TISSUE ANTIGENS
IMMUNE RESPONSE GENETICS

Tissue Antigens ISSN 0001-2815

BRIEF COMMUNICATION

Haplotype determination of the upstream regulatory region and the second exon of the *BoLA-DRB3* gene in Holstein cattle

D. E. Goszczynski[1,†], M. V. Ripoli[1], S.-N. Takeshima[2], L. Baltian[3], Y. Aida[2] & G. Giovambattista[1]

1 Instituto de Genética Veterinaria (IGEVET), CCT La Plata – CONICET, Facultad de Ciencias Veterinarias, Universidad Nacional de La Plata, La Plata, Argentina
2 Viral Infectious Diseases Unit, RIKEN, Wako, Japan
3 Dpto de Producción, FCV, UNLPam, General Pico, Argentina

Key words
BoLA-DRB3; linkage disequilibrium; polymorphism; selection

Correspondence
Dr Guillermo Giovambattista
Instituto de Genética Veterinaria (IGEVET)
CCT La Plata – CONICET
School of Veterinary Medicine
National University of La Plata
60 y 118 S/N, CC 296
La Plata 1900
Argentina
Tel: +54 221 4211799
Fax: +54 221 4211799
e-mail: ggiovam@fcv.unlp.edu.ar

Received 28 August 2013; revised 27 November 2013; accepted 17 December 2013

doi: 10.1111/tan.12293

Abstract

Polymorphisms of the BoLA-DRB3 gene are located primarily in the second exon [antigen binding site (ABS)] and, to a lesser extent, in the upstream regulatory region (URR). It can be hypothesised that exon 2 and the URR are under different types of natural selection. The aim of this work was to determine the URR-exon 2 haplotypes; 34 Holstein samples were genotyped by direct sequencing. A total of 7 URR alleles and 23 exon 2 alleles were detected, and 3 of the URR alleles were novel. Our results may suggest that no relationship exists between the URR and exon 2 of the BoLA-DRB3 gene (linkage disequilibrium P value > 0.05), most likely due to recombination over time. Our results also suggest that both regions of class II genes may be included in the development of new genotyping methods based on next-generation DNA sequencing technologies.

BoLA-DRB3 is a class II gene that plays a major role in the immune response (1, 2). BoLA-DRB3 polymorphisms, which are located primarily in the second exon that encodes the ABS of the molecule (1, 3), have been extensively studied in various cattle breeds (4–6). A number of these polymorphisms have been associated with resistance/susceptibility to infectious and autoimmune diseases, such as leukaemia virus, dermatophilosis and mastitis (1, 6–11).

Additional variation has also been detected in other gene regions, such as the upstream regulatory region (URR) (12–16). Polymorphisms found in conserved consensus sequences could affect DNA–protein interactions (17, 18). Consequently, these mutations may confer allelic differences in expression, inducibility and/or tissue specificity of class II molecules, affecting the transcriptional levels and the immune response.

It can be hypothesised that the second exon and the URR are under different types of natural selection. The unusually high levels of polymorphism found in the second exon may have been maintained by overdominance or balancing selection (19); this possibility is supported by the high number of non-synonymous changes in the ABS (3). Conversely, the URR presents a moderate degree of polymorphism that may be purged by positive selection. For this reason, the aim this present work was to determine the URR-exon 2 haplotypes of the BoLA-DRB3 gene.

Blood samples from 34 Holstein cattles were obtained from various Argentine dairy farms, and the sample size was sufficient to detect the most frequent BoLA-DRB3 exon 2 alleles in this breed. Genomic DNA was extracted using Wizard® Genomic DNA Purification kits (Promega, Madison, WI) following the manufacturer's instructions. The regions were genotyped by sequencing based typing methods [polymerase chain reaction-sequence based typing (PCR-SBT)], as described by Ripoli et al. (13) and Takeshima et al. (20). The alleles were identified using ASSIGN 400 ATF ver 1.0.2.41. Linkage disequilibrium (LD) among the single nucleotide polymorphisms (SNPs) of the URR was determined by PHASE

[†]Fellows from CONICET.

Table 1 Polymorphic sites that defined the upstream regulatory region (URR) haplotypes of the BoLA-DRB3 gene detected in the studied Holstein population, and their gene frequencies

	Polymorphic sites (bp)						
Haplotype	−161	−138	−135	−27	−20	Gene frequency	Previous allele
URR-DRB3-Hol*1	C	G	T	A	A	0.25	DRB-H5-R-U[a]
URR-DRB3-Hol*2	C	G	T	A	G	0.03	DRB-H8-R-U[a]
URR-DRB3-Hol*3	C	G	T	G	A	0.40	New
URR-DRB3-Hol*4	C	G	C	A	A	0.16	BoLA-DRB-D[b]
URR-DRB3-Hol*5	C	A	T	A	A	0.12	New
URR-DRB3-Hol*6	C	A	T	A	G	0.01	New
URR-DRB3-Hol*7	T	G	T	G	A	0.03	BoLA-DRB-B[b], DRB-H19-R-U[a]

[a]Wang et al. (16).
[b]Ripoli et al. (13).

(21, 22) and visualised on HAPLOVIEW (23), while URR-exon 2 haplotypes were determined by PHASE. *P* value of LD was performed using the G-test implemented in GENEPOP 4.0 (24).

The observed (h_o) and unbiased expected heterozygosity (h_e) were estimated according to Nei (25), using ARLEQUIN 3.5 (26), and deviations from Hardy–Weinberg equilibrium (HWE) were estimated by F_{IS} statistics (27), using the exact test included in GENEPOP 4.0. The Ewens–Watterson–Slatkin exact neutrality test was estimated using the method described by Slatkin (28) and was implemented in the ARLEQUIN 3.5 program. Genetic diversity at the DNA sequence level was estimated through nucleotide diversity (π) and the mean number of pairwise differences (NPD) using ARLEQUIN 3.5. Pairwise genetic distances between BoLA-DRB3 exon 2 DNA sequences were estimated on the basis of Kimura's two-parameter model (29). The gene tree was constructed from a distance matrix that was based on the neighbor joining (NJ) method of Saitou and Nei (30).

The URR-BoLA-DRB3 DNA sequences obtained in this study and those previously reported (AF510446, AJ488500, AY040327, AY364454, AY364455, AY550181, AY550183, AY550184, AY550185, AY550186, AY550187, AY550188, AY550189, AY550190, AY550191, AY550192,AY570362, AY858800,FQ482091, FQ482110, JN803939, JN803940, JN803941, JN803942, JN803943, JN803944, and JN803945, JN803946) were aligned to identify polymorphic sites. This comparison allowed us to detect five SNPs in the studied Holstein population; however, none of them corresponded to new polymorphic sites. Previous studies showed that the URR is composed of highly conserved sequence motifs that include from 5′ to 3′, the W, X, Y, CCAAT and TATA-like boxes (12, 13, 17, 18). Although these motifs are highly conserved among mammals, polymorphisms have been reported within and between boxes (12, 17). However, all SNPs detected in this work were transitions located in interconsensus regulatory regions (Figure S1, *Supporting Information*).

LD analysis revealed seven URR haplotypes. Four of them matched with haplotypes previously reported: BoLA-DRB-B (accession number: AY364454; also named DRB-H19-R-U, accession number: JN803945), BoLA-DRB-D (accession number: AY570362), DRB-H8-R-U (accession number: JN80 3941) and DRB-H5-R-U (accession number: JN803939). The remaining three corresponded to new haplotypes (accession number: KF576968, KF576969, KF576970) (Figures S1 and S2, and Table 1). The gene frequencies of the detected haplotypes varied from 0.015 (URR-DRB3-Hol*6) to 0.397 (URR-DRB3-Hol*3) (Table 1).

A distinguishing feature of the second exon of the DRB gene is the high degree of polymorphism in combination with a relatively even distribution of allele frequencies. Furthermore, a significant excess of heterozygous genotypes was reported in different populations (9, 19, 31). It had been proposed that this extensive polymorphism is maintained by a selective mechanism(s), such as overdominace or balancing selection, maternal–foetal interactions and non-random mating (19, 31). In contrast, mutations on most of DNA sequences, such as URR, are expected to be purged by positive selection (13, 32). For this reason, a different degree of diversity is expected depending on the studied DNA region: URR or exon 2 (ABS). Comparison of diversity values at both allele and nucleotide levels showed a considerably higher diversity in exon 2 than in URR, although a moderate degree of variability was detected in URR (Table 2). In this way, 23 alleles were detected in exons 2 and 7 in the URR. Nucleotide diversity values were 0.010 for URR and 0.103 for exon 2, and for the NPD, 1.14 and 25.75. HWE and neutrality tests neither showed significant deviation from theoretical proportions, nor overdominace or balancing selection in URR and exon 2 (Table 2).

To analyse the LD between the URR and the second exon and to determine whether certain groups of alleles of exon 2 were associated with haplotypes of the URR, a NJ tree of the second exon was constructed, and the associated URR haplotypes were overlapped (Figure S3). This analysis showed that each URR haplotype was associated with a range of exon 2 variants that varied from 1 to 15. The URR haplotypes were spread among the tree and linked to exon 2 alleles

Table 2 Standard and molecular diversity were estimated through the number of alleles (n_a), observed (h_o) and expected heterozygosities (h_e), nucleotide diversity (π), mean number of pairwise differences (NPD), for the upstream regulatory region (URR) and the second exon in the studied Holstein population. Hardy–Weinberg equilibrium (HWE), measure through F_{IS}, and Slatkin's exact neutrality test were calculated for the cattle breeds studied.

Gene region	n_a	h_o	(h_e)	π	NPD	HWE F_{IS} – P value	Slatkin's exact P value
URR	7	0.18	0.75	0.103	1.14	0.767–< 0.00	0.608
Exon 2	22	0.912	0.935	0.005	25.75	0.025–0.900	0.191

Table 3 BoLA-DRB3 exon 2 alleles associated with each BoLA-DRB3-URR haplotypes

URR haplotype	Exon 2 haplotypes
URR-DRB3-Hol*1	BoLA-DRB3*0101, 0301, 1001, 14011, 1601, 1801, 2502, 2703, 2707
URR-DRB3-Hol*2	BoLA-DRB3*0701
URR-DRB3-Hol*3	BoLA-DRB3*0101, 0201, 0601, 0801, 0901, 0902, 1001, 1101, 1102, 1104, 1201, 1501, 2402, 2703, 3601
URR-DRB3-Hol*4	BoLA-DRB3*0101, 0301, 1101, 14011, 2006
URR-DRB3-Hol*5	BoLA-DRB3*0301, 0601, 0901, 0902, 1001
URR-DRB3-Hol*6	BoLA-DRB3*0902
URR-DRB3-Hol*7	BoLA-DRB3*0902, 1201

that belonged to different clusters. Moreover, several exon 2 alleles were linked to more than one URR haplotype (Table 3; Figure S3). G-test confirmed the lack of LD (P value of LD > 0.05).

In large effective size populations, even though recombination between tightly linked SNPs occurs slowly, LD is maintained through generations over short distances, given that the LD represents a balance between mutation, drift and recombination (33). Recently, whole genome LD studies showed that the mean block size varied from 5.7 to 15.7 kb across breeds (with a mean block size of 10.3 kb over all breeds), and the r^2 values were high only at small distances (33–35). Because the URR and exon 2 of BoLA-DRB3 are located approximately 8 kb away (GeneID:282530), URR – second exon of BoLA-DRB3 LD is expected to be maintained over the time. However, our results suggest that no relationship occurs between the URR and exon 2 of the BoLA-DRB3 gene, most likely due to recombination.

In conclusion, these results increase our knowledge of the co-evolution of two DNA sequences under different types of selection. Description of the URR-exon 2 LD would be useful for further studies of resistance/susceptibility to infection and autoimmune diseases because the immune response might depend on the major histocompatibility complex (MHC) molecule – antigen binding affinity and the level of MHC gene expression. For this reason, both regions of class II genes may be included in the development of new genotyping methods based on next-generation DNA sequencing technologies.

Acknowledgments

The studies were supported by Grants-in-Aid for Scientific Research (A, B and C) from the Japan Society for the Promotion of Science (JSPS), and by a grant from the Program for the Promotion of Basic and Applied Research for Innovations in Bio-oriented Industry, and was also supported by ANPCYT, CONICET and UNLP from Argentine.

Conflict of interest

The authors have declared no conflicting interests.

References

1. Rothschild MF, Skow L, Lamont SJ. The major histocompatibility complex and its role in disease resistance and immune responsiveness. In: Axford RFE, Bishop SC, Nichols FW, Owen JB, eds. *Breeding for Disease Resistance in Farm Animals*, *2nd edn. Wallingford*: *CABI publishing*, *2000*, *73*.
2. Takeshima SN, Aida Y. Structure, function and disease susceptibility of the bovine major histocompatibility complex. *Anim Sci J* 2006: **77**: 138–50.
3. Takeshima SN, Sarai Y, Saitou N, Aida Y. MHC class II DR classification based on antigen-binding groove natural selection. *Biochem Biophys Res Commun* 2009: **385**: 137–42.
4. Takeshima SN, Saitou N, Morita M, Inoko H, Aida Y. The diversity of bovine MHC class II DRB3 genes in Japanese Black, Japanese Shorthorn, Jersey and Holstein cattle in Japan. *Gene* 2003: **316**: 111–8.
5. Miyasaka T, Takeshima SN, Sentsui H, Aida Y. Identification and diversity of bovine major histocompatibility complex class II haplotypes in Japanese Black and Holstein cattle in Japan. *J Dairy Sci* 2012: **95**: 420–31.
6. Giovambattista G, Takeshima SN, Ripoli MV et al. Characterization of bovine MHC DRB3 diversity in Latin American Creole cattle breeds. *Gene* 2013: **519**: 150–8.
7. Xu A, van Eijk MJ, Park C, Lewin HA. Polymorphism inBoLA-DRB3 exon 2 correlates with resistance to persistentlymphocytosis caused by bovine leukemia virus. *J Immunol* 1993: **151**: 6977–85.
8. Maillard JC, Chantal I, Berthier D, Thevenon S, Sidibe I, Razafindraibe H. Molecular immunogenetics in susceptibility to bovine dermatophilosis: a candidate gene approach and a concrete field application. *Ann N Y Acad Sci* 2002: **969**: 92–6.
9. Takeshima SN, Matsumoto Y, Chen J, Yoshida T, Mukoyama H, Aida Y. Evidence for cattle major histocompatibility complex (BoLA) class II DQA1 gene heterozygote advantage

against clinical mastitis caused by Streptococci and Escherichia species. *Tissue Antigens* 2008: **72**: 525–31.
10. Yoshida T, Mukoyama H, Furuta H et al. Association of the amino acid motifs of BoLA-DRB3 alleles with mastitis pathogens in Japanese Holstein cows. *Anim Sci J* 2009: **80**: 510–9.
11. Miyasaka T, Takeshima SN, Jimba M et al. Identification of bovine leukocyte antigen class II haplotypes associated with variations in bovine leukemia virus proviral load in Japanese Black cattle. *Tissue Antigens* 2013: **81**: 72–82.
12. Ripoli MV, Díaz S, Peral-García P, Giovambattista G. Nucleotide sequence of the upstream regulatory region of the BoLA-DRB. *Eur J Immunogenet* 2002: **29**: 537–40.
13. Ripoli MV, Peral-García P, Dulout FN, Giovambattista G. Polymorphism in the Bovine BoLA-DRB3 Upstream Regulatory Regions detected through PCR-SSSP and DNA sequencing. *Gene* 2004: **339**: 71–8.
14. Russell GC, Smith JA, Oliver RA. Structure of the BoLA-DRB3 gene and promoter. *Eur J Immunogenet* 2004: **31**: 145–51.
15. Ripoli MV, Villegas Castagnasso EE, Peral-García P, Giovambattista G. New polymorphisms for the BoLA-DRB3 upstream regulatory region. *Tissue Antigens* 2005: **66**: 136–7.
16. Wang K, Sun DX, Li KY, Wang XQ, Zhang F. Identification of four novel alleles of the BoLA-DRB3 upstream regulatory region in Chinese yellow cattle. *Tissue Antigens* 2012: **80**: 58–60.
17. Singal DP, Qiu X. Polymorphism in both X and Y motifs controls level of expression of HLA-DRB1 genes. *Immunogenetics* 1996: **43**: 50–6.
18. Emery P, Mach B, Reith W. The different level of expression of HLA-DRB1 and -DRB3 genes is controlled by conserved isotopic differences in promoter sequence. *Hum Immunol* 1993: **38**: 137–47.
19. Hedrick PW. Balancing selection and MHC. *Genetica* 1998: **104**: 207–14.
20. Takeshima SN, Matsumoto Y, Miyasaka T et al. A new method for typing bovine major histocompatibility complex class II DRB3 alleles by combining two established PCR sequence-based techniques. *Tissue Antigens* 2011: **78**: 208–13.
21. Crawford DT, Bhangale N, Li G et al. Evidence for substantial finne-scale variation in recombination rates across the human genome. *Nat Genet* 2004: **36**: 700–6.
22. Li N, Stephens M. Modelling linkage disequilibrium, and identifying recombination hotspots using SNP data. *Genetics* 2003: **165**: 2213–33.
23. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005: **21**: 263–5.
24. Rousset F. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour* 2008: **8**: 103–6.
25. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 1978: **89**: 583–90.
26. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 2010: **10**: 564–7.
27. Weir B, Cockerham C. Estimating F-statistics for the analysis of population structure. *Evolution* 1984: **38**: 1358–70.
28. Slatkin M. A correction to the exact test based on the Ewens sampling distribution. *Genet Res* 1996: **68**: 259–60.
29. Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 1980: **16**: 111–20.
30. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987: **4**: 406–25.
31. Penn DJ, Damjanovich K, Potts WK. MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc Natl Acad Sci U S A* 2002: **99**: 11260–4.
32. Klein J, Gutknecht J, Fischer N. The major histocompatibility complex and human evolution. *Trends Genet* 1990: **6**: 7–11.
33. Goddard ME, Hayes BJ. Genome-wide association studies and linkage diequilibrium in cattle. In: Womack JE, ed. *Bovine Genomics*. *Chapter 3*. *Oxford*: *Wiley-Blackwell*, *2012*, *211–33*.
34. Villa-Angulo R, Matukumalli LK, Gill CA, Choi J, Van Tassell CP, Grefenstette JJ. High-resolution haplotype block structure in the cattle genome. *BMC Genet* 2009: **10**: 19. DOI: 10.1186/1471-2156-10-19.
35. Qanbari S, Pimentel ECG, Tetens J et al. The pattern of linkage disequilibrium in German Holstein cattle. *Genetics* 2010: **41**: 346–56.

Supporting Information

The following supporting information is available for this article:

Figure S1. Alignment of the nucleotide sequences of the upstream regulatory region (URR) of BoLA-DRB3 previously reported and obtained in this work (Hol*1 to Hol*7) URR-BoLA-DRB3.The A of ATG is designated as +1. Dots indicate nucleotide identity to the consensus sequence, and dashes (−) represent gaps introduced to achieve the best alignment. Boxed sequences represent the W, X, Y, CCAAT and TATA regions. Arrows indicate the location of the primers.

Figure S2. Linkage disequilibrium (r^2) plot obtained in upstream regulatory region of BoLA-DRB3 gene with HAPLOVIEW 3.31 (21). Solid lines mark the block identified.

Figure S3. Neighbour joining tree of the BoLA-DRB3 exon 2 alleles. Arrows indicate the BoLA-DRB3-URR associated haplotypes.