

Article

A DNA Regulatory Element Haplotype at Zinc Finger Genes Is Associated with Host Resilience to Small Ruminant Lentivirus in Two Sheep Populations

Alisha T. Massa ¹, Michelle R. Mousel ^{2,3}, Codie J. Durfee ², Maria K. Herndon ¹, Kaneesha M. Hemmerling ¹, J. Bret Taylor ⁴, Holly L. Neibergs ^{5,6} and Stephen N. White ^{2,6,*}

¹ Department of Veterinary Microbiology and Pathology, Washington State University, Pullman, WA 99163, USA; alisha.massa@wsu.edu (A.T.M.); mkmeyer@wsu.edu (M.K.H.); kaneesha.hemmerling@wsu.edu (K.M.H.)

² Animal Disease Research Unit, Agricultural Research Service, United States Department of Agriculture, Pullman, WA 99163, USA; Michelle.Mousel@usda.gov (M.R.M.); Codie.Durfee@usda.gov (C.J.D.)

³ Paul G. Allen School for Global Animal Health, Washington State University, Pullman, WA 99163, USA

⁴ Range Sheep Production Efficiency Research, Agricultural Research Service, United States Department of Agriculture, Dubois, ID 83423, USA; bret.taylor@usda.gov

⁵ Department of Animal Science, Washington State University, Pullman, WA 99163, USA; neibergs@wsu.edu

⁶ Center for Reproductive Biology, Washington State University, Pullman, WA 99163, USA

* Correspondence: stephen.white@usda.gov



Citation: Massa, A.T.; Mousel, M.R.; Durfee, C.J.; Herndon, M.K.; Hemmerling, K.M.; Taylor, J.B.; Neibergs, H.L.; White, S.N. A DNA Regulatory Element Haplotype at Zinc Finger Genes Is Associated with Host Resilience to Small Ruminant Lentivirus in Two Sheep Populations. *Animals* **2021**, *11*, 1907. <https://doi.org/10.3390/ani11071907>

Academic Editors: Emilia Bagnicka and Aldona Kawęcka

Received: 28 April 2021

Accepted: 22 June 2021

Published: 26 June 2021

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



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

Simple Summary: Sheep are affected by a viral infection that causes an incurable and difficult to treat lung, joint and brain disease that decreases production efficiency. This small ruminant lentivirus is closely related to human immunodeficiency virus (HIV) that causes AIDS in humans. Differences in breed susceptibility to disease are known in sheep that indicate genetic, or hereditary, resilience factors exist. Our objective was to study the source of this hereditary advantage so that it can eventually be translated into a tool for sheep breeders to improve herd health. Previously, one such hereditary region was detected, but little was known about possible mechanisms or nearby mutations. Here, we report several mutations that may underlie this hereditary mechanism and are in regions of DNA that are known to affect genes by increasing or decreasing gene expression, akin to gene “on/off switches.” These mutations were strongly associated with a predictor of disease severity in live animals and had a greater predicted effect on the degree of disease than previously studied mutations. Statistical association ($p < 0.05$) of disease was demonstrated in two different groups of sheep that were reared in different environments, which indicates increased likelihood that a genetic factor is producing this effect.

Abstract: Small ruminant lentivirus (SRLV) causes Maedi-Visna or Ovine Progressive Pneumonia in sheep and creates insidious livestock production losses. This retrovirus is closely related to human immunodeficiency virus and currently has no vaccines or cure. Genetic marker assisted selection for sheep disease resiliency presents an attractive management solution. Previously, we identified a region containing a cluster of zinc finger genes that had association with ovine SRLV proviral concentration. Trait-association analysis validated a small insertion/deletion variant near *ZNF389* (rs397514112) in multiple sheep breeds. In the current study, 543 sheep from two distinct populations were genotyped at 34 additional variants for fine mapping of the regulatory elements within this locus. Variants were selected based on ChIP-seq annotation data from sheep alveolar macrophages that defined active cis-regulatory elements predicted to influence zinc finger gene expression. We present a haplotype block of variants within regulatory elements that have improved associations and larger effect sizes (up to 4.7-fold genotypic difference in proviral concentration) than the previously validated *ZNF389* deletion marker. Hypotheses for the underlying causal mutation or mutations are presented based on changes to in silico transcription factor binding sites. These variants offer alternative markers for selective breeding and are targets for future functional mutation assays.

Keywords: small ruminant lentivirus; sheep; post-infection control; genetic fine map; validation

1. Introduction

Ovine lentivirus (OvLV) exists as a collection of viral subtypes in sheep also termed small ruminant lentiviruses (SRLV) [1]. These are retroviruses which infect both sheep and goats and belong to the same genus as human immunodeficiency virus (HIV). Conserved host factors have been discovered that effect viral restriction of lentiviruses in both humans and small ruminants, such as TRIM5 α [2]. The SRLV retrovirus is macrophage-tropic and incorporates into the host DNA through reverse transcription that results in pervasive life-long infection [3,4]. In sheep, the disease is referred to as Ovine Progressive Pneumonia in the U.S. and Maedi-Visna in many other countries [5], because it targets the lungs, central nervous system and joints. SRLV was documented to cause production losses by increased lamb mortality [6], reduced lamb weights in diseased older ewes [7], reduced fertility [8], and early culling of breeding stock [9]. These losses can be difficult to quantify and are frequently underestimated in individual flocks since producers may not test all animals or maintain quantitative records on all potential measures of lost production [10]. Nonetheless, economic and production efficiency consequences of SRLV are significant since 36% to 66% of commercial flocks in the U.S. contained positive sheep on serological studies [11–13]. Complete eradication is impractical for many flocks since retroviruses have a very small minimum infectious dose [14], and serology is the standard diagnostic testing method available with imperfect sensitivity [15,16]. Once infected there is no cure, and no approved vaccine is commercially available to prevent infection [17]. Selective breeding for sheep that better control the virus, after infection, is an option to combat economic losses and welfare issues associated with SRLV without requiring complete eradication from the flock.

Previously, a genome-wide association study identified a locus containing zinc finger genes that were associated with host control of the virus and lesion severity after infection [18,19]. Proviral load in the blood has been documented to correlate with histological lesion severity in sheep [19] and is likely a useful phenotype to predict disease resiliency in sheep since this is well documented in humans with HIV [20]. Early efficient control of viral replication is generally associated with reduced disease severity for HIV and other lentiviruses [21]. Proviral load in humans is associated with disease severity, clinical progression, treatment response and is documented to be very low to exceptionally low in slow progressors and elite controllers infected with HIV [20–24]. In sheep, we are defining the phenotype of reduced proviral load as resiliency to small ruminant lentivirus because it is a marker of control of the pathogen after infection potentially involving aspects of disease tolerance and is not considered resistance, which is defined as imperviousness to infection or elimination of infection [25]. Breeding for tolerance or resiliency may create less selective pressure on the pathogen than breeding for strict resistance [26]. The locus contains four zinc finger genes *ZNF389*, *ZKSCAN8*, *ZSCAN16*, and *ZNF165* [27–29] that had not been otherwise documented to affect viral infections in vivo. However, other zinc finger proteins, such as ZAP, have been shown to restrict retroviruses [30]. In fact, this innate host factor has been shown to restrict viral replication in a variety of retroviruses from several species [1]. Zinc finger genes have also been demonstrated to coevolve in the host with endogenous and exogenous retroviruses [31,32]. Eight markers across the locus, mostly within genes, were tested in previous work, and a single insertion/deletion marker (rs397514112) near *ZNF389* was validated in three sheep populations [33]. This marker and the haplotype it tracks were found to have no negative associations with production traits [34], meaning this locus could be used for selective breeding without obvious unintended negative effects.

The four zinc finger genes in the region associated with lentiviral resiliency in sheep are the Cys2/His2-type (C2H2) and include Kruppel associated box (KRAB) domains with

known repressive function [35]. C2H2 zinc finger genes contain repeated domains that form finger-like protein structures of 28–30 amino acids which bind the DNA helix. C2H2 zinc finger proteins are one of the largest families of transcription factors in eukaryotic genomes, however those with greater than three zinc finger domains at the C-terminus such as *ZNF389*, *ZKSCAN8*, *ZSCAN16* and *ZNF165*, are the most poorly understood [36], and prediction of their DNA binding sequence becomes less accurate when there are higher numbers of zinc finger domains. Domestic cattle and pigs, other artiodactyl species related to sheep, have annotated orthologs of *ZNF389*; *ZKSCAN8*, *ZSCAN16*, and *ZNF165* are annotated in many mammalian species including humans in the current Refseq annotations [37–40]. In humans, a pseudogene *ZKSCAN8P1*, has been implicated as *ZNF389* but it is currently unclear if this is an orthologous gene to *ZNF389* annotated in sheep. At the time of the association analysis of the zinc finger gene region in sheep [33], the specific functional implications for the intergenic validated marker near *ZNF389* and the zinc finger genes were unknown.

Reported here, further assessment of polymorphisms in the region were completed to identify the functional importance of these variants and to examine additional markers that may be useful for genetic marker-assisted selection. DNA regulatory elements were previously annotated by chromatin immunoprecipitation and high throughput sequencing (ChIP-seq) for histone post-translational modification marks including H3K27ac and H3K4me3 in sheep [41] and used to focus the search for polymorphic variants of interest to the SRLV resilient phenotype. These regulatory elements are associated with reproducible biological functions such as gene enhancers and promoters. For instance, the histone modification H3K27ac denoted active chromatin regions typically associated with enhancers [42,43]. Chromatin that has the histone modification H3K4me3 and H3K27ac near the transcription start site of a gene defines active promoter elements that indicate active gene expression [44,45].

Associations were tested across 34 variants within the eight regulatory elements recently annotated [41] in this chromatin domain. A haplotype within experimentally determined active DNA regulatory elements from previously determined ChIP-seq data in sheep was associated with a resilient phenotype against small ruminant lentivirus. Variants within active cis-regulatory elements such as promoters and enhancers often have functional consequences for transcription factor binding and gene expression which will be explored further. These results are consistent with the model that most causal mutations of complex disease phenotypes are within noncoding regulatory elements rather than coding regions of genes and influence gene expression [46–48].

2. Materials and Methods

2.1. Ethics Statement

Animals were cared for and handled according to protocols approved by the Washington State University Institutional Animal Care and Use Committee (ASAF 4618 and 6632) or by the U.S. Sheep Experiment Station Institutional Animal Care and Use Committee (protocols: 10-06 and 10-07).

2.2. DNA Extractions and Genotyping

Whole blood was collected via jugular venipuncture into EDTA-vacutainer tubes for DNA extraction as previously described [31] from two sheep populations in a total of 543 animals. Rambouillet purebred ewes (population 1) were sampled at the U.S. Sheep Experiment Station in Idaho and crossbred (Rambouillet, Columbia) ewes (population 2) from a privately owned production flock in Montana. Population 1 consisted of 164 small ruminant lentivirus infected ewes between one to five years of age. Flock SRLV prevalence for population 1 was 42.2% [33]. Population 2 consisted of 379 ewes, between one to six years of age, selected by a random number generator from the 533 SRLV infected animals in the flock. Older infected animals aged seven to eight years were excluded for estimability due to low numbers of animals in these groups since age was included in the model. Popu-

lation 2 SRLV prevalence was 87.2%. Additional characteristics of the populations were reported in previous publications [18,33]. No animals were shared between population groups. The GeneCatcher gDNA Blood Kit (Invitrogen, Life Technologies, Carlsbad, CA) was used according to manufacturer's instructions for DNA extraction.

Four individual Rambouillet animals were selected from population 1 based on alternate homozygous genotype status at the *ZNF389* validated marker for whole genome high-throughput sequencing (Rambouillet) in order to select variants within regulatory elements for further genotyping in all 543 sheep. These four Rambouillet were selected to fit the overall population trend, in that the two insertion homozygous animals had low proviral concentrations and two deletion homozygous animals had high proviral concentrations. Two additional uninfected Rambouillet from the same flock as population 1 and two crossbred sheep (Suffolk, Polypay, Rambouillet, Targhee) from a Washington research flock [41], were sequenced at high depth and included in evaluation for DNA variants to screen for rare alleles (crossbred sheep). Paired-end sequencing (Illumina) produced between 600–700 million reads per animal that passed read quality and adapter trimming with BBDuk from the BBTools suite (Bushnell, 2021, <http://sourceforge.net/projects/bbmap/>) (accessed on 1 November 2020). Reads were aligned to the Rambouillet reference genome (Oar_Rambouillet_v1.0, GCF_002742125.1) [27] with BWA [49]. Aligned reads were quality filtered, sorted, and indexed with SAMtools [50].

The region on chromosome 20 between 32.92 Mb and 33.01 Mb was evaluated for variants in the sequenced Rambouillet sheep by comparison to the Rambouillet reference genome and the sequenced crossbred sheep. Variants were selected by sequence inspection of aligned reads in IGV [51] that appeared to be on the same or similar haplotypes as the validated marker and were within the active regulatory elements as determined by ChIP-seq. ChIP-seq annotations in the zinc finger region were utilized from those previously published [41] and with attention to regions annotated as transcription start sites from cap analysis gene expression (CAGE) data [27]. ChIP-seq signal consensus BED files combined from two animals for the histone modifications H3K4me3 (promoters) and H3K27ac (active enhancers and promoters) and raw ChIP-seq signal BigWig tracks in each individual animal were used for regulatory element annotation. CAGE data that shows transcription start sites for genes from publicly available annotation data was overlaid with the active *cis*-regulatory elements of interest [27,41]. ChIP-seq BigWig signal tracks for CTCF (insulators) and H3K27me3 (silencers/heterochromatin) enrichment were used to define the boundaries of the chromatin domain and limit which regulatory elements and genes around the previously validated marker were included in assessment.

Taqman Genotyping Assays were performed according to manufacturer's instructions (Applied Biosystems, Foster City, CA USA) utilizing 20 ng of genomic DNA in 10 microliter reactions for all 543 animals in both populations. Dried genomic DNA samples were incubated in liquid Polymerase Chain Reaction (PCR) reagents for 15 min at 37 °C immediately prior to PCR cycling. The PCR thermal cycler program was polymerase activation at 95 °C for 10 min, then repeating 50 cycles of 95 °C denaturing for 15 s and 60 °C annealing/extension for 1 min. Primers and probes were designed for variants by the manufacturer's website tool <https://www.thermofisher.com/order/custom-genomic-products/tools/genotyping/> (accessed on 1 November 2020). Assay identification codes and sequences for primers and fluorescent probes are provided in supplementary Table S1. Automated design for the initially selected variants failed at 11% of the sites, mostly due to a high density of additional variants nearby. Genotyping was completed at 36 variants with successful genotyping assay cluster differentiation.

2.3. Small Ruminant Lentivirus Phenotypes

Proviral concentration was assessed on all animals in triplicate qPCR assays as previously described [16] utilizing 1 ug of DNA extracted from peripheral blood. Only animals infected with SRLV were included in the downstream association analyses. The two pop-

ulations of sheep were chosen for high SRLV prevalence within the flocks as previously described [33].

2.4. Statistical Analysis

Association analyses were completed in JMP Genomics version 7.1 (SAS Institute, Cary, NC USA). A standard least squares model was employed with independent variables set as genotype and age in years, and sire treated as a random effect. The dependent variable was \log_{10} -transformed SRLV proviral concentration. Age in years was treated as a categorical (nominal) variable to account for non-linearity of proviral concentration as sheep age. p -values less than 0.05 were considered significant. SRLV phenotypes were reverse-transformed (10^x) to viral copies/ug DNA scale for indicated results. Genotypes were assessed for Hardy-Weinberg Equilibrium (HWE) with chi-squared analysis ($p < 0.05$) which yielded 34 variants for association testing. HWE p -values are displayed for each variant in supplementary Table S2.

2.5. Linkage Disequilibrium and Haplotype Analyses

Linkage disequilibrium measures were derived using Haploview v4.2 [52] r^2 and D' values. Haplotypes with less than 1% allele frequency were considered rare and removed from further analysis. PHASE v2.1.1 [53,54] was used to estimate population frequency of haplotypes, and genotypes were assigned for individual animals based on calculated haplotypes. Animals with haplotype assignment probabilities of less than 95% were dropped from haplotype association analysis. A standard least squares model as described above was used to evaluate haplotype association with SRLV phenotypes.

2.6. Transcription Factor Binding Analysis

The MATCH tool [55] was used with the TRANSFAC database (version 2021.1, geneXplain GmbH, Qiagen, Germantown, MD) of binding motif sequences to predict allelic difference in binding affinity around each selected variant. Settings for motif scanning included immune cell specific with a core and matrix match above 0.9, or vertebrate non-redundant with minimize false positives and minimize the sum of false negative and false positive results. Any transcription factor binding motif that was predicted to be created or abolished between the two alleles was recorded.

3. Results

3.1. ZNF Gene Region in Detail

The chromatin domain around the previously validated marker, rs397514112 [25], was delineated by evaluation of boundary elements from ChIP-seq for in the relevant cell type of alveolar macrophages. Evaluation of both CTCF insulators and H3K27me3 silencers delimited the domain as extending from approximately 32.92 Mb to 33.01 Mb on Rambouillet_v1.0 chromosome 20. These boundaries were affirmed by the presence of annotated tRNA genes.

Four annotated genes were found within the region, *ZNF389* (NCBI Gene ID 101104612), *ZKSCAN8* (NCBI Gene ID 101121465), *ZSCAN16* (NCBI Gene ID 101104351) and *ZNF165* (NCBI Gene ID 106990266), each containing C2H2 zinc finger repeats at the C-terminus and KRAB or SCAN domains near the N-terminus. Each gene had a proximal 5' active *cis*-regulatory element defined by both H3K4me3 and H3K27ac enrichment of the genome. *ZNF389* also had two downstream, within-gene, smaller regulatory elements. CAGE data indicated there were three transcription start sites for *ZNF389*, one within each of the ChIP-seq regulatory elements. There were two additional active, proximal regulatory elements upstream of two predicted lncRNAs, *ZSCAN16-AS1* (NCBI Gene ID 105603808) and LOC114108620 (NCBI Gene ID 114108620). The regions searched for variants included approximately 14,500 bp of DNA in the eight total active regulatory elements. Taqman genotyping assays were performed and analyzed for 34 variants including 30 single nucleotide polymorphisms (SNP), one multiple nucleotide polymorphism (MNP) and three

small insertion/deletion variants that varied from 1–3 bp. Greater than 95% of animals were successfully genotyped, and all variants passed HWE analysis ($p > 0.05$), see supplementary Table S2 for chi-squared results for population 1 and supplementary Table S3 for population 2. In addition, the previously validated insertion/deletion marker (rs397514112) was included for reference. Twenty-four of the variants were within the two regulatory elements that immediately flanked the previously validated marker (Figure 1). The remainder of the variants were distributed amongst five additional regulatory elements within 57.5 Kb in the 3' direction and within 10.7 Kb in the 5' direction from the validated marker.

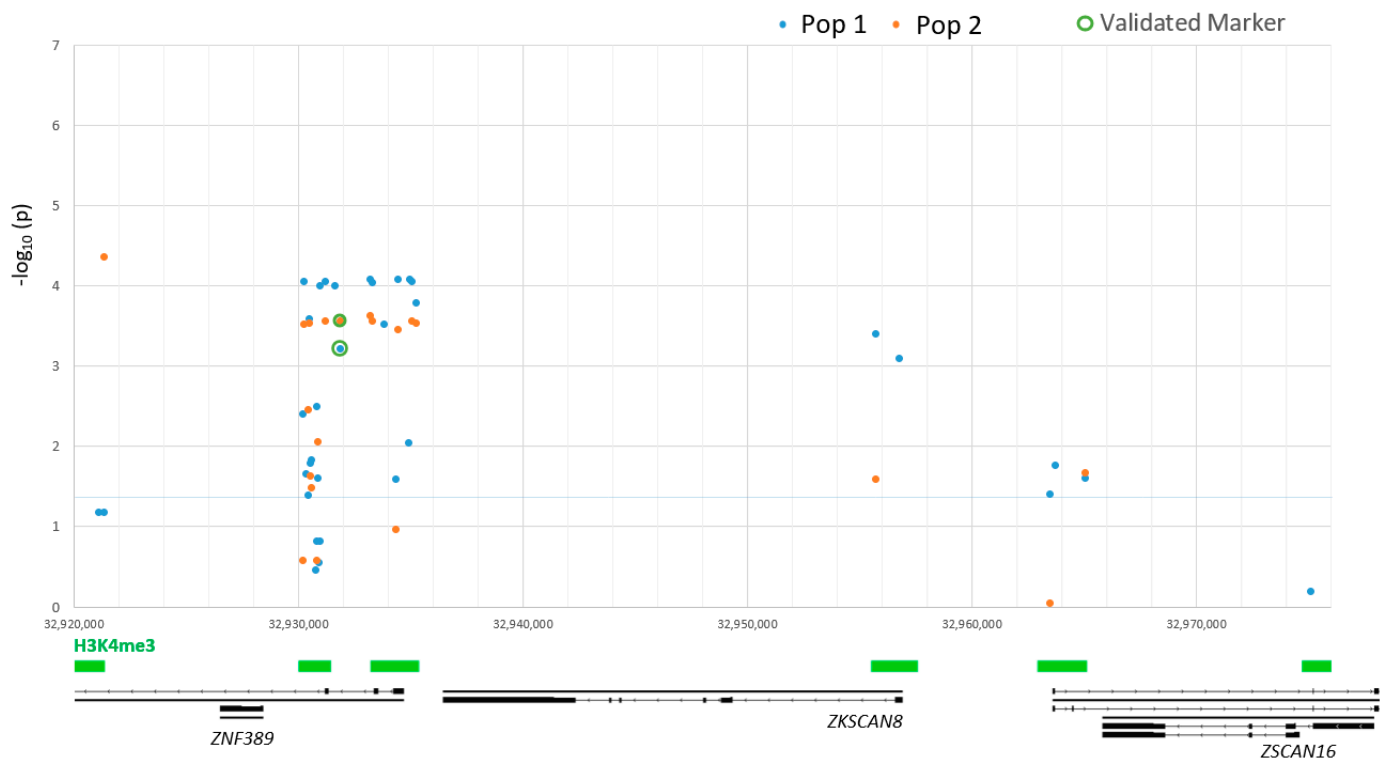


Figure 1. Association with small ruminant lentivirus proviral concentration at variants within regulatory elements for zinc finger genes on ovine chromosome 20. ChIP-seq annotations for H3K4me3 enrichment are shown as green bars at the bottom, above Refseq gene annotations in black. Population 1 (purebred Rambouillet) are identified by blue dots and population 2 (crossbred Rambouillet, Columbia) are identified by orange dots. The previously validated marker (rs397514112) is highlighted by green circles for each population.

3.2. Population One Association Analysis

Within population one, 164 SRLV positive animals were included in the association analyses. Overall, twenty-eight of the variants had significant association with SRLV proviral concentration ($p < 0.05$) shown in Figure 1. Variant minor allele frequencies ranged from 0.48 to 0.02 in purebred Rambouillet sheep (Table 1). For all variants with a minor allele frequency greater than 0.05, the heterozygous animals adjusted mean proviral concentration was not significantly different from the resilient homozygote adjusted mean.

At the previously validated marker, the association with SRLV remained ($p = 6.00 \times 10^{-4}$) with a proviral concentration log-transformed difference of 0.803 (Table 2) between animals with a homozygous deletion and those with a homozygous insertion. This corresponded to a reverse-transformed adjusted mean of 24.4 viral copies/ug of DNA in the insertion homozygotes and 155.2 viral copies/ug of DNA in the more susceptible deletion homozygotes. Additionally, 14 variants tested had improved p -values ($p = 4.00 \times 10^{-4}$ to $p = 5.08 \times 10^{-7}$) and greater effect sizes as calculated by the \log_{10} proviral concentration difference between homozygous genotypes when compared to the previous validated marker as shown in Table 2. Variants within three separate regulatory elements for ZNF389 and ZNF192

(ZKSCAN8) had the highest log-transformed adjusted mean proviral concentration difference between genotypes of 0.838 to 0.967 in population 1. The reverse-transformed adjusted mean proviral concentrations were 32.7 viral copies/ug DNA and 303.4 viral copies/ug DNA for alternate homozygous genotypes at the most extreme variant (rs420471584) which was 24 Kb 3' from the previous validated marker.

Table 1. Minor allele frequency at each variant in the sheep population 1 purebred Rambouillet and population 2 crossbred (Rambouillet, Columbia) sheep. Allele nucleotide bases are shown for SNPs: adenine (A), cytosine (C), guanine (G), and thymine (T), I = insertion and D = deletion for small indels. Total animals genotyped for each population (*n*). See supplementary Table S2 for animal counts by genotype. * Previously validated marker.

Marker	Population 1 <i>n</i> = 164		Population 2 <i>n</i> = 379	
	Minor Allele	Freq	Minor Allele	Freq
rs193645606	C	0.44	T	0.45
rs418789060	A	0.35	A	0.47
rs161334222	T	0.44	C	0.44
rs398476053	C	0.42	C	0.20
rs414155747&rs425583788	TT	0.43	TT	0.20
rs427575002	T	0.36	T	0.47
rs407355422	G	0.44	C	0.44
rs397514112 *	D	0.44	I	0.44
rs599110985	G	0.44	T	0.44
rs411076283	C	0.44	T	0.44
rs407841455	T	0.44	G	0.44
rs161334287	T	0.44	A	0.44
rs598937573	I	0.44	D	0.44
rs406431156	C	0.33	C	0.47

Table 2. Genotypes at each regulatory element marker tested that yielded a significant association ($p < 0.05$) with SRLV phenotype of resilience in sheep. The most extreme genotypes associated with resiliency and susceptibility to SRLV and log-transformed adjusted mean proviral concentration difference between genotypes (Genotypic log₁₀ Conc. Diff.) for Population 1 (purebred Rambouillet) is shown. The *p*-value calculated for the association analysis in population 1 and linkage disequilibrium (r^2) between each newly tested marker and the previously validated marker for both population 1 and population 2 is also shown. Genotype allele nucleotide bases are shown for SNPs: adenine (A), cytosine (C), guanine (G), and thymine (T), I = insertion and D = deletion for small indels. See supplementary Table S4 for additional data on population 2. * Previously validated marker.

Marker	Resilient Genotype	Susceptible Genotype	Genotypic Log ₁₀ Conc. Diff.	<i>p</i> -Value	LD * Pop. 1 (r^2)	LD * Pop. 2 (r^2)
rs193645606	T/T	C/C	0.857	8.95×10^{-5}	0.950	0.989
rs418789060	C/C	A/A	0.537	4.06×10^{-2}	0.664	0.705
rs161334222	C/C	T/T	0.830	2.57×10^{-4}	0.926	0.989
rs398476053	C/C	A/A	0.434	1.64×10^{-2}	0.537	0.312
rs414155747&rs425583788	TT/TT	GC/GC	0.445	1.46×10^{-2}	0.556	0.313
rs427575002	A/A	T/T	0.572	2.46×10^{-2}	0.690	0.678
rs407355422	C/C	G/G	0.859	8.70×10^{-5}	0.950	0.989
rs397514112 *	I/I ¹	D/D	0.803	6.00×10^{-4}	-	-
rs599110985	T/T	G/G	0.857	8.39×10^{-5}	0.950	0.989
rs411076283	T/T	C/C	0.856	9.20×10^{-5}	0.950	0.989
rs407841455	G/G	T/T	0.857	8.39×10^{-5}	0.950	0.989
rs161334287	A/A	T/T	0.857	8.72×10^{-5}	0.950	0.989
rs598937573	D/D ²	I/I	0.838	1.62×10^{-4}	0.950	0.989
rs406431156	T/T	C/C	0.739	2.50×10^{-2}	0.389	0.604

¹ The insertion allele at the validated marker is AAT and the deletion allele is A. ² The deletion allele is G and the insertion allele is GAAT.

3.3. Population Two Association Analysis

Population two was genotyped at 20 variants which had association with SRLV resiliency in population one. In total, 378 animals were retained in analyses as one animal was removed for high genotyping failure rate (>10% of variants tested). Minor allele frequency in the crossbred population ranged from 0.47 to 0.04 (Table 1). Fifteen variants across five different regulatory elements were associated with SRLV proviral concentration ($p < 0.05$), see Table 2. Eight of the variants in the haplotype had a larger adjusted mean log-transformed proviral concentration difference between homozygous genotypes than the previous validated marker, and nine of the variants had greater significance (lower p -values), as displayed in Table 2. For nearly all variants tested in this population, the adjusted mean proviral concentration in heterozygous animals was not different from the susceptible homozygote ($p > 0.05$).

3.4. Significant Variants in Multiple Populations Were in Two Regulatory Elements

In both populations, the most significant and extreme variants were located within the two regulatory elements that flank the *ZNF389* insertion/deletion marker (Figure 2). These two active regulatory elements were located at 32.931 Mb and 32.935 Mb on chromosome 20. The larger element was approximately 2100 base pairs in length and the smaller downstream element was 1400 bp in length. The larger regulatory region was predicted to contain approximately 29 total variants within Rambouillet, including rare variants and those for which Taqman allelic discrimination assays could not be designed. The smaller regulatory element was predicted to contain 17 total variants within Rambouillet. The alleles across the markers tested in these elements are displayed in Table 2 according to association with resiliency or susceptibility to SRLV. The same alleles were associated with resiliency and susceptibility in both populations.

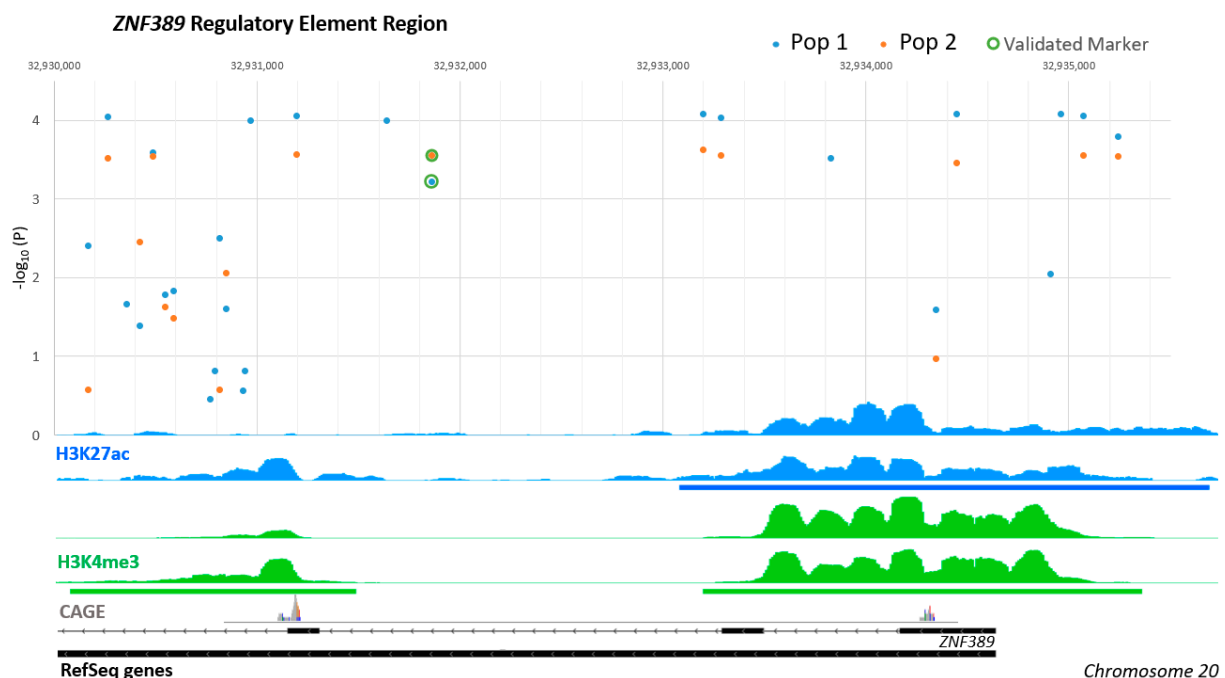


Figure 2. Association with SRLV proviral concentration at higher resolution for variants within the two regulatory elements near *ZNF389* that flank the previously validated insertion/deletion marker (rs397514112) on chromosome 20. Below, the ChIP-seq signal tracks for two animals enriched for H3K27ac (active proximal enhancers) and H3K4me3 (active promoters) is displayed. Transcription start site data from cap analysis gene expression (CAGE) from alveolar macrophages is also displayed [27]. Peak calls significant in both animals are indicated by the solid blue or green bars. Refseq gene annotation is displayed in black along the bottom.

3.5. Haplotypes and Phenotypic Association with Common Haplotypes

Linkage disequilibrium analysis showed that 7 of the 15 significant markers were in exceptionally strong LD with the previously validated marker ($r^2 > 0.95$) in population 1 (see Table 2) and 8 of the 15 significant markers were in exceptionally strong LD with the previously validated marker in population 2. Haplotypes were assessed across the 15 significant markers in each population and yielded three common haplotypes that were present for both animal groups in greater than 10% of the population (Figure 3). Three shortened common haplotypes were determined with PHASE analysis that consisted of the markers in the two key regulatory elements: rs414155747, rs425583788, rs407355422 and rs397514112 (validated marker). Haplotype 1 had alleles GCCD (D = deletion, A and I = insertion, AAT) across these four markers and was most common in both populations, with a frequency of 0.436 in population 1 and 0.554 in population 2. Haplotype 2 had alleles TTCI with a frequency of 0.427 in population 1 and 0.201 in population 2. Haplotype 3 was composed of the alleles GCCI with a frequency of 0.125 in population 1 and 0.241 in population 2. Haplotype 1 was highly associated with an SRLV susceptible phenotype in population 1 ($p = 2.72 \times 10^{-5}$) in a dose dependent manner (0, 1 or 2 copies of the haplotype) and in population 2 ($p = 5.82 \times 10^{-4}$). The reverse transformed adjusted mean proviral concentration for population 1 was 226 viral copies/ μg DNA for two copies of the haplotype, 49.8 viral copies/ μg DNA for one copy of the haplotype and 29.4 viral copies/ μg DNA for zero copies of haplotype 1. Haplotype 2 was also associated with SRLV proviral load, but was protective with one or two copies of the haplotype associated with a similar resiliency effect ($p = 2.10 \times 10^{-3}$ for population 1 and $p = 2.94 \times 10^{-2}$ for population 2). Haplotype 3 was not significantly associated with SRLV phenotype ($p = 0.09$ and $p = 0.11$ for populations 1 and 2).

** *+	TCCCTTACITTTGADT	0.389	→	** *+	CATAGCTGDGCTTIC	0.450
	CATAGCTGDGCTTIC	0.287	→		TCCAGCACITTTGADT	0.230
	TCCAGCACDTTGADT	0.133	→		TCCCTTACITTTGADT	0.200
	CCTAGCAGDGCTTIT	0.079			CCTAGCAGDGCTTIT	0.085
	CATAGCTGDGCTTIT	0.064			TCCAGCACITTTGADC	0.011
	TCCCTTACITTTGADC	0.029			CATAGCTGDGCTTIT	0.009
	TCCCTTACDTTGADT	0.006			CATAGCAGDGCTTIT	0.009
	CCCAGCTGDGCTTIC	0.006				

Figure 3. Haplotypes and population frequency as determined by PHASE. On the left are the haplotypes within population 1 (purebred Rambouillet) and on the right are haplotypes for population 2 (crossbred, Rambouillet, Columbia). The same three haplotypes were the most common in both populations. The fifteen loci are shown in columns listed in the same order as in Tables 1 and 2, with the previously validated marker listed as +. An asterisk (*) is above each locus in the two key regulatory elements upstream of *ZNF389*. For loci 09 and 14, D represents the deletion allele and I represents the insertion allele.

3.6. TRANSFAC Predictions

The variants associated with SRLV proviral concentration that were validated in both populations were analyzed using the TRANSFAC database to identify transcription factor binding motifs. Ten of the variants including a single insertion/deletion variant and a single MNP (2 bp) had predicted differences in transcription factor binding based on the allelic sequence. All variants, except one, were within the two *ZNF389* regulatory elements that flank the previous validated marker. Transcription factor proteins implicated in unique binding at one of the alleles included GATA3, c-JUN, TCF-1, LEF-1, Ikaros, ZNF333, AP-2alpha, NFATc2 and HNF-3beta. Ubiquitous proteins such as AP-1 and SP100 were also implicated in predicted binding differences. At the rs427575002 variant, the repressive transcription factors BEN and Kaiso had predicted differences in binding affinity. The variant that is predicted to bind these repressors was found within the smaller downstream

ZNF389 regulatory element which had an annotated difference in ChIP-seq signal in the region between alternate homozygote animals (see Figure 2).

4. Discussion

The zinc finger chromatin domain examined in sheep represents a synteny block conserved from humans to artiodactyls. However, one major difference in humans compared to sheep is that *ZNF389* is annotated as a pseudogene paralog of *ZKSCAN8* (*ZNF192*) in humans. *ZNF389* is an interesting gene target for a novel viral restriction factor as it may have functionally diverged in the caprine subfamily in response to the recent emergence of small ruminant lentiviruses [29,32,56]. Maedi-visna, ovine progressive pneumonia and caprine arthritis and encephalitis are caused by multiple subtypes of the same slow viral disease in sheep and goats that have emerged in the last century [57] and cause multisystemic wasting characterized by chronic lymphocytic pneumonia, arthritis, mastitis, and encephalitis. The poly-zinc finger transcription factor gene family consists of over 700 members and examination across host clades indicates that these genes arose from repeated duplication events followed by functional divergence [32,58]. Generally, positive selection for new transcriptional repression activity in response to transposable retroelements and retroviruses has been found or hypothesized [59,60]. Zinc finger transcription factors also have emerging roles in modulating the innate immune response [61] and therefore viral restriction factors in one species may be protective against retroviruses from another species. There are other well-known examples of host viral restriction factors that affect retroviruses from multiple species, such as APOBEC3, Tetherin, ZAP and TRIM5 α [30,56,62].

Through fine mapping of this region on ovine chromosome 20, we identified several candidates for selection markers which may prove more useful for predictive breeding decisions in other breeds of sheep than the single previously validated marker. This region appears to contain a haplotype block of variants that are significantly associated with SRLV resiliency within two active regulatory elements upstream of the *ZNF389* gene. This could indicate that these regulatory elements are functionally linked and influence gene expression together or that more than one sequence variant may be playing a causal role in the phenotype. These variants validated in two populations of sheep from different breed backgrounds and under different environmental conditions and had greater significance than the previously validated marker. Studies have also highlighted variability in utility of a single genetic marker for trait selection in different breeds or environments [5], thus it is important to mention that the related breed composition of the two populations in this study is a limitation to broader assumptions on SRLV association with these markers. In particular, the *TMEM154* mutations for odds of infection have been shown to associate with reduced risk [63,64] in American and European sheep and with improved control post-infection in some U.S. sheep [65]. However, *TMEM154* mutations were not associated with SRLV infection in different breeds in several populations [66].

Importantly, the effect size is greater at some of these markers (MNP: rs414155747 and rs425583788, rs407355422) compared to the previously validated marker (rs397514112). We hypothesize that these variants would cause a change in expression of the zinc finger genes that restrict viral replication. The discovery of this haplotype within the regulatory elements of zinc finger genes lends further support that one or multiple host zinc finger genes may function as a viral restriction factor in sheep.

Genome-wide association studies in humans and livestock indicate that the majority of functional variants lie within DNA regulatory elements [46,67–69]. Annotation of these regulatory elements in relevant cell- and tissue- types will be of increasing importance to elucidate the biology behind functional mutations in food producing species. Published alveolar macrophages ChIP-seq data [41] was leveraged to evaluate active regulatory elements in the chromosome region and identify the boundaries of the putative chromatin domain since these cells are the primary phagocyte of the lungs. The distance searched for variants was also limited based on decline of linkage disequilibrium with distance in the same population of sheep ($r^2 < 0.25$ at 35 Kb) [18]. Out of eight total regulatory

elements evaluated within the region, the most significant variants were present within two adjacent *cis*-regulatory elements nearest to *ZNF389* flanking the previously validated marker (Figures 1 and 2).

Determination of regulatory elements with trait associations is important because polymorphisms have been detected in intergenic regions associated with critical economic traits in livestock such as polledness [70]. The vast wealth of human literature indicates that it is most common for genome-wide association studies to find significant polymorphisms for traits within promoters and enhancers rather than within coding regions of genes [46,71]. Breaking research within livestock species indicates a similar trend emphasizing the importance of studying these regulatory elements in many species [69]. *Cis*-acting DNA regulatory elements control gene expression by serving as structural elements for many *trans*-acting proteins to promote or disrupt gene expression [72]. Promoters marked by H3K4me3 are often present at many genes in the cell type that are not actively expressed [41,44]. However, if the region is also marked at nucleosomes by H3K27ac, this is highly associated with gene expression. The zinc finger gene regulatory elements identified to contain the haplotype cluster had enrichment for both H3K27ac and H3K4me3 in macrophages indicating they may be expressed in this key cell type for lentiviral infection. Examination of CAGE data [27] showed that each of the two *ZNF389* regulatory elements had a narrow region of active mRNA transcription. Other members of this zinc finger family have been implicated as transcriptional repressors; we hypothesize that increased expression of one of these genes due to promoter variants may cause increased transcriptional repression to the viral genome or provirus. ChIP-seq data in sheep indicate potential insulators enriched by CTCF around *ZNF389* and *ZNF192* which would pose the most likely gene effectors. However, these four, and many other zinc finger genes on the same chromosome have been linked to similarly located active *cis*-regulatory elements in humans from the ENCODE cell line data [73].

Evaluation of predicted *in silico* transcription factor binding motifs between alleles for each variant resulted in discovery of putative functional differences in these two active regulatory elements. This method may reveal functional variants that cause changes in transcription factor protein binding [74,75], as described here, and found in livestock at gene promoters for other traits of economic significance [48]. Analysis of transcription factor binding changes yielded hypotheses for function of ten of the associated variants from within the two *ZNF389* regulatory elements that were validated in both sheep populations. Several of the transcription factors with predicted changes have been implicated to various degrees in altering the host response to viruses or altering viral replication. Some of these transcription factors have been directly shown to interact with viruses such as LEF-1 in the case of HIV [76]. However, binding of transcription factors to host promoters would be an indirect effect. We hypothesize transcription factors function through expression alteration in zinc finger genes which then mediate a downstream direct effect on the virus. In mice, Gata3 has been implicated in altered anti-viral immunity [77], which was one of the transcription factors we predicted to have differential binding affinity to sheep promoters in the region. Functional testing for molecular effects of these variants were outside the scope of this research study but will be performed in the future such as electrophoretic mobility shift assays and promoter reporter assays.

Regulation of gene expression appears to be complex within this region. For example, there is an active regulatory element for both *ZSCAN16* and *ZSCAN16-AS1* which is an antisense lncRNA. Antisense RNA can act as a sponge for miRNA and protect the complementary mRNA transcript from the effects of miRNA, creating several layers of ncRNA regulation for a single gene. However, this study cannot rule out long distance interactions of these regulatory elements or genes with distant *cis*-regulatory elements along the chromosome. Chromatin folding has occasionally been implicated in functional association of gene-enhancer loops at distances of up to 1 Mb. Additionally, since several of these variants are in high linkage disequilibrium with the previously validated marker, this study does not fully resolve linkage association from biological function in the tested populations.

5. Conclusions

Overall, a haplotype of variants within two active *cis*-regulatory elements for *ZNF389* was identified that improved the association and/or effect size for small ruminant lentivirus proviral concentration as a live-animal measure of lesion severity in Rambouillet and related crossbreeds from the Western U.S. This provision of additional variants will prove useful in selective breeding decisions in the future, although the utility of phenotype predictive value in sheep from different breeds or geographical locations remains unknown. Ten variants were identified that had predicted transcription factor binding differences between alleles. One or more of these variants can be incorporated into a commercial test to combat SRLV infection which ultimately has potential to improve efficiency of meat, milk, and wool production. Since several viral restriction factors previously known affect lentiviruses from multiple species, a novel viral restriction factor identified in sheep may have implications for interactions between host and retroviruses in additional species.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ani11071907/s1>, supplementary Table S1: additional details on genotyping reagents; Table S2: Hardy-Weinberg Equilibrium analysis and animal counts by genotype for population 1; Table S3: Hardy-Weinberg Equilibrium analysis and animal counts by genotype for population 2; Table S4: Genotypes at each regulatory element marker tested that yielded a significant association ($p < 0.05$) with SRLV phenotype of resilience in sheep in population 2.

Author Contributions: Conceptualization, A.T.M., M.R.M. and S.N.W.; data curation, A.T.M. and C.J.D.; formal analysis, A.T.M.; funding acquisition, M.R.M. and S.N.W.; investigation, A.T.M., M.R.M., C.J.D., M.K.H., K.M.H., J.B.T., H.L.N. and S.N.W.; methodology, A.T.M., M.R.M. and S.N.W.; project administration, A.T.M., M.R.M., J.B.T. and H.L.N. and S.N.W.; resources, A.T.M., M.R.M., K.M.H., J.B.T. and S.N.W.; supervision, M.R.M., J.B.T., H.L.N. and S.N.W.; validation, A.T.M. and C.J.D.; visualization, A.T.M.; writing—original draft, A.T.M.; writing—review and editing, A.T.M., M.R.M., C.J.D., M.K.H., K.M.H., J.B.T., H.L.N. and S.N.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by USDA-ARS 2090-32000-036-00D.

Institutional Review Board Statement: The study was conducted according to the guidelines approved by the Institutional Animal Care and Use Committee of Washington State University and the U.S. Sheep Experiment Station of USDA ARS.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here <https://www.ebi.ac.uk/ena/browser/view/PRJEB40528> (accessed on 1 February 2021) for ChIP-seq [41]. Data for CAGE can be found here <https://www.ebi.ac.uk/ena/browser/view/PRJEB34864> [27] (accessed on 20 February 2021).

Acknowledgments: Thank you to the USSES farm crew and ADRU staff for initial blood collection and SRLV phenotypes. Thank you to David Herndon for assistance with whole genome sequencing. Thank you to Caylee Birge and ADRU staff for technical assistance with genotyping. Thank you to Maggie Highland for important consultation. This work was supported by USDA-ARS 2090-32000-036-00D. ATM was supported by the Seattle Chapter of ARCS and Connie Niva, the Washington State University Department of Veterinary Microbiology and Pathology and the American Sheep Industry Heritage Memorial scholarship. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

1. De Pablo-Maiso, L.; Doménech, A.; Echeverría, I.; Gómez-Arrebola, C.; De Andrés, D.; Rosati, S.; Gómez-Lucia, E.; Reina, R. Prospects in Innate Immune Responses as Potential Control Strategies against Non-Primate Lentiviruses. *Viruses* **2018**, *10*, 435. [CrossRef]
2. Jáuregui, P.; Crespo, H.; Glaria, I.; Luján, L.; Contreras, A.; Rosati, S.; de Andrés, D.; Amorena, B.; Towers, G.J.; Reina, R. Ovine TRIM5 α can restrict visna/maedi virus. *J. Virol.* **2012**, *86*, 9504–9509. [CrossRef]

3. Minguijón, E.; Reina, R.; Pérez, M.; Polledo, L.; Villoria, M.; Ramírez, H.; Leginagoikoa, I.; Badiola, J.J.; García-Marín, J.F.; de Andrés, D.; et al. Small ruminant lentivirus infections and diseases. *Vet. Microbiol.* **2015**, *181*, 75–89. [[CrossRef](#)]
4. Blacklaws, B.A. Small ruminant lentiviruses: Immunopathogenesis of visna-maedi and caprine arthritis and encephalitis virus. *Comp. Immunol. Microbiol. Infect. Dis.* **2012**, *35*, 259–269. [[CrossRef](#)]
5. White, S.N.; Knowles, D.P. Expanding possibilities for intervention against small ruminant lentiviruses through genetic marker-assisted selective breeding. *Viruses* **2013**, *5*, 1466–1499. [[CrossRef](#)]
6. Arsenault, J.; Dubreuil, P.; Girard, C.; Simard, C.; Bélanger, D. Maedi-visna impact on productivity in Quebec sheep flocks (Canada). *Prev. Vet. Med.* **2003**, *59*, 125–137. [[CrossRef](#)]
7. Keen, J.E.; Hungerford, L.L.; Littledike, E.T.; Wittum, T.E.; Kwang, J. Effect of ewe ovine lentivirus infection on ewe and lamb productivity. *Prev. Vet. Med.* **1997**, *30*, 155–169. [[CrossRef](#)]
8. Dohoo, I.R.; Heaney, D.P.; Stevenson, R.G.; Samagh, B.S.; Rhodes, C.S. The effects of maedi-visna virus infection on productivity in ewes. *Prev. Vet. Med.* **1987**, *4*, 471–484. [[CrossRef](#)]
9. Peterhans, E.; Greenland, T.; Badiola, J.; Harkiss, G.; Bertoni, G.; Amorena, B.; Eliaszewicz, M.; Juste, R.A.; Krassnig, R.; Lafont, J.-P.; et al. Routes of transmission and consequences of small ruminant lentiviruses (SRLVs) infection and eradication schemes. *Vet. Res.* **2004**, *35*, 257–274. [[CrossRef](#)]
10. Lago, N.; López, C.; Panadero, R.; Cienfuegos, S.; Pato, J.; Prieto, A.; Díaz, P.; Mourazos, N.; Fernández, G. Seroprevalence and risk factors associated with Visna/Maedi virus in semi-intensive lamb-producing flocks in northwestern Spain. *Prev. Vet. Med.* **2012**, *103*, 163–169. [[CrossRef](#)]
11. Redden, R.R.; Schmidt, R.J.; Kirsch, J.D. Ovine Progressive Pneumonia Virus Infection Rate and Incidence of Genetic Susceptibility Diplotype in North Dakota Sheep Flocks. Available online: <https://www.ag.ndsu.edu/HettingerREC/sheep/individual-articles-from-2014-sheep-research-report/ovine-progressive-pneumonia-virus-infection-rate-and-incidence-of-genetic-susceptibility-diplotype-in-north-dakota-sheep-flocks> (accessed on 8 April 2021).
12. APHIS. Info Sheet: Ovine Progressive Pneumonia Awareness, Management, and Seroprevalence. Available online: http://www.aphis.usda.gov/animal_health/nahms/sheep/downloads/sheep01/Sheep01_is_OPP.pdf (accessed on 8 April 2021).
13. Cutlip, R.C.; Lehmkuhl, H.D.; Sacks, J.M.; Weaver, A.L. Seroprevalence of ovine progressive pneumonia virus in sheep in the United States as assessed by analyses of voluntarily submitted samples. *Am. J. Vet. Res.* **1992**, *53*, 976–979. [[PubMed](#)]
14. Herrmann-Hoesing, L.M.; Lehmkuhl, H.D.; Cutlip, R.C. Minimum intravenous infectious dose of ovine progressive pneumonia virus (OPPV). *Res. Vet. Sci.* **2009**, *87*, 329–331. [[CrossRef](#)] [[PubMed](#)]
15. Herrmann-Hoesing, L.M.; Noh, S.M.; Snekvik, K.R.; White, S.N.; Schneider, D.A.; Truscott, T.; Knowles, D.P. Ovine progressive pneumonia virus capsid antigen as found in CD163- and CD172a-positive alveolar macrophages of persistently infected sheep. *Vet. Pathol.* **2010**, *47*, 518–528. [[CrossRef](#)] [[PubMed](#)]
16. Herrmann-Hoesing, L.M.; White, S.N.; Lewis, G.S.; Mousel, M.R.; Knowles, D.P. Development and validation of an ovine progressive pneumonia virus quantitative PCR. *Clin. Vaccine Immunol.* **2007**, *14*, 1274–1278. [[CrossRef](#)]
17. Kalogianni, A.I.; Bossis, I.; Ekateriniadou, L.V.; Gelasakis, A.I. Etiology, Epizootiology and Control of Maedi-Visna in Dairy Sheep: A Review. *Animals* **2020**, *10*, 616. [[CrossRef](#)] [[PubMed](#)]
18. White, S.N.; Mousel, M.R.; Herrmann-Hoesing, L.M.; Reynolds, J.O.; Leymaster, K.A.; Neibergs, H.L.; Lewis, G.S.; Knowles, D.P. Genome-Wide Association Identifies Multiple Genomic Regions Associated with Susceptibility to and Control of Ovine Lentivirus. *PLoS ONE* **2012**, *7*, e47829. [[CrossRef](#)] [[PubMed](#)]
19. Herrmann-Hoesing, L.M.; Noh, S.M.; White, S.N.; Snekvik, K.R.; Truscott, T.; Knowles, D.P. Peripheral ovine progressive pneumonia provirus levels correlate with and predict histological tissue lesion severity in naturally infected sheep. *Clin. Vaccine Immunol.* **2009**, *16*, 551–557. [[CrossRef](#)] [[PubMed](#)]
20. Woldemeskel, B.A.; Kwaa, A.K.; Blankson, J.N. Viral reservoirs in elite controllers of HIV-1 infection: Implications for HIV cure strategies. *EBioMedicine* **2020**, *62*, 103118. [[CrossRef](#)] [[PubMed](#)]
21. Re, M.C.; Vitone, F.; Biagetti, C.; Schiavone, P.; Alessandrini, F.; Bon, I.; de Crignis, E.; Gibellini, D. HIV-1 DNA proviral load in treated and untreated HIV-1 sero-positive patients. *Clin. Microbiol. Infect.* **2010**, *16*, 640–646. [[CrossRef](#)]
22. Lathey, J.L.; Hughes, M.D.; Fiscus, S.A.; Pi, T.; Jackson, J.B.; Rasheed, S.; Elbeik, T.; Reichman, R.; Japour, A.; D’Aquila, R.T.; et al. Variability and Prognostic Values of Virologic and CD4 Cell Measures in Human Immunodeficiency Virus Type 1-Infected Patients with 200–500 CD4 Cells/mm³ (ACTG 175). *J. Infect. Dis.* **1998**, *177*, 617–624. [[CrossRef](#)] [[PubMed](#)]
23. Tierney, C.; Lathey, J.L.; Christopherson, C.; Bettendorf, D.M.; D’Aquila, R.T.; Hammer, S.M.; Katzenstein, D.A. Prognostic Value of Baseline Human Immunodeficiency Virus Type 1 DNA Measurement for Disease Progression in Patients Receiving Nucleoside Therapy. *J. Infect. Dis.* **2003**, *187*, 144–148. [[CrossRef](#)]
24. Pereyra, F.; Addo, M.M.; Kaufmann, D.E.; Liu, Y.; Miura, T.; Rathod, A.; Baker, B.; Trocha, A.; Rosenberg, R.; Mackey, E.; et al. Genetic and Immunologic Heterogeneity among Persons Who Control HIV Infection in the Absence of Therapy. *J. Infect. Dis.* **2008**, *197*, 563–571. [[CrossRef](#)]
25. Rauw, W. Immune response from a resource allocation perspective. *Front. Genet.* **2012**, *3*, 267. [[CrossRef](#)] [[PubMed](#)]
26. Stear, M.J.; Bishop, S.C.; Mallard, B.A.; Raadsma, H. The sustainability, feasibility and desirability of breeding livestock for disease resistance. *Res. Vet. Sci.* **2001**, *71*, 1–7. [[CrossRef](#)] [[PubMed](#)]

27. Salavati, M.; Caulton, A.; Clark, R.; Gazova, I.; Smith, T.P.L.; Worley, K.C.; Cockett, N.E.; Archibald, A.L.; Clarke, S.M.; Murdoch, B.M.; et al. Global analysis of transcription start sites in the new ovine reference genome (Oar rambouillet v1.0). *Front. Genet.* **2020**. [[CrossRef](#)] [[PubMed](#)]
28. Clark, E.L.; Bush, S.J.; McCulloch, M.E.B.; Farquhar, I.L.; Young, R.; Lefevre, L.; Pridans, C.; Tsang, H.G.; Wu, C.; Afrasiabi, C.; et al. A high resolution atlas of gene expression in the domestic sheep (*Ovis aries*). *PLoS Genet.* **2017**, *13*, e1006997. [[CrossRef](#)] [[PubMed](#)]
29. Jiang, Y.; Xie, M.; Chen, W.; Talbot, R.; Maddox, J.F.; Faraut, T.; Wu, C.; Muzny, D.M.; Li, Y.; Zhang, W.; et al. The sheep genome illuminates biology of the rumen and lipid metabolism. *Science* **2014**, *344*, 1168–1173. [[CrossRef](#)]
30. Zhu, Y.; Chen, G.; Lv, F.; Wang, X.; Ji, X.; Xu, Y.; Sun, J.; Wu, L.; Zheng, Y.T.; Gao, G. Zinc-finger antiviral protein inhibits HIV-1 infection by selectively targeting multiply spliced viral mRNAs for degradation. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 15834–15839. [[CrossRef](#)]
31. Lukic, S.; Nicolas, J.-C.; Levine, A.J. The diversity of zinc-finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. *Cell Death Differ.* **2014**, *21*, 381–387. [[CrossRef](#)]
32. Thomas, J.H.; Schneider, S. Coevolution of retroelements and tandem zinc finger genes. *Genome Res.* **2011**, *21*, 1800–1812. [[CrossRef](#)]
33. White, S.N.; Mousel, M.R.; Reynolds, J.O.; Herrmann-Hoesing, L.M.; Knowles, D.P. Deletion variant near ZNF389 is associated with control of ovine lentivirus in multiple sheep flocks. *Anim. Genet.* **2014**, *45*, 297–300. [[CrossRef](#)] [[PubMed](#)]
34. White, S.N.; Mousel, M.R.; Gonzalez, M.V.; Highland, M.A.; Herrmann-Hoesing, L.M.; Taylor, J.B.; Knowles, D.P. Association analysis of variant near ZNF389 with ewe cumulative production in three sheep breeds. *Anim. Genet.* **2014**, *45*, 613–614. [[CrossRef](#)] [[PubMed](#)]
35. Huntley, S.; Baggott, D.M.; Hamilton, A.T.; Tran-Gyamfi, M.; Yang, S.; Kim, J.; Gordon, L.; Branscomb, E.; Stubbs, L. A comprehensive catalog of human KRAB-associated zinc finger genes: Insights into the evolutionary history of a large family of transcriptional repressors. *Genome Res.* **2006**, *16*, 669–677. [[CrossRef](#)] [[PubMed](#)]
36. Fedotova, A.A.; Bonchuk, A.N.; Mogila, V.A.; Georgiev, P.G. C2H2 Zinc Finger Proteins: The Largest but Poorly Explored Family of Higher Eukaryotic Transcription Factors. *Acta Nat.* **2017**, *9*, 47–58. [[CrossRef](#)]
37. Consortium, B.G.S.A.; Elsik, C.G.; Tellam, R.L.; Worley, K.C.; Gibbs, R.A.; Muzny, D.M.; Weinstock, G.M.; Adelson, D.L.; Eichler, E.E.; Elnitski, L.; et al. The genome sequence of taurine cattle: A window to ruminant biology and evolution. *Science* **2009**, *324*, 522–528. [[CrossRef](#)]
38. Groenen, M.A.M.; Archibald, A.L.; Uenishi, H.; Tuggle, C.K.; Takeuchi, Y.; Rothschild, M.F.; Rogel-Gaillard, C.; Park, C.; Milan, D.; Megens, H.-J.; et al. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **2012**, *491*, 393–398. [[CrossRef](#)]
39. O’Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufu, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **2016**, *44*, D733–D745. [[CrossRef](#)] [[PubMed](#)]
40. Consortium, I.H.G.S. Finishing the euchromatic sequence of the human genome. *Nature* **2004**, *431*, 931–945. [[CrossRef](#)]
41. Massa, A.T.; Mousel, M.R.; Herndon, M.K.; Herndon, D.R.; Murdoch, B.M.; White, S.N. Genome-Wide Histone Modifications and CTCF Enrichment Predict Gene Expression in Sheep Macrophages. *Front. Genet.* **2021**, *11*, 1658. [[CrossRef](#)]
42. Won, K.-J.; Chepelev, I.; Ren, B.; Wang, W. Prediction of regulatory elements in mammalian genomes using chromatin signatures. *BMC Bioinform.* **2008**, *9*, 547. [[CrossRef](#)]
43. Creighton, M.P.; Cheng, A.W.; Welstead, G.G.; Kooistra, T.; Carey, B.W.; Steine, E.J.; Hanna, J.; Lodato, M.A.; Frampton, G.M.; Sharp, P.A.; et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 21931–21936. [[CrossRef](#)]
44. Barski, A.; Cuddapah, S.; Cui, K.; Roh, T.-Y.; Schones, D.E.; Wang, Z.; Wei, G.; Chepelev, I.; Zhao, K. High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* **2007**, *129*, 823–837. [[CrossRef](#)]
45. Wang, Z.; Zang, C.; Rosenfeld, J.A.; Schones, D.E.; Barski, A.; Cuddapah, S.; Cui, K.; Roh, T.-Y.; Peng, W.; Zhang, M.Q.; et al. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* **2008**, *40*, 897–903. [[CrossRef](#)]
46. Maurano, M.T.; Humbert, R.; Rynes, E.; Thurman, R.E.; Haugen, E.; Wang, H.; Reynolds, A.P.; Sandstrom, R.; Qu, H.; Brody, J.; et al. Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. *Science* **2012**, *337*, 1190–1195. [[CrossRef](#)] [[PubMed](#)]
47. Hindorff, L.A.; Sethupathy, P.; Junkins, H.A.; Ramos, E.M.; Mehta, J.P.; Collins, F.S.; Manolio, T.A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9362–9367. [[CrossRef](#)] [[PubMed](#)]
48. Xiang, R.; Berg, I.v.d.; MacLeod, I.M.; Hayes, B.J.; Prowse-Wilkins, C.P.; Wang, M.; Bolormaa, S.; Liu, Z.; Rochfort, S.J.; Reich, C.M.; et al. Quantifying the contribution of sequence variants with regulatory and evolutionary significance to 34 bovine complex traits. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 19398–19408. [[CrossRef](#)] [[PubMed](#)]
49. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
50. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
51. Thorvaldsdóttir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **2013**, *14*, 178–192. [[CrossRef](#)]

52. Barrett, J.C.; Fry, B.; Maller, J.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **2005**, *21*, 263–265. [[CrossRef](#)]
53. Stephens, M.; Smith, N.J.; Donnelly, P. A New Statistical Method for Haplotype Reconstruction from Population Data. *Am. J. Hum. Genet.* **2001**, *68*, 978–989. [[CrossRef](#)]
54. Stephens, M.; Donnelly, P. A Comparison of Bayesian Methods for Haplotype Reconstruction from Population Genotype Data. *Am. J. Hum. Genet.* **2003**, *73*, 1162–1169. [[CrossRef](#)] [[PubMed](#)]
55. Kel, A.E.; Gössling, E.; Reuter, I.; Cheremushkin, E.; Kel-Margoulis, O.V.; Wingender, E. MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.* **2003**, *31*, 3576–3579. [[CrossRef](#)] [[PubMed](#)]
56. Ito, J.; Gifford, R.J.; Sato, K. Retroviruses drive the rapid evolution of mammalian *APOBEC3* genes. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 610–618. [[CrossRef](#)] [[PubMed](#)]
57. Cutlip, R.C.; Lehmkuhl, H.D.; Schmerr, M.J.F.; Brogden, K.A. Ovine progressive pneumonia (maedi-visna) in sheep. *Vet. Microbiol.* **1988**, *17*, 237–250. [[CrossRef](#)]
58. Najafabadi, H.S.; Mnaimeh, S.; Schmitges, F.W.; Garton, M.; Lam, K.N.; Yang, A.; Albu, M.; Weirauch, M.T.; Radovani, E.; Kim, P.M.; et al. C2H2 zinc finger proteins greatly expand the human regulatory lexicon. *Nat. Biotechnol.* **2015**, *33*, 555–562. [[CrossRef](#)]
59. Wang, G.; Zheng, C. Zinc finger proteins in the host-virus interplay: Multifaceted functions based on their nucleic acid-binding property. *FEMS Microbiol. Rev.* **2020**. [[CrossRef](#)]
60. Wolf, G.; Greenberg, D.; Macfarlan, T.S. Spotting the enemy within: Targeted silencing of foreign DNA in mammalian genomes by the Krüppel-associated box zinc finger protein family. *Mob. DNA* **2015**, *6*, 17. [[CrossRef](#)]
61. Fu, M.; Blackshear, P.J. RNA-binding proteins in immune regulation: A focus on CCCH zinc finger proteins. *Nat. Rev. Immunol.* **2017**, *17*, 130–143. [[CrossRef](#)]
62. Crespo, H.; Bertolotti, L.; Proffitti, M.; Cascio, P.; Cerruti, F.; Acutis, P.L.; de Andrés, D.; Reina, R.; Rosati, S. Low proviral small ruminant lentivirus load as biomarker of natural restriction in goats. *Vet. Microbiol.* **2016**, *192*, 152–162. [[CrossRef](#)]
63. Heaton, M.P.; Clawson, M.L.; Chitko-Mckown, C.G.; Leymaster, K.A.; Smith, T.P.; Harhay, G.P.; White, S.N.; Herrmann-Hoesing, L.M.; Mousel, M.R.; Lewis, G.S.; et al. Reduced lentivirus susceptibility in sheep with TMEM154 mutations. *PLoS Genet.* **2012**, *8*, e1002467. [[CrossRef](#)]
64. Bowles, D.; Carson, A.; Isaac, P. Genetic distinctiveness of the Herdwick sheep breed and two other locally adapted hill breeds of the UK. *PLoS ONE* **2014**, *9*, e87823. [[CrossRef](#)]
65. Alshanbari, F.A.; Mousel, M.R.; Reynolds, J.O.; Herrmann-Hoesing, L.M.; Highland, M.A.; Lewis, G.S.; White, S.N. Mutations in *Ovis aries* TMEM154 are associated with lower small ruminant lentivirus proviral concentration in one sheep flock. *Anim. Genet.* **2014**, *45*, 565–571. [[CrossRef](#)]
66. Ramírez, H.; Echeverría, I.; Benito, A.A.; Glaria, I.; Benavides, J.; Pérez, V.; de Andrés, D.; Reina, R. Accurate Diagnosis of Small Ruminant Lentivirus Infection Is Needed for Selection of Resistant Sheep through TMEM154 E35K Genotyping. *Pathogens* **2021**, *10*, 83. [[CrossRef](#)] [[PubMed](#)]
67. Naval-Sanchez, M.; Nguyen, Q.; McWilliam, S.; Porto-Neto, L.R.; Tellam, R.; Vuocolo, T.; Reverter, A.; Perez-Enciso, M.; Brauning, R.; Clarke, S.; et al. Sheep genome functional annotation reveals proximal regulatory elements contributed to the evolution of modern breeds. *Nat. Commun.* **2018**, *9*, 859. [[CrossRef](#)] [[PubMed](#)]
68. Fang, L.; Liu, S.; Liu, M.; Kang, X.; Lin, S.; Li, B.; Connor, E.E.; Baldwin, R.L.; Tenesa, A.; Ma, L.; et al. Functional annotation of the cattle genome through systematic discovery and characterization of chromatin states and butyrate-induced variations. *BMC Biol.* **2019**, *17*, 68. [[CrossRef](#)] [[PubMed](#)]
69. Xiang, R.; MacLeod, I.M.; Daetwyler, H.D.; de Jong, G.; O'Connor, E.; Schrooten, C.; Chamberlain, A.J.; Goddard, M.E. Genome-wide fine-mapping identifies pleiotropic and functional variants that predict many traits across global cattle populations. *Nat. Commun.* **2021**, *12*, 860. [[CrossRef](#)] [[PubMed](#)]
70. Schuster, F.; Aldag, P.; Frenzel, A.; Hader, K.-G.; Lucas-Hahn, A.; Niemann, H.; Petersen, B. CRISPR/Cas12a mediated knock-in of the Polled Celtic variant to produce a polled genotype in dairy cattle. *Sci. Rep.* **2020**, *10*, 13570. [[CrossRef](#)]
71. Albert, F.W.; Kruglyak, L. The role of regulatory variation in complex traits and disease. *Nat. Rev. Genet.* **2015**, *16*, 197. [[CrossRef](#)] [[PubMed](#)]
72. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74. Available online: <https://www.nature.com/articles/nature11247#supplementary-information> (accessed on 20 February 2021). [[CrossRef](#)] [[PubMed](#)]
73. Fishilevich, S.; Nudel, R.; Rappaport, N.; Hadar, R.; Plaschkes, I.; Iny Stein, T.; Rosen, N.; Kohn, A.; Twik, M.; Safran, M.; et al. GeneHancer: Genome-wide integration of enhancers and target genes in GeneCards. *Database* **2017**, *2017*. [[CrossRef](#)] [[PubMed](#)]
74. Miller, D.E.; Patel, Z.H.; Lu, X.; Lynch, A.T.; Weirauch, M.T.; Kottyan, L.C. Screening for Functional Non-coding Genetic Variants Using Electrophoretic Mobility Shift Assay (EMSA) and DNA-affinity Precipitation Assay (DAPA). *JoVE* **2016**, e54093. [[CrossRef](#)]
75. Mathelier, A.; Shi, W.; Wasserman, W.W. Identification of altered cis-regulatory elements in human disease. *Trends Genet.* **2015**, *31*, 67–76. [[CrossRef](#)] [[PubMed](#)]
76. Sheridan, P.L.; Sheline, C.T.; Cannon, K.; Voz, M.L.; Pazin, M.J.; Kadonaga, J.T.; Jones, K.A. Activation of the HIV-1 enhancer by the LEF-1 HMG protein on nucleosome-assembled DNA in vitro. *Genes Dev.* **1995**, *9*, 2090–2104. [[CrossRef](#)] [[PubMed](#)]
77. Sato, F.; Kawai, E.; Martinez, N.E.; Omura, S.; Park, A.-M.; Takahashi, S.; Yoh, K.; Tsunoda, I. T-bet, but not Gata3, overexpression is detrimental in a neurotropic viral infection. *Sci. Rep.* **2017**, *7*, 10496. [[CrossRef](#)]