Title: Reservoir host immunology and life history shape virulence evolution in zoonotic viruses

Authors: Cara E. Brook[1,2*], Carly Rozins[3], Sarah Guth[1], Mike Boots[1,4]

Affiliations:
[1] Department of Integrative Biology, University of California, Berkeley, Berkeley, CA, USA
[2] Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA
[3] Department of Science and Technology Studies, York University, Toronto, Canada
[4] Biosciences, University of Exeter, Penryn Campus, UK. TR10 9FE
* Corresponding author. Email: cbrook@uchicago.edu

Abstract: Future pandemic risk management requires better understanding of the mechanisms that determine the virulence of emerging zoonotic viruses. Bats host viruses that cause higher case fatality rates upon spillover to humans than those derived from any other mammal, suggesting that reservoir host immunological and life history traits may be important drivers of cross-species virulence. Using a nested population-level and within-host modelling approach, we generate virulence predictions for viral zoonoses derived from diverse mammalian reservoirs, successfully recapturing corresponding virus-induced human mortality rates from the literature. Our work offers a mechanistic explanation for the virulence of bat-borne zoonoses and, more generally, demonstrates how key differences in reservoir host longevity, tolerance, and population density impact the evolution of viral traits that generate severe disease following spillover to humans. We provide a theoretical framework that offers a series of testable questions and hypotheses designed to stimulate future work comparing cross-species virulence evolution in zoonotic viruses derived from diverse mammalian hosts.

Main Text:

The devastating impacts of the SARS-CoV-2 pandemic highlight the extreme public health outcomes that can result upon cross-species emergence of zoonotic viruses. Predicting the virulence of an emerging virus is a complex problem since chance will always play a role in dictating the initial spillover that precedes selection[1], virulence upon emergence may be maladaptive in novel hosts[2,3], and patterns in available data may be muddled by attainment bias if avirulent infections go underreported[1]. Nonetheless, reservoir host immunological and life history traits may be important drivers of cross-species virus virulence; bats, in particular, host viruses that cause higher human case fatality rates than zoonoses derived from other mammals and birds[4,5]. Understanding the mechanisms that select for the evolution of unique viral traits in bats compared to those selected in other mammalian reservoirs is essential to accurately predicting virulence for future zoonotic threats.

Although the disproportionate frequency with which the Chiropteran order may source viral zoonoses remains debated[6,7], the extraordinary human pathology induced by many bat-borne zoonoses—including Ebola and Marburg filoviruses, Hendra and Nipah henipaviruses, and SARS, MERS, and SARS-CoV-2 coronaviruses[8]—is not contested. Remarkably, bats demonstrate limited clinical pathology from infection with viruses that cause extreme morbidity and mortality in other hosts[8]. This may be due, in part, to unique bat resistance mechanisms, which include constitutive expression of antiviral cytokines in some species[9] and enhanced autophagy[10] and heat-shock protein expression in others[11]. While these protracted inflammatory responses would be pathological in most mammals, bats have evolved unique mechanisms of mitigating inflammation, including loss of PYHIN[12,13] and downregulation of NLRP3[14] inflammasome-forming gene families, dampened STING-dependent interferon activation[15], and diminished caspase-1 inflammatory signaling[16]. As an apparent by-product of these adaptations that enable flight, bats maintain long lifespans[17] and tolerate inflammation incurred by the recruitment of immune cells to viral infections[8]. The extent to which these unique features of bat immunology and physiology modulate the evolution of the viruses they host, however, remains largely unexplored.

General theory suggests that because viral 'tolerance' mitigates virulence (defined here as the virus-induced mortality rate) without reducing viral load, a host strategy of tolerance will select for higher growth rate pathogens that achieve gains in between-host transmission without causing damage to the original host[18,19] (**Fig. 1**). Bat viral tolerance thus predicts the evolution of enhanced viral growth rates, which—though avirulent to bats—may cause significant pathology upon spillover to hosts lacking bat immunology and physiology. Indeed, recent work demonstrates how high virus growth rates easily tolerated in bat cells cause significant pathology in cells derived from hosts lacking bat antiviral defenses[20]. While tolerance reduces pathogen-induced mortality for a single host, tolerant host populations, with limited checks on virus transmission, may demonstrate high pathogen prevalence. If imperfectly tolerant hosts still experience some virus-induced pathology, high prevalence can consequently elevate total population-level mortality for infected hosts—a phenomenon known as the 'tragedy of tolerance'[18]. Reports of virus-induced mortality in bats are rare[21], suggesting that bat virus tolerance is likely very effective. We explore the extent to which the immunological and life-history traits of mammalian reservoirs can explain variation in the virulence of zoonotic viruses emerging into human hosts.

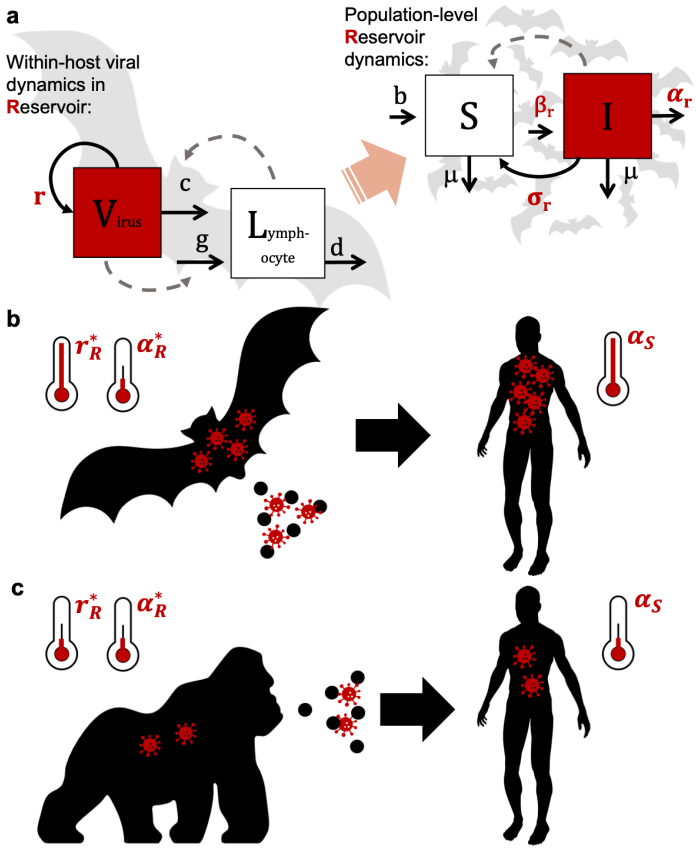


Figure 1. Conceptual mechanistic framework to predict zoonotic virus virulence from reservoir immunology and life history traits. (**A**) A within-host predator-prey model of lymphocyte-virus dynamics is embedded in a population-level transmission model for reservoir hosts. Between-host rates of transmission (β_r), virulence (α_r), and recovery (σ_r) are expressed as functions of within-host dynamics to derive optimal virus growth rates in the reservoir (r_R^*). (**B**) Virus growth rates (r_R^*) evolve to maximize gains in between-host transmission over virulence incurred on the reservoir (α_R^*). Unique bat immunology and physiology allows high r_R^* viruses to evolve without incurring virulence on the bat (α_R^*). Bat-optimized viruses are likely to generate pathology upon spillover (α_S) to secondary hosts (i.e. humans) lacking bat life history traits. (**C**) Viruses evolved in taxa with physiologies similar to humans (i.e. primates) should optimize at lower r_R^* values less likely to cause human pathology (α_S).

To elucidate how immunological and life-history traits of mammalian hosts combine to drive zoonotic virus virulence, we adopted a nested modeling approach, embedding a within-host model of viral and immune system dynamics within an epidemiological, population-level framework (**Fig. 1**). We examined how the life history traits of a primary reservoir drive the evolution of viral traits likely to cause pathology in a secondary, spillover host—chiefly, a human. Using our nested modeling framework, we first computed the optimal within-host virus growth rate (r^*), the corresponding virulence at that growth rate (α^*), the population-level infection prevalence, and the total population-level infected host mortality for a virus evolved under differing reservoir host-specific cellular and immunological parameter values (**Fig. 2**, Extended Data Fig. 1-2, Extended Data Table 1). Our analyses highlight several critical drivers of virus evolution likely to generate significant pathology following spillover to a secondary host (Fig. 2): higher within-host virus growth rates (r^*) are selected in reservoir hosts with higher background mortality (μ), more rapid virus clearance (c), and more rapid lymphocyte activation (g) upon infection. Higher $r^*$is also selected in host populations that exhibit higher ρ, corresponding to slower saturation of transmission per virion. Critically, higher tolerance of both

virus-induced pathology (T_v) and immunopathology (T_w) also select for higher r^*, but the effects of T_w are proportionally greater—suggesting that bat tolerance of immunopathology facilitates the evolution of virulent zoonoses. Of note, changes in the majority of within-host parameters drive corresponding increases in r^* and α^*, such that reservoir hosts evolving viruses with high growth rates also experience high pathogen-induced mortality. The rate of virus clearance by lymphocytes (c) and the two modeled mechanisms of tolerance (T_v and T_w), diverge in this regard, permitting evolution of high r^* viruses that incur minimal virulence (α^*) on reservoir hosts. By extension, we predict that viruses evolving optimal r^* in reservoir hosts with high c, T_v, or T_w might incur substantial pathology upon spillover to secondary hosts exhibiting lower values for these same parameters. Unique among mammals, at least two species of bat constitutively express the antiviral cytokine, IFN-α[9], which would manifest as elevated c in our model, and several sources indicate heightened tolerance of immunopathology in bats, modeled here as high T_w[12,15,16,22].

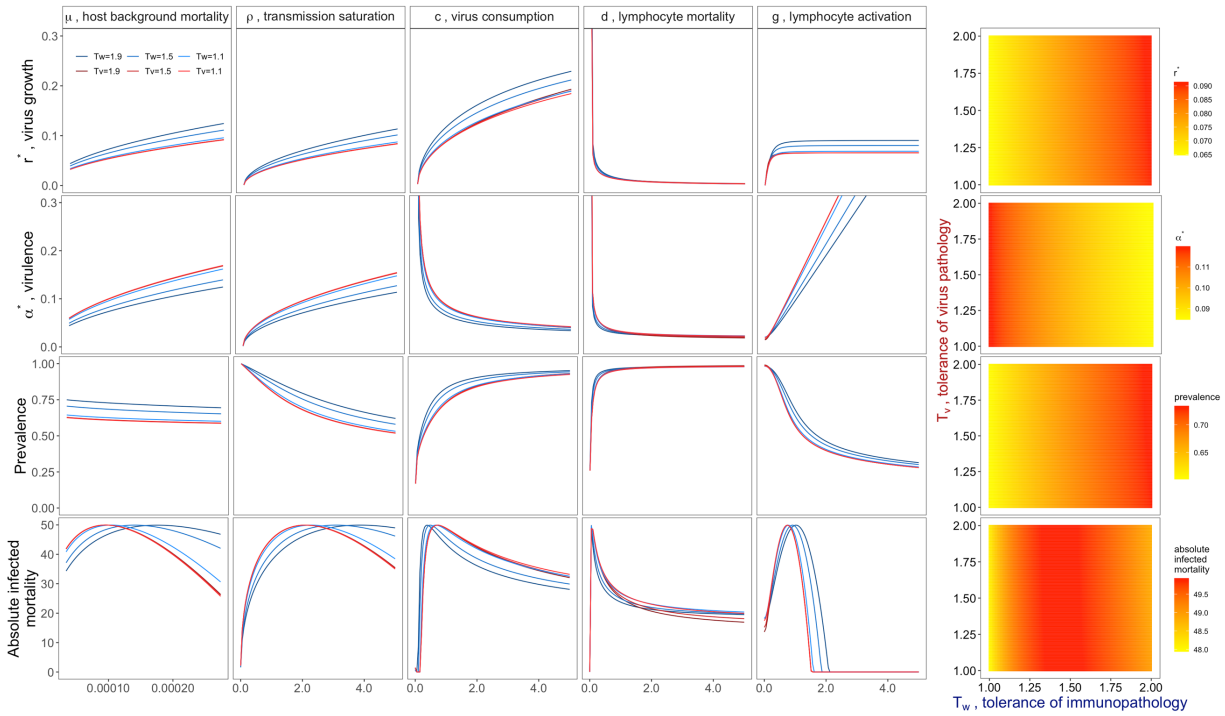


Figure 2. Optimal virus growth rates vary across reservoir host immunological and life history parameters. Rows (top-down) indicate the evolutionarily optimal within-host virus growth rate (r^*) and the corresponding virus-induced mortality rate (α^*), population-level infected prevalence, and population-level absolute infected mortality for a reservoir host infected with a virus at r^*. Columns demonstrate the dependency of these outcomes on variable reservoir host parameters: background mortality (μ), transmission saturation (ρ), virus consumption by lymphocytes (c), lymphocyte mortality (d), and lymphocyte activation upon viral contact (g). Darker colored lines depict outcomes at higher tolerance of virus pathology (T_v, red) or immunopathology (T_w, blue), assuming no tolerance of the opposing type. Heat maps demonstrate how T_v and T_w interact to produce each outcome. Parameter values can be viewed in Extended Data Table 1. Figure assumes tolerance in the 'constant' form (see Extended Data Fig. 2 for 'complete').

We next made broad predictions of the evolution of zoonotic virus virulence across eight mammalian reservoir orders, based on order-specific variation in four key parameters from our within-host nested model: the reservoir host background mortality rate (μ), the rate at which between-reservoir host transmission saturates with increasing viral load (ρ), the reservoir host

tolerance of immunopathology (T_w), and the spillover (human) host tolerance of pathology induced by a virus derived from a non-human reservoir (T_{vS}). Within-host immunological data needed to quantify these parameters is lacking across most mammalian orders; thus, we used order-specific life history traits to proxy these metrics for the eight mammalian orders of interest, while holding all other parameters constant across reservoir taxa (**Fig. 3**, Extended Data Fig. 3, Extended Data Tables 1-3). Host background mortality rates (μ) were quantified directly and summarized by mammalian order, but the lack of available cross-species immunological data in the literature necessitated significant assumptions for the other three parameters. We modeled ρ as a function of reservoir host population size: since wildlife viruses transmit via predominantly density-dependent mechanisms, large wildlife host populations offer more opportunities for onward transmission at a given viral load, meaning that population-level transmission rates will saturate more slowly with increasing virus density. We represented T_w as proportional to the residual of observed host lifespan (again, summarized across orders) from that predicted by basal metabolic rate (BMR): since longevity is linked to resilience to inflammation[23], we hypothesized that extraordinarily long-lived hosts would demonstrate tolerance of immunopathological inflammation, as well. Finally, we represented T_{vS} as proportional to the phylogenetic distance between a given reservoir host order and Primates, assuming that human immune systems would be better adapted to tolerate viruses evolved in phylogenetically-related hosts, as has been suggested in the literature[5,6]. Though assumptions, these parameter representations enabled us to compare the impact of broad reservoir host life-history trait variation on zoonotic virulence: among mammals considered, orders Eulipotyphla and Rodentia exhibited the fasted background mortality rates (μ); orders Primate, Chiroptera, and Carnivora demonstrated the longest lifespans per BMR, corresponding to higher tolerance of immunopathology in our model (T_w); orders Chiroptera and Cetartiodactyla had the largest group sizes, equating to slowest transmission rate saturation (high ρ); and order Diprotodontia was most distantly related to humans, resulting in low human tolerance of viral pathology (T_{vS}). High estimates for $T_w, \mu,$ and ρ and low estimates for T_{vS} all predict high spillover virulence to humans (α_S), and virulence rankings between orders are sensitive to the scaling of these within-host parameters. Nonetheless, bats demonstrate uniquely long lifespans per BMR and large group sizes; when combined, as in our analysis, to represent high T_w and high ρ, these reservoir host traits elevate predicted r^* and α_S beyond other orders. Thus, our approach successfully recovers the key result from the zoonosis literature: bat-derived zoonoses yield higher rates of virus-induced mortality upon spillover to humans (α_S) than do viruses derived from any other mammals (**Fig. 3d**, Extended Data Fig. 4)[24].

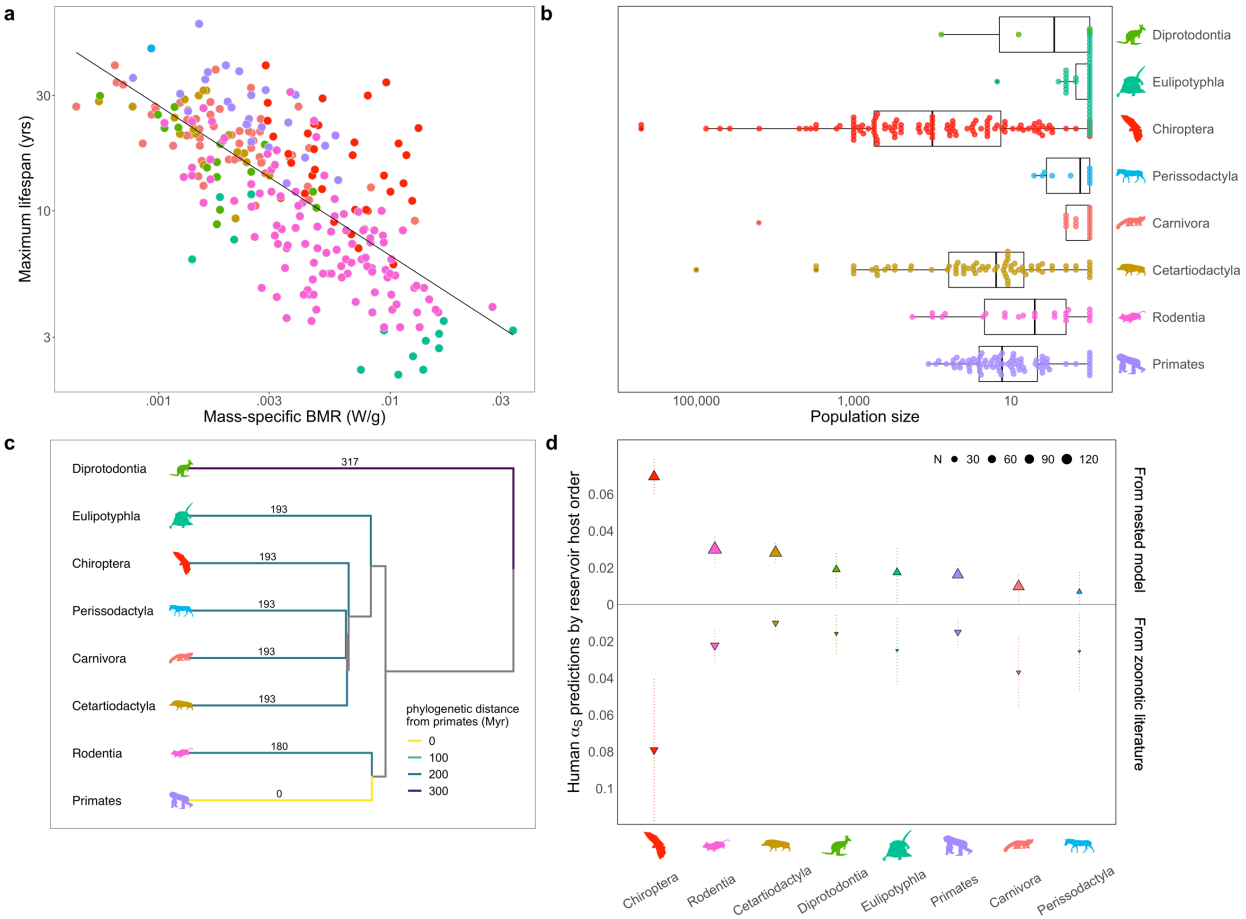


Figure 3. Reservoir host life history traits predict evolution of zoonotic virus virulence. (**a**) Regression of mammal maximum lifespan vs. mass-specific basal metabolic rate (Extended Data Table 3; data derived from[32]). Points are colored by mammalian order, corresponding to records of host population size in (**b**) (data derived from[33]) and phylogenetic distance from Primates (η_R, indicated by color and specified atop each branch) on a timescaled phylogeny in (**c**) (data derived from[34]). (**d**) Comparison of human spillover virulence (α_S) predictions for zoonoses from nested life history model with estimates from the literature. Top panel depicts human α_S predictions by order, derived from within-host model, using order-specific parameters for μ, T_w, ρ, T_{vS} computed from life history traits in (a), (b), and (c). Bottom panel depicts human α_S estimates derived from case fatality rates and infection duration reported in the zoonotic literature[4]. Point size corresponds to the average of the number of species-level datapoints (N) used to generate each of the four variable parameters impacting α_S for top panel and to the number of independent host-virus associations from which estimates were determined for bottom panel. 95% confidence intervals were computed by standard error; for top panel, these reflect the upper and lower confidence intervals of the optimal virus growth rate. Extended Data Table 2 lists order-level values for r^*, $\rho, \mu, T_w,$ and T_v; all other parameters involved in calculation of α_S are listed in Extended Data Table 1. Figure assumes 'constant' tolerance (see Extended Data Fig. 4 for 'complete').

Compared with the data[4], our model underpredicts cross-order comparative virulence rankings for Carnivora-derived viruses and overpredicts rankings for viruses derived from Rodentia and Cetartiodactyla. For order Carnivora, the discrepancy in our predictions vs. those from the literature is resolved when rabies lyssavirus is excluded from data comparisons (Extended Data Fig. 5)[4]: though most rabies zoonoses are sourced from domestic dogs, lyssaviruses are Chiropteran by origin[25], and viral traits responsible for rabies' virulence may reflect its bat evolutionary history more than that of any intermediate carnivore host. For order

Rodentia, comparatively lower spillover virulence reported in the literature suggests that attainment bias may be less extreme for this clade: rodent-borne hantaviruses and arenaviruses represent some of the best-studied zoonoses, and multiple avirulent species within these families are referenced in our dataset[4]. Records of avirulent zoonoses are rare for other taxa, excepting Primates, which are also well-studied, and Cetartiodactyla, for which existing data on zoonotic viruses are derived entirely from domesticated livestock[4]. Our model's overprediction of virus virulence from Cetartiodactyla squares with hypotheses that the low virulence of zoonoses recorded for this clade may reflect higher research effort and sampling rates for livestock pathogens (thus minimizing attainment bias)—or that centuries of human-livestock cohabitation and development of human immunity to livestock pathogens could have muted the virulence of Cetartiodactyla viruses in human hosts[4].

Our model provides a mechanistic explanation for the biological processes that underpin cross-species patterns in the evolution of virus virulence—a major advance for efforts to evaluate zoonotic risk. Of note, we are able to recapture broad patterns in the literature while holding constant many basic immunological parameters, including rates of virus consumption by lymphocytes (c) and lymphocyte activation (g), which almost certainly vary across taxa. By summarizing within-host traits across eight mammalian orders, we also generalize substantial within-clade diversity that likely contributes to heterogeneous patterns in available data. Bats alone make up more than 1,400 species and account for some 20% of mammalian diversity[26]; only a subset of bats are long-lived[27] or gregarious[28], suggesting considerable variation in immunopathological tolerance or transmission saturation rates—and consequently, the evolution of virus virulence—which we largely disregard here. As more data becomes available, modeling could be fine-tuned to make more specific, species-level predictions of highly virulent disease risk.

Currently, our work emphasizes the uniqueness of bats as flying mammals and, in consequence, as tolerant viral reservoirs. For the first time, we offer a mechanism for the evolution of bat-derived viruses, which demonstrate significant pathology upon spillover to non-bat, particularly human, hosts. In providing a theoretical framework to explain this phenomenon, we generate a series of testable questions and hypotheses for future comparative immunological studies, to be carried out at *in vitro* and *in vivo* scales. Empirical work should aim to determine whether virus growth rates are truly higher when evolved in bat immune systems, whether anti-inflammatory mechanisms in bat cells are equally effective at mitigating virus-induced pathology and immunopathology, and whether comparative taxonomic predictions of virus tolerance, resistance, and virulence evolution apply to non-viral pathogens, too. In light of the emergence of SARS-CoV-2, the field of bat molecular biology has echoed the call for mechanistic understanding of bat immunology[29], and the NIH has responded[30], soliciting research on the development of tools needed to test the predictions outlined here. We offer a bottom-up mechanistic framework enabling the prediction of emerging virus virulence from the basic immunological and life-history traits of zoonotic hosts.

255 **Main References**

256 1. Bonneaud, C. & Longdon, B. Emerging pathogen evolution. *EMBO Rep.* **21**, e51374 (2020).
257 2. André, J.-B. & Hochberg, M. E. Virulence evolution in emerging infectious diseases.
258 *Evolution (N. Y).* **59**, 1406 (2005).
259 3. Bolker, B. M., Nanda, A. & Shah, D. Transient virulence of emerging pathogens. *J. R.*
260 *Soc. Interface* **7**, 811–822 (2010).
261 4. Guth, S. E. *et al.* Bats host the most virulent--but not the most dangerous--zoonotic
262 viruses. *bioRxiv* (2021). doi:10.1101/2021.07.25.453574
263 5. Guth, S., Visher, E., Boots, M. & Brook, C. E. Host phylogenetic distance drives trends in
264 virus virulence and transmissibility across the animal – human interface. *Phil. Trans. R.*
265 *Soc B* **374**, 20190296 (2019).
266 6. Olival, K. J. *et al.* Host and viral traits predict zoonotic spillover from mammals. *Nature*
267 (2017). doi:10.1038/nature22975
268 7. Mollentze, N. & Streicker, D. G. Viral zoonotic risk is homogenous among taxonomic
269 orders of mammalian and avian reservoir hosts. *Proc. Natl. Acad. Sci.* 1–8 (2020).
270 doi:10.1073/pnas.1919176117
271 8. Irving, A. T., Ahn, M., Goh, G., Anderson, D. E. & Wang, L.-F. Lessons from the host
272 defences of bats, a unique viral reservoir. *Nature* **589**, 363–370 (2021).
273 9. Zhou, P. *et al.* Contraction of the type I IFN locus and unusual constitutive expression of
274 IFN-α in bats. *Proc. Natl. Acad. Sci.* **113**, 2696–2701 (2016).
275 10. Laing, E. D. *et al.* Enhanced autophagy contributes to reduced viral infection in black
276 flying fox cells. *Viruses* **11**, (2019).
277 11. Chionh, Y. T. *et al.* High basal heat-shock protein expression in bats confers resistance to
278 cellular heat/oxidative stress. *Cell Stress Chaperones* **24**, 835–849 (2019).
279 12. Zhang, G. *et al.* Comparative analysis of bat genomes provides insight into the evolution
280 of flight and immunity. *Science* **339**, 456–60 (2013).
281 13. Santillan, D. M. *et al.* Large-scale genome sampling reveals unique immunity and
282 metabolic adaptations in bats. *Mol. Ecol.* (2021). doi:10.1111/mec.16027
283 14. Ahn, M. *et al.* Dampened NLRP3-mediated inflammation in bats and implications for a
284 special viral reservoir host. *Nat. Microbiol.* **4**, (2019).
285 15. Xie, J. *et al.* Dampened STING-dependent interferon activation in bats. *Cell Host Microbe*
286 **23**, 1–5 (2018).
287 16. Goh, G. *et al.* Complementary regulation of caspase-1 and IL-1 β reveals additional
288 mechanisms of dampened inflammation in bats. *Proc. Natl. Acad. Sci.* (2020).
289 doi:10.1073/pnas.2003352117
290 17. Austad, S. N. & Fischer, K. E. Mammalian aging, metabolism, and ecology: evidence
291 from the bats and marsupials. *J. Gerontol.* **46**, B47-53 (1991).
292 18. Miller, M. R., White, A. & Boots, M. The evolution of parasites in response to tolerance
293 in their hosts: The good, the bad, and apparent commensalism. *Evolution (N. Y).* **60**, 945
294 (2006).
295 19. Anderson, R. M. & May, R. M. Coevolution of hosts and parasites. *Parasitology* **85**, 411–
296 426 (1982).
297 20. Brook, C. E. *et al.* Accelerated viral dynamics in bat cell lines, with implications for
298 zoonotic emergence. *eLife* **9**, e48401 (2020).
299 21. Brook, C. E. & Dobson, A. P. Bats as 'special' reservoirs for emerging zoonotic
300 pathogens. *Trends Microbiol.* **23**, 172–180 (2015).
301 22. Pavlovich, S. S. *et al.* The Egyptian Rousette genome reveals unexpected features of bat

antiviral immunity. *Cell* **173**, 1–13 (2018).
23. Caruso, C., Lio, D. & Cavallone, L. Aging, longevity, inflammation, and cancer. *Ann. New York Acad. Sci.* **1028**, 1–13 (2004).
24. Visher, E. *et al.* The three Ts of virulence evolution during zoonotic emergence. *Proc. R. Soc. B Biol. Sci.* **288**, 20210900 (2021).
25. Rupprecht, C. E., Turmelle, A. & Kuzmin, I. V. A perspective on lyssavirus emergence and perpetuation. *Curr. Opin. Virol.* **1**, 662–670 (2011).
26. Simmons, N. & Cirranello, A. Bat species of the world: A taxonomic and geographic database. (2020). Available at: https://batnames.org/. (Accessed: 11th July 2020)
27. Wilkinson, G. S. & Adams, D. M. Recurrent evolution of extreme longevity in bats. *Biol. Lett.* **15**, 20180860 (2019).
28. Luis, A. D. *et al.* Network analysis of host – virus communities in bats and rodents reveals determinants of cross-species transmission. *Ecol. Lett.* (2015). doi:10.1111/ele.12491
29. Wang, L.-F., Gamage, A. M. & Chan, W. O. Y. Decoding bat immunity: the need for a coordinated research approach. *Nat. Rev. Immunol.* **0123456789**, 0–2 (2021).
30. National Institutes of Health. Research on Bat Immunology. Available at: https://grants.nih.gov/grants/guide/pa-files/PAR-21-184.html.
31. Boots, M., Donnelly, R. & White, A. Optimal immune defence in the light of variation in lifespan. *Parasite Immunol.* **35**, 331–338 (2013).
32. Healy, K. *et al.* Ecology and mode-of-life explain lifespan variation in birds and mammals. *Proc. R. Soc. B* (2014).
33. Jones, K. E. *et al.* PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90**, 2648–2648 (2009).
34. Kumar, S., Stecher, G., Suleski, M. & Hedges, S. B. TimeTree: a resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017).

Methods

1. Population-level dynamics and R_0 under SIS assumptions.

To begin, we follow Miller et al. (2006)[18], assuming a simple population dynamics system which allows for both infection-induced mortality and recovery from infection. Since we are interested in examining virus evolution in reservoir hosts, we assume that hosts maintain long infectious periods, before eventually recovering to return to the susceptible class.

We assume that hosts are born at rate b and die of natural death at rate μ, where $b > \mu$. All hosts are assumed to be born susceptible and population density is regulated via a crowding term (q). Because our goal in this endeavor is to allow for the evolution of rates that link to within-host dynamics (Section 4), we represent transmission, virulence, and recovery as functions of the virus causing the infection (respectively, β_r, α_r and σ_r), where r denotes the intrinsic virus growth rate. Our system is described by the following system of Ordinary Differential Equations, where $N = S + I$.

$$\frac{dS}{dt} = N(b - qN) - \beta_r SI - \mu S + \sigma_r I \qquad [1a]$$

$$\frac{dI}{dt} = \beta_r SI - \mu I - \alpha_r I - \sigma_r I \qquad [1b]$$

Solving the equilibria of this system establishes that there is both a disease-free equilibrium (DFE) at $(I^*_{DFE}, S^*_{DFE}) = (0, \frac{b-\mu}{q})$, as well as an endemic infection equilibrium, (I^*_{EE}, S^*_{EE}) (see below, *[2a], [2b]*).

$$S^*_{EE} = \frac{\alpha_r + \mu + \sigma_r}{\beta_r} \qquad [2a]$$

$$I^*_{EE} = \frac{\sqrt{\beta_r[\beta_r(b-\mu-\alpha_r)^2 + 4\,\alpha_r q(\mu+\alpha_r+\sigma_r)]} + \beta_r(b-\mu-\alpha_r) - 2q(\mu+\alpha_r+\sigma_r)}{2\,\beta_r q} \qquad [2b]$$

We study the dynamics of *[1]* in the following feasibility region:

$$\Omega = \{(S, I) \in R^2_+ \,|S \geq 0, I \geq 0, S + I \leq \frac{b-\mu}{q}\}$$

This region is positively invariant with respect to *[1]*. From *[1]*, we can then formulate the Jacobian matrix, J, to decipher the local stability of the DFE:

$$J = \begin{bmatrix} b - 2qS - 2qI - \beta_r I - \mu & b - 2qS - 2qI - \beta_r S + \sigma_r \\ \beta_r I & \beta_r S - \mu - \sigma_r - \alpha_r \end{bmatrix} \qquad [3a]$$

Next, we evaluate the Jacobian at the DFE:

$$J_{DFE} = \begin{bmatrix} -b - \mu & -b + 2\mu - \frac{\beta_r(b-\mu)}{q} + \sigma_r \\ 0 & \frac{\beta_r(b-\mu)}{q} - (\mu + \sigma_r + \alpha_r) \end{bmatrix} \qquad [3b]$$

The DFE is locally unstable when at least one of the eigenvalues of the Jacobian is non-negative. The eigenvalues can be read off the diagonal (J_{DFE} is upper triangular). Clearly, the first eigenvalue is negative: $-b - \mu < 0$. A little algebra will establish that the second eigenvalue is negative when:

$$R_0 = \frac{\beta_r(b-\mu)}{q(\mu+\sigma_r+\alpha_r)} < 1 \qquad [4]$$

Therefore, if $R_0 > 1$, the DFE is no longer locally asymptotically stable, resulting in an increase in infected hosts within the population.

When $R_0 > 1$, the EE is globally asymptotically stable in the interior of Ω. We demonstrate this by first establishing the local asymptotic stability of the EE, then by determining that there are no limit cycles in the region, Ω.

If we evaluate the Jacobian, once again, at the EE, we will find that the trace *(Trace(J_{EE}))* of that matrix is negative:

$$Trace(J_{EE}) = \frac{-(2q+\beta_r)\sqrt{\beta_r(\beta_r(b-\mu-\alpha_r)^2+4\alpha q(\mu+\alpha_r+\sigma_r)}-\beta_r(\beta_r(b-\mu-\alpha_r)-2q(\mu+2\alpha_r+\sigma_r)}{2q\beta_r} < 0 \qquad [5]$$

Note that $b > \mu + \alpha_r$ (a necessary condition for the existence of the EE). In addition, the determinant of the J_{EE} is positive:

$$Det(J_{EE}) = \frac{\sqrt{\beta(\beta(b-\mu-\alpha)^2+4\alpha q(\mu+\alpha+\sigma))}(1-2q(\mu+\alpha+\sigma))+\beta(b-\mu-\alpha)}{2q\beta} > 0 \qquad [6]$$

Note that $q \ll 1$.

Therefore, the eigenvalues of the Jacobian matrix (J_{EE}) have negative real parts, and therefore the EE (when it exists) is locally asymptotically stable.

We now use the Dulac criterion to establish that no limit cycles exist in Ω. Consider the Dulac function: $\varphi(S,I) = 1/(\beta SI)$, and let the function *F(S,I)* represent the right hand side of *[1a]*, and let the function *G(S,I)* represent the right hand side of *[1b]*. Then we have:

$$\frac{\partial(\Phi F)}{\partial S} + \frac{\partial(\Phi G)}{\partial I} = -\frac{(b+\sigma)I}{\beta S^2 I} + q\frac{I^2-S^2}{\beta S^2 I} < 0 \qquad [7]$$

since $q \ll 1$. Thus, system *[1]* has no limit cycles present in Ω. Therefore, the EE is globally asymptotically stable in Ω when $R_0 > 1$.

<u>2. Invasion of a mutant parasite under SIS assumptions.</u>

Next, we aim to establish under what conditions a mutant pathogen can invade this system. We accomplish this task by first re-writing our system *[1]* to include hosts infected with a resident wild-type pathogen, I_1, or a novel mutant pathogen, I_2 where, now, $N = S + I_1 + I_2$:

$$\frac{dS}{dt} = N(b - qN) - \beta_{r1}SI_1 - \beta_{r2}SI_2 - \mu S + \sigma_{r1}I_1 + \sigma_{r2}I_2 \qquad [8a]$$

$$\frac{dI_1}{dt} = \beta_{r1}SI_1 - \mu I_1 - \alpha_{r1}I_1 - \sigma_{r1}I_1 \qquad [8b]$$

$$\frac{dI_2}{dt} = \beta_{r2}SI_2 - \mu I_2 - \alpha_{r2}I_2 - \sigma_{r2}I_2 \qquad [8c]$$

Notice that the novel mutant pathogen has a different intrinsic viral growth rate than that of the resident (wild-type) pathogen ($r_1 \neq r_2$). It then follows from *[1]* that there exists a mutant-free equilibrium (MFE) which recapitulates the endemic equilibrium in system *[1]*. To evaluate the local stability of the MFE we can again construct the Jacobian matrix of the new system, and evaluate the Jacobian at the MFE (Note: at the MFE $I_2^* = 0$ while $S^*$and I_1^* replicate the endemic equilibrium values in *[2a]* and *[2b]*):

$$J_{MFE} = \begin{bmatrix} b-2qS-2qI_1-\beta_{r1}I_1-\mu & b-2qS-2qI_1-\beta_{r1}S+\sigma r_1 & b-2qS-2qI_1-\beta_{r2}+\sigma r_2 \\ \beta_{r1}I_1 & \beta_{r1}S-\mu-\sigma r_1-\alpha_{r1} & 0 \\ 0 & 0 & \beta_{r2}S-\mu-\sigma_{r2}-\alpha_{r2} \end{bmatrix} \qquad [9]$$

Notice that the upper left four quadrats of this matrix exactly replicate the 2x2 Jacobian (*[3a]*) evaluated at the EE, which we now refer to as the resident equilibrium. J_{MFE} is a block-triangular matrix, which we can break down into component parts:

$$J = \begin{bmatrix} J_{res} & V \\ 0 & J_{mut} \end{bmatrix} \quad [10]$$

Stability of the MFE requires that all eigenvalues of the Jacobian, evaluated at the MFE, must be negative. Since the matrix is block diagonal, the eigenvalues of the Jacobian are given by the eigenvalues of the two matrices on the diagonal, J_{res} and J_{mut}. Since we have already established that the EE is globally asymptotically stable when $R_0 > 1$, we know that the eigenvalues of J_{res} under this condition must be negative. The local stability of the MFE then depends on the sign of the eigenvalue of J_{mut}, where $J_{MFE} = \beta_2 S_{EE}^* - \mu - \sigma_2 - \alpha_2$. J_{mut} thus corresponds to the growth rate of a rare mutant in an environment with an endemic resident strain. If the mutant pathogen is able to grow within the population, then:

$$\beta_{r2}\frac{\mu+\sigma_{r1}+\alpha_{r1}}{\beta_{r1}} - \mu - \sigma_{r2} - \alpha_{r2} > 0 \quad [11]$$

We can rearrange terms to include mutant terms on one side of the equation and resident terms on the other:

$$\frac{\beta_{r2}}{\mu+\sigma_{r2}+\alpha_{r2}} > \frac{\beta_{r1}}{\mu+\sigma_{r1}+\alpha_{r1}} \quad [12]$$

It is evident that the MFE is unstable (allowing for invasion of the mutant strain) when the ratio of transmission over the duration of infection is larger for the mutant strain than the resident strain. In general, a pathogen will evolve to maximize this expression, which is known as the invasion fitness, and is given by:

$$\frac{\beta_r}{\mu+\sigma_r+\alpha_r} \quad [13]$$

3. Re-evaluating the invasion fitness under SIR assumptions.

We next re-evaluate the population-level dynamics shown in Sections 1 and 2 above, under SIR assumptions to show that different representations of the population-level dynamics will not affect the selection gradient on our within-host parameters in Section 4. Assuming SIR dynamics, the system in the absence of a mutant takes the following form, where $N = S + I + R$:

$$\frac{dS}{dt} = N(b - qN) - \beta_r SI - \mu S \quad [14a]$$

$$\frac{dI}{dt} = \beta_r SI - \mu I - \alpha_r I - \sigma_r I \quad [14b]$$

$$\frac{dR}{dt} = \sigma_r I - \mu R \quad [14c]$$

We proceed as we did in Sections 1-2. Likewise the EE in model *[14]* is globally asymptotically stable when $R_0 > 1$, and we can introduce a mutant pathogen as we did in *[8]* to evaluate the local stability of the MFE. Doing so, we find that the invasion fitness remains unchanged between SIS and SIR model assumptions:

$$\frac{\beta_{r2}}{\mu+\sigma_{r2}+\alpha_{r2}} > \frac{\beta_{r1}}{\mu+\sigma_{r1}+\alpha_{r1}} \quad [15]$$

Therefore, just as before, the evolutionary stable pathogen, or the strain that will avoid invasion by any other strain, is the strain that maximizes:

$$\frac{\beta_r}{\mu+\sigma_r+\alpha_r} \quad [16]$$

4. Within-host dynamics.

Now that we have determined the invasion fitness for a mutant virus, we can write the population-level rates dependent on within-host dynamics in their within-host terms (transmission: $\beta_{r1/r2}$, virulence: $\alpha_{r1/r2}$, and recovery: $\sigma_{r1/r2}$). This will allow us to derive a selection gradient on the viral growth rate, an intrinsic virus property likely to be carried from one host to another. To do this, we establish a simple within-host model that represents the

dynamics of infection *within* each I_1 and I_2 host as outlined above. We assume the simplest Lotka-Volterra predator-prey relationship between the virus population (V) and the lymphocyte population (L) within each infected host:

$$\frac{dV}{dt} = rV\left(1 - \frac{V}{K}\right) - cVL \qquad [17a]$$

$$\frac{dL}{dt} = gVL - dL \qquad [17b]$$

Where r corresponds to the intrinsic virus growth rate, K gives the carrying capacity for virus within its reservoir host, c corresponds to the attack efficacy of the immune system upon contact with the virus, and g signifies the recruitment rate of immune cells with virus density. The parameter d gives the natural lymphocyte death rate.

This system has three equilibria: a virus/immune cell free trivial equilibrium at $V^* = 0$ and $L^* = 0$, a virus at carrying capacity/immune cell free trivial equilibrium at $V^* = K$ and $L^* = 0$, and a non-trivial endemic virus equilibrium at $V^* = \frac{d}{g}$ and $L^* = \frac{r(gK-d)}{cgK}$. Assuming $K \gg c, g, d, r$, then the non-trivial endemic virus equilibrium is a stable spiral (through a linear stability analysis).

Building from above, we can then rewrite the component terms of the population-level transmission, virulence, and recovery rates dependent on within-host dynamics ($\beta_{r1/r2}$, $\alpha_{r1/r2}$, and $\sigma_{r1/r2}$) in terms of their within-host components, assuming that the within host dynamics are fast relative to the population-level dynamics.

In line with previous work, we assume transmission to be a bounded increasing function of the virus growth rate[18,19,35,36], which we represent as:

$$\beta_r = \frac{\zeta r V^*}{rV^* + \rho} \qquad [18]$$

where ζ is the max transmission rate and ρ corresponds to the rate at which transmission approaches ζ as the virus growth rate (rV^*) increases (smaller values of ρ correspond to faster rates of transmission saturation).

We can also assume that:

$$\alpha_r = \frac{vrV^*}{T_v} + \frac{wgL^*}{T_w} \qquad [19]$$

by which infection-induced host mortality ('virulence') results from both virus-induced pathology, a function of the intrinsic virulence of the parasite, v, and the parasite growth rate rV^*, and immunopathology, which we model as proportional to the growth rate of lymphocytes, gL^*, upon infection, multiplied by the host's susceptibility to immunopathology and inflammation, given as w. The terms T_v and T_w correspond to host tolerance of virus-induced pathology and immunopathology, respectively.

Expression *[19]* assumes a form of "constant" tolerance, by which both virus pathology and immunopathology are reduced by a constant proportion across the course of infection. After Miller et al. 2006[18], we can also model these relationships assuming "complete tolerance" *[20]*, whereby virus pathology and immunopathology are completely eliminated up to a threshold value, beyond which pathology scales proportionally with virus and immune cell growth.

$$\alpha_r = (v - T_v)rV^* + (w - T_w)gL^* \qquad [20]$$

We consider both constant and complete forms of tolerance in all subsequent analyses. It is important to note that the predictions of the complete form of tolerance are not dependent on the assumption of complete elimination: a more gradual function shows equivalent results [18].

Lastly, we can represent the recovery rate from infection σ_r as proportional to the equilibrium immune cell population. This assumes the following expression:

$$\sigma_r = \gamma L^* \qquad [21]$$

where γ scales the rapidity of recovery with the immune cell density.

We can then allow the virus to evolve its within-host growth rate (r) to maximize its reproductive success at the invasion fitness (*[13]* for the SIS model and *[16]* for the SIR model, above). In within-host terms, the invasion fitness takes on the following general form, under assumptions of constant tolerance:

$$\frac{cdgK\zeta rT_vT_w}{(dr+g\rho)[cKT_w(g\mu T_v+drv)-rT_v(d-gK)(\gamma T_w+gw)]} \qquad [22]$$

And under assumptions of complete tolerance:

$$\frac{cdgKr\zeta}{(dr+g\rho)[cK(g\mu+dr(v-T_v))-r(d-gK)(\gamma+g(w-T_w))]} \qquad [23]$$

<u>5. Selection gradient on within-host virus growth rate.</u>

We can then determine the selection gradient on r by differentiating *[22]* and *[23]* with respect to r. This gives us the following expression under assumptions of constant tolerance (differentiating *[22]*):

$$\frac{cdgKT_vT_w[cKg^2\mu\rho T_vT_w+d^2r^2(\gamma T_vT_w-cKvT_w+gwT_v)-dKgr^2T_v(\gamma T_w+gw)]}{(dr+g\rho)^2[cKT_w(g\mu T_v+drv)-rT_v(d-gK)(\gamma T_w+gw)]^2} \qquad [24]$$

And assuming complete tolerance (differentiating *[23]*):

$$\frac{cdgK\zeta[cKg^2\mu\rho+dKgr^2(g(T_w-w)-\gamma)+d^2r^2(\gamma+cK(T_v-v)+g(w-T_w))]}{(dr+g\rho)^2[cK(g\mu+dr(v-T_v))-r(d-gK)(\gamma+g(w-T_w))]^2} \qquad [25]$$

We then set these derivatives equal to 0 and solve for *r*, which we now refer to as r^*. r^* corresponds to the optimal within-host virus growth rate, here under assumptions of constant tolerance:

$$r^* = g\sqrt{\frac{c\mu\rho KT_vT_w}{d[gKT_v(\gamma T_w+gw)-d(\gamma T_vT_w-cKT_wv+gT_vw)]}} \qquad [26]$$

And here under assumptions of complete tolerance:

$$r^* = g\sqrt{\frac{c\mu\rho K}{d[gK(\gamma+g(w-T_w))-d(\gamma+cK(T_v-v)+g(w-T_w))]}} \qquad [27]$$

We can check that this optimum is a locally stable maximum (not a minimum) by proving that the second derivative of the invasion fitness (*[22]* and *[23]*) with respect to r is negative at the respective values for r^* derived in *[26]* and *[27]*. When we do this, we find that this is, indeed, true, under assumptions of both constant and complete tolerance. This means that *[26]* and *[27]* represent true evolutionarily stable strategies (ESS).

Furthermore, we can compute a pairwise invasibility plot to visually demonstrate that the derived expression for r^* represents an *ESS* (Extended Data Fig. 1).

In Extended Data Fig. 1, we see that all points along the vertical line through the point of intersection fall within the unshaded regions, corresponding to the conditions for a locally stable

ESS[37]. Additionally, we see that higher degrees of tolerance shift r^* higher under both constant and complete tolerance assumptions.

Finally, we can express *[19]* and *[20]* in within-host terms to calculate the virulence which a virus incurs on the reservoir host when evolved to its optimal within-host virus growth rate (r^*). When we do this, we find that virus-induced virulence at r^* (which we call α^*) can be expressed as the following, under assumptions of complete tolerance:

$$\alpha^* = \frac{\sqrt{\mu\rho}[cKd(v-T_v)+g(d-gk)(T_w-w)]}{\sqrt{cdK[gK(\gamma+g(w-T_w))-d(\gamma+cK(T_v-v)+g(w-T_w))]}} \qquad [28]$$

And the following, under assumptions of constant tolerance:

$$\alpha^* = \frac{\sqrt{\mu\rho}(cdKT_w v+gwT_v(gK-d))}{\sqrt{cdKT_vT_w[gKT_v(\gamma T_w+gw)-d(\gamma T_vT_w-cKvT_w+gwT_v)]}} \qquad [29]$$

There are several important insights to come out of these simple expressions for optimal viral growth r^* and virulence α^* in a reservoir host. We highlight the relationships between r^* and α^* and specific within-host parameters under assumptions of constant tolerance in Fig. 2 and under assumptions of complete tolerance in Extended Data Fig. 2.

<u>6. Spillover virulence in the secondary host.</u>

Our primary interest in understanding the evolution of the optimal virus growth rate in a reservoir host is to predict the virulence that a virus is likely to cause upon spillover to a novel, particularly human, host. Assuming a virus evolves to an optimal growth rate in its reservoir host and then emerges as a zoonosis in a human, we can adapt *[19]* to describe the "spillover virulence", using the following expression, here under assumptions of constant tolerance:

$$\alpha_S = \frac{r_R^*[c_S d_S v K_S T_{wS}+g_S w_S(g_S K_S-d_S)T_{vS}(\eta_R)]}{c_S g_S K_S T_{wS} T_{vS}(\eta_R)} \qquad [30]$$

where α_S corresponds to virulence incurred by the virus in a secondary, spillover host. In this expression, the within-host parameters of immune cell death rate d_S, growth rate g_S, and clearance rate c_S, as well as tolerance to immunopathology, T_{wS},are expressed as features of spillover host physiology (indicated by subscript "*S*"), while the intrinsic virus virulence, v, remains unchanged from one host to the next. By contrast, we represent the virus growth rate as optimized on the reservoir host (at least for recent spillovers), now called r_R^* to clarify its selection on the reservoir (indicated by subscript "*R*"). Finally, we model spillover host tolerance of virus-induced pathology (T_{vS}) as a function of phylogenetic distance between the reservoir and secondary host, here given as η_R. This term echoes recent work demonstrating increasing zoonotic virus virulence with phylogenetic distance between reservoir hosts and humans[5].

The "spillover virulence" remains largely unchanged under assumptions of complete tolerance:

$$\alpha_S = \frac{r_R^*[c_S d_S K_S(v-T_{vS}(\eta_R))+g_S(g_S K_S-d_S)(w_S-T_{wS})]}{c_S g_S K_S} \qquad [31]$$

After calculating spillover virulence, we then follow Day 2002[38] to relate virulence to case fatality rates in the human host, as most widely reported in the literature, following the equation:

$$CFR_S = \frac{\alpha_S}{\alpha_S+\sigma_S} \qquad [32]$$

where σ_S corresponds to the recovery rate from the spilled-over-virus in the human host. While we previously modeled reservoir host recovery rate as evolved to an optimum that responded to the within-host virus growth rate (as in *[18]*), this will not be the case for infections of recent spillover hosts. More likely, spillover recovery will scale inversely with $r_R{}^*$. When linking to data (Section 7), we control for this variation by accounting for total duration of infection, D_{IS}:

$$D_{IS} = \frac{1}{\alpha_S + \sigma_S} \quad [33]$$

From *[32]* and *[33]*, we can then derive expressions for spillover virulence (α_S) and spillover recovery (σ_S), in terms of data commonly reported in the literature (CFR_S and D_I):

$$\alpha_S = \frac{CFR_S}{D_{IS}} \quad [34]$$

$$\sigma_S = \frac{CFR_S - 1}{D_{IS}} \quad [35]$$

<u>7. Linking to the data to establish order-level parameters.</u>

In Fig. 3, we illustrate within-host parameters in possible mammalian reservoir species which could influence the evolution of optimal virus growth rates in the native reservoir host (r_R^*) and the corresponding virulence upon spillover to a secondary, human host (α_S). In the case of many of these parameters, we currently lack knowledge of these features across a majority of mammals; in particular, to our knowledge, no standardized database of mammalian immune cell parameters (i.e. g, c, and d) exists.

As such, we here focus on four parameters for which some data are available in the literature: the reservoir host mortality rate (μ), the reservoir host tolerance of immunopathology (T_w), the reservoir host transmission saturation rate per virion (ρ), and the phylogenetic distance between the reservoir host and human spillover host (η_R), which, for zoonoses, could influence human tolerance of an animal-derived virus (T_{vS}). Fig. 3 captures both natural host mortality (μ) and tolerance of immunopathology (T_w, the scaled order effect size on the regression residuals) in panel a, the reservoir host population size (a proxy for ρ) in panel b, and the reservoir host-human phylogenetic distance by mammalian order (a proxy for T_{vS}) in panel c. We summarize data compiled in Fig. 3 into order-level predictions for the within-host parameters, μ, T_w, ρ, and T_{vS}, in Extended Data Fig. 3 and Extended Data Tables 2-3.

To generate order-level summary terms for μ, we simply took the inverse of host maximum lifespan across all species, then applied a generalized additive model[39] to the raw data, modeling mammalian order as a random effect in a single smoothing term (Extended Data Table 3). The order-level prediction for μ is then given as the result of the predict.gam function applied to the fitted model (Extended Data Fig. 3).

To generate order-level summary terms for T_w, we fit a linear mixed effect model to the $\log_{10}$ of maximum lifespan as a response variable with the $\log_{10}$ of BMR as the predictor variable, incorporating a random effect of host order (Extended Data Fig. 3, Extended Data Table 3)[40]. To maintain the magnitude of differences among orders, we permitted only linear transformations of resulting data. Thus, to obtain order-level estimates of T_w on par with those explored in Fig. 2, we added the absolute value of the lower confidence limit of the minimum estimated partial effect size per order to all partial effect sizes to obtain complete tolerance estimates for T_w (ranging from 0 to 1), such that the least tolerant orders assumed values for T_w close to 0 and the most tolerant orders assumed values for T_w close to 1. For constant tolerance assumptions, we simply added 1 to all estimates, such that T_w ranged from close to 1 for poor tolerance to close to 2 for highly tolerant hosts.

To generate order-level summary terms for ρ, we fit another generalized additive model to the response variable of $\log_{10}$ of the reservoir host population size (from Fig. 3b), modeling

642 mammalian order as a random effect in a single smoothing term (Extended Data Fig. 3/Table 3).
643 We then generated an order-level prediction for the reservoir host population size by applying
644 the predict.gam function to the resulting fitted model. We converted this order-level prediction of
645 group size to an estimate of ρ on par with values explored in Fig. 2 by dividing all group size
646 predictions by the upper confidence limit of the maximum estimated group size x100. This
647 resulted in the prediction that reservoir hosts with larger base population sizes had larger values
648 for ρ and a wider range of transmission capacities per virion. These relationships are consistent
649 with our assumptions of density-dependent transmission rates at the population level: more
650 virions generate more infections in large, gregarious host populations.
651
652 Using these order-level summary terms, we then generated an order-level prediction for r_R^*
653 across all mammalian host orders, following *[26]* and *[27]* for, respectively, assumptions of
654 constant and complete tolerance. All other parameters (c, g, d, γ, w, v, and T_v) were held
655 constant across all taxa at values listed in Extended Data Table 1.
656
657 Once we had constructed an order-level prediction for r_R^*, we finally sought to explore the effects
658 of these reservoir-evolved viruses upon spillover to humans, after *[30]* and *[31]*. As with the
659 reservoir host taxa, we held immunological parameters, c, g, d, and w, constant in humans (at
660 the same values listed above) due to a lack of informative data to the contrary. Then, to generate
661 an order-level estimate for virus virulence on humans (α_S), we combined our order-level
662 predictions for r_R^* with order-specific values for T_{vS}, the human tolerance of an animal-derived
663 virus, which we scaled such that more closely-related orders evolved more easily-tolerable
664 viruses for humans. Specifically, we represented T_{vS} as the scaled inverse of η_R, the cophenetic
665 phylogenetic distance of each mammalian order from Primates. No summarizing was needed for
666 η_R because, following Mollentze and Streicker 2020[7], we used a composite time-scaled reservoir
667 phylogeny derived from the TimeTree database, which estimated a single mean divergence date
668 for all clades. Thus, all species within a given mammalian order were assigned identical
669 divergence times from the order Primates. To convert η_R to reasonable values for T_{vS}, we
670 divided all order-level values of η_R by the largest (to generate a fraction), then subtracted that
671 fraction from 1 for assumptions of complete tolerance and from 2 for assumptions of constant
672 tolerance.
673
674 See github.com/carabrook/spillover-virulence for detailed code documenting parameter
675 derivation for $\mu, T_w, \rho,$ and T_{vS}.
676
677 ***<u>8. Comparing spillover virulence predictions with estimates from the literature</u>***
678 To compare estimates of spillover virulence generated from our nested modeling approach (α_S)
679 with estimates from the literature, we next followed *[32]-[35]* to obtain empirical estimates of α_S
680 for each mammalian order, using case fatality rates (CFR_S) reported in the literature[4] and
681 durations of human infection (D_{IS}) which we collected ourselves by searching the primary
682 literature. Infection durations and associated references are listed in the dataset hosted on our
683 publicly available GitHub repository: github.com/carabrook/spillover-virulence. To obtain a
684 composite order-level prediction for α_S from D_{IS} and CFR_S as reported in[4], we adopted the best-
685 fit generalized additive model used by the authors in[4] to estimate CFR_S to here estimate α_S
686 instead (Extended Data Table 3), then predicted α_S from the fitted model while excluding the
687 predictive effects of viral family.
688

In Fig. 3d, we then compare predictions of α_S in humans derived from our nested modeling approach with empirical estimates from the zoonosis literature in[4], under assumptions of constant tolerance. Constant tolerance offers a reliable baseline assumption under which virulence is proportionally reduced by a constant extent. From this comparison, we see that both methods predict the highest values for (α_S) by far from bats (Order Chiroptera), a pattern recapitulated under assumptions of complete tolerance in Extended Data Fig. 4. Our life history method appears to overpredict α_S from the orders Rodentia and Cetartiodactyla, relative to the data in Guth et al. (2021)[4]. In Extended Data Fig. 5, we compare our order-level predictions of α_S with a version recomputed from[4] where rabies lyssavirus is excluded from the dataset. This greatly reduces the estimated order-level average CFR for Carnivora, making the data align even more closely to our modeled predictions.

Data availability:
All data used in this analysis is publicly available at the following GitHub repository: github.com/carabrook/spillover-virulence

Code availability: All code used in this analysis is publicly available at the following GitHub repository: github.com/carabrook/spillover-virulence

Additional Methods References:

35. Baalen, M. Van & Sabelis, M. W. The dynamics of multiple infection and the evolution of virulence. *Am. Nat.* **146**, 881–910 (1995).
36. Restif, O. & Koella, J. C. Shared control of epidemiological traits in a coevolutionary model of host-parasite interactions. *Am. Nat.* **161**, 827–836 (2003).
37. Geritz, S. A. H., Kisdi, É., Meszéna, G. & Metz, J. A. J. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* **12**, 35–57 (1998).
38. Day, T. On the evolution of virulence and the relationship between various measures of mortality. *Proc. R. Soc. B* **269**, 1317–1323 (2002).
39. Wood, S. N. mgcv: GAMs and Generalized Ridge Regression for R. *R News* **1/2**, 20–24 (2001).
40. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

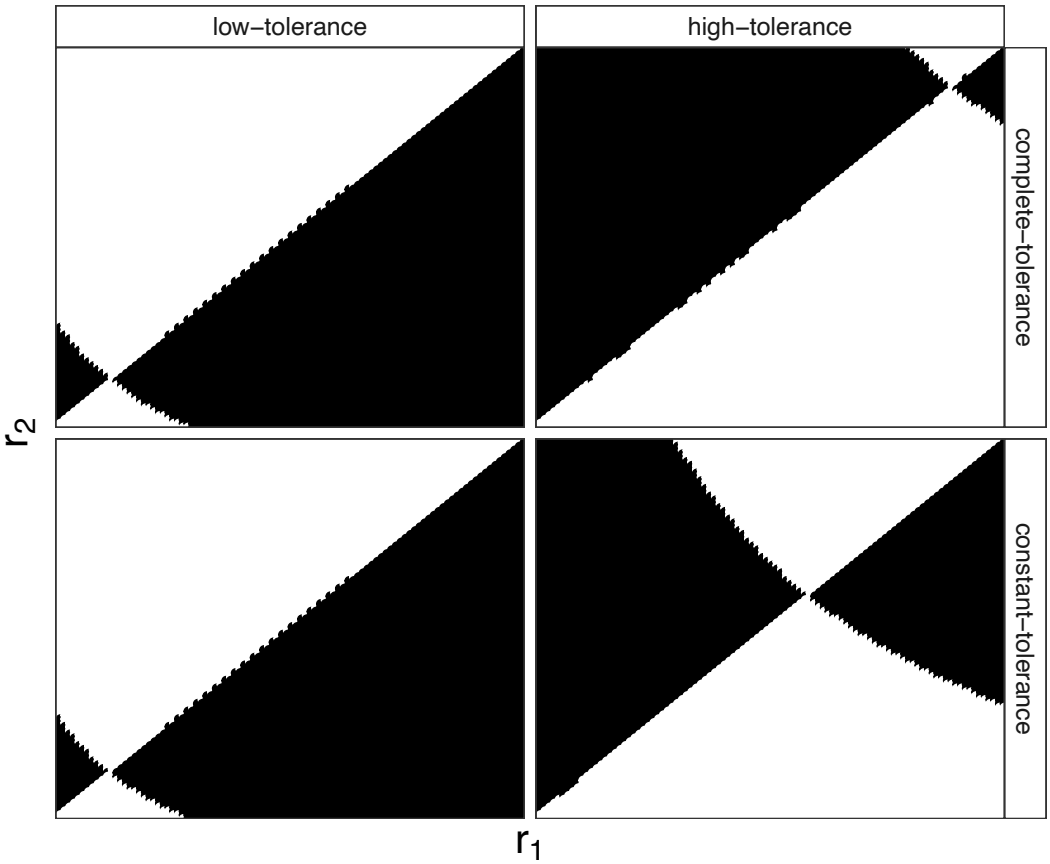
Acknowledgments:
The authors thank the Boots Lab at UC Berkeley for helpful comments on this manuscript and acknowledge funding from a National Institutes of Health grant 1R01AI129822-01 to CEB, an Adolph C. and Mary Sprague Miller Institute for Basic Research fellowship to CEB, a Branco Weiss Science in Society fellowship to CEB, a Loréal-USA for Women in Science fellowship to CEB, and a Defense Advanced Research Projects Agency PREEMPT Program subgrant D18AC00031 to CEB.

Author contributions:
CEB and MB conceptualized the research. CEB carried out the modeling and analysis with contributions from CR and SG. CEB wrote the first draft of the manuscript, which all authors edited and revised.

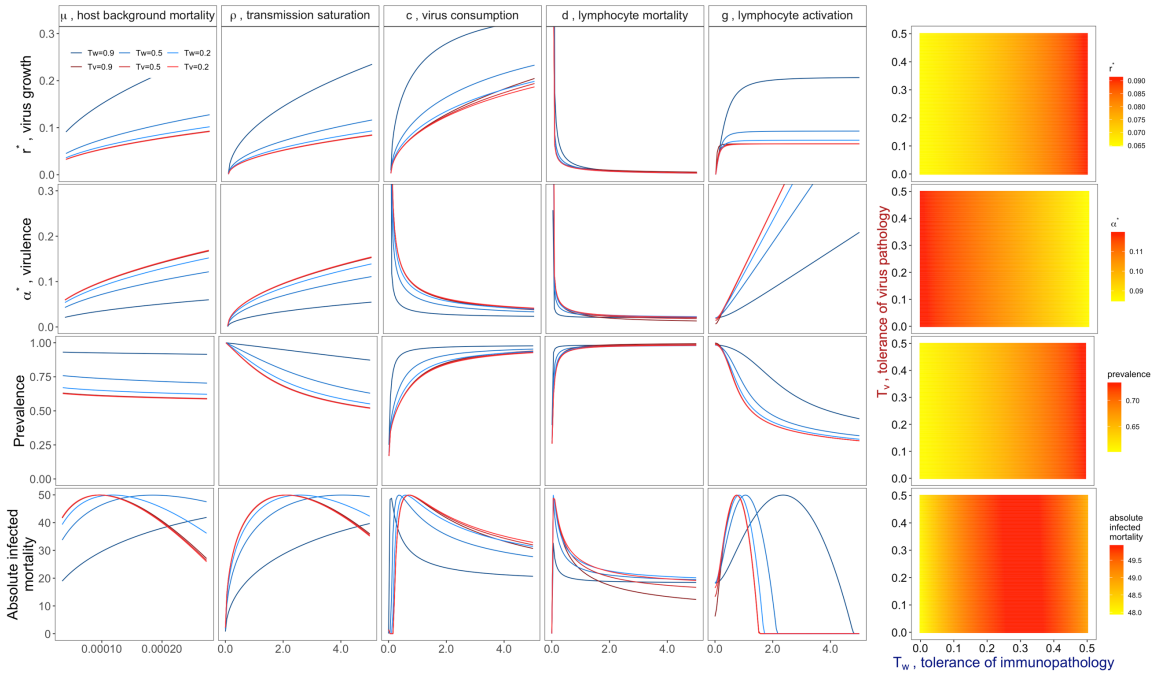
Competing interests:
Authors declare that they have no competing interests.

Materials and correspondence:
All correspondence should be addressed to corresponding author Cara E. Brook at cbrook@uchicago.edu.

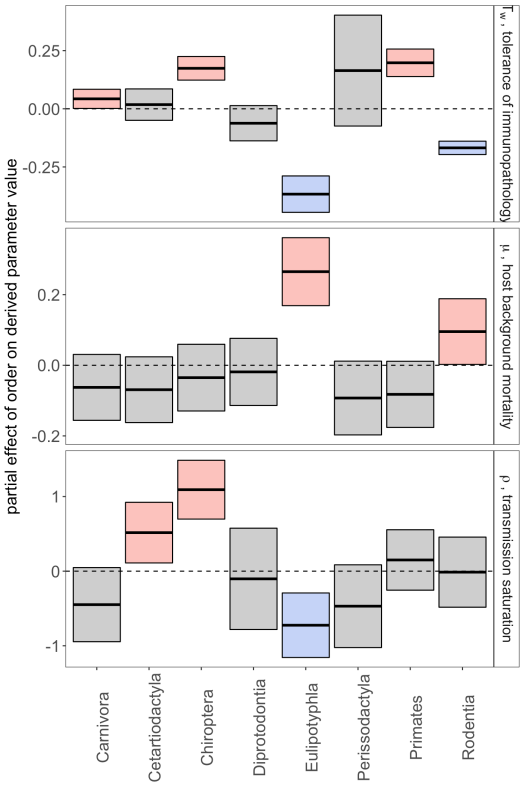
Extended Data Figures:



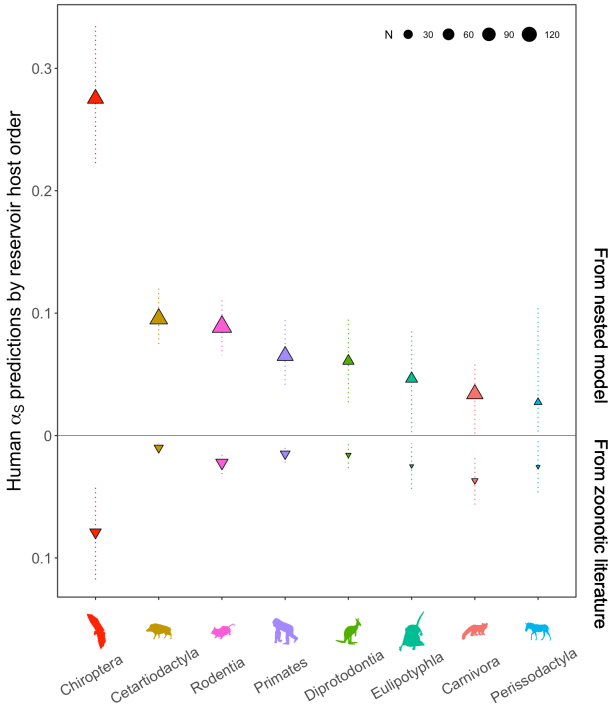
Extended Data Figure 1. Pairwise invasibility plots demonstrate optimal viral growth rate, under low and high tolerance conditions. Invading growth rates (r_2) will displace resident growth rates (r_1) at values indicated by the shaded regions. Tolerance of immunopathology (T_w) is modeled as low in left column (.1 and 1.1 for row 1 and 2, respectively) and high in right column (.9 and 5 for row 1 and 2), assuming a complete (row 1) or a constant form (row 2). T_v is fixed at .1 and 1.1 for all complete (row 1) and constant tolerance scenarios (row 2). For this visualization, r_1 and r_2 span from 0.05 to 2. All other parameters involved in computation of r^* (see *[26]* and *[27]*) were fixed at values listed in Extended Data Table 1.



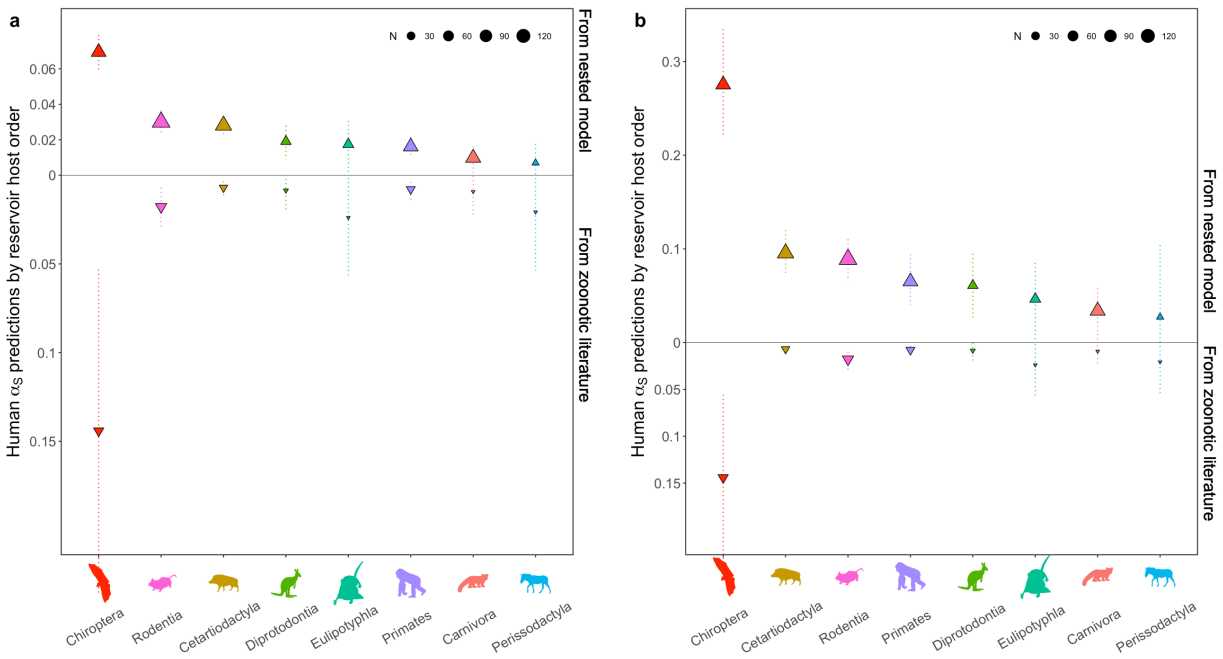
Extended Data Figure 2. Optimal virus growth rates vary across reservoir host immunological and life history parameters. Figure replicates Fig. 2 (main text) under assumptions of complete tolerance. Rows (top-down) indicate the evolutionarily optimal within-host virus growth rate (r^*) and the corresponding virus-induced mortality rate (α^*), population-level infected prevalence, and population-level absolute infected mortality for a reservoir host infected with a virus at r^*. Columns demonstrate the dependency of these outcomes on variable reservoir host parameters: background mortality rate (μ), transmission saturation scalar (ρ), virus consumption rate by lymphocytes (c), lymphocyte mortality rate (d), and lymphocyte activation rate upon viral contact (g). Darker colored lines depict outcomes at higher values for tolerance of virus pathology (T_v, red) or immunopathology (T_w, blue), assuming no tolerance of the opposing type. Heat maps demonstrate how T_v and T_w interact to produce each outcome. Outcome ranges differ between lineplots (y-axes) and heatmaps (scalebars). Parameter values are listed in Extended Data Table 1.



Extended Data Figure 3. Partial effects of host order on derived within-host nested-model parameters. Outputs from linear regression and generalized additive models used to generate within-host parameters across eight mammalian orders (from top: T_w, tolerance of immunopathology, linked to lifespan per BMR; μ, natural host mortality; ρ, transmission saturation scalar linked to group size). Orders with significant partial effects (computed following methods established in Mollentze and Streicker 2020[7]) on the resulting parameter are indicated by standard error confidence intervals which do not cross zero (dashed line). Significantly positive partial effects are shaded red, and significantly negative effects are shaded blue. Significantly positive effects for all three parameters drive r^* higher, while significantly negative effects drive r^* lower for viruses derived from each order. The thick line in each boxplot corresponds to the median effect of each order on the corresponding parameter value. Resulting parameter values are summarized in Extended Data Table 2 and model outputs summarized in Extended Data Table 3.



Extended Data Figure 4. Comparing human spillover virulence (α_S) predictions for zoonoses from nested life history model with estimates from the literature, under complete tolerance assumptions. Figure replicates Fig. 3d (main text) under assumptions of complete tolerance. Top panel depicts α_S in human predictions for zoonoses by order, derived from our nested within-host, life history model, while bottom panel depicts human α_S estimates derived from case fatality rates and infection duration reported in the zoonotic literature[4]. For predictions, point size corresponds to the average of the number of species-level datapoints (N) used to generate each of the four variable parameters (Fig. 3, Extended Data Fig. 3) used in human α_S predictions. For estimates from the literature, N corresponds to the number of independent host-virus associations from which estimates were determined. 95% confidence intervals were computed by standard error; for predictions from nested model, these reflect the upper and lower confidence intervals of the optimal virus growth rate. Order-specific parameter values for r^*, ρ, μ, T_w, T_v, are listed in Extended Data Table 2; all other parameters involved in calculation of α_S are listed in Extended Data Table 1.



Extended Data Figure 5. Comparing human case fatality rate predictions for zoonoses from nested life history model with estimates from the literature, excluding rabies. Figure replicates Fig. 3d (main text) under assumptions of (**a**) constant and (**b**) complete tolerance but excluding rabies lyssavirus from the zoonotic data (bottom panels). Rank-order predictions of virulence are more consistent with order Carnivora further down the rankings. As in Fig. 3d and Extended Data Fig. 4, order-specific parameter values for r^*, ρ, μ, T_w, T_v, are listed in Extended Data Table 2 ; all other parameters involved in calculation of α_S are listed in Extended Data Table 1.

860 **Extended Data Tables:**

861

Extended Data Table 1. Default parameter values for population-level and within-host mode

Parameter	Definition	Units	Value*	Scale
$\boldsymbol{b}$	host natural birth rate	days^{-1}	0.2	population-level
$\boldsymbol{q}$	crowding term to regulate density-dependent growth	days^{-1}	0.0002	population-level
$\boldsymbol{\mu}^{\dagger}$	host background mortality rate	days^{-1}	$\frac{1}{20 * 365}$	population-level
$\boldsymbol{c}$	rate of virus consumption by lymphocyte	days^{-1}	0.5	within-host
$\boldsymbol{d}$	background mortality rate for lymphocytes	days^{-1}	$\frac{1}{21}$	within-host
$\boldsymbol{g}$	rate of lymphocyte activation by virus	days^{-1}	0.9	within-host
K	carrying capacity of virions per host	virions	100,000	within-host
$\boldsymbol{\zeta}$	maximum between-host transmission rate	days^{-1}	0.2	within-host
$\boldsymbol{\rho}^{\dagger}$	transmission saturation scalar (rapidity by which β approaches ζat increasing virion density)	unitless	3	within-host
$\boldsymbol{\gamma}$	scalar of host recovery rate with immune cell density	(lymphocytes*days)$^{-1}$	0.001	within-host
$\boldsymbol{w}$	host susceptibility to immunopathology	lymphocytes^{-1}	1	within-host
$\boldsymbol{v}$	intrinsic virus virulence	virions^{-1}	1	within-host
$\boldsymbol{T_w}^{\dagger}$	tolerance of immunopathology (constant or complete)	unitless	*variable* constant: $1-2$ complete: $0-1$	within-host
$\boldsymbol{T_{vS}}^{\dagger}$	human tolerance of viral pathology (constant or complete)	unitless	*variable* constant: $1-2$ complete: $0-1$	between-host spillover

*Value column gives parameter values fixed in Fig. 2 and Extended Data Fig. 1, excepting instances when parameter was modulated along the x-axis or internally (T_w, T_v) as indicated in the figure legend. Parameter values are comparable to the ranges explored in Miller et al. 2006[18] and Boots et al. 2013[31].

†Parameters μ, ρ, T_w, and T_{vS} were modeled as functions of mammalian order-specific life history traits in Fig. 4 and Extended Data Fig. 4 and 5, as outlined Methods, Section 7 and in the main text.

862

863

864

865

866

867

Extended Data Table 2. Order-specific parameter values for within-host nested model.*

	Carnivora	Cetartiodactyla	Chiroptera	Diprotodontia	Eulipotyphla	Perissodactyla	Primates	Rodentia
$\boldsymbol{\mu}$, host background mortality	0.063 [0.046-0.08]	0.056 [0.041-0.072]	0.091 [0.067-0.11]	0.11 [0.08-0.13]	0.39 [0.36-0.42]	0.033 [0-0.09]	0.043 [0.024-0.063]	0.22 [0.21-0.24]
$\boldsymbol{\rho}$, transmission saturation	0.00028 [0-6e-04]	0.0026 [0.0023-0.0028]	0.0098 [0.0095-0.01]	0.00062 [0.00018-0.0011]	0.00015 [0-0.00042]	0.00027 [0-0.00063]	0.0011 [0.00086-0.0014]	0.00077 [0.00046-0.0011]
$\boldsymbol{T_w}$, tolerance of immunopathology (constant)	1.5 [1.4-1.5]	1.5 [1.4-1.5]	1.6 [1.6-1.7]	1.4 [1.3-1.5]	1.1 [1-1.2]	1.6 [1.4-1.8]	1.6 [1.6-1.7]	1.3 [1.2-1.3]
$\boldsymbol{T_w}$, tolerance of immunopathology (complete)	0.49 [0.45-0.53]	0.46 [0.4-0.53]	0.62 [0.57-0.67]	0.38 [0.31-0.46]	0.078 [0-0.16]	0.61 [0.37-0.85]	0.64 [0.58-0.7]	0.28 [0.25-0.31]
$\boldsymbol{r^*}$, optimal virus growth rate (constant)	0.0053 [0-0.0088]	0.015 [0.012-0.018]	0.038 [0.032-0.043]	0.01 [0.0048-0.015]	0.0095 [0-0.017]	0.0038 [0-0.0096]	0.0088 [0.0058-0.012]	0.016 [0.012-0.02]
$\boldsymbol{r^*}$, optimal virus growth rate (complete)	0.019 [0-0.032]	0.052 [0.04-0.066]	0.15 [0.12-0.18]	0.033 [0.015-0.051]	0.025 [0-0.046]	0.015 [0-0.057]	0.036 [0.022-0.052]	0.049 [0.036-0.06]
$\boldsymbol{T_{vS}}$, human tolerance of viral pathology (constant)	1.4	1.4	1.4	1	1.4	1.4	2	1.4
$\boldsymbol{T_{vS}}$, human tolerance of viral pathology (complete)	0.39	0.39	0.39	0	0.39	0.39	1	0.43
$\boldsymbol{\alpha_S}$, spillover virulence (constant)	0.0098 [0-0.016]	0.028 [0.023-0.033]	0.069 [0.059-0.079]	0.019 [0.0089-0.028]	0.017 [0-0.031]	0.0069 [0-0.018]	0.016 [0.011-0.022]	0.03 [0.023-0.037]
$\boldsymbol{\alpha_S}$, spillover virulence (complete)	0.034 [0-0.058]	0.095 [0.073-0.12]	0.28 [0.22-0.33]	0.061 [0.027-0.095]	0.046 [0-0.085]	0.027 [0-0.1]	0.065 [0.04-0.094]	0.089 [0.066-0.11]

*Derived parameter values (depicted as *mean [lower confidence interval – upper confidence interval]*) computed from linear mixed effects regression and generalized additive models are shown across all eight mammalian orders (μ, ρ, T_w, T_v), in addition to the r^* and α_S values computed from their inputs. Confidence intervals are 95% derived via standard error. All other parameters involved in the calculation of r^* and α_S are listed in Extended Data Table 1.

Extended Data Table 3. Model summary outputs for within-host parameter estimation.*

A. Summary of linear mixed effects regression model to estimate $\boldsymbol{T_w}$

Formula	Term	Value	[lci-uci]	Statistic	P-value	Degrees of freedom
log_{10}(max_lifespan_yr) ~ log_{10}(BMR) + (1|order)	intercept	-0.0512	[-0.278 – 0.176]	*T:* -0.443	0.661	35.2
	$\log_{10}$(BMR)	-0.483	[-0.552 – -0.415]	*T:* -13.8	<.001***	264

B. Summary of generalized additive model to estimate $\boldsymbol{\mu}$

Formula	Term	Value	[lci-uci]	Statistic	P-value	Effective degrees of freedom
$\boldsymbol{\mu}$ *~ s(order, bs = "re")*	intercept	0.126	[.0334 – .218]	*T:* 2.67	<.001***	
	smoothing: order			*F:* 94.1	<.001***	6.9

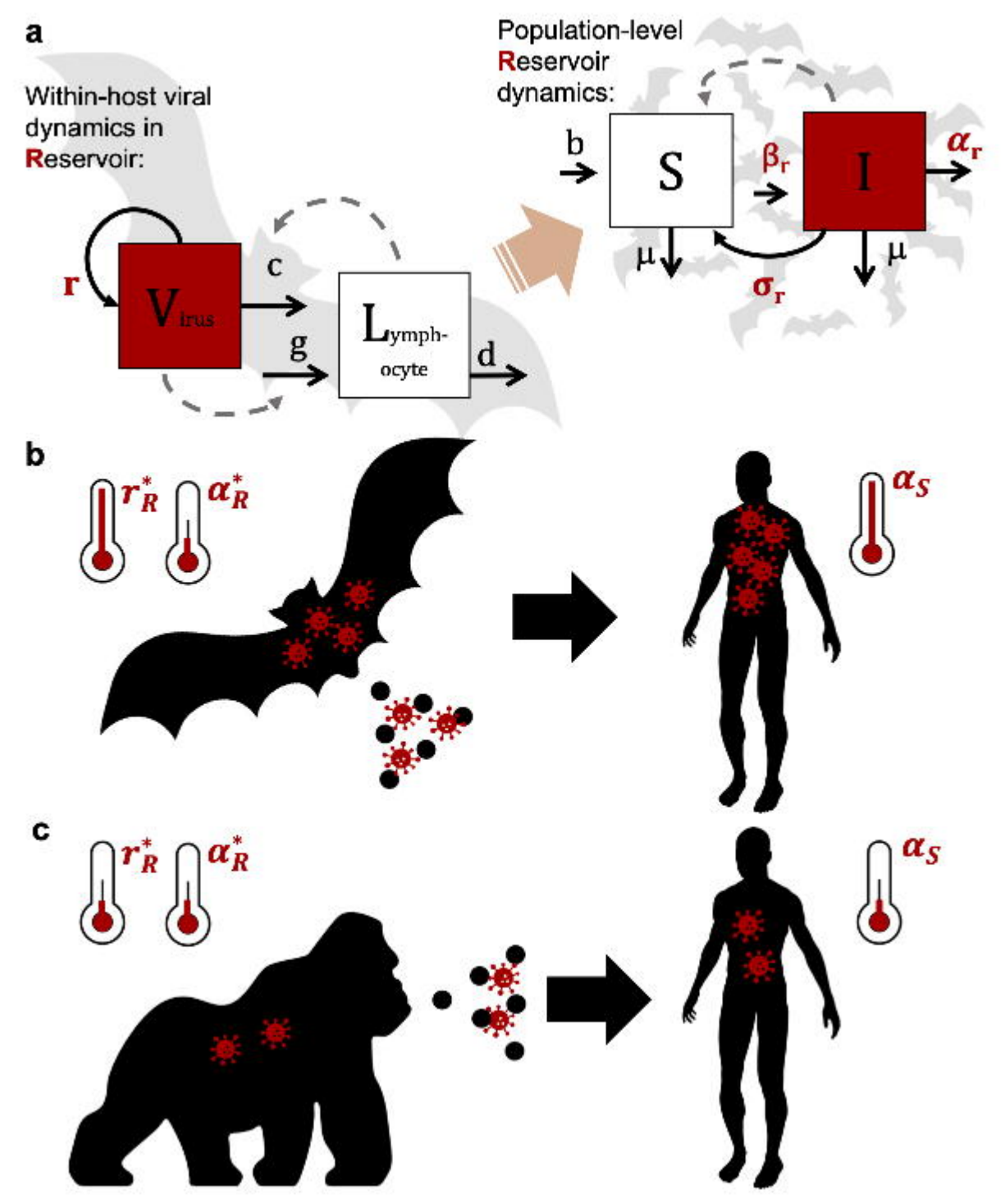
C. Summary of generalized additive model to estimate $\boldsymbol{\rho}$

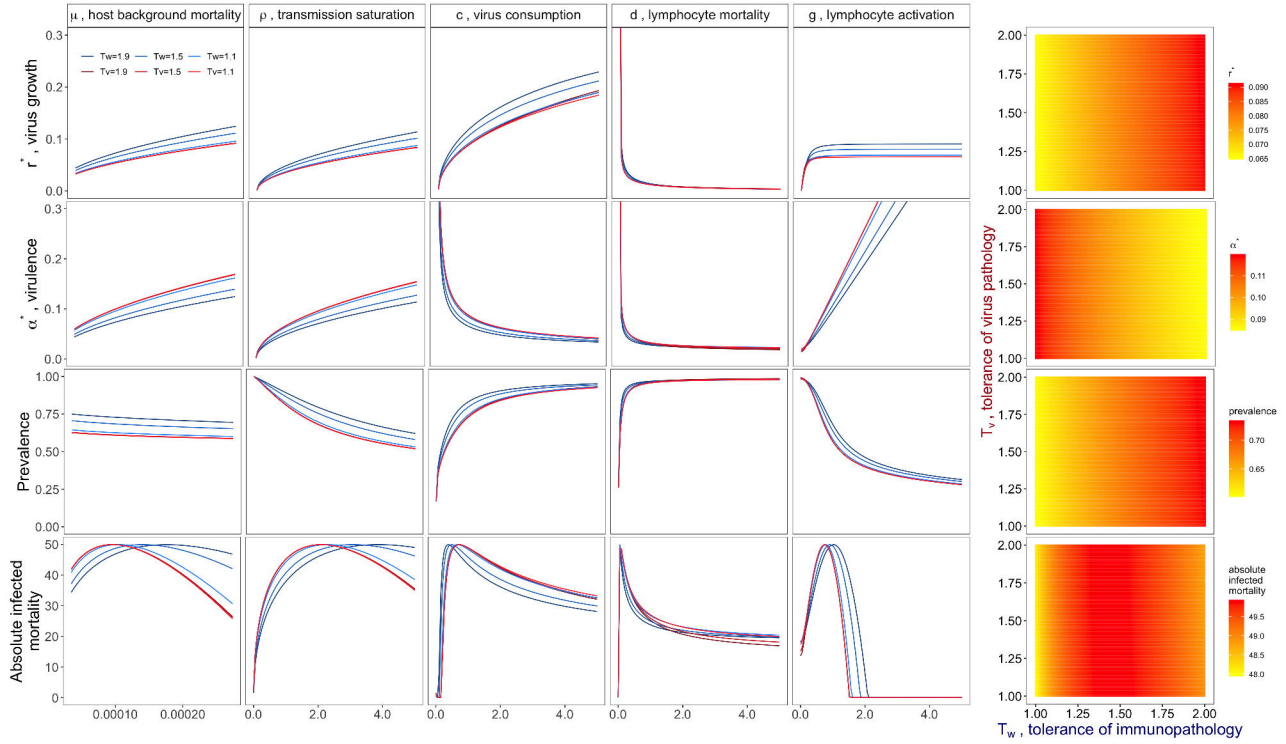
Formula	Term	Value	[lci-uci]	Statistic	P-value	Effective degrees of freedom
log_{10}(pop_group_size) ~ s(order, bs = "re")	intercept	0.886	[0.517 – 1.25]	*T:* 4.71	<.001***	
	smoothing: order			*F:* 25.2	<.001***	5.93

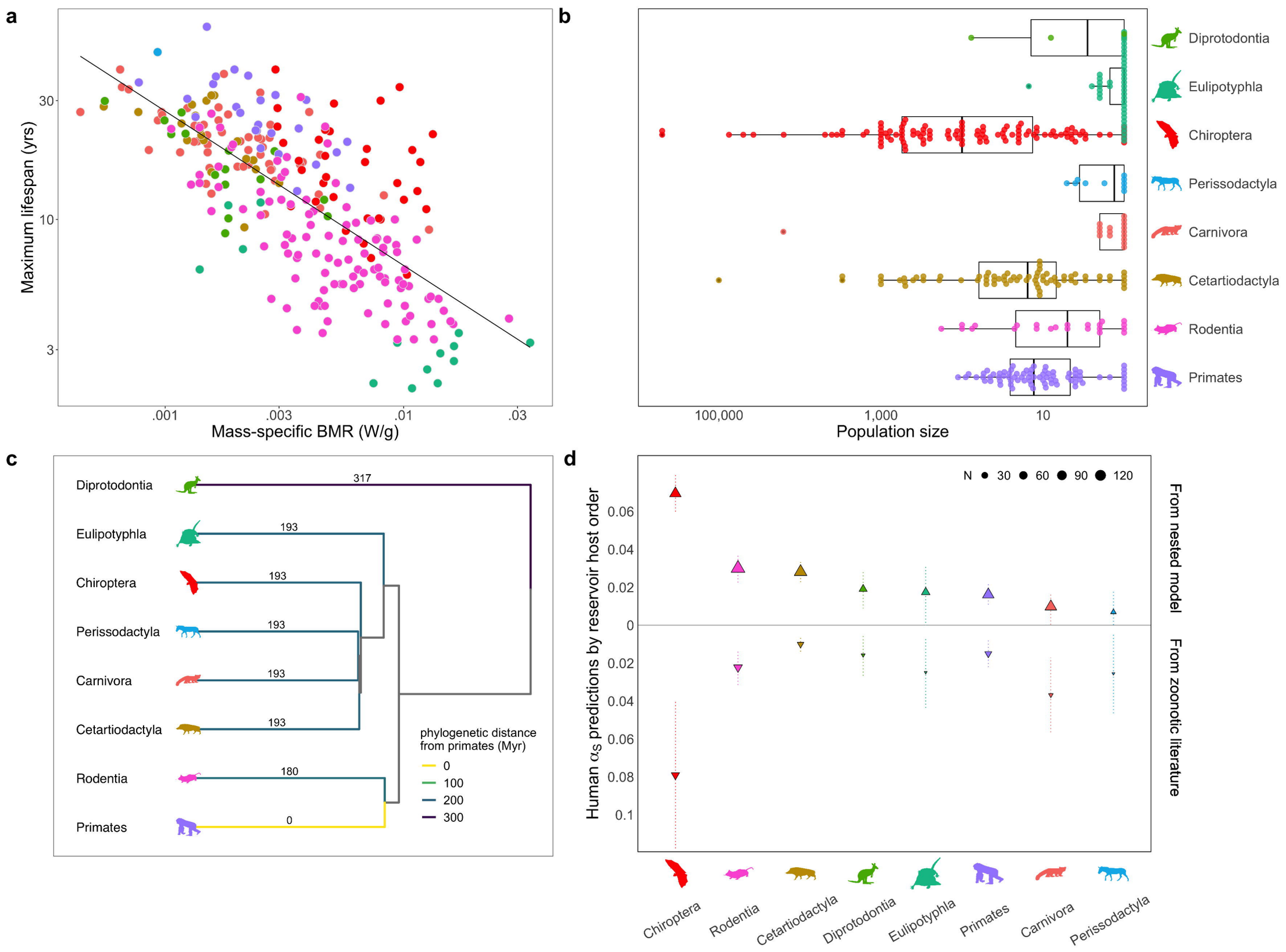
D. Summary of generalized additive model to summarize $\boldsymbol{\alpha_S}$ by order from literature

Formula	Term	Value	[lci-uci]	Statistic	P-value	Effective degrees of freedom
alpha ~ s(vFamily, bs = "re") + s(hOrder, bs = "re") + s(VirusSppPubs, k = 7, bs = "tp") + s(spill_type, bs = "re") + s(IsVectorBorne, bs = "re")	Intercept	-4.55	[-6.15 – -2.95]	*z:* -5.57	<.001***	
	smoothing:					
	virus family			χ^2: 612	.00387**	9.50
	order			χ^2: 1161	<.001***	4.74
	virus spp. pub count			χ^2: 5.77	.0163*	1.00
	spillover type			χ^2: 67.9	.00856**	0.867
	vector-borne status			χ^2: 98.9	.0249*	0.806

*Outputs from fitted linear mixed effect regression models and fitted generalized additive models used to generate within-host parameter values. Order-level random effects are visualized in Extended Data Fig. 3 and resulting parameters summarized in Extended Data Table 2. Significance at p < 0.001***, 0.01***, 0.05*.


a
Within-host viral dynamics in Reservoir:
r
Virus
c
g
Lymph-ocyte
d
Population-level Reservoir dynamics:
b
S
β_r
I
α_r
μ
σ_r
μ
b
r_R^*
α_R^*
α_S
c
r_R^*
α_R^*
α_S


μ , host background mortality
ρ , transmission saturation
c , virus consumption
d , lymphocyte mortality
g , lymphocyte activation
Tw=1.9 Tw=1.5 Tw=1.1
Tv=1.9 Tv=1.5 Tv=1.1
r* , virus growth
α* , virulence
Prevalence
Absolute infected mortality
r*
α*
prevalence
absolute infected mortality
Tv , tolerance of virus pathology
Tw , tolerance of immunopathology


a
Maximum lifespan (yrs)
30
10
3
.001
.003
.01
.03
Mass-specific BMR (W/g)
b
Diprotodontia
Eulipotyphla
Chiroptera
Perissodactyla
Carnivora
Cetartiodactyla
Rodentia
Primates
100,000
1,000
10
Population size
c
Diprotodontia
Eulipotyphla
Chiroptera
Perissodactyla
Carnivora
Cetartiodactyla
Rodentia
Primates
317
193
193
193
193
193
180
0
phylogenetic distance from primates (Myr)
0
100
200
300
d
N 30 60 90 120
Human α_S predictions by reservoir host order
0.06
0.04
0.02
0
0.02
0.04
0.06
0.08
0.1
From nested model
From zoonotic literature
Chiroptera
Rodentia
Cetartiodactyla
Diprotodontia
Eulipotyphla
Primates
Carnivora
Perissodactyla