UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

POLIMORFISMOS ASSOCIADOS À CARACTERÍSTICAS DE LEITE EM BUBALINOS E À MORFOLOGIA ESPERMÁTICA EM BOVINOS

Ana Cláudia de Freitas

Zootecnista

2018

UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

POLIMORFISMOS ASSOCIADOS À CARACTERÍSTICAS DE LEITE EM BUBALINOS E À MORFOLOGIA ESPERMÁTICA EM BOVINOS

Ana Cláudia de Freitas

Orientador: Prof. Dr. Humberto Tonhati

Coorientadora: Dra. Nedenia Bonvino Stafuzza

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutor em Genética e Melhoramento Animal.

2018

Freitas, Ana Cláudia de
F866p Polimorfismos associados à características de leite em bubalinos e à morfologia espermática em bovinos / Ana Cláudia de Freitas. – – Jaboticabal, 2018
iv, 41 p. : il. ; 29 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2018
Orientador: Humberto Tonhati
Coorientadora: Nedenia Bonvino Stafuzza
Banca examinadora: Vera Fernanda Martins Hossepian de Lima, Larissa Fernanda Simielli Fonseca Noronha, Marina Mortati Dias Barbero, Luciana Correia de Almeida Regitano
Bibliografia

1. Espermatozoide. 2. Mastite. 3. Polimorfismo. 4. Reprodução I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 636.082.22

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação – Diretoria Técnica de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: POLIMORFISMOS ASSOCIADOS À CARACTERÍSTICAS DE LEITE EM BUBALINOS E À MORFOLOGIA ESPERMÁTICA EM BOVINOS

AUTORA: ANA CLÁUDIA DE FREITAS
ORIENTADOR: HUMBERTO TONHATI
COORIENTADORA: NEDENIA BONVINO STAFUZZA

Aprovada como parte das exigências para obtenção do Título de Doutora em GENÉTICA E MELHORAMENTO ANIMAL, pela Comissão Examinadora:

Pós-doutoranda NEDENIA BONVINO STAFUZZA
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal

Profa. Dra. VERA FERNANDA MARTINS HOSSEPIAN DE LIMA
Departamento de Medicina Veterinária Preventiva e Reprodução Animal / FCAV / UNESP - Jaboticabal

Pesquisadora Dra. LUCIANA CORREIA DE ALMEIDA REGITANO
Empresa Brasileira de Pesquisa Agropecuária / EMBRAPA / São Carlos/SP

Pós-doutoranda LARISSA FERNANDA SIMIELLI FONSECA NORONHA
Departamento de Zootecnia / FCAV / UNESP - Jaboticabal

Profa. Dra. MARINA MORTATI DIAS BARBERO
Departamento de Produção Animal / Universidade Federal Rural do Rio de Janeiro

Jaboticabal, 27 de fevereiro de 2018

DADOS CURRICULARES DO AUTOR

ANA CLÁUDIA DE FREITAS - nascida em 8 de março de 1988, na cidade de Araraquara – SP, filha de Cláudio de Freitas e Ana Rita Seno de Freitas. Iniciou em fevereiro de 2007 o curso de graduação em Zootecnia na Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista "Júlio de Mesquita Filho", campus de Jaboticabal obtendo o título de Zootecnista em fevereiro de 2011. Durante a graduação, foi bolsista de Iniciação Científica da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) sob orientação do Prof. Dr. Humberto Tonhati. Em agosto de 2012, ingressou no Programa de Pós-graduação em Genética e Melhoramento Animal da Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista "Júlio de Mesquita Filho", campus de Jaboticabal como bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), sob orientação do Prof. Dr. Humberto Tonhati, obtendo o título de Mestre em 25 de fevereiro de 2014. Em março de 2014 iniciou o doutorado na mesma instituição e no mesmo programa de pós-graduação, com orientação do Prof. Dr. Humberto Tonhati e coorientação da Dra. Nedenia Bonvino Stafuzza. Atualmente é bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e realizou parte do seu doutorado (doutorado sanduíche) na Universidade de Queensland – Austrália, sob orientação da Profa. Dra. Marina R. S. Fortes.

“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar.
Mas o mar seria menor se lhe faltasse uma gota”.

Madre Teresa de Calcutá

AGRADECIMENTOS

À minha família e principalmente aos meus pais, sempre presentes em minha jornada. Sem eles nada disso faria sentido. Serei eternamente grata.

Ao meu orientador de tantos anos, Prof. Humberto Tonhati. Obrigada sempre pela oportunidade e ensinamentos. Tenho sorte por ter um amigo e orientador na mesma pessoa.

À minha coorientadora Nedenia, uma das peças principais para que essa pesquisa acontecesse. Obrigada pela paciência, dedicação e pelos conselhos. Você colaborou com meu crescimento pessoal e profissional.

Ao Daniel Jordan, por ter participado do projeto e me ajudado principalmente com as análises estatísticas. Obrigada pelo suporte e ensinamentos.

Aos membros da banca de qualificação, Prof. Henrique Nunes, Prof. Fernando Baldi, Dra. Larissa Simielli, Dra. Marina Mortati e aos atuais membros da banca de defesa. As correções e sugestões dadas por vocês serão de grande valor para esse trabalho.

À Dra. Marina Fortes, que tão bem me recebeu na University of Queensland e em sua equipe. Foi uma experiência maravilhosa trabalhar com alguém tão cativante.

Ao Prof. Danísio Prado Munari e à Profa. Lucia Galvão de Albuquerque que tão bem conduzem nosso programa de pós-graduação em Genética e Melhoramento Animal.

Aos meus amigos de Jaboticabal, em especial Raphael, Mariana, Camila, Thaise e Fernanda. Vocês foram meu suporte nos dias difíceis e conseguiram deixar tudo mais leve.

Ao meu namorado Guilherme e sua família. Sempre atenciosos e carinhosos comigo.

À FAPESP, processo 2014/25309-7 por todos os recursos disponibilizados para a realização desse projeto.

À CAPES pela bolsa de estudos durante todo o meu período de doutorado, no Brasil e no exterior, processo 88881.133033/2016-01.

À FCAV, que tem sido a minha segunda casa desde que iniciei minha graduação em 2007.

Muito obrigada!

SUMÁRIO

POLIMORFISMOS ASSOCIADOS À CARACTERÍSTICAS DE LEITE EM BUBALINOS E À MORFOLOGIA ESPERMÁTICA EM BOVINOS

RESUMO – O melhoramento genético animal utiliza ferramentas de cunho genômico e molecular a fim de auxiliar na seleção dos animais geneticamente superiores. No presente trabalho foram abordadas duas técnicas distintas para a obtenção de informações sobre características de interesse em bubalinos leiteiros e bovinos de corte. O objetivo do primeiro trabalho, realizado com búfalas da raça Murrah, foi identificar SNP nos genes *DRB2*, *DRB3*, *DMA* e *DMB* do MHC, os quais foram utilizados em estudos de associação com contagem de células somáticas, produção e qualidade do leite. Para tanto, foram utilizados DNA proveniente de 200 animais, que tiveram as regiões alvo dos genes candidatos sequenciadas. Posteriormente, foi feita a identificação dos SNP, cálculo das frequências alélicas e genotípicas, análises de variância dos efeitos dos SNP, bem como o teste do desequilíbrio de ligação. A partir dos resultados encontrados, sugerimos que os haplótipos podem ser utilizados como marcadores moleculares, uma vez que apresentaram associação com as características estudadas. O objetivo do segundo trabalho, realizado com bovinos de corte, foi identificar QTL associados a características de fertilidade de machos, como anormalidades na porção média dos espermatozoides em animais das raças Brahman (n = 1.023) e Composto Tropical (n = 1.648). Após a realização do GWAS, o SNP com maior valor significativo nos animais da raça Brahman foi identificado no gene *AKAP* (cromossomo X), já nos animais do Composto Tropical, o valor com maior associação significativa foi encontrado no SNP posicionado em uma região flanqueada pelos genes *SERINC5* e *GNPNAT1* (cromossomo 10). Os resultados indicam, portanto, que esses são potenciais SNP para auxiliar no melhoramento de características associadas à fertilidade em machos bovinos.

Palavras-chave: Espermatozoide, mastite, reprodução, SNP

POLYMORPHISMS ASSOCIATED WITH MILK TRAITS IN BUFFALOES AND SPERMATIC MORPHOLOGY IN CATTLE

ABSTRACT – Animal genetic improvement uses genomic and molecular tools to aid in the selection of genetically superior animals. In the present study two different techniques were approached to obtain information about characteristics of interest in dairy buffaloes and beef cattle. The objective of the first study with Murrah buffaloes, was to identify SNP in the MHC in *DRB2, DRB3, DMA* and *DMB* genes, which were verified its association with somatic cell count, milk production and quality. DNA from 200 animals was used, which had the target regions of the candidate genes sequenced. Subsequently, SNP were identified, allele and genotype frequencies were calculated, analyzes of variance of the effects of SNP, as well as the linkage disequilibrium test. From the results found, we suggest that the haplotypes can be used as molecular markers, once they were associated with the traits studied. The objective of the second study was to identify QTL associated with male fertility traits in beef cattle, such as mid-piece abnormalities in the spermatozoa in Brahman (n = 1,023) and Tropical Composites (n = 1,648). After GWAS, the most significant SNP in Brahman was identified in the *AKAP* gene (chromosome X), whereas in the Tropical Composites, the most significant association was found in the SNP positioned in a region flanked by *SERINC5* and *GNPNAT1* genes (chromosome 10). The results therefore indicate that these are potential SNP to assist in the improvement of characteristics associated with fertility in bovine males.

Keywords: *Spermatozoa*, mastitis, reproduction, SNP

Lista de Abreviações

AF: Allele frequency
AGL: Animal Genetics Laboratory
AKAP/ AKAP14: A-kinase-anchoring protein
BAC: Bacterial Artificial Chromosome
BoLA: Bovine Leukocyte Antigen
BuLA: Buffalo Leukocyte Antigen
CD4 +/ CD8+: Cluster of differentiation antigen 4 or 8
CNV: Copy number variation
dbSNP: Single Nucleotide Polymorphism database
DRB2: DR Beta 2
DRB3: DR Beta 3
DMA: DM Alpha
DMB: DM Beta
DNA: Deoxyribonucleic acid
EGF-HW: Expected and genotype frequencies for the Hardy-Weinberg equilibrium
FCAV: School of Agricultural and Veterinarian Sciences
F%: Fat percentage
FP: Fat production
GWAS: Genome-wide association study
GNPNAT1: Glucosamine-phosphate N-acetyltransferase 1
LD: Linkage disequilibrium
MAF: Minor allele frequency
MHC: Major Histocompatibility Complex
MOZ: Mozzarella cheese production
MP: Milk production
MPA: Mid piece abnormality
mRNA: Messenger ribonucleic acid
NK: Natural killer cell
OGF: Observed genotype frequencies
PCR: Polymerase Chain Reaction
P%: Protein percentage
PP: Protein production
PKA: Protein kinase A
QTL: Quantitative trait locus
SERINC5: Serine incorporator 5
SCC: Somatic cell count
SCCt: Somatic cell count in logarithmic scale
SNP: Single Nucleotide Polymorphism
TBE: Tris-borate-EDTA
TNF: Tumor necrosis factor
UNESP: São Paulo State University
χ^2 : Chi-square test

CAPÍTULO 1 – Considerações Gerais

O estudo de genes candidatos permite a associação de genes de ação biológica conhecida com características de interesse econômico na produção animal, sendo que esses genes estão envolvidos com o desenvolvimento ou processos fisiológicos responsáveis pela expressão dessas características. A abordagem do gene candidato pode proporcionar um entendimento mais direto da base genética das diferenças fenotípicas observadas entre os indivíduos (BRYNE; MCMULLEN, 1996; PAZ et al., 2004; LIU et al., 2008).

O sequenciamento pelo método de Sanger foi descrito em 1975, e desde então tem contribuído para um crescimento importante na área do melhoramento genético animal. Com essa técnica foi possível identificar a sequências dos nucleotídeos do DNA, o que culminou no início do mapeamento genômico dos seres vivos (CARVALHO, 2016). Porém, a rápida evolução e popularização das tecnologias da genética molecular tem permitido acessar e manipular o genoma das espécies de diferentes maneiras. A evolução tecnológica das plataformas de genotipagem com polimorfismos de nucleotídeos únicos (SNP, “Single Nucleotide Polymorphisms”) possibilitou a criação de novas abordagens para investigar os componentes genéticos subjacentes à fenótipos complexos, denominadas estudos de associação genômica ampla (GWAS, “Genome-Wide Association Studies”) (MATOS, 2013; MENEZES et al., 2013). A utilização do GWAS não requer conhecimento prévio profundo sobre os mecanismos fisiológicos da característica desejada, uma vez que polimorfismos espalhados por todo o genoma (SNP) são analisados, os quais são associados com características de interesse, como à suscetibilidade a uma doença ou a maiores ou menores níveis de produção (MENEZES et al., 2013).

Revisão de literatura

Identificação de polimorfismos em genes relacionados ao sistema imune de búfalas leiteiras por sequenciamento

O complexo principal de histocompatibilidade (MHC, “Major Histocompatibility Complex”) é caracterizado por um grande conjunto de genes intimamente ligados, que desempenham papel fundamental em processos como resistência/suscetibilidade a doenças infecciosas e resposta imune inata e adaptativa (PENN, 2002).

Em bovinos, o MHC (BoLA, “Bovine Leukocyte Antigen”) está localizado no cromossomo 23 e em bubalinos (BuLA, “Buffalo Leukocyte Antigen”), está localizado no braço curto do cromossomo 2 (STAFUZZA et al., 2013). Seus genes estão organizados em três classes distintas, agrupadas de acordo com suas funções: I, II e III (MOSAFER et al., 2012).

Os genes da classe I expressam moléculas responsáveis pela apresentação de peptídeos antígenos aos linfócitos CD8+ por meio de receptores de membrana das células T, os quais destroem as células infectadas (BEHL et al., 2012). Os genes da classe II são expressos exclusivamente nas células que contém antígenos, como macrófagos, linfócitos B e células dendríticas, e têm como função apresentar antígenos extracelulares aos linfócitos CD4+ (DE et al., 2002). As proteínas expressas pelos genes da classe II também degradam peptídeos de patógenos intracelulares nas células T-helpers, como parte de um mecanismo que identifica antígenos externos iniciando a resposta imune (BROWN et al., 1993). No complexo BoLA e BuLA, os genes da classe II de bovídeos encontram-se divididos em duas regiões separadas do cromossomo em vez de um grupamento único de genes, como é visto na maioria dos mamíferos. Essas duas partes são conhecidas como IIa e IIb, com classe IIa intimamente associada com a classe I e regiões de classe III (IANNUZZI et al., 1993; STAFUZZA et al., 2013). Já os genes da classe III do MHC codificam um grupo de proteínas com variadas funções, destacando-se interações com os receptores das células NK (“Natural Killer Cell”), além de codificar os chamados fatores de necrose tumoral (TNF, “Tumour Necrosis Factor”) os quais também estão associados à resposta imune, causando a morte de células infectadas pelo processo de apoptose (TROWSDALE, 2002; BEHL et al., 2012).

A classe IIa possui duas sub-regiões, uma composta pelos genes da família DQ (*DQA* e *DQB*) e outra pelos genes da família DR (*DRA* e *DRB*). Os genes DQ são os responsáveis por codificar moléculas que tornam as células CD4+ aptas para o reconhecimento de patógenos (NORIMINE; BROWN, 2005). Em bovinos é observada uma variedade no número de loci DQ entre os indivíduos devido a um processo de duplicação gênica, gerando um nível de complexidade adicional a essa região do sistema imune (ANDERSSON; RASK, 1988; SCOTT et al., 1991; BEHL et al., 2012).

A sub-região composta pelos genes DR é a mais amplamente estudada devido sua importância funcional, apresentando um grande número de polimorfismos identificados (DONGXIAO; YUAN, 2003). Em bovinos, o gene *DRA* possui apenas um

alelo que codifica a cadeia α da molécula DR, ou seja, trata-se de um gene monomórfico, já os genes da família DRB que codificam a cadeia β da molécula DR são altamente polimórficos (AMILLS et al., 1998). Em bovinos, existem pelo menos três genes da família DRB descritos, o *DRB1*, considerado um pseudogene; o *DRB2*, pouco expresso e o gene *DRB3*, descrito como altamente expresso e polimórfico (KUMAR et al., 2011). Em bovinos, o gene *DRB3* está relacionado com resistência/ suscetibilidade a várias doenças, como por exemplo, neospirose (SCHWAB et al., 2009), dermatofilia (MAILLARD et al., 1996) e mastite (IBRAHIM et al., 2012). Em bubalinos, diversos trabalhos na literatura vêm descrevendo polimorfismos do gene *DRB3* associados com doenças em animais de interesse econômico, como a mastite, nas raças Murrah (KUMAR et al., 2011) e Nili-Ravi (KUMAR et al., 2008).

A região da classe IIb do MHC, compreende os genes *DM, DOB, DNA, LM* e *TAP*. A sub-região composta pela família *DM*, é a mais estudada dessa classe, apresentando estudos de mapeamento em bovinos (BALLINGALL et al., 2004) e bubalinos (STAFUZZA et al., 2013). A caracterização de genes *DM* em diferentes organismos vertebrados pode fornecer informações importantes sobre a evolução do MHC, em particular no que diz respeito ao estabelecimento da função dos genes da classe IIb. Muitos estudos foram relatados com as mais variadas espécies, como humanos e macacos (ALVAREZ et al., 1998), ratos (KURTH et al., 1997) e bovinos (NIIMI et al., 1995). De maneira geral, as proteínas provenientes dos genes *DMA e DMB* estão sempre relacionadas aos processos imunes do hospedeiro.

A necessidade de melhorar os métodos de controle de doenças em animais de interesse econômico, principalmente para o desenvolvimento de novas vacinas e seleção de animais resistentes, faz com que tais genes da família *DR* e *DM* despertem o interesse dos pesquisadores (NIRANJAN et al., 2010), cujo o conhecimento das possíveis variações presentes nesses genes, poderá auxiliar nas estratégias de seleção dos programas de melhoramento genético animal.

Estudo de associação genômica ampla (GWAS) para identificar regiões de QTL associadas a características de fertilidade em bovinos de corte

A metodologia de estudos de associação genômica ampla (GWAS) tem como principal objetivo, a utilização de SNP para o mapeamento de QTL, ou seja, identificar regiões do genoma relacionadas com a expressão de características de importância

econômica ou com a regulação de alguma rota metabólica que atue diretamente no fenótipo de interesse. Os GWAS estão embasados na suposição de que uma mutação causal capaz de condicionar um fenótipo está em desequilíbrio de ligação (DL) com marcadores adjacentes em indivíduos de uma população, mesmo após diversas gerações e eventos recombinação (VAN TASSELL et al., 2008).

É interessante notar que os cromossomos sexuais são excluídos das análises na maioria dos trabalhos publicados de GWAS, uma vez que não são diploides em machos. Wise et al. (2013) levantou várias razões para não ocorrer a inclusão do cromossomo X nas análises, e estas incluem uma menor proporção de genes nesse cromossomo e uma menor cobertura do cromossomo X nas atuais plataformas de genotipagem em comparação com a cobertura autossômica. Esses autores também descreveram uma série de obstáculos técnicos, como complicações na genotipagem, imputação e na seleção de testes estatísticos. No entanto, esses cromossomos parecem ter funções bastante importantes para características reprodutivas e com a exclusão, sua influência não é levada em consideração (CAMARGO, 2015). No estudo de Fortes et al. (2012) e também abordado no capítulo 3 dessa tese, foi verificada uma grande influência do cromossomo X na fertilidade de machos bovinos. Os GWAS para características de fertilidade e puberdade em bovinos, são importantes visto que precocidade sexual ainda é um problema nos sistemas produtivos e sua melhora pode promover maior retorno econômico (CAMARGO, 2015; FORTES et al., 2012).

Uma produção pecuária eficiente requer fertilidade adequada, e os touros têm um efeito importante na disseminação de mérito genético superior (DAHLEN et al., 2014; KUES et al., 2008). Assim, indicadores de fertilidade masculina, como a susceptibilidade à fragmentação, a morfologia do esperma ou a circunferência escrotal, são utilizados no contexto clínico para avaliar a capacidade de reprodução em espécies bovinas (EVENSON, 2013; HOLROYD et al., 2002). Características de fertilidade em raças bovinas tropicais vêm sendo estudadas e dentre os fatores que podem influenciar na qualidade do sêmen, estão as anormalidades na peça intermediária do espermatozoide (FORTES et al., 2012). Essa região está localizada entre a cabeça e a cauda do espermatozoide, e é caracterizada pela presença de um grande número de mitocôndrias arranjadas de forma espiralada (bainha mitocondrial), sendo a principal responsável pela propulsão do espermatozoide (BEDFORD; HOSKINS, 1990). Os defeitos da peça intermediária incluem o pescoço "curvado" (onde o pescoço e a cauda formam um ângulo maior que 90% para o eixo longo da

cabeça), a inserção assimétrica da parte intermediária na cabeça, peça intermediária grossa ou irregular, peça intermediária anormalmente fina (sem bainha mitocondrial), bem como qualquer combinação destes defeitos (SUN et al., 2006).

Problemas na peça intermediária do espermatozoide são bastante comuns e influenciam diretamente a motilidade deste (FORDYCE et al., 2006). Com a melhoria de características relacionadas à qualidade seminal, podem ser elevados os índices de concepção na inseminação artificial por exemplo, fator de alto retorno econômico na bovinocultura de corte (SAMARAJEEWA et al., 2012; WOLFOVA et al. 2010). Portanto, a utilização de GWAS para características reprodutivas em machos bovinos se faz importante para identificar regiões de QTL que influenciam diretamente características economicamente importantes e que poderão ser incorporados na avaliação genética desses animais.

Objetivos

Identificar polimorfismos nos genes *DRB2, DRB3, DMA* e *DMB* por meio de sequenciamento e realizar análise de associação entre estes polimorfismos e características de qualidade e produção de leite em búfalas da raça Murrah.

Identificar QTL associados à característica de fertilidade em bovinos das raças Brahman e Composto Tropical por meio de GWAS.

Referências bibliográficas

ALVAREZ, M.; RECIO, M. J.; MARTINEZ-LASO, J.; PÉREZ-BLAS, M.; GARCIA-DE-LA-TORRE, C.; VARGAS-ALARCON, G.; ALEGRE, R.; GOMEZ-CASADO, E.; ARNAIZ-VILLENA, A. Allelic diversity at the primate MHC-DMB locus: Presence of a conserved tyrosine inhibitory motif in the cytoplasmictail. **Tissue Antigens**, v. 51, p. 174–182, 1998.

AMILLS, M.; RAMIYA, V.; NORIMINE, J.; LEWIN, H. A. The major histocompatibility complex of ruminants. **Revue Scientifique Technique**, v. 17, p. 108-120, 1998.

ANDERSSON, L.; RASK, L. Characterization of the MHC class II region in cattle. The number of DQ genes varies between haplotypes. **Immunogenetics**, v. 27, p. 110-120,1988.

BALLINGALL, K.T.; ELLIS, S.A.; MACHUGH, N.D.; ARCHIBALD, S.D.; MCKEEVER, D.J. The DY genes of the cattle MHC: expression and comparative analysis of an unusual class II MHC gene pair. **Immunogenetics**, v. 55, p. 748-755, 2004.

BEDFORD, J. M.; HOSKINS, D. D. The mammalian spermatozoon: Morphology, biochemistry and physiology In: Marshall's Physiology of Reproduction. **Male Reproduction Lamming**, G.E. (Ed), London: Churchil Livingstone,1990. v. 2, p. 379-568.

BEHL, J.D.; VERMA, N.K.; TYAGI, N.; MISHRA, P.; BEHL, R.; JOSHI, B.L. The Major Histocompatibility Complex in Bovines: A Review. **ISRN Veterinary Science**, v. 2012, p. 1-12, 2012.

BROWN, J. H.; JARDETZKY, T. S.; GORGA, J. C.; STERN, L. J.; URBAN, R.G.; STROMINGER, J. L.; WILEY, D. C. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. **Nature**, v. 364, p. 33-39,1993.

BRYNE, P. F.; MCMULLEN, M. D. Defining genes for agricultural traits: QTL analysis and the candidate gene approach. **Probe**, v. 7, p. 24-27, 1996.

CAMARGO, G.M.F. **Mutações putativo-causais em genes candidatos associadas à fertilidade de bovinos de corte e bubalinos**. Tese (doutorado). 2015. Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, "Júlio de Mesquita Filho", Jaboticabal, 2015.

CARVALHO, R. A. **Avaliação do método de sequenciamento de nova geração no diagnóstico genético de neoplasia endócrina múltipla tipo 1.** Dissertação (Mestrado). 2016. Faculdade de Medicina da Universidade de São Paulo, São Paulo, 2016.

DAHLEN, C.; LARSON, J.; LAMB, G. Cliff. Impacts of reproductive technologies on beef production in the United States. In: **Current and Future Reproductive Technologies and World Food Production**. Springer New York, 2014. p. 97-114.

DONGXIAO, S.; YUAN, Z. Polymorphisms of the Second Exon of MHC-DRB gene in Chinese local sheep and goat. **Biochemical Genetics**, v. 42, p. 385-390, 2004.

EVENSON, D. P. Sperm chromatin structure assay (SCSA®). **Spermatogenesis: Methods and Protocols**, p. 147-164, 2013.

FORDYCE, G.; ENTWISTLE, K.; NORMAN, S.; PERRY, V.; GARDINER, B.; FORDYCE, P. Standardising bull breeding soundness evaluations and reporting in Australia. **Theriogenology**, v. 66, n. 5, p. 1140-1148, 2006.

FORTES, M.R.S. LEHNERT, S. A., BOLORMAA, S., REICH, C., FORDYCE, G., CORBET N. J., WHAN V.; HAWKEN, R. J.; REVERTER, A. Finding genes for economically important traits: Brahman cattle puberty. **Animal Production Science**, v. 52, p.143-150, 2012.

HOLROYD, R. G.; DOOGAN, V. J.; FAVERI, J. D.; FORDYCE, G.; MCGOWAN, M. R.; BERTRAM, J. D.; VANKAN, D. M.; FITZPATRICK, L. A.; JAYAWARDHANA, G. A.; MILLER, R. G. Bull selection and use in northern Australia: 4. Calf output and predictors of fertility of bulls in multiple-sire herds. **Animal Reproduction Science**, v. 71, p. 67-79, 2002.

IANNUZZI, L.; GALLAGHER, D. S.; WOMACK, J. E.; DI MEO, G. P.; SKOW, L. C.; FERRARA, L. Chromosomal localization of the major histocompatibility complex in cattle and river buffalo by fluorescent in situ hybridization. **Hereditas**, v. 118, p. 187-190, 1993.

IBRAHIM, E. A.; ALLAM, N. A. T.; KOTB, E. E. Z.; EL-RAFEY, G. A.; EL-DEEN, M. M. A.; FADLALLAH, M. G. Sequence-based typing-study on the relationship between subclinical mastitis and BoLA-DRB3.2 allelic polymorphism in Egyptian cows. **Global Veterinaria**, v. 9, p. 8-22, 2012.

KUES, W. A.; RATH, D.; NIEMANN, H. Reproductive biotechnology in farm animals goes genomics. **CAB Reviews**, v. 3, n. 36, p. 1-18, 2008.

KUMAR, S.; SANGWAN, M.L.; RUPENDER. Polymorphism in DRB3 exon 2 by PCR-RFLP and its association with mastitis in Nili-Ravi breed. **Indian Journal of Biotechnology**, v. 7, p. 398-400, 2008.

KUMAR, S.; SANGWAN, M.L.; AHLAWAT, S.; BARWAR, A. Polymorphism in DRB3 exon by PCR-RFLP and its association with mastitis in Murrah buffaloes. **Indian Journal of Biotechnology**, v. 10, p. 232-234, 2011.

KURTH, D.; NEUMANN, J.; DEMLEITNER, K.; HILDMANN, J.; MEHLIG, M.; SCHEICHER, C.; WENDLING, U.; SUN, D.; RESKE, K. Full length cDNA of rat RT1.DMa and RT1.DMb and expression of RT1.DM genes in dendritic and Langerhans cells. **The Journal of Biological Chemistry**, v. 378, p.1005–1012, 1997.

LIU, M.; PENG, J.; XU, D.Q.; ZHENG, R.; LI, F.; LI, J.L.; ZUO, B.; LEI, M.G.; XIONG, Y.Z.; DENG, C.Y.; JIANG, S.W. Association of MYF5 and MYOD1 gene polymorphisms and meat quality traits in Large White X Meishan F2 pig populations. **Biochemical Genetics**, v. 46, p. 720–732, 2008.

MAILLARD, J. C.; MARTINEZ, D.; BENSAID, A. An amino acid sequence coded by the exon 2 of the BoLA DRB3 gene associated with a BoLA class I specificity constitutes a likely genetic marker of resistance to dermatophilosis in Brahman zebu cattle of Martinique (FWI). **Annals of the New York Academy of Sciences**, v. 791, p. 185-197, 1996.

MATOS, M.C. **Associação genômica ampla para características reprodutivas em bovinos da raça nelore**. Dissertação (mestrado). 2013. Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, "Júlio de Mesquita Filho", Jaboticabal, 2013.

MENEZES, G.R.O., REGITANO, L.C.A., SILVA, M.V.G.B., CARDOSO, F.F., SILVA, L.O.C., SIQUEIRA, F., EGITO, A.A. Genômica aplicada ao melhoramento genético de gado de corte. In: Melhoramento genético aplicado em gado de corte. Brasília: EMBRAPA, 2013. cap. 17, 213-226.

MOSAFER, J.; HEYDARPOUR, M.; MANSHAD, E.; RUSSELL, G.; SULIMOVA, G.E. Distribution of BoLA-DRB3 allelic frequencies and identification of two new alleles in Iranian buffalo breed. **The Scientific World Journal**, v. 2012, 2012.

NIIMI, M.; NAKAI, Y.; AIDA, Y. Nucleotide sequences and the molecular evolution of the DMA and DMB genes of the bovine major histocompatibility complex. **Biochemical and Biophysical Research Communications**, v. 217, p. 522-528, 1995.

NIRANJAN, S. K.; DEB, S. M.; KUMAR, S.; MITRA, A.; SHARMA, A.; SAKARAM, D.; NASKAR, S.; SHARMA, D.; SHARMA, S. R. Allelic diversity at MHC class II DQ loci in buffalo (Bubalus bubalis): Evidence for duplication. **Veterinary Immunology and Immunopathology**, v. 138, p. 206-212, 2010.

NORIMINE, J.; BROWN, W. C. Intrahaplotype and interhaplotype pairing of bovine leukocyte antigen DQA and DQB molecules generate functional DQ molecules important for priming CD4(+) Tlymphocyte responses. **Immunogenetics**, v. 57, p. 750-762, 2005.

PAZ, C.C.P.; PACKER, I.U.; FREITAS, A.R; TAMBASCO-TALHARIS, D.; REGITANO, L.C.A., ALENCAR, M.M.; RODRIGUES, A.A. Influência de polimorfismos genéticos

sobre os parâmetros da curva de crescimento em bovinos de corte. **Revista Brasileira de Zootecnia**, v. 33, p. 858-69, 2004.

PENN, D.J. Major Histocompatibility Complex (MHC). **Encyclopedia of Life Sciences**, p. 1-7, 2002.

SAMARAJEEWA, S.; HAILU, G.; JEFFREY, S. R.; BREDAHL, M. Analysis of production efficiency of beef cow/calf farms in Alberta. **Applied Economics**, v. 44, p. 313-322, 2012.

SCHWAB, A. E.; GEARY, T. G.; BAILLARGEON, P.; SCHWAB, A. J.; FECTEAU, G. Association of BoLA DRB3 and DQA1 alleles with susceptibility to Neospora caninum and reproductive outcome in Quebec Holstein cattle. **Veterinary Parasitology**, v. 165, p. 136-140, 2009.

SCOTT, P. C.; GOGOLIN-EWENS, K. J.; ADAMS, T. E.; BRANDON, M. R. Nucleotide sequence, polymorphism, and evolution of ovine MHC class II DQA genes**.** **Immunogenetics**, v. 24, p. 69-79, 1991.

STAFUZZA, N.B.; GRECO, A.J.; GRANT, J.R.; ABBEY, C.A.; GILL, C.A.; RAUDSEPP, T.; SKOW, L.C.; WOMACK, J.E.; RIGGS, P.K.; AMARAL, M.E.J. A high-resolution radiation hybrid map of the river buffalo major histocompatibility complex and comparison with BoLA. **Animal Genetics**, v. 44, p. 369-376, 2013.

SUN, F.; KO, E.; MARTIN, R. H. Is there a relationship between sperm chromosome abnormalities and sperm morphology?. **Reproductive Biology and Endocrinology**, v. 4, n. 1, p. 1, 2006.

TROWSDALE, J. The gentle art of gene arrangement: the meaning of gene clusters. **Genome Biology**, v. 3, p. 1-5, 2002.

VAN TASSELL, C. P.; SMITH, T. P.; MATUKUMALLI, L. K.; TAYLOR, J. F.; SCHNABEL, R. D.; LAWLEY, C. T. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. **Nature Methods**, v. 5, n. 3, p. 247-252, 2008.

WISE, A. L.; GYI, L.; MANOLIO T. A. 2013. eXclusion: toward integrating the X chromosome in genome-wide association analyses. **The American Journal of Human Genetics**, v. 92, n. 5, p. 643-647, 2013.

WOLFOVÁ, M.; WOLF, J.; ZAHRÁDKOVÁ, R.; PŘIBYL, J.; DAŇO, J.; KRUPA, E.; KICA, J. Breeding objectives for beef cattle used in different production systems - 2. Model application to production systems with the Charolais breed. **Livestock Production Science**, v.95, p.217-230, 2005

CHAPTER 2 - Polymorphisms in major histocompatibility complex genes and its associations with milk quality in Murrah buffaloes[a]

Ana Cláudia de Freitas[1], Nedenia Bonvino Stafuzza[2], Marina Mortati Dias Barbero[1], Daniel Jordan de Abreu Santos[1], Marina Rufino Salinas Fortes[3] and Humberto Tonhati[1]

[1]Department of Animal Science, School of Agricultural and Veterinarian Sciences (FCAV), São Paulo State University (UNESP), (anafreitas.unesp@gmail.com.br, daniel_jordan2008@hotmail.com, tonhati@fcav.unesp.br)
[2]Department of Exact Sciences, School of Agricultural and Veterinarian Sciences (FCAV), São Paulo State University (UNESP), (nedeniabs@gmail.com)
[3]School of Chemistry and Molecular Biosciences, The University of Queensland, Saint Lucia Campus, Australia, (m.fortes@uq.edu.au)

*Corresponding author: anafreitas.unesp@gmail.com.br
Address: Department of Animal Science, São Paulo State University (FCAV/UNESP), Via de Acesso Prof. Paulo Donato Castellane s/n, 14884-900 Jaboticabal, São Paulo, Brazil
Phone: +55 16 3209-7436

Abstract

Background: Genetic improvement programs have used molecular genetic tools as an auxiliary method to select animals with superior genetic merit for traits such as milk production and milk quality as well as disease resistance. The current work investigated SNP and haplotypes in the *DRB2, DRB3, DMA* and *DMB* genes of the major histocompatibility complex (MHC) in buffaloes as potential tools for genetic selection. Two hundred DNA samples from Murrah buffaloes were used in this study. The target regions of candidate genes were amplified by PCR followed by sequencing and identification of polymorphisms. Allele and genotype frequencies as well as linkage disequilibrium between SNP were calculated. Genotypes were used in association analyses with milk production and quality traits.

Results: Except for the *DMA* gene which was identified as monomorphic, the other genes presented several polymorphisms. The *DMB*, *DRB2*, and *DRB3* genes presented two, six and seven SNP, respectively. A total of 57 haplotype blocks were

[a] Article submitted to "Journal of Animal Science and Biotechnology"

constructed from 15 SNP identified, which was used in association analyses. All the studied traits had at least one associated haplotype.

Conclusion: It is suggested that the haplotypes found herein can be associated with milk, fat, protein and mozzarella production, fat percentage and protein and somatic cell counts.

Keywords: *Bubalus bubalis*, milk fat, milk protein, somatic cell coun

Background

Economic losses caused by alterations in milk production and composition, often caused by mastitis, make this disease the most harmful and prevalent illness in dairy cows (USDA, 2007). Mastitis is defined as an inflammation of the mammary glands caused by metabolic and physiologic alterations, traumas and microorganism invasions (Zhao & Lacasse, 2008). Mastitis has been categorized as clinical or subclinical, in reference to the disease intensity and evolution of the clinical presentation. Clinical mastitis is characterized by evident clinical signs with systemic illness, gland swelling and abnormal milk. Subclinical mastitis is the most prevalent form of the disease and responsible for most of the associated economic loss. Cows with subclinical mastitis produce apparently normal milk with increase in somatic cell count (SCC) as previously described (El Nahas et al., 2017).

A reference genome is still not available for *Bubalus bubalis*, an important species for dairy production. To develop molecular genetic tools, studies focusing on major candidate genes are important. Currently candidate gene studies are the basis for the construction of buffalo-based genetic mapping (El Nahas et al., 2017). Genes of the major histocompatibility complex (MHC) have been proposed as candidate

genes associated with mastitis and milk production traits because they are known to be contribute to the innate and adaptive immune responses (Penn & Ilmonen, 2005).

MHC class II genes can be considered important candidates genes since they were associated with diseases and are extremely polymorphic in vertebrates (Trowsdale, 1995). MHC class II have been studied in bovines and buffaloes and presented results highlighting how genetic variation can influence host resistance to infection or quality and milk production. SNP in *DRB3* gene presented association with milk production traits in dairy cattle (Sermyagin et al., 2016), and could be markers for susceptibility/resistance to mastitis in dairy cows (Ibrahim et al., 2012).

Therefore, the identification of polymorphisms in MHC genes in buffalo, such as *DRB2, DRB3*, *DMA* and *DMB* genes, and its association with SCC, milk yield and quality traits may provide relevant information on the contribution of these genes to the studied traits. Allelic and genotypic variability for the tested population might be related to mastitis resistance, as indicated by SCC. If identified, associated polymorphisms could be useful as markers for these economically important traits related to milk production and quality.

Methods

Samples

The Murrah buffaloes used in this research belong to the Tapuio Farm, located in Taipu, RN, Brazil. The herd has approximately 1,000 heads, of which 468 are dams, with milk production of around 3,500 liters per day. Milk recording and sampling were performed periodically to determine fat and protein contents as well as SCC. This farm implements the dairy buffalo milk-recording program supplied by the Department of Animal Science of School of Agricultural and Veterinarian Sciences (FCAV), São Paulo State University (UNESP), Brazil.

Hair follicle samples from 200 Murrah buffaloes were collected and stored at 4°C until DNA extraction. Genomic DNA was extracted by the phenol-chloroform-isoamyl alcohol method (Sambrook et al., 1989). The DNA samples were quantified using a NanoDrop® 1000 spectrophotometer (Thermo Scientific, USA) and its concentrations were adjusted to 70 ng/μL.

Primer designer and Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction (PCR) primers were used to amplify DNA segments of *DMA, DMB, DRB2* and *DRB3* genes. Primers were designed using the Primer3Plus program (Untergasser et al., 2007), and their quality was verified using the OligoAnalyzer 3.1 tool (http://www.idtdna.com/calc/analyzer). Primers were design using as a reference the gene sequences from the *Bos taurus* UMD_3.1 reference genome assembly: *DMB* (AC_000180.1, 7,186,801-7,193,645 bp), *DRB2* (AC_000180.1, 25,544,945-25,555,504 bp) and *DRB3* gene (AC_000180.1, 25,472,161-25,476,885 bp). For the *DMA* gene, were used a genomic sequence obtained from the sequencing and molecular characterization of one clone of a buffalo BAC (Bacterial Artificial Chromosome) library reported previously (Stafuzza et al., 2012). The details of each primer pair are described in Table 1.

Table 1. Primers designed to amplify by specific regions of the *DMA, DMB, DRB2* and *DRB3* genes, annealing region, annealing temperature (T ºC) and fragment size.

Gene	Primers (5'→3')	T (ºC)	Size (bp)	Region
DMA	F: ATCTGAGCTTCCTTCCCTCA R: GCCACTGTGTCTGGCTCTCC	63	700	exon 2, intron 2, exon 3 and intron 3
DMB	F: GATGGAGTGGGAGACAGGAA R: GCAAGTGGAAGTGAAGAATC	60	579	intron 3 and exon 4
DRB3	F: TGAAAGGCAGCTAACCAAGG R: CCCGGGGGGTGTAGGTATTT	63	581	exon2 and intron 2
DRB2	F: TGACTCACTTTGTCTTGCTCACTG R: GAGGATTTCTATCACAGGATAAAGG	60	500	intron 2, exon 3 and intron 3

The genomic DNA samples were amplified by PCR in a final volume reaction of 50 µL containing 70 ng DNA, 5.5 µL of each primer (forward and reverse), and 26 µL GoTaq® Colorless Master Mix (Promega, USA). The annealing temperature was selected for a PCR performed in a T100™ Thermal Cycler (Bio-Rad, USA) according to the following steps: initial denaturation at 94 °C for 2 min, followed by 35 cycles at 94 °C for 30 s (denaturation), 60 to 63 ºC for 30 s (annealing) depending of each primer requirement (Table 1), 72 °C for 30 s (extension), followed by a final extension at 72 °C for 6 min.

In order to check the amplification by PCR of the desired fragment, a total of 2 µL of each PCR product was electrophoresed on 1.5% agarose gels stained with GelRed™ (Biotium, USA) in 1X TBE (Tris-borate-EDTA) buffer at 90 V, for 50 min. The agarose gels were visualized in ultraviolet light, the gel photographed by the Gel-Doc ™ apparatus (BioRad, USA), and analyzed with the Image Analysis software (Kodak, USA).

Sequencing and Single Nucleotide Polymorphism detection

The PCR products were purified with the Wizard® SV Gel and PCR Clean-Up System kit (Promega, USA) as described by the manufacturer. The samples were sequenced using the ABI PRISM® BigDye® Terminator v3.0 Cycle Sequencing Ready Reaction Kit (Applied Biosystems, USA) in an ABI 3730xl automatic sequencer (Applied Biosystems, USA). Each sequencing reaction was carried out using 2 µL Save Money buffer (200 mM Tris, pH 9.0, 0.5 mM $MgCl_2$), 1 µL primer (2.5 pmol / µL), 2 µL BigDye and 70 ng DNA. Sequencing reaction comprised denaturation at 95 °C for 20 s, annealing at 60 to 63°C for 15 s and extension at 60 °C for 1 min, in a total of 30 cycles.

DNA sequences obtained from *DMA, DMB, DRB2* and *DRB3* genes were analyzed using CodonCode Aligner software (http://www.codoncode.com/aligner) to detect polymorphisms. All SNP identified in this study were submitted to dbSNP (Single Nucleotide Polymorphism database) (https://www.ncbi.nlm.nih.gov/snp/).

Statistical Analysis

Allelic and Genotypic Frequency

Allelic frequencies of the genotyped SNP were determined by simple allele counting (Falconer and Mackay, 1996) using the PROC FREQ in SAS (SAS 9.2, SAS Institute, Cary, NC, USA). Possible deviations from the genotypic and allelic frequencies were tested using the chi-square test to verify whether the genotyped population was in Hardy-Weinberg equilibrium.

Linkage Disequilibrium

Linkage disequilibrium (r^2) was estimated using the Plink v1.07 software (Purcell et al., 2007) to check which SNP are segregating together. Considering two loci with two alleles for each locus (A1/A2 and B1/B2), we have the following equation (Hill & Robertson, 1966):

$$r^2 = \frac{D^2}{[freq(A_1) * freq(A_2) * freq(B_1) * freq(B_2)]}$$

where (Hill, 1981):

$$D = freq(A_1\_B_1) * freq(A_2\_B_2) - freq(A_1\_B_2) * freq(A_2\_B_1)$$

Haplotype Construction

Haplotypes were constructed using the Haploview v4.1 statistical software (Barrett et al., 2005). Linkage disequilibrium analyses (LD) was established based on r^2 parameter (Hill and Robertson, 1968). The confidence intervals method described

by Gabriel et al. (2002) was used to define haplotype blocks and haplotypes were estimated using an expectation maximization algorithm based on the maximum likelihood. SNP that were not part of any haplotype were considered as individual markers.

Analyses of genetic associations between haplotypes and traits were estimated using the statistical program SAS/STAT 9.3 (SAS Institute, Inc., Cary, NC, USA).

Association Analysis

The traits utilized in this association study refer to first lactation records obtained between 2010 and 2016 for milk production up to 270 days of lactation (MP), fat production (FP), protein production (PP), *mozzarella* cheese production (MOZ), fat percentage (F%), protein percentage (P%) and somatic cells count (SCC).

The relationship of each haplotype with the production traits, was stablished using the following statistical model:

$$Y_{ijklm} = \mu + GC_i + S_j + A_k + M_l + \beta_1\left(I_{ijkl} - \bar{I}\right) + \beta_2\left(I_{ijkl} - \bar{I}\right)^2 + e_{ijklm}$$

where $\mathrm{Y_{ijkl}}$ is the observed value of the trait in the lactation *m* of the animal *k*, daughter of sire *j*, in contemporary group *i*; *μ* is overall average; $\mathrm{CG_i}$ is effect of contemporary group; $\mathrm{S_j}$ is random effect of sire, with mean 0 and variance σ_s^2; $\mathrm{A_k}$ is random effect of the animal, with mean 0 and variance σ_A^2; $\mathrm{M_l}$ is fixed effect of genotype for SNP l or haplotype 1; β_1 is linear regression coefficient of the observation $\mathrm{Y_{ijklm}}$ in relation of cow age at measurement; β_2 is quadratic regression coefficient of observation $\mathrm{Y_{ijklm}}$ in relation of cow age at measurement; $\mathrm{I_{ijkl}}$ is cow age evaluated at measurement; $\bar{\mathrm{I}}$ is cows average age evaluated at measurement; and $\mathrm{e_{ijklm}}$ is random residual effect associated with trait $\mathrm{Y_{ijkl}}$ with mean 0 and variance σ_e^2. The contemporary group was

considered the concatenation of year and calving season. Estimates of variance were performed using the PROC MIXED (SAS 9.2, SAS Institute, Cary, NC, USA).

The distribution of SCC was not normal so, therefore, the data was transformed to a logarithmic scale (SCCt), using the following function described by Dabdoub and Shook (1984):

$$SCCt = \left(log_2 \left(\frac{SCC}{100{,}000} \right) \right) + 3$$

The mean MP, FP, PP, F%, P% and SCC for the different genotypes were analyzed using the Tukey test ($P<0.05$) and, after Bonferroni correction was applied at 5% significance for all traits.

The *mozzarella* production was estimated according to Altiero et al. (1989):

$$\mathrm{MOZ(kg)} = \mathrm{MP(Kg)} \times \left[\frac{3{,}5(\mathrm{P\%}) + 1{,}23(\mathrm{F\%}) - 0{,}88}{100} \right]$$

where: MOZ is *mozzarella* production; MP is milk production; P% is protein percentage and F% is fat percentage.

Results

Phenotype variation in the studied population was observed. Descriptive statistics such as mean, minimum (Min) and maximum (Max), for the seven traits measured are provided in Table 2.

Table 2. Descriptive statistics of the production traits.

Trait	Mean	Min	Max
MP (Kg)	2352.29	2127.48	2577.10
FP (Kg)	160.53	145.00	176.06
PP (Kg)	126.10	98.17	154.03
MOZ (Kg)	545.44	506.46	584.43
F%	6.33	6.14	6.53
P%	4.12	3.98	4.27

Trait	Mean	Min	Max
SCC (x 1000 cel/mL)	48.93	5.94	91.93

*MP= Milk production, FP= Fat production, PP= Protein production, MOZ= *Mozzarella* production, F%=Fat percentage, P%= Protein percentage and SCC= Somatic cells count

Polymorphisms

The sequence alignments revealed a total of 15 polymorphisms, of which two were described in *DMB*, six in *DRB2* and seven in *DRB3*. No polymorphism was identified in *DMA*. Polymorphism descriptions, positions and possible amino acid change were described in the Table 3. Most polymorphisms detected in this study result in non-synonymous SNP that lead to amino acid changes in the coded proteins.

Table 3. Gene symbol, position, intron or exon region, SNP and amino acid change in the specific regions of candidate genes.

Gene	SNP_ID	Position	Mutation	Region	Consequence	Amino acid	Polarity
DMB	ss2137539310	414121	C/T	Exon 4	Synonymous	Leu / Leu	-
	ss3029150226	414134	G/C	Exon 4	Synonymous	Val / Val	-
DRB2	ss3029150227	20400	A/G	Intron 3	-	-	-
	ss3029150228	20458	G/T	Exon 3	Non-synonymous	Ala / Ser	Apolar/Polar
	ss3029150229	20569	G/T	Exon 3	Non-synonymous	Gly / Cys	Polar/Polar
	ss3029150230	20570	A/C	Exon 3	Non-synonymous	Gly / Ser	Polar/Polar
	ss3029150231	20572	G/A	Exon 3	Non-synonymous	Gly / Ser	Polar/Polar
	ss3029150232	20693	C/T	Exon 3	Synonymous	Tyr / Tyr	-
DRB3	ss3029150233	43895	G/A	Exon 2	Non-synonymous	Trp / STP	Apolar/STP
	ss3029150234	43914	G/A	Exon 2	Non-synonymous	Gly / Arg	Polar/Polar
	ss3029150235	43925	G/T	Exon 2	Non-synonymous	Ser / Ile	Polar/Apolar
	ss3029150236	44046	A/G	Exon 2	Non-synonymous	Met / Val	Apolar/Apolar
	ss3029150237	44092	G/A	Exon 2	Non-synonymous	Thr / Ile	Polar/Apolar
	ss3029150238	44093	C/T	Exon 2	Non-synonymous	Thr / Ile	Apolar/Polar
	ss3029150239	44160	T/C	Exon 2	Non-synonymous	STP / Gln	STP/Polar

Most allele frequencies have values greater than 0.10 which contributes to the association analyses, since they will not be affected by sharp discrepancy in frequency

distribution. The chi-square test result was 3.84 ($P < 0.05$), revealing that all the SNP analyzed had expected frequencies (Table 4). Therefore, the population is in Hardy-Weinberg equilibrium.

Table 4. SNP description according the animals genotyped: allele frequency (AF), observed genotype frequencies (OGF) expected and genotype frequencies for the Hardy-Weinberg equilibrium (EGF-HW), and observed chi-square test (χ^2) for all the SNP identified.

Gene	SNP	AF	OGF	(EGF-HW)	χ^2*
DMB	ss3029150226	G = 0.389 C = 0.611	GG = 0 GC = 0.777 CC = 0.223	GG = 0.151 GC = 0.475 CC = 0.374	0.4037
	ss2137539310	T = 0.328 C = 0.672	TT = 0 CT = 0.655 CC = 0.345	TT = 0.107 CT = 0.441 CC = 0.452	0.2376
DRB2	ss3029150232	C = 0.425 T = 0.575	CC = 0.151 CT = 0.548 TT = 0.301	CC = 0.180 CT = 0.489 TT = 0.331	0.0147
	ss3029150231	G = 0.137 A = 0.863	GG = 0.034 GA = 0.206 AA = 0.760	GG = 0.019 GA = 0.236 AA = 0.745	0.0171
	ss3029150230	A = 0.572 C = 0.428	CC = 0.151 AC = 0.555 AA = 0.294	CC = 0.183 AC = 0.490 AA = 0.327	0.0177
	ss3029150229	G = 0.606 T = 0.394	TT = 0.151 GT = 0.486 GG = 0.363	TT = 0.155 GT = 0.478 GG = 0.367	0.0003
	ss3029150228	G = 0.164 T = 0.836	GG = 0.027 GT = 0.274 TT = 0.699	GG = 0.027 GT = 0.275 TT = 0.698	7.465×10^{-6}
	ss3029150227	A = 0.428 G = 0.572	AA = 0.096 AG = 0.664 GG = 0.240	AA = 0.183 AG = 0.490 GG = 0.327	0.1273
DRB3	ss3029150239	T = 0.833 C = 0.167	CC = 0.141 TC = 0.051 TT = 0.808	CC = 0.028 TC = 0.278 TT = 0.694	0.6648
	ss3029150238	C = 0.340 T = 0.660	CC = 0.276 CT = 0.128	CC = 0.115 CT = 0.449	0.5101

		TT = 0.596	TT = 0.436	
ss3029150237	G = 0.859 A = 0.141	AA = 0.064 GA = 0.154 GG = 0.782	AA = 0.020 GA = 0.242 GG = 0.738	0.1332
ss3029150236	A = 0.202 G = 0.798	AA = 0.147 AG = 0.109 GG = 0.744	AA = 0.041 AG = 0.322 GG = 0.637	0.4381
ss3029150235	G = 0.599 T = 0.401	TT = 0.353 GT = 0.096 GG = 0.551	TT = 0.161 GT = 0.480 GG = 0.359	0.6396
ss3029150234	G = 0.833 A = 0.167	AA = 0.096 GA = 0.141 GG = 0.763	AA = 0.028 GA = 0.278 GG = 0.694	0.2423
ss3029150233	G = 0.125 A = 0.875	GG = 0.051 GA = 0.148 AA = 0.801	GG = 0.015 GA = 0.219 AA = 0.766	0.1062

* χ^2: observed chi-square, values above 3.84 are significant ($P \leq 0.05$)

Linkage disequilibrium

Linkage disequilibrium (LD) occurs when r^2 values are greater than 0.33 (Ardlie et al., 2002). Estimated LD between SNP in this population ranged from 0.00000292 to 0.8901 (Table 5). As expected, SNP within the same gene were more likely to be inherited together, presenting an LD higher than 0.33.

Table 5. Linkage disequilibrium between the 15 identified SNP in the *DMB*, *DRB2* and *DRB3* genes.

Gene and SNP	***DMB* ss2137539310**	***DRB2* ss3029150232**	***DRB2* ss3029150231**	***DRB2* ss3029150230**	***DRB2* ss3029150229**	***DRB2* ss3029150228**	***DRB2* ss3029150227**	***DRB3* ss3029150239**	***DRB3* ss3029150238**	***DRB3* ss3029150237**	***DRB3* ss3029150236**	***DRB3* ss3029150235**	***DRB3* ss3029150234**	***DRB3* ss3029150233**
***DMB* ss3029150226**	0.4495	0.0018	0.0004	0.0004	0.0023	1.22E-05	0.003	0.0003	0.0013	0.0136	0.0007	0.0002	0.0001	0.0012
***DMB* ss2137539310**		1,00E-04	0.001	0.0009	0.0004	0.002	0.0022	0.0001	0.0015	0.0006	0.0002	0.0015	0.0006	0.0018
***DRB2* ss3029150232**			0.0688	0.891	0.877	0.0752	0.6975	0.0067	0.0089	0.0603	0.072	0.0283	0.0168	0.0002
***DRB2* ss3029150231**				0.0839	0.0683	0.5674	0.0705	0.0531	0.0483	0.0007	0.015	0.0786	0.0002	0.0232
***DRB2* ss3029150230**					0.8082	0.0797	0.6482	0.0077	0.0075	0.0585	0.0696	0.0187	0.0156	8.58E-05
***DRB2* ss3029150229**						0.0908	0.6307	0.0021	0.0125	0.0559	0.0741	0.0369	0.01	0.0024
***DRB2* ss3029150228**							0.0927	0.0386	0.0676	0.0196	0.0369	0.0598	0.0109	0.0144
***DRB2* ss3029150227**								2.92E-06	0.0212	0.0351	0.0453	0.0397	0.0139	0.0028
DRB3* ss*3029150239									0.4851	0.0079	0.0075	0.3663	0.0252	0.0003
DRB3* ss*3029150238										0.2106	0.2414	0.6314	0.1927	0.0244
***DRB3* ss3029150237**											0.357	0.1359	0.3957	0.0107
***DRB3* ss3029150236**												0.3927	0.3916	0.3103
***DRB3* ss3029150235**													0.1598	0.1758
***DRB3* ss3029150234**														0.0167

Haplotypes

All the 15 SNP identified in this study were utilized to construct a total of 57 haplotypes, termed H1 to H57. Thirteen haplotypes presented 12 significant associations with the traits studied (Table 6). The H1 and H2 haplotypes contained the SNP of *DMB* gene. The H3 and H4 haplotypes contained the SNP of *DMB* and *DRB3* genes. The H5 and H13 haplotypes contained the SNP of gene *DRB2* and the H6, H7, H8, H9, H10, H11 and H12 haplotypes contained the SNP of gene *DRB3*.

Haplotypes were tested for association with the studied traits. Only H6 haplotype presented significant association with MP e FP traits. For PP and MOZ, we observed association with H2, H6 and H8 haplotypes. For F%, we identified association with the H10 and H12 haplotypes. The phenotype P% was associated with H3, H4, H5, H7 and H9 haplotypes. For SCC, haplotypes H1, H8, H10, H11 and H13 were significant.

Table 6. Haplotype association with important economic traits.

Haplotypes		Gene	Traits						
			MP	**FP**	**PP**	**MOZ**	**F%**	**P%**	**SCC**
H1	GT	*DMB*	P=0.8568	P=0.4265	P=0.5716	P=0.4637	P=0.9623	P=0.2513	P=0.0268*
H2	GC	*DMB*	P=0.1766	P=0.0892	P=0.0231*	P=0.0198*	P=0.7537	P=0.9632	P=0.3968
H3	TC	*DMB, DRB3*	P=0.1532	P=0.3105	P=0.4683	P=0.4761	P=0.5003	P=0.0239*	P=0.5127
H4	TT	*DMB, DRB3*	P=0.0866	P=0.1178	P=0.1779	P=0.1310	P=0.8128	P=0.0169*	P=0.5464
H5	AT	*DRB3*	P=0.3895	P=0.1085	P=0.9106	P=0.4268	P=0.3574	P=0.0233*	P=0.6267
H6	TAAGGA	*DRB2*	P=0.0079*	P=0.0280*	P=0.0092*	P=0.0102*	P=0.7931	P=0.1677	P=0.4753
H7	CCGATGA	*DRB3*	P=0.5977	P=0.9944	P=0.9237	P=0.9195	P=0.3342	P=0.0499*	P=0.4874
H8	CCGATAG	*DRB3*	P=0.0764	P=0.0589	P=0.0272*	P=0.0307*	P=0.8437	P=0.3810	P=0.0006*
H9	TTGGGAA	*DRB3*	P=0.4731	P=0.6293	P=0.9844	P=0.8691	P=0.6908	P=0.0188*	P=0.4238
H10	TTGATGG	*DRB3*	P=0.3836	P=0.1790	P=0.6731	P=0.5021	P=0.0154*	P=0.8884	P=0.0448*
H11	TTGGTGG	*DRB3*	P=0.1334	P=0.1776	P=0.4167	P=0.2640	P=0.1100	P=0.5392	P=0.0409*
H12	TTAATGG	*DRB3*	P=0.5632	P=0.1970	P=0.6800	P=0.4049	P=0.0292*	P=0.4869	P=0.8308
H13	TAAGG	*DRB2*	P=0.8093	P=0.6900	P=0.8923	P=0.7598	P=0.2782	P=0.5233	P= 0.0436*

MP = milk production; FP = fat production; PP = protein production; MOZ= *mozzarella* production; F% = fat percentage; P% = protein percentage and SCC = Somatic cell count

* Significant values ($P \leq 0.05$)

Discussion

MHC genes are admittedly highly polymorphic genes in vertebrate genomes (Wiseman et al., 2013). MHC genes play a central role in the immune system, and therefore these genes are established candidate for susceptibility and/or resistance to diseases and pathogens. Polymorphisms in the MHC complex genes facilitate the immune recognition process for various pathogens. Evolutionary selection pressure contribute to high genetic variation in MHC loci (Zinkernagel and Doherty, 1979).

Non-synonymous mutations can cause changes in the amino acid sequence and then alter the protein configuration and maybe its biological function (Cargill et al., 1999). Synonymous mutations do not change the amino acid sequence, and then the protein sequence is unaffected. Nonetheless, synonymous mutations may change transcription rate or transcripts turnover, affecting gene expression (Flint and Woolliams, 2008). As reviewed by Chamary et al. (2006), synonymous mutations can also affect mRNA splicing and/or stability.

In bovine species, the MHC (BoLA – *Bovine Leukocyte Antigen*) loci was mapped to chromosome 23 and in buffaloes (BuLA – *Buffalo Leukocyte Antigen*) it was mapped to the short arm of chromosome 2 (Stafuzza et al., 2013). Genotyping of BoLA (Ibrahim et al., 2012) and BuLA is relatively complex because the genes within this family are extremely polymorphic. The MHC genes are organized in three distinct classes, grouped according to their function: I, II e III (Mosafer et al., 2012). Class I MHC genes express molecules capable of destroying cells infected by invaders (Behl et al., 2012). Class II MHC genes are expressed exclusively in antigen-containing cells, such as macrophages, lymphocytes and dendritic cells and have the function of presenting extracellular antigens to CD4 + (Cluster of differentiation antigen 4) (De et al., 2002). Proteins expressed by genes in class II also degrade peptides from intracellular pathogens in T-helper cells as part of a mechanism that identifies foreign

antigens, initiating the immune response (Brown et al., 1993), while class III MHC genes encode a group of proteins with varied functions, highlighting the interactions with natural killer cell (NK), besides encoding tumor necrosis factor (TNF), both are associated with the immune response since they cause the death of cells infected by the apoptosis (Trowsdale, 2002; Behl et al., 2012).

The *DRB* and *DM* loci are promising regions of MHC to use as candidate genes in association with SSC because it is related to the immune processes of the host (Ibrahim et al, 2012). The locus *DRB* are in IIa class in the MHC, where associations with certain haplotypes and reduced SCC and/or clinical mastitis have been described before in bovine (Ibrahim et al, 2012; Fernandez et al, 2008). Kumar et al. (2011) studied *DRB3* gene polymorphisms in Murrah buffaloes and found greater numbers of genotype and alleles among healthy individuals than those with mastitis and they suppose that this heterogeneity may be the reason for resistance to mastitis in these animals. The characterization of *DM* genes in different vertebrate organisms can provide important information about the evolution of MHC, mainly the establishment of IIb gene class function. In this study, we verified that the most significant haplotypes for SCC included SNP mapped to *DRB* genes: three within *DRB3* gene (H8, H10 and H11), one within *DRB2* gene (H13) and one within *DMB* gene (H1).

The H6 haplotype formed by SNP in *DRB2* gene was associated with the highest number of traits: MP, FP, PP and MOZ. On the other hand, the traits that had the largest number of haplotype associated were P% and SCC, each with five haplotypes. When the haplotype was associated with P% it was not associated with SCC, probably this happens because there is a negative correlation between these traits. Litwinczuk et al. (2011) studied Polish Holstein-Friesian, Simmental and Jersey cattle breeds observed a distinct negative relationship between milk casein (major

fraction of milk protein) and SCC which was confirmed by a value of correlation coefficient ($r = -0.591$, $P < 0.001$). Negative values can indicate that the decrease of milk proteins is likely to be higher in more advanced stages of udder disease with more destruction of the secretory parenchyma and less secretion of caseins by the mammary gland cells (Litwińczuk et al., 2011). Whilst Ramos et al. (2015) observed that the high SCC may increase the proteolysis and breakdown of casein micelles. Therefore, regardless of mechanism theories, milk quality is directly and negatively affected by high SCC (Ramos et al., 2015; Cerón-Muñoz et al., 2002).

If SNP and haplotypes associations are established with economical important traits, SNP can be used to form a panel and then assist with selection to decrease SCC and to increase milk production and quality traits. Thus, the SNP and haplotypes tested in the current study could be proposed as markers to select buffaloes at a young age, before actual expression of the traits of interest. This selection approach will not only improve selection intensity, but also decrease generation interval, increase accuracy of prediction leading to enhanced genetic progress per unit of time and cost of production (Kumar et al., 2011; Van Eijk et al., 1992; Ryder et al., 1981).

Conclusions

This is the first work that present polymorphism study association between *DMA*, *DMB*, *DRB2* and *DBR3* genes with production traits and milk quality in buffaloes. The results suggest that the haplotypes can be used how molecular markers, since was verified associations with the traits.

References

Altiero V, Moio L, Addeo F. Previsione della resa in mozzarella sulla base del contenuto in grasso e proteine del latte di bufala. Science Tecnology Latte. 1989;40:425-433.

Ardlie KG, Kruglyak L, Seielstad M. Patterns of linkage disequilibrium in the human genome. Nat Rev Genet. 2002;3 Suppl 4:299-309.

Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21 Suppl 2:263-265.

Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, et al. Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nat Genet. 1999;22 Suppl 3:231-238.

Cerón-Muñoz M, Tonhati H, Duarte J, Oliveira J, Muñoz-Berrocal M, Jurado-Gámez H. Factors affecting somatic cell counts and their relations with milk and milk constituent yield in buffaloes. J Dairy Sci. 2002;85 Suppl 11:2885-2889.

Chamary JV, Parmley JL, Hurst LD. Hearing silence: non-neutral evolution at synonymous sites in mammals. Nat Rev Genet. 2006;7 Suppl 2:98–108.

Dabdoub SAM, Shook GE. Phenotypic relations among milk yield, somatic cell count and clinical mastitis. J Dairy Sci. 1984;67:163-164.

De S, Singh RK, Butchaiah G. MHC-DRB exon 2 allele polymorphism in Indian river buffalo (Bubalus bubalis). Anim Genet. 2002;33 Suppl 3:215-219.

El Nahas SM, Mossallem AA, Abdelhamid MI, Warda M. A study on IL8RB gene polymorphism as a potential immuno-compromised adherent in exaggeration of

parenteral and mammo-crine oxidative stress during mastitis in buffalo. J Adv Res. 2017;8 Suppl 6:617-625.

Falconer DS, Mackay TFC. Introduction to quantitative genetics. 4th ed. Edinburg:Longman Group Limited; 1996.

Fernandez IG, Ramirez JGR, Vazquez AG, Arvizu RU, Morales RAA. Polymorphism of locus DRB3.2 in populations of Creole cattle from northern Mexico. Gen Mol Biol. 2008;31 Suppl 4:880-886.

Flint APF, Woolliams JA. Precision animal breeding. Philos Trans R Soc Lond B Biol Sci. 2008;363:573-590.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, Defelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. Science. 2002;296 Suppl 5576:2225-2229.

Hill HG. Estimation of effective population size from data on linkage disequilibrium. Genet Res. 1981;38:209-216.

Hill HG, Robertson A. Linkage disequilibrium in finite populations. TAG Theo Appl Genet. 1968;38 Suppl 6:226-231.

Hill HG, Robertson A. The effect of linkage on limits to artificial selection. Genet Res.

1966;8:269-294.

Ibrahim EA, Allam NAT, Kotb EEZ, El-Rafey GA, El-Deen MMA, Fadlallah MG. Sequence-based typing-study on the relationship between subclinical mastitis and BoLA-DRB3.2 allelic polymorphism in Egyptian cows. Global Vet. 2012;9:8-22.

Kumar, S.; Sangwan, M.L. Polymorphism in DRB3 exon 2 by PCR-RFLP and its association with mastitis in Nili-Ravi breed. Indian Journal of Biotechnology. 2008;7:398-400.

Litwińczuk Z, Król J, Brodziak A, Barłowska J. Changes of protein content and its fractions in bovine milk from different breeds subject to somatic cell count. J Dairy Sci. 2011;94 Suppl 2:684-691.

Mosafer J, Heydarpour M, Manshad E, Russell G, Sulimova GE. Distribution of BoLA-DRB3 Allelic Frequencies and Identification of Two New Alleles in Iranian Buffalo Breed. ScientificWorldJournal. 2012;2012:1-6.

Penn DJ, Ilmonen P. Major Histocompatibility Complex (MHC). Encyclopedia of Life Sciences. 2005; doi:10.1038/npg.els.0003986.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81 Suppl 3:559-75.

Ramos TM, Costa FF, Pinto ISB, Pinto SM, Abreu LR. Effect of Somatic Cell Count on Bovine Milk Protein Fractions. J Anal Bioanal Tech. 2015;6 Suppl 269:1-7.

Ryder LP, Svejgaard A, Dausset J. Genetics of HLA disease association. Ann Rev Genet. 1981;15 Suppl 1:169-187.

Sambrook J, Fristch EF, Maniatis T. Molecular Cloning: A Laboratory Manual. 2nd ed. New York: Cold spring harbor laboratory press; 1989.

Sermyagin AA, Kovalyuk NV, Ermilov AN, Yanchukov IN, Satsuk VF, Dotsev AV, Deniskova TE, Brem G, Zinovieva NA. Associations of BoLA-DRB3 Genotypes with Breeding Values for Milk Production Traits in Russian Dairy Cattle Population. Agric Biol. 2016;51 Suppl 6:775-781.

Stafuzza NB, Abbey CA, Gill CA, Womack JE, Amaral MEJ. Construction and preliminary characterization of a river buffalo bacterial artificial chromosome library. Genet Mol Res. 2012;11 Suppl 3:3013-3019.

Stafuzza NB, Greco AJ, Grant JR, Abbey CA, Gill CA, Raudsepp T, Skow LC, Womack JE, Riggs PK, Amaral MEJ. A high-resolution radiation hybrid map of the river buffalo major histocompatibility complex and comparison with BoLA. Anim Genet. 2013;44:369-376.

Trowsdale J. "Both man & bird & beast": comparative organization of MHC genes.

Immunogenetics. 1995;41 Suppl 1:1-17.

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JA. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 2007;35 Suppl 2:W71-W74.

USDA. Dairy 2007. In: Reference of dairy cattle health and management practices in the United States. 2007. https://www.aphis.usda.gov/animal_health/nahms/dairy/downloads/dairy07/Dairy07_dr_PartI.pdf . Acessed 20 Dec 2017.

Van Eijk MJT, Stewart-Haynes JA, Lewin HA. Extensive polymorphism of the BOLA-DRB3 gene distinguished by PCR-RFLP. Anim Genet. 1992;23:483-496.

Wiseman RW, Karl JA, Bohn PS, Nimityongskul FA, Starrett GJ, O'Connor DH. Haplessly hoping: Macaque major histocompatibility complex made easy. ILAR J. 2013;54:196-210.

Zhao X, Lacasse P. Mammary tissue damage during bovine mastitis: causes and control. J Anim Sci. 2008;86 Suppl 13:57-65.

Zinkernagel RM, Doherty PC. MHC-Restricted Cytotoxic T Cells: Studies on the Biological Role of Polymorphic Major Transplantation Antigens Determining T-Cell Restriction-Specificity, Function, and Responsiveness. Adv Immunol. 1979;27:51-177.

CHAPTER 3 - Mapping genetic determinants of sperm mid-piece abnormalities in tropical bulls[b]

Marina R. S. Fortes[1], Laercio R. Porto-Neto[2], Nana Satake[3], Loan T. Nguyen[1], Gry Boe-Hansen[3], Ana Claudia Freitas[4] & Antonio Reverter[2]

[1]School of Chemistry and Molecular Biosciences, The University of Queensland, Saint Lucia Campus, Qld 4072, Australia m.fortes@uq.edu.au (corresponding author)
[2]CSIRO Agriculture and Food, Saint Lucia, Qld 4067, Australia
[3]School of Veterinary Sciences, The University of Queensland, Gatton Campus, Qld 4343, Australia
[4]Universidade Estadual Paulista, Campus Jaboticabal, SP 14884-900, Brazil

Summary

Viable sperm is a requirement for cattle reproduction. Sperm quality is related to fertilization ability and used to indicate bull fertility. Sperm morphology assessment, which details a number of sperm abnormalities, is a standard procedure to evaluate bulls prior to their use in breeding programs. The percentage of sperm with morphological abnormalities is an indicator of semen quality and a predictor of calf output in the field. Mid-piece abnormalities are relatively common and considered a major sperm defect as it interferes with sperm progressive motility. Here, we used the percentage of sperm with mid-piece abnormalities as a phenotype in a high-density genome-wide association study (n = 2,671 bulls, across two breeds). The heritability for the phenotype were 0.39 in Brahman and 0.38 in Tropical Composites. We identify regions of the X chromosome that are associated with mid-piece abnormalities in both breeds. In Brahman, the strongest QTL signal was mapped to the X chromosome near the gene *AKAP14*. In Tropical Composites, the strongest association was mapped to chromosome 10 near the gene *SERINC5*. Candidate genes and mutations are discussed in terms of their position and biological function.

Keywords: AKAP14, Bovine, gene association, mid-piece abnormalities, Sperm morphology

[b] Article submitted to “World Congress on Genetics Applied to Livestock Production”

Introduction

Bull fertility impacts livestock breeding as it influences the efficiency of on-farm production. Bull fertility indicators, such as sperm morphology or scrotal circumference, are used to evaluate breeding capacity (Holroyd *et al.* 2002). Sperm mid-piece abnormalities are one of the observed morphological defects in the Bull Breeding Soundness Evaluation, routine assessment for sires (Fordyce *et al.* 2006; Beggs *et al.* 2013). The percentage of morphologically normal sperm was the best predictor of calf-output in multiple-sire mating cattle operations in Northern Australia (Fitzpatrick *et al.* 2002; Holroyd *et al.* 2002). In one of these studies of tropically adapted bulls, mid-piece abnormalities (MPA) were a common sperm defect (Fitzpatrick *et al.* 2002). The mid-piece is formed by a filamentous core surrounded by many mitochondria that produce the necessary energy for the tail movements. MPA can cause serious problems for sperm motility leading to decreased fertility potential in humans (Chemes and Alvarez Sedo 2012). Potential causes for MPA are testicular and epididymis environment and genetic susceptibility, these are hypotheses to be tested.

The objectives of the current work were two-fold: 1) to provide evidence and quantify the genetic inheritance of sperm MPA; 2) to identify genomic regions and genes associated with this phenotype. To these aims, we performed genome-wide association studies (GWAS) in two breeds: Brahman and Tropical Composites.

Materials and Methods

Animals

Animal Care and Use Committee approval was not required for this study because data and samples were obtained from existing databases and storage banks of the

Cooperative Centre for Beef Genetic Technologies. Bulls were either Brahman (*Bos indicus*, n = 1,023) or Tropical Composites (*Bos indicus* and *Bos taurus* crossbreds, n = 1,648). Details about the cattle and semen analyses were published before (Burns *et al.* 2013).

Phenotypes

Classic indicators of bull fertility were measured as described in the project design (Burns *et al.* 2013). For this experiment, we used the records of MPA, described as the percentage of sperm presenting a morphological abnormal mid-piece. This percentage included the subcategories abaxial, broken necks, bent, distal reflex, Dag defect, and segmental aplasia. A minimum 100 spermatozoa per sample were evaluated by examining a thin cover-slip preparation of semen using differential interference contrast microscopy (magnification at x1000). Spermatozoa were classified according to Fordyce *et al.* (2006) by an accredited sperm morphologist.

Genotypes and Genome-wide association study (GWAS)

Genotyping services were provided by Animal Genetics Laboratory (AGL), at the University of Queensland, Gatton. Genotyping with Illumina Infinium® chemistry – ~50,000 SNP for all bulls and high-density chip for selected animals (~770,000 SNP) – quality control, and imputation were carried as described earlier (Fortes *et al.* 2012; Fortes *et al.* 2013). This earlier work reported initial GWAS on classic fertility indicators (Fortes *et al.* 2012; Fortes *et al.* 2013), but did not detail associations for any specific sperm defect. Additional bulls from the same populations were genotyped for the current study with the GeneSeek® Genomic Profiler™ chip, based on Illumina Infinium® chemistry, which features ~78,000 SNP. Quality control follow standard practices

(Fortes *et al.* 2012; Fortes *et al.* 2013), except that SNP were not excluded on the basis of low minor allele frequency (MAF) to accommodate differences in MAF between breeds and facilitate comparison of results. MAF is reported for the most significant SNP for informed interpretation of association results. Imputation to high-density used Beagle (Browning and Browning 2009). After quality control, SNP effects for 729,069 polymorphisms were calculated via single-SNP association analysis, carried within breed, since we previously ascertained that QTLs can be breed specific for these populations (Fortes *et al.* 2012; Fortes *et al.* 2013). Additive effect of each SNP on each trait was calculated by regression analysis, using genomic mixed models that corrected for contemporary group effects and used a genomic relationship matrix. Solutions to the model were estimated using SNP & Variation Suite software (Release 8.3.0, Golden Helix, Inc., 2014). We also estimated the genomic-based heritability for the trait using Qxpak (Perez-Enciso, M. & Misztal, I. 2011) in a bivariate analysis across multiple fertility traits, which included MAP and other sperm morphology defects, as well as scrotal circumference and other fertility indicators; the term “genomic-based heritability” reflects that estimates are based on genomics (without using any pedigree knowledge).

Results

The average MAP was 0.10 in both breeds, with a standard deviation of 0.11 in Brahman and 0.10 in Tropical Composites. The minimum MPA was zero in both breeds with the maximum being 0.81 in Brahman and 0.89 in Tropical Composites.

In Brahman, the genomic-based heritability for MPA was 0.39; while in Tropical Composites it was 0.38. Association analyses resulted in 201 SNP that were significant in both breeds ($P < 0.01$). These SNP were mapped to chromosomes 1, 3, 5, 6, 7, 8, 10, 11, 12, 13, 15, 16, 18, 21, 25, 26, 29 and X, with the majority (133) mapped to the

X chromosome. In Brahman, the strongest association was in the X chromosome: SNP identified as BovineHD3000001446, mapped to position 4,095,564 base pairs and associated with a *P* of 3.13 x 10^{-17} (MAF = 0.16). This same SNP in Tropical Composites had a *P* of 0.042 (MAF = 0.14). The strongest signal (with a MAF > 0.05) in Tropical Composites was the SNP identified as BovineHD1000003788 mapped to chromosome 10, position 11,100,385 base pairs, which had a *P* of 5.89 x 10^{-7}. This SNP was not significant in Brahman (*P* = 0.94 and MAF = 0.09).

We detected QTL evidence for MPA in the X chromosome in Brahman (Figure 1). This QTL was mapped to the vicinity of the gene *AKAP14*. The most significant SNP in Tropical Composites was in the intergenic region flanked by *SERINC5* and *GNPNAT1*.

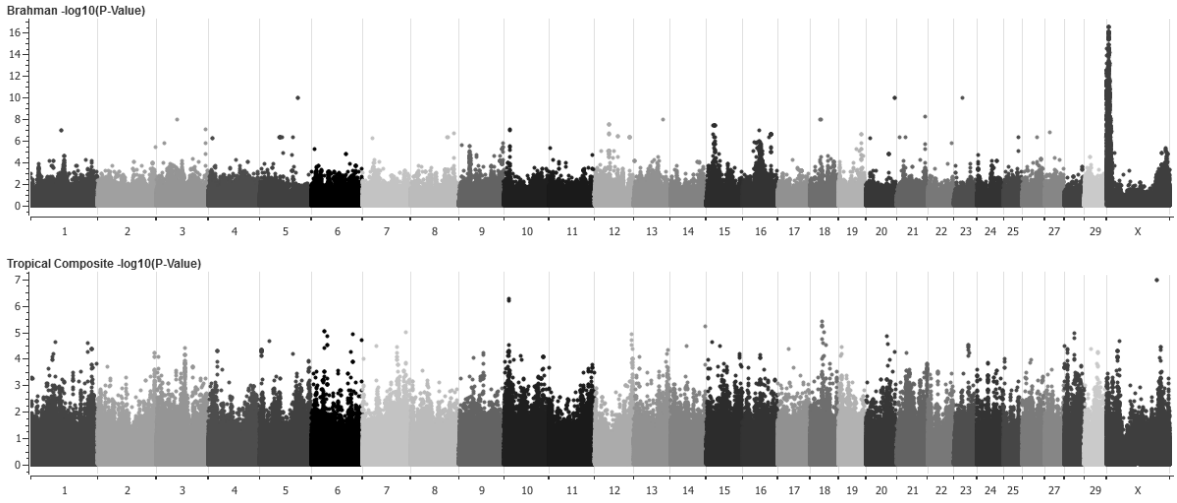

Figure 1. Manhattan plots showing the single-nucleotide polymorphisms associations found in Brahman (on top) and in Tropical Composites (on bottom) for the phenotype percentage of mid-piece abnormalities present in sperm. The y-axis displays –log(P) of the association and the x-axis is organized in chromosomal positions from chromosome 1 to X.

Discussion

The X chromosome is thought to be highly conserved among mammals (Delgado *et al.* 2009). A recent comparison between mouse and human X chromosomes concluded that most genes that were not in common between species were highly expressed in testis (Mueller *et al.* 2013). More specifically, these species-divergent genes were expressed in the germline of testis and hence Mueller and colleagues (2013) suggests that the X chromosome is an interesting place to look for genes associated to spermatogenesis and male fertility.

The SNP with the most significant *P* for MPA in Brahman was mapped to chromosome X located in the vicinity of the gene *AKAP14*; which is an A-kinase-anchoring protein. Proteins from this family are known to have a role in ciliary movements in humans. For example, tissues containing cilia or flagella express *AKAP28* in humans (*AKAP28* is the common name for *AKAP14*) and its suggested role is to modulate ciliary beat frequency (Kultgen *et al.* 2002). The *AKAP* family in general is known to have a role in gametogenesis and sperm specific isoforms have been implicated in pathways that influence sperm motility (Luconi *et al.* 2011). The function of *AKAP* proteins is to regulate post-transcriptionally signalling cascades; mainly through interaction with PKA (protein kinase) in a space-time precise manner. For example, some sperm *AKAP* isoforms localize specifically to the mitochondria, guiding its organization in the mid-piece during spermatogenesis (Luconi *et al.* 2011). In the *AKAP14* exon closest to the association peak found in Brahman, there are two splice site variants, nine missense mutations, one stop codon gain variant and a synonymous variant (UMD 3.1 assembly). These mutations are now strong positional candidates to explain the identified QTL. The most significant SNP in Tropical Composites was mapped to the intergenic region flanked by *SERINC5* and *GNPNAT1*. Two copy number variation (CNV) structural variants are located in this intergenic

region (UMD3.1 assembly). CNV and mutations present in the nearby genes could be explored to further elucidate the factors associated with MPA.

List of References

Beggs, D., J. Bertram, P. Chenoweth, K. Entwistle, G. Fordyce, H. Johnston, P. Johnston, M. McGowan, G. Niethe, S. Norman, & V. Perry, 2013. Veterinary Bull Breeding Soundness Evaluation. Austalian Cattle Veterinarians, 2404 Logan Rd Eight Mile Plains QLD 4113.

Browning, B.L., & S.R. Browning 2009. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Gen. 84.

Burns, B.M., N.J. Corbet, D.H. Corbet, J.M. Crisp, B.K. Venus, D.J. Johnston, Y. Li, M.R. McGowan, & R.G. Holroyd 2013. Male traits and herd reproductive capability in tropical beef cattle. 1. Experimental design and animal measures. Anim. Prod. Sc. 53, 87-100.

Chemes, H.E., & C. Alvarez Sedo 2012. Tales of the Tail and Sperm Head Aches Changing concepts on the prognostic significance of sperm pathologies affecting the head, neck and tail. Asian J. Androl. 14, 14-23.

Delgado, C.L.R., P.D.Waters, C. Gilbert, T.J. Robinson, & J.A.M. Graves 2009. Physical mapping of the elephant X chromosome: conservation of gene order over 105 million years. Chr. Res. 17, 917-926.

Fitzpatrick, L.A., G. Fordyce, M.R. McGowan, J.D. Bertram, V.J. Doogan, J. De Faveri, & R.G. Miller & R.G Holroyd 2002. Bull selection and use in northern Australia. Part 2. Semen traits. Anim. Reprod. Sc. 71, 39-49.

Fordyce, G., K. Entwistle, S. Norman, V. Perry, B. Gardiner, & P. Fordyce, 2006. Standardising bull breeding soundness evaluations and reporting in Australia. Theriogenol. 66, 1140-1148.

Fortes, M.R.S., A. Reverter, R.J. Hawken, S. Bolormaa, & S.A. Lehnert 2012. Candidate genes associated with testicular development, sperm quality, and hormone levels of inhibin, luteinizing hormone, and insulin-like growth factor 1 in Brahman bulls. Biol. Reprod. 87.

Fortes, M.R.S., A. Reverter, M. Kelly, R. McCulloch, & S.A. Lehnert 2013. Genome-wide association study for inhibin, luteinizing hormone, insulin-like growth factor 1, testicular size and semen traits in bovine species. Androl. 1, 644-650.

Holroyd, R.G., W. Doogan, J. De Faveri, G. Fordyce, M.R. McGowan, J.D. Bertram, D.M. Vankan, L.A. Fitzpatrick, G.A. Jayawardhana, & R.G. Miller, 2002. Bull selection and use in northern Australia. 4. Calf output and predictors of fertility of bulls in multiple-sire herds. Anim. Reprod. Sc. 71, 67-79.

Kultgen, P.L., S.K. Byrd, L.E. Ostrowski, & S.L. Migram 2002. Characterization of an A-kinase anchoring protein in human ciliary axonemes. Mol. Biol. Cell 13, 4156-4166.

Luconi, M., G. Cantini, E. Baldi, & G. Forti 2011. Role of a-kinase anchoring proteins (AKAPs) in reproduction. Front. Biosc.-Land. 16, 1315-1330.

Mueller, J.L., H. Skaletsky, L.G. Brown, S. Zaghlul, S. Rock, T. Graves, K. Auger, W.C. Warren, R.K. Wilson, D.C. Page 2013. Independent specialization of the human and mouse X chromosomes for the male germ line. Nat. Genet. 45, 1083-7.