



Communication

# Analysis of Single Nucleotide Variants (SNVs) Induced by Passages of Equine Influenza Virus H3N8 in Embryonated Chicken Eggs

Wojciech Rozek <sup>1,\*</sup>, Malgorzata Kwasnik <sup>1</sup>, Wojciech Socha <sup>1</sup>, Pawel Sztromwasser <sup>2</sup>  and Jerzy Rola <sup>1</sup> 

<sup>1</sup> Department of Virology, National Veterinary Research Institute, Al. Partyzantow 57, 24-100 Pulawy, Poland; malgorzata.kwasnik@piwet.pulawy.pl (M.K.); wojciech.socha@piwet.pulawy.pl (W.S.); jrola@piwet.pulawy.pl (J.R.)

<sup>2</sup> Department of Omics Analyses, National Veterinary Research Institute, Al. Partyzantow 57, 24-100 Pulawy, Poland; pawel.sztromwasser@gmail.com

\* Correspondence: wojciech.rozek@piwet.pulawy.pl

**Abstract:** Vaccination is an effective method for the prevention of influenza virus infection. Many manufacturers use embryonated chicken eggs (ECE) for the propagation of vaccine strains. However, the adaptation of viral strains during subsequent passages can lead to additional virus evolution and lower effectiveness of the resulting vaccines. In our study, we analyzed the distribution of single nucleotide variants (SNVs) of equine influenza virus (EIV) during passaging in ECE. Viral RNA from passage 0 (nasal swabs), passage 2 and 5 was sequenced using next generation technology. In total, 50 SNVs with an occurrence frequency above 2% were observed, 29 of which resulted in amino acid changes. The highest variability was found in passage 2, with the most variable segment being IV encoding hemagglutinin (HA). Three variants, HA (W222G), PB2 (A377E) and PA (R531K), had clearly increased frequency with the subsequent passages, becoming dominant. None of the five nonsynonymous HA variants directly affected the major antigenic sites; however, S227P was previously reported to influence the antigenicity of EIV. Our results suggest that although host-specific adaptation was observed in low passages of EIV in ECE, it should not pose a significant risk to influenza vaccine efficacy.

**Keywords:** equine influenza virus; vaccines; genetic variants; NGS; host adaptation; quasispecies



**Citation:** Rozek, W.; Kwasnik, M.; Socha, W.; Sztromwasser, P.; Rola, J. Analysis of Single Nucleotide Variants (SNVs) Induced by Passages of Equine Influenza Virus H3N8 in Embryonated Chicken Eggs. *Viruses* **2021**, *13*, 1551. <https://doi.org/10.3390/v13081551>

Academic Editor: Romain Paillot

Received: 22 June 2021

Accepted: 2 August 2021

Published: 5 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Influenza A virus (IAV) is a member of the *Orthomyxoviridae* family characterized by a wide range of hosts, including humans, birds, pigs, horses, marine mammals and others. Among the four types of influenza virus, IAV has the highest rate of mutations and causes the most severe diseases [1]. The genome of IAV is composed of eight segments of single-stranded, negative-sense RNA, packed in separate viral ribonucleoprotein complexes. The eight RNA segments (I–VIII) encode at least 10 proteins: polymerase basic 2 (PB2) (segment I), polymerase basic 1 (PB1) (segment II), polymerase acidic (PA) (segment III), hemagglutinin (HA) (segment IV), nucleoprotein (NP) (segment V), neuraminidase (NA) (segment VI), matrix protein 1 (M1) and 2 (M2) (segment VII), non-structural protein 1 (NS1) and 2 (NS2/NEP) (segment VIII) [2]. More recently, additional proteins that are encoded by the IAV genome have been identified: PB1-F2, PB1-N40, PA-X, PA-N155, PA-N182, M42 and NS3 [3–7]. HA and NA glycoproteins are the major determinants of antigenicity because they induce neutralizing antibody production. As with other RNA viruses, influenza exhibits high evolutionary dynamics through several molecular mechanisms including the genetic reassortment of segments, genetic recombination and point mutation, which often results in an antigenic shift or antigenic drift. An antigenic shift is a rapid change in the virus genome and antigenicity due to a reassortment of RNA segments of different

viral strains producing new subtypes and in the case of influenza, causing recurring pandemics [8]. Antigenic drift is associated with point mutations occurring as a result of the low fidelity of viral RNA polymerases. As a consequence of their high mutation rate, influenza viruses exist as a large group of non-identical but closely related viral genomes called quasispecies, which are constantly competing and undergoing selection [9,10]. The high variability rate of IAV is a challenge for vaccination programs at both the vaccine design and formulation stages. It is well established that vaccine mismatch caused by the antigenic drift of the virus leads to the reoccurrence of influenza epidemics and creates the need for regular updates of the vaccine content [11]. A less studied aspect of the risk associated with genetic drift are the mutations induced during the manufacturing processes. Currently, most manufacturers use embryonated chicken eggs (ECE) for vaccine strain propagation, as this allows for the production of large amounts of high-titer viruses (<https://www.cdc.gov/flu/prevent/how-fluvaccine-made.htm>, accessed on 4 February 2021). Previous studies based on Sanger sequencing reported that the adaptation of influenza to chicken embryos may result in changes in hemagglutinin (HA) that could affect the receptor-binding specificity of the virus [12,13]. The goal of our studies was to analyze the distribution of the single nucleotide variants (SNVs) across the whole genome of the equine influenza virus (EIV) during low passages in ECE in order to evaluate viral variability resulting from host adaptation. The use of next generation sequencing (NGS) technology allowed us to determine of the relative frequency of SNVs within a quasispecies population.

## 2. Materials and Methods

### 2.1. Virus Propagation

Equine influenza virus A/equine/Pulawy/1/2005 (H3N8) was isolated in our laboratory and had been characterized previously [14]. The virus was propagated in 10-day-old embryonated specific pathogen-free (SPF) chicken eggs. Briefly, 0.2 mL of PCR-positive nasal swab solution was inoculated into the allantoic cavity of ECE. After 72 h of incubation at 37 °C, the amniotic-allantoic fluids were collected. Five consecutive passages were carried out, and the three fluids with the highest hemagglutination titers were combined and used for inoculation of 15 eggs or RNA extraction (p0, p2 and p5). Fluids showing incomplete hemagglutination ( $HA < 4$ ) after the first passage were combined and used undiluted for the inoculation (p2). Fluids positive in p2 ( $HA \geq 32$ ) were combined and used for the next passage in  $10^{-2}$  dilution in PBS, and for p4 and p5 ( $HA \geq 128$ ), the virus was diluted to  $10^{-3}$ . The amniotic-allantoic fluids from inoculated ECE were stored at  $-70$  °C between passages.

### 2.2. Hemagglutination (HA) and Hemagglutination Inhibition (HI) Assays

The HA and HI assays were performed according to the OIE Manual ([https://www.oie.int/fileadmin/Home/eng/Health\\_standards/tahm/3.06.07\\_EQ\\_INF.pdf](https://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.06.07_EQ_INF.pdf); accessed on 11 February 2021). The HA titer was determined for all collected amniotic-allantoic fluids. Viruses obtained in p2 and p5 were tested in HI with the following reference sera: South Africa/4/03 (FL1 lineage), Newmarket/1/93 (American lineages) and Kentucky/81 (Pre-divergent).

### 2.3. Next Generation Sequencing and Data Analysis

Swab samples taken directly from a naturally infected horse and designated passage 0 (p0) and the amniotic-allantoic fluids from passage 2 (p2) and passage 5 (p5) were selected for further analysis. In order to compare the viral diversity generated at subsequent passages, samples were subjected to deep sequencing. Briefly, RNA extraction was performed using the RNeasy mini kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Each sample was treated with DNase and reverse transcribed with the 5'AGCAAAGCAGG/3 UNI12 primer [15] using Transcriptor Reverse Transcriptase (Roche, Basel, Switzerland). Viral RNA was sequenced on the Illumina MiSeq platform using a  $2 \times 250$  read length with a MiSeq Reagent Kit v3 (Illumina, San Diego, CA, USA).

The adapters were removed by merging overlapping read-pairs using bbmerge (<https://jgi.doe.gov/data-and-tools/bbtools/>; accessed on 19 December 2019) and trimming non-overlapping read-pairs with Trimmomatic [16]. Contigs were assembled using SPAdes 3.9.0 [17]. Reads were mapped to the A/equine/Richmond/1/2007 (H3N8) reference genome using the Burrows–Wheeler alignment BWA MEM algorithm [18]. The coverage depth for each base and segment was calculated using Samtools software (Galaxy version 1.9) with no depth limit set and MQ > 0 [19]. Variant calling was performed using FreeBayes [20]. Single nucleotide variants with a frequency greater than 2% in any virus passage were considered.

In order to assess the diversity of the virus populations, Shannon entropy was calculated for each RNA segment separately and also for whole viral genomes. The following formula was used, where  $f_i$  was the frequency of the nucleotide A, C, G or T at position  $i$  and  $N$  was the total length of the segment [21,22].

$$SE = -\frac{1}{N} \sum_{i=1}^N (f_{iA} \ln f_{iA} + f_{iC} \ln f_{iC} + f_{iT} \ln f_{iT} + f_{iG} \ln f_{iG}) \quad (1)$$

### 3. Results

Full-length nucleotide sequences of all influenza virus RNA segments were obtained from nasal swabs of the naturally infected horse (p0) and amniotic-allantoic fluid samples (p2 and p5). The average coverage depths were 948.54 for p0, 7447.38 for p2 and 15,013.54 for p5 (Table S1). The sequences were submitted to Genbank: p0 (MZ373289-MZ373296), p2 (MZ364009-MZ364016) and p5 (MZ364001-MZ364008).

NGS sequencing data analysis allowed for the identification of 50 SNVs within eight segments of the influenza virus genome. Among them, 29 induced a change in an amino acid, 20 were synonymous, and 1 variant was in the non-coding region of the segment. The positions with variations were distributed as follows: segment I divulged 5 amino acid changes in 11 variants (I—5/11), segment II—6/9, segment III—4/6, segment IV—5/9, segment V—1/3, segment VI—3/4, segment VII—3/4, and segment VIII—2/4 (Table 1, Figure 1). A permanent change in the dominant amino acid variant was observed between p0 and p5 at three positions: segment I (PB2 (A377E)), segment III (PA (R531K)) and segment IV (HA (W222G)).

**Table 1.** Single nucleotide variants of equine influenza virus H3N8 strain A/equine/Pulawy/2005 passaged in embryonated chicken eggs.

Segment (Gene)	Nucleotide			Amino Acid			Passage 0		Passage 2		Passage 5		
	Position	Ref <sup>1</sup>	Alt <sup>2</sup>	Position	Ref	Alt	Alt (%)	Ref/Alt	Alt (%)	Ref/Alt	Alt (%)	Ref/Alt	
I (PB2)	260	G	T	78	W	L	10.78	662/80	0.00		0.41	8699/36	
	556	C	T	177	I	syn <sub>3</sub>	0.00		2.89	6289/187	0.71	6613/47	
	556	A	C	177	I	L	0.00		3.01	6281/195	0.72	6612/48	
	1105	T	G	360	Y	D	0.00		26.93	2936/1082	2.52	3668/95	
	1125	C	T	366	V	syn	48.02	184/170	0.43	3919/17	0.30	3634/11	
	1157	C	A	377	A	E	0.00		62.45	1458/2452	97.54	83/3297	
	1323	T	C	432	H	syn	0.00		3.79	2790/110	10.48	2075/243	
	1347	G	A	440	K	syn	5.26	252/14	2.29	2641/62	2.62	2084/56	
	1348	A	G	441	N	D	5.26	252/14	2.00	2640/54	2.60	2063/55	
	1953	G	A	642	G	syn	0.00		3.30	1406/48	1.79	1375/25	
	2133	G	A	702	K	syn	0.00		99.44	9/1611	99.96	2/4884	
	II (PB1)	437	C	T	138	P	L	0.00		2.62	7832/211	0.00	
		862	G	A	280	A	T	5.51	343/20	7.80	4147/351	4.48	7402/347
1071		G	A	349	A	syn	40.20	180/121	0.24	2919/7	0.27	4808/13	
1140		G	A	372	M	I	0.00		46.87	1521/1342	31.29	3261/1485	
1759		C	A	579	L	M	0.00		0.00		5.98	2264/144	
1770		G	A	582	Q	syn	35.05	126/68	0.00		0.00		
1979		C	G	652	A	G	0.00		9.67	1729/185	3.31	1780/61	
1988		T	G	655	M	R	0.00		26.30	1864/665	20.54	2100/543	
2307		G	A	-	non coding	-	41.05	56/39	0.00		0.00		

Table 1. Cont.

Segment (Gene)	Nucleotide			Amino Acid			Passage 0		Passage 2		Passage 5	
	Position	Ref <sup>1</sup>	Alt <sup>2</sup>	Position	Ref	Alt	Alt (%)	Ref/Alt	Alt (%)	Ref/Alt	Alt (%)	Ref/Alt
III (PA/PA-X)	366	G	A	114	E	syn	8.01		0.00		0.00	
(PA)	1070	A	G	349	E	G	0.00		0.00		3.44	5748/205
	1080	A	G	352	E	syn	0.00		25.53	4805/1647	31.00	3888/1747
	1489	A	T	489	S	C	28.92	644/262	30.21	7470/3234	21.38	11,306/3074
	1616	G	A	531	R	K	40.98	743/516	99.80	26/12,685	99.89	23/21,153
	2122	G	C	700	V	L	1.88	313/6	3.21	5546/184	0.00	
IV (HA-signal)	47	T	C	6	I	syn	0.00		99.76	20/8254	99.54	75/16404
(HA1)	89	C	A	5	I	syn	5.20	1731/95	0.00		0.00	
	665	A	G	197	Q	syn	0.00		0.00		5.42	10,604/608
	738	T	G	222	W	G	0.00		68.88	1812/4010	99.25	73/9723
	753	T	C	227	S	P	0.00		23.73	4422/1376	0.00	
(HA2)	1042	T	C	323	V	A	0.00		0.00		3.63	9924/374
	1078	T	A	335	I	K	35.47	464/255	32.84	9714/4749	10.84	9728/1183
	1082	G	A	336	A	syn	42.65	464/345	32.81	9853/4812	10.73	9755/1172
	1676	T	A	534	F	L	0.00		2.08	3299/70	0.00	
V (NP)	189	A	G	48	K	syn	2.33	882/21	99.79	20/9610	99.87	19/14,348
	711	T	G	222	I	M	44.32	250/199	0.26	4953/13	0.14	8296/12
	1062	G	A	339	E	syn	6.33	296/20	0.00		0.00	
VI (NA)	51	G	A	11	G	R	0.00		41.89	4702/3389	3.52	13,578/496
	205	T	A	62	I	N	0.00		3.57	8696/322	0.00	
	500	A	G	160	K	syn	0.00		7.30	7501/591	1.56	13,249/210
	1207	A	G	396	N	S	0.00		0.00		27.88	6035/2333
VII (M1)	147	C	T	41	A	V	0.00		11.42	14,515/1871	12.23	25,778/3593
	272	G	A	83	A	T	0.00		0.00		2.71	16,057/448
(M2)	367	A	G	114	E	syn	0.15	646/1	45.42	7291/6067	81.64	3202/14,237
	894	A	G	61	R	G	4.06	307/13	0.00		0.00	
VIII (NS1)	216	A	G	64	I	V	0.00		2.73	12,270/344	0.66	19,306/129
	437	C	T	137	I	syn	0.00		5.59	9605/569	0.43	17,447/76
	445	A	G	140	K	R	0.00		5.66	9393/564	0.22	17,413/38
	452	G	A	142	E	syn	0.00		5.78	9184/563	0.40	16,970/68

<sup>1</sup> Dominant variant in passage 0, A/equine/Pulawy/1/2005; <sup>2</sup> Alternative variant; <sup>3</sup> Synonymous variant.

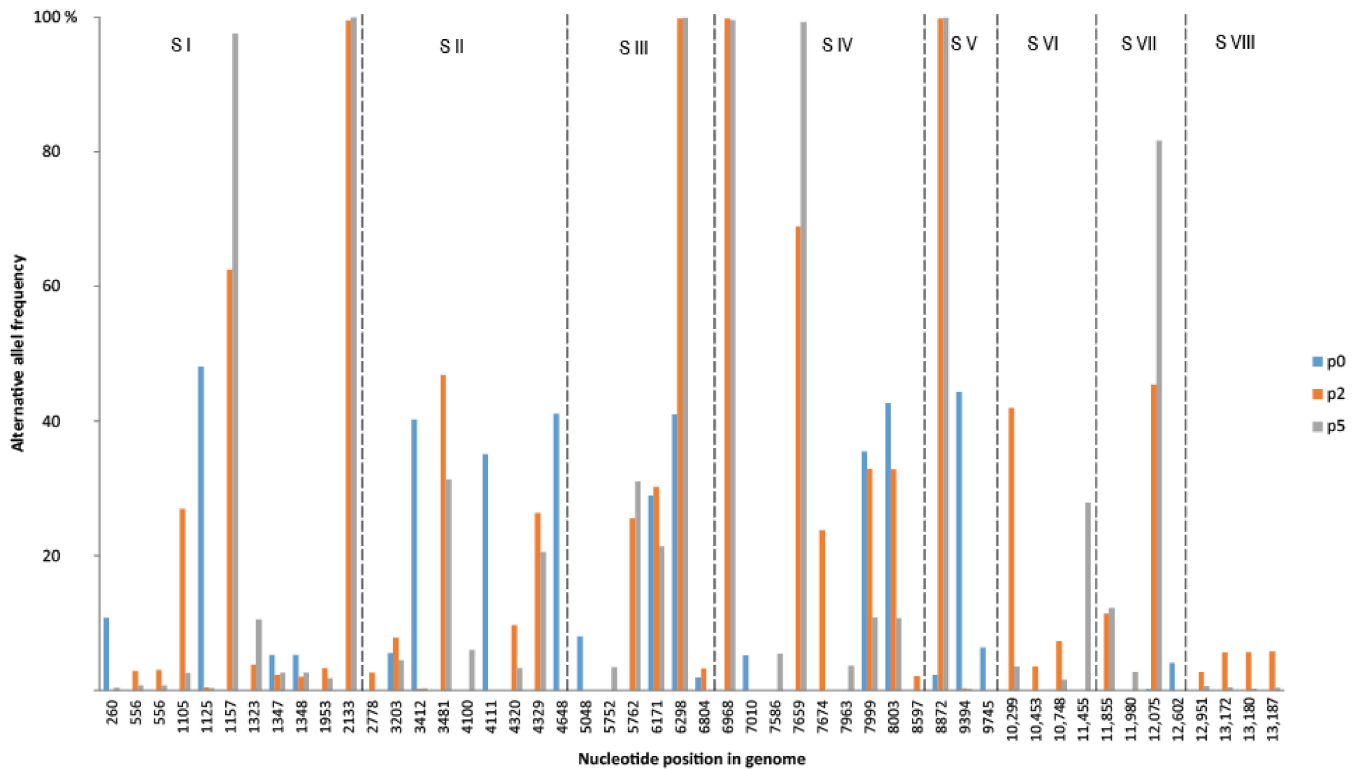
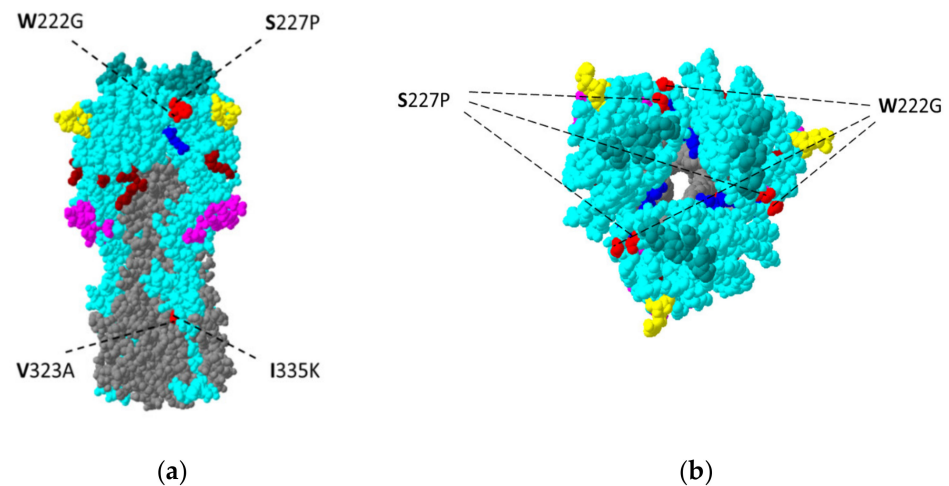


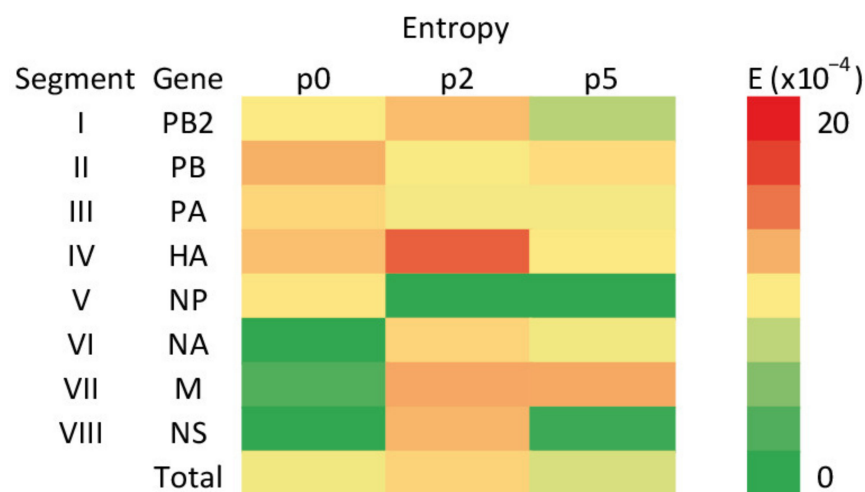
Figure 1. The frequency and distribution of single nucleotide variants of the A/equine/Pulawy/2005 detected in nasal swabs (passage 0) and passages in embryonated chicken eggs (passage 2 and passage 5). S I–S VIII genome segments.

Four of the five amino acid variants—W222G, S227P, V323A and I335K—observed in segment IV (HA) are depicted in Figure 2. None of them were located within any of the previously defined major antigenic regions: A (aa 140–146), B (aa 155–160/187–196), C (aa 44–54/275–280), D (aa 207–212), E (aa 62–64/83/260–264). However, two of them, W222G and S227P, were located within the *loop220* domain [23,24].



**Figure 2.** Single nucleotide variants (red) detected in the hemagglutinin protein of the A/equine/Pulawy/1/2005 (H3N8) strain during passages in embryonated chicken eggs. Regions of hemagglutinin are distinguished by color: teal—HA1; gray—HA2; as are five antigenic sites: yellow—antigenic site A; green—antigenic site B; pink—antigenic site C; dark blue—antigenic site D; and brown—antigenic site E. (a) Side view; (b) top view. The three-dimensional model is based on the crystal structure of A/equine/Richmond/1/2007 (H3N8) HA (PDB accession no. 4UO0) and visualized using Swiss-PdbViewer 4.1.0. [25].

In order to characterize the differences in the complexity of the viral population, the per-site Shannon entropy was calculated, considering the frequencies of nucleotide substitutions across each of the genome segments, and visualized with a heatmap (Figure 3, Table S2).



**Figure 3.** Heatmap of Shannon entropy for EIV genome segments and whole genomes, calculated from the frequencies of variants present in the virus population in nasal swabs (p0), passage 2 (p2) and passage 5 (p5).

Fluctuations of entropy measured between passages that were observed for each of the segments with the overall genome complexity peaking in p2. The highest values were

observed for segment IV ( $9.7 \times 10^{-4}$  on average) with the peak of complexity occurring in p2 ( $14.3 \times 10^{-4}$ ). Additionally, in segments I, VI and VIII, p2 was the passage in which the highest entropy was observed. In contrast, complexity was reduced to 0 for segment V starting from p2.

Viruses obtained in p2 and p5 were tested in HI with the reference sera previously referred to. There was no difference in p2 and p5 reactivity: HI 1/32 for South Africa/4/03 -, HI 1/32 for Newmarket/1/93 and HI 1/256 for Kentucky/81. The virus from p0 showed HA < 4, and therefore the HI assay was not feasible.

#### 4. Discussion

Vaccination is an effective method used for the prevention of influenza virus infections. The efficacy of the vaccine may be reduced as a result of antigenic differences between the vaccine and the field virus strains. The antigenic evolution of EIV is closely monitored by OIE experts and the genetic background of these changes is studied. In addition to the genetic variability of the virus in the population, mutations induced during vaccine production may also compromise vaccine efficacy. The use of ECE for the multiplication of mammalian influenza viruses may engender the risk of changes related to adaptation to a new host, since influenza viruses propagated with this method have been known to undergo phenotypic changes affecting their pathogenic profile since the 1940s [26]. Said et al. (2011) used Sanger sequencing to identify amino acid positions substituted upon adaptation of EIV H3N8 to the chicken host [27]. However, in the case of the influenza virus, it seems to be more reliable to analyze the viral quasispecies populations rather than focus on consensus sequences based on dominant variants. This has become possible with the application of NGS technology. By using this method, we studied the effect of low passages of EIV in ECE in the context of viral quasispecies. It can be assumed that the viral heterogeneity observed at passage 0 reflected the primary host–virus equilibrium, whereas the variability observed in consecutive passages involved changes associated with the adaptation to ECE. In the study by Said et al. (2011), six amino acid substitutions were identified in the course of multiple passages of EIV [27]. Those changes were located not only in the highly variable surface glycoprotein HA, but also in the internal proteins M1, NP and PA. The next generation sequencing performed in our study identified variable positions in each segment of the viral genome. The Shannon entropy calculations for genetic sequences of virus from p0, p2 and p5 gave a measure of viral genetic diversity. The highest average entropy for the whole genome was found in p2, whereas the lowest was observed in p5. The most diverse segment was IV (HA). In segments I (PB2), IV (HA), VI (NA), VII (M) and VIII (NS), entropy increased in p2, then decreased or remained constant in p5. The change in host was accompanied by an increase in the viral variability in p2, while a stabilization and a decrease were observable in p5. A similar tendency was found in a previous study, when changing the host from turkeys to chickens led to an early increase in the overall genetic variability of influenza virus strains followed by a decrease in this indicator by the fifth passage [28].

Out of the 50 identified SNVs, 29 resulted in changes in amino acids. Among non-synonymous SNVs, only 14 occurred with frequencies over 10%. Hemagglutinin and neuraminidase (NA) proteins are the major targets of neutralizing antibodies induced by infection or vaccination and determinants of pathogenicity and host specificity. The former protein binds to sialoglycoprotein on the surface of the host cell and mediates virus attachment, and the latter removes sialic acid from viral receptors and allows the release of progeny virus from host cells [29]. Adaptation to novel host species by the influenza virus requires the readjustment of the functional balance of the sialic acid receptor–binding HA and the receptor-destroying NA to the sialoglycan receptor of the new host [30]. Mutations affecting antigenicity and associated with adaptation investigated in eggs have been identified within HA of the human H3N2 influenza virus strains H156Q [31], L194P [32], and T160K [33]. In our analysis, we found five positions with amino acid substitutions within the HA sequence: W222G, S227P, V323A, I335K and F534L (Table 1). All of these

changes were located outside of the major antigenic sites of HA. The study of the antigenic characterization of viruses obtained during passaging would require the preparation of specific cross-reactive sera against p0, p2 and p5 viruses. To address the issue to some extent, we tested p2 and p5 viruses by HI using three reference sera. There were no differences in the reactivity of p2 and p5 viruses, which seems to be consistent with our SNV analysis showing no variants to be induced strictly at the major antigenic sites of HA. This supports our assumption that low passages in ECE should not pose a significant risk to vaccine production.

Substitutions in HA at positions V323A and F534L appeared only in p2 or p5 and did so with very low frequencies (3.63 and 2.08%). The change at I335K within the fusion peptide of HA2 occurred in p0, p2 and p5 with a frequency above 10% but remained a minority variant. Changes at residues 222 and 227 seemed to be important, as these residues are both located within the receptor-binding site *loop220* of hemagglutinin [23]. In our experiment, the W222G substitution became more fixed in subsequent passages, starting from 0% in p0 but reaching 99.25% in p5. This position has been postulated as affecting the adaptation to the new host [27]. Casalegno et al. (2014) confirmed that changes at position 222 affect the intensity of binding for SAa-2,6 and alter the HA–NA balance, which may influence viral fitness [34]. Mutation W222L could facilitate the viral adaptation of EIV H3N8 to dogs [35]. Yang et al. (2019) described finding predominately tryptophan (W) at residue 222 in the HA of avian H3N2, whereas they observed leucine (L) at this position in canine IAV [12]. The W222G substitution observed in our study seemed to be linked to the propagation of the virus in ECE as it became dominant. Other studies have also reported that the mutation at residue 222 in HA can affect influenza viral infectivity in mice [36]. The same substitution was reported by Ilobi et al. (1994) as being associated with the adaptation of equine H3N8 to Madin–Darby canine kidney cells [37]. The S227P substitution appeared in p2 (where the frequency was 23.73%) but was not observed in p5. Woodward et al. (2015) postulated that mutation at this position may alter not only receptor-binding activity, but also the antigenicity of the virus [38]. This position is also speculated to be critical for the HI reactivity induced by H5 influenza vaccines. However, whether the effect of residue 227 on immunogenicity can be generalized for different subtypes is unclear [39]. In a study analyzing H5N2 AIV strains, Hiono et al. (2016) proved that the glycan-binding specificity of HA was determined by amino acid residues at positions 222 and 227, together with NA [40]. However, in our study, among three nonsynonymous mutations within NA—G11R, I62N, and N396S—none were directly involved in substrate binding.

Proteins of the influenza virus polymerase complex (PB2, PB1 and PA) have been postulated to be involved in the adaptation of avian IAVs to mammalian hosts [41]. The PB2 protein plays a role in viral transcription and has been found to affect the host range as well as the virulence of the influenza viruses [42–44]. We observed that the A377E alteration appeared in p2 (62.45%), and its frequency had increased in p5 (97.54%). This substitution is located in the cap-binding domain (PB2<sub>319–481</sub>) and may be important in the context of mammalian/avian adaptation. The region of PB2 between amino acids 362 and 581 is believed to promote replication in mammalian cells, and single amino acid changes in this sequence contribute to increased virulence in mice [45]. The PA subunit of avian influenza virus may play a major role in the host adaptation. The pandemic H1N1pdm09 virus acquired multiple mutations in the PA gene that promotes polymerase activity in mammalian cells, even in the absence of previously identified host adaptive mutations in other polymerase genes [46]. The substitution in PA observed in our experiment, which is R531K, is located within the large PA-C domain, and this binds the 15 first residues of the PB1 subunit, which is also involved in viral mRNA transcription [47]. The frequency of the substituted R531K variant increased from 40.98% in p0 to over 99% in p2 and p5.

It was speculated that the M1 protein may be involved in host tropism [48]. We found one M1 SNV—A41V, with the rate of the alternative variant exceeding 10% of that which appeared in p2. This mutation was previously described as being associated with a morphological change in the shape of influenza virus particles from a predominantly

filamentous to spherical form. It was conjectured that the filamentous form could present some advantages for the virus in a natural host [49].

## 5. Conclusions

Our results have shown that the passaging of EIV in ECE affects the composition of quasispecies populations. Genetic variants were detected in all EIV segments, and this may reflect the mechanisms of host adaptation. In general, variability was found to be higher in passage 2 than in passage 5, with the most variable being segment IV, encoding hemagglutinin (HA). In most cases, alternative variants appeared with low frequency or did not alter amino acid composition. The frequency of three variants—HA (W222G), PB2 (A377E) and PA (R531K)—clearly increased with the number of passages, becoming dominant. None of the SNVs in the HA gene were located within the regions encoding major antigenic sites; however, variant S227P may affect viral antigenicity. In summary, our results suggest that although host specific adaptation was observed in low passages of EIV in ECE, it seems unlikely that it could create a risk for influenza vaccines' efficacy.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/v13081551/s1>, Table S1: Quantitation of equine influenza virus specific reads and sequence coverage; Table S2: Shannon entropy calculation.

**Author Contributions:** Conceptualization, W.R. and M.K.; methodology, W.R., P.S.; software, P.S., W.S.; data analysis, W.R., M.K., W.S., P.S.; investigation, W.R., M.K., W.S.; writing—original draft preparation, W.R., M.K., W.S., P.S.; writing—review and editing, W.R., J.R.; visualization, W.S., P.S.; supervision, J.R.; project administration, J.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** The study was supported by KNOW, Ministry of Science and Higher Education, Poland, No. 05-1/KNOW2/2015 (K/02/1.0).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data supporting the reported results can be requested from the corresponding author.

**Acknowledgments:** The authors would like to thank Urszula Bocian and Weronika Pietrasiak for technical assistance.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Taubenberger, J.K.; Morens, D.M. The Pathology of Influenza Virus Infections. *Annu. Rev. Pathol. Mech. Dis.* **2008**, *3*, 499–522. [[CrossRef](#)] [[PubMed](#)]
2. Bouvier, N.M.; Palese, P. The Biology of Influenza Viruses. *Vaccine* **2008**, *26*, D49–D53. [[CrossRef](#)]
3. Jagger, B.W.; Wise, H.M.; Kash, J.C.; Walters, K.-A.; Wills, N.M.; Xiao, Y.-L.; Dunfee, R.L.; Schwartzman, L.M.; Ozinsky, A.; Bell, G.L.; et al. An Overlapping Protein-Coding Region in Influenza A Virus Segment 3 Modulates the Host Response. *Science* **2012**, *337*, 199–204. [[CrossRef](#)] [[PubMed](#)]
4. Wise, H.M.; Foeglein, A.; Sun, J.; Dalton, R.M.; Patel, S.; Howard, W.; Anderson, E.C.; Barclay, W.S.; Digard, P. A Complicated Message: Identification of a Novel PB1-Related Protein Translated from Influenza A Virus Segment 2 mRNA. *J. Virol.* **2009**, *83*, 8021–8031. [[CrossRef](#)] [[PubMed](#)]
5. Muramoto, Y.; Noda, T.; Kawakami, E.; Akkina, R.; Kawaoka, Y. Identification of Novel Influenza A Virus Proteins Translated from PA mRNA. *J. Virol.* **2013**, *87*, 2455–2462. [[CrossRef](#)] [[PubMed](#)]
6. Wise, H.M.; Hutchinson, E.C.; Jagger, B.W.; Stuart, A.D.; Kang, Z.H.; Robb, N.; Schwartzman, L.M.; Kash, J.C.; Fodor, E.; Firth, A.E.; et al. Identification of a Novel Splice Variant Form of the Influenza A Virus M2 Ion Channel with an Antigenically Distinct Ectodomain. *PLoS Pathog.* **2012**, *8*, e1002998. [[CrossRef](#)]
7. Selman, M.; Dankar, S.K.; Forbes, N.E.; Jia, J.-J.; Brown, E.G. Adaptive Mutation in Influenza A Virus Non-Structural Gene Is Linked to Host Switching and Induces a Novel Protein by Alternative Splicing. *Emerg. Microbes Infect.* **2012**, *1*, 1–10. [[CrossRef](#)]
8. Dadonaite, B.; Gilbertson, B.; Knight, M.L.; Trifkovic, S.; Rockman, S.; Laederach, A.; Brown, L.E.; Fodor, E.; Bauer, D.L.V. The Structure of the Influenza A Virus Genome. *Nat. Microbiol.* **2019**, *4*, 1781–1789. [[CrossRef](#)]



9. Lauring, A.S.; Andino, R. Quasispecies Theory and the Behavior of RNA Viruses. *PLoS Pathog.* **2010**, *6*, e1001005. [[CrossRef](#)] [[PubMed](#)]
10. Van den Hoecke, S.; Verhelst, J.; Vuylsteke, M.; Saelens, X. Analysis of the Genetic Diversity of Influenza A Viruses Using Next-Generation DNA Sequencing. *BMC Genom.* **2015**, *16*, 79. [[CrossRef](#)]
11. Carrat, F.; Flahault, A. Influenza Vaccine: The Challenge of Antigenic Drift. *Vaccine* **2007**, *25*, 6852–6862. [[CrossRef](#)]
12. Yang, L.; Cheng, Y.; Zhao, X.; Wei, H.; Tan, M.; Li, X.; Zhu, W.; Huang, W.; Chen, W.; Liu, J.; et al. Mutations Associated with Egg Adaptation of Influenza A(H1N1)Pdm09 Virus in Laboratory Based Surveillance in China, 2009–2016. *Biosaf. Health* **2019**, *1*, 41–45. [[CrossRef](#)]
13. Robertson, J.S.; Nicolson, C.; Bootman, J.S.; Major, D.; Robertson, E.W.; Wood, J.M. Sequence Analysis of the Haemagglutinin (HA) of Influenza A (H1N1) Viruses Present in Clinical Material and Comparison with the HA of Laboratory-Derived Virus. *J. Gen. Virol.* **1991**, *72*, 2671–2677. [[CrossRef](#)]
14. Rozek, W.; Purzycka, M.; Polak, M.P.; Gradzki, Z.; Zmudzinski, J.F. Genetic Typing of Equine Influenza Virus Isolated in Poland in 2005 and 2006. *Virus Res.* **2009**, *145*, 121–126. [[CrossRef](#)]
15. 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](#)]
16. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
17. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Pribelski, A.D.; et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [[CrossRef](#)] [[PubMed](#)]
18. Li, H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
19. Li, H. A Statistical Framework for SNP Calling, Mutation Discovery, Association Mapping and Population Genetical Parameter Estimation from Sequencing Data. *Bioinformatics* **2011**, *27*, 2987–2993. [[CrossRef](#)] [[PubMed](#)]
20. Garrison, E.; Marth, G. Haplotype-Based Variant Detection from Short-Read Sequencing. *arXiv* **2012**, arXiv:1207.3907.
21. Lin, J. Divergence Measures Based on the Shannon Entropy. *IEEE Trans. Inform. Theory* **1991**, *37*, 145–151. [[CrossRef](#)]
22. Milani, A.; Fusaro, A.; Bonfante, F.; Zamperin, G.; Salviato, A.; Mancin, M.; Mastrorilli, E.; Hughes, J.; Hussein, H.A.; Hassan, M.; et al. Vaccine Immune Pressure Influences Viral Population Complexity of Avian Influenza Virus during Infection. *Vet. Microbiol.* **2017**, *203*, 88–94. [[CrossRef](#)]
23. Wilson, I.A.; Skehel, J.J.; Wiley, D.C. Structure of the Haemagglutinin Membrane Glycoprotein of Influenza Virus at 3 Å Resolution. *Nature* **1981**, *289*, 366–373. [[CrossRef](#)] [[PubMed](#)]
24. Wiley, D.C.; Wilson, I.A.; Skehel, J.J. Structural Identification of the Antibody-Binding Sites of Hong Kong Influenza Haemagglutinin and Their Involvement in Antigenic Variation. *Nature* **1981**, *289*, 373–378. [[CrossRef](#)] [[PubMed](#)]
25. Guex, N.; Peitsch, M.C. SWISS-MODEL and the Swiss-Pdb Viewer: An Environment for Comparative Protein Modeling. *Electrophoresis* **1997**, *18*, 2714–2723. [[CrossRef](#)]
26. Burnet, F.; Bull, D.H. Changes in influenza virus associated with adaptation to passage in chick embryos. *Aust. J. Exp. Biol. Med.* **1943**, *21*, 55–69. [[CrossRef](#)]
27. Said, A.W.A.; Kodani, M.; Usui, T.; Fujimoto, Y.; Ito, T.; Yamaguchi, T. Molecular Changes Associated with Adaptation of Equine Influenza H3N8 Virus in Embryonated Chicken Eggs. *J. Vet. Med. Sci.* **2011**, *73*, 545–548. [[CrossRef](#)]
28. Świętoń, E.; Olszewska-Tomczyk, M.; Giza, A.; Śmietanka, K. Evolution of H9N2 Low Pathogenic Avian Influenza Virus during Passages in Chickens. *Infect. Genet. Evol.* **2019**, *75*, 103979. [[CrossRef](#)]
29. Sakai, T.; Nishimura, S.I.; Naito, T.; Saito, M. Influenza A Virus Hemagglutinin and Neuraminidase Act as Novel Motile Machinery. *Sci. Rep.* **2017**, *7*, 45043. [[CrossRef](#)] [[PubMed](#)]
30. De Vries, E.; Du, W.; Guo, H.; de Haan, C.A.M. Influenza A Virus Hemagglutinin–Neuraminidase–Receptor Balance: Preserving Virus Motility. *Trends Microbiol.* **2020**, *28*, 57–67. [[CrossRef](#)]
31. Jin, H.; Zhou, H.; Liu, H.; Chan, W.; Adhikary, L.; Mahmood, K.; Lee, M.-S.; Kemble, G. Two Residues in the Hemagglutinin of A/Fujian/411/02-like Influenza Viruses Are Responsible for Antigenic Drift from A/Panama/2007/99. *Virology* **2005**, *336*, 113–119. [[CrossRef](#)]
32. Chen, Z.; Zhou, H.; Jin, H. The Impact of Key Amino Acid Substitutions in the Hemagglutinin of Influenza A (H3N2) Viruses on Vaccine Production and Antibody Response. *Vaccine* **2010**, *28*, 4079–4085. [[CrossRef](#)]
33. Zost, S.J.; Parkhouse, K.; Gumina, M.E.; Kim, K.; Diaz Perez, S.; Wilson, P.C.; Treanor, J.J.; Sant, A.J.; Cobey, S.; Hensley, S.E. Contemporary H3N2 Influenza Viruses Have a Glycosylation Site That Alters Binding of Antibodies Elicited by Egg-Adapted Vaccine Strains. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12578–12583. [[CrossRef](#)]
34. Casalegno, J.-S.; Ferraris, O.; Escuret, V.; Bouscambert, M.; Bergeron, C.; Linès, L.; Excoffier, T.; Valette, M.; Frobort, E.; Pillet, S.; et al. Functional Balance between the Hemagglutinin and Neuraminidase of Influenza A(H1N1)Pdm09 HA D222 Variants. *PLoS ONE* **2014**, *9*, e104009. [[CrossRef](#)]
35. Wen, F.; Blackmon, S.; Olivier, A.K.; Li, L.; Guan, M.; Sun, H.; Wang, P.G.; Wan, X.-F. Mutation W222L at the Receptor Binding Site of Hemagglutinin Could Facilitate Viral Adaptation from Equine Influenza A(H3N8) Virus to Dogs. *J. Virol.* **2018**, *92*, e01115-18. [[CrossRef](#)] [[PubMed](#)]

36. Bradley, K.C.; Galloway, S.E.; Lasanajak, Y.; Song, X.; Heimburg-Molinaro, J.; Yu, H.; Chen, X.; Talekar, G.R.; Smith, D.F.; Cummings, R.D.; et al. Analysis of Influenza Virus Hemagglutinin Receptor Binding Mutants with Limited Receptor Recognition Properties and Conditional Replication Characteristics. *J. Virol.* **2011**, *85*, 12387–12398. [[CrossRef](#)]
37. Ilobi, C.P.; Henfrey, R.; Robertson, J.S.; Mumford, J.A.; Erasmus, B.J.; Wood, J.M. Antigenic and Molecular Characterization of Host Cell-Mediated Variants of Equine H3N8 Influenza Viruses. *J. Gen. Virol.* **1994**, *75*, 669–673. [[CrossRef](#)] [[PubMed](#)]
38. Woodward, A.; Rash, A.S.; Medcalf, E.; Bryant, N.A.; Elton, D.M. Using Epidemics to Map H3 Equine Influenza Virus Determinants of Antigenicity. *Virology* **2015**, *481*, 187–198. [[CrossRef](#)]
39. Hu, Z.; Shi, L.; Zhao, J.; Gu, H.; Hu, J.; Wang, X.; Liu, X.; Hu, S.; Gu, M.; Cao, Y.; et al. Role of the Hemagglutinin Residue 227 in Immunogenicity of H5 and H7 Subtype Avian Influenza Vaccines in Chickens. *Avian Dis.* **2020**, *64*, 445–450. [[CrossRef](#)]
40. Hiono, T.; Okamatsu, M.; Igarashi, M.; McBride, R.; de Vries, R.P.; Peng, W.; Paulson, J.C.; Sakoda, Y.; Kida, H. Amino Acid Residues at Positions 222 and 227 of the Hemagglutinin Together with the Neuraminidase Determine Binding of H5 Avian Influenza Viruses to Sialyl Lewis X. *Arch. Virol.* **2016**, *161*, 307–316. [[CrossRef](#)] [[PubMed](#)]
41. Boivin, S.; Cusack, S.; Ruigrok, R.W.H.; Hart, D.J. Influenza A Virus Polymerase: Structural Insights into Replication and Host Adaptation Mechanisms. *J. Biol. Chem.* **2010**, *285*, 28411–28417. [[CrossRef](#)] [[PubMed](#)]
42. Engelhardt, O.G.; Smith, M.; Fodor, E. Association of the Influenza A Virus RNA-Dependent RNA Polymerase with Cellular RNA Polymerase II. *J. Virol.* **2005**, *79*, 5812–5818. [[CrossRef](#)]
43. Labadie, K.; Dos Santos Afonso, E.; Rameix-Welti, M.-A.; van der Werf, S.; Naffakh, N. Host-Range Determinants on the PB2 Protein of Influenza A Viruses Control the Interaction between the Viral Polymerase and Nucleoprotein in Human Cells. *Virology* **2007**, *362*, 271–282. [[CrossRef](#)]
44. 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.* **1993**, *67*, 1761–1764. [[CrossRef](#)]
45. Yao, Y.; Mingay, L.J.; McCauley, J.W.; Barclay, W.S. Sequences in Influenza A Virus PB2 Protein That Determine Productive Infection for an Avian Influenza Virus in Mouse and Human Cell Lines. *J. Virol.* **2001**, *75*, 5410–5415. [[CrossRef](#)]
46. Lutz, M.M.; Dunagan, M.M.; Kurebayashi, Y.; Takimoto, T. Key Role of the Influenza A Virus PA Gene Segment in the Emergence of Pandemic Viruses. *Viruses* **2020**, *12*, 365. [[CrossRef](#)]
47. Pflug, A.; Guilligay, D.; Reich, S.; Cusack, S. Structure of Influenza A Polymerase Bound to the Viral RNA Promoter. *Nature* **2014**, *516*, 355–360. [[CrossRef](#)] [[PubMed](#)]
48. Furuse, Y.; Suzuki, A.; Kamigaki, T.; Oshitani, H. Evolution of the M Gene of the Influenza A Virus in Different Host Species: Large-Scale Sequence Analysis. *Virol. J.* **2009**, *6*, 67. [[CrossRef](#)] [[PubMed](#)]
49. Elleman, C.J.; Barclay, W.S. The M1 Matrix Protein Controls the Filamentous Phenotype of Influenza A Virus. *Virology* **2004**, *321*, 144–153. [[CrossRef](#)] [[PubMed](#)]