viruses

MDPI

Article

Characterization of EIAV *env* Quasispecies during Long-Term Passage In Vitro: Gradual Loss of Pathogenicity

Cong Liu [1,†], Xue-Feng Wang [1,†], Yan Wang [1], Jie Chen [2], Zhaohua Zhong [3,*], Yuezhi Lin [1,3,*] and Xiaojun Wang [1,*]

1 State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute of Chinese Academy of Agricultural Sciences, Harbin 150069, China; liucong@caas.cn (C.L.); wangxuefeng@caas.cn (X.-F.W.); 17545106209@163.com (Y.W.)

2 Key Laboratory of Special Animal Epidemic Disease of Ministry of Agriculture, Institute of Special Animal and Plant Sciences, Chinese Academy of Agricultural Sciences, Changchun 130112, China; chenjie@caas.cn

3 Department of Microbiology, Harbin Medical University, Harbin 150069, China

* Correspondence: zhongzh@hrbmu.edu.cn (Z.Z.); linyuezhi@caas.cn (Y.L.); wangxiaojun@caas.cn (X.W.); Tel.: +86-86685122 (Z.Z.); +86-45151051746 (Y.L.); +86-45151051749 (X.W.)

† These authors contribute equally to the work.

Received: 15 February 2019; Accepted: 17 April 2019; Published: 24 April 2019

Abstract: As the only widely used live lentiviral vaccine, the equine infectious anima virus (EIAV) attenuated vaccine was developed by in vitro passaging of a virulent strain for 121 generations. In our previous study, we observed that the attenuated vaccine was gradually selected under increased environmental pressure at the population level (termed a quasispecies). To further elucidate the potential correlation between viral quasispecies evolution and pathogenesis, a systematic study was performed by sequencing *env* using several methods. Some key mutations were identified within Env, and we observed that increased percentages of these mutations were accompanied by an increased passage number and attenuated virulence. Phylogenetic analysis revealed that *env* mutations related to the loss of virulence might have occurred evolutionarily. Among these mutations, deletion of amino acid 236 in the V4 region of Env resulted in the loss of one N-glycosylation site that was crucial for virulence. Notably, the 236-deleted sequence represented a "vaccine-specific" mutation that was also found in wild $EIAV_{LN40}$ strains based on single genome amplification (SGA) analysis. Therefore, our results suggest that the EIAV attenuated vaccine may originate from a branch of quasispecies of $EIAV_{LN40}$. Generally, the presented results may increase our understanding of the attenuation mechanism of the EIAV vaccine and provide more information about the evolution of other lentiviruses.

Keywords: EIAV; SGA (single genome amplification); quasispecies; virulence

1. Introduction

As the simplest member of the lentivirus family, equine infectious anemia virus (EIAV) shares similar features with other lentiviruses, including its genomic structure, life cycle, cell tropism and antigen evolution [1,2]. The uniqueness of EIAV lies in the fact that most infected horses become asymptomatic carriers after recurrent febrile episodes, which are associated with a high viral load and anemia, and thus the horses achieve natural immunologic control [3]. Notably, an attenuated vaccine strain has been developed that successfully controls EIAV prevalence in China. The vaccine was generated by serial passage of a virulent EIAV strain ($EIAV_{LN40}$) for 16 generations in vivo, followed by passaging for as many as 121 passages in vitro [4,5]. Due to these peculiarities, the EIAV

attenuation system may provide an ideal model for elucidating the correlates of genomic evolution and immunologic control.

For RNA viruses, mutation swarms are generated rapidly through error-prone viral replication due to inaccuracy of the RNA polymerase [6]. This mutant swarm, which is also termed a viral quasispecies, appears as a whole population with a mutation-selection balance and is significantly involved in viral pathogenesis [7,8]. Viral quasispecies are also essential for viral survival because they yield beneficial viral phenotypes under in vivo selection. Interestingly, the mutation spectra are higher after long-term passage in vitro, which causes a change in virulence [9–11]. Hence, the relationship between RNA viral quasispecies evolution and viral pathogenesis is an intriguing problem.

As discussed previously, multi-position mutations located in the LTR and in *env* appeared during the EIAV attenuation process in vitro, especially in the viral *env* gene, which developed eight hyper variable regions located in the V3 and V4 regions [1,12–14]. Further analysis showed that these variations were related to viral pathogenesis during the EIAV attenuation process. A virulence-correlated parallel in *env* variation was also observed during the EIAV attenuation process [4]. In another study, we observed that the LTR showed a similar pattern at the population level [15]. However, the LTR is a noncoding region which could not reflect the predominant antigen gene of the EIAV evolutionary pattern. Hence, *env* was chosen as the target of evolutionary selection in this study to address this important question. The aim of this study was to characterize *env* quasispecies evolution and further investigate the related mechanism between virulence attenuation and *env* quasispecies. Our results will be of great interest for understanding the evolutionary mechanism of the EIAV attenuated vaccine and defining vaccine development strategies for other lentiviruses.

2. Materials and Methods

2.1. Study Subjects

All samples were stored at the Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences (CAAS). These samples included a virulent strain ($EIAV_{LN40}$) that caused an infection in horses with 100% mortality at a dose of 1×10^5 TCID50; this strain was initially isolated from an EIA positive horse and passed for 16 generations in horses, resulting in three representative strains ($EIAV_{DLV34}$, $EIAV_{DLV62}$, and $EIAV_{DLV92}$) with 100%, 100%, and 9.1% mortality that stemmed from EIAV passage in vitro for 34, 62, and 92 generations, respectively (Figure S1) [16]. Additionally, we included a vaccine strain ($EIAV_{DLV121}$) that provided 85% protection against $EIAV_{LN40}$ challenge and a full-length infectious molecular clone (pLGFD3-8) constructed from a vaccine strain and then passaged in fetal donkey dermal (FDD) cells for 3 generations. Animal experiments showed that this infectious clone was avirulent [17].

2.2. Viral RNA Extraction and cDNA Synthesis

Viral RNA was extracted from 140 µL of the virus samples using the Viral RNA Mini Kit (QIAamp, Dusseldorf, Germany) according to the manufacturer's protocol. Reverse transcription of RNA to single-stranded cDNA was performed using the SuperScript^{TM} IV Reverse Transcriptase System (Life Technologies, Carlsbad, USA). First, 1 µg of RNA, dNTPs (0.5 mM each), and 0.5 µM of the *env* NR primer (5′-CAGCTACAATGGCAGCTATTATAGCAG-3′; nucleotides (nt) 6702 to 6676 of the EIAV sequence) were incubated for 5 min at 65 °C to denature the RNA secondary structure. First-strand cDNA synthesis was carried out in 20 µL reaction mixtures with 5× SSIV buffer, 5 mM DTT, 2 U/µL of an RNase inhibitor (RNaseOUT) (Life Technologies, Carlsbad, USA), and 10 U/µL of SuperScript^{TM} IV (Life Technologies, Carlsbad, USA). The reaction mixture was incubated at 55°C for 15 min and heated to 80°C for 10 min to complete the reverse transcription reaction, followed by RNase H (Life Technologies, Carlsbad, USA) digestion at 37 °C for 20 min. The synthesized cDNA was used immediately for PCR or stored at −80 °C.

2.3. Bulk PCR

The full-length hypervariable region of the *env* cassette was amplified by nested PCR from the viral cDNA. The specific method was as follows. First, 0.6 μL of bulk cDNA was used for the first-round PCR in a 20 μL volume. The PCR was performed using the KOD FX (Toyobo, Osaka, Japan) system, which included 2× PCR buffer, 0.4 mM dNTPs, and 0.1 μM of the *env* NF (5′-GAAGGCCATCAGGGAGGGAAG-3′; (nt) 5240 to 5260) and *env* NR primers. The cycling conditions were as follows: 94 °C for 2 min, followed by 35 cycles of 94 °C for 30 s, 58 °C for 30 s, and 68 °C for 1 min 30 s, and a final extension of 68 °C for 2 min. The second-round PCR was performed using 0.6 μL of the first-round PCR product and the V3-1 (5′-ACAAACATATACAGGACATCT-3′; (nt) 5808 to 5828) and V4-1 (5′-CTCCAATATTCCAAGAAATAC-3′; (nt) 6218 to 6198) primers under the same conditions used for the first-round PCR. The final PCR products were analyzed by 1% agarose gel electrophoresis. The recycled products were ligated into pMD18-T (TaKaRa, Dalian, China). The constructed plasmids were transformed into E. coli DH5α (TaKaRa, Dalian, China) and spread onto a plate. A total of 98 positive clones identified by PCR were sequenced by COME (Jilin, China).

2.4. Probe Design and Real-Time PCR

Two specific probes and primers were synthesized by General Biosystems (Anhui, China). The first probe was the 6-carboxy-fluorescein (FAM) probe (probe-FAM: 5′-TGCAGCAAAGTAAAAAC ACTTGGAT-3′; (nt) 6010 to 6034), which targeted the *env*-Δ236D-phenotype sequence, and the second was the hexachloride fluorescein (HEX) probe (probe-HEX: 5′-TGCAGCAAAGCGATAAT AACACTTG-3′; (nt) 6010 to 6034), which targeted the *env*-236D-phenotype sequence. The *env*-Δ236D-phenotype and *env*-236D-phenotype sequences were named depending on the presence of the 236D deletion in the EIAV Env. The DNA fragments of *env*-Δ236D-phenotype and *env*-236D-phenotype were amplified with primer F (5′-ATAGGAGGTAGACTAAATGGTTCAGG-3′; (nt) 5634-5668) and primer R (5′-AAACAAAAAGAATGGAGGTTGGACA-3′; (nt) 6143-6119) and ligated into the pMD-18T (TaKaRa, Dalian, China) vector. Then, the two plasmids were serially diluted to 10^2 to 10^9 copies/μL as standards to construct real-time PCR standard curves. The experiments were performed using 1 μM of each primer, 10 μL of premix Ex Taq (TaKaRa, Dalian, China), 0.25 μM of the HEX-labeled or FAM-labeled probe, 2 μL of the cDNA samples and RNase-free water to a final volume of 20 μL. The samples were amplified using TaqMan-based real-time PCR on the ABI Prism 7500 sequence detection system (Applied Biosystems GmbH, Thermo Fisher Scientific, Wilmington, USA) with one cycle at 95 °C (5 min), followed by 40 cycles at 95 °C (15 s), 58 °C (30 s), and 72 °C (20 s). The calculated efficiencies for all primers were determined by dilution experiments and ranged from 97% to 98%; thus, the target sequences were amplified with similar efficiencies. All samples were run with at least three duplicates.

2.5. SGA (Single Genome Amplification)

According to the Poisson distribution [18], the DNA dilution yielding positive PCR products in no more than 30% of the wells will contain one amplicon resulting from single molecule amplification in over 80% of these wells. The viral cDNA was serially diluted to yield a single copy by nested PCR as described for the bulk PCR. All products derived from cDNA dilutions conforming to the above description were sequenced with the V3-1 and V4-1 primers by GENEWIZ (Suzhou, China).

2.6. Data Processing and Analysis

The sequences generated in this study have been submitted to GenBank (GenBank acc. no. MK268242 to MK268339 for bulk PCR and MK278920 to MK279296 for SGA). A preliminary sequence analysis was conducted using the SeqMan and MegAlign programs from the Lasergene DNA & Protein analysis software (version 7.0, DNASTAR Inc., Madison, WI, USA), the BioEdit Sequence Alignment

Editor (version 7.2.5), and ClustalX (version 2.1). A maximum likelihood tree was constructed using MEGA 7.0 [19]. The amino acid alignment was automatically generated by MargFreq (version 1.02).

3. Results

3.1. General Characteristics of EIAV env Variability during the Course of EIAV Attenuation Using Bulk PCR

To verify the characteristics of EIAV *env* variation, some *env* genes derived from EIAV strains during the critical stages of the attenuation process were analyzed using bulk PCR [20,21]. A phylogenetic tree was constructed based on 98 *env* sequences amplified from EIAV strains with differing virulence ($EIAV_{LN40}$, $EIAV_{DLV34}$, $EIAV_{DLV62}$, $EIAV_{DLV92}$, and $EIAV_{DLV121}$) (the morbidity of all strains showed in Figure S1). The general features of *env* variability are shown in Figure 1. Obviously, individual clusters accompanied by a decrease in virulence were generated. In other words, the evolutionary pattern of the *env* gene was consistent with the loss of virulence. Notably, the sequences derived from $EIAV_{DLV62}$ displayed a separated distribution between $EIAV_{LN40}$ and $EIAV_{DLV121}$. This result suggested that $EIAV_{DLV62}$ was an important viral evolutionary state that was associated with EIAV pathogenicity. The details of the sequences that also displayed several mutations in each cluster are presented in Table 1. The isolates from each generation displayed the same mutation types (insertions and deletions) but had different mutation frequencies and sites. We observed that the proportion of strains carrying this mutation gradually increased with the serial passage number in vitro ($EIAV_{DLV34}$, $EIAV_{DLV62}$, $EIAV_{DLV92}$, and $EIAV_{DLV121}$ = 6.25%, 18.5%, 42.1%, and 100%, respectively). We also noted that most mutations were random except for a successful amino acid deletion (236D) located in the V4 domain of Env compared to the reference strain sequence ($EIAV_{LN40}$), as shown in Figure 2. This mutation was also stable in horses inoculated with $EIAV_{DLV121}$ [22]. Thus, we speculated that this site might represent a vaccine-specific mutation.

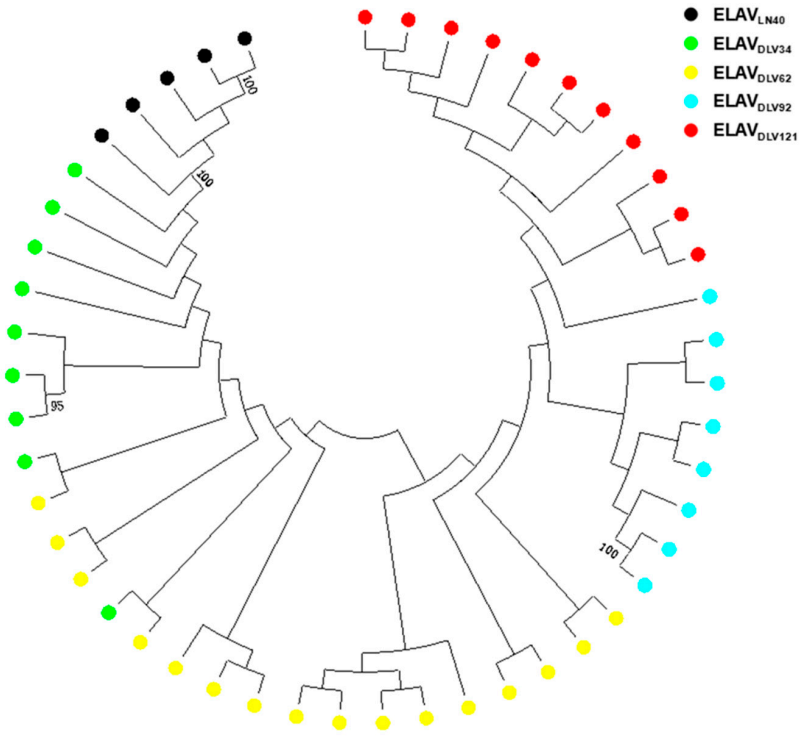

Figure 1. The evolutionary history was inferred by applying MEGA 7.0 program based on the general time reversible model. Whole genome sequences were obtained by bulk PCR. Bootstrap values above 80% are indicated. $EIAV_{LN40}$, $EIAV_{DLV34}$, $EIAV_{DLV62}$, $EIAV_{DLV92}$, and $EIAV_{DLV121}$ are presented with black circles, black triangles, gray circles, blue circles and red circles, respectively.

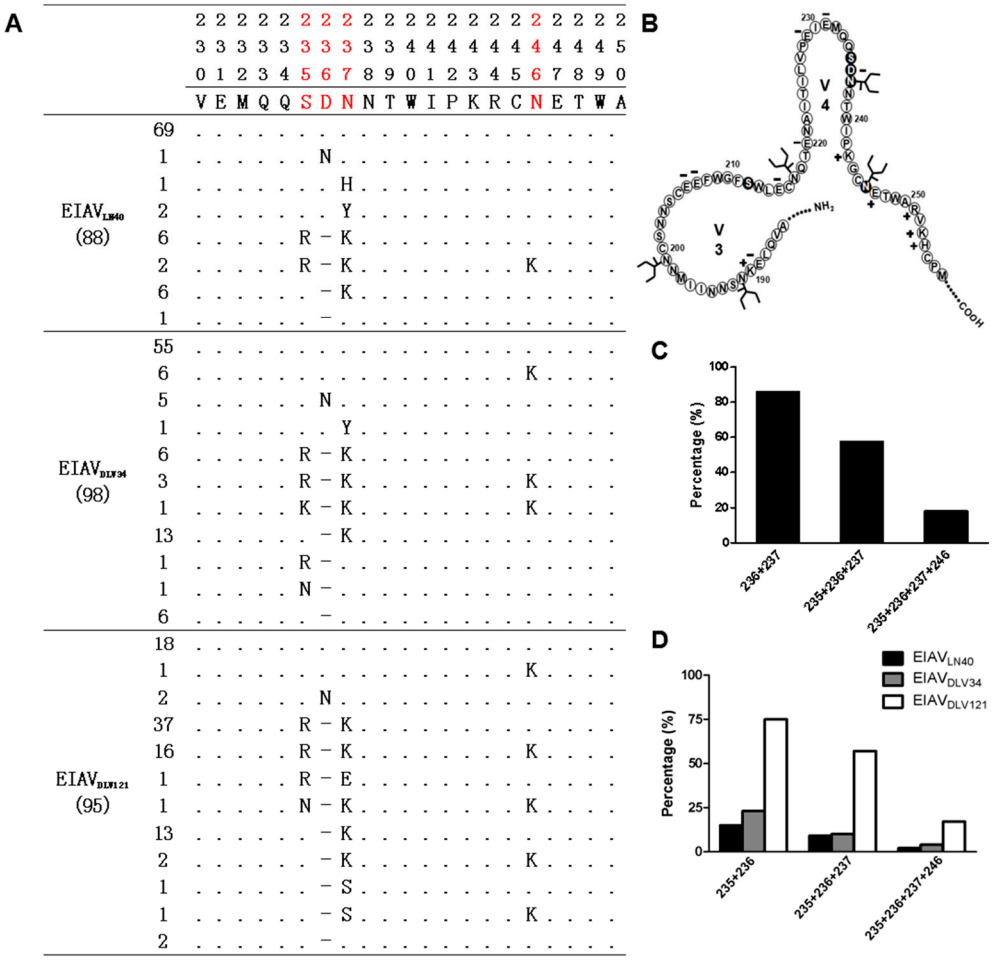

Figure 2. Distribution of the consensus mutations located in the V3 and V4 regions and schematic structural representations. (**A**) Red letters are defined as critical consensus mutation sites displayed within the Env region during the EIAV attenuation process. (**B**) Schematic figure of the $EIAV_{DLV121}$ V3 and V4 regions. N-Glycosylation sites are shown as branched lines. (**C**) Comparison of the proportions of several combined mutations at 235, 236, 237, and 246 for the whole mutants. (**D**) The percentages of these combined mutations at the key stages of vaccine development ($EIAV_{LN40}$, $EIAV_{DLV34}$, and $EIAV_{DLV121}$).

Table 1. Amino acid differences in Env V4 among different Equine Infectious Anima Virus (EIAV) strains obtained using bulk PCR.

	221	**222**	**223**	**224**	**225**	**226**	**227**	**228**	**229**	**230**	**231**	**232**	**233**	**234**	**235**	**236**	**237**	**238**	**239**	**240**	**241**	**242**	**243**	**244**	**245**	**246**	**247**	**248**	**249**	**250**	**251**	**252**	**Mutation Rate (236D/-)**	**Morbidity (Horses)**
$EIAV_{LN40}$	N_{29}	A_{29}	I_{29}	T_{29}	I_{29}	L_{29}	V_{29}	P_{29}	E_{29}	V_{29}	E_{29}	M_{28} I_1	Q_{29}	Q_{28} R_1	S_{29}	D_{29}	N_{28} S_1	N_{29}	T_{29}	W_{29}	I_{29}	P_{29}	K_{29}	R_{29}	C_{27} Y_2	N_{29}	E_{29}	T_{29}	W_{29}	A_{29}	R_{29}	V_{29}	0.00%	100%
$EIAV_{DLV34}$	S_9 N_7	A_{16}	I_{15} V_1	T_{15} A_1	I_{16}	L_{16}	V_{16}	P_{16}	E_{16}	V_{10} I_6	E_{16}	M_{16}	Q_{16}	Q_{14} K_1 R_1	S_{16}	D_{14} N_1 $-_1$	N_{16}	N_{16}	T_{16}	W_{16}	I_{16}	P_{16}	K_{16}	G_8 R_8	C_{16}	N_{14} K_2	E_{13} K_3	T_{16}	W_{16}	A_{16}	R_{12} K_4	V_{16}	6.25%	100%
$EIAV_{DLV62}$	N_9 S_7	A_{16}	I_{16}	T_{16}	I_{16}	L_{16}	V_{16}	P_{16}	E_{16}	V_9 I_7	E_{16}	M_{16}	Q_{16}	Q_{14} R_2	S_{16}	D_9 N_4 $-_3$	N_{16}	N_{16}	T_{16}	W_{16}	I_{16}	P_{16}	K_{16}	R_{11} G_5	C_{16}	N_{13} K_3	E_{13} K_2 G_1	T_{16}	W_{16}	A_{16}	R_{12} K_4	V_{16}	18.5%	100%
$EIAV_{DLV92}$	N_{11} S_8	A_{19}	I_{19}	T_{16} A_3	I_{19}	L_{19}	V_{19}	P_{19}	E_{19}	V_{18} I_1	E_{19}	M_{19}	Q_{19}	Q_{17} R_1 K_1	S_{19}	D_3 N_8 $-_8$	N_{13} K_5 R_1	N_{18} D_1	T_{16} A_3	W_{19}	I_{18} V_1	P_{19}	K_{19}	R_{17} G_2	C_{19}	N_{14} K_5	E_{18} K_1	T_{18} A_1	W_{19}	A_{19}	R_{12} K_7	V_{19}	42.1%	9.1%
$EIAV_{DLV121}$	N_{11} S_7	A_{17} T_1	I_{18}	T_{17} A_1	I_{18}	L_{18}	V_{18}	P_{18}	E_{17} K_1	V_{17} I_1	E_{17} K_1	M_{18}	Q_{18}	Q_{18}	R_{15} S_3	$-_{18}$	K_{17} S_1	N_{18}	T_{18}	W_{18}	I_{17} V_1	P_{18}	K_{18}	R_{17} G_1	C_{18}	K_9 N_8 E_1	E_{10} K_8	T_{18}	W_{18}	A_{18}	K_{13} R_5	V_{17} I_1	100%	0

The additional amino acid D, which consistently appeared in virulent EIAV strains, is marked with dark grey, and N-linked glycosylation sites are marked with light grey.

3.2. Precise Characteristics of the EIAV env Mutation Distribution during the Attenuation Process

Once the vaccine-specific mutation was verified, we wanted to investigate whether this mutation could be detected during serial passage using a higher sensitivity method. These data will further benefit our understanding of the relationship between *env* quasispecies and virulence. Firstly, we adopted double-probe real-time PCR to accurately determine the proportion of strains with the *env*-Δ236D-phenotype and *env*-236D-phenotype sequences (which refer to virus sequences without and with the 236D mutation) in different EIAV generations [23,24]. An *env*-236D-phenotype-specific probe (HEX) and an *env*-Δ236D-phenotype-specific probe (FAM) were designed based on the V4 domain sequences (Figure S2). To avoid the influence of the initial viral copy number on the detection sensitivity, samples containing different initial viral copy numbers were analyzed in parallel (10^2 copies/μL, 10^3 copies/μL and 10^4 copies/μL). As presented in Figure S3, the *env*-Δ236D-phenotype sequence was detected with low copy numbers from all virulent strains ($EIAV_{DLV34}$ and $EIAV_{DLV62}$) except for $EIAV_{LN40}$, and the proportions of strains with the *env*-Δ236D-phenotype gradually increased with the passage number (from 2.47% to 67.5% at the 10^3 copies/μL concentration), whereas the proportion of strains with the *env*-236D-phenotype decreased concomitantly.

Although a gradual increase in the percentage of vaccine-specific sequences was observed using double-probe real-time PCR, the distribution of this mutation in $EIAV_{LN40}$ was still unclear. The SGA method can avoid the recombination events induced by the polymerase and better detect minor nucleotide changes at the same position; thus, this technique has been performed recently for viral quasispecies analysis [25–27]. Thus, we were interested in using SGA to resequence *env* to obtain more information about *env* quasispecies during long-term passage. Samples obtained separately from $EIAV_{LN40}$ (n = 88), $EIAV_{DLV34}$ (n = 98), $EIAV_{DLV121}$ (n = 95) and infectious clone pLGFD3-8 (n = 100) were amplified by SGA and aligned. As expected, the 236D mutation was found in the EIAV strains derived from each generation. The proportion of sequences with the 236D mutation best matched the results derived from the bulk PCR (17.0%, 30.6%, 77.9%, and 100%, respectively) (Table 2). Importantly, we detected 236D (15/72) in the $EIAV_{LN40}$ reservoir, which was not observed using bulk PCR. Additionally, the SGA method exhibited a relatively higher mutation detection rate than bulk PCR. Hence, more variants located at the same position in the EIAV population were observed; for instance, multiple substitutions (E, V, K or Q) instead of a single substitution (E) were detected at position 229 in $EIAV_{LN40}$ using the SGA method. In particular, in addition to position 236, we found that three other positions, which were separately located at S235R, N237K, and N246K, presented gradual and successful mutation accumulation with the increasing passage number, as presented in Table 2. Further analysis showed that the deletion of amino acid 237D resulted in the loss of a N-glycosylation site (237NNTW240) and that the N237K/N246K mutations would also be involved in N-glycosylation site formation (Figure 2A,B). We also noted that these key mutations (235, 236, 237, and 247) were absent or present in combination; for instance, the frequency of the combined mutations at sites 236 and 237 was 85.8% (109/127). The proportions of mutation that concurred at 235, 236, and 237 was 57.5% (73/127). Meanwhile, the sequences with mutations at 235, 236, 237, and 246 showed a lower frequency than the other sequences, as shown in Figure 2C. Moreover, the proportion of these combined mutations at sites 235–236, 235–236–237 and 235–236–237–246 gradually increased as the virulence of the strains decreased, especially in 235–236 (shown in Figure 2D). Hence, we speculated the existence of a link between the sites 236 and 237. According to these results, we speculated that long-term passage in vitro not only had a significant effect on the *env* quasispecies compositions but also might contribute to viral attenuation.

Table 2. Amino acid differences in Env V4 among different EIAV strains detected using single genome amplification (SGA).

	221	**222**	**223**	**224**	**225**	**226**	**227**	**228**	**229**	**230**	**231**	**232**	**233**	**234**	**235**	**236**	**237**	**238**	**239**	**240**	**241**	**242**	**243**	**244**	**245**	**246**	**247**	**248**	**249**	**250**	**251**	**252**	**Mutation Rate (236D/-)**
$EIAV_{LN40}$	N_{86} S_{2}	A_{88}	I_{88}	T_{88}	I_{88}	L_{88}	V_{88}	P_{88}	E_{85} V_{1} K_{1} Q_{1}	V_{87} I_{1}	E_{88}	M_{88}	Q_{87} H_{1}	Q_{88}	S_{80} R_{8}	D_{72} $-_{15}$ N_{1}	N_{71} K_{14} H_{1}	N_{88}	T_{88}	W_{88}	I_{88}	P_{88}	K_{87} N_{1}	R_{80} G_{8}	C_{88}	N_{86} K_{2}	E_{87} K_{1}	T_{87} I_{1}	W_{88}	A_{88}	R_{81} K_{7}	V_{88}	17.0%
$EIAV_{DLV34}$	N_{51} S_{47}	A_{97} T_{1}	I_{98}	T_{98}	I_{98}	L_{97} V_{1}	V_{98}	P_{97} L_{1}	E_{94} K_{3} D_{1}	V_{60} I_{38}	E_{96} K_{2}	M_{98}	Q_{98}	Q_{82} K_{8} R_{8}	S_{86} R_{10} K_{1} N_{1}	D_{63} $-_{30}$ N_{5}	N_{75} K_{22} Y_{1}	N_{98}	T_{98}	W_{95} S_{1} C_{2}	I_{95} S_{2} C_{1}	P_{97} L_{1}	K_{96} Q_{1} P_{1}	R_{57} G_{41}	C_{97} W_{1}	N_{85} K_{10} S_{3}	E_{79} K_{17} G_{2}	T_{96} A_{1} I_{1}	W_{94} C_{3} R_{1}	A_{94} C_{2} S_{1} T_{1}	R_{75} K_{20} T_{3}	V_{94} I_{4}	30.6%
$EIAV_{DLV121}$	N_{76} S_{18} K_{1}	A_{95}	I_{95}	T_{92} A_{3}	I_{95}	L_{95}	V_{94} I_{1}	P_{95}	E_{92} K_{1} D_{1} Q_{1}	I_{10} V_{84} L_{1}	E_{93} K_{1} V_{1}	M_{94} V1	Q_{94} H_{1}	Q_{91} R_{3} K_{1}	R_{54} S_{40} N_{1}	$-_{74}$ D_{19} N_{2}	K_{69} N_{23} S_{2} E_{1}	N_{95}	T_{95}	W_{95}	I_{95}	P_{95}	K_{95}	R_{68} G_{25} K_{2}	C_{95}	N_{74} K_{21}	E_{83} K_{11} G_{1}	T_{94} A_{1}	W_{95}	A_{93} T_{2}	K_{55} R_{40}	V_{95}	77.9%
pLGFD3-8	N_{100}	A_{100}	I_{100}	T_{100}	I_{100}	L_{100}	V_{100}	P_{100}	E_{99} Q_{1}	V_{97} I_{3}	E_{99} K_{1}	M_{100}	Q_{98} H_{1} R_{1}	Q_{100}	S_{87} R_{13}	$-_{100}$	K_{10}	N_{10}	T_{97} I_{3}	W_{98} C_{2}	I_{99} V_{1}	P_{100}	K_{100}	G_{88} R_{12}	C_{99} G_{1}	K_{99} N_{1}	K_{99} E_{1}	T_{10}	W_{99} C_{1}	A_{100}	R_{97} K_{2} M_{1}	V_{100}	100%

The additional amino acid D, which consistently appeared in virulent EIAV strains, is marked with dark grey, and N-linked glycosylation sites are marked with light grey.

3.3. The Relationship between EIAV Attenuation In Vitro and env Gene Quasispecies

To better analyze the relationship between virulence and *env* quasispecies, we constructed a maximum likelihood tree using the MEGA 7.0 program based on the SGA results. As presented in Figure 3, two obvious major groups of *env* sequences have evolved, as represented by the *env*-236D-phenotype (indicated with solid circles) and *env*-Δ236D-phenotype (indicated with hollow circles). Then, serial passaging events continued to lead to accumulation of the *env*-Δ236D-phenotype in the whole population, which made the variant distribution even more complex (i.e., the variant distribution of $EIAV_{DLV34}$). Over the course of 121 generations in vitro, the cluster distribution of related sequences taken from $EIAV_{DLV121}$ showed a pattern similar to that of $EIAV_{LN40}$ except for reversal of the predominant phenotype from *env*-236D to *env*-Δ236D at the individual population level. We observed that two typical *env* phenotypes displayed a gradual transition pattern. These results demonstrated that a predominant phenotype transition occurred in the natural virus population under selection, which was consistent with the LTR investigation results [15].

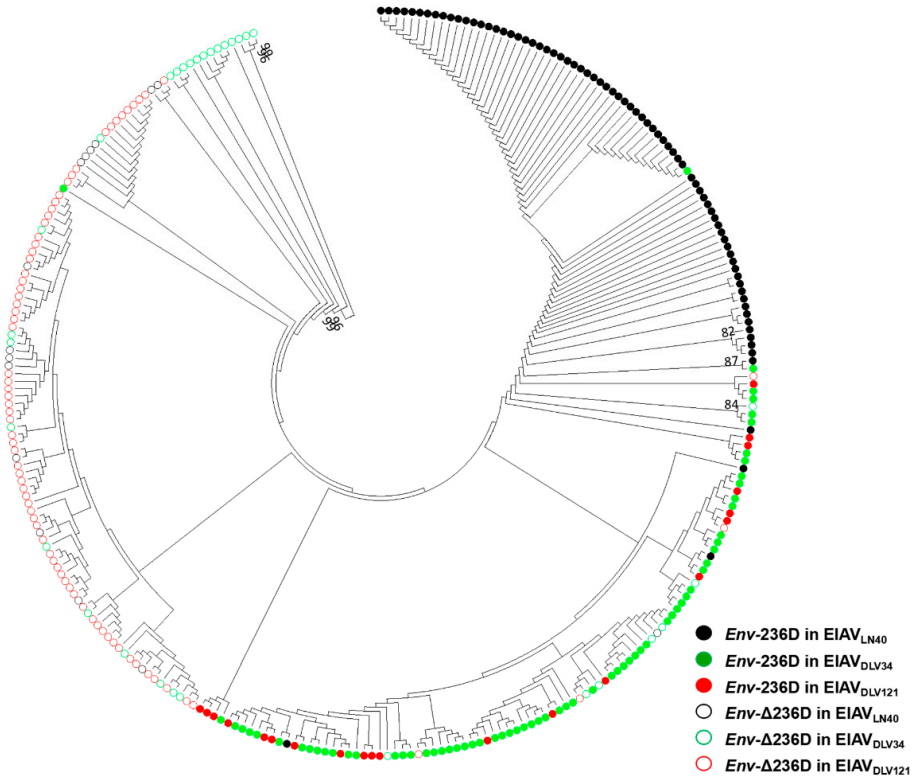

Figure 3. Construction of a maximum likelihood tree using the MEGA 7.0 program based on the Tamura-Nei model. Bootstrap values above 80% are indicated. $EIAV_{LN40}$, $EIAV_{DLV34}$, and $EIAV_{DLV121}$ are shown in black, green and red, respectively. Solid and hollow circles represent the *env*-236D-phenotype and *env*-Δ236D-phenotype sequences, respectively. All samples were collected from three virus strains ($EIAV_{LN40}$, $EIAV_{DLV34}$, and $EIAV_{DLV121}$). A total of 381 sequences were analyzed.

4. Discussion

In contrast to the use of artificial mutagenesis to induce RNA viral attenuation, the EIAV live-attenuated vaccine was developed based on natural evolution through consecutive passaging in vitro. There is growing evidence that analyzing virulence determinants from an evolutionary perspective rather than nucleotide mutations in specific virulence genes is very important, particularly for viruses undergoing consecutive passages in vitro [28,29]. Although a quasispecies is well accepted as a key factor for viral virulence [30], the relationship between these factors remains an open question.

In other words, not all viruses will lose their virulence even through serial passage in vitro, such as the CDV Rockborn strain generated in canine macrophage cells [31].

In our previous study, we observed that the EIAV attenuated vaccine also existed as a complex quasispecies [15]. Furthermore, our results suggested that the attenuated vaccine might be a result driven by selection of the dominant EIAV variant at the whole population level. However, whether a similar evolutionary pattern existed in the predominant virulence gene of EIAV was not clear. Here, we used the EIAV attenuation system to investigate the relationship between virulence and EIAV *env* gene variation. In the present study, we used three different methods to answer this question. Regardless of the method used, some stable mutations located in the V3 and V4 regions of Env were verified. Compared with those of the other methods, SGA showed the highest detection sensitivity for minor nucleotide changes; for instance, the mutations located at positions 235, 236, 237, 246, 248, and 250 could be verified only using the SGA method. In other words, more variants were detected at the same positions. Several key mutations, including S235R, N237K, and N246K, attracted our attention. Not only did the percentages of these mutations parallel with the increased passage number but these mutations (237NNTW240 and N237K/N246K) could also result in loss of an N-glycosylation site or formation of an adjacent N-glycosylation site. We confirmed that this position in combination with three other N-glycosylation site mutations in Env resulted in reversion of virulence and influenced the neutralizing activity [32]. Further analysis showed that inter-linkages existed among these key mutations. Based on these results, we speculated that the observed pattern of Env variation was consistent with the loss of viral virulence.

In particular, we noted that the *env*-Δ236D-phenotype, which could present as a vaccine-specific mutation, was present in its parental virus reservoir ($EIAV_{LN40}$) at a lower level. Hence, our observations indicate that the EIAV vaccine strain (*env*-Δ236D-phenotype) may preexist in the EIAV population as a minority population; indeed, another study from our laboratory detected vaccine-specific LTR sequences in horses infected with virulent strains [15]. According to these results, we proposed a model describing *env* evolution during the attenuation process, as shown in Figure 4. In this model, the *env*-236D-phenotype is the predominant hypervirulent EIAV quasispecies that causes typical EIA symptoms in infected horses and induces 100% mortality. The *env*-Δ236D-phenotype is an avirulent quasispecies that may be a beneficial phenotype for new culture environments (equine macrophages), but the underlying reason is complicated (possibly due to its high fitness or harboring a selectable trait) [6]. To adapt to new culture conditions, the avirulent phenotype selectively accumulated [8,33]. The shape of the phylogenetic tree further reflected this phenomenon in which the *env*-Δ236D-phenotype gradually became predominant within the whole population but progressively lost its virulence. Correspondingly, the *env*-Δ236D-phenotype became dominant with ongoing passages, and thus the strain lost its virulence in animals.

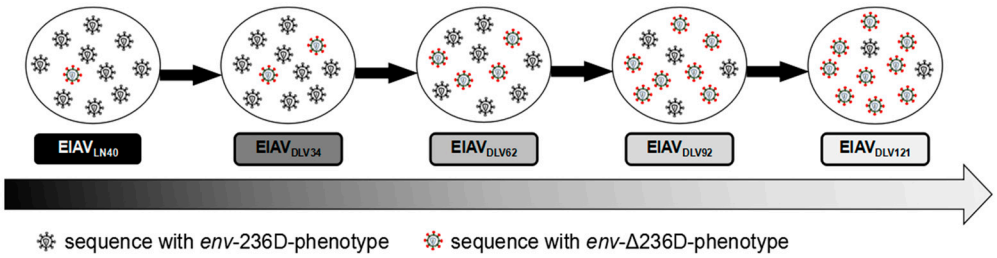

Figure 4. A schematic evolutionary model comparing the EIAV virulent strain to the vaccine strain during long-term passage in vitro. The virulent virus strains are presented with black circles, and the avirulent virus strains are presented with red circles. In this model, the *env*-Δ236D-phenotype sequences preexisted in $EIAV_{LN40}$ as a minority population. With serial passages in vitro, sequences containing the *env*-Δ236D-phenotype gradually became the predominant EIAV quasispecies and generated high replicative fidelity under the new environmental selection conditions. Virulence was simultaneously attenuated as a result, as discussed in the text.

In summary, our study characterized EIAV *env* quasispecies and observed a virulence-correlated parallel in *env* variation during EIAV passage in vitro. However, this strain in the context of the ensemble of the predominant population progressively lost its virulence in vivo. Hence, our data support the view of quasispecies theory and provide a role for the *env* quasispecies in EIAV pathogenesis, which partly explains the attenuation-related mechanism of EIAV.

Supplementary Materials: The following are available online at http://www.mdpi.com/1999-4915/11/4/380/s1, Figure S1: The morbidity of varied EIAV strains. Figure S2: WebLogo presentation of variability in EIAV *env*. Figure S3: Quantification of the *env*-Δ236D-phenotype and *env*-236D-phenotype sequences located in the V4 region of *env* by double-probe real-time PCR.

Author Contributions: Conceived and designed the experiments: Y.L. and X.W. Performed the experiments: C.L., X.-F.W., Y.W., and J.C. Analyzed the data: C.L. Wrote the paper: Z.Z., Y.L. and X.W.

Funding: This research was funded by National Natural Science Foundation of China (31672533, 31302066, and 31672578).

Acknowledgments: We thank Han GZ for technical assistance on evolutionary analysis.

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

References

1. Yuan, Q.; Liu, C.; Liang, Z.; Chen, X.; Diao, D.; Kong, X. The comparison of genetic variation in the envelope protein between various immunodeficiency viruses and equine infectious anemia virus. *Virol. Sin.* **2012**, *27*, 241–247. [CrossRef]
2. Cook, R.F.; Leroux, C.; Issel, C.J. Equine infectious anemia and equine infectious anemia virus in 2013: A review. *Vet. Microbiol.* **2013**, *167*, 181–204. [CrossRef]
3. Ma, J.; Shi, N.; Jiang, C.G.; Lin, Y.Z.; Wang, X.F.; Wang, S.; Lv, X.L.; Zhao, L.P.; Shao, Y.M.; Kong, X.G.; et al. A proviral derivative from a reference attenuated eiav vaccine strain failed to elicit protective immunity. *Virology* **2011**, *410*, 96–106. [CrossRef] [PubMed]
4. Wang, X.; Wang, S.; Lin, Y.; Jiang, C.; Ma, J.; Zhao, L.; Lv, X.; Wang, F.; Shen, R.; Kong, X.; et al. Genomic comparison between attenuated chinese equine infectious anemia virus vaccine strains and their parental virulent strains. *Arch. Virol.* **2011**, *156*, 353–357. [CrossRef]
5. Wang, X.F.; Lin, Y.Z.; Li, Q.; Liu, Q.; Zhao, W.W.; Du, C.; Chen, J.; Wang, X.; Zhou, J.H. Genetic evolution during the development of an attenuated EIAV vaccine. *Retrovirology* **2016**, *13*, 9. [CrossRef]
6. Bingham, R.J.; Dykeman, E.C.; Twarock, R. RNA virus evolution via a quasispecies-based model reveals a drug target with a high barrier to resistance. *Viruses* **2017**, *9*, 347. [CrossRef] [PubMed]
7. Colizzi, E.S.; Hogeweg, P. Evolution of functional diversification within quasispecies. *Genome Biol. Evol.* **2014**, *6*, 1990–2007. [CrossRef]
8. Park, J.M.; Niestemski, L.R.; Deem, M.W. Quasispecies theory for evolution of modularity. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **2015**, *91*, 012714. [CrossRef]
9. Domingo, E.; Sheldon, J.; Perales, C. Viral quasispecies evolution. *Microbiol. Mol. Biol. Rev.* **2012**, *76*, 159–216. [CrossRef]
10. Andino, R.; Domingo, E. Viral quasispecies. *Virology* **2015**, *479–480*, 46–51. [CrossRef] [PubMed]
11. Domingo, E.; Martin, V.; Perales, C.; Grande-Perez, A.; Garcia-Arriaza, J.; Arias, A. Viruses as quasispecies: Biological implications. *Curr. Top. Microbiol.* **2006**, *299*, 51–82.
12. Craigo, J.K.; Ezzelarab, C.; Cook, S.J.; Chong, L.; Horohov, D.; Issel, C.J.; Montelaro, R.C. Envelope determinants of equine lentiviral vaccine protection. *PLoS ONE* **2013**, *8*, e66093. [CrossRef]
13. Du, J.S.; Wang, X.F.; Ma, J.; Wang, J.X.; Qin, Y.Y.; Zhu, C.H.; Liu, F.; Shao, Y.M.; Zhou, J.H.; Qiao, W.T.; et al. Structural and biochemical insights into the v/i505t mutation found in the eiav gp45 vaccine strain. *Retrovirology* **2014**, *11*. [CrossRef]
14. Duan, L.; Du, J.; Wang, X.; Zhou, J.; Wang, X.; Liu, X. Structural and functional characterization of EIAV gp45 fusion peptide proximal region and asparagine-rich layer. *Virology* **2016**, *491*, 64–72. [CrossRef]
15. Wang, X.F.; Liu, Q.; Wang, Y.H.; Wang, S.; Chen, J.; Lin, Y.Z.; Ma, J.; Zhou, J.H.; Wang, X. Characterization of equine infectious anemia virus long terminal repeat quasispecies in vitro and in vivo. *J. Virol.* **2018**, *92*. [CrossRef]

16. Lin, Y.Z.; Shen, R.X.; Zhu, Z.Y.; Deng, X.L.; Cao, X.Z.; Wang, X.F.; Ma, J.; Jiang, C.G.; Zhao, L.P.; Lv, X.L.; et al. An attenuated eiav vaccine strain induces significantly different immune responses from its pathogenic parental strain although with similar in vivo replication pattern. *Antiviral Res.* **2011**, *92*, 292–304. [CrossRef]
17. Jiang, C.G.; Gao, X.; Ma, J.; Lin, Y.Z.; Wang, X.F.; Zhao, L.P.; Hua, Y.P.; Liu, D.; Zhou, J.H. C-terminal truncation of the transmembrane protein of an attenuated lentiviral vaccine alters its in vitro but not in vivo replication and weakens its potential pathogenicity. *Virus Res.* **2011**, *158*, 235–245. [CrossRef]
18. Hayat, M.J.; Higgins, M. Understanding poisson regression. *J. Nurs. Educ.* **2014**, *53*, 207–215. [CrossRef]
19. Kumar, S.; Stecher, G.; Tamura, K. Mega7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef]
20. Craigo, J.K.; Montelaro, R.C. Eiav envelope diversity: Shaping viral persistence and encumbering vaccine efficacy. *Curr. HIV Res.* **2010**, *8*, 81–86. [CrossRef]
21. Etemad, B.; Ghulam-Smith, M.; Gonzalez, O.; White, L.F.; Sagar, M. Single genome amplification and standard bulk pcr yield hiv-1 envelope products with similar genotypic and phenotypic characteristics. *J. Virol. Methods* **2015**, *214*, 46–53. [CrossRef] [PubMed]
22. Wang, X.F.; Wang, S.; Liu, Q.; Lin, Y.Z.; Du, C.; Tang, Y.D.; Na, L.; Wang, X.; Zhou, J.H. A unique evolution of the s2 gene of equine infectious anemia virus in hosts correlated with particular infection statuses. *Viruses* **2014**, *6*, 4265–4279. [CrossRef]
23. Ntziora, F.; Paraskevis, D.; Haida, C.; Magiorkinis, E.; Manesis, E.; Papatheodoridis, G.; Manolakopoulos, S.; Beloukas, A.; Chryssoy, S.; Magiorkinis, G.; et al. Quantitative detection of the m204v hepatitis B virus minor variants by amplification refractory mutation system real-time PCR combined with molecular beacon technology. *J. Clin. Microbiol.* **2009**, *47*, 2544–2550. [CrossRef]
24. Chen, Q.; Belmonte, I.; Buti, M.; Nieto, L.; Garcia-Cehic, D.; Gregori, J.; Perales, C.; Ordeig, L.; Llorens, M.; Soria, M.E.; et al. New real-time-pcr method to identify single point mutations in hepatitis C virus. *World J. Gastroenterol.* **2016**, *22*, 9604–9612. [CrossRef] [PubMed]
25. Salazar-Gonzalez, J.F.; Bailes, E.; Pham, K.T.; Salazar, M.G.; Guffey, M.B.; Keele, B.F.; Derdeyn, C.A.; Farmer, P.; Hunter, E.; Allen, S.; et al. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J. Virol.* **2008**, *82*, 3952–3970. [CrossRef]
26. Geng, Q.M.; Li, H.P.; Bao, Z.Y.; Liu, Y.J.; Zhuang, D.M.; Li, L.; Liu, S.Y.; Li, J.Y. Indinavir resistance evolution in one human immunodeficiency virus type 1 infected patient revealed by single-genome amplification. *Virol. Sin.* **2010**, *25*, 316–328. [CrossRef]
27. Guinoiseau, T.; Moreau, A.; Hohnadel, G.; Ngo-Giang-Huong, N.; Brulard, C.; Vourc'h, P.; Goudeau, A.; Gaudy-Graffin, C. Deep sequencing is an appropriate tool for the selection of unique hepatitis C virus (HCV) variants after single genomic amplification. *PLoS ONE* **2017**, *12*, e0174852. [CrossRef]
28. Presloid, J.B.; Novella, I.S. Rna viruses and rnai: Quasispecies implications for viral escape. *Viruses* **2015**, *7*, 3226–3240. [CrossRef] [PubMed]
29. Liu, F.; Wu, X.; Li, L.; Zou, Y.; Liu, S.; Wang, Z. Evolutionary characteristics of morbilliviruses during serial passages in vitro: Gradual attenuation of virus virulence. *Comp. Immunol. Microbiol. Infect. Dis.* **2016**, *47*, 7–18. [CrossRef] [PubMed]
30. Vignuzzi, M.; Stone, J.K.; Arnold, J.J.; Cameron, C.E.; Andino, R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **2006**, *439*, 344–348. [CrossRef]
31. Martella, V.; Blixenkrone-Moller, M.; Elia, G.; Lucente, M.S.; Cirone, F.; Decaro, N.; Nielsen, L.; Banyai, K.; Carmichael, L.E.; Buonavoglia, C. Lights and shades on an historical vaccine canine distemper virus, the rockborn strain. *Vaccine* **2011**, *29*, 1222–1227. [CrossRef]
32. Han, X.; Zhang, P.; Yu, W.; Xiang, W.; Li, X. Amino acid mutations in the env gp90 protein that modify n-linked glycosylation of the chinese eiav vaccine strain enhance resistance to neutralizing antibodies. *Virus Genes* **2016**, *52*, 814–822. [CrossRef] [PubMed]
33. Glebova, O.; Knyazev, S.; Melnyk, A.; Artyomenko, A.; Khudyakov, Y.; Zelikovsky, A.; Skums, P. Inference of genetic relatedness between viral quasispecies from sequencing data. *BMC Genomics* **2017**, *18*, 918. [CrossRef]