Journal of Animal Science, 2023, **101**, 1–7
https://doi.org/10.1093/jas/skad081
Advance access publication 18 March 2023
Short Communications

Short communication: diversity of endogenous avian leukosis virus subgroup E elements in 11 chicken breeds

Ziyi Wang,[†,1] Yiming Yuan[†,1] Gang Zheng,[†] Meng Sun,[†] Qinyuan Wang,[†] Junfeng Wu,[†] Junying Li,[†] Congjiao Sun,[†] Yongqiang Wang,[‡] Ning Yang,[†] and Ling Lian[†,2]

[†]National Engineering Laboratory for Animal Breeding, Beijing Key Laboratory for Animal Genetic Improvement, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China
[‡]Department of Preventive Veterinary Medicine, College of Veterinary Medicine, China Agricultural University, Beijing, 100193, China

[1]These authors contributed equally to this work.
[2]Corresponding author: lianlinglara@126.com

Abstract

Avian leukosis virus subgroup E (ALVE) as a kind of endogenous retroviruses extensively exists in chicken genome. The insertion of ALVE has some effects on chicken production traits and appearance. Most of the work on ALVEs has been done with commercial breeds. We present here an investigation of ALVE elements in seven Chinese domestic breeds and four standard breeds. Firstly, we established an ALVE insertion site dataset by using the obsERVer pipeline to identify ALVEs from whole-genome sequence data of eleven chicken breeds, seven Chinese domestic breeds, including Beijing You (BY), Dongxiang (DX), Luxi Game (LX), Shouguang (SG), Silkie (SK), Tibetan (TB) and Wenchang (WC), four standard breeds, including White Leghorn (WL), White Plymouth Rock (WR), Cornish (CS), and Rhode Island Red (RIR). A total of 37 ALVE insertion sites were identified and 23 of them were novel. Most of these insertion sites were distributed in intergenic regions and introns. We then used locus-specific PCR to validate the insertion sites in an expanded population with 18~60 individuals in each breed. The results showed that all predicted integration sites in 11 breeds were verified by PCR. Some ALVE insertion sites were breeds specific, and 16 out of 23 novel ALVEs were found in only one Chinese domestic chicken breed. We randomly selected three ALVE insertions including ALVE_CAU005, ALVE_ros127, and ALVE_ros276, and obtained their insertion sequences by long-range PCR and Sanger sequencing. The insertion sequences were all 7525 bp, which were full-length ALVE insertion and all of them were highly homologous to ALVE1 with similarity of 99%. Our study identified the distribution of ALVE in 11 chicken breeds, which expands the current research on ALVE in Chinese domestic breeds.

Lay Summary

Avian leukemia virus subgroup E (ALVE) is an endogenous retrovirus, which is extensively integrated with the chicken genome, and has some effects on chicken production traits and appearance. Most of the current studies on ALVE insertion sites were conducted in standard breeds. In this study, we performed a comprehensive analysis of ALVE insertion sites in seven Chinese domestic breeds and four standard breeds using whole genome sequencing data. A total of 37 ALVE insertion sites were identified and all of them were verified by PCR. Twenty-three of the insertion sites were novel. Some ALVE insertion sites were breeds specific, and 16 out of 23 novel ALVEs were found in only one Chinese domestic chicken breed. In addition, the whole sequences of three ALVE insertions were collected by long-range PCR and Sanger sequencing. We found all the insertion sequences were 7525 bp, which were full-length ALVE insertions and all of them were highly homologous to ALVE1 with similarity of 99%. These results provide a theoretical basis for further studies on the effects of ALVE on production traits and disease resistance traits in chickens.

Distribution of ALVE insertions in the genome of 11 chicken breeds.

Key words: chicken, Chinese domestic breeds, Avian leukosis virus subgroup E

Abbreviations: ALVE, Avian leukosis virus subgroup E; BY, Beijing You chicken; CS, Cornish; DX, Dongxiang chicken; env, envelope gene; ERV, endogenous retrovirus; gag, capsid gene; LTR, long terminal repeats; LX, Luxi Game chicken; NCBI, National Center for Biotechnology Information; RIR, Rhode Island Red; SG, Shouguang chicken; SK, Silkie chicken; TB, Tibetan chicken; TSD, target site duplication; WC, Wenchang chicken; WGS, Whole genome sequencing; WL, White Leghorn; WR, White Plymouth Rock

Introduction

Retrovirus is well known as causative agent of tumorous disease. For vertebrate, retroviruses are a persistent challenge and stress (Doolittle et al., 1989; Patel et al., 2011). Retroviruses can integrate into the host genome during the replication process and they exist in the genome as "provirus" and are transmitted vertically according to Mendelian fashion (Doolittle et al., 1989; Macfarlan et al., 2012; Stoye, 2012; Johnson, 2019). About 3% of the chicken genome is comprised of endogenous retrovirus (ERV) (Mason et al., 2016).

Avian leukosis viruses are divided into 6 subgroups (A, B, C, D, E, and J). The ALV A, B, C, D, and J are exogenous retrovirus, and ALVE is endogenous retrovirus that is integrated into chicken genome (Coffin et al., 1983; Fadly, 1997; Payne and Nair, 2012). The structure of full-length ALVE is similar to exogenous ALV that includes

Received December 15, 2022 Accepted March 16, 2023.

envelope gene (env), capsid gene (gag), polymerase gene, and genes are flanked by the long terminal repeats (5'LTR and 3'LTR) (Bacon et al., 2000). Up to date, more than 322 ALVE integration sites are identified in chickens by next-generation sequencing and 19 ALVE sequences can be accessed in the NCBI (National Center for Biotechnology Information) database (Mason et al., 2020b). Although the full length of ALVE is 7.5 kb, the fragments from different insertion sites may have different lengths. Until now, the insertion sequences in full length were observed from nine insertion sites, including ALVE1, ALVE-B5, ALVE_ros001, ALVE_ros003, ALVE_ros004, ALVE_ros008, ALVE-TYR, ALVE21, and ALVE-NSAC7. Incomplete proviruses sequence insertion has been observed in the 10 integration sites, including ALVE3, ALVE6, ALVE9, ALVE15, ALVE-B9, ALVE-B10, ALVE-B11, ALVE-NSAC1, ALVE_ros005, and ALVE_ros007 (Benkel, 1998; Mason et al., 2020b). For example, ALVE3 is lack of gag- polymerase gene region and ALVE15 has only the LTR region. Some of ALVEs are actively transcribed, such as ALVE3 and ALVE6 in chicken LMH cell lines (Ronfort et al., 1995), but others are silent (Baker et al., 1981).

Previously, the insertions of ALVE into the genomes of their host cells were thought to be a random phenomenon. However, with further research, some researchers found that the insertion sites of ALVEs were not completely random and showed preferred sites of integration. Almost one-half of ALVEs were integrated on chromosome 1 (Tereba, 1983; Mason et al., 2020b). By comparing the sequences of the ALVE insertion site, repeating sequences with six nucleotides as the motif were found at upstream and downstream of each insertion site. Each ALVE insertion site had a unique motif that was known as target site duplication (TSD) (Hishinuma et al., 1981; Benkel, 1998; Mason et al., 2020b). In addition, some ALVE insertion sites prefer to locate around repetitive-sequence elements. For example, ALVE-B10, ALVE-NSAC2, and ALVE16 are located within different CR1 elements, while ALVE21 and ALVE-NSAC3 are located downstream and upstream of CR1 elements, respectively (Benkel, 1998; Smith and Benkel, 2009; Rutherford and Benkel, 2013).

Studies have shown that the insertions of ALVE have some effects on chicken production traits and exterior. A study showed that selection for egg production led to the increased frequencies of ALVE4, ALVE7, and ALVE8, and a decreased frequency of ALVE9. Such ALVE insertion sites might surround the locations of the genes that affect production traits (Kuhnlein et al., 1989). The presence of ALVE10, ALVE19, and ALVE12 was associated with reduction in annual egg production rate, egg weight, and egg-specific gravity. ALVE3 has negative or positive impact on egg production traits in different White Leghorn lines. Although ALVE15 only has LTR, it can impact eggshell color by acting as the promoter of *GRIK2* gene (Gavora et al., 1991; Fulton et al., 2021). In White Plymouth Rock chickens with low and high body weight, the expression of env gene of ALVE was significantly higher in lower body weight for the females (Ka et al., 2009). ALVE are also associated with feather color and feathering traits. ALVE-TYR, the completely avian retroviral sequence, inserts in intron 4 of tyrosinase gene, leading to aberrant transcripts of the exon 5 which results in the recessive white mutation in chickens (Chang et al., 2006). Sex-linked gene K and k + are associated with late feathering and rapid feathering, respectively. ALVE21 and K gene are tightly linked, ALVE21 was integrated into one of two large homologous segments that located on the Z chromosome of late feathering chickens (Bacon et al., 1988; Iraqi and Smith, 1995). Detection of ALVE21 can distinguish late and rapid feathering traits in White Leghorn and some broiler lines (Smith and Levin, 1991; Lu et al., 2009).

Until now, most ALVE works have been done with standard breeds. In China, there are abundant resources of domestic chickens, however, comprehensive investigation for ALVE insertion sites in Chinese domestic chickens are limited. In this study, we used the obsERVer bioinformatic pipeline and locus-specific PCR to detect and verify ALVEs in seven Chinese chicken breeds and four standard breeds.

Materials and Methods

Data source

Whole-genome sequencing (WGS) data of 11 chicken breeds (one individual for each breed) were downloaded from NCBI (accession number: PRJNA232548) and the depth of datasets were 10X. These datasets contained seven Chinese indigenous breeds, including: Beijing You (BY), Dongxiang (DX), Luxi Game (LX), Shouguang (SG), Silkie (SK), Tibetan (TB) and Wenchang (WC), four standard breeds, including White Leghorn (WL), White Plymouth Rock (WR), Cornish (CS), and Rhode Island Red (RIR).

Identification and validation of ALVE integration sites

All sequencing reads were quality checked by FastQC v0.11.9. We removed sequencing adapters and low-quality reads by Trimmomatic v0.39. The ObsERVer bioinformatic pipeline (https://github.com/andrewstephenmason/obsERVer) (Mason et al., 2020b) was used to detect ALVE integration sites from the WGS datasets. Putative ALVE integration sites were inspected in IGV v2.8.10. New ALVEs were named as following format: ALVE_CAU001, in which "CAU" was abbreviation of China Agricultural University, and "001" stood for the first new site identified in this study.

We verified ALVE insertion sites in an expanded population with 18~60 individuals for each breed as follows: BY ($N = 60$), DX ($N = 50$), LX ($N = 50$), SG ($N = 46$), SK ($N = 49$), TB ($N = 52$) WC ($N = 30$), WL ($N = 53$), WR ($N = 50$), CS ($N = 18$), and RIR ($N = 48$).

Total DNA of BY、DX、LX、SG、SK、TB、WC、RIR and WL were extracted from chicken blood. WR were extracted from liver, and CS were extracted from muscle. Locus-specific PCR was performed by three primers (Forward, Reverse, and LTR) for each site (Benkel, 1998). The primers were listed in Supplementary Table S1. The products were analyzed on agarose gels. Individuals with two bands were heterozygous birds, with one shorter band were homozygous ALVE insertion birds, and with one longer band were non-insertion chickens. The volume for PCR was 25 μL, containing 1.1 × T3 Super PCR Mix (Tsingke Biotechnology Co., Ltd., Beijing, China), 10 μM of forward/reverse/LTR primer, and 100 ng of genome DNA. The PCR conditions were as follows: 2 min of denaturation at 98 °C, 35 cycles for 10 s of denaturation at 98 °C, 10 s of annealing, and 20s for extension at 72 °C, followed by an extra 2 min of extension at 72 °C. Electrophoresis on 2% agarose gel was performed to distinguish ALVE insertion.

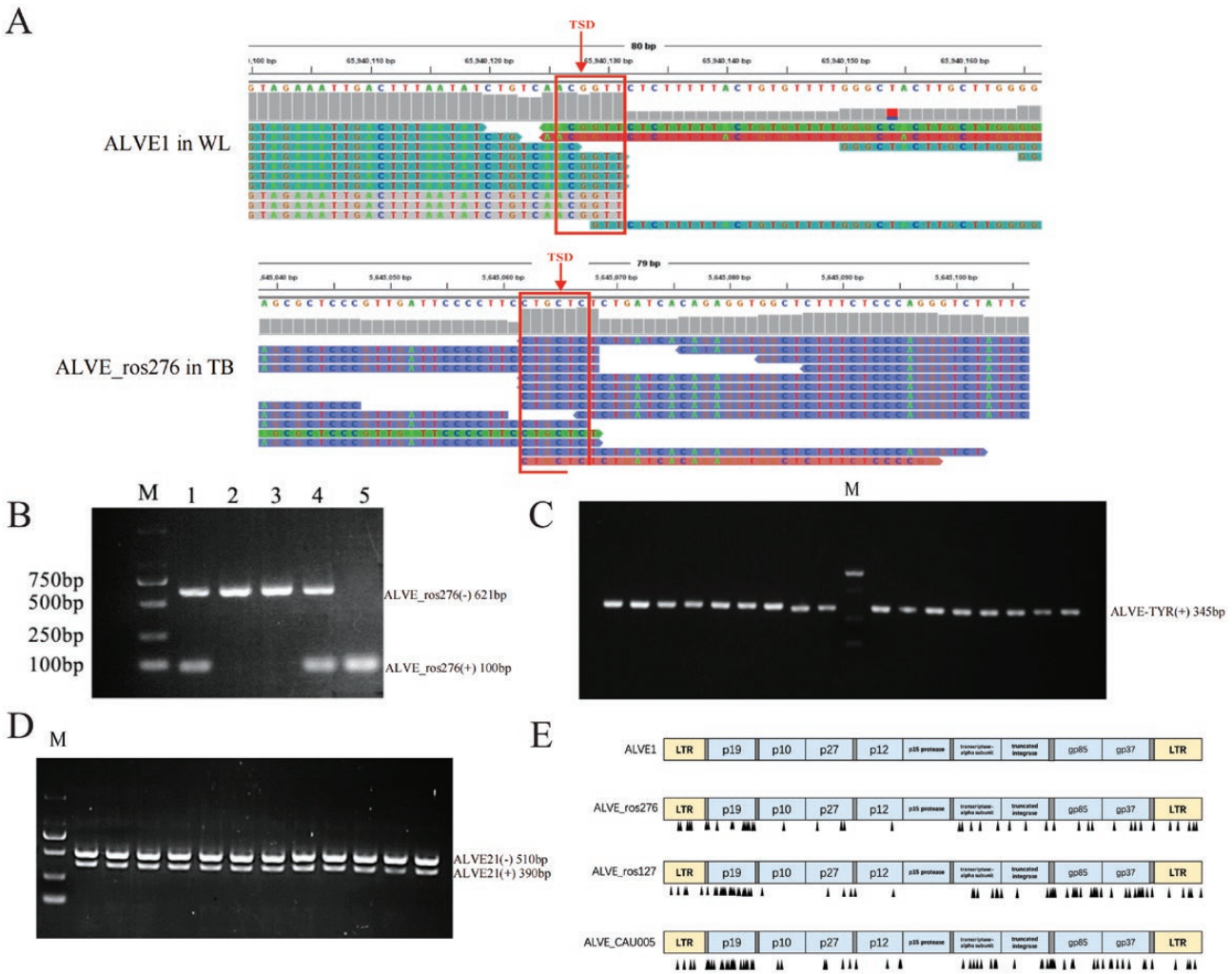

Figure. 1 IGV alignment view of ALVE1 and ALVE_ros276 in White Leghorn and Tibetan, respectively(A). TSD: Target site duplication. (B) PCR diagnostic for ALVE_ros276. M: Marker. (–) 621 bp, (+) 100 bp. lane 1 and 4 was a heterozygous insertion, lane 2 and 3 was no insertion, and lane 5 was a homozygous insertion. (C) PCR diagnostic for ALVE TYR in Silkies, Cornish, and White Plymouth Rock. (+) 345 bp, all lanes were homozygous insertion. (D) The insertion of ALVE21 in White Leghorn. (+) 390 bp, (–) 510 bp, all lanes were heterozygous insertion. (E) The diagram for comparison of sequences from ALVE_ros276, ALVE_ros127, and ALVE_CAU005 with that from ALVE1. SNPs were marked by arrows. The intergenics represented non-coding regions.

Sequencing of PCR products for ALVE insertion fragments from three ALVE integrations

Insertion sequences from ALVE_ros276, ALVE_ros127, and ALVE_CAU005 were amplified by long-range PCR followed by the standard of Takara PrimeSTAR Max DNA Polymerase protocol. PCR was conducted in 50 μL reaction volumes. Each PCR reaction began with activation at 98 °C for 1 min, then 35 cycles of 10 s denaturing at 98 °C, 4 min extension at 68 °C, and then finished with a 4 min final extension at 68 °C. The primers were listed in Supplementary Table S2. PCR products were sequenced at Beijing Genomics institution (Beijing, China). The sequences were aligned using the BLAST toolkit in the NCBI database.

Results

Identification of ALVE integration sites in 11 chicken populations

We used the obsERVer pipeline reported by Mason et al. (2020b) and IGV v2.8.10 to detect the integration of ALVE in the WGS datasets. There were obvious TSD at ALVE integration sites when inspecting the result BAM files with IGV (Figure 1A). We identified a total of 37 different ALVE in the 11 breeds studies, of which 23 had not been previously reported. The chromosome 1, 2, and 3 were the top three chromosomes with most ALVE insertions, and there were 14, 9, and 11 integration sites in Chr1, Chr2, and Chr3, respectively. Each ALVE site has unique motif of TSD (Table 1). We analyzed surrounding or overlapped genes of 37 ALVE sites based on Gallus gallus 6.0 reference genome (GRCg6a, GenBank assembly accession: GCA_000 002315.5). The results showed that 20 ALVEs were inserted in introns, 16 ALVEs were inserted in intergenic regions, and ALVE_ros276 was detected in 3'-UTR of *NT5C1A* gene.

The verification of ALVE integration sites by locus-specific PCR

The locus-specific PCR for 37 ALVEs were performed to distinguish heterozygous, homozygous, and non-inserted individuals by bands. Take ALVE_ros276 as an example, PCR products with only a 100 bp band were homozygous,

Table 1 ALVEs of 11 chicken breeds detected by the obsERVer pipeline

Breed	Name	Location	Hexamer	Overlapped genes
BY	ALVE_CAU001	3:32,243,447	CTCAAT	NO
	ALVE_CAU002	3:89,260,091	GCTAGT	CSMD1(intron1)
	ALVE_ros273	20:3,074,677	GCCCAC	PTPRT (intron7)
	ALVE_ros276	23:5,645,068	CTGCTC	NT5C1A(3'UTR)
CS	ALVE-TYR	1:189,153,674	ACACTG	TYR(intron4)
DX	ALVE-B5	1:10,787,714	GGTGGT	LOC112532986(intron3)
	ALVE_CAU003	2:71,978,681	GAGGAG	NO
	ALVE_ros127	2:147,349,291	GGAGGC	TSNARE1(intron2)
	ALVE_CAU004	4:54,377,551	GGGGAC	LOC112532303(intron5)
	ALVE_CAU005	6:7,270,339	GGCAGT	PCDH15(intron27)
LX	ALVE_ros100	2:54,009,821	GTAGGC	CNTNAP2(intron14)
	ALVE_CAU006	2:104,963,623	GATGCT	NO
	ALVE_CAU007	2:113,949,658	GCAAAG	NO
	ALVE_CAU008	3:11,026,451	CTCACT	NO
	ALVE_CAU009	3:43,153,318	TTCTCT	PDE10A(intron1)
	ALVE_CAU010	3:76,499,348	CAAGTC	NO
	ALVE_ros276	23:5,645,068	CTGCTC	NT5C1A(3'UTR)
RIR	ALVE-B5	1:10,787,714	GGTGGT	LOC112532986(intron3)
	ALVE_ros001	1:102,980,011	GTTGTG	NO
	ALVE_ros004	2:123,787,338	CTTGAC	NO
SG	ALVE_CAU011	1:39,162,579	TTCAGC	LOC101749434(intron2)
	ALVE_CAU012	1:65,996,184	CAGCGC	SOX5(intron3)
	ALVE_CAU013	1:182,903,121	ACTAAA	NO
	ALVE_CAU014	3:950,253	ACACAC	FANCL(intron7)
	ALVE_CAU015	8:10,388,286	CCCCAC	NO
SK	ALVE_CAU016	1:102,187,653	ATCTAC	NCAM2(intron1)
	ALVE_CAU017	1:143,035,486	AGTATT	NO
	ALVE-TYR	1:189,153,674	ACACTG	TYR(intron4)
	ALVE_CAU018	3:110,009,243	GTCTGC	CD2AP(intron2)
	ALVE_CAU019	4:42,984,321	GAAATC	GALNTL6(intron4)
	ALVE_CAU020	Z:304,564	ATGAAT	NO
TB	ALVE_CAU021	1:129,602,884	GACCTA	NO
	ALVE_ros240	9:12,295,680	TGTAAA	LOC112533006(intron4)
	ALVE_ros276	23:5,645,068	CTGCTC	NT5C1A(3'UTR)
WC	ALVE_NSAC5	3:73,083,966	GGCTGA	NO
	ALVE_CAU022	9:12,011,455	AGCCTA	NO
	ALVE21	Z:10,681,604	GGGTAG	NO
WL	ALVE1	1:65,940,132	ACGGTT	SOX5(intron2)
	ALVE21	Z:10,681,604	GGGTAG	NO
WR	ALVE-B5	1:10,787,714	GGTGGT	LOC112532986(intron3)
	ALVE-TYR	1:189,153,674	ACACTG	TYR(intron4)
	ALVE_CAU023	5:21,052,270	CCGAGT	LRRC4C(intron1)
	ALVE-B1	5:23,474,345	GTTATT	C5H11orf49(intron3)
	ALVE_ros220	7:14,856,972	AGTTAT	ZNF385B(intron2)
	ALVE_CAU022	9:12,011,455	AGCCTA	NO

those with both 100 and 621 bp bands were heterozygous, and those with only a 621 bp band were non-insertion (Figure 1B). All the predicted integration sites in 11 breeds were 100% verified by PCR. PCR results of 23 novel ALVEs were shown in Supplementary Figure 1. We analyzed the distribution of 37 ALVEs in 11 breeds (Table 2). The results showed that some of the ALVE integration sites were breed specific, such as ALVE_CAU001 and ALVE_CAU002, were only found in Beijing You and ALVE_CAU007, ALVE_CAU008, ALVE_CAU009 were only detected in Luxi Game. Some of the ALVE insertion sites, including ALVE_CAU006, ALVE_CAU022, ALVE_CAU023, ALVE_ros276, and ALVE_ros100

Table 2 The results of ALVE insertions in 11 breeds verified by PCR

ALVE	BY	CS	DX	LX	RIR	SG	SK	TB	WL	WR	WC
ALVE_CAU001	√										
ALVE_CAU002	√										
ALVE_CAU003			√								
ALVE_CAU004			√								
ALVE_CAU005			√								
ALVE_CAU006	√	√	√	√	√	√	√	√	√	√	√
ALVE_CAU007				√							√
ALVE_CAU008				√							
ALVE_CAU009				√							
ALVE_CAU010				√							
ALVE_CAU011						√					
ALVE_CAU012						√					
ALVE_CAU013						√					
ALVE_CAU014						√					
ALVE_CAU015						√					√
ALVE_CAU016							√				
ALVE_CAU017							√				
ALVE_CAU018							√				
ALVE_CAU019	√	√	√				√				
ALVE_CAU020							√				
ALVE_CAU021								√			√
ALVE_CAU022	√	√	√	√	√		√	√		√	√
ALVE_CAU023	√				√	√			√	√	√
ALVE-B1										√	
ALVE_NSAC5						√					√
ALVE_ros001					√						
ALVE_ros004					√						
ALVE_ros100				√			√	√			
ALVE_ros127			√								
ALVE_ros220										√	
ALVE_ros240								√			
ALVE_ros273	√										
ALVE _ros276	√					√		√	√		
ALVE1									√		
ALVE21									√		
ALVE-B5		√	√		√	√			√		
ALVE-TYR		√					√			√	

were more common, and these ALVE insertion sites were found in at least three breeds.

ALVE-TYR insertion was found in all individuals of SK (N = 49), CS (N = 18), and WR (N = 50). All three breeds were white plumage, and all individuals were homozygous insertion (Figure 1C). Heterozygous ALVE21 integration was found in all WL chickens (N = 53) (Figure 1D).

Inserted sequence analysis from ALVE_ros127, ALVE_CAU005, and ALVE_ros276

The complete sequence of ALVE_ros127, ALVE_CAU005, and ALVE_ros276 were amplified and sequenced by Sanger sequencing. The full length of ALVE_ros127, ALVE_CAU005, and ALVE_ros276 were first reported in this study, and they were full-length ALVE insertions with 7525 bp. The sequence of ALVE_ros127, ALVE_CAU005, and ALVE_ros276 had been uploaded to the NCBI database, and the GenBank accession number are ON063488, ON063489, and ON063490, respectively. The sequences of ALVE_CAU005, ALVE_ros276, and ALVE_ros127 were highly homologous to ALVE1 with similarity of 99% (Figure 1E).

Discussion

ALVEs could be identified by locus-specific PCR (Benkel, 1998). Although locus-specific PCR detection saves costs, it is difficult to find new ALVE integration sites. To date, next-generation sequencing can be easily accessed in public databases, which allows investigation insertion sites by bioinformatic methods. In this study, we used the obsERVer pipeline to detect ALVE integrations from the WGS data. All the

integration sites could be successfully verified by locus-specific PCR, which demonstrated the obsERVer was a reliable and accurate pipeline for detecting comprehensive ALVE insertions in WGS datasets. However, the obsERVer also had a limitation to identify some insertion sites, such as ALVE6 (Mason et al., 2020a, 2020b), because ALVE6 is located near the chromosome 1 p arm telomere which was a poorly assembled region (Benkel and Rutherford, 2014; Mason et al., 2020a). In our study, we did not identify ALVE6 from the WGS data. However, through locus-specific PCR, we found that ALVE6 was inserted in many WL individuals. Some studies also reported that ALVE6 commonly exists in standard chicken lines, such as WL and Brown Leghorn (Benkel, 1998; Mason et al., 2020a). Besides the obsERVer, a software tool called Vermillion, which was published in 2016 by Rutherford et al., can also be used to target ALVE from next-generation sequencing of chicken (Rutherford et al., 2016).

It has been reported that ALVE insertions were associated with some exterior phenotypes. For example, ALVE-TYR insertion was causative mutation for recessive white. In this study, we found ALVE-TYR insertion existed in all individuals of SK, CS, and WR, which are consistence with their recessive white trait. In some chicken populations, ALVE21 was previously regarded as potential causation for late feathering because it was inserted in K gene which was tightly linked with late feathering. However, there are some exceptions, some studies reported that ALVE21 can be detected in both late feathering and rapid feathering RIR layers. Smith and Fadly (1988) reported that no ALVE21 was detected in few late feathering White Leghorn chickens (Smith and Fadly, 1988; Boulliou et al., 1992; Takenouchi et al., 2018). Therefore, the causative mutation of rapid and late feathering has not been determined clearly yet. In this study, we found that ALVE21 existed in all WL, which were consistent with their late feathering phenotypes at one-day old. It suggests that ALVE21 was still a molecular marker to distinguish late and fast feathering phenotypes in WL population that we used.

Collectively, we used the WGS data to identify ALVE insertion sites within 11 chicken breeds. We subsequently expanded the population and verified these insertions using locus-specific PCR. Our identification of 23 new ALVEs extended the current research on ALVEs in Chinese domestic breeds.

Supplementary Data

Supplementary data are available at *Journal of Animal Science* online.

Acknowledgments

This research was supported by National Key Research and Development Program of China (2022YFF1000200 and 2021YFD1300600), National Natural Science Foundation of China (32272865 and U1901206), China Agricultural University Basic scientific research (2022TC150), National Germplasm Center of Domestic Animal Resources, Agriculture Research Systems (CARS-40), the Young Scientist Supporting Project, Program for Changjiang Scholars and Innovative Research Team in University (IRT_15R62), the Beijing Key Laboratory for Animal Genetic Improvement. We thank Dr. Mason for his technical assistant for running the obsERVer pipeline.

Conflict of interest statement

The authors declare that they have no competing interests.

Literature Cited

Bacon, L. D., H. D. Hunt, and H. H. Cheng. 2000. A review of the development of chicken lines to resolve genes determining resistance to diseases. *Poult. Sci*. 79:1082–1093. doi: 10.1093/ps/79.8.1082

Bacon, L. D., E. Smith, L. B. Crittenden, and G. B. Havenstein. 1988. Association of the slow feathering-(K) and an endogenous viral (Ev21) gene on the Z-chromosome of chickens. *Poult. Sci*. 67(2):191–197. doi: 10.3382/ps.0670191

Baker, B., H. Robison, H. E. Varmus, and J. M. Bishop. 1981. Analysis of endogenous avian retrovirus DNA and RNA: viral and cellular determinants of retrovirus gene expression. *Virology*. 114:8–22. doi: 10.1016/0042-6822(81)90248-8

Benkel, B. F. 1998. Locus-specific diagnostic tests for endogenous avian leukosis-type viral loci in chickens. *Poult. Sci*. 77:1027–1035. doi: 10.1093/ps/77.7.1027

Benkel, B., and K. Rutherford. 2014. Endogenous avian leukosis viral loci in the Red Jungle Fowl genome assembly. *Poult. Sci*. 93:2988–2990. doi: 10.3382/ps.2014-04309

Boulliou, A., J. P. Lepennec, G. Hubert, R. Donal, and M. Smiley. 1992. The endogenous retroviral Ev21 locus in commercial chicken lines and its relationship with the slow-feathering phenotype (K). *Poult. Sci*. 71(1):38–46. doi: 10.3382/ps.0710038

Chang, C. M., J. L. Coville, G. Coquerelle, D. Gourichon, A. Oulmouden, and M. Tixier-Boichard. 2006. Complete association between a retroviral insertion in the tyrosinase gene and the recessive white mutation in chickens. *BMC Genom*. 7. doi: 10.1186/1471-2164-7-19

Coffin, J. M., P. N. Tsichlis, K. F. Conklin, A. Senior, and H. L. Robinson. 1983. Genomes of endogenous and exogenous avian retroviruses. *Virology*. 126:51–72. doi: 10.1016/0042-6822(83)90461-0.

Doolittle, R. F., D. F. Feng, M. S. Johnson, and M. A. Mcclure. 1989. Origins and evolutionary relationships of retroviruses. *Q. Rev. Biol*. 64:1–30. doi: 10.1086/416128

Fadly, A. M. 1997. Avian retroviruses. *Vet. Clin. North Am.-Food Anim. Pract*. 13(1):71–85. doi: 10.1016/s0749-0720(15)30365-0

Fulton, J. E., A. S. Mason, A. Wolc, J. Arango, P. Settar, A. R. Lund, and D. W. Burt. 2021. The impact of endogenous Avian Leukosis Viruses (ALVE) on production traits in elite layer lines. *Poult. Sci*. 100:101121. doi: 10.1016/j.psj.2021.101121

Gavora, J. S., U. Kuhnlein, L. B. Crittenden, J. L. Spencer, and M. P. Sabour. 1991. Endogenous viral genes - association with reduced egg-production rate and egg size in white leghorns. *Poult. Sci*. 70(3):618–623. doi: 10.3382/ps.0700618

Hishinuma, F., P. J. Debona, S. Astrin, and A. M. Skalka. 1981. Nucleotide-sequence of acceptor site and termini of integrated avian endogenous provirus Ev1 - integration creates a 6-Bp repeat of host DNA. *Cell*. 23:155–164. doi: 10.1016/0092-8674(81)90280-4

Iraqi, F., and E. J. Smith. 1995. Organization of the sex-linked late-feathering haplotype in chickens. *Anim. Genet*. 26:141–146. doi: 10.1111/j.1365-2052.1995.tb03153.x

Johnson, W. E. 2019. Origins and evolutionary consequences of ancient endogenous retroviruses. *Nat. Rev. Microbiol*. 17:355–370. doi: 10.1038/s41579-019-0189-2

Ka, S., S. Kerje, L. Bornold, U. Liljegren, P. B. Siegel, L. Andersson, and F. Hallbook. 2009. Proviral integrations and expression of endogenous Avian leucosis virus during long term selection for high and low body weight in two chicken lines. *Retrovirology*. 6:(1): 1–13. doi: 10.1186/1742-4690-6-68

Kuhnlein, U., M. Sabour, J. S. Gavora, R. W. Fairfull, and D. E. Bernon. 1989. Influence of selection for egg-production and Mareks-disease resistance on the incidence of endogenous viral genes in white leghorns. *Poult. Sci*. 68:1161–1167. doi: 10.3382/ps.0681161

Lu, X. Q., J. R. Han, X. F. Liu, T. H. Lin, Y. L. Li, X. X. Hu, and N. Li. 2009. The LTR of endogenous retrovirus ev21 retains promoter

activity and exhibits tissue specific transcription in chicken. *Chin. Sci. Bull*. 54:4664–4670. doi: 10.1007/s11434-009-0547-y

Macfarlan, T. S., W. D. Gifford, S. Driscoll, K. Lettieri, H. M. Rowe, D. Bonanomi, A. Firth, O. Singer, D. Trono, and S. L. Pfaff. 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature*. 487(7405):57. doi: 10.1038/nature11244

Mason, A. S., J. E. Fulton, P. M. Hocking, and D. W. Burt. 2016. A new look at the LTR retrotransposon content of the chicken genome. *BMC Genom*. 17. doi: 10.1186/s12864-016-3043-1

Mason, A. S., J. E. Fulton, and J. Smith. 2020a. Endogenous avian leukosis virus subgroup E elements of the chicken reference genome. *Poult. Sci*. 99:2911–2915. doi: 10.1016/j.psj.2019.12.074

Mason, A. S., A. R. Lund, P. M. Hocking, J. E. Fulton, and D. W. Burt. 2020b. Identification and characterisation of endogenous Avian Leukosis Virus subgroup E (ALVE) insertions in chicken whole genome sequencing data. *Mobile DNA*. 11(1):1–13. doi: 10.1186/s13100-020-00216-w

Patel, M. R., M. Emerman, and H. S. Malik. 2011. Paleovirology - ghosts and gifts of viruses past. *Curr. Opin. Virol*. 1:304–309. doi: 10.1016/j.coviro.2011.06.007

Payne, L. N., and V. Nair. 2012. The long view: 40 years of avian leukosis research. *Avian Pathol*. 41:11–19. doi: 10.1080/03079457.2011.646237

Ronfort, C., Y. Chebloune, F. L. Cosset, C. Faure, V. M. Nigon, and G. Verdier. 1995. Structure and expression of endogenous retroviral sequences in the permanent Lmh chicken-cell line. *Poult. Sci*. 74(1):127–135. doi: 10.3382/ps.0740127

Rutherford, K., and B. F. Benkel. 2013. Characterization of insertion sites and development of locus-specific assays for three broiler-derived subgroup E avian leukosis virus proviruses. *Avian Pathol*. 42:373–378. doi: 10.1080/03079457.2013.809694

Rutherford, K., C. J. Meehan, M. G. I. Langille, S. G. Tyack, J. C. McKay, N. L. McLean, K. Benkel, R. G. Beiko, and B. Benkel. 2016. Discovery of an expanded set of avian leukosis subroup E proviruses in chickens using Vermillion, a novel sequence capture and analysis pipeline. *Poult. Sci*. 95:2250–2258. doi: 10.3382/ps/pew194

Smith, A., and B. F. Benkel. 2009. Novel avian leukosis virus-related endogenous proviruses from layer chickens: characterization and development of locus-specific assays. *Poult. Sci*. 88:1580–1585. doi: 10.3382/ps.2009-00148

Smith, E. J., and A. M. Fadly. 1988. Influence of congenital transmission of endogenous virus-21 on the immune-response to avian-leukosis virus-infection and the incidence of tumors in chickens. *Poult. Sci*. 67(12):1674–1679. doi: 10.3382/ps.0671674

Smith, E. J., and I. Levin. 1991. Application of a locus-specific DNA hybridization probe in the analysis of the slow-feathering endogenous virus complex of chickens. *Poult. Sci*. 70(9):1957–1964. doi: 10.3382/ps.0701957

Stoye, J. P. 2012. Studies of endogenous retroviruses reveal a continuing evolutionary saga. *Nat. Rev. Microbiol*. 10:395–406. doi: 10.1038/nrmicro2783

Takenouchi, A., M. Toshishige, N. Ito, and M. Tsudzuki. 2018. Endogenous viral gene ev21 is not responsible for the expression of late feathering in chickens. *Poult. Sci*. 97:403–411. doi: 10.3382/ps/pex345

Tereba, A. 1983. Asymmetric chromosomal distribution of endogenous retrovirus loci in chickens and mice. *Curr. Top. Microbiol. Immunol*. 107:29–50. doi: 10.1007/978-3-642-69075-4_2