Cross-Species Transmission of Bat Coronaviruses in the Americas: Contrasting Patterns between Alphacoronavirus and Betacoronavirus

Diego A. Caraballo[a,b]

aCONICET-Universidad de Buenos Aires, Instituto de Ecología, Genética y Evolución de Buenos Aires (IEGEBA), Ciudad Universitaria-Pabellón II, Buenos Aires, Argentina

bUniversidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Buenos Aires, Argentina

ABSTRACT Bats harbor the largest number of coronavirus (CoV) species among mammals, serving as major reservoirs of alphaCoVs and betaCoVs, which can jump between bat species or to different mammalian hosts, including humans. Bat-CoV diversity is correlated with host taxonomic diversity, with the highest number of CoV species found in areas with the highest levels of bat species richness. Although the Americas harbor a unique and distinctive CoV diversity, no cross-species transmission (CST) or phylogeographic analysis has yet been performed. This study analyzes a large sequence data set from across the Americas through a Bayesian framework to understand how codivergence and cross-species transmission have shaped long-term bat-CoV evolution and ultimately identify bat hosts and regions where the risk of CST is the highest. Substantial levels of CST were found only among alphaCoVs. In contrast, cospeciation prevailed along the evolution of betaCoVs. Brazil is the center of diversification for both alpha and betaCoVs, with the highest levels of bat species richness. The bat family Phyllostomidae has played a key role in the evolution of American bat-CoVs, supported by the highest values of host transition rates. Although the conclusions drawn from this study are supported by biological/ecological evidence, it is likely that novel lineages will be discovered, which could also reveal undetected CSTs given that sequences are available from 11 of the 35 countries encompassing the Americas. The findings of this study can be useful for conducting targeted discovery of bat-CoVs in the region, especially in countries of the Americas with no reported sequences.

IMPORTANCE Coronaviruses (CoVs) have a strong zoonotic potential due to their high rates of evolvability and their capacity for overcoming host-specific barriers. Bats harbor the largest number of CoV species among mammals, with the highest CoV diversity found in areas with the highest levels of bat species richness. Understanding their origin and patterns of cross-species transmission is crucial for pandemic preparedness. This study aims to understand how bat-CoVs diversify in the Americas, circulate among and transmit between bat families and genera, and ultimately identify bat hosts and regions where the risk of CoV spillover is the highest.

KEYWORDS bats, codivergence, coronavirus, cross-species transmission, host shift, phylogeny, spillover, virus

Bats (order Chiroptera) include more than 1,400 species representing 20% of mammal diversity (1). They are also among the most widespread and abundant mammals. Together with their unique capacity for flying, some of their characteristics, such as food choices, population structure, seasonal migration, and roosting behaviors, make them particularly suitable hosts and spreaders of viruses and other disease agents (2).

RNA viruses have high mutation rates that allow especially fast evolution and can produce sporadic human epidemics overcoming host-specific barriers, resulting in the following different outcomes: single spillover events such as rabies virus (RABV) infections in humans (3), transient endemic outbreaks such as Nipah virus (4), transient epidemics such as the severe acute respiratory syndrome coronavirus (SARS-CoV) (5), or pandemic outbreaks such as SARS-CoV-2 (6). Cross-species transmissions (CSTs) are relatively frequent events that depend only on prevalence and contact rates between zoonotic host and nonhost species. However, the vast majority of CSTs are single spillover events, which epidemiologically are dead ends. Long-term establishment (when the virus is maintained by a novel host species) are rare events and depend not only on prevalence and contact rates but also on the biological interaction (and evolvability) of the virus and the new host (7).

Coronaviruses (order, *Nidovirales*; family, *Coronaviridae*), single-stranded positive-sense RNA viruses with the largest nonsegmented RNA viral genomes (16 to 31 kb), are divided into two subfamilies, *Orthocoronavirinae* and *Letovirinae*. The former includes the genera *Alphacoronavirus* (alphaCoV); *Betacoronavirus* (betaCoV), coronaviruses related to Middle East respiratory syndrome CoV (MERS-CoV), SARS-CoV, and SARS-CoV-2; *Gammacoronavirus* (gammaCoV); and *Deltacoronavirus* (deltaCoV). Alpha and betaCoVs are found in mammals and likely originated in bats, while the latter two are found and originated in birds (8). Coronaviruses are able to jump cross-species barriers and rapidly adapt to new hosts, likely due to their large genome size, high recombination rates, and genomic plasticity (9).

Bats harbor the largest number of CoV species among mammals, with alphaCoVs more widespread and abundant than betaCoVs (10). Bat-CoV diversity is correlated with host taxonomic diversity, with the highest CoV diversity found in areas with the highest levels of bat species richness (11, 12). Bat-CoVs are able to jump to different mammalian hosts, including humans (8). The emergence of SARS-CoV-2 highlights the relevance of bat-CoVs to global health and why understanding their origin and patterns of CST is crucial for pandemic preparedness (13). In the Americas, bat-CoVs have been detected in 11 countries as of 2021 (14). Both alpha and betaCoVs have been reported, with a predominance of the former (15–20). Although the Americas harbor a unique and distinctive CoV diversity, no cross-species transmission or phylogeographic analysis has yet been performed.

This study analyzes all bat-CoV sequences from the Americas available up to 2021. A phylogeographic Bayesian statistical framework was applied to model virus transmission between bat host species and the spatial location of bat-CoVs over time. This study aims to understand how bat-CoVs diversify, circulate among, and transmit between bat families and genera and ultimately identify bat hosts and regions where the risk of CoV CST is the highest.

RESULTS

American bat-CoVs in the context of global CoV diversity. A total of 7 well-supported reciprocally monophyletic groups were retrieved for bat alphaCoVs in the Americas, all of them characterized for circulating in one or two main host species (Fig. 1). These clades are interspersed with other global alphaCoV lineages. Except for the sequence isolated from *Eumops glaucinus*, the American bat betaCoV lineage is monophyletic, supporting a common origin for this group.

Alpha coronaviruses. The American bat alphaCoVs were divided into seven well-supported clades, each of them defined by their main hosts as follows: *Molossus* (A), *Myotis* (B, C), and several Phyllostomidae genera (D to G) (*Anoura*, *Artibeus*, *Carollia*, *Desmodus*, *Glossophaga*, *Mesophylla*, *Phyllostomus*, and *Sturnira*). Two types of CSTs can be identified as follows: (i) long-term CSTs, which can be identified as monophyletic groups composed of a novel host; and (ii) dead ends, which are typically single sequences representing transmission events that did not produce long-term establishment in a novel host. Within clade G, involving Phyllostomidae, the following long-term CSTs can be identified: *Artibeus→Carollia*, *Artibeus→Phyllostomus*, *Phyllostomus→Desmodus*,
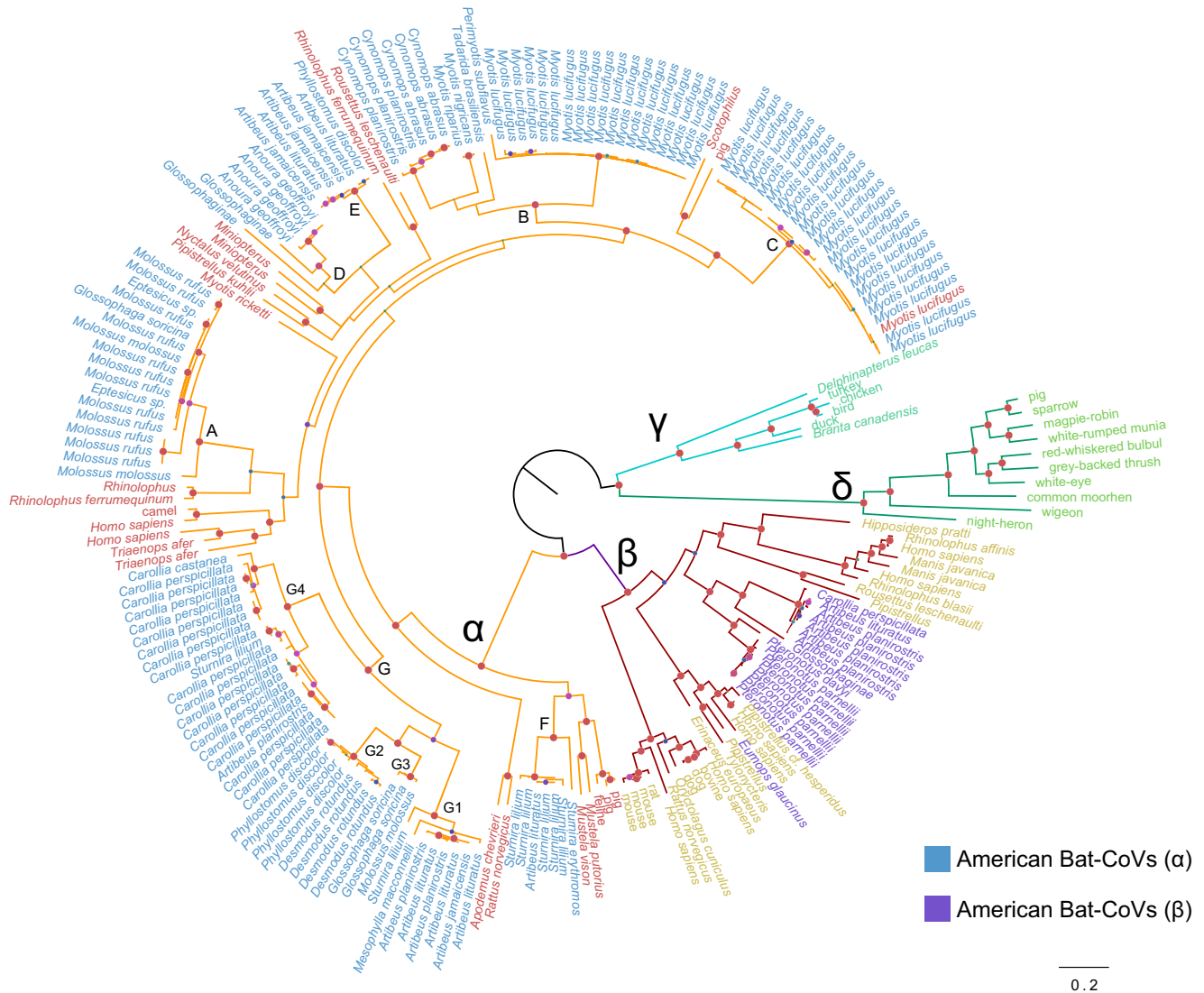
FIG 1 American bat-CoVs in the context of global CoV diversity. Bayesian phylogeny showing global alpha, beta, delta, and gammacoronavirus. Sequences of American bat-CoVs are shown in blue (alphaCoVs) and purple (betaCoVs). Seven monophyletic American bat alphaCoVs are shown with letters A to G, and subclades of the major lineage G are shown with numbers 1 to 4. Nodes with Bayesian posterior probability >0.9 are shown in red. The scale bar is expressed in substitutions/site.

Artibeus→Glossophaga (Fig. 2). Another long-term CST can be identified from *Myotis* to *Cynomops* (clade B), which is also supported by the geographic location of the isolates since both donor and recipient sequences are from Brazil (see Supplementary Figure 1 and Supplementary Table 1 posted at https://doi.org/10.6084/m9.figshare.c.6029390.v1).

The majority of CSTs were observed in terminal branches, which likely correspond to dead ends in the transmission chain. All of these CSTs took place between species of the same (or neighboring) countries (Fig. 1; see also Supplementary Figure 1 posted at https://doi.org/10.6084/m9.figshare.c.6029390.v1), an expected result since a spillover would require physical contact. In clade A, *Molossus→Eptesicus* and *Molossus→Glossophaga* CSTs took place. It is worth noting that species from different families are involved in these events, the donor being Molossidae, while the recipients being Vespertilionidae and Phyllostomidae. These CSTs occurred between Brazilian bats. In clade B, *Myotis→Perimyotis* and *Myotis→Tadarida* CSTs took place. The donor sequences are from Canada, while the recipient sequences are from the United States.

A number of CSTs were also found among Phyllostomidae. The *Glossophaga→Anoura* CST (clade D) is plausible since there are sequences of both genera from the same region
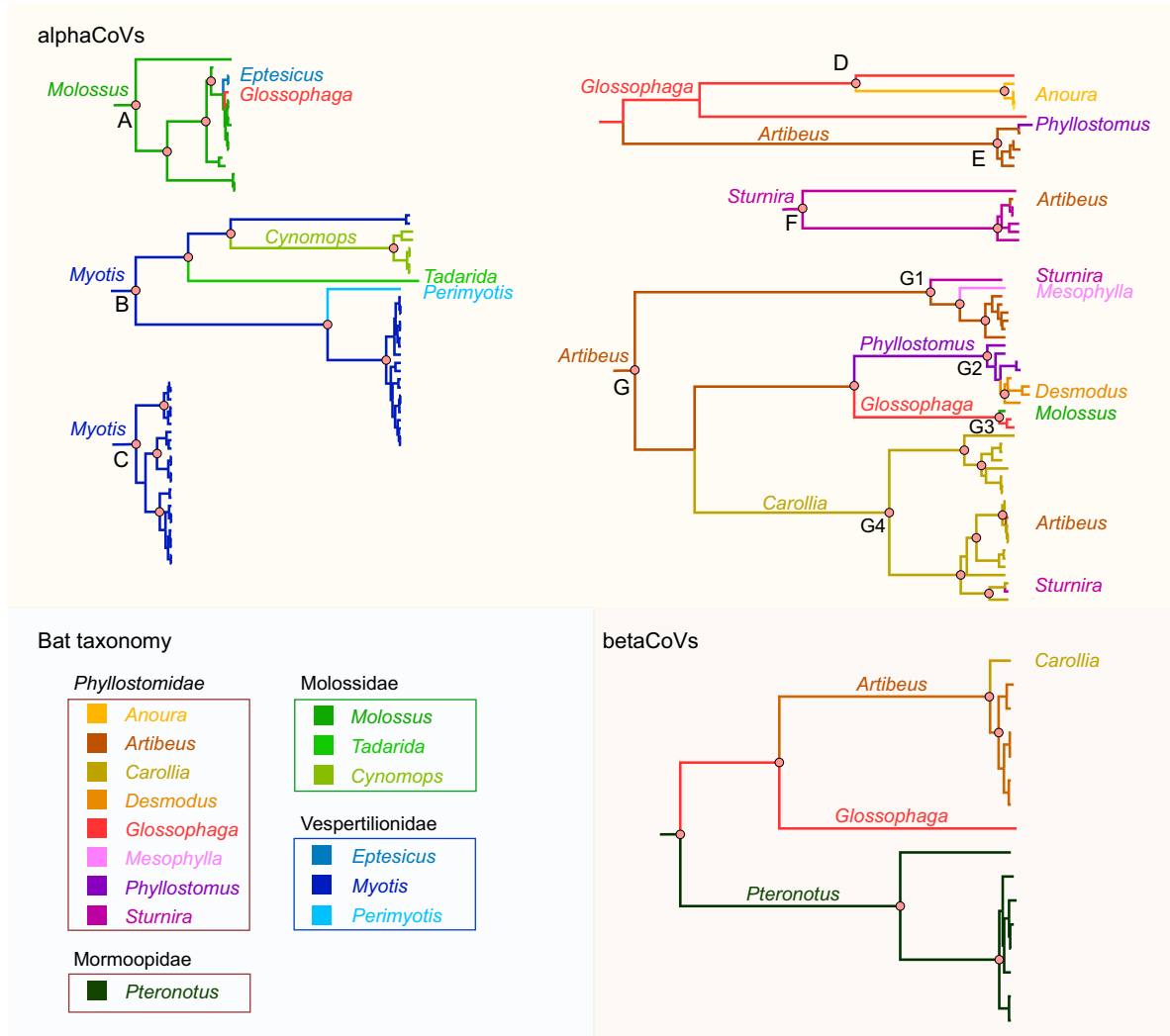
FIG 2 Ancestral host reconstruction of alphaCoVs and betaCoVs. Maximum clade credibility trees annotated using partial RdRp sequences and bat host genus as discrete character state. Branch colors correspond to the inferred ancestral bat genus with the highest posterior probability. Relevant nodes with posterior probability >0.9 are indicated with pink circles.

(Costa Rica). In clade E, an *Artibeus*→*Phyllostomus* CST was identified. It is interesting to note that in this clade, there are sequences from Brazil, Panama, and Costa Rica, but the *Artibeus* and *Phyllostomus* sequences involved in the CST are both from Brazil. In clade F, *Sturnira* sequences are split into two groups reflecting geographic divergence. The *Sturnira*→*Artibeus* CST found in this clade occurred between sequences isolated from Brazilian bats. In clade G1, *Artibeus*→*Mesophylla* and *Artibeus*→*Sturnira* CSTs were found, supported by the isolation of variants from members of the three genera in Brazil. In clade G3, a *Glossophaga*→*Molossus* CST was found, involving Brazilian bats, and was the only interfamilial spillover found in the Phyllostomidae clades. Two additional CSTs were found in clade G4, *Carollia*→*Sturnira* and *Carollia*→*Artibeus*. All *Carollia* sequences split into two groups, one from Costa Rica, and the other from Brazil and Bolivia, reflecting the geographic structure of bat alphaCoVs. The CST events occurred between bats from Brazil (Fig. 2).

The cophylogeny shows both deep and terminal CSTs (Fig. 3). AlphaCoV clades A and B are, in general, congruent with Molossidae/Vespertilionidae phylogenetic relationships. Clades D to G include exclusively (except for terminal CSTs) variants found among Vespertilionidae. The fact that many species (genera *Artibeus*, *Desmodus*, *Sturnira*, and
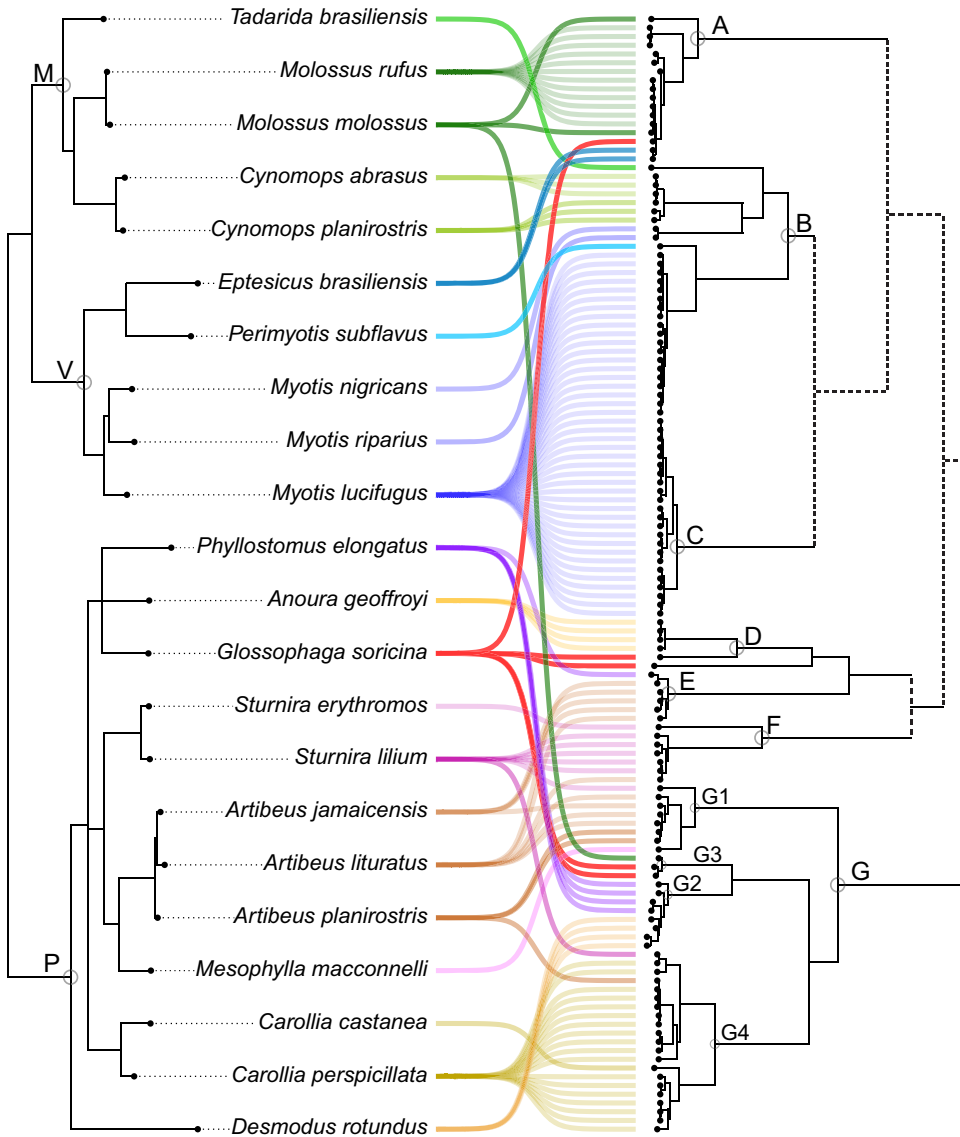
FIG 3 Cophylogeny plot of bat species phylogeny and alphaCoVs tree. Cyt *b*-based bat species phylogeny (left) and alphaCoVs phylogeny (right). Gray circles indicate bat families (M, Molossidae; V, Vespertilionidae; P, Phyllostomidae) and main alphaCoV lineages and sublineages (A to G). Bat-CoV links are colored by each bat family. Dotted lines in the alphaCoV phylogeny depict nonsupported clusterings.

Glossophaga) have alphaCoVs from more than one clade (E, F, and G) indicates that these species maintain different alphaCoV variants.

The Bayes factor (BF) test enabled the distinction of three host-switching rate categories (Fig. 4). The transitions *Glossophaga→Molossus, Artibeus→Sturnira, Artibeus→Phyllostomus, Carollia→Artibeus,* and *Molossus→Eptesicus* received the highest support (BF > 100), all of them reflecting terminal CSTs. *Glossophaga→Molossus* is the host transition with the highest BF support, which is an expected result since CSTs in both directions were found (Fig. 2 and 3). The deep CST *Phyllostomus→Desmodus* also received high support (BF > 100).

There are six additional host switching rates that received strong BF support (10 < BF < 100), *Myotis→Perimyotis, Carollia→Sturnira, Artibeus→Mesophylla, Myotis→Tadarida, Myotis→Cynomps,* and *Glossophaga→Anoura.* Except for the latter two, which correspond to deep CSTs, the remaining are terminal CSTs. One last transition rate, with moderate BF support (3 < BF < 10), *Artibeus→Glossophaga,* can be considered. It is notable that *Artibeus* and *Molossus* function as both donor and receptor species in terminal CSTs.
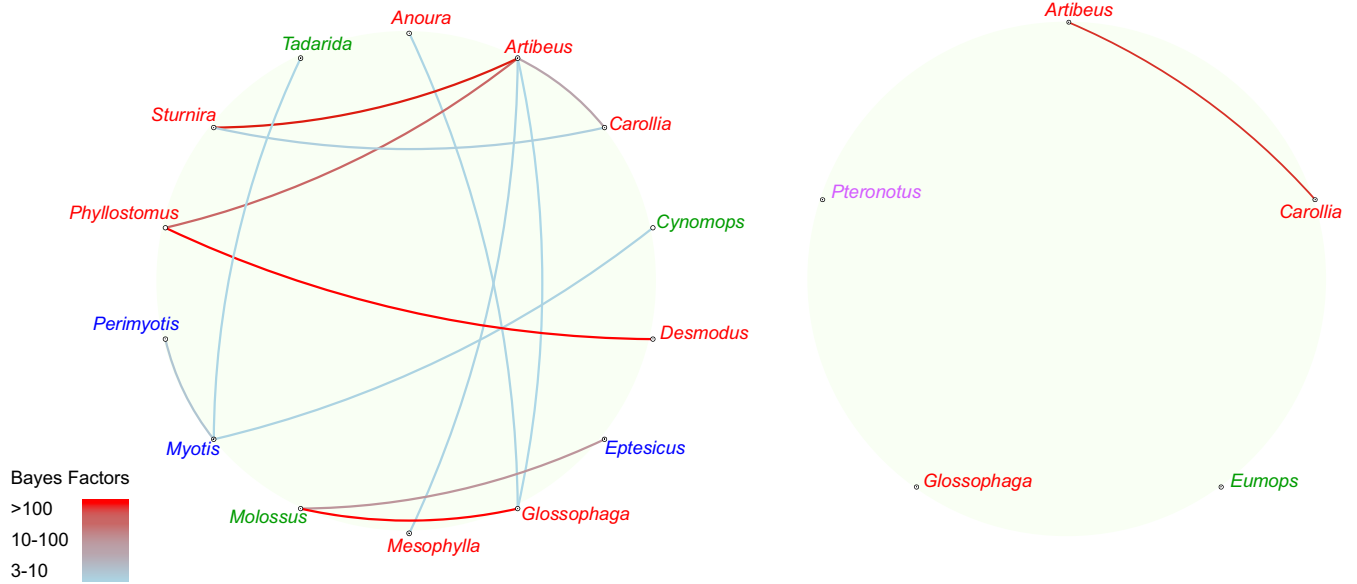
FIG 4 Host transition rates among American bat genera. Strongly supported host switches between bat genera for alphaCoVs (left) and betaCoVs (right). Line color indicates the switch significance level according to the Bayes factor test. Genus color indicates host family (green, Molossidae; red, Phyllostomidae; blue, Vespertilionidae; purple, Mormoopidae).

Betacoronaviruses. The phylogenetic structure of the betaCoV tree shows a general correspondence with the host species phylogeny (Fig. 2). It can be divided into two well-supported clades defined by different host families, Phyllostomidae and Mormoopidae, also reflecting internal phylogenetic relationships. There are no observed terminal CSTs. The cophylogeny reflects a total correspondence between host and virus trees (Fig. 5), while only one host switching rate, *Artibeus→Carollia*, received strong BF support (10 < BF < 100), denoting low levels of CSTs found among betaCoVs (Fig. 3).

The cospeciation test yielded two maximum parsimony reconciliations (MPRs), one under each cost set (Fig. 6). Cost set *a* (2, 6, and 1; for duplications, transfers, and losses, respectively), invokes 5 cospeciation events, 10 duplications, and 1 loss. Cost set *b* (4, 2, and 1), hypothesizes 5 cospeciation events, 9 duplications, and 1 transfer. The *P* value of the randomization tests was <0.01 under cost set *a*, refuting the null hypothesis that host and virus trees are similar by chance. The MPR of condition *b* was not statistically different from randomized associations. These results indicate that the diversification of betaCoVs and American bats can be fully explained by cospeciation (and duplication/loss) events in the absence of host shifts.

DISCUSSION

This study presents the first phylogenetic and phylodynamic analysis of bat-CoVs at a continental scale in the Americas. High levels of CoV diversity were found, which is expected, given that 75% of bat living genera are found in this region (1) and that bat-CoV diversity correlates with host taxonomic diversity. This is further supported by the fact that Brazil harbors the richest diversity of bat-CoVs found to date in the region (14; this study) as well as the highest levels of bat species richness at a global scale (21) (Fig. 7).

The results presented in this study confirm that the long-term evolution of bats and CoVs result from the interaction between codivergence and cross-species transmission, consistent with previous studies of bat-CoVs in other regions of the World (22–24). The phylogenetic analysis shows that, in both alpha and betaCoVs, there is at least some degree of species- and genus-specific tropism (25). However, both deep and shallow
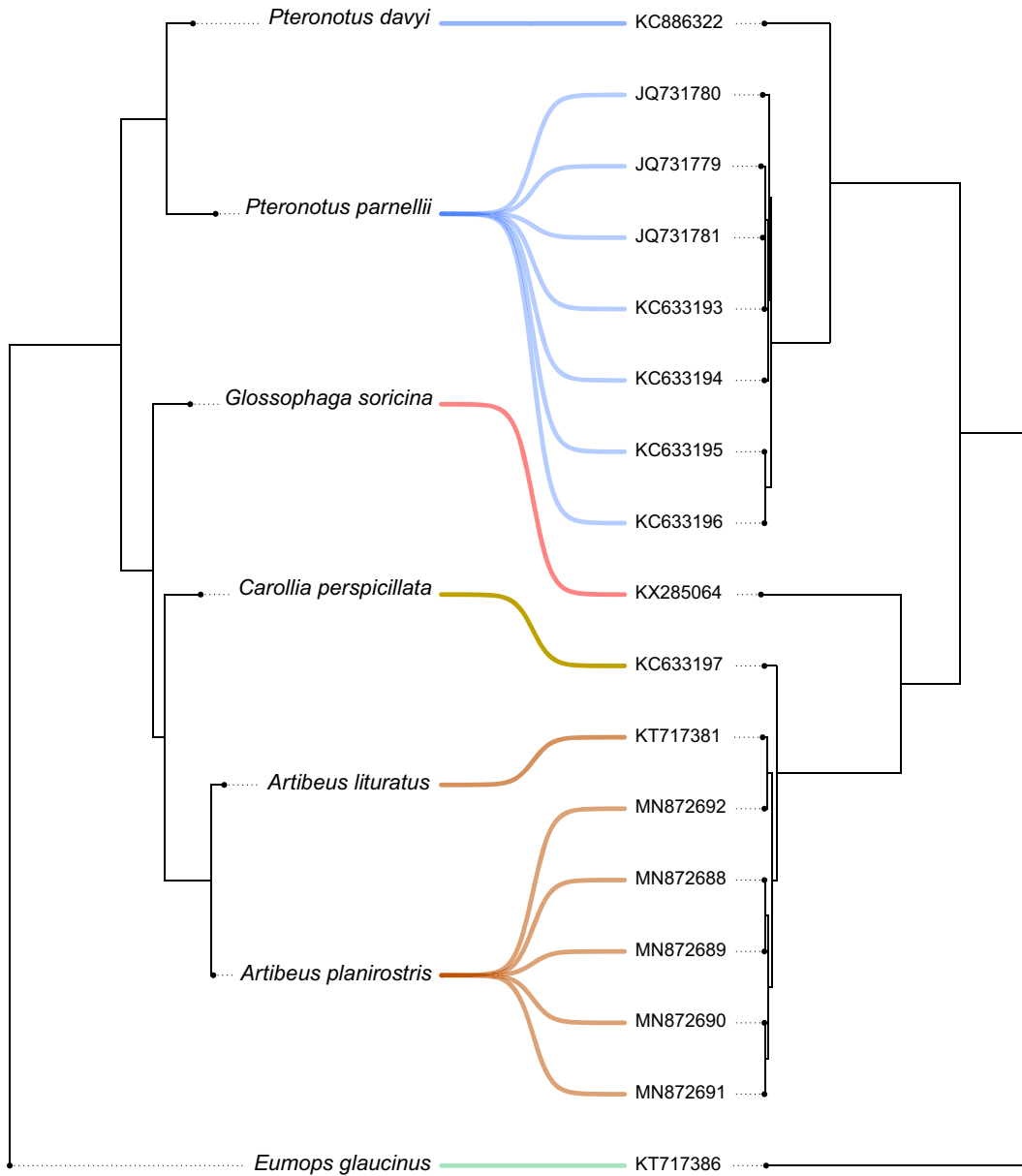
FIG 5 Cophylogeny plot of bat species phylogeny and betaCoVs tree. Cyt *b*-based bat species phylogeny (left) and alphaCoVs phylogeny (right). Bat-CoV links are colored by each bat family (blue, Mormoopidae; green, Molossidae; red-yellow-brown, Phyllostomidae). Dotted lines in the betaCoV phylogeny depict nonsupported clusterings.

CSTs prevailed among alphaCoVs, as evidenced by the lack of monophyly of these sequences within the global CoV tree (Fig. 1) and by a high number of terminal CSTs between nonrelated species (Fig. 2). In contrast, the existence of genetically related viruses found among related bat species (or conspecifics) in distant sampling locations reflects codivergence between CoVs and bat species. This is the case of *Myotis*, *Sturnira*, *Phyllostomus*, *Carollia*, and *Artibeus* alphaCoVs and *Pteronotus* betaCoVs (see Supplementary Figure 1 posted at https://doi.org/10.6084/m9.figshare.c.6029390.v1).

The *Betacoronavirus* cophylogeny is entirely congruent, showing that this virus genus is less prone to host switching compared to *Alphacoronavirus*, which is in agreement with previous findings (23). The diversification of betaCoVs accompanied the diversification of host species, at the family, genus, and species levels (Fig. 2 and 5). Only cospeciation and duplication/loss events were recovered in the MPR, underlying the prevailing role of codiversification and discarding any CST in this viral lineage.
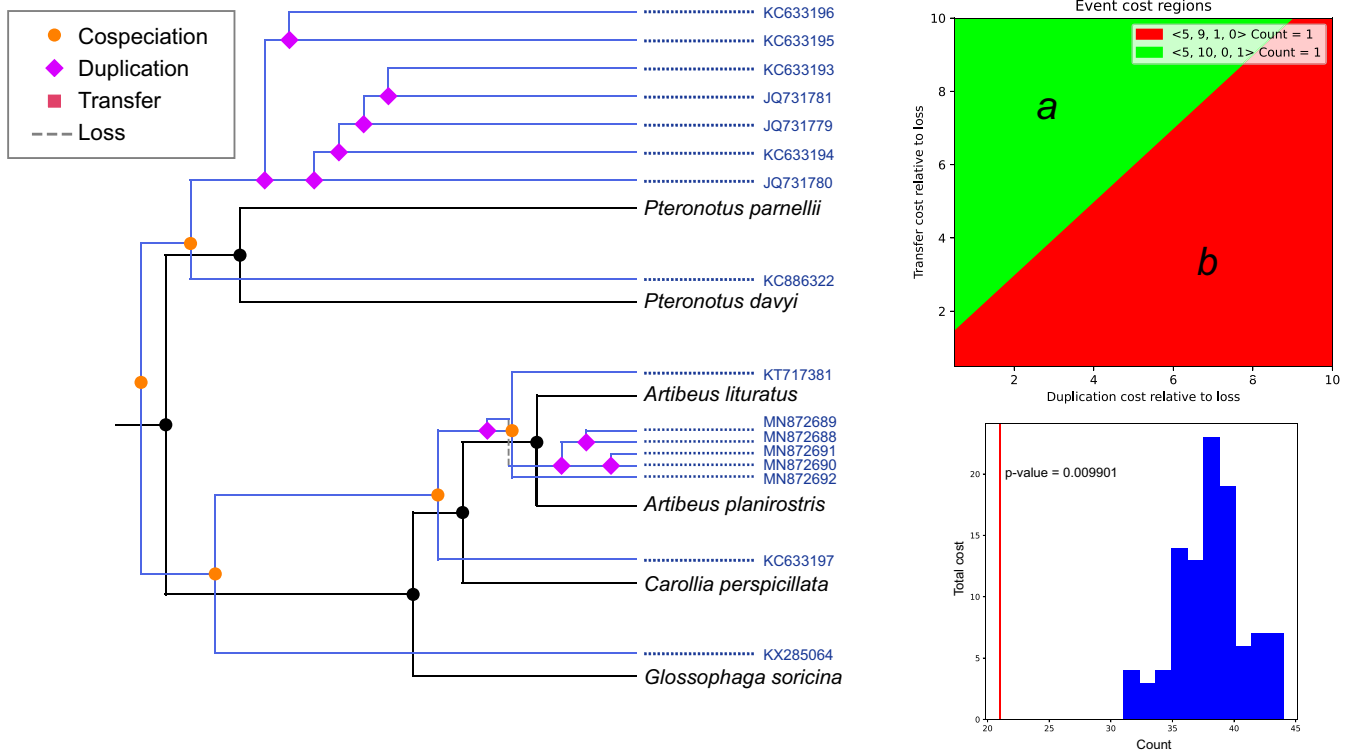
FIG 6 Phylogenetic reconciliation between betaCoVs and bat hosts trees. Maximum parsimony reconciliation (left) proposed for the cost region *a* (upper right) and its significance by randomization tests (lower left). Cospeciations and duplications/losses but not host switching took place in the evolution of betaCoVs. The MPR obtained under scheme *b* (not shown) was not significantly different from the randomized MPR distribution.

It is interesting to note that inferred CSTs took place between species distributed in the same region. Deep CSTs occurring between bats from different families are unambiguous host shifts. Such host shifts were found in alphaCoVs between *Myotis* (Vespertilionidae) and *Cynomops* (and probably *Tadarida*) (Molossidae) (Fig. 2). In addition, several host shifts have been identified among Phyllostomidae as mentioned above (Fig. 2). Terminal CSTs took place both between closely related host species and between members of different bat families (Fig. 2). At least 10 terminal CSTs were detected among 128 alphaCoV sequences, denoting the high promiscuity of this virus genus in bats. These high CST rates were statistically corroborated; all bat genera included in the analysis depicted significant host-switching rates (Fig. 4). Except for *Glossophaga→Molossus*, all highly supported host switching rates were found between members of Phyllostomidae (Fig. 4), indicating that this family played a key role in the evolution, diversification, and cross-species transmission history of alphaCoVs in the Americas. It is notable that members of this family, such as *Carollia*, *Phyllostomus*, and *Glossophaga*, have been shown to share mixed-species roosts between them and with other species (26). The propensity of interspecific communal roosting has been attributed to benefits, such as protection against predation, information transfer on the location of foraging sites, and reduced thermoregulatory costs, but these benefits have the counterpart of higher chances for CST (27; this study).

Despite its large scale, this study has some limitations that should be considered. First, it is based on partial RNA-dependent RNA polymerase (RdRp) sequences since the majority of bat-CoV sequences available in public databases were generated with the primers designed by Watanabe et al. (28), which produce a fragment of 440 bp. The RdRp gene is a suitable marker for phylogenetic and phylogeographic analysis since it reflects vertical ancestry and is less prone to recombination than other regions of the CoV genome (29). However, the short length of the sequences is less than half of the 816 bp necessary for RGU (RdRp-based grouping units) classification (30) and is detrimental to phylogenetic resolution. It is worth noting that the general structure of the viral phylogenies inferred in

Bat species

alphaCoV host species

betaCoV host species
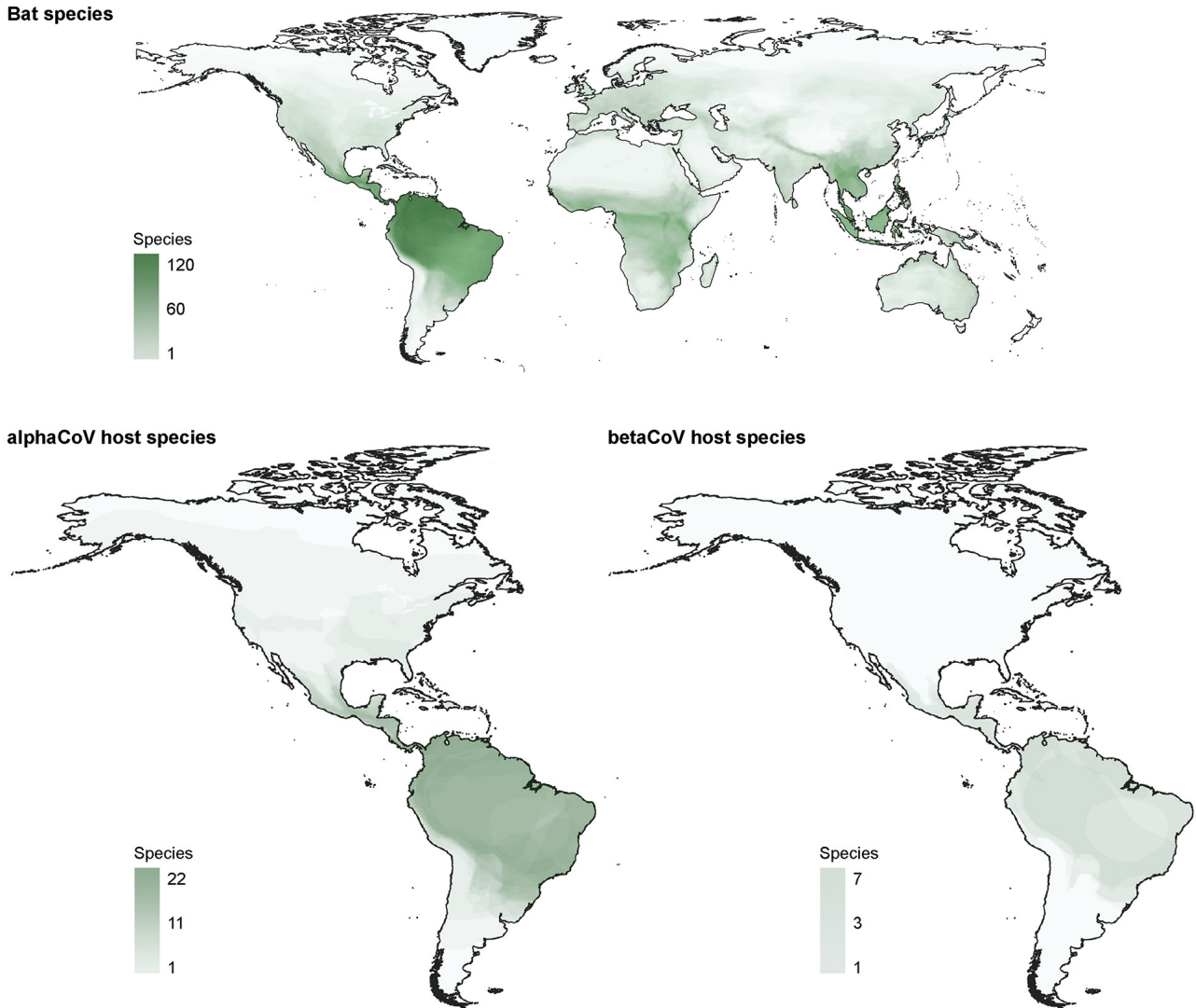
FIG 7 Bat species distributions. Global bat species distributions (top) and distribution of bat-CoV hosts in the Americas (bottom). Geographical ranges of bat species were obtained from the International Union for Conservation of Nature (IUCN). The plots display the number of bat species based on overlapping geographical ranges from a total of 1,317 species with IUCN range data as of March 2022.

this study showed high levels of resolution with highly supported nodes. Another aspect that should be taken into consideration is the uncertainty in host species classification due to procedural errors or to the cryptic nature of several bat species that cannot be unambiguously identified from their external morphology (31). This issue was solved by centering the analysis on genus and family levels (Fig. 2, Fig. 4) but keeping the species-level resolution in the cophylogeny (Fig. 3, Fig. 5).

In this study, both host taxa and geographic regions that define the hot spot of CoV phylogenetic diversity in the Americas were identified. Namely, for both alpha and betaCoVs, Brazil is the center of diversification, with the highest number of lineages as well as the highest level of bat species richness (Fig. 7; see also Supplementary Figure 1 posted at https://doi.org/10.6084/m9.figshare.c.6029390.v1). Brazil has a humid tropical and subtropical climate, which has been demonstrated to promote higher viral evolutionary rates than other climates in other RNA viruses, such as rabies, but also in CoVs (23, 32). These regions are characterized by rapidly urbanizing human populations as well as high levels of livestock and poultry production, which, in addition to higher levels of CSTs and viral evolution, could favor disease emergence (33). Although

the conclusions drawn from this study are supported by biological/ecological evidence (e.g., CSTs occurred between species cooccurring in the same region, and higher rates were found between species sharing mixed-species roosts), this may not represent the whole picture since there are sequences available from 11 of the 35 countries encompassing the Americas. It is likely that novel lineages will be discovered, circulating in both previously identified as well as in novel hosts, which could also reveal undetected CSTs. Therefore, substantially wider sampling efforts will be required to encompass the complete diversity of CoVs in bats, especially in countries of the Americas that have not reported sequences yet.

As pointed out, Phyllostomidae has played a key role in the long-term evolution of bat-CoVs in the Americas, with the highest observed host transition rates. These findings can be useful for conducting targeted discovery of bat-CoVs in Brazil and neighboring countries as well as for the implementation of surveillance programs for early detection of outbreaks in people or livestock. In particular, the bat genera *Phyllostomus*, *Artibeus*, *Sturnira*, and *Glossophaga* deserve special attention in alphaCoV discovery. Coronavirus research in bats and other wildlife or domestic species will continue to be important for the identification of natural reservoirs and potential sources of spillover into the human population.

MATERIALS AND METHODS

Data acquisition and virus typing. A total of 195 CoV nucleotide sequences were retrieved from the Virus Pathogen Database and Analysis Resource (ViPR) (34) through the website http://www.viprbrc.org/, filtering for *Orthocoronavirinae*, bats, and the Americas in the subfamily, host, and location fields, respectively. The resulting sequences covered the period of 2010 to 2019 and were mostly partial sequences encoding for the RNA-dependent RNA polymerase (RdRp) protein (see Supplementary Table 1 posted at https://doi .org/10.6084/m9.figshare.c.6029390.v1). All supplemental material is available at figshare (https://doi.org/10 .6084/m9.figshare.c.6029390.v1). The majority of the sequences were obtained with the primers designed by Watanabe et al. (28), frequently used in molecular diagnostic protocols. A total of 146 sequences encompassing this fragment were aligned (minimum length, 229 bp; maximum length, 2,779 bp; mean length, 462 bp) with Clustal Omega (35) and tested for phylogenetic affinity using reference sequences for alpha, beta, gamma, and delta CoVs representing the global CoV diversity based on a recently published data set (36). To this end, 1E8 Markov chain Monte Carlo (MCMC) generations were run in MrBayes v3.2.7 (37), under the GTR+I+G model selected with MrModeltest v2 for each codon position separately (38), discarding the first 25% of the run as burn-in. Trees were visualized in FigTree v1.4.4 (39). All tested sequences belonged to either alphaCoV (129 sequences) or betaCoV (17 sequences). The final data set included 145 sequences with their associated host species (Appendix; see also Supplementary Table 1 posted at https://doi.org/10.6084/ m9.figshare.c.6029390.v1). Host, country, and collection date metadata associated with each sequence was retrieved using the rentrez package (40), using the R environment (41) and RStudio (42).

Phylogeographic and host-switching analysis. The spatial and CST dynamics of alphaCoVs (128 sequences) and betaCoVs (17 sequences) were reconstructed using the software BEAST v1.10.4 (43). The American bat alphaCoV data set was divided into seven reciprocally monophyletic groups, while 16 of 17 American bat betaCoV sequences were monophyletic (Fig. 1). Each monophyletic group (except clades D and E, which comprise a single higher-order clade) was treated separately in the cross-species transmission analyses.

Sampling dates were used for molecular clock calibrations. Selected substitution models were HKY+I+G and HKY+I for alphaCoVs and betaCoVs, respectively. Base frequencies, substitution rate parameters, and rate heterogeneity models were unlinked across codon positions. A symmetric substitution model was selected for host and country ancestral state reconstruction, while a Bayesian stochastic search variable selection (BSSVS) procedure was chosen to identify hosts transition rates that adequately explain the phylogenetic diffusion process. A strict clock and a constant population size coalescent prior were used for these analyses (Appendix). Convergence and mixing of the MCMC were analyzed using Tracer v1.7.2 (44), combining two independent runs with LogCombiner v1.10.4. Trees were summarized with the maximum clade credibility (MCC) option, discarding 10% of the run as burn-in, using TreeAnnotator v1.10.4. Effective sample sizes (ESS) were >200 for all parameters.

The program SpreaD3 v0.9.7.1 (45) was used to visualize the output of the Bayesian phylogeographic and host-switching analyses. The MCC trees obtained were used to analyze host transition rates through a Bayes factor test.

Cophylogeny. Host phylogenies were obtained based on cytochrome *b* (Cyt *b*) nucleotide sequences retrieved from GenBank, one for alphaCoVs host species and another for betaCoVs host species (Appendix; see also Supplementary Table 2 posted at https://doi.org/10.6084/m9.figshare.c.6029390.v1). The cophylogenies of alphaCoVs, betaCoVs, and their respective hosts were obtained using the function cophylo included in the R package phytools v0.7-70 (46).

The American bat betaCoV tree was statistically tested for cospeciation using the software eMPRess (47), which implements the duplication-transfer-loss (DTL) model for maximum parsimony reconciliation (MPR). This software is advantageous as it allows setting the event cost landscape as a two-dimensional plot, and different regions can be established for cost combinations that yield the same MPRs. For each

cost, a variable number of MPRs can be obtained, which in turn can be clustered to minimize the distance between different sets of solutions. The betaCoV data set yielded two cost regions. Reconciliations were computed for both cost regions (cost set *a*, 2, 6, and 1; cost set *b*, 4, 2, and 1 [for duplications, transfers, and losses, respectively]). The resulting MPRs were statistically tested through a randomization test, consisting of 100 replicates permuting tip associations between hosts and viral sequences.

Species richness maps. Bat species richness maps were built using the software QGIS 3.16.0 Hannover (48), based on the IUCN Red List of Threatened Species (49) distribution shapes of all Chiroptera and selecting host species for the alphaCoVs and betaCoVs included in this study.

ACKNOWLEDGMENTS

APPENDIX

Final CoV data set. The final data set included 128 alphaCoV and 17 betaCoV sequences with their associated host species. The sequence MT734810 (alpha coronavirus) was removed from the analysis because it had an unspecified host. Two sequences were annotated as isolated from Glossophaginae species. These were treated as *Glossophaga soricina* since this was the only representative of Glossophaginae found in other sequences. Two sequences were annotated as isolated from *Eptesicus* species. These were treated as *Eptesicus brasiliensis* for being the most widespread and abundant species in Brazil, the country where these sequences were obtained (see Supplementary Table 1 posted at https://doi.org/10.6084/m9.figshare.c.6029390.v1).

Molecular clock prior setting. An initial analysis was run for 1E7 MCMC to check parameter mixing. The default clock prior was quite diffuse, so the posterior of this first run was used to set a more informative prior in subsequent analysis. In the final analysis, a normal distribution was chosen, with a mean of 3E−3 substitutions/site/year (s/s/y) and a standard deviation (SD) of 2E−4 s/s/y, truncated between 2.5 and 3.5 E−3 s/s/y.

Cyt *b* data curation. The Cyt *b* sequence of *Perimyotis subflavus* is annotated as *Pipistrellus subflavu*s (a synonym of the former). *Phyllostomus discolor* had partial sequences with several stop codons, probably being a pseudogene. The sequence of *Phyllostomus elongatus* was used to represent this genus.

Host phylogenies. The alphaCoV host tree included 22 bat species corresponding to the families Vespertilionidae, Molossidae, and Phyllostomidae. A total of 1E7 MCMC generations were run in MrBayes, using HKY+I+G and GTR+I+G for the first+second and the third codon positions, respectively, setting the burn-in to 2.34% of the run, according to the convergence criterion mentioned above.

The betaCoV host tree included seven species belonging to the families Mormoopidae, Phyllostomidae, and Molossidae (see Supplementary Table 2 posted at https://doi.org/10.6084/m9.figshare.c.6029390.v1). A total of 1E7 MCMC generations were run in MrBayes, using HKY+G and HKY+I+G for the first+second and the third codon positions, respectively, setting the burn-in to 3.8% of the run.

REFERENCES

1. Wilson DE, Mittermeier RA. 2019. Handbook of the mammals of the world. Bats, vol 9. Lynx Edicions, Catalonia, Spain.
2. Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T. 2006. Bats: important reservoir hosts of emerging viruses. Clin Microbiol Rev 19:531–545. https://doi.org/10.1128/CMR.00017-06.
3. Messenger SL, Smith JS, Rupprecht CE. 2002. Emerging epidemiology of bat-associated cryptic cases of rabies in humans in the United States. Clin Infect Dis 35:738–747. https://doi.org/10.1086/342387.
4. Chua KB, Goh KJ, Wong KT, Kamarulzaman A, Tan PSK, Ksiazek TG, Zaki SR, Paul G, Lam SK, Tan CT. 1999. Fatal encephalitis due to Nipah virus among pig-farmers in Malaysia. Lancet 354:1257–1259. https://doi.org/10.1016/S0140-6736(99)04299-3.
5. Song H-D, Tu C-C, Zhang G-W, Wang S-Y, Zheng K, Lei L-C, Chen Q-X, Gao Y-W, Zhou H-Q, Xiang H, Zheng H-J, Chern S-WW, Cheng F, Pan C-M, Xuan H, Chen S-J, Luo H-M, Zhou D-H, Liu Y-F, He J-F, Qin P-Z, Li L-H, Ren Y-Q, Liang W-J, Yu Y-D, Anderson L, Wang M, Xu R-H, Wu X-W, Zheng H-Y, Chen J-D, Liang G, Gao Y, Liao M, Fang L, Jiang L-Y, Li H, Chen F, Di B, He L-J, Lin J-Y, Tong S, Kong X, Du L, Hao P, Tang H, Bernini A, Yu X-J, Spiga O, Guo Z-M, et al. 2005. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. Proc Natl Acad Sci U S A 102:2430–2435. https://doi.org/10.1073/pnas.0409608102.

6. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Wang W, Song H, Huang B, Zhu N, Bi Y, Ma X, Zhan F, Wang L, Hu T, Zhou H, Hu Z, Zhou W, Zhao L, Chen J, Meng Y, Wang J, Lin Y, Yuan J, Xie Z, Ma J, Liu WJ, Wang D, Xu W, Holmes EC, Gao GF, Wu G, Chen W, Shi W, Tan W. 2020. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 395:565–574. https://doi.org/10.1016/S0140-6736(20)30251-8.

7. Streicker DG, Turmelle AS, Vonhof MJ, Kuzmin IV, McCracken GF, Rupprecht CE. 2010. Host phylogeny constrains cross-species emergence and establishment of rabies virus in bats. Science 329:676–679. https://doi.org/10.1126/science.1188836.

8. Woo PCY, Lau SKP, Lam CSF, Lau CCY, Tsang AKL, Lau JHN, Bai R, Teng JLL, Tsang CCC, Wang M, Zheng B-J, Chan K-H, Yuen K-Y. 2012. Discovery of seven novel mammalian and avian coronaviruses in the genus *Deltacoronavirus* supports bat coronaviruses as the gene source of *Alphacoronavirus* and *Betacoronavirus* and avian coronaviruses as the gene source of *Gammacoronavirus* and *Deltacoronavirus*. J Virol 86:3995–4008. https://doi.org/10.1128/JVI.06540-11.

9. Graham RL, Baric RS. 2010. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. J Virol 84:3134–3146. https://doi.org/10.1128/JVI.01394-09.

10. Wong ACP, Li X, Lau SKP, Woo PCY. 2019. Global epidemiology of bat coronaviruses. Viruses 11:174. https://doi.org/10.3390/v11020174.

11. Anthony SJ, Johnson CK, Greig DJ, Kramer S, Che X, Wells H, Hicks AL, Joly DO, Wolfe ND, Daszak P, Karesh W, Lipkin WI, Morse SS, Mazet JAK, Goldstein T. 2017. Global patterns in coronavirus diversity. Virus Evol 3:vex012. https://doi.org/10.1093/ve/vex012.

12. Mollentze N, Streicker DG. 2020. Viral zoonotic risk is homogenous among taxonomic orders of mammalian and avian reservoir hosts. Proc Natl Acad Sci U S A 117:9423–9430. https://doi.org/10.1073/pnas.1919176117.

13. Pekar JE, Magee A, Parker E, Moshiri N, Izhikevich K, Havens JL, Gangavarapu K, Malpica S, Lorena M, Crits-Christoph A, Matteson NL, Zeller M, Levy JI, Wang JC, Hughes S, Lee J, Park H, Park M-S, Yan KCZ, Tzer Pin Lin R, Isa MNM, Noor YM, Vasylyeva TI, Garry RF, Holmes EC, Rambaut A, Suchard MA, Andersen KG, Worobey M, Wertheim JO. 26 February 2022. SARS-CoV-2 emergence very likely resulted from at least two zoonotic events. Zenodo. https://doi.org/10.5281/ZENODO.6291628.

14. Hernández-Aguilar I, Lorenzo C, Santos-Moreno A, Naranjo EJ, Navarrete-Gutiérrez D. 2021. Coronaviruses in bats: a review for the Americas. Viruses 13:1226. https://doi.org/10.3390/v13071226.

15. Bittar C, Machado RRG, Comelis MT, Bueno LM, Beguelini MR, Morielle-Versute E, Nogueira ML, Rahal P. 2020. Alphacoronavirus detection in lungs, liver, and intestines of bats from Brazil. Microb Ecol 79:203–212. https://doi.org/10.1007/s00248-019-01391-x.

16. Corman VM, Rasche A, Diallo TD, Cottontail VM, Stöcker A, Souza BFDCD, Corrêa JI, Carneiro AJB, Franke CR, Nagy M, Metz M, Knörnschild M, Kalko EKV, Ghanem SJ, Morales KDS, Salsamendi E, Spínola M, Herrler G, Voigt CC, Tschapka M, Drosten C, Drexler JF. 2013. Highly diversified coronaviruses in neotropical bats. J Gen Virol 94:1984–1994. https://doi.org/10.1099/vir.0.054841-0.

17. Góes LGB, Ruvalcaba SG, Campos AA, Queiroz LH, de Carvalho C, Jerez JA, Durigon EL, Dávalos LII, Dominguez SR. 2013. Novel bat coronaviruses, Brazil and Mexico. Emerg Infect Dis 19:1711–1713. https://doi.org/10.3201/eid1910.130525.

18. Góes LGB, de Campos AC, de Carvalho C, Ambar G, Queiroz LH, Cruz-Neto AP, Munir M, Durigon EL. 2016. Genetic diversity of bats coronaviruses in the Atlantic Forest hotspot biome, Brazil. Infect Genet Evol 44:510–513. https://doi.org/10.1016/j.meegid.2016.07.034.

19. Lima FEDS, Campos FS, Kunert Filho HC, Batista HBDCR, Carnielli P, Cibulski SP, Spilki FR, Roehe PM, Franco AC. 2013. Detection of Alphacoronavirus in velvety free-tailed bats (Molossus molossus) and Brazilian free-tailed bats (Tadarida brasiliensis) from urban area of Southern Brazil. Virus Genes 47:164–167. https://doi.org/10.1007/s11262-013-0899-x.

20. Moreira-Soto A, Taylor-Castillo L, Vargas-Vargas N, Rodríguez-Herrera B, Jiménez C, Corrales-Aguilar E. 2015. Neotropical bats from Costa Rica harbour diverse coronaviruses. Zoonoses Public Health 62:501–505. https://doi.org/10.1111/zph.12181.

21. Alves DMCC, Diniz-Filho JAF, da Silva e Souza K, Gouveia SF, Villalobos F. 2018. Geographic variation in the relationship between large-scale environmental determinants and bat species richness. Basic Appl Ecol 27:1–8. https://doi.org/10.1016/j.baae.2017.12.002.

22. Joffrin L, Goodman SM, Wilkinson DA, Ramasindrazana B, Lagadec E, Gomard Y, Le Minter G, Dos Santos A, Corrie Schoeman M, Sookhareea R, Tortosa P, Julienne S, Gudo ES, Mavingui P, Lebarbenchon C. 2020. Bat coronavirus phylogeography in the Western Indian Ocean. Sci Rep 10:6873. https://doi.org/10.1038/s41598-020-63799-7.

23. Latinne A, Hu B, Olival KJ, Zhu G, Zhang L, Li H, Chmura AA, Field HE, Zambrana-Torrelio C, Epstein JH, Li B, Zhang W, Wang LF, Shi ZL, Daszak P. 2020. Origin and cross-species transmission of bat coronaviruses in China. Nat Commun 11:4235. https://doi.org/10.1038/s41467-020-17687-3.

24. Leopardi S, Holmes EC, Gastaldelli M, Tassoni L, Priori P, Scaravelli D, Zamperin G, De Benedictis P. 2018. Interplay between co-divergence and cross-species transmission in the evolutionary history of bat coronaviruses. Infect Genet Evol 58:279–289. https://doi.org/10.1016/j.meegid.2018.01.012.

25. Cui J, Han N, Streicker D, Li G, Tang X, Shi Z, Hu Z, Zhao G, Fontanet A, Guan Y, Wang L, Jones G, Field HE, Daszak P, Zhang S. 2007. Evolutionary relationships between bat coronaviruses and their hosts. Emerg Infect Dis 13:1526–1532. https://doi.org/10.3201/eid1310.070448.

26. Graham GL. 1988. Interspecific associations among Peruvian bats at diurnal roosts and roost sites. J Mammal 69:711–720. https://doi.org/10.2307/1381626.

27. Kelm DH, Toelch U, Jones MM. 2021. Mixed-species groups in bats: non-random roost associations and roost selection in neotropical understory bats. Front Zool 18:53. https://doi.org/10.1186/s12983-021-00437-6.

28. Watanabe S, Masangkay JS, Nagata N, Morikawa S, Mizutani T, Fukushi S, Alviola P, Omatsu T, Ueda N, Iha K, Taniguchi S, Fujii H, Tsuda S, Endoh M, Kato K, Tohya Y, Kyuwa S, Yoshikawa Y, Akashi H. 2010. Bat coronaviruses and experimental infection of bats, the Philippines. Emerg Infect Dis 16:1217–1223. https://doi.org/10.3201/eid1608.100208.

29. Gouilh MA, Puechmaille SJ, Gonzalez J-P, Teeling E, Kittayapong P, Manuguerra J-C. 2011. SARS-Coronavirus ancestor's foot-prints in South-East Asian bat colonies and the refuge theory. Infect Genet Evol 11:1690–1702. https://doi.org/10.1016/j.meegid.2011.06.021.

30. Drexler JF, Gloza-Rausch F, Glende J, Corman VM, Muth D, Goettsche M, Seebens A, Niedrig M, Pfefferle S, Yordanov S, Zhelyazkov L, Hermanns U, Vallo P, Lukashev A, Müller MA, Deng H, Herrler G, Drosten C. 2010. Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. J Virol 84:11336–11349. https://doi.org/10.1128/JVI.00650-10.

31. Caraballo DA, Montani ME, Martínez LM, Antoniazzi LR, Sambrana TC, Fernández C, Cisterna DM, Beltrán FJ, Colombo VC. 2020. Heterogeneous taxonomic resolution of cytochrome b gene identification of bats from Argentina: implications for field studies. PLoS One 15:e0244750. https://doi.org/10.1371/journal.pone.0244750.

32. Streicker DG, Lemey P, Velasco-Villa A, Rupprecht CE. 2012. Rates of viral evolution are linked to host geography in bat rabies. PLoS Pathog 8:e1002720. https://doi.org/10.1371/journal.ppat.1002720.

33. Allen T, Murray KA, Zambrana-Torrelio C, Morse SS, Rondinini C, Di Marco M, Breit N, Olival KJ, Daszak P. 2017. Global hotspots and correlates of emerging zoonotic diseases. Nat Commun 8:1124. https://doi.org/10.1038/s41467-017-00923-8.

34. Pickett BE, Sadat EL, Zhang Y, Noronha JM, Squires RB, Hunt V, Liu M, Kumar S, Zaremba S, Gu Z, Zhou L, Larson CN, Dietrich J, Klem EB, Scheuermann RH. 2012. ViPR: an open bioinformatics database and analysis resource for virology research. Nucleic Acids Res 40:D593–D598. https://doi.org/10.1093/nar/gkr859.

35. Sievers F, Higgins DG. 2018. Clustal Omega for making accurate alignments of many protein sequences: Clustal Omega for many protein sequences. Protein Sci 27:135–145. https://doi.org/10.1002/pro.3290.

36. Kirichenko AD, Poroshina AA, Sherbakov DY, Sadovsky MG, Krutovsky KV. 2022. Comparative analysis of alignment-free genome clustering and whole genome alignment-based phylogenomic relationship of coronaviruses. PLoS One 17:e0264640. https://doi.org/10.1371/journal.pone.0264640.

37. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542. https://doi.org/10.1093/sysbio/sys029.

38. Nylander J. 2004. Mr Modeltest 2.2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.

39. Rambaut A. 2014. FigTree v1.4.2. http://tree.bio.ed.ac.uk/software/figtree/.

40. Winter DJ. 2017. rentrez: an R package for the NCBI eUtils API. R J 9:520–526. https://doi.org/10.32614/RJ-2017-058.

41. R Development Core Team. 2021. R: a language and environment for statistical computing (R version 4.0.3). R Foundation for Statistical Computing, Vienna, Austria.

42. RStudio Team. 2020. RStudio: integrated development environment for R. RStudio, PBC, Boston, MA. http://www.rstudio.com/.

43. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol 4:vey016. https://doi.org/10.1093/ve/vey016.

44. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol 67:901–904. https://doi.org/10.1093/sysbio/syy032.

45. Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A, Lemey P. 2016. SpreaD3: interactive visualization of spatiotemporal history and trait evolutionary processes. Mol Biol Evol 33:2167–2169. https://doi.org/10.1093/molbev/msw082.

46. Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 3:217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x.

47. Santichaivekin S, Yang Q, Liu J, Mawhorter R, Jiang J, Wesley T, Wu Y-C, Libeskind-Hadas R. 2021. eMPRess: a systematic cophylogeny reconciliation tool. Bioinformatics 37:2481–2482. https://doi.org/10.1093/bioinformatics/btaa978.

48. QGIS Development Team. 2021. QGIS geographic information system. QGIS Association. https://www.qgis.org.

49. IUCN. 2021. The IUCN Red List of Threatened Species. Version 2021-3. https://www.iucnredlist.org. Accessed on 21 December 2021.