Copy number variant-based genome wide association study reveals immune-related genes associated with parasite resistance in a heritage sheep breed from the United States

Zaira M. Estrada-Reyes\(^1,2\) | Ibukun M. Ogunade\(^3\) | Andres A. Pech-Cervantes\(^4\) | Thomas H. Terrill\(^4\)

\(^1\)Department of Animal Sciences, North Carolina A&T State University, Greensboro, North Carolina, USA
\(^2\)Department of Animal Sciences, University of Florida, Gainesville, Florida, USA
\(^3\)Division of Animal and Nutritional Science, West Virginia University, Morgantown, West Virginia, USA
\(^4\)College of Agricultural, Family Sciences, and Technology, Fort Valley State University, Fort Valley, Georgia, USA

Correspondence
Zaira M. Estrada-Reyes, Department of Animal Sciences, North Carolina A&T State University, Greensboro, NC 27411, USA.
Email: zmestradareyes@ncat.edu

Funding information
Neogen® Genomics

Abstract
Florida Native is a heritage sheep breed in the United States and expresses superior ability to regulate gastrointestinal nematodes. The objective of the present study was to investigate the importance of copy number variants (CNVs) on resistance to natural Haemonchus contortus infections. A total of 300 Florida Native sheep were evaluated. Phenotypic records included fecal egg count (FEC, eggs/gram), FAMACHA© score, percentage cell volume (PCV, %), body condition score (BCS) and average daily gain (ADG, kg). Sheep were genotyped using the GGP Ovine 50K single nucleotide polymorphism (SNP) chip. Log ratios from 45.2 k SNP markers spanning the entire genome were utilized for CNV detection. After quality control, 261 animals with CNVs and phenotypic records were used for the association testing. Association tests were carried out using correlation-trend test and principal component analysis correction to identify CNVs associated with FEC, FAMACHA© score, percentage cell volume (PCV, %), body condition score (BCS) and average daily gain (ADG, kg). Significant CNVs were detected when their adjusted \(p\)-value was <.05 after FDR correction. A total of 8124 CNVs were identified, which gave 246 non-overlapping CNVs. Fourteen CNVs were significantly associated with FEC and phenotypic records were used for the association testing. Association tests were carried out using correlation-trend test and principal component analysis correction to identify CNVs associated with FEC, FAMACHA©, PCV, BCS and ADG. Significant CNVs were detected when their adjusted \(p\)-value was <.05 after FDR correction. A total of 8124 CNVs were identified, which gave 246 non-overlapping CNVs. Fourteen CNVs were significantly associated with FEC and PCV. CNVs associated with FEC overlapped 14 Quantitative Trait Locus previously associated with \(H.\ contortus\) resistance. Our study demonstrated for the first time that CNVs could be potentially involved with parasite resistance in Florida Native sheep. Immune-related genes such as \(CCL1\), \(CCL2\), \(CCL8\), \(CCL11\), \(NOS2\), \(TNF\), \(CSF3\) and \(STAT3\) genes could play an important role for controlling \(H.\ contortus\) resistance. These genes could be potentially utilized as candidate markers for selection of parasite resistance in this breed.

Keywords
copy number variants (CNVs), Florida Native sheep, parasite resistance

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. Parasite Immunology published by John Wiley & Sons Ltd.
1 | INTRODUCTION

The greatest constraint for sustainability of small ruminant systems is infection with gastrointestinal nematodes (GINs) in the southern United States (US). Several sheep breeds have been classified as resistant to GINs using fecal egg count (FEC), FAMACHA© score, and percentage cell volume (PCV).\textsuperscript{1–6} Within these breeds, Florida Native and Gulf Coast Native sheep have expressed superior control of GIN burdens.\textsuperscript{7–14} In Florida Native sheep, carcass weight and meat quality attributes at slaughter are negatively affected by infection with Haemonchus contortus when individuals are parasite susceptible (Figure 1). Haemonchus contortus, is a highly pathogenic blood-feeding GIN that is particularly devastating for young stock during times of stress, such as at weaning, and for pregnant ewes.\textsuperscript{15} For this reason, identification of parasite resistance sheep represents an important procedure in sheep breeding programs.

Previous work with Florida Native sheep has utilized single-nucleotide polymorphisms (SNPs) to identify potential candidate variants and genes associated with H. contortus resistance.\textsuperscript{13,14} However, other DNA variants, such as copy number variants (CNVs), may play an important role in phenotypic variation associated with parasite resistance. These variants include duplications and deletions that modify gene structure and dosage, alternate gene regulation and expose recessive alleles.\textsuperscript{16–18} These genetic variants are considered important genetic markers for disease susceptibility\textsuperscript{18} and can provide a better understanding of diseases that result from complex genetic patterns of inheritance.\textsuperscript{19}

In cattle, evidence from a previous study suggest that a deletion polymorphism on chromosome 7 is associated with resistance to GINs.\textsuperscript{20} For sheep, a recent study identified 31 CNVs associated with milk production traits,\textsuperscript{21} but parasite resistance has not been evaluated. In the present study, we investigated for the first time CNVs associated with parasite resistance in sheep.

2 | MATERIALS AND METHODS

2.1 | Sampling

The research protocol was approved by the University of Florida Institutional Animal Care and Use Committee (Approval number 201810108). For this study, a total of 300 sheep (3–5 months old) from a commercial farm in Ocala, Florida were used. Animals were naturally exposed to parasites since birth, kept under natural grazing conditions and grouped based on age. At the beginning of the study, initial FEC, FAMACHA© score, weight, PCV and body condition score (BCS) were measured, and animals were dewormed with levamisole (18 mg per kg of body weight) the same day as initial screening. Then, animals returned to normal grazing conditions in the commercial farm. Ten days post deworming, reduction of FEC was verified, and FAMACHA© score, weight, and BCS were recorded. This day was utilized as baseline for the study. Then, the same phenotypic variables were evaluated at 28 days post-baseline. Average daily gain (ADG) was determined based on weight and days under study. Identification of gastrointestinal parasite eggs was carried out using fresh fecal samples collected directly from the rectum of evaluated sheep, and McMaster’s FEC were performed as described by Zajac et al.\textsuperscript{22} at initial screening and during baseline (Day 0) and 28 days post baseline (Day 28). Briefly, approximately 10 g of feces were collected in plastic bags directly from the rectum of each animal, and transported from the commercial farm to the laboratory in a refrigerated container at 4°C. The samples were individually analyzed using fecal flotation and the modified McMaster method. FEC was a continuous variable expressed as the number of eggs/g feces. Egg morphology was determined using the procedures described by Zajac et al.\textsuperscript{22} and a confocal microscope. Only eggs from H. contortus were recorded for this study.

The Shapiro–Wilk test was used to test continuous (FEC, PCV) and discrete phenotypic variables (FAMACHA and BCS) for normality. Box-Cox transformation was performed to obtain a normal distribution of values and carry out logarithmic transformation (log [x + 1]) for FEC (LFEC). For all the traits (FEC, FAMACHA, PCV and BCS), data from initial screening and Day 28 was included in the association analysis. The descriptive statistics for these traits are presented in Table 1.

2.2 | Genotyping and quality control

About 4 ml of blood was collected from the jugular vein using vacutainer tubes containing ethylenediamine tetraacetic acid as anticoagulant...
from each animal. Genomic DNA was extracted from blood samples using DNeasy Blood & Tissue Kit (Qiagen) according to the manufacturer’s instructions. The DNA yield was calculated from a spectrophotometric measurement at 260 nm (NanoDrop-1000, Thermo Scientific), and the purity was assessed using a ratio of 260/280 nm.

Animals were genotyped with the commercial GGP Ovine 50K SNP chip (GeneSeek, Inc.) and intensity values from 45,205 SNPs were available for quality control procedures using Golden Helix SNP & Variation Suite (SVS) 8.7.0 software (Golden Helix, Inc.; www.goldenhelix.com). Log R ratio (LRR) values from each SNP were imported into SVS and unmapped SNPs and sex chromosomes were removed from the analysis. Derivative log ratio spread (DLRS) analysis was used to evaluate noisiness in log ratio data. Wave detection was performed to identify genomic waves in the log ratio data using the wave correction algorithm of the SVS. Samples with a median DLRS above 0.2953 and an Abs Wave Factor above 0.079 were removed. After these procedures, only 261 animals remained for further analysis. Principal component analysis (PCA) was applied to detect and correct for the presence of batch effects and to correct the LRR values.

2.3 | CNV segmentation and association testing

CNV segmentation and association testing was performed using the Golden Helix SNP & Variation Suite (SVS) 8.7.0 software. Optimal segmenting was performed in the copy number analysis module (CNAM) of the SVS software using the univariate method. The segmentation algorithm of the CNAM provides high quality results and considers only one sample at a time to detect rare or large CNVs. For this procedure, univariate outlier removal, maximum number of 100 segments per 20,000 markers, a minimum of 1 marker per segment, and 2000 permutations per pair with a p-value cutoff of .05 were used. After CNV segmentation, total number of segments for each animal in the data was examined and the sample with the highest count was removed. Then segment means were discretized using a three state (−1,0,1) model to approximate the copy number calls to potential deletions, duplications or neutral CNVs, and to reduce the influence of outliers (extremely small or large logR values) if present. Discretized segment means of CNVs were then utilized for association analysis.

Association tests were carried out using the correlation-trend test plugin in the SVS 8.7.0 software with PCA correction to identify CNVs associated with LFEC, FAMACHA©, PCV and BCS at initial screening and Day 28. Significant CNVs were detected when their adjusted p-value was <.05 after false discovery rate (FDR) correction.

2.4 | Gene functional annotation

The genes with significant CNVs associated with the traits evaluated in this study were identified using the Ovis aries v3.1 and v4 in the Genome Data Viewer genome browser (https://www.ncbi.nlm.nih.gov/genome/?term=ovis+aries). The Quantitative Trait Locus (QTL) Database (https://www.animalgenome.org/cgi-bin/QTLdb/OA/index) was used to identify QTL regions (QTLs) that overlapped or were closer (<5 Mb) to significant CNVs. We investigated gene function using Gene Ontology (GO).
RESULTS

3.1 | CNV segmentation and association testing

A total of 8124 CNVs were detected using the univariate method with an average length and median size of 88.60 and 72.18 Mb, respectively (Data S1). The CNVs covered 2.92% of the sheep autosomal genome and 2.46% of the total genome length. After aggregating the overlapping CNVs, 245 non-redundant genomic regions were identified. The average number of CNVs involved in CNVRs was 29.86 per individual. The average number of markers within each segment was 1463. Out of 245 CNVs, only 15 segment means included potential gains (duplications) and 65 segment means included potential losses (deletions). Fourteen SNPs were significantly associated with initial LFEC, Day 28 LFEC, initial PCV and PCV Day 28, respectively. Significant CNVs are shown as circular Manhattan plots in Figures 2 and 3, respectively. No significant CNVs were identified for FAMA-CHA©, BCS and ADG in this study.

For initial LFEC, 2 CNVs (one deletion and one duplication, respectively) were associated with this trait (Table 2, Figure 2). The deletion CNV and the duplication CNV were in chromosome 10 and included approximately 50 reported genes. This region also overlapped a QTL previously associated with Strongyle FEC.

For Day 28 LFEC, two deletion CNVs and one duplication were associated with this trait (Table 3, Figure 3). The first deletion CNV was in chromosome 11 and overlapped 14 QTLs previously associated with H. contortus resistance. This CNV included approximately 62 reported genes. The second deletion CNV and the duplication CNV were both identified in chromosome 26 and included XKR5, ZNF596 genes and LOC101106907 gene, respectively. No QTL previously associated with FEC, or parasite resistance overlapped these CNVs.

Nine significant CNVs identified in chromosomes 4, 10, 12 and 17 were simultaneously associated with initial PCV and PCV Day 28 (Tables 4 and 5, Figures 4 and 5). These CNVs included six potential deletions and three potential duplications. The CNVs in chromosome 10 overlapped the same QTL that contained the two CNVs for Initial FEC. The deletion CNV in chromosome 12 was located closer to a QTL that was previously associated with H. contortus resistance.

Overall, results from GO analysis (Table 6) showed that functions of the proteins encoded by these genes included several biological events such as immune system process, response to stimulus, signaling, biological regulation, cellular process, interspecies interaction...
### TABLE 3  
Significant CNVs associated with Day 28 LFEC in Florida Native sheep naturally infected with *Haemonchus contortus*

| Trait       | CNV p-Value | Chr | Start position | End position | Length (bp) | No. markers | Segment mean | Genes                                                                 | QTL                                                                 |
|-------------|-------------|-----|----------------|--------------|-------------|-------------|--------------|----------------------------------------------------------------------|----------------------------------------------------------------------|
| Day 28 LFEC | 3           | 7.8E-4 | 11 15,165,032 | 66,805,603   | 51,640,572  | 991         | −0.029       | CCL1, CCL8, CCL11, CCL2, NOS2, TNAIP1, TRAF4, CSF3, ITGA3, ITGA2B, ITGB3, CCR7, STAT5B, STAT5A, STAT3, GDHC, IFI35, CD7, LOC101117947, LOC101118202, LOC101102156, LOC101120341, CRLF3, CDK5R1, PSMD11, ZNF207, RNF135, TEFM, SUZ12, UTP6, RAB11FIP4, NF1, EVI2A, OMG, LOC101117683, LYRM9, NLR, SEBOX, VTN, SLC46A1, SLC13A, UNC1192, SCL10A2, TMEM255B, STRADA | IgA (50); 180542, 180544, 180551, 180504, 180516, 180528, 180541, 180543, 180556, 180505, 180530, 180557, 180545, 180529: H. contortus resistance |

| Trait       | CNV p-Value | Chr | Start position | End position | Length (bp) | No. markers | Segment mean | Genes                                                                 | QTL                                                                 |
|-------------|-------------|-----|----------------|--------------|-------------|-------------|--------------|----------------------------------------------------------------------|----------------------------------------------------------------------|
| Day 28 LFEC | 4           | 5.7E-4 | 26 5,600,812  | 5,629,890    | 29,079      | 5           | −0.274       | XKR5, ZNF596                                                          |                                                                      |

4  | 5.5E-4 | 26 10,633,898  | 10,633,898  | 1 | 3 | 0.441 | LOC101106907  |

Abbreviations: CNV, copy number variant; FEC, fecal egg count; LFEC, log-transformed fecal egg count; QTL, Quantitative Trait Locus.

### TABLE 4  
Significant CNVs associated with initial PCV in Florida Native sheep naturally infected with *Haemonchus contortus*

| Trait       | CNV p-Value | Chr | Start position | End position | Length (bp) | No. markers | Segment mean | Genes                                                                 | QTL                                                                 |
|-------------|-------------|-----|----------------|--------------|-------------|-------------|--------------|----------------------------------------------------------------------|----------------------------------------------------------------------|
| Initial PCV | 6           | 9.2E-18 | 4 114,722,181 | 114,742,686  | 20,506      | 3           | −3.674       | LOC101116394  |

7  | 2.3E-7 | 4 54,630 | 127,201,684  | 127,147,055  | 1913        | 0.035       | LOC101108909, LOC101109171, LOC101109440, LOC101109699 |

8  | 8.9E-7 | 10 71,082,310 | 94,127,923  | 23,045,614  | 304         | −0.035       | LOC101108321, LOC101108582, LOC101109105, LOC101109890, CLDNL10, DZIP1, DNAJC3, UGGT2, HS6ST3, OXCR1, M6B1L2, RAP2A, IPO5, STK24, PEPT1, SLC15A1, DOCK9, GPR18, LOC101111720, ITGB1L, NALCN, FGF14, TPP2, POGLUT2, BIVM, LOC101114712, SLC10A2, EFNB2, LG4, ABHD13, TNF513B, MYO16, COL4A1, COL4A2, ING1, ANKRD10, TUBGCP3, F7, F10, PCID2, COL4A, LAMP1, GPR1, TMEM255B, GAS6, RAS3A, CDC16 |

9  | 5.6E-5 | 10 71,168,171 | 71,168,220  | 50          | 2           | −2.432       | LOC101108582  |

10 | 9.1E-5 | 10 71,168,220 | 71,180,524  | 12,305      | 5           | 1.114        | LOC101108582, LOC101109370  |

11 | 5.9E-4 | 10 70,905,390 | 94,127,923  | 23,222,534  | 306         | 0.006        | LOC101108321, LOC101108843, LOC101109105, LOC101109370, LOC101107612, LOC101109890, |

(Continues)
### TABLE 4 (Continued)

| Trait | CNV p-Value | Chr | Start position | End position | Length (bp) | No. markers | Segment mean | Genes                                                                 | QTL |
|-------|-------------|-----|----------------|--------------|-------------|-------------|--------------|----------------------------------------------------------------------|-----|
|       |             |     | 25,228,638     | 25,228,638   | 1           | 2           | -4.589       | CAPN2                                                                | 180546: H. contortus resistance* |
|       |             |     | 4,832,850      | 5,979,442    | 1,146,593   | 52          | -0.061       | TMEM154, FBXW7, LOC10112092, -PET112, FAM160A1                          |     |
|       |             |     | 4,832,878      | 4,832,881    | 4           | 5           | -0.631       | TMEM154                                                                |     |

Abbreviations: CNVs, copy number variants; PCV, percentage cell volume; QTL, Quantitative Trait Locus.

*QTL is located close to the significant CNV.

### TABLE 5

Significant CNVs associated with PCV Day 28 in Florida Native sheep naturally infected with *Haemonchus contortus*

| Trait | CNV p-Value | Chr | Start position | End position | Length (bp) | No. markers | Segment mean | Genes                                                                 | QTL |
|-------|-------------|-----|----------------|--------------|-------------|-------------|--------------|----------------------------------------------------------------------|-----|
| Initial PCV | 6 | 8.6E-16 | 4 | 114,722,181 | 114,742,686 | 20,506 | 3 | -3.674 | LOC101116394 | - |
| | 7 | 5.1E-7 | 4 | 54,630 | 127,201,684 | 127,147,055 | 1913 | 0.035 | LOC101108909, LOC101109171, LOC101109440, LOC101109699 | - |
| | 8 | 1.6E-6 | 10 | 71,082,310 | 94,127,923 | 23,045,614 | 304 | -0.035 | LOC101108321, LOC101108582, LOC101109890, CLDN10, DZIP1, DNAJC3, UGTT2, H56ST3, OXGR1, MBNL2, RAP2A, IP05, STK24, PEPT1, SLCL15A1, DOCK9, GPR18, LOC10111720, ITGLB1, NALCN, FGF14, TPP2, POGLUT2, BIVM, LOC101114712, SLCL10A2, EFNB2, LIG4, ABHD13, TNFSF13B, MYO16, COL4A1, COL4A2, ING1, ANKR10, TUBGCP3, F7, F10, PCID2, CUL4A, LAMP1, GRTP1, TMEM255B, GAS6, RASA3, CDC16 | FEC (50): 193038: Strongyle FEC |
| | 9 | 7.7E-5 | 10 | 71,168,171 | 71,168,220 | 50 | 2 | -2.432 | LOC101108582 | - |
| | 10 | 1.2E-4 | 10 | 71,168,220 | 71,180,524 | 12,305 | 5 | 1.114 | LOC101108582, LOC101109370 | - |
| | 11 | 6.5E-4 | 10 | 70,905,390 | 94,127,923 | 23,222,534 | 306 | 0.006 | LOC101108321, LOC10110843, LOC101109105, LOC101109370, LOC101107612, LOC101109890, CLDN10, | - |
between organisms, localization, metabolic process, multicellular organismal process, integrin signaling, molecular regulation, blood coagulation, biological regulation, cellular process, and developmental process.

| Trait | CNV | p-Value | Chr | Start position | End position | Length (bp) | No. markers | Segment mean | Genes |
|-------|-----|---------|-----|----------------|--------------|-------------|-------------|--------------|-------|
| 12    | 5.9E-13 | 12 | 25,228,638 | 25,228,638 | 1 | 2 | -4.589 | DZIP1, LOC101110424, DNAJC3, UGGT2, HS6ST3, OXGR1, MBNL2, RAP2A, IPO5, STK24, PEPT1, SLC15A1, DOCK9, GPR18, LOC10111720, ITGBL1, NALCN, FGFL, TPP2, POGLUT2, BVVM, LOC10111712, SLC10A2, EFNB2, LIG4, ABHD13, TNFSF13B, MYO16, COL4A1, COL4A2, ING1, ANKRIT, TUBGCP3, F7, F10, PCD2, CUL4A, LAMMP1, GRTP1, TMEM2558, GAS5, RASA3, CDC16 |
| 13    | 1.8E-20 | 17 | 4,832,850 | 5,979,442 | 1,146,593 | 52 | -0.061 | TMEM154, FBXW7, LOC1011102092, PET112, FAM160A1 |
| 14    | 8.9E-11 | 17 | 4,832,878 | 4,832,881 | 4 | 5 | -0.631 | TMEM154 |

Abbreviations: CNVs, copy number variants; PCV, percentage cell volume; QTL, Quantitative Trait Locus.

\( ^a \)QTL is located close to the significant CNV.

**TABLE 5** (Continued)

**FIGURE 4** Circular Manhattan plot for significant copy number variants associated with initial PCV in Florida Native sheep naturally infected with *Haemonchus contortus*

**FIGURE 5** Circular Manhattan plot for significant copy number variants associated with PCV Day 28 in Florida Native sheep naturally infected with *Haemonchus contortus*

**4 | DISCUSSION**

Several of the SNP-chip arrays from Illumina and Affymetrix are commonly utilized for detection of SNPs and CNVs simultaneously and
and CNV-based GWA studies have observed several parasite resistance.

Studies with pigs have reported 35 CNVs associated with milk production traits (daily milk yield, milk fat percentage, protein percentage, protein yield, and milk somatic cell count) in dairy sheep.21 However, CNVs controlling parasite resistance.

In sheep, it is estimated that approximately 6.9% of the sheep genome includes CNVs and, a recent study identified 31 CNVs associated with milk production traits (daily milk yield, milk fat percentage, protein percentage, protein yield, and milk somatic cell count).

Detection of CNVs from SNP data and CNV-based GWA studies have been commonly utilized in livestock studies20,21 to provide valuable insights on the genetic control of complex traits. In cattle, it has been estimated that approximately 2% to 7% of the genome includes CNVs.23 and CNV-based GWA studies have observed several variants associated with milk production24 and parasite resistance.20,25,26 Studies with pigs have reported 35 CNVs associated with high and low fertility.27 and CNVs within immune response genes have been related to resistance to PRRS virus.28 A recent publication with broiler chickens utilized CNV-based GWAS to reveal potential candidate genes that may regulate performance traits.29

Consider ‘unSNPable’ genome probes that allow the study of areas in the genome with high segmental duplication. In our study, we utilized the GGP Ovine 50K array which includes 15,000 SNPs from the GGP Ovine 15K, over 10,000 SNPs from the existing Illumina 50K array, more than 12,000 SNPs from Illumina HD content, and over 1700 SNPs to fill in spatial gaps within the genome. In addition, the GGP Ovine 50K includes 9000 SNPs from the new novel ovine sequence not previously included in earlier genotyping products. This array is suitable for identification of rare or large CNVs and provides an increased coverage for SNP genotyping. While the data from this study is informative, the large and rare CNVs identified in this work may not represent all the CNVs distributed in the sheep genome.

### Table 6

*Gene ontology terms for the significant copy number variants (CNVs) associated with initial fecal egg count (initial FEC), fecal egg count at Day 28, and parasite resistance.*

| Trait | Genes | GO term | GO term name |
|-------|-------|---------|--------------|
| **Initial FEC** | LOCI01108321, LOCI01108843, LOCI01109015, LOCI01109370, LOCI01107612, LOCI01109890, CLDN10, DZIP1, LOC101110424, DNAJC3, UGGT2, H56ST3, OXR1, MBNL2, RAP2A, IPO5, STK24, PEP1, SLCL154A1, DOCK9, GPR18, LOCI01111720, ITGB1L, NALCN, FGF14, TTP2, POGLUT2, BIVM, LOCI01114712, SLC10A2, ENFB2, LI4G, ABHD13, TNFSF13B, MYO16, COL4A1, COL4A2, ING1, ANKRD10, TUBGCP3, F7, F10, PCID2, CUL4A, LAMP1, GRTPI, TMEM255B, GASA, RASA3, CDC16 | GO:0005488 | Binding/catalytic activity, structural molecule activity, integrin signaling, molecular function regulator, blood coagulation, biological regulation, cellular process, developmental process, localization, immune system process, response to stimulus, signaling, locomotion, metabolic process, multicellular organismal process |
| PCV | LOCI01108321, LOCI01108843, LOCI01109015, LOCI01109370, LOCI01107612, LOCI01109890, CLDN10, DZIP1, LOC101110424, DNAJC3, UGGT2, H56ST3, OXR1, MBNL2, RAP2A, IPO5, STK24, PEP1, SLCL154A1, DOCK9, GPR18, LOCI01111720, ITGB1L, NALCN, FGF14, TTP2, POGLUT2, BIVM, LOCI01114712, SLC10A2, ENFB2, LI4G, ABHD13, TNFSF13B, MYO16, COL4A1, COL4A2, ING1, ANKRD10, TUBGCP3, F7, F10, PCID2, CUL4A, LAMP1, GRTPI, TMEM255B, GASA, RASA3, CDC16, CAPN2, TMEM154, FBXW7, LOCI01102092, PET112, FAM160A1, LOCI01108940, LOCI01109699, LOCI01116394 | GO:0005488 | Binding/catalytic activity, structural molecule activity, integrin signaling, molecular function regulator, blood coagulation, biological regulation, cellular process, developmental process, localization, immune system process, response to stimulus, signaling, locomotion, metabolic process, multicellular organismal process |

Note: (FEC Day 28) and PCV in Florida Native sheep.
resistance have not been elucidated in this species. Our study identified for the first time CNVs significantly associated with FEC and PCV, respectively. We identified 245 CNV regions with a small count of CNV-overlapping genes and more abundant copy losses than gains. Similarly, more CNV losses than gains have been previously observed in samples derived from the sheep HapMap populations.

For initial FEC, it is possible that genes within significant CNVs such as LOC101110424, DOCK7, ITGBL1, ITGBM, TNFSF13B, ING1, F7, F10, PCID2 and GAS6 genes could have important effects on immune response mechanisms against *H. contortus*. For example, the ferritin heavy chain-like (LOC101110424) gene encodes a protein that binds B, CD4+ and CD8+ T-lymphocytes and regulates proliferation. The ITGBL1 gene expression is associated with infiltration of immune cells. The TNFSF13B (B cell activation factor) proteins can function like cytokines that modulate immune cells and inflammatory response. The F7 and F10 genes play a key role in blood coagulation initiation and defense against pathogens. The PCID2 gene has been implicated in regulation of B-cell development, and GAS6 gene promotes sequestration of leukocytes in endothelium. Further studies are required to evaluate the role of these two genes in sheep infected with *H. contortus*.

For LFEC at Day 28, significant CNVs included genes related to immune response, such as CCL1, CCL2, CCL8, CC11, CRLF3, CCR7, NOS2, TNFAIP1, TNF, STAT3, STAT5A, CSF3, GHDC, IFI35, ITGA3, ITGA2B and ITGB3 genes. For example, the CCL1, CCL2 and CCL8 genes are chemokine genes that have been proposed as candidate markers for parasite resistance in Scottish Blackface lambs. The CCL11 gene is part of the eotaxin chemokines and promotes migration of activated eosinophils. Eosinophilia is a common event observed in sheep infected with *H. contortus* and it is utilized as a phenotypic marker for parasite resistance. The CCR7 (C-C motif chemokine receptor 7) gene encodes a receptor in T lymphocytes, which regulate lymphocyte mobilization during inflammatory processes. The NOS2 (nitric oxide synthase 2) is a gene under directional selection in *H. contortus* resistant sheep and promotes production of nitric oxide to kill invading microbes during classical macrophage activation (Figure 4). The tumor necrosis factor (TNF) gene encodes a proinflammatory cytokine important for immunity and homeostasis and represents a potential genetic marker for *H. contortus* resistance due to its previous association with neutrophil count in Florida Native sheep. The STAT3, STAT5 and STAT5A genes belong to the STAT family genes known as signal transducers and activators of transcription. Previous studies with Florida Native sheep have observed that polymorphisms within STAT3 gene are associated with FEC in Florida Native sheep (Figure 6). The CSF3 (colony stimulating factor 3) gene encodes a cytokine that promotes production of granulocytes and differentiation into Th2/Treg responses and it is under directional selection in sheep selected for *H. contortus* resistance. Additionally, three galectin genes (LOC101117947, LOC101118202 and LOC101102156) genes within the CNV
associated with Day 28 LFEC. Galectins are proteins involved in the immune response to GIN infection in sheep and are upregulated during H. contortus infection. Some of these galectins, such as galectin-11, can regulate larval growth and development by binding L4 and adult H. contortus.

For initial PCV and PCV Day 28, the same CNV regions were associated with these traits. Most of the genes contained in these CNVs were related to cellular processes and several encoded the multidrug resistance proteins. These proteins are normally expressed in CD3/CD4 T-cells from peripheral blood and can regulate mucosal inflammation.

Most of the significant CNVs overlapped QTLs associated with FEC and IgA in adult sheep. Our study also confirmed previous candidate genes (CCL1, CCL2, CCL8, CCL11, NOS2, TNF, CSF3 and STAT3) associated with parasite resistance in Florida Native and other sheep breeds. Since this Florida Native sheep population has been selected for parasite resistance, it is possible that the CNVs identified in our study could be segregating in the population, but validation of our findings is required. Our results may contribute to the development of new strategies for improving parasite resistance and to promote selective breeding and marker-assisted selection in this heritage sheep breed.

5 | CONCLUSION

This study confirmed for the first time that CNVs could be potentially involved with parasite resistance in Florida Native sheep. Immune-related genes such as CCL1, CCL2, CCL8, CCL11, NOS2, TNF, CSF3 and STAT3 genes could play an important role for controlling H. contortus resistance. These genes could be potentially utilized as candidate markers for selection of parasite resistance in this breed.

AUTHOR CONTRIBUTIONS

Zaira M. Estrada-Reyes conceived and designed the project. Zaira M. Estrada-Reyes, Ibukun M. Ogunade, Andres A. Pech-Cervantes and Thomas H. Terrill analyzed and interpreted the data. All authors prepared the draft and approved the final manuscript.

ACKNOWLEDGEMENTS

We gratefully acknowledge Fairmeadow farm for providing the animals under study. We acknowledge the College of Veterinary Medicine and the Department of Animal Sciences at the University of Florida for assisting during sample collection. We thank Neogen Genomics for providing free genotyping using the GGP Ovine SNP array.

FUNDING INFORMATION

This research was partially supported by Neogen Genomics.

CONFLICT OF INTEREST

The authors declare that there were no conflicting interests that could have influenced the conduct and reporting of this study.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/pim.12943.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

REFERENCES

1. Notter DR, Burke JM, Miller JE, Morgan JLM. Association between FAMACHA scores and fecal egg counts in Katahdin lambs. J Anim Sci. 2017;95(3):1118-1123. doi:10.2527/jas.2016.1248
2. Kaplan RM, Burke JM, Terrill TH, et al. Validation of the FAMACHA© eye color chart for detecting clinical anemia in sheep and goats on farms in the southern United States. Vet Parasitol. 2004;123(1–2):105-120. doi:10.1016/j.vetpar.2004.06.005
3. Vaninimetti HB, Greiner SP, Zajac AM, Notter DR. Performance of hair sheep composite breeds: resistance of lambs to Haemonchus contortus. J Animal Sci. 2004;82(2):595-604. doi:10.2527/2004.822595x
4. MacKinnon KM, Zajac AM, Kooyman FNJ, Notter DR. Differences in immune parameters are associated with resistance to Haemonchus contortus in Caribbean hair sheep. Parasite Immunol. 2010;32(7):484-493. doi:10.1111/j.1365-3024.2010.01211.x
5. Bowdridge SA, Zajac AM, Notter DR. St. Croix sheep produce a rapid and greater cellular immune response contributing to reduced establishment of Haemonchus contortus. Vet Parasitol. 2015;208(3–4):204-210. doi:10.1016/j.vetpar.2015.01.019
6. Burke JM, Notter DR, Miller JE. Sire’s estimated breeding value (EBV) for fecal egg count (FEC) influences BW, FEC, and anemia measures of offspring in Katahdin Sheep. J Anim Sci. 2017;95:52. doi:10.2527/ ssasas2017.0105/
7. Miller JE, Bahirathan M, Lemarie SL, Hembry FG, Kearney MT, Barras SR. Epidemiology of gastrointestinal nematode parasitism in Suffolk and Gulf Coast Native sheep with special emphasis on relative susceptibility to Haemonchus contortus infection. Vet Parasitol. 1998;74(1):55-74. doi:10.1016/S0304-4017(97)00094-0
8. Amarante AF, Craig TM, Ramsey WS, El-Sayed NM, Desouki AY, Bazer FW. Comparison of naturally acquired parasite burdens among Florida Native, Rambouillet and crossbreed ewes. Vet Parasitol. 1999a;85(1):61-69. doi:10.1016/S0304-4017(99)00010-4
9. Amarante AF, Craig TM, Ramsey WS, Davis SK, Bazer FW. Nematode burdens and cellular responses in the abomasal mucosa and blood of Florida Native, Rambouillet and crossbreed lambs. Vet Parasitol. 1999;80(4):311-324. doi:10.1016/S0304-4017(98)00229-5
10. Miller JE, Bishop SC, Cockett NE, McGraw RA. Segregation of natural and experimental gastrointestinal nematode infection in F2 progeny of susceptible Suffolk and resistant Gulf Coast Native sheep and its usefulness in assessment of genetic variation. Vet Parasitol. 2006;140(1–2):83-89. doi:10.1016/j.vetpar.2006.02.043
11. Shakya KP, Miller JE, Horohov DW. A Th2 type of immune response is associated with increased resistance to Haemonchus contortus in naturally infected Gulf Coast Native lambs. Vet Parasitol. 2009;163(1–2):57-66. doi:10.1016/j.vetpar.2009.03.052
12. Shakya KP, Miller JE, Lomax LG, Burnett DD. Evaluation of immune response to artificial infections of Haemonchus contortus in Gulf Coast Native compared with Suffolk lambs. Vet Parasitol. 2011;181(2–4):239-247. doi:10.1016/j.vetpar.2011.03.051
13. Estrada-Reyes ZM, Rae O, Postley C, Jiménez Medrano MB, Leal Gutiérrez JL, Mateescu RG. Association study reveals Th17, Treg, and Th2 loci related to resistance to Haemonchus contortus in Florida Native sheep. Vet Parasitol. 2009;163(1–2):57-66. doi:10.1016/j.vetpar.2009.03.052
50. Atlija M, Arranz JJ, Martínez-Valladares M, Gutiérrez-Gil B. Detection and replication of QTL underlying resistance to gastrointestinal nematodes in adult sheep using the ovine 50K SNP array. Genet Sel Evol. 2016;48(1):1-6. doi:10.1186/s12711-016-0182-4

SUPPORTING INFORMATION
Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Estrada-Reyes ZM, Ogunade IM, Pech-Cervantes AA, Terrill TH. Copy number variant-based genome wide association study reveals immune-related genes associated with parasite resistance in a heritage sheep breed from the United States. Parasite Immunol. 2022;44(11):e12943. doi:10.1111/pim.12943