ORIGINAL ARTICLE

Veterinary Research Forum. 2023; 14 (10) 531 - 539

doi: 10.30466/vrf.2022.552748.3454

Journal Homepage: vrf.iranjournals.ir

Veterinary Research Forum

Differences in CpG island distribution between exogenous and endogenous jaagsiekte sheep retrovirus strains

Qiang Liu*, Yue-Yang Yu, Huai-Yu Wang

Department of Agricultural Science and Technology, Nanchong Vocational and Technical College, Nanchong, China.

Article Info

Article history:

Received: 28 April 2022
Accepted: 16 October 2022
Available online: 15 October 2023

Keywords:

CpG islands
env
gag
Jaagsiekte sheep retrovirus
LTR

Abstract

The jaagsiekte sheep retrovirus (JSRV), belonging to the betaretrovirus genus of the retroviridae family, includes both exogenous and endogenous jaagsiekte sheep retroviruses (exJSRV and enJSRV, respectively). At the proviral genome level, exJSRV and enJSRV strains have a high degree of similarity with their main variation regions being the LTR, *gag*, and *env* genes. In this study, for the first time, we investigated and compared the distribution of CpG islands between these enJSRV and exJSRV strains. Specifically, we analyzed a total of 42 full-length JSRV genomic sequences obtained from the GenBank® database to identify CpG islands in the exJSRV and enJSRV genomes using the MethPrimer software. Our results showed that the CpG islands in the two JSRV strains were mainly distributed in the LTR, *gag*, and *env* genes. In exJSRVs, 66.66% (6/9), 33.33% (3/9), and 100% (9/9) of the sequences presented at least one CpG island in LTR, *gag*, *env* genes, respectively, and for enJSRVs, 84.84% (28/33), 57.57% (19/33), and 96.96% (32/33) of the sequences presented at least one CpG island in the LTR, *gag*, and *env* genes. These findings suggested that the distribution, length, and genetic traits of CpG islands were different for the exJSRV and enJSRV strains. In future, it would be necessary to demonstrate the biological significance of CpG islands within these genes in exJSRV and enJSRV genomes. This will enhance understanding regarding the potential role of CpG islands in epigenetic regulation.

© 2023 Urmia University. All rights reserved.

Introduction

Ovine pulmonary adenocarcinoma (OPA) is a chronic, progressive and contagious pulmonary tumor disease in sheep caused by jaagsiekte sheep retrovirus (JSRV).[1] The disease is characterized by dyspnea, cough, emaciation, the discharge of large amounts of serous nasal fluid and neoplastic hyperplasia of type II alveolar epithelial cells and non-ciliated bronchiolar epithelial cells.[1] It has also been observed in goats[2] and was first reported in South Africa in the 19th century.[1,3] Subsequently, it was reported in the United Kingdom, Germany, France, the United States, Mexico, Canada, Ireland and other countries.[1] Presently, OPA exists in almost all countries and regions with a developed sheep industry except for Australia, New Zealand, and Iceland.[1,4,5] Further, in China, OPA was first reported in 1951, with Xinjiang and Inner Mongolia being the most affected endemic areas.[6-8]

On the one hand, exogenous JSRV (exJSRV), which is an etiological agent for naturally occurring and experimentally induced OPA in sheep, belongs to the betaretrovirus genus of the family retroviridae and is phylogenetically related to the enzootic nasal tumor virus (ENTV), the causative agent of enzootic nasal adenocarcinoma, a contagious tumor of the mucosal nasal glands in sheep and goats.[1] On the other hand, endogenous jaagsiekte sheep retrovirus (enJSRV), highly related to exJSRV/ENTV and is present in sheep and goats, is derived from the integration of exJSRV in host genomes followed by genetic stabilization via the accumulation of mutations.[8] Both exJSRV and enJSRV have a similar gene structure comprising a linear, single- and positive-stranded RNA with a total length of approximately 7,500 bp.[4,9,10] The RNA-coding regions of these strains is mainly composed of four overlapping viral structural proteins encoded by the *gag*, *pro*, *pol*, and *env* genes.[4,9,10] Further, at the 3′ end of the *pol* gene there is a special open reading frame (ORF), the overlapping region ORF-X, which is approximately 500 bp long.[4,9,10] Additionally, two repetitive (R) regions are located at both ends of the genome[4,9,10] and the 5′ and 3′ ends of the R

***Correspondence:**

Qiang Liu. DSc
Department of Agricultural Science and Technology, Nanchong Vocational and Technical College, Nanchong, China
E-mail: liuqiang_yyy@163.com

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) which allows users to read, copy, distribute and make derivative works for non-commercial purposes from the material, as long as the author of the original work is cited properly.

region are the U3 region (U3-R) and U5 region (R-U5), respectively, both of which constitute the non-coding region.[4,9,10] When JSRV enters susceptible ovine cells such as bronchial epithelial cells, under the action of virus-encoded reverse transcriptase, its RNA is used as a template for DNA synthesis.[4,9,10] The synthesized DNA is then integrated into the infected ovine cells causing the rapid proliferation of ovine Type II alveolar epithelial cells and non-ciliated bronchiolar epithelial cells resulting in the formation of small adenomas.[4,9,10] Additionally, JSRV exists temporarily or long-term in the sheep genome in the form of proviral DNA "*LTR-gag-pro-pol(orf-x)-env-LTR*".[9,10]

Studies have confirmed that some retroviruses including avian myeloblastosis virus, Rous sarcoma virus, murine leukemia virus, bovine leukemia virus and human T cell lymphotropic virus have tumorigenic effects in humans and animals.[11-15] Specifically, a study showed that JSRV has a tumorigenic effect and that JSRV envelope (Env) protein plays a similar viral oncogenic role.[16,17] It has also been reported that three signal transduction pathways are mainly involved in cell transformation mediated by JSRV Env.[18-20] Additionally, approximately 30 copies of enJSRV sequences that be closely related to exJSRV are present in all studied normal sheep and goat genomes.[21] It has also been suggested that enJSRV plays a crucial role in pregnancy and placenta formation and may be involved in local immune regulation during pregnancy mediating cell-to-cell interactions to prevent fetal infection with exJSRV.[22,23]

The DNA methylation is induced by DNA methyl-transferase which transfers a methyl group from S-adenosylmethionine to cytosine position 5 thereby forming 5-methylcytosine which is involved in gene expression regulation and cell differentiation.[24,25] Further, DNA methylation sites predominantly include CpG islands or regions rich in CpG islands[26,27] which are primarily located in the gene promoter region and partly in the first exon. It has also been observed that CpG island methylation can directly lead to the epigenetic silencing of related genes.[26] This regulatory effect mainly represses gene expression at the transcriptional level.[28] Also, several studies have confirmed that DNA methylation is involved in the occurrence and development of tumors, e.g., the methylation status of tumor suppressor genes can lead to tumor suppressor gene inactivation.[29] However, it is still unclear whether there is methylation in the JSRV genome especially in the *env* gene that may contribute to JSRV pathogenicity such as tumorigenicity. Therefore, the aim of this study was to analyze the number, distribution and length of CpG islands within and between exJSRV and enJSRV provirus sequences. The results obtained may help clarify the viral characteristics of these two JSRV strains from an epigenetic perspective.

Materials and Methods

Sequence sources. The nucleotide sequences of exJSRV and enJSRV strain genomes were searched up to March 09,2022 in GenBank® database of National Center for Bio-technology Information, using "Jaagsiekte sheep retrovirus" and "complete genome" as search terms. Complete genomes with sizes in the range 2,494 - 7,989 bp were included.

CpG island analysis. The identified genomic sequences of exJSRV and enJSRV were then analyzed as shown in Table 1. Further, MethPrimer software (http:// www.urogene.org/cgi-bin/methprimer/methprimer.cgi) was used to the locate CpG islands in the gene sequences of the JSRV strains. Specifically, the CpG islands were located by measuring the GC contents and observed-to-expected CpG dinucleotide ratios of the strains over a sequence window. The standard settings for CpG islands were: sequence window length ≥ 100 bases, GC content ≥ 50.00% and observed-to-expected CpG ratio ≥ 60.00%.

Results

Phylogenetic analysis based on the complete genome sequences of exJSRV and enJSRV strains. Phylogenetic analysis of known exJSRVs and enJSRVs (Table 1) with sequences in the GenBank® database was performed using the neighbor-joining method in MEGA Software (version 5.0; Biodesign Institute, Tempe, USA). In the phylogenetic tree constructed using complete genome sequences, all the exJSRV strains (except strains KX817296 and DQ838494) were clustered in an independent clade (Fig. 1), while all the enJSRV strains predominantly clustered into two clades. Among the enJSRVs strains, the EF680311, EF680309, EF680305, EF680319, EF680307, EF680315, EF680312, and EF680314 strains were clustered into one clade (Clade A, which represents the oldest provirus clade as previously reported),[21] while the DQ838493, AF136224, AF136225, AF153615, MF175067, MF175068, MF175069, MF175070, MF175071, EF680310, EF680317, EF680298, EF680306, EF680316, EF680303, EF680296, EF680308, EF680299, EF680300, EF680301, EF680304, EF680302, EF680313, EF680318, and EF680297 strains clustered into a different clade (Clade B as reported in a previous study) in the phylogenetic tree based on the complete genome sequences of known enJSRVs.[21] Further, strains KX817296 and DQ838494 which were recognized as exogenous viral strains in a previous study, shared the closest genetic relationship with enJSRV strains in Clades A and B, respectively.[6,21] Additionally, the AF136225 cluster in an independent clade in the phylogenetic tree was constructed using complete genome sequences.

Characteristics of CpG islands in exJSRV strains. In this study, the number, location, length and distribution characteristics of the CpG islands of the proviruses of different exJSRV strains were determined using the MethPrimer software. The results obtained in this regard showed that the A7 provirus of the exJSRV strain had four CpG islands located in the LTR (1), *gag* (2), and *pol*/*env* (1) genes, with lengths 104, 114, 108 and 170 bp, respectively. The A2 and A5 proviruses each had three CpG islands, each located in the LTR, *gag*, and *pol*/*env* genes with lengths 119, 113 and 269 bp (A2) and 134, 114, and 205 bp (A5), respectively. Further, the A4, A6, A8 and A9 proviruses each had two CpG islands located in the LTR and *pol*/*env* genes with lengths 122 and 205 bp (A4), 120 and 205 bp (A6) 122 and 205 bp (A8), and 120 and 205 bp (A9), respectively. Proviruses A1 and A3 each had only one CpG island located in the *pol*/*env* gene with lengths 183 and 170 bp, respectively (Table 2).

Characteristics of CpG islands in enJSRV strains. The results confirmed that the B11 provirus of the enJSRV strain had four CpG islands, located in the LTR (1), *gag* (1), and *pol*/*env* (2) genes with lengths 133, 108, 169, and 100 bp, respectively. The B2, B5–B9, B16, B17, B20, B22–B24, B26, B27 and B33 proviruses each had three CpG islands each located in the LTR, *gag*, and *pol/env* genes with varying lengths in the range 104–200 bp.

Table 1. JSRV strains used for sequence analyses.

Group	Strains	Year	Accession No.	Length (bp)	Origin
Exogenous JSRV					
A1	Jaagsiekte sheep retrovirus isolate HH13	2013	KX817296	6,943	China
A2	Jaagsiekte sheep retrovirus isolate sheep/C1/China/2013/G1	2013	KP691837	7,456	China
A3	Jaagsiekte sheep retrovirus	2006	DQ838494	7,430	China
A4	Ovine pulmonary adenocarcinoma virus JSRV21	1998	AF105220	7,455	UK
A5	JS7	2001	AF357971	11,791 (7,841)	UK
A6	Jaagsiekte sheep retrovirus	1992	M80216	7,462	South Africa
A7	Jaagsiekte sheep retrovirus strain DL37	2017	MN161849	7,460	India
A8	Patent WO2004104032	2004	CQ964469	7,455	USA
A9	JSRV-USA	2000	NC001494	7,462	USA
Endogenous JSRV					
B1	*Ovis aries* endogenous virus Jaagsiekte sheep retrovirus	2000	DQ838493	7,942	China
B2	enJSRV 5f16	2000	AF136224	6,916	South Africa
B3	enJSRV 5.9A1	2000	AF136225	6,696	South Africa
B4	enJSRV 5.6A1	2000	AF153615	7,940	South Africa
B5	*Ovis aries* isolate HamJ1 endogenous virus Jaagsiekte sheep retrovirus	2017	MF175067	7,941	USA
B6	*Ovis aries* isolate HamJ2 endogenous virus Jaagsiekte sheep retrovirus	2017	MF175068	7,941	USA
B7	*Ovis aries* isolate HamM endogenous virus Jaagsiekte sheep retrovirus	2017	MF175069	7,941	USA
B8	*Ovis aries* isolate KarJ endogenous virus Jaagsiekte sheep retrovirus	2017	MF175070	7,941	USA
B9	*Ovis aries* isolate KarM endogenous virus Jaagsiekte sheep retrovirus	2017	MF175071	7,941	USA
B10	*Ovis aries* strain enJSRV-1 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680311	7,989	USA
B11	*Ovis aries* strain enJSRV-2 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680310	7,937	USA
B12	*Ovis aries* strain enJSRV-3 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680309	7,835	USA
B13	*Ovis aries* strain enJSRV-4 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680317	3,294	USA
B14	*Ovis aries* strain enJSRV-5 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680305	7,921	USA
B15	*Ovis aries* strain enJSRV-6 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680319	5,914	USA
B16	*Ovis aries* strain enJSRV-7 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680298	7,941	USA
B17	*Ovis aries* strain enJSRV-8 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680306	7,942	USA
B18	*Ovis aries* strain enJSRV-9 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680316	5,593	USA
B19	*Ovis aries* strain enJSRV-10 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680307	7,934	USA
B20	*Ovis aries* strain enJSRV-11 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680303	7,945	USA
B21	*Ovis aries* strain enJSRV-13 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680296	7,939	USA
B22	*Ovis aries* strain enJSRV-14 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680308	7,946	USA
B23	*Ovis aries* strain enJSRV-15 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680299	7,943	USA
B24	*Ovis aries* strain enJSRV-16 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680300	7,941	USA
B25	*Ovis aries* strain enJSRV-17 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680315	3,973	USA
B26	*Ovis aries* strain enJSRV-18 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680301	7,941	USA
B27	*Ovis aries* strain enJSRV-19 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680304	7,936	USA
B28	*Ovis aries* strain enJSRV-20 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680302	7,941	USA
B29	*Ovis aries* strain enJSRV-21 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680312	7,936	USA
B30	*Ovis aries* strain enJSRV-23 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680313	7,549	USA
B31	*Ovis aries* strain enJSRV-24 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680318	2,494	USA
B32	*Ovis aries* strain enJSRV-25 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680314	7,581	USA
B33	*Ovis aries* strain enJSRV-26 endogenous virus Jaagsiekte sheep retrovirus	2007	EF680297	7,875	USA

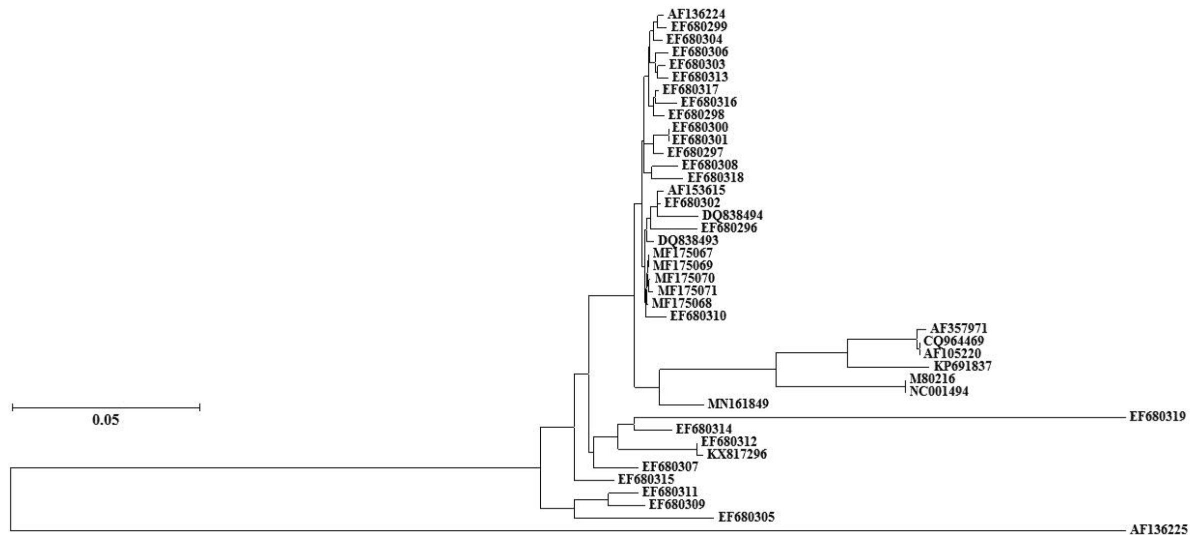

Fig. 1. Evolutionary relationships of exJSRV and enJSRV strains. Phylogenetic analysis based on 42 nucleotide sequences of previously published exJSRV and enJSRV proviruses was conducted using the neighbor-joining method (MEGA 5.0). The tree is drawn to scale with branch lengths provided in the same units as the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the maximum composite likelihood method. Further, bootstrap values were estimated for 1,000 replicates.

Further, the B14, B19 and B32 proviruses each had three CpG islands each located in the LTR, *gag*/*pol*, and *env* genes, with varying lengths in the range 100–203 bp. Furthermore, the B1, B3, B4, B10, B15, B18, B21, B28 and B29 proviruses each had two CpG islands located in the LTR and *pol*/*env* genes with varying lengths in the range 105–183 bp. Our analysis also showed that the B12 and B30 proviruses each had two CpG islands located in the *gag* and *pol*/*env* genes with varying lengths in the range 107–164 bp and the B13, B25, and B31 proviruses each had only one CpG island located in the *env*, *gag*, and *env* genes, respectively, with lengths 162, 164, and 200 bp, respectively (Table 3).

Discussion

Based on the genetic evolutionary tree analysis in this study, as previously described, the DQ838494 strain and enJSRVs belonged to the same branch suggesting a close relationship between them.[21] However, the typical ScaI restriction site in the upstream U5 and downstream U3 regions of the LTR region has been identified as an infectious exJSRV-specific enzyme cleavage map feature.[30] Additionally, it has been reported that the transmembrane (TM) region encoded by the *env* gene has a specific "YXXM" motif, which is a unique sequence of infectious exJSRV.[31] Therefore, the DQ838494 strain was classified as an exogenous virus based on the molecular features of its proviral genome. Phylogenetic analysis also showed that strain KX817296 from Heilongjiang Province in China was closely related to enJSRVs in Clade A, however, it contained the "LHMKYXXM" motif, but not the ScaI enzyme site, indicating that it could be an exogenous virus.[6]

In this study, we also observed that the CpG islands in the two types of strains were primarily distributed in the LTR, *gag*, and *env* genes. At the proviral genome level, the exJSRV and enJSRV strains are very similar.[2,32] Their two variable regions (VRs) are mainly located in the U3 regions of the LTR, *gag* (including two variable regions, VR1\VR2), and *env* genes.[32,33] In particular, it has been reported that the *env* gene encodes the TM region (VR3) of proteins.[32] Further, we observed that nucleotide variation at the genome level explained the difference in the number and location of CpG islands between exJSRV and enJSRV strains to some extent. Reportedly, enJSRVs, which according to a previous study,[21] represent the oldest proviruses, was clustered into clade A, and most of them except for B25 (only 3,973 bp long) showed a defect in the CpG island within the *gag* gene. Additionally, abnormal DNA methylation is closely related to genomic abnormalities such as gene mutation and deletion.[34]

A possible explanation for this observation is that methylation leads to gene instability.[35-37] Particularly, previous studies have shown that DNA methylation increases gene mutation rate and that 5-methylcytosine can spontaneously deaminate to form thymine. [38] This process is more frequent than cytosine conversion to thymine.[38] In addition, the types and active states of infected cells (e.g., clone expansion) and the chromatin environment (e.g., spatial structure, gene density, GC content, and repetitive elements) near the retroviral integration sites in the host genome may also be considered to influence differences in the number and location of CpG islands between the exJSRV and enJSRV strains.[39,40] This can be further investigated in future *in vitro* and *in vivo* epigenetic experiments.

Table 2. Location in figure (bp), gene location (bp) and lengths (bp) of CpG islands within sequences of exogenous JSRV strains. Group representing the strains are shown in Table 1.

Groups	Islands 1			Islands 2			Islands 3			Islands 4		
	Location	Gene location	Length	Location	Gene location	Length	Location	Gene location	Length	Location	Gene location	Length
A1										5,051–5,233	2,857–5,199 (pol) 5,096–6,943 (env)	183
A2	50–168	14–127 (LTR)	119	866–978	262–2,097 (gag)	113				5,301–5,569	2,824–5,451 (pol) 5,348–7,195 (env)	269
A3										5,323–5,492	3,119–5,422 (pol) 5,362–7,197 (env)	170
A4	51–172	14–128 (LTR)	122							5,300–5,504	2,827–5,451 (pol) 5,348–7,195 (env)	205
A5	312–445	1–399 (LTR)	134	1,130–1,243	534–2,369 (gag)	114				5,568–5,772	3,097–5,718 (pol) 5,594–7,462 (env)	205
A6	50–169	14–126 (LTR)	120							5,304–5,508	3,108–5,441 (pol) 5,350–7,197 (env)	205
A7	49–152	(LTR)	104	850–963	264–2,112 (gag)	114	1,824–1,931	2,64–2,112(gag)	108	5,280–5,449	3,119–5,422 (pol) 5,362–7,197 (env)	170
A8	51–172	(LTR)	122							5,300–5,504	(pol), (env)	205
A9	50–169	14–126 (LTR)	120							5,304–5,508	3,108–5,441 (pol) 5,350–7,197 (env)	205

The blanks represent the absence of CpG island.

Table 3. Location in figure (bp), gene location (bp) and lengths (bp) of CpG islands within sequences of endogenous JSRV strains. Groups representing the strains shown in Table 1.

Groups	Islands 1			Islands 2			Islands 3			Islands 4			Islands 5		
	Location	Gene location	Length	Location	Gene location	Length	Location	Gene location	Length	Location	Gene location	Length	Location	Gene location	Length
B1[a]	366–488	1–435 (LTR)	123							5,637–5,806	3,440–5,742 (pol) 5,682–7,517 (env)	170			
B2	382–490	(LTR)	109	2,182–2,288	583–2,433(gag)	107				4,611–4,772	4,656–6,491 (env)	162			
B3	346–471	(LTR)	126							5,623–5,795	(env)	173			
B4	366–492	(LTR)	127							5,634–5,797	3,440–5,740 (pol) 5,680–7,515 (env)	164			
B5	345–487	1–446 (LTR)	143	2,180–2,287	582–2,432(gag)	108				5,636–5,805	3,439–5,784 (pol) 5,681–7,516 (env)	170			
B6	359–487	1–446 (LTR)	129	2,180–2,287	582–2,432(gag)	108				5,636–5805	3,439-5,784 (pol) 5681–7516 (env)	170			
B7	359–487	1–446 (LTR)	129	2,180–2,287	582–2,432 (gag)	108				5636–5,805	3439–5784 (pol) 5,681–7,516 (env)	170			
B8	347–487	1–446 (LTR)	141	2,180–2,287	582–2,432 (gag)	108				5,636–5,805	3,439–5,784 (pol) 5,681–7,516 (env)	170			
B9	359–487	1–446 (LTR)	129	2,180–2,287	582–2,432 (gag)	108				5,636–5,805	3,439–5,784 (pol) 5,681–7,516 (env)	170			
B10	340–444	1–404 (LTR)	105							5,728–5,904	(pol) (env)	177			
B11	358–490	1–446 (LTR)	133	2,176–2,283	(gag)	108				5,633–5,801	(pol) (env)	169	6,164–6,263	(env)	100
B12				2,136–2,243	536–2,386 (gag)	108				5,575–5,738	3,379–5,724 (pol)(env)	164			
B13										989–1,150	1,034–2,869 (env)	162			
B14	364–477	1–441 (LTR)	114				4,788–4,887	(pol)	100	5,633–5,812	(pol) (env)	180			

Table 3. Continued.

Groups	Islands 1			Islands 2			Islands 3			Islands 4			Islands 5		
	Location	Gene location	Length	Location	Gene location	Length	Location	Gene location	Length	Location	Gene location	Length	Location	Gene location	Length
B11	358–490	1–446 (LTR)	133	2,176–2,283	(gag)	108				5,633–5,801	(pol) (env)	169	6,164–6,263	(env)	100
B12				2,136–2,243	536–2,386 (gag)	108				5,575–5,738	3,379–5,724 (pol)(env)	164			
B13										989–1,150	1,034–2,869 (env)	162			
B14	364–477	1–441 (LTR)	114				4,788–4,887	(pol)	100	5,633–5,812	(pol) (env)	180			
B15	364–489	1–444 (LTR)	126							3,623–3,778	3,724–5,490 (env)	156			
B16	384–487	1–446 (LTR)	104	2,181–2,287	582–2,432 (gag)	107				5,636–5,797	3,439–5,772 (pol) 5,681–7,516 (env)	162			
B17	382–487	1–446 (LTR)	106	2,180–2,287	582–2,432 (gag)	108				5,636–5,742	3,439–5,772 (pol) 5,681–7,516 (env)	107			
B18	382–487	1–446 (LTR)	106							3,288–3,448	3,333–5,168 (env)	161			
B19	345–487	1–444 (LTR)	143							5,630–5,801	(pol) (env)	172	6,223–6,353	(env)	131
B20	361–488	1–447 (LTR)	128	2,182–2,289	(gag)	108				5,639–5,745	5,684–7,519 (env)	107			
B21	361–486	1–446 (LTR)	126							5,640–5,803	5,679–7,514 (env)	164			
B22	383–487	1–446 (LTR)	105	2,181–2,287	582–2,432 (gag)	107				5,636–5,835	3,439–5,808 (pol) 5,681–7,516 (env)	200			
B23	365–488	1–447 (LTR)	124	2,181–2,288	583–2,433 (gag)	108				5,637–5,798	3,440–5,773(pol) 5,682–7,517 (env)	162			
B24	361–487	1–446 (LTR)	127	2,180–2,287	582–2,432 (gag)	108				5,636–5,797	3,439–5,772 (pol) 5,681–7,516 (env)	162			
B25				2,012–2,175	(gag)	164									
B26	361–487	1–446 (LTR)	127	2,180–2,287	582–2,432 (gag)	108				5,636–5,797	3,439–5,772 (pol) 5,681–7,516 (env)	162			
B27	381–491	1–446 (LTR)	111	2,175–2,287	582–2,432 (gag)	113				5,637–5,792	4,937–5,437 (pol) 5,676–7,511 (env)	156			
B28	365–491	1–446 (LTR)	127							5,636–5,811	3,439–5,784 (pol) 5,681–7,516 (env)	176			
B29	346–485	1–446 (LTR)	140							5,631–5,813	(pol) (env)	183			
B30				1,788–1,895	(gag)	108				5,244–1,350	3,047–5,380 (pol) 5,289–7,124 (env)	107			
B31										189–388	234–2069 (env)	200			
B32	345–485	1–446 (LTR)	141							5,634–5,836	3,439–5,784 (pol) 5,681–7,516 (env)	203	6,148–6,254	(env)	107
B33	348–454	1–413 (LTR)	107	2,147–2,254	549–2,399 (gag)	108				5,603–5,778	3,406–5,751 (pol) 5,648–7,483 (env)	176			

The blanks represent the absence of CpG islands.

The number of CpG islands located in the LTR region and *gag* gene (especially the latter) differ among exJSRV strains. The study has confirmed that the U3 region of the LTR and env proteins of JSRV are synergistically responsible for this unique tissue tropism.[41] Notably, the LTRs of JSRVs contain viral promoter and enhancer elements are needed to drive transcription and numerous studies have demonstrated several enhancer elements in the LTR of JSRVs.[42,43] The difference in the number of CpG islands among exJSRV strains may influence the transcriptional specificity of exJSRV LTRs in viral infected cells and thereafter determine their pathogenicity and disease specificity. The *gag* gene encodes several structural proteins including nucleocapsid, matrix and capsid which is required to encapsulate the viral RNA genome and form the viral particle core.[43] Our study showed that the *gag* gene of some exJSRV strains had one to two CpG islands. Further, the difference between the exJSRV strains with respect to the number of CpG islands within the *gag* gene possibly affected the synthesis of structural proteins which contributed to the assembly of progeny viral particles. In future studies, it would be necessary to investigate whether DNA methylation is involved in the regulation of JSRV LTR transcriptional activity and the synthesis of *gag*-encoded proteins, thereby, affecting the pathogenicity of the virus.

The env protein of exJSRVs is a major factor for inducing cell transformation and OPA tumorigenesis *in vitro*.[44,45] In this study, our analysis revealed that all exJSRV and enJSRV strains had at least one CpG island within the *env* gene. It has also been observed that the *env* gene encodes the viral Env protein, a glycoprotein containing approximately 615 amino acids that can be cleaved into surface and a TM protein by cellular furin protease.[1] Further, exJSRV TM proteins contain a 45-amino acid cytoplasmic tail (CT) region comprising the "LHMKYXXM" motif, where "YXXM" is required for exJSRV-induced cell transformation. Furthermore, no enJSRV env protein contains the "LHMKYXXM" motif in the CT region of the TM protein.[6] Previous studies have shown that oncogene activation and tumor suppressor gene inactivation are mainly mediated by changes in DNA sequences caused by gene mutation and deletion.[46] However, tumor research, mutations or deletions have not yet been detected in many important tumor genes and it is generally considered that abnormal gene expression primarily results from DNA methylation.[47,48] Whether DNA methylation is involved in cell transformation induced by the exJSRV Env protein remains to be further verified.

Acknowledgments

This study was supported by grants from the Nanchong Vocational and Technical College for basic scientific research (no. ZRA1904 and no. NZYBZ2002), the Nanchong High-Level Talents Special Support Program (Scientific Elite of Ten-Thousands Talents Program) and the Science and Technology Planning Project of Nanchong (no. 22XCZX0003).

Conflict of interest

The authors declare no conflicts of interest.

References

1. Leroux C, Girard N, Cottin V, et al. Jaagsiekte sheep retrovirus (JSRV): from virus to lung cancer in sheep. Vet Res 2007; 38(2): 211-228.
2. De las Heras M, Ortín A, Cousens C, et al. Enzootic nasal adenocarcinoma of sheep and goats. Curr Top Microbiol Immunol 2003; 275: 201-223.
3. York DF, Querat G. A history of ovine pulmonary adenocarcinoma (jaagsiekte) and experiments leading to the deduction of the JSRV nucleotide sequence. Curr Top Microbiol Immunol 2003; 275: 1-23.
4. De Las Heras M, Borobia M, Ortín A. Neoplasia-associated wasting diseases with economic relevance in the sheep industry. Animals (Basel) 2021; 11(2): 381. doi: 10.3390/ani11020381.
5. Zhang K, Kong H, Liu Y, et al. Diagnosis and phylogenetic analysis of ovine pulmonary adenocarcinoma in China. Virus Genes 2014; 48(1): 64-73.
6. Shi W, Jia S, Guan X, et al. A survey of jaagsiekte sheep retrovirus (JSRV) infection in sheep in the three northeastern provinces of China. Arch Virol 2021; 166(3): 831-840.
7. Yang S, Liang T, Zhao Q, et al. Pathological diagnoses and whole-genome sequence analyses of the jaagsiekte sheep retrovirus in Xinjiang, China [Chinese]. Bing Du Xue Bao 2015; 31(3): 217-225.
8. Chiu ES, VandeWoude S. Endogenous retroviruses drive resistance and promotion of exogenous retroviral homologs. Annu Rev Anim Biosci2021; 9: 225-248.
9. Youssef G, Wallace WA, Dagleish MP, et al. Ovine pulmonary adenocarcinoma: a large animal model for human lung cancer. ILAR J 2015; 56(1): 99-115.
10. Gray ME, Meehan J, Sullivan P, et al. Ovine pulmonary adenocarcinoma: a unique model to improve lung cancer research. Front Oncol 2015; 9: 335. doi: 10.3389/fonc.2019.00335
11. Lanigan LG, Hildreth BE 3rd, Dirksen WP, et al. In vivo tumorigenesis, osteolytic sarcomas, and tumorigenic cell lines from transgenic mice expressing the human t-lymphotropic virus type 1 (HTLV-1) Tax viral oncogene. Am J Pathol 2021; 191(2): 335-352.
12. Durkin K, Rosewick N, Artesi M, et al. Characterization of novel bovine leukemia virus (BLV) antisense transcripts by deep sequencing reveals constitutive

expression in tumors and transcriptional interaction with viral microRNAs. Retrovirology 2016; 13(1): 33. doi: 10.1186/s12977-016-0267-8.
13. Voisin V, Barat C, Hoang T, et al. Novel insights into the pathogenesis of the Graffi murine leukemia retrovirus. J Virol 2006; 80(8): 4026-4037.
14. Irby RB, Mao W, Coppola D, et al. Activating SRC mutation in a subset of advanced human colon cancers. Nat Genet 1999; 21(2): 187-190.
15. Pecenka V, Pajer P, Karafiat V, et al. *HRAS, EGFR,* MET, and *RON* genes are recurrently activated by provirus insertion in liver tumors induced by the retrovirus myeloblastosis-associated virus 2. J Virol 2017; 91(20): e00467-17. doi: 10.1128/JVI.00467-17.
16. Wootton SK, Halbert CL, Miller AD. Sheep retrovirus structural protein induces lung tumours. Nature 2005; 434(7035): 904-907.
17. Maeda N, Palmarini M, Murgia C, et al. Direct transformation of rodent fibroblasts by jaagsiekte sheep retrovirus DNA. Proc Natl Acad Sci U S A 2001; 98(8): 4449-4454.
18. Palmarini M, Maeda N, Murgia C, et al. A phosphatidylinositol 3-kinase docking site in the cytoplasmic tail of the jaagsiekte sheep retrovirus transmembrane protein is essential for envelope-induced transformation of NIH 3T3 cells. J Virol 2001; 75(22): 11002-11009.
19. Rai SK, Duh FM, Vigdorovich V, et al. Candidate tumor suppressor HYAL2 is a glycosylphosphatidylinositol (GPI)-anchored cell-surface receptor for jaagsiekte sheep retrovirus, the envelope protein of which mediates oncogenic transformation. Proc Natl Acad Sci U S A 2001; 98(8): 4443-4448.
20. Maeda N, Fu W, Ortin A, et al. Roles of the Ras-MEK-mitogen-activated protein kinase and phosphatidylinositol 3-kinase-Akt-mTOR pathways in jaagsiekte sheep retrovirus-induced transformation of rodent fibroblast and epithelial cell lines. J Virol 2005; 79(7): 4440-4450.
21. Arnaud F, Caporale M, Varela M, et al. A paradigm for virus-host coevolution: sequential counter-adaptations between endogenous and exogenous retroviruses. PLoS Pathog 2007; 3(11): e170. doi: 10.1371/journal. ppat.0030170.
22. Lavialle C, Cornelis G, Dupressoir A, et al. Paleovirology of '*syncytins*', retroviral *env* genes exapted for a role in placentation. Philos Trans R Soc Lond B Biol Sci 2013; 368(1626): 20120507. doi: 10.1098/rstb.2012.0507.
23. Spencer TE, Palmarini M. Endogenous retroviruses of sheep: a model system for understanding physiological adaptation to an evolving ruminant genome. J Reprod Dev 2012; 58(1): 33-37.
24. Xiao J, Liu P, Wang Y, et al. A novel cognition of decitabine: insights into immunomodulation and antiviral effects. Molecules 2022; 27(6):1973. doi: 10.3390/molecules27061973.
25. Kausar S, Abbas MN, Cui H. A review on the DNA methyltransferase family of insects: Aspect and prospects. Int J Biol Macromol 2021; 186: 289-302.
26. Cain JA, Montibus B, Oakey RJ. Intragenic CpG islands and their impact on gene regulation. Front Cell Dev Biol 2022; 10: 832348. doi: 10.3389/fcell.2022.832348.
27. Mittelstaedt NN, Becker AL, de Freitas DN. DNA methylation and immune memory response. Cells 2021; 10(11): 2943. doi: 10.3390/cells10112943.
28. Mazloumi Z, Farahzadi R, Rafat A, et al. Effect of aberrant DNA methylation on cancer stem cell properties. Exp Mol Pathol 2022; 125: 104757. doi: 10.1016/j.yexmp.2022.104757.
29. Nishiyama A, Nakanishi M. Navigating the DNA methylation landscape of cancer. Trends Genet 2021; 37(11): 1012-1027.
30. Palmarini M, Holland MJ, Cousens C, et al. Jaagsiekte retrovirus establishes a disseminated infection of the lymphoid tissues of sheep affected by pulmonary adenomatosis. J Gen Virol 1996; 77 (Pt 12): 2991-2998.
31. Hofacre A, Fan H. Multiple domains of the jaagsiekte sheep retrovirus envelope protein are required for transformation of rodent fibroblasts. J Virol 2004; 78(19): 10479-10489.
32. Palmarini M, Hallwirth C, York D, et al. Molecular cloning and functional analysis of three type D endogenous retroviruses of sheep reveal a different cell tropism from that of the highly related exogenous jaagsiekte sheep retrovirus. J Virol 2000; 74(17): 8065-8076.
33. Hallwirth C, Maeda N, York D, et al. Variable regions 1 and 2 (VR1 and VR2) in JSRV gag are not responsible for the endogenous JSRV particle release defect. Virus Genes 2005; 30(1): 59-68.
34. Yang T, Liu X, Kumar SK, et al. Decoding DNA methylation in epigenetics of multiple myeloma. Blood Rev 2022; 51: 100872. doi: 10.1016/j.blre.2021.100872.
35. Shrestha KS, Tuominen MM, Kauppi L. Mlh1 heterozygosity and promoter methylation associates with microsatellite instability in mouse sperm. Mutagenesis 2021; 36(3): 237-244.
36. Radpour R, Barekati Z, Haghighi MM, et al. Correlation of telomere length shortening with promoter methylation profile of p16/Rb and p53/p21 pathways in breast cancer. Mod Pathol 2010; 23(5): 763-772.
37. Meng H, Cao Y, Qin J, et al. DNA methylation, its mediators and genome integrity. Int J Biol Sci 2015; 11(5): 604-617.
38. Chen CL, Rappailles A, Duquenne L, et al. Impact of replication timing on non-CpG and CpG substitution rates in mammalian genomes. Genome Res 2010; 20(4): 447-457.
39. Boltz VF, Ceriani C, Rausch JW, et al. CpG methylation

profiles of HIV-1 Pro-viral DNA in individuals on ART. Viruses 2021; 13(5): 799. doi: 10.3390/v13050799.
40. Yan H, Zhang D, Liu H, et al. Chromatin modifications and genomic contexts linked to dynamic DNA methylation patterns across human cell types. Sci Rep 2015; 5: 8410. doi: 10.1038/srep08410.
41. Palmarini M, Datta S, Omid R, et al. The long terminal repeat of jaagsiekte sheep retrovirus is preferentially active in differentiated epithelial cells of the lungs. J Virol 2000; 74(13): 5776-5787.
42. McGee-Estrada K, Fan H. In vivo and in vitro analysis of factor binding sites in jaagsiekte sheep retrovirus long terminal repeat enhancer sequences: roles of HNF-3, NF-I, and C/EBP for activity in lung epithelial cells. J Virol 2006; 80(1): 332-341.
43. McGee-Estrada K, Fan H. Comparison of LTR enhancer elements in sheep beta retroviruses: insights into the basis for tissue-specific expression. Virus Genes 2007; 35(2): 303-312.
44. Hofacre A, Fan H. Jaagsiekte sheep retrovirus biology and oncogenesis. Viruses 2010; 2(12): 2618-2648.
45. Hull S, Lim J, Hamil A, et al. Analysis of jaagsiekte sheep retrovirus (JSRV) envelope protein domains in transformation. Virus Genes 2012; 45(3): 508-517.
46. Rosales Gerpe MC, van Lieshout LP, Domm JM, et al. The U3 and Env proteins of jaagsiekte sheep retrovirus and enzootic nasal tumor virus both contribute to tissue tropism. Viruses 2019; 11(11):1061. doi: 10.3390/v11111061.
47. Chen L, Liu S, Tao Y. Regulating tumor suppressor genes: post-translational modifications. Sig Transduct Target Ther 2020; 5: 90. doi: 10.1038/s41392-020-0196-9
48. Ehrlich M. DNA hypermethylation in disease: mechanisms and clinical relevance. Epigenetics 2019; 14(12): 1141-1163.