Article

Analysis of the Variability in the Non-Coding Regions of Influenza A Viruses

Jessica Benkaroun [1] [ID], Gregory J. Robertson [2], Hugh Whitney [1] and Andrew S. Lang [1,*] [ID]

[1] Department of Biology, Memorial University of Newfoundland, 232 Elizabeth Ave., St. John's, NL A1B 3X9, Canada; jb5163@mun.ca (J.B.); hughwhitneynl@gmail.com (H.W.)

[2] Wildlife Research Division, Environment and Climate Change Canada, 6 Bruce St., Mount Pearl, NL A1N 4T3, Canada; greg.robertson@canada.ca

[*] Correspondence: aslang@mun.ca; Tel.: +1-709-864-7517

check for updates

Abstract: The genomes of influenza A viruses (IAVs) comprise eight negative-sense single-stranded RNA segments. In addition to the protein-coding region, each segment possesses 5′ and 3′ non-coding regions (NCR) that are important for transcription, replication and packaging. The NCRs contain both conserved and segment-specific sequences, and the impacts of variability in the NCRs are not completely understood. Full NCRs have been determined from some viruses, but a detailed analysis of potential variability in these regions among viruses from different host groups and locations has not been performed. To evaluate the degree of conservation in NCRs among different viruses, we sequenced the NCRs of IAVs isolated from different wild bird host groups (ducks, gulls and seabirds). We then extended our study to include NCRs available from the National Center for Biotechnology Information (NCBI) Influenza Virus Database, which allowed us to analyze a wider variety of host species and more HA and NA subtypes. We found that the amount of variability within the NCRs varies among segments, with the greatest variation found in the HA and NA and the least in the M and NS segments. Overall, variability in NCR sequences was correlated with the coding region phylogeny, suggesting vertical coevolution of the (coding sequence) CDS and NCR regions.

Keywords: viral evolution; phylogeny; avian influenza

1. Introduction

Influenza A viruses (IAVs) are most well known for their circulation in humans, causing yearly seasonal epidemics and occasional pandemics, and for outbreaks in poultry. However, the viral strains responsible for these infections represent only a small portion of the total known IAV diversity, most of which is maintained in the natural wild bird reservoir.

IAVs are members of the family Orthomyxoviridae and have segmented negative-sense single-stranded RNA genomes. The viral genome is composed of eight segments that encode at least 11 proteins, depending on the strain. Each segment has non-coding regions (NCRs) at both ends. The 3′ NCR contains 12 conserved nucleotides followed by a segment-specific region of variable length. The 5′ NCR contains 13 conserved nucleotides and a segment-specific region of variable length. The two conserved regions are partially complementary [1–3] and form a panhandle-like structure [4–6] involved in the transcription and replication steps of the viral life cycle [7–9]. Mutagenesis studies identified residues within the NCRs that are required for the viral polymerase complex to bind and initiate transcription [10,11]. However, a binding preference for the 5′ NCR was demonstrated [12]. There is also a stretch of uridines at the 5′ end that is involved in the polyadenylation of viral mRNAs [13–16]. Further studies have revealed the presence of some variability within the two NCRs at different positions. There is variability in the fourth nucleotide (C or U) in the 3′ end of

all RNA segments [1], which affects the promoter and could play a role in differential regulation of transcription and replication [17]. Non-conserved nucleotides located in the segment-specific regions have also been shown to play an important role in viral replication [18]. Mutagenesis studies have revealed the importance of NCRs for the incorporation of the eight different gene segments during viral assembly [19–22]. These packaging signals can influence the reassortment of genes during co-infections of different strains by preventing their incorporation during the virus genome packaging [20,23].

Amplification of complete IAV genomes with universal primers was initially designed based on the conserved nature of NCRs [24], but amplification biases have been reported [25,26]. To date, a thorough analysis of IAV NCRs from different hosts and across segments has not been performed. We hypothesized that the NCRs within both the conserved and the segment-specific regions may vary according to the original host species. To test our hypothesis, we characterized the true NCRs of IAV genomes from different wild bird species by the rapid amplification of cDNA ends (RACE) method and then expanded our analyses to include additional viral subtypes, hosts and geographic origins.

2. Materials and Methods

2.1. NCR Sequence Determination

NCR sequences from previously identified and sequenced viruses originating from three different wild bird host groups (ducks, gulls and seabirds) were determined: A/domestic duck/Newfoundland/MW668/2010(H1N1), A/American black duck/Newfoundland/GR679/2011(H3N2), A/herring gull/Newfoundland/GR578/2011(H13N6), A/herring gull/Newfoundland/GR848/2011(H13N6), A/common murre/Newfoundland/AB318/2011(H1N2), and A/common murre/Newfoundland/AB324/2011(H1N2). From these six viruses, we determined 92 NCR sequences, comprising 46 3′ and 46 5′ NCRs.

Viral genomic RNA was extracted from allantoic fluids of specific pathogen free eggs inoculated with the different viruses using the Trizol reagent (Thermo Fisher Scientific, Waltham, MA, USA). Primers were designed to efficiently amplify the 3′ and 5′ NCRs of the targeted viruses based on the full genomic sequences of viruses that were available from previous studies (Supplementary Table S1).
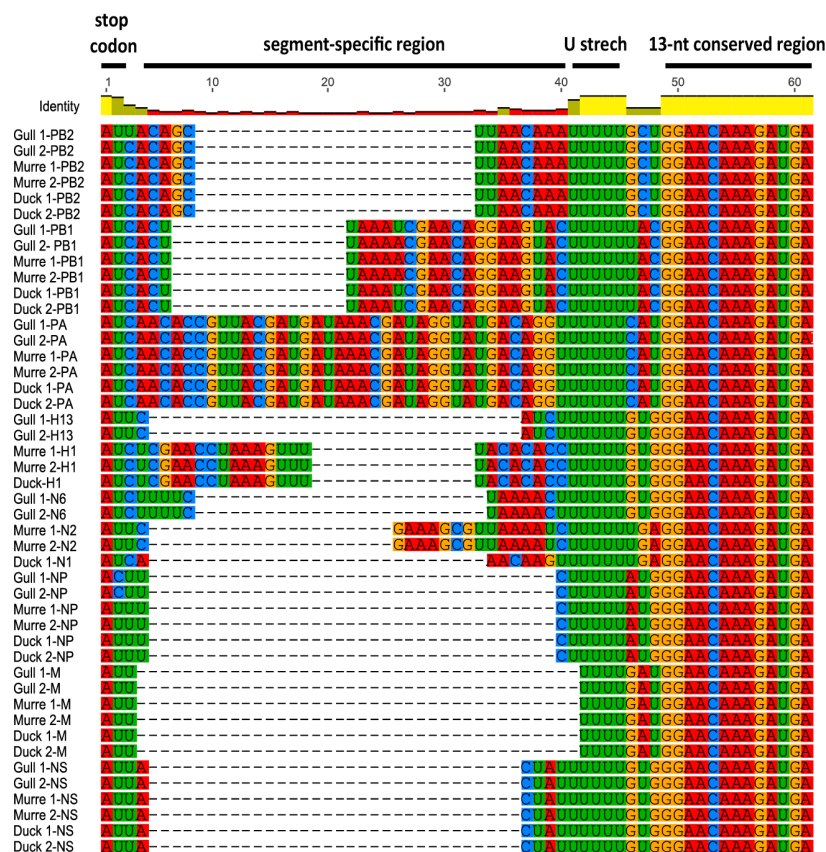
For the 3′ NCRs, genomic RNAs were polyadenylated with poly(A) polymerase (2 U μL^{-1}) (Thermo Fisher Scientific) in the presence of $MnCl_2$ (2.5 mM) and ATP (1 mM) in 1X reaction buffer and a final volume of 50 μL. Reactions were incubated at 37 °C for 30 min and the resulting RNA used as the input for the 3′ RACE System for Rapid Amplification of cDNA Ends (Thermo Fisher Scientific). For the 5′ NCRs, the 5′ RACE System for Rapid Amplification of cDNA Ends (Thermo Fisher Scientific) was used according to the manufacturer's recommendations. The resulting amplicons were purified and sequenced using Sanger chemistry at The Center for Applied Genomics (Toronto, ON, Canada).

2.2. Sequence Analyses

The 92 NCRs determined from our samples were compared with NCRs retrieved from the influenza resource database (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html), chosen to represent different viral subtypes, host species, geographic origins, and collection dates. Where possible, all available or a high number of nucleotide sequences of complete gene segments were downloaded from the database. The number of NCRs analyzed are indicated in Figure 3. Sequence alignments were performed using MUSCLE, implemented in MEGA version 6 [27]. Multiple alignments were then edited in Geneious version 8 [28]. The CDS regions of segments that had complete NCRs were removed from the sequence, with only the NCRs kept for further analysis. Complete segments without NCRs were removed from the analysis.

Pairwise genetic distance matrices determined with the Maximum Composite Likelihood method, generated by MEGA, were used for classical multi-dimensional analysis (MDS) in R [29]. The viral segments were assigned geographic origins based on nucleotide BLAST searches [30] and neighbor-joining phylogenetic analyses in MEGA. A small number of sequences was added in each

multiple alignment figures due to space limitation, but the exact number of NCRs originally aligned with MUSCLE are indicated in Figure 3.

3. Results

3.1. Characterization of NCRs from Different Wild Bird Viruses

Comparison of the 3′ and 5′ NCRs we determined by the RACE method showed that the 12- and 13-nt conserved regions, respectively, contained no novel substitutions in any of the different wild bird viruses (Figures 1 and 2), which agrees with previous studies [1–3]. Comparison of the 3′ NCRs (Figure 1) showed that the 4th position in the 12-nucleotide conserved region displays the expected variability, with the polymerase gene segments, PB2, PB1 and PA, containing a cytosine residue while the other segments contain a uridine residue, as previously observed [1,17].

Figure 1. 3′ NCRs of wild bird virus segments determined in this study. The multiple alignment of 3′ NCRs was generated with MUSCLE, implemented in Geneious version 8. The RNA sequences are displayed in the 3′ to 5′ orientation corresponding to the packaged RNAs.

Figure 2. 5′ NCRs of wild bird virus segments determined in this study. The multiple alignment of 5′ NCRs was generated with MUSCLE implemented in Geneious version 8. The RNA sequences are displayed in the 3′ to 5′ orientation corresponding to the packaged RNAs.

For the 3′ segment-specific regions (Figure 1), some of the segments (PB1, M and NS) had identical sequences across all six viruses, but the others showed some differences. The murre and gull PB2 and NP segments were identical whereas the gull viruses differed from the others for these segments. The PA segment was identical in the gull and duck viruses but differed from those in the murre viruses. The HA sequences were identical within subtypes, even for the H1 viruses that came from murres and a duck. Similarly, the NA sequences were identical within the same subtype.

Comparison of the 5′ NCRs showed the conservation of a stretch of five or six uridine residues upstream of the 13-nt conserved region among the different wild bird viruses (Figure 2). Similar to the 3′ NCRs, the segment-specific regions varied among segments, with some of the individual segments' NCRs conserved among the six viruses and others varying among the viruses (Figure 2). All PB2, PA, M, and NS segments' NCRs were identical among the viruses. There was a single variation in the PB1 segment, with the murre and one gull virus differing from the other gull virus and the duck viruses. As observed for the 3′ NCRs, the HA and NA segments were the same within subtypes, and the H1 and H2 segments were also identical.

3.2. The Variability in Segment-Specific NCR Sequences Differs among Segments

To further investigate the differences observed in the NCRs of our limited set of six viruses, we expanded our analysis of NCR genetic diversity to include viral sequences from a variety of host species and geographic origins retrieved from the National Center for Biotechnology Information (NCBI) Influenza Virus Database (Supplementary Tables S2 and S3). There is no apparent relationship between the number of sequences analyzed and the NCR sequence diversity (Figure 3). The greatest amount of variability in NCRs was observed in the HA segments (Figure 3A), but there is variability

among the segments and between 3′ and 5′ NCRs for some subtypes. The H4, H6, H7, H8, H10, H12, and H14 NCRs showed very little variability overall. The H5, H9 and H13 subtypes show greater variability within their 3′NCRs relative to their 5′ NCRs, whereas the H1, H2, H3, H11 and H16 were equally variable at the two ends.
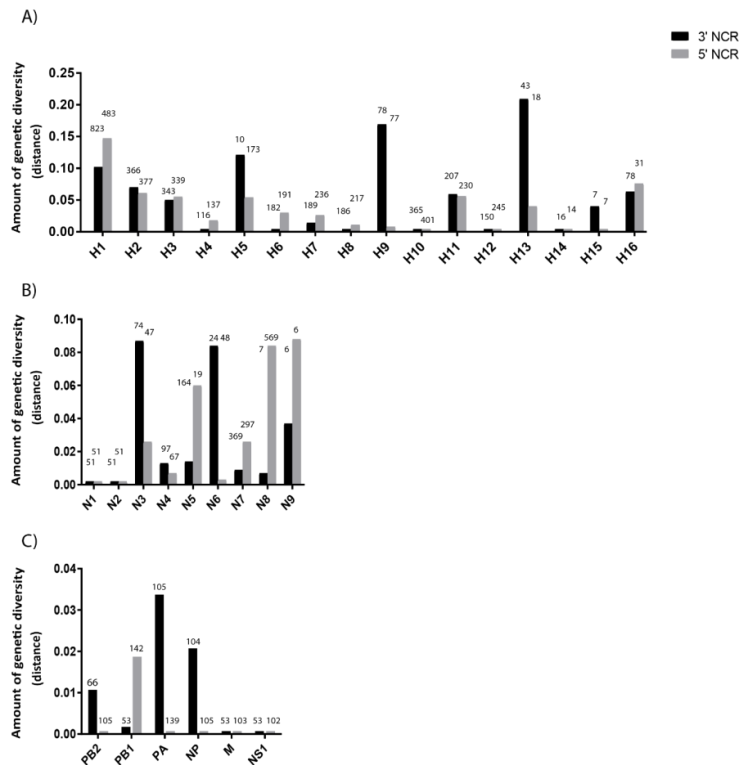
Figure 3. Genetic diversity of 3′ and 5′ NCRs for different segments. (**A**) HA segments of the different subtypes. (**B**) NA segments of the different subtypes. (**C**) "Internal" protein-coding segments. The genetic diversity for each of the NCR regions for the different segments was calculated as mean genetic distance for the NCRs included in our analyses. The number above each bar represents the number of sequences used to calculate the amount of genetic diversity for each segment.

The genetic diversity of different NA subtypes' NCRs (Figure 3B), which originated from a large number of host species and geographic locations, showed that the N3 and N6 subtypes have more diversity at their 3′ ends while the N5, N8, and N9 are more diverse at the 5′ ends. The N1 and N2 viruses showed very little diversity, despite the large assortment of hosts from which viruses were included.

Matching the general pattern observed for IAV segment coding regions, the NCRs of the remaining segments were overall less variable than found for HA and NA (Figure 3C). The M and NS segments showed the lowest diversity, in agreement with the slower evolution rates for these segments' CDS (e.g., References [31,32]). These differences have been observed across many viruses which is not surprising given that a previous analysis of NCRs, from a limited dataset of human H3N2 viruses circulating worldwide since 1968, found that NCR variability differed among the segments [33].

3.3. Patterns of Variability within NCRs Can Be Explained by Viral Host Species and Geographic Origins

To better understand the origin of the segment-specific NCR diversity, we assessed whether the NCRs of viruses from different hosts and geographic origins showed similar patterns of relationships of the respective coding regions (CDS). We used a multi-dimensional scaling (MDS) analysis based on pairwise genetic distance matrices for the NCR sequences, which allowed us to visualize the

relationships among NCRs relative to the viral hosts and geographic origins. The results of these analyses were then compared to MDS analyses performed on the respective coding regions. We focused this analysis on segments that showed greater NCR diversity, as described above.
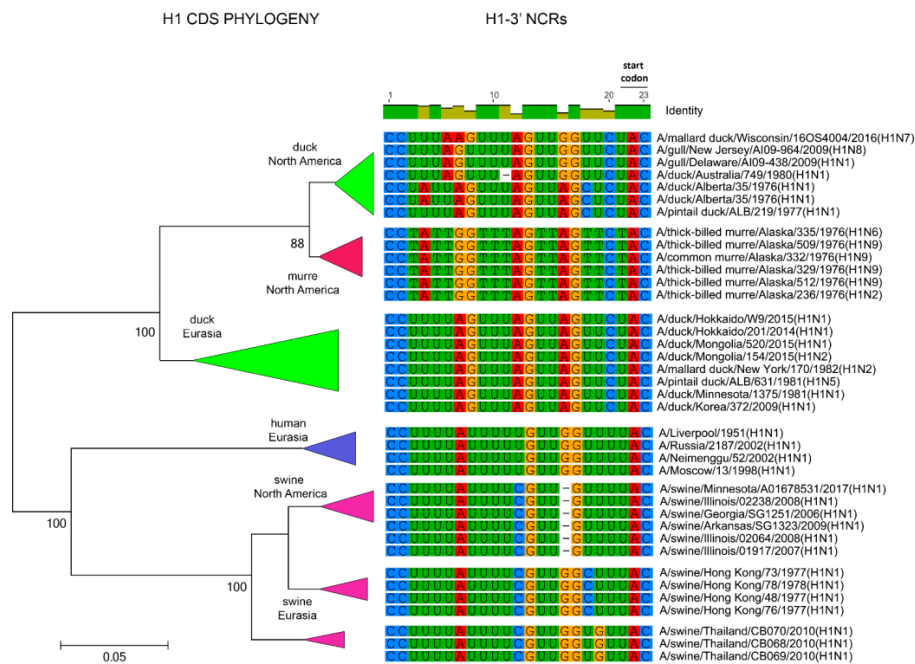
The relationships among the 3′ NCRs and CDS regions of a set of H9 segments were compared (Figure 4). The 3′ NCRs and CDS regions showed similar grouping in the MDS analyses (Figure 4A), corresponding to their CDS phylogenetic clades (Figure 4B). The two North American clade sequences grouped more closely and were more distant from the Eurasian and North African sequences (Figure 4A). This matches the patterns for the CDS phylogeny and 3′ NCR alignments (Figure 4B). The same patterns were observed with the H13 subtype 3′ NCRs in the MDS analyses (Figure 5A), and there was again a clear correspondence between the CDS phylogeny and the NCR sequence relationships (Figure 5B). For the H1 subtype, the CDS phylogeny contained larger host-specific (swine, human and avian) clades that separated into geographic-specific (Eurasian and North American) clades (Figure 6). The correspondence between CDS phylogeny and 3′ NCR similarities was again found for these H1 sequences (Figure 6). A similar pattern was also found for the N6 NA (Figure 7) and NP sequences (Figure 8), where the CDS phylogenies also determined the 3′ NCR relationships. This was true even when the viruses falling within specific clades originated from different hosts (e.g., gulls and ducks) or different continental origins (Eurasia and North America). Overall, these patterns suggest vertical coevolution of the CDS and NCRs.
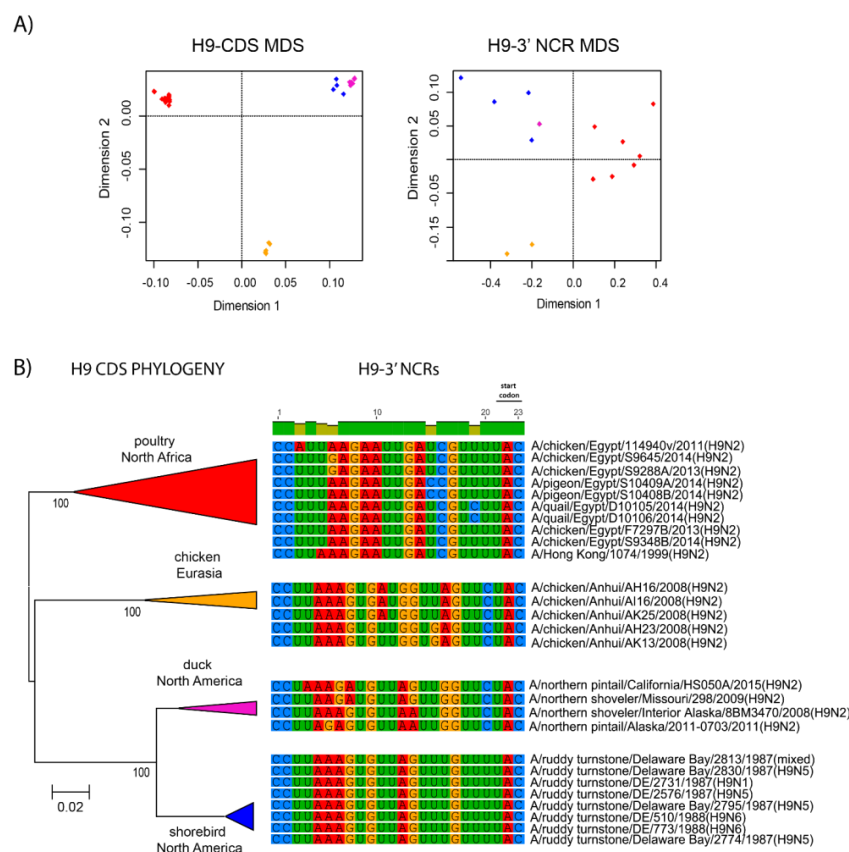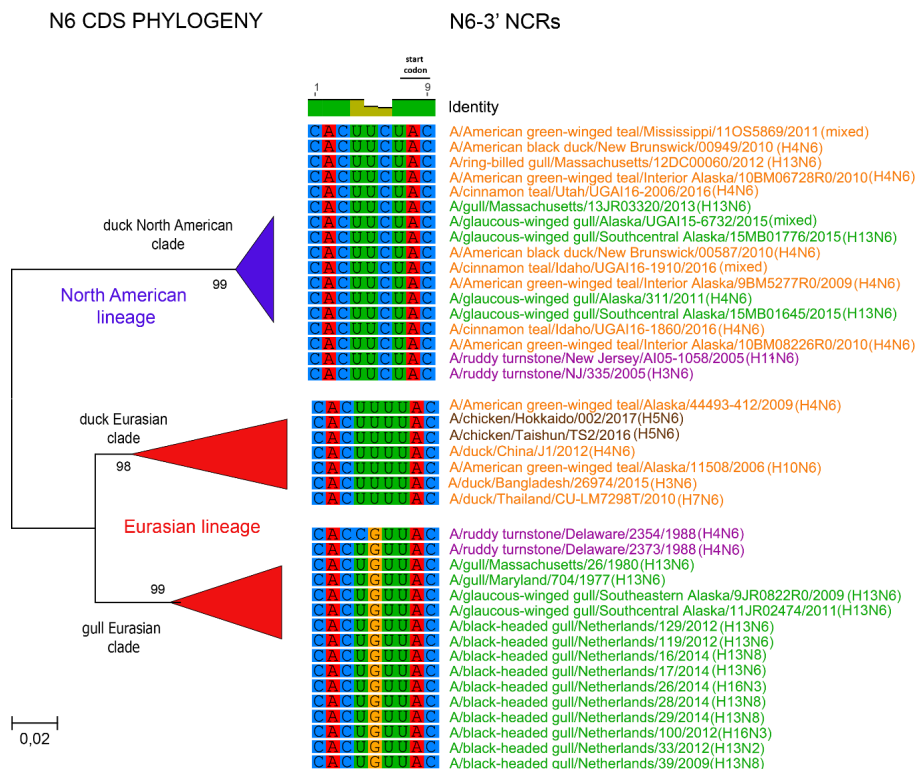

Figure 4. Relationships among H9 3′ NCR and CDS regions for viruses from different host species and geographic locations. The multi-dimensional analysis (MDS) scatter plots (**A**) are based on the pairwise genetic distance matrices from the multiple alignments of complete CDS regions and 3′NCRs. Each dot represents a virus, colored according to their origins as indicated on the CDS-based phylogenetic tree (**B**). Alignments of the 3′ NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses.
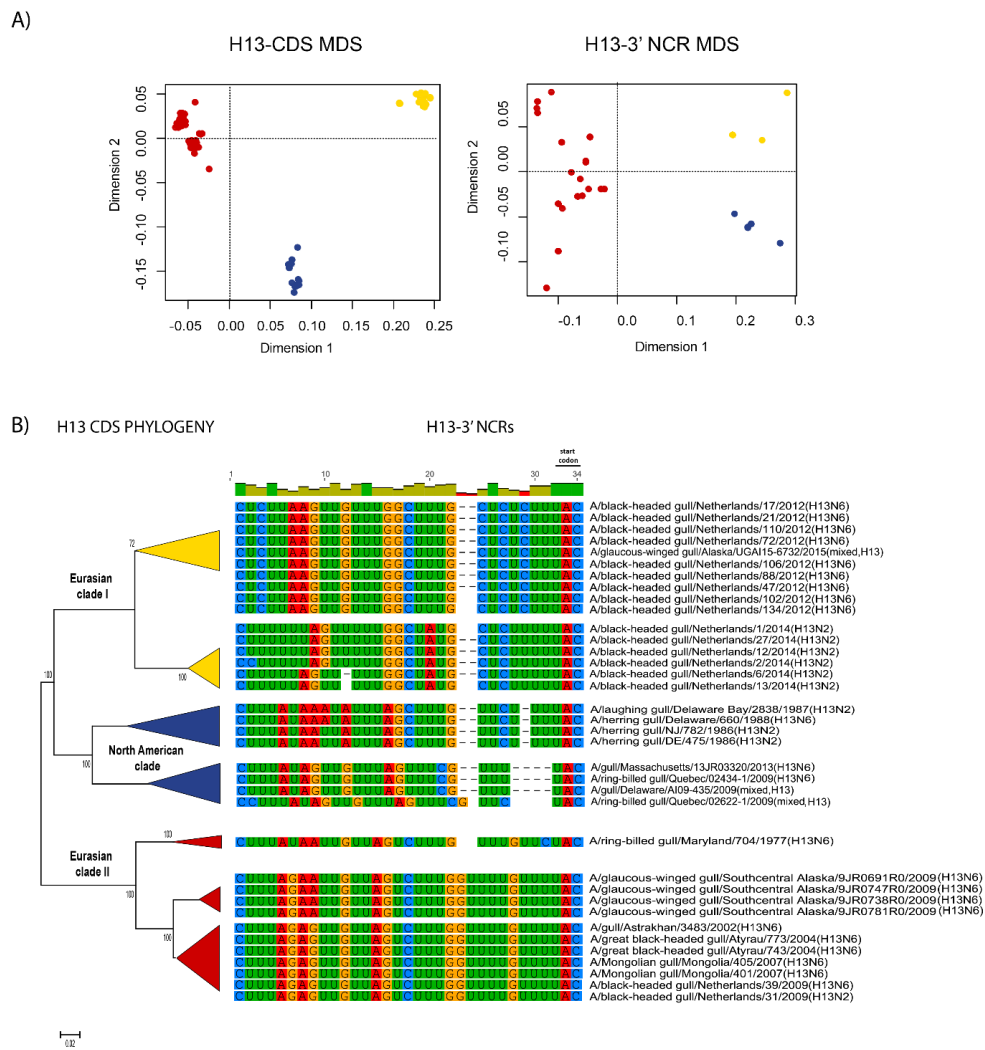
Figure 5. Relationships among H13 3′ NCR and CDS regions for viruses from different host species and geographic locations. The MDS scatter plots (**A**) are based on the pairwise genetic distance matrices from the multiple alignments of complete CDS regions and 3′ NCRs. Each dot represents a virus, colored according to their origins as indicated on the CDS-based phylogenetic tree (**B**). Alignments of the 3′ NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses.
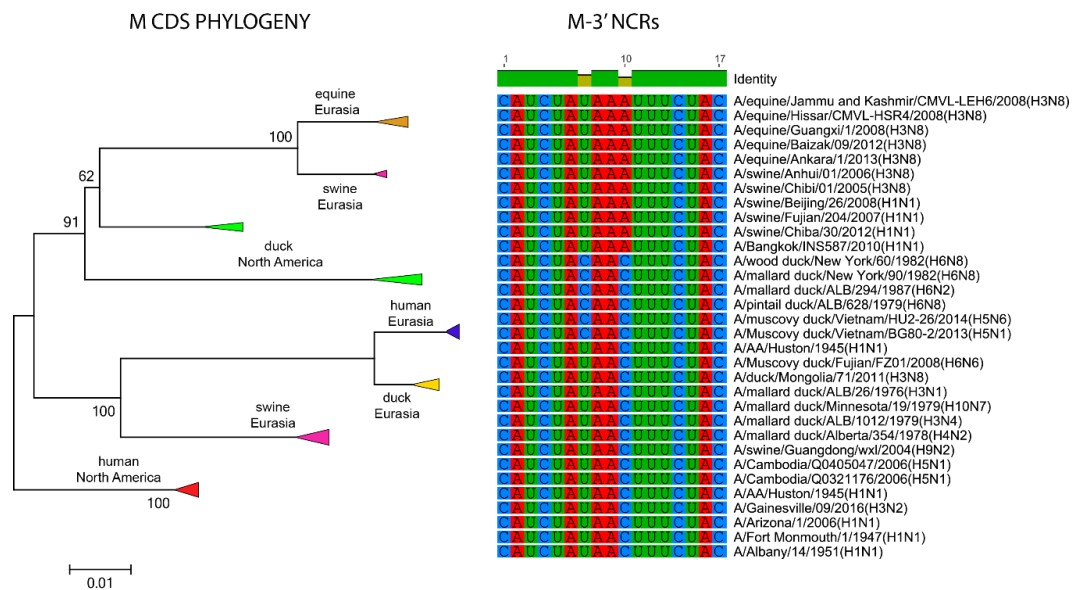
Figure 6. Relationships among H1 3′ NCR and CDS regions for viruses from different host species and geographic locations. Alignments of the 3′ NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses.

Figure 7. Relationships among N6 3′ NCR and CDS regions for viruses from different host species and geographic locations. Alignments of the 3′ NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses. To facilitate the distinction of viruses from different hosts, gull viruses are in green, duck viruses are in orange, shorebird viruses are in purple, and chicken viruses are in brown.

Figure 8. Relationships among NP 3′ NCR and CDS regions for viruses from different host species and geographic locations. Alignments of the 3′ NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses.

Lastly, we looked at the M and NS segments which overall, showed very low NCR diversity (Figure 3C). These two segments share very similar 3′ NCRs amongst different viruses, regardless of the host species or geographic origins (Figure 9, Supplementary Figure S1). Some variations were found for the M segment, for a total of three different M 3′ NCR sequences (Figure 9), but they did not show the same correspondence to the phylogeny of their CDS as observed for the other segments. Comparison of a subset of NS NCRs from various hosts and geographic locations found these all to be identical (Supplementary Figure S1). The conservation of the M and NS NCRs, as mentioned earlier, may be linked to the slower evolutionary rate of the coding regions of these segments.

Figure 9. Relationships among M 3′ NCR and CDS regions for viruses from different host species and geographic locations. Alignments of the 3′ NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses.

4. Conclusions

We first hypothesized that NCRs may vary according to the host species. Our initial evaluation of the 3′ and 5′ NCRs from viruses isolated from different wild bird hosts showed that the conserved regions are identical among different hosts, but the segment-specific regions showed some variability among the viruses for some segments. This prompted a larger analysis including more sequences, which showed that the amount of variability in the NCRs varies among the different segments and can also vary between the 3′ and 5′ ends, depending on the segment. The overall variability is highest for the HA and NA segments. Examination of 3′ NCRs from HA and NA subtypes with higher variability indicates that the relationships among NCR sequences matches the coding region phylogenies, which generally follow patterns based on the viral host species and geographic origins, and possibly reflects a lack of recombination in these segments. The most conserved NCRs are found on the M and NS segments, which also show the slowest coding sequence evolutionary rates. The exact evolutionary pressures acting on the NCRs and the consequences of sequence changes in these regions, remain to be fully defined.

References

1. Robertson, J.S. 5′ and 3′ terminal nucleotide sequences of the RNA genome segments of influenza virus. *Nucleic Acids Res.* **1979**, *6*, 3745–3757. [CrossRef] [PubMed]
2. Skehel, J.J.; Hay, A.J. Nucleotide sequences at the 5′ termini of influenza virus RNAs and their transcripts. *Nucleic Acids Res.* **1978**, *5*, 1207–1219. [CrossRef] [PubMed]
3. Desselberger, U.; Racaniello, V.R.; Zazra, J.J.; Palese, P. The 3′ and 5′-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. *Gene* **1980**, *8*, 315–328. [CrossRef]
4. Hsu, M.T.; Parvin, J.D.; Gupta, S.; Krystal, M.; Palese, P. Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8140–8144. [CrossRef] [PubMed]
5. Bergmann, M.; Muster, T. Mutations in the nonconserved noncoding sequences of the influenza A virus segments affect viral vRNA formation. *Virus Res.* **1996**, *44*, 23–31. [CrossRef]
6. Kim, H.J.; Fodor, E.; Brownlee, G.G.; Seong, B.L. Mutational analysis of the RNA-fork model of the influenza A virus vRNA promoter in vivo. *J. Gen. Virol.* **1997**, *78*, 353–357. [CrossRef] [PubMed]
7. Seong, B.L.; Brownlee, G.G. Nucleotides 9 to 11 of the influenza A virion RNA promoter are crucial for activity in vitro. *J. Gen. Virol.* **1992**, *73*, 3115–3124. [CrossRef] [PubMed]
8. Fodor, E.; Pritlove, D.C.; Brownlee, G.G. The influenza virus panhandle is involved in the initiation of transcription. *J. Virol.* **1994**, *68*, 4092–4096. [PubMed]
9. Flick, R.; Neumann, G.; Hoffmann, E.; Neumeier, E.; Hobom, G. Promoter elements in the influenza vRNA terminal structure. *RNA* **1996**, *2*, 1046–1057. [PubMed]
10. Fodor, E.; Seong, B.L.; Brownlee, G.G. Photochemical cross-linking of influenza A polymerase to its virion RNA promoter defines a polymerase binding site at residues 9 to 12 of the promoter. *J. Gen. Virol.* **1993**, *74*, 1327–1333. [CrossRef] [PubMed]
11. Li, X.; Palese, P. Mutational analysis of the promoter required for influenza virus virion RNA synthesis. *J. Virol.* **1992**, *66*, 4331–4338. [PubMed]
12. Pritlove, D.C.; Fodor, E.; Seong, B.L.; Brownlee, G.G. In vitro transcription and polymerase binding studies of the termini of influenza A virus cRNA: Evidence for a cRNA panhandle. *J. Gen. Virol.* **1995**, *76*, 2205–2213. [CrossRef] [PubMed]
13. Luo, G.X.; Luytjes, W.; Enami, M.; Palese, P. The polyadenylation signal of influenza virus RNA involves a stretch of uridines followed by the RNA duplex of the panhandle structure. *J. Virol.* **1991**, *65*, 2861–2867. [PubMed]
14. Li, X.; Palese, P. Characterization of the polyadenylation signal of influenza virus RNA. *J. Virol.* **1994**, *68*, 1245–1249. [PubMed]
15. Poon, L.L.; Pritlove, D.C.; Fodor, E.; Brownlee, G.G. Direct evidence that the poly(A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template. *J. Virol.* **1999**, *73*, 3473–3476. [PubMed]
16. Pritlove, D.C.; Poon, L.L.; Fodor, E.; Sharps, J.; Brownlee, G.G. Polyadenylation of influenza virus mRNA transcribed in vitro from model virion RNA templates: Requirement for 5′ conserved sequences. *J. Virol.* **1998**, *72*, 1280–1286. [PubMed]
17. Lee, M.K.; Bae, S.H.; Park, C.J.; Cheong, H.K.; Cheong, C.; Choi, B.S. A single-nucleotide natural variation (U4 to C4) in an influenza A virus promoter exhibits a large structural change: Implications for differential viral RNA synthesis by RNA-dependent RNA polymerase. *Nucleic Acids Res.* **2003**, *31*, 1216–1223. [CrossRef] [PubMed]
18. Zheng, H.; Palese, P.; Garcia-Sastre, A. Nonconserved nucleotides at the 3′ and 5′ ends of an influenza A virus RNA play an important role in viral RNA replication. *Virology* **1996**, *217*, 242–251. [CrossRef] [PubMed]
19. Crescenzo-Chaigne, B.; Barbezange, C.V.; Léandri, S.; Roquin, C.; Berthault, C.; Van Der Werf, S. Incorporation of the influenza A virus NA segment into virions does not require cognate non-coding sequences. *Sci. Rep.* **2017**, *7*, 43462. [CrossRef] [PubMed]
20. Gao, Q.; Palese, P. Rewiring the RNAs of influenza virus to prevent reassortment. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15891. [CrossRef] [PubMed]

21. Fujii, K.; Fujii, Y.; Noda, T.; Muramoto, Y.; Watanabe, T.; Takada, A.; Goto, H.; Horimoto, T.; Kawaoka, Y. Importance of both the coding and the segment-specific noncoding regions of the influenza A virus NS segment for its efficient incorporation into virions. *J. Virol.* **2005**, *79*, 3766–3774. [CrossRef] [PubMed]

22. Zhao, L.; Peng, Y.; Zhou, K.; Cao, M.; Wang, J.; Wang, X.; Jiang, T.; Deng, T. New Insights into the Nonconserved Noncoding Region of the Subtype-Determinant Hemagglutinin and Neuraminidase Segments of Influenza A Viruses. *J. Virol.* **2014**, *88*, 11493–11503. [CrossRef] [PubMed]

23. White, M.C.; Steel, J.; Lowen, A.C. Heterologous Packaging Signals on Segment 4, but Not Segment 6 or Segment 8, Limit Influenza A Virus Reassortment. *J. Virol.* **2017**, *91*. [CrossRef] [PubMed]

24. 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.* **2001**, *146*, 2275–2289. [CrossRef] [PubMed]

25. Widjaja, I.; de Vries, E.; Rottier, P.J.; de Haan, C.A. Competition between Influenza A Virus Genome Segments. *PLoS ONE* **2012**, *7*, e47529. [CrossRef] [PubMed]

26. Zhou, B.; Donnelly, M.E.; Scholes, D.T.; George, K.S.; Hatta, M.; Kawaoka, Y.; Wentworth, D.E. Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza a viruses. *J. Virol.* **2009**, *83*, 10309–10313. [CrossRef] [PubMed]

27. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef] [PubMed]

28. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef] [PubMed]

29. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017.

30. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]

31. Qu, Y.; Zhang, R.; Cui, P.; Song, G.; Duan, Z.; Lei, F. Evolutionary genomics of the pandemic 2009 H1N1 influenza viruses (pH1N 1v). *Virol. J.* **2011**, *8*, 250. [CrossRef] [PubMed]

32. Bahl, J.; Vijaykrishna, D.; Holmes, E.C.; Smith, G.J.; Guan, Y. Gene flow and competitive exclusion of avian influenza A virus in natural reservoir hosts. *Virology* **2009**, *390*, 289–297. [CrossRef] [PubMed]

33. Furuse, Y.; Oshitani, H. Evolution of the influenza A virus untranslated regions. *Infect. Genet. Evol.* **2011**, *11*, 1150–1154. [CrossRef] [PubMed]