# Comparing Class II MHC DRB3 Diversity in Colombian Simmental and Simbrah Cattle Across Worldwide Bovine Populations

Diego Ordoñez[1,2†], Michel David Bohórquez[3,4†], Catalina Avendaño[1] and Manuel Alfonso Patarroyo[3,5,6]*

[1]Animal Science Faculty, Universidad de Ciencias Aplicadas y Ambientales (U.D.C.A), Bogotá, Colombia, [2]PhD Program in Tropical Health and Development, Universidad de Salamanca, Salamanca, Spain, [3]Molecular Biology and Immunology Department, Fundación Instituto de Inmunología de Colombia (FIDIC), Bogotá, Colombia, [4]MSc Program in Microbiology, Universidad Nacional de Colombia, Bogotá, Colombia, [5]Health Sciences Division, Main Campus, Universidad Santo Tomás, Bogotá, Colombia, [6]Microbiology Department, Faculty of Medicine, Universidad Nacional de Colombia, Bogotá, Colombia

The major histocompatibility complex (MHC) exerts great influence on responses to infectious diseases and vaccination due to its fundamental role in the adaptive immune system. Knowledge about MHC polymorphism distribution among breeds can provide insights into cattle evolution and diversification as well as population-based immune response variability, thus guiding further studies. Colombian Simmental and Simbrah cattle's *BoLA-DRB3* genetic diversity was compared to that of taurine and zebuine breeds worldwide to estimate functional diversity. High allele richness was observed for Simmental and Simbrah cattle; nevertheless, high homozygosity was associated with individual low sequence variability in both the β1 domain and the peptide binding region (PBR), thereby implying reduced MHC-presented peptide repertoire size. There were strong signals of positive selection acting on *BoLA-DRB3* in all populations, some of which were poorly structured and displayed common alleles accounting for their high genetic similarity. PBR sequence correlation analysis suggested that, except for a few populations exhibiting some divergence at PBR, global diversity regarding potential MHC-presented peptide repertoire could be similar for the cattle populations analyzed here, which points to the retention of functional diversity in spite of the selective pressures imposed by breeding.

Keywords: MHC, BoLA-DRB3, genetic diversity, peptide-binding region, genetic resistance, cattle

## INTRODUCTION

The major histocompatibility complex (MHC) plays an important effector role in the adaptive immune response (Rock et al., 2016). This particular system offers a unique opportunity for addressing functional and evolutionary diversity issues in many species (Trtkova et al., 1995). The MHC is formed by a group of loci encoding specific cell surface glycoproteins which are necessary for T-lymphocyte antigen peptide recognition (Rock et al., 2016). MHC class I proteins are expressed by all nucleated cells and are related to presenting antigens to CD8[+] T-lymphocytes processed in the intracellular compartment, thereby eliciting cytotoxic responses (Neefjes et al., 2011; Rock et al., 2016). By contrast, MHC class II proteins are expressed by professional antigen-presenting cells and are associated with the presentation of extracellular antigen peptides to CD4[+] T-lymphocytes,

triggering cellular or humoral responses against various pathogens (Neefjes et al., 2011; Rock et al., 2016).

The MHC in cattle has been called the bovine leukocyte antigen (BoLA) and the genes encoding the expression of class II antigen presentation (DR and DQ)-related molecules are located in chromosome 23 *IIa* subregion (Ellis, 2004). BoLA-DR consists of the *DRA* monomorphic locus and three *DRB* loci, of which *BoLA-DRB3*—characterized by a high degree of polymorphism, with 330 different alleles reported to date—is the only one which has been described as functional (Maccari et al., 2017). Such polymorphisms are concentrated in the second exon which encodes the peptide binding region (PBR) β1 domain and has been used for determining *BoLA-DRB3* alleles (Sigurdardóttir et al., 1991). Such high variability determines the amino acids (aa) forming PBR binding pockets, influencing the peptides presented on MHC for different alleles and setting different repertoires modulating the immune response (Ellis, 2004; Baxter et al., 2009). *BoLA-DRB3* diversity could be used for estimating potential peptide-binding repertoire size, based on the assumption that highly divergent alleles are associated with broader peptide spectra (Klein et al., 2007).

Assessing MHC genetic variability is of interest for animal science researchers, breeders, and evolutionary biologists. Variations regarding susceptibility to infectious diseases (Dietz et al., 1997; Acosta-Rodríguez et al., 2005; Martinez et al., 2006; Nascimento et al., 2006; Kulberg et al., 2007; Juliarena et al., 2008; Nikbakht Brujeni et al., 2016; Carignano et al., 2017), vaccine response (Bacon and Witter, 1995; García-Briones et al., 2000; Rupp et al., 2007; Baxter et al., 2009; Gowane et al., 2013) and production traits (Sharif et al., 1999; Rupp et al., 2007) have been associated with different *BoLA-DRB3* alleles, and information regarding their frequency has thus been used for running infectious disease control programs (Maillard et al., 2003). MHC variability patterns reflect evolutionary processes such as adaptation, selection (natural, sexual or artificial) and genetic diversity within and between populations (Trtkova et al., 1995; Edwards and Hedrick, 1998; Uinuk-Ool et al., 2002; Ujvari and Belov, 2011). Some studies have shown that decreased MHC variability (often higher rates than of neutral loci (Sutton et al., 2011)) might be caused by population bottlenecks (Bollmer et al., 2011; Mason et al., 2011; Babik et al., 2012; Taylor et al., 2012; Zhang et al., 2016), while others have shown that a high level of diversity could be maintained by pathogen-driven balancing selection or other mechanisms, despite extreme population decline (Mikko et al., 1997; Edwards and Hedrick, 1998; Garrigan and Hedrick, 2001; Aguilar et al., 2004; Borg et al., 2011; Galaverni et al., 2013; Moutou et al., 2013; Newhouse and Balakrishnan, 2015). Such information can be used for prioritizing other types of study, such as immunopeptidomic or binding affinity assays used for developing peptide-MHC *in silico* interaction predictive algorithms that could provide feedback regarding immune system response and evolution (Nielsen et al., 2008; Nielsen et al., 2018).

The Simmental is a cattle (*Bos taurus*) breed that was selected in North America and Europe, mainly for increasing meat production efficiency (Amaya et al., 2020). The introduction of Simmental to Colombia 5 decades ago aimed to increase both

milk and beef production by artificial insemination-based genetic improvement schemes using semen from proven bulls in North America and Europe (Amaya et al., 2020). Simbrah is considered a composite breed developed to combine Brahman cattle (*Bos indicus*) adaptability, maternal instinct, hardiness and disease resistance with Simmental fertility, milk production and beef quality (Goszczynski et al., 2018; Amaya et al., 2020). Most tropical countries where Simbrah cattle occur have chosen a different breeding strategy, producing animals with different percentages of Zebuine genes, ranging from 1/4 (25%) to 5/8 (62.5%) based on the requirement of particular features, such as better adaptation to humid environments (Agung et al., 2016).

Despite recent advances in exploring *BoLA-DRB3* genetic diversity in cattle, a significant amount of breeds and crossbreeds still remain uncharacterized (Takeshima et al., 2003; Giovambattista et al., 2013; Takeshima et al., 2014; Takeshima et al., 2015; Takeshima et al., 2018). The aim of this study, therefore, was to describe for the first time *BoLA-DRB3* genetic diversity in the Colombian Simmental breed and its common zebuine cross, Simbrah, comparing it with that of worldwide taurine and zebuine breeds to assess the impact on potential peptide-binding repertoire size and divergence. Such new MHC diversity information will assist in introducing appropriate breeding schemes, guiding further MHC studies.

## MATERIALS AND METHODS

### Study Population and DNA Extraction

Whole blood was collected from the coccygeal or jugular veins of 130 Simmental (N = 67; 5 farms) and Simbrah cattle (N = 60; 5 farms) (Supplementary Data S1), stored in EDTA-containing vacutainer tubes. Bovines were selected from extensive production systems from Colombia's main breeding regions characterized by a reduced number of purebred animals per farm. The herds and purebred animals analyzed were randomly sampled avoiding related individuals. Genomic DNA (gDNA) was extracted using the PureLink Genomic DNA Mini Kit (Invitrogen, Carlsbad, CA, United States) and following the manufacturer's instructions. Previous allelic richness data (Greenbaum et al., 2014) of 14 taurine and zebuine populations from Asia, South America and Europe were included for comparison (**Table 1**). This study was carried out following the protocol approved by the *Universidad de Ciencias Aplicadas y Ambientales*' (U.D.C.A) Animal Research Ethics Committee (minutes No.201901).

### DNA Amplification and Sequencing

*BoLA-DRB3* exon 2 was amplified with primers DRB3F (5′-TCC CGCATTGGTGGGTGT-3′) and DRB3R (5′-CTCCACACT GGCCGTCCAC-3′) (Ledwidge et al., 2001). The PCR mixture contained 1X Pfx amplification buffer, 300 mM of each dNTP, 0.45 mM of each primer, 1 mM MgSO$_4$, 1 U Platinum Pfx DNA Polymerase (Invitrogen) and 50 ng gDNA, in a 50 ml final volume. Two independent reactions were performed for each sample, following previous recommendations to avoid chimeric product formation (Lenz and Becker, 2008). The thermal profile

**TABLE 1 |** General information regarding the 16 cattle populations analyzed in this study.

| Breed | Acronym | N | Farms/places | Country | Type | Reference |
|---|---|---|---|---|---|---|
| Simmental | SmtCo | 67 | 5 | Colombia | taurine | This study |
| Simbrah | SbhCo | 60 | 5 | Colombia | taurine | This study |
| Normande | NorCo | 111 | 14 | Colombia | taurine | Bohórquez et al. (2020) |
| Morucha | MorSp | 54 | 15 | Spain | taurine | Bohórquez et al. (2020) |
| Nellore | NeBo | 116 | 2 | Bolivia | zebuine | Takeshima et al. (2018) |
| Nellore x Brahman | NeBrPe | 195 | 1 | Peru | zebuine | Takeshima et al. (2018) |
| Holstein | HolAr | 413 | 4 | Argentina | taurine | Takeshima et al. (2015) |
| Holstein | HolBo | 153 | 2 | Bolivia | taurine | Takeshima et al. (2015) |
| Holstein | HolPa | 127 | 5 | Paraguay | taurine | Takeshima et al. (2015) |
| Holstein | HolPe | 132 | 2 | Peru | taurine | Takeshima et al. (2015) |
| Holstein | HolCh | 113 | 5 | Chile | taurine | Takeshima et al. (2015) |
| Yacumeño | YacBo | 100 | 4 | Bolivia | taurine | Giovambattista et al. (2013) |
| Holstein | HolJa | 101 | Random collection | Japan | taurine | Takeshima et al. (2003) |
| Brahman | BrPh | 233 | 2 | Philippines | zebuine | Takeshima et al. (2014) |
| Native x Brahman | NaBrPh | 131 | 4 | Philippines | zebuine | Takeshima et al. (2014) |
| Native | NaPh | 480 | 4 | Philippines | zebuine | Takeshima et al. (2014) |

consisted of a denaturation step at 94°C for 5 min followed by 30 cycles of 94°C for 30 s, 64°C for 30 s and 68°C for 1 min, with no final extension. Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, United States) were used for purifying the amplicons according to the manufacturer's instructions prior to sequencing both directions using the BigDye Terminator Kit.

## Sequence Data Analysis

CLC Main Workbench (CLC bio, Aarhus, Denmark) was used for assembling and editing each sequence independently. Polymorphic positions were recognized for producing a final consensus sequence containing IUPAC ambiguity codes. HAPLOFINDER (Miltiadou et al., 2003) was used for assigning each animal's genotype by comparing the obtained sequences to the *BoLA-DRB3* allele sequences reported in the IPD-MHC database (Maccari et al., 2017), as previously indicated (Baxter et al., 2008).

## Sequence Diversity, Hardy-Weinberg Equilibrium and Selection Signatures

The number of alleles ($N_a$) and allele frequencies were manually obtained by direct counting using the maximum likelihood method for estimating standard errors for allele frequencies according to Li (Li, 1976). Nei and Chesser's method (Nei, 1978) was used for calculating the observed heterozygosity ($h_o$) and ARLEQUIN v.3.5 (Excoffier and Lischer, 2010) for estimating the unbiased expected heterozygosity ($h_e$). The correlation between sample size and these estimators was used for assessing their dependence. Allele richness (a measure of the average amount of alleles per locus) was also used for comparing the number of alleles found in each population independently from sample size (Greenbaum et al., 2014). The $F_{IS}$ index (Weir and Cockerham, 1984) was estimated for determining potential departures from Hardy-Weinberg equilibrium using the exact test of significance implemented in GENEPOP v.4.5.0 (Rousset, 2008). GENEDOC v.2.7 (Nicholas, 1997) was used for calculating identity and similarity percentages (as assessed by the BLOSUM62 substitution matrix) for all genotypes observed within populations for the whole β1 domain and PBR positions [considering the previously reported 31 putative positions constituting the MHC-DRB PBR (Suárez et al., 2006)]. PBR sequence logos, along with PBR sequence correlation between populations, were analyzed to further evaluate differences in potential MHC-presented peptide repertoire among cattle populations. MEGA X (Kumar et al., 2018) was used for calculating the average amount of synonymous ($d_S$) and nonsynonymous ($d_N$) substitutions per site by Nei-Gojobori's method with Jukes-Cantor correction. The Z-test was used for assessing $d_N/d_S$ ratio significance. Codons subject to positive selection were inferred using maximum likelihood (FEL (Kosakovsky Pond and Frost, 2005)) and Bayesian (MEME (Murrell et al., 2012)) methods using the Datamonkey web server (Delport et al., 2010), using a p-value <0.1 as significant threshold.

## Population Structure, Genetic Differentiation and PBR Similarity Correlation

Population structure and genetic differentiations between populations were evaluated by calculating pairwise $F_{ST}$ statistics (Weir and Cockerham, 1984) using ARLEQUIN. POPTREE v.2 (Takezaki et al., 2010) was used for calculating genetic distances ($D_A$) (Nei, 1978) from allele frequency data and PAST v.3.2 (Hammer et al., 2001) for generating an allele frequency-based principal component analysis (PCA). WebLogo web server (Crooks et al., 2004) was used to construct a logo for PBR positions for each population using BLOSUM62 substitution matrix for each position including all alleles observed within populations. A 620-coordinate vector (31 positions x 20 possible aa genotypes) representing each PBR population was used for calculating Pearson correlation coefficients using R package *amap* (Ihaka and Gentleman, 1996), while R package *ape* was used to construct an UPGMA dendrogram (Paradis et al., 2004). Mitochondrial DNA D-loop sequences from individuals of several of the populations analyzed here were recovered from GenBank (**Supplementary Data S2**) and used for evaluating MHC allele distribution based on genetic

**TABLE 2 |** Allele frequency and standard error for *BoLA-DRB3* alleles common to Simmental and Simbrah cattle.

| Allele | Simbrah | Simmental |
|---|---|---|
| 001:01 | 0.008 ± 0.008 | 0.037 ± 0.016 |
| 002:01 | **0.058 ± 0.021** | **0.082 ± 0.023** |
| 003:01 | 0.017 ± 0.011 | 0.022 ± 0.012 |
| 005:01 | 0.025 ± 0.014 | **0.06 ± 0.02** |
| 005:03 | **0.058 ± 0.021** | **0.06 ± 0.02** |
| 009:01 | 0.017 ± 0.011 | 0.022 ± 0.012 |
| 010:01 | 0.033 ± 0.016 | **0.06 ± 0.02** |
| 011:01 | 0.025 ± 0.014 | 0.022 ± 0.012 |
| 012:01 | **0.117 ± 0.029** | **0.052 ± 0.019** |
| 013:01 | 0.025 ± 0.014 | **0.052 ± 0.019** |
| 014:01:01 | 0.017 ± 0.011 | 0.022 ± 0.012 |
| 016:01 | 0.017 ± 0.011 | **0.097 ± 0.025** |
| 017:03 | 0.033 ± 0.016 | 0.007 ± 0.007 |
| 018:01 | 0.017 ± 0.011 | 0.015 ± 0.01 |
| 019:02 | 0.017 ± 0.011 | 0.007 ± 0.007 |
| 020:01:01 | 0.025 ± 0.014 | 0.007 ± 0.007 |
| 021:01 | 0.025 ± 0.014 | 0.007 ± 0.007 |
| 026:01 | 0.008 ± 0.008 | 0.037 ± 0.016 |
| 027:03 | 0.008 ± 0.008 | 0.007 ± 0.007 |
| 030:01 | 0.017 ± 0.011 | **0.052 ± 0.019** |

*Alleles having >5% frequency are highlighted in bold.*

affinity between populations. A maximum likelihood phylogeny was built using Tamura-3-parameter model (best fit model) for mitochondrial DNA and JTT model for *BoLA-DRB3* aa sequences using MEGA X.

# RESULTS

## *BoLA-DRB3* Allele Distribution in Colombian Simmental and Simbrah Cattle

Sixty *BoLA-DRB3* alleles were identified in both focal populations: 37 in Simmental and 43 in Simbrah cattle (GenBank accession numbers OM100952-OM101010). No new alleles were detected. Three alleles (*BoLA-DRB3*002:01*, *005:03* and *012:01*) occurred with >5% frequency in both populations and seven (*BoLA-DRB3*005:01*, *008:01*, *010:01*, *013:01*, *015:01*, *016:01* and *022:01*) only in the Simmental breed. Four alleles (*BoLA-DRB3*002:01*, *005:03*, *012:01* and *022:01*) occurred with >5% frequency in Simbrah, accounting for 31.67% of cumulated frequency, whereas 10 alleles (*BoLA-DRB3*002:01*, *005:01*, *005:03*, *008:01*, *010:01*, *012:01*, *013:01*, *015:01*, *016:01* and *030:01*) with >5% frequency, accounting for 61.94% cumulative frequency, occurred in Simmental. Twenty alleles were shared by both cattle populations, accounting for 64.7% of their mean cumulative allele frequency (summation of weighted mean allele frequency for all shared alleles) (**Table2**).

## *BoLA-DRB3* Genetic Diversity, Hardy-Weinberg Equilibrium and Selection Pattern

The number of alleles corrected for sample size effect ($Rs$) showed that Simmental and Simbrah had a high genetic diversity, similar

to what was observed for Colombian Normande and Philippine populations (**Table 3**). Sample size appeared to have little effect on both $h_o$ ($r = 0.36$) and $h_e$ ($r = 0.06$), highlighting these estimates as being good proxy of *BoLA-DRB3* diversity in the populations analyzed here. Colombian Simmental, Colombian Normande, Spanish Morucha and Colombian Simbrah had the lowest $h_o$ and the highest $h_e$ values, in contrast with the results for other breeds (**Table 3**; **Supplementary Data S1**). Consequently, the highest departures from Hardy-Weinberg equilibrium were observed in these four breeds, as evidenced by the statistically significant $F_{IS}$ fixation index (**Table 3**). These results indicated significant heterozygote deficiency regarding the *BoLA-DRB3* locus in Colombian Simmental and Colombian Simbrah cattle.

Whole β1 domain identity ranged from 82.45% (Philippine Brahman) to 88.95% (Colombian Normande), with Colombian Simmental (88.79%) and Simbrah (86.33%) displayed some of the highest values (**Table 4**). Bolivian Holstein and Colombian Normande had the lowest (87.45%) and the highest (92.5%) whole β1 domain similarity, respectively. PBR identity and similarity differences were broader than those for the whole β1 domain. The former ranged from 65.87% (Peruvian Holstein) to 80.81% (Spanish Morucha), whereas the latter ranged from 75.18% (Bolivian Holstein) to 87.59% (Spanish Morucha). Colombian Simmental and Simbrah displayed some of the highest identity and similarity values regarding both whole β1 domain and PBR.

Average $d_N$ was significantly higher than average $d_S$ for Colombian Simmental and Simbrah populations, similar to what was observed across all the other populations (**Table 3**). Moreover, codons 10, 11, 12, 26, 30, 32, 37, 57, 70, 71, 74 and 78 were identified as sites under diversifying selection, most of which (11, 12, 26, 30, 37, 57, 70, 71, 74 and 78) were PBR-related. Mitochondrial DNA clusters were mostly formed by individuals of one to four breeds. Infrequently occurring allele lineages (i.e., *BoLA-DRB3*025:01:02*, *025:02*, *025:01:01*, *037:01* and *039:01* in most populations) or moderately occurring ones (i.e., *BoLA-DRB3*030:01*, *030:02*, *036:01*, *023:01* and *044:01*) were identified in populations that are not closely related (**Supplementary Data S3** and **Supplementary Data S4**).

## Population Structure and Genetic Differentiation Based on *BoLA-DRB3* Gene

Pairwise $F_{ST}$ values ranged from −0.0009 (Chilean Holstein with Peruvian Holstein) to 0.1185 (Peruvian Holstein with Bolivian Nellore) (**Figure 1A**: **Supplementary Data S5**). All comparisons were statistically significant, except for Chilean with Peruvian Holstein and Argentinian with Paraguayan Holstein. Two groups with low $F_{ST}$ values were mainly observed (**Figure 1A**); the first consisted of Holstein populations (0.0088 mean $F_{ST}$ value), while the second was formed by Philippine, Colombian Simmental, Colombian Simbrah, Bolivian Yacumeño and Colombian Normande populations (0.0208 mean $F_{ST}$ value). Differences between Colombian Simmental and Simbrah were also observed regarding Argentinian, Peruvian, Chilean and Paraguayan Holstein.

**TABLE 3 |** Genetic diversity estimates and Hardy-Weinberg equilibrium within cattle populations.

| Population | N | $N_a$ | $R_s$ | $h_o$ | $h_e$ | $F_{IS}$ (S.E.) | dN | dS | dN/dS |
|---|---|---|---|---|---|---|---|---|---|
| SmtCo | 67 | 37 | 34 | 0.642 | 0.959 | **0.333 (0.0005)**\*\* | 0.107 | 0.03 | *3.57* |
| SbhCo | 60 | 43 | 41.6 | 0.783 | 0.966 | **0.191 (0.0004)**\*\* | 0.116 | 0.026 | *4.46* |
| NorCo | 111 | 53 | 38.2 | 0.667 | 0.955 | **0.303 (<0.0001)**\*\* | 0.111 | 0.028 | *3.96* |
| MorSp | 54 | 29 | 29 | 0.667 | 0.913 | **0.271 (<0.0001)**\*\* | 0.105 | 0.026 | *4.04* |
| HolAr | 413 | 31 | 20.5 | 0.833 | 0.908 | **0.082 (0.0001)**\*\* | 0.115 | 0.025 | *4.6* |
| HolBo | 153 | 21 | 18 | 0.928 | 0.895 | −0.038 (0.0054) | 0.124 | 0.031 | *4* |
| HolPa | 127 | 26 | 19.6 | 0.835 | 0.891 | **0.064 (0.0140)** \* | 0.117 | 0.024 | *4.88* |
| HolPe | 132 | 19 | 16.6 | 0.902 | 0.886 | −0.018 (0.0151) | 0.12 | 0.032 | *3.75* |
| HolJa | 101 | 18 | 15.7 | 0.922 | 0.902 | −0.022 (0.0324) | 0.115 | 0.027 | *4.26* |
| HolCh | 113 | 21 | 17.2 | 0.841 | 0.893 | **0.059 (0.0067)**\* | 0.113 | 0.028 | *4.04* |
| YacBo | 100 | 33 | 30.1 | 0.91 | 0.947 | 0.04 (0.0399) | 0.116 | 0.023 | *5.04* |
| NeBrPe | 195 | 33 | 23.5 | 0.759 | 0.855 | **0.113 (0.0015)**\* | 0.116 | 0.027 | *4.3* |
| BrPh | 233 | 57 | 36.3 | 0.884 | 0.95 | **0.07 (0.0019)**\* | 0.112 | 0.026 | *4.31* |
| NeBo | 116 | 26 | 21.9 | 0.784 | 0.87 | 0.099 (0.0412) | 0.117 | 0.023 | *5.09* |
| NaBrPh | 131 | 56 | 40.8 | 0.908 | 0.966 | 0.06 (0.0201) | 0.113 | 0.023 | *4.91* |
| NaPh | 480 | 71 | 36.8 | 0.915 | 0.959 | **0.044 (0.0040)**\* | 0.112 | 0.024 | *4.67* |

*Number of individuals (N), Number of alleles ($N_a$), allelic richness ($R_s$), observed ($h_o$) and expected ($h_e$) heterozygosity, non-synonymous/synonymous substitution ratio (dN/dS) and Hardy-Weinberg equilibrium as evaluated by $F_{IS}$, coefficients along with standard errors (SE). Statistical significance is indicated in bold and \* or \*\* for p < 0.05 and p < 0.001, respectively.*

**TABLE 4 |** Genetic diversity at sequence level within cattle populations.

| Population | β1 domain | | PBR | |
|---|---|---|---|---|
| | Identity | Similarity | Identity | Similarity |
| SmtCo | 88.79 (9.6) | 92.33 (6.9) | 78.67 (17.7) | 84.61 (13.1) |
| SbhCo | 86.33 (8.6) | 91.2 (5.9) | 73.68 (16.9) | 82.48 (11.7) |
| NorCo | 88.95 (8.8) | 92.5 (6.2) | 77.75 (17.4) | 84.59 (12.5) |
| MorSp | 88.35 (9.5) | 92.22 (6.9) | 80.81 (15.7) | 87.59 (10.7) |
| HolAr | 85.34 (7.6) | 89.66 (5.7) | 69.67 (16) | 78.03 (12.3) |
| HolBo | 82.71 (6.6) | 87.45 (5.3) | 66.31 (12.2) | 75.18 (10.1) |
| HolPa | 84.87 (7.8) | 89.29 (5.7) | 68.61 (16.4) | 77.04 (12.5) |
| HolPe | 83.89 (7.1) | 88.77 (5.5) | 65.87 (15) | 75.72 (11.9) |
| HolJa | 83.44 (6.1) | 88.56 (4.5) | 66.15 (12.6) | 76.38 (9.2) |
| HolCh | 84.44 (7.7) | 88.83 (5.9) | 67.4 (16.4) | 75.91 (12.9) |
| YacBo | 83.15 (7.2) | 87.9 (5.4) | 66.87 (13.7) | 75.7 (11.1) |
| NeBrPe | 85.66 (10.1) | 90.36 (7.5) | 73.43 (17.6) | 82.62 (12.3) |
| BrPh | 82.44 (8.1) | 87.71 (6.1) | 69.81 (14) | 79.15 (10.1) |
| NeBo | 85.41 (8.9) | 90.85 (6.1) | 71.96 (16.6) | 83.58 (11) |
| NaBrPh | 82.53 (7.5) | 88.14 (5.8) | 68.68 (12.9) | 78.96 (9.3) |
| NaPh | 83.13 (7.1) | 88.71 (5.4) | 68.96 (12.9) | 79.23 (9.5) |

*Similarity was calculated according to BLOSUM62 substitution matrix. PBR, peptide binding region.*

Colombian Simbrah was differentiated from native Philippine, Colombian Normande, Bolivian Yacumeño and Holstein populations to a greater extent than Colombian Simmental. On the contrary, Colombian Simmental was more differentiated from Philippine Native-Brahman, Philippine Brahman, Spanish Morucha, Peruvian Nellore-Brahman and Bolivian Nellore than Colombian Simbrah (**Figure 1A**: **Supplementary Data S5**).

Genetic distance $D_A$ clustering was similar to that based on pairwise $F_{ST}$ among Holstein and the Philippine populations (**Figure 1B**). The latter group was well-differentiated from Colombian Simmental, Colombian Simbrah, Bolivian Yacumeño and Colombian Normande populations on the basis of $D_A$ but not on pairwise $F_{ST}$. Mean genetic distance

values for these breeds indicated that common alleles could explain a large amount of their cumulative allele frequency. Thus, the Holstein population group (0.078 mean $D_A$ distance) had 28 alleles in common, accounting for 95.2% of their mean cumulative allele frequency. Ten of these common alleles (*BoLA-DRB3\*001:01, 002:01, 006:01, 009:02, 010:01, 011:01, 012:01, 14: 01:01, 015:01* and *027:03*) occurred with >5% frequency in at least one of these populations. The Philippine group (0.169 mean $D_A$ distance) had 65 common alleles accounting for 93.1% of their mean cumulative allele frequency, of which 8 (*BoLA-DRB3\*002: 01, 003:01, 012:01, 015:01, 022:01, 030:01, 036:01* and *041:01*) were alleles with >5% frequency. The group formed by Colombian and Bolivian Yacumeño populations (0.387 mean $D_A$ distance) had 49 alleles in common, accounting for 69.1% of their mean cumulative allele frequency. Seventeen (*BoLA-DRB3\*001:01, 002:01, 005:01, 005:03, 007:01, 008:01, 009:02, 010:01, 012:01, 013:01, 014:01:01, 015:01, 016:01, 018:01, 022:01, 030:01* and *048:02*) of these shared alleles had >5% frequency.

Groups identified in the first two principal component (PC) plots were consistent with those identified in $F_{ST}$ and $D_A$ distance analysis (**Figure 2**: **Supplementary Data S6**). The first PC (42.7% variance) differentiated four main groups. The Holstein population group was characterized by high *BoLA-DRB3\*015: 01, 011:01, 001:01, 027:03, 010:01, 012:01, 014:01:01* and *009:02* allele frequency, while, group formed by Normande, Simmental and Yacumeño by intermediate frequency regarding the same alleles; both groups had intermediate *BoLA-DRB3\*006:01, 017:01, 016:01, 009:01* and *002:01* allele frequencies. Bolivian Nellore and Peruvian Nellore-Brahman along with a group formed by Philippine Native, Philippine Native-Brahman, Colombian Simbrah, Philippine Brahman and Spanish Morucha were at the other extreme of the first PC. These two groups were characterized by high or intermediate *BoLA-DRB3\*048:02, 030: 01, 028:01* and *022:01* allele frequency. Spanish Morucha was remarkably differentiated in the second PC (20.3% variance) due to high *BoLA-DRB3\*048:02, 003:01* and *005:01* frequency,
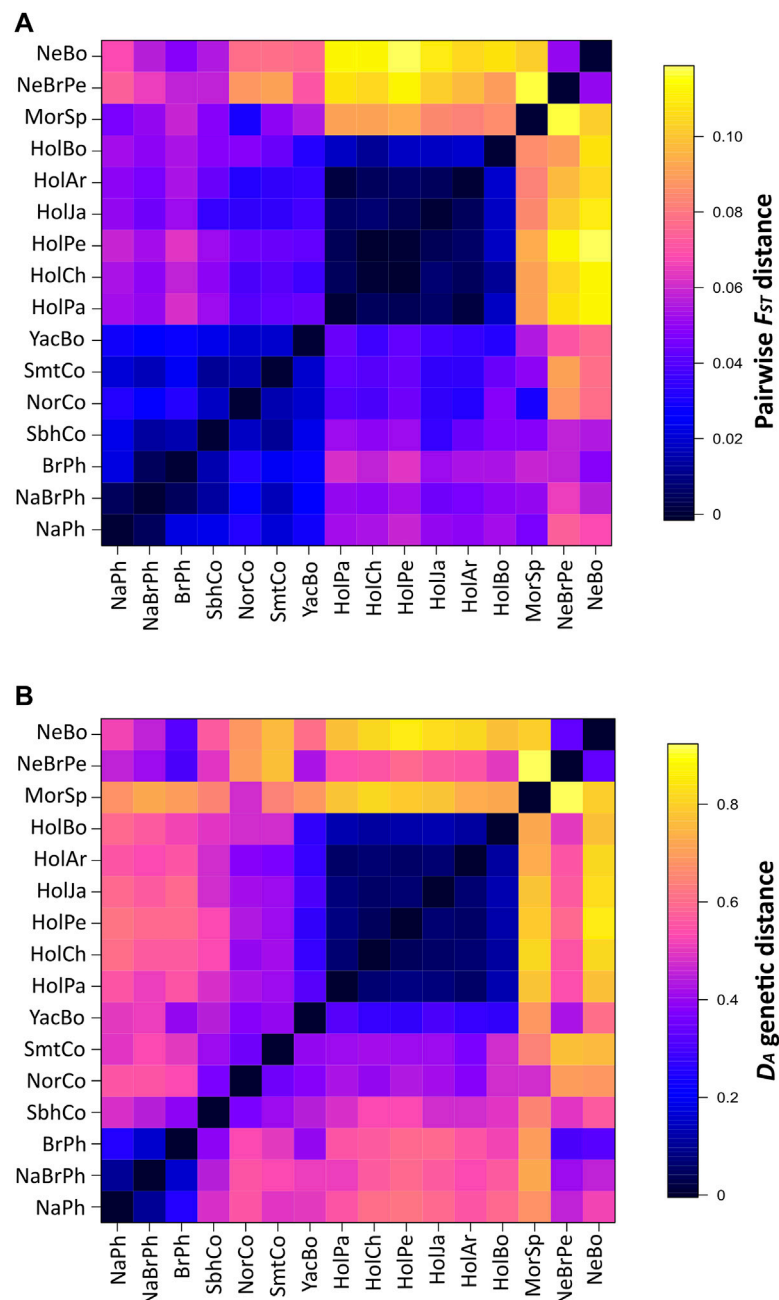
**FIGURE 1 |** Pairwise $F_{ST}$ and $D_A$ genetic distances between populations. Symmetrical heat maps for pairwise $F_{ST}$ values **(A)** and $D_A$ genetic distance **(B)** between cattle populations based on *BoLA-DRB3* alleles. All pairwise $F_{ST}$ comparisons (except for Chilean with Peruvian Holstein and Argentinian with Paraguayan Holstein) were significant at 0.05 level.

whereas Nellore cattle had high *BoLA-DRB3*028:01*, *009:02* and *022:01* frequencies.

## PBR Sequence Similarity and Correlation Between Populations

PBR logo representation showed that positions 70, 71, 74, 11, 13, 30, 67, 37, and 57 were highly variable and tended to

accumulate non-conservative changes, while positions 82, 83, 14, 15, 40, 72, 73, 79, 29, 64, 47, 9, and 38 were invariable or only displayed conservative changes (**Figure 3A**). Some highly variable sites were found under positive selection (11, 30, 37, 57, 70, 71, and 74). PBR logos had very similar substitution patterns for all populations. Likewise, Pearson correlation coefficients (PCC) were remarkably high, having high global correlation (PCC = 0.987), thereby indicating low
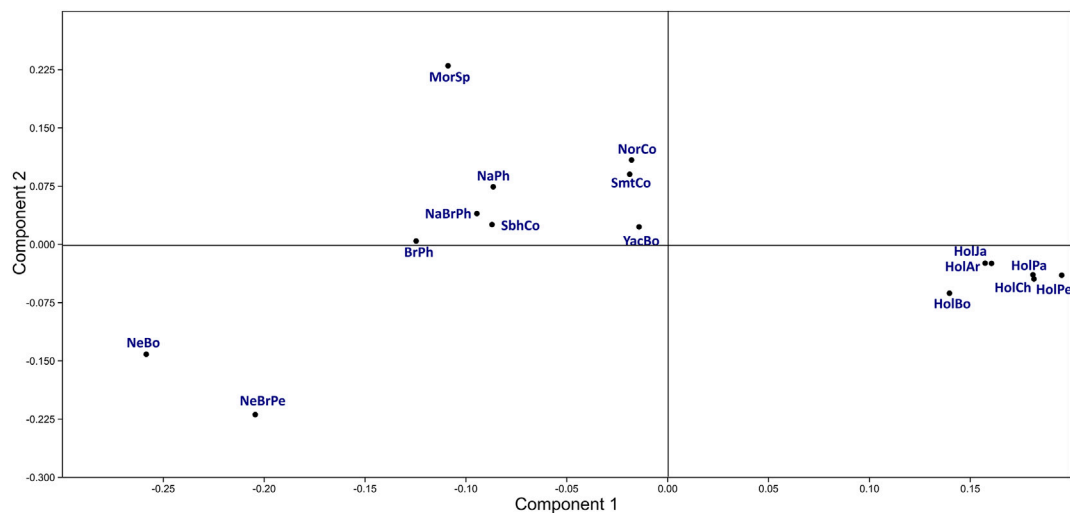
**FIGURE 2 |** Principal component analysis using *BoLA-DRB3* allele frequencies for cattle populations.

variation for aa frequency for each PBR position among all populations.

PBR correlation analysis identified four major groups (**Figures 3B,C**). The first one consisted of Colombian Normande, Colombian Simbrah, Spanish Morucha and Colombian Simmental (0.994 mean PCC). It is worth mentioning that the most notable differences between Colombian Simmental and Simbrah occurred regarding aa frequency for positions 70, 30 and 57. Paraguayan Holstein and Peruvian Holstein clustered together (PCC = 0.996). The third group was formed by native Philippine, Philippine Brahman, Bolivian Holstein, Japanese Holstein and Argentinian Holstein (0.996 mean PCC). The fourth group included native Philippine-Brahman, Peruvian Nellore-Brahman and Bolivian Nellore (0.995 mean PCC). Divergence between population groups was mainly due to PBR Pockets 6 (PCC = 0.981), 7 (PCC = 0.979) and 9 (PCC = 0.978), while Pockets 1 (PCC = 0.992) and 4 (PCC = 0.992) were less variable.

## DISCUSSION

The MHC influences susceptibility and resistance to infectious diseases, vaccine responses and productions traits. MHC allele distribution information can be used to guide resource-consumption studies (such as immunopeptidomic or binding affinity assays) aimed at identifying MHC-associated peptides and developing *in silico* binding predictive algorithms that can be used for understanding and predicting immune response patterns (Nielsen et al., 2018; Rappazzo et al., 2020). MHC high polymorphism can also provide insights into populations evolutionary history. Nevertheless, MHC diversity has only been explored in a few cattle breeds to date (Takeshima et al., 2003; Giovambattista et al., 2013; Takeshima et al., 2014; Takeshima et al., 2015; Takeshima et al., 2018). In this study, we have first characterized *BoLA-DRB3* genetic diversity in the

taurine Simmental breed and in its most common cross with zebuine cattle in tropical regions, the Simbrah in Colombia. Considering the recent origin of Colombian Simmental and Simbrah cattle, a large percentage of highly related animals is expected since small herds are derived from few parents, thus reflecting a potentially reduced genetic diversity, a condition previously found for other pure cattle breeds in the country (Amaya et al., 2020; Bohórquez et al., 2020).

Cattle populations varied considerably in terms of allele richness and Hardy-Weinberg equilibrium. It is worth noting that unequal sample size have no significant impact on genetic diversity estimates such as $h_o$ (unbiased parametric value estimator, mainly determined by the sampling method used (Nei and Chesser, 1983; Bohórquez et al., 2020)), $h_e$ or $R_s$, also when the pertinent corrections were applied (Nei, 1978; Leberg, 2002). Simmental and Simbrah $R_s$ and $h_e$ were among the highest values, similar to those of Normande and Philippine populations. However, differences in $R_s$ (or $N_a$) associated with similar $h_e$ values across populations (as observed for Simmental and Simbrah) indicated a marked allele frequency distribution variation, while $h_o$ values were the lowest for Colombian cattle, with $F_{IS}$ indices being thereby the highest for these populations. Population sub-structuring may have reduced $h_o$ [by means of the Wahlund effect, i.e., reduced observed heterozygosity in a population caused by subpopulation structure (De Meeûs, 2017)], thus magnifying allele frequency differences compared to those found in other populations.

Natural selection and random genetic drift (notoriously exerting higher effects in populations with smaller effective population size) are factors affecting allele frequency distribution (Nei and Tajima, 1981; Tajima and Nei, 1984; Akashi et al., 2012; Husemann et al., 2016), which in the case of the MHC is expected to reflect balancing selection with heterozygote excess (Hughes and Nei, 1988; Hughes and Yeager, 1998; Takeshima et al., 2008). Even though $F_{IS}$ indicated heterozygote deficiency for *BoLA-DRB3*, the
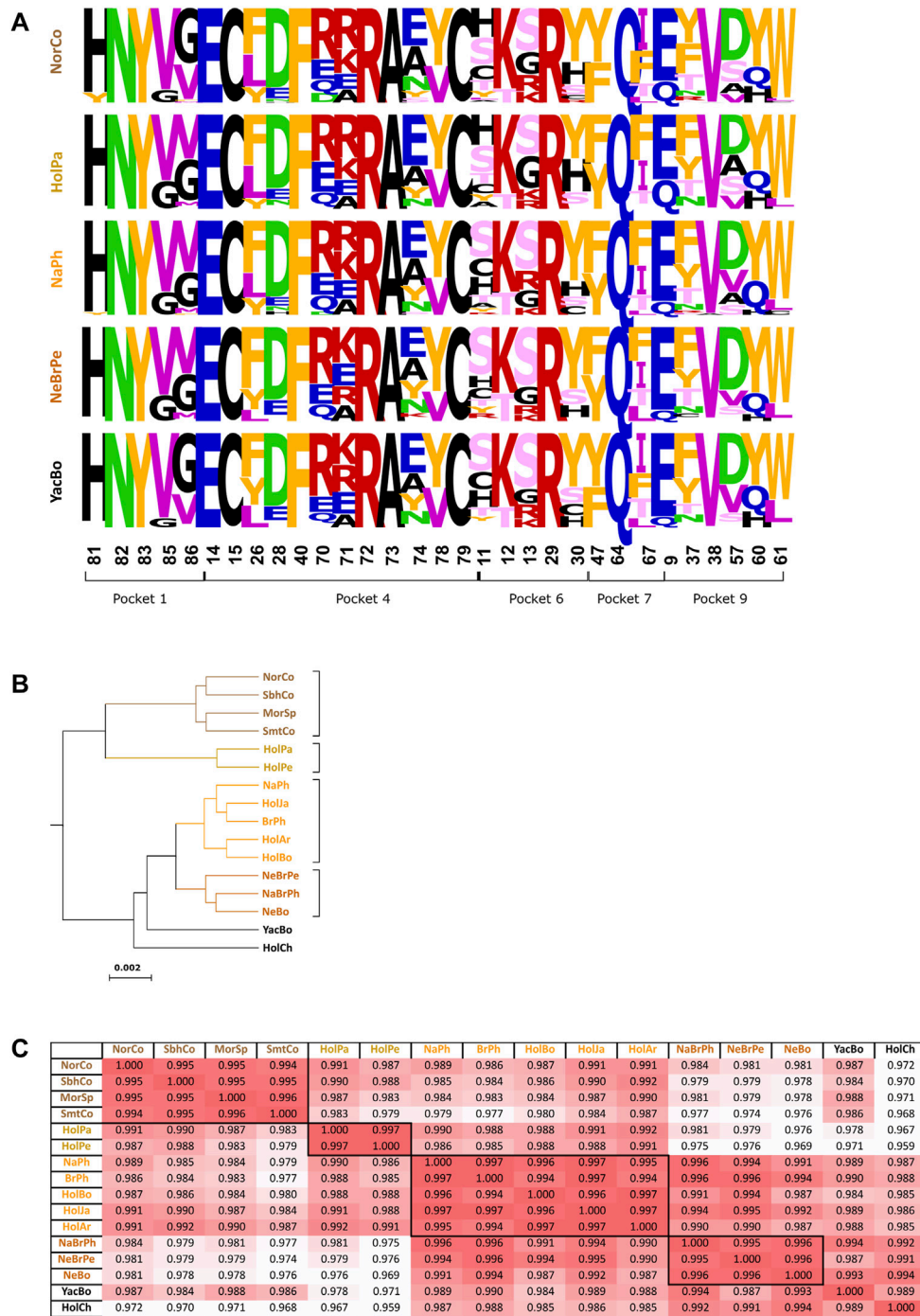
**FIGURE 3 |** Peptide binding region similarity and correlation. **(A)** *BoLA-DRB3* allele peptide binding region (PBR) logos representative of each cattle population group. Conservative aa changes according to the BLOSUM 62 substitution matrix are shown in the same color; different colors indicate non-conservative changes: FYW (yellow), MILV (purple), RK (red), EQ (blue), DN (green), ST (pink), HCPAG (black). PBR Pockets and their aa forming positions are indicated below the logos. **(B)** Clustering analysis based on correlation analysis. **(C)** Pearson correlation coefficient matrix for populations based on PBR sequence variability. Population group colors are the same for panels **(A,B,C)**.

persistence of identical or similar alleles (allelic lineages) in spite of the overall genetic differentiation points to the action of balancing selection. Moreover, the highly significant $d_N/d_S$ values and codons identified under positive selection suggest that aa variability preference has not been eroded in these cattle. Although drift and non-random mating are well-known

factors leading to increased homozygosity (Crow, 2010) as well as small effective population sizes and inbreeding, the use of only one locus hampered from singling out the underlying cause of such pattern. Likewise, the high $F_{IS}$ values for some of these populations, indicating higher homozygote percentages than those expected for Hardy-Weinberg equilibrium, points to the occurrence of evolutionary forces acting on these populations. Therefore, further studies based on genome-wide data including both neutral and non-neutral loci, are necessary to get a comprehensive picture of the evolutionary forces acting on this system, and of their relative contributions [in Simmental, selection based on production and its small effective population size should be taken into account (Amaya et al., 2020; de Araujo Neto et al., 2020)].

Alleles fell into different categories based on their distribution throughout the populations tested. The first category consisted of alleles widely distributed in populations from different continents and often displaying relatively high frequencies. These alleles, such as *BoLA-DRB3\*011:01*, are possibly present in all populations, or absent in a few of them, such as *BoLA-DRB3\*010:01*, *012:01* and *014:01:01*. Considering that taurine and zebuine cattle were domesticated in more than two independent events (Beja-Pereira et al., 2006; Decker et al., 2014), these alleles probably predate *Bos primigenius* divergence which gave rise to these cattle types. Such alleles might have been either present in just a subset of the founder populations or ubiquitous before undergoing secondary loss due to random genetic drift and/or natural selection (alternatively, their very low frequencies impaired their sampling) (Barton, 1996; Sutton et al., 2011). Another category consisted of alleles, such as *BoLA-DRB3\*017:01*, *006:01*, *009:01* and *027:10*, found predominantly in taurine or zebuine cattle with their presence in the other type of cattle populations being possibly indicative of admixture. The last category includes the alleles found exclusively in some populations and displaying low frequencies, possibly representing the most recently arisen ones (Uinuk-Ool et al., 2002). Despite forming just a moderate proportion of known *BoLA-DRB3* alleles found in and Simbrah cattle (67 and 60 out of 330, respectively) (Maccari et al., 2017), they occur with significant frequency in other cattle populations, representing the major allele variants (Bohórquez et al., 2020) (**Supplementary Data S3**, **Supplementary Data S4** and **Supplementary Data S6**). Nevertheless, some alleles contributing towards the distinction of Colombian Simmental from Simbrah were also significant in differentiating zebuine from taurine cattle, such as *BoLA-DRB3\*015:01* and *022:01* frequently occurring in taurine and zebuine cattle, respectively (Takeshima et al., 2018). Alleles private to Simmental or Simbrah further contributed to their differentiation.

Although most of the target bovine populations analyzed here were genetically well-differentiated based on *BoLA-DRB3*, others had a shallow structure due to the sharing of several alleles occurring with high or intermediate frequency. These populations clustered into five groups according to the measures of differentiation used. Low mean distance values indicated high genetic affinity for these populations and more detailed analysis showed that commonly occurring alleles

accounted for a large percentage of their mean cumulative allele frequency. Several factors may result in weak structure. For instance, the limited genetic distance between Colombian Simmental and Normande might be due to sample origin, as geographical dispersal patterns in cattle reflects those of exportation and co-migration in humans (Decker et al., 2014), as well as similar selection pressure (Bohórquez et al., 2020). Furthermore, as zebuine introgression occurred independently in American and Indian cattle (Decker et al., 2014), crossbreeding with Brahman may have led to low genetic differentiation between Simbrah and Philippine populations. These results contrasted with the weak differentiation in Holstein cattle, such breed forming a very compact group in spite of multiple sample origins (Takeshima et al., 2015), possibly as the result of intense selective pressure regarding milk production traits (Bohórquez et al., 2020) and a high level of gene flow *via* genetic improvement strategies, thereby leading to a high degree of homogenization (Spieth, 1974; Leroy et al., 2013).

MHC PBR positions hosted the highest β1 domain variability associated with the peptides to which an allele could bind (Hughes and Nei, 1989; Stern et al., 1994). The similarity/identity matrix of the MHC-DRB PBR position across cattle populations suggested potential differences in MHC-presented peptide repertoire size (**Table 4**). Colombian populations had the highest identity and similarity values in both β1 domain and PBR. This could be due to high overall homozygosity but also to a limited variability at these specific loci and suggested that these animals had smaller individual MHC-presented peptide repertoires. Nevertheless, the good correlation regarding PBR aa sequence and logo analysis may suggest that potential population-related MHC-presented peptide repertoire diversity could be equivalent among all cattle populations analyzed here, with just a few groups displaying a much higher variability. This implies that breeding is unlikely to have decreased functional MHC variability, which bears important implications for peptide-based vaccine design, so that different cattle populations could be targeted using similar peptide combinations. Although it has been shown that decreased MHC variability might be caused by population bottlenecks (Bollmer et al., 2011; Zhang et al., 2016), balancing selection driven by pathogens can still maintain a high degree of diversity (Aguilar et al., 2004; Newhouse and Balakrishnan, 2015). Moutou *et al.*, have shown that functional polymorphism may be lower than genetic polymorphism in pigs (Moutou et al., 2013). The highest similarities within and between porcine populations were mainly due to high correlation in PBR Pocket 1 and 4, whereas a higher divergence was observed for Pocket 6, 7 and 9. This could have arisen from Pocket 1, 4 and 6 aa preferences as anchor positions (Rappazzo et al., 2020). This work has shown that, in spite of high values for some genetic diversity measures regarding *BoLA-DRB3* (such as allele richness and expected heterozygosity), Colombian Simmental, Simbrah and other cattle populations may have a limited potential MHC-presented peptide repertoire diversity could be similar among all cattle populations analyzed and that breeding did not decrease functional diversity. Additional analyses directly addressing peptide repertoire diversity are needed to confirm these results.

# DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

# ETHICS STATEMENT

The animal study was reviewed and approved by the Universidad de Ciencias Aplicadas y Ambientales' (U.D.C.A) Animal Research Ethics Committee (minute No. 201901). Written informed consent was obtained from the owners for the participation of their animals in this study.

# AUTHOR CONTRIBUTIONS

DO, MB, CA, and MP: conceptualisation. DO and MB: data curation. MB, DO and CA: formal analysis. MP: funding acquisition. MB, DO and CA: investigation. DO: resources. MP: supervision. MB: visualisation. MB and DO: writing – original draft preparation. MB, DO, CA, and MP: writing—review and editing.

# SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.772885/full#supplementary-material

# REFERENCES

Acosta-Rodríguez, R., Alonso-Morales, R., Balladares, S., Flores-Aguilar, H., García-Vazquez, Z., and Gorodezky, C. (2005). Analysis of BoLA Class II Microsatellites in Cattle Infested with Boophilus Microplus Ticks: Class II Is Probably Associated with Susceptibility. *Vet. Parasitol.* 127 (3-4), 313–321. doi:10.1016/j.vetpar.2004.10.007

Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D., and Wayne, R. K. (2004). High MHC Diversity Maintained by Balancing Selection in an Otherwise Genetically Monomorphic Mammal. *Proc. Natl. Acad. Sci.* 101 (10), 3490–3494. doi:10.1073/pnas.0306582101

Agung, P. P., Saputra, F., Septian, W. A., Zein, M. S., Sulandari, S., Anwar, S., et al. (2016). Study of Genetic Diversity Among Simmental Cross Cattle in West Sumatra Based on Microsatellite Markers. *Asian-Australas J. Anim. Sci.* 29 (2), 176–183. doi:10.5713/ajas.15.0155

Akashi, H., Osada, N., and Ohta, T. (2012). Weak Selection and Protein Evolution. *Genetics.* 192 (1), 15–31. doi:10.1534/genetics.112.140178

Amaya, A., Martínez, R., and Cerón-Muñoz, M. (2020). Population Structure and Genetic Diversity in Colombian Simmental Cattle. *Trop. Anim. Health Prod.* 52 (3), 1133–1139. doi:10.1007/s11250-019-02111-w

Babik, W., Kawałko, A., Wójcik, J. M., and Radwan, J. (2012). Low Major Histocompatibility Complex Class I (MHC I) Variation in the European Bison (Bison Bonasus). *J. Hered.* 103 (3), 349–359. doi:10.1093/jhered/ess005

Bacon, L. D., and Witter, R. L. (1995). Efficacy of Marek's Disease Vaccines in Mhc Heterozygous Chickens: Mhc Congenic × Inbred Line F1 Matings. *J. Hered.* 86 (4), 269–273. doi:10.1093/oxfordjournals.jhered.a111580

Barton, N. H. (1996). Natural Selection and Random Genetic Drift as Causes of Evolution on Islands. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351 (1341), 785–795. doi:10.1098/rstb.1996.0073

Baxter, R., Craigmile, S. C., Haley, C., Douglas, A. J., Williams, J. L., and Glass, E. J. (2009). BoLA-DR Peptide Binding Pockets Are Fundamental for Foot-And-Mouth Disease Virus Vaccine Design in Cattle. *Vaccine.* 28 (1), 28–37. doi:10.1016/j.vaccine.2009.09.131

Baxter, R., Hastings, N., Law, A., and Glass, E. J. (2008). A Rapid and Robust Sequence-Based Genotyping Method for BoLA-DRB3 alleles in Large Numbers of Heterozygous Cattle. *Anim. Genet.* 39 (5), 561–563. doi:10.1111/j.1365-2052.2008.01757.x

Beja-Pereira, A., Caramelli, D., Lalueza-Fox, C., Vernesi, C., Ferrand, N., Casoli, A., et al. (2006). The Origin of European Cattle: Evidence from Modern and Ancient DNA. *Proc. Natl. Acad. Sci.* 103 (21), 8113–8118. doi:10.1073/pnas.0509210103

Bohórquez, M. D., Ordoñez, D., Suárez, C. F., Vicente, B., Vieira, C., López-Abán, J., et al. (2020). Major Histocompatibility Complex Class II (DRB3) Genetic Diversity in Spanish Morucha and Colombian Normande Cattle Compared to Taurine and Zebu Populations. *Front. Genet.* 10 (1293), 1293. doi:10.3389/fgene.2019.01293

Bollmer, J. L., Hull, J. M., Ernest, H. B., Sarasola, J. H., and Parker, P. G. (2011). Reduced MHC and Neutral Variation in the Galápagos Hawk, an Island Endemic. *BMC Evol. Biol.* 11 (143), 143–2148. doi:10.1186/1471-2148-11-143

Borg, Å. A., Pedersen, S. A., Jensen, H., and Westerdahl, H. (2011). Variation in MHC Genotypes in Two Populations of House Sparrow (*Passer domesticus*) with Different Population Histories. *Ecol. Evol.* 1 (2), 145–159. doi:10.1002/ece3.13

Carignano, H. A., Beribe, M. J., Caffaro, M. E., Amadio, A., Nani, J. P., Gutierrez, G., et al. (2017). BOLA-DRB3 gene Polymorphisms Influence Bovine Leukaemia Virus Infection Levels in Holstein and Holstein × Jersey Crossbreed Dairy Cattle. *Anim. Genet.* 48 (4), 420–430. doi:10.1111/age.12566

Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: A Sequence Logo Generator: Figure 1. *Genome Res.* 14 (6), 1188–1190. doi:10.1101/gr.849004

Crow, J. F. (2010). Wright and Fisher on Inbreeding and Random Drift. *Genetics.* 184 (3), 609–611. doi:10.1534/genetics.109.110023

de Araujo Neto, F. R., Vieira, D. A., Santos, D. J. d. A., Pessoa, M. C., Borquis, R. R. A., de Oliveira, H. N., et al. (2020). Population Structure of Simmental Beef Cattle Using Pedigree Analysis. *Trop. Anim. Health Prod.* 52 (3), 1513–1517. doi:10.1007/s11250-019-02102-x

De Meeûs, T. (2017). Revisiting FIS, FST, Wahlund Effects, and Null Alleles. *J. Hered.* 109 (4), 446–456. doi:10.1093/jhered/esx106

Decker, J. E., McKay, S. D., Rolf, M. M., Kim, J., Molina Alcalá, A., Sonstegard, T. S., et al. (2014). Worldwide Patterns of Ancestry, Divergence, and Admixture in Domesticated Cattle. *Plos Genet.* 10 (3), e1004254. doi:10.1371/journal.pgen.1004254

Delport, W., Poon, A. F. Y., Frost, S. D. W., and Kosakovsky Pond, S. L. (2010). Datamonkey 2010: a Suite of Phylogenetic Analysis Tools for Evolutionary Biology. *Bioinformatics.* 26 (19), 2455–2457. doi:10.1093/bioinformatics/btq429

Dietz, A. B., Detilleux, J. C., Freeman, A. E., Kelley, D. H., Stabel, J. R., and Kehrli, M. E. (1997). Genetic Association of Bovine Lymphocyte Antigen DRB3 Alleles

with Immunological Traits of Holstein Cattle. *J. Dairy Sci.* 80 (2), 400–405. doi:10.3168/jds.s0022-0302(97)75950-2

Edwards, S. V., and Hedrick, P. W. (1998). Evolution and Ecology of MHC Molecules: from Genomics to Sexual Selection. *Trends Ecol. Evol.* 13 (8), 305–311. doi:10.1016/s0169-5347(98)01416-5

Ellis, S. (2004). The Cattle Major Histocompatibility Complex: Is it Unique? *Vet. Immunol. Immunopathol.* 102 (1-2), 1–8. doi:10.1016/j.vetimm.2004.06.007

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin Suite Ver 3.5: a New Series of Programs to Perform Population Genetics Analyses Under Linux and Windows. *Mol. Ecol. Resour.* 10 (3), 564–567. doi:10.1111/j.1755-0998.2010.02847.x

Galaverni, M., Caniglia, R., Fabbri, E., Lapalombella, S., and Randi, E. (2013). MHC Variability in an Isolated Wolf Population in Italy. *J. Hered.* 104 (5), 601–612. doi:10.1093/jhered/est045

García-Briones, M. M., Russell, G. C., Oliver, R. A., Tami, C., Taboga, O., Carrillo, E., et al. (2000). Association of Bovine DRB3 Alleles with Immune Response to FMDV Peptides and protection against Viral challenge. *Vaccine.* 19 (9-10), 1167–1171. doi:10.1016/s0264-410x(00)00313-3

Garrigan, D., and Hedrick, P. (2001). Class I MHC Polymorphism and Evolution in Endangered California Chinook and Other Pacific salmon. *Immunogenetics.* 53 (6), 483–489. doi:10.1007/s002510100352

Giovambattista, G., Takeshima, S.-n., Ripoli, M. V., Matsumoto, Y., Franco, L. A. A., Saito, H., et al. (2013). Characterization of Bovine MHC DRB3 Diversity in Latin American Creole Cattle Breeds. *Gene.* 519 (1), 150–158. doi:10.1016/j.gene.2013.01.002

Goszczynski, D. E., Corbi-Botto, C. M., Durand, H. M., Rogberg-Muñoz, A., Munilla, S., Peral-Garcia, P., et al. (2018). Evidence of Positive Selection Towards Zebuine Haplotypes in the BoLA Region of Brangus Cattle. *Animal.* 12 (2), 215–223. doi:10.1017/s1751731117001380

Gowane, G. R., Sharma, A. K., Sankar, M., Narayanan, K., Das, B., Subramaniam, S., et al. (2013). Association of BoLA DRB3 Alleles with Variability in Immune Response Among the Crossbred Cattle Vaccinated for Foot-And-Mouth Disease (FMD). *Res. Vet. Sci.* 95 (1), 156–163. doi:10.1016/j.rvsc.2013.03.001

Greenbaum, G., Templeton, A. R., Zarmi, Y., and Bar-David, S. (2014). Allelic Richness Following Population Founding Events-Aa Stochastic Modeling Framework Incorporating Gene Flow and Genetic Drift. *PLoS ONE.* 9, e115203–23. doi:10.1371/journal.pone.0115203

Hammer, Ø., Harper, D., and Ryan, P. D. (2001). PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontol. Electronica.* 4, 1–9.

Hughes, A. L., and Yeager, M. (1998). Natural Selection and the Evolutionary History of Major Histocompatibility Complex Loci. *Front. Biosci.* 3 (3), d509–16. doi:10.2741/a298

Hughes, A. L., and Nei, M. (1988). Pattern of Nucleotide Substitution at Major Histocompatibility Complex Class I Loci Reveals Overdominant Selection. *Nature.* 335 (6186), 167–170. doi:10.1038/335167a0

Hughes, A. L., and Nei, M. (1989). Nucleotide Substitution at Major Histocompatibility Complex Class II Loci: Evidence for Overdominant Selection. *Proc. Natl. Acad. Sci.* 86 (3), 958–962. doi:10.1073/pnas.86.3.958

Husemann, M., Zachos, F. E., Paxton, R. J., and Habel, J. C. (2016). Effective Population Size in Ecology and Evolution. *Heredity.* 117 (4), 191–192. doi:10.1038/hdy.2016.75

Ihaka, R., and Gentleman, R. (1996). R: A Language for Data Analysis and Graphics. *J. Comput. Graphical Stat.* 5 (3), 299–314. doi:10.1080/10618600.1996.10474713

Juliarena, M. A., Poli, M., Sala, L., Ceriani, C., Gutierrez, S., Dolcini, G., et al. (2008). Association of BLV Infection Profiles with Alleles of theBoLA-DRB3.2gene. *Anim. Genet.* 39 (4), 432–438. doi:10.1111/j.1365-2052.2008.01750.x

Klein, J., Sato, A., and Nikolaidis, N. (2007). MHC, TSP, and the Origin of Species: from Immunogenetics to Evolutionary Genetics. *Annu. Rev. Genet.* 41, 281–304. doi:10.1146/annurev.genet.41.110306.130137

Kosakovsky Pond, S. L., and Frost, S. D. W. (2005). Not so Different after All: a Comparison of Methods for Detecting Amino Acid Sites under Selection. *Mol. Biol. Evol.* 22 (5), 1208–1222. doi:10.1093/molbev/msi105

Kulberg, S., Heringstad, B., Guttersrud, O. A., and Olsaker, I. (2007). Study on the Association of BoLA-DRB3.2 Alleles with Clinical Mastitis in Norwegian Red Cows. *J. Anim. Breed. Genet.* 124 (4), 201–207. doi:10.1111/j.1439-0388.2007.00662.x

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35 (6), 547–1549. doi:10.1093/molbev/msy096

Leberg, P. L. (2002). Estimating Allelic Richness: Effects of Sample Size and Bottlenecks. *Mol. Ecol.* 11 (11), 2445–2449. doi:10.1046/j.1365-294x.2002.01612.x

Ledwidge, S. A., Mallard, B. A., Gibson, J. P., Jansen, G. B., and Jiang, Z. H. (2001). Multi-primer Target PCR for Rapid Identification of Bovine DRB3 Alleles. *Anim. Genet.* 32 (4), 219–221. doi:10.1046/j.1365-2052.2001.00761.x

Lenz, T. L., and Becker, S. (2008). Simple Approach to Reduce PCR Artefact Formation Leads to Reliable Genotyping of MHC and Other Highly Polymorphic Loci-Iimplications for Evolutionary Analysis. *Gene.* 427 (1-2), 117–123. doi:10.1016/j.gene.2008.09.013

Leroy, G., Mary-Huard, T., Verrier, E., Danvy, S., Charvolin, E., and Danchin-Burge, C. (2013). Methods to Estimate Effective Population Size Using Pedigree Data: Examples in Dog, Sheep, Cattle and Horse. *Genet. Sel Evol.* 45 (1), 1. doi:10.1186/1297-9686-45-1

Li, C. C. *First Course in Population Genetics1976 Pacific Grove.* California, USA: Boxwood Press, 631.

Maccari, G., Robinson, J., Ballingall, K., Guethlein, L. A., Grimholt, U., Kaufman, J., et al. (2017). IPD-MHC 2.0: an Improved Inter-species Database for the Study of the Major Histocompatibility Complex. *Nucleic Acids Res.* 45 (D1), D860–D864. doi:10.1093/nar/gkw1050

Maillard, J. C., Berthier, D., Chantal, I., Thevenon, S., Sidibé, I., Stachurski, F., et al. (2003). Selection Assisted by a BoLA-DR/DQ Haplotype against Susceptibility to Bovine Dermatophilosis. *Genet. Sel Evol.* 35 (Suppl. 1), S193–S200. doi:10.1186/1297-9686-35-S1-S193

Martinez, M. L., Machado, M. A., Nascimento, C. S., Silva, M. V., Teodoro, R. L., Furlong, J., et al. (2006). Association of BoLA-DRB3.2 Alleles with Tick (Boophilus Microplus) Resistance in Cattle. *Genet. Mol. Res.* 5 (3), 513–524.

Mason, R. A. B., Browning, T. L., and Eldridge, M. D. B. (2011). Reduced MHC Class II Diversity in Island Compared to mainland Populations of the Black-Footed Rock-Wallaby (Petrogale Lateralis Lateralis). *Conserv Genet.* 12 (1), 91–103. doi:10.1007/s10592-009-9993-y

Mikko, S., Spencer, M., Morris, B., Stabile, S., Basu, T., Stormont, C., et al. (1997). A Comparative Analysis of Mhc DRB3 Polymorphism in the American Bison (*Bison bison*). *J. Hered.* 88 (6), 499–503. doi:10.1093/oxfordjournals.jhered.a023144

Miltiadou, D., Law, A. S., and Russell, G. C. (2003). Establishment of a Sequence-Based Typing System for BoLA-DRB3 Exon 2. *Tissue Antigens.* 62 (1), 55–65. doi:10.1034/j.1399-0039.2003.00080.x

Moutou, K. A., Koutsogiannouli, E. A., Stamatis, C., Billinis, C., Kalbe, C., Scandura, M., et al. (2013). Domestication Does Not Narrow MHC Diversity in *Sus scrofa*. *Immunogenetics.* 65 (3), 195–209. doi:10.1007/s00251-012-0671-8

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., and Kosakovsky Pond, S. L. (2012). Detecting Individual Sites Subject to Episodic Diversifying Selection. *Plos Genet.* 8 (7), e1002764. doi:10.1371/journal.pgen.1002764

Nascimento, C. S. d., Machado, M. A., Martinez, M. L., Silva, M. V. G. B. d., Guimarães, M. F. M., Campos, A. L., et al. (2006). Association of the Bovine Major Histocompatibility Complex (BoLA) BoLA-DRB3 Gene with Fat and Protein Production and Somatic Cell Score in Brazilian Gyr Dairy Cattle (*Bos indicus*). *Genet. Mol. Biol.* 29, 641–647. doi:10.1590/s1415-47572006000400011

Neefjes, J., Jongsma, M. L. M., Paul, P., and Bakke, O. (2011). Towards a Systems Understanding of MHC Class I and MHC Class II Antigen Presentation. *Nat. Rev. Immunol.* 11 (12), 823–836. doi:10.1038/nri3084

Nei, M., and Chesser, R. K. (1983). Estimation of Fixation Indices and Gene Diversities. *Ann. Hum. Genet.* 47 (3), 253–259. doi:10.1111/j.1469-1809.1983.tb00993.x

Nei, M. (1978). Estimation of Average Heterozygosity and Genetic Distance from a Small Number of Individuals. *Genetics.* 89 (3), 583–590. doi:10.1093/genetics/89.3.583

Nei, M., and Tajima, F. (1981). Genetic Drift and Estimation of Effective Population Size. *Genetics.* 98 (3), 625–640. doi:10.1093/genetics/98.3.625

Newhouse, D. J., and Balakrishnan, C. N. (2015). High Major Histocompatibility Complex Class I Polymorphism Despite Bottlenecks in Wild and Domesticated Populations of the Zebra Finch (*Taeniopygia guttata*). *BMC Evol. Biol.* 15 (265), 265–0546. doi:10.1186/s12862-015-0546-3

Nicholas, K. B. (1997). GeneDoc : Analysis and Visualization of Genetic Variation. *Embnew. News.* 4, 14.

Nielsen, M., Lundegaard, C., Blicher, T., Peters, B., Sette, A., Justesen, S., et al. (2008). Quantitative Predictions of Peptide Binding to Any HLA-DR Molecule of Known Sequence: NetMHCIIpan. *Plos Comput. Biol.* 4 (7), e1000107. doi:10.1371/journal.pcbi.1000107

Nielsen, M., Connelley, T., and Ternette, N. (2018). Improved Prediction of Bovine Leucocyte Antigens (BoLA) Presented Ligands by Use of Mass-Spectrometry-Determined Ligand and *In Vitro* Binding Data. *J. Proteome Res.* 17 (1), 559–567. doi:10.1021/acs.jproteome.7b00675

Nikbakht Brujeni, G., Ghorbanpour, R., Esmailnejad, A., and Association of BoLA-Drb3 (2016). Association of BoLA-DRB3.2 Alleles with BLV Infection Profiles (Persistent Lymphocytosis/Lymphosarcoma) and Lymphocyte Subsets in Iranian Holstein Cattle. *Biochem. Genet.* 54 (2), 194–207. doi:10.1007/s10528-016-9712-6

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R Language. *Bioinformatics.* 20 (2), 289–290. doi:10.1093/bioinformatics/btg412

Rappazzo, C. G., Huisman, B. D., and Birnbaum, M. E. (2020). Repertoire-scale Determination of Class II MHC Peptide Binding via Yeast Display Improves Antigen Prediction. *Nat. Commun.* 11 (1), 4414. doi:10.1038/s41467-020-18204-2

Rock, K. L., Reits, E., and Neefjes, J. (2016). Present Yourself! by MHC Class I and MHC Class II Molecules. *Trends Immunol.* 37 (11), 724–737. doi:10.1016/j.it.2016.08.010

Rousset, F. (2008). genepop'007: a Complete Re-implementation of the Genepop Software for Windows and Linux. *Mol. Ecol. Resour.* 8 (1), 103–106. doi:10.1111/j.1471-8286.2007.01931.x

Rupp, R., Hernandez, A., and Mallard, B. A. (2007). Association of Bovine Leukocyte Antigen (BoLA) DRB3.2 with Immune Response, Mastitis, and Production and Type Traits in Canadian Holsteins. *J. Dairy Sci.* 90 (2), 1029–1038. doi:10.3168/jds.s0022-0302(07)71589-8

Sharif, S., Mallard, B. A., Wilkie, B. N., Sargeant, J. M., Scott, H. M., Dekkers, J. C. M., et al. (1999). Associations of the Bovine Major Histocompatibility Complex DRB3 (BoLA-DRB3) with Production Traits in Canadian Dairy Cattle. *Anim. Genet.* 30 (2), 157–160. doi:10.1046/j.1365-2052.1999.00459.x

Sigurdardóttir, S., Borsch, C., Gustafsson, K., and Andersson, L. (1991). Cloning and Sequence Analysis of 14 DRB Alleles of the Bovine Major Histocompatibility Complex by Using the Polymerase Chain Reaction. *Anim. Genet.* 22 (3), 199–209. doi:10.1111/j.1365-2052.1991.tb00670.x

Spieth, P. T. (1974). Gene Flow and Genetic Differentiation. *Genetics.* 78 (3), 961–965. doi:10.1093/genetics/78.3.961

Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L., et al. (1994). Crystal Structure of the Human Class II MHC Protein HLA-DR1 Complexed with an Influenza Virus Peptide. *Nature.* 368 (6468), 215–221. doi:10.1038/368215a0

Suárez, C. F., Patarroyo, M. E., Trujillo, E., Estupiñán, M., Baquero, J. E., Parra, C., et al. (2006). Owl Monkey MHC-DRB Exon 2 Reveals High Similarity with Several HLA-DRB Lineages. *Immunogenetics.* 58 (7), 542–558. doi:10.1007/s00251-006-0127-0

Sutton, J. T., Nakagawa, S., Robertson, B. C., and Jamieson, I. G. (2011). Disentangling the Roles of Natural Selection and Genetic Drift in Shaping Variation at MHC Immunity Genes. *Mol. Ecol.* 20 (21), 4408–4420. doi:10.1111/j.1365-294x.2011.05292.x

Tajima, F., and Nei, M. (1984). Note on Genetic Drift and Estimation of Effective Population Size. *Genetics.* 106 (3), 569–574. doi:10.1093/genetics/106.3.569

Takeshima, S.-N., Giovambattista, G., Okimoto, N., Matsumoto, Y., Rogberg-Muñoz, A., Acosta, T. J., et al. (2015). Characterization of Bovine MHC Class II DRB3 Diversity in South American Holstein Cattle Populations. *Tissue Antigens.* 86 (6), 419–430. doi:10.1111/tan.12692

Takeshima, S., Matsumoto, Y., Chen, J., Yoshida, T., Mukoyama, H., and Aida, Y. (2008). Evidence for Cattle Major Histocompatibility Complex (BoLA) Class IIDQA1gene Heterozygote Advantage against Clinical Mastitis Caused byStreptococciandEscherichiaspecies. *Tissue Antigens.* 72 (6), 525–531. doi:10.1111/j.1399-0039.2008.01140.x

Takeshima, S. N., Corbi-Botto, C., Giovambattista, G., and Aida, Y. (2018). Genetic Diversity of BoLA-DRB3 in South American Zebu Cattle Populations. *BMC Genet.* 19 (1), 33–0618. doi:10.1186/s12863-018-0618-7

Takeshima, S. N., Miyasaka, T., Polat, M., Kikuya, M., Matsumoto, Y., Mingala, C. N., et al. (2014). The Great Diversity of Major Histocompatibility Complex Class II Genes in Philippine Native Cattle. *Meta Gene* 2, 176–190. doi:10.1016/j.mgene.2013.12.005

Takeshima, S., Saitou, N., Morita, M., Inoko, H., and Aida, Y. (2003). The Diversity of Bovine MHC Class II DRB3 Genes in Japanese Black, Japanese Shorthorn, Jersey and Holstein Cattle in Japan. *Gene.* 316, 111–118. doi:10.1016/s0378-1119(03)00744-3

Takezaki, N., Nei, M., and Tamura, K. (2010). POPTREE2: Software for Constructing Population Trees from Allele Frequency Data and Computing Other Population Statistics with Windows Interface. *Mol. Biol. Evol.* 27 (4), 747–752. doi:10.1093/molbev/msp312

Taylor, S. S., Jenkins, D. A., and Arcese, P. (2012). Loss of Mhc and Neutral Variation in Peary Caribou: Genetic Drift Is Not Mitigated by Balancing Selection or Exacerbated by Mhc Allele Distributions. *Plos One.* 7 (5), e36748. doi:10.1371/journal.pone.0036748

Trtkova, K., Mayer, W. E., O'hUign, C., and Klein, J. (1995). Mhc-DRB Genes and the Origin of New World Monkeys. *Mol. Phylogenet. Evol.* 4 (4), 408–419. doi:10.1006/mpev.1995.1038

Uinuk-Ool, T. S., Takezaki, N., Sukernik, R. I., Nagl, S., and Klein, J. (2002). Origin and Affinities of Indigenous Siberian Populations as Revealed by HLA Class II Gene Frequencies. *Hum. Genet.* 110 (3), 209–226. doi:10.1007/s00439-001-0668-0

Ujvari, B., and Belov, K. (2011). Major Histocompatibility Complex (MHC) Markers in Conservation Biology. *Int. J. Mol. Sci.* 12 (8), 5168–5186. doi:10.3390/ijms12085168

Weir, B. S., and Cockerham, C. C. (1984). Estimating F -Statistics for the Analysis of Population Structure. *Evolution.* 38 (6), 1358–1370. doi:10.1111/j.1558-5646.1984.tb05657.x

Zhang, X., Lin, W., Zhou, R., Gui, D., Yu, X., and Wu, Y. (2016). Low Major Histocompatibility Complex Class II Variation in the Endangered Indo-Pacific Humpback Dolphin (Sousa Chinensis): Inferences about the Role of Balancing Selection. *Jhered.* 107 (2), 143–152. doi:10.1093/jhered/esv138