

ARTICLE

Open Access

Evidence of a fixed internal gene constellation in influenza A viruses isolated from wild birds in Argentina (2006–2016)

Agustina Rimondi¹, Ana S. Gonzalez-Reiche^{2,7}, Valeria S. Olivera¹, Julieta Decarre³, Gabriel J. Castresana⁴, Marcelo Romano⁵, Martha I. Nelson⁶, Harm van Bakel⁷, Ariel J. Pereda^{1,8}, Lucas Ferreri², Ginger Geiger² and Daniel R. Perez²

Abstract

Wild aquatic birds are the major reservoir of influenza A virus. Cloacal swabs and feces samples ($n = 6595$) were collected from 62 bird species in Argentina from 2006 to 2016 and screened for influenza A virus. Full genome sequencing of 15 influenza isolates from 6 waterfowl species revealed subtypes combinations that were previously described in South America (H1N1, H4N2, H4N6 ($n = 3$), H5N3, H6N2 ($n = 4$), and H10N7 ($n = 2$)), and new ones not previously identified in the region (H4N8, H7N7 and H7N9). Notably, the internal gene segments of all 15 Argentine isolates belonged to the South American lineage, showing a divergent evolution of these viruses in the Southern Hemisphere. Time-scaled phylogenies indicated that South American gene segments diverged between ~ 30 and ~ 140 years ago from the most closely related influenza lineages, which include the avian North American (PB1, HA, NA, MP, and NS-B) and Eurasian lineage (PB2), and the equine H3N8 lineage (PA, NP, and NS-A). Phylogenetic analyses of the hemagglutinin and neuraminidase gene segments of the H4, H6, and N8 subtypes revealed recent introductions and reassortment between viruses from the Northern and Southern Hemispheres in the Americas. Remarkably and despite evidence of recent hemagglutinin and neuraminidase subtype introductions, the phylogenetic composition of internal gene constellation of these influenza A viruses has remained unchanged. Considering the extended time and the number of sampled species of the current study, and the paucity of previously available data, our results contribute to a better understanding of the ecology and evolution of influenza virus in South America.

Introduction

Influenza A Viruses (IAVs) belong to the family *Orthomyxoviridae*. The genome of IAVs is composed of 8 segments of negative-sense, single-stranded RNA. Wild aquatic birds of the order Anseriformes and Charadriiformes are considered the natural hosts of IAVs. Based on the antigenic characteristics of their surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA),

IAVs are divided into subtypes. To date, 16 HA subtypes (H1 to H16) and 9 NA subtypes (N1 to N9) have been recognized in wild birds. Due to the segmented nature of their genomes, IAVs frequently exchange gene segments, which has led to the identification of many subtype combinations in waterfowl species¹. Regardless of the subtype, IAV infections in wild aquatic birds occur mostly in the intestinal tract in an apparently asymptomatic fashion. The virus is excreted in large amounts in the feces, which allows for an efficient fecal-oral route mode of transmission among birds.

Many of the efforts to understand the ecology of IAVs in their natural reservoir have been driven by a desire to track the origin of pandemic strains; prompted in

Correspondence: Agustina Rimondi (rimondi.agustina@inta.gob.ar) or Daniel R. Perez (dperez1@uga.edu)

¹Instituto de Virología CICVyA – Instituto Nacional de Tecnología Agropecuaria (INTA), CC25 (1712) Castelar, Buenos Aires, Argentina

²Poultry Diagnostic and Research Center, College of Veterinary Medicine, University of Georgia, 953 College Station Rd, Athens, GA 30602, USA
Full list of author information is available at the end of the article.

© The Author(s) 2018



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

particular by realization that the pandemic IAV viruses of the 1957 Asian flu (H2N2) and of the 1968 Hong Kong flu (H3N2) carried gene segments (HA, NA, and PB1; and HA and PB1, respectively) similar to IAV strains from ducks². Renewed interest in understanding the ecology of IAVs in natural hosts resulted from the emergence of the HPAIVs of the H5 subtype in China with the ability to cause exacerbated and often lethal respiratory disease in humans. The HA gene segment of these viruses has continued to evolve and diversify into genetically distinct clades and sub-clades and have successfully spread across 4 different continents^{3,4}. Historically, long-term active IAV surveillance in wild birds has been circumscribed to the United States, Canada, Western Europe, Southeast Asia, and China. In other parts of the world, surveillance efforts have been mostly passive, in response to disease outbreaks^{5,6}. As a result, only 0.1% of all available sequences from IAV isolated from wild birds originate from South America⁷. The first evidence of IAV in birds in South America was associated to an outbreak of highly pathogenic avian influenza virus (HPAIV) of the H7 subtype in commercial chickens and turkeys in Chile in 2002⁸.

The evolution of IAVs in the natural reservoir is largely determined by the geographical separation and migratory patterns of the avian hosts. Thus, IAVs are broadly separated into lineages that correspond to vast geographic regions. IAVs of North American (NAM), Eurasian (EAs), and Australian lineages are well defined due to the significant amount of sequence information of IAV isolates available from those regions^{9–11}. More recently, additional lineages were proposed for South America^{12–14} and Antarctica⁷ based on phylogenetic analyses of a limited number of IAV isolates obtained from wild birds.

Our group has been conducting active IAV surveillance in wild birds in Guatemala (Central America) and in Argentina (South America) as part of efforts to establish early warning systems for the potential introduction of IAVs from Asia into the Americas and/or from wild birds into commercial poultry. In the present study, we performed full genomic characterization and phylogenetic analyses of 15 low pathogenic avian influenza viruses (LPAIVs) isolated from 6 different avian species in Argentina between 2006 and 2016. We inferred time-scaled phylogenies using a previously described host specific local clock model¹⁵ using currently available sequences of IAVs from South America. Our studies provide further support for the existence of a South American (SAM) IAV lineage circulating in wild birds in Argentina. Interestingly, while most HA and NA subtypes obtained from the IAVs from Argentina follow unique evolutionary pathways within the SAM lineage, the HA and NA gene segments of three subtypes, H4, H6, and N8, appear to be recent introductions from IAVs of the NAM

lineage. These analyses provide novel insights into the evolutionary patterns of these viruses. Taken together, our study highlights the importance of continuous surveillance of IAVs in South America as an integral part of worldwide efforts to better understand the ecology of these viruses and to better monitor the circulation of strains with zoonotic and/or pandemic concern.

Results

IAVs isolated from wild birds in Argentina with subtype combinations and from hosts not previously described in the region

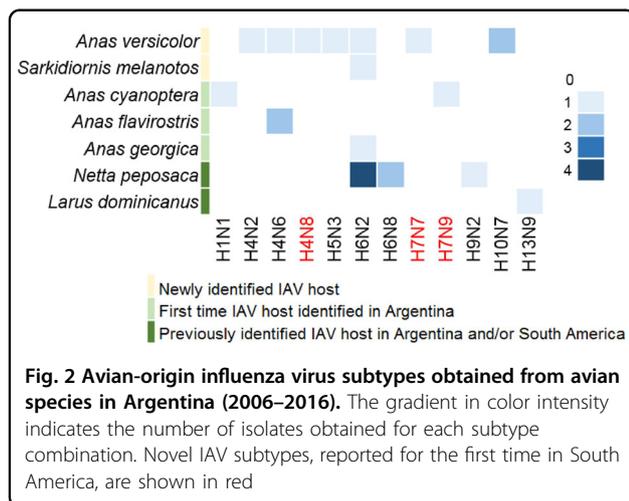
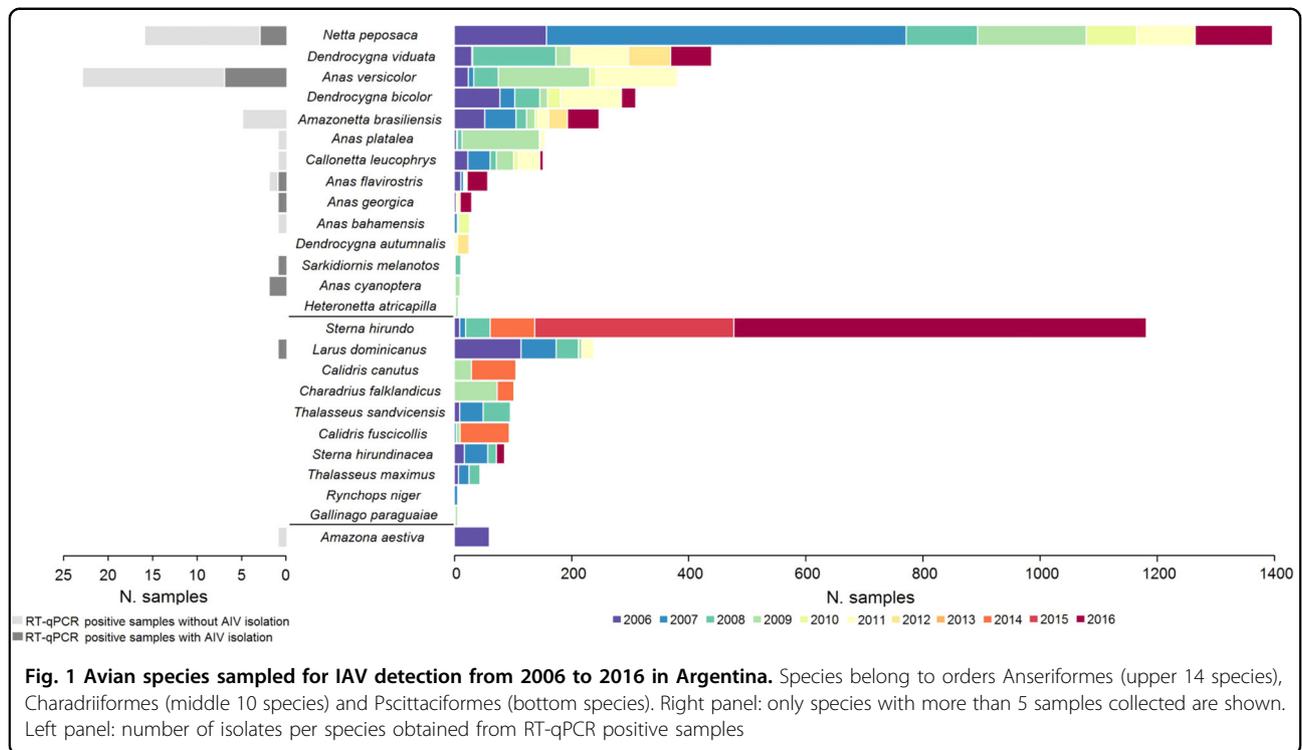
In total, 6525 cloacal swabs and 70 fecal samples representing 62 different bird species were collected and tested for the presence of IAV (Fig. 1 and STable 1). Birds of the order Anseriformes ($n = 3258$; 49.4%), Charadriiformes ($n = 1973$; 29.9%) and Sphenisciformes ($n = 744$; 11.3%) comprised the majority of the birds sampled. Of these samples, 55 tested positive for IAV by RT-qPCR (0.83%), in their majority from ducks ($n = 53$), one from gull (*Larus dominicanus*) and one from parrot (*Amazona aestiva*) (STable 2).

Viable IAVs were isolated from twenty-two RT-qPCR positive samples representing seven different bird species (Fig. 2), including *Anas versicolor* (silver teal, $n = 8$), *Netta peposaca* (rosy-billed pochard, $n = 7$), *Anas cyanoptera* (cinnamon teal $n = 2$), *Anas flavirostris* (yellow-billed teal, $n = 2$), *Anas georgica* (yellow-billed pintail, $n = 1$), *Sarkidiornis melanotos* (comb duck, $n = 1$) and *Larus dominicanus* (kelp gull, $n = 1$). Of these isolates, 7 have been previously described (subtypes H6N2 (3), H6N8 (2), H9N2 and H13N9, STable 3)^{12–14}. Full genome sequences of the additional 15 IAV isolates (HA subtypes H1, H4, H5, H6, H7, and H10, and NA subtypes N1, N2, N3, N6, N7, N8, and N9) were generated for further characterization (STable 3). The subtype combinations H4N8, H7N7 and H7N9 have not been reported previously in South America (shown in red in Fig. 2). Previously unknown IAV hosts were also identified including *Anas versicolor* (silver teal) and *Sarkidiornis melanotos* (comb duck). Isolates from *Anas cyanoptera* (cinnamon teal), *Anas flavirostris* (yellow-billed teal), and *Anas georgica* (yellow-billed pintail) were obtained for the first time in Argentina (Fig. 2).

Together, the fifteen Argentinean IAVs account for nearly seventeen percent of IAVs obtained from wild birds throughout South America. With these data, Argentina is the country with the second-highest number of IAV isolates in the Southern Hemisphere (Fig. 3).

Phylogenetic analyses of internal gene segment constellations

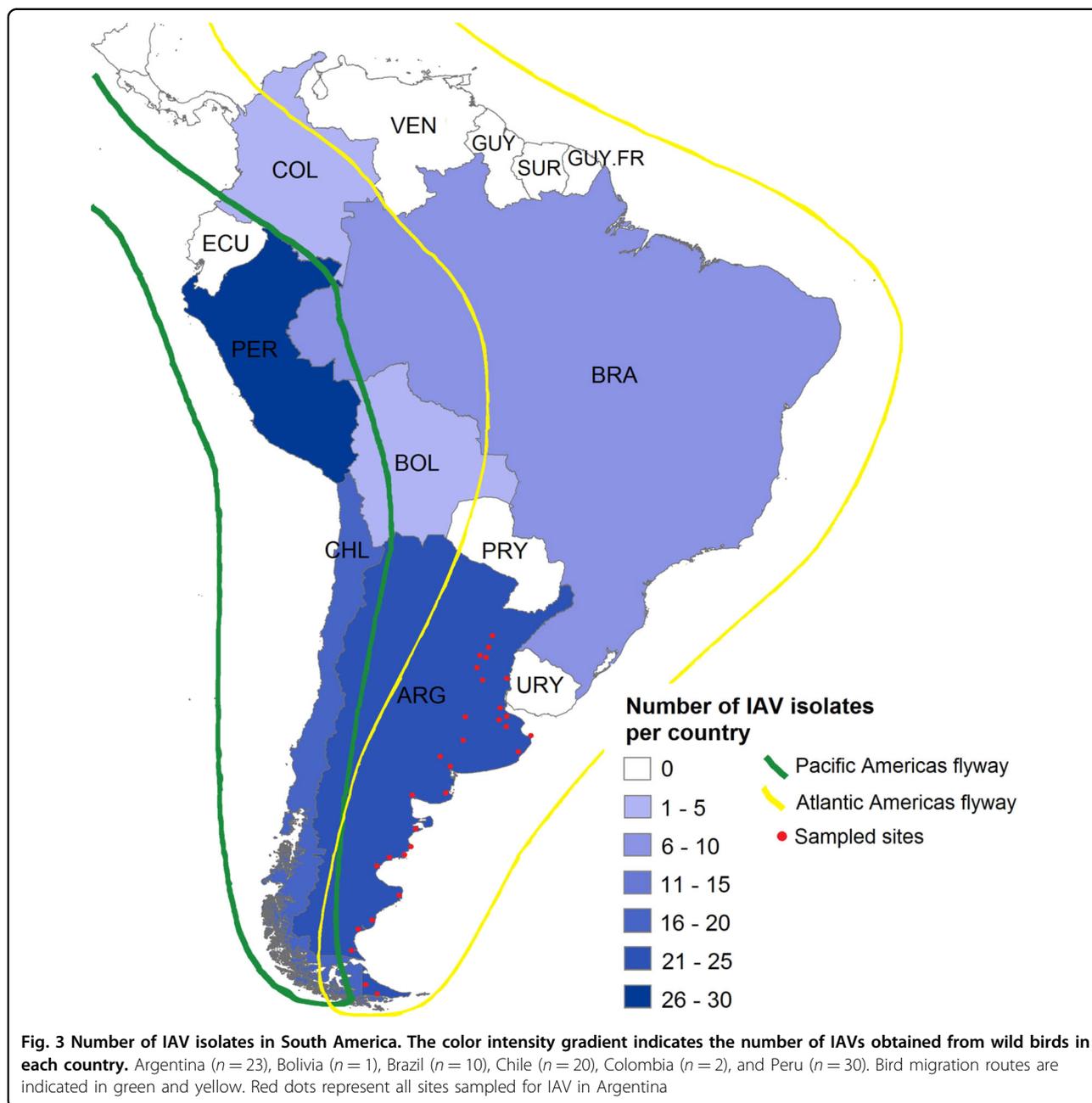
The major open reading frame (ORF) sequence of each internal gene segment was used for phylogenetic analyses.



Given prior evidence of possible common ancestry, sequences derived from equine IAVs were included in our analyses^{13,16,17}. Prototypic nucleotide IAV sequences from other hosts (bats, swine, and human—except 2009 pandemic H1N1) were also included. The internal gene segments of the new 15 IAVs analyzed clustered together and are highly similar to other previously reported IAVs of avian origin from South America (SFig. 1–6). We used molecular clock analyses that incorporate host-specific evolutionary rates of IAV gene segments¹⁵ to estimate the time of the most recent common ancestor (tMRCA) for

the internal gene segments of the IAV isolates from Argentina (Fig. 4 and STable 4), which indicated that these viruses diverged from other main IAV lineages between ~30 and ~140 years ago. A closer examination of each tree shows different topologies for each gene segment that indicate that the modern SAM lineage arose from multiple reassortment events, the exact timing of which remain unclear. The PB2 segment of the IAV strains from Argentina was introduced from an EAs-like IAV ancestor (Fig. 4 and SFig. 1) in the late 1880's (tMRCA_{PB2} ~ 1887), in agreement with the previous analyses¹⁵. For this segment, the SAM avian clade diverged from the EAs lineage sometime between 1885 (tMRCA between EAs and SAM clades) and 1990 (tMRCA for SAM clade), implying a ~100-year gap of unknown diversity for the ancestors of the contemporary SAM clade. In contrast, the PB1 segment (Fig. 4 and SFig. 2) diverged more recently from a NAM lineage-like ancestor (tMRCA_{PB1} ~ 1988). Interestingly, the PA (tMRCA_{PA} ~ 1943, Fig. 4 and SFig. 3) and NP (tMRCA_{NP} ~ 1951, Fig. 4 and SFig. 4) segments diverged from the West equine-like (H3N8) lineage (Fig. 4 and STable 4) and share common ancestry with other IAVs from South America, Antarctica, and other equine-origin IAVs^{7,18}. Similar to the PB1, the M segment diverged more recently from a NAM lineage ancestor (tMRCA_M ~ 1971, Fig. 4 and SFig. 5) (STable 4).

Phylogenetic analysis of the NS genes showed that 13 of the IAVs from Argentina contain the NS allele A (NS-A),



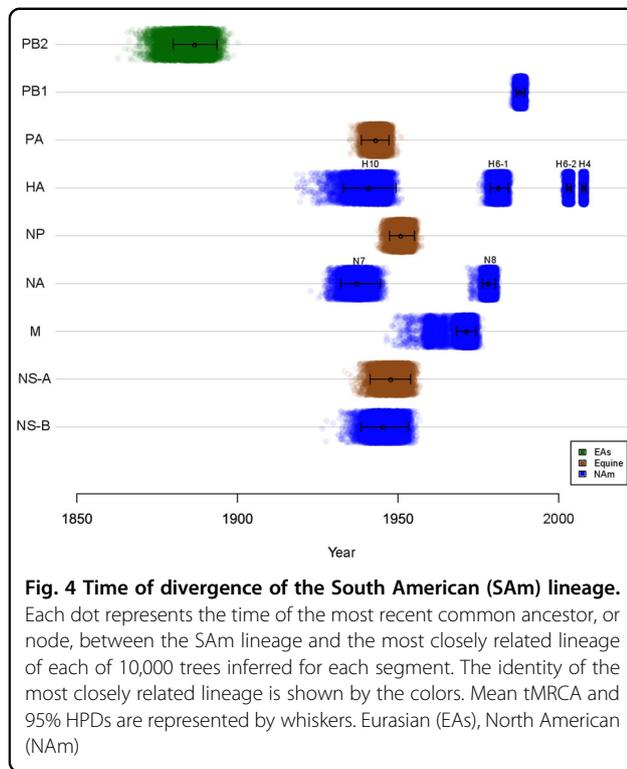
whereas only two carry allele B (NS-B) (SFig. 6). Both of these alleles diverged from the main IAV lineages around the same time, between 1939 and 1954 (Fig. 4 and STable 4). Like the PA and NP segments, the NS-A gene segments form a sister group with IAVs of equine origin^{13,16,17}.

Analysis of the internal gene segment constellations for all Argentinean LPAIV strains and other viruses obtained from wild aquatic birds and poultry in South America revealed that all viruses from Argentina are positioned exclusively within the South American lineage. In

contrast, viruses from Bolivia, Brazil, Chile, Colombia, and Peru include reassortants with IAVs of other genetic lineages (Fig. 5, STable 5 and S Figs. 1–6).

Phylogenetic analyses of the surface gene segments

Overall the HA gene segments of Argentinean IAVs clustered with other South American IAVs (H1, H5, H7, H4, and a subset of H6, named H6-2), with the exception of the H10 and a subset of H6 (H6-1) gene segments that formed a cluster of IAVs exclusively from Argentina. We denoted all these clades as SAm lineage. Following is a



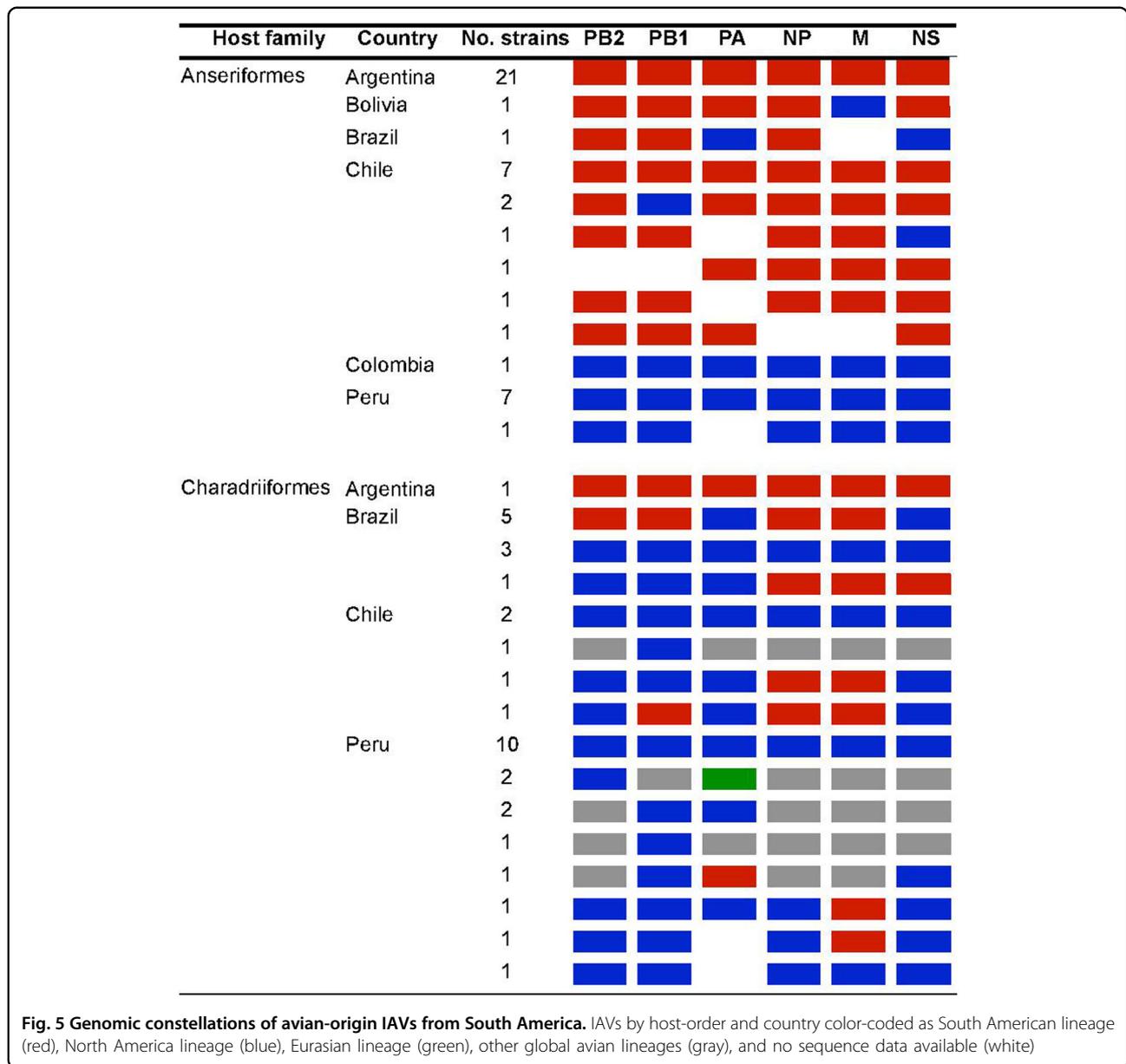
description of the main findings for each HA subtype. The H1 and H5 isolates from Argentina grouped within the NA lineage (Fig. 6 and S Figs. 7 and 8). The H1 subtype SA lineage viruses were recovered from resident wild birds (Anseriformes and Tinaniformes) from Argentina and Chile (Fig. 6 and S Fig. 7). The other H1 subtype IAVs identified in Peru (Charadriiformes and Gruiformes) belongs to the NA lineage (Fig. 6 and S Fig. 7). The H5 subtype SA lineage contains only two sequences from IAVs from resident duck species (silver teal and yellow-billed pintail) in Argentina and Chile. There were other two highly similar H5 sequences in the database (from Chile), but they were excluded from the analysis because they were only partial fragments of the H5. In addition, there are other H5 subtype IAVs isolated from different wild duck spp. in South America (Chile and Colombia), but they cluster in a different position within the NA lineage and are less divergent compared to the Argentinean isolate (Fig. 6 and S Fig. 8). The H7 subtype from IAVs from the SA lineage was basal to the main NA lineage, showing two clades (S Fig. 9). One clade contains the two H7 IAVs from Argentina and other IAVs from Chile (all IAV isolates from resident wild duck spp.). The second clade contains the HPAIV poultry isolates from the outbreak in Chile in 2002 and their closest wild birds LPAIV ancestor from Bolivia (cinnamon teal) (refs. 8,17). Although all H7 from Argentina are LPAIV, they share a common ancestor with H7 HPAIVs from

2002. Similar to the H5 subtype, the other H7 subtype identified from a wild duck (unknown spp.) in Peru belongs to the NA lineage (Fig. 6 and S Fig. 9).

The H10 subtype has been previously reported in Anseriformes and Charadriiformes in Peru¹⁹. For this subtype, the two H10N7 IAV isolates from Argentina grouped together and formed a unique branch that is clearly separated from H10 gene segments from viruses from other geographic locations, and remarkably distant from the H10N7 viruses from Peru (S Fig. 10, red branches). As this was the first isolation of a SA H10 subtype, we inferred an MCC tree to estimate the time of divergence of these viruses. Both H10 Argentinean viruses, isolated from silver teal, have a common ancestor with H10 viruses from NA lineage, with an estimated tMRCA between 1933 and 1949 (95% HPD) (Fig. 4 and S Table 4).

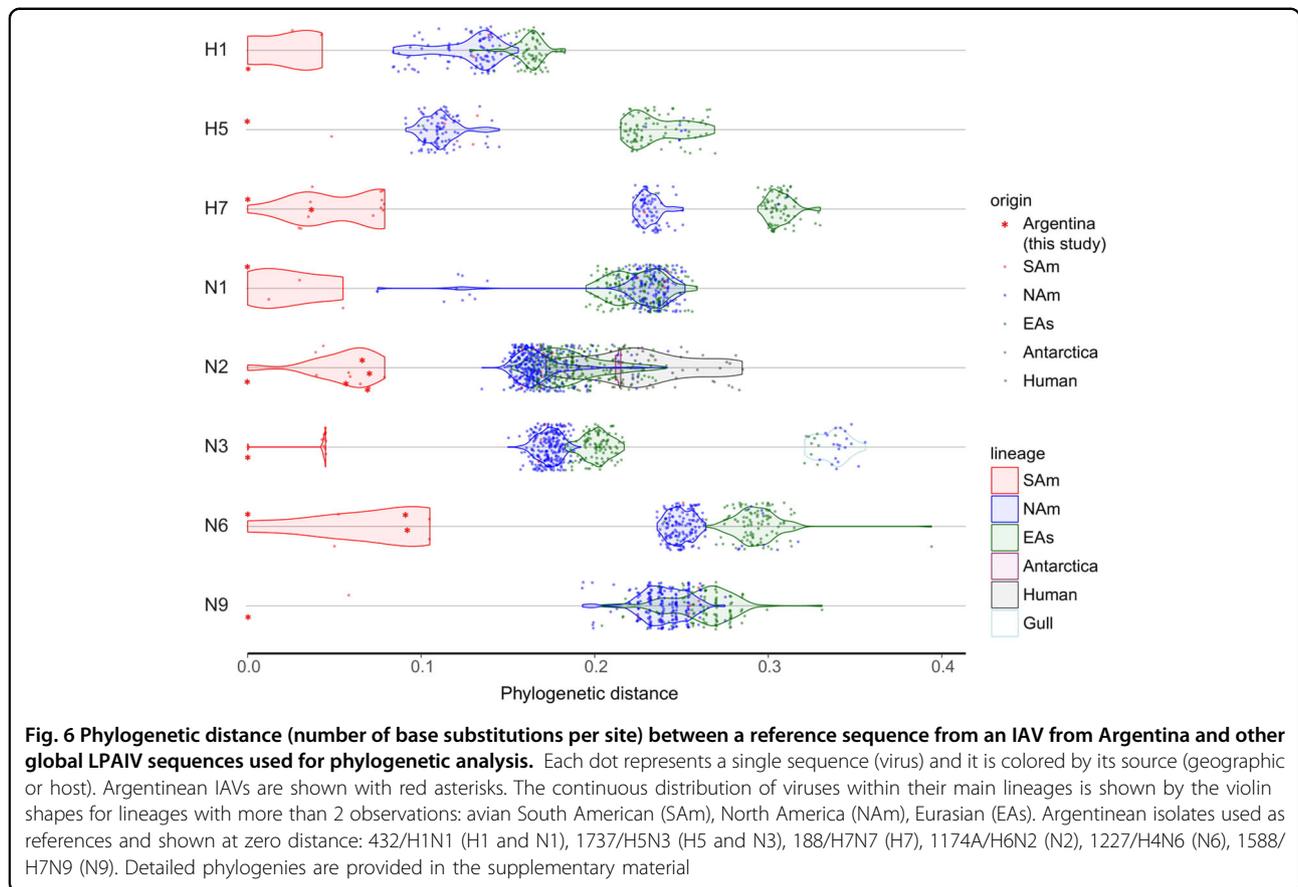
Phylogenetic analysis of H4 HA gene segment showed that the IAVs from Argentina (four isolates from 2011 and one from 2016) grouped within the NA lineage and were closely related to other H4 viruses identified in Anseriformes in Chile in 2013 and 2015, suggesting dissemination of NA H4 subtype IAVs in the region (S Fig. 11). As this seemed to be a recent introduction to South America we estimated the time of introduction for this clade. Time-scaled phylogenetic analysis estimated that the introduction of these NA H4 viruses in South America occurred between 2007–2009 (95% HPD) (Fig. 4 and S Table 4). Similarly, the MCC tree of H6 HA gene showed at least two separate introductions of H6 viruses in Argentina (identified as H6-1 and H6-2 in Fig. 4 and S Fig. 12). The first introduction (H6-1) includes two new H6 isolates (272/H6N2 and 1174A/H6N2) that clustered together with four H6 IAVs from Argentina published previously¹³. In this case, we estimated that the SA clade (as well as the most closely related NA clade) has diverged from the EAs lineage sometime between 1978 and 1984 (95% HPD, Fig. 4 and S Table 4). The second most recent introduction (H6-2) includes other two new H6 isolates (49/H6N2 and 52/H6N2) that grouped together with one of our previous H6 subtype IAVs (1977/H6N2) and three IAVs from Brazil isolated in 2012 (S Fig. 12). On this occasion, the SA clade seems to have diverged from the NA lineage sometime between 2002 and 2004 (95% HPD, Fig. 4 and S Table 4). The new Argentinean H6 IAVs were obtained from four wild duck spp. (comb duck, rosy-billed pochard, silver teal, and yellow-billed pintail), while Brazilian H6 IAVs were obtained from a small shorebird (white-rumped sandpiper, a Charadriiforme).

For the NA gene segments, phylogenetic analysis of the N1, N2, N3, N6, and N9 gene segments grouped them into the unique SA clades as previously observed^{12–14} (Fig. 6 and S Figs. 13–17). The N1 subtype SA lineage



contains four similar sequences, three are paired with the H1 HA subtype from IAVs isolated in Argentina and Chile (from Tinaniformes and Anseriformes respectively), and one is paired with the H6 HA subtype from IAVs isolated from white-rumped sandpiper in Brazil (SFig. 13). These four viruses form a sister group with some IAVs isolated from shorebirds in Delaware Bay and New Jersey in a cluster that is basal to the NAM and EAs lineages. The N1 of other IAVs identified from Charadriiformes in Brazil and Peru belong to the NAM lineage. The N2 phylogenetic tree shows IAVs from SAM lineages share ancestry with avian IAVs from diverse origins that are basal to the main NAM and EAs lineages. The SAM viruses grouped in two clades (SFig. 14), one clade

contains IAVs isolated exclusively from Anseriformes in Argentina and Chile between 2007 and 2014 ($n = 9$, four new sequences from this report). The second clade contains one N2 subtype from an IAV isolated in Argentina in 2016 (this report) and N2 sequences from IAVs from Chile (2015). The other N2 IAVs identified in Chile, Colombia, and Peru belong to the NAM lineage. The N3 subtype from IAVs from the SAM lineage also shows two clades (SFig. 15). One clade contains only two H5N3 IAVs isolated from wild ducks, one from Argentina (this report) and the other from Chile. There is a 6-year gap between these IAVs (2009 and 2015) and they were isolated from different species—silver teal and yellow-billed pintail, respectively. The second clade contains only



H7N3 HPAIV isolate from poultry in Chile in 2002 (ref. ⁸) (SFig. 15). These two clades shared ancestry with non-contemporary NAm viruses, that are basal to NAm lineage. The N6 subtype SAM lineage, basal to the NAm lineage, contains the three newly identified N6 IAVs from Argentina and other sequences from Chile, all from resident wild bird spp. (SFig. 16). The only other N6 IAV identified in South America, in Peru, belongs to the NAm lineage and was isolated from a gull. Interestingly, our last N2 and N6 IAV isolates from 2016 (1174A/H6N2 and 1227/H4N6 respectively) are more distant from previous isolates from Argentina and more closely related to recent N2 and N6 segments from IAVs isolated in Chile (SFig. 14 and 16 respectively). With respect to the N9 gene segments, those obtained from IAVs in Argentina ($n = 2$) were basal to the main EAs and NAm lineages, whereas N9 subtype sequences from other IAV isolates from Brazil, Chile and Peru clearly belong to the NAm lineage (SFig. 17).

The N7 subtype phylogeny shows that sequences of those IAVs isolated in Argentina, all from silver teals, are unique and share no relationship with those obtained from wild bird isolates in Chile²⁰ or Peru¹⁹ (SFig. 18, red branches). We estimated the time of divergence of these

distinct viruses. The N7 subtype sequences from IAVs from Argentina have a NAm ancestor with a tMRCA between 1932–1945 (95% HPD), similar to the H10 HA subtype sequences (Fig. 4 and STable 4).

Finally, three N8 IAVs from Argentina clustered together as a distinct clade within the NAm lineage. Time-scaled phylogenetic analysis estimated that the time of introduction of the N8 subtype in Argentina was between 1976–1980 (95% HPD) (Fig. 4 and STable 4), which overlaps with the first introduction of H6 HA in SAM, suggesting common origins in a NAm ancestor. The N8 NA phylogenetic tree revealed that previously described sequences from IAVs from wild birds (not Anseriformes) in Peru¹⁹ also belong to the NAm lineage but represent a more recent, independent introduction, different from Argentina (red branches SFig. 19).

Molecular analyses of virulence markers

The diversity of HA subtypes found in IAVs from Argentina confirm the circulation of group 1 (H1, H5, H6, H9, and H13) and group 2 (H4, H7, and H10) of HAs previously proposed²¹. Previous studies have demonstrated the amino acid changes in the receptor binding site at positions 190, 225, 226, and 228 of the HA protein

(H3 numbering) in a variety of subtype, including H1, H2, H3, H4, H5, H6, and H7^{22–25}. Analysis of predicted HA amino acid sequences showed that all Argentinean viruses isolated from 2006 to 2016, regardless of subtype, contain amino acid residues consistent with binding to “avian-like” α 2,3 sialic acid receptors (E190, G225, Q226, and G228, data not shown). In addition, 16 out of 22 Argentinean IAV strains shared potential glycosylation sites of asparagine (Asn) along the HA1 domain (H3 numbering) with IAVs previously isolated in other regions: Asn129²⁶ for 432/H1N1 strain; Asn6, Asn22, Asn165, Asn196, and Asn483²⁷ for all H4 strains ($n=5$); Asn158²⁸ for 1737/H5N3 strain; and Asn21, Asn33, Asn170, Asn291, and Asn296²⁹ for all H6 strains ($n=9$). Instead, the 188/H7N7 and 1588/H7N9 strains maintain Asn133 but have the mutation N158T that abolishes a potential glycosylation site.

The HA gene segments of the H1, H4 ($n=5$), H5, H6 ($n=9$), H9 and H13 subtypes, have cleavage site motifs typical of LPAIV (Table 1). The HA gene segment of the H7 ($n=2$) and H10 ($n=2$) subtypes, have cleavage sites different from the consensus sequences described by Baron et al.³⁰ (Table 1). In addition to the arginine at the P1 position, the H7 HA cleavage site contains basic amino acids (lysine) at the P3 and P5 positions, and the H10 HA cleavage site carries histidine at the P3 position. It remains to be determined if such cleavage site motifs modulate the virulence of these viruses in wild birds and/or poultry.

No mammalian-associated virulence markers in NS1 protein (T92E)³¹ or PB2 (E627K and D701N)^{32–34} were found in the viruses from Argentina (STable 6). None of the viruses contained mutations associated with drug resistance in the NA and M2 proteins^{35,36}, except for the 432/H1N1 virus, which carries the mutation NA H274Y predictive of oseltamivir resistance³⁷. No mutations associated with enhanced transmission in mammals were found in the PB2 and PA proteins^{34,38}. In comparison to other lineages (NAM and EAs), none of the amino acid residues associated with drug resistance and virulence described here were unique to the Argentinean viruses or the SAM clade (STable 6). High degree of sequence homology at the protein level of the SAM lineage with other lineages was previously observed¹³.

The IAVs from Argentina contain the typical PB1-F2 ORF of 90 amino acids in length commonly found in IAVs from ducks^{39,40}. There are 45 amino acid differences among the PB1-F2 proteins in the IAVs from Argentina (SFig. 20). Interestingly, the PB1-F2 of Argentinean strains carry the N66S mutation implicated in increase virulence for mammals⁴¹, with the exception of the 52/H6N2 strain. In addition, the PB1-F2 protein of the 1227/H4N6 strain contains three amino acid markers, 51T, 56V, and 87E, associated with virulence in mallards of a highly pathogenic H5N1 avian influenza virus⁴². Other Argentinean

strains contain the 56V and 87E markers, the 1977/H6N2 contains only the 87E and the 1174A/H6N2 and 559/H9N2 strains contain only the 56V marker. Since the effects of PB1-F2 are virus- and host-dependent, further studies are needed to better understand the role of these PB1-F2 variants in the IAVs from Argentina^{40,41,43}.

The IAVs from Argentina also contain the predicted full-length PA-X ORF of 252 amino acids in length. There are 24 amino acid differences among the PA-X proteins in the IAVs from Argentina (SFig. 21). As is the case for PB1-F2, the effects of PA-X on virus replication and pathogenesis are strain- and host-dependent^{44,45}, and further studies are necessary to determine the biological significance of these mutations in birds.

Discussion

Long-term IAV surveillance in wild waterfowl in Argentina between 2006 and 2016 resulted in the recovery of 22 isolates, 15 of which are newly reported here. Despite limited and opportunistic sampling across host species and locations within Argentina throughout the 10 years of the study (shown as red dots in Fig. 3), we consistently observed that all internal segments of the Argentinean viruses were closely related to those from viruses previously isolated in Argentina, and exclusive to this geographical location in South America. Taken together, the results provide additional support for the existence of a unique SAM avian lineage with local independent evolution for all internal viral gene segments^{12–14,17}.

The phylogenetic analysis of the internal gene segments from the Argentinean IAVs showed shared ancestry with equine-origin IAVs (PA, NP, NS-A), and avian-origin IAVs of Eurasia (PB2), North America (PB1, M, NS-B) and Antarctica (PA, M), demonstrating that their genetic diversity has been shaped by multiple introductions and reassortment occurring as early as the end of the 19th century¹⁵. Interestingly, the IAVs isolated from wild birds in Argentina have all internal gene segments of SAM origin, while many Chilean and Peruvian viruses have either all or most of their internal genes segments of NAM origin, and close to those from Central America^{19,20,46}. More recently, H5N5 isolates from Antarctica were shown to carry gene constellations of NAM origin^{7,18,47}. While reassortment of internal gene segments from IAV from different lineages do occur in parts of South America, the Argentinean IAVs circulate mostly in isolation with limited gene segment exchange with viruses from other latitudes, regardless of their subtype.

Phylogenies of the H1, H5, H7, and H10 HA subtypes and the N1, N2, N3, N6, N7, and N9 NA subtypes demonstrated that both genes follow unique evolutionary patterns in wild birds in Argentina. Although the primordial source of these viruses remains unknown, the

Table 1 Cleavage site motifs of avian-origin influenza A viruses isolated in Argentina

Virus	Amino acid consensus sequence at the HA cleavage site (subtype) ^a	HA cleavage site ^b	Amino acid position in HA0	Virulence ^c
432/H1N1 ^d	PSIQ S RG (H1)	PSIQ A RG	339–345	Avirulent
32/H4N2 ^d	PEK A S/TRG (H4)	PEK A TRG	338–344	Avirulent
48/H4N6 ^d		PEK A TRG	338–344	Avirulent
91/H4N6 ^d		PEK A TRG	338–344	Avirulent
1227/H4N6a		PEK A TRG	338–345	Avirulent
25/H4N8 ^d		PEK A TRG	338–344	Avirulent
1737/H5N3 ^d	PQRE T RG (H5)	PQRE T RG	337–343	Avirulent
272/H6N2 ^d	PQIE T RG (H6)	PQIE T RG	339–345	Avirulent
557/H6N2		PQIE T RG	339–345	Avirulent
925/H6N2		PQIE T RG	339–345	Avirulent
1977/H6N2		PQIE T RG	339–345	Avirulent
49/H6N2 ^d		PQIE T RG	339–345	Avirulent
52/H6N2 ^d		PQIE T RG	339–345	Avirulent
1174 A/H6N2 ^d		PQIE T RG	339–345	Avirulent
269/H6N8		PQIE T RG	339–345	Avirulent
575/H6N8		PQIE T RG	339–345	Avirulent
188/H7N7 ^d	PEI/NPKG/ T RG (H7)	PEK <u>P</u> K T RG	333–340	Unkown
1588/H7N9 ^d		PEK <u>P</u> K T RG	333–340	Unkown
559/H9N2	PAR/K/XS/LX/ S RG (H7)	PAAS N RG	333–339	Avirulent
171/H10N7 ^d	PEI/VM/VQG/ E RG (H10)	PEI <u>H</u> E RG	334–341	Unkown
175/H10N7 ^d		PEI <u>H</u> E RG	334–341	Unkown
LDC4/H13N9	PAIS N /TRG (H13)	PAIS N RG	344–350	Avirulent

^aFrom Table 2 in Baron J et al., *J Virol.* 2013 Feb; 87(3): 1811–1820. The basic amino acid at the cleavage site is shown in bold

^bThe basic amino acid at the cleavage site is shown in bold. Amino acid at the HA cleavage site different from amino acid consensus sequences are underlined

^cVirulence phenotype inferred from protein sequence analysis

^dNew fifteen IAVs isolates from Argentina from this study

results suggest that the Argentinean strains split from their ancestors, either Eurasian or North American, ≤ 30 years ago and continue an independent evolution pathway. It must be noted that our surveillance efforts in Argentina have covered limited areas of the country and a limited subset of avian species. Together with the scarce data available from IAVs in South America, it is difficult to confirm if the basal lineages of SAM viruses are ancestral to other main IAV lineages (for example NAM lineage) or whether they are a consequence of earlier virus introductions (from any lineage) that then diverged separately in this region. Expansion into broader geographic areas, animal species, and number of samples are indispensable as a first step to better understand the ecology and evolution of IAVs in wild birds in Argentina. Sequence analysis of the Argentinean IAVs are consistent with typical LPAIVs; however, there are remarkably

divergent variants (e.g. H5, H7, H10) that warrant further experimental characterization in the laboratory to deepen our knowledge in the pathogenicity and host-range of these particular SAM IAVs.

In this study, we traced the origin of the Argentinean H4, H6, and N8 subtypes to North America, and detected at least four independent introduction events within the past 40 years. Particularly, the results of the H6 analyses further support the presence of two independent H6 subtype populations circulating in Argentina¹³. In addition, these two Argentinean H6 gene clusters are different from the H6 subtype population that circulates in Charadriiformes and Pelecaniformes in Peru, indicating distinct ways of dissemination of NAM H6 subtype IAVs along the region. The IAV gene influx from North America into South America could be driven by the migration of different bird species that arrive to the

Southern Hemisphere through the Atlantic and Pacific flyways (Fig. 3). With the available data, however, the bird species involved in this process and the frequency of such events remains unclear.

At the level of internal gene constellations, the lack of reassortment between the viruses from Argentina and other latitudes is notable (Fig. 5 and STable 5), but may be the result of sampling bias in geographical locations and/or host species. Several IAVs from Chile and Peru were obtained from gulls and shorebirds that perform long migrations that serve as important circuits for the dissemination of IAV^{48,49}. In contrast, Argentinean isolates were obtained almost exclusively from resident wild duck species with distribution more restricted to South America⁵⁰. Systematic surveillance of a larger number of IAV isolates obtained from South America will be necessary to trace the true origin of the SAM lineage, and further understanding of population densities and regional movement among migratory and resident birds in Argentina and other SAM countries is crucial to understand the dynamics of IAV transmission. Monitoring the movement of wild birds by geolocators, and identifying species subpopulations through analysis of stable isotopes and molecular genotyping^{51–53}, have the potential to establish the possible points of contact in which wild birds from North and South America overlap to allow virus transmission and reassortment. Such studies would also serve to elucidate whether geographic barriers, such as the Andes Mountains (natural border between Argentina and Chile), restrict the movement of wild bird populations and spillover of viruses from other genetic lineages (e.g., N.Am) into resident species, explaining the absence of reassortant virus strains in Argentina.

Materials and methods

Sample collection

Trained biologists, veterinarians, and park rangers performed sampling between April–July (waterfowl and resident birds) and October–March (shorebirds and migratory species), from 2006 to 2016. Samples were obtained from a variety of seabirds (gulls, cormorants, penguins, terns, and shorebirds) captured along the Argentine Atlantic coast, and hunter-killed ducks in the Lower Paraná River Valley, as described previously^{12,13}.

Cloacal swabs were collected using single-use sterile polyester swabs and then stored separately in single plastic cryovials, containing 2 ml of phosphate buffer solution (PBS) with 50% glycerol and penicillin 10,000 IU/ml, streptomycin 5 mg/ml, gentamicin Sulfate 1 mg/ml, kanamycin sulfate 700 µg/ml and amphotericin B 10 µg/ml (Sigma Chemical Co, St. Louis, MO, USA). Samples were frozen in liquid nitrogen and transported on dry ice. Once in the laboratory, all samples were stored at –80 °C until processing for molecular diagnosis and virus isolation.

Virus detection

Viral RNA was extracted from 140 µl of suspension from cloacal swabs using a QIAamp Viral RNA Mini Kit (Qiagen Inc., Valencia, CA, USA). RNA was eluted in a final volume of 60 µl and stored at –80 °C until further use. Viral cDNA was prepared using 15 µl of viral RNA and random hexamers in a final volume of 30 µl using a High Capacity cDNA Archive kit (Applied Biosystems, Foster City, CA, USA). The cDNA was tested for IAV by real-time reverse transcriptase PCR (RT-qPCR) using TaqMan Universal PCR Master Mix (Applied Biosystems) directed to the matrix (M) gene. Samples collected from 2006 to 2012 were tested using primers and probe previously described by Spackman et al.⁵⁴, and samples collected since 2013 were tested using primers (InfA Forward and InfA Reverse) and probe (InfA Probe) described in WHO/CDC protocol titled “CDC protocol of real time RTPCR for swine influenza A(H1N1)” and “28 April 2009 revision 1 (30 April 2009).” The PCR reaction was performed on an ABI Prism 7500 SDS (Applied Biosystems). Both methods were validated for sensitivity against avian influenza viruses and no differences were obtained with respect to IAV detection (data not shown).

Virus isolation

Swab samples that tested positive by RT-qPCR were inoculated into 9–11-day old specific-pathogen-free (SPF) embryonated chicken eggs. Briefly, 200 µl of PBS suspension from cloacal swab samples was injected into the allantoic cavity of the eggs, incubated for 72 h, and harvested in accordance with standardized protocols as suggested by the World Health Organization (http://www.wpro.who.int/emerging_diseases/documents/manual_on_animal_ai_diagnosis_and_surveillance/en/) and following Argentine regulations.

Full genome sequencing and assembly by Next Generation Sequencing

PCR was performed using specific primers for each gene segment as described previously⁵⁵. PCR products were purified with a QIAQuick PCR purification kit (Qiagen). Sequencing was performed using the BigDye Terminator v3.1 Cycle Sequencing Kit on an ABI PRISM 3700 DNA Analyzer (Applied Biosystems) following the manufacturer’s instructions. A subset of isolates (marked with “#” in STable 3) was re-sequenced by Next Generation Sequencing and nucleotide sequence identity from Sanger sequencing was verified. Briefly, products from a multisegment-RT-PCR strategy⁵⁶ were cleaned by 0.45X of Agencourt AMPure XP Magnetic Beads (Beckman Coulter) according to manufacturer’s protocol. Eluate concentration was measured using the Qubit High Sensitivity dsDNA kit (Fisher Scientific) in the Qubit 3.0 fluorometer (ThermoFisher) and normalized to 0.2 ng/ul.

Adaptors were added by tagmentation using the Nextera XT DNA library preparation kit (Illumina). Samples were purified using 0.7X of Agencourt AMPure XP Magnetic Beads and fragment size distributions were analyzed on a Bioanalyzer using the High Sensitivity DNA kit (Agilent). Next, samples were normalized to 4 nM and pooled. The final loading concentration of the pooled libraries was 15pM. Libraries were sequenced in a paired-end run using the MiSeq v2, 300 cycle reagent kit (Illumina).

Genome assembly was performed using a customized pipeline developed at the Icahn School of Medicine at Mount Sinai⁵⁶.

Phylogenetic and molecular analysis

The consensus amino acid and nucleotide sequences for all 8 gene segments of the viruses were generated using Megalign (DNASTAR, Madison, WI, USA). Phylogenetic analyses were performed using additional influenza virus sequence data available in the Influenza Research Database (<https://www.fludb.org/brc/home.spg?decorator=influenza>), and were run with Maximum clade credibility in BEAST or Maximum Likelihood (ML) in MEGA (see below).

Time-scaled phylogenetic analyses were performed in the program BEAST v.1.8.2. For the virus internal gene segments, and surface glycoproteins N7 and N8, previously published datasets were downloaded from the Dryad repository¹⁵ with permission from the authors. Additional virus nucleotide sequences (2011–2016) were added to the original alignments, including all sequences from Argentina. Representative strains from South America were selected after inferring Maximum Likelihood phylogenetic trees for all internal gene segments available in IRD. The new alignments for PB2 (total number of sequences = 435), PB1 ($n = 413$), PA ($n = 462$), NP ($n = 508$), M ($n = 519$), NS ($n = 442$), N7 ($n = 213$) and N8 ($n = 463$) were performed with MUSCLE as implemented in the program MEGA 6.0. The MCC trees were inferred using a host-specific local clock model as previously described¹⁵, in which all host clades are permitted individual substitution rates. The parameters for the time-scaled phylogenetic analysis were set in the program Beauti, with manual edits for assignment of host-specific rates as specified in the original xml files¹⁵. We used the SRD06 model of nucleotide substitution⁵⁷, and the coalescent tree prior GMRF Bayesian Skyride. For the H4, H6, and H10, the inferred NJ phylogenies were analyzed in the program Path-O-Gen to identify and remove outlier sequences. The final alignments of H4 ($n = 264$), H6 ($n = 469$) and H10 ($n = 258$) were used for phylogenetic analyses. The MCC trees were then inferred under the SRD06 model of nucleotide substitution, using the uncorrelated lognormal clock and a constant population size. For the mean rate, a prior of 0.001 was used in all

cases. For each data set the MCMC chain was run for 20–100 million steps until convergence of parameters was observed in the program Tracer (v.1.6.0). For each chain, the first 10% states were removed (10% burn-in). Then the results of at least three independent runs were merged in the program LogCombiner. The MCC trees were summarized and annotated in TreeAnnotator, visualized and printed in FigTree (v.1.4.2). The times of divergence for the viruses from South America and Argentina were obtained from the node ages (and their corresponding 95% HPD) on the final MCC trees.

Maximum likelihood (ML) phylogenetic trees for H1 ($n = 172$), H5 ($n = 183$), H7 ($n = 173$), N1 ($n = 339$), N2 ($n = 637$), N3 ($n = 380$) and N6 ($n = 264$) and N9 ($n = 297$) gene segments were constructed by using MEGA 6.0 (<https://www.megasoftware.net/>), including the viruses from Argentina, other IAVs from South America and other background sequences representative of the LPAIV global diversity, available in Influenza Research Database (<https://www.fludb.org/brc/home.spg?decorator=influenza>). Estimates of the phylogenies were calculated by performing 1000 bootstrap replicates under a general reversible model with gamma distribution (GTR+G). The FigTree (v.1.4.2) program was used for visualization and printing of phylogenetic trees.

Nucleotide sequence accession numbers

The nucleotide sequences obtained in this study are available from GenBank under accession numbers MK070525 to MK070532 and MK071345 to MK071456.

Acknowledgements

The National Institute of Allergy and Infectious Diseases (NIAID) Center for Research on Influenza Pathogenesis (CRIP) (contracts HHSN266200700010C and HHSN272201400008C) and the Instituto Nacional de Tecnología Agropecuaria (INTA) (PNSA 1115052 and PNSA 1115056) funded this study. This study was supported in part by resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia's Office of the Vice President for Research and Office of the Vice President for Information Technology. This work was also supported by PICT 2008 – 0922 from MINCYT and CONICET, through a fellowship gave to Agustina Rimondi during part of her work. The content is solely the responsibility of the authors and does not represent official views of the National Institutes of Health.

Author details

¹Instituto de Virología CICVyA – Instituto Nacional de Tecnología Agropecuaria (INTA), CC25 (1712) Castelar, Buenos Aires, Argentina. ²Poultry Diagnostic and Research Center, College of Veterinary Medicine, University of Georgia, 953 College Station Rd, Athens, GA 30602, USA. ³Instituto de Recursos Biológicos CIRN – Instituto Nacional de Tecnología Agropecuaria (INTA), CC25 (1712) Castelar, Buenos Aires, Argentina. ⁴Dirección de Áreas Naturales Protegidas, Organismo Provincial para el Desarrollo Sostenible (O.P.D.S), Gobierno de la provincia de Buenos Aires, General Conesa, Buenos Aires, Argentina. ⁵Centro de Investigaciones en Biodiversidad y Ambiente, Rosario (ECOSUR), Rosario, Santa Fe, Argentina. ⁶Fogarty International Center, National Institutes of Health, Bethesda, MD 20894, USA. ⁷Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ⁸Present address: Instituto de Patobiología CICVyA – Instituto Nacional de Tecnología Agropecuaria (INTA), CC25 (1712) Castelar, Buenos Aires, Argentina.

Author contributions

Author contributions were as follows: Study design: A.R., A.P. and D.R.P. Samples and data collection: A.R., V.S.O., J.D., G.C. and M.R. Virus detection and virus sequencing: A.R., V.S.O., A.S.G., A.P., L.F., G.G. and H.B. Phylogenetic analyses: A.R. (maximum likelihood phylogenetic inference) and A.S.G. (time-scaled Bayesian phylogenetic analyses). Data analysis and interpretation: A.R., A.S.G. and M.N. Writing of the manuscript: A.R., A.S.G. and D.R.P. All authors approved the manuscript before submission.

Conflict of interest

The authors declare that they have no conflict of interest.

Publisher's note

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

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41426-018-0190-2>).

Received: 26 July 2018 Revised: 26 October 2018 Accepted: 31 October 2018

Published online: 28 November 2018

References

- Fouchier, R. A. M. & Munster, V. J. Epidemiology of low pathogenic avian influenza viruses in wild birds. *Rev. Sci. Tech.* **28**, 49–58 (2009).
- Kawaoka, Y., Krauss, S. & Webster, R. G. Avian-to-Human Transmission of the PB1 Gene of Influenza A Viruses in the 1957 and 1968 Pandemics. *J. Virol.* **63**, 4603–4608 (1989).
- WHO/OIE/FAO H5N1 Evolution Working Group. Toward a unified nomenclature system for highly pathogenic avian influenza virus (H5N1). *Emerg. Infect. Dis.* **14**, e1 (2008).
- Lee, D. H. et al. Intercontinental Spread of Asian-Origin H5N8 to North America through Beringia by Migratory Birds. *J. Virol.* **89**, 6521–6524 (2015).
- Vergne, T. et al. Avian influenza vaccination of poultry and passive case reporting, Egypt. *Emerg. Infect. Dis.* **18**, 2076–2078 (2012).
- Waziri, N. E. et al. Evaluating a surveillance system: live-bird market surveillance for highly pathogenic avian influenza, a case study. *Pan Afr. Med. J.* **18**, 11 (2014).
- Hurt, A. C. et al. Detection of evolutionarily distinct avian influenza A viruses in Antarctica. *mBio* **5**, e01098–e01114 (2014).
- Suarez, D. L. et al. Recombination resulting in virulence shift in avian influenza outbreak, Chile. *Emerg. Infect. Dis.* **10**, 693–699 (2004).
- Obenauer, J. C. et al. Large-Scale Sequence Analysis of Avian Influenza Isolates. *Science* **311**, 1576–1580 (2006).
- Bulach, D. et al. Molecular analysis of H7 avian influenza viruses from Australia and New Zealand: genetic diversity and relationships from 1976 to 2007. *J. Virol.* **84**, 9957–9966 (2010).
- Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M. & Kawaoka, Y. Evolution and ecology of influenza A viruses. *Microbiol. Rev.* **56**, 152–179 (1992).
- Pereda, A. J. et al. Avian influenza virus isolated in wild waterfowl in Argentina: evidence of a potentially unique phylogenetic lineage in South America. *Virology* **378**, 363–370 (2008).
- Rimondi, A. et al. Phylogenetic analysis of H6 influenza viruses isolated from rosy-billed pochards (*Netta peposaca*) in Argentina reveals the presence of different HA gene clusters. *J. Virol.* **85**, 13354–13362 (2011).
- Xu, K. et al. Isolation and characterization of an H9N2 influenza virus isolated in Argentina. *Virus Res.* **168**, 41–47 (2012).
- Worobey, M., Han, G. Z. & Rambaut, A. A synchronized global sweep of the internal genes of modern avian influenza virus. *Nature* **508**, 254–257 (2014).
- Spackman, E., McCracken, K. G., Winker, K. & Swayne, D. E. An avian influenza virus from waterfowl in South America contains genes from North American avian and equine lineages. *Avian Dis.* **51**, 273–274 (2007).
- Spackman, E., McCracken, K. G., Winker, K. & Swayne, D. E. H7N3 avian influenza virus found in a South American wild duck is related to the Chilean 2002 poultry outbreak, contains genes from equine and North American wild bird lineages, and is adapted to domestic turkeys. *J. Virol.* **80**, 7760–7764 (2006).
- Hurt, A. C. et al. Evidence for the introduction, reassortment, and persistence of diverse influenza A viruses in Antarctica. *J. Virol.* **90**, 9674–9682 (2016).
- Nelson, M. I. et al. The genetic diversity of influenza A viruses in wild birds in Peru. *PLoS ONE* **11**, e0146059 (2016).
- Jiménez-Bluhm, P. et al. Wild birds in Chile Harbor diverse avian influenza A viruses article. *Emerg. Microbes Infect.* **7**, 4–7 (2018).
- Nobusawa, E. et al. Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. *Virology* **182**, 475–485 (1991).
- Shi, Y., Wu, Y., Zhang, W., Qi, J. & Gao, G. F. Enabling the “host jump”: structural determinants of receptor-binding specificity in influenza A viruses. *Nat. Rev. Microbiol.* **12**, 822–831 (2014).
- Wang, F. et al. Adaptation of avian influenza A (H6N1) virus from avian to human receptor-binding preference. *EMBO J.* **34**, 1661–1673 (2015).
- Tzarum, N. et al. Structure and receptor binding of the hemagglutinin from a human H6N1 influenza virus. *Cell. Host. Microbe* **17**, 369–376 (2015).
- Song, H. et al. Avian-to-human receptor-binding adaptation by influenza A virus hemagglutinin H4. *Cell Rep.* **20**, 1201–1214 (2017).
- Deom, C. M., Caton, A. J. & Schulze, I. T. Host cell-mediated selection of a mutant influenza A virus that has lost a complex oligosaccharide from the tip of the hemagglutinin. *Proc. Natl Acad. Sci. USA* **83**, 3771–3775 (1986).
- Liang, L. et al. Genetics, receptor binding, replication, and mammalian transmission of H4 Avian influenza viruses isolated from live poultry markets in China. *J. Virol.* **90**, 1455–1469 (2016).
- Wang, W. et al. Glycosylation at 158N of the hemagglutinin protein and receptor binding specificity synergistically affect the antigenicity and immunogenicity of a live attenuated H5N1 A/Vietnam/1203/2004 vaccine virus in ferrets. *J. Virol.* **84**, 6570–6577 (2010).
- Wang, G. et al. H6 influenza viruses pose a potential threat to human health. *J. Virol.* **88**, 3953–3964 (2014).
- Baron, J. et al. Matriptase, HAT, and TMPRSS2 activate the hemagglutinin of H9N2 influenza A viruses. *J. Virol.* **87**, 1811–1820 (2013).
- Heui, S. S., Hoffmann, E. & Webster, R. G. Lethal H5N1 influenza viruses escape host anti-viral cytokine responses. *Nat. Med.* **8**, 950–954 (2002).
- Subbarao, E. K., London, W. & Murphy, B. R. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *J. Virol.* **67**, 1761–1764 (1993).
- Hatta, M., Gao, P., Halfmann, P. & Kawaoka, Y. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. *Science* **293**, 1840–1842 (2001).
- Shinya, K. et al. PB2 amino acid at position 627 affects replicative efficiency, but not cell tropism, of Hong Kong H5N1 influenza A viruses in mice. *Virology* **320**, 258–266 (2004).
- McKimm-Breschkin, J. L. Resistance of influenza viruses to neuraminidase inhibitors—a review. *Antivir. Res.* **47**, 1–17 (2000).
- Abed, Y., Goyette, N. & Boivin, G. Generation and characterization of recombinant influenza A (H1N1) viruses harboring amantadine resistance mutations. *Antimicrob. Agents Chemother.* **49**, 556–559 (2005).
- Baz, M., Abed, Y., Simon, P., Hamelin, M. & Boivin, G. Effect of the neuraminidase mutation H274Y conferring resistance to oseltamivir on the replicative capacity and virulence of old and recent human influenza A(H1N1) viruses. *J. Infect. Dis.* **201**, 740–745 (2010).
- Gabriel, G. et al. The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proc. Natl Acad. Sci. USA* **102**, 18590–18595 (2005).
- Krumbholz, A. et al. Current knowledge on PB1-F2 of influenza A viruses. *Med. Microbiol. Immunol.* **200**, 69–75 (2011).
- Košik, I., Holly, J. & Russ, G. PB1-F2 expedition from the whole protein through the domain to aa residue function. *Acta Virol.* **57**, 138–148 (2013).
- Conenello, G. M., Zamarin, D., Perrone, L. A., Tumpey, T. & Palese, P. A single mutation in the PB1-F2 of H5N1 (HK/97) and 1918 influenza A viruses contributes to increased virulence. *PLoS Pathog.* **3**, 1414–1421 (2007).
- Marjuki, H. et al. Three amino acid changes in PB1-F2 of highly pathogenic H5N1 avian influenza virus affect pathogenicity in mallard ducks. *Arch. Virol.* **155**, 925–934 (2010).
- Zamarin, D., Ortigoza, M. B. & Palese, P. Influenza A virus PB1-F2 protein contributes to viral pathogenesis in mice. *J. Virol.* **80**, 7976–7983 (2006).

44. Jagger, B. W. et al. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* **337**, 199–204 (2012).
45. Gao, H. et al. The contribution of PA-X to the virulence of pandemic 2009 H1N1 and highly pathogenic H5N1 avian influenza viruses. *Sci. Rep.* **5**, 8262 (2015).
46. Gonzalez-Reiche, A. S. et al. Evidence of intercontinental spread and uncommon variants of low-pathogenicity avian influenza viruses in ducks overwintering in Guatemala. *mSphere* **2**, e00362–16 (2017).
47. Barriga, G. P. et al. Avian influenza virus H5 strain with north American and Eurasian lineage genes in an Antarctic penguin. *Emerg. Infect. Dis.* **22**, 2221–2223 (2016).
48. Krauss, S. et al. Coincident ruddy turnstone migration and horseshoe crab spawning creates an ecological “hot spot” for influenza viruses. *Proc. Biol. Sci.* **277**, 3373–3379 (2010).
49. De Araujo, J. et al. Avian Influenza Virus (H11N9) in Migratory Shorebirds Wintering in the Amazon Region, Brazil. *PLoS ONE* **9**, e110141 (2014).
50. Kear, J. (eds). Ducks, geese, and swans (Oxford University Press, 2005).
51. Hobson, K. A. Stable isotopes and the determination of avian migratory connectivity and seasonal interactions. *Am. Ornithol. Soc.* **122**, 1037–1048 (2005).
52. Ruegg, K. C. et al. Identifying migrant origins using genetics, isotopes, and habitat suitability. *bioRxiv* <https://doi.org/10.1101/085456> (2013).
53. Bridge, E. S. et al. Bird migration and avian influenza: A comparison of hydrogen stable isotopes and satellite tracking methods. *Ecol. Indic.* **45**, 266–273 (2014).
54. Spackman, E. et al. Development of a real-time reverse transcriptase PCR assay for type A influenza virus and the avian H5 and H7 hemagglutinin subtypes. *J. Clin. Microbiol.* **40**, 3256–3260 (2002).
55. Hoffmann, E., Stech, J., Guan, Y., Webster, R. G. & Perez, D. R. Universal primer set for the full-length amplification of all influenza A viruses. *Arch. Virol.* **146**, 2275–2289 (2001).
56. Mena, I. et al. Origins of the 2009 H1N1 influenza pandemic in swine in Mexico. *eLife* **5**, e16777 (2016). pii.
57. Shapiro, B., Rambaut, A. & Drummond, A. J. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* **23**, 7–9 (2005).