scientific reports

Check for updates

OPEN

Genomics of novel influenza A virus (H18N12) in bats, Caribe Colombia

Daniel Echeverri-De la Hoz[1], Caty Martínez-Bravo[1], Bertha Gastelbondo-Pastrana[1,2,4], Ricardo Rivero[3], Yesica López[1], Valeria Bertel[1], Maira Alemán-Santos[1], Evelin Garay[1], Richard Hoyos[1], Germán Arrieta[1,4], Juan David Ramírez[5,6] & Salim Mattar[1]✉

Influenza viruses are highly capable of mutating and adapting in mammalian hosts. While these viruses have been extensively studied in birds, research on their presence in bats has been limited. However, influenza viruses circulating in bats have shown notable molecular divergence. The present study aimed to characterize the phylogenetic, evolutionary, and antigenic relationships of an influenza A virus detected in the fishing bat *Noctilio albiventris*. As part of a pathogen surveillance study of public health interest, 159 rectal samples were collected from bats in the Colombian Caribbean. The samples were sequenced using RNA-Seq. A genome (eight viral contigs) associated with the *Orthomyxoviridae* family was identified in a pool. Most segments showed approximately 90% similarity with H18N11, except for the neuraminidase. Analysis of the N protein shows that occupies a basal position relative to the N11 subtype, with its divergence date estimated to be approximately 50 years earlier than the earliest reported N11 sequence. 3D modeling identified three mutations (K363R, T242K, and I139V), which may enhance interaction with the HLA-DR of bats. The analyses and antigenic divergence observed in the N protein of *N. albiventris* suggests the existence of a new subtype (H18N12) with unknown pathogenicity, which requires further investigation.

Keywords Neuraminidase, *Noctilio albiventris*, Neotropical bats, Adaptability, Ortomyxoviridae

Bats are subject to epidemiological surveillance because they are hosts of viruses such as coronaviruses, paramyxoviruses, filoviruses[1], Venezuelan Equine Encephalitis[2]and dengue virus[3,4]. Although bats have not yet been established as reservoirs of influenza viruses, a new genomic sequence of influenza A virus (IAV) designated as H17N10 was detected in fruit bats in Guatemala[5]. Later, another genome classified as H18N11 was characterized in bats in Peru[6], Bolivia, and Brazil[7]. In addition, a virus similar to avian strains was detected in Egyptian bats[8]. This evidence suggests new hypotheses about the origins of Influenza A viruses and their potential public health impact. They highlight the need for further research on bats as reservoirs and the implications for influenza control strategies[5,6].

The origin of bat IAVs needs to be better understood, and only a few studies have clarified their origin. Phylogenetic and evolutionary analyses have established that they originated from a common ancestor with avian IAV subtypes[9]. However, divergence is a feature that has drawn the attention of researchers. IAVs probably split into two branches due to geographic separation and multiple early spread events, or IAVs may have undergone drastic changes to adapt to bats[10].

Crystallographic structures of the hemagglutinin (H) and neuraminidase (N) proteins have been generated to characterize bat IAVs[6,11,12]. The structures were similar to avian IAV strains but with specific molecular modifications as an adaptability mechanism to bats. IAVs do not use sialic acid, but molecules of the major histocompatibility complex class II (MHC-II) as cell entry receptors[13,14]. This mechanism suggests adaptability and possible inter-species jumping, since MHC-II is expressed in immune system cells and epithelial tissue in many animals, including pigs, mice, and chickens[10].

In contrast, the N protein of bat IAV appears to have no catalytic activity, and its function is an enigma for researchers. These results indicate that the N protein could induce low expression of MHC-II molecules by an

[1]Instituto de Investigaciones Biológicas del Trópico, Universidad de Córdoba, Street, Montería 230002, Córdoba, Colombia. [2]Grupo de Investigaciones Microbiológicas y Biomédicas de Córdoba-GIMBIC, Universidad de Córdoba, Montería, Colombia. [3]Paul G. Allen School for Global Health, Washington State University, Pullman, Washington 99164, United States of America. [4]Grupo de Salud Pública y Auditoría en Salud, Corporación Universitaria del Caribe-CECAR, Sincelejo, Colombia. [5]Centro de Investigaciones en Microbiología y Biotecnología-UR (CIMBIUR), Universidad del Rosario, Bogotá, Colombia. [6]Molecular Microbiology Laboratory, Department of Pathology, Molecular and Cell-Based Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ✉email: smattar@correo.unicordoba.edu.co

unknown mechanism[10]. Confirmation of these data would demonstrate that the surface glycoproteins of bat IAVs have receptor binding and destruction activities. Therefore, bat IAVs would carry out the same infection and release processes of viral particles as avian IAVs[9]. The search for IAVs in bats has motivated researchers to investigate the possible role of bats as hosts and to delve deeper into the evolution of these viruses. Current knowledge does not precisely establish the adaptability of IAVs.

The present study aimed to characterize the phylogenetic, evolutionary, and antigenic relationships of an influenza A virus detected in the fishing bat *Noctilio albiventris*.

Results

Bat capture, sample, and sequencing analysis

A total of 159 bats were randomly captured in four municipalities in the Colombian Caribbean: Talaigua Nuevo (Bolívar), Santa Ana (Magdalena), Moñitos (Córdoba), and Colosó (Sucre). All samples (26 pools) processed and sequenced by RNA-Seq belonged to individuals of the *Phyllostomidae*, *Molossidae*, *Noctilionidae*, and *Emballonuridae* families. Only a pool of four individuals of the fishing bat *Noctilio albiventris* captured in Talaigua Nuevo, Bolívar (Fig. 1) yielded contigs associated with *Orthomyxoviridae*.

Genomic, phylogenetic and evolutionary analysis of influenza A virus detected in *N. albiventris*

Initially, three contigs with a similarity of 90% were assigned to the H18N11 subtype. The reference genome (H18N11 Peru CY125942-CY125949) map yielded seven segments corresponding to the PB1, PB2, PA, HA, NS, NP, and M genes. The search effort for the NA gene was increased by decreasing the alignment and similarity criteria at the seed level. Eight segments were obtained (Supplementary Fig. S1) that corresponded to 0.12% (22,256/18'129,755) of the total reads with a depth of 100X, and this virus was designated A/ fishing bat/ Colombia/2023 (A/bat/Colombia/23) (SRA: PRJNA1162262). The genome comparison analysis shows greater similarity with the H18N11 subtype sequence recorded in Peru (93%). The percentage of similarity of each segment varied between 92 and 98%, and the N protein gene of A/bat/Colombia/23 showed greater divergence (Table 1). Figure 2 shows the similarity patterns between A/bat/Colombia/23 concerning the H1N1, H2N2, H3N3, H17N10, and H18N11 subtypes.

Phylogenetic analyses indicate that the A/bat/Colombia/23 segments are related to IAVs detected in bats and form a divergent clade to avian IAVs (Supplementary Fig. S2). The HA segment is closely related to node H18 from Peru, Brazil, and Bolivia (Fig. 3A). However, the PB2, PA, and M segments form a basal branch

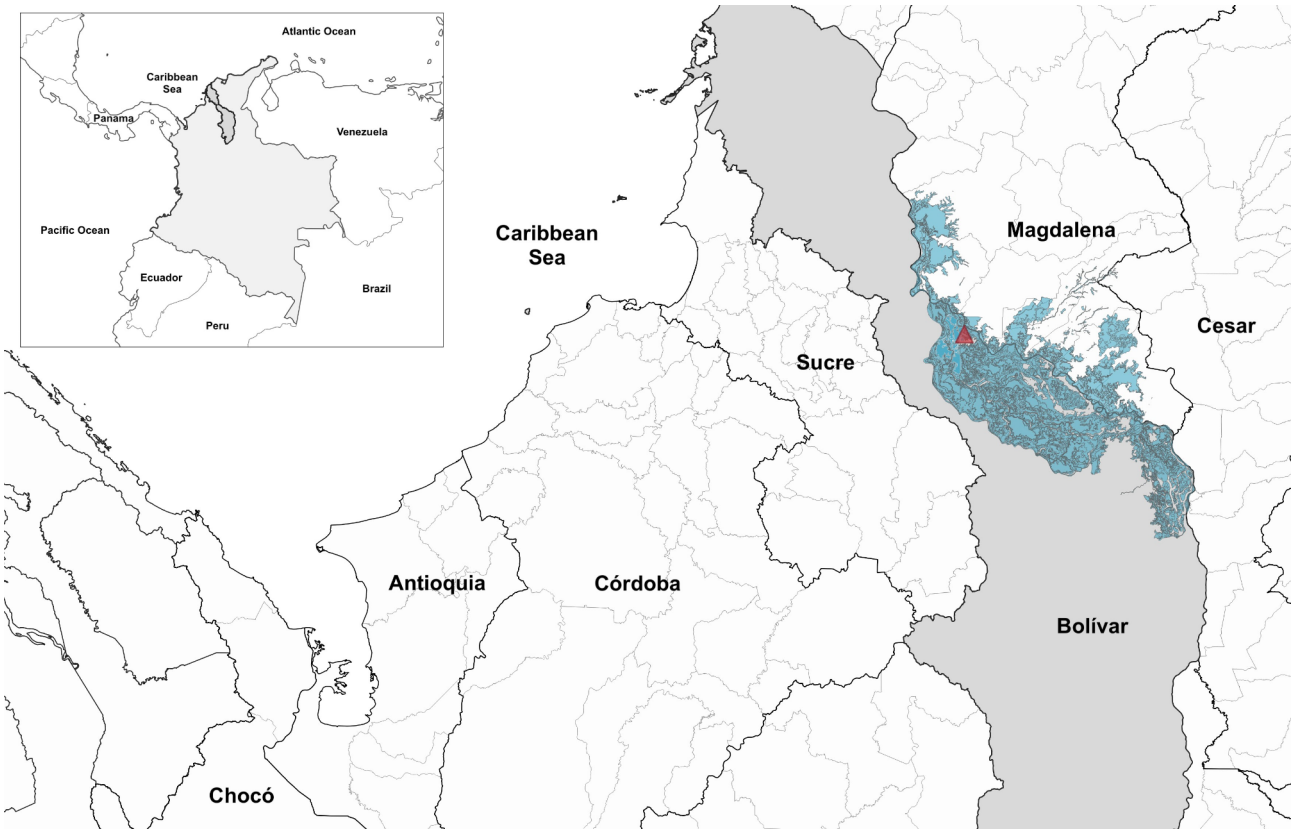

Fig. 1. Geographic location of the department of Bolívar, Colombia. The red triangle shows the sampling point (9°18'28"N- 74°36'56"O). The map was generated with QGIS 3.32.2 (QGIS.org, 2021. QGIS Geographic Information System. QGIS Association. http://www.qgis.org).

Gene	H18N11-Peru		H18N11-Bolivia	
	Nucleotides	Amino acid exchange site	Nucleotides	Amino acid exchange site
PB2	92.67%	V76I, I105V, K156R, R471K, N472S, T473N, V476I, V478I, I570V, S702N	92.32%	V76I, K156R, R471K, N472S, T473N, V476I, V478I, I570V, S702N
PB1	95.28%	I172T, A401V, F551L	94.97%	I54V, V64I, S104N, I172T, I243T, A401V, R429K, F551L, V645I
PA	97.31%	N227H	97.27%	N227H, I311M, G683E
HA	98.74%	None	96.21%	None
NP	94.73%	I436V	98.49%	None
NA	81%	I15L, N39E, V82L, I83L, L89F, N92T, S95N, I96V, D128N, M138F, K161S, I171L, Q188H, S201T, I202V, I207V, I208V, T215S, D231E, T242K, T245K, A246T, L254I, G282R, V284I, K333R, G349A, N351S, I358T, G361S, K363R, L372M, V381A, I394V	80%	N39E, V82L, I83L, L89F, N92T, S95N, I96V, D128N, M138F, K161S, I171L, Q188H, S201T, I202V, I207V, I208V, T215S, D231E, T242K, T245K, A246T, I255V, G282R, V284I, K333R, G349A, N351S, I358T, G361S, K363R, L372M, V381A, I394V, T419A
M	98.81%	None	91.03%	None
NS	96.35%	A193T	98.45%	None

Table 1. Comparison and percentages of similarity of A/bat/Colombia/23 segments concerning the H18N11 subtypes obtained in Peru and Bolivia.

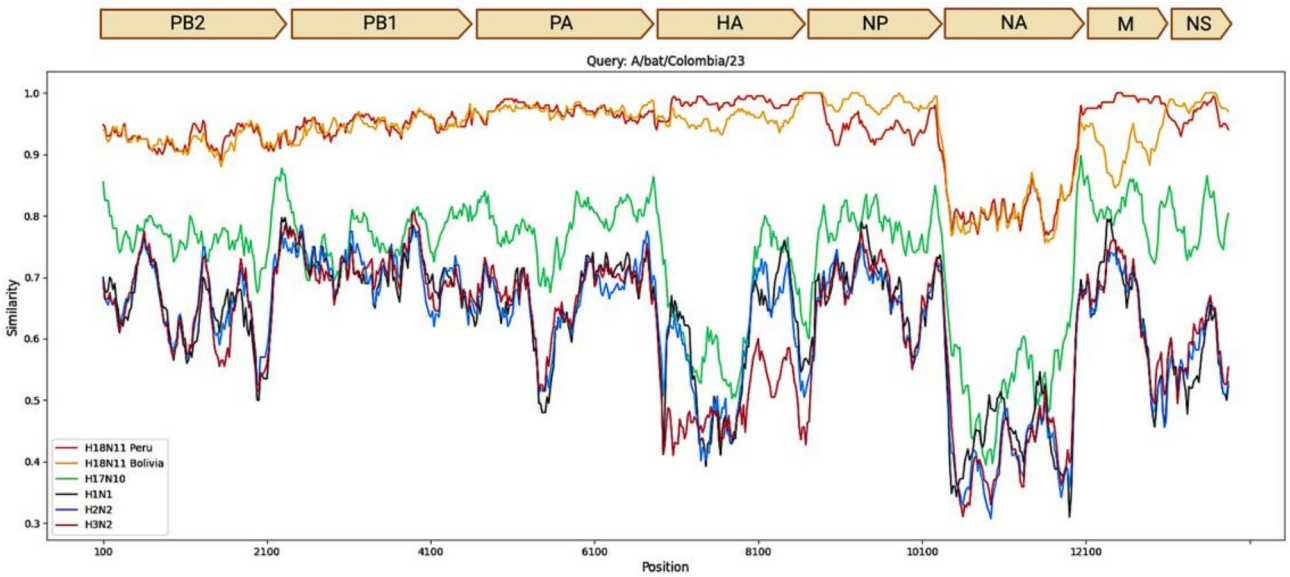

Fig. 2. Similarity patterns between A/bat/Colombia/23 concerning the H1N1, H2N2, H3N3, H17N10, and H18N11 subtypes. The comparison was made with the amino acid sequences for each segment. A 500 bp visualization window with 50 bp steps and Kimura as a distance model was established.

(Supplementary Fig. S2). In addition, the NA gene is the most phylogenetically distant (Fig. 3B). Based on phylogenetic analysis and NA gene divergence, a time to the most recent common ancestor (TMCRA) analysis was conducted. This analysis, with a chain length of 100,000,000 interactions indicated that TMCRA gave rise to the bat NA segment in 1420. The A/bat/Colombia/23 NA segment came from a node in 1940, and the node that gave rise to the N11 subtype dates to around 1990 (Fig. 3C). This demonstrates that the previously described N11 segment is phylogenetically related to A/bat/Colombia/23NA but is estimated to be approximately 50 years earlier than the earliest reported N11 sequence.

Because of the evolutionary implications associated with the phylogenetic and evolutionary divergence of some segments of A/bat/Colombia/23, possible genetic reassortment events with HA and NA segments were inferred. The antigenic variability and selective pressure to which these genes are subjected allow us to demonstrate the processes involved in adaptability to new hosts and genetic diversity[15]. Interestingly, the analysis suggested that the HA segment of bat IAVs was acquired from canonical IAV strains, indicating a possible evolutionary interaction between the two groups (Fig. 4). The NA segment originates from the interaction between strains circulating in neotropical bats. On average, there was one reassortment event every 31 years in the HA and NA segments. The low reassortment rate may reflect differences in genetic compatibility between segments or ecological restrictions that limit co-infections. However, bats also influence the evolution of IAVs.

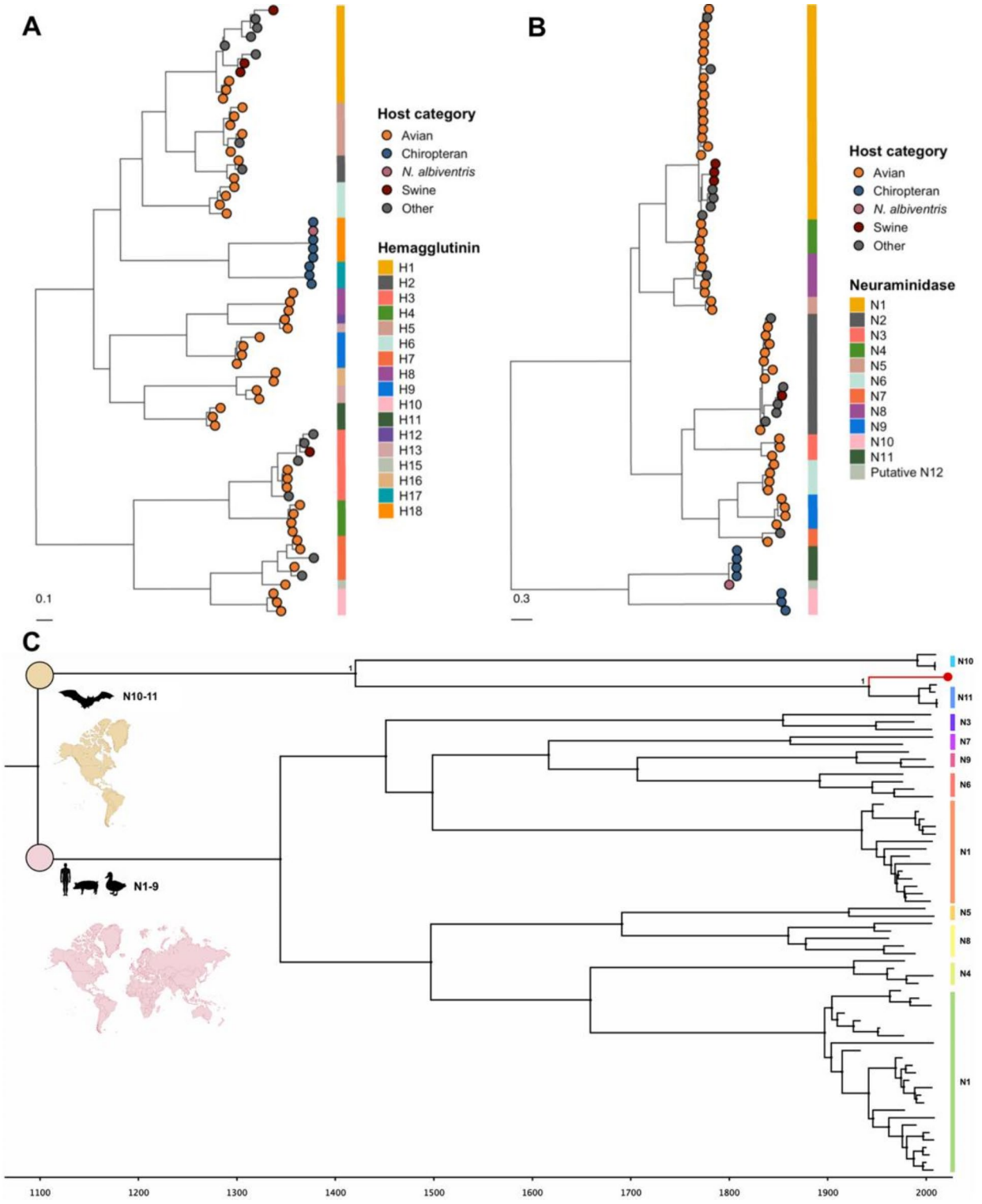

Fig. 3. Phylogenetic tree of A/bat/Colombia/23 HA (**A**) and NA (**B**) genes. Bat influenza A branches and the new sequence are presented in red, which signifies their unique evolutionary path. The phylogenetic tree was constructed using the maximum likelihood method. Bootstrap values (1000) were given at the relevant nodes. The branch scale represents the number of amino acid substitutions along the tree branches. Reference sequences from all reported IAV subtypes (HA1-16 and NA1-9) were included in all trees. (**C**) Molecular clock to determine the TMCRA of the NA segment. The branch of the NA segment of A/bat/Colombia/23 is represented in red and arises from a node dating back to 1940. The scale represents the estimated time from the common ancestor to the current sequences. The posterior probability values were given to the nodes of interest. The phylogenetic trees were constructed using complete gene sequences.

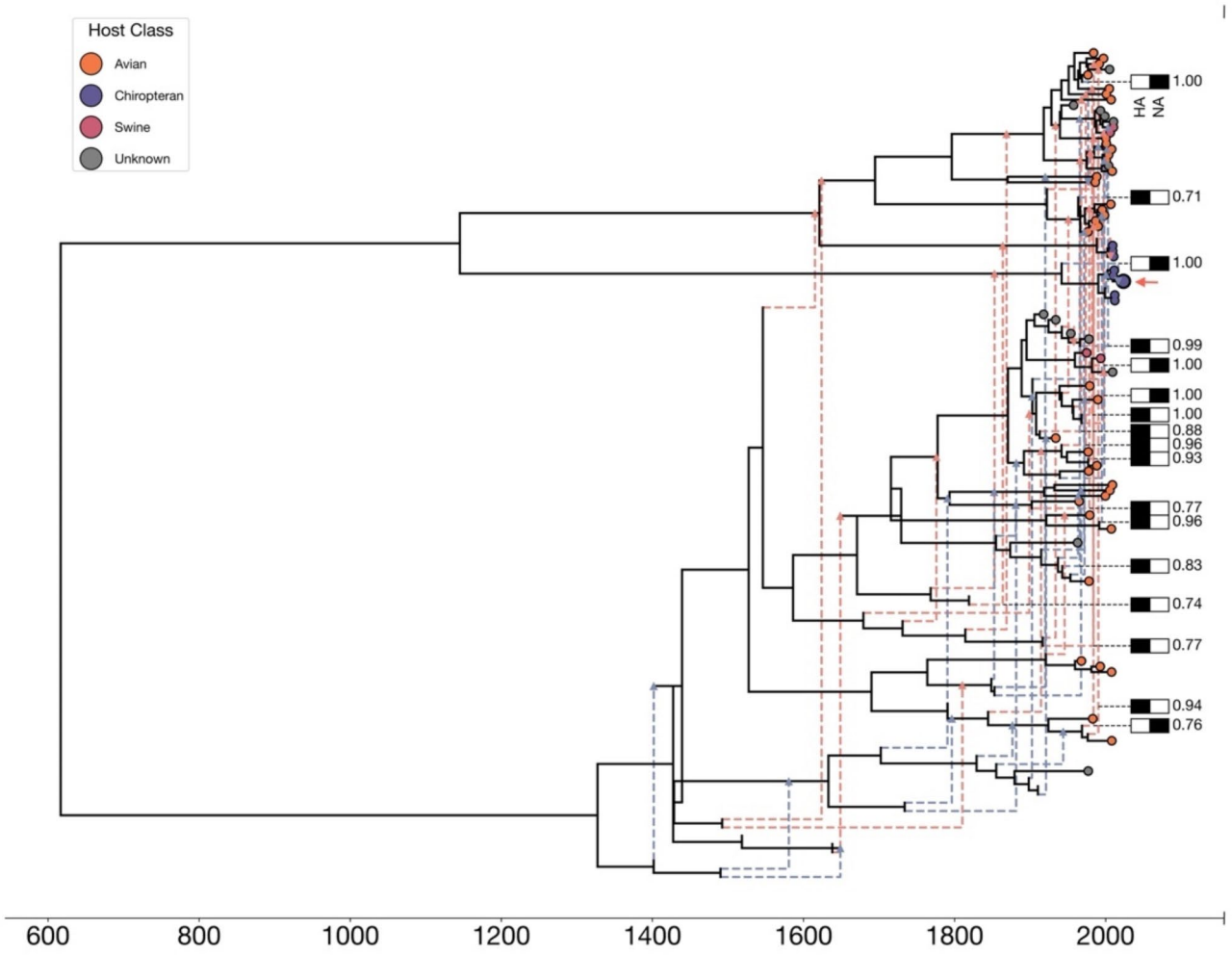

Fig. 4. Estimates of the genetic recombination networks of the HA and NA segments of canonical IAV and bat IAV. The analysis was carried out with 300,000,000 interactions. The recombination rate was 0.0321 (95% HPD: 0.0206–0.0445). The purple and red dashed lines indicate rearrangement events for the NA and HA segment, respectively. The red flake indicates the HA and NA segment of A/bat/Colombia/23. The posterior probability values were given to the nodes of interest. The phylogenetic tree was constructed using complete gene sequences.

Antigenic structure and molecular docking of the N protein

Amino acid sequence analysis shows a domain conserved at the level of the *Alphainfluenzavirus* genus and another of sialidase enzymes (Supplementary Fig. S3). The 3D structure for the N protein was generated from the crystallographic structure of the N11 protein of the H18N11 subtype from Peru (4K3Y) (Supplementary Table S1 and Fig. S4). Computational modeling generated a tetramer structurally similar to the N protein of avian strains that contains six antiparallel β-sheets in a helix-like arrangement. It also maintains six residues (R118, W178, S179, R224, E276 and E425) conserved in the active site. The alignment with the H18N11 subtype from Peru and Bolivia has yielded points of divergence at the level of the transmembrane and distal domains of the protein (Fig. 5A). The most significant mutations are highlighted in the hypothetical active site region (Fig. 5B). At the conformational level, our analysis of the crystal structure of the N11 protein from Peru (Fig. 6A) and the A/bat/Colombia/23 N protein (Fig. 6B) has revealed critical structural differences. The hypothetical active site pocket of the N11 protein is broader than that of the IAV and influenza B virus N proteins due to the movements of loops 150 and 430[6] (Fig. 5A). Conversely, the hypothetical active site pocket of the A/bat/Colombia/23 N protein is narrow due to the plasticity of loop 150 and the K363R mutation (Fig. 6B)[6]. This unique structural feature has led to a statistically significant interaction between the A/bat/Colombia/23 N protein and HLA-DR (MHC-II) of bats (8JRJ) (Fig. 6C), with nine residues involved in hydrogen bond formation being identified (Fig. 6D). Notably, three of the five mutations (Ser361, Ar363, and Lys242) have been found to increase the binding with the S1 subunit (110–165) of the α2 chain of bat HLA-DR, forming bonds with high specificity and strength ($^{\circ}A2 > 1500$). The movement of loop 150, which reduces the space of the active site, allows for a more significant contact and interaction surface. These findings significantly impact our understanding of viral

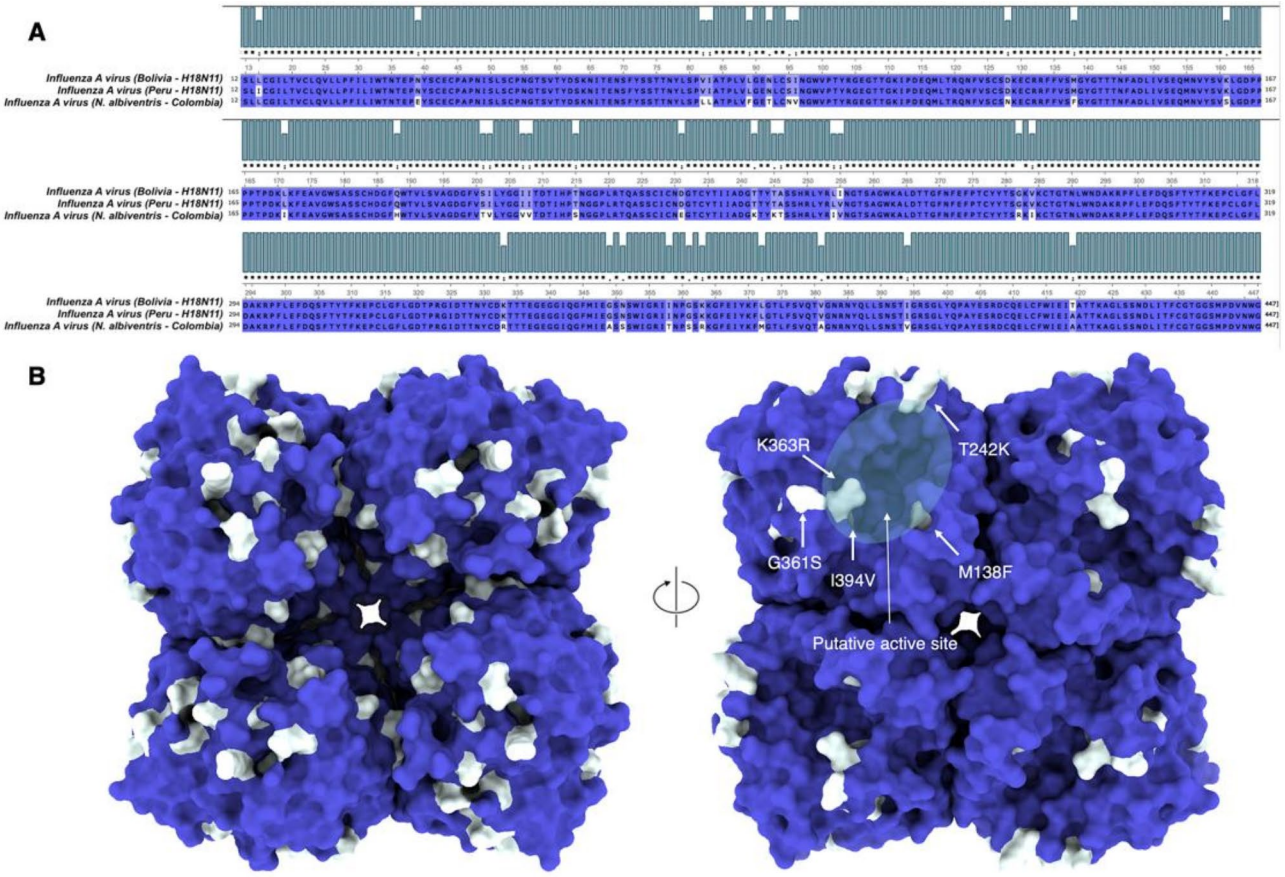

Fig. 5. Homology modeling of the A/bat/Colombia/23 N protein. (**A)** Alignment of the amino acid sequence of the A/bat/Colombia/23 NA protein with the N11 subtypes from Peru and Bolivia. (**B)**3D structure of the tetramer from a bottom transmembrane view and a top view showing the putative active site of the protein with five mutations (K363R, T242K, M138F, G361S, and I139V) relative to the N11 sequence from Peru and Bolivia. Conserved regions are shown in purple, and divergent amino acids in white. The image of the 3D model was generated with ChimeraX 1.7.1[42] (https://www.cgl.ucsf.edu/chimerax/).

protein interactions and could potentially inform future research and drug development, particularly in the design of antiviral drugs targeting the N protein.

Discussion

This study represents a significant contribution to the current pool of known IAV genomes, with a sequence detected for the first time in *Noctilio albiventris*from Colombia. Phylogenetic analysis unveils a divergent clade of IAVs from bats in seven of the eight segments, a discovery that aligns with previous studies[5,6]. Notably, the HA gene is the only segment related to all 16 known avian subtypes and is found within group one. Likewise, the results indicated that four segments (PB1, PA, M, and NA) of A/bat/Colombia/23 originated from a rearrangement event between viral strains circulating in bat populations. This provides a comprehensive understanding of viral adaptation to different hosts.

The clade of bat IAVs diverged about 300–500 years ago[5]. Evolutionary analyses indicate that HA was acquired from canonical IAV strains and then made changes to bind to bat MHC-II[13,14], but retained characteristics of primitive nodes. This finding has profound implications for the evolution and transmission of IAV. Undoubtedly, the hemagglutinin (H) of avian strains have tropism for the enteric tract due to the abundance of α−2,3 sialic acid that presents a "linear" structure that improves fusion with the cell, while the strains that infect humans bind to α−2,6 that has a "bent" structure and is mainly found in the respiratory tract[16]. However, the hemagglutinin of bat IAVs is highly divergent and binds to MHC-II[5,6]. Using this approach, the binding capacity of bat IAVs should not be analyzed based on the knowledge generated by avian IAVs[12]; instead, it should be analyzed based on the phenomenon of adaptability and the ability to infect new hosts[15].

The N protein of A/bat/Colombia/23 is the most divergent among the eight segments compared to the bat and bird IAV sequences. In addition, the new sequence does not share TMCRA with the N11 subtype. The molecular divergence of the NA segment raises uncertainty and raises hypotheses about the origin of this sequence. Discrepancy analysis with the NA and HA segments suggests that there were reassortment events

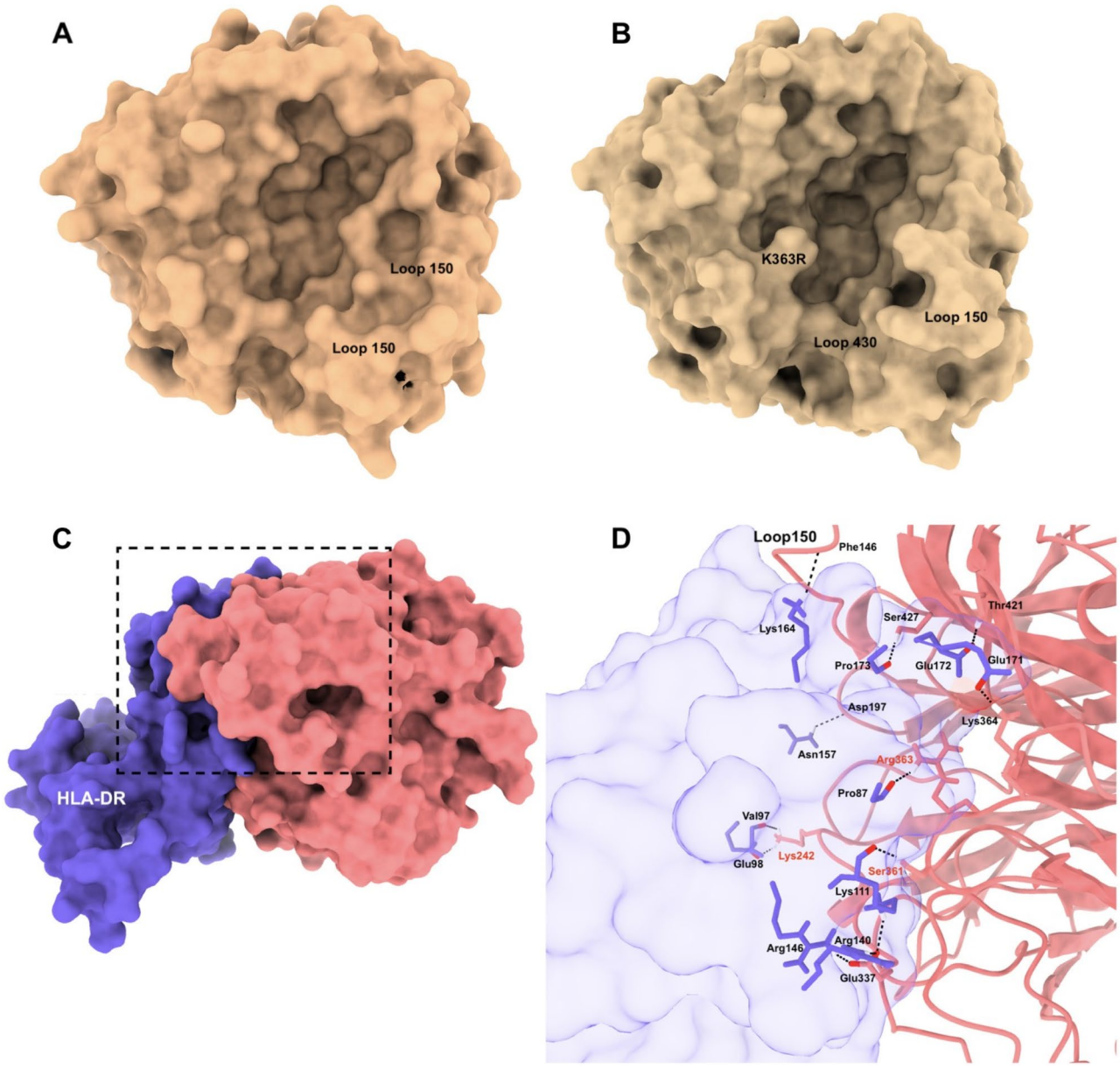

Fig. 6. Molecular analysis and docking of the N protein. (**A)** A/bat/Peru N11 protein (4K3Y). (**B)** N protein of A/bat/Colombia/23. (**C)** Monomer of the N protein of A/bat/Colombia/23 bound to the bat HLA-DR. The contact region (2935.1 ˚A2) between the two proteins is indicated. The Van der Waal energy was 67.8. D. Binding details between residues of the N protein of A/bat/Colombia/23 and bat HLA-DR. (**D**)The binding of the three mutations Ser361, Ar363, and Lys242 of the hypothetical active site of the N protein with Lys111, pro87, and Val97/Glu98, respectively, with the bat HLA-DR. The image of the 3D model was generated with ChimeraX 1.7.1[42] (https://www.cgl.ucsf.edu/chimerax/).

between circulating strains in neotropical bats. However, the NA segment was acquired from a bat IAV ancestor that has not yet been described. This divergence suggests a new subtype called H18N12[17].

A comparison of the antigenic architecture of the N11 protein in Peru showed critical structural differences. The movement of loop 150 affected the general morphology of the protein, suggesting a structural plasticity that could influence the efficiency and specificity of neuraminidase to perform its function during viral release[18]. In addition, a reduction in the size of the hypothetical active site implies changes in the stability of the structure[5,6].

Three of the five mutations (K363R, T242K, and G361S) near the hypothetical active site increased the possibility of binding to HLA-DR in bats. This finding generates uncertainty regarding the functions of this protein and the implications that such changes could have on viral replication or transmission. The changes generated in the protein strengthen the interactions with cellular receptors and support the theory of downregulation of MHC-II by bat IAV N protein[10]. The change in *Lys* to *Arg* favors new interactions by increasing the number of hydrogen bonds due to the presence of a guanidino group in its side chain. Meanwhile, *Met* for *Phe* allows

binding with the hydrophobic regions of the HLA-DR of bats[19,20]. It was shown that there is an interaction of loop 150 (Phe146), a crucial area for the enzymatic activity of this viral protein since it participates in the recognition and cleavage of receptors on cell surfaces. However, these data must be corroborated by in vitro assays.

Another aspect is the biological implications and ability of IAV to bind to MHC-II in other mammals. A recent study showed that catalytic activity can be acquired again with significant mutations in the amino acid sequence[21]. F144C and T342A changes in the N11 protein increased viral particles in MDCK II cells compared to cells infected with the wild-type virus. Furthermore, IAV has broader tissue tropism in the airways of mice (BALB/c) and ferrets (*Mustela putorius furo*)[19]. This is because, in the three-dimensional structure of N11, residue 144 is located at the outer edge of the hypothetical active site, whereas residue 342 is close to the binding site. Both sites are critical for sialidase activity of the N protein and are present in avian IAVs[20,22,23]. Therefore, the zoonotic potential and interspecies jump of bat IAVs cannot be ruled out.

Studies indicate that H18N11 subtypes have inefficient transmission and infection in non-bat hosts without generating critical structural changes[10]. For A/bat/Colombia/23, hemagglutinin was 100% identical to the H18 subtype. However, mutations in neuraminidase could improve the infection process and transmission to new species. The genetic plasticity of IAVs allows them to adapt to different species because, in the case of the bat H9N2 subtype, it has been shown that it can replicate and be transmitted between ferrets. In addition, they can efficiently infect human lung cell cultures. It can evade antiviral inhibition by MxA in B6 transgenic mice and generate cross-reaction with N2 specific antibodies in human sera[24]. Bat IAVs demonstrate the ability of influenza viruses to transmit and adapt to new hosts[24]. Therefore, new sequences must be studied to understand whether they can generate significant epidemiological outbreaks.

In conclusion, the characterization of viruses with zoonotic potential, such as IAV in bats, is crucial for understanding the role of these wild animals in the evolution of the IAV virus. The virus identified in this study has sequences that are new to science, with genetic and structural changes in the N protein that possibly propose a new subtype (H18N12). Experimental trials are needed to provide more information on influenza virus adaptation in bat hosts. The detection of IAVs in members of *Noctilionidae* expands the range of hosts in which these viruses can be detected. The fishing bat *Noctilio albiventris* is associated with aquatic ecosystems, and bodies of water can be a means of interaction between bats and birds. These findings underscore the urgent need for further research to understand the potential epidemiological implications of these new sequences, and the importance of ongoing research in this field.

Methods

Type of study, location, and ethical aspects

A prospective, cross-sectional descriptive study was conducted with rectal swab samples collected from bats captured between January and December 2023 in the Colombian Caribbean. The collected samples are part of an epidemiological surveillance study of emerging viruses in bats and mosquitoes developed by the Universidad de Córdoba, Colombia.

Ethical statement

The Instituto de Investigaciones Biológicas del Trópico and the Ethics Committee of the Universidad de Córdoba approved the project with permission from the National Environmental Licensing Authority (ANLA) of Colombia (Resolution 00914 of August 4, 2017). The animals captured in the study were released. Samples were taken in accordance with CDC Guidelines for Safe Work Practices in Human and Animal Medical Diagnostic Laboratories[25]. All methods were performed in accordance with the relevant guidelines and regulations.

Capture of bats and sample collection

The capture was done through mist nets (6 × 2 m). Animals were taxonomically identified using dichotomous keys based on morphometry[26]. Then, two to three rectal swab samples were taken, which were placed in a viral transport medium stored in N_2 and transported to the Instituto de Investigaciones Biológicas del Trópico of the Universidad de Córdoba, where they were kept at −80 °C until processing.

RNA extraction, purification, and sequencing

Rectal samples were vortexed for 30s, and then a pool was made by locality and species with the supernatant of each sample. Pool suspension was filtered through a 0.45 µm filter by centrifugation at 2000 x g in a microcentrifuge. RNA extraction was done from 200 μl of the supernatant using the GeneJET Viral DNA/RNA Purification Kit (Thermo Fisher Scientific™). The RNA was subjected to degradation of contaminating DNA with DNAse I (Promega™). RNA from the host was not removed. Then, it was purified and concentrated with the GeneJET RNA Cleanup and Concentration Kit (Thermo Fisher Scientific™). The concentration and integrity of the RNA were determined by fluorometry with Qubit® (Thermo Fisher Scientific™). Finally, they were processed with the Paired-End FCL 150 MGIEasy Fast RNA Library Prep Set™ under the high-throughput sequencing methodology based on DNA nanobeads (DNB) from MGI Tech™. Metatranscriptomic sequencing was performed on the MGI-G50® (Shenzhen, China) (Supplementary material).

Bioinformatics analysis

The sequences were subjected to quality assessment and reads elimination (< Q20) using Fastp[27]. A *de novo* assembly with a minimum length of 300 nucleotides was performed using MEGAHIT[28]. The contigs were compared with the non-redundant (nr) protein database of the National Center for Biotechnology Information (NCBI) with DIAMOND[29]to optimize the search for CDS regions encoding viral proteins. The maximum expected value is used to obtain an alignment of 0.001. The files obtained were processed and analyzed using MEGAN6[29]. BLASTn and BLASTx compared sequences of interest with more than 80% similarity. To confirm

the viral contigs, reads were mapped to the reference sequence of the H18N11 subtype from Peru (CY125942-CY125949) with Bowtie2[30]. The confirmed segments were used as reference sequences to obtain the segment coverage and depth using Bowtie2 and SAMTools[31]. The aligned reads were visualized and analyzed using UGENE[32]. The genome was annotated using Prokka[33].

Phylogenetic and evolutionary analysis

Reference genomes were obtained from GenBank (Supplementary Table S2). Aligning each segment's amino acid sequences was done with MAFFT[34], and manual editing was done in UGENE. A maximum likelihood tree was then created with a bootstrap of 1000 using IQ-TREE[35]and rooted, taking into account the substitution rate measured in time with Treetime[36]. Identity patterns with reference genomes were performed with SimPlot++[37]. Evolutionary tracking in time was executed in Bayesian Evolutionary Analysis Sampling Trees (BEAST)[38]. The substitution rate was estimated using TreeTime from nucleotide sequences aligned in MAFFT. The evolutionary model and options for the MCMC analysis were built using the BEAUti tool, where tests were carried out with different priors.

A strict molecular clock with uniform distribution was established. The analysis was executed with a GTR substitution model with Γ_4 distribution, partitioning at the codon positions (3 partitions: positions 1, 2, 3), and tree construction under an analysis coalescent with constant population size prior. The results were visualized and analyzed in Tracer[39], where an Effective Sample Size (ESS) > 200 was taken into account in all statistical analyses and the convergence of the Markov chains. Then, the tree information was summarized with TreeAnnotator[38] and visualized in FigTree (http://tree.bio.ed.ac.uk/software/figtree/). Finally, the rearrangement networks were inferred using a coalescing model with CoalRe in BEAST2[40]. The evolutionary model was built using the BEAUti tool where the priors were established as described above. The tree was constructed under a Coalescent Analysis with Reassortment Constant Population. The summary of the distribution of networks maximizing the credibility of the clade was made with networktree[40]. A burn-in of 10% was applied in all cases.

Analysis of the antigenic structure and molecular docking of the N protein

The 3D model of the N protein of A/bat/Colombia/23 was generated by homology with SWISS-MODEL server[41]. The template was selected according to its identity, coverage, Global Model Quality Estimation (GMQE), and Quaternary Structure Quality Estimation (QSQE). The model was visualized using ChimeraX[42]. Subsequently, the mutations concerning the N11 subtype of Peru and Bolivia were identified and mapped onto the 3D structure of the N protein of *N. albiventris*. The antigenic architecture of the N protein was then compared to the N11 crystallographic structure from Peru (4K3Y). For docking analysis, charge addition of the N protein of A/bat/Colombia/23 and bat HLA-DR (8JRJ) was performed in ChimeraX. Then, an information-driven flexible docking approach was used for modeling biomolecular complexes using HADDOCK[43]. Van der Waals energy values, Buried Surface Area, and Z score mainly evaluated the models. Finally, the contact points were visualized and predicted using ChimeraX.

Data availability

The datasets generated during the current study are available in Genbank (SRA data-BioProject ID: PRJNA1162262, Temporary SubmissionID: SUB14733853). All the data and scripts required to reproduce this analysis are available from: https://github.com/dmecheverri/Flu_bat_H18N12.

Received: 16 September 2024; Accepted: 18 February 2025
Published online: 22 February 2025

References

1. Gupta, P. et al. Bats and viruses: a death-defying friendship. *VirusDisease* 32, 467–479 URL (2021). https://link.springer.com/article/10.1007/s13337-021-00716-0
2. Guzmán, C. et al. Molecular and cellular evidence of natural Venezuelan equine encephalitis virus infection in frugivorous bats in Colombia. *Veterinary World*. **13**, 495–501 (2020). https://pubmed.ncbi.nlm.nih.gov/32367955/
3. Calderón, A. et al. Dengue virus in bats from C´ordoba and Sucre, Colombia. *VectorBorne Zoonotic Dis.* **19**, 747–751 (2019). https://www.liebertpub.com/doi/10.1089/vbz.2018.2324
4. Calderón, A. et al. Two cases of natural infection of dengue-2 virus in bats in the colombian caribbean. *Tropical Medicine and Infectious Disease* 6, 35 URL (2021). https://www.mdpi.com/2414-6366/6/1/35
5. Tong, S. et al. A distinct lineage of influenza a virus from bats. *Proceedings of the National Academy of Sciences of the United States of America* 109, 4269–4274 (2012). https://doi.org/10.1073/pnas.1116200109
6. Tong, S. et al. New world bats harbor diverse influenza a viruses. *PLOS Pathogens* 9, e1003657 URL (2013). https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1003657
7. Campos, A., Góes, B., Moreira-Soto, A., de Carvalho, C., Ambar, G., Sander, A. L.,… Drexler, F. Bat influenza A (HL18NL11) virus in fruit bats, Brazil. *Emerging infectious diseases* 25 (2), 333–337 (2019). URL https://pmc.ncbi.nlm.nih.gov/articles/PMC6346480/.
8. Kandeil, A. et al. Isolation and characterization of a distinct influenza a virus from egyptian bats. *Journal of Virology* 93 URL https://journals.asm.org/doi/full/ (2019). https://doi.org/10.1128/JVI.01059-18
9. Ciminski, K., Pfaff, F., Beer, M. & Schwemmle, M. Bats reveal the true power of influenza a virus adaptability. *PLOS Pathogens* 16, e1008384 URL (2020). https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1008384
10. Ciminski, K. et al. Bat influenza viruses transmit among bats but are poorly adapted to non-bat species. *Nature Microbiology* 4, 2298–2309 URL (2019). https://www.nature.com/articles/s41564-019-0556-9
11. Zhu, X. et al. Crystal structures of two subtype n10 neuraminidase-like proteins from bat influenza a viruses reveal a diverged putative active site. *Proceedings of the National Academy of Sciences of the United States of America* 109, 18903– 18908 URL https://www.pnas.org/doi/abs/ (2012). https://doi.org/10.1073/pnas.1212579109
12. Sun, X. et al. Bat-derived influenza hemagglutinin h17 does not bind canonical avian or human receptors and most likely uses a unique entry mechanism. *Cell. Rep.* **3**, 769–778 (2013). https://www.cell.com/cell-reports/fulltext/S2211-1247(13)00032-6?script=true

13. Karakus, U. et al. Mhc class Ii proteins mediate cross-species entry of Bat influenza viruses. *Nature* **567**, 109–112 (2019). https://www.nature.com/articles/s41586-019-0955-3
14. Giotis, E. S. et al. Entry of the bat influenza H17N10 virus into mammalian cells is enabled by the mhc class ii hla-dr receptor. *Nature Microbiology* 4, 2035–2038 URL (2019). https://www.nature.com/articles/s41564-019-0517-3
15. Müller, N. F., Stolz, U., Dudas, G., Stadler, T. & Vaughan, T. G. Bayesian inference of reassortment networks reveals fitness benefits of reassortment in human influenza viruses. *Proceedings of the National Academy of Sciences* 117, 17104–17111 URL https://www.pnas.org/doi/ (2020). https://doi.org/10.1073/pnas.1918304117
16. Dou, D., Revol, R., Ostbye, H., Wang, H. & Daniels, R. Influenza a virus cell¨ entry, replication, virion assembly and movement. *Frontiers in Immunology* 9, 383042 URL https://www.frontiersin.org/journals/immunology/articles/ (2018). https://doi.org/10.3389/fimmu.2018.01581/full
17. Wu, Y., Wu, Y., Tefsen, B., Shi, Y. & Gao, G. F. Bat-derived influenza-like viruses H17N10 and H18N11. *Trends Microbiol.* **22**, 183–191 (2014). https://www.cell.com/trends/microbiology/fulltext/S0966-842X(14)00022-5
18. Guy, C., Ratcliffe, J. M. & Mideo, N. The influence of bat ecology on viral diversity and reservoir status. *Ecology and Evolution* 10, 5748–5758 URL (2020). https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.6315.
19. Solanki, A. et al. The role of hydrophobicity in peptide-MHC binding. *Lecture Notes Comput. Sci. (including Subser. Lecture Notes Artif. Intell. Lecture Notes Bioinformatics)*. **13060 LNBI**, 24–37 (2021). https://link.springer.com/chapter/10.1007/978-3-030-91241-3_3
20. Varghese, J. N. & Colman, P. M. Three-dimensional structure of the neuraminidase of influenza virus a/tokyo/3/67 at 2.2 a resolution. *Journal of molecular biology* 221, 473–486 URL (1991). https://pubmed.ncbi.nlm.nih.gov/1920428/
21. Zhong, G. et al. Mutations in the neuraminidase-like protein of bat influenza H18N11 virus enhance virus replication in mammalian cells, mice, and ferrets. *Journal of Virology* 94 URL (2020). https://journals.asm.org/doi/10.1128/jvi. 01416-19.
22. Varghese, J. et al. (ed, N.) Structural evidence for a second Sialic acid binding site in avian influenza virus neuraminidases. *Proc. Natl. Acad. Sci. U.S.A.* **94** 11808–11812 https://pubmed.ncbi.nlm.nih.gov/9342319/ (1997).
23. Smith, B. J. et al. Structure of a calcium-deficient form of influenza virus neuraminidase: implications for substrate binding. *Acta crystallographica. Section D, Biological crystallography* 62, 947–952 URL (2006). https://pubmed.ncbi.nlm.nih.gov/16929094/
24. Halwe, N. J. et al. Bat-borne H9N2 influenza virus evades MxA restriction and exhibits efficient replication and transmission in ferrets. *Nature Communications* 15, 1–8 URL (2024). https://www.nature.com/articles/s41467-024-47455-6
25. Miller, J. M. et al. Guidelines for safe work practices in human and animal medical diagnostic laboratories recommendations of a CDC-convened, biosafety blue ribbon panel centers for disease control and prevention MMWR editorial and production staff MMWR editorial board. *Centers Dis. Control Prev. Morb Mortal. Wkly. Rep.* **61**, 105 (2012). https://pubmed.ncbi.nlm.nih.gov/22217667/
26. Diaz, M., Solari, S., Aguirre, L., Aguiar, L. & Barquez, R. Clave de identificación de los murciélagos de SudaméricaEditorial Magna, (2016).
27. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one fastq preprocessor. *Bioinformatics* 34, i884–i890 (2018). https://doi.org/10.1093/bioinformatics/bty560
28. Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. Megahit: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinf. (Oxford England).* **31**, 1674–1676 (2015). https://pubmed.ncbi.nlm.nih.gov/25609793/
29. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using diamond. *Nature Methods* 12, 59–60 URL (2014). https://www.nature.com/articles/nmeth.3176
30. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with bowtie 2. *Nature Methods* 9, 357–359 URL (2012). https://www.nature.com/articles/nmeth.1923
31. Li, H. et al. The sequence alignment/map format and samtools. *Bioinformatics (Oxford, England)* 25, 2078–2079 URL (2009). https://pubmed.ncbi.nlm.nih.gov/19505943/
32. Okonechnikov, K. et al. Unipro ugene: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166–1167. https://doi.org/10.1093/bioinformatics/bts091 (2012).
33. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinf. (Oxford England).* **30**, 2068–2069 (2014). https://pubmed.ncbi.nlm.nih.gov/24642063/
34. Kuraku, S., Zmasek, C. M., Nishimura, O. & Katoh, K. aleaves facilitates ondemand exploration of metazoan gene family trees on mafft sequence alignment server with enhanced interactivity. *Nucleic Acids Research* 41, W22–W28 (2013). https://doi.org/10.1093/nar/gkt389
35. Nguyen, L. T., Schmidt, H. A., Haeseler, A. V. & Minh, B. Q. Iq-tree: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32, 268–274 (2015). https://doi.org/10.1093/molbev/msu300
36. Sagulenko, P., Puller, V., Neher, R. A. & Treetime Maximum-likelihood phylodynamic analysis. *Virus Evolution* 4 (2018). https://doi.org/10.1093/ve/vex042
37. Samson, S., Etienne, L. & Makarenkov, V. Simplot++: a python application´ for representing sequence similarity and detecting recombination. *Bioinformatics (Oxford, England)* 38, 3118–3120 URL (2022). https://pubmed.ncbi.nlm.nih.gov/35451456/
38. Drummond, A. J. & Rambaut, A. BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 1–8 (2007). https://bmcecolevol.biomedcentral.com/articles/10.1186/1471-2148-7-214
39. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in bayesian phylogenetics using tracer 1.7. *Systematic Biology* 67, 901–904 (2018). URL https://dx.doi.org/10.1093/sysbio/syy032.
40. Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., … Drummond,A. J. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology* 10, e1003537 (2014). URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003537.
41. Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. Swiss-model: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381 (2003). URL /pmc/articles/PMC168927//pmc/articles/PMC168927/?report= abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC168927/.
42. Goddard, T. D. et al. Ucsf chimerax: Meeting modern challenges in visualization and analysis. *Protein science: a publication of the Protein Society* 27, 14–25 URL (2018). https://pubmed.ncbi.nlm.nih.gov/28710774/
43. Honorato, R. V. et al. Structural biology in the clouds: The wenmr-eosc ecosystem. *Frontiers in Molecular Biosciences* 8, 729513 URL https://www.frontiersin.org/journals/molecular-biosciences/articles/ (2021). https://doi.org/10.3389/fmolb.2021.729513/full

Acknowledgements

The authors are grateful for the work done by the laboratory staff of the University of Cordoba, Instituto de Investigaciones Biológicas del Trópico. Ministerio de Ciencia, Tecnología e Innovación de Colombia (MINCIENCIAS).

Author contributions

D.E.D contributed to the protein structure prediction, molecular docking and writing of the manuscript. D.E.D and R.R contributed to the quality control of the genome, bioinformatic and phylogenetic analyses. D.E.D,

C.M.B, R.H, M.A, E.G, and V.B carried out the sample collection, RNA sequencing and data analysis. D.E.D, B.G.P, Y.L, G.A, J.D.R and S.M contributed to the methodological design of the study, results analysis and discussion. All the authors reviewed and approved the manuscript.

Funding

This research was founded by a grant of the Ministry of Science, Technology and Innovation of Colombia. Code 91722-contract 601–2022. Project: "Strengthening public health research capacities: Metagenomics of infectious agents in mosquitoes and bats from four departments of the Colombian Caribbean." R.R was supported by funding to Verena (viralemergence.org) from the U.S. National Science Foundation, including NSF BII 2021909 and NSF BII 2213854.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-91026-8.

Correspondence and requests for materials should be addressed to S.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.