

**UNIVERSIDADE ESTADUAL PAULISTA “JULIO DE MESQUITA FILHO”**

**FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS  
CÂMPUS DE JABOTICABAL**

**CONSTRUÇÃO DE UMA BIBLIOTECA BAC E AVALIAÇÃO DE  
MARCADORES PARA CARACTERIZAÇÃO DE REGIÕES ALVO  
DO GENOMA DO BÚFALO**

**Nedenia Bonvino Stafuzza**  
Bióloga

JABOTICABAL – SÃO PAULO – BRASIL  
Novembro de 2010

**UNIVERSIDADE ESTADUAL PAULISTA “JULIO DE MESQUITA FILHO”**

**FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS  
CÂMPUS DE JABOTICABAL**

**CONSTRUÇÃO DE UMA BIBLIOTECA BAC E AVALIAÇÃO DE  
MARCADORES PARA CARACTERIZAÇÃO DE REGIÕES ALVO  
DO GENOMA DO BÚFALO**

**Nedenia Bonvino Stafuzza**

**Orientadora: Profa. Dra. Maria Elisabete Jorge Amaral**

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

**JABOTICABAL – SÃO PAULO – BRASIL  
Novembro de 2010**

Stafuzza, Nedenia Bonvino  
S779c Construção de uma biblioteca BAC e avaliação de marcadores  
para caracterização de regiões alvo do genoma do búfalo / Nedenia  
Bonvino Stafuzza. -- Jaboticabal, 2010  
ix, 142 f. : il. ; 28 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de  
Ciências Agrárias e Veterinárias, 2010  
Orientadora: Maria Elisabete Jorge Amaral  
Banca examinadora: Humberto Tonhati, Vera Fernanda Martins  
Hossepian de Lima, Beatriz da Costa Aguiar Alves Reis, Luciane  
Madureira de Almeida  
Bibliografia

1. *Bubalus bubalis*. 2. Biblioteca genômica. 3. Mapeamento. I.  
Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 636.293:636.082

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



UNIVERSIDADE ESTADUAL PAULISTA

CAMPUS DE JABOTICABAL

FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS DE JABOTICABAL

### CERTIFICADO DE APROVAÇÃO

**TÍTULO:** CONSTRUÇÃO DE UMA BIBLIOTECA BAC E AVALIAÇÃO DE MARCADORES PARA CARACTERIZAÇÃO DE REGIÕES ALVO DO GENOMA DO BÚFALO

**AUTORA:** NEDENIA BONVINO STAFUZZA

**ORIENTADORA:** Profa. Dra. MARIA ELISABETE JORGE AMARAL

**CO-ORIENTADORA:** Profa. Dra. VERA FERNANDA MARTINS HOSSEPIAN DE LIMA

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

  
Profa. Dra. MARIA ELISABETE JORGE AMARAL

Departamento de Biologia / Instituto de Biociências, Letras e Ciências Exatas de São José do Rio Preto

  
Prof. Dr. HUMBERTO TONHATI

Departamento de Zootecnia / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal

  
Profa. Dra. VERA FERNANDA MARTINS HOSSEPIAN DE LIMA

Departamento de Medicina Veterinária Preventiva e Reprodução Animal / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal

  
Profa. Dra. BEATRIZ DA COSTA AGUIAR ALVES REIS

Centro de Pesquisas Em Biotecnologia / Universidade de São Paulo / São Paulo/SP

  
Profa. Dra. LUCIANE MADUREIRA DE ALMEIDA

Departamento Agronomia e Engenharia Florestal / Universidade Estadual de Goiás / Goiania/GO

Data da realização: 12 de novembro de 2010.

## DADOS CURRICULARES DA AUTORA

**NEDENIA BONVINO STAFUZZA** - nascida em 11 de setembro de 1980, na cidade de São José do Rio Preto - SP, Brasil, filha de Maria Celia Stafuzza e Moacir Bonvino Júnior. Iniciou em Fevereiro de 2000 o curso de Ciências Biológicas no Instituto de Biociências, Letras e Ciências Exatas de São José do Rio Preto-SP (IBILCE – UNESP), onde em 29 de novembro de 2003 e em 02 de Julho de 2005 obteve os títulos de Licenciada e Bacharel em Ciências Biológicas, respectivamente. Em março de 2005 ingressou no Programa de pós-graduação em Genética (IBILCE – UNESP) como aluna de mestrado, bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), orientada pela Profa. Dra. M. Elisabete Amaral, cuja dissertação intitulada “Construção de um Mapa Comparativo Preliminar do Cromossomo 6 do búfalo de rio (*Bubalus bubalis*) utilizando células somáticas híbridas irradiadas” foi defendida em fevereiro de 2007. Ingressou no curso de doutorado em março de 2007, no mesmo programa de pós-graduação e com bolsa da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), também sob orientação da Profa. Dra. M. Elisabete Amaral. Em dezembro de 2008, transferiu o doutorado para o Programa de Pós-graduação em Genética e Melhoramento Animal (FCAV – UNESP), sob a mesma orientação, onde a tese de doutorado foi concluída. Atuou como professora substituta da disciplina de Biologia Molecular no IBILCE-UNESP (Campus de São José do rio Preto-SP) nos anos de 2007 e 2008. As atividades desenvolvidas ao longo da carreira acadêmica resultaram em sete artigos publicados em revistas internacionais, além de 28 resumos publicados em anais de congressos nacionais e internacionais. Também durante o período, exerceu a co-orientação de seis alunos de iniciação científica e/ou estágio básico.

“Há um tempo em que é preciso abandonar as roupas usadas, que já tem a forma do nosso corpo, e esquecer os nossos caminhos, que nos levam sempre aos mesmos lugares. É o tempo da travessia e, se não ousarmos fazê-la, teremos ficado para sempre à margem de nós mesmos.”

Fernando Pessoa

## **AGRADECIMENTOS**

Aos então coordenadores do Programa de Pós-graduação em Genética e Melhoramento Animal da Faculdade de Ciências Agrárias e Veterinárias (FCAV-UNESP), Dra. Lúcia Galvão de Albuquerque e Dr. Humberto Tonhati, pela oportunidade de realizar o Doutorado em um programa de excelente qualidade.

À FAPESP pela concessão da bolsa de estudos durante 42 meses, deixo aqui expresso meus agradecimentos.

À Dra. Elisabete Amaral, por esses quase 10 anos de orientação. Tudo que aprendi na vida acadêmica teve sua participação e orientação. Obrigada por ter me ensinado a fazer Ciência, sempre com ética e responsabilidade. Agradeço sua disponibilidade irrestrita, seu modo exigente, crítico e criativo de arguir as idéias, sempre me instigando a pensar “cientificamente” para alcançar meus objetivos. Obrigada pela sua compreensão, por sua amizade, pela alegria que tem sido trabalhar ao seu lado e, principalmente, pela confiança depositada.

À Dra. Vera Hossepian de Lima pela orientação/co-orientação durante a fase de transição do credenciamento da Dra. Elisabete Amaral junto ao programa de pós-graduação em Genética e Melhoramento Animal. Obrigada por ter me aceito como uma das suas “bruxas”. Jamais esquecerei o carinho e a atenção prestada.

Aos professores Dr. Humberto Tonhati, Dra. Lucia Galvão de Albuquerque, Dra. Maria Inês Tiraboshi Ferro e Dra. Eliana Gertrudes Macedo Lemos pelas sugestões prestadas no Exame de Qualificação da presente Tese.

Aos membros da comissão examinadora da presente tese de doutorado Dra. Luciane Madureira de Almeida, Dra. Beatriz da Costa Aguiar Alves Reis, Dra. Vera Fernanda Martins Hossepian de Lima e Dr. Humberto Tonhati por terem aceitado o convite de prontidão. Agradeço as valiosas sugestões e/ou correções prestadas. A participação de vocês foi de extrema importância para deixar esse trabalho mais completo.

Aos professores do programa de Pós graduação em Genética e Melhoramento Animal (FCAV - UNESP) por terem auxiliado na minha formação.

Aos colegas de laboratório que tive o prazer de conviver durante o doutorado, cujo tempo de convivência foi muito menor do que eu gostaria, porém o suficiente para criarmos laços de amizade: André Santana, Anelise Russo, Bruna Naressi, Carolina Stucchi, Edson Rodrigues-Filho, Larissa Fornitano, Mário Azevedo, Mariana Victoretti, Melissa Miziara e Vanderlei Pelai.

A todos os colaboradores, que não foram poucos, já que essa tese é um produto coletivo, onde várias pessoas contribuíram direta ou indiretamente para que este trabalho chegasse ao bom termo. A todos os colaboradores, registro aqui minha eterna gratidão.

Aos eternos amigos por estarem sempre comigo, mesmo distantes! Vocês fazem muita falta! Agradeço em especial a grande amiga Paola, por todos esses anos de amizade, respeito e confiança. Foram inúmeros os momentos em que exerceu a paciência, a preocupação, a tolerância e o respeito, itens que só cabem a uma pessoa de espírito iluminado! Obrigada pelos conselhos e incentivos, sempre me colocando em pé durante meus tombos!

À todos os membros da família Stafuzza, em especial a minha mãe por ter me apoiado e permitido a realização de um sonho. Sem você eu não teria chegado até aqui!

Ao Luciano, meu eterno namorado, por esses 10 anos de respeito, amizade e companheirismo!

Obrigada a todos pelo carinho, respeito e amizade!

## SUMÁRIO

	Página
CAPÍTULO 1 - CONSIDERAÇÕES GERAIS .....	1
Referências bibliográficas .....	13
CAPÍTULO 2 - CONSTRUÇÃO DE UMA BIBLIOTECA BAC DE BUFALO DE RIO ....	17
Resumo .....	18
Introdução .....	19
Objetivos .....	28
Material e Métodos .....	29
Resultados e Discussão .....	46
Conclusões .....	55
Referências bibliográficas .....	56
CAPÍTULO 3 - AVALIAÇÃO DE MARCADORES DE DNA DO GENOMA BOVINO PARA SELEÇÃO DE CLONES DA BIBLIOTECA BAC DE BÚFALO .....	60
Resumo .....	61
Introdução .....	62
Objetivos .....	77
Material e Métodos .....	78
Resultados e Discussão .....	81
Conclusões .....	97
Referências bibliográficas .....	98
CAPÍTULO 4 - GERAÇÃO DE MARCADORES DERIVADOS DO GENOMA DO BÚFALO PARA O GENE CSN3 .....	107
Resumo .....	108
Introdução .....	109
Objetivos .....	114
Material e Métodos .....	115
Resultados e Discussão .....	119
Conclusões .....	126
Referências bibliográficas .....	127
APÊNDICES .....	130
Apêndice 1 - Seqüência do Vetor pBeloBAC11.....	131
Apêndice 2 - Soluções utilizadas na construção da biblioteca BAC de búfalo .....	134
Apêndice 3 - Seqüência consenso do gene CSN3 de búfalo .....	137
Apêndice 4 – Alinhamento da seqüência consenso do gene CSN3 de búfalo com as seqüências do gene CSN3 bovino .....	138

## CONSTRUÇÃO DE UMA BIBLIOTECA BAC E AVALIAÇÃO DE MARCADORES PARA CARACTERIZAÇÃO DE REGIÕES ALVO DO GENOMA DO BÚFALO

**RESUMO** - O crescimento da população bubalina em território brasileiro está relacionado ao interesse dos produtores nesse animal, como uma alternativa para a produção de carne, leite e seus derivados. Diante deste panorama, torna-se necessário o aprimoramento de programas de melhoramento genético, visando a seleção de animais geneticamente superiores para a reprodução. Por essa razão, o conhecimento do genoma da espécie é de extrema valia para o setor, uma vez que gera informações necessárias à identificação e avaliação de genes associados com características de interesse econômico. Desse modo, o presente trabalho envolveu a construção de uma nova ferramenta genômica para búfalo, denominada biblioteca BAC, a qual permitirá um novo direcionamento nos estudos moleculares do genoma bubalino, destacando-se a definição da estrutura e organização de genes de interesse econômico, fornecendo informações em nível de seqüência de DNA para o entendimento dos mecanismos e regulação dos mesmos. Com a disponibilidade dessa ferramenta, as primeiras regiões a serem caracterizadas serão aquelas que contêm genes de interesse econômico, as quais se destacam as regiões com genes relacionados com produção e qualidade do leite, regiões com genes relacionados com resposta imune e adaptativa, e regiões com genes da família das lipocalinas, os quais estão envolvidos com características de produção e reprodução. Porém, para que esses genes possam ser caracterizados por meio da biblioteca BAC de búfalo, há a necessidade de gerar marcadores para identificar e isolar clones específicos para esses genes. Marcadores derivados do genoma bovino têm sido utilizados com sucesso em estudos de mapeamento do genoma bubalino, os quais podem ser uma fonte importante de marcadores. Porém, para famílias gênicas, há a necessidade da geração de marcadores específicos para búfalo, uma vez que apresentam seqüências de DNA muito semelhantes. A disponibilidade de marcadores torna possível a caracterização dessas regiões-alvo no genoma do búfalo, fornecendo informações em nível de seqüência de DNA, que quando comparadas com a seqüência correspondente em outros bovídeos, podem revelar diferenças significativas de determinadas características entre as espécies.

**Palavras-chave:** *Bubalus bubalis*, caseínas, lipocalinas, mapeamento, resposta imune, seqüenciamento

## **CONSTRUCTING A BAC GENOMIC LIBRARY AND EVALUATING MARKERS TO CHARACTERIZE TARGET REGIONS OF THE BUFFALO GENOME**

**SUMMARY** - The increase of the buffalo population in Brazil is the result of the great interest of the producers as an alternative source for the production of meat, milk and dairy products. Thus, it becomes necessary genetic improvement programs more effective, in order to select genetically superior animals for breeding. The knowledge of the buffalo genome is valuable in this regard, since it generates information to identify and evaluate genes associated with economically traits. In this project we constructed a new genomic tool for buffalo - a genomic BAC library - which can be used in molecular studies of buffalo genome especially those related with the definition of the molecular structure and organization of economically important genes. Once this genomic tool is available, the first target regions of the buffalo genome to be characterized are those containing genes related to milk production, adaptive and innate immune response, and regions lipocalin genes involved in reproduction and production traits. However, to characterize these regions using the BAC library is necessary to have molecular markers to be able to identify and isolate the specific clones. Markers derived from the bovine genome have been successfully used in buffalo genome mapping studies, showing to be an important source of markers. On the another hand, for those genes found in the genome as gene families, there is a need for buffalo specific markers. The evaluation of new markers will contribute to the characterization of those target regions of the buffalo genome, providing information about the genomic architecture of this specie when compared with other bovid.

**Keywords:** *Bubalus bubalis*, caseins, immune response, lipocalins, mapping, sequencing

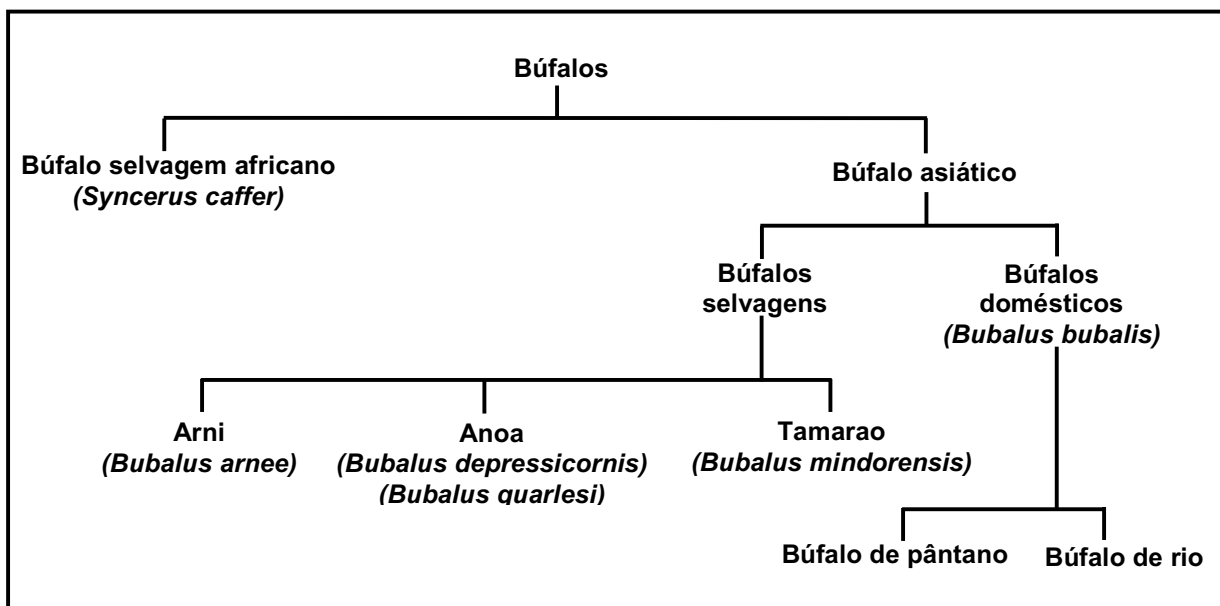
# Capítulo 1

## Considerações Gerais



## CAPÍTULO 1 - CONSIDERAÇÕES GERAIS

A família Bovidae é a mais diversa das nove famílias da ordem *Artiodactyla*, cuja subsubfamília Bovinae compreende seis gêneros: *Angra*, *Bibos*, *Bison*, *Bos*, *Bubalus* e *Syncerus*. Os búfalos estão divididos em dois gêneros: os búfalos selvagens de origem africana (*Syncerus*) e os búfalos de origem asiática (*Bubalus*). Do gênero *Bubalus*, são ainda encontrados em estado selvagem os búfalos Arni na Índia e Nepal (*B. arnee*), os búfalos Anoa das planícies e montanhas da Indonésia (*B. depressicornis* e *B. quarlesi*) e os búfalos Tamarao das Filipinas (*B. mindorensis*). Já os búfalos criados com objetivos econômicos (búfalos domésticos) são apenas os da espécie *B. bubalis*, como ilustrado na Figura 1 (LENSTRA & BRADLEY, 1999).

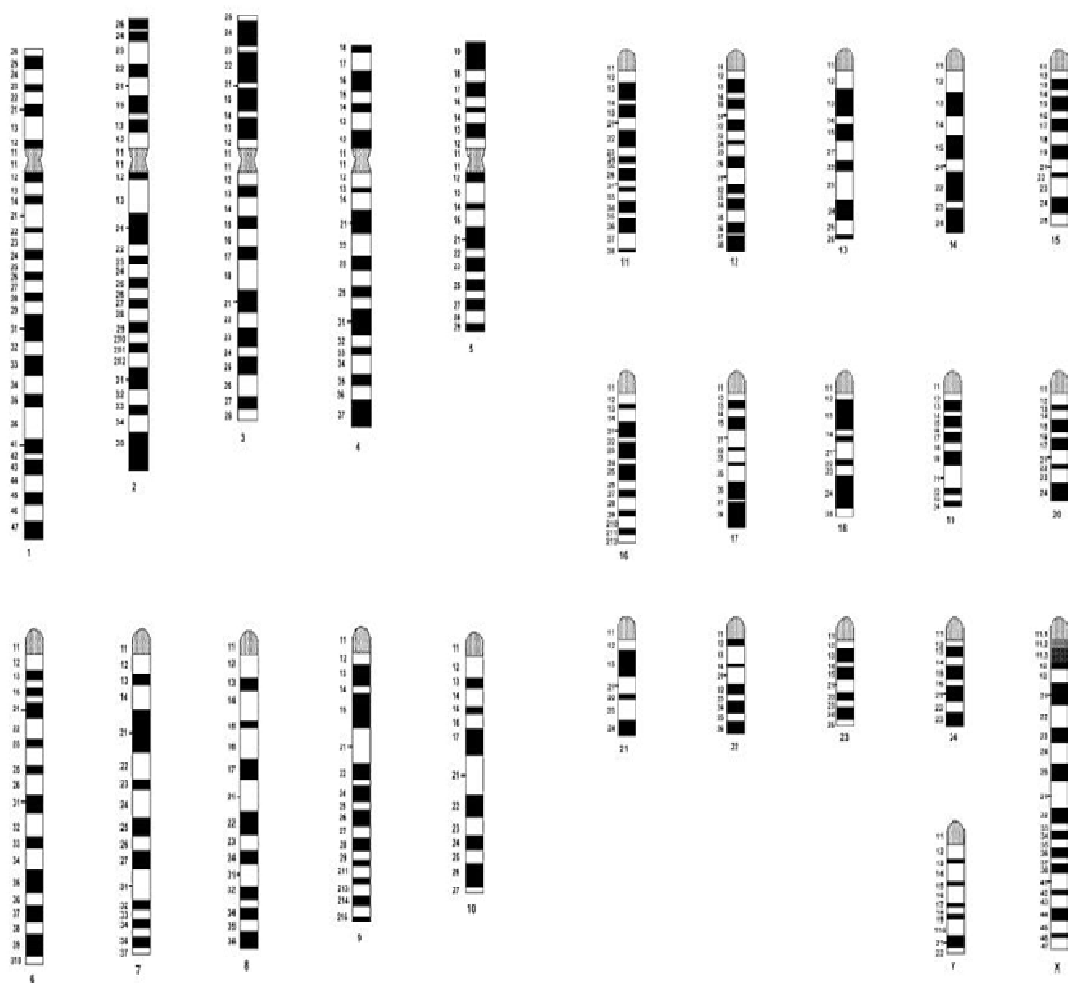


**Figura 1.** Búfalos africanos (*Syncerus*) e asiáticos (*Bubalus*) com as respectivas espécies selvagens (*S. caffer*, *B. mindorensis*, *B. depressicornis*, *B. quarlesi* e *B. arnee*) e domesticada (*B. bubalis*).

De acordo com estudos morfológicos (MACGREGOR, 1939), bioquímicos (BARKER et al., 1997a; AMANO 1983), genéticos (IANNUZZI, 1994; BERARDINO & IANNUZZI, 1981; FISCHER & ULBRICH 1968) e moleculares (LAU et al., 1998; BARKER et al., 1997b; KIKKAWA et al., 1997), os búfalos da espécie *Bubalus bubalis* podem ser divididos em dois grupos: os búfalos de pântano ( $2n=48$ ) e os búfalos de rio ( $2n=50$ ). Estimativas do tempo de divergência desses grupos dependem da calibração do relógio molecular, uma vez que extrapolações baseadas em taxas de variações de microssatélites apontam para uma divergência entre 10 e 15 mil anos (BARKER et al., 1997b), enquanto que estudos de seqüências de DNA mitocondrial revelam uma divergência entre 28 e 87 mil anos (LAU et al., 1998).

Existem diversas raças de búfalo, com maior ou menor aptidão para produção de leite, carne ou para transporte de carga e tração, uma vez que este animal é considerado de tripla aptidão (GARCIA et al., 2005; OLIVEIRA, 2005; NANDA & NAKAO, 2003).

O cariótipo do búfalo de rio consiste em cinco pares de cromossomos autossomos submetacêntricos, 19 cromossomos acrocêntricos e um par de cromossomos sexuais (X e Y), onde os cinco cromossomos submetacêntricos são resultantes de fusões entre os cromossomos acrocêntricos do genoma bovídeo ancestral (IANNUZZI, 1994). A divergência em número de cromossomos em relação ao búfalo de pântano se deu pela fusão do braço curto do cromossomo 4 com o cromossomo 9 (translocação telômero-centrômero) (IANNUZZI, 1994). A Figura 2 ilustra o ideograma do cariótipo padrão do búfalo de rio, estabelecido em 1994 por IANNUZZI.



**Figura 2.** Ideograma dos cromossomos do genoma do búfalo de rio, representado por cinco pares de cromossomos autossomos submetacêntricos, 19 pares acrocêntricos e um par de cromossomos sexuais (Adaptado de IANNUZZI, 1994).

Segundo estimativas da FAO (*Food and Agriculture Organization of the United Nations*), a espécie *Bubalus bubalis* se difundiu por todo o mundo, a partir do continente asiático, com uma população atual que ultrapassa 180 milhões de animais distribuídos em 42 países (FAO, 2008). No Brasil, os búfalos foram introduzidos no século XIX, em pequenos lotes originários da Ásia, Europa e Caribe, apresentando hoje o maior rebanho das Américas, estimado em 1,1 milhões de animais (FAO, 2008; BERNARDES, 2007).

Devido à maneira com que o sistema estatístico nacional coleta os dados agropecuários, os búfalos em muitas situações tem sido registrados como bovinos. Assim, acredita-se que haja uma subestimação da real dimensão do rebanho bubalino brasileiro, já que a Associação Brasileira de Criadores de Búfalos (ABCB), por levantamentos indiretos e avaliações de abate/desfrute, aponta um rebanho com cerca de 3,5 milhões de animais e um crescimento anual de 3% (BERNARDES, 2007). Resultados do Censo Agropecuário do Instituto Brasileiro de Geografia e Estatística (IBGE) mostram o rebanho brasileiro distribuído em aproximadamente 64,5% dos animais na região Norte, 12,8% na região Sul, 9,1% no Sudeste, 9,1% no Nordeste e 4,5% no Centro-Oeste (IBGE, 2006).

Três raças de búfalo de rio são reconhecidas pela ABCB - Mediterrâneo, Murrah e Jafarabadi – e uma raça de búfalo de pântano, denominada Carabao (ABCB, 2010). O padrão racial estabelecido pela ABCB orienta os criadores na identificação de animais quando não há registros de origem da formação de um plantel, no julgamento e na seleção de reprodutores (machos e fêmeas), bem como no estabelecimento da aptidão produtiva de acordo com a conformação e o tipo do animal (ANDRADE & GARCIA, 2005).

A adaptabilidade dos búfalos aos mais variados ambientes e sua elevada longevidade produtiva, contribuíram para o interesse dos produtores em criar os búfalos como uma alternativa para a produção de carne, leite e derivados, gerando um crescimento significativo do rebanho bubalino brasileiro (OLIVEIRA, 2005). Esse crescente aumento da bubalinocultura nacional também está relacionado com as vantagens produtivas inerentes da espécie, que elevam cada vez mais o seu desempenho econômico no cenário agropecuário brasileiro e mundial.

Diante deste panorama, torna-se necessário um aprimoramento dos programas de melhoramento genético, visando maximizar a reprodução de animais geneticamente

superiores por meio do uso de biotecnologias reprodutivas. Por isso, o conhecimento do genoma da espécie é de extrema valia para o setor, uma vez que gera informações necessárias à identificação e avaliação de genes associados com características economicamente importantes.

O mapeamento é a área mais avançada das pesquisas da genômica animal. Genes codificantes de proteínas têm sido adicionados aos mapas dos genomas dos animais da pecuária nos últimos dez anos, os quais resultaram em uma gama de testes de DNA aplicados na avaliação de populações comerciais de bovinos e suínos, por exemplo. Os dados de mapeamento associados às seqüências dos genomas formam a base para o grande desafio do momento da pesquisa genômica, que é identificar como as variações encontradas nas seqüências de um determinado gene afetam o fenótipo e, como a rede de genes e de seus transcritos interagem.

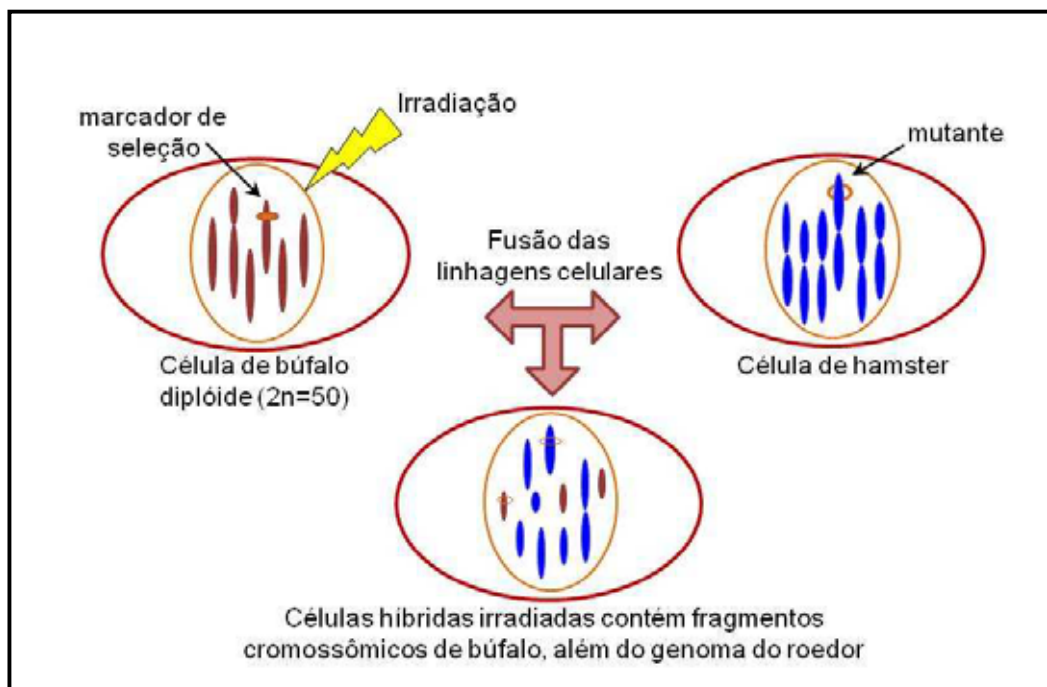
Mapas do genoma contendo genes podem ser comparados com aqueles de outras espécies para selecionar genes e/ou regiões candidatas potencialmente associadas com a segregação de características de interesse.

A mais recente estratégia de mapeamento aplicada no estudo dos genomas das principais espécies de animais de interesse econômico foi descrita por GOSS e HARRIS (1975) e redescoberta no início dos anos 90 para estudos do genoma humano (COX et al., 1990). Essa estratégia de mapeamento envolve a utilização de painéis de clones celulares híbridos (painéis RH), cujas células contêm fragmentos do genoma da espécie de interesse incorporados aleatoriamente aos cromossomos de uma célula receptora sendo, no caso de mamíferos, uma célula de roedor (hamster).

Para a obtenção dos clones celulares do painel, uma dose letal de radiação é aplicada nas células do genoma de interesse, com a função de quebrar, ao acaso, seus

cromossomos em múltiplos fragmentos (COX et al., 1990). Estes fragmentos são posteriormente incorporados de maneira aleatória ao genoma da célula receptora mutante, por meio de um processo de fusão celular como ilustrado na Figura 3.

Cada clone celular gerado a partir do processo de fusão celular contém um único conjunto de fragmentos cromossômicos da espécie de interesse incorporado ao genoma da célula receptora, atingindo estabilidade cromossômica após sucessivos processos de divisão celular.



**Figura 3.** Processo de fusão celular, ocorrido na construção de um painel de células híbridas irradiadas. Adaptação da figura gentilmente cedida pelo Dr. James E. Womack.

Estudos iniciados em 2004 no Laboratório de Genômica Comparativa - IBILCE/UNESP resultaram na construção de um painel RH para o genoma bubalino (BBURH<sub>5000</sub>) o qual foi utilizado com sucesso na geração de mapas RH de cromossomos específicos (STAFUZZA et al., 2009; IANELLA et al., 2008; RODRIGUES-FILHO et al.,

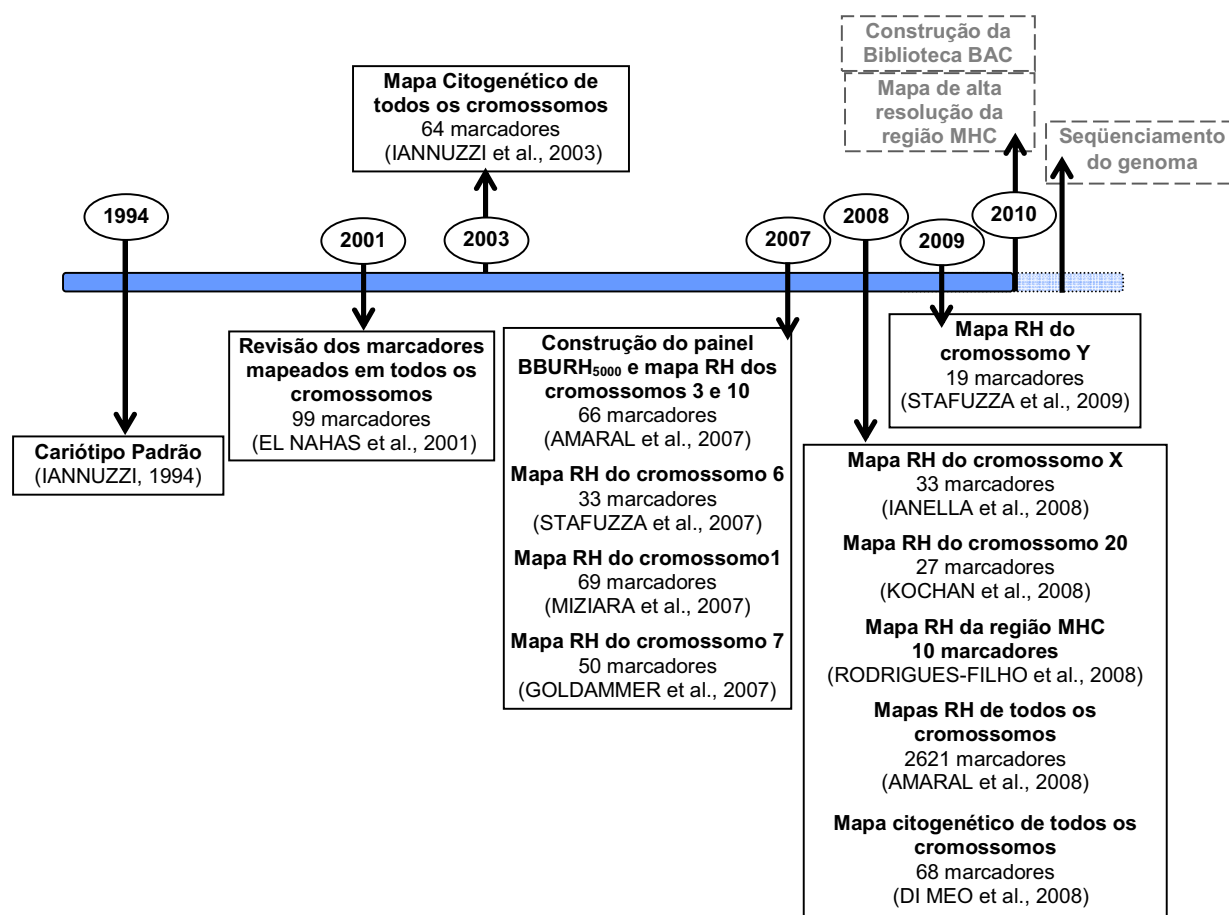
2008; AMARAL et al., 2007; STAFUZZA et al., 2007; GOLDAMMER et al., 2007; MIZIARA et al., 2007), além da primeira geração de mapas de todos os cromossomos deste genoma (AMARAL et al., 2008).

O painel BBURH<sub>5000</sub> foi utilizado em conjunto com ferramentas genômicas de última geração, como os *chips* de DNA do genoma bovino, para um mapeamento extensivo do genoma do búfalo, onde foram mapeados quase 3000 marcadores, incluindo microssatélites, genes e SNPs (do inglês *Single Nucleotide Polymorphism*) derivados do genoma bovino (AMARAL et al., 2008). Esse elevado número de marcadores resultou em mapas altamente informativos do genoma bubalino, principalmente quando comparados com o genoma bovino, fornecendo resultados relevantes para ambas espécies.

Além disso, a tecnologia de mapeamento RH inseriu o búfalo no panorama mundial das espécies de interesse econômico com ferramentas e informações genômicas disponíveis (<http://www.ncbi.nlm.nih.gov/projects/genome/guide/buffalo/>). Os mapas de todos os cromossomos do búfalo estão disponíveis no banco de dados americano *MapView*, no endereço eletrônico [http://www.ncbi.nlm.nih.gov/projects/\\_mapview/map\\_search.cgi?taxid=89462](http://www.ncbi.nlm.nih.gov/projects/_mapview/map_search.cgi?taxid=89462).

Grupos de pesquisa de diversos países estão se organizando para unir esforços visando o seqüenciamento do genoma do búfalo. Aproximadamente 67 mil seqüências de DNA de búfalo já estão depositadas no banco de dados GenBank (<http://www.ncbi.nlm.nih.gov/nucleotide/>), das quais menos de duas mil são seqüências de genes nucleares (MICHELIZZI et al., 2010). Diante dessa realidade, os mapas RH de todos os cromossomos do búfalo são de extrema importância, pois servem de orientação para organização e alinhamento das seqüências de DNA geradas.

A Figura 4 ilustra uma linha do tempo com os principais trabalhos publicados sobre os estudos do genoma do búfalo.



**Figura 4.** Principais trabalhos publicados de estudos do genoma do búfalo, onde os trabalhos de mapeamento estão representados com os respectivos números de marcadores mapeados. Os trabalhos da construção da biblioteca BAC e do mapa de alta resolução da região MHC e da Segunda Geração de mapas RH para todos os cromossomos do genoma bubalino estão previstos para serem finalizados ainda em 2010. Já o seqüenciamento do genoma bubalino total está previsto para ser concluído em 2011.

Na Figura 4 observa-se o avanço das pesquisas do genoma dessa importante espécie econômica após os estudos resultantes do uso do painel BBURH<sub>5000</sub>. Também se encontram ilustrados, os trabalhos em preparação, como por exemplo, a construção da biblioteca BAC e o mapa de alta resolução da região MHC (do inglês, *Major*

*Histocompatibility Complex*), os quais estão descritos nos capítulos 2 e 3 da presente tese, respectivamente.

Outra ferramenta genômica importante para os estudos do genoma bubalino é a biblioteca genômica do tipo BAC (do inglês *Bacterial Artificial Chromosome*), a qual recebe esta denominação devido ao tipo de vetor de clonagem utilizado. Esse tipo de ferramenta também foi construída para os principais animais de interesse econômico como suíno (LIU et al., 2010; FAHRENKRUG et al., 2001), eqüino (GODARD et al., 1998), bovino (EGGEN et al., 2001; BUITKAMP et al., 2000; WARREN et al., 2000; ZHU et al., 1999; CAI et al., 1995), caprino (SCHIBLER et al., 1998) e ovino (VAIMAN et al., 1999).

A biblioteca BAC de búfalo permitirá um novo direcionamento nos estudos moleculares do genoma bubalino destacando-se: isolamento de genes e/ou grupos de genes a partir de uma região de interesse (CHIU et al., 2004); mapeamento físico de cromossomos inteiros (CHEN et al., 2004); elucidação da organização gênica (STRONG et al., 1999); clonagem posicional (DONOVAN et al., 2000); seqüenciamento e re-seqüenciamento do genoma ou de regiões-alvo do mesmo e, identificação de elementos reguladores (KRZYWINSKI et al., 2004).

Com a disponibilidade da biblioteca BAC, as primeiras regiões do genoma do búfalo a serem caracterizadas serão aquelas que contêm os genes relacionados com produção e qualidade do leite, as quais englobam os genes codificantes das caseínas (alfa-s1, alfa-s2, beta e kapa), da beta-lactoglobulina e da alfa-lactoalbumina, presentes nos cromossomos bubalinos 7, 12 e 4 respectivamente. Outra região-alvo se encontra no braço curto do cromossomo 2 bubalino (BBU2p), onde estão localizados os genes do MHC, muitos dos quais desempenham um papel importante na resposta imune inata e adaptativa. Por fim, os genes da família das lipocalinas também se localizam em regiões importantes a serem

caracterizadas devido às diversas funções desempenhadas por essas proteínas, das quais algumas já foram relacionadas com características de produção, reprodução e resposta imune e inflamatória. Alguns genes das lipocalinas encontram-se alocados nos cromossomos bubalinos 2, 3, 4, 9, 12, 15 e 23.

Para que essas regiões possam ser caracterizadas com a biblioteca BAC de búfalo, há a necessidade de gerar marcadores de DNA para identificar e isolar clones específicos da biblioteca para esses genes. Marcadores derivados do genoma bovino têm sido utilizados com sucesso em estudos de mapeamento do genoma bubalino demonstrando ser uma opção viável (AMARAL et al., 2008). Porém, para genes com seqüências de DNA muito similares, principalmente aqueles genes oriundos de famílias gênicas, surge a necessidade da geração de marcadores específicos de búfalo.

A associação entre as informações geradas pelos mapas RH e a biblioteca BAC, possibilitará gerar conhecimentos que poderão ser aplicados no aperfeiçoamento da utilização e conservação da variabilidade genética em búfalos, na implementação dos programas de melhoramento genético desses animais e na avaliação de populações comerciais, contribuindo sobremaneira para um maior retorno econômico para o setor.

## Objetivos

Esse trabalho teve como objetivo geral ampliar as ferramentas genômicas disponíveis no Brasil para o estudo do genoma do búfalo.

Os objetivos específicos se encontram descritos em cada capítulo da presente tese, os quais podem ser sumarizados em:

- ✓ Construir uma biblioteca BAC com o genoma do búfalo de rio (*Bubalus bubalis*);
- ✓ Avaliar marcadores de DNA derivados de genes bovino para posterior aplicação na seleção dos clones da biblioteca BAC de búfalo;
- ✓ Gerar marcadores de DNA derivados do genoma do búfalo visando identificar clones na biblioteca BAC gerada.

## Referências Bibliográficas\*

ABCB. **Associação Brasileira dos Criadores de Búfalos**. Disponível em: <http://www.bufalo.com.br/>. Acesso em: Junho de 2010.

AMANO, T. Genetic differences between swamp and river buffaloes in biochemical and immunological characteristics. In: **World Conference on Animal Production**, 5., Tsukuba, Japão, Japanese Society of Zootechnical Sciences. **Anais...** 1983. p. 131-135.

AMARAL, M.E.J. et al. Construction of a river buffalo (*Bubalus bubalis*) whole-genome radiation hybrid panel and preliminary RH mapping of chromosomes 3 and 10. **Animal Genetics**, v. 38, p. 311-314, 2007.

AMARAL, M.E.J. et al. A first generation whole genome RH map of the river buffalo with comparison to domestic cattle. **BMC Genomics**, v. 9, 2008.

ANDRADE, V.J.; GARCIA, S.K. Padrões raciais e registro de bubalinos. **Revista Brasileira de Reprodução Animal**, v. 29, p. 39-45, 2005.

BARKER, J.S.F. et al. Genetic variation within and relationships among populations of Asian water buffalo (*Bubalus bubalis*). **Animal Genetics**, v. 28, p. 1-13, 1997a.

BARKER, J.S.F. et al. Genetic diversity of Asian water buffalo (*Bubalus bubalis*): microsatellite variation and a comparison with protein coding loci. **Animal Genetics**, v. 28, p. 103-115, 1997b.

BERARDINO, D.D.; IANNUZZI, L. Chromosome banding homologies in Swamp and Murrah buffalo. **The Journal of Heredity**, v. 72, p. 183-188, 1981.

BERNARDES, O. Bubalinocultura no Brasil: situação e importância econômica. **Revista Brasileira de Reprodução Animal**, v. 31, p. 293-298, 2007.

BUITKAMP, J. et al. Construction and characterization of a gridded cattle BAC library. **Animal Genetics**, v. 31, p. 347-351, 2000.

CAI, L. et al. Construction and characterization of a bovine bacterial artificial chromosome library. **Genomics**, v. 29, p. 413-425, 1995.

CHEN, R. et al. Dynamic building of a BAC clone tiling path for the Rat Genome Sequencing Project. **Genome Research**, v. 14, p. 679-684, 2004.

CHIU, C.H. et al. Bichir *HoxA* cluster sequence reveals surprising trends in ray-finned fish genomic evolution. **Genome Research**, v. 14, p. 11-17, 2004.

COX, D.R. et al. Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. **Science**, v. 250, p. 245-250, 1990.

\*Associação Brasileira de Normas Técnicas (ABNT). NBR6023:2002, Informação e documentação - Referências - Elaboração, Rio de Janeiro, Agosto de 2002.

DI MEO, G.P. et al. An extended river buffalo (*Bubalus bubalis*, 2n = 50) cytogenetic map: assignment of 68 autosomal loci by FISH-mapping and R-banding and comparison with human chromosomes. **Chromosome Research**, v. 16, p. 827-837, 2008.

DONOVAN, A. et al. Positional cloning of zebrafish ferroportin 1: identifies a conserved vertebrate iron exporter. **Nature**, v. 403, p. 776-781, 2000.

EGGEN, A. et al. Construction and characterization of a bovine BAC library with four genome-equivalent coverage. **Genetics, Selection, Evolution**, v. 33, p. 543-548, 2001.

EL NAHAS, S.M.; DE HONT, H.A.; WOMACK, J.E. Current status of river buffalo (*Bubalus bubalis* L.) gene map. **The Journal of Heredity**, v. 92, p. 221-225, 2001.

FAHRENKRUG, S.C. et al. A porcine BAC library with tenfold genome coverage: a resource for physical and genetic map integration. **Mammalian Genome**, v. 12, p. 472-474, 2001.

FAO. **Food and Agriculture Organization of the United Nations**. Faostat agriculture data. Disponível em: <http://faostat.fao.org/>. Acesso em: Agosto de 2010.

FISCHER, H.; ULBRICH, F. Chromosomes of the Murrah buffalo and its crossbreds with the asiatic swamp buffalo (*Bubalus bubalis*). **Zeitschrift für Tierzüchtung und Züchtungsbiologie**, v. 84, p. 110-114, 1968.

GARCIA, S.K.; AMARAL, A.; SALVADOR, D.F. **Situação da bubalinocultura mineira. Revista Brasileira de Reprodução Animal**, v.29, p.18-27, 2005.

**GenBank**. Disponível em: <http://www.ncbi.nlm.nih.gov/nucleotide/>. Acesso em: Julho de 2010.

GODARD, S. et al. Construction of a horse BAC library and cytogenetical assignment of 20 type I and type II markers. **Mammalian Genome**, v. 9, p. 633-637, 1998.

GOLDAMMER, T. et al. A radiation hybrid map of river buffalo (*Bubalus bubalis*) chromosome 7 and comparative mapping to the cattle and human genomes. **Cytogenetic and Genome Research**, v. 119, p. 235-241, 2007.

GOSS, S.J.; HARRIS, H. New method for mapping genes in human chromosomes. **Nature**, v. 255, p. 680-684, 1975

IANELLA, P. et al. First radiation hybrid map of the river buffalo X chromosome (BBUX) and comparison with BTAX. **Animal Genetics**, v. 39, p. 196-200, 2008.

IANNUZZI, L. Standard karyotype of the river buffalo (*Bubalus bubalis* L., 2n = 50). Report of the committee for the standardization of banded karyotypes of the river buffalo. **Cytogenetics and Cell Genetics**, v. 67, p. 102-113, 1994.

IANNUZZI, L. et al. The river buffalo (*Bubalus bubalis*, 2n=50) cytogenetic map: assignment of 64 loci by fluorescence in situ hybridization and R-banding. **Cytogenetic and Genome Research**, v. 102, p. 65-75, 2003.

IBGE. Instituto Brasileiro de Geografia e Estatística, Censo Agropecuário 2006. Disponível em: <http://www.ibge.gov.br/home/estatistica/economia/agropecuaria/censoagro/2006/agropecuario.pdf>. Acesso em: Julho de 2010.

KIKKAWA, Y. et al. Analysis of genetic diversity of domestic water buffaloes and anoas based on variations in the mitochondrial gene by cytochrome b. **Animal Genetics**, v. 28, p. 195-201, 1997.

KOCHAN, K.J. et al. Application of dissociation curve analysis to radiation hybrid panel marker scoring: generation of a map of river buffalo (*B. bubalis*) chromosome 20. **BMC Genomics**, v. 9, p. 1-12, 2008.

KRZYWINSKI, M. et al. Integrated and sequence ordered BAC- and YAC-based physical maps for the rat genome. **Genome Research**, v. 14, p. 765-779, 2004.

LAU, C.H. et al. Genetic diversity of Asian water buffalo (*Bubalus bubalis*): mitochondrial DNA D-loop and cytochrome b sequence variation. **Animal Genetics**, v. 29, p. 253-64, 1998.

LENSTRA, J.A.; BRADLEY, D.G. Systematics and phylogeny of cattle. **The genetics of cattle**, CAB International, New York, pp. 1–14, 1999.

LIU, L. et al. Construction of a bacterial artificial chromosome library for the Rongchang pig breed and its use for the identification of genes involved in intramuscular fat deposition. **Biochemical and Biophysical Research Communications**, v. 391, p. 1280-1284, 2010.

MACGREGOR, R. **The Domestic Buffalo**. Tese apresentada no Royal College of Veterinary Surgeons. The Royal College of Veterinary Surgeons, Londres, p. 1-47, 1939.

**Map Viewer**. Disponível em: <http://www.ncbi.nlm.nih.gov/mapview/>. Acesso em: Julho de 2010.

MICHELIZZI, V.N. et al. Water buffalo genome science comes of age. **International Journal of Biological Sciences**, v. 6, p. 333-349, 2010.

MIZIARA, M.N. et al. A radiation hybrid map of river buffalo (*Bubalus bubalis*) chromosome 1 (BBU1). **Cytogenetic and Genome Research**, v. 119, p. 100-104, 2007.

NANDA, A.S.; NAKAO, T. Role of buffalo in the socioeconomic development of rural Asia: Current status and future prospectus. **Animal Science Journal**, v. 74, p. 443-455, 2003.

OLIVEIRA, A.L. Búfalos: produção, qualidade de carcaça e de carne. Alguns aspectos quantitativos, qualitativos e nutricionais para promoção do melhoramento genético. **Revista Brasileira de Reprodução Animal**, v. 29, p. 122-134, 2005.

RODRIGUES-FILHO, E.A. et al. Mapping MHC genes in river buffalo. **Developments in Biologicals**, v. 132, p. 343-346, 2008.

SCHIBLER, L. et al. Construction and extensive characterization of a goat bacterial artificial chromosome library with threefold genome coverage. **Mammalian Genome**, v. 9, p. 119-124, 1998.

STAFUZZA, N.B. et al. Preliminary radiation hybrid map for river buffalo chromosome 6 and comparison to bovine chromosome 3. **Animal Genetics**, v. 38, p. 406-409, 2007.

STAFUZZA, N.B. et al. Comparative RH maps of the river buffalo and bovine Y chromosomes. **Cytogenetic and Genome Research**, v. 126, p.132-138, 2009.

STRONG, S.J. et al. A novel multigene family encodes diversified variable regions. **PNAS**, v. 96, p. 15080-15085, 1999.

VAIMAN, D. et al. Construction and characterization of a sheep BAC library of three genome equivalents. **Mammalian Genome**, v. 10, p. 585-587, 1999.

WARREN, W. et al. Construction and characterization of a new bovine bacterial artificial chromosome library with 10 genome-equivalent coverage. **Mammalian Genome**, v. 11, p. 662-663, 2000.

WOMACK, J.E. et al. A whole-genome radiation hybrid panel for bovine gene mapping. **Mammalian Genome**, v. 8, p. 854-856, 1997.

ZHU, B. et al. A 5x genome coverage bovine BAC library: production, characterization, and distribution. **Mammalian Genome**, v. 10, p. 706-709, 1999.

## Capítulo 2

Construção de uma biblioteca BAC  
de búfalo de rio



## CAPÍTULO 2 - CONSTRUÇÃO DE UMA BIBLIOTECA BAC DE BUFALO DE RIO

**RESUMO:** Estudos do genoma do búfalo no Brasil se iniciaram em 2004 com a construção de um painel de células para o mapeamento genômico da espécie. Visando dar continuidade aos estudos genômicos do búfalo, este trabalho teve como objetivo a construção de uma outra ferramenta genômica consistindo de uma biblioteca BAC. A biblioteca BAC de búfalo foi gerada utilizando a enzima de restrição HindIII para digestão parcial do DNA bubalino, o vetor pBeloBAC11 para a ligação dos fragmentos de DNA digeridos e a bactéria DH10B para o processo de transformação com a molécula de DNA recombinante. A biblioteca BAC gerada, denominada BBU\_LG, possui 50.000 clones, dos quais 240 foram analisados quanto ao tamanho do inserto, revelando um tamanho médio de inserto de 92 kb, o que fornece uma representatividade de 1,5X o genoma bubalino. A disponibilidade dessa nova ferramenta genômica para a espécie *B. bubalis* permitirá um novo direcionamento nos estudos moleculares do genoma bubalino, uma vez que a biblioteca BAC poderá ser utilizada para o isolamento e caracterização de genes a partir de uma região de interesse, no mapeamento físico, na elucidação da organização gênica, na clonagem posicional, no seqüenciamento e re-sequenciamento de regiões-alvo ou de cromossomos inteiros, além da identificação de elementos reguladores inter e intra-gênicos. A associação das informações geradas pelos mapas com a biblioteca BAC tornará possível o estabelecimento da estrutura molecular de regiões alvo do genoma do búfalo. Os resultados poderão ser aplicados em um melhor entendimento da conservação e da variação genética dos búfalos nos rebanhos brasileiros contribuindo para o desenvolvimento de novas estratégias para os programas de melhoramento genético desta importante espécie.

**Palavras-chave:** Biblioteca genômica, *Bubalus bubalis*, clonagem, cultura celular

## Introdução

### 1. A importância da construção de uma biblioteca BAC

Uma biblioteca de DNA é uma coleção de clones obtidos “*in vitro*” contendo fragmentos de DNA (insertos) inseridos em um vetor de clonagem, o qual apresenta os elementos necessários para sua replicação em uma célula hospedeira, geralmente uma bactéria. Esse DNA pode ser originado de duas fontes: do genoma da espécie (bibliotecas genômicas) ou de cDNA (DNA complementar) gerado a partir da cópia de moléculas de mRNA extraídas de um tecido específico. Muitas estratégias vêm sendo desenvolvidas e utilizadas na construção de bibliotecas de DNA de animais, plantas, insetos e microorganismos, uma vez que se trata de uma poderosa ferramenta de pesquisa que possibilita o estudo tanto de genes quanto de genomas inteiros (ZHANG, 2000).

O número de clones (N) requeridos para uma biblioteca genômica depende da cobertura desejada do genoma (C), do tamanho do genoma da espécie em questão (G) e da média do tamanho dos insertos da biblioteca (I), cuja relação é representada pela seguinte fórmula:  $C = (N \times I)/G$  (PATERSON, 1996). O valor da cobertura do genoma indica, em média, quantas vezes uma determinada seqüência está presente na biblioteca (MCCUBBIN & ROALSON, 2005).

A probabilidade (P) de um fragmento particular do genoma estar representado na biblioteca é determinada teoricamente pela seguinte fórmula:  $N = \ln(1-P)/\ln(1-[I/G])$  (CLARKE & CARBON, 1976).

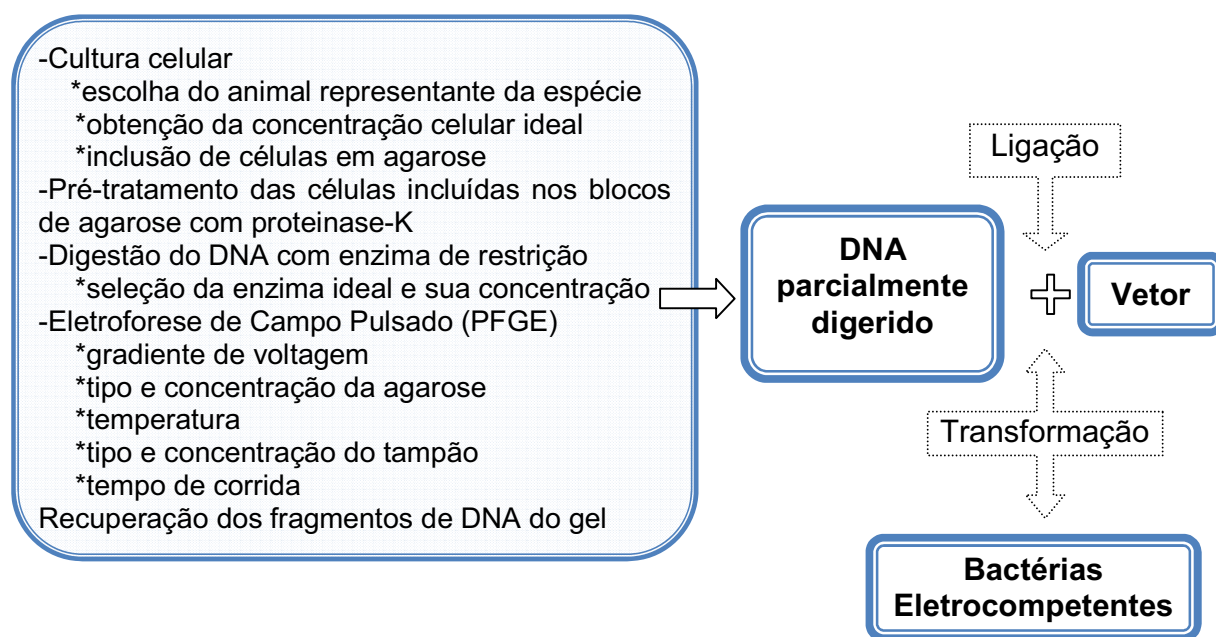
As bibliotecas genômicas do tipo BAC são aquelas que utilizam como vetor de clonagem o BAC (do inglês *Bacterial Artificial Chromosome*). Esses vetores possuem origem de replicação (fator-F) e gene para resistência ao antibiótico cloranfenicol. Esse

vetor permite a propagação estável de insertos maiores que 100 Kb, permitindo a construção de bibliotecas de alta integridade (AMENIYA & ZON, 1999). Os clones BAC são estáveis e fáceis de manipular, além de não apresentarem quimerismo (WOO et al., 1994). Clones quiméricos apresentam o inserto de DNA oriundos de diferentes partes do genoma, os quais foram combinados em um único clone acidentalmente (LIBERT et al. 1993). Esse é um grave obstáculo para o mapeamento físico, porque a verdadeira localização do gene pode não ser facilmente determinada (GILL et al., 1999).

Diferentes bibliotecas BAC foram construídas a partir do genoma dos principais animais de interesse econômico como bovino (CAI et al., 1995; ZHU et al., 1999; BUITKAMP et al., 2000; WARREN et al., 2000; EGGEN et al., 2001; FUJISAKI et al., 2002), caprino (SCHIBLER et al., 1998), eqüino (GODARD et al., 1998), suíno (SUZUKI et al., 2000; FAHRENKRUG et al., 2001; JEON et al., 2003; LIU et al., 2010) e ovino (VAIMAN et al., 1999). Essas bibliotecas mostraram-se como uma ferramenta genômica indispensável para diferentes estudos, tais como o isolamento de genes isolados ou agrupados (*clusters*) (CHIU et al., 2004; ZENG et al., 2007; LIU et al., 2010); mapeamento físico de cromossomos únicos ou de genomas inteiros (ROGEL-GAILLARD et al., 2001; CHEN et al., 2004); elucidação da organização gênica (STRONG et al., 1999; SAZANOV et al., 2006); estudos de clonagem posicional (DONOVAN et al., 2000), seqüenciamento de genomas (KRZWINSKI et al., 2004) e identificação de elementos reguladores intra e inter-gênicos (SUMIYAMA & RUDDLE, 2003).

## 2. Principais etapas envolvidas na construção de uma biblioteca BAC

O esquema a seguir (Figura 1) aponta as principais etapas da construção de uma biblioteca BAC, que envolve: a obtenção de um DNA íntegro representando o genoma de interesse, a digestão parcial deste DNA com enzima de restrição para gerar fragmentos que serão clonados (preferencialmente com um tamanho médio de 100kb), a escolha do vetor de clonagem (BAC) para a inserção dos fragmentos de DNA e transformação por eletroporação de células eletrocompetentes para inserção da molécula recombinante (DNA + vetor) na célula hospedeira (bactéria).



**Figura 1.** Representação esquemática das principais etapas e tecnologias utilizadas na construção de uma biblioteca BAC.

A seguir serão discutidos os principais aspectos técnicos relacionados com a construção de uma biblioteca BAC, um projeto sempre de grandes dimensões, realizado pela primeira vez no Brasil com o genoma de um mamífero de interesse econômico.

## **2.1. Obtenção do DNA-HMW (do inglês *High Molecular Weight*)**

Para a construção de uma biblioteca genômica é necessário determinar a concentração ideal de células/mL antes da inclusão das mesmas em blocos de agarose, para a manutenção do DNA de alta integridade. Uma revisão da literatura mostra que a concentração de células/mL é bastante variável conforme a espécie em estudo. Por exemplo, para a construção da biblioteca de bovino foi utilizada uma concentração de  $2 \times 10^7$  cel/ml (CAI et al., 1995); para rã  $2,4 \times 10^7$  cel/ml (ISHII et al., 2004); humano (ASAKAWA et al. 1997) e bovino (FUJISAKI et al., 2002)  $3 \times 10^7$  cel/ml; para salmão (THORSEN et al., 2005) e gato (BECK et al., 2001)  $5 \times 10^7$  cel/ml; para caprino  $6 \times 10^7$  cel/ml (SCHIBLER et al., 1998). Já para galinha (ZIMMER & VERRINDER GIBBINS, 1997) e urso panda (ZENG et al., 2007) foi utilizada uma concentração de  $1 \times 10^8$  cel/ml.

Outro fator importante é o tipo celular utilizado como fonte de DNA. A maioria das bibliotecas BAC de mamíferos foram construídas utilizando DNA proveniente de células brancas do sangue, como descrito para as bibliotecas de bovino (CAI et al., 1995; ZHU et al., 1999; BUITKAMP et al., 2000; WARREN et al., 2000), cachorro (LI et al., 1999), coelho (ROGEL-GAILLARD et al., 2001) e suíno (JEON et al., 2003). Bibliotecas BAC construídas com células brancas do sangue podem conter os genes das imunoglobulinas rearranjados, não sendo indicadas para estudos desses genes (FUJISAKI et al., 2002).

Informações da literatura sobre bibliotecas BAC de mamíferos construídas utilizando DNA de fibroblasto incluem aquelas do genoma bovino (FUJISAKI et al., 2002), caprino (SCHIBLER et al., 1998), eqüino (GODARD et al., 1998), gato (BECK et al., 2001) e humano (KIM et al., 1996). A principal vantagem da utilização de fibroblastos se deve à imortalização da linhagem celular, a qual pode ser congelada em nitrogênio líquido e utilizada também para outros estudos.

## **2.2. Digestão parcial do DNA com enzima de restrição**

Segundo MCCUBIN e ROALSON (2005), a enzima de restrição ideal a ser utilizada na digestão parcial do DNA deve fragmentar a molécula em intervalos uniformes de 40 a 60 Kb, o que varia conforme o DNA da espécie.

As enzimas mais utilizadas em bibliotecas BAC são EcoRI e HindIII. Algumas bibliotecas BAC que utilizaram a enzima EcoRI foram: bovino (ZHU et al., 1999; WARREN et al., 2000), cachorro (LI et al., 1999), camundongo (OSOEGAWA et al., 2000) e suíno (FAHRENKRUG et al., 2001). A enzima de restrição HindIII, por sua vez, foi utilizada na construção de bibliotecas BAC de bovino (CAI et al., 1995; EGGEN et al., 2001), caprino (SCHIBLER et al., 1998), eqüino (GODARD et al., 1998), coelho (ROGEL-GAILLARD et al., 2001), galinha (ZIMMER & VERRINDER GIBBINS, 1997), humano (KIM et al., 1996; ASAKAWA et al., 1997) ovino (VAIMAN et al., 1999) e urso panda (ZENG et al., 2007).

A enzima HindIII apresenta uma particularidade que faz com que seu uso seja vantajoso na construção de uma biblioteca BAC, que é o fato do DNA centromérico ser resistente à ação desta enzima. Sendo assim, sua utilização evita que essas seqüências altamente repetitivas sejam incorporadas nos clones gerados. A enzima HindIII é produzida a partir do microorganismo *Haemophilus influenzae* e reconhece uma seqüência palindrômica composta por seis nucleotídeos (5'-A<sup>↓</sup>AGCTT-3').

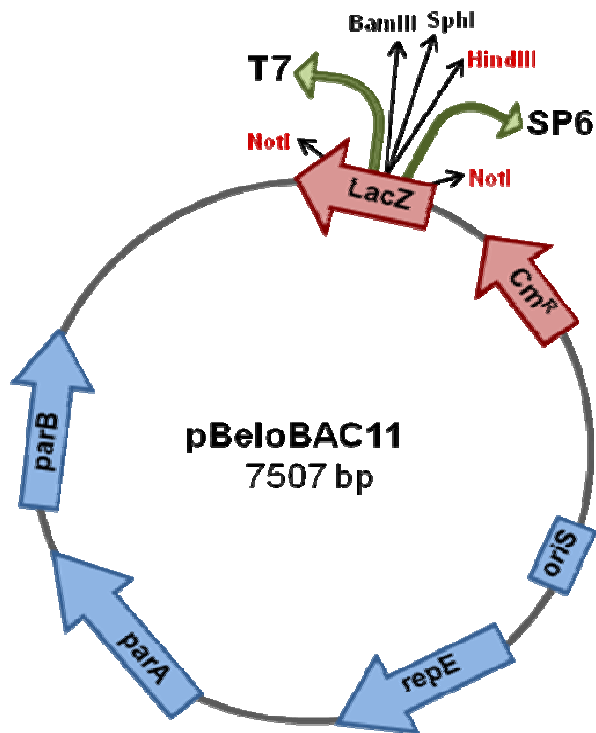
## **2.3. Vetor de clonagem do tipo BAC**

Os vetores BAC disponíveis para a construção de bibliotecas genômicas apresentam os padrões para seleção característicos dos plasmídeos como resistência a antibiótico e um sítio de policlonagem com um gene repórter. Segundo a literatura, o vetor BAC mais utilizado é o pBeloBAC11, como descrito para diversos animais de interesse econômico

como, por exemplo, galinha (ZIMMER & VERRINDER GIBBINS, 1997), eqüino (GODARD et al., 1998), caprino (SCHIBLER et al., 1998), ovino (VAIMAN et al., 1999), bovino (EGGEN et al., 2001) e coelho (ROGEL-GAILLARD et al., 2001).

O vetor pBeloBAC11 foi desenvolvido por KIM e colaboradores em 1996 e representa a segunda geração de vetores BAC. Esse vetor, desenvolvido a partir do vetor pBAC108L, contém o gene LacZ o qual permite a identificação dos clones por meio do fenótipo de colônias brancas (clones positivos) e colônias azuis (clones negativos). Sítios de clonagem para três enzimas (BamHI, SphI e HindIII) estão presentes no vetor, os quais são flanqueados pelos promotores T7 e SP6. Sítios de restrição adicionais para seis enzimas (NotI, EagI, XmaI, SmaI, BglI e SfiI), ricos em citosina e guanina, podem ser utilizados para a excisão do inserto de DNA do vetor.

O vetor pBeloBAC11 apresenta ainda o gene de resistência ao antibiótico cloranfenicol (CM<sup>R</sup>), o qual é utilizado na seleção dos clones transformados. Esse vetor também apresenta um sistema de regulação da sua replicação, denominado fator-F, que controla seu número de cópias por célula. Os genes regulatórios presentes no vetor incluem os genes oriS (origem de replicação), repE, parA e parB, sendo que os dois primeiros são mediadores da replicação unidirecional do fator-F e os dois últimos mantêm o número de cópias entre uma ou duas moléculas por célula. A Figura 2 ilustra o vetor pBeloBAC11 com seus genes e principais seqüências reguladoras, e no Apêndice 1 está apresentada a seqüência linear completa (em nucleotídeos) do referido vetor.



**Figura 2.** Vetor de clonagem pBeloBAC11, ilustrando os principais genes (regiões em vermelho) e as principais regiões reguladoras (regiões em azul). As setas verdes indicam os promotores T7 e SP6 e as setas pretas alguns sítios para enzimas de restrição. Adaptado de KIM et al., 1996.

#### 2.4. Eletroforese de Campo Pulsado (PFGE)

Em 1984, SCHWARTZ e CANTOR desenvolveram a técnica de eletroforese de campo pulsado (PFGE), tornando possível a separação de grandes fragmentos de DNA com tamanhos de até 2 Mb. No entanto, esta tecnologia também pode ser aplicada com sucesso na separação de fragmentos de DNA com tamanhos variando de 1 Kb a 6 Mb (GARDINER 1991).

Há muitas variáveis que influenciam a eletroforese de campo pulsado, como por exemplo, o gradiente de voltagem, o tipo e a concentração da agarose utilizada no gel, o tipo de tampão utilizado para a corrida, a temperatura e o tempo de duração da corrida (LAI & BIRREN 1995).

O gradiente de voltagem é a diferença entre o potencial elétrico dos eletrodos, representando a força que dirige as moléculas de DNA através do gel, a qual é expressa em unidades de voltagem por distância. Mudanças no gradiente de voltagem afetam diretamente a velocidade de migração e o tamanho das moléculas a serem separadas no gel (BIRREN & LAI, 1993).

Uma variedade de tipos de agarose encontra-se disponível para uso em eletroforese, cujas principais propriedades incluem a sua capacidade de resistência à rupturas, a carga molecular e o nível de pureza. A agarose com baixo ponto de fusão resulta em um gel com poros menores quando comparado com géis de agarose comum. Seu uso é recomendado quando se almeja resgatar fragmentos de DNA do gel após a eletroforese. Mudanças na concentração da agarose em um gel submetido à eletroforese de campo pulsado afetam diretamente a velocidade de separação dos fragmentos e sua resolução. Para separar fragmentos de DNA de 1 Kb a 2,5 Mb é aconselhável utilizar géis de agarose com baixo ponto de fusão a 1% (BIRREN & LAI, 1993).

Outro ponto crucial para a separação com sucesso de grandes fragmentos de DNA é a escolha da temperatura utilizada durante a eletroforese, sendo as temperaturas entre 12-15°C as mais utilizadas. Com o aumento da temperatura, as moléculas de DNA tendem a migrar mais rapidamente, porém há uma diminuição na resolução da separação desses fragmentos no gel. A temperatura de 14°C é tida como ideal por manter um equilíbrio entre velocidade de separação dos fragmentos de DNA e resolução (BIRREN & LAI, 1993).

A mobilidade das moléculas através do gel, assim como sua resolução também dependem da composição do tampão utilizado durante a eletroforese de campo pulsado, já que o DNA tende a migrar mais rapidamente em tampões com baixa força iônica. Tris-acetato (TAE) e Tris-borato (TBE) são os tampões mais utilizados. Fragmentos de DNA

migram 25% mais rápido em 0.5X TBE do que em 1X TBE e cerca de 10% mais rápido em 0.5X TAE do que em 1X TAE. Embora fragmentos de DNA migrem de maneira mais rápida em TAE do que em TBE, o tampão TAE perde suas propriedades iônicas após um período de corrida de 24 a 36 horas (BIRREN et al., 1997).

## Objetivos

Visando ampliar as ferramentas genômicas existentes no Brasil para os estudos do genoma do búfalo, o presente trabalho teve como objetivos:

- ✓ Construir uma biblioteca genômica do tipo BAC para o búfalo de rio (*Bubalus bubalis*) para posterior utilização em estudos relacionados com a caracterização da estrutura molecular de regiões-alvo do genoma bubalino;
- ✓ Avaliar parte dos clones obtidos na biblioteca BAC quanto ao tamanho médio do inserto de DNA e a cobertura do genoma.

## **Material e Métodos**

Todas as soluções citadas no item Material e Métodos estão descritas em detalhe no Apêndice 2, em ordem alfabética. Toda metodologia descrita a seguir foi realizada no Laboratório de Genômica Comparativa do Instituto de Biociências, Letras e Ciências Exatas de São José do Rio Preto-SP (IBILCE-UNESP).

### **1. Estabelecimento e manutenção da cultura celular**

Células de fibroblasto foram utilizadas como fonte de DNA do genoma bubalino para a construção da biblioteca BAC. Para o estabelecimento da cultura de fibroblasto, foram utilizados fragmentos de cartilagem de orelha de um animal macho da raça Murrah.

Os fragmentos do tecido foram desprovidos dos pêlos e da camada externa da pele com o auxílio de um bisturi. Durante o processo de limpeza, os fragmentos de tecido foram mantidos em solução tampão, livre de cloreto de cálcio, sulfato de magnésio e carbonato de sódio, conhecida como HANKS 1X (do inglês *Hanks' balanced salt solution* – *Sigma-Aldrich* - EUA), na qual foram adicionados 10% de um composto antibiótico e antimicótico da *GIBCO* (EUA) a 100X (10.000 U/mL de penicilina sódica, 10 mg/mL de sulfato de estreptomicina e 25 µg/ml de anfotericina B).

Para o estabelecimento da cultura, os fragmentos foram cortados em pedaços de aproximadamente 1 mm, e transferidos para frascos de cultura com superfície de 25 mm<sup>2</sup> contendo 5 mL de meio de cultura DMEM (do inglês *Dulbecco's Modified Eagle's Medium* - *GIBCO* - EUA) enriquecido com 10% de soro fetal bovino estéril (*Cripion* - BR), 1% de antibiótico e antimicótico (*GIBCO* - EUA) e 1% de piruvato de sódio a 100mM (*GIBCO* - EUA). As culturas celulares foram mantidas em estufa à 37°C, 5% de CO<sub>2</sub> e 90% de

umidade. Durante os 10 primeiros dias de cultura, o meio foi trocado diariamente. Após esse período inicial, o meio de cultura foi trocado conforme a confluência apresentada pelas células.

Diariamente, a proliferação do número de células foi acompanhada visando a obtenção de uma camada uniforme de células por toda a superfície de cultura do frasco. O manuseio das células, troca do meio de cultura e posteriores subculturas foram realizadas em condições de assepsia em fluxo laminar.

As subculturas foram realizadas na fase de confluência das células, ou seja, no momento em que a superfície de cultura do frasco encontrava-se inteiramente coberta por células, o que ocorreu após 16 dias da manutenção da cultura celular. Tal procedimento foi realizado removendo-se o meio antigo, lavando-se as células em cultura com 1 mL de solução HANKS 1X e em seguida, promovendo a liberação das células do frasco por ação da tripsina. Este passo foi realizado adicionando 1 mL de solução de tripsina 1X (*Sigma-Aldrich* - EUA) diluída em solução HANKS. Após 5 minutos de incubação com tripsina, as células foram ressuspendidas em 1 mL de solução HANKS 1X, e posteriormente transferidas para novos frascos de cultura providos com meio DMEM (5 mL quando em frascos com 25 mm<sup>2</sup> de superfície e 25 mL quando em frascos com 150 mm<sup>2</sup>).

Os fibroblastos em cultura foram utilizados em três diferentes procedimentos: armazenamento em nitrogênio líquido para estoque; preparações citogenéticas para avaliação da integridade do genoma bubalino e contagem celular para inclusão em blocos de agarose.

## **2. Preparações citogenéticas para avaliação da integridade do genoma bubalino**

### 2.1. Obtenção das células metafásicas a partir da cultura celular

Cada frasco de cultura contendo uma monocamada confluyente de células foi submetido a subcultura utilizando 0,5 mL de suspensão celular, 9 mL de meio DMEM enriquecido com 1% de antibiótico, 1% de piruvato e 10% de soro fetal bovino, acrescido de 0,5 mL de BrdU (Bromodexodiuridina a 1 mg/mL – *Sigma-Aldrich* - EUA). Esses frascos foram incubados em estufa a 37°C, 5% de CO<sub>2</sub> e 90% de umidade, até o aparecimento de células mitóticas, que se deu entre 18 e 20 horas após a subcultura.

A partir do momento em que foram encontradas 10 a 20 células mitóticas nos frascos, o meio de cultura foi removido e as células foram lavadas duas vezes com solução HANKS 1X. Em seguida adicionou-se 10 mL de meio DMEM e 0,1 mL de solução de timidina a 300 µg/mL (*Sigma-Aldrich* - EUA), seguido de incubação em estufa por 6 horas.

Posteriormente ao período de incubação, adicionou-se 0,1 mL de colcemid (*GIBCO* - EUA) por 10, 15 e 30 minutos para promover o acúmulo de células em mitose e impedir que as células em metáfase completem o processo de divisão celular. Em seguida, o meio foi removido e adicionou-se 5 mL de solução hipotônica de KCl (75 mM). Essa solução de células foi transferida para um tubo cônico, o qual foi submetido à centrifugação durante 5 minutos a 1000 rpm.

O concentrado de células foi ressuspendido em 0,5 mL de solução de KCl e em seguida foi adicionada a solução fixadora contendo metanol anidro-ácido acético glacial (3:1), no qual as células foram mantidas a 4°C por 10 minutos. Em seguida as células foram novamente centrifugadas (condições descritas previamente) e o concentrado de

células foi ressuspensão com o fixador e mantidos a 4°C durante 12 horas. Por fim, promoveu-se a lavagem das células com o fixador por mais duas vezes e seguiu-se com a montagem das lâminas.

## 2.2. Preparo e montagem das lâminas para microscopia óptica

As lâminas para microscopia foram limpas, secas e mantidas no freezer (-20°C) até o momento de sua utilização. A solução de células em fixador foi gotejada nas lâminas ainda geladas de uma distância aproximada de 30 cm para obtenção de um bom espalhamento dos cromossomos. Em seguida, as lâminas foram lavadas em ácido acético 70%; rapidamente flambadas em lamparina, secas em temperatura ambiente e armazenadas em freezer -80°C.

Visando determinar a integridade do cariótipo do animal, ou seja, a ausência de qualquer tipo de alteração cromossômica numérica e/ou estrutural como deleção, duplicação, inversão e/ou translocação, as lâminas com os cromossomos em metáfase foram submetidas à diferentes técnicas citogenéticas, tais como coloração com Giemsa e Ag-NOR, além de bandamentos C e G. As análises foram realizadas sob microscópio óptico (objetiva de imersão - 100X) e foto documentadas com câmera digital Cyber-shot DSC5700 (Sony).

## 2.3. Coloração com Giemsa

As lâminas envelhecidas em estufa a 60°C por um dia, foram hidrolisadas em HCl 1N a 60°C por 7 minutos. A hidrólise foi interrompida mergulhando-se as lâminas em água deionizada gelada. Posteriormente foram lavadas em água destilada por 10 minutos e em solução tampão Sorensen por um período de 1 a 2 minutos. Na seqüência, as lâminas

foram mergulhadas em uma solução de Giemsa 6% (*GIBCO* - EUA) por 5 minutos e finalmente lavadas em água deionizada e secas em temperatura ambiente. Esta técnica de coloração permite a visualização e classificação dos cromossomos por grupos de acordo com o tamanho e a posição do centrômero.

#### 2.4. Bandamento G

As bandas G foram produzidas pelo tratamento da lâmina com tripsina e subsequente coloração com Giemsa. Para isso, foram montados 5 boréis, sendo um com solução de tripsina, dois com tampão Sorensen, um com Giemsa a 6% e um com água deionizada.

As lâminas foram envelhecidas em estufa, seguindo o mesmo procedimento descrito na técnica de coloração com Giemsa. Em seguida, foram colocadas por 20 segundos na solução de tripsina, banhadas no primeiro borel com tampão Sorensen e incubadas por 2 minutos no segundo borel com tampão Sorensen. Finalmente foram coradas em solução Giemsa 6% durante 5 minutos sob agitação, lavadas com água deionizada e secas em temperatura ambiente.

#### 2.5. Bandamento C

As lâminas foram mergulhadas em HCl 0,2 N por 5 minutos, lavadas em água gelada e secas a temperatura ambiente. Em seguida, foram incubadas de 5 a 10 minutos em solução de bário 5%. Posteriormente as lâminas foram lavadas em solução gelada de HCl 3%, em água gelada e secas a temperatura ambiente. Depois de secas foram incubadas a temperatura ambiente em solução 2X SSC por 10 minutos e em seguida, em solução 2X SSC a 60°C por uma hora. Posteriormente foram lavadas em água gelada e secas a

temperatura ambiente. Quando secas, as lâminas foram submetidas a várias incubações em álcool para promover a desidratação, sendo duas incubações em álcool 70% por 5 minutos cada e duas incubações em álcool 95% por 5 minutos cada. Posteriormente foram secas e coradas com uma solução tamponada de Giemsa 6% por 30 minutos, lavadas em água destilada e secas a temperatura ambiente.

### 2.6. Coloração Ag-NOR

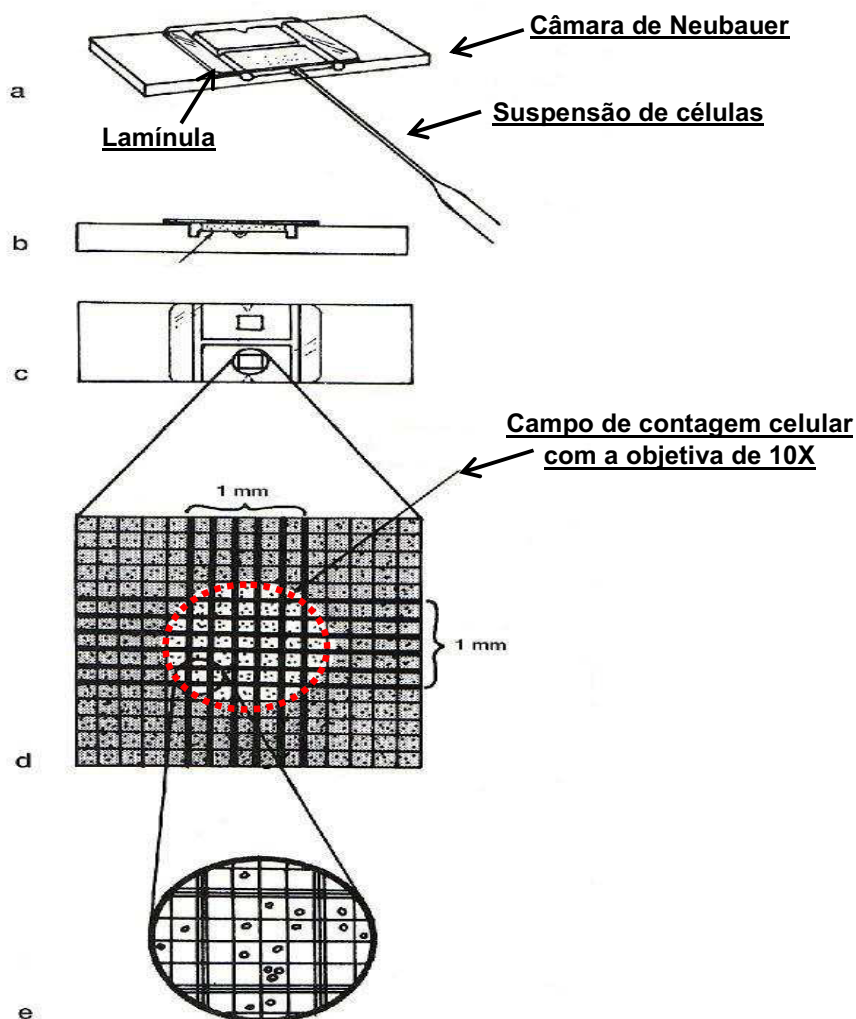
Três gotas de solução coloidal reveladora foram adicionadas sobre as lâminas, seguindo-se com a adição de 3 gotas da solução de nitrato de prata 50%, as quais foram misturadas com o auxílio de uma pipeta. Cada lâmina foi coberta com uma lamínula e incubada em câmara úmida por 7 minutos a 70°C. Em seguida, as lâminas foram lavadas duas vezes com água destilada, coradas em Giemsa 6% por 3 minutos, lavadas em água destilada e secas a temperatura ambiente.

### **3. Contagem e inclusão das células em blocos de agarose**

Na construção de uma biblioteca BAC, o processo de inclusão das células em blocos de agarose é primordial para preservar a integridade do DNA até o momento da digestão parcial e clonagem.

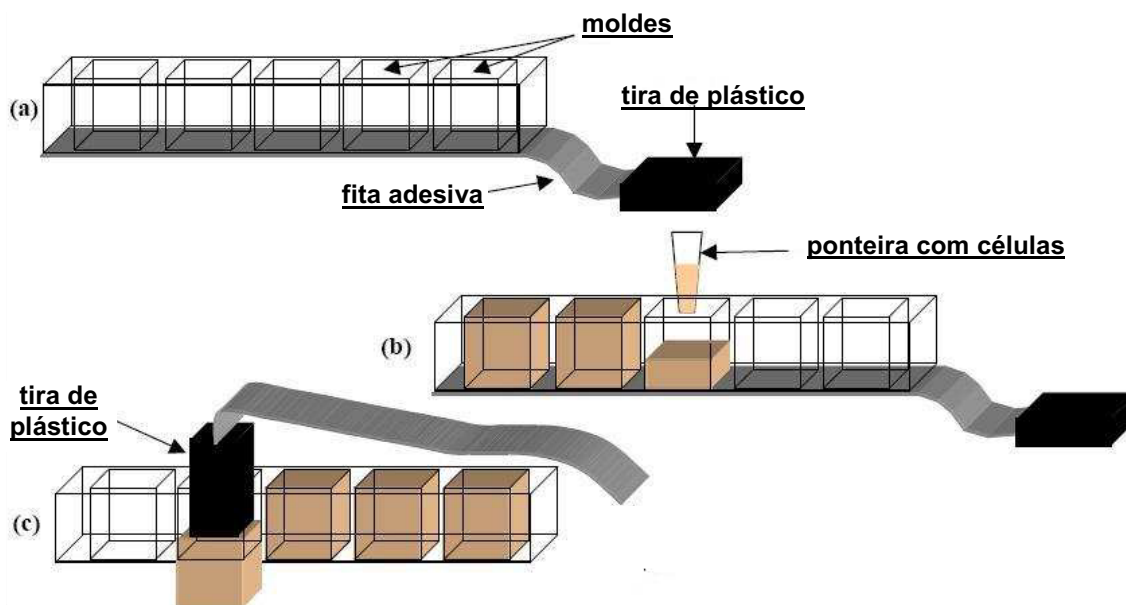
Para a contagem celular e posterior inclusão das células em blocos de agarose, frascos de cultura de 150 mm<sup>2</sup> de superfície tiveram suas células recolhidas, como descrito previamente. A determinação da concentração celular foi estimada por meio de contagem direta em câmara de *Neubauer* utilizando microscópio óptico invertido, seguindo a metodologia desenvolvida por Freshney (2000), ilustrada na Figura 3.

Por capilaridade foram introduzidos 20  $\mu\text{L}$  da solução de células entre o contador *Neubauer* e uma lamínula de vidro. Com a objetiva de 10X do microscópio óptico foi localizado o campo de contagem (Figura 3d). Após a contagem das células em ambos os campos, foi calculada a concentração do número de células, por meio da fórmula:  $C = N/V$  ( $C$ = concentração;  $N$ = número de células;  $V$ = volume em  $\text{mL} = 0,1\text{mm}^3$ ).



**Figura 3.** Câmara de Neubauer. **a)** Adição da suspensão celular na câmara de Neubauer. **b)** Seção longitudinal da câmara de Neubauer mostrando a posição da amostra celular em 0,1 mm de profundidade. **c)** Câmara de Neubauer vista de cima. **d)** Visão da área total da lâmina, cuja área central iluminada pela luz de um microscópio comum com a objetiva de 10X representa  $1\text{mm}^2$  de área. **e)** Visão ampliada de um dos 25 quadrados menores cercados por três linhas paralelas que juntos formam a área de  $1\text{mm}^2$ . Ilustração modificada de FRESHNEY, 2000.

Uma vez determinada a concentração, as células em suspensão foram mantidas em banho-maria a 42°C e misturadas com solução de agarose com baixo ponto de fusão 1% na proporção 1:1. Em seguida, 100 µl dessa solução foi aplicada nos moldes (Figura 4) para a confecção dos blocos, e mantidos a 4°C por 2 horas para a solidificação. Em seguida os blocos de agarose foram retirados dos moldes e armazenados em tubos cônicos contendo TE (Tris-EDTA pH 8.0), e mantidos a 4°C.



**Figura 4.** Preparação de blocos de agarose com DNA. **a)** Ilustração dos componentes de um molde para blocos; **b)** A mistura de células com agarose de baixo ponto de fusão é depositada dentro de cada um dos poços do molde; **c)** Depois de solidificado o conteúdo, a fita é removida e os blocos de agarose com as células são retirados dos moldes com auxílio de uma tira plástica. Esquema modificado de PETERSON e colaboradores, 2000.

Segundo a literatura, as bibliotecas BAC bovino (FUJISAKI et al., 2002) e de cabra (SCHIBLER et al., 1998) foram construídas a partir de células de fibroblastos, nas quais foram utilizadas as concentrações celulares de  $3 \times 10^7$  e  $6 \times 10^7$  cel/ml, respectivamente, na confecção de blocos de agarose. Assim, também optou-se pela geração de blocos de

agarose para búfalo com as mesmas concentrações celulares utilizadas para esses bovídeos.

#### **4. Tratamento das células contidas nos blocos de agarose com proteinase K**

Comparando-se os procedimentos estabelecidos para as diversas espécies, verifica-se a necessidade de um protocolo específico para cada espécie e tipo celular utilizado. Desse modo, o protocolo desenvolvido para a manipulação de blocos de agarose contendo células de fibroblasto de búfalo encontra-se descrito a seguir.

Os blocos de agarose com as células foram colocados em tubo cônico contendo 0,5 mg/mL de proteinase K (*Promega Corporation* – EUA) e 3 mL de uma solução de laurilsarcosina sódica-EDTA (pH 9.0) , os quais foram incubados a 50°C, sob agitação de 25 rpm, durante 12 horas para a lise da membrana nuclear e degradação das proteínas. Em seguida os tubos foram colocados no gelo para equilíbrio da temperatura. Esse processo de incubação de 12 horas foi repetido por mais duas vezes, com novas soluções de laurilsarcosina sódica-EDTA e proteinase K. Os blocos foram transferidos para um novo tubo cônico contendo 30 mL de TE (Tris-EDTA pH 8.0) acrescido de 1 mM de PMSF (fluoreto de fenilmetilsulfonil) e incubados por uma hora a temperatura ambiente, sendo invertidos ocasionalmente. Essa etapa de equilíbrio dos blocos com TE + PMSF foi repetida por mais duas vezes para remover todos os vestígios de proteinase K. Em seguida, os blocos foram transferidos para outro tubo cônico contendo 30 mL de TE e incubados por uma hora a temperatura ambiente. Essa etapa foi repetida por mais duas vezes para remover os traços de PMSF. Após as sucessivas lavagens dos blocos com TE, os mesmos foram transferidos para uma nova solução de TE e armazenados a 4°C.

A qualidade do DNA obtido foi verificada por eletroforese de campo pulsado em gel de agarose 1%, antes do processo de digestão parcial com enzima de restrição. Os parâmetros utilizados na eletroforese foram: 14°C, 120°, 6V/cm, 1,4s-13,5s por 16 horas em tampão TAE 1X.

## **5. Digestão parcial do DNA nos blocos de agarose com enzima de restrição HindIII e recuperação dos fragmentos de DNA de interesse**

A digestão parcial do DNA nos blocos de agarose é outro aspecto crucial no processo de construção de uma biblioteca BAC, uma vez que cada genoma apresenta uma exigência particular. A concentração de células assim como a concentração de enzima de restrição necessária para uma digestão parcial do DNA é específica para cada genoma.

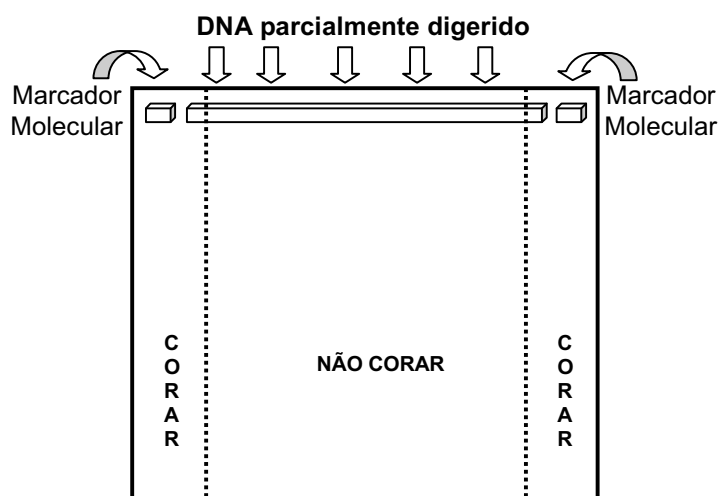
Para se obter o DNA parcialmente digerido, há quatro metodologias descritas que se baseiam no controle de componentes e variáveis da reação de digestão do DNA. São elas: variação na concentração da enzima de restrição (BURKE & OLSON, 1991), variação no tempo de incubação com a enzima de restrição (ANAND et al., 1989), variação na concentração de co-fatores da enzima como, por exemplo,  $Mg^{+2}$  (ALBERTSEN et al., 1989), e variação da proporção de enzimas e metilases a serem adicionadas na reação de digestão do DNA (LARIN et al., 1991). Para búfalo, optou-se pela utilização de duas variáveis - a concentração de enzima de restrição e tempo de incubação - por serem de mais fácil controle, resultando em uma menor porcentagem de erros durante a manipulação.

Para o teste de digestão, cada bloco de agarose com as células foi cortado em 4 partes iguais. Cada  $\frac{1}{4}$  do bloco foi fatiado 4 vezes na direção horizontal e 4 vezes na

direção vertical, cujos fragmentos foram colocados em microtubos contendo 5 µl do tampão da enzima HindIII (*New England Biolabs* - EUA), 5 µl de espermidina a 40 mM (*Sigma-Aldrich* - EUA), 0,5 µl de BSA a 10 mg/ml (*Sigma-Aldrich* - EUA) e 34,5 µl de água. Seguiu-se a incubação em gelo por 60 minutos. A digestão ocorreu em banho-maria a 37°C por um período de 2 a 12 minutos, com a adição de 5 µl da enzima HindIII (*New England Biolabs* - EUA) em diferentes concentrações, variando de 0,25U a 12U. A interrupção da ação da enzima de restrição foi realizada adicionando-se 5 µl de EDTA 0,5 M (pH 8.0).

Com o auxílio de um micropipetador, as amostras foram colocadas em um gel de agarose 1%. A eletroforese em campo pulsado foi realizada de acordo com os seguintes parâmetros: 6V/cm, 120°, 50s-50s por 18 horas em tampão TAE 1X a 14°C.

Após 18 horas da corrida de eletroforese em campo pulsado, a seleção dos fragmentos de DNA de interesse (100 a 300 Kb) foi realizada a partir da coloração com brometo de etídio das laterais do gel contendo o marcador molecular Lambda (*New England Biolabs* - EUA). As regiões coradas indicaram a região do gel com os fragmentos de DNA de interesse, a qual foi excisada e submetida ao processo de digestão da agarose para a liberação dos fragmentos de DNA para a clonagem. A Figura 5 ilustra de maneira esquemática o processo.



**Figura 5.** Partes do gel submetido à PFGE que devem ser recortadas e coradas para guiar a retirada da fatia do gel contendo os fragmentos de DNA com tamanho desejado.

## 6. Purificação dos fragmentos de DNA utilizando a enzima beta-agarase

A purificação dos fragmentos de DNA de interesse a partir do gel de agarose foi realizada utilizando-se a enzima beta-agarase (GELase<sup>TM</sup> - *Epicentre*) que digere a agarose produzindo moléculas de carboidratos, os quais não interferem na reação de ligação dos fragmentos de DNA com o vetor.

Os procedimentos de digestão da agarose pela enzima beta-agarase, foram realizados de acordo com as indicações do fabricante, os quais estão descritos resumidamente a seguir. A fatia do gel contendo os fragmentos de DNA de interesse foi transferida para um microtubo, o qual foi pesado para o cálculo do volume da enzima e do tampão a serem adicionados. Em seguida, o microtubo foi levado ao banho-maria a 68°C por 5 minutos para o derretimento da agarose. Posteriormente adicionou-se o tampão da enzima beta-agarase 1X em um volume correspondente a 3X o peso da fatia do gel. O tubo foi transferido imediatamente para o banho-maria a 45°C por 5 minutos para equilíbrio da temperatura antes da adição da enzima beta-agarase. Por fim, adicionou-se 1

$\mu\text{L}$  da enzima beta-agarase (1U/ $\mu\text{L}$ ) para cada 300 mg de agarose, incubando-se em banho-maria a 45°C por 30 minutos.

Após o processo de digestão da agarose, a concentração do DNA foi estimada submetendo-se 10  $\mu\text{L}$  da solução obtida em gel de agarose 1%, contendo uma diluição seriada de DNA lambda (*Promega Corporation* - EUA) de 10 a 50 ng. Esse gel foi submetido à eletroforese horizontal em tampão TBE 0,5X a 80V por 30 minutos.

### **7. Preparação do vetor de clonagem pBeloBAC11**

O vetor de clonagem pBeloBAC11 é uma molécula de DNA circular. Portanto, deve ser clivada e desfosforizada para poder se ligar ao fragmento de DNA de interesse, formando assim uma molécula de DNA recombinante (vetor + inserto de DNA). O vetor utilizado no presente trabalho foi doado pela Dra. Clare Gill (Department of Animal Science - Texas A&M University - EUA) como parte da colaboração estabelecida.

Foram utilizadas 50 unidades da enzima HindIII (*New England Biolabs* - EUA) para cortar 1  $\mu\text{g}$  do vetor pBeloBAC11. A reação de digestão do vetor foi realizada em um volume total de 100  $\mu\text{L}$ , sendo 2,5  $\mu\text{L}$  de DNA (400 ng/ $\mu\text{L}$ ), 2,5  $\mu\text{L}$  de HindIII (20 U/ $\mu\text{L}$ ), 10  $\mu\text{L}$  de 10X tampão fosfatase ácida e 85  $\mu\text{L}$  de água ultra pura. A reação de digestão foi incubada a 37°C por 6 horas, seguida de uma incubação a 65°C por 20 minutos para desativar a enzima HindIII. Posteriormente foi adicionado  $\text{CaCl}_2$  fornecido com a fosfatase alcalina em uma concentração final de 5 mM, sendo incubado a 30°C por 5 minutos.

A reação de desfosforização do vetor pBeloBAC11 foi realizada em um volume total de 108  $\mu\text{L}$ , sendo 100  $\mu\text{L}$  do vetor de clonagem, 5  $\mu\text{L}$  de  $\text{CaCl}_2$  e 3  $\mu\text{L}$  de fosfatase alcalina. Essa reação foi incubada a 30°C por 2 horas, seguida de incubação a 65°C por 20 minutos para desativar a fosfatase alcalina.

Em seguida a concentração do vetor foi determinada por eletroforese horizontal em gel de agarose 1% com tampão TBE 0,5X a 80V por 30 minutos. Por fim, o vetor foi diluído em água ultrapura para a concentração de uso (10 ng/ $\mu$ L) e armazenado em freezer (-20°C).

## **8. Clonagem**

O processo de clonagem envolveu duas etapas. A primeira inclui a ligação do fragmento de DNA bubalino com o vetor de clonagem pBeloBAC11. A segunda é o processo de transformação bacteriana, ou seja, a entrada do vetor contendo o DNA bubalino nas células eletrocompetentes por meio de eletroporação.

### 8.1. Reação de ligação do vetor de clonagem pBeloBAC11 com o fragmento de DNA bubalino

A reação de ligação foi realizada em um volume total de 60  $\mu$ L, sendo 1  $\mu$ L do vetor pBeloBAC11 (10 ng/ $\mu$ L), 50  $\mu$ L dos fragmentos de DNA bubalino (2 ng/ $\mu$ L), 0,6  $\mu$ L da enzima T4 DNA ligase (*Promega Corporation* - EUA), 6  $\mu$ L do tampão da enzima T4 DNA ligase e 2,4  $\mu$ L de água ultrapura. A reação de ligação ocorreu a temperatura ambiente por 12 horas, sendo posteriormente armazenada em freezer (-20°C).

### 8.2. Transformação bacteriana por eletroporação

Para o processo de transformação foi utilizado 1,5  $\mu$ L do produto de ligação e 20  $\mu$ L da bactéria eletrocompetente ElectroMax<sup>TM</sup> DH10B<sup>TM</sup> (*Invitrogen Corporation* - EUA), os quais foram adicionados a uma cubeta de eletroporação com 1 mm de espaçamento e submetidos à eletroporação com 2000 V no aparelho Electroporator 2510 (*Eppendorf-*

EUA). Após a eletroporação, as células foram imediatamente transferidas para um tubo de fundo U, contendo 1 mL de meio SOC e levadas ao agitador a 37°C por 50 minutos a 200 rpm.

Em seguida, 200 µL da cultura foi distribuída em placas com meio LB-ágar acrescido de antibiótico cloranfenicol (12,5 mg/mL – *SERVA* - Alemanha), 20 µL de IPTG a 20% (Isopropil-β-D-tiogalactosídeo - *Eppendorf* - EUA) e 100 µL de XGal a 2% (5-bromo-4-cloro-3-indolil-β-D-galactopiranosídeo - *Invitrogen Corporation* - EUA), sendo levados a estufa a 37°C por 12 horas.

### **9. Preparação dos BACs para verificação dos tamanhos dos fragmentos de DNA bubalino**

Cada clone foi colocado em um tubo de fundo cônico contendo 5 mL de meio LB acrescido de cloranfenicol (12,5 mg/mL), sendo levados ao agitador a 37°C, 200 rpm por 18 horas para propagação celular. Após este período, foram então centrifugados a 3000 rpm por 15 minutos a 4°C, sendo descartado o sobrenadante. O concentrado de células foi ressuspenso em 200 µL de solução de lisoenzima com auxílio de uma pipeta descartável e transferido para um microtubo de 1,5 mL, o qual foi mantido no gelo por 5 minutos.

Em seguida, foram adicionados 400 µL de solução de NaOH-SDS e novamente incubados em gelo por 5 minutos. Foram adicionados 300 µL de solução de acetato de potássio e incubadas a -80°C por 15 minutos, e em seguida à temperatura ambiente por 15 minutos. Seguiu-se com centrifugação a 12000 rpm por 15 minutos à temperatura de 4°C. O sobrenadante foi transferido para novos microtubos, onde foram adicionados 600 µL de isopropanol gelado (-20°C), mantendo-se os mesmos a -80°C por 20 minutos. Uma nova centrifugação foi realizada a 12000 rpm por 20 minutos a 4°C. Por fim, o

sobrenadante foi removido e o concentrado de BAC foi ressuspendido em 1 mL de etanol 70% gelado (-20°C), seguido de centrifugação por 5 minutos a 12000 rpm a 4°C. Os microtubos foram colocados para secar em temperatura ambiente. Depois de secos, os BACs foram ressuspendidos em 30 µL de TE (Tris-EDTA pH 8.0) e seguiu-se a digestão do vetor BAC para a liberação do inserto.

A reação de digestão do BAC para a liberação do inserto de DNA foi realizada em um volume total de 21 µL, sendo 10 µL do BAC\_DNA; 3 µL do tampão da enzima NotI 10X (*New England Biolabs* - EUA); 0,3 µL de 100X BSA; 0,5 µL da enzima Not I (10U/ µL - *New England Biolabs* - EUA) e 17,2 µL de água. A reação ocorreu em banho-maria a 37°C por 2 horas. O produto de digestão foi submetido a eletroforese de campo pulsado em gel de agarose 1% seguindo-se os parâmetros de: 6V/cm, 120°, 1,4s-13,5s por 16 horas em tampão TBE 0,5X a 14°C. O gel foi corado com brometo de etídio e fotodocumentado com câmera digital (DC290 Kodak™).

A partir da avaliação do tamanho dos insertos de parte dos clones, foi calculada a média de tamanho de inserto na biblioteca BAC. Também foi calculada a cobertura do genoma, utilizando a fórmula  $C = (N \times I)/G$ , onde a cobertura do genoma (C) depende do número de clones (N), da média do tamanho dos insertos da biblioteca (I), além do tamanho do genoma da espécie (G).

## **10. Congelamento e armazenamento dos clones**

Após a obtenção do tamanho dos insertos de DNA desejado, os clones foram retirados da placa com meio sólido, um a um, com auxílio de um palito de dente, e depositados em placas de 384 poços contendo 70 µL de uma solução composta por meio LB acrescido de uma solução de congelamento (36 mM de K<sub>2</sub>HPO<sub>4</sub>, 13,2 mM de KH<sub>2</sub>PO<sub>4</sub>,

1,7 mM de  $\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$ , 0,4 mM de  $\text{MgSO}_4$ , 68 mM de  $(\text{NH}_4)_2\text{SO}_4$  e 4,4% de glicerol) na proporção 9:1.

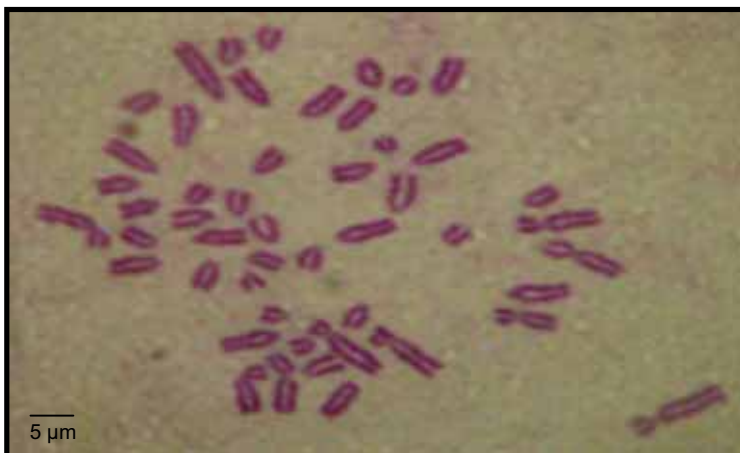
As placas de 384 poços contendo os clones foram então levadas à estufa a 37°C por 6 horas para propagação dos clones. Em seguida, foram transferidas para o freezer -80°C para armazenamento.

## Resultados e Discussão

### 1. Avaliação da integridade do genoma

Os resultados obtidos a partir das análises citogenéticas permitiram avaliar o cariótipo quanto à ausência de alterações cromossômicas numéricas e/ou estruturais.

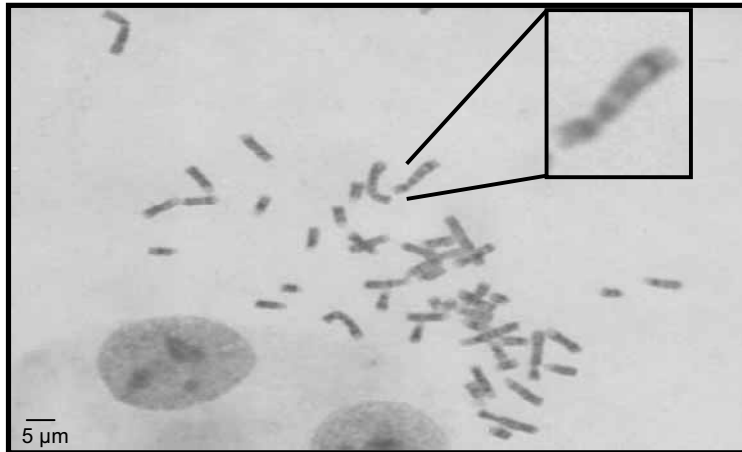
A partir da coloração com Giemsa observou-se a presença de um cariótipo típico da espécie *Bubalus bubalis* ( $2n=50$ ), composto por cinco pares de cromossomos autossomos submetacêntricos, 19 pares de cromossomos autossomos acrocêntricos e um par de cromossomos sexuais (IANNUZZI, 1994). A Figura 6 ilustra uma metáfase corada com uma solução de Giemsa 6%.



**Figura 6.** Fotomicrografia óptica de uma metáfase obtida a partir de um animal macho (*Bubalus bubalis*,  $2n=50$ ) com um aumento de 1000X, corada com Giemsa, mostrando o cariótipo padrão da espécie,

As metáfases submetidas à técnica de bandamento G apresentaram cromossomos com bandas distintas e contrastantes entre regiões claras e escuras, e com um mínimo de distorção da morfologia cromossômica, como pode ser observado na Figura 7. Optou-se por fotografar as metáfases em preto e branco para facilitar a identificação do padrão de bandas obtido. A partir do bandamento G, quando comparado com o bandamento G

padrão da espécie descrito por IANNUZZI em 1994, não foi detectada deleção e/ou duplicação de material genético.



**Figura 7.** Fotomicrografia óptica de uma metáfase obtida a partir de um animal macho (*Bubalus bubalis*,  $2n=50$ ) submetida à técnica de bandamento G com um aumento de 1000X. No detalhe observa-se um cromossomo acrocêntrico em maior aumento, evidenciando o padrão do bandamento G correspondente ao descrito por IANNUZZI em 1994.

Como descrito por MIYAKI e colaboradores (1980), resultados de bandamento C em *Bubalus bubalis* somente podem ser observados nos cromossomos acrocêntricos, uma vez que esta técnica revela somente heterocromatina constitutiva. O fato dos cromossomos submetacêntricos não apresentarem as bandas C pode ser explicado pela provável perda de heterocromatina constitutiva durante o processo de fusão cêntrica na formação dos cromossomos submetacêntricos (IANNUZZI, 1994).

Como pode ser observado na Figura 8, os cromossomos mantiveram sua morfologia íntegra, apresentando as regiões centroméricas e de constrição secundária muito coradas, contrastando com o restante do cromossomo.



**Figura 8.** Fotomicrografia óptica de uma metáfase obtida a partir de um animal macho (*Bubalus bubalis*,  $2n=50$ ), com um aumento de 1000X, submetida à técnica de bandamento C, cujos cromossomos acrocêntricos apresentam a região centromérica fortemente corada.

A coloração Ag-NOR evidenciou pontos escuros em determinados cromossomos indicando a localização de um tipo particular de DNA, o DNA ribossômico (rDNA), que codifica o RNA ribossômico (rRNA), principal componente dos ribossomos.

O rDNA caracteriza-se por sua natureza repetitiva, sendo constituído por seqüências ou unidades de repetições em regiões específicas de determinados cromossomos. As regiões que apresentam esse DNA são denominadas de regiões organizadoras de nucléolos (NORs). Apenas as NORs que estiverem ativas na interfase podem ser visualizadas por essa coloração, que por intermédio do nitrato de prata, impregna às proteínas não histônicas remanescentes da atividade transcricional dessas regiões, contrastando com a cromatina fracamente corada no restante do cromossomo.

A Figura 9 ilustra uma metáfase submetida ao tratamento Ag-NOR, onde podem ser observadas as marcações nos cromossomos 3, 19, 22, 25, 26 e 28, que estão de acordo com o padrão estabelecido para a espécie *Bubalus bubalis* (IANNUZZI et al., 1996).

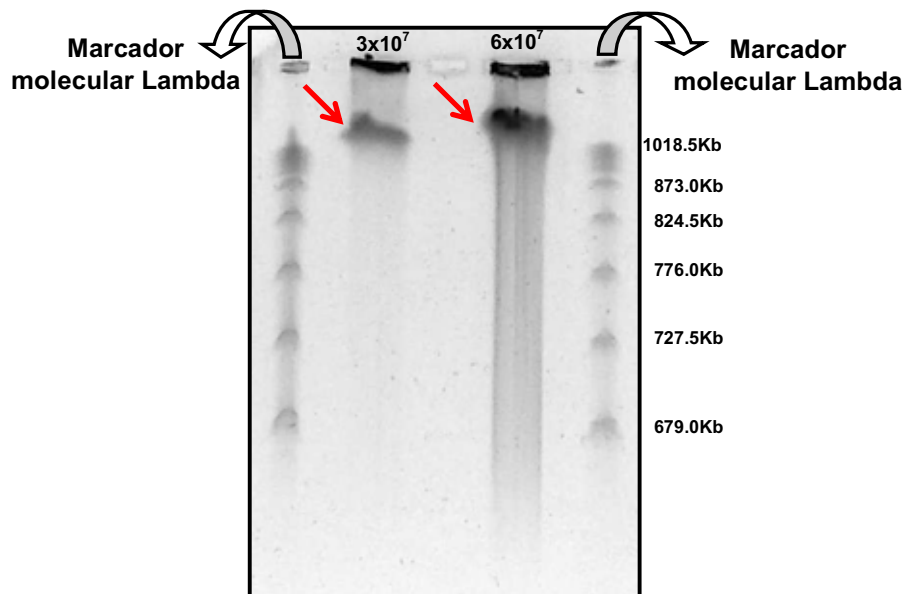


**Figura 9.** Fotomicrografia óptica de uma metáfase obtida a partir de um animal macho (*Bubalus bubalis*,  $2n=50$ ), com um aumento de 1000X, submetida ao tratamento Ag-NOR, cujos cromossomos 3, 19, 22, 25, 26 e 28 apresentam as regiões organizadoras nucleolares. A seta aponta para uma dessas regiões organizadoras nucleolares.

## 2. Construção da biblioteca BAC

A partir da contagem de células utilizando-se a câmara de Neubauer e posterior inclusão das mesmas em agarose com baixo ponto de fusão, foram produzidos 96 blocos de agarose com concentração celular de  $3 \times 10^7$  células/mL e 58 blocos com  $6 \times 10^7$  células/mL, os quais foram armazenados em Tris-EDTA (pH 8.0) a  $4^\circ\text{C}$ .

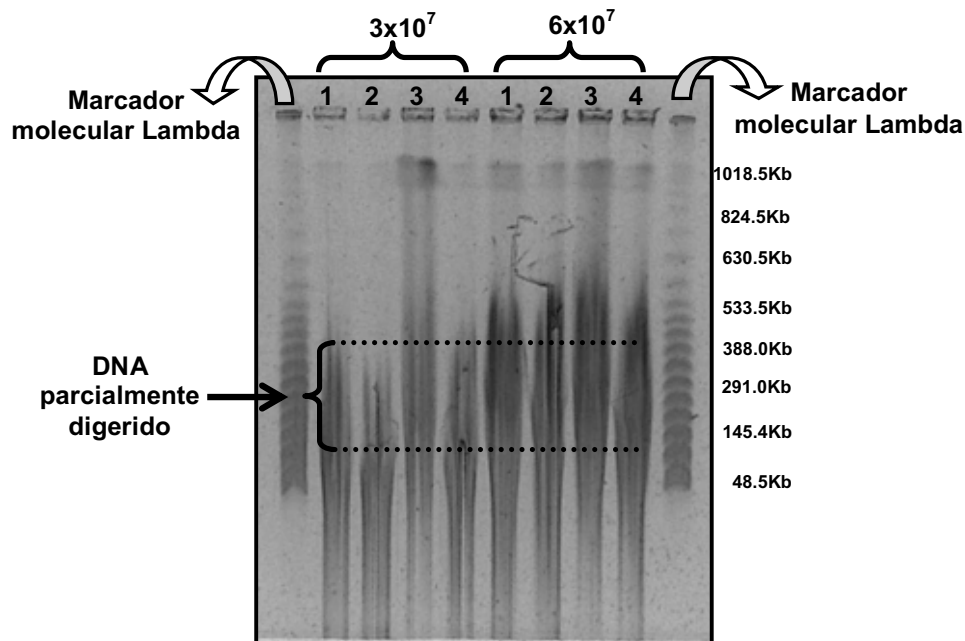
As células contidas nos blocos foram submetidas ao tratamento com proteinase K para lise da membrana nuclear e degradação das proteínas. A Figura 10 ilustra um gel de agarose 1% corado com brometo de etídio, submetido à PFGE, mostrando a qualidade do DNA nos blocos de agarose com concentrações de  $3 \times 10^7$  e  $6 \times 10^7$  células/mL, após o tratamento com proteinase K. Pode ser observada a presença de um DNA íntegro indicando que ambas as concentrações celulares apresentaram a qualidade exigida para a construção de uma biblioteca BAC.



**Figura 10.** Gel de agarose 1% corado com brometo de etídio e submetido à PFGE, ilustrando a qualidade das células nos blocos de agarose com as diferentes concentrações celulares ( $3 \times 10^7$  células/mL e  $6 \times 10^7$  células/mL) após os tratamentos com proteinase K. Marcador molecular lambda (*New England Biolabs* - EUA). As setas vermelhas indicam o DNA íntegro.

Após verificada a qualidade do DNA dos blocos de agarose, realizou-se o teste de digestão parcial do DNA com diferentes concentrações da enzima de restrição HindIII (de 0,25U a 12U) para definir a concentração ideal da mesma. Para os blocos de agarose com concentração de  $3 \times 10^7$  células/mL a concentração ideal de HindIII foi de 0,75U por 5 minutos a  $37^\circ\text{C}$ . Já para os blocos de agarose com concentração de  $6 \times 10^7$  células/mL, a concentração ideal de HindIII foi de 1,5U também por 5 minutos.

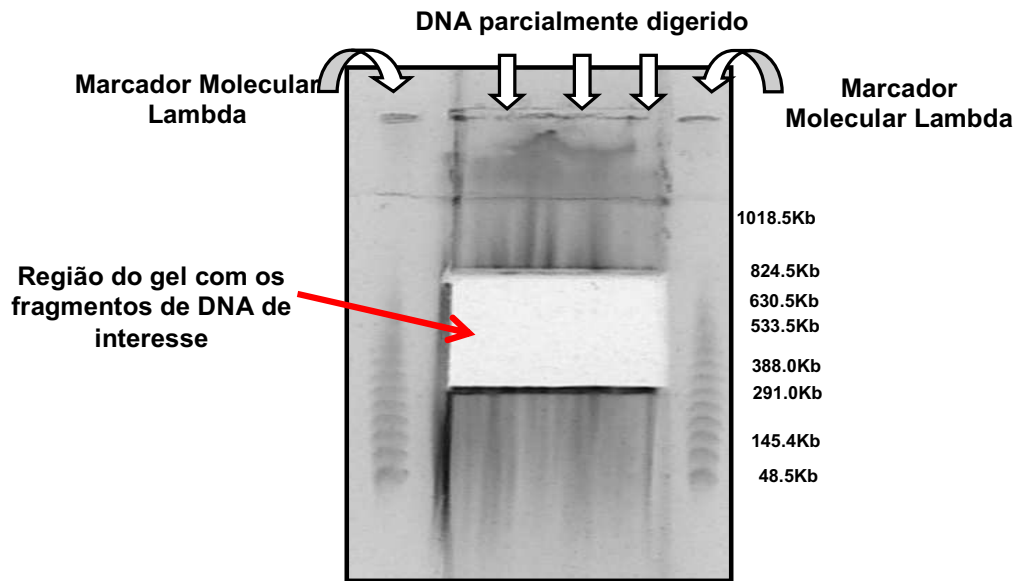
A Figura 11 ilustra um gel de agarose 1% corado com brometo de etídio, submetido à PFGE, mostrando os resultados obtidos com diferentes concentrações da enzima de restrição HindIII e diferentes tempos de digestão.



**Figura 11.** Gel de agarose 1% corado com brometo de etídio, submetido à PFGE, ilustrando um dos testes de digestão parcial realizada com diferentes concentrações da enzima HindIII nos blocos de agarose com concentração de  $3$  e  $6 \times 10^7$  células/ml. **3 e  $6 \times 10^7$  células/ml:** **1.** 0,25U de HindIII por 10 minutos a  $37^\circ\text{C}$ ; **2.** 0,25U de HindIII por 12 minutos a  $37^\circ\text{C}$ ; **3.** 0,5U de HindIII por 3 minutos a  $37^\circ\text{C}$ ; **4.** 0,5U de HindIII por 5 minutos a  $37^\circ\text{C}$ . Marcador molecular lambda (New England Biolabs - EUA).

Com a determinação da concentração ideal da enzima HindIII para as duas concentrações celulares ( $3$  e  $6 \times 10^7$  células/mL), seguiu-se a digestão do DNA em larga escala para posterior isolamento e recuperação dos fragmentos para o processo de clonagem.

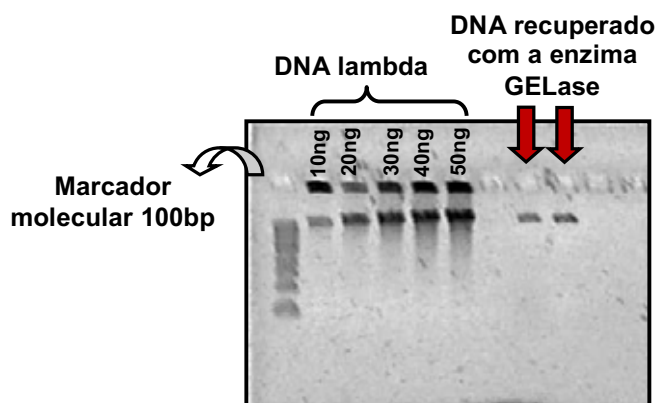
A Figura 12 ilustra um gel de agarose a 1%, corado com brometo de etídio, contendo o DNA proveniente de dois blocos de agarose com  $3 \times 10^7$  células/mL, parcialmente digeridos com 0,75U da enzima HindIII por 5 minutos a  $37^\circ\text{C}$ , submetido à PFGE.



**Figura 12.** Gel de agarose de baixo ponto de fusão a 1% corado com brometo de etídio, submetido à PFGE, ilustrando o resultado da digestão parcial em larga escala com 0,75U de HindIII por 5 minutos de dois blocos de agarose com  $3 \times 10^7$  células/ml. Marcador molecular lambda (*New England Biolabs* - EUA).

A fatia excisada do gel contendo os fragmentos de DNA de interesse foi submetida à ação da enzima beta-agarase para a digestão da agarose com baixo ponto de fusão. A concentração do DNA recuperado foi estimada em 2 ng/ $\mu$ L por meio de gel de agarose contendo diluição seriada de DNA lambda e submetido à eletroforese horizontal.

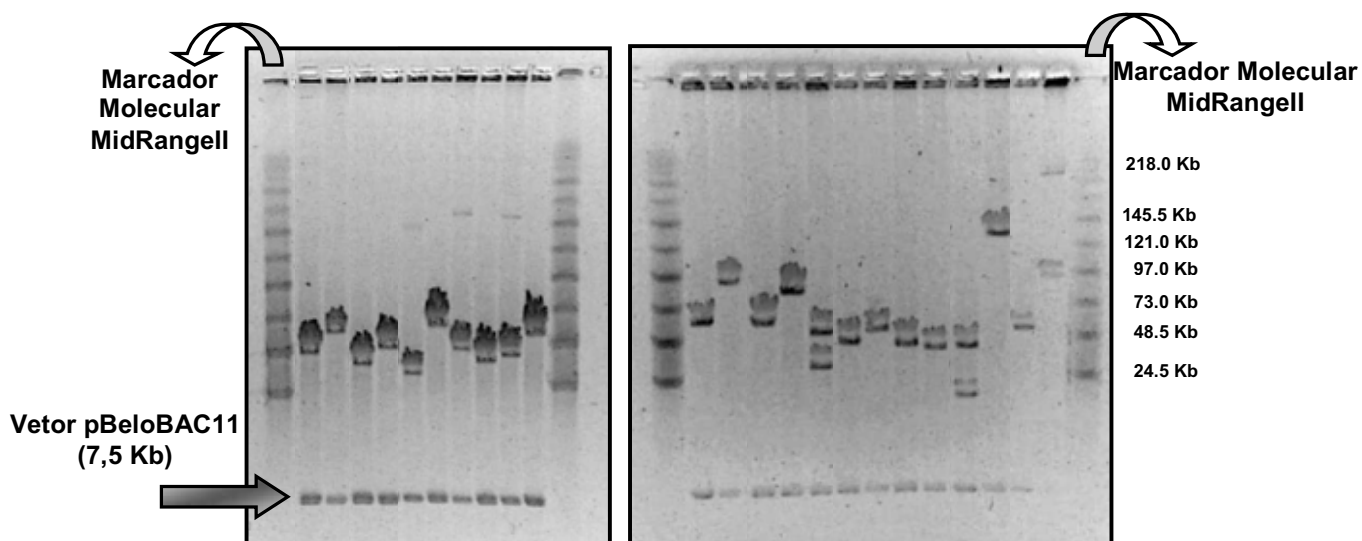
A Figura 13 ilustra um gel de agarose 2%, submetido à eletroforese horizontal e corado com brometo de etídio, onde pode ser observada a qualidade do DNA recuperado pela ação da enzima beta-agarase.



**Figura 13.** Gel de agarose 1% corado com brometo de etídio, submetido à eletroforese horizontal, ilustrando a recuperação do DNA parcialmente digerido após ação da enzima beta-agarase dos fragmentos de gel dos blocos de  $3 \times 10^7$  células/ml. O DNA lambda padrão (*Promega Corporation* - EUA) foi diluído nas concentrações de 10, 20, 30, 40 e 50ng, respectivamente.

As reações de ligação do DNA bubalino com o vetor pBeloBAC11 foram realizadas utilizando uma proporção de 8:1 entre vetor e DNA. Após a reação de ligação, seguiu-se com a transformação por eletroporação das bactérias comerciais ElectroMAX™ DH10B™ (*Invitrogen Corporation* - EUA), as quais foram espalhadas em placas com meio LB-ágar acrescido do antibiótico cloranfenicol (*Serva* - Alemanha), XGal (*Invitrogen Corporation* - EUA) e IPTG (*Eppendorf* - EUA).

Cerca de 50 mil clones foram gerados, dos quais 240 foram analisados quanto ao tamanho do inserto de DNA, revelando uma média de tamanho de inserto de 92 Kb, variando de 50 a 300 Kb. A Figura 14 ilustra um gel de agarose 1% submetido à PFGE, ilustrando o tamanho dos insertos de DNA contidos em um conjunto de 23 clones da biblioteca.



**Figura 14.** Gel de agarose 1%, submetido à PFGE e corado com brometo de etídio, ilustrando os tamanhos de insertos de DNA presente em 23 clones da biblioteca BAC. Marcador molecular MidRangell (*New England Biolabs* - EUA).

Os dados obtidos quanto ao tamanho dos insertos da biblioteca BAC de búfalo estão de acordo com o descrito para as bibliotecas construídas para bovino: 103 Kb (FUJISAKI et al., 2002), 104 Kb (BUIKAMP et al., 2000) e 105 Kb (ZHU et al., 1999). Quando comparado com bibliotecas construídas para outros mamíferos, nossos dados continuam dentro do tamanho de inserto esperado, como por exemplo, a biblioteca de cavalo apresentou uma média de tamanho de inserto de 110 Kb (GODARD et al., 1998), de suíno de 116 Kb (LIU et al., 2010), de ovelha de 123 Kb (VAIMAN et al., 1999) e de urso panda apresentou uma média de inserto de 97 Kb (ZENG et al., 2007).

Com o tamanho do genoma haplóide do búfalo estimado em três bilhões de nucleotídeos, a cobertura da biblioteca BAC, gerada com 50 mil clones e tamanho médio de inserto de 92 Kb, foi estimada em 1,5X.

## Conclusões

Com a construção da biblioteca BAC, agora está disponível para a comunidade científica do Brasil e do exterior, mais uma ferramenta para estudos do genoma do búfalo, denominada BBU\_LGC. Essa biblioteca se encontra disponível no Laboratório de Genômica Comparativa, IBILCE-UNESP, Campus de São José do Rio Preto-SP.

O presente trabalho estabeleceu todos os parâmetros de obtenção, manutenção e manipulação do DNA-HWM, muitos dos quais se mostraram exclusivos para a espécie *Bubalus bubalis*, além daqueles parâmetros envolvidos na clonagem em larga escala, tornando possível a geração de clones com tamanho de inserto apropriado para uma cobertura do genoma satisfatória com a biblioteca.

A biblioteca BAC de búfalo possui 50 mil clones com um tamanho médio de inserto de 92 Kb, o que fornece uma cobertura do genoma de aproximadamente 1,5X.

A disponibilidade dessa nova ferramenta genômica para a espécie *B. bubalis* permitirá um novo direcionamento nos estudos moleculares do genoma bubalino, uma vez que a biblioteca BBU\_LGC poderá ser utilizada para o isolamento e caracterização de genes ou grupos de genes a partir de uma região de interesse, no mapeamento físico, na elucidação da organização gênica, na clonagem posicional, no seqüenciamento e re-seqüenciamento de regiões-alvo ou de cromossomos inteiros, além da identificação de elementos reguladores inter e intra-gênicos. Os resultados poderão ser aplicados em um melhor entendimento da conservação e da variação genética dos búfalos nos rebanhos brasileiros, contribuindo para o desenvolvimento de novas estratégias para os programas de melhoramento genético desta importante espécie.

## Referências Bibliográficas\*

ALBERTSEN, H.M. et al. Improved control of partial DNA restriction enzyme digest in agarose using limiting concentrations of Mg<sup>++</sup>. **Nucleic Acids Research**, v. 17, p. 808, 1989.

ALBERTSEN, H.M. et al. Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. **PNAS**, v. 87, p. 4256-4260, 1990.

AMENIYA, C.T.; ZON, L.I. Generation of a zebrafish P1 artificial chromosome library. **Genomics**, v. 58, p. 211-213, 1999.

ANAND, R.; VILLASANTE, A.; TYLER-SMITH, C. Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed-field gel electrophoresis. **Nucleic Acids Research**, v. 17, p. 3425-3433, 1989.

ASAKAWA, S. et al. Human BAC library: construction and rapid screening. **Gene**, v. 191, p. 67-79, 1997.

BECK, T.W. et al. Comparative feline genomics: a BAC/PAC contig map of the major histocompatibility complex class II region. **Genomics**, v. 71, p. 282-295, 2001.

BIRREN, B.W.; LAI, E. **Pulsed Field Gel Electrophoresis: A Practical Guide**. Elsevier Science & Technology Books, p. 129-140, 1993.

BIRREN, B.; MANCINO, V.; SHIZUYA, H. Bacterial Artificial Chromosomes. **Genome analysis: a laboratory manual**, Cold Spring Harbor Laboratory Press, v. 3, p. 241-295, 1997

BUITKAMP, J. et al. Construction and characterization of a gridded cattle BAC library. **Animal Genetics**, v. 31, p. 347-351, 2000.

BURKE, D.T.; OLSON, M.V. Preparation of clone libraries in yeast artificial-chromosome vectors. **Methods in Enzymology**, v. 194, p. 251-270, 1991.

CAI, L. et al. Construction and characterization of a bovine bacterial artificial chromosome library. **Genomics**, v. 29, p. 413-425, 1995.

CHEN, R. et al. Dynamic building of a BAC clone tiling path for the Rat Genome Sequencing Project. **Genome Research**, v. 14, p. 679-684, 2004.

CHIU, C.H. et al. Bichir *HoxA* cluster sequence reveals surprising trends in ray-finned fish genomic evolution. **Genome Research**, v. 14, p. 11-17, 2004.

\*Associação Brasileira de Normas Técnicas (ABNT). NBR6023:2002, Informação e documentação - Referências - Elaboração, Rio de Janeiro, Agosto de 2002.

CLARKE, L.; CARBON, J. A colony bank containing synthetic Col EI hybrid plasmids representative of the entire *E. coli* genome. **Cell**, v. 9, p. 91-99, 1976.

DONOVAN, A. et al. Positional cloning of zebrafish ferroportin 1: identifies a conserved vertebrate iron exporter. **Nature**, v. 403, p. 776-781, 2000.

EGGEN, A. et al. Construction and characterization of a bovine BAC library with four genome-equivalent coverage. **Genetics, Selection, Evolution**, v. 33, p. 543-548, 2001.

FAHRENKRUG, S.C. et al. A porcine BAC library with tenfold genome coverage: a resource for physical and genetic map integration. **Mammalian Genome**, v. 12, p. 472-474, 2001.

FRESHNEY, R.I. **Culture of Animal Cells**. Ed. Willey Liss, n.4, 557p., 2000.

FUJISAKI, S. et al. Construction of a bovine bacterial artificial chromosome library from fibroblasts used for cloned cattle. **Animal Genetics**, v. 33, p. 379-381, 2002.

GARDINER, K. Pulsed field gel electrophoresis. **Analytical Chemistry**, v. 63, p. 658-665, 1991.

GILL, C.A. et al. Construction and characterization of an ovine bacterial artificial chromosome library. **Mammalian Genome**, v. 10, p. 1108-1111, 1999.

GODARD, S. et al. Construction of a horse BAC library and cytogenetical assignment of 20 type I and type II markers. **Mammalian Genome**, v. 9, p. 633-637, 1998.

IANNUZZI, L. Standard karyotype of the river buffalo (*Bubalus bubalis* L., 2n = 50). Report of the committee for the standardization of banded karyotypes of the river buffalo. **Cytogenetics and Cell Genetics**, v.67, p.102-113, 1994.

IANNUZZI, L.; DI MEO, G.P.; PERUCATTI, A. Identification of nucleolus organizer chromosomes and frequency of active NORs in river buffalo (*Bubalus bubalis* L.). **Caryologia**, v. 49, p. 27-34, 1996.

ISHII, Y. et al. Construction of BAC library for the amphibian *Xenopus tropicalis*. **Genes & Genetic Systems**, v. 79, p. 49-51, 2004.

JEON, J.T. et al. A large-insert porcine library with sevenfold genome coverage: a tool for positional cloning of candidate genes for major quantitative traits. **Molecules and Cells**, v. 16, p. 113-116, 2003.

KIM, U.J. et al. Construction and characterization of a human bacterial artificial chromosome library. **Genomics**, v. 34, p. 213-218, 1996.

KRZYWINSKI, M. et al. Integrated and sequence-ordered BAC- and YAC-based physical maps for the rat genome. **Genome Research**, v. 14, p. 765-779, 2004.

LAI, E.; BIRREN, B.W. Use a secondary pulse field gel electrophoresis in separation of large DNA. **Analytical Biochemistry**, v. 224, p. 68-74, 1995.

LARIN, Z.; MONACO, A.P.; LEHRACH, H. Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. **PNAS**, v. 88, p. 4123-4127, 1991.

LI, R. et al. Construction and characterization of an eightfold redundant dog genomic bacterial artificial chromosome library. **Genomics**, v. 15, p. 9-17, 1999.

LIBERT, F. et al. Construction of a bovine genomic library of large yeast artificial chromosome clones. **Genomics**, v. 18, p. 270-276, 1993.

LIU, L. et al. Construction of a bacterial artificial chromosome library for the Rongchang pig breed and its use for the identification of genes involved in intramuscular fat deposition. **Biochemical and Biophysical Research Communications**, v. 391, p. 1280-1284, 2010.

MCCUBBIN, A.G.; ROALSON, E.H. Construction of bacterial artificial chromosome libraries for use in phylogenetic studies. **Methods in Enzymology**, v. 395, p. 384-400, 2005.

MIYAKI, Y.; KANAGAWA, H.; ISHIKAWA, T. A chromosomal analysis based on the G and C band staining techniques of the buffalo (*Bubalus bubalis*). **The Japanese Journal of Veterinary Research**, v. 28, p. 122-128, 1980.

OSOEGAWA, K. et al. Bacterial artificial chromosome libraries for mouse sequencing and functional analysis. **Genome Research**, v. 10, p. 116-128, 2000.

PATERSON, A.H. The DNA revolution. **Genome Mapping in Plants**. Academic Press, San Diego, CA, p. 7-21, 1996.

PETERSON, D.G. et al. Construction of plant bacterial artificial chromosome (BAC) libraries: an illustrated guide. **Journal of Agricultural Genomics**, v.5, p. 1-100, 2000.

ROGEL-GAILLARD, C. et al. Construction of a rabbit bacterial artificial chromosome (BAC) library: application to the mapping of the major histocompatibility complex to position 12q.1.1. **Mammalian Genome**, v. 12, p. 253-255, 2001.

SAZANOV, A.A. et al. Characterization of the CHORI-240 BAC clones containing the bovine CSN1S1, CSN2, STATH, CSN1S2 and CSN3 genes. **Journal of Applied Genetics**, v. 47, p. 243-245, 2006.

SCHIBLER, L. et al. Construction and extensive characterization of a goat bacterial artificial chromosome library with threefold genome coverage. **Mammalian Genome**, v. 9, p. 119-124, 1998.

SCHWARTZ, D.C.; CANTOR, C.R. Separation of yeast chromosome-sized DNA by pulse field gradient gel electrophoresis. **Cell**, v. 37, p. 67-75, 1984.

STRONG, S.J. et al. A novel multigene family encodes diversified variable regions. **PNAS**, v. 96, p. 15080-15085, 1999.

SUMIYAMA, K.; RUDDLE, F.H. Regulation of Dix3-7 gene expression in visceral arches by evolutionary conserved enhancer elements. **PNAS**, v. 100, p. 4030-4034, 2003.

SUZUKI, K. et al. Construction and evaluation of a porcine bacterial artificial chromosome library. **Animal Genetics**, v. 31, p. 8-12, 2000.

THORSEN, J. et al. A highly redundant BAC library of Atlantic salmon (*Salmo salar*): an important tool for salmon projects. **BMC Genomics**, v. 6, p.50, 2005.

VAIMAN, D. et al. Construction and characterization of a sheep BAC library of three genome equivalents. **Mammalian Genome**, v. 10, p. 585-587, 1999.

WARREN, W. et al. Construction and characterization of a new bovine bacterial artificial chromosome library with 10 genome-equivalent coverage. **Mammalian Genome**, v. 11, p. 662-663, 2000.

WOO, S.S. et al. Construction and characterization of a bacterial artificial chromosome library of Sorghum bicolor. **Nucleic Acids Research**, v. 22, p. 4922-4931, 1994.

ZENG, C. et al. Giant panda BAC library construction and assembly of a 650-kb contig spanning major histocompatibility complex class II region. **BMC Genomics**, v. 8, 2007.

ZHANG , H.B. Construction and manipulation of large-insert bacterial clone libraries. **Manual**, Texas A&M University, TX, USA, 2000.

ZHU, B. et al. A 5x genome coverage bovine BAC library: production, characterization, and distribution. **Mammalian Genome**, v. 10, p. 706-709, 1999.

ZIMMER, R.; VERRINDER GIBBINS, A.M. Construction and characterization of a large-fragment chicken bacterial artificial chromosome library. **Genomics**, v. 42, p. 217-226, 1997.

## Capítulo 3

Avaliação de marcadores de DNA  
do genoma bovino para seleção de clones da  
biblioteca BAC de búfalo



### **CAPÍTULO 3 – AVALIAÇÃO DE MARCADORES DE DNA DO GENOMA BOVINO PARA SELEÇÃO DE CLONES DA BIBLIOTECA BAC DE BÚFALO**

**RESUMO:** Com a disponibilidade de uma biblioteca BAC de búfalo, surge a necessidade de marcadores de DNA para resgate de clones específicos da biblioteca. Diversos trabalhos têm mostrado que marcadores provenientes do genoma bovino podem ser utilizados com sucesso em estudos do genoma bubalino. Desse modo, o presente trabalho avaliou 58 iniciadores para PCR provenientes de estudos do genoma bovino, os quais amplificam genes relacionados com produção e qualidade de leite (6), genes relacionados com resposta imune e adaptativa (43) e genes pertencentes à família das lipocalinas (9) os quais desempenham uma variedade de funções destacando-se aquelas relacionadas com lactação e características de reprodução. Todos os 58 iniciadores para PCR foram submetidos à reações de PCR com DNA bubalino, os quais tiveram suas temperaturas de anelamento determinadas. Uma vez otimizados com DNA bubalino, podem ser considerados marcadores do genoma desta espécie, podendo ser utilizados em diversos tipos de estudos como, por exemplo, o mapeamento e a caracterização desses genes utilizando a biblioteca BAC. Devido ao grande número de marcadores do MHC que se mostraram adequados para estudos do genoma bubalino e à importância econômica desses genes, seguiu-se com os experimentos de mapeamento utilizando o painel BBURH<sub>5000</sub>. Foi gerado um mapa com 53 marcadores, o qual foi comparado com o mapa da região MHC de bovino, revelando 33 marcadores em comum. Algumas alterações quanto à ordem dos marcadores foram observadas, os quais sugerem uma rápida evolução desses genes após a divergência do ancestral comum, as quais podem ter surgido devido à diferentes pressões seletivas de parasitas e infecções a quais essas espécies estão sujeitas.

**Palavras-chave:** BBURH<sub>5000</sub>, caseínas, genômica comparativa, lipocalinas, mapeamento, MHC

## Introdução

O búfalo (*Bubalus bubalis*), assim como o bovino, pertence à família Bovidae, a mais diversa das nove famílias da ordem *Artiodactyla*, possuindo 24 pares de cromossomos autossomos (19 acrocêntricos e cinco submetacêntricos) e um par de cromossomos sexuais (2N=50) (IANNUZZI, 1994). O bovino (*Bos taurus*), por sua vez, possui 29 pares de cromossomos autossomos acrocêntricos e um par de cromossomos sexuais (2N=60).

Estudos citogenéticos dos genomas bovino e bubalino estabeleceram que os cinco pares de cromossomos submetacêntricos do búfalo (BBU) são resultantes de translocações por fusões cêntricas entre cromossomos acrocêntricos do genoma bovino (BTA), sendo elas: BBU1 = BTA1/BTA27; BBU2 = BTA2/BTA23; BBU3 = BTA8/BTA19; BBU4 = BTA5/BTA28; BBU5 = BTA16/BTA29 (EL NAHAS et al., 2001; IANNUZZI et al., 2003). Portanto, essas translocações explicam o fato de cada braço dos cromossomos do genoma bubalino apresentar correspondência com os cromossomos acrocêntricos de bovino (EL NAHAS et al., 2001).

Devido a essa homologia cromossômica, é possível utilizar as informações do genoma bovino disponíveis para estudos envolvendo o genoma bubalino, ampliando os conhecimentos acerca deste. Tal estratégia é denominada genômica comparativa.

Mapas genômicos foram construídos para cromossomos específicos do genoma bubalino utilizando iniciadores para PCR derivados do genoma bovino: BBU3 e BBU10 (AMARAL et al., 2007), BBU1 (MIZIARA et al., 2007), BBU6 (STAFUZZA et al., 2007), BBU7 (GOLDAMMER et al., 2007), BBU20 (KOCHAN et al., 2008), BBUX (IANELLA et al., 2008) e BBUY (STAFUZZA et al., 2009). Em 2008, foi construída a primeira geração de mapas para todos os cromossomos do búfalo, onde foram mapeados 2621 marcadores

derivados do genoma bovino, os quais, a partir de então, se tornaram marcadores de DNA também para o genoma bubalino (AMARAL et al., 2008).

Análises comparativas entre os mapas do genoma bovino e bubalino evidenciaram regiões-alvo contendo genes e marcadores moleculares potencialmente associados com características economicamente importantes como, por exemplo, regiões contendo genes relacionados com produção e qualidade do leite, genes da família das lipocalinas e aquelas com genes associados a resposta imune inata e adaptativa. A partir da disponibilidade de uma biblioteca BAC de búfalo, torna-se possível a caracterização molecular dessas regiões do genoma do búfalo.

A seguir estão descritas informações geradas sobre essas regiões-alvo do genoma, assim como os genes envolvidos. Essas serão as primeiras regiões a serem caracterizadas em nível molecular do genoma do búfalo, utilizando os clones da biblioteca BAC.

### **1. Regiões contendo genes relacionados com produção e qualidade de leite**

Ruminantes vêm sendo utilizados para produção de leite há aproximadamente 10.000 anos. Hoje, o leite e seus derivados são os maiores componentes da dieta humana em muitas partes do mundo. A facilidade com que o leite pode ser convertido em uma larga diversidade de produtos é uma característica crucial para a indústria. Nesse aspecto, as proteínas do leite são provavelmente os mais importantes constituintes devido a suas propriedades fundamentais associadas na fabricação de produtos derivados como queijos e iogurtes (HUPPERTZ et al., 2006).

Segundo a Federação Internacional de Laticínios (IDF, 2009), a produção mundial de leite bubalino, em 2008, foi de 87,5 milhões de toneladas, com um crescimento anual de

3,5%, constituindo cerca de 13% do volume total de leite produzido no mundo. No Brasil, a produção de leite de búfalas é uma atividade em ascensão nos últimos anos, particularmente nos estados da região Sudeste, onde o leite destina-se, quase que na sua totalidade, para a produção de queijo tipo *mozzarella*, tendo mercado assegurado e com preços compensatórios (MADELLA-OLIVEIRA et al., 2005).

O leite é constituído por uma combinação de água e diversos elementos sólidos (lipídios, carboidratos, proteínas, sais minerais e vitaminas), cuja composição desses elementos sólidos pode variar de acordo com a raça, o estágio de lactação, a alimentação do animal, a temperatura do ambiente, o manejo e o intervalo entre as ordenhas, e até mesmo a presença ou não de infecções nas glândulas mamárias do animal (AMARAL et al., 2005).

As proteínas do leite são sintetizadas e excretadas pelas células epiteliais das glândulas mamárias durante a fase de lactação, fornecendo elementos nutritivos cruciais para os bezerros, além de proteção imunológica (CLARE & SWAISGOOD, 2000).

As principais proteínas constituintes do leite compreendem quatro caseínas (caseína alfa-s1, caseína alfa-s2, beta-caseína e kapa-caseína), além da alfa-lactoalbumina e da beta-lactoglobulina. Essas proteínas em conjunto representam 95% do total das proteínas presentes no leite, cujos genes representados pelas siglas CSN1S1, CSN1S2, CSN2, CSN3, LALBA e LGB respectivamente, são os responsáveis por codificá-las (THREADGILL & WOMACK, 1990). A nomenclatura oficial para esses genes em bovino, assim como símbolos e nomenclaturas alternativas utilizadas para outros mamíferos estão descritos na Tabela 1.

**Tabela 1.** Nomenclatura dos genes relacionados com produção e qualidade do leite em bovino, com os respectivos símbolos e nomenclaturas alternativas utilizadas em outras espécies de mamíferos.

<b>Símbolo oficial dos gene em boi</b>	<b>Nome oficial em <i>B. taurus</i></b>	<b><i>Bos taurus</i> Gene ID</b>	<b>Símbolos alternativos</b>	<b>Nomes alternativos em outras espécies de mamíferos</b>
CSN1S1	<i>casein alpha s1</i>	ID: 282208	CASA, CSN1, CSNA	<i>alpha S1 casein</i>
CSN2	<i>casein beta</i>	ID: 281099	CASB, CSNB	<i>beta casein B; beta-casein; milk protein</i>
CSN1S2	<i>casein alpha s2</i>	ID: 282209	--	<i>alpha S2 casein</i>
CSN3	<i>casein kappa</i>	ID: 281728	CSN10, CSNK, KCA	<i>k-casein; kappa casein; milk protein; kappa-CN; casoplatelin; casoxin C</i>
LALBA	<i>lactalbumin, alpha-</i>	ID: 281894	a-LACTA, alfaLA	<i>alpha-lactalbumin; alpha-lactalbumin A; alpha-lactalbumin B; lactalbumin, alpha; lactose synthase B protein</i>
LGB	<i>lactoglobulin, beta</i>	ID: 280838	PAEP	<i>beta lactoglobulin D; beta-LG; beta-lactoglobulin; progesterone-associated endometrial protein</i>

Informações disponíveis no banco de dados Entrez Gene (<http://www.ncbi.nlm.nih.gov/gene/>)

As caseínas compreendem cerca de 80% das proteínas presentes no leite e são responsáveis por aumentarem a concentração de cálcio e fosfato, formando micelas com esses minerais (JENG et al., 1997). As caseínas alfa-s1, alfa-s2 e beta-caseína precipitam em baixas concentrações de cálcio, mantendo a estabilidade na micela em suspensão a partir da interação com a proteína kapa-caseína, a qual é responsável pela iniciação das agregações que resultam na formação do coalho (THREADGILL & WOMACK, 1990).

Os genes codificantes das caseínas se encontram organizados no genoma como um agrupamento denominado *cluster*, o qual tem sido bem caracterizado em diversas espécies de mamíferos, compreendendo uma região com tamanho variando de 250 a 350 Kb, dependendo da espécie em questão (RIJNKELS, 2002). Em bovino, os quatro genes das caseínas estão bem descritos, os quais se encontram arranjados no cromossomo 6, na seguinte ordem: CSN1S1, CSN2, CSN1S2 e CSN3 (LEMAY et al., 2009). De acordo com a seqüência do genoma bovino depositada no banco de dados americano NCBI (<http://www.ncbi.nlm.nih.gov/>), o *contig* número NC007304 abrange os quatro genes das

caseínas, onde os genes CSN1S1, CSN2, CSN1S2 e CSN3 apresentam 17536 bp (19 éxons), 9355 bp (9 éxons), 18480 bp (18 éxons) e 13065 bp (5 éxons), respectivamente.

Em búfalo, ainda não se sabe como esses genes estão distribuídos no genoma. Dados de mapeamento do gene CSN3 indicaram sua localização no cromossomo 7 bubalino (BBU7) (GOLDAMMER et al., 2007). Há também, disponíveis no banco de dados do NCBI diversas seqüências parciais dos quatro genes das caseínas de búfalo, as quais em sua grande maioria são seqüências codificantes parciais ou seqüências de éxons únicos, não havendo seqüências genômicas depositadas até o momento. A Tabela 2 apresenta os números de acesso no banco de dados do NCBI de todas as seqüências disponíveis para os quatro genes das caseínas de búfalo.

**Tabela 2.** Número de acesso das seqüências disponíveis no banco de dados do NCBI\* correspondentes aos quatro genes das caseínas (CSN1S1, CSN1S2, CSN2 e CSN3)

<b>Gene</b>	<b>Número de acesso das seqüências disponíveis no banco de dados NCBI para búfalo</b>
CSN1S1	FJ392261; EF025978; EF025979; EF025981; EF025980; EF025982; EF025983
CSN1S2	FM865619; FM865618; FM865621; FM865620; EF025986; EF025985; EF025984; EF025989; EF025988; EF025987
CSN2	GQ259487; GQ259486; GQ259485; GQ259484; GQ259483; GQ259482; GQ176293; GQ176292; GQ176291; GQ176290; GQ176289; GQ176288; GQ176287; AY352050; EF066481; DQ317447; AY599833; FN424088; FM986648; FM946182; EF115306; DQ631829; DQ191172; DQ191171; DQ191170; AJ005432; AJ005165; Y17839; Y17838; Y17837; Y17836
CSN3	AY750857; DQ645429; DQ191174; DQ191173; FJ770200; AM900443; EF133463; EF066482; D14370; AJ628346; AJ011387

\*Acesso em setembro de 2010

As proteínas beta-lactoglobulina e a alfa-lactoalbumina são as principais proteínas presentes no soro do leite, apresentando inúmeras propriedades nutricionais e fisiológicas (CLARE & SWAISGOOD, 2000). A beta-lactoglobulina é a maior proteína encontrada no soro do leite de ruminantes, cuja função está relacionada ao transporte de pequenos ligantes hidrofóbicos, como o retinol, essenciais para o crescimento do recém-nascido

(KONTOPIDIS et al., 2004; THREADGILL & WOMACK, 1990). A proteína alfa-lactoalbumina exerce a função de transporte de cálcio influenciando a síntese de lactose pelas glândulas mamárias durante o período de lactação. Trata-se de uma das proteínas mais bem caracterizadas na literatura, apresentando também atividade antitumor e antibactericida (PERMYAKOV et al., 2005). Dentre os genes codificantes das proteínas do soro, apenas o gene da beta-lactoglobulina (LGB) foi mapeado no genoma de búfalo, o qual se encontra no cromossomo 12 (AMARAL et al., 2008).

Devido à importância das proteínas do leite para a indústria biotecnológica e de laticínios, torna-se possível, com a disponibilidade de uma biblioteca BAC de búfalo, não só a caracterização da estrutura molecular desses genes no genoma do búfalo, como também a comparação de suas estruturas com aquelas correspondentes no genoma de outros bovídeos.

## **2. Regiões contendo genes da família das lipocalinas**

As lipocalinas formam uma família grande e diversa de proteínas presentes em vertebrados, invertebrados, plantas e bactérias, sendo tipicamente pequenas proteínas secretadas, que apresentam diferentes propriedades de reconhecimento molecular, tais como habilidade em ligar pequenas moléculas (principalmente hidrofóbicas), capacidade de ligação a receptores de superfície celular e capacidade de formação de complexos macro-moleculares (AKERSTROM et al., 2000; GUTIÉRREZ et al., 2000).

Os membros da família das lipocalinas exibem grande diversidade estrutural e funcional tanto intraespecífica quanto interespecífica. Embora tenham sido classificadas no passado como proteínas de transporte, hoje sabe-se que os membros desta família estão envolvidos em uma variedade de funções diferentes incluindo transporte de retinol,

coloração críptica, olfação, transporte de feromônio, síntese enzimática de prostaglandinas, regulação da resposta imune, mediação da homeostase celular e participação na proliferação de células cancerosas, além de estarem envolvidas em características reprodutivas (AKERSTROM et al., 2000; FLOWER, 1996).

Nos recentes mapas construídos para todos os cromossomos do genoma bubalino, quatro genes pertencentes à família das lipocalinas tiveram suas posições indicadas em cromossomos específicos: APOA1 (do inglês *apolipoprotein A-I*) no BBU16, BDA20 (do inglês *major allergen BDA20*) no BBUX, CRABP2 (do inglês *cellular retinoic acid binding protein 2*) no BBU6 e RBP3 (do inglês *retinol binding protein 3, interstitial*) no BBU4 (AMARAL et al., 2008). Com a disponibilidade de uma biblioteca BAC de búfalo, os genes da família das lipocalinas são interessantes para serem caracterizados em nível molecular. Outros genes da família das lipocalinas merecem destaque pelas suas funções, os quais estão descritos a seguir.

O gene AGP (do inglês *alpha-1 acid glycoprotein*) faz parte da subfamília das Imunocalinas, um grupo de genes cujos membros estão envolvidos na modulação da resposta imune e inflamatória (LOGDBERG & WESTER, 2000). Sua proteína é produzida principalmente pelo fígado e secretada no plasma em resposta a infecções e lesões, desempenhando um importante papel na regulação de reações inflamatórias, reduzindo os danos aos tecidos devido à alguma reação descontrolada (TILG et al., 1993), ou ao excesso de apoptose (VAN MOLLE et al., 1997). Além disso, essa proteína é capaz de reduzir o estado de ativação de neutrófilos e aumentar a expressão de citocinas anti-inflamatórias pelos macrófagos (HOCHEPIED et al., 2003). Estudos de imunohistoquímica têm mostrado que esse gene é expresso também em todos os tecidos da

glândula mamária durante a fase aguda, agindo potencialmente como um regulador da imunidade inata (CECILIANI et al., 2007).

O gene AMBP (do inglês *alpha-1-microglobulin/bikunin precursor*) também faz parte da subfamília das Imunocalinas. Esse gene é transcrito intensamente no parênquima, pâncreas, fígado e epitélio intestinal (SANCHEZ et al., 2002). Estudos sugerem que sua proteína pode proteger o óvulo de processos inflamatórios que ocorrem antes da ovulação, além de proteger o feto do processo de rejeição que acontece durante a gravidez (BERGGARD et al., 1999; ODUM & NIELSEN, 1997). Embora a expressão desse gene já tenha sido identificada em vários tecidos, a totalidade das funções fisiológicas ainda não está bem estabelecida (LARSSON et al., 2004). Em bovinos, o gene AMBP tem sido descrito como um biomarcador para tuberculose bovina provocada pela micobactéria *Mycobacterium bovis*, devido ao aumento no nível de expressão durante o processo de infecção (SETH et al., 2009).

O gene APOM (do inglês *apolipoprotein M*) codifica uma proteína secretada diretamente na membrana plasmática, envolvida no transporte de lipídeos e expressa predominantemente no fígado e no rim (ZHANG et al., 2003). Trata-se de uma apolipoproteína presente em todos os mamíferos, apresentando uma grande conservação, sugerindo alguma função importante (DUAN et al., 2001). Estudos demonstraram que a concentração da proteína APOM no plasma está correlacionada positivamente com os níveis de leptina e negativamente com os níveis de colesterol (LUO et al., 2004). Em processos inflamatórios e infecciosos, os níveis de expressão do gene APOM são reduzidos consideravelmente (FEINGOLD et al., 2008). Em suínos, esse gene está relacionado com deposição de gordura intramuscular (PAN et al., 2010).

As proteínas oriundas dos genes FABPs (do inglês *fatty acid binding proteins*) de

mamíferos estão localizadas em uma variedade de tipos celulares. Embora os aspectos estruturais dessas proteínas em tecidos de mamíferos já estejam bem definidas, avanços significativos em relação às funções fisiológicas têm sido mais lentos. O gene FABP1 (do inglês *fatty acid binding protein 1, liver*) codifica uma proteína carregadora exclusivamente intracelular, a qual tem sido relacionada com vários processos celulares, incluindo o transporte intracelular e armazenamento de ácidos graxos, modulação da atividade enzimática, transdução de sinal, diferenciação e regulação do crescimento (BRATT, 2000; WOLFRUM et al., 2000).

O gene FABP3 (do inglês *fatty acid binding protein 3, muscle and heart*) tem sido relacionado com espessura e deposição de gordura intramuscular em suíno (CHO et al., 2010; SWITONSKI et al., 2010). Em bovino, os genes FABP3 e FABP5 (do inglês *fatty acid binding protein 5*) são super expressos em glândulas mamárias durante o período de lactação, enquanto que os genes FABP1 e FABP6 (do inglês *fatty acid binding protein 6, ileal*) apresentam baixa expressão nas glândulas mamárias durante esse período (BIONAZ & LOOR, 2008a). Recentes estudos têm demonstrado que o gene FABP3 desempenha um papel fundamental na regulação de ácidos graxos para uma síntese abundante de gordura no leite de bovinos durante o período de lactação (QANBARI et al., 2010; BIONAZ & LOOR 2008b).

O gene PMP2 (do inglês *peripheral myelin protein 2*) tem sua expressão predominante nos nervos periféricos, sendo um gene candidato para neuropatias periféricas (BESANÇON et al., 2001). Embora a estrutura da proteína tenha sido determinada para bovinos desde 1980, suas funções ainda não estão estabelecidas (KITAMURA et al., 1980).

As proteínas oriundas dos genes RBPs (do inglês *retinol binding proteins*) também

merecem destaque por estarem relacionadas com o transporte de retinol, uma molécula essencial para diversos processos biológicos, tais como resposta imune, desenvolvimento embrionário, crescimento e diferenciação celular (BLOMHOFF & BLOMHOFF, 2006). Na reprodução de mamíferos, a importância da vitamina A e de seus metabólitos como retinol e ácido retinóico (coletivamente conhecidos como retinóides) já se encontram bem estabelecidas. As proteínas RBPs atuam durante o período de gestação, representando o principal componente do histotrofo no útero (MCNEILL et al., 2006). São expressas no endométrio bovino durante todo o ciclo estral e prenhez, cuja regulação é temporal e regulada por ações contrárias ao estrogênio e progesterona (MACKENZIE et al., 1997; MACKENZIE et al., 1996). Há trabalhos que relatam que as proteínas RBPs e outros elementos da via de sinalização de retinóides estão presentes em embriões bovinos pré-implantados (MOHAN et al., 2001; 2002; BROWN et al., 2003). Particularmente, o gene RPB4 (do inglês *retinol binding protein 4, plasma*) tem sido relacionado com resistência a insulina e obesidade (GRAHAM et al., 2006). Em bovinos, um estudo recente tem denominado o gene RBP4 como um marcador para produção de carne e outras características de desempenho produtivo e reprodutivo (WANG et al., 2010).

A nomenclatura oficial para os genes da família das lipocalinas em bovino, assim como símbolos e nomenclaturas alternativas estão descritas na Tabela 3.

**Tabela 3.** Nomenclatura dos genes da família das lipocalinas em bovino, com os respectivos símbolos e nomenclaturas alternativas utilizadas em outras espécies de mamíferos.

<b>Símbolo oficial em <i>B. taurus</i></b>	<b>Nome oficial em <i>B. taurus</i></b>	<b><i>Bos taurus</i> Gene ID</b>	<b>Símbolos alternativos</b>	<b>Nomes alternativos em outras espécies de mamíferos</b>
AGP	<i>alpha-1 acid glycoprotein</i>	ID: 497200	AGP1, AGP-A, OMD1, ORM, ORM1	<i>orosomucoïd; orosomucoïd 1</i>
AMBP	<i>alpha-1-microglobulin/bikunin precursor</i>	ID: 280996	A1M, EDC1, HCP, HI30, IATIL, ITI, ITIL, ITILC, UTI	<i>bikunin; complex-forming glycoprotein heterogeneous in charge; growth-inhibiting protein 19; inter-alpha-trypsin inhibitor light chain; trypstatin; uristatin; uronic-acid-rich protein</i>
APOM	<i>apolipoprotein M</i>	ID: 505830	--	<i>apo-M</i>
FABP1	<i>fatty acid binding protein 1, liver</i>	ID: 327700	L-FABP, LBFABP	<i>L-FABP; fatty acid-binding protein, liver; liver-type fatty acid-binding protein</i>
FABP3	<i>fatty acid binding protein 3, muscle and heart (mammary-derived growth inhibitor)</i>	ID: 281758	FABP, FABP-3, H-FABP	<i>fatty acid-binding protein, heart; heart fatty acid-binding protein; heart-type fatty acid-binding protein</i>
FABP5	<i>fatty acid binding protein 5 (psoriasis-associated)</i>	ID: 281760	E-FABP, PAFABP	<i>differentiation-associated lipid-binding protein LP2; epidermal fatty acid-binding protein; epidermal-type fatty acid-binding protein; fatty acid-binding protein, epidermal</i>
FABP6	<i>fatty acid binding protein 6, ileal</i>	ID: 514650	IBABP, I-BALB, I-BAP, ILBP, ILBP3, ILLBP	<i>fatty acid binding protein 6, ileal (gastrotropin); fatty acid-binding protein 6; gastrotropin; ileal lipid-binding protein</i>
PMP2	<i>peripheral myelin protein 2</i>	ID: 506062	FABP8, M-FABP, MP2, P2	<i>myelin P2 protein</i>
RBP4	<i>retinol binding protein 4, plasma</i>	ID: 281444	PRBP; RBP	<i>plasma retinol-binding protein</i>

Informações disponíveis no banco de dados Entrez Gene (<http://www.ncbi.nlm.nih.gov/entrez>)

### 3. Regiões contendo genes relacionados com resposta imune inata e adaptativa

As doenças infecciosas que acometem os búfalos são muito parecidas com as que afetam os bovinos. Ambos são suscetíveis a um espectro similar de agentes infecciosos, sendo que muitas das vacinas e quimioterápicos desenvolvidos para uso em bovinos têm mostrado eficácia em búfalo, como por exemplo, medicamentos que combatem brucelose, leptospirose, antraz, peste bovina e febre aftosa (MARQUES, 2000; RAHMAN et al., 1987; TOMAR & TRIPATHI, 1987; NATIONAL RESEARCH COUNCIL, 1981).

No entanto, as necessidades sanitárias de búfalos, especialmente a dos animais jovens, são distintas das apresentadas pelos bovinos. Por exemplo, o comportamento de chafurdar expõe esses animais ao contato com patógenos de veiculação hídrica, aumentando o contato com vetores transmissores de doenças, o que normalmente não ocorre com o bovino. Além disso, divergências genéticas entre bovinos e bubalinos podem ser responsáveis por diferenças existentes na susceptibilidade/resistência a determinadas doenças entre esses animais (DAVIS et al., 2001).

O MHC (do inglês *Major Histocompatibility Complex*) consiste em uma região genômica com um denso agrupamento de genes altamente polimórficos, muitos dos quais estão relacionados à resposta imune inata e adaptativa, desempenhando um papel importante na resposta do hospedeiro à patógenos (KELLEY et al., 2005; LEWIN et al., 1999).

Na maioria dos mamíferos, esses genes estão organizados em *clusters* (CHILDERS et al., 2005), os quais transcrevem proteínas que sinalizam infecção, apresentando o patógeno aos receptores dos linfócitos CD8<sup>+</sup> e CD4<sup>+</sup> (ELLIS, 2004). Via de regra, os genes do MHC estão organizados por regiões com funções similares, contendo genes de classe I, genes de classe II e genes de classe III (KLEIN & SATO, 1998), além daqueles considerados de classe I e classe II estendidas (STEPHENS et al., 1999; HERBERG et al., 1998; TOTARO et al., 1996). O número de genes e a presença/ausência de cada uma das regiões variam conforme a espécie, uma vez que durante a evolução novos genes são criados pelo processo de duplicação gênica, podendo ser mantidos no genoma, deletados ou tornarem-se não funcionais devido à mutações (NEI & ROONEY, 2005; TAKAHASHI et al., 2000; GU & NEI, 1999; NEI et al., 1997; KLEIN et al., 1993).

Os genes de classe I geralmente expressam moléculas responsáveis pela apresentação de peptídeos dos antígenos aos linfócitos CD8 por meio dos receptores de membrana de célula T (BRINKMEYER-LANGFORD et al., 2009). Também podem atuar como ligantes para receptores nas células NK (do inglês *Natural Killer Cell*) (KELLEY et al., 2005).

Os genes de classe II são em maior número e os mais bem caracterizados em mamíferos (KELLEY et al., 2005). Esses genes são expressos exclusivamente nas células apresentadoras de antígenos incluindo macrófagos, células dendríticas e linfócitos B (DE et al., 2002). Além disso, intracelularmente nas células *T-helpers*, as proteínas expressas pelos genes de classe II degradam os peptídeos de patógenos (BROWN et al., 1993).

Os genes pertencentes à classe III codificam um diverso grupo de proteínas (incluindo as citocinas) com funções variadas, destacando-se interações com os receptores das células NK (TROWSDALE, 2002). Porém, nem todos os genes pertencentes à classe III estão envolvidos com a resposta imune (BRINKMEYER-LANGFORD et al., 2009; XIE et al., 2003; MILNER & CAMPBELL, 2001).

Devido ao importante papel que os genes do MHC desempenham na resposta imune, essa é uma das regiões mais bem estudadas em diversas espécies de mamíferos, demonstrando uma grande complexidade genética, não só com um número variável de genes e diferentes haplótipos, como também uma grande diversidade alélica (BIRCH et al., 2006; DEAKIN et al., 2006; KELLEY et al., 2005; TAKAHASHI et al., 2000).

Na maioria das espécies de mamíferos, os genes de classe II se apresentam como um segmento único, como por exemplo, em humano (MHC SEQUENCING CONSORTIUM, 1999), camundongo (MOUSE GENOME SEQUENCING CONSORTIUM, 2002), cachorro (WAGNER, 2003), gato (YUHKI et al., 2008), suíno (LUNNEY et al., 2009)

e eqüino (GUSTAFSON et al., 2003). Embora exista evidências de uma clara homologia dos genes desta região entre todos os mamíferos, a organização do MHC em bovídeos se apresenta de uma maneira diferente dos demais mamíferos, com os genes da classe II distribuídos em dois *clusters* denominados IIa e IIb (LIU et al., 2010; BRINKMEYER-LANGFORD et al., 2009; TRAUJL et al., 2005; AMILLS et al., 1998).

Em bovinos, os dois *clusters* de genes da classe II estão localizados cerca de 20 cM de distância, separados por uma inversão cromossômica quando comparados com o MHC de outros mamíferos não bovídeos (BAND et al., 1998). Genes da classe IIa estão intimamente associados com os genes das classe I e III, enquanto que os genes da classe IIb estão posicionados em um *cluster* mais próximo ao centrômero. Essa separação única dos *loci* de classe II, de função relacionada e intimamente ligados em outras espécies, torna o estudo desta parte do genoma de alta prioridade para a compreensão dos processos envolvidos na regulação gênica, na estrutura e evolução deste complexo gênico.

Um mapa de alta resolução do MHC bovino foi construído utilizando um painel RH bovino/hamster de 12.000 *rads*, ordenando 77 marcadores, dos quais 14 foram mapeados na região de classe I, 27 na classe IIa, 17 na classe IIb e 19 na classe III (BRINKMEYER-LANGFORD et al., 2009).

Comparado com outras espécies de interesse econômico, os genes do MHC de búfalo são pouco estudados, uma vez que a literatura apresenta poucos trabalhos dos genes presentes nessa região, sendo quase que na sua totalidade trabalhos relacionados com polimorfismos (SENA et al., 2003; DE et al., 2002; ARAVINDAKSHAN et al., 2000).

Os genes do MHC de búfalo (*Bubalus bubalis*) estão localizados no braço curto do cromossomo 2 (BBU2p), o qual é homólogo ao cromossomo bovino 23 (BTA23)

(IANNUZZI et al., 1993). Estudos citogenéticos indicam a presença do gene DRA (classe IIa) na banda 22 e do gene DYA (classe IIb) na banda 13, fornecendo indícios sobre uma organização em *clusters*, similar àquela apresentada em bovinos (DI MEO et al., 2008; IANNUZZI et al., 2003).

Em 2008 foi publicado o mapa RH preliminar do MHC em búfalos gerado com o painel de células BBURH<sub>5000</sub>, contendo 10 genes de classe II mapeados, sendo seis da classe IIa (BoLA-DRB, BoLA-DRB2, BoLA-DRB3, BoLA-DRA, BoLA-DQA e BoLA-DQA2) e quatro da classe IIb (BoLA-DYA, BoLA-DYB, PSMB9 e BoLA-DMA) (RODRIGUES FILHO et al., 2008). Esse mapa preliminar do MHC em búfalos corroborou os dados citogenéticos de que os genes da classe II se encontravam divididos em dois *clusters*, uma vez que apresentou os genes da classe IIa e IIb alocados em dois grupos de ligação distintos.

Uma vez que o búfalo apresenta uma resposta a agentes infecciosos de maneira diferente do bovino, é evidente que as necessidades do búfalo não serão sempre atendidas por meio de vacinas e regimes de gestão concebidos para o bovino. Para solucionar problemas exclusivos aos búfalos, será necessário compreender as variáveis envolvidas na susceptibilidade diferencial à determinada doença. Tal conhecimento não só contribuirá para a melhoria da gestão dos búfalos, mas também fornecerá uma visão para os mecanismos responsáveis por diferenças na resistência a doenças de importância econômica (SHARMA, 1987; SHARMA et al., 1978).

Um passo importante no estudo da resposta imune em búfalo envolve a caracterização de genes do MHC, elucidando as semelhanças e diferenças existentes na organização dos genes de resposta imune, quando comparadas com a região correspondente em bovinos.

## **Objetivos**

Visando ampliar o número de marcadores moleculares para o estudo do genoma do búfalo, este trabalho teve como objetivos:

- ✓ Selecionar iniciadores para PCR de marcadores provenientes de estudos do genoma bovino, incluindo aqueles relacionados com produção e qualidade de leite, da família das lipocalinas e do complexo principal de histocompatibilidade;
- ✓ Mapear parte dos marcadores selecionados no genoma bubalino, utilizando o painel BBURH<sub>5000</sub>.

## Material e Métodos

### 1. Seleção dos iniciadores para PCR do genoma bovino e sua otimização com DNA bubalino

Os iniciadores para PCR foram selecionados de artigos científicos e/ou das informações disponíveis no banco de dados americano NCBI (<http://www.ncbi.nlm.nih.gov>).

Para a otimização das condições da reação de PCR com DNA de búfalo, as mesmas foram realizadas em um volume total de 10 µl, contendo 50 ng de DNA; 0,2 mM de cada oligonucleotídeo iniciador; 100 mM de dNTP; 10 mM de Tris-HCl; 1,5 mM de MgCl<sub>2</sub> e 0,5U da DNA polimerase AmpliTaqGold™ (*Applied Biosystems*). Para a otimização da temperatura de anelamento dos iniciadores para PCR foi utilizado um termociclador do tipo gradiente (Veriti™ – *Applied Biosystems*), com bloco para 96 amostras distribuídas em 8 linhas e 12 colunas. No gradiente de temperatura, cada coluna do termociclador segue uma temperatura de anelamento determinada, sendo a temperatura mínima utilizada de 50°C e a máxima de 65°C.

Para os experimentos de mapeamento dos genes, foram utilizados além do DNA bubalino, DNA de hamster e DNA bovino como controles da reação de PCR uma vez que os iniciadores para PCR são derivados do genoma bovino e o painel BBURH<sub>5000</sub> utilizado no mapeamento foi construído com células de hamster como a célula receptora para a formação das células híbridas.

Os parâmetros utilizados para a amplificação foram: desnaturação inicial de 94°C por 10 minutos, seguida de 35 ciclos com desnaturação a 94°C por 30 segundos, anelamento

de 50-65°C por 30 segundos e extensão a 72°C por 30 segundos, seguida de uma extensão final a 72°C por 6 minutos.

Os produtos de PCR foram visualizados em gel de agarose 2%, corado com brometo de etídio, submetido à eletroforese horizontal em tampão TBE 0,5X por 100V a 30 minutos. Os resultados foram fotodocumentados com câmera digital do tipo DC290 Kodak™.

## **2. Mapeamento utilizando o painel BBURH<sub>5000</sub>**

Após a otimização da temperatura de anelamento dos iniciadores para PCR, seguiu-se com a genotipagem das linhagens do painel BBURH<sub>5000</sub> (AMARAL et al., 2007). Este painel é composto por 90 linhagens híbridas irradiadas resultantes da fusão de fragmentos do genoma bubalino com o genoma de células de hamster. Cada linhagem celular gerada a partir do processo de fusão contém um conjunto de fragmentos cromossômicos da célula doadora incorporado ao genoma da célula receptora, a qual atinge a estabilidade após sucessivos processos de divisão celular (AMARAL et al., 2007).

Cada marcador foi genotipado duas vezes para evitar a ocorrência de falsos positivos e/ou negativos, sendo que, nos casos duvidosos, o marcador foi genotipado uma terceira vez. O padrão de presença e ausência do produto de PCR de cada marcador nas linhagens do painel foi registrado em uma tabela do tipo *Excel*. A análise estatística dos dados foi realizada com o pacote estatístico CarthaGene (DE GIVRY et al., 2005), o qual emprega a função de máxima verossimilhança, descrita por LANGE e colaboradores (1995), para estimar as distâncias e avaliar a ordem dos marcadores ao longo dos cromossomos, utilizando o parâmetro LOD.

Para a construção do mapa RH, os dados da genotipagem foram analisados com o pacote estatístico CarthaGene. Os marcadores genotipados com o painel BBURH<sub>5000</sub> foram distribuídos em grupos de ligação (LGs) usando um LOD mínimo de 8.0. Um mapa *framework* 1000:1 foi gerado com os marcadores agrupados por meio do comando “*buildfw*”, o qual gerou um mapa através da inserção de *locus*. Todos os demais marcadores foram inseridos, um a um, em todas as posições possíveis até a posição mais provável ser encontrada. O mapa gerado foi então aperfeiçoado, utilizando o algoritmo “*flip*” para testar todas as ordens possíveis para os marcadores dentro de um intervalo. Finalmente, o comando “*polish*” foi usado para testar a confiabilidade do mapa, deslocando cada marcador em todos os possíveis intervalos. O *software* MapChart<sup>®</sup> versão 2.0 (VORRIPS, 2002) foi utilizado para desenhar os mapas obtidos.

## Resultados e Discussão

### 1. Seleção dos iniciadores para PCR do genoma bovino e sua otimização com DNA bubalino

O processo de seleção resultou na coleta de informações de iniciadores para PCR para seis genes relacionados com produção e qualidade do leite (CSN1S1, CSN1S2, CSN2, CSN3, LALBA e LGB), nove genes da família das lipocalinas (AGP, AMBP, APOM, FABP1, FABP3, FABP5, FABP6, PMP2 e RBP4), e 43 genes do complexo principal de histocompatibilidade (MHC). As informações obtidas para esses marcadores, incluindo a seqüência dos iniciadores para PCR, os cromossomos bovino (BTA) e bubalino (BBU) e a referência, encontram-se descritos na Tabela 4.

**Tabela 4.** Marcadores selecionados do genoma bovino, incluindo a seqüência dos iniciadores para PCR, a localização dos genes no genoma bovino (BTA), o cromossomo homólogo em búfalo (BBU) e a referência bibliográfica.

Símbolo do marcador	BTA	BBU	Seqüência senso (F) e anti-senso (R) dos iniciadores para PCR (5'-3')	Referência Bibliográfica
CSN1S1	6	7	F: CCTTACCACTTCACTATTGCCAC R: CTTTCTTAAAGCATAGAGCATATCC	UniSTS:253703*
CSN1S2	6	7	F: TCTCCTCCTAGGATTGAAAAGA R: TTGAGACGGTTTGGGAATCT	SAZANOV et al., 2006
CSN2	6	7	F: ACAGCCTCCCACAAAACATC R: AGACTGGAGCAGAGGCAGAG	SAZANOV et al., 2006
CSN3	6	7	F: CGCTGTGAGAAAGATGAAAGATT R: TTTGGGCAGGTGACCTAACT	UniSTS:519151*
LALBA	5	4q	F: CTCTGTGAGAAGTTGTGAACACC R: CATCGAGCAAGGGTCAAAAAGTCC	ALEYASIN & BARENDSE, 1999
LGB	11	12	F: TGTGCTGGACACCGACTACAAAAAG R: GCTCCCGGTATATGACCACCCTCT	UniSTS:279479*
AGP	8	3q	F: TGGAGAAGCAGCATGAGGAG R: TTGATGCAACCGAGGGAECT	DORROCH et al., 2001
AMBP	8	3q	F: GGCCTGTAACCTCCCATAG R: CAGGAATGCCACAGTACTCC	ITOH et al. 2005
APOM	23	2p	F: AAGGAGGAGTTGGCGACTTT R: TCTGTGCTCCCATCAGACAG	UniSTS:269475*

**Tabela 4.** *Continuação...*

<b>Símbolo do marcador</b>	<b>BTA</b>	<b>BBU</b>	<b>Seqüência senso e anti-senso dos iniciadores para PCR</b>	<b>Referência Bibliográfica</b>
FABP1	11	12	F: GTTCATCATCACCGCTGGCT R: CCACTGCCTTGATCTTCTCCC	BIONAZ & LOOR, 2008
FABP3	2	2q	F: GTTCCTTCTCAGGCAGTGGCTT R: CAGTGGTTCTGTGACAGGGTA	EVESTS-VAN DER WIND et al., 2004
FABP5	14	15	F: GGGATGGAAAGGAAAGCACAAAT R: CTTGTTTCATCCTCGCAGCTAAT	UniSTS:278364
FABP6	7	9	F: ATCTCGGAGATACAGCAGGACG R: CTTGAACCTTCTGTTTCCCACG	BIONAZ & LOOR, 2008
PMP2	14	15	F: TGGCCAAACCCAGAGTGATCATCA R: GTCTCAAAGCAGTCGATGGAACCA	JANN et al., 2006
RBP4	26	23	F: ACCACTGGATCATCGACACGG R: CCTGGCCAGGCACAGCTC	UniSTS:254003
10.00	23	2p	F: TGTCTATCCTGTCCCAGAGCTA R: CGATTTCTCTCATGCTTACCTG	BRINKMEYER-LANGFORD et al., 2009
10.05	23	2p	F: GTGCCTACGTTACAGCTCATTC R: CTCTCTTTTTCTTGTTCACAGGT	BRINKMEYER-LANGFORD et al., 2009
10.20	23	2p	F: CTGTAACATATCGCCATCAGGAG R: CTCAAGTAAGCTGCATCCATTC	BRINKMEYER-LANGFORD et al., 2009
10.25	23	2p	F: ACAGCATGTGGGATGTTAGTTC R: GCCTTCACACTTAGAGCAACAC	BRINKMEYER-LANGFORD et al., 2009
10.30	23	2p	F: CAGAAAGGAAAGCTGTTGACAC R: CAGTTTGAAGCCCATTACTCAG	BRINKMEYER-LANGFORD et al., 2009
10.40	23	2p	F: TCAAGCTCTACCTGTCCAAGAA R: CATCTTCACTGTTGCTTTCTCTC	BRINKMEYER-LANGFORD et al., 2009
10.45	23	2p	F: AGTTTTCCAGGCTTCTTACAGG R: GTACACTCACCCACGATGTCTT	BRINKMEYER-LANGFORD et al., 2009
10.50	23	2p	F: AAGTTGAAGACCACGACTCAGA R: CTGTGAGGTCAATGTTTCAGGAT	BRINKMEYER-LANGFORD et al., 2009
10.55	23	2p	F: ATGAGTTAGGGAAGAAGGGTCA R: TCACCGTCTATATCAAGCATCC	BRINKMEYER-LANGFORD et al., 2009
10.60	23	2p	F: TCCCCTACCTCACCAGATAGTT R: ACCCACTAGGAAGAAGGTTTCA	BRINKMEYER-LANGFORD et al., 2009
11.00	23	2p	F: TAACCGCAGCAACTAGAGAGAG R: GAATAGGCAGGTGGGAGAATAC	BRINKMEYER-LANGFORD et al., 2009
11.05	23	2p	F: GTTTCCTCGGATGGAGATACA R: AGGCAGGATGTAGAGACTCTGA	BRINKMEYER-LANGFORD et al., 2009
11.10	23	2p	F: GGACAATATCATCAGTGGGCTA R: CAGGAGACTTGCGTAGAAGAGA	BRINKMEYER-LANGFORD et al., 2009
12.05	23	2p	F: CTAAGGTGACAGATGTGGGTGT R: AAAGCGAAGTAGGGTAGTGAGG	BRINKMEYER-LANGFORD et al., 2009
13.00	23	2p	F: CTCACCGTTTTAAGAGCTTCT R: CTTTCCCATTAACTGTTCTCTC	BRINKMEYER-LANGFORD et al., 2009
13.05	23	2p	F: CAAGCTCCACCACAGAACTAAC R: ATCAGCTATCCCAGAGGAAGAG	BRINKMEYER-LANGFORD et al., 2009
48.00	23	2p	F: ACCTGGATAGAGCTGACACAGA R: GTTCCTCAGTAGTCTGGTGCTG	BRINKMEYER-LANGFORD et al., 2009

**Tabela 4.** *Continuação...*

<b>Símbolo do marcador</b>	<b>BTA</b>	<b>BBU</b>	<b>Seqüência senso e anti-senso dos iniciadores para PCR</b>	<b>Referência Bibliográfica</b>
48.05	23	2p	F: CCAGAGGGTTAGGTTTAGGATG R: CTACCTCCCTTCAGCTTTCCT	BRINKMEYER-LANGFORD et al., 2009
48.15	23	2p	F: GGAAATAGGAACTGTGGGTCTC R: CCTGAGATACGGGTGTTGATAA	BRINKMEYER-LANGFORD et al., 2009
48.25	23	2p	F: AGACTGATGAATGGATGACCAG R: TAAGGGCAGGGAAAAGAGATAG	BRINKMEYER-LANGFORD et al., 2009
50.15	23	2p	F: GCAAAGAATTGATTTCCACACA R: AAACAGTCGGATGGCATTACTT	BRINKMEYER-LANGFORD et al., 2009
52.00	23	2p	F: TTAACAAAGGATCAGAGCAGCA R: AGGCAGTTGTATAGCGGTTGTT	BRINKMEYER-LANGFORD et al., 2009
52.55	23	2p	F: ATGGCAATATTTGTTCTCCTGGAC R: TTCCCCTTCAAGTGTATGGTCT	BRINKMEYER-LANGFORD et al., 2009
53.05	23	2p	F: TTTGACATGACACACTTCACCA R: AAGCATTCTTTGAAAGGACGAG	BRINKMEYER-LANGFORD et al., 2009
53.25	23	2p	F: CAGGAGGAACTCTAGCTCCAAA R: AGTTCCTGACTTACGCGTCTTC	BRINKMEYER-LANGFORD et al., 2009
55.00	23	2p	F: GTCACACAAATAGAGCAGCACA R: CAGTCACAGGCTTAGATTTTGC	BRINKMEYER-LANGFORD et al., 2009
55.05	23	2p	F: GCCATATTTTCAGGGTCTCTAGG R: CCAGCTTTGTTTCTTCTCACAC	BRINKMEYER-LANGFORD et al., 2009
55.15	23	2p	F: CTTCTGGAAGTCTTATCATGC R: TGACAGTGCATACAGAGGTGAG	BRINKMEYER-LANGFORD et al., 2009
55.30	23	2p	F: AGTTCCTCCAGCTCTCCTCATC R: GAATATCCTGCTTTCAGCCTCT	BRINKMEYER-LANGFORD et al., 2009
56.00	23	2p	F: GGTCTTAAAATCTGGGGACTTG R: TCTCTGCCACTGTCCACTACTT	BRINKMEYER-LANGFORD et al., 2009
57.00	23	2p	F: ACTCTACCCAGCACTCTCAACA R: GTCTCTGCCTCCTACCAAGACT	BRINKMEYER-LANGFORD et al., 2009
57.10	23	2p	F: CTAAGTTGCAGGGATAGCAGTG R: CATTGCGCAGTTTCTACTGTTC	BRINKMEYER-LANGFORD et al., 2009
57.20	23	2p	F: GGACACAGGACACAGGTTAAAA R: AAACACGAGTCTGGAAGCTACA	BRINKMEYER-LANGFORD et al., 2009
57.30	23	2p	F: AGATGTGACTATTGAGGCATGG R: CCTTTCTAGGTGATCTGCTGTG	BRINKMEYER-LANGFORD et al., 2009
57.40	23	2p	F: ATGATGACTCCAGGTGATTCTG R: ACATACTAGGAAAAGGGGTGGA	BRINKMEYER-LANGFORD et al., 2009
58.00	23	2p	F: CCTGAAACTCAGCCTTTTGC R: GCCTGGAGAGTCAGACCAAG	BRINKMEYER-LANGFORD et al., 2009
59.05	23	2p	F: ATCCTCTCCTCTGCGACTTTAC R: CTGGTATCCATGACCTCACACT	BRINKMEYER-LANGFORD et al., 2009
61.00	23	2p	F: TCCTCAGAAACCAGAGTGTGTT R: GGCCTTGGTGAGTGAAGCTT	BRINKMEYER-LANGFORD et al., 2009
65.00	23	2p	F: GGCTTCTGCTCCTAAGTATGGT R: CATAGTTCTTGACTGCTTTGG	BRINKMEYER-LANGFORD et al., 2009
65.05	23	2p	F: CACACTGATTGCACTGATATGG R: AGACCTTCTGGGCTCATTAGAG	BRINKMEYER-LANGFORD et al., 2009

**Tabela 4.** *Continuação...*

Símbolo do marcador	BTA	BBU	Seqüência senso e anti-senso dos iniciadores para PCR	Referência Bibliográfica
65.10	23	2p	F: CACACGTACACACACATTCTCC R: CTATGGTGATGCTGATCTGGTT	BRINKMEYER-LANGFORD et al., 2009
65.15	23	2p	F: CCTGCTTATCTTCCTCCCTCTA R: GTGGTAAATGGGAACTCCAGAT	BRINKMEYER-LANGFORD et al., 2009
BoLA-DNA	23	2p	F: ACACAGCTGACCACATGGGC R: CGGTACCGTTGGGGGCTCTG	UniSTS:519084*

\*UniSTS refere-se ao número de identificação da respectiva seqüência de iniciadores para PCR no banco de dados NCBI (<http://www.ncbi.nlm.nih.gov/unists/>)

Os 58 pares de iniciadores para PCR selecionados foram submetidos a reações de PCR com gradiente de temperatura de anelamento, cujas temperaturas selecionadas para os estudos de mapeamento estão descritos na Tabela 5, juntamente com o tamanho aproximado do produto de PCR obtido com DNA de búfalo em pares de base (bp).

**Tabela 5.** Marcadores derivados do genoma bovino para exploração do genoma bubalino por meio da biblioteca BAC e/ou mapeamento.

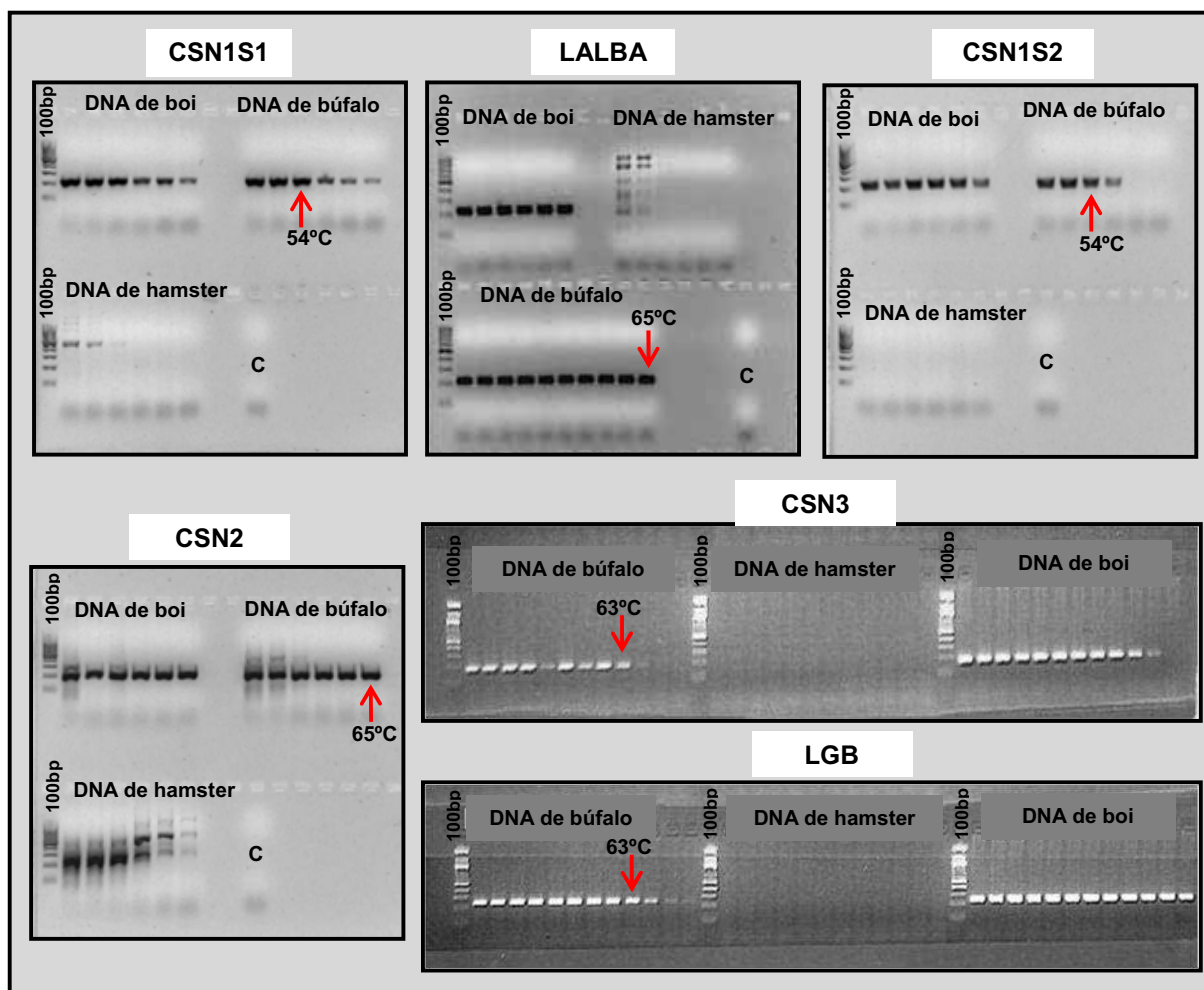
Símbolo do Marcador	Temperatura de anelamento	Produto de PCR em búfalo	Símbolo do Marcador	Temperatura de anelamento	Produto de PCR em búfalo
CSN1S1	54°C	220 bp	13.00	65°C	300 bp
CSN1S2	54°C	220 bp	13.05	65°C	300 bp
CSN2	65°C	200 bp	48.00	65°C	320 bp
CSN3	63°C	220 bp	48.05	65°C	300 bp
LALBA	65°C	220 bp	48.15	65°C	200 bp
LGB	63°C	250 bp	48.25	65°C	280 bp
AGP	59°C	120 bp	50.15	54°C	150 bp
AMBP	65°C	150 bp	52.00	54°C	150 bp
APOM	65°C	500 bp	52.55	54°C	150 bp
FABP1	65°C	100 bp	53.05	56°C	200 bp
FABP3	65°C	180 bp	53.25	56°C	200 bp
FABP5	65°C	1000 bp	55.00	65°C	280 bp
FABP6	65°C	120 bp	55.05	65°C	280 bp
PMP2	65°C	180 bp	55.15	65°C	350 bp
RBP4	56°C	160 bp	55.30	65°C	320 bp
10.00	65°C	280 bp	56.00	65°C	280 bp
10.05	65°C	300 bp	57.00	58°C	300 bp
10.20	56°C	300 bp	57.10	60°C	250 bp
10.25	65°C	250 bp	57.20	60°C	300 bp
10.30	65°C	300 bp	57.30	56°C	150 bp
10.40	58°C	280 bp	57.40	58°C	200 bp
10.45	56°C	300 bp	58.00	58°C	250 bp

**Tabela 5.** *Continuação...*

<b>Símbolo do Marcador</b>	<b>Temperatura de anelamento</b>	<b>Produto de PCR em búfalo</b>	<b>Símbolo do Marcador</b>	<b>Temperatura de anelamento</b>	<b>Produto de PCR em búfalo</b>
10.50	65°C	250 bp	59.05	55°C	300 bp
10.55	65°C	300 bp	61.00	56°C	300 bp
10.60	65°C	300 bp	65.00	65°C	250 bp
11.00	65°C	350 bp	65.05	65°C	320 bp
11.05	56°C	280 bp	65.10	65°C	300 bp
11.10	65°C	350 bp	65.15	65°C	300 bp
12.05	65°C	320 bp	BoLA-DNA	58°C	260 bp

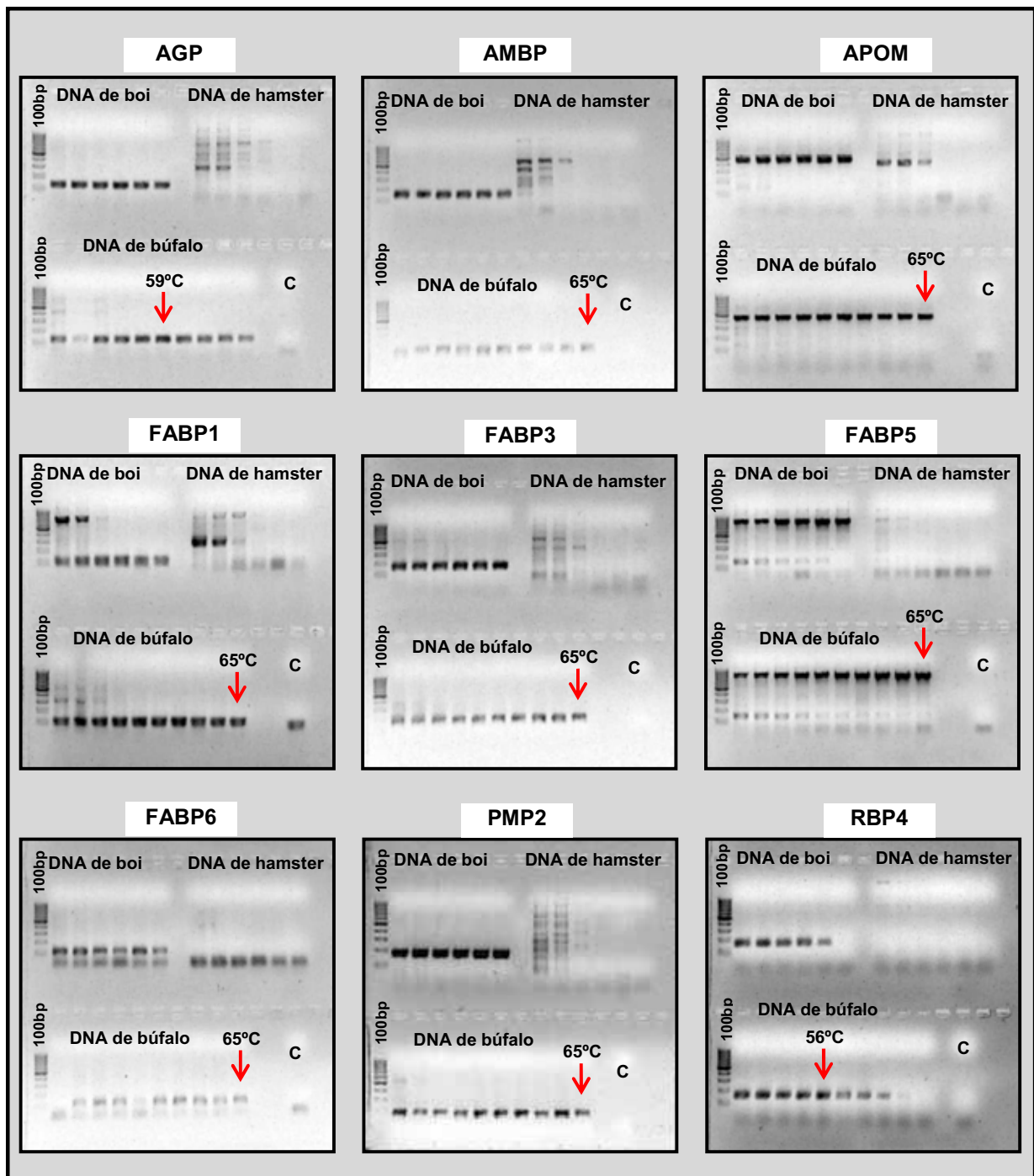
A Figura 1 representa o resultado da amplificação por PCR obtida a partir de experimentos com gradiente de temperatura de anelamento com os iniciadores para PCR correspondentes aos genes CSN1S1, CSN1S2, CSN2, CSN3, LALBA e LGB.

Como pode ser observado, todos os iniciadores para PCR apresentaram um produto de PCR único em búfalo e ausente em hamster na temperatura de anelamento selecionada (indicada por uma seta vermelha na Figura 1). Para o marcador CSN1S1 foi obtido um produto de PCR único em búfalo e ausente em hamster na temperatura de anelamento de 54°C. Quanto ao gene CSN1S2, observa-se que nas temperaturas de 50°C a 59°C foi possível obter produto de PCR com o DNA bubalino e ausência de produto de PCR com o DNA de hamster, cuja temperatura selecionada para esse marcador também foi de 54°C. Os iniciadores para PCR dos genes CSN2 e LALBA amplificaram produto de PCR com DNA de búfalo em todas as temperaturas de anelamento testadas, onde optou-se pela temperatura mais elevada (65°C). Já para os marcadores CSN3 e LGB, apresentaram produto de PCR robusto com o DNA de búfalo nas temperaturas de 50°C a 63°C, sendo selecionada esta última.



**Figura 1.** Géis de agarose 2% corado com brometo de etídio, mostrando o resultado das reações de PCR com gradiente de temperatura de anelamento com os iniciadores para PCR derivado dos genes CSN1S1, CSN1S2, CSN2, CSN3, LALBA e LGB. **DNA de boi:** controle positivo da reação, uma vez que os iniciadores para PCR foram desenhados a partir do genoma bovino. **DNA de hamster:** DNA controle para a seleção dos marcadores para o mapeamento utilizando o painel BBURH<sub>5000</sub>. **C:** controle negativo da reação de PCR (sem adição de DNA). As setas vermelhas indicam as temperaturas de anelamento selecionadas para cada par de iniciadores para PCR: 54°C para CSN1S1 e CSN1S2, 65°C para LALBA e CSN2, e 63°C para CSN3 e LGB.

A Figura 2 representa o resultado da amplificação por PCR obtida com os pares de iniciadores para PCR correspondentes aos genes da família das lipocalinas (AGP, AMBP, APOM, FABP1, FABP3, FABP5, FABP6, PMP2 e RBP4) também a partir de experimentos com gradiente de temperatura de anelamento variando de 50°C a 65°C.



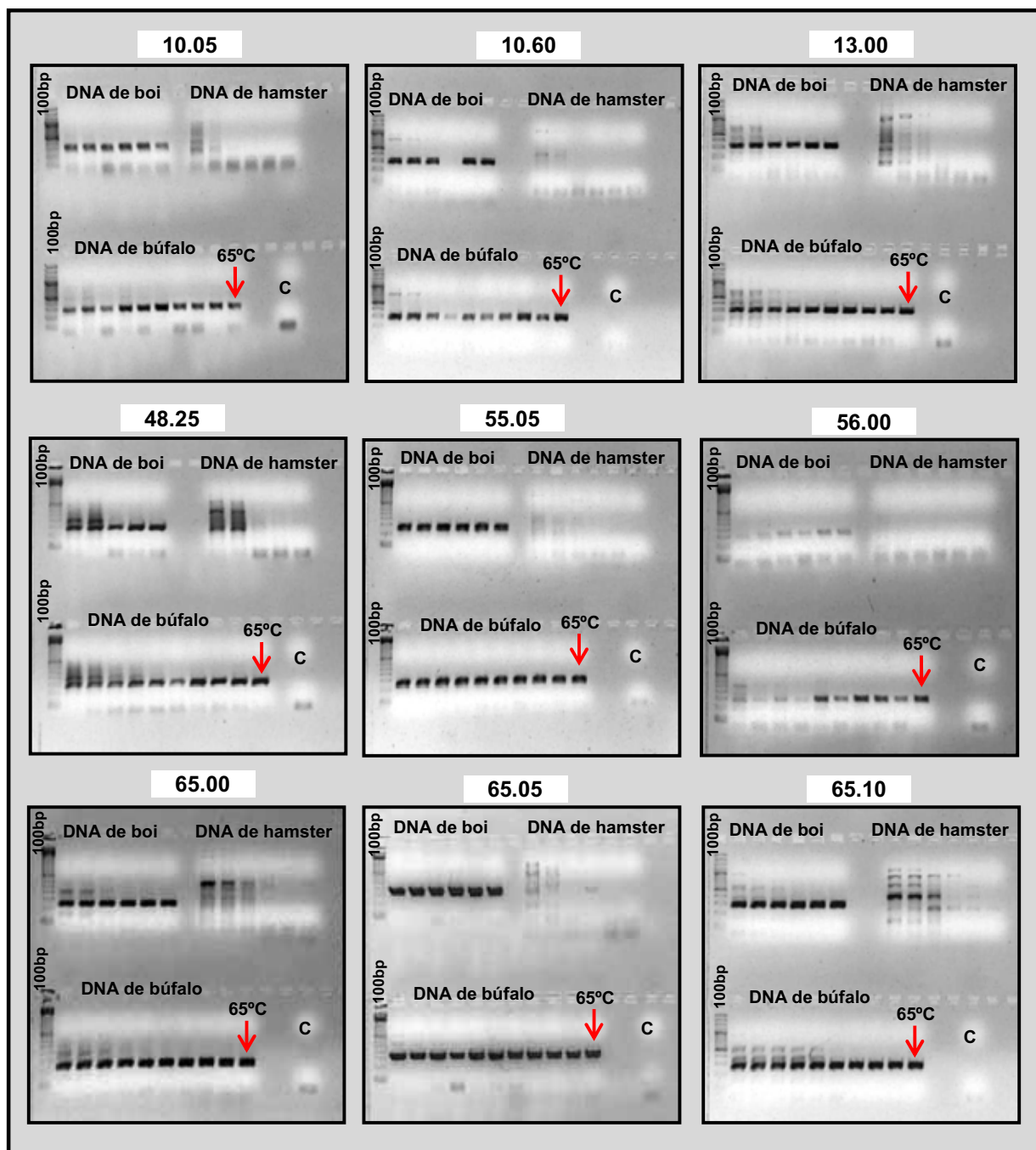
**Figura 2.** Géis de agarose 2% corado com brometo de etídio, mostrando o resultado das reações de PCR com gradiente de temperatura de anelamento com os iniciadores para PCR dos genes da família das lipocalinas. **DNA de boi:** controle positivo da reação, uma vez que os iniciadores para PCR foram obtidos a partir do genoma bovino. **DNA de hamster:** DNA controle para a seleção dos marcadores para o mapeamento utilizando o painel BBURH<sub>5000</sub>. **C:** controle negativo da reação de PCR (sem adição de DNA). As setas vermelhas indicam as temperaturas de anelamento selecionadas para cada par de iniciadores para PCR: 59°C para AGP, 56°C para RBP4 e 65°C para AMBP, APOM, FABP1, FABP3, FABP5, FABP6 e PMP2.

Todos os pares de iniciadores para PCR dos genes das lipocalinas amplificam um produto de PCR único em búfalo e ausente em hamster, sendo possível selecionar uma temperatura de anelamento para estudos posteriores.

Os marcadores AGP e RBP4 apresentaram um produto de PCR robusto com o DNA de búfalo nas temperaturas de 59°C e 56°C respectivamente. Os demais marcadores (AMBP, APOM, FABP1, FABP3, FABP5, FABP6 e PMP2) embora tenham amplificado produto de PCR com DNA de búfalo em todas as temperaturas testadas, optou-se por selecionar a temperatura mais elevada (65°C) para posteriores estudos.

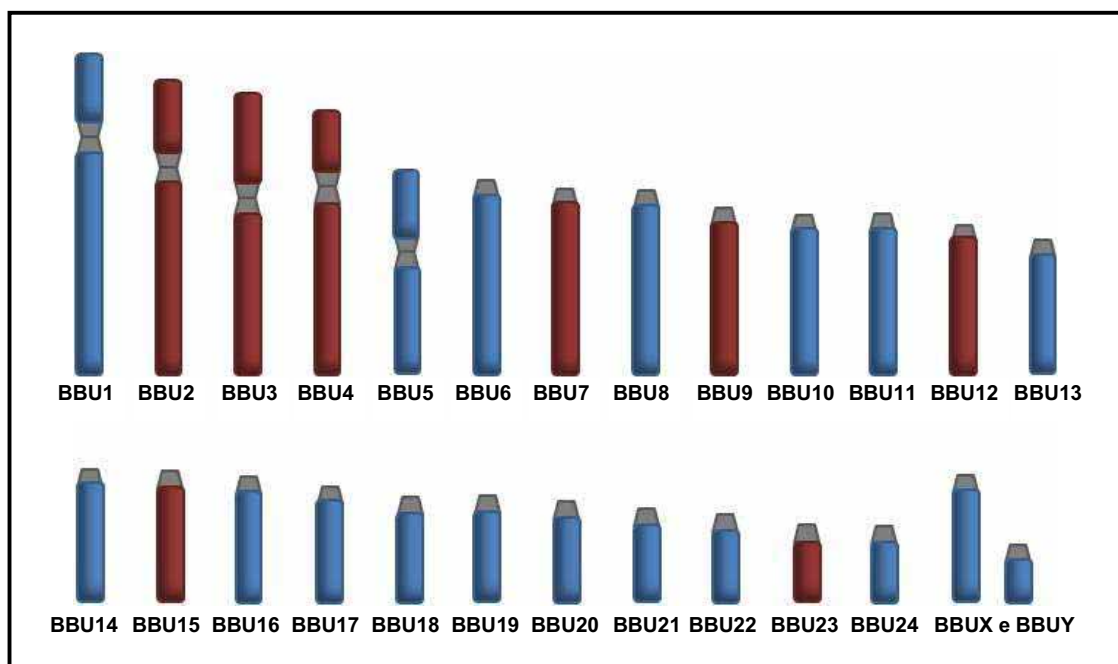
A Figura 3 representa o resultado dos experimentos com gradiente de temperatura de anelamento com pares de iniciadores para PCR que amplificam genes do MHC. As temperaturas de anelamento também foram selecionadas de acordo com o padrão dos produtos de PCR descrito para os genes anteriores. As temperaturas selecionadas estão indicadas pelas setas em vermelho na Figura 3.

Todos os 58 pares de iniciadores para PCR provenientes de estudos do genoma bovino, uma vez otimizados com DNA bubalino tornaram-se também marcadores para búfalo, podendo ser utilizados em diversos tipos de estudos como, por exemplo, o mapeamento utilizando o painel BBURH<sub>5000</sub> e a caracterização desses genes utilizando a biblioteca BAC de búfalo.



**Figura 3.** Géis de agarose 2% corado com brometo de etídio, mostrando o resultado das reações de PCR com gradiente de temperatura de anelamento com os iniciadores para PCR de marcadores da região MHC. **DNA de boi:** controle positivo da reação, uma vez que os iniciadores para PCR foram desenhados a partir do genoma bovino. **DNA de hamster:** DNA controle para a seleção dos marcadores para o mapeamento utilizando o painel BBURH<sub>5000</sub>. **C:** controle negativo da reação de PCR (sem adição de DNA). As setas vermelhas indicam a temperatura de anelamento de 65°C selecionadas para esses pares de iniciadores para PCR.

A Figura 4 ilustra um ideograma do genoma bubalino, mostrando os primeiros cromossomos bubalinos que podem ser explorados com a biblioteca BAC (em vermelho), a partir da seleção e otimização dos marcadores descritos. Os marcadores relacionados com produção e qualidade de leite estão descritos como pertencentes aos cromossomos bubalinos 4, 7 e 12; os genes das lipocalinas nos cromossomos 2, 3, 9, 12, 15 e 23) e os genes do MHC no braço curto do cromossomo 2 (BBU2p).

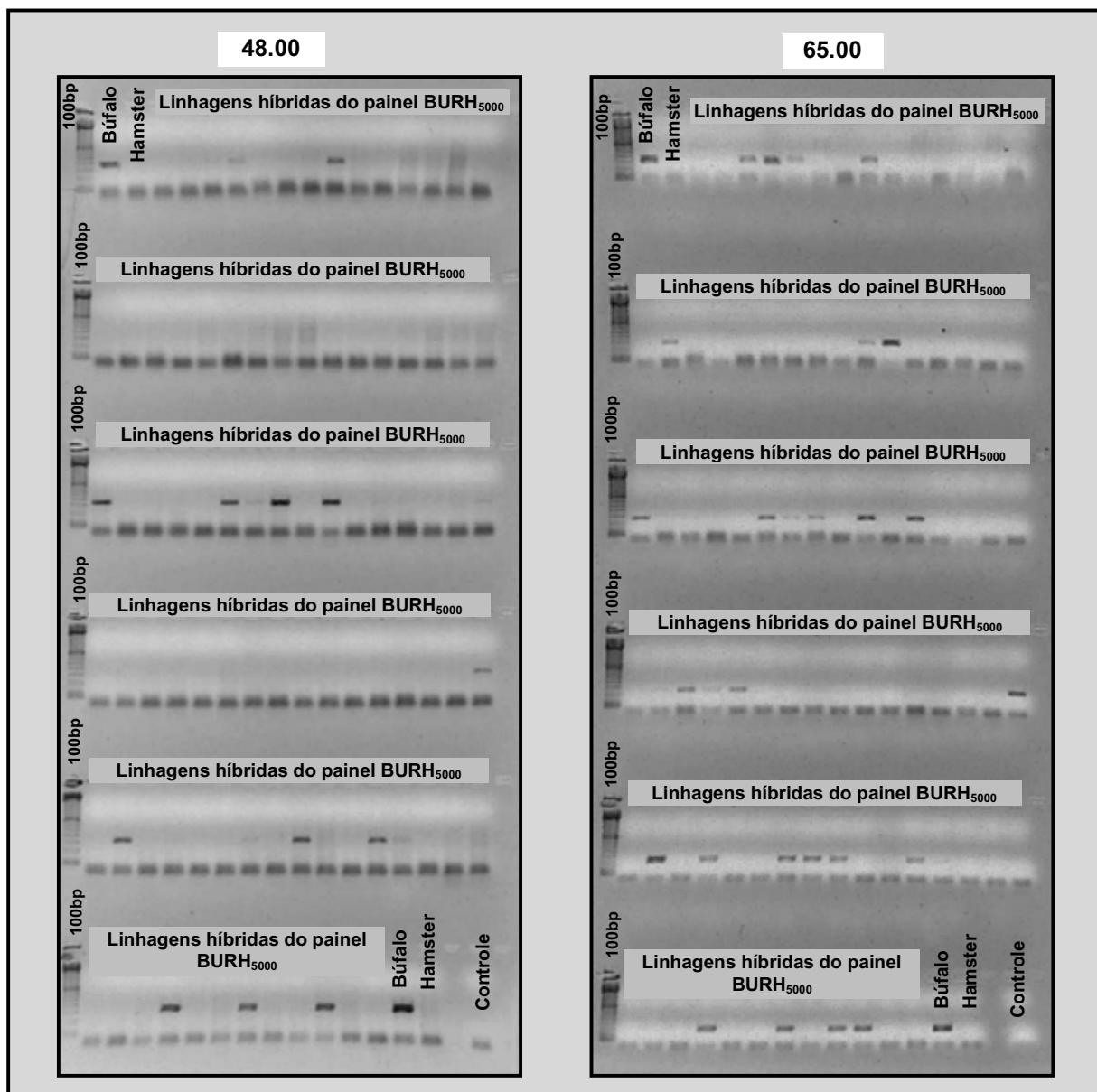


**Figura 4.** Ideograma do cariótipo do búfalo, onde os primeiros cromossomos a terem regiões caracterizadas na biblioteca BAC de búfalo se encontram em vermelho. **BBU2:** genes do MHC, APOM e FABP3; **BBU3:** AMBP e AGP; **BBU4:** LALBA; **BBU7:** caseínas CSN1S1, CSN1S2, CSN2 e CSN3; **BBU9:** FABP6; **BBU12:** LGB e FABP1; **BBU15:** FAPB5 e PMP2; **BBU23:** RBP4.

## 2. Mapeamento utilizando o painel **BBURH<sub>5000</sub>**

Devido ao grande número de marcadores do MHC que se mostraram adequados para estudos no genoma bubalino e a importância econômica desses genes para a espécie, seguiu-se com os experimentos de mapeamento desses marcadores com o

painel BBURH<sub>5000</sub>. A Figura 5 ilustra o resultado das reações de PCR dos marcadores 48.00 e 65.00 com o DNA das 90 linhagens do painel BBURH<sub>5000</sub>.



**Figura 5.** Géis de agarose 2% corado com brometo de etídio, mostrando o resultado das amplificações por PCR dos marcadores 48.00 e 65.00 nas 90 linhagens do painel BBURH<sub>5000</sub> com temperatura de anelamento de 65°C. O controle da reação de PCR consiste dos componentes necessários para uma reação de PCR sem adição de DNA.

O padrão de amplificação do produto de PCR nas 90 linhagens do painel BBURH<sub>5000</sub> foi registrado em uma tabela do tipo Excel, a qual foi analisada com o pacote estatístico

CarthaGene (DE GIVRY et al., 2005), adicionando-se também os dez genes previamente mapeados nessa região (DRB, DRB2, DRB3, DRA, DQA, DQA2, DYA, DYB, PSMB9 e DMA), incluídos no mapa preliminar (RODRIGUES FILHO et al., 2008).

O mapa de alta resolução do MHC em búfalo foi construído com um total de 53 marcadores, onde todos tiveram suas posições definidas em *centiRays* (cR) com um *lod score* maior ou igual a 8 (LOD 8). Os 53 marcadores foram distribuídos em dois grupos de ligação (LG1 e LG2). O grupo de ligação 1 (LG1) com 199,9 cR<sub>5000</sub>, apresenta 32 marcadores, dos quais oito são genes de classe I e classe I estendida (65.10, 65.15, 65.05, 65.00, 48.00, 48.05, 48.15 e 48.25), 16 da classe IIa e IIa estendida (53.25, 53.05, 52.55, 52.00, 50.15, DQA, DQA2, DRA, DRB, DRB3, 55.00, 55.05, 55.15, 55.30, 56.00 e DRB2) e oito da classe III (61.00, 57.40, 59.05, 57.20, 57.10, 57.30, 57.00 e 58.00). Já o grupo de ligação 2 (LG2) está representado por 21 marcadores, todos genes de classe IIb (10.20, 10.40, 10.45, 11.05, 11.00, DNA, 13.00, 13.05, DMA, 12.05, 11.10, 10.60, PSMB9, DYB, 10.30, DYA, 10.25, 10.05, 10.00, 10.50 e 10.55), totalizando 269,6 cR<sub>5000</sub>.

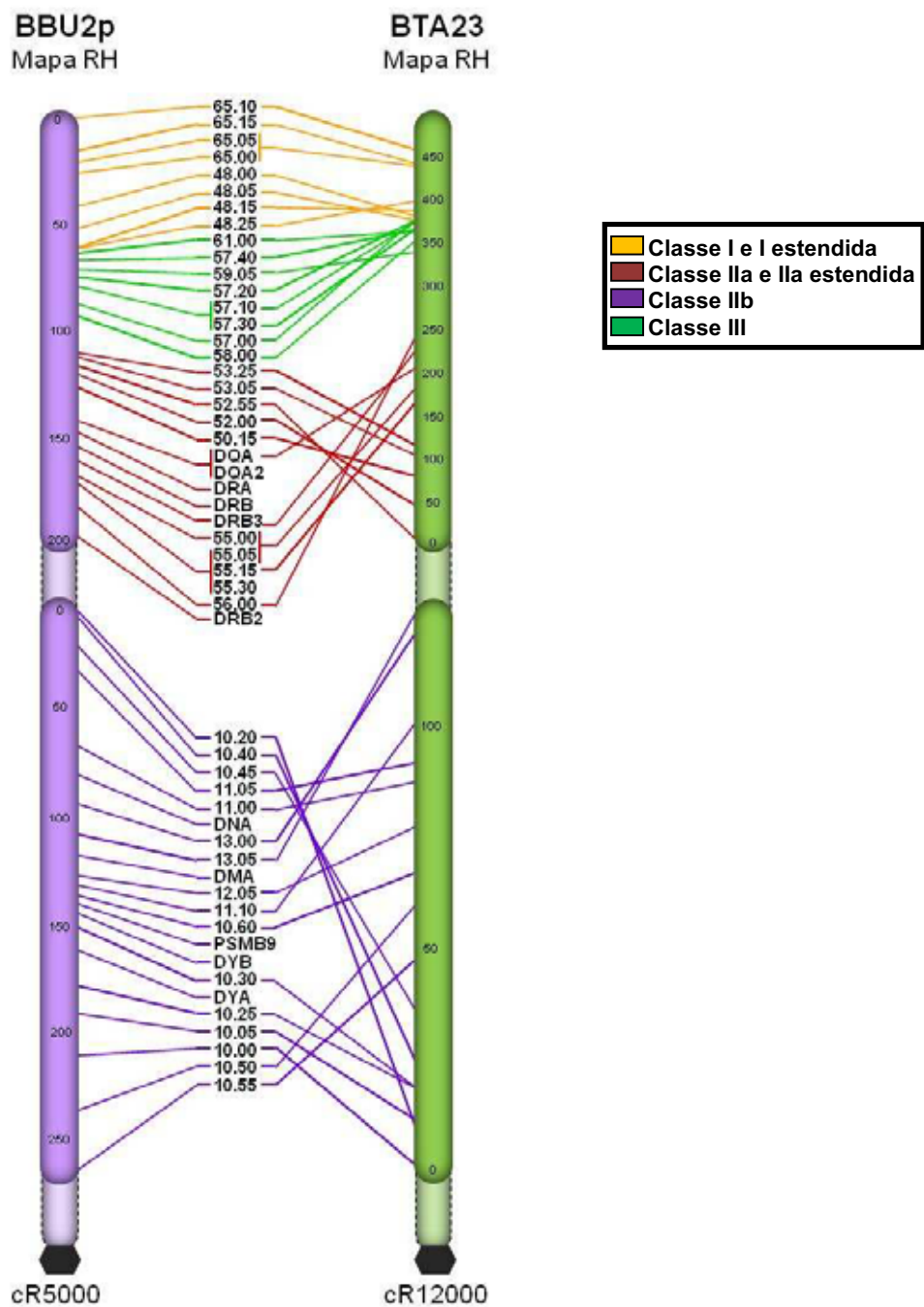
A frequência de retenção para cada marcador foi calculada pela razão entre o número de linhagens positivas do painel BBURH<sub>5000</sub> para o produto de PCR e o número total de linhagens do mesmo (90 linhagens). O valor da frequência de retenção para cada marcador apresentou variação de 25,5% (marcador 56.00) a 57,8% (marcador 12.05).

A tabela 6 apresenta as posições (indicadas em *centiRays*) de cada um dos marcadores mapeados, além da frequência de retenção (FR%) apresentada pelo marcador no painel BBURH<sub>5000</sub> e o grupo de ligação (LG) que cada marcador pertence.

**Tabela 6.** Marcadores do MHC mapeados no braço curto do cromossomo 2 bubalino, com suas respectivas posições em *centiRay* (cR), grupo de ligação (LG) e frequência de retenção (FR%)

<b>Marcador</b>	<b>Posição (cR)</b>	<b>LG</b>	<b>FR (%)</b>	<b>Marcador</b>	<b>Posição (cR)</b>	<b>LG</b>	<b>FR (%)</b>
10.00	208,4	2	50.0	55.15	171,7	1	31.5
10.05	189,3	2	55.5	55.30	171,7	1	31.5
10.20	0	2	34.4	56.00	183,7	1	25.5
10.25	176,0	2	54.4	57.00	88,2	1	36.6
10.30	149,3	2	52.2	57.10	78,8	1	34.4
10.40	4,4	2	33.3	57.20	75,8	1	34.4
10.45	19,5	2	37.8	57.30	78,8	1	34.4
10.50	233,8	2	38.9	57.40	66,6	1	33.3
10.55	269,6	2	31.1	58.00	92,8	1	40.0
10.60	134,3	2	56.7	59.05	71,2	1	32.2
11.00	66,4	2	42.2	61.00	63,0	1	36.6
11.05	32,3	2	37.8	65.00	25,9	1	35.5
11.10	130,0	2	55.5	65.05	20,6	1	41.1
12.05	125,6	2	57.8	65.10	0	1	30.0
13.00	92,7	2	42.2	65.15	15,9	1	36.7
13.05	106,2	2	46.7	BoLA-DMA	115,5	2	50.0
48.00	40,9	1	26.7	BoLA-DNA	79,2	2	25.6
48.05	51,5	1	26.7	BoLA-DQA	143,1	1	34.4
48.15	61,8	1	32.2	BoLA-DQA2	143,1	1	34.4
48.25	61,9	1	35.5	BoLA-DRA	148,3	1	33.3
50.15	127,2	1	38.8	BoLA-DRB	154,6	1	30.0
52.00	120,0	1	41.1	BoLA-DRB2	199,9	1	25.5
52.55	117,1	1	37.7	BoLA-DRB3	161,9	1	32.2
53.05	112,5	1	44.4	BoLA-DYA	159,8	2	48.8
53.25	110,2	1	37.7	BoLA-DYB	143,2	2	51.1
55.00	169,2	1	32.2	PSMB9	138,8	2	51.1
55.05	171,7	1	31.5				

A Figura 6 representa a análise comparativa realizada entre os mapas de alta resolução do MHC de búfalo e boi (BRINKMEYER-LANGFORD et al., 2009), os quais apresentaram 43 marcadores em comum.



**Figura 6.** Comparação do mapa RH de alta resolução da região MHC de búfalo (BBU2p) com o mapa RH bovino (BTA23). Os marcadores em comum de ambos os mapas estão ligados, onde as linhas amarelas ligam os marcadores de classe I e classe I estendida, as linhas vermelhas ligam os marcadores de classe IIa e IIa estendida, as linhas roxas ligam os marcadores de classe IIb e as linhas verdes ligam os marcadores de classe III. As posições no BBU2p e BTA23 estão em *centiRay*.

Os genes do MHC em búfalo apresentam uma organização molecular similar à descrita em bovinos, uma vez que os genes da classe II estão separados em dois grandes *clusters*: IIa e IIb. Os genes do *cluster* IIa se encontram associados com genes das classes I e III, enquanto que os genes do *cluster* IIb se encontram separados dos demais genes do MHC.

Algumas alterações quanto à ordem dos marcadores são observadas dentro de cada uma das classes. Esses pequenos rearranjos observados na organização gênica sugerem uma rápida evolução desses genes após a divergência de um ancestral comum (KELLEY et al., 2005). Após a divergência partindo de um ancestral comum, os genes do MHC sofrem várias alterações rápidas por meio de duplicação gênica e do surgimento de polimorfismos, apresentando uma variação inclusive em espécies próximas (KULSKI et al., 2002; SHIINA et al., 1999; WATKINS 1995). Diferentes pressões de parasitas e infecções a quais essas espécies estão sujeitas, podem ser responsáveis por essas rápidas diversificações (HURT et al., 2004; VOGEL et al., 1999).

Análise detalhada do MHC de vertebrados vem revelando uma conservação de sintenia, embora haja diferenças em tamanho, organização e número de genes entre as espécies (KELLEY et al., 2005). A organização do MHC em *clusters* poderia ser explicada por uma possível vantagem para co-evolução ou regulação (TROWSDALE, 2002). Já a organização dos genes dentro dos respectivos *clusters*, como pode ser muito diferente entre as espécies, sugere que houve uma rápida evolução dos genes MHC após a divergência de um ancestral comum (KELLEY et al., 2005).

Tem sido sugerido que a organização molecular do MHC reflete um significado funcional. A alta resolução obtida com mapeamento de RH será uma ferramenta valiosa para o mapeamento comparativo com outras espécies de mamíferos e para o isolamento

de genes candidatos para características economicamente importantes (BAND et al., 1998). Com o seqüenciamento do genoma bubalino iniciado, o mapa RH de alta resolução do MHC de búfalo também poderá auxiliar na montagem e alinhamento das seqüências dessa região.

## **Conclusões**

O presente trabalho gerou 58 novos marcadores para estudos no genoma bubalino, sendo seis genes relacionados com produção e qualidade de leite, nove genes da família das lipocalinas e 43 do complexo principal de histocompatibilidade. Esses novos marcadores podem ser utilizados tanto para estudos de mapeamento quanto para a caracterização desses genes com a biblioteca BAC de búfalo.

Devido ao grande número de marcadores do MHC que amplificaram produto de PCR com o DNA de búfalo, foi gerado um mapa de alta resolução contendo 53 marcadores. A comparação com o mapa MHC bovino evidenciou algumas alterações quanto à ordem dos marcadores dentro de cada uma das classes.

O mapa comparativo descrito no presente trabalho pode auxiliar na compreensão dos processos evolutivos envolvendo o MHC bubalino. Em conjunto com estudos funcionais, poderá proporcionar uma melhor compreensão sobre as diferenças entre búfalos e bovinos em relação à resposta imune, auxiliando o desenvolvimento de estratégias de controle de doenças.

## Referências Bibliográficas\*

AKERSTROM, B.; FLOWER, D. R.; SALIER, J-P. Lipocalins: unity in diversity. **Biochimica et Biophysica Acta**, v. 1482, p. 1-8, 2000.

AMARAL, M.E.J. et al. A first generation whole genome RH map of the river buffalo with comparison to domestic cattle. **BMC Genomics**, v. 9, 2008.

AMARAL, M.E.J. et al. Construction of a river buffalo (*Bubalus bubalis*) whole-genome radiation hybrid panel and preliminary RH mapping of chromosomes 3 and 10. **Animal Genetics**, v. 38, p. 311-314, 2007.

AMARAL, F.R. et al. Qualidade do leite de búfalas: composição. **Revista Brasileira de Reprodução Animal**, v. 29, p. 106-110, 2005.

AMILLS, M. et al. The major histocompatibility complex of ruminants. **Revue Scientifique Technique**, v. 17, p. 108-120, 1998.

ARAVINDAKSHAN, T.V.; NAINAR, A.M.; SIVASELVAM, S.N. Polymorphism in exon 2 of the BuLA-DRB3 gene in Indian buffalo (*Bubalus bubalis* var. *indicus*) detected by PCR-RFLP. **Animal Science**, v. 70, p. 221-226, 2000.

BAND, M. et al. A radiation hybrid map of BTA23: identification of a chromosomal rearrangement leading to separation of the cattle MHC class II subregions. **Genomics**, v. 53, p. 269-275, 1998.

BESANÇON, R. et al. Mutation and Polymorphism Report. **Human Mutation**, v. 17, 2001.

BERGGARD, T. et al. Histologic distribution and biochemical properties of alpha 1-microglobulin in human placenta. **American Journal of Reproductive Immunology and Microbiology**, v. 41(1), p. 52-60, 1999.

BIONAZ, M.; LOOR, J.J. ACSL1, AGPAT6, FABP3, LPIN1, and SLC27A6 are the most abundant isoforms in bovine mammary tissue and their expression is affected by stage of lactation. **The Journal of Nutrition**, v. 138, p. 1019-1024, 2008a.

BIONAZ M.; LOOR J. Gene networks driving bovine milk fat synthesis during the lactation cycle. **BMC Genomics**, v. 9, 2008.

BIRCH, J. et al. Generation and maintenance of diversity in the cattle MHC class I region. **Immunogenetics**, v. 58, p. 670-679, 2006.

BLOMHOFF, R.; BLOMHOFF, H.K. Overview of retinoid metabolism and function. **Journal of Neurobiology**, v.66, p.606-630, 2006.

\*Associação Brasileira de Normas Técnicas (ABNT). NBR6023:2002, Informação e documentação - Referências - Elaboração, Rio de Janeiro, Agosto de 2002.

- BORRIELLO, G. et al. Genetic resistance to *Brucella abortus* in the water buffalo (*Bubalus bubalis*). **Infection and Immunity**, v. 74, p. 2115-2120, 2006.
- BRATT, T. Lipocalins and cancer. **Biochimica et Biophysica Acta**, v. 1482, p. 318-326, 2000.
- BRINKMEYER-LANGFORD, C.L. et al. A high resolution RH map of the bovine major histocompatibility complex. **BMC Genomics**, v. 10, p. 1-10, 2009.
- BROWN, J.A. Expression of retinol-binding protein and cellular retinol-binding protein in the bovine ovary. **Molecular Reproduction and Development**, v. 64, p. 261-269, 2003.
- BROWN, J.H. et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. **Nature**, v. 364, p. 33-39, 1993.
- CARDOSO, K.I. et al. Occurrence of rabies in buffaloes (*Bubalus bubalis*) in the city of Mazagão, state of Amapá Brazil – Case report. In: WORLD BUFFALO CONGRESS, 9., Buenos Aires, Argentina. **Anais...** 2010. p. 473-474.
- CECILIANI, F. et al. Differential expression and secretion of  $\alpha$ 1-acid glycoprotein in bovine milk. **Journal of Dairy Research**, v. 74, p. 374-380, 2007.
- CHILDERS, C.P. et al. Comparative analysis of the bovine MHC class II sequence identifies inversion breakpoints and three unexpected genes. **Animal Genetics**, v. 37, p. 121-129, 2005.
- CHO, K.H. et al. Association of genetic variants for FABP3 gene with back fat thickness and intramuscular fat content in pig. **Molecular Biology Reports**, in press.
- CLARE, D.A.; SWAISGOOD, H.E. Bioactive Milk Peptides: A Prospectus. **Journal of Dairy Science**, v. 83, p. 1187-1195, 2000.
- DAVIS, W.C. et al. The use of crossreactive monoclonal antibodies to characterize the immune system of the water buffalo (*Bubalus bubalis*). **Journal of Veterinary Science**, v. 2, p. 103-109, 2001.
- DE, S.; SINGH, R.K.; BUTCHAIHAH, G. MHC-DRB exon 2 allele polymorphism in Indian river buffalo (*Bubalus bubalis*). **Animal Genetics**, v. 33, p. 215-219, 2002.
- DE GIVRY, S. et al. CarthaGene: multipopulation integrated genetic and radiated hybrid mapping. **Bioinformatics**, v. 21, p. 1703-1704, 2005.
- DEAKIN, J.E. et al. Evolution and comparative analysis of the MHC Class III inflammatory region. **BMC Genomics**, v. 7, 2006.

DI MEO, G.P. et al. An extended river buffalo (*Bubalus bubalis*, 2n = 50) cytogenetic map: assignment of 68 autosomal loci by FISH-mapping and R-banding and comparison with human chromosomes. **Chromosome Research**, v. 16, p. 827-837, 2008.

DORROCH, U. et al. Isolation and characterization of hepatic and intestinal expressed sequence tags potentially involved in trait differentiation between cows of different metabolic type. **Mammalian Genome**, v. 12, p. 528-537, 2001.

DUAN, J.; DAHLBACK, B.; VILLOUTREIX, B. O. Proposed lipocalin fold for apolipoprotein M based on bioinformatics and site-directed mutagenesis. **FEBS Letters**, v. 499, p. 127-132, 2001.

EL NAHAS, S.M.; DE HONT, H.A.; WOMACK, J.E. Current status of river buffalo (*Bubalus bubalis* L.) gene map. **The Journal of Heredity**, v. 92, p. 221-225, 2001.

ELLIS, S. The cattle major histocompatibility complex: is it unique?. **Veterinary Immunology and Immunopathology**, v. 102, p. 1-8, 2004.

FEINGOLD, K.R. et al. Infection and inflammation decrease apolipoprotein M expression. **Atherosclerosis**, v. 199, p. 19-26, 2008.

FLOWER, D.R. The lipocalin protein family: structure and function. **Biochemical Journal**, v. 318, p. 1-14, 1996.

GOLDAMMER, T. et al. A radiation hybrid map of river buffalo (*Bubalus bubalis*) chromosome 7 and comparative mapping to the cattle and human genomes. **Cytogenetic and Genome Research**, v.119, p. 235-241, 2007.

GRAHAM, T.E. et al. Retinol-binding protein 4 and insulin resistance in lean, obese, and diabetic subjects. **The New England Journal of Medicine**, v. 354, p. 2552-2563, 2006.

GU, X.; NEI, M. Locus specificity of polymorphic alleles and evolution by a birth-and-death process in mammalian MHC genes. **Molecular Biology and Evolution**, v. 16, p. 147-156, 1999.

GUANZIROLI, M.C. et al. Isolation of *Mycobacterium bovis* from water buffalo (*Bubalus bubalis*, Linne 1758) of the north east of Argentina. In: **World Buffalo Congress**, 9., Buenos Aires, Buenos Aires, Argentina. **Anais...** 2010. p. 467-469.

GUSTAFSON, A.L. et al. An ordered BAC contig map of the equine major histocompatibility complex. **Cytogenetic and Genome Research**, v. 102, p. 189-195, 2003.

GUTIÉRREZ, G.; GANFORNINA, M. D.; SÁNCHEZ, D. Evolution of the lipocalin family as inferred from a protein sequence phylogeny. **Biochimica et Biophysica Acta**, v. 1482, p. 35-45, 2000.

HERBERG, J.A.; BECK, S.; TROWSDALE, J. TAPASIN, DAXX, RGL2, HKE2 and four new genes (BING1, 3 to 5) form a dense cluster at the centromeric end of the MHC. **Journal of Molecular Biology**, v. 277, p. 839-857, 1998.

HOCHEPIED, T. et al.  $\alpha$ 1-Acid glycoprotein: an acute phase protein with inflammatory and immunomodulatory properties. **Cytokine and Growth Factors Reviews**, v. 54, p. 25-34, 2003.

HUPPERTZ, T. et al. High pressure-induced changes in bovine milk proteins: A review. **Biochimica et Biophysica Acta**, v. 1764, p. 593-598, 2006.

HURT, P. et al. The genomic sequence and comparative analysis of the rat major histocompatibility complex. **Genome Research**, v. 14, p. 631-639, 2004.

IANELLA, P. et al. First radiation hybrid map of the river buffalo X chromosome (BBUX) and comparison with BTAX. **Animal Genetics**, v. 39, p. 196-200, 2008.

IANNUZZI, L. et al. The river buffalo (*Bubalus bubalis*, 2n=50) cytogenetic map: assignment of 64 loci by fluorescence in situ hybridization and R-banding. **Cytogenetic and Genome Research**, v. 102, p. 65-75, 2003.

IANNUZZI, L. Standard karyotype of the river buffalo (*Bubalus bubalis* L., 2n=50). Report of the committee for the standardization of banded karyotypes of the river buffalo. **Cytogenetic and Cell Genetics**, v. 67, p. 2-113, 1994.

IANNUZZI, L. et al. Chromosomal localization of the major histocompatibility complex in cattle and river buffalo by fluorescent in situ hybridization. **Hereditas**, v. 118, p. 187-190, 1993.

JENG, S. et al. Characterization and Partial Purification of Bovine  $\alpha$ -Lactalbumin and  $\beta$ -Casein Produced in Milk of Transgenic Mice. **Journal of Dairy Science**, v. 80, p. 3167-3175, 1997.

KELLEY, J.; WALTER, L.; TROWSDALE, J. Comparative genomics of major histocompatibility complexes. **Immunogenetics**, v. 56, p. 683-695, 2005.

KITAMURA, K. et al. The complete amino acid sequence of the p2 protein in bovine peripheral nerve myelin. **FEBS Letters**, v. 115, p. 27-30, 1980.

KLEIN, J.; SATO, A. Birth of the Major Histocompatibility Complex. **Scandinavian Journal of Immunology**, v. 47, p. 199-209, 1998.

KLEIN, J. et al. The accordion model of MHC evolution. **Progress in Immunology**, v. 8, p. 137-143, 1993.

KOCHAN, K.J. et al. Application of dissociation curve analysis to radiation hybrid panel marker scoring: generation of a map of river buffalo (*B. bubalis*) chromosome 20. **BMC Genomics**, v. 9, p. 1-12, 2008.

KONTOPIDIS, G. et al. Invited Review:  $\beta$ -Lactoglobulin: Binding Properties, Structure, and Function. **Journal of Dairy Science**, v. 87, p. 785-796, 2004.

KOSTIA, S. Genomic evolution and diversity in Artiodactyla. Dissertação em Genética, Department of Biosciences, Division of Genetic University of Helsinki, p. 1-47, 2000.

KULSKI, J.K. et al. Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity, and complexity from shark to man. **Immunological Review**, v. 190, p. 95-122, 2002.

LANGE, K. et al. Statistical methods for polyploid radiation hybrid mapping. **Genome Research**, v. 5, p. 136-150, 1995.

LARSSON, J. et al. The lipocalin alpha(1)-microglobulin binds heme in different species. **Archives of Biochemistry and Biophysics**, v. 432, p. 196-204, 2004.

LEMAY, D.G. et al. The bovine lactation genome: insights into the evolution of mammalian milk. **Genome Biology**, v. 10:R43, 2009.

LEWIN, H.; RUSSEL, G.; GLASS, E. Comparative organization and function of the major histocompatibility complex of domesticated cattle. **Immunological Review**, v. 167, p. 145-158, 1999.

LOGDBERG, L.; WESTER, L. Immunocalins: a lipocalin subfamily that modulates immune and inflammatory responses. **Biochimica et Biophysica Acta**, v. 1482, p. 284-297, 2000.

LIU, K. et al. Closing a gap in the physical map of the physical map of the ovine major histocompatibility complex. **Animal Genetics**, 2010.

LUNNEY, J.K. et al. Molecular genetics of the swine major histocompatibility complex, the SLA complex. **Developmental and Comparative Immunology**, v. 33, p. 362-374, 2009.

LUO, G. et al. Apolipoprotein M. **Lipids in Health and Disease**, v. 3, 2004.

MACKENZIE, S.H. et al. Estradiol and progesterone regulation of retinol-binding protein (RBP) secretion and mRNA expression in bovine endometrial tissue. **Biology of Reproduction**, v. 54, p. 79, 1996.

MACKENZIE, S.H. et al. Bovine endometrial retinol-binding protein secretion, messenger ribonucleic acid expression, and cellular localization during the estrous cycle and early pregnancy. **Biology of Reproduction**, v. 57, p. 1445-1450, 1997.

MADELLA-OLIVEIRA, A.F. et al. Aspectos da comercialização de carne e leite de bubalinos na região Norte Fluminense. **Revista Brasileira de Reprodução Animal**, Belo Horizonte, v. 29, n. 1, p. 53-54, 2005.

MARQUES, J.R.F. **Búfalos: o produtor pergunta, a Embrapa responde**. Brasília: Embrapa, 176p., 2000.

MCNEILL, R. et al. Associations between milk progesterone concentration on different days and with embryo survival during the early luteal phase in dairy cows. **Theriogenology**, v. 65, p. 1435-1441, 2006.

MHC SEQUENCING CONSORTIUM. Complete sequence and gene map of a human major histocompatibility complex. **Nature**, v. 401, p. 921-923, 1999.

MILNER, C.M.; CAMPBELL, R.D. Genetic organization of the human MHC class III region. **Frontiers in Bioscience**, v. 6, p. 914-926, 2001.

MIZIARA, M.N. et al. A radiation hybrid map of river buffalo (*Bubalus bubalis*) chromosome 1 (BBU1). **Cytogenetic and Genome Research**, v. 119, p. 100-104, 2007.

MOHAN, M. et al. Expression of retinol-binding protein messenger RNA and retinoic acid receptors in preattachment bovine embryos. **Molecular Reproduction and Development**, v. 60, p. 289-296, 2001.

MOHAN, M. et al. Expression patterns on retinoid X receptors, retinaldehyde dehydrogenase, and peroxisome proliferator activated receptor gamma in bovine preattachment embryos. **Biology of Reproduction**, v. 66, p. 692-700, 2002.

MOUSE GENOME SEQUENCING CONSORTIUM. Initial sequencing and comparative analysis of the mouse genome. **Nature**, v. 420, p. 520-562, 2002.

NATIONAL RESEARCH COUNCIL. The water buffalo: new prospects for an underutilized animal. Washington DC: National Academy Press, 111 p., 1981.

NEI, M.; GU, X.; SITNIKOVA, T. Evolution by the birth-and-death process in multigene families of vertebrate immune system. **PNAS**, v. 94, p. 7799-7806, 1997.

NEI, M.; ROONEY, A.P. Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics*, v. 39, p. 121-152, 2005.

ODUM, L.; NIELSEN, H.W. Bikunin and alpha 1-microglobulin in human zona pellucida and connective tissue. **Histochemical Journal**, v. 29(3), p. 199-203, 1997.

PAN, G. et al. Molecular characterization, expression profile and association analysis with fat deposition traits of the porcine APOM gene. **Molecular Biology Reports**, v. 37, p. 1363-1371, 2010.

PERMYAKOV, S.E. et al. How to improve nature : study of the electrostatic properties of the surface for  $\alpha$ -lactalbumin. **Protein Engineering, Design & Selection**, v. 18, p. 425-433, 2005.

PRIYANTHA, M.A.R. et al. Serological detection of *Leptospira* serovars from aborted water buffaloes in Sri Lanka. In: WORLD BUFFALO CONGRESS, 9., Buenos Aires, Argentina. **Anais...** 2010. p. 480-483.

QANBARI, S. et al. A genome-wide scan for signatures of recent selection in Holstein cattle. **Animal Genetics**, v. 41, p. 377-389, 2010.

RAHMAN, M.M. et al. Immune response of buffalo to foot-and-mouth disease subtype "A22" vaccine. **Indian Journal of Animal Science**, v. 57, p. 619-622, 1987.

RIJNKELS, M. Multispecies comparison of the casein gene loci and evolution of casein gene family. **Journal of Mammary Gland Biology and Neoplasia**, v. 7, p. 327-345, 2002.

RODRIGUES FILHO, E.A. et al. Mapping MHC genes in river buffalo. **Developments in Biologicals**, v. 132, p. 343-346, 2008.

ROTHSCHILD, M.F.; SKOW, L.; LAMONT, S.J. The major histocompatibility complex and its role in disease resistance and immune responsiveness. In: Breeding for disease resistance in farm animals. CAB International, 2<sup>nd</sup> Edition, p. 243-252, 1999.

SANCHEZ, D. et al. Expresión of the AMBP gene transcript and its two protein products, alpha (1)-microglobulin and bikunin, in mouse embryogenesis. **Mechanisms of Development**, v. 117, p. 293-298, 2002.

SENA, L. et al. Polymorphisms in MHC-DRA and -DRB alleles of water buffalo (*Bubalus bubalis*) reveal different features from cattle DR alleles. **Animal Genetics**, v. 34, p. 1-10, 2003.

SETH, M. Biomarker discovery in subclinical mycobacterial infections of cattle. **PloS One**, v. 4, e5478, 2009.

SHARMA, S.P. et al. Anaplasma marginale infection in Indian water buffalo (*Bubalus bubalis*). **Indian Journal of Animal Science**, v. 17, p. 105-110, 1978.

SHARMA, S.P. Characterization of Anaplasma marginale infection in buffaloes. **Indian Journal of Animal Science**, v. 57, p. 76-78, 1987.

SHIINA, T. et al. Molecular dynamics of MHC genesis unraveled by sequence analysis of the 1,796,938-bp HLA class I region. **PNAS**, v. 96, p. 13282-13287, 1999.

STAFUZZA, N.B. et al. Preliminary radiation hybrid map for river buffalo chromosome 6 and comparison to bovine chromosome 3. **Animal Genetics**, v. 38, p. 406-409, 2007.

STAFUZZA, N.B. et al. Comparative RH maps of the river buffalo and bovine Y chromosomes. **Cytogenetic and Genome Research**, v. 126, p.132-138, 2009.

STEAR, M.J. et al. The relationships among ecto- and endoparasite levels, class I antigens of the bovine major histocompatibility system, immunoglobulin E levels and weight gain. **Veterinary Parasitology**, v. 34, p. 303-321, 1990.

STEPHENS, R. et al. Gene organization, sequence variation and isochore structure at the centromeric boundary of the human MHC. **Journal of Molecular Biology**, v. 291, p. 789-799, 1999.

SWITONSKI, M. et al. Genetics of fat tissue accumulation in pigs: a comparative approach. **Journal of Applied Genetics**, v. 51, p. 153-168, 2010.

TAKAHASHI, K.; ROONEY, A.P.; NEI, M. Origins and divergence times of mammalian class II MHC gene clusters. **The Journal of Heredity**, v. 91, p. 198-205, 2000.

THREADGILL, D.W.; WOMACK, J.E. Genomic analysis of the major bovine milk protein genes. **Nucleic Acids Research**, v. 18, p. 6935-6942, 1990.

TILG, H. et al. Antiinflammatory properties of hepatic acute phase proteins: preferential induction of interleukin 1 (IL-1) receptor antagonist over IL-1 beta synthesis by human peripheral blood mononuclear cells. **Journal of Experimental Medicine**, v. 178, p. 1629-1636, 1993.

TOMAR, S.S.; TRIPATHI, V.N. Disease spectrum in a herd of Murrah buffaloes. **Indian Veterinary Journal**, v. 64, p. 683-688, 1987.

TOTARO, A. et al. Hereditary hemochromatosis: generation of a transcription map within a refined and extended map of the HLA class I region. **Genomics**, v. 31, p. 319-326, 1996.

TRAUL, D.L. et al. Characterization of Bison bison major histocompatibility complex class IIa haplotypes. **Immunogenetics**, v. 57, p. 845-854, 2005.

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

VOGEL, T.U. et al. Major histocompatibility complex class I genes in primates: co-evolution with pathogens. **Immunological Reviews**, v. 167, p. 327-337, 1999.

VOORRIPS, R.E. MapChart: software for the graphical presentation of linkage maps and QTLs. **The Journal of Heredity**, v. 93, p. 77-78, 2002.

YUHKI, N. et al. Sequences, Annotation and Single Nucleotide Polymorphism of the Major Histocompatibility Complex in the Domestic Cat. **PLoS ONE**, v. 3, e2674, 2008.

WAGNER, J.L. Molecular organization of the canine major histocompatibility complex. **Journal of Heredity**, v. 94, p. 23-26, 2003.

WANG, M. et al. Two novel single nucleotide polymorphisms (SNPs) and 4-bp deletion mutation of *RBP4* gene in Chinese cattle. **Journal of Genetics**, v. 89, p. 233-236, 2010.

WATKINS, D.I. The evolution of major histocompatibility class I genes in primates. **Critical Reviews in Immunology**, v. 15, p. 1-29, 1995.

WOLFRUM, C. et al. Binding of fatty acids and peroxisome proliferators to orthologous fatty acid binding proteins from human, murine, and bovine liver. **Biochemistry**, v. 39, p. 1469-1474, 2000.

VAN MOLLE, W et al. Alpha 1-acid glycoprotein and alpha 1-antitrypsin inhibit TNF-induced but not anti-Fas-induced apoptosis of hepatocytes in mice. **Journal of Immunology**, v. 159, p. 3555-3564, 1997.

XIE, T. et al. Analysis of the gene-dense major histocompatibility complex class III region and its comparison to mouse. **Genome Research**, v. 13, p. 2621-2636, 2003.

ZHANG, X. Y. et al. Specific tissue expression and cellular localization of human apolipoprotein M as determined by in situ hybridization. **Acta Histochemica**, v. 105, p. 67-72, 2003.

ZIMMER, P.A. et al. Brucellosis in buffalos in Formosa, Argentina. In: WORLD BUFFALO CONGRESS, 9., Buenos Aires, Argentina. **Anais...** 2010. p. 441-443.

## Capítulo 4

Geração de marcadores derivados do genoma do búfalo para o gene CSN3



## **CAPÍTULO 4 – GERAÇÃO DE MARCADORES DERIVADOS DO GENOMA DO BÚFALO PARA O GENE CSN3**

**RESUMO:** A produção de leite vem sendo explorada pelo homem há aproximadamente 10 mil anos, representando hoje, junto com seus derivados, um dos principais componentes da dieta humana em várias partes do mundo. As caseínas representam cerca de 80% das proteínas do leite, sendo que a kapa-caseína desempenha um papel fundamental na estabilização das micelas durante o processo de coagulação do leite para produção de queijos. Visando gerar marcadores específicos para búfalo e avaliar potenciais polimorfismos do tipo SNP (do inglês *Single Nucleotide Polymorphism*), uma região de 4,3 Kb desse gene foi seqüenciada em três animais da raça Murrah. Foram produzidos 10 pares de iniciadores para PCR a partir da seqüência gerada, os quais poderão ser utilizados na busca dos clones na biblioteca BAC de búfalo. O seqüenciamento revelou 18 potenciais polimorfismos do tipo SNP, os quais podem ser utilizados em estudos envolvendo rebanhos comerciais, na busca por associação com características de interesse econômico. A seqüência consenso do gene CSN3 gerada no presente trabalho foi comparada com as seqüências dos alelos A e B do gene CSN3 bovino, revelando com ambas as seqüências, uma similaridade de 97% (*e-value* 0.0), apresentando 168 nucleotídeos divergentes entre boi e búfalo. Embora as diferenças observadas sejam exclusivas em íntrons e região reguladora (5'UTR), não pode ser descartada a hipótese de haver elementos reguladores em regiões intrônicas, os quais podem estar associados com as diferenças existentes entre os bóvidos.

**Palavras-chave:** *Bubalus bubalis*, kapa-caseína, polimorfismos, seqüenciamento, SNP

## Introdução

As caseínas são proteínas secretadas nas células das glândulas mamárias, representando as principais proteínas do leite. Suas variantes têm sido extensivamente estudadas por estarem associadas com o desempenho na lactação, composição do leite e eficiência na produção de queijos (ALEANDRI et al., 1990).

A kapa-caseína constitui aproximadamente 12% das caseínas presentes no leite bovino, desempenhando um importante papel na estabilização da estrutura micelar nos coágulos do leite (AZEVEDO et al., 2008; SORIA et al., 2003).

O gene codificante da kapa-caseína (CSN3), embora não relacionado evolutivamente com os genes que codificam as demais caseínas (CSN1S1: alfa-s1, CSN1S2: alfa-s2 e CSN2: beta-caseína) encontra-se fisicamente ligado à esses genes, formando um *cluster* no cromossomo 6 bovino (THREADGILL & WOMACK, 1990).

Mutações de ponto no gene CSN3 em bovinos, encontram-se em sua grande maioria no éxon 4, resultando em onze alelos para esse gene: A, B, C, E, F<sup>1</sup>, F<sup>2</sup>, G<sup>1</sup>, G<sup>2</sup>, H, I e J (FARRELL et al., 2004). A Tabela 1 ilustra as trocas de aminoácidos que ocorrem na proteína kapa-caseína bovina, resultante destas mutações.

**Tabela 1.** Alterações nas posições dos aminoácidos da proteína kapa-caseína, codificada pelos diferentes alelos do gene CSN3 bovino (adaptada de FARREL et al., 2004).

Alelo	Posição da alteração de aminoácido na proteína em relação ao alelo A						
	10	97	104	135	136	148	155
A	Arg	Arg	Ser	Tre	Tre	Asp	Ser
B					Ile	Ala	
C		His					
E							Gli
F <sup>1</sup>						Val	
F <sup>2</sup>	His				Ile	Ala	
G <sup>1</sup>		Cis			Ile	Ala	
G <sup>2</sup>						Ala	
H					Ile		
I			Ala				
J					Ile	Ala	Arg

Ala=alanina; Arg=arginina; Asp=ácido aspártico; Cis=cisteína; Gli=glicina; His=histidina; Ile=isoleucina; Ser=serina; Tre=treonina e Val=valina

Os alelos A e B foram descritos há quase 40 anos por MERCIER e colaboradores (1973), e por serem os mais freqüentes nas populações são, conseqüentemente, os mais amplamente estudados (PRINZENBERG et al., 1999). Em geral, o alelo B tem sido associado com um aumento da concentração da proteína kapa-caseína no leite (HALLÉN et al., 2007; TSIARAS et al., 2005; IKONEN et al., 1997; VAN DEN BERG et al., 1992). Em particular, o aumento da concentração dessa proteína gera micelas menores, o que resulta em uma consistência mais firme do coágulo, contribuindo com um aumento na produção de queijos de aproximadamente 10% (WALSH et al., 1998; MCLEAN, 1987; LIN et al., 1986). Desse modo, para a produção de queijos parece ser mais eficiente aumentar a proporção de kapa-caseína no leite do que a concentração das demais caseínas (HALLÉN et al., 2008).

Em diversos programas de melhoramento, o teste de progênie é realizado selecionando os genótipos favoráveis (BB e AB) para a inseminação artificial (AZEVEDO et al., 2008; MITRA et al., 1998). Genótipos específicos do gene CSN3 têm mostrado um

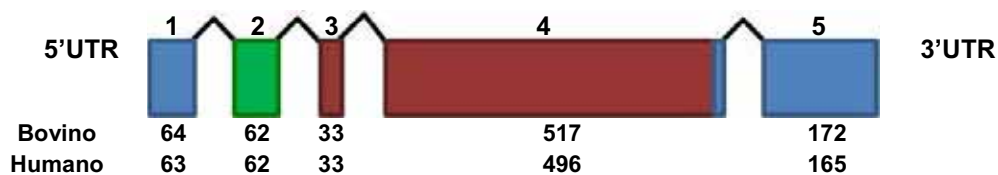
efeito significativo na composição das proteínas do leite em diversos estudos, porém a significância dos demais alelos nas populações bovinas ainda não está bem determinada (HALLÉN et al., 2008; CARDAK et al., 2003; GRAML & PIRCHNER, 2003). Estudos recentes têm associado o alelo E com uma baixa concentração da proteína kapa-caseína, resultando em efeitos adversos sobre as propriedades de coagulação do leite (HALLÉN et al., 2008; 2007).

Além das mutações em regiões codificantes que levam ao surgimento dos diferentes alelos do gene CSN3 em bovino, também tem sido relatada a presença de polimorfismos na região não codificante desse gene, as quais afetam diretamente a taxa de transcrição do mesmo (ROBITAILLE et al., 2005; SZYMANOWSKA et al., 2004a). Um particular interesse nessas seqüências é a presença de elementos regulatórios, os quais influenciam nos processos transcricionais e pós transcricionais dos genes (MAZUMBER et al., 2003; WEDEMEYER et al., 2000).

Polimorfismos tem sido identificados em regiões 5'UTR (do inglês *Untranslated Region*) do gene CSN3 bovino, com alguns deles localizados em seqüências regulatórias, como por exemplo, em sítios de ligação de fatores de transcrição, podendo influenciar a sua expressão (SZYMANOWSKA et al., 2004b). Há sugestão de que a expressão diferenciada dos alelos do gene CSN3 em bovino pode ser resultado de uma combinação entre a variante protéica e polimorfismos nas regiões regulatórias desse gene (VAN EENENNAAM & MEDRANO, 1991).

A organização do gene CSN3 é considerada conservada nas espécies que já possuem esse gene totalmente caracterizado, como por exemplo, em humano, camundongo, bovino e rato. Em todas as espécies mencionadas, esse gene possui cinco éxons e apresenta um tamanho total de 15 Kb, aproximadamente. Porém, a parte

codificante do gene CSN3 apresenta menos de 600 bp, a qual se encontra quase que em sua totalidade no éxon 4. A Figura 1 possui uma representação esquemática da estrutura do gene CSN3 de bovino e humano, com informação em número de nucleotídeos para cada um dos cinco éxons.



**Figura 1.** Estrutura dos éxons do gene CSN3. As porções em azul correspondem às regiões 5' e 3'UTR, a porção verde corresponde a seqüência codificante do peptídeo sinal e às porções em vermelho representam as porções codificantes do peptídeo maduro. Os números acima indicam o número do éxon e os números abaixo indicam o tamanho dos éxons em bovino e humano. Esquema adaptado de RIJNKELS, 2002.

Devido a essa conservação na estrutura do gene CSN3, há relatos de que a pressão seletiva sobre o mesmo parece ser mais forte em comparação com aquela exercida nos outros genes das caseínas, o que pode ser explicado pelo seu importante papel na estrutura das micelas do leite (RIJNKELS, 2002).

Até o momento, há 11 seqüências disponíveis no banco de dados americano NCBI (<http://www.ncbi.nlm.nih.gov/>) para o gene CSN3 de búfalo (AY750857, DQ645429, DQ191174, DQ191173, FJ770200, AM900443, EF133463, EF066482, D14370, AJ628346 e AJ011387), sendo a maioria seqüências de éxons parciais ou completos.

A seqüência dos éxons do gene CSN3 em búfalo já se encontra bem definida, apresentando também cinco éxons, como definido para os outros mamíferos, com um tamanho total de 848 bp. Os éxons 1, 2, 3, 4 e 5 apresentam tamanhos de 64 bp, 62 bp,

33 bp, 517 bp e 172 bp, respectivamente. Porém não há seqüências depositadas de íntrons do gene CSN3 de búfalo. No gene CSN3 bovino, os íntrons 1, 2, 3 e 4 possuem 2424, 5915, 2005 e 1864 nucleotídeos, respectivamente.

A disponibilidade de uma biblioteca BAC de búfalo torna possível não só a caracterização da estrutura molecular desse gene em búfalo, mas também a comparação de sua estrutura com aquelas correspondentes no genoma de outros bovídeos. Porém, para que essa região possa ser caracterizada utilizando a biblioteca BAC, há uma necessidade de gerar marcadores de DNA para identificar e isolar clones específicos da biblioteca. Embora marcadores derivados do genoma bovino tenham sido utilizados com sucesso em estudos de mapeamento do genoma bubalino (AMARAL et al., 2008), genes com seqüências de DNA muito similares, principalmente aqueles genes pertencentes a famílias gênicas, requerem a geração de marcadores específicos de búfalo.

Particularmente, a comparação dos genes codificantes das proteínas do leite entre bovino e bubalino resultará em dados valiosos para o entendimento do funcionamento dos mesmos, permitindo traçar os elementos responsáveis por diferenças tão significativas na qualidade e produção de leite.

## **Objetivos**

Visando gerar marcadores moleculares específicos para o gene CSN3 de búfalo, o presente trabalho teve como objetivos:

- ✓ Seqüenciar regiões não codificantes do gene da CSN3 de búfalo em três animais para a geração de uma seqüência consenso, identificando potenciais regiões de polimorfismos;
- ✓ Comparar a seqüência consenso de búfalo as de bovino depositadas nos bancos de dados públicos.

## Material e Métodos

### 1. Seqüenciamento

O seqüenciamento foi realizado a partir do produto de PCR obtido para três fêmeas da raça Murrah (#739, #764 e #766), provenientes de Vermont (EUA). As reações de PCR foram realizadas em um volume total de 50 µl utilizando 50 ng de DNA; 0,2 mM de cada iniciador para PCR; 100 mM de dNTP; 10 mM de Tris-HCl; 1,5 mM de MgCl<sub>2</sub> e 1U da DNA polimerase AmpliTaqGold™ (*Applied Biosystems* - EUA).

Os ciclos para amplificação foram: desnaturação inicial de 94°C por 10 minutos, seguida de 35 ciclos com desnaturação a 94°C por 30 segundos, anelamento de 50-66°C por 30 segundos (dependendo do par de iniciadores para PCR) e extensão a 72°C por 30 segundos, seguida de uma extensão final a 72°C por 6 minutos. A qualidade dos produtos de PCR obtidos foram visualizados em gel de agarose 2%, corado com brometo de etídio e submetido à eletroforese horizontal por 30 minutos a 80V.

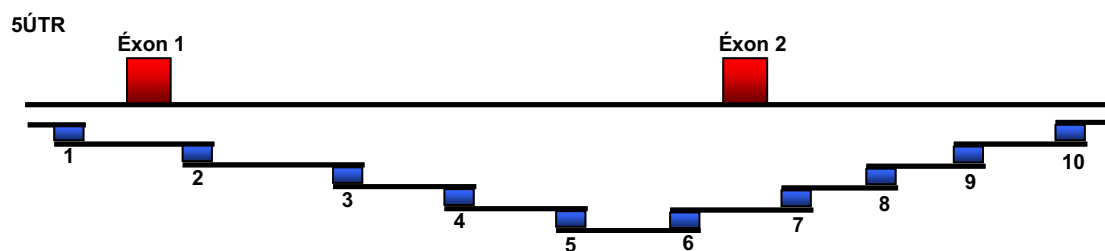
As reações de PCR foram purificadas com o kit PSIΨClone (*Princeton Separations* - EUA), seguindo as indicações do fabricante. As concentrações (ng/µL) foram determinadas por meio de espectrofotômetro (Nanodrop ND1000™ -*Thermo Scientific* - EUA).

As reações de seqüenciamento foram realizadas utilizando o kit BigDye® Terminator v.1.1 (*Applied Biosystems* - EUA). Para cada reação de seqüenciamento foram utilizados 30 ng do DNA amplificado por PCR, 10 µM de iniciador para PCR (senso e anti-senso em reações separadas) e 1,5 µl de tampão de seqüenciamento (5X), os quais foram submetidos à desnaturação no termociclador ABI 2720 (*Applied Biosystems* - EUA) durante 2 minutos a 96°C, sendo posteriormente colocados no gelo por 2 minutos. Em

seguida, foram adicionados 1 µl de Big Dye® em cada uma das reações, às quais foram levadas ao termociclador por 25 ciclos de 96°C por 10 segundos, 50°C por 5 segundos e 60°C por 4 minutos. As reações foram armazenadas em freezer -80°C até o momento da análise no seqüenciador de capilar ABI3130xl (*Applied Biosystems* - EUA) com o software Sequence Analysis 5.3.1.

## 2. *Primer Walking*

As seqüências do gene CSN3 de búfalo foram geradas por *primer walking*, um método passo-a-passo desenvolvido para seqüenciar moldes longos de DNA de uma extremidade a outra, superando o limite da técnica de terminação em cadeia de Sanger, que gera a leitura de somente poucas bases em uma única reação de seqüenciamento. Uma primeira etapa de seqüenciamento é realizada a partir de uma seqüência conhecida e próxima de uma das extremidades do DNA molde de interesse, a qual é utilizada no desenho de um novo iniciador para a próxima reação de seqüenciamento. Este processo é repetido até que a seqüência de interesse inteira seja determinada, conforme ilustrado na Figura 2.



**Figura 2.** Metodologia de *primer walking*, utilizada no sequenciamento do íntron 1 e parte do íntron 2 do gene CSN3 de búfalo. Dez iniciadores estão representados pelos retângulos azuis, os quais estão enumerados de 1-10, enquanto que as seqüências geradas estão representadas pelas linhas que unem os iniciadores. Esquema adaptado de CHARIZOPOULOU et al., 2004.

Os primeiros iniciadores para PCR foram gerados a partir da região promotora 5'UTR e do éxon 1 disponíveis para búfalo, cujos respectivos números de acesso no banco de dados do NCBI são AJ628346 e AY750857. Posteriormente, os iniciadores para PCR foram desenhados a partir das seqüências obtidas.

O programa Primer3 (versão 0.4.0), disponível gratuitamente no endereço eletrônico <http://frodo.wi.mit.edu/primer3/>, foi utilizado para desenhar os iniciadores para PCR, obedecendo aos seguintes parâmetros:

a) quantidade de citosina e guanina em torno de 50% e sempre que possível com uma proporção semelhante de cada um dos nucleotídeos;

b) diferença de temperatura de anelamento entre os iniciadores para PCR inferior a 5°C, cuja temperatura ideal se encontra em torno de 55°C e 60°C;

c) iniciadores para PCR com três ou mais nucleotídeos idênticos seqüenciais foram evitados, principalmente citosinas e guaninas;

d) foram descartados os iniciadores para PCR com mais de duas citosinas ou guaninas nas extremidades 3' e 5', assim como os que apresentaram mais de uma auto-complementaridade na extremidade 3' e com mais de três auto-complementaridade na extremidade 5', evitando assim a formação de dímeros;

e) o tamanho do produto a ser amplificado deve ser inferior a 750 bp, devido a limitações da técnica de seqüenciamento;

Após a construção dos iniciadores para PCR, os mesmos foram analisados com a ferramenta BLAST (do inglês *Basic Local Alignment Search Tool*), disponível no endereço eletrônico <http://blast.ncbi.nlm.nih.gov/Blast.cgi>, para verificação da especificidade com a seqüência alvo.

### 3. Análise dos dados

As seqüências geradas para os três animais foram analisadas com o software Sequencher® 4.10.1 (<http://www.genecodes.com>). Esse programa permitiu a avaliação da qualidade das seqüências geradas, assim como a montagem da seqüência. Também foi realizado o alinhamento múltiplo dessas três seqüências (uma de cada animal), para a identificação de possíveis polimorfismos.

As três seqüências geradas foram também alinhadas para a obtenção de uma seqüência consenso para o gene CSN3 de búfalo. Essa seqüência consenso foi alinhada com as seqüências do gene CSN3 bovino disponíveis no banco de dados americano NCBI, utilizando a ferramenta Clustal W (THOMPSON et al., 1994).

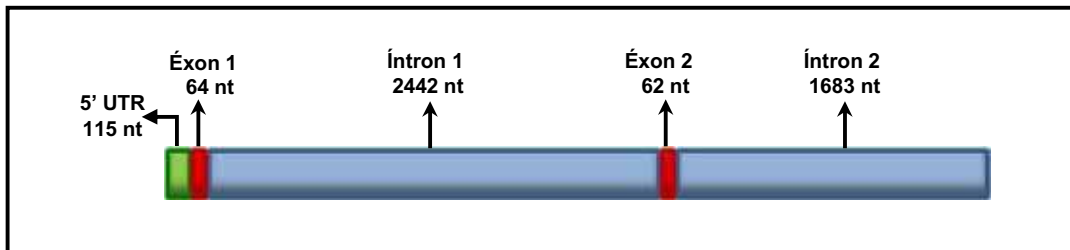
Variações de nucleotídeos entre as seqüências senso e anti-senso em uma mesma posição foram nomeadas de acordo com o código de bases degeneradas NC-IUB (*Nomenclature Committee of the International Union of Biochemistry*), estabelecido há mais de vinte anos (CORNISH-BOWDEN, 1985), o qual se encontra descrito na Tabela 2.

**Tabela 2.** Nomenclatura de variações entre as bases em uma mesma posição, durante o alinhamento de seqüências.

Duas bases		Três bases		Quatro bases
R = A ou G	S = C ou G	V = A, C ou G	D = A, G ou T	N = A,C,G ouT
M = A ou C	Y = C ou T	H = A, C ou T	B = C, G ou T	
W = A ou T	K = G ou T			

## Resultados e Discussão

No presente trabalho foi seqüenciada uma região de aproximadamente 4,3 Kb do gene CSN3 de búfalo para três animais representantes da raça Murrah. Essas seqüências foram alinhadas, gerando uma seqüência consenso com 4366 nucleotídeos, dos quais 115 nucleotídeos são da região 5'UTR, 64 do éxon 1, 2442 do íntron 1, 62 do éxon 2 e 1683 do íntron 2, como ilustrado na Figura 3. Cabe salientar que a região 5'UTR e o éxon 2 foram seqüenciados parcialmente.

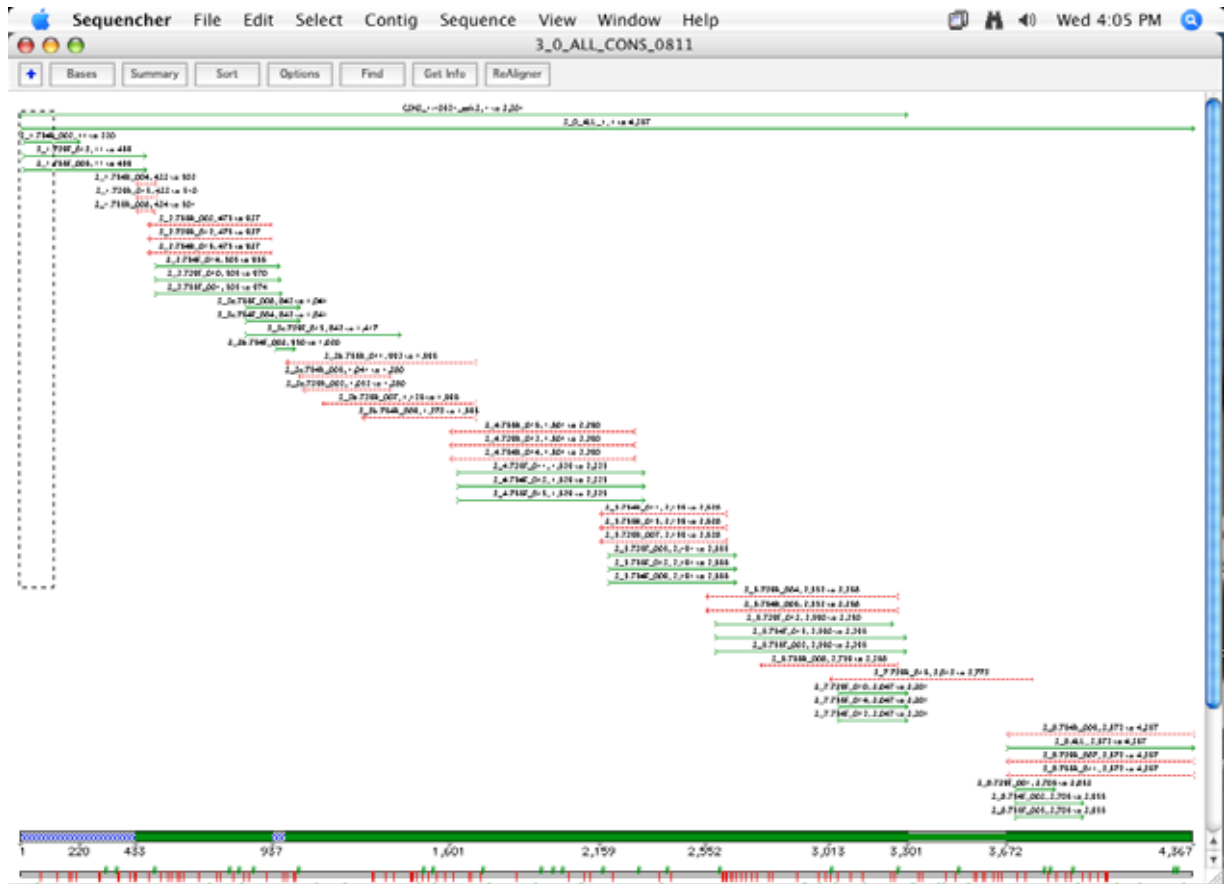


**Figura 3.** Região do gene CSN3 gerada no presente trabalho, com o número de nucleotídeos das respectivas regiões: 5'UTR (115 nt), éxon 1 (64 nt), íntron 1 (2442 nt), éxon 2 (62 nt) e íntron 2 (1683 nt). **nt**=nucleotídeos

Os éxons 1 e 2 apresentam os mesmos tamanhos dos éxons correspondentes do gene CSN3 bovino. Pela primeira vez foi descrita a seqüência completa do íntron 1 e uma seqüência parcial do íntron 2 do gene CSN3 de búfalo. O íntron 1 apresentou 18 nucleotídeos a mais que o íntron 1 bovino. Como as seqüências geradas da região 5'UTR e do íntron 2 foram parciais, não foi possível fazer comparação do número de nucleotídeos dessas regiões com as correspondentes em bovino.

A Figura 4 ilustra a análise e o alinhamento das seqüências geradas para os três animais da raça Murrah, formando uma seqüência consenso para o gene CSN3 de búfalo,

utilizando o software Sequencher® 4.10.1. A seqüência consenso do gene CSN3 de búfalo gerada no presente trabalho se encontra no Apêndice 3.



**Figura 4.** Alinhamento de seqüências geradas do gene CSN3 com o software Sequencher® 4.10.1, formando uma seqüência consenso de 4,3 Kb.

Durante o seqüenciamento foram gerados 10 pares de iniciadores para PCR para o gene CSN3 de búfalo, os quais poderão ser utilizados tanto na busca dos clones da biblioteca BAC que contém o gene CSN3, como para estudos populacionais de variabilidade genética nos rebanhos comerciais. As seqüências dos iniciadores para PCR gerados, assim como o tamanho dos produtos de PCR gerados e as respectivas temperaturas de anelamento estão descritas na Tabela 3.

**Tabela 3.** Pares de iniciadores para PCR gerados durante o seqüenciamento de uma região de 4,3 Kb do gene CSN3 de búfalo.

<b>Marcador</b>	<b>Seqüência dos iniciadores para PCR (5'--3')</b>	<b>Tamanho do amplificado</b>	<b>Posição na seqüência consenso</b>	<b>T*</b>
CSN3_1	F: GCATTCCATTAACCGAGACTGA R: ATTAAGGGATCAATCTGGATTATGCT	598bp	2011-2608bp	66°C
CSN3_2	F: CAAGACAGTGAATCTATTCTACAGAC R: CTTGATGATTACAATACAGTACAGTGA	498bp	2518-3015bp	64°C
CSN3_3	F: CTACAATCTTAAAAGGTAGCCAGT R: ATCTGCTTCTCCAGTAATCATAA	839bp	2945-3783bp	64°C
CSN3_4	F: CTACCCTATACTTGAGTCTTTAGAA R: GCTTATTAAGGTGAATTGGTC	733bp	3631-4363bp	58°C
CSN3_5	F: GCAAGTCCCACCTATATGCTATGTC R: GGTTAATGCCAGGATAGTCACAAC	519bp	4186-4704bp	64°C
CSN3_6	F: CATTGCTCCAAGTCACTCCTA R: GCACTTATTTGCCAACAGTAAGA	755bp	4595-5349bp	64°C
CSN3_7	F: CCATAGAGGAAGATCAGAA R: TCAGCCATTGGTATGTTAT	817bp	5056-5872bp	58°C
CSN3_8	F: CCATGTGTAAAATTAGATAGCCAGT R: GAGGTACAGGTCAGCAGTTGATT	726bp	5716-6441bp	64°C
CSN3_9	F: CAAGTCTGGTTCAGTCTCTT R: GGTATTCCCATCTCTTTATGA	568bp	6214-6781bp	60°C
CSN3_10	F: CACCCTTCTACCATCTTGCT R: GGATGCCATGATCTTCGT	569bp	6570-7138bp	60°C

\*Temperatura de anelamento dos iniciadores para PCR

Foram encontradas 18 alterações na análise das seqüências do gene CSN3 obtidas para os três animais da raça Murrah, as quais estão descritas na Tabela 4. Esses potenciais polimorfismos do tipo SNP são interessantes para serem estudados nas populações bubalinas comerciais, buscando associações com características economicamente importantes.

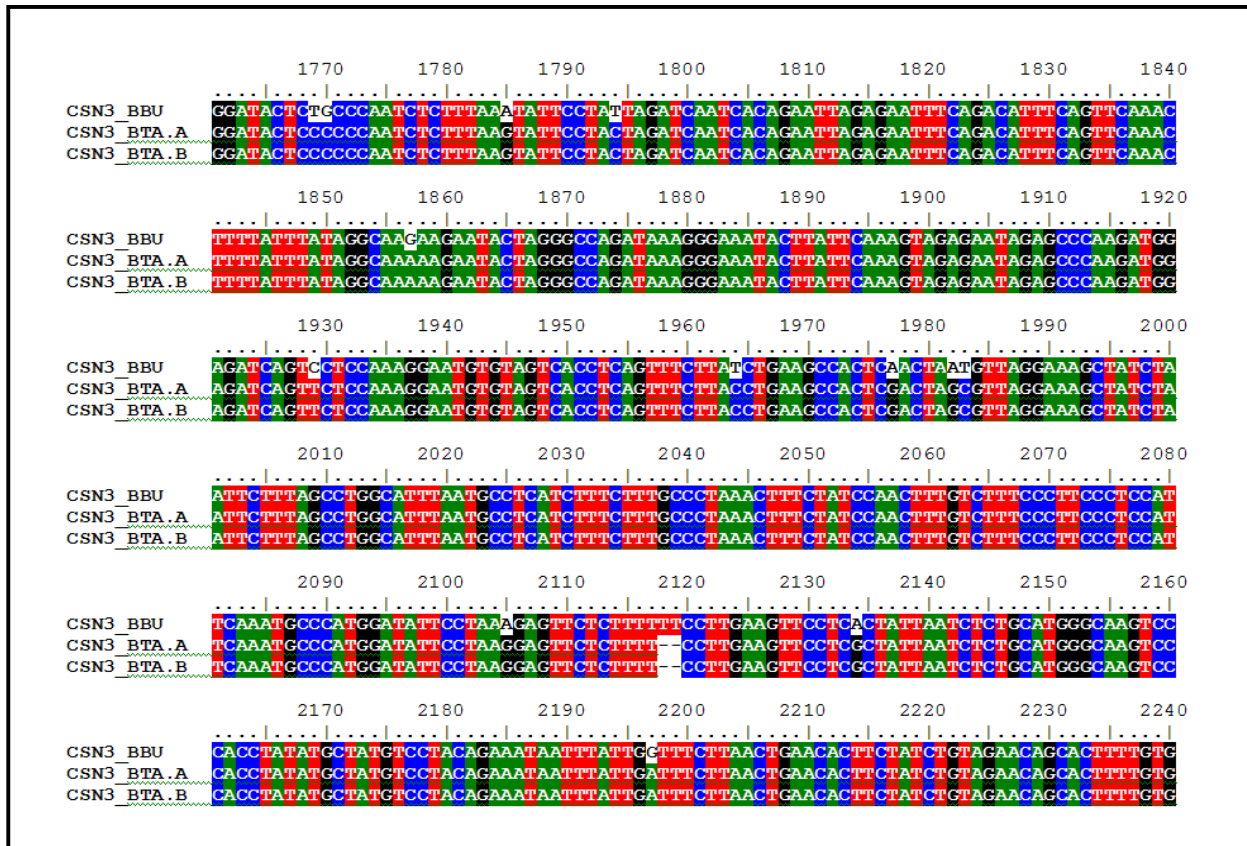
Embora os 18 potenciais SNPs descritos no presente trabalho estejam localizados em íntrons, não pode ser descartada a hipótese de haver elementos regulatórios nessas regiões, uma vez que a seqüência completa do íntron 1 e a seqüência parcial do íntron 2 estão sendo descritas pela primeira vez em búfalo.

**Tabela 4.** Descrição das diferenças encontradas entre as seqüências obtidas para os três animais (#739, #764 e #766), evidenciando os potenciais polimorfismos encontrados e o consenso para esses potenciais polimorfismos.

Posição	#739	#764	#766	Consenso
538	G	deleção	G	G
671	A/G	G	A/G	R
1008	A/C	A	A/C	M
1010	A/C	C	A/C	M
1011	A/C	A	A/C	M
1012	A/C	C	A/C	M
1019	A/G	deleção	A/G	R
1020	A/G	G	A/G	R
1170	G	A/G	A/G	R
1294	G	A	A/G	R
1363	T	T	T/C	Y
1365	C	C	deleção	C
2876	A	A/G	A	R
2889	A/G	G	A/G	R
3221	G ou C/G	G ou C/G	C/G	S
4047	A	A/G	A/G	R
4108	C	C/T	C	Y
4350	T	deleção	T	T

A seqüência consenso do gene CSN3 de búfalo foi comparada com as seqüências disponíveis no banco de dados do NCBI para bovino. Os alinhamentos mais significativos foram obtidos com as seqüências do alelo A (AY380228.1) e do alelo B (AY380229.1) do gene CSN3 bovino, ambos com 97% de similaridade e *e-value* 0.0.

Os alinhamentos da seqüência do gene CSN3 (CSN3\_BBU) de búfalo com as seqüências dos alelos A (CSN3\_BTA.A) e B (CSN3\_BTA.B) do gene bovino se encontra detalhado no Apêndice 4. Parte desse alinhamento se encontra ilustrado na Figura 5.



**Figura 5.** Parte do alinhamento do gene CSN3 de búfalo (CSN3\_BBU) com os alelos A (CSN3\_BTA.A) e B (CSN3\_BTA.B) do gene CSN3 bovino, gerado com o programa Clustal W.

Na Tabela 5 estão apresentadas as diferenças em nível de nucleotídeos observadas entre boi (alelos A e B) e búfalo para o gene CSN3.

**Tabela 5.** Diferenças observadas entre a seqüência do gene CSN3 de búfalo gerada no presente trabalho (CSN3\_BBU), com as correspondentes em bovino para o alelo A (CSN3\_BTA.A) e o alelo B (CSN3\_BTA.B).

*Pos.	CSN3_BBU	CSN3_BTA.A	CSN3_BTA.B	*Pos.	CSN3_BBU	CSN3_BTA.A	CSN3_BTA.B	*Pos.	CSN3_BBU	CSN3_BTA.A	CSN3_BTA.B
62	gap	T	T	1367	C	gap	gap	2939	T	G	G
63	gap	C	C	1374	A	G	G	2979	T	C	C
218	C	A	A	1376	A	G	G	3021	T	G	G
317	C	G	G	1402	G	C	C	3079	A	G	G
344	A	T	T	1455	T	C	C	3105	G	A	A
353	T	C	C	1461	G	gap	gap	3111	A	T	T
357	G	A	A	1462	A	gap	gap	3163	A	G	G
370	G	A	A	1463	T	gap	gap	3217	C	T	T
432	T	A	A	1474	G	A	A	3223	S	C	C
461	G	gap	gap	1497	C	T	T	3262	G	C	C

Tabela 5. Continuação...

*Pos.	CSN3_BBU	CSN3_BT.A.A	CSN3_BT.A.B	*Pos.	CSN3_BBU	CSN3_BT.A.A	CSN3_BT.A.B	*Pos.	CSN3_BBU	CSN3_BT.A.A	CSN3_BT.A.B
462	A	gap	gap	1508	A	G	A	3274	T	C	C
464	A	A	gap	1545	T	gap	gap	3388	A	G	G
539	G	C	C	1574	A	C	C	3395	T	C	C
540	G	gap	gap	1670	T	G	G	3426	gap	T	T
653	A	C	C	1671	G	T	T	3427	gap	A	A
654	C	A	A	1672	A	T	T	3428	gap	G	G
673	R	G	G	1680	A	T	T	3429	gap	T	T
725	C	A	A	1769	T	C	C	3430	gap	A	A
735	A	G	G	1770	G	C	C	3431	gap	C	C
747	T	gap	gap	1785	A	G	G	3432	gap	A	A
748	A	gap	gap	1794	T	C	C	3433	gap	T	T
749	A	gap	gap	1857	G	A	A	3434	gap	T	T
750	A	gap	gap	1929	C	T	T	3435	gap	A	A
751	G	gap	gap	1964	T	C	C	3450	Y	T	T
755	A	G	G	1977	A	G	G	3462	A	gap	gap
791	G	A	A	1982	A	G	G	3463	G	gap	gap
806	G	A	A	1983	T	C	C	3520	A	G	G
913	G	T	T	2105	A	G	G	3545	G	A	A
920	G	A	A	2118	T	gap	gap	3675	C	T	T
923	A	C	C	2119	T	gap	gap	3724	A	G	G
975	T	G	G	2134	A	G	G	3747	G	A	A
979	G	A	A	2197	G	A	A	3782	A	G	G
980	C	A	A	2255	C	gap	gap	3804	C	G	G
1007	M	A	A	2256	T	gap	gap	3844	A	G	G
1009	M	C	C	2331	T	G	G	3900	C	T	T
1010	M	A	A	2338	A	G	G	3910	T	A	A
1011	M	C	C	2367	A	G	G	3914	T	A	A
1018	R	G	G	2370	C	A	A	3937	T	C	C
1019	R	A	A	2386	T	C	C	3959	C	T	T
1059	G	A	A	2398	C	T	T	3963	C	G	G
1079	C	T	T	2439	G	A	A	3968	G	A	A
1080	A	G	G	2462	T	G	G	3994	G	A	A
1083	A	G	G	2520	A	T	T	4023	G	T	T
1132	C	T	T	2536	T	C	C	4034	A	G	G
1146	C	T	T	2548	G	T	T	4054	G	A	A
1172	R	G	G	2591	T	C	C	4059	R	A	A
1177	C	T	T	2626	A	G	G	4067	T	G	G
1178	A	G	G	2719	A	T	A	4078	C	T	T
1219	G	A	A	2726	T	T	A	4082	T	C	C
1231	A	G	G	2862	T	T	C	4103	T	C	C
1261	G	A	A	2878	R	A	A	4104	G	A	A
1296	R	A	A	2886	C	T	T	4120	Y	C	C
1331	A	G	G	2891	R	A	A	4147	A	G	G
1355	G	T	T	2896	G	A	G	4176	T	C	C
1359	A	G	G	2916	C	T	T	4211	C	T	T
1365	Y	T	T	2924	G	A	A	4364	T	gap	gap

\*Pos. corresponde à posição do nucleotídeo na seqüência consenso do gene CSN3 de búfalo.

A seqüência consenso apresentou 168 nucleotídeos diferentes das seqüências dos alelos A e B bovino, das quais duas são exclusivas com relação ao alelo A (posição 464 e 2862) e duas são exclusivas com relação ao alelo B (posição 1508 e 2896).

Analisando as 168 diferenças encontradas entre boi e búfalo por região, dois *gaps* foram encontrados na seqüência parcial da região 5'UTR gerada. Em relação aos íntrons, foi verificado a presença de 100 nucleotídeos diferentes no íntron 1 e 66 nucleotídeos diferentes na seqüência parcial do íntron 2. Nenhuma divergência foi encontrada nas porções correspondentes aos éxons 1 e 2, o que já era esperado, uma vez que os éxons são seqüências conservadas, principalmente em espécies da mesma família, como ocorre entre boi e búfalo (família Bovidae).

A caracterização molecular completa do gene CSN3 de búfalo e sua comparação com o gene bovino será de grande valor para o entendimento da arquitetura genômica do búfalo, traçando os elementos responsáveis por diferenças tão significativas nesses traços entre os ruminantes. Outro aspecto a ser considerado é a contribuição que esses dados proporcionarão para o entendimento dos processos evolutivos que ocorreram dentro da família Bovidae para este gene.

## Conclusões

O presente trabalho gerou 10 novos pares de iniciadores para PCR do gene CSN3 de búfalo. Esses iniciadores para PCR poderão ser utilizados na biblioteca BAC de búfalo (BBU\_LGC) para localizar os clones que possuem o gene CSN3 para sua posterior caracterização.

No presente trabalho foi seqüenciada uma região de 4366 nucleotídeos do gene CSN3 de búfalo, sendo 115 nucleotídeos da região 5'UTR, 64 do éxon 1, 2442 do íntron 1, 62 do éxon 2 e 1683 do íntron 2. A seqüência completa do íntron 1 e a seqüência parcial do íntron 2 foram descritas pela primeira vez para a espécie *Bubalus bubalis*.

Com a comparação das seqüências obtidas para os três animais da raça Murrah, o presente trabalho evidenciou 18 potenciais polimorfismos do tipo SNPs, todos em nas regiões dos íntrons 1 e 2. Embora todos os potenciais SNPs descritos no presente trabalho estejam localizados nos íntrons, não pode ser descartada a hipótese de haver elementos regulatórios nessas regiões. Esses potenciais SNPs são interessantes para serem estudados nos rebanhos comerciais, tanto em estudos populacionais quanto na busca por associação com alguma característica de interesse econômico, por meio de *chips* de DNA.

A seqüência do gene CSN3 de búfalo gerada no presente trabalho foi comparada com a correspondente em bovino para dois alelos (A e B), revelando uma diferença de 168 nucleotídeos. Todas as diferenças encontradas entre as seqüências de boi e búfalo foram nas regiões não codificantes, sendo duas na região 5'UTR, 100 no íntron 1 e 66 no íntron2.

## Referências Bibliográficas\*

ALEANDRI, R. et al. The effects of milk protein polymorphisms on milk components and cheese-producing ability. **Journal of Dairy Science**, v. 73, p. 241-255, 1990.

AMARAL, M.E.J. et al. A first generation whole genome RH map of the river buffalo with comparison to domestic cattle. **BMC Genomics**, v. 9, p. 1-11, 2008.

AZEVEDO, A.L.S. et al. Genetic polymorphism of the kappa-casein gene in Brazilian cattle. **Genetics and Molecular Research**, v. 7, p. 623-630, 2008.

BLAST. **Basic Local Alignment Search Tool**. Disponível em: <http://blast.ncbi.nlm.nih.gov/Blast.cgi>. Acesso em Julho de 2010.

CARDAK, A.D.; BARTENSCHLAGER, H.; GELDERMANN, H. Effects of polymorph milk proteins on the individual milk protein content of Holstein-Friesian and Simmental cows. **Milk Science International**, v. 58, p. 235-238, 2003.

CHARIZOPOULOU, N. et al. Instability of the insertional mutation in CfrTgH(neoim)Hgu cystic fibrosis mouse model. **BMC Genetics**, v. 5, 2004.

CORNISH-BOWDEN, A. Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. **Nucleic Acids Research**, v. 13, p. 3021-3030, 1985.

FARRELL, H.M. et al. Nomenclature of the proteins of cow's milk – sixth revision. **Journal of Dairy Science**, v. 87, p.1641-1674, 2004.

GRAML, R.; PIRCHNER, F. Effects of milk protein loci on content of their proteins. **Archives Animal Breeding**, v. 46, p. 331-340, 2003.

HALLÉN, E. et al. Effect of  $\beta$ -casein,  $\kappa$ -casein and  $\beta$ -lactoglobulin genotypes on concentration of milk protein variants. **Journal of Animal Breeding and Genetics**, v. 125, p. 119-129, 2008

HALLÉN, E. et al. Effect of genetic polymorphism of milk proteins on rheology of chymosin induced milk gels. **International Dairy Journal**, v, 17, p. 791-799, 2007.

IKONEN, T.; OJALA, M.; SYVÄOJA, E.L. Effects of composite casein and  $\beta$ -lactoglobulin genotypes on renneting properties and composition of bovine milk by assuming an animal model. **Agricultural and Food Science in Finland**, v. 6, p. 283-294, 1997.

LIN, C.Y. et al. Effects of milk protein loci on first lactation production in dairy cattle. **Journal of Dairy Science**, v. 69, p. 704-712, 1986.

MAZUMBER, B.; SESHADRI, V.; FOX, P.L. Translational control by the 3'UTR: the ends specify the means. **Trends in Biochemical Sciences**, v. 28, p. 91-98, 2003.

\*Associação Brasileira de Normas Técnicas (ABNT). NBR60232002, Informação e documentação - Referências - Elaboração, Rio de Janeiro, Agosto de 2002.

MCLEAN, D.M. Influence of milk protein variants on milk composition, yield and cheese making properties. **Animal Genetics**, v. 18, p. 100-102, 1987.

MERCIER, J.C.; BRIGNON, G.; RIBADEAU-DUMAS, B. Structure primaire de la caseine k-B bovine sequence complete. **European Journal of Biochemistry**, v. 35, p. 222-235, 1973.

MITRA, A. et al. Kappa-casein polymorphisms in indian dairy cattle and buffalo: a new genetic variant in buffalo. **Animal Biotechnology**, v. 9, p. 81-87, 1998.

NCBI. **National Center for Biotechnology Information**. Disponível em: <http://www.ncbi.nlm.nih.gov/sites/entrez>. Acesso em Julho de 2010.

**Primer3** (versão 0.4.0). Disponível em: <http://frodo.wi.mit.edu/primer3/>. Acesso em Junho de 2010.

PRINZENBERG, E.M.; KRAUSE, I.; ERHARDT, G. SSCP analysis at the bovine CSN3 locus discriminates six alleles corresponding to known protein variants (A, B, C, E, F, G) and three new DNA polymorphisms (H, I, A<sup>1</sup>). **Animal Biotechnology**, v. 10, p. 49-62, 1999.

RIJNKELS, M. Multispecies comparison of the casein gene loci and evolution of casein gene family. **Journal of Mammary Gland Biology and Neoplasia**, v. 7, p. 327-345, 2002.

ROBITAILLE, G. et al. Polymorphism in the bovine  $\kappa$ -casein (CSN3) gene and the 5'-flanking region: sequence analysis of CSN3 A and B alleles. **Animal Genetics**, v. 36, p. 184-185, 2005.

Sequencher® 4.10.1. **DNA Sequence Assembly Software**. Disponível em: <http://www.genecodes.com>

SORIA, L.A. A PCR-RFLP test to detect allelic variants of the bovine kappa-casein gene. **Animal Biotechnology**, v. 14, p. 1-5, 2003.

SZYMANOWSKA, M. et al. Association of nucleotide sequence polymorphism in the 5'-flanking regions of bovine casein genes with casein content in cow's milk. **Le Lait**, v. 84, p. 579-590, 2004a.

SZYMANOWSKA, M. et al. Transcription factor binding to variable nucleotide sequences in 5'-flanking regions of bovine casein genes. **International Dairy Journal**, v. 14, p. 103-115, 2004b.

THOMPSON, J.D.; HIGGINS, D.G.; GIBSON, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v. 22, p. 4673-4680, 1994.

THREADGILL, D.W.; WOMACK, J.E. Genomic analysis of the major bovine milk protein genes. **Nucleic Acids Research**, v. 18, p. 6935-6942, 1990.

TSIARAS, A.M. et al. Effect of kappa-casein and beta-lactoglobulin loci on milk production traits and reproductive performance of Holstein cows. **Journal of Dairy Science**, v. 88, p. 327-334, 2005.

VAN DEN BERG, G. et al. Genetic polymorphism of  $\kappa$ -casein and  $\beta$ -lactoglobulin in relation to milk composition and processing properties. **Netherlands Milk and Dairy Journal**, v. 46, p. 145-168, 1992.

VAN EENENNAAM, A.L.; MEDRANO, J.F. Differences in allelic protein expression in the milk of heterozygous  $\kappa$ -casein cows. **Journal of Dairy Science**, v. 74, p. 1491-1496, 1991.

WALSH, C.D. et al. Influence of  $\kappa$ -casein genetic variant on rennet gel microstructure, cheddar cheese-making properties and casein micelle size. **International Dairy Journal**, v. 8, p. 707-714, 1998.

WEDEMEYER, N. et al. Conservation of the 3'-untranslated region of the *Rab1a* gene in amniote vertebrates: exceptional structure in marsupials and possible role for postranscriptional regulation. **FEBS Letters**, v. 477, p. 49-54, 2000.

# Apêndices



**Apêndice 1 – Sequência do Vetor pBeloBAC11, disponível no banco de dados americano NCBI (<http://www.ncbi.nlm.nih.gov/nucore>) com o número de acesso U51113.1.**

**Regiões:**

436- 56 lacZ  
 312- 329 Promotor T7  
 414- 398 Promotor SP6  
 1425- 766 Resistência ao antibiótico cloranfenicol (CM<sup>R</sup>)  
 2765-3520 repE (repA)  
 2370-2436 oriS (ori2)  
 4108-5274 sopA (parA)  
 5274-6245 sopB (parB)

**Tamanho: 7507 nucleotídeos**

```

1 GCGGCCGCAA GGGGTTTCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG
51 CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG
101 GTGTGAAATA CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGGCGCC
151 ATTCGCCATT CAGGCTGCGC AACTGTTGGG AAGGGCGATC GGTGCGGGCC
201 TCTTCGCTAT TACGCCAGCT GGCGAAAGGG GGATGTGCTG CAAGGCGATT
251 AAGTTGGGTA ACGCCAGGGT TTTCCAGTC ACGACGTTGT AAAACGACGG
301 CCAGTGAATT GTAATACGAC TCACTATAGG GCGAATTCGA GCTCGGTACC
351 CGGGGATCCT CTAGAGTCGA CCTGCAGGCA TGCAAGCTTG AGTATTCTAT
401 AGTGTACCT AAATAGCTTG GCGTAATCAT GGTTCATAGCT GTTTCCTGTG
451 TGAAATTGTT ATCCGCTCAC AATTCCACAC AACATACGAG CCGGAAGCAT
501 AAAGTGTAAG GCCTGGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG
551 CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG GAAACCTGTC GTGCCAGCTG
601 CATTAATGAA TCGGCCAACG CGAACCCTT GCGGCCGCC GGGCCGTCGA
651 CCAATTCTCA TGTTTGACAG CTTATCATCG AATTTCTGCC ATTCATCCGC
701 TTATTATCAC TTATTCAGGC GTAGCAACCA GGCGTTTAAG GGCACCAATA
751 ACTGCCTTAA AAAAATTACG CCCC GCCCTG CCACTCATCG CAGTACTGTT
801 GTAATTCATT AAGCATTCTG CCGACATGGA AGCCATCACA AACGGCATGA
851 TGAACCTGAA TCGCCAGCGG CATCAGCACC TTGTGCGCCTT GCGTATAATA
901 TTTGCCCATG GTGAAAACGG GGGCGAAGAA GTTGTCCATA TTGGCCACGT
951 TTAAATCAAA ACTGGTGAAG CTCACCCAGG GATTGGCTGA GACGAAAAAC
1001 ATATTCTCAA TAAACCCTTT AGGGAAATAG GCCAGGTTTT CACCGTAACA
1051 CGCCACATCT TGCGAATATA TGTGTAGAAA CTGCCGAAA TCGTCGTGGT
1101 ATTCACTCCA GAGCGATGAA AACGTTTCAG TTTGCTCATG GAAAACGGTG
1151 TAACAAGGGT GAACACTATC CCATATCACC AGCTCACCGT CTTTCATTGC
1201 CATA CGGAAT TCCGGATGAG CATTATCAG GCGGGCAAGA ATGTGAATAA
1251 AGGCCGGATA AAAC TTGTGC TTATTTTTCT TTACGGTCTT TAAAAAGGCC
1301 GTAATATCCA GCTGAACGGT CTGGTTATAG GTACATTGAG CAACTGACTG
1351 AAATGCCTCA AAATGTTCTT TACGATGCCA TTGGGATATA TCAACGGTGG
1401 TATATCCAGT GATTTTTTTC TCCATTTTAG CTTCCTTAGC TCCTGAAAAT
1451 CTCGATAACT CAAAAAATAC GCCCGGTAGT GATCTTATTT CATTATGGTG
1501 AAAGTTGGAA CCTCTTACGT GCCGATCAAC GTCTCATTTT CGCCAAAAGT
1551 TGGCCAGGG CTTCCCGGTA TCAACAGGGA CACCAGGATT TATTTATTCT
1601 GCGAAGTGAT CTTCCGTCAC AGGTATTTAT TCGCGATAAG CTCATGGAGC
1651 GGCGTAACCG TCGCACAGGA AGGACAGAGA AAGCGCGGAT CTGGGAAGTG
1701 ACGGACAGAA CGGTCAGGAC CTGGATTGGG GAGGCGGTTG CCGCCGCTGC
1751 TGCTGACGGT GTGACGTTCT CTGTTCCGGT CACACCACAT ACGTTCGGCC
1801 ATTCTATGTC GATGCACATG CTGTATGCCG GTATACCGCT GAAAGTTCTG

```

1851 CAAAGCCTGA TGGGACATAA GTCCATCAGT TCAACGGAAG TCTACACGAA  
1901 GGTTTTTTCG CTGGATGTGG CTGCCCCGCA CCGGGTGCAG TTTGCGATGC  
1951 CGGAGTCTGA TGCGGTTGCG ATGCTGAAAC AATTATCCTG AGAATAAATG  
2001 CCTTGGCCTT TATATGGAAA TGTGGAAGT AGTGGATATG CTGTTTTTGT  
2051 CTGTTAAACA GAGAAGCTGG CTGTTATCCA CTGAGAAGCG AACGAAACAG  
2101 TCGGGAAAAT CTCCCATTAT CGTAGAGATC CGCATTATTA ATCTCAGGAG  
2151 CCTGTGTAGC GTTTATAGGA AGTAGTGTTT TGTCATGATG CCTGCAAGCG  
2201 GTAACGAAAA CGATTTGAAT ATGCCTTCAG GAACAATAGA AATCTTCGTG  
2251 CGGTGTACG TTGAAGTGGA GCGGATTATG TCAGCAATGG ACAGAACAAC  
2301 CTAATGAACA CAGAACCATG ATGTGGTCTG TCCTTTTACA GCCAGTAGTG  
2351 CTCGCCGAG TCGAGCGACA GGGCGAAGCC CTCGAGTGAG CGAGGAAGCA  
2401 CCAGGGAACA GCACTTATAT ATTCTGCTTA CACACGATGC CTGAAAAAAC  
2451 TTCCCTTGGG GTTATCCACT TATCCACGGG GATATTTTTA TAATTATTTT  
2501 TTTTATAGTT TTTAGATCTT CTTTTTTAGA GCGCCTTGTA GGCCTTTATC  
2551 CATGCTGGTT CTAGAGAAGG TGTTGTGACA AATTGCCCTT TCAGTGTGAC  
2601 AAATCACCTT CAAATGACAG TCCTGTCTGT GACAAATTGC CCTTAACCCT  
2651 GTGACAAATT GCCCTCAGAA GAAGCTGTTT TTTCACAAAG TTATCCCTGC  
2701 TTATTGACTC TTTTTTATTT AGTGTGACAA TCTAAAAACT TGTCACACTT  
2751 CACATGGATC TGTCATGGCG GAAACAGCGG TTATCAATCA CAAGAAACGT  
2801 AAAAAATAGC CGCGAATCGT CCAGTCAAAC GACCTCACTG AGGCGGCATA  
2851 TAGTCTCTCC CGGGATCAAA AACGTATGCT GTATCTGTTT GTTGACCAGA  
2901 TCAGAAAATC TGATGGCACC CTACAGGAAC ATGACGGTAT CTGCGAGATC  
2951 CATGTTGCTA AATATGCTGA AATATTCGGA TTGACCTCTG CGGAAGCCAG  
3001 TAAGGATATA CGGCAGGCAT TGAAGAGTTT CGCGGGGAAG GAAGTGGTTT  
3051 TTTATCGCCC TGAAGAGGAT GCCGGCGATG AAAAAGGCTA TGAATCTTTT  
3101 CCTTGGTTTT TCAAACGTGC GCACAGTCCA TCCAGAGGGC TTTACAGTGT  
3151 ACATATCAAC CCATATCTCA TTCCCTTCTT TATCGGGTTA CAGAACCGGT  
3201 TTACGTTAGT TCGGCTTAGT GAAACAAAAG AAATCACCAA TCCGTATGCC  
3251 ATGCGTTTAT ACGAATCCCT GTGTCAGTAT CGTAAGCCGG ATGGTCTCAG  
3301 CATCGTCTCT CTGAAAATCG ACTGGATCAT AGAGCGTTAC CAGCTGCCTC  
3351 AAAGTTACCA GCGTATGCCT GACTTCCGCC GCCGCTTCCT GCAGGTCTGT  
3401 GTTAATGAGA TCAACAGCAG AACTCCAATG CGCCTCTCAT ACATTGAGAA  
3451 AAAGAAAGGC CGCCAGACGA CTCATATCGT ATTTTCCTTC CGCGATATCA  
3501 CTTCCATGAC GACAGGATAG TCTGAGGGTT ATCTGTCACA GATTTGAGGG  
3551 TGGTTCGTCA CATTTGTTCT GACCTACTGA GGGTAATTTG TCACAGTTTT  
3601 GCTGTTTTCT TCAGCCTGCA TGGATTTTCT CATACTTTT GAACTGTAAT  
3651 TTTTAAAGGAA GCCAAATTTG AGGGCAGTTT GTCACAGTTG ATTTCCCTTCT  
3701 CTTTCCCTTC GTCATGTGAC CTGATATCGG GGGTTAGTTC GTCATCATTG  
3751 ATGAGGGTTG ATTATCACAG TTTATTACTC TGAATTGGCT ATCCGCGTGT  
3801 GTACCTCTAC CTGGAGTTTT TCCCACGGTG GATATTTCTT CTGCGCTGA  
3851 GCGTAAGAGC TATCTGACAG AACAGTTCTT CTTTGCTTCC TCGCCAGTTC  
3901 GCTCGCTATG CTCGGTTACA CGGCTGCGGC GAGCGCTAGT GATAATAAGT  
3951 GACTGAGGTA TGTGCTCTTC TTATCTCCTT TTGTAGTGTT GCTCTTATTT  
4001 TAAACAACCT TGCGGTTTTT TGATGACTTT GCGATTTTGT TGTTGCTTTG  
4051 CAGTAAATTG CAAGATTTAA TAAAAAACG CAAAGCAATG ATTTAAAGGAT  
4101 GTTCAGAATG AAACCTATGG AAACACTTAA CCAGTGCATA AACGCTGGTC  
4151 ATGAAATGAC GAAGGCTATC GCCATTGCAC AGTTTAAATGA TGACAGCCCG  
4201 GAAGCGAGGA AAATAACCCG GCGCTGGAGA ATAGGTGAAG CAGCGGATTT  
4251 AGTTGGGGTT TCTTCTCAGG CTATCAGAGA TGCCGAGAAA GCAGGGCGAC  
4301 TACCGCACCC GGATATGGAA ATTCGAGGAC GGGTTGAGCA ACGTGTGGT  
4351 TATACAATTG AACAAATTA TCCATATGCGT GATGTGTTTG GTACGCGATT  
4401 GCGACGTGAT GAAGACGTAT TTCCACCGGT GATCGGGGTT GCTGCCATA  
4451 AAGGTGGCGT TTACAAAACC TCAGTTTCTG TTCATCTTGC TCAGGATCTG  
4501 GCTCTGAAGG GGCTACGTGT TTTGCTCGTG GAAGGTAACG ACCCCAGGG  
4551 AACAGCCTCA ATGTATCACG GATGGGTACC AGATCTTCAT ATTCATGCAG  
4601 AAGACTCTCT CCTGCCTTTC TATCTTGGGG AAAAGGACGA TGTCACCTAT  
4651 GCAATAAAGC CCACTTGCTG GCCGGGGCTT GACATTATTC CTTCCTGTCT

4701 GGCTCTGCAC CGTATTGAAA CTGAGTTAAT GGGCAAATTT GATGAAGGTA  
4751 AACTGCCAC CGATCCACAC CTGATGCTCC GACTGGCCAT TGAAACTGTT  
4801 GCTCATGACT ATGATGTCAT AGTTATTGAC AGCGCGCCTA ACCTGGGTAT  
4851 CGGCACGATT AATGTCGTAT GTGCTGCTGA TGTGCTGATT GTTCCCACGC  
4901 CTGCTGAGTT GTTTGACTAC ACCTCCGCAC TGCAGTTTTT CGATATGCTT  
4951 CGTGATCTGC TCAAGAACGT TGATCTTAAA GGGTTCGAGC CTGATGTACG  
5001 TATTTTGCTT ACCAAATACA GCAATAGTAA TGGCTCTCAG TCCCCGTGGA  
5051 TGGAGGAGCA AATTTCGGGAT GCCTGGGGAA GCATGGTTCT AAAAAATGTT  
5101 GTACGTGAAA CGGATGAAGT TGGTAAAGGT CAGATCCGGA TGAGAACTGT  
5151 TTTTGAACAG GCCATTGATC AACGCTCTTC AACTGGTGCC TGGAGAAATG  
5201 CTCTTTCTAT TTGGGAACCT GTCTGCAATG AAATTTTCGA TCGTCTGATT  
5251 AAACCACGCT GGGGAGATTAG ATAATGAAGC GTGCGCCTGT TATTCAAAA  
5301 CATACGCTCA ATACTCAACC GGTTGAAGAT ACTTCGTTAT CGACACCAGC  
5351 TGCCCCGATG GTGGATTCTG TAATTGCGCG CGTAGGAGTA ATGGCTCGCG  
5401 GTAATGCCAT TACTTTGCCT GTATGTGGTC GGGATGTGAA GTTTACTCTT  
5451 GAAGTGCTCC GGGGTGATAG TGTTGAGAAG ACCTCTCGGG TATGGTCAGG  
5501 TAATGAACGT GACCAGGAGC TGCTTACTGA GGACGCACTG GATGATCTCA  
5551 TCCCTTCTTT TCTACTGACT GGTCAACAGA CACCGCGCTT CGGTCGAAGA  
5601 GTATCTGGTG TCATAGAAAT TGCCGATGGG AGTCGCCGTC GTAAAGCTGC  
5651 TGCACTTACC GAAAGTGATT ATCGTGTTCT GGTGGCGAG CTGGATGATG  
5701 AGCAGATGGC TGCATTATCC AGATTGGGTA ACGATTATCG CCCAACAAAGT  
5751 GCTTATGAAC GTGGTCAGCG TTATGCAAGC CGATTGCAGA ATGAATTTGC  
5801 TGGAAATATT TCTGCGCTGG CTGATGCGGA AAATATTTCA CGTAAGATTA  
5851 TTACCCGCTG TATCAACACC GCCAAATTGC CTAAATCAGT TGTGCTCTT  
5901 TTTTCTCACC CCGGTGAACT ATCTGCCCGG TCAGGTGATG CACTTCAAAA  
5951 AGCCTTTACA GATAAAGAGG AATTACTTAA GCAGCAGGCA TCTAACCTTC  
6001 ATGAGCAGAA AAAAGCTGGG GTGATATTTG AAGCTGAAGA AGTTATCACT  
6051 CTTTAACTT CTGTGCTTAA AACGTCATCT GCATCAAGAA CTAGTTAAG  
6101 CTCACGACAT CAGTTTGCTC CTGGAGCGAC AGTATTGTAT AAGGGCGATA  
6151 AAATGGTGCT TAACCTGGAC AGGTCTCGTG TTCCAAGTGA GTGTATAGAG  
6201 AAAATTGAGG CCATTCTTAA GGAAGTTGAA AAGCCAGCAC CCTGATGCGA  
6251 CCACGTTTTA GTCTACGTTT ATCTGTCTTT ACTTAATGTC CTTTGTTACA  
6301 GGCCAGAAAG CATAACTGGC CTGAATATTC TCTCTGGGCC CACTGTTCCA  
6351 CTTGTATCGT CGGTCTGATA ATCAGACTGG GACCACGGTC CCACTCGTAT  
6401 CGTCGGTCTG ATTATTAGTC TGGGACCACG GTCCCACCTG TATCGTCCGT  
6451 CTGATTATTA GTCTGGGACC ACGGTCCCAC TCGTATCGTC GGTCGATAA  
6501 TCAGACTGGG ACCACGGTCC CACTCGTATC GTCGGTCTGA TTATTAGTCT  
6551 GGGACCATGG TCCCCTCGT ATCGTCCGTC TGATTATTAG TCTGGGACCA  
6601 CGGTCCCCT CGTATCGTCT GTCTGATTAT TAGTCTGGAA CCACGGTCCC  
6651 ACTCGTATCG TCGGTCTGAT TATTAGTCTG GGACCACGGT CCCACTCGTA  
6701 TCGTCCGTCT GATTATTAGT CTGGGACCAC GATCCCACCT GTGTTGTCCG  
6751 TCTGATTATC GGTCTGGGAC CACGGTCCCA CTTGTATTGT CGATCAGACT  
6801 ATCAGCGTGA GACTACGATT CCATCAATGC CTGTCAAGGG CAAGTATTGA  
6851 CATGTCGTCG TAACCTGTAG AACGGAGTAA CCTCGGTGTG CGGTTGTATG  
6901 CCTGCTGTGG ATTGCTGCTG TGTCTGCTT ATCCACAACA TTTTGCACAC  
6951 GGTTATGTGG ACAAATACC TGGTTACCCA GGCCGTGCCG GCACGTTAAC  
7001 CGGGCTGCAT CCGATGCAAG TGTGTCGCTG TCGACGAGCT CGCGAGCTCG  
7051 GACATGAGGT TGCCCCGTAT TCAGTGTCCG TGATTTGTAT TGTCTGAAGT  
7101 TGTTTTTACG TTAAGTTGAT GCAGATCAAT TAATACGATA CCTGCGTCAT  
7151 AATTGATTAT TTGACGTGGT TTGATGGCCT CCACGCACGT TGTGATATGT  
7201 AGATGATAAT CATTATCACT TTACGGGTCC TTTCCGGTGA TCCGACAGGT  
7251 TACGGGGCGG CGACCTCGCG GGTTCCTGCT ATTTATGAAA ATTTTCCGGT  
7301 TTAAGGCGTT TCCGTTCTTC TTCGTCATAA CTTAATGTTT TTATTTAAAA  
7351 TACCCTCTGA AAAGAAAGGA AACGACAGGT GCTGAAAGCG AGCTTTTGG  
7401 CCTCTGTCTG TTCCTTCTC TGTTTTTGTG CGTGGAATGA ACAAATGGAAG  
7451 TCCGAGCTCA TCGCTAATAA CTTCTGATAG CATAATTAT ACGAAGTTAT  
7501 ATTCGAT

## Apêndice 2 – Soluções utilizadas na construção da biblioteca BAC de búfalo

- **BSA (Albumina sérica bovina) - 10 mg/mL**
  - 100 mg de BSA
  - completar com água para 10 mL
  
- **Cloranfenicol - 12,5 mg/mL**
  - 0,5 g de cloranfenicol
  - completar com etanol absoluto para 40 mL
  
- **EDTA 0,5 M pH 8.0**
  - 18,1 g de Na<sub>2</sub>EDTA.2H<sub>2</sub>O
  - 2 g de NaOH
  - completar com água para 100 mL
  
- **Espermidina - 40 mM**
  - 1,16 g de espermidina
  - completar com água para 200 mL
  
- **Hidróxido de Bário a 5%**
  - 5 g de Hidróxido de Bário (Ba(OH)<sub>2</sub>)
  - completar com água para 100 ml à 60°C
  
- **IPTG 20% (200 mg/mL)**
  - 2 g de IPTG
  - completar com água para 10 mL
  
- **Meio LB**
  - 5 g de bacto triptona
  - 2,5 g de extrato de levedura
  - 5 g de NaCl
  - completar com água para 500 mL
  
- **Meio LB-ágar**
  - 5 g de bacto triptona
  - 2,5 g de extrato de levedura
  - 5 g de NaCl
  - 7,5 g de ágar
  - completar com água para 500 mL

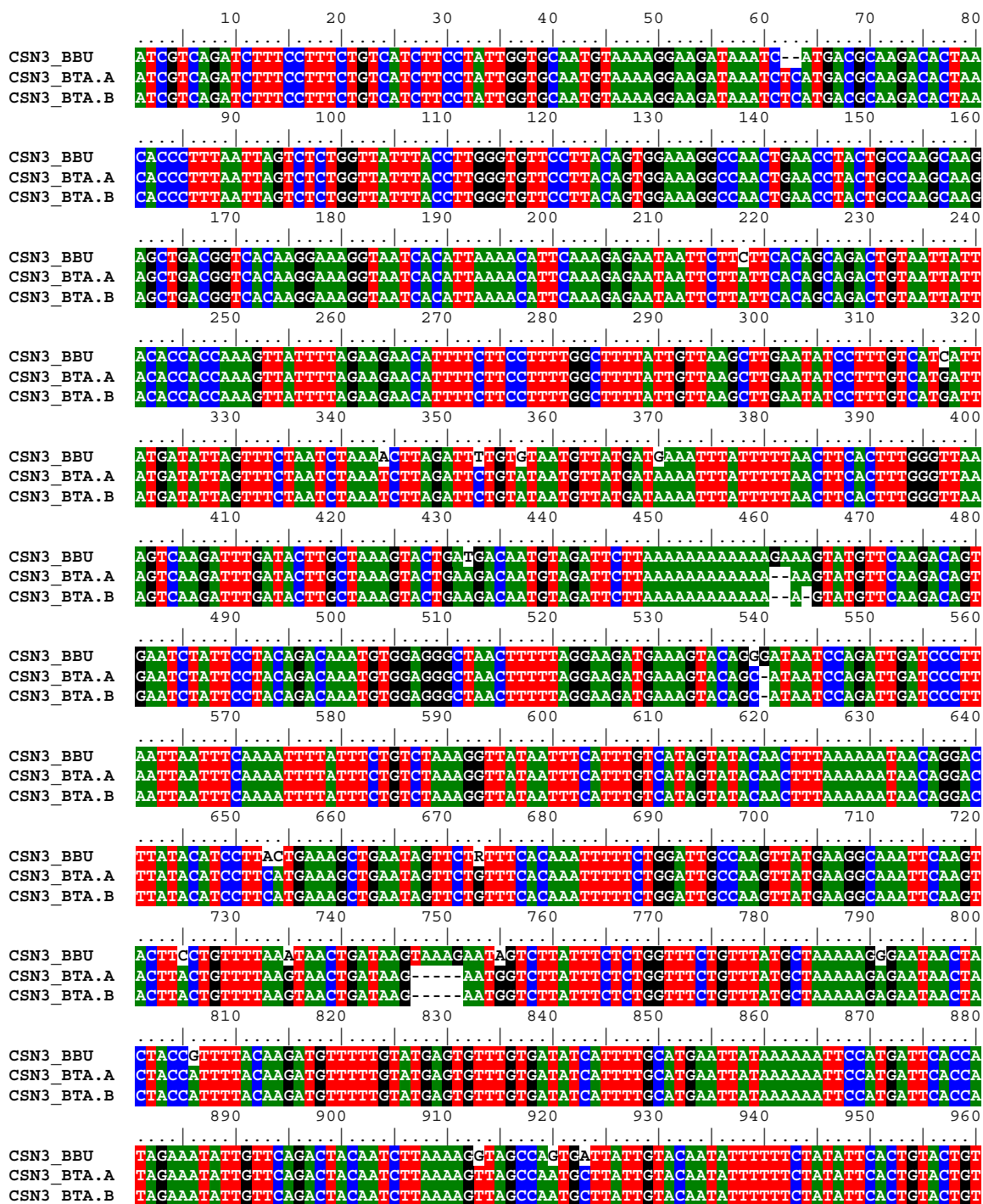
- **Meio SOC**
  - 2% de Bactotripton
  - 0,5% de extrato de levedura
  - 10 mM de NaCl
  - 2,5 mM de KCl
  - 10 mM de MgCl
  - 10 mM de MgSO<sub>4</sub>
  - 20 mM de glicose
  
- **Solução Coloidal Reveladora**
  - 0,5 g de gelatina
  - 0,25 mL de ácido fórmico
  - completar com água para 25 mL
  
- **Solução de acetato de potássio**
  - 147 g de KC<sub>2</sub>H<sub>3</sub>O<sub>2</sub>
  - 142,5 mL de ácido acético glacial
  - completar com água para 500mL
  
- **Solução de congelamento dos clones**
  - 36 mM de K<sub>2</sub>HPO<sub>4</sub>
  - 13,2 mM de KH<sub>2</sub>PO<sub>4</sub>
  - 1,7 mM de Na<sub>3</sub>C<sub>6</sub>H<sub>5</sub>O<sub>7</sub>
  - 0,4 mM de MgSO<sub>4</sub>
  - 68 mM de (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>
  - 4,4% de glicerol
  
- **Solução de Giemsa a 6%**
  - 6 mL de Giemsa
  - 94 ml de tampão sorensen
  
- **Solução de laurilsarcosina sódica-EDTA pH 9.0**
  - 25 mM de EDTA
  - 1% de laurilsarcosina sódica
  
- **Solução de lisoenzima**
  - 4,5 g de glicose
  - 10 mL de EDTA 0,5 M pH 8.0
  - 12,5mL de Tris-HCl 1M
  - completar com água para 500 mL
  
- **Solução de NaOH-SDS**
  - 1 mL de NaOH 4N
  - 1 mL de SDS 20%
  - 18 mL de água

- **Solução de Nitrato de Prata 50%**
  - 1 g de AgNO<sub>3</sub>
  - completar com água para 2 mL
  
- **Solução Salina (2X SSC)**
  - 4,38 g de NaCl
  - 2,20 g de C<sub>6</sub>H<sub>5</sub>Na<sub>3</sub>O<sub>7</sub>.2H<sub>2</sub>O
  - completar com água para 500 mL
  
- **Solução de Tripsina**
  - 12 g de tripsina
  - 50 mL de tampão sorensen
  
- **Tampão Sorensen pH 6.8**
  - 49 mL de Na<sub>2</sub>HPO<sub>4</sub> a 0,2 M
  - 51 mL de KH<sub>2</sub>PO<sub>4</sub> a 0,2 M
  - completar com água para 200 mL
  
- **Tampão TAE (Tris-Acetato-EDTA) - 1X**
  - 40 mM de Tris-base
  - 20 mM de ácido acético glacial
  - 2 mM de EDTA
  
- **Tampão TBE 0,5X**
  - 44,5 mM de Tris-base
  - 44,5 mM de ácido bórico
  - 1 mM de EDTA
  
- **Tampão TE (Tris-EDTA) pH 8.0**
  - 10 mM de Tris-HCl
  - 1 mM EDTA
  
- **XGal 2% (20 mg/mL)**
  - 200 µL de XGal a 100 mg/mL
  - 800 µL de dimetilformamida

**Apêndice 3 - Seqüência gerada no presente trabalho para o gene CSN3 de búfalo com 4366 nucleotídeos, cujas seqüências grifadas em vermelho correspondem aos éxons 1 e 2.**

```
>CSN3_BBU_CONSENSO
ATCGTCAGATCTTTCCCTTCTGTGCATCTTCCTATTGGTGCAATGTAAGGAAAGATAAATCATGACGCAAGACACTAACACCCTTTAAT
TAGTCTCTGGTTATTTACCTTGGGTCTCCTTACAGTGGAAAGGCCAACTGAACCTACTGCCAAGCAAGAGCTGACGGTCACAAGGAAA
GGTAATCACATTAAAACATTCAAAGAGAATAAATCTTCTTCCACAGCAGACTGTAATTATTACACCACCAAAGTTATTTTAGAAGAACAT
TTTCTTCCCTTTGGCTTTTATTGTTAAGCTTGAATATCCTTTGTGCATCATTATGATATTAGTTTCTAATCTAAAACCTTAGATTTTGTGT
AATGTTATGATGAAATTTATTTTAACTTCACTTTGGGTTAAAGTCAAGATTGATACTTGCTAAAAGTACTGATGACAATGTAGATTCT
TAAAAAAGAAAGTATGTTCAAGACAGTGAATCTATTCTACAGACAAATGTGGAGGGCTAACTTTTGTAGGAGATGAAAGTAC
AGGGATAATCCAGATTGATCCCTTAATTAATTTCAAATTTTATTTCTGTCTAAAGGTTATAATTTCAATTTGTGCATAGTATACAACCTT
AAAAATAACAGGACTTATACATCCTTACTGAAAGCTGAATAGTTCTRTTTCACAAATTTTCTGGATTGCCAAGTTATGAAGGCAAT
TCAAGTACTTCTGTTTTAAATAACTGATAAGTAAAGAATAGTCTTATTTCTCTGGTTTCTGTTTATGCTAAAAAGGGAAATACTACTA
CCGTTTTACAAGATGTTTTGTATGAGTGTGTGTGATATCATTGTCATGAATTATAAAAAATTCATGATTACCATAGAAATATTTGT
TCAGACTACAATCTTAAAGGTAGCCAGTGATTATGTACAATATTTTCTATATTCACTGTACTGTATGTAATCATCAATCTTGGCT
AAGAGACTAGATATTAATTATCCAMCMMMAAAAARRAATGATAAATTATGTGACATGACAGAGTTGTTAATTATCAGTACAATTGCAA
TCAAATTACAATACACAAATACATTAATAAACATGTTAGACACATTAAATTAACACAATGCTATATATTAATCCATTTCAACTAAGA
AAAATTTCACTGRCAAGCAAAGAATTCATAAGGAGCTCCATTCTATTCAATGAAATTTGGGGAACAATCAATGCTGTAGGTTGCTCAA
ATTTGAACCTCAGTTAAGTTCTTTCCATTCACTGATATTTCCATCAGRGGTTCTGAAGTCAGTTTAGTAGACAGCCTTTCTAATCATTC
TTACCTGGAAATCAAAGTGAATTGAGYCTGGGTGACACTTTTATCAAGCTGAACCTTAAGTCTGTGACCCCTAGTCTATAAAAAC
CAGTTTGCAGGTTTCAAGTCTATACAGTATGATGATAACAATGGCTGTAACAAGATGTAGTATTTAATTCATAAAATGCAAAATTTATT
AATATTTTAAAAATCTAAAATGTTATATGTATCTTTCTAAAAAGACAATTTAGCTTTTAAAAATCAAATAAATAGGTATTACTACCCTAT
ACTTGAGTCTTTAGAAAAGATAAGGTTAATTTATTTCTCAAATATTTTCATCCAGAAAACCTGTAATTGACATATCTAACATTTTTTCAAT
AAATTGATCAAGCTAATTATATACTTATGATTACTGGAGAAGCAGATTATGAAAAATGCAGTTTCAGGATACTCTGCCCAATCTCTTT
AAATATCTTATAGATCAATCACAGAATTAGAGAATTTAGACATTTCAAGTTCAAATTTTATTTATAGGCAAGAAGAATACTAGGGC
CAGATAAAGGGAAATACTTATTCAAAGTAGAGAATAGAGCCCAAGATGGAGATCAGTCTCCAAAGGAATGTGTAGTCACTCAGTTTC
TTATCTGAAGCCACTCAACTAATGTTAGGAAAGCTATCTAATCTTTAGCCTGGCATTTAATGCCTCATCTTTCTTTGCCCTAACTTT
CTATCCAATTTGTCTTTCCCTCCCTCCATTCAAATGCCATGGATATTCCTAAAGAGTTCTCTTTTTCTTGAAGTTCTCACTAT
TAATCTCTGCATGGGCAAGTCCCACCTATATGCTATGTCTACAGAAAATAATTTATTGGTTTCTTAACTGAAACACTTCTATCTGTAGAA
CAGCACTTTTGTGACACTTTCTACTCTCTTTTATCTAACAGTGAATTTGTGTACATTGCAACTCTTCTAAAAATTTATGTTACTTGACCA
ATTCACCTTAATAATCTCAACTGACACACAGAGCTTTGTCTACTACATCTTTGGTACATATGAAATGAATACAGGATCAACCTTG
TAATGACTCATTGAGTAAACATAACCAAAAATTAGATTATATATTTCTACCATTTTGGTTTTCTAAAGCTTTTAAATAAATAAGATTTT
TCACATTTGGCTATATCTCCCCCTCAATCTGGCTTTTCCACTCTAAATCAAAGAAAACATTGCTCCAAGTCACTCTAAATATCAGC
AACAGATTCTGTTAATCTAATGTTTTAATTTAATTTTAGGTACAATGATGAAGAGTTTTTCTAGTTGTGACTATCTGGCATTAAAC
CCTGCCATTTTGGTGAGTTAATTTCACTAATCAGTTTGTATCAAATTTTTTAAATTTATGAAGGGCTTTAATTGTGTAATTTCTA
AATACAGTTATATTTTATAGAACAAAATAGCTGCATAAGTAATTTATTTTACATTTATTTTCTTATTAAGTCAATTTCTTCTGGAT
ATTCTTCTGGTTTCTAGGAATAAAGAARTCACACTATCTAGAGAAAACCTTTAAAAATCTCATATGATGGAGAAAATACATTTT
GGGGAAAAATCCTGAAAATATTTTAAATATAATATGTACTTAAAAATAATTCACAAAAATAAATATATAAACCATAGATGAAGATC
AGAAATGTTTTTAGTTAAATCTTATTTTATGCAAACTCCAATTTGTTTCAAACTAAATCCTTAAATAATAGGAGCTTGTATAAGCT
TTTTTACCAATTTCAAATAACTTGACATGTACTTTAATATTTAAATGTATTGTAACCTTTGTTATAATATTTCCATTGTGTCACTCC
CTCTCAGGGACTATATSCATAAACACCTCTTTGATTTTATAGGCAACATCTGAAAGCATCAGATGTTTTATTTCTGTTCTTACTGTTG
GCAATAAGTGCCCTGGGATCCCCAAATGAACATGAAAATAAACAGACCAATAGCCAAAAACAATGAGGAAAACCTAGTCAAAAAATTA
AAGAAAAAATCCCAAACCTTTTCAATTTACATGTAATTACAGAGTGTAATGCCYAAAAATCATAGAGTTATTTATGGATACAATCCC
TGAGAGGGAGTTGACAAAATATGAAGATTATAACAAAAGTTACAATATATTTGAATGTTAAGCATAATCAAGTTATTTGAAATTTGG
TAAAAAGAAAAAGAAAACTTAAACAACATCCTAATTTTTAAGAATTTAGTCTCTAAATAGACTGGAGTTTGTGTGAGATAGGAAT
TAATTAACAACATCATAATACCATGTGTAATAATAGATAGCCAGTGGAATTTGCTGTATAATGCAGGGAGCTCAAACCTAGGGCTC
TGTGACAACCTAGAGGGGTAGGATGGGGTGAAGGTGGAAGGGAGGTTCAAGGCGGAGGGGACATAACATAACCAATGGCTGATTCATAT
TAATATATGGCAGAAACCAACAATATTTGTAAGCAATTTCTTCAACTAAAAATAACAATTTAAATTTTAAATATCTGATGCTG
TATGACCTTCTCTGTTTTGATGGGCAGGGCCATCTCAGTAAATCTTTAATCCAATTTTCTGTTGATGGGCAGGGCTGTGTTCCCTCC
CTGTTGTTTGTGACCTGAAGCCAAATTTAGGTGGAGGTGATGARGATAATGTCTGGCGACCTCTTTAAAAAGGTCCCATGCAAACTGCT
GCACTCAGTGCCYAGCCTGCAGCAGGCCACCTCTGACCAATGCCTCTGCCAGAGACTCTGGACACTTATGGGCAAGTCTGGTTCCAG
TCTCTTGTGGGGTCACTGCTTCTTTCTTGGGTCTGTTGCACACAAGCTTTTGTGTTGTGCCCTCCAAGAGTCTGTTTCCCTCAGTCT
GTGTAAGTTCTAGAGGCTCTATGGTGGGGTAAATGGTGACCTCTTCAAGAGGGCTTATGCCATACCAGGTCTACTGCTACCCAGAGC
TCCTG
```

**Apêndice 4 – Alinhamento da seqüência consenso do gene CSN3 gerada no presente trabalho para búfalo, com as seqüências do gene CSN3 bovino depositadas no banco de dados NCBI, para o alelo A (AY380228.1) e o alelo B (AY380229.1), utilizando a ferramenta Clustal W.**



970 980 990 1000 1010 1020 1030 1040

CSN3\_BBU ATTGTAATCATCAATCTTGC TAAGAGACTAGATATTAATTTATCCAMCMMMAAAAAARRAATGATAAATATGTGACATGA

CSN3\_BTA.A ATTGTAATCATCAAGCTTAA TAAGAGACTAGATATTAATTTATCCAAACCACAAAAAGAAATGATAAATATGTGACATGA

CSN3\_BTA.B ATTGTAATCATCAAGCTTAA TAAGAGACTAGATATTAATTTATCCAAACCACAAAAAGAAATGATAAATATGTGACATGA

1050 1060 1070 1080 1090 1100 1110 1120

CSN3\_BBU CAGAGTTGTTAATTATCAGTACAATTGCAATCAAATTACAATAACAAAATACATTAAATAAACATGTTAGACACATTAAA

CSN3\_BTA.A CAGAGTTGTTAATTATCAATACAATTGCAATCAAATTATGATGCACAAAATACATTAAATAAACATGTTAGACACATTAAA

CSN3\_BTA.B CAGAGTTGTTAATTATCAATACAATTGCAATCAAATTATGATGCACAAAATACATTAAATAAACATGTTAGACACATTAAA

1130 1140 1150 1160 1170 1180 1190 1200

CSN3\_BBU TTTAACACAATGCTATATATTTAAATCTATTTCAAATAAGAAAAATTTCACTGRCAGCAAGAATTCAATAAGGAGCTCCAT

CSN3\_BTA.A TTTAACACAATGTTATATATTTAAATCTATTTCAAATAAGAAAAATTTCACTGGCAAGTGAAGAATTCATAAGGAGCTCCAT

CSN3\_BTA.B TTTAACACAATGTTATATATTTAAATCTATTTCAAATAAGAAAAATTTCACTGGCAAGTGAAGAATTCATAAGGAGCTCCAT

1210 1220 1230 1240 1250 1260 1270 1280

CSN3\_BBU TCTATTCAATGAAATTTGGGAAACAATCAATGCTGTAGGTTGCTCAAATTTGAACTCACGTTAAGTTCTTTTCCATTCA

CSN3\_BTA.A TCTATTCAATGAAATTTGAGGAAACAATCAGTGTCTGTAGGTTGCTCAAATTTGAACTCACATTAAAGTTCTTTTCCATTCA

CSN3\_BTA.B TCTATTCAATGAAATTTGAGGAAACAATCAGTGTCTGTAGGTTGCTCAAATTTGAACTCACATTAAAGTTCTTTTCCATTCA

1290 1300 1310 1320 1330 1340 1350 1360

CSN3\_BBU GTGATATTCATCAGRGGTTCTGAAGTCAGTTTAGTAGACAGCCTTTCTAATCATTCTTACCAGAAATCAAAAATGAAAT

CSN3\_BTA.A GTGATATTCATCAGAGGTTCTGAAGTCAGTTTAGTAGACAGCCTTTCTAGTCATTCTTACCAGAAATCAAAAATGAGT

CSN3\_BTA.B GTGATATTCATCAGAGGTTCTGAAGTCAGTTTAGTAGACAGCCTTTCTAGTCATTCTTACCAGAAATCAAAAATGAGT

1370 1380 1390 1400 1410 1420 1430 1440

CSN3\_BBU TGAGYCTGGTGTACACTCTTTATCAAGCCTGAACCTTCTAAGCTGTGTGACCCCTAGTCTATAAACTCAGTTTGCAGGTTT

CSN3\_BTA.A TGAGTCTGGTGTGCGCTCTTTATCAAGCCTGAACCTTCTAAGCTGTGTGACCCCTAGTCTATAAACTCAGTTTGCAGGTTT

CSN3\_BTA.B TGAGTCTGGTGTGCGCTCTTTATCAAGCCTGAACCTTCTAAGCTGTGTGACCCCTAGTCTATAAACTCAGTTTGCAGGTTT

1450 1460 1470 1480 1490 1500 1510 1520

CSN3\_BBU CAAGTCTATACAAGTATGATGATAACAATGGCTGTAAACAAGATGTAGTATTTAATTCATAAAATGCAAAATTTATTAATAT

CSN3\_BTA.A CAAGTCTATACAAGCATGAT---AACAATGGCTATAACAAGATGTAGTATTTAATTTATAAAATGCAAGATTATTAATAT

CSN3\_BTA.B CAAGTCTATACAAGCATGAT---AACAATGGCTATAACAAGATGTAGTATTTAATTTATAAAATGCAAAATTTATTAATAT

1530 1540 1550 1560 1570 1580 1590 1600

CSN3\_BBU TTTAAAAATCTAAAAATGTTATATGATCTTTCTAAAAAGCAAATTTAGCTTTTAAAAATCAAAATAAATAGGTATTTACTACC

CSN3\_BTA.A TTTAAAAATCTAAAAATGTTATATG---ATCTTTCTAAAAAGCAAATTTAGCTTTTCAAATCAAAATAAATAGGTATTTACTACC

CSN3\_BTA.B TTTAAAAATCTAAAAATGTTATATG---ATCTTTCTAAAAAGCAAATTTAGCTTTTCAAATCAAAATAAATAGGTATTTACTACC

1610 1620 1630 1640 1650 1660 1670 1680

CSN3\_BBU CTATACCTTGAGTCTTTAGAAAAGATAAGGTTAATTTATTTCTTCAAATAATTTCAATCCAGAAAACTGTAATGTTCAATCTT

CSN3\_BTA.A CTATACCTTGAGTCTTTAGAAAAGATAAGGTTAATTTATTTCTTCAAATAATTTCAATCCAGAAAACTGTAATGTTCAATCTT

CSN3\_BTA.B CTATACCTTGAGTCTTTAGAAAAGATAAGGTTAATTTATTTCTTCAAATAATTTCAATCCAGAAAACTGTAATGTTCAATCTT

1690 1700 1710 1720 1730 1740 1750 1760

CSN3\_BBU ACATTTTTTCAATTAATTTGATCAAGCTAATTAATAACTTTATGATTACTGGAGAACAGATTATGAAAAATGCAGTTTCA

CSN3\_BTA.A ACATTTTTTCAATTAATTTGATCAAGCTAATTAATAACTTTATGATTACTGGAGAACAGATTATGAAAAATGCAGTTTCA

CSN3\_BTA.B ACATTTTTTCAATTAATTTGATCAAGCTAATTAATAACTTTATGATTACTGGAGAACAGATTATGAAAAATGCAGTTTCA

1770 1780 1790 1800 1810 1820 1830 1840

CSN3\_BBU GGATACTCTGCCCAATCTCTTTAAATATTCTTACTAGATCAATCACAGAATTAGAGAATTTAGAGATTTTCAAGTTCAAAC

CSN3\_BTA.A GGATACTCCCCCAATCTCTTTAAATATTCTTACTAGATCAATCACAGAATTAGAGAATTTAGAGATTTTCAAGTTCAAAC

CSN3\_BTA.B GGATACTCCCCCAATCTCTTTAAATATTCTTACTAGATCAATCACAGAATTAGAGAATTTAGAGATTTTCAAGTTCAAAC

1850 1860 1870 1880 1890 1900 1910 1920

CSN3\_BBU TTTTATTTATAGGCAAGAAATACTAGGGCCAGATAAAGGGAAATACTTATTCAAAGTAGAGAATAGAGCCCAAGATGG

CSN3\_BTA.A TTTTATTTATAGGCAAAAAGAAATACTAGGGCCAGATAAAGGGAAATACTTATTCAAAGTAGAGAATAGAGCCCAAGATGG

CSN3\_BTA.B TTTTATTTATAGGCAAAAAGAAATACTAGGGCCAGATAAAGGGAAATACTTATTCAAAGTAGAGAATAGAGCCCAAGATGG

1930 1940 1950 1960 1970 1980 1990 2000

CSN3\_BBU AGATCAGTCTCCAAAGGAATGTGTAGTCACCTCAGTTTCTTATCTGAAGCCACTCACTAATGTTAGGAAAGCTATCTA

CSN3\_BTA.A AGATCAGTTCTCCAAAGGAATGTGTAGTCACCTCAGTTTCTTACTCTGAAGCCACTCGACTAGCGTTAGGAAAGCTATCTA

CSN3\_BTA.B AGATCAGTTCTCCAAAGGAATGTGTAGTCACCTCAGTTTCTTACTCTGAAGCCACTCGACTAGCGTTAGGAAAGCTATCTA

2010 2020 2030 2040 2050 2060 2070 2080

CSN3\_BBU ATTCTTTAGCCTGGCAATTTAATGCCATCTTTCTTTGCCCTAACTTTCTATCCAACTTTGCTTTCCCTTCCCTCCAT

CSN3\_BTA.A ATTCTTTAGCCTGGCAATTTAATGCCATCTTTCTTTGCCCTAACTTTCTATCCAACTTTGCTTTCCCTTCCCTCCAT

CSN3\_BTA.B ATTCTTTAGCCTGGCAATTTAATGCCATCTTTCTTTGCCCTAACTTTCTATCCAACTTTGCTTTCCCTTCCCTCCAT

	2090	2100	2110	2120	2130	2140	2150	2160
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	TCAAATGCCATGGATATTCTCTAAAGAGTTCTCTTTTTCTTTGAAAGTTCCTCACTATTAATCTCTGCATGGGCAAGTCC							
CSN3_BTA.B	TCAAATGCCATGGATATTCTCTAAAGAGTTCTCTTTTT--CCTTGAAAGTTCCTCGCTATTAATCTCTGCATGGGCAAGTCC							
	2170	2180	2190	2200	2210	2220	2230	2240
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	CACCTATATGCTATGTCCTACAGAAATAAATTATTGTTTCTTAAGTGAACACTTCTATCTGTAGAACAGCACATTTTGTG							
CSN3_BTA.B	CACCTATATGCTATGTCCTACAGAAATAAATTATTGATTTCCTTAAGTGAACACTTCTATCTGTAGAACAGCACATTTTGTG							
	2250	2260	2270	2280	2290	2300	2310	2320
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	ACACTTTCTACTCTCTTTTATCTAACAGTGAATTTGTGTACATTGCAACTCTTCTAAAAATTTATGTTACTTACCAATTC							
CSN3_BTA.B	ACACTTTCTACTCT--TTTATCTAACAGTGAATTTGTGTACATTGCAACTCTTCTAAAAATTTATGTTACTTACCAATTC							
	2330	2340	2350	2360	2370	2380	2390	2400
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	ACCTTAATAAATCTCAACCTGACACACAGAGCTTTGTTTCTACTACATTTCTTGGTACATATGAAATGAATACAGGATCAA							
CSN3_BTA.B	ACCTTAATAAAGCTCAACGCTGACACACAGAGCTTTGTTTCTACTACGTTATTGGTACATATGAAACGAATACAGGATTA							
	2410	2420	2430	2440	2450	2460	2470	2480
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	CCTTGTAATGACTCATTGAGTAACTAAACCAAAAATTAGATTCAATATTTCCACCATTGTTTTCTAAAGCTTTTA							
CSN3_BTA.B	CCTTGTAATGACTCATTGAGTAACTAAACCAAAAATTAAATTCATATATTTCCACCATTGTTTTCTAAAGCTTTTA							
	2490	2500	2510	2520	2530	2540	2550	2560
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	ATAAAATAGATTTTTACATTGGCTATATCTCCCCCTCAATCTGGCTTTTCCCACCTTAAAACAAATAAAACATTGCTC							
CSN3_BTA.B	ATAAAATAGATTTTTACATTGGCTATATCTCCCCCTCAATCTGGCTTTTCCCACCTTAAAACAAATAAAACATTGCTC							
	2570	2580	2590	2600	2610	2620	2630	2640
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	CAAGTCACTCCAAAATATCAGCAACAGATCTGTTAATCTAATGTTTTTAATTTAATTTTAGGTAACAATGATGAAGAGT							
CSN3_BTA.B	CAAGTCACTCCAAAATATCAGCAACAGATCCTGTTAATCTAATGTTTTTAATTTAATTTTAGGTAACAATGATGAAGAGT							
	2650	2660	2670	2680	2690	2700	2710	2720
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	TTTTTCTAGTTGTGACTATCCTGGCATTAAACCTGCCATTTTTGGTGAAGTAAATTCATCTAATCAGTTTGTATCAAAAT							
CSN3_BTA.B	TTTTTCTAGTTGTGACTATCCTGGCATTAAACCTGCCATTTTTGGTGAAGTAAATTCATCTAATCAGTTTGTATCAAAAT							
	2730	2740	2750	2760	2770	2780	2790	2800
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	TTTTTTAATTTATTGAAGGGCTTTAATTTGTAAATTTCTAAATACAGTTATATTTTTATAGAACAAAATAGCTGCATAAG							
CSN3_BTA.B	TTTTTTAATTTATTGAAGGGCTTTAATTTGTAAATTTCTAAATACAGTTATATTTTTATAGAACAAAATAGCTGCATAAG							
	2810	2820	2830	2840	2850	2860	2870	2880
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	TAATTTAATTTTACATTTATTTTCTTATTAAAGTCATATCTTCTGGATATTCTTCTGGTTTCTAGGAATAAAGAAATC							
CSN3_BTA.B	TAATTTAATTTTACATTTATTTTCTTATTAAAGTCATATCTTCTGGATATTCTTCTGGTTTCTAGGAATAAAGAAATC							
	2890	2900	2910	2920	2930	2940	2950	2960
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	ACACTCTATCTCTAGAGAAAAACCTTTAAAAATCTCATATGATGAGAAAAATACATTTGGGGGAAAAATCCTGAAAAATA							
CSN3_BTA.B	ACACTTTATCATCTAGAGAAAAACCTTTAAAAATCTTATATGATAGAGAAAAATACATTTGGGGGAAAAATCCTGAAAAATA							
	2970	2980	2990	3000	3010	3020	3030	3040
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	TTTTTTAAATTAATAATGTACTTAAAAATAATCACAAAAATAAATATAAACCATAGATGAAGATCAGAAATGTTTTTT							
CSN3_BTA.B	TTTTTTAAATTAATAATGTACTTAAAAATAATCACAAAAATAAATATAAACCATAGAGGAAGATCAGAAATGTTTTTT							
	3050	3060	3070	3080	3090	3100	3110	3120
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	AGTTAAATTTCTATTTTCATGCAAACTCCAATTTGTTTCAAACTAAATCCTTAAAAATTTAGGAGCTTGTATAAGCTTTT							
CSN3_BTA.B	AGTTAAATTTCTATTTTCATGCAAACTCCAATTTGTTTCAAACTAAATCCTTAAAAATTTAGGAGCTTGTATAAGCTTTT							
	3130	3140	3150	3160	3170	3180	3190	3200
CSN3_BBU	..... ..... ..... ..... ..... ..... ..... ..... .....							
CSN3_BTA.A	TTACCAATTTCAAATAACTTGACATGTACTTTTAAATTTTAAATGTATTGTAACCTTTTGTATAAATTTCCATTGTGTC							
CSN3_BTA.B	TTACCAATTTCAAATAACTTGACATGTACTTTTAAATTTTAAAGTATTGTAACCTTTTGTATAAATTTCCATTGTGTC							

	3210	3220	3230	3240	3250	3260	3270	3280
CSN3_BBU	AACTCCCTCTCAGGGA	CTATAT	SCATAAAACACCTCTTTGATTTTTAGGCCAAACATCTGAAAGCATCAGATGTTTATTTTC					
CSN3_BTA.A	AACTCCCTCTCAGGGA	TTATAT	CCATAAAACACCTCTTTGATTTTTAGGCCAAACATCTGAAACCATCAGATGTTTCAATTTTC					
CSN3_BTA.B	AACTCCCTCTCAGGGA	TTATAT	CCATAAAACACCTCTTTGATTTTTAGGCCAAACATCTGAAACCATCAGATGTTTCAATTTTC					
	3290	3300	3310	3320	3330	3340	3350	3360
CSN3_BBU	CTGTTCTTACTGTTGGCAAA	TAAAGTGCTGGGAT	CCCCAAATGAACATGAAAA	TAAAAAGACCAATAGCCAAAA	CAATG			
CSN3_BTA.A	CTGTTCTTACTGTTGGCAAA	TAAAGTGCTGGGAT	CCCCAAATGAACATGAAAA	TAAAAAGACCAATAGCCAAAA	CAATG			
CSN3_BTA.B	CTGTTCTTACTGTTGGCAAA	TAAAGTGCTGGGAT	CCCCAAATGAACATGAAAA	TAAAAAGACCAATAGCCAAAA	CAATG			
	3370	3380	3390	3400	3410	3420	3430	3440
CSN3_BBU	AGGAAAACTAGTCAAAAAAA	TTAAAAAGAAAAAA	CCCCAAACCTTTTTCATATTA	CATGTAATTA				CAGAG
CSN3_BTA.A	AGGAAAACTAGTCAAAAAAA	TTAAAAAGAAAAAA	CCCCAAACCTTTTTCATATTA	CATGTAATTA	TATAGTACATTACAGAG			
CSN3_BTA.B	AGGAAAACTAGTCAAAAAAA	TTAAAAAGAAAAAA	CCCCAAACCTTTTTCATATTA	CATGTAATTA	TATAGTACATTACAGAG			
	3450	3460	3470	3480	3490	3500	3510	3520
CSN3_BBU	TGTAATGCCYGAAAAAT	CATAGAGTTATTTATGGATACAATCCCTGAGAGGGAGTTGACAAAA	TATGAAGATTA	TAAACAA				
CSN3_BTA.A	TGTAATGCCYGAAAAAT	CATAGAGTTATTTATGGATACAATCCCTGAGAGGGAGTTGACAAAA	TATGAAGATTA	TAAACAA				
CSN3_BTA.B	TGTAATGCCYGAAAAAT	CATAGAGTTATTTATGGATACAATCCCTGAGAGGGAGTTGACAAAA	TATGAAGATTA	TAAACAA				
	3530	3540	3550	3560	3570	3580	3590	3600
CSN3_BBU	AGTTACAATATATTTGAAT	GTTAAAGCATACATATCAAGTTATTTGAAAT	TGGTAAAAAGAAAAAGAAAA	CTTAAACAA				
CSN3_BTA.A	AGTTACAATATATTTGAAT	GTTAAAGCATACATATCAAGTTATTTGAAAT	TGGTAAAAAGAAAAAGAAAA	CTTAAACAA				
CSN3_BTA.B	AGTTACAATATATTTGAAT	GTTAAAGCATACATATCAAGTTATTTGAAAT	TGGTAAAAAGAAAAAGAAAA	CTTAAACAA				
	3610	3620	3630	3640	3650	3660	3670	3680
CSN3_BBU	CATCCTAATATTTTAA	GAATTTAGTCTCTAAA	TAGACTGGAGTTTGTGTGAGATAGGAATTTTAA	TAAAAACATCATACA				
CSN3_BTA.A	CATCCTAATATTTTAA	GAATTTAGTCTCTAAA	TAGACTGGAGTTTGTGTGAGATAGGAATTTTAA	TAAAAACATCATACA				
CSN3_BTA.B	CATCCTAATATTTTAA	GAATTTAGTCTCTAAA	TAGACTGGAGTTTGTGTGAGATAGGAATTTTAA	TAAAAACATCATACA				
	3690	3700	3710	3720	3730	3740	3750	3760
CSN3_BBU	TTACCATGTGTA	AAAAATTAGATAGCCAGTGGAAATTTGCTGTATATATGCAGGGAGCTCAAACCTAGGGCTCTGTGACAACC						
CSN3_BTA.A	TTACCATGTGTA	AAAAATTAGATAGCCAGTGGAAATTTGCTGTATATATGCAGGGAGCTCAAACCTAGGGCTCTGTGACAACC						
CSN3_BTA.B	TTACCATGTGTA	AAAAATTAGATAGCCAGTGGAAATTTGCTGTATATATGCAGGGAGCTCAAACCTAGGGCTCTGTGACAACC						
	3770	3780	3790	3800	3810	3820	3830	3840
CSN3_BBU	TAGAGGGTAGGATGGGGT	TGGAGGTGGAAGGGAGGTTCAAGGGGGAGGGGACATAACATACCAATGGCTGATTCATATT						
CSN3_BTA.A	TAGAGGGTAGGATGGGGT	TGGAGGTGGAAGGGAGGTTCAAGGGGGAGGGGACATAACATACCAATGGCTGATTCATATT						
CSN3_BTA.B	TAGAGGGTAGGATGGGGT	TGGAGGTGGAAGGGAGGTTCAAGGGGGAGGGGACATAACATACCAATGGCTGATTCATATT						
	3850	3860	3870	3880	3890	3900	3910	3920
CSN3_BBU	AATATATGGCAGAAACCAACACA	CAATATTGTAAGCAATTTATCCTCCAACATAAAAA	TAAACACATTTAAATTTTAAATATC					
CSN3_BTA.A	AATATATGGCAGAAACCAACACA	CAATATTGTAAGCAATTTATCCTCCAACATAAAAA	TAAACACATTTAAATTTTAAATATC					
CSN3_BTA.B	AATATATGGCAGAAACCAACACA	CAATATTGTAAGCAATTTATCCTCCAACATAAAAA	TAAACACATTTAAATTTTAAATATC					
	3930	3940	3950	3960	3970	3980	3990	4000
CSN3_BBU	TGATGCTGTATGACCTCCTCT	TTGTTTTGATGGGCAGGGCCATCCTCAGTAAATCTTTAATCCAATTTTCTGTTGATGGGC						
CSN3_BTA.A	TGATGCTGTATGACCTCCTCT	TTGTTTTGATGGGCAGGGCCATCCTCAGTAAATCTTTAATCCAATTTTCTGTTGATGGGC						
CSN3_BTA.B	TGATGCTGTATGACCTCCTCT	TTGTTTTGATGGGCAGGGCCATCCTCAGTAAATCTTTAATCCAATTTTCTGTTGATGGGC						
	4010	4020	4030	4040	4050	4060	4070	4080
CSN3_BBU	AGGGCTGTGTTCCCTCCCTGTT	TTTTGACCTGAAAGCCAAATTATGGTGGAGGTGATGARGATAATGTTCTGGCGACCTCCT						
CSN3_BTA.A	AGGGCTGTGTTCCCTCCCTGTT	TTTTGACCTGAAAGCCAAATTATGGTGGAGGTGATGARGATAATGTTCTGGCGACCTCCT						
CSN3_BTA.B	AGGGCTGTGTTCCCTCCCTGTT	TTTTGACCTGAAAGCCAAATTATGGTGGAGGTGATGARGATAATGTTCTGGCGACCTCCT						
	4090	4100	4110	4120	4130	4140	4150	4160
CSN3_BBU	TTAAAAAGTCCCATGCAAAAC	CTGCTGCACTCAGTGCCCYCAGCCTGCAGCAGGCCCTCTGACCAATGCCCTCTGCCAG						
CSN3_BTA.A	TTAAAAAGTCCCATGCAAAAC	CTGCTGCACTCAGTGCCCYCAGCCTGCAGCAGGCCCTCTGACCAATGCCCTCTGCCAG						
CSN3_BTA.B	TTAAAAAGTCCCATGCAAAAC	CTGCTGCACTCAGTGCCCYCAGCCTGCAGCAGGCCCTCTGACCAATGCCCTCTGCCAG						
	4170	4180	4190	4200	4210	4220	4230	4240
CSN3_BBU	AGACTCCTGGACACTTATGGG	CAAGTCTGGTTCACTCTCTTGTGGGGTCACTGCTTCTTTCTCTTGGTCCCTGGTGACA						
CSN3_BTA.A	AGACTCCTGGACACTTATGGG	CAAGTCTGGTTCACTCTCTTGTGGGGTCACTGCTTCTTTCTCTTGGTCCCTGGTGACA						
CSN3_BTA.B	AGACTCCTGGACACTTATGGG	CAAGTCTGGTTCACTCTCTTGTGGGGTCACTGCTTCTTTCTCTTGGTCCCTGGTGACA						
	4250	4260	4270	4280	4290	4300	4310	4320
CSN3_BBU	CAAGCTTTTGTGTTGTGCC	CTCAAGAGTCTGTTTCTCAGTCCCTGTGTAAGTTCTAGAGGCTCTATGTTGGGTTAATGG						
CSN3_BTA.A	CAAGCTTTTGTGTTGTGCC	CTCAAGAGTCTGTTTCTCAGTCCCTGTGTAAGTTCTAGAGGCTCTATGTTGGGTTAATGG						
CSN3_BTA.B	CAAGCTTTTGTGTTGTGCC	CTCAAGAGTCTGTTTCTCAGTCCCTGTGTAAGTTCTAGAGGCTCTATGTTGGGTTAATGG						

4330 4340 4350 4360 4370  
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....  
CSN3\_BBU TGACCTCTTCTAAGAGGGCTTATGCCATACCCAGGTCTACTGCTACCCAGAGCTCCTG  
CSN3\_BTA.A TGACCTCTTCTAAGAGGGCTTATGCCATACCCAGGTCTACTGCACCCAGAGCTCCTG  
CSN3\_BTA.B TGACCTCTTCTAAGAGGGCTTATGCCATACCCAGGTCTACTGCACCCAGAGCTCCTG