PNAS RESEARCH ARTICLE | MICROBIOLOGY OPEN ACCESS

Enveloped viruses show increased propensity to cross-species transmission and zoonosis

Ana Valero-Rello[a] and Rafael Sanjuán[a,1]

Edited by Eugene Koonin, NIH, Bethesda, MD; received September 12, 2022; accepted November 7, 2022

The transmission of viruses between different host species is a major source of emerging diseases and is of particular concern in the case of zoonotic transmission from mammals to humans. Several zoonosis risk factors have been identified, but it is currently unclear which viral traits primarily determine this process as previous work has focused on a few hundred viruses that are not representative of actual viral diversity. Here, we investigate fundamental virological traits that influence cross-species transmissibility and zoonotic propensity by interrogating a database of over 12,000 mammalian virus–host associations. Our analysis reveals that enveloped viruses tend to infect more host species and are more likely to be zoonotic than nonenveloped viruses, while other viral traits such as genome composition, structure, size, or the viral replication compartment play a less obvious role. This contrasts with the previous notion that viral envelopes did not significantly impact or even reduce zoonotic risk and should help better prioritize outbreak prevention efforts. We suggest several mechanisms by which viral envelopes could promote cross-species transmissibility, including structural flexibility of receptor-binding proteins and evasion of viral entry barriers.

zoonosis | cross-species transmission | viral emergence | viral host breadth | virus discovery

Viral cross-species transmission is at the origin of an increasing number of emerging diseases. This includes well-known zoonoses, such as HIV/AIDS, influenza, Zika, Ebola, monkeypox, and COVID, in addition to other animal and plant diseases. Understanding and ultimately predicting viral emergence has therefore become a research priority (1–4). Nearly 90% of known viral zoonoses originate from wild or domesticated mammals (5). A number of risk factors have been identified, such as biodiversity loss (6), species invasions (7), wildlife trade (8), viral host plasticity (9, 10), life history traits of reservoir hosts (11), and host proximity to humans in terms of phylogenetic distance (12) and interaction frequency (13). Additionally, information about host tropism and zoonotic propensity can be extracted from viral genomes by analyzing features such as codon or dinucleotide usage biases and to what extent these biases resemble those found in host gene transcripts (12, 14, 15).

Despite these advances, it remains unclear how cross-species transmissibility and zoonosis depend on the fundamental properties of a virus. Previous work has emphasized that RNA viruses should in principle be more prone to host jumps than DNA viruses owing to their extensive genetic diversity and fast adaptability (16, 17). Some recent studies have supported this view, whereas others have identified seemingly more relevant features, such as viral replication in the cytoplasm (18–20). Some analyses have also suggested an effect of viral genome size or genome segmentation, but these findings have not been supported by others (9, 13, 18, 21).

A limitation of previous studies is that they rely essentially on well-known viruses approved by the International Committee on Taxonomy of Viruses (ICTV), which constitute a small and nonrepresentative subset of the mammalian virosphere. Importantly, viruses with socioeconomic implications, such as those causing human disease, have been preferentially investigated, critically biasing inferences on viral host usage, cross-species transmissibility, and zoonotic propensity (3). However, the advent of omics technologies and the implementation of major initiatives such as PREDICT and the Global Virome Project (22) have revolutionized viral discovery (23). These advances have provided a still limited but less biased picture of viruses in nature and have revealed a large number of new virus–host associations (24). In many cases, only a few short sequence reads are available for newly described viruses. Although this precludes certain analyses, these sequences inform about fundamental viral features, such as the nature, size, and structure of their genetic material. It is therefore possible to leverage this information to examine how these features influence the ability of viruses to infect multiple host species and cause zoonoses.

Significance

The cross-species transmission of viruses from wild or domestic animals to humans (zoonosis) has produced major epidemics, but our understanding of this complex process is still very limited. Some risk factors have been identified, such as ecological perturbations and the nature of viral reservoir species. However, it remains unclear how cross-species transmission and zoonosis depend on fundamental viral features. Using large amounts of data released by recent viral discovery initiatives, here we show that enveloped viruses tend to infect more host species and are more likely to cause zoonotic infections than nonenveloped viruses, while other basic viral features examined play less obvious roles. These findings challenge previous views in the field and will help guide viral outbreak surveillance.

Author contributions: R.S. designed research; A.V.-R. and R.S. performed research; A.V.-R. and R.S. analyzed data; and R.S. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: rafael.sanjuan@uv.es.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2215600119/-/DCSupplemental.

Published December 6, 2022.

Results

Dataset. We extracted from the Global Virome in One Network (VIRION) database 5,149 viruses belonging to 36 families and 1,599 host species from 20 orders comprising in total 12,888 virus–host associations (Dataset S1). All viruses were ratified by the National Center for Biotechnology Information (NCBI) taxonomy site and were assigned a viral family. Most (77.8%) were not ICTV-approved viruses and thus may be genuine species, higher-order taxa, or viral subtypes. Approximately half of the viruses (52.6%) had a single sequence record in NCBI Entrez, whereas some had thousands (Fig. 1*A*). As expected, the number of host species in which a virus was found depended strongly on the number of sequence records available for that virus, N (Fig. 1*B*). We therefore included N as a covariate in all subsequent analyses.

We considered five dichotomous variables and one continuous variable defining fundamental viral features: whether the genome is made of RNA or DNA, single or double stranded, segmented or nonsegmented, whether the virus replicates in the cytoplasm or the nucleus, whether the virus is enveloped or nonenveloped, and the viral genome size. These features are conserved within viral families, and hence, their imputation was straightforward even if only few short sequences were available without further characterization.

Cross-Species Transmissibility. Since zoonosis is a special case of viral cross-species transmission, we first examined the host breadth of each virus and measured the number of host species in which a virus has been found, excluding humans to reduce bias. An exploratory analysis revealed that the observed number of host species increased more rapidly with N for enveloped viruses than for nonenveloped viruses, being approximately twice as high in the former group (Fig. 1*B*). This difference was also detectable when we combined the envelope factor with each of the other dichotomous viral features considered (*SI Appendix*, Fig. S1).

To formally assess the effects of different viral features on host species counts, we performed a negative binomial regression analysis. For this, we initially set a cutoff number of sequences N ≥ 5 to avoid poorly explored or ill-defined viruses, a condition that was met for 1,305 viruses, of which nearly half (42.4%) were not ICTV-approved. The model explained 53.0% of the deviance relative to an intercept-only model and revealed a strongly significant effect of the viral envelope on the number of host species per virus ($P < 0.0001$). We also detected an increase in host species counts associated with viruses that replicate in the cytoplasm ($P = 0.013$), whereas all other factors showed no significant effects (*SI Appendix*, Table S1 and Fig. 2*A*). As suggested previously (21), viruses with smaller genomes tended to be found in fewer host species, but this effect was accounted for by the fact that these are typically nonenveloped viruses.

We checked that our conclusions were robust to the cutoff N used by repeating the above analysis with all 5,149 viruses, those satisfying N ≥ 2, N ≥ 3, and so on for all cutoff values until the sample size became too small to detect any effect (Fig. 3). We also inspected whether differences between enveloped and nonenveloped viruses were driven by specific viral families (Fig. 4). To assess this, we included each viral family as a categorical factor in the binomial regression analysis using an N ≥ 5 cutoff. Despite the loss of statistical power, the difference in average number of host species between enveloped and nonenveloped viruses was confirmed ($P = 0.008$; *SI Appendix*, Fig. S2 and Table S2).

The above regression analysis was complicated by the strong overdispersion shown by host species counts (*SI Appendix*, Fig. S3). To take a simpler and probably more robust approach, we converted these counts into a binary variable indicating whether a virus was found in multiple host species (multihost virus). This dichotomous response variable was analyzed using binary logistic regression, which again confirmed that enveloped viruses were more likely to undergo cross-species transmission than nonenveloped viruses ($P < 0.0001$), whereas all other viral features examined were either not significant or marginally significant ($P > 0.04$; *SI Appendix*, Fig. S4 and Table S3).

Zoonotic Propensity. We then focused on whether these mammalian viruses were zoonotic, i.e., capable of infecting humans. As above, the fraction of zoonotic viruses for a given N tended to be higher for enveloped viruses than for nonenveloped viruses (Fig. 1*C*). A binary logistic regression analysis for viruses with N ≥ 5 sequence records confirmed a strongly significant effect of the viral envelope, with an estimated 2.5-fold increase in zoonotic propensity compared with nonenveloped viruses (Fig. 2*B* and *SI Appendix*, Table S4). Statistical significance was robust to the cutoff N used, as evidenced when we considered all 5,419 viruses in the dataset, those with N ≥ 2, N ≥ 3, and up to N ≥ 65 (Fig. 3). In addition, we found that viruses replicating in the cytoplasm were 1.9 times more likely to be zoonotic than those replicating in the nucleus, in accordance with previous studies (18, 19). We also detected a slightly higher propensity of segmented viruses to zoonosis compared with nonsegmented viruses, also consistent with previous suggestions (13), and a decreasing zoonotic

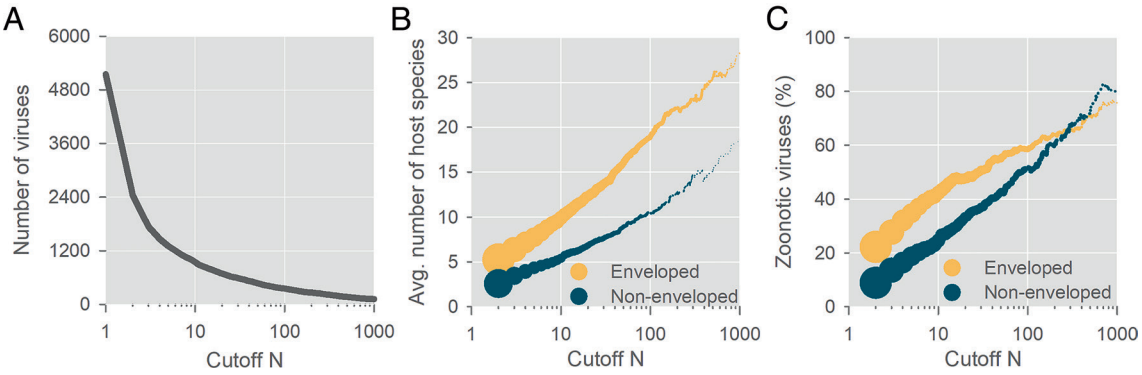


Fig. 1. Viral discovery patterns. (*A*) Most recently discovered viruses have a small number of deposited sequence records. The X axis shows the minimal number of sequences required for inclusion in our analysis (cutoff N), and the number of viruses meeting this condition is shown on the Y axis. (*B*) The average number of host species per virus increases more rapidly with the cutoff N for enveloped viruses than for nonenveloped viruses. Dot sizes are proportional to the number of different viruses included in each cutoff. The dots are shown for N ≥ 2, N ≥ 3, and so on. Dots for N ≥ 1 were too large for visualization and are omitted. (*C*) Enveloped viruses tend to be more zoonotic than nonenveloped viruses. For the small fraction of viruses with many available sequences, nonenveloped viruses are similarly or even more zoonotic than enveloped viruses, reproducing previous findings (see text).

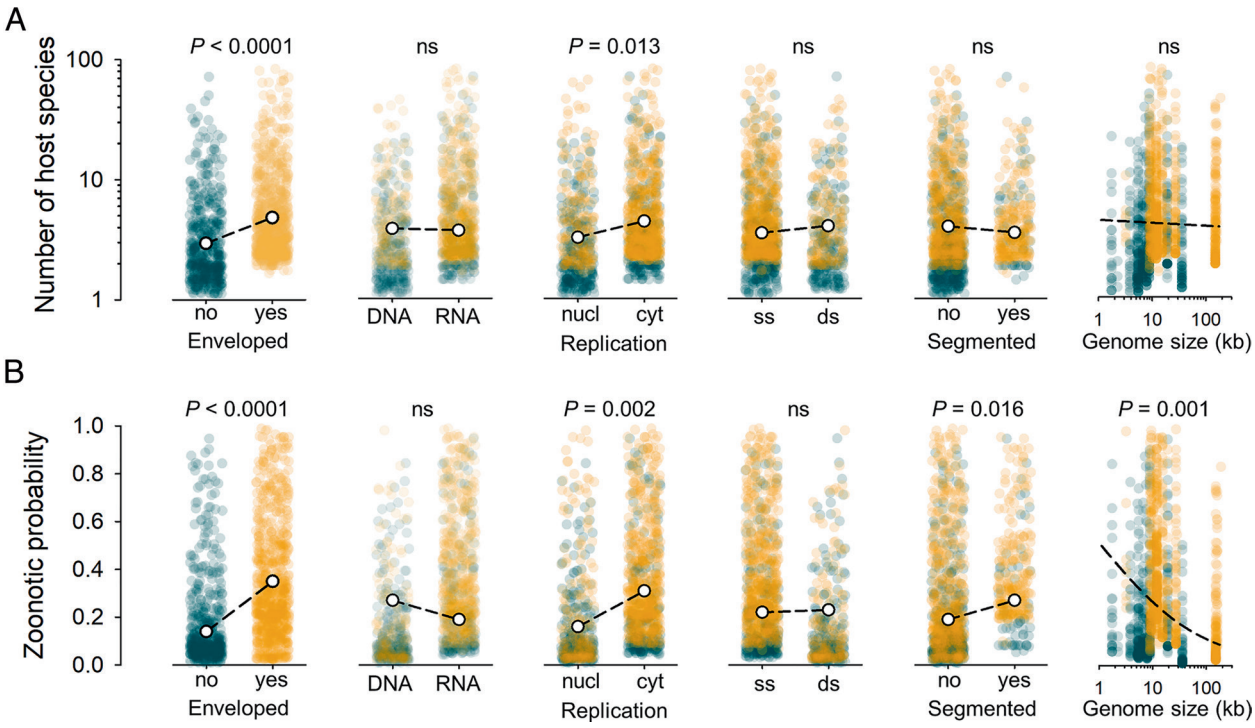


Fig. 2. Regression analysis for viruses with N ≥ 5 available sequences. (*A*) Scatterplots show the number of host species per virus predicted by a negative binomial regression for each of the 1,305 viruses included in this analysis. Orange dots indicate enveloped viruses, and green dots indicate nonenveloped viruses. White dots and dashed lines indicate the marginal means predicted by the model (i.e., setting all other variables to their mean values). *P* values for each predictor variable are shown. N was included as a covariate in the analysis but is not shown. A summary of the model statistics is provided in *SI Appendix*, Table S1. (*B*) Same plots for zoonosis probability. This response variable was analyzed using binary logistic regression. A summary of the model statistics is provided in *SI Appendix*, Table S4.

probability for viruses with larger genomes. Since these two features had no significant effects on cross-species transmissibility at large, their effect on zoonotic propensity might be due to human-specific factors or, more likely, to biases in human-infective virus datasets.

Discussion

Our results contrast with previous work suggesting that enveloped viruses were similarly or even less prone to zoonosis than nonenveloped viruses and that the viral feature most clearly associated with zoonosis was replication in the cytoplasm (9, 13, 18, 21). Our analyses do capture the effect of replicating in the cytoplasm on host breadth and zoonotic propensity but suggest a stronger influence of the viral envelope. This discrepancy could be due to the fact that previous analyses were focused on approximately 350 highly studied viruses, which may nevertheless be unrepresentative of the actual viral diversity. Indeed, replication in the cytoplasm became the only feature with a significant effect on zoonotic propensity when our logistic regression analysis was performed only with the 353 viruses that had N ≥ 100 available sequence records, reproducing previous findings (18).

Another possible caveat with the effects of cytoplasmic replication is that this feature is typical of RNA viruses (93.1% and 96.5% coincidence for ICTV viruses and our dataset, respectively), making it difficult to separate the contribution of these two traits. Therefore, we suggest caution in concluding whether the factor that promotes transmissibility between species is cytoplasmic replication or having an RNA genome.

Previous work has also suggested that enveloped viruses tend to be less transmissible among humans than nonenveloped viruses,

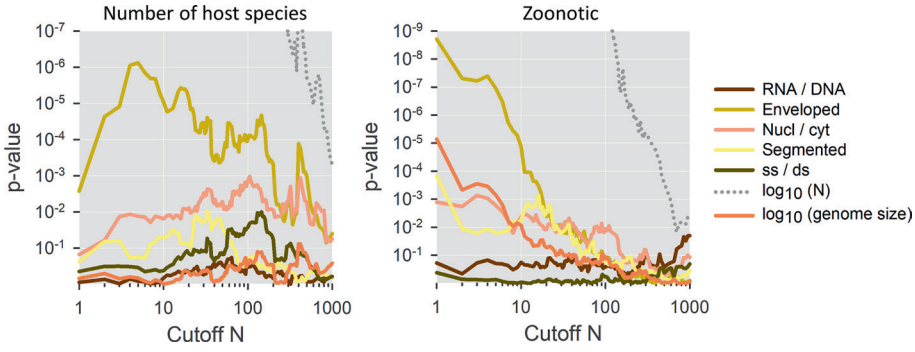
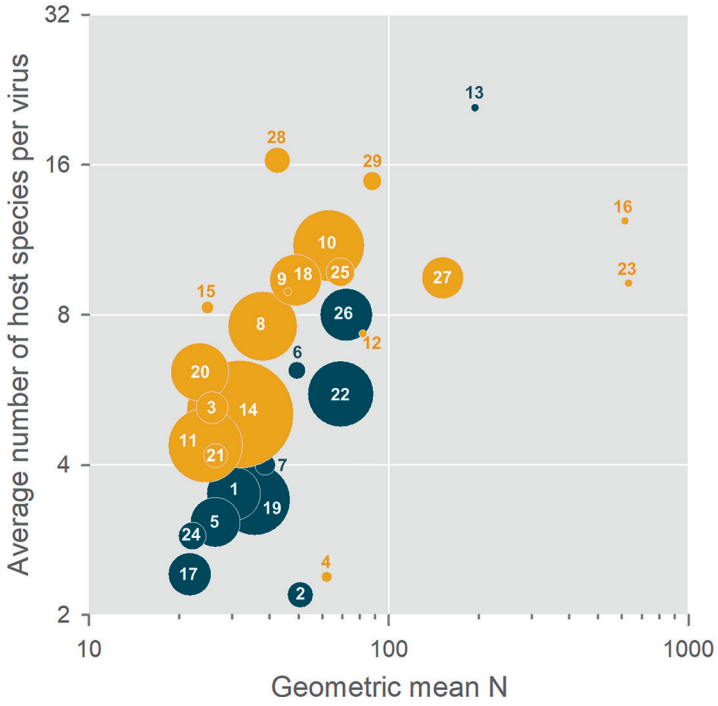


Fig. 3. Results of the regression analyses using different cutoff N values. The Y axis shows the *P* value obtained for each of the explanatory variables. The analysis was run for viruses with N ≥ 1, 2, ..., 1,000 sequences. The *P* values for the covariate log(N) are shown out of scale to help visualize relevant *P* values. *Left*: number of host species; *Right*: zoonotic viruses.



	Family	Enveloped	Genetic material	Replication site	Genome Strands	Segmented	Genome size
1.	*Adenoviridae*	no	DNA	nuc	ds	no	36
2.	*Anelloviridae*	no	DNA	nuc	ss	no	4
3.	*Arenaviridae*	yes	RNA	cyt	ss	yes	10
4.	*Arteriviridae*	yes	RNA	cyt	ss	no	15
5.	*Astroviridae*	no	RNA	cyt	ss	no	7
6.	*Caliciviridae*	no	RNA	cyt	ss	no	8
7.	*Circoviridae*	no	DNA	nuc	ss	no	2
8.	*Coronaviridae*	yes	RNA	cyt	ss	no	27
9.	*Filoviridae*	yes	RNA	cyt	ss	no	19
10.	*Flaviviridae*	yes	RNA	cyt	ss	no	9
11.	*Hantaviridae*	yes	RNA	cyt	ss	yes	12
12.	*Hepadnaviridae*	yes	DNA	nuc	ds	no	3
13.	*Hepeviridae*	no	RNA	cyt	ss	no	7
14.	*Herpesviridae*	yes	DNA	nuc	ds	no	151
15.	*Nairoviridae*	yes	RNA	cyt	ss	yes	19
16.	*Orthomyxoviridae*	yes	RNA	nuc	ss	yes	13
17.	*Papillomaviridae*	no	DNA	nuc	ds	no	8
18.	*Paramyxoviridae*	yes	RNA	cyt	ss	no	15
19.	*Parvoviridae*	no	DNA	nuc	ss	no	6
20.	*Peribunyaviridae*	yes	RNA	cyt	ss	yes	12
21.	*Phenuiviridae*	yes	RNA	cyt	ss	yes	12
22.	*Picornaviridae*	no	RNA	cyt	ss	no	9
23.	*Pneumoviridae*	yes	RNA	cyt	ss	no	15
24.	*Polyomaviridae*	no	DNA	nuc	ds	no	5
25.	*Poxviridae*	yes	DNA	cyt	ds	no	160
26.	*Reoviridae*	no	RNA	cyt	ds	yes	19
27.	*Retroviridae*	yes	RNA	nuc	ss	no	9
28.	*Rhabdoviridae*	yes	RNA	cyt	ss	no	12
29.	*Togaviridae*	yes	RNA	cyt	ss	no	12

Fig. 4. Average number of host species per virus for each viral family as a function of N. The 29 families with five or more different viruses, each having N ≥ 5 sequence records are shown and identified with numbers. Enveloped viruses are colored in orange, and nonenveloped viruses are colored in dark green. The size of the bubbles is proportional to the number of different viruses belonging to each family. The largest bubble corresponds to *Herpesviridae*, which includes 127 different viruses with N ≥ 5, whereas the smallest corresponds to *Hepeviridae*, with five viruses. On the right is provided the correspondence between numbers and families, as well as the six viral features considered in our analyses. A summary of the model statistics is provided in *SI Appendix*, Table S2.

potentially due to their lower stability in the environment (25). We suggest that this observation can be viewed as a consequence of the increased cross-species transmissibility exhibited by enveloped viruses. Hosts harbor a mixture of ancient vertically inherited viruses and viruses more recently acquired by cross-species transmission (26). According to our results, enveloped viruses should cause spillover infections more frequently than nonenveloped viruses. Moreover, even if spillovers are frequent, only a small fraction of these events result in successful transmission in the new host. It follows that the fraction of human-infective viruses that do not achieve sustained human–human transmission should be higher for enveloped viruses.

A recent work ranked 889 wildlife animal viruses according to the risk of animal-to-human spillover based on a systematic analysis of expert opinion (13). Interestingly, although the viral envelope was given a negative score in this analysis, 47 of the 50 top-listed viruses were enveloped, despite the fact that these represented only about half of the total viruses considered. This emphasizes how the importance of enveloped viruses has been overlooked. Indeed, the majority of remarkable zoonotic viruses in human history are enveloped, including poxviruses (e.g., monkeypox), morbilliviruses (e.g., measles), rhabdoviruses (e.g., rabies), coronaviruses (e.g., COVID), flaviviruses (e.g., Zika), orthomyxoviruses (e.g., influenza), and retroviruses (e.g., HIV).

Not surprisingly, we found a large variation in cross-species transmissibility among viral families. This could be in part due to methodological reasons. For instance, arenaviruses and coronaviruses have been subject to intense surveys in certain mammalian species but not others, potentially biasing down host breadth estimates. However, differences between families may also obey biological reasons. For example, host species counts were particularly elevated among hepeviruses. Although these viruses have been traditionally classified as nonenveloped, they also produce quasi-enveloped viral particles derived from the exosomal pathway (27). Although not addressed here, cross-species transmissibility and zoonotic potential can also vary strongly among viruses within a given family. For example, Middle East respiratory syndrome coronavirus appears to have jumped hundreds of times from camels to humans, whereas other coronaviruses have crossed the species barrier only once or a few times (28, 29). As mentioned above, zoonotic propensity depends on multiple factors in addition to purely virological features such as, for instance, availability of an intermediate domesticated host.

The mechanistic basis for the larger host breadth displayed by enveloped viruses remains to be investigated. Receptor-mediated viral entry is a critical stage in viral infection and cross-species transmission (30). Envelope proteins should be structurally less constrained than capsid proteins, and this might allow enveloped viruses to bind cellular receptors from different host species in a more flexible manner, to bind a greater number of alternative receptors, or to better accommodate host-switch mutations without compromising other functions. In addition, enveloped viruses can enter host cells through apoptotic mimicry, a process by which viral particles are engulfed by cells camouflaged as apoptotic bodies with a defined membrane lipid conformation (31). This process is relevant for many enveloped viruses, including alphaviruses, ebolaviruses, and poxviruses, all of which display a broad host range. Finally, it is also possible that envelopes contribute to cross-species transmissibility by helping viruses evade host immunity (32). Future work might elucidate whether these or other processes drive the increased ability of enveloped viruses to infect different hosts and cause zoonoses.

Materials and Methods

Dataset and Data Curation. Viruses associated with mammalian hosts were obtained from the VIRION database (24) (www.viralemergence.org/virion), last accessed in July 2022. This database uses different sources of evidence to assign hosts to viruses, but the vast majority of records in which a detection method is specified come from PCR- or sequence-based virus identification (97.5%); the remainder coming from direct virus isolations or serological tests. Duplicates were removed, resulting in unique virus–host associations. Viruses that were not classified at the family level, with unresolved NCBI taxonomy, with missing names or named after viral family only (-viridae sp.) were also filtered out. The total counts of nucleotide sequences available for each virus (including any lower taxonomic level) were obtained from the NCBI database, and viruses without nucleotide records were removed. The vast majority of viruses with a single sequence record originated from a single report and were associated with a single host species, although 1.2% were multihost viruses because evidence of infection was obtained by other methods (serological and PCR). Hosts not resolved at the species level or with uncertainty in their identification according to the VIRION database (tagged as HostFlagID = TRUE, not binomial scientific name) were also removed. Human-exclusive viruses were not included. Further manual curation was done to clear isolates suspected to infect only nonmammalian hosts according to the literature, such as viruses belonging to the family *Picobirnaviridae*, which were initially believed to infect animals but were later suggested to be bacteriophages (33). The family *Smacoviridae* was also removed due to its poor characterization, which did not allow assigning all features unambiguously.

Viral Features. Since the sequence information available for each virus was sufficient for taxonomical classification at the family level, fundamental features such as the genetic material (RNA/DNA), the presence of an envelope, the replication site (cytoplasmic/nuclear), genome strands (single/double), genome segmentation, and genome size could be assigned to each virus and were obtained from either ViralZone or ICTV. For genome size, we took the value corresponding to a prototypical member of the family (*SI Appendix*, Table S5). The family *Hepeviridae* was assumed to be nonenveloped, although these viruses also produce quasi-enveloped particles.

Statistical Analysis. The total count of host species was obtained for each virus, from which the binary response variable multihost was calculated. Zoonotic viruses were defined sensu lato as for those found in humans and at least one additional mammal species. The database contained no specific information about the direction and timing of host jump events. Binary predictors were DNA/RNA genome, enveloped/nonenveloped, nuclear/cytoplasmic replication, single/double stranded, and segmented/ nonsegmented. The covariates N and genome size were log-transformed. Generalized linear models using different distribution families and link functions were benchmarked (*SI Appendix*, Table S6). For host species counts, negative binomial models were judged more convenient than Poisson models due to data dispersion (*SI Appendix*, Fig. S3). For these models, square root and log link functions performed similarly, but we favored the canonical link function log. For the binary response variables multihost and zoonotic, we used binomial regression models, and we also selected the canonical link function log, making the model equivalent to a binary logistic regression. The selected link function performed similarly well as the cloglog and square root functions. We verified that, in all examined models, the presence/absence of an envelope was the viral trait that explained each of the three response variables with the highest significance ($P < 0.0001$ in all cases). We also explored generalized additive models, but these did not provide an obvious improvement in performance over generalized linear models and had the drawback of being less interpretable. For the negative binomial regression on host species counts that incorporated the viral family as a predictor, the 29 families having at least five viral species were included, and each family was treated as a binary factor, nested within the interaction term between the DNA/RNA genome, enveloped/nonenveloped, nuclear/cytoplasmic replication, single/double stranded, and segmented/nonsegmented factors. The marginal effect of each binary factor i, which could take values 0 or 1, was calculated as $M_i(0) = \exp(a + \sum_{j \neq i} b_j E(x_j))$ and $M_i(1) = \exp(a + \sum_{j \neq i} b_j E(x_j) + b_i)$, where a is the intercept term, b_i and b_j are the regression coefficients for factor i and another factor/covariate j, x_j is the value taken by factor/covariate j, and $E(x_j)$ is the average value of factor/covariate j. Notice that, when i was a viral family, all other binary factors were constant (i.e., $E(x_j) = 0$ or $E(x_j) = 1$) but not the N covariate, which was averaged. The exponential term appears because a log link function was used. Statistical analyses were performed with R functions glm, glm. nb (R package MASS), and gam (R package mgcv) and SPSS v22.

Data, Materials, and Software Availability. All study data are included in the article and/or *SI Appendix*. Open-source statistical packages (MASS and mgcv) can be accessed through CRAN repository. The SPSS software is available commercially.

ACKNOWLEDGMENTS. We thank Jérémy Dufloo, Pilar Domingo-Calap, and Iván Andreu-Moreno for helpful comments. This work was financially supported by Advanced Grant 101019724–EVADER and grant PID2020-118602RB-I00–ZooVir from the Spanish Ministerio de Ciencia e Innovación to R.S.

Author affiliations: [a]Institute for Integrative Systems Biology, Consejo Superior de Investigaciones Científicas-Universitat de València 46980, Paterna, València, Spain

1. C. J. Carlson, From PREDICT to prevention, one pandemic later. *Lancet Microbe* **1**, e6–e7 (2020).
2. C. J. Carlson *et al.*, The future of zoonotic risk prediction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200358 (2021).
3. M. Wille, J. L. Geoghegan, E. C. Holmes, How accurately can we assess zoonotic risk? *PLoS Biol.* **19**, e3001135 (2021).
4. R. K. Plowright *et al.*, Pathways to zoonotic spillover. *Nat. Rev. Microbiol.* **15**, 502–510 (2017).
5. M. Woolhouse, F. Scott, Z. Hudson, R. Howey, M. Chase-Topping, Human viruses: Discovery and emergence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 2864–2871 (2012).
6. F. Keesing, R. S. Ostfeld, Impacts of biodiversity and biodiversity loss on zoonotic diseases. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2023540118 (2021).
7. L. Zhang *et al.*, Biological invasions facilitate zoonotic disease emergences. *Nat. Commun.* **13**, 1762 (2022).
8. K. N. Shivaprakash, S. Sen, S. Paul, J. M. Kiesecker, K. S. Bawa, Mammals, wildlife trade, and the next global pandemic. *Curr. Biol.* **31**, 3671–3677.e3 (2021).
9. S. Guth, E. Visher, M. Boots, C. E. Brook, Host phylogenetic distance drives trends in virus virulence and transmissibility across the animal-human interface. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **374**, 20190296 (2019).
10. C. Kreuder Johnson *et al.*, Spillover and pandemic properties of zoonotic viruses with high host plasticity. *Sci. Rep.* **5**, 14830 (2015).
11. B. A. Han, J. P. Schmidt, S. E. Bowden, J. M. Drake, Rodent reservoirs of future zoonotic diseases. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7039–7044 (2015).
12. S. A. Babayan, R. J. Orton, D. G. Streicker, Predicting reservoir hosts and arthropod vectors from evolutionary signatures in RNA virus genomes. *Science* **362**, 577–580 (2018).
13. Z. L. Grange *et al.*, Ranking the risk of animal-to-human spillover for newly discovered viruses. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2002324118 (2021).
14. N. Mollentze, S. A. Babayan, D. G. Streicker, Identifying and prioritizing potential human-infecting viruses from their genome sequences. *PLoS Biol.* **19**, e3001390 (2021).
15. J. M. Bartoszewicz, A. Seidel, B. Y. Renard, Interpretable detection of novel human viruses from genome sequencing data. *NAR Genom. Bioinform.* **3**, lqab004 (2021).
16. S. Cleaveland, M. K. Laurenson, L. H. Taylor, Diseases of humans and their domestic mammals: Pathogen characteristics, host range and the risk of emergence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **356**, 991–999 (2001).
17. C. R. Parrish *et al.*, Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiol. Mol. Biol. Rev.* **72**, 457–470 (2008).
18. K. J. Olival *et al.*, Host and viral traits predict zoonotic spillover from mammals. *Nature* **546**, 646–650 (2017).
19. J. R. C. Pulliam, J. Dushoff, Ability to replicate in the cytoplasm predicts zoonotic transmission of livestock viruses. *J. Infect. Dis.* **199**, 565–568 (2009).
20. K. Wells, S. Morand, M. Wardeh, M. Baylis, Distinct spread of DNA and RNA viruses among mammals amid prominent role of domestic species. *Glob. Ecol. Biogeogr.* **29**, 470–481 (2020).
21. R. E. Grewelle, Larger viral genome size facilitates emergence of zoonotic diseases. bioRxiv [Preprint] (2020). https://doi.org/10.1101/2020.03.10.986109, last accessed on July 15, 2022.
22. D. Carroll *et al.*, The global virome project. *Science* **359**, 872–874 (2018).
23. E. Harvey, E. C. Holmes, Diversity and evolution of the animal virome. *Nat. Rev. Microbiol.* **20**, 321–334 (2022).
24. C. J. Carlson *et al.*, The global virome in one network (VIRION): An atlas of vertebrate-virus associations. *mBio* **13**, e0298521 (2022).
25. J. L. Geoghegan, A. M. Senior, F. Di Giallonardo, E. C. Holmes, Virological factors that increase the transmissibility of emerging human viruses. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4170–4175 (2016).
26. J. L. Geoghegan, S. Duchêne, E. C. Holmes, Comparative analysis estimates the relative frequencies of co-divergence and cross-species transmission within viral families. *PLoS Pathog.* **13**, e1006215 (2017).
27. S. Nagashima *et al.*, Characterization of the quasi-enveloped hepatitis E virus particles released by the cellular exosomal pathway. *J. Virol.* **91**, e00822–17 (2017).
28. G. Dudas, L. M. Carvalho, A. Rambaut, T. Bedford, MERS-CoV spillover at the camel-human interface. *Elife* **7**, e31257 (2018).
29. N. D. Rochman, Y. I. Wolf, E. V. Koonin, Molecular adaptations during viral epidemics. *EMBO Rep.* **23**, e55393 (2022).
30. C. J. Warren, S. L. Sawyer, How host genetics dictates successful viral zoonosis. *PLoS Biol.* **17**, e3000217 (2019).
31. A. Amara, J. Mercer, Viral apoptotic mimicry. *Nat. Rev. Microbiol.* **13**, 461–469 (2015).
32. J. P. Buchmann, E. C. Holmes, Cell walls and the convergent evolution of the viral envelope. *Microbiol. Mol. Biol. Rev.* **79**, 403–418 (2015).
33. S. R. Krishnamurthy, D. Wang, Extensive conservation of prokaryotic ribosomal binding sites in known and novel picobirnaviruses. *Virology* **516**, 108–114 (2018).