

# Host Restrictions of Avian Influenza Viruses: *In Silico* Analysis of H13 and H16 Specific Signatures in the Internal Proteins

Ragnhild Tønnessen<sup>1\*</sup>, Anna G. Hauge<sup>2</sup>, Elisabeth F. Hansen<sup>1</sup>, Espen Rimstad<sup>1</sup>, Christine M. Jonassen<sup>2</sup>

<sup>1</sup> Department of Food Safety and Infection Biology, Norwegian School of Veterinary Science, Oslo, Norway, <sup>2</sup> Department of Laboratory Services, Norwegian Veterinary Institute, Oslo, Norway

## Abstract

Gulls are the primary hosts of H13 and H16 avian influenza viruses (AIVs). The molecular basis for this host restriction is only partially understood. In this study, amino acid sequences from Eurasian gull H13 and H16 AIVs and Eurasian AIVs (non H13 and H16) were compared to determine if specific signatures are present only in the internal proteins of H13 and H16 AIVs, using a bioinformatics approach. Amino acids identified in an initial analysis performed on 15 selected sequences were checked against a comprehensive set of AIV sequences retrieved from Genbank to verify them as H13 and H16 specific signatures. Analysis of protein similarities and prediction of subcellular localization signals were performed to search for possible functions associated with the confirmed signatures. H13 and H16 AIV specific signatures were found in all the internal proteins examined, but most were found in the non-structural protein 1 (NS1) and in the nucleoprotein. A putative functional signature was predicted to be present in the nuclear export protein. Moreover, it was predicted that the NS1 of H13 and H16 AIVs lack one of the nuclear localization signals present in NS1 of other AIV subtypes. These findings suggest that the signatures found in the internal proteins of H13 and H16 viruses are possibly related to host restriction.

**Citation:** Tønnessen R, Hauge AG, Hansen EF, Rimstad E, Jonassen CM (2013) Host Restrictions of Avian Influenza Viruses: *In Silico* Analysis of H13 and H16 Specific Signatures in the Internal Proteins. PLoS ONE 8(4): e63270. doi:10.1371/journal.pone.0063270

**Editor:** Matthias Johannes Schnell, Thomas Jefferson University, United States of America

**Received:** November 24, 2012; **Accepted:** April 2, 2013; **Published:** April 30, 2013

**Copyright:** © 2013 Tønnessen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the Norwegian Research Council (www.forskningradet.no) (project number 178275), Animalia Meat and Poultry Research Centre (www.animalia.no) and the Norwegian Food Safety Authority (www.mattilsynet.no). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: \*ragnhild.tonnessen@vetinst.no

‡ Current address: Department of Laboratory Services, Norwegian Veterinary Institute, Oslo, Norway

## Introduction

Influenza A viruses (family *Orthomyxoviridae*) have a genome of eight negative sense single-stranded RNA segments that encodes up to 13 proteins [1]. To date, 16 hemagglutinin (HA) (H1-H16) and 9 neuraminidase (NA) (N1-N9) subtypes have been detected in wild birds [2,3]. Ducks, gulls and shorebirds constitute the primary natural reservoir of avian influenza viruses (AIVs) [2,3]. Influenza A viruses that cause disease in humans and animals have evolved from AIVs [2].

Limited overlap between the flyways used by migratory birds in Eurasia and America has led to the development of two major phylogenetic AIV lineages [4]. AIVs of the H13 and H16 subtypes are mainly detected in gulls, and for these viruses there are separate American and Eurasian phylogenetic lineages [4,5]. Genetic reassortment frequently occurs for AIVs within continents. More rarely, birds move single or multiple viral segments across continents [5–13]. Nevertheless, the number of intercontinental reassortment events found among H13 and H16 AIVs is relatively high [5,14], and long-distance migrating gull species are believed to play an important role as vectors for geographically reassorted viruses [5]. Eurasian gene segments are often detected in influenza viruses from gulls in America [5,12]. On the contrary, American gene segments have only occasionally been found in

AIVs from gulls in Eurasia [7,14]. However, the total number of sequenced AIV genomes from gulls in Eurasia is still low [5].

Gulls may get infected by most AIV subtypes found in waterfowl [15]. In contrast, H13 and H16 viruses are extremely rarely detected in surveillance samples from waterfowl [16], and experimentally challenged ducks are resistant to several strains of H13 AIVs [17,18]. The molecular basis for this host restriction is only partially understood.

HA is an important determinant of host specificity. The HA glycoprotein must interact with a cellular receptor for the virus to get endocytosed. AIVs preferentially bind to Neu5Ac  $\alpha$ 2-3Gal-terminated receptors [19]. Influenza viruses from ducks and gulls differ in their receptor specificity [20], also in their ability to recognize the more distal parts of the receptor [19]. Several amino acid residues in the receptor-binding site of HA are shared by H13 and H16 AIVs, but not by influenza viruses of other subtypes [3,20].

Host range is a polygenic trait [2] and previous studies have suggested that several amino acids also in the internal viral proteins of H13 AIVs are associated with host adaptation [2,21–25]. The majority of these studies were performed in the 1990s and the early 2000s on a limited number of sequences from a restricted geographical area. But in a recent study, the occurrence of specific signatures in the internal viral proteins of

both H13 and H16 viruses was examined [26]. However, the analysis was performed on American H13 and H16 strains, and only one H16 virus was included. The amount of available AIV sequences has increased vastly the last decade [5]. Therefore, to gain more insight about host restriction of both H13 and H16 AIVs, and to provide a basis for functional studies, we reexamined these AIV subtypes for H13 and H16 specific signatures in the internal viral proteins, using a bioinformatics approach.

In the present study, whole genome sequence data from Eurasian gull AIVs (subtypes H13 and H16) were compared with a selection of Eurasian non H13 and H16 AIVs to search for H13 and H16 specific amino acids in the internal proteins. Five AIV genomes from Norway and 10 AIV genomes from the Influenza Virus Resource were included in this initial analysis. To verify the identified amino acids as H13 and H16 specific signatures, they were checked against a comprehensive set of AIV sequences in GenBank, including all available H13 and H16 sequences [27,28]. To search for potential functions associated with the identified signatures, *in silico* analyses, including protein similarity searches and prediction of subcellular localization signals were performed. The obtained results suggest that signatures found in the internal proteins of H13 and H16 viruses may be related to host restriction.

## Materials and Methods

### Ethics statement

The samples characterized in this study were obtained from the Norwegian surveillance program for AIV in wild birds. All samples were collected by voluntary, licensed hunters from hunter-harvested birds during the permitted hunting season. No special permit is required during this period.

### Virus isolates

Virus isolates were obtained from samples collected through the surveillance program for AIV in wild birds in Norway during the period 2006–2009. Cloacal and tracheal swabs were collected by hunters from hunter-harvested birds during the ordinary hunting season (August – December). The samples were sent to the Norwegian Veterinary Institute where they were analyzed as previously described [15,29]. In brief, RNA was extracted [15] and analyzed for AIV by real-time RT-PCR. AIV positive samples were HA and NA subtyped by sequencing of HA2 [30] and full-length NA genes [31]. Virus isolation was performed in 10-day-old embryonated chicken eggs for a selection of real-time RT-PCR positive samples [32,33]. Virus isolates were stored at  $-80^{\circ}\text{C}$ .

Five virus isolates, all originating from samples collected in Rogaland county on the south-west coast of Norway, were used in the study. One H13, one H16 and one non H13/H16 isolate from common gull (*Larus canus*) were chosen: A/common gull/Norway/1313/2009(H13N2), A/common gull/Norway/1617/2006(H16N3) and A/common gull/Norway/1602/2009(H6N8). In addition, two randomly selected AIVs from mallard were included: A/mallard/Norway/779/2009(H3N8) and A/mallard/Norway/1537/2009(H9N2). Allantoic fluids from first (H6, H13), second (H3, H16) and third (H9) egg passages were used for sequencing.

### Sequencing and sequence analysis

For full length amplification of the viral segments, a two-step RT-PCR was performed, using SuperScript<sup>®</sup> III (Invitrogen, Carlsbad, CA, USA) and the Advantage 2 PCR kit (Clontech, Mountain View, CA, USA). Cycling conditions [29] and primers [31] used were as previously described. For some polymerase basic protein 2 (PB2) and polymerase basic protein 1 (PB1) genes,

additional primers (Table S1) were used to obtain full length sequences. Amplicons were purified, using Zymoclean<sup>™</sup> Gel DNA recovery kit (Zymo Research, Orange, CA), and cloned into pCR2.1 TOPO vectors (Invitrogen). Sequencing was performed by a commercial sequencing service (GATC Biotech, Constance, Germany). For large amplicons, primers were designed for internal sequencing (Table S2). For each viral segment, three to six clones were sequenced. Initial sequence analysis, including translation of nucleotide sequences into putative amino acid sequences, was performed using Vector NTI Advance ver.11 (Invitrogen). All sequences have been submitted to the EMBL Nucleotide Sequence Database under the accession numbers: HE802704-HE802743.

### Available complete AIV genomes from gulls in Eurasia

A search in the Influenza Virus Resource database [28] for complete AIV genomes from gulls in Europe and Asia identified 13 genomes (as of February 2012) (Table S3). Three of these were excluded from further analysis due to subtype (HPAI H5N1), suboptimal sequence quality (A/slaty-backed gull/Japan/6KS0191/2006(H4N8)) or because it was genetically distinct from all the other Eurasian avian sequence (A/slaty-backed gull/Japan/6KS0185/2006(H4N8)) [34].

### Initial characterization of potentially H13 and H16 specific amino acids

To generate hypotheses regarding the existence of H13 and H16 specific amino acids in the internal viral proteins of H13 and H16 AIVs, the deduced amino acid sequences from the 15 Eurasian viral genomes (Table S3) were aligned using Clustal W in MEGA 4 (data not shown) [35]. The full-length sequences from the following internal viral proteins were analyzed: PB2, PB1, polymerase basic 1 frame 2 (PB1-F2), polymerase acidic (PA), nucleoprotein (NP), matrix 1 (M1), matrix 2 (M2), non-structural protein 1 (NS1) and nuclear export protein (NEP). In the initial analysis, the sequences from each protein were sorted into two groups called Eurasian avian and Eurasian gull AIVs. Eurasian avian viruses were defined as Eurasian AIVs of all subtypes except H13 and H16, while Eurasian gull viruses were defined as Eurasian AIVs from gulls of the H13 and H16 subtypes. All amino acids displaying heterogeneity between these two groups were plotted in Excel spread sheets (Microsoft Excel<sup>®</sup> 2010, Microsoft Corporation). Amino acids showing a distinct pattern between the Eurasian avian and the Eurasian gull viruses were identified as potentially H13 and H16 specific amino acids (Figures S1–S9) and further analyzed.

### Signature validation

To confirm or reject the initially identified amino acids (Figures S1–S9) as H13 and H16 specific signatures, similarity searches were performed on the protein sequences from the two Norwegian H13 and H16 isolates, using tblastn with up to 10000 target sequences of influenza A virus from any species [36]. The target sequences were converted to FASTA format using FASTA BLAST Scan 2.1 [37], and the potentially H13 and H16 specific amino acids (Figures S1–S9) were compared to the amino acids of the target sequences using BioEdit version 7.1.3 [38]. Amino acids mostly or entirely found in H13 and H16 AIVs, were defined as H13 and H16 specific signatures. If several consecutive or nearby amino acids seemed to be involved in a longer motif they were searched together, including any amino acids positions in between, to identify larger H13 and H16 specific motifs where single amino acid positions would be missed when searched individually. The

analysis was repeated on all AIV sequences available (after collapsing identical sequences) at the NCBI Influenza Virus Resource as of April 2012. AIV sequences of all subtypes and from all avian species, both wild and domestic, were included in this search. Furthermore, all H13 and H16 AIV sequences available at the Influenza Virus Resource as of August 2012 were used to finally verify the signatures, and to determine if different signatures were present in H13 and H16 AIVs from Eurasia and America. AIV proteins encoded by reassorted gene segments that most likely originated from subtypes other than H13 and H16, [5], were ignored for the purpose of the latter analysis.

### Prediction of protein similarity and localization

In order to search for potential functions of the verified H13 and H16 specific signatures, protein motif and similarity searches were performed for each of the internal gene products of A/common gull/Norway/1313/2009(H13N2), A/common gull/Norway/1617/2006(H16N3) and A/mallard/Norway/779/2009(H3N8), using Motif search [39], including PROSITE, ProDom, BLOCKS, PRINTS and Pfam. The sequences were compared to previously characterized proteins. When protein family signatures were identified in H13 and H16 viruses but not in AIVs of other subtypes, using A/mallard/Norway/779/2009(H3N8) as reference, the searches were repeated with *in silico* modifications of the H13 and H16 specific signatures to confirm H13 and H16 specific amino acid motifs associated with the proposed protein family signatures. Further, prediction of sub-cellular localization signals within the amino acid sequences was assessed using PSORT [40].

## Results

### Amino acids initially characterized as potentially H13 and H16 specific

Among the 15 AIV genomes initially analyzed for amino acid heterogeneity in the internal proteins, the highest percentage of variable amino acid sites was found in PB1-F2 (54.4%), followed by NS1 (29.6%), NEP (14.0%), M2 (12.4%), NP (6.6%), PB2 (5.9%), PA (4.9%), PB1 (4.2%) and M1 (2.4%). All amino acid positions in the internal proteins that displayed variation between the 15 Eurasian avian and Eurasian gull AIVs included in the initial analysis are shown in Figures S1–S9.

### H13 and H16 specific amino acid signatures

Several of the amino acids assumed to be H13 and H16 specific in the initial analysis were rejected as signatures after validation against the sequences in the NCBI database. Among the verified signatures, some were mostly or entirely found in H13 and H16 AIVs, while others were primarily identified in H13 and H16 AIVs either from Eurasia or America. A few signatures were only present in subgroups of Eurasian or American gull viruses. The amino acids in the internal AIV proteins that were finally verified as H13 and H16 specific signatures are listed in Table 1. H13 and H16 specific signature variants that were only found among the NCBI sequences, and not among the 15 initially analyzed sequences, were also included in Table 1.

The polymerase (PB2, PB1 and PA) and M (M1 and M2) proteins were conserved across AIV subtypes and few signatures were found to be present (Table 1). Signatures highly specific for H13 and H16 viruses were identified in PB2 at positions 470, 559–560 and 674 (Table 1). In PB1 and PA, signatures were found in subgroups of Eurasian gull viruses (Table 1). In PB1-F2, only one signature proved to be specific for a few Eurasian H13 AIVs (Table 1). Similarly, the signatures in the M proteins were also

**Table 1.** Amino acid signatures found mostly or entirely in the internal proteins of Eurasian and American H13 and H16 AIVs from gulls and shorebirds, validated against a comprehensive set of AIV sequences in GenBank.

Protein	Amino acid position(s)	Signatures <sup>1,2</sup>	
		Eurasian H13, H16	American H13, H16
PB2	470	N	N
PB2	559–560	VI/(FI)	MI
PB2	674	S	S
PB1	391	(N)	
PB1-F2	33–37	(LSLTR/LSLTL)	
PA	387–388	(IN)	
NP	87	R	R
NP	105–109	VRELV	VRELV
NP	269	A	A
NP	350–353	ARVL	TRVL
NP	375	S	N
NP	432–433	(SA)	(NP)
M1	63	(I)	
M1	207	G	(G)
M2	55–58	IKYE/IKYD	
NS1	52	I/M	I/M
NS1	60	(S)	
NS1	65–70	IERILD/IERILK	IERILD
NS1	90–95	VDTMTP/MTDMTP/SP	VDTMTP/(VIDMTP)
NS1	111–114	FAGP	FAGP
NS1	124–127	IDKD/IDKE/(IDRE)	LDKN
NS1	152	S/N/G	S
NS1	162	(Q)	Q
NS1	194–197	(ISEA/ISEN)	
NS1	214–216	FAP/LTS/FTP/(LAP)	FAP
NS1	219–222	EQKL/KQKL/(EPKL)	ERKM
NEP	22–26	ESSSG/(VSSSG)	ESSSG
NEP	63–64	RN/KN/KS	
NEP	86–89	KLKT	KLKT

<sup>1</sup>Geographical origin of the AIV segments is based on analyses by Wille et al. (2011) [5].

<sup>2</sup>Signatures in brackets were only found in one or a few H13 or H16 sequences. doi:10.1371/journal.pone.0063270.t001

mostly found in Eurasian gull viruses (Table 1). Specific amino acids that could be used to differentiate between Eurasian and American gull AIVs were found in PB2, NP and NS1 (Table 1).

The majority of H13 and H16 specific signatures were located in NS1, NP and NEP (Table 1). For most internal proteins, identical signatures were present in viruses of the H13 and the H16 subtype. In the initial analysis, variation in specific amino acids between H13 and H16 AIVs seemed to be present in NS1 at positions 52, 70, 90, 127, 197 and 215 (Figure S8) and in NEP at position 63 (Figure S9), respectively. However, validation against H13 and H16 protein sequences at NCBI showed that these were signatures characteristic for two subgroups of Eurasian H13 and H16 AIVs. One of these subgroups contained a large number of geographically reassorted American gull viruses.

## Prediction of protein similarity and localization

The *in silico* analysis predicted that a specific signature was present in NEP of the Eurasian H13 and H16 AIVs sequenced in this study. A motif with similarity to a conserved bacterial protein of unknown function (DUF2316), belonging to the Helix-Turn-Helix protein clan, was predicted in the Eurasian gull isolates by Pfam (independent E-value of 0.07), and was dependent on RN, KN or KS at positions 63–64 rather than GK, and on KLKT at positions 86–89. Changing these key positions of the H13 and H16 AIVs into Eurasian avian specific amino acids abolished the predicted function.

The PSORT analysis identified a nuclear localization signal (NLS) in the NS1 sequences of most AIVs (non H13 and H16 viruses), with a 7 amino acid pattern, starting at position 215 or 216. Neither H13 nor H16 NS1 sequences were predicted to harbor a NLS at this site. Further, in M1, the G residue at position 207 in H13 and H16 isolates abolished the predicted coiled-coil region at position 186–216 that were identified in most avian AIV M1 proteins. In M2, a signature at positions 55–58, located in the amphipathic helix within the C-terminal cytoplasmic domain [41], was observed in the Eurasian gull viruses.

## Discussion

In this paper, we searched for H13 and H16 specific signatures in the internal proteins that may be involved in host restriction of H13 and H16 viruses to gulls and shorebirds. The 15 AIVs used in the initial screening were isolated within a limited period of time (2005–2009) from scattered areas in Northern Eurasia. Eurasian gull AIVs were chosen for the initial analysis since a larger number of American gull viruses are reassortants [5]. The validity of the amino acids first identified as potentially H13 and H16 specific was subsequently assessed towards a large number of sequences in the NCBI database. Therefore, our approach should have been able to detect most signatures, including those present in American H13 and H16 viruses. The signatures were similar between H13 and H16 viruses, but might be biased towards the H13 subtype, since a lower number of complete H16 AIV genomes have been sequenced.

The final verification of the identified signatures against all available H13 and H16 sequences in the NCBI database was complicated by several factors. Firstly, a high proportion of reassorted gene segments were present among the H13 and H16 AIVs, causing a dilution effect on the H13 and H16 specific amino acids [5]. Moreover, several of the signatures were only found to be specific for subgroups of H13 and H16 AIVs. Secondly, adaptive mutations obtained through replication in a new or multiple host species or during virus cultivation, might be present in the studied viral sequences and have led to random changes in some of the signatures, and may explain why some of the H13 and H16 specific signatures were also exceptionally found in AIVs of other subtypes obtained from non-gull hosts. Thirdly, since the classical AIV sequences available from domestic poultry and waterfowls are more numerous and diverse than the H13 and H16 sequences from gulls, the importance of some H13 and H16 AIV signatures could be underestimated.

In the present study, signatures specific for H13 and H16 AIVs were identified in all nine internal viral proteins of the Eurasian gull AIVs. Since gulls are the primary hosts of AIVs of the H13 and H16 subtypes, some of the identified signatures could be important for host restriction, and may be responsible for the poor susceptibility of non-charadriiform avian species to these virus subtypes [42,43]. Among internal proteins, PB2, NP, NS1 and NEP contained most signatures. This is consistent with findings by

Fouchier et al. (2005) who reported that European gull viruses can be phylogenetically distinguished from other influenza A viruses based on their PB2, NP and NS genes [3].

We found H13 and H16 specific signatures at positions 470, 559, 560 and 674 in PB2. These sites have previously been reported to contain amino acids unique for H13 gull PB2s [21]. Our analysis found amino acids VI to be common at positions 559–560 in addition to MI [21], and that MI and VI predominated in American and Eurasian gull AIVs, respectively.

Few H13 and H16 specific signatures were present in PB1 and PA, which is in accordance with previous findings [44]. All sequenced PA and PB1 segments in American AIVs, as of today, have been of Eurasian origin [45] explaining why we only found H13 and H16 specific signatures in a selection of the Eurasian viruses. In Europe, most sequenced H13 and H16 genomes are composed exclusively of gull-like segments. In America, where shorebirds have been suggested to play an important role in the AIV dynamics [4], H13 and H16 AIVs contain avian-like polymerase segments [5]. A polymerase with avian-like PB1 and PA subunits may increase the overall viral fitness of the American gull viruses and facilitate virus transmission between gulls and shorebirds. Gulls and shorebirds belong to the same avian order (i.e. Charadriiformes), which may ease the transmission of H13 and H16 AIVs between these birds. Also in the other viral proteins, a lower number of H13 and H16 specific signatures were found in the American viruses compared to that in the Eurasian viruses. This may be explained by transmission in different host species within each continent, or alternatively represent viral adaptations to genetically different bird populations.

In this study, the M proteins were shown to be conserved across AIV subtypes as previously reported [46], suggesting a minor role of the M proteins in host restriction of AIVs between avian species. In contrast, previous studies have indicated that several amino acid residues in NP are important for the host specificity of H13 AIVs [23]. We identified six signatures in NP of which several had already been described as H13-specific [23]. In addition to the H13-specific 350T, 353L, 375N and 433P reported by Gorman et al. (1990) [23], we found that 350–353 ARVL/TRVL, 375 S/N and 432–433 SA/NP were specific for Eurasian and American H13 and H16 viruses, respectively. NP participates in NP-NP interactions, plays a role during transcription and binds to RNA, PB1, PB2, M1 and to several cellular proteins that are important for nuclear import and export of ribonucleoprotein [47,48]. It was recently shown that NP of human influenza A viruses determines the sensitivity of Mx proteins [49]. The antiviral activity of Mx proteins varies between and within species, and a wide range of intra- and interspecies variation in the Mx gene has been found in wild ducks [50]. Different signatures in NP may contribute to the host restriction of H13 and H16 AIVs to gulls.

Our results show that most H13 and H16 AIV specific signatures are present in the NS1 protein [51]. Some of these signatures are located in the N-terminal RNA binding domain whereas the majority are found in the C-terminal effector domain. The NS1 protein is only present in infected cells [51] and its major function is to counter the innate host antiviral defense primarily by antagonizing type I interferon (IFN) response. Both the RNA binding domain and the effector domain of NS1 are required for optimal suppression of the host immune responses during influenza infection [52]. NS1 proteins of influenza A viruses have multiple ways to exert the IFN-antagonism and differences in these actions have been found to be strain-specific [53]. This suggests that the H13 and H16 AIVs have developed specific signatures in the NS1 protein that enable them to replicate in gulls. Furthermore, the NS1 proteins of H13 and H16 AIVs may be

less adapted to suppress the innate immune response in ducks and thus not be able to efficiently replicate in duck cells.

The effector domain of the NS1 protein facilitates viral replication by interaction with proteins involved in innate cellular antiviral response [54]. One of the H13 and H16 AIV specific signatures was located in the effector domain at amino acid positions 214–216. In a previous study a Src homology (SH) binding motif, PPLPPK, called SH3, was mapped at NS1 amino acid positions 212–217 of most AIVs subtypes [54,55]. This motif is essential for binding to cellular CRK and CRKL proteins, and P212, P215 and K217 are critical for this binding [55]. Binding ability of NS1 to CRK/CRKL may have evolved in virus strains that over-activate the cellular JNK–ATF2 signaling pathway leading to apoptosis [54]. Thus, AIVs containing PPLPPK replicate more efficiently. Our database searches confirm that the SH3 motif, PPLPPK, is conserved among the NS1 proteins of most AIV subtypes. However, this SH3 motif was not found in H13 and H16 AIVs, and it is also uncommon in human influenza A viruses [54]. In human influenza viruses the P in position 215 of NS1 is replaced by T, which results in failure to bind CRK/CRKL proteins [55]. In the NS1 of H13 and H16 AIVs, we found that T or A were present at position 215. T215 in human influenza A viruses is phosphorylated, and replacement with an A residue reduces viral replication [56]. The disrupted SH3 motifs in H13 and H16 AIVs suggest that their NS1s also lack CRK/CRKL binding activity.

NS1 has previously been shown to harbor one conserved NLS, NLS1, at position 34–38. In addition, several influenza viruses harbor a second NLS, NLS2, in the C-terminal part of the NS1 protein (216–221 PKQKRRK). Both NLSs are involved in binding  $\alpha$ -importin, which facilitates nuclear import [57–59]. We found the predicted NLS2 in most avian NS1 proteins, involving parts of the 214–216 and the 219–222 signatures, while NLS2 was not present in the NS1 protein of Eurasian H13 and H16 AIVs. However, as in other AIVs, NLS1 is conserved in the H13 and H16 viruses, suggesting that NLS1 is sufficient for nuclear import of NS1 in these viruses.

For most internal viral proteins, the signatures were found to be similar among H13 and H16 viruses. However, different signatures between the H13 and the H16 subtype seemed to be present in NS1 and NEP in the initial analysis. It has previously been shown that all the internal genes of H13 and H16 viruses from European black-headed gull (*Chroicocephalus ridibundus*), except NS, are closely related [3]. Our initial analysis only included two complete H16 genomes. In addition, one of the selected H13N8 viruses in the initial *in silico* screening, isolated from herring gull (*Larus argentatus*) in Mongolia, had the same motifs as the H16 viruses. However, these features are shared by most NS1 proteins from H16 sequences deposited in GenBank, while only few of the H13 sequences have these combinations of amino acids in NS1. The signature patterns in NS1 turned out to be characteristic for two subgroups of Eurasian H13 and H16 AIVs, of which one contained a large number of American reassorted gull viruses. These results suggest that different genotypes of NS1 are present in these two groups of H13 and H16 AIVs. Alternatively, it might reflect reassortment between H13 and H16 AIVs.

The NS1 is a multifunctional viral protein implying that a multifaceted selection pressure is exerted on this protein, which is mirrored in a large nucleotide sequence variation [51]. The multifunctionality might also bring along strain variations and host-specific adaptation.

In this study, we examined the internal proteins of H13 and H16 AIVs for presence of specific signatures. The surface proteins, the HA in particular, are important determinants of host

specificity of influenza A viruses [20]. The HAs of H13 and H16 AIVs have been found to be resistant to cleavage by several proteases including trypsin, impairing further fusion process *in vitro* [60,61]. The crystal structure of H16 HA0 shows the presence of an  $\alpha$ -helix covering the cleavage site which may be involved in the protease resistance [61]. Furthermore, it has previously been reported that H13 and H16 AIVs replicate poorly in Madin Darby Canine Kidney (MDCK) cells in the presence of trypsin, but efficiently in embryonated chicken eggs, the latter also as recombinants with H13 or H16 subtypes in an A/PR/8/34 backbone, when cell suspension rather than supernatant is inoculated [62]. The presence of HA activating proteases in target species is thus likely an important factor in the host range of influenza A viruses [60]. The identified signatures in our study indicate that the host specificity of H13 and H16 AIVs may not solely be determined by receptor recognition and interaction between HA and host proteases, but that it may be partially explained by the internal viral proteins, and NS1 and NP in particular.

## Conclusions

To summarize, several H13 and H16 AIV specific amino acid signatures were identified in the internal proteins by bioinformatics analysis. The NS1 and NP proteins contained most specific signatures. In particular, H13 and H16 AIVs were predicted to lack a NLS in NS1 present in AIVs of other subtypes. The number of available H13 and H16 AIV sequences in the public databases was much lower compared to that from other AIV subtypes. Sequence data from more geographically dispersed areas is therefore needed to generalize the importance of the identified signatures. Further experimental data will be needed to confirm the *in silico* proposed signatures, and our results provide a basis for future functional studies to get a better understanding of host adaptation and epidemiology of AIVs in wild birds.

## Supporting Information

### Figure S1 Amino acid heterogeneity in the PB2 protein.

Amino acid heterogeneity in the PB2 protein of the 15 Eurasian avian (non H13 and H16) and Eurasian gull (H13 and H16) avian influenza viruses (AIVs) included in the initial analysis. Reassorted gene segments of Eurasian avian origin are indicated by two asterisks.

(TIF)

### Figure S2 Amino acid heterogeneity in the PB1 protein.

Amino acid heterogeneity in the PB1 protein of the 15 Eurasian avian (non H13 and H16) and Eurasian gull (H13 and H16) avian influenza viruses (AIVs) included in the initial analysis.

(TIF)

### Figure S3 Amino acid heterogeneity in the PB1-F2 protein.

Amino acid heterogeneity in the PB1-F2 protein of the 15 Eurasian avian (non H13 and H16) and Eurasian gull (H13 and H16) avian influenza viruses (AIVs), included in the initial analysis.

(TIF)

### Figure S4 Amino acid heterogeneity in the PA protein.

Amino acid heterogeneity in the PA protein of the 15 Eurasian avian (non H13 and H16) and Eurasian gull (H13 and H16) avian influenza viruses (AIVs) included in the initial analysis. Reassorted gene segments of Eurasian avian origin are indicated by two asterisks.

(TIF)

**Figure S5 Amino acid heterogeneity in the NP protein.**

Amino acid heterogeneity in the NP protein of the 15 Eurasian avian (non H13 and H16) and Eurasian gull (H13 and H16) avian influenza viruses (AIVs) included in the initial analysis. (TIF)

**Figure S6 Amino acid heterogeneity in the M1 protein.**

Amino acid heterogeneity in the M1 protein of the 15 Eurasian avian (non H13 and H16) and Eurasian gull (H13 and H16) avian influenza viruses (AIVs) included in the initial analysis. (TIF)

**Figure S7 Amino acid heterogeneity in the M2 protein.**

Amino acid heterogeneity in the M2 protein of the 15 Eurasian avian (non H13 and H16) and Eurasian gull (H13 and H16) avian influenza viruses (AIVs) included in the initial analysis. (TIF)

**Figure S8 Amino acid heterogeneity in the NS1 protein.**

Amino acid heterogeneity in the NS1 protein of the 15 Eurasian avian (non H13 and H16) and Eurasian gull (H13 and H16) avian influenza viruses (AIVs) included in the initial analysis. (TIF)

**Figure S9 Amino acid heterogeneity in the NEP protein.**

Amino acid heterogeneity in the NEP protein of the 15 Eurasian

avian (non H13 and H16) and Eurasian gull (H13 and H16) avian influenza viruses (AIVs) included in the initial analysis. (TIF)

**Table S1 Primers used in the PCR amplification of PB1 from AIV H13N2\* and H9N2\*\* and of PB2 from AIV H3N8\*\*\*.**

(DOCX)

**Table S2 Primers used for internal sequencing of the PA, PB1 and PB2 genes.**

(DOCX)

**Table S3 Accession numbers for complete AIV genomes from gulls in Eurasia, available from the Influenza Virus Resource at NCBI as of February 2012\*.**

(DOC)

**Author Contributions**

Conceived and designed the experiments: RT CMJ. Performed the experiments: RT CMJ EFH AGH. Analyzed the data: RT CMJ ER AGH. Contributed reagents/materials/analysis tools: AGH CMJ ER. Wrote the paper: RT AGH ER CMJ EFH.

**References**

- Jagger BW, Wise HM, Kash JC, Walters KA, Wills NM, et al. (2012) An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* 337: 199–204.
- Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y (1992) Evolution and ecology of influenza A viruses. *Microbiol Rev* 56: 152–179.
- Fouchier RAM, Munster V, Wallensten A, Bestebroer TM, Herfst S, et al. (2005) Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls. *J Virol* 79: 2814–2822.
- Olsen B, Munster VJ, Wallensten A, Waldenstrom J, Osterhaus AD, et al. (2006) Global patterns of influenza A virus in wild birds. *Science* 312: 384–388.
- Wille M, Robertson GJ, Whitney H, Bishop MA, Runstadler JA, et al. (2011) Extensive geographic mosaicism in avian influenza viruses from gulls in the northern hemisphere. *PLoS One* 6: e20664.
- Wallensten A, Munster VJ, Elmerg J, Osterhaus AD, Fouchier RA, et al. (2005) Multiple gene segment reassortment between Eurasian and American lineages of influenza A virus (H6N2) in Guillemot (*Uria aalge*). *Arch Virol* 150: 1685–1692.
- Lebarbenchon C, Chang CM, Gauthier-Clerc M, Thomas F, Renaud F, et al. (2008) H9N2 avian influenza virus in a Mediterranean gull. *J Mol Genet Med* 3: 121–123.
- Krauss S, Obert CA, Franks J, Walker D, Jones K, et al. (2007) Influenza in migratory birds and evidence of limited intercontinental virus exchange. *PLoS Pathog* 3: e167.
- Koehler AV, Pearce JM, Flint PL, Franson JC, Ip HS (2008) Genetic evidence of intercontinental movement of avian influenza in a migratory bird: the northern pintail (*Anas acuta*). *Mol Ecol* 17: 4754–4762.
- Dugan VG, Chen R, Spiro DJ, Sengamalay N, Zaborsky J, et al. (2008) The evolutionary genetics and emergence of avian influenza viruses in wild birds. *Plos Pathog* 4: e1000076.
- Manzoor R, Sakoda Y, Mwene A, Tsuda Y, Kishida N, et al. (2008) Phylogenetic analysis of the M genes of influenza viruses isolated from free-flying water birds from their Northern Territory to Hokkaido, Japan. *Virus Genes* 37: 144–152.
- Ramey AM, Pearce JM, Ely CR, Sheffield Guy LM, Irons DB, et al. (2010) Transmission and reassortment of avian influenza viruses at the Asian-North American interface. *Virology* 406: 352–359.
- Wille M, Robertson GJ, Whitney H, Ojick D, Lang AS (2011) Reassortment of American and Eurasian genes in an influenza A virus isolated from a great black-backed gull (*Larus marinus*), a species demonstrated to move between these regions. *Arch Virol* 156: 107–115.
- Van Borm S, Rosseel T, Vangeluwe D, Vandenbussche F, van den Berg T, et al. (2012) Phylogeographic analysis of avian influenza viruses isolated from Charadriiformes in Belgium confirms intercontinental reassortment in gulls. *Arch Virol* 157: 1509–1522.
- Germundsson A, Madsljen KI, Hjortaa MJ, Handeland K, Jonassen CM (2010) Prevalence and subtypes of influenza A viruses in wild waterfowl in Norway 2006–2007. *Acta Vet Scand* 52: 28.
- Munster VJ, Baas C, Lexmond P, Waldenstrom J, Wallensten A, et al. (2007) Spatial, temporal, and species variation in prevalence of influenza A viruses in wild migratory birds. *PLoS Pathog* 3: e61.
- Hinshaw VS, Air GM, Gibbs AJ, Graves L, Prescott B, et al. (1982) Antigenic and genetic characterization of a novel hemagglutinin subtype of influenza A viruses from gulls. *J Virol* 42: 865–872.
- Brown JD, Poulson RL, Carter D, Lebarbenchon C, Pantin-Jackwood M, et al. (2012) Susceptibility of avian species to north american H13 low pathogenic avian influenza viruses. *Avian Dis* 56: 969–975.
- Gambaryan A, Yamnikova S, Lvov D, Tuzikov A, Chinarev A, et al. (2005) Receptor specificity of influenza viruses from birds and mammals: new data on involvement of the inner fragments of the carbohydrate chain. *Virology* 334: 276–283.
- Yamnikova SS, Gambaryan AS, Tuzikov AB, Bovin NV, Matrosovich MN, et al. (2003) Differences between HA receptor-binding sites of avian influenza viruses isolated from Laridae and Anatidae. *Avian Dis* 47: 1164–1168.
- Gorman OT, Donis RO, Kawaoka Y, Webster RG (1990) Evolution of influenza A virus PB2 genes: implications for evolution of the ribonucleoprotein complex and origin of human influenza A virus. *J Virol* 64: 4893–4902.
- Widjaja L, Krauss SL, Webby RJ, Xie T, Webster RG (2004) Matrix gene of influenza A viruses isolated from wild aquatic birds: ecology and emergence of influenza A viruses. *J Virol* 78: 8771–8779.
- Gorman OT, Bean WJ, Kawaoka Y, Webster RG (1990) Evolution of the nucleoprotein gene of influenza A virus. *J Virol* 64: 1487–1497.
- Kawaoka Y, Gorman OT, Ito T, Wells K, Donis RO, et al. (1998) Influence of host species on the evolution of the nonstructural (NS) gene of influenza A viruses. *Virus Res* 55: 143–156.
- Bean WJ (1984) Correlation of influenza A virus nucleoprotein genes with host species. *Virology* 133: 438–442.
- Obenauer JC, Denson J, Mehta PK, Su X, Mukatira S, et al. (2006) Large-scale sequence analysis of avian influenza isolates. *Science* 311: 1576–1580.
- NCBI (2012) Influenza Virus Resource. Available: <http://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi?go=database>. Accessed 2012 Feb 1.
- Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, et al. (2008) The influenza virus resource at the National Center for Biotechnology Information. *J Virol* 82: 596–601.
- Jonassen CM, Handeland K (2007) Avian influenza virus screening in wild waterfowl in Norway, 2005. *Avian Dis* 51: 425–428.
- Phipps LP, Essen SC, Brown IH (2004) Genetic subtyping of influenza A viruses using RT-PCR with a single set of primers based on conserved sequences within the HA2 coding region. *J Virol Methods* 122: 119–122.
- Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR (2001) Universal primer set for the full-length amplification of all influenza A viruses. *Arch Virol* 146: 2275–2289.
- Toennessen R, Germundsson A, Jonassen CM, Haugen I, Berg K, et al. (2011) Virological and serological surveillance for type A influenza in the black-legged kittiwake (*Rissa tridactyla*). *Virol J* 8: 21.
- World Health Organization (WHO) (2002) Manual of Animal Influenza Diagnosis and Surveillance. Accessed 2012 Nov 11.

34. Bui VN, Ogawa H, Xinigen, Karibe K, Matsuo K, et al. (2012) H4N8 subtype avian influenza virus isolated from shorebirds contains a unique PB1 gene and causes severe respiratory disease in mice. *Virology* 423: 77–88.
35. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596–1599.
36. BLAST: Basic Local Alignment Search Tool (2012) Available: <http://blast.ncbi.nlm.nih.gov/>. Accessed 2012 Apr 1.
37. FASTA BLAST Scan 2.1 (2012) Available: [http://www.bioinformatics.org/annhyb/fasta\\_blast\\_scan/](http://www.bioinformatics.org/annhyb/fasta_blast_scan/). Accessed 2012 Apr 10.
38. BioEdit version 5.0.9 (2012) Available: <http://www.mbio.ncsu.edu/bioedit/bioedit.html>. Accessed 2012 Apr 1.
39. Motif (2012) Available: <http://www.genome.jp/tools/motif/>. Accessed 2012 Apr 10.
40. PSORT (2012) Available: <http://psort.hgc.jp/form2.html>. Accessed 2012 Oct 1.
41. Pielak RM, Chou JJ (2011) Influenza M2 proton channels. *Biochim Biophys Acta* 1808: 522–529.
42. Tonnessen R, Valheim M, Rimstad E, Jonassen CM, Germundsson A (2011) Experimental inoculation of chickens with gull-derived low pathogenic avian influenza virus subtype H16N3 causes limited infection. *Avian Dis* 55: 680–685.
43. Daoust PY, van de Bildt M, van Riel D, van Amerongen G, Bestebroer T, et al. (2012) Replication of 2 Subtypes of Low-Pathogenicity Avian Influenza Virus of Duck and Gull Origins in Experimentally Infected Mallard Ducks. *Vet Pathol*. In press.
44. Kawaoka Y, Krauss S, Webster RG (1989) Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. *J Virol* 63: 4603–4608.
45. Bahl J, Vijaykrishna D, Holmes EC, Smith GJ, Guan Y (2009) Gene flow and competitive exclusion of avian influenza A virus in natural reservoir hosts. *Virology* 390: 289–297.
46. Ito T, Gorman OT, Kawaoka Y, Bean WJ, Webster RG (1991) Evolutionary analysis of the influenza A virus M gene with comparison of the M1 and M2 proteins. *J Virol* 65: 5491–5498.
47. Li Z, Watanabe T, Hatta M, Watanabe S, Nanbo A, et al. (2009) Mutational analysis of conserved amino acids in the influenza A virus nucleoprotein. *J Virol* 83: 4153–4162.
48. Portela A, Digard P (2002) The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. *J Gen Virol* 83: 723–734.
49. Zimmermann P, Manz B, Haller O, Schwemmler M, Kochs G (2011) The viral nucleoprotein determines Mx sensitivity of influenza A viruses. *J Virol* 85: 8133–8140.
50. Dillon D, Runstadler J (2010) Mx gene diversity and influenza association among five wild dabbling duck species (*Anas* spp.) in Alaska. *Infect Genet Evol* 10: 1085–1093.
51. Hale BG, Randall RE, Ortin J, Jackson D (2008) The multifunctional NS1 protein of influenza A viruses. *J Gen Virol* 89: 2359–2376.
52. Tisoncik JR, Billharz R, Burmakina S, Belisle SE, Proll SC, et al. (2011) The NS1 protein of influenza A virus suppresses interferon-regulated activation of antigen-presentation and immune-proteasome pathways. *J Gen Virol* 92: 2093–2104.
53. Kochs G, Garcia-Sastre A, Martinez-Sobrido L (2007) Multiple anti-interferon actions of the influenza A virus NS1 protein. *J Virol* 81: 7011–7021.
54. Hrinicus ER, Wixler V, Wolff T, Wagner R, Ludwig S, et al. (2010) CRK adaptor protein expression is required for efficient replication of avian influenza A viruses and controls JNK-mediated apoptotic responses. *Cell Microbiol* 12: 831–843.
55. Heikkinen LS, Kazlauskas A, Melen K, Wagner R, Ziegler T, et al. (2008) Avian and 1918 Spanish influenza A virus NS1 proteins bind to Crk/CrkL Src homology 3 domains to activate host cell signaling. *J Biol Chem* 283: 5719–5727.
56. Hale BG, Knebel A, Botting CH, Galloway CS, Precious BL, et al. (2009) CDK/ERK-mediated phosphorylation of the human influenza A virus NS1 protein at threonine-215. *Virology* 383: 6–11.
57. Greenspan D, Palese P, Krystal M (1988) Two nuclear location signals in the influenza virus NS1 nonstructural protein. *J Virol* 62: 3020–3026.
58. Melen K, Kinnunen L, Fagerlund R, Ikonen N, Twu KY, et al. (2007) Nuclear and nucleolar targeting of influenza A virus NS1 protein: striking differences between different virus subtypes. *J Virol* 81: 5995–6006.
59. Boulo S, Akarsu H, Ruigrok RW, Baudin F. (2007) Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. *Virus Res* 124: 12–21.
60. Galloway SE, Reed ML, Russell CJ, Steinhauer DA. (2013) Influenza HA Subtypes Demonstrate Divergent Phenotypes for Cleavage Activation and pH of Fusion: Implications for Host Range and Adaptation. *PLoS Pathog* 9: e1003151.
61. Lu X, Shi Y, Gao F, Xiao H, Wang M, et al. (2012) Insights into avian influenza virus pathogenicity: the hemagglutinin precursor HA0 of subtype H16 has an alpha-helix structure in its cleavage site with inefficient HA1/HA2 cleavage. *J Virol* 86: 12861–12870.
62. Keawcharoen J, Spronken MI, Vuong O, Bestebroer TM, Munster VJ, et al. (2010) Repository of Eurasian influenza A virus hemagglutinin and neuraminidase reverse genetics vectors and recombinant viruses. *Vaccine* 28: 5803–5809.