coincides  
285 with viral excretion and transmission. The duration of clinical signs is therefore crucial in  
286 determining whether an index host can transmit to conspecifics. Notably in the case of rabies virus,  
287 the clinical period coincides with the onset of signs such as aggression that facilitate onward spread  
288 through biting. It is also relatively short and invariably ends in death of the infected host,  
289 terminating transmission opportunities (Hanlon, 2013). Since many animals in our dataset were  
290 euthanised at some point after symptoms appeared (meaning the duration of their clinical period  
291 was right-censored) it was difficult to estimate effect sizes for the clinical period with enough  
292 precision to clearly distinguish them from zero (Figure 2A). However, there was some evidence  
293 that increased phylogenetic distances between virus reservoirs and inoculated species reduced the  
294 duration of clinical periods. Such increased virulence would mean that onward transmission  
295 becomes increasingly unlikely following cross-species transmission between more distant  
296 relatives. This is consistent with previous work showing that the number of successful rabies virus  
297 host shifts among North American bats decreases with phylogenetic distance (Streicker et al.,  
298 2010). This effect of phylogenetic distance on virulence and infectious period also provides further  
299 evidence that rabies virus adapts to specific reservoir species by altering disease progression.

300 Because rabies virus is generally transmitted via bite, the amount of virus excreted in the salivary  
301 glands will affect the probability of transmission from the index case. Further, the overall strong  
302 effects of dose we observed suggest that the amount of virus transferred to secondary cases will be  
303 a primary determinant of disease progression in secondary cases and hence further transmission in  
304 the new host population. Several authors have noted a positive correlation between the length of  
305 rabies virus incubation periods and subsequent virus excretion levels (Baer and Bales, 1967; Davis  
306 et al., 2013; Fekadu et al., 1982). Our results confirm this correlation across a wide range of host  
307 species combinations, and show that it can be partially explained by inoculation doses, with higher  
308 doses leading to both shorter incubation periods and decreased virus excretion. The remaining  
309 correlation is explained by clustering of disease progression kinetics on the inoculated species  
310 phylogeny, with related species having similar incubation periods and excreting similar amounts  
311 of virus (Figure 3A–C). After removing these sources of correlation, non-reservoirs tended to have  
312 lower virus titres in their salivary glands than established rabies virus reservoirs (Figure 3D & F),

313 which may explain why the virus remains restricted to a relatively small number of reservoir hosts  
314 despite frequent spillovers to other species (Mollentze et al., 2014). Our results also provide  
315 tentative evidence of shorter incubation periods in non-reservoirs (Figure 2A). The exact  
316 mechanisms underlying these differences between known rabies virus reservoirs and other species  
317 remain unexplained, but are likely to be evolutionarily conserved, given the phylogenetic clustering  
318 in excretion levels apparent in our analysis.

319 Many of our results point to adaptation of the rate of disease progression to match individual host  
320 species. Studies of rabies virus host shifts have thus far failed to find sites in the virus genome  
321 which consistently change during host adaptation (Kuzmin et al., 2012; Streicker et al., 2012;  
322 Troupin et al., 2016). This has led to the suggestion that host adaptation can be achieved through  
323 numerous sets of molecular changes (Streicker et al., 2012), which would indeed be the case if the  
324 requirement is to balance disease progression to the point where onward transmission becomes  
325 likely. In some host-pathogen systems, such adaptation is explained by a trade-off between  
326 selection for faster growth to maximize viral load and thereby infectiousness, and selection for  
327 reduced host damage (virulence) to maximize transmission opportunities (de Roode et al., 2008;  
328 Fraser et al., 2007). In contrast, we observed a positive correlation between salivary gland titres  
329 and incubation period duration, implying that slower host damage leads to greater infectiousness.  
330 This, in turn, suggests a trade-off between faster replication and/or spread (speeding up disease  
331 progression) and the ability to reach the salivary glands via functioning neural pathways. Such a  
332 trade-off is supported by the observation that the correlation between incubation period duration  
333 and salivary gland virus titre was modulated in part by dose (Figure 3C), with higher amounts of  
334 virus reducing both.

335 The experimental data analysed here offer a unique view on index infection dynamics following  
336 cross-species transmission. By revealing the complex links between dose, physiological  
337 differences between hosts, disease progression and virus excretion, our analyses bring us closer to  
338 being able to model and predict the process of disease emergence and host shifts. The large dataset  
339 of controlled infections further enabled us to generate disease progression parameter distributions  
340 for all observed combinations of within and cross-species transmissions which can be directly  
341 applied in future efforts to model rabies transmission dynamics (supplemental dataset S1). Of note  
342 – many effects observed here were eventually overcome by high doses (Figure 2B–D & Figure 3F;  
343 the median dose for experiments was 40,000 mouse LD50s), and future infection studies should

344 aim to utilize doses closer to those of natural exposures. Following cross-species transmission,  
345 rabies virus shows increased virulence (i.e., more rapid death) in more distantly related species, to  
346 the point that opportunities for transmission are likely to be markedly reduced. At the same time,  
347 a mismatch in host physiological features (including features not strongly correlated with  
348 phylogeny, such as body temperature) can alter both infectivity and disease progression, with  
349 implications for onward transmission. Thus, the picture that emerges is one of a potential virulence  
350 mismatch in index infections, that may explain why – despite having the ability to infect all  
351 mammals and frequent involvement in cross-species transmission events – rabies virus remains  
352 restricted to a relatively small number of species-specific maintenance cycles.

353 While the determinants of cross-species transmission have been the subject of intense research  
354 (reviewed in Lloyd-Smith et al., 2009; Plowright et al., 2014, 2017), the very next step, i.e. what  
355 happens during the initial infection to determine the likelihood of onward transmission, has  
356 remained relatively unexplored. Our results show that meta-analyses of cross-species infection  
357 experiments provide a tractable means of investigating this process. Expanding such analyses to  
358 other viruses may allow us to identify general rules which predict the outcome of cross-species  
359 transmissions. More work is needed to understand the host features that affect the probability of  
360 infection upon exposure, the within-host mechanisms driving virulence, and the epidemiological  
361 consequences of differences in disease progression and virus excretion. Our findings illustrate how  
362 understanding these mechanisms will be key to predicting which cross-species exposures are most  
363 likely to lead to future host shifts of rabies virus, and of zoonotic diseases more broadly.

## 364 **Methods**

### 365 Literature search and data collection

366 A search for published rabies virus infection studies was performed as described in supplementary  
367 document S1. Searching across the PubMed and Web of Science databases yielded 2279 records  
368 on 16 January 2015. These records were reviewed according to the criteria listed in table 1 to select  
369 studies for inclusion in the meta-analysis.

370 From each study, we recorded individual-level data on the species inoculated, the dose, inoculation  
371 route, and the reservoir host species of the virus used. Response variables, when available, included  
372 the observed incubation and clinical period durations, and the titre of virus present in the salivary



373 glands post-mortem. Incubation and clinical period data comprised a mixture of exact times,  
374 interval censored times (i.e. studies only reported ranges for groups of animals), and right-censored  
375 observations (i.e. deaths unrelated to rabies before the conclusion of the study or euthanasia of  
376 survivors at the end of each study). Because animals are generally euthanised once symptoms  
377 appeared, the majority of data on the duration of clinical periods were right-censored, with exact  
378 durations reported in only a small number of studies, generally performed before euthanasia  
379 became common practice in animal experiments.

380 Taxonomic classifications were updated to match Wilson and Reeder (2005), by matching the  
381 scientific and common names given in each publication against the Integrated Taxonomic  
382 Information System database<sup>1</sup>. Data from our meta-analysis were supplemented with species-level  
383 data from the PanTHERIA database, along with body-temperature data from the AnAge database  
384 (Jones et al., 2009; Tacutu et al., 2013). Because not all records were resolved to sub-species level,  
385 and external data sources only contained data at the species level, information on the specific  
386 subspecies involved was ignored in the analyses described here. This resulted in two pairs of  
387 subspecies being clustered together, while two domesticated species were analysed using species-  
388 level data for their wild ancestor (Supplementary table S 1). Further data cleaning and validation  
389 steps are described in supplementary document S1.

### 390 Accelerated failure time model

391 The durations of incubation and clinical periods were modelled using independent generalized  
392 linear mixed models on the censored event times, in this context more frequently termed  
393 accelerated failure time models. These models perform a regression on the waiting time to some  
394 specific event (e.g. the appearance of clinical signs, signifying the end of the incubation period),  
395 with coefficients acting to increase or decrease the time to the event. We assumed a log-normal  
396 distribution for the event times  $\mathbf{T}$ .

397 Thus, the duration of the incubation period of each individual  $i$  of species  $r$  (the inoculated species),  
398 inoculated with a virus from species  $d$  (the source or reservoir species) in experiment  $j$  was  
399 modelled as:

---

<sup>1</sup> [www.itis.gov](http://www.itis.gov)

$$400 \quad f(Y_{i,d,r,j} | \mu_{i,d,r,j}, \sigma_\epsilon) \sim \mathcal{N}(\mu_{i,d,r,j}, \sigma_\epsilon^2)$$
$$\mu_{i,d,r,j} = \mathbf{X}'_i \boldsymbol{\beta} + p_d + q_r + m_d + n_r + o_j$$

401 where  $\mathbf{Y} = \log(\mathbf{T})$ , while  $\mu_{i,d,r,j}$  and  $\sigma_\epsilon$  are the mean and standard deviation of a normal  
402 distribution, respectively. Coefficients are represented by  $\boldsymbol{\beta}_i$ , with  $\mathbf{X}'_i$  representing a vector of data  
403 on potential explanatory variables. Finally,  $(\mathbf{p}, \mathbf{q})$  and  $(\mathbf{m}, \mathbf{n}, \mathbf{o})$  respectively represent  
404 phylogenetic and non-phylogenetic random effects for the source species (the virus reservoir,  $\mathbf{p}$   
405 and  $\mathbf{m}$ ), inoculated species ( $\mathbf{q}$  and  $\mathbf{n}$ ), and experiment ( $\mathbf{o}$ ).

406 Phylogenetic random effects were drawn from a multivariate normal distribution taking the form

$$407 \quad \mathbf{p} \sim \mathcal{N}_k(\mathbf{0}, \sigma_p^2 \mathbf{A}_p)$$
$$\mathbf{q} \sim \mathcal{N}_l(\mathbf{0}, \sigma_q^2 \mathbf{A}_q),$$

408 where  $\mathbf{0}$  is a vector of zeros (of length  $k$  or  $l$ , equal to the number of source or inoculated species,  
409 respectively), and  $\sigma_p^2$  and  $\sigma_q^2$  are variance parameters.  $\mathbf{A}_p$  and  $\mathbf{A}_q$  represent correlation matrixes  
410 for all source and inoculated species, respectively. These matrixes were calculated from a  
411 composite time-scaled phylogeny generated by timetree.org (Kumar et al., 2017) assuming a  
412 Brownian model of trait evolution using version 3.5 of the APE package in R (Hadfield and  
413 Nakagawa, 2010; Paradis et al., 2004). These random effects adjust for potential correlation in the  
414 response variables due to relatedness. Similar results were obtained when using the mammalian  
415 supertree of Bininda-Emonds et al. (2007), but this supertree had a slightly lower resolution than  
416 the timetree.org phylogeny.

417 The non-phylogenetic random effects took the form

$$418 \quad m_d \sim \mathcal{N}(0, \sigma_m^2)$$
$$n_r \sim \mathcal{N}(0, \sigma_n^2)$$
$$o_j \sim \mathcal{N}(0, \sigma_o^2),$$

419 where  $\sigma_m^2$  and  $\sigma_n^2$  respectively measure the variance between source and inoculated species not  
420 captured by the Brownian model (Longdon et al., 2011), while  $\sigma_o^2$  measures the variance between  
421 experiments. A similar model was used for the duration of clinical periods, except that data were  
422 pooled across virus reservoirs by removing the random effects for reservoir species and reservoir  
423 phylogeny ( $\mathbf{p}$  and  $\mathbf{m}$  above). This was necessary because the clinical period data involved viruses

424 from just four reservoir species, making it impossible to accurately estimate the variance between  
425 observations associated with different reservoirs.

426 To accommodate censoring, the vector of event times,  $\mathbf{T}$ , was treated as a latent variable. When  
427 only the range of incubation or clinical period durations was given for a specific group of animals,  
428 data was treated as interval censored, i.e.  $T_{i,d,r,j} \in [L_{i,d,r,j}, U_{i,d,r,j}]$ , where  $\mathbf{L}$  and  $\mathbf{U}$  represent the  
429 lower and upper boundaries of observed event times (supplementary document S1). When an  
430 animal left the dataset without experiencing an event (e.g. when it was euthanised), the time of  
431 censoring was recorded as a lower boundary, reflecting our knowledge that the event of interest  
432 had not yet occurred by this point, but could have occurred at any time after this. In this case  
433  $T_{i,d,r,j} \in [L_{i,d,r,j}, \infty)$ .

434 In this context, exact observations can be recorded simply by setting  $L_{i,d,r,j} = U_{i,d,r,j}$ .

#### 435 Multi-response models

436 In an independent model, the amount of virus detected in salivary glands *post mortem* was modelled  
437 jointly with incubation period durations, to allow estimation of the amount of residual correlation  
438 between incubation period duration and the amount of virus in the salivary glands. Several authors  
439 have noted a link between these measures (Baer and Bales, 1967; Davis et al., 2013; Fekadu et al.,  
440 1982), but it remains unexplained.

441 This regression was similar to the model above, except that the normal distribution on  
442 log(observations) was replaced with a multivariate normal distribution. Thus, in the full model, the  
443 observed value  $Y$  of response variable  $v$  for individual  $i$  of species  $r$  in experiment  $j$  was modelled  
444 as

$$445 \quad \begin{aligned} f(Y_{v,i,r,j} | \mu_{v,i,r,j}, \sigma_{\epsilon}) &\sim \mathcal{N}_2(\mu_{v,i,r,j}, \sigma_{\epsilon}) \\ \mu_{v,i,r,j} &= \mathbf{X}'_i \boldsymbol{\beta}_v + q_{v,r} + o_{v,j} \end{aligned}$$

446 where  $v = 1$  represents the incubation period (i.e.  $\mathbf{Y}_{v=1} = \log(\mathbf{T})$ ) and  $v = 2$  is the virus titre in  
447 the salivary glands. Virus titre was thus also modelled as log-normal, i.e.  $\mathbf{Y}_{v=2} = \log(1 + \mathbf{W})$ ,  
448 where  $\mathbf{W}$  represents the observed titres, which may be 0 if no virus was detected in the salivary  
449 glands).  $\boldsymbol{\beta}_v$  is a vector of coefficients unique to each response variable, while  $q_{v,r}$  and  $o_{v,r}$  are  
450 random effects for inoculated species phylogeny and experiment, respectively.

451 In these models,  $\sigma_{\varepsilon}$  is a variance-covariance matrix of the form

452 
$$\sigma_{\varepsilon} = \begin{bmatrix} \sigma_{\varepsilon,v=1}^2 & \sigma_{\varepsilon,v=1,v=2} \\ \sigma_{\varepsilon,v=1,v=2} & \sigma_{\varepsilon,v=2}^2 \end{bmatrix}$$

453 where  $\sigma_{\varepsilon,v=1}^2$  is the residual variance in response variable 1 (incubation period durations), and  
454  $\sigma_{\varepsilon,v=1,v=2}$  is the residual covariance between incubation period durations and salivary gland titres.  
455 From this, the Pearson correlation between incubation periods and salivary gland titres can be  
456 calculated as

457 
$$\frac{\sigma_{\varepsilon,v=1,v=2}}{\sigma_{\varepsilon,v=1} \sigma_{\varepsilon,v=2}}.$$

#### 458 Explanatory variables

459 Variables measuring differences between the reservoir and the inoculated species were included to  
460 assess the influence of previous virus adaptation on the outcome of infection in heterologous host  
461 species. These included the phylogenetic distance between the reservoir and the inoculated species,  
462 measured as patristic distances along the same composite time-scaled phylogeny generated by  
463 timetree.org used above. As above, similar results were obtained when using the mammalian  
464 supertree of Bininda-Emonds et al. (2007). We also included the difference in typical body  
465 temperatures between the reservoir and inoculated species, as an example of a physiological  
466 difference which does not appear to follow phylogenetic constraints (Clarke and Rothery, 2008),  
467 because temperature is known to affect rabies virus infectivity *in vitro* (Morimoto et al., 1996).  
468 Finally, a binary variable distinguishing viruses derived from bat and carnivore reservoirs was  
469 included, because differences in the clinical presentation of bat- and carnivore-associated rabies  
470 virus infection in humans have been noted (Begeman et al., 2018).

471 Features of the inoculated host species were accommodated primarily through random effects for  
472 species and inoculated species phylogeny. However, we also included a measure of the typical  
473 body mass of the inoculated species, since larger species may be proportionally more resistant to  
474 the effects of a given dose of virus. Because only some species maintain rabies virus transmission  
475 endemically, for reasons that are not well understood, a binary variable distinguishing known  
476 reservoirs of rabies virus from other inoculated species was also included.

477 Differences between experiments were accommodated by including variables for dose, the  
478 inoculation site, and whether the inoculum consisted of brain material or was derived from salivary  
479 glands/saliva, along with a random effect distinguishing between experiments to accommodate any  
480 remaining differences. Because the doses encountered in these experiments differed over several  
481 orders of magnitude and the effects of increasing dose is assumed to decrease (saturate) at very  
482 large doses, this variable was included in its log-transformed form. The varying inoculation routes  
483 encountered were summarised as a ‘proportional inoculation distance’, representing the relative  
484 distance between the inoculation site and the brain (the primary site of rabies virus replication).  
485 This distance was calculated by classifying inoculation routes by body part (head, neck, torso or  
486 limbs) and depth (intracranial, intramuscular, or subcutaneous) and was expressed as a proportion,  
487 where 1 indicates the furthest and shallowest possible inoculation site relative to the brain (sub-  
488 cutaneous inoculation of a limb), while 0 indicates intracerebral inoculation (Supplementary table  
489 S 2 & Supplementary table S 3). Such proportional scaling means this variable is independent of  
490 the differing body sizes of the inoculated species. Finally, because larger doses may compensate  
491 for any decreases in infectivity caused by features of the inoculated species and/or physiological  
492 differences between the inoculated species and the reservoir to which the virus was adapted, we  
493 included interactions between dose and all host and virus effects above.

#### 494 Model fitting

495 Models were fit using version 2.25 of the MCMCglmm package in R version 3.5.1 (Hadfield, 2010;  
496 R Core Team, 2018). All coefficients and the residual variance parameter received the default prior  
497 distributions used by MCMCglmm, while parameter-expanded priors were used for the variance  
498 parameters of all random effects (Gelman, 2006). For each dataset, models were fitted using  
499 10 million MCMC steps, saving every 1000th sample. The first 10% of samples in each chain were  
500 discarded as burn-in. Results were inspected and summarised using version 0.18-1 of the coda  
501 package in R (Plummer et al., 2006). Effective sample sizes were checked to ensure efficient  
502 sampling was achieved, and chains were visually inspected for convergence.

#### 503 **Acknowledgments**

504 We thank Paul Johnson, Mafalda Viana, Dan Haydon and Ben Longdon for helpful discussions  
505 and advice on statistical analyses, and Andrew Yates for comments on a draft on this manuscript.

506 NM was funded by a Lord Kelvin – Adam Smith studentship from the University of Glasgow.  
507 PRM was funded by the Medical Research Council (MC\_UU\_12014/9). DGS was supported by a  
508 Sir Henry Dale Fellowship, jointly funded by the Wellcome Trust and Royal Society  
509 (102507/Z/13/Z) and a Wellcome Senior Research Fellowship (217221/Z/19/Z). KH was  
510 supported by the Wellcome Trust (207569/Z/17/Z & 095787/Z/11/Z).

## 511 **Competing interests**

512 The authors declare no competing interests.

## 513 **References**

514 Baer GM, Bales GL. 1967. Experimental rabies infection in the Mexican freetail bat. *J Infect Dis*  
515 **117**:82–90.

516 Begeman L, GeurtsvanKessel C, Finke S, Freuling CM, Koopmans M, Müller T, Ruigrok TJH,  
517 Kuiken T. 2018. Comparative pathogenesis of rabies in bats and carnivores, and implications for  
518 spillover to humans. *Lancet Infect Dis* **18**:e147–e159. doi:10.1016/S1473-3099(17)30574-1

519 Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, Grenyer R, Price SA,  
520 Vos RA, Gittleman JL, Purvis A. 2007. The delayed rise of present-day mammals. *Nature*  
521 **446**:507–512. doi:10.1038/nature07347

522 Chancey C, Grinev A, Volkova E, Rios M. 2015. The Global Ecology and Epidemiology of West  
523 Nile Virus. *BioMed Res Int*. doi:10.1155/2015/376230

524 Charlton KM, Casey GA, Campbell JB. 1987. Experimental rabies in skunks: immune response  
525 and salivary gland infection. *Comp Immunol Microbiol Infect Dis* **10**:227–35.

526 Clarke A, Rothery P. 2008. Scaling of body temperature in mammals and birds. *Funct Ecol* **22**:58–  
527 67. doi:10.1111/j.1365-2435.2007.01341.x

528 Davis AD, Jarvis JA, Pouliott CE, Morgan SMD, Rudd RJ. 2013. Susceptibility and pathogenesis  
529 of little brown bats (*Myotis lucifugus*) to heterologous and homologous rabies viruses. *J Virol*  
530 **87**:9008–9015. doi:10.1128/JVI.03554-12

- 531 de Roode JC, Yates AJ, Altizer S. 2008. Virulence-transmission trade-offs and population  
532 divergence in virulence in a naturally occurring butterfly parasite. *Proc Natl Acad Sci U S A*  
533 **105**:7489–7494. doi:10.1073/pnas.0710909105
- 534 Fekadu M, Chandler F, Harrison A. 1982. Pathogenesis of rabies in dogs inoculated with an  
535 Ethiopian rabies virus strain. Immunofluorescence, histologic and ultrastructural studies of the  
536 central nervous system. *Arch Virol* **71**:109–126.
- 537 Fisher CR, Streicker DG, Schnell MJ. 2018. The spread and evolution of rabies virus: Conquering  
538 new frontiers. *Nat Rev Microbiol* **16**:241–255. doi:10.1038/nrmicro.2018.11
- 539 Fraser C, Hollingsworth TD, Chapman R, de Wolf F, Hanage WP. 2007. Variation in HIV-1 set-  
540 point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc Natl Acad Sci U S*  
541 *A* **104**:17441–6. doi:10.1073/pnas.0708559104
- 542 Gelman A. 2006. Prior distribution for variance parameters in hierarchical models. *Bayesian Anal*  
543 **1**:515–533. doi:10.1214/06-BA117A
- 544 Gilbert GS, Webb CO. 2007. Phylogenetic signal in plant pathogen-host range. *Proc Natl Acad Sci*  
545 **104**:4979–4983. doi:10.1073/pnas.0607968104
- 546 Hadfield JD. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The  
547 **MCMCglmm** R Package. *J Stat Softw* **33**. doi:10.18637/jss.v033.i02
- 548 Hadfield JD, Nakagawa S. 2010. General quantitative genetic methods for comparative biology:  
549 Phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J Evol*  
550 *Biol* **23**:494–508. doi:10.1111/j.1420-9101.2009.01915.x
- 551 Hampson K, Coudeville L, Lembo T, Sambo M, Kieffer A, Attlan M, Barrat J, Blanton JD, Briggs  
552 DJ, Cleaveland S, Costa P, Freuling CM, Hiby E, Knopf L, Leanes F, Meslin F-X, Metlin A,  
553 Miranda ME, Müller T, Nel LH, Recuenco S, Rupprecht CE, Schumacher C, Taylor L, Vigilato  
554 MAN, Zinsstag J, Dushoff J, on behalf of the Global Alliance for Rabies Control Partners for  
555 Rabies Prevention. 2015. Estimating the Global Burden of Endemic Canine Rabies. *PLoS Negl*  
556 *Trop Dis* **9**:e0003709. doi:10.1371/journal.pntd.0003709

- 557 Hanlon CA. 2013. Chapter 5 - Rabies in Terrestrial Animals In: Jackson AC, editor. Rabies.  
558 Boston: Academic Press. pp. 179–213. doi:10.1016/B978-0-12-396547-9.00005-5
- 559 Hemachudha T, Laothamatas J, Rupprecht CE. 2002. Human rabies: a disease of complex  
560 neuropathogenetic mechanisms and diagnostic challenges. *Lancet Neurol* **1**:101–109.  
561 doi:10.1016/S1474-4422(02)00041-8
- 562 Jones KE, Bielby J, Cardillo M, Fritz SA, O'Dell J, Orme CDL, Safi K, Sechrest W, Boakes EH,  
563 Carbone C, Connolly C, Cutts MJ, Foster JK, Grenyer R, Habib M, Plaster CA, Price SA, Rigby  
564 EA, Rist J, Teacher A, Bininda-Emonds ORP, Gittleman JL, Mace GM, Purvis A. 2009.  
565 PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently  
566 extinct mammals. *Ecology* **90**:2648–2648. doi:10.1890/08-1494.1
- 567 Kobayashi Y, Sato G, Kato M, Ito T, Cunha EMS, Silva MV, Mota CS, Ito FH, Sakai T. 2007.  
568 Genetic diversity of bat rabies viruses in Brazil. *Arch Virol* **152**:1995. doi:10.1007/s00705-007-  
569 1033-y
- 570 Kumar S, Stecher G, Suleski M, Hedges SB. 2017. TimeTree: A Resource for Timelines,  
571 Timetrees, and Divergence Times. *Mol Biol Evol* **34**:1812–1819. doi:10.1093/molbev/msx116
- 572 Kuzmin IV, Shi M, Orciari LA, Yager PA, Velasco-Villa A, Kuzmina NA, Streicker DG, Bergman  
573 DL, Rupprecht CE. 2012. Molecular inferences suggest multiple host shifts of rabies viruses from  
574 bats to mesocarnivores in Arizona during 2001-2009. *PLoS Pathog* **8**:e1002786.  
575 doi:10.1371/journal.ppat.1002786
- 576 Lloyd-Smith JO, George D, Pepin KM, Pitzer VE, Pulliam JRC, Dobson AP, Hudson PJ, Grenfell  
577 BT. 2009. Epidemic dynamics at the human-animal interface. *Science* **326**:1362–7.  
578 doi:10.1126/science.1177345
- 579 Longdon B, Brockhurst MA, Russell CA, Welch JJ, Jiggins FM. 2014. The evolution and genetics  
580 of virus host shifts. *PLoS Pathog* **10**:e1004395. doi:10.1371/journal.ppat.1004395
- 581 Longdon B, Hadfield JD, Day JP, Smith SCL, Mcgonigle E, Cogni R, Cao C, Jiggins FM. 2015.  
582 The Causes and Consequences of Changes in Virulence following Pathogen Host Shifts. *PLoS*  
583 *Pathog* **11**:e1004728. doi:10.1371/journal.ppat.1004728



- 584 Longdon B, Hadfield JD, Webster CL, Obbard DJ, Jiggins FM. 2011. Host phylogeny determines  
585 viral persistence and replication in novel hosts. *PLoS Pathog* **7**:e1002260.  
586 doi:10.1371/journal.ppat.1002260
- 587 Mollentze N, Biek R, Streicker DG. 2014. The role of viral evolution in rabies host shifts and  
588 emergence. *Curr Opin Virol* **8**:68–72. doi:10.1016/j.coviro.2014.07.004
- 589 Morimoto K, Patel M, Corisdeo S, Hooper DC, Fu ZF, Rupprecht CE, Koprowski H, Dietzschold  
590 B. 1996. Characterization of a unique variant of bat rabies virus responsible for newly emerging  
591 human cases in North America. *Proc Natl Acad Sci U S A* **93**:5653–5658.
- 592 Obregón-Morales C, Aguilar-Setién Á, Perea Martínez L, Galvez-Romero G, Martínez-Martínez  
593 FO, Aréchiga-Ceballos N. 2017. Experimental infection of *Artibeus intermedius* with a vampire  
594 bat rabies virus. *Comp Immunol Microbiol Infect Dis* **52**:43–47. doi:10.1016/j.cimid.2017.05.008
- 595 Ogbu O, Ajuluchukwu E, Uneke CJ. 2007. Lassa fever in West African sub-region: an overview  
596 11.
- 597 Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in R  
598 language. *Bioinformatics* **20**:289–290. doi:10.1093/bioinformatics/btg412
- 599 Plowright RK, Eby P, Hudson PJ, Smith IL, Westcott D, Bryden WL, Middleton D, Reid PA,  
600 McFarlane RA, Martin G, Tabor GM, Skerratt LF, Anderson DL, Crameri G, Quammen D, Jordan  
601 D, Freeman P, Wang LF, Epstein JH, Marsh GA, Kung NY, McCallum H. 2014. Ecological  
602 dynamics of emerging bat virus spillover. *Proc R Soc B Biol Sci* **282**. doi:10.1098/rspb.2014.2124
- 603 Plowright RK, Parrish CR, McCallum H, Hudson PJ, Ko AI, Graham AL, Lloyd-Smith JO. 2017.  
604 Pathways to zoonotic spillover. *Nat Rev Microbiol* **15**:502–510. doi:10.1038/nrmicro.2017.45
- 605 Plummer M, Best N, Cowles K, Vines K. 2006. CODA: convergence diagnosis and output analysis  
606 for MCMC. *R News* **6**:7–11. doi:10.1159/000323281
- 607 R Core Team. 2018. R: A Language and Environment for Statistical Computing. Vienna: R  
608 Foundation for Statistical Computing.

- 609 Rupprecht CE, Hanlon CA, Hemachudha T. 2002. Rabies re-examined. *Lancet Infect Dis* **2**:327–  
610 343.
- 611 Streicker D, Altizer S, Velasco-Villa A, Rupprecht C. 2012. Variable evolutionary routes to host  
612 establishment across repeated rabies virus host shifts among bats. *Proc Natl Acad Sci U S A*  
613 **109**:19715–19720. doi:10.1073/pnas.1203456109
- 614 Streicker DG, Turmelle AS, Vonhof MJ, Kuzmin IV, McCracken GF, Rupprecht CE. 2010. Host  
615 Phylogeny Constrains Cross-Species Emergence and Establishment of Rabies Virus in Bats.  
616 *Science* **329**:676–679. doi:10.1126/science.1188836
- 617 Tacutu R, Craig T, Budovsky A, Wuttke D, Lehmann G, Taranukha D, Costa J, Fraifeld VE, De  
618 Magalhães JP. 2013. Human ageing genomic resources: Integrated databases and tools for the  
619 biology and genetics of ageing. *Nucleic Acids Res* **41**:D1027–D1033. doi:10.1093/nar/gks1155
- 620 Troupin C, Dacheux L, Tanguy M, Sabeta C, Blanc H, Bouchier C, Vignuzzi M, Duchene S,  
621 Holmes EC, Bourhy H. 2016. Large-scale phylogenomic analysis reveals the complex evolutionary  
622 history of rabies virus in multiple carnivore hosts. *PLoS Pathog* **12**:1–20.  
623 doi:10.1371/journal.ppat.1006041
- 624 Wilson DE, Reeder DM, editors. 2005. Mammal species of the world: a taxonomic and geographic  
625 reference, 3rd ed. Baltimore: Johns Hopkins University Press.
- 626 Wolfe ND, Dunavan CP, Diamond J. 2007. Origins of major human infectious diseases. *Nature*  
627 **447**:279–283. doi:10.1038/nature05775
- 628

629

## Supplementary material

630 **Supplementary tables & figures**

631 Supplementary table S 1: Known sub-species in the dataset, which were analysed at the species

632 level

<b>Species</b>	<b>Subspecies</b>	<b>Common name</b>
<i>Antrozous pallidus</i>	<i>pallidus</i>	Pallid bat
<i>Canis lupus</i>	<i>familiaris</i>	Domestic dog
<i>Mustela putorius</i>	<i>furo</i>	Domestic ferret
<i>Tadarida brasiliensis</i>	<i>mexicana</i>	Mexican free-tailed bat
	<i>brasiliensis</i>	Brazilian free-tailed bat
<i>Vulpes vulpes</i>	<i>crucigera</i>	European red fox
	<i>fulves</i>	Eastern North American red fox

633

634 Supplementary table S 2: Classification of inoculation routes by body zone (rows) and depth  
 635 (columns)\*

	<i>Intracranial</i>	<i>Intramuscular</i>	<i>Subcutaneous/peripheral</i>
<i>Head</i>	Intracerebral, Optic nerve, Intracranial, Intracisternal, IC, IC intercranial	IM-masseter, IM-temporal, IM occipital, Submandibular, IM-masseter, Masseter, IM masseter, IM-Masseter, IM-masseter(both.sides), IM-masseter(both), left & right masseter, IM Right Masseter, IM M.masseter, IM Left Masseter, IM-maseter	Intranasal, Intranasally, Subcutaneous-lips
<i>Neck</i>		IM-neck, IM-cervical, IM-cervival(left), IM-deap-neck	Subcutaneous-neck, subcutaneous(neck)
<i>Torso</i>		Deltoid, IM-deltoid, IM-Deltoid, IM-pectoral, IM-Dorsal	SC-deltoid
<i>Limbs</i>		IM-quadriceps femoris (right), IM-right hindleg, IM-hind leg, IM-leg, IM right thigh, IM-LongDigitalExtensor(Pelvic), IM-AbductorDigitiQuinti(Pelvic), IM-gastrocnemius, Right gastrocnemius muscle, Gastrocnemicus , IM-crural, IM-HindFoot	Peripheral-footpad, SC-HindLimb

636 \*The inoculation routes shown are free-text entries present in the raw data, where IC = intracranial,

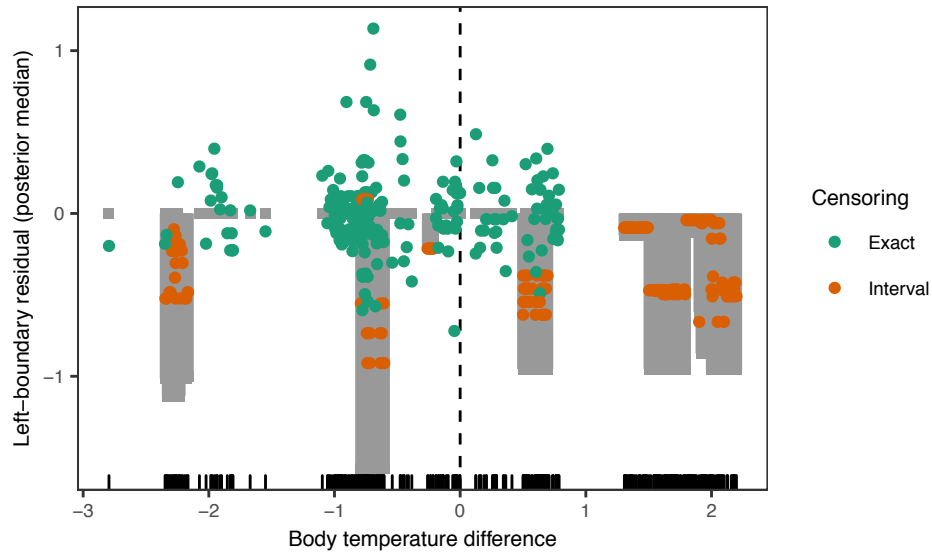
637 IM = intramuscular, and SC = subcutaneous.

638

639 Supplementary table S 3: Distances assigned to each inoculation route category to express relative  
640 distance to the brain\*

	<i>Intracranial (0)</i>	<i>Intramuscular (1/7)</i>	<i>Subcutaneous (3/14)</i>
<i>Head (1/14)</i>	0	$1/14 + 1/7$	$1/14 + 3/14$
<i>Neck (1/7)</i>	-	$1/7 + 1/7$	$1/7 + 3/14$
<i>Torso (3/7)</i>	-	$3/7 + 1/7$	$3/7 + 3/14$
<i>Limbs (7/7)</i>	-	$7/7 + 1/7$	$7/7 + 3/14$

641 \*Distances were chosen to ensure intramuscular inoculations to the neck received the same value  
642 as subcutaneous inoculations to the head, and were subsequently scaled to range between 0  
643 (intracranial, head) and 1 (subcutaneous, head). See supplementary document S1 for details.



644

645 Supplementary figure S 1: Posterior median residuals as a function of body temperature difference.

646 The observed data was subject to censoring, but the lower-boundary was always known. Residuals

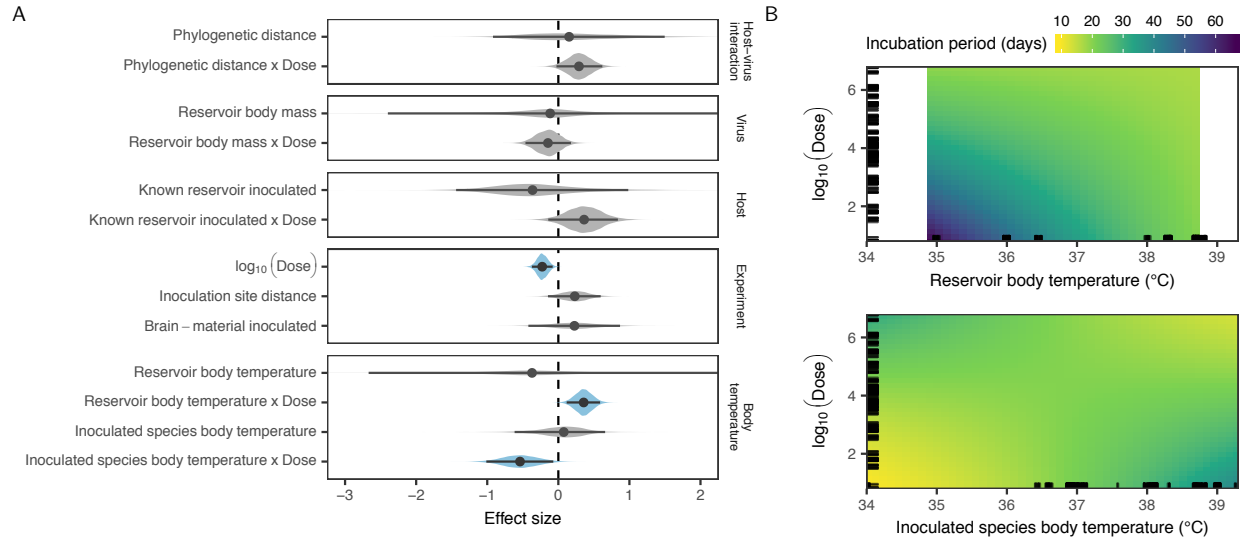
647 were calculated as the observed lower limit minus the fitted value, with grey-shaded areas showing

648 the expected range of each observation. In the case of exact observations, a perfect fit would result

649 in residuals equal to 0, while the residuals for interval-censored observations are expected to lie in

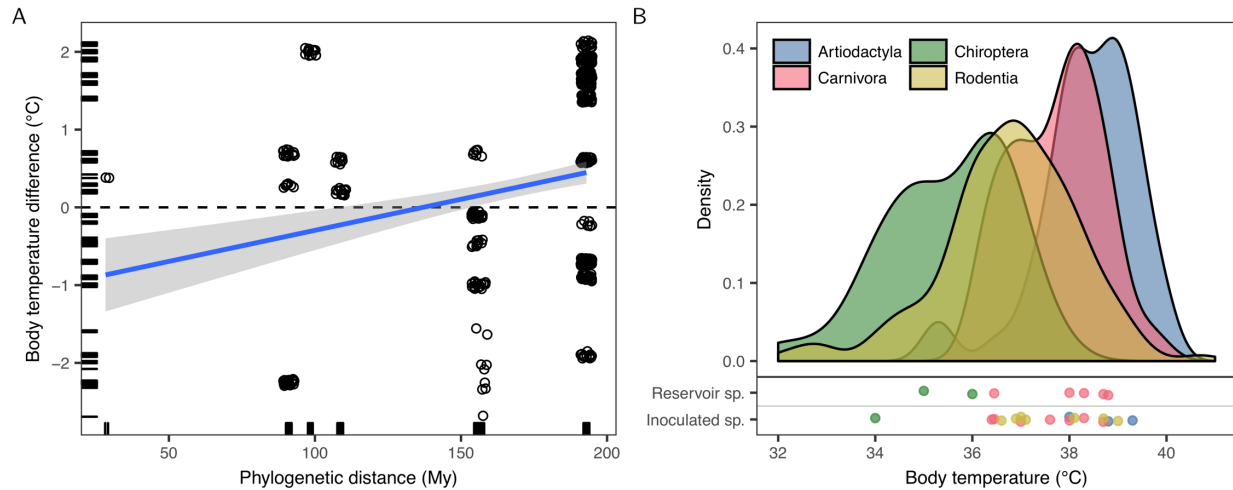
650  $[0, \text{lower} - \text{upper})$ . No systematic bias in residuals was observed, supporting the conclusion that the

651 effect of body temperature distance remained linear for both positive and negative values.



652

653 Supplementary figure S 2: Coefficient estimates and predicted effects when fitting separate effects  
654 for the typical body temperature of the inoculated species and virus reservoir. A: Coefficient  
655 estimates, with lines showing the 95% highest posterior distribution and points showing the  
656 posterior median. Compared to the model shown in the main text, this model does not contain an  
657 effect for inoculated species body mass (which had no effect in the full model, and is correlated  
658 with inoculated species body temperature) and contains an effect for reservoir body mass, replacing  
659 the variable separating bat-associated viruses from carnivore-associated viruses (which was  
660 correlated with inoculated species phylogeny). B: Posterior median predicted incubation periods  
661 as a function of reservoir (top) or inoculated species body temperature (bottom) and dose.  
662 Predictions are shown within the range of observed values in the dataset.



663  
664 Supplementary figure S 3: Relationship between phylogenetic distance and host body temperature.  
665 A: Relationship between phylogenetic distance (in millions of years) and body temperature  
666 difference in the incubation period dataset. Points are shown jittered for clarity, while the blue line  
667 represents the best linear fit. B: Distribution of body temperature in the taxonomic orders of hosts  
668 included in this study, based on all species for which body temperature data are available in the  
669 AnAge database. Points indicate the body temperatures for species included in the incubation  
670 period dataset as either reservoirs or inoculated species, jittered vertically to reduce overlap.



## 671 **Supplementary document S1: Supplementary methods**

### 672 Literature search

673 To allow optimisation of the literature search strategy, a set of publications of known relevance  
674 (according to the abstract review criteria, *Supplementary table S 4*) was obtained by manual  
675 searching and from citations in several rabies textbooks. A script implementing the BioPython  
676 library in Python 3 was used to find a search query that returned the greatest possible number of  
677 these publications in a PubMed search, whilst excluding the most common sources of false  
678 positives – notably annual veterinary reports and vaccine studies. This resulted in the following  
679 query for Web of Science:

```
680 TS=(rabies AND
681     (inoculated OR inoculate OR
682     ((intranasal* OR intracerebral* OR
683     intracranial* OR intramuscular* OR
684     subcutaneous* OR intraperitoneal*))
685     AND (inoculat* OR challenge OR inject*) NOT vaccine) OR
686     "experimental infection" OR
687     "experimental rabies infection" OR
688     (experiment* NEAR transmission))
689     NOT (review OR "fixed virus" OR "fixed rabies virus" OR
690         "annual report")
```

691 The same query was used to search the PubMed database, except that all instances of the NEAR  
692 operator were replaced with AND (PubMed does not support proximity searching). The PubMed  
693 version of this query returned 46 out of the 53 publications of known importance confirmed to be  
694 in the PubMed database. Of the remaining seven, five publications did not have searchable abstracts  
695 in the PubMed database.

696 Combining the results from Web of Science and PubMed yielded 2501 records on 16 January 2015.  
697 The search results were further filtered to exclude records where the keyword rabies did not occur  
698 in either the English title or abstract, although records where either of these fields was empty were  
699 also retained. The abstracts of the resulting 2279 records were reviewed according to the criteria in  
700 *Supplementary table S 4*. The abstracts of 35 records (1.54% of the filtered search results) could

701 not be obtained in time for consideration. A total of 412 records (18.36% of the remaining search  
702 results) were selected for full text review, and the full texts of 382 publications (92.72%) were  
703 successfully obtained.

704 Publications in languages other than English were subjected to optical character recognition (where  
705 needed) using VietOCR version 4.0, a graphical user interface for the Tesseract OCR engine  
706 (version 3.30RC), optimised for each language being recognised. The digitised text was manually  
707 corrected and machine-translated using the statistical machine translation version of Google  
708 translate. All full texts were then reviewed using the criteria in *Supplementary table S 4*, resulting  
709 in the selection of 63 publications for inclusion in this study.

710

711 Supplementary table S 4: Criteria used to select publications

<b>Review stage</b>	<b>Criteria</b>
Abstract review	Involves experimental inoculations with rabies virus Does not solely focus on laboratory-adapted virus strain
Full text review	Source and inoculated species clearly identified Source host species is a known maintenance host of rabies virus (passaging in non-maintenance hosts allowed) Inoculated host species is a mammal Inoculated host species is wild-type or outbred Inoculated dose stated (in mouse LD <sub>50</sub> ) Healthy, unvaccinated animals used Multiple experimental groups or a control group present

712

### 713 Data cleaning and validation

714 Individual-level data on all inoculated animals present in the selected publications were recorded.  
715 Following data-entry, the raw data was extensively checked for accuracy and consistency. The  
716 scientific names of all inoculated and source species were manually checked for current validity  
717 using the ITIS database and standardised to match the taxonomy of Wilson and Reeder (2005). A  
718 further literature search was performed to obtain data on whether each of the included species is a  
719 known maintenance host of rabies virus. Records where the source species was not a known  
720 maintenance host were validated against the original publication to check for passaging and

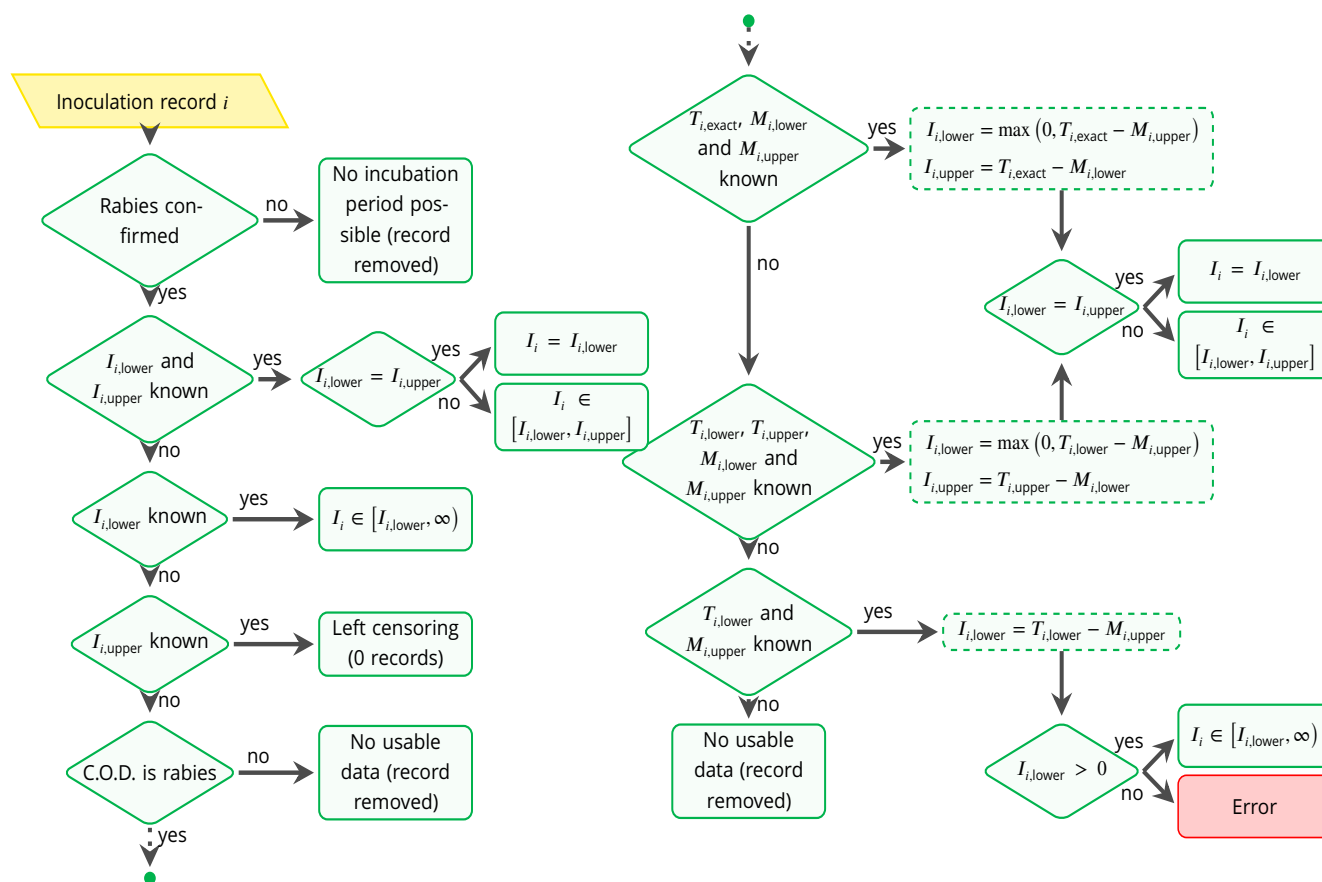
721 information on the maintenance host associated with the virus inoculum used. All virus inoculum  
722 identifiers were manually checked to identify the re-use of viruses in multiple publications, and  
723 were corrected to reflect this shared origin.

724 Some data were excluded from the current study, but retained in the database for potential future  
725 studies. A total of 197 records were excluded because the method of titration used to determine the  
726 dose inoculated was either not known (generally negative control animals) or because their reported  
727 inoculation dose involved a tissue culture-based rather than an *in vivo* method of titration. All  
728 remaining records listed virus doses obtained by intra-cerebral inoculation of mice to determine  
729 the 50% lethal dose. A further 31 records with a reported dose of 0, representing non-inoculated  
730 control animals, were also excluded. A total of 131 records involving five source species were  
731 excluded because the *source* species was either not a known maintenance host of rabies virus or  
732 could not be resolved to the species level, and 17 records involving one inoculated species were  
733 excluded because the *inoculated* species could not be resolved to species level.

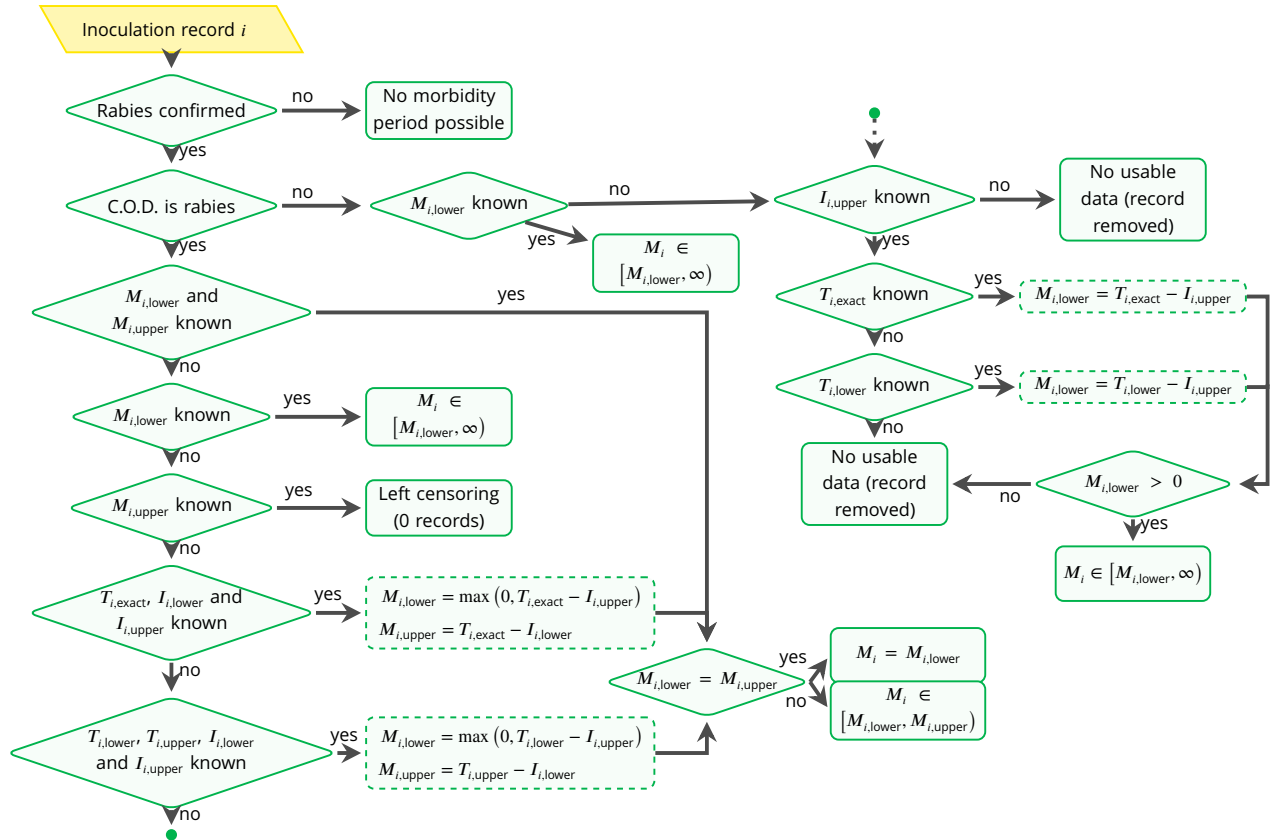
734 A derived variable was created to unify diverse descriptions of inoculation routes and sites, based  
735 on the proportional distance of each inoculation site from the brain. Inoculation routes were  
736 classified into four zones on the body and three depths, and distances were assigned to each  
737 category (*Supplementary table S 2 & Supplementary table S 3*). These distances were normalised  
738 to lie in [0,1]. Inoculation routes that were not specific enough to be classified into this scheme  
739 were checked against the original publications, and 115 records where the inoculation route  
740 information could not be improved were excluded.

741 All available data on the timing of disease progression and the total number of days survived  
742 following inoculation were recorded. When timing data was reported for groups of animals, the  
743 event times for each animal in the group was recorded as known only to be within the interval  
744 reported for that group (a form of interval censoring). Data on the total number of days survived  
745 was recorded in both exact ( $T_{\text{exact}}$ ) and interval censored ( $T_{\text{lower}}$  and  $T_{\text{upper}}$ ) forms, while the  
746 length of incubation and clinical periods were recorded in interval censored form only (with both  
747 columns given the same value when an exact time was available). Although only two of these  
748 timing variables needs to be known to calculate the third, data was recorded only in the form  
749 reported in each study.

750 The data were then processed as described in supplementary figures 12 - 13, and timing data were  
 751 checked for internal consistency. This involved consolidating information on the cause of death,  
 752 results of clinical observations and post-mortem diagnostic tests and the timing of disease  
 753 progression stages associated with each animal into variables describing the lower and upper limits  
 754 of the incubation and clinical period duration. When rabies was confirmed for a given animal  
 755 through either clinical diagnosis or post-mortem diagnostic tests, it was assumed to have rabies. In  
 756 many cases, data could be calculated for unreported timing variables by using some combination  
 757 of reported data on either the number of days survived post inoculation, the duration of the  
 758 incubation period, and/or the duration of the clinical period. For example, if only the earliest day  
 759 that an animal  $i$  could have died ( $T_{i,lower}$ ) as well as the upper limit of the clinical period for that  
 760 animal ( $M_{i,upper}$ ) is known, the lower limit of the incubation period can be calculated as  $I_{i,lower} =$   
 761  $T_{i,lower} - M_{i,upper}$  (supplementary figure S12).



762  
 763 Supplementary figure S12: Overview of the algorithm used to consolidate data on incubation  
 764 periods.



766

767 Supplementary figure S13: Overview of the algorithm used to consolidate data on clinical periods.