

**UNIVERSIDADE FEDERAL DA GRANDE DOURADOS**  
**FACULDADE DE CIÊNCIAS BIOLÓGICAS E AMBIENTAIS**  
**ANDRÉ VIEIRA DO NASCIMENTO**

**ESTUDO DE ASSOCIAÇÃO GENÔMICA AMPLA PARA CARACTERÍSTICA  
REPRODUTIVA EM FÊMEAS DA RAÇA NELORE**

Dourados-MS

2014

ANDRÉ VIEIRA DO NASCIMENTO

**ESTUDO DE ASSOCIAÇÃO GENÔMICA AMPLA PARA CARACTERÍSTICA  
REPRODUTIVA EM FÊMEAS DA RAÇA NELORE**

Trabalho de conclusão de curso de graduação  
apresentado para obtenção do título de  
Bacharel em Biotecnologia – Faculdade de  
Ciências Biológicas e Ambientais –  
Universidade Federal da Grande Dourados.  
Orientadora: Prof<sup>ª</sup> Dr<sup>ª</sup> Alexéia Barufatti  
Grisolia.

Dourados-MS

2014

ANDRÉ VIEIRA DO NASCIMENTO

**ESTUDO DE ASSOCIAÇÃO GENÔMICA AMPLA PARA CARACTERÍSTICA  
REPRODUTIVA EM FÊMEAS DA RAÇA NELORE**

Trabalho de Conclusão de Curso aprovado como requisito parcial para obtenção do título de Bacharel em Biotecnologia da Faculdade de Ciências Biológicas da Universidade Federal da Grande Dourados, pela comissão formada por:

---

Prof. Dr<sup>a</sup> Alexéia Barufatti Grisolia

---

Dr<sup>a</sup> Márcia Cristina Matos

---

Dr<sup>a</sup> Fabiane Siqueira

Dourados-MS

2014

## **Agradecimentos**

*Agradeço à minha família, especialmente a minha mãe Lícia e minha irmã Amanda, por sempre acreditarem em mim e me apoiarem. São exemplos de força e solidez pra mim e tenho uma enorme admiração por vocês.*

*Aos meus grandes amigos (quase irmãos) Anderson, Rafael e Rodrigo, pela amizade, fidelidade e inúmeros momentos de descontração.*

*Aos amigos do Laboratório de Biotecnologia Aplicada a Produção Animal (Alexandre, Bruno, Dani, Joyce, Jussara e Lara), por todo o conhecimento adquirido, crescimento profissional, amizade e companheirismo. Admiro muito vocês, são exemplos de profissionais.*

*A todos os amigos de minha turma, em especial ao Alexandre, Ana (Nerd), Ewerton, Isabela, Marcela e Tales, que tive a sorte de conhecer e conviver durante os anos de graduação.*

*À minha namorada Ândrea por todo o carinho, incentivo e companheirismo.*

*À Prof<sup>a</sup> orientadora Alexéia pela orientação e oportunidade de conviver e aprender constantemente.*

*À Dr<sup>a</sup> Marcia por todo o aprendizado e colaboração para o acontecimento deste trabalho.*

*À Universidade Federal da Grande Dourados pelo apoio logístico e oportunidade de cursar a graduação em Biotecnologia.*

*Aos órgãos de fomento CNPq e Fundect, pelo apoio financeiro à pesquisa e concessão de bolsa de iniciação científica.*

*A todos que não citei, mas que de alguma forma contribuíram para o meu desenvolvimento acadêmico e possibilitou a realização deste trabalho.*

*"A mente que se abre a uma nova ideia jamais voltará ao seu tamanho original."*

*(Albert Einstein)*

*Dedico este trabalho a minha mãe Lícia,  
por todo o amor a mim dedicado e por  
sempre confiar, compreender e apoiar  
minhas decisões.*

## Sumário

Capítulo I.....	1
Considerações gerais .....	2
Referências .....	5
Capítulo II .....	7
Estudo de associação genômica ampla para característica reprodutiva em fêmeas da raça Nelore .....	8
Resumo.....	8
<i>Abstract</i> .....	8
Abreviaturas .....	9
Introdução.....	9
Material e Métodos.....	11
Animais e dados fenotípicos .....	11
Colheita de sangue e extração de DNA.....	12
Análise de SNP.....	12
Controle de qualidade.....	13
Teste de associação .....	14
Exploração dos SNP significativos .....	14
Resultados e Discussão .....	15
Controle de Qualidade.....	15
Teste de associação .....	18
Conclusão .....	24
Agradecimentos.....	24
Referências .....	24
Anexos.....	28
Normas da revista.....	28

# **Capítulo I**

## **Considerações gerais**

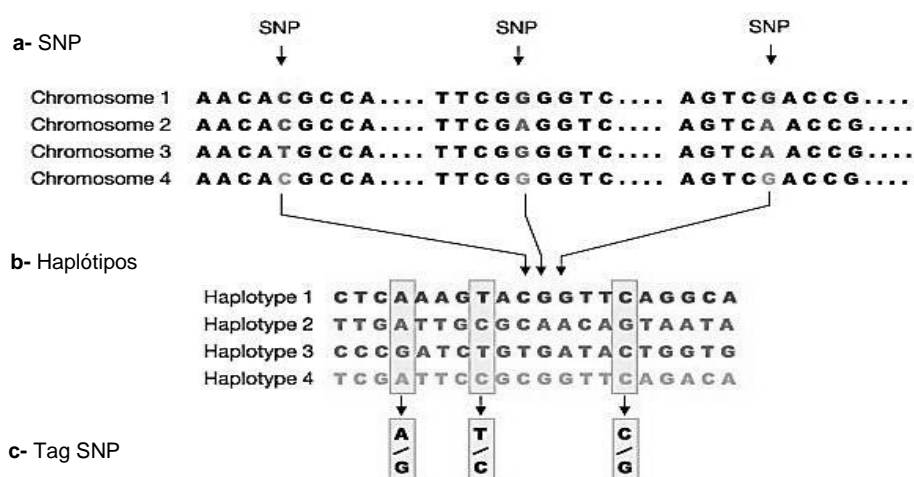


## Considerações gerais

Os marcadores moleculares são fundamentados em variações ocorrentes no genoma dos espécimes que, por padrão de hereditariedade Mendeliana, segregam nas gerações subsequentes (FERREIRA e GRATTAPAGLIA, 1996).

Dentre os marcadores moleculares, o SNP (do inglês, *single nucleotide polymorphism*) ou polimorfismo de nucleotídeo de sítio único destaca-se como responsável por mais de 80% das variações genéticas de um indivíduo. Estes são concebidos na permuta de uma única base nitrogenada na sequência de DNA (do inglês, *deoxyribonucleic acid*). Quando ocorre troca pelo mesmo tipo de base nitrogenada, ou seja, pirimidina por pirimidina (citosina por timina ou vice-versa) ou purina por purina (adenina por guanina ou vice-versa), chama-se transição. As substituições por base de tipo diferente, pirimidina por purina ou vice-versa, são conhecidas como transversões (PASSOS-BUENO e MOREIRA, 2004).

Variantes genéticas próximas umas das outras tendem a serem herdadas juntas, de modo que indivíduos portadores de um SNP específico, localizado em uma determinada coordenada cromossômica, têm as variantes genéticas idênticas nas proximidades desta região (INTERNATIONAL HAPMAP CONSORTIUM, 2003). Estas regiões com polimorfismos ligados são conhecidas como haplótipos (Figura 1).



**Figura 1.** Representação da conformação de SNP, haplótipos e tagSNP. a) SNP: polimorfismos de nucleotídeo de sítio único em uma região do DNA de múltiplos indivíduos. b) Haplótipos: formação dos haplótipos de acordo com polimorfismos próximos. c) TagSNP:

identificação e escolha de três tagSNP suficientemente capazes de distinguir os quatro haplótipos (adaptado de INTERNATIONAL HAPMAP CONSORTIUM, (2003)).

Atualmente, os marcadores do tipo SNP vem sendo massivamente utilizados em diferentes estudos, devido estes serem facilmente identificados, abundantes por toda a extensão do código genético e bialélicos, possibilitando a diferenciação de alelos do mesmo gene. Estas variações pontuais podem estar em regiões intrônicas, intergênicas ou exônicas (sinônimas ou não sinônimas). Quando localizadas em regiões codificadoras não sinônimas podem resultar em trocas aminoacídicas, gerando proteínas modificadas e, conseqüentemente, fenótipos diferenciados (FALEIRO, 2007).

Um dos marcos científicos de grande importância e responsável por inúmeros avanços tecnocientíficos foi o desenvolvimento do Projeto Genoma Humano. O referido projeto formou um consórcio na década de 1990 e envolveu pesquisadores de diversos países para a realização do sequenciamento do genoma *Homo sapiens*. O projeto custou bilhões de dólares e um trabalho extensivo de 13 anos. Alguns dos objetivos do referido projeto foram: identificação dos genes humanos, desenvolvimento de ferramentas de análise de dados, sequenciamento de cerca de 3.2 bilhões de pares de bases que compõem o genoma e transferência de tecnologia (WATSON, 1990; SAWICKI et al., 1993; U.S. DEPARTMENT OF ENERGY OFFICE OF SCIENCE, 2013).

Os avanços tecnológicos no sequenciamento genômico e em bioinformática foram fundamentais na geração de informação biológica e na maior compreensão do código genético, abrindo oportunidade ao surgimento de novas tecnologias na identificação de polimorfismos (CAETANO, 2009). Uma delas foi o desenvolvimento dos painéis de genotipagem de alta densidade, que constituem lâminas de vidro contendo na sua superfície milhares ou milhões de sondas (fragmentos específicos de DNA) capazes de interrogar polimorfismos pontuais, partindo do princípio de hibridação molecular (PASSOS-BUENO e MOREIRA, 2004).

Os painéis de SNP estão promovendo uma revolução na genômica animal, pois permitem varredura do genoma para milhares desses marcadores ao mesmo tempo, a um menor custo e de maneira mais rápida, quando comparado a técnicas de genotipagem como PCR-RFLP (do inglês, *polymerase chain reaction - restriction fragment length polymorphism*) e detecção de microssatélites. O desenvolvimento de metodologias para

genotipagem massal de dezenas de milhares até um milhão de SNP em um único ensaio, levou os custos de geração de dados de US\$ 0.10 a US\$ 0.001 por SNP genotipados, quando comparado à técnica de RFLP (CAETANO, 2009).

Atualmente, existe diversas plataformas de genotipagem com capacidades distintas de geração de dados genéticos e espécie do qual foi originada (Tabela 1). Dentre as empresas fabricantes, a Illumina e Affymetrix se destacam no mercado mundial, as duas tiveram seus primeiros painéis destinados a humanos (LI et al, 2008).

**Tabela 1.** Principais painéis de genotipagem de SNP atualmente disponíveis no mercado.

Identificação	Organismo	Número de SNP	Empresa
HumanOmni5-Quad	Humano	4.3 Milhões	Illumina
BovineHD BeadChip	Bovino	777,000	Illumina
Axiom <sup>®</sup> Genome-Wide Genotyping Array	Frango	580,000	Affymetrix
Axiom <sup>®</sup> Salmon Genotyping Array	Salmão	130,000	Affymetrix
Axiom <sup>®</sup> Strawberry Genotyping Array	Morango	95,062	Affymetrix
Axiom <sup>®</sup> Buffalo Genotyping Array	Bubalino	90,000	Affymetrix
PorcineSNP60 BeadChip	Suíno	65,000	Illumina
BovineSNP50 BeadChip	Bovino	54,609	Illumina
EquineSNP50 BeadChip	Equino	54,602	Illumina
OvineSNP50 BeadChip	Ovino	54,241	Illumina
MaizeSNP50 BeadChip	Milho	50,000	Illumina
GeneChip <sup>®</sup> Rice 44K SNP Genotyping Array	Arroz	44,100	Affymetrix

Fonte: Affymetrix e Illumina.

Os SNP-chips para humanos são os mais robustos e, conseqüentemente, reportam uma quantidade maior de informações genômicas, podendo interrogar cerca de 4.3 milhões de marcadores, alguns ainda permitem avaliar até 500 mil SNP adicionais customizados, possibilitando ao pesquisador adequar o painel as suas necessidades. Esses painéis são voltados principalmente à área da saúde em estudos oncológicos e também de exoma (HEDGES et al., 2009; ILLUMINA, 2014).

Na agropecuária existem painéis destinados às espécies animais (bovino, bubalino, frango, ovino, suíno, equino) e vegetais (arroz, milho, morango) (AFFYMETRIX, 2013; ILLUMINA, 2014). A aplicabilidade desta ferramenta genômica nesta área busca por indivíduos geneticamente superiores ou polimorfismos associados a características fenotípicas de importância econômica, visando maior produtividade e custos reduzidos com melhoramento genético (RESENDE et al., 2008; CORNELIS et al., 2010).

As plataformas de genotipagem de SNP promovem uma varredura genômica ampla com acurácia elevada e rapidez. Proporciona o estudo de características complexas de herança poligênica e mecanismos genômicos desconhecidos, devido sua densa cobertura, de modo que muitos genes de um caráter quantitativo fiquem em desequilíbrio de ligação com ao menos uma parte dos marcadores e, assim, auxiliando na compreensão dos efeitos de diversos genes para uma característica quantitativa (RESENDE, 2008; LEE et al., 2012).

Os painéis SNP-chip se apresentam uma tecnologia promissora na identificação de variações genéticas, genes candidatos e funções biológicas correlacionadas com características de interesse principalmente no homem, vegetais e animais de interesse agropecuário.

## Referências

- Affymetrix (2013) Product literature. <http://www.affymetrix.com/estore>. Accessed 26 Jan. 2014
- Caetano AR (2009) Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. *Revista Brasileira de Zootecnia* 38 (8):64-71
- Cornelis MC, Agrawal A, Cole JW, Hansel NN, Barnes KC et al (2010) The Gene, Environment Association Studies consortium (GENEVA): maximizing the knowledge obtained from GWAS by collaboration across studies of multiple conditions. *Genetic epidemiology* 34 (4):364-372
- Faleiro F (2007) Marcadores genético-moleculares aplicados a programas de conservação e uso de recursos genéticos. 1 edn. Embrapa Cerrados, Planaltina, DF
- Ferreira ME, Grattapaglia D (1996) Introdução ao uso de marcadores moleculares em análise genética. 2 edn. Embrapa Cenargem

- Hedges DJ, Burges D, Powell E, Almonte C, Huang J, Young S, Boese B, Schmidt M, Pericak-Vance MA, Martin E, Zhang X, Harkins TT, Zuchner S (2009) Exome sequencing of a multigenerational human pedigree. *PLoS One* 4 (12):e8232
- Illumina (2014) Product literature. <http://www.illumina.com/>. Accessed 26 Jan. 2014
- International HapMap Consortium (2003) The International HapMap Project. *Nature* 426 (6968):789-796
- Lee JB, Yoo CK, Jung EJ, Hwang JH, Seo BY et al (2012) A missense mutation (c.1963A<G) of the complementary component 2 (C2) gene is associated with serum Ca(+)(+) concentrations in pigs. *Molecular biology reports* 39 (10):9291-9297
- Li M, Li C, Guan W (2008) Evaluation of coverage variation of SNP chips for genome-wide association studies. *European journal of human genetics: EJHG* 16 (5):635-643
- Passos-Bueno MRS, Moreira ES (2004) Ferramentas básicas da genética molecular humana. In: Mir L (ed) *Genômica*. Atheneu, São Paulo, pp 43-69
- Resende MDV, Lopes PS, da Silva RL, Pires IE (2008) Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. *Pesquisa Florestal Brasileira* (56):63-77
- Sawicki MP, Samara G, Hurwitz M, Passaro E, Jr. (1993) Human Genome Project. *American journal of surgery* 165 (2):258-264
- U.S. DEPARTMENT OF ENERGY OFFICE OF SCIENCE. HUMAN GENOME PROJECT. (2013). <http://www.ornl.gov/hgmis>. Accessed 27 Jan. 2014
- Watson JD (1990) The human genome project: past, present, and future. *Science* 248 (4951):44-49

## **Capítulo II**

### **Estudo de associação genômica ampla para característica reprodutiva em fêmeas da raça Nelore**

Manuscrito em elaboração a ser submetido:

**Molecular Biology Reports**

**ISSN: 1573-4978**

## **Estudo de associação genômica ampla para característica reprodutiva em fêmeas da raça Nelore**

**André Vieira do Nascimento · Alexéia Barufatti Grisolia**

A.V. Nascimento · A.B. Grisolia

Faculdade de Ciências Biológicas e Ambientais, Universidade Federal da Grande Dourados, Dourados 79804-970, Brasil

e-mail: andrevn16@gmail.com

**Resumo** A idade ao primeiro parto (IPP) é uma característica economicamente importante em gado de corte. Com o objetivo de avaliar a utilização de uma ferramenta genômica para associação de polimorfismo de nucleotídeo de sítio único (SNP) distribuídos no genoma bovino com a característica reprodutiva de IPP em vacas Nelore, o estudo foi conduzido em 55 animais genotipados em painel de alta densidade para analisar mais de 777.000 marcadores. O SNP rs110169757 localizado no BTA2:6151365 foi considerado o mais significativo e possivelmente correlacionado com o fenótipo de IPP. Na região de 500 Kb para cada lado da posição do SNP foram encontrados 19 *loci* de características quantitativas (QTL) anteriormente descritos na literatura, sendo três destes relacionados com fenótipos reprodutivos de Intervalo para o primeiro estro pós-parto, fixação do úbere e *strenght*. Os resultados apresentados apontam que o uso da ferramenta genômica foi eficiente na avaliação e análise de milhares de SNP presentes no DNA de novilhas Nelore com fenótipo de idade ao primeiro parto, obtendo alta densidade de genótipos acurados. Sendo possível a verificação regiões de QTLs relacionados à características reprodutivas.

**Palavras-chave** *Bos indicus* · SNP · Idade ao primeiro parto · Precocidade sexual

**Abstract** The age at first calving (AFC ) is an economically important trait in beef cattle. Aiming to evaluate the use of a genomic tool for association of single nucleotide polymorphism site ( SNP ) distributed in the bovine genome with reproductive characteristic of IPP in Nellore cows the study was conducted in 55 animals genotyped at high density panel to analyze more than 777,000 markers . The SNP located in rs110169757 BTA2 : 6151365 was considered the most significant and possibly correlated with the phenotype of IPP . In the region of 500 Kb for each side of the SNP position of 19 quantitative trait loci (QTL)

previously described in the literature were found , three of these being related to reproductive phenotypes interval to first postpartum estrus , udder attachment and strength . These results suggest that the use of genomics tool was effective in the evaluation and analysis of thousands of SNPs present in the DNA of Nelore heifers phenotype of age at first calving , obtaining high density accurate genotypes . With the possible regions of QTLs for reproductive traits verification.

**Keywords** *Bos indicus* · SNP · Age at first calving · Sexual precocity

### Abreviaturas

CR	Call Rate (% de determinação)
DNA	Desoxyribonucleic Acid (Ácido desoxirribonucleico)
GWAS	Genome-wide association study (Estudo de associação genômica Ampla)
HWE	Hardy-Weinberg Equilibrium (Equilíbrio de Hardy-Weinberg)
IBS	Identity By State (Identidade por estado)
IPP	Idade ao Primeiro Parto
Kb	Kilo base ( $10^3$ pares de bases)
MAF	Minor allele frequency (Frequência do menor alelo)
Mb	Mega base ( $10^6$ pares de bases)
QC	Quality control (Controle de qualidade)
QTL	Quantitative trait <i>loci</i> ( <i>loci</i> de característica quantitativa)
SNP	Single nucleotide polymorphism (Polimorfismo de nucleotídeo de sítio único)

### Introdução

As raças zebuínas são extensivamente utilizadas nacionalmente, chegando a representar cerca de 80% do rebanho de corte, desta população 90% são animais da raça Nelore. Os principais fatores para esta preferência se deve às suas características adaptativas as condições climáticas do Brasil, maior capacidade de aproveitamento de alimento e resistência a parasitas (ACNB, 2014). Porém, para se consolidar no cenário do comércio mundial e atender ao mercado consumidor constantemente ascendente e cada vez mais exigente é necessário o desenvolvimento de tecnologias que possibilitem melhorar o aproveitamento dos rebanhos, pois as áreas produtivas são limitadas. (ABIEC, 2014).



Os aspectos reprodutivos interferem diretamente na lucratividade do rebanho na pecuária de corte. Devido a esse fato, pesquisas relacionadas à área de biotecnologia animal que possibilitem investigar tecnologias que permitam identificar no rebanho indivíduos com maior eficiência reprodutiva são de grande interesse econômico (VAICIUNAS, 2008).

A puberdade é definida como o período transitório entra a fase pós-natal e a fase adulta, durante o qual a maturidade do sistema reprodutivo é atingida, capacitando uma fêmea para gerar descendentes (TERASAWA e FERNANDEZ, 2001). As fêmeas zebuínas são mais adaptadas ao clima tropical e resistentes a parasitas, entretanto apresentam um retardo para atingirem a puberdade, em média de 6 a 12 meses, quando comparadas às fêmeas taurinas, característica que traz prejuízos ao produtor (WILTBANK et al., 1969; DOBSON et al., 1986; PIRLO, 2000).

A idade ao primeiro parto (IPP) é um parâmetro que marca o ingresso da novilha no sistema produtivo de gado de corte, fenótipo facilmente coletável. A redução na IPP leva a maior eficiência produtiva, pois resulta em animais com melhor desempenho reprodutivo (PIRLO et al., 2000; VALLE et al., 2000; SASAKI et al., 2013)

Estudos relacionados à identificação de genes ligados a fenótipos de interesse econômico obtiveram grandes avanços com o desenvolvimento de *chips* de SNP de alta densidade, de modo que a genotipagem feita por esta ferramenta analisa dezenas de milhares simultaneamente (BOVINE HAPMAP CONSORTIUM, 2009; MATUKUMALLI et al., 2009).

Os painéis de alta densidade podem ser utilizados para estudo de associação genômica amplas ou GWAS (do inglês, *Genome Wide Association Study*) ou para seleção genômica ampla ou GWS (do inglês, *Genome Wide Selection*). A GWS pode proporcionar ao produtor a seleção precoce direta de animais com genótipo superior para produção, pois atua em genes que serão expressos na fase adulta, genes estes muitas vezes para uma característica quantitativa. O GWAS permite identificar polimorfismos que possam estar associados com característica fenotípica de interesse comercial (HIRSCHHORN e DALY, 2005; RESENDE et al., 2010)

Atualmente, características mensuráveis diretamente nas fêmeas tem ganhado atenção, pois estas influenciam o sistema pecuário de corte (SANTANA et al., 2012). Assim a redução da idade ao primeiro parto objetivando maior produtividade é uma das grandes preocupações de pesquisadores e produtores (NOGUEIRA, 2004). No entanto, a estimativa de

herdabilidade para IPP são de baixas a moderadas, variando entre 0,09 a 0,19 (PEREIRA et al., 2001; PEREIRA et al., 2002; DIAS et al., 2004).

Em décadas anteriores, os marcadores microssatélite foram intensamente utilizados na prospecção de *loci* de características quantitativas, denominados QTLs (do inglês, *Quantitative Trait Locus*) em bovinos, resultando na identificação de diversos QTLs ao longo do genoma bovino. Atualmente existem 8,305 QTLs identificados para características de interesse econômico para gado de corte e leite (CattleQTLdb, 2014). Neste contexto, o GWAS tem demonstrado resultados satisfatórios na averiguação de bases genéticas dos processos reprodutivos.

Este trabalho teve como objetivo avaliar a utilização de uma ferramenta genômica na associação de marcadores SNP distribuídos no genoma bovino com característica reprodutiva de IPP em vacas Nelore.

## **Material e Métodos**

### **Animais e dados fenotípicos**

O grupo amostral foi composto por 55 fêmeas da raça Nelore, nascidas entre os anos de 1996 e 2004. Esses animais foram selecionados dentre todo o rebanho comercial de fêmeas ativas da propriedade São Jorge do Maracay (Iguatemi – MS), tendo como critério de seleção a presença dos dados de Idade ao Primeiro Parto (IPP) e paternidade conhecida. A fazenda adota manejo alimentar exclusivamente a pasto.

No período de 1995 a 2004, foi estabelecida na propriedade uma estação de acasalamento de novilhas denominada “estação de desafio”, ocorrendo independente do peso e condição corporal do animal. As fêmeas avaliadas foram desafiadas aproximadamente aos 18 meses de idade e as “estações de desafio” aconteceram entre os meses de janeiro e fevereiro com duração de 45 dias. As novilhas foram expostas às mesmas condições de manejo e em monta natural com um grupo de touros jovens, na proporção de um reprodutor para 25 matrizes.

Os 55 indivíduos foram separados em dois agrupamentos, denominados de Grupo de Prenhas (P) para novilhas com prenhez confirmada aos 18 meses, e Grupo de Não Prenhas (NP) para não gestantes aos 18 meses, com número amostral de 30 e 25, respectivamente.

Dessa forma, a característica analisada foi característica binária, onde os grupos P e NP assumiram os valores de 1 e 0, respectivamente.

A estatística descritiva para a característica de Idade ao Primeiro Parto dos grupos P e NP pode ser visualizada na Tabela 2.

**Tabela 2.** Estatística descritiva da idade ao primeiro parto (IPP) em meses para os grupos de prenhas (P) e não prenhas (NP) e o total.

Grupo	N	Média	Desvio-padrão	Mediana	Min.	Máx.
P	35	26.85	0.93	26.84	24.11	28.55
NP	20	33.27	0.69	33.61	31.90	33.94
Total	55	29.19	3.22	27.5	24.11	33.94

#### Colheita de sangue e extração de DNA

Para as análises laboratoriais foram coletados 4.5 mL de sangue periférico de cada animal por punção da veia jugular. Essas amostras foram armazenadas em tubos do tipo Vacuntainer com EDTA potássico à -20°C até o processamento. O material genético dos animais foi extraído com auxílio do kit comercial de extração de DNA e tecido (DNeasy Blood & Tissue Kit, QIAGEN, Espanha). Em seguida as amostras foram avaliadas quanto a sua qualidade e concentração por meio de espectrofotometria de micro-volume (NanoDrop 2000, Thermo Fisher Scientific Inc., EUA). As concentrações médias das amostras ficaram em torno de 50 ng/ $\mu$ L e a razão 260:280 em 1.8 ( $\pm$  0.2).

#### Análise de SNP

Os animais foram genotipados no painel de alta densidade BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA), que contém sondas para interrogar mais de 777.000 SNP com intervalo médio de 3.4 e mediana de 2.68 kb (kilobase, 1kb = 1,000 pares de bases) entre marcadores. Este painel foi construído a partir de amostras DNA obtidas a partir de animais taurinos, zebuínos e híbridos, dentre raças destinadas para corte e produção de leite como: Angus, Jersey, Hereford, Holstein, Brahman, Gir, Nelore, Brangus, Santa Gertrudis, entre outras (ILLUMINA, 2012).

O procedimento de genotipagem e análise de SNP foram realizados em laboratório comercial, seguindo o protocolo preconizado pelo fabricante (ILLUMINA, 2014).

O agrupamento e a determinação dos genótipos foram realizados no *software* GenomeStudio v2011.1 (Illumina Inc., EUA). O programa faz a conversão das intensidades fluorescentes para cada amostra em genótipos de SNP. Estima o parâmetro denominado “*GenCall score*” (GCScore) que indica a confiabilidade dos genótipos determinados a partir de algoritmos que analisam os dados de normalidade, agrupamento e determinação dos genótipos a partir dos dados gerados da leitura do painel. O cálculo do GCScore é baseado na dispersão da amostra em relação ao centro da área de distribuição esperada para intensidade e frequência luminosa normalizada daquele genótipo (AA, AB ou BB), variando de 0 (ponto mais distante do centro) a 1 (centro do cluster). Assim de acordo com o fabricante, valores inferiores a 0.15 são designados como genótipos não determinados, o intervalo de 0.15 a 0.7 considerado como uma faixa sensível e valores acima de 0.7 indicam genótipos excelentes.

Neste estudo os genótipos com sucesso foram determinados pelo GCScore > 0.7, exportados como arquivo de texto e importados para o pacote estatístico R (versão 3.0.2), utilizando-se de arquivos do programa Plink. As análises de QC e teste de associação foram realizadas utilizando funções customizadas em R e biblioteca GenABEL (versão 1.8-0) (AULCHENKO et al., 2007; PURCELL et al., 2007; AULCHENKO et al., 2011).

As coordenadas dos SNP no ensaio foram baseadas na versão de montagem do genoma *Bos taurus* UMD 3.1 (acesso GenBank DAAA00000000.2).

## Controle de qualidade

Antes das análises de associação os dados genômicos foram submetidos a um controle de qualidade (QC) inicial, onde os critérios de exclusão de marcadores foram: não mapeados, mapeados para a mesma coordenada genômica e não autossômicos, ou seja, os SNP presentes nos cromossomos X, Y e DNA mitocondrial. Em seguida foi realizada a análise de IBS (do inglês, *Identity By State*) para avaliação de amostras possivelmente duplicadas, onde é calculada a proporção de alelos compartilhados por estado entre os indivíduos, assim amostras com IBS > 95% foram removidas.

Com a intenção de excluir marcadores e amostras potencialmente problemáticos, foram removidos os SNP levando em consideração os seguintes critérios: 1) MAF (do inglês,

*Minor Allele Frequency*), marcadores com frequência alélica menor ou igual a 2%; 2) HWE (do inglês, *Hardy Weinberg Equilibrium*), SNP com valor  $P < 10^{-5}$  para o Teste Exato de Fisher (ou seja, com desvios extremos do HWE, sugerindo potencial erro de genotipagem); 3) “*Call rate*” ( $CR_{SNP}$ ), polimorfismo que não se encontra ao menos em 98% da população. Após a verificação dos SNP foi realizado os QC por indivíduo, “*Call rate*” ( $CR_{IND}$ ), onde amostras com menos de 90% de genótipos determinados foram desconsideradas.

O procedimento de controle de qualidade dos genótipos foi realizado por meio de uso de *scripts* customizados e da biblioteca GenABEL (AULCHENKO et al., 2007) do pacote estatístico R.

### Teste de associação

A função “*qtscore*” da biblioteca GenABEL foi utilizada no teste de escore para associação entre a característica e o polimorfismo genético. Resíduos foram estimados após o ajustamento para covariáveis (grupo de contemporâneos (GC) (ano e trimestre de nascimento) e paternidade). Neste caso, os efeitos foram aproximados aos *Odds Ratios* (ORs) esperados em um modelo de regressão logística.

O cálculo do fator de inflação  $\lambda$  foi estimado para verificar a confiabilidade dos resultados. Valores de  $\lambda$  acima de 1.1 sugerem uma inflação generalizada no modelo. Com propósito de controle suplementar, foi plotado um gráfico quantil-quantil (Q-Q) para verificar o ajuste dos valores de  $P$  para uma distribuição Qui-quadrado.

Os resultados foram apresentados na forma de “Manhattan plot” com o co-logaritmo dos valores de  $P$ , fazendo uso de Bonferroni, o qual é um método para correção de testes múltiplos, onde  $\alpha$  é definido por  $0,05/N$ , sendo  $N$  o número de marcadores testados (BENJAMIN e HOCHBERG, 1995) e, com intuito de explorar SNP possivelmente associáveis, foi estabelecida a linha de significancia ( $-\log_{10}(P) = 5$ ).

### Exploração dos SNP significativos

Para a verificação da região genômica para o SNP mais significativo, foi explorada uma janela de 500 Kb (kilobase) de cada lado do marcador usando a ferramenta *BioMart* (KINSELLA et al., 2011) do banco de dados Ensembl, com base no assembly UMD v3.1. O

banco de dados *Cattle QTLdb* (HU et al., 2013) foi utilizado para examinar se o SNP mais significativo se encontrava dentro de regiões de QTLs de espécie bovina já catalogado na literatura e genes mais próximos.

## Resultados e Discussão

### Controle de Qualidade

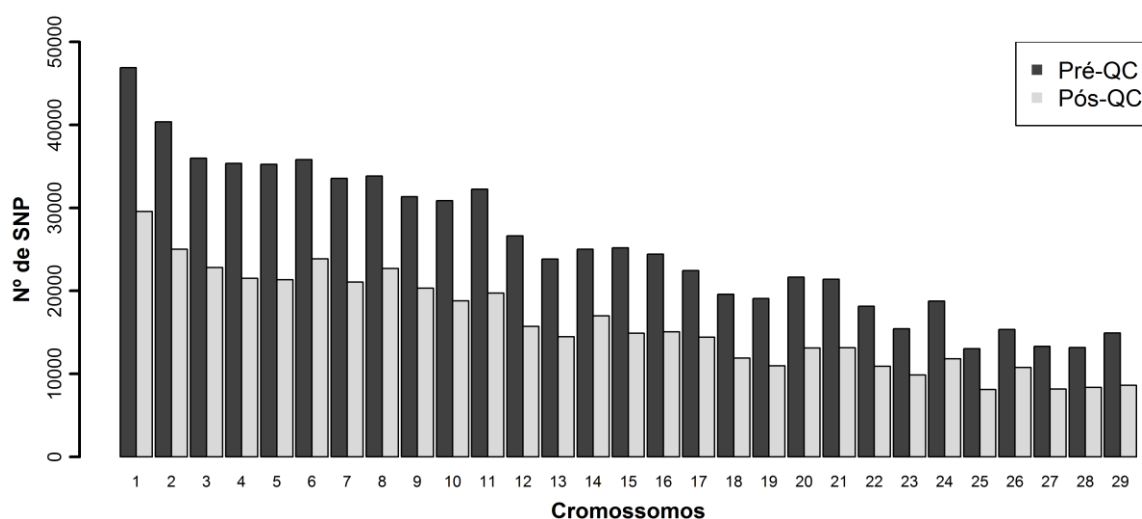
De um montante inicial de 786,798 SNP, 43,889 (5.5%) não eram marcadores autossômicos. Cinquenta e quatro marcadores autossômicos (0.01%) apresentaram as mesmas coordenadas genômicas e foram excluídos. A análise de IBS não detectou nenhuma amostra possivelmente duplicada. Um total de 203,561 (28.8%) SNP foram excluídos por apresentarem  $MAF < 2\%$ . O número de marcadores removidos por  $CR_{SNP} < 98\%$  e teste exato de Fisher  $P$ -value para HWE  $< 10^{-5}$  foi 44,418 (6.3%) e 772 (0.1%) respectivamente. O  $CR_{IND}$  por indivíduo não removeu nenhum animal. Ao fim do QC o grupo amostral foi de 55 animais e 464,005 marcadores.

O número amostral, espécie e painel de genotipagem são fatos considerados importantes na escolha do critério de “*Call rate*” a serem aplicados no estudo. Jiang et al. (2013) adotaram um  $CR_{IND}$  rigoroso (99.9%) com grupo amostral constituído de 96 animais da raça Holstein, que nesse caso, apresenta grande quantidade de marcadores polimórficos no painel (BovineHD Genotyping BeadChip), quando comparado à raça Nelore. Em outro estudo utilizando animais da raça Japanese Black, o  $CR_{IND}$  usado foi de 95%, um valor mais flexível considerando o número 1156 de indivíduos, entretanto o SNP-chip em questão apresentava 54,001 marcadores (BovineSNP50 BeadChip) (NISHIMURA et al., 2012). Assim o uso de  $CR_{IND}$  de 90%, se deve ao fato do presente estudo apresentar um grupo amostral de 55 animais e estes serem zebuínos.

Utsunomiya et al. (2013) em estudo com bovinos Nelore utilizando painel de alta densidade (777,961 marcadores) usando os mesmos parâmetros no QC obtiveram proporções de exclusão de SNP, não autossômicos e  $MAF$  similares ao presente trabalho, 5.5% e 30.4% respectivamente, sendo a  $MAF$  responsável pelo maior número de marcadores excluídos. Os parâmetros de  $CR_{SNP}$  (16.7%) e teste exato de Fisher  $P$ -value para HWE (1.8%) teve

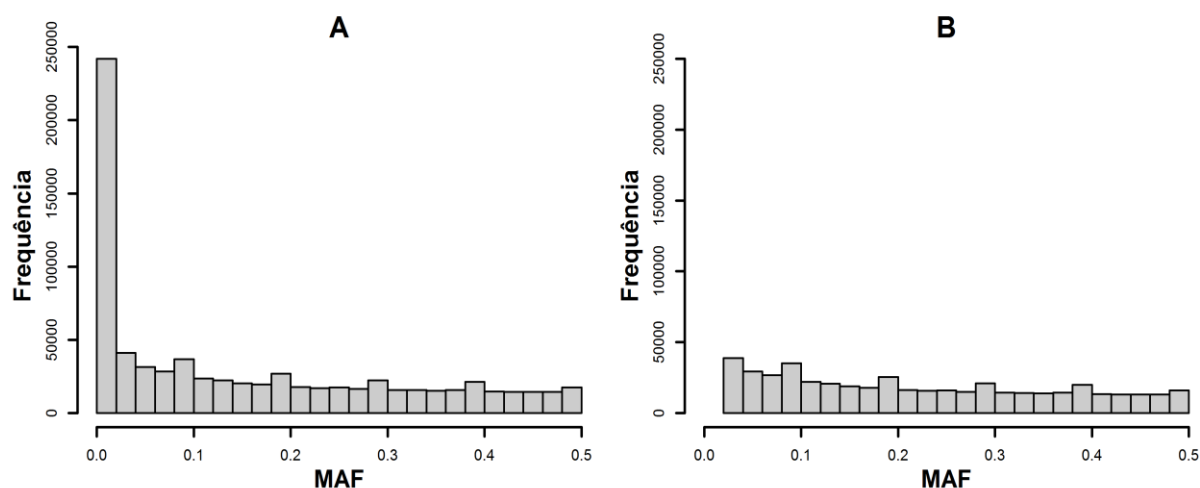
proporções distintas, possivelmente isto está relacionado com o número amostral maior (654 animais).

A Figura 2 mostra o número de SNP autossômicos antes e após o QC, onde houve redução de 37.5% da quantidade de marcadores. Esse resultado corrobora com o trabalho de Utsunomiya et al (2013), que utilizou o painel de genotipagem BovineHD Genotyping BeadChip (777,961 SNP) em Nelore e obteve uma exclusão de marcadores por volta de 40.9%. Entretanto, o mesmo painel aplicado em bovinos da raça Holstein foi possível obter melhor aproveitamento de SNP, com taxa de remoção por volta de 30.7% (SHIN et al., 2013), isto devido as sondas desta ferramenta serem construídas predominantemente com base no genoma de animais taurinos (452 indivíduos) quando comparado aos zebuínos (104), acarretando assim na existência de grande quantidade de polimorfismo monomórficos para zebuínos (ILLUMINA, 2012).



**Figura 2.** Representação do número de marcadores por cromossomo antes e depois da realização do controle de qualidade dos dados genotípicos.

As distribuições das MAFs dos marcadores pré e pós-controle de qualidade podem ser observadas na Figura 3, onde os marcadores remanescentes apresentam MAF superior a 2%.

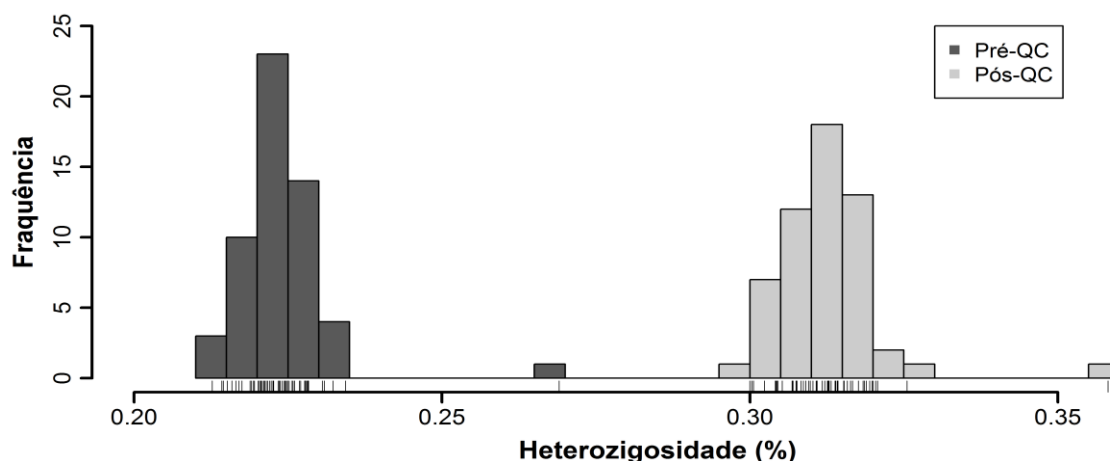


**Figura 3.** Distribuição da frequência do menor alelo (MAF) dos SNP antes (A) e após (B) o controle de qualidade.

Em estudos com bovinos os valores usando para a MAF variam de 1% a 10% (NISHIMURA et al., 2012; SHIN et al., 2013; UTSUNOMIYA et al., 2014), de maneira que o critério de MAF é mais rigoroso quando este percentual for maior, o que acarreta em elevadas reduções no número de marcadores nas análises. Conforme Anderson et al. (2010), os valores de MAF em trabalhos com humanos costumam variar de 1% a 2%, indicando que o presente estudo está de acordo com pesquisas anteriores. Marcadores com MAF inferior a 1% são raros na população e não apresenta poder estatístico, de modo que a exclusão destes do grupo de dados é recomendada (TURNER et al., 2011).

A distribuição da heterozigosidade dos SNP apresentou médias de 22% e 31%, para antes e depois do controle de qualidade, respectivamente. A distribuição da heterozigosidade dos marcadores pode ser observada na Figura 4.





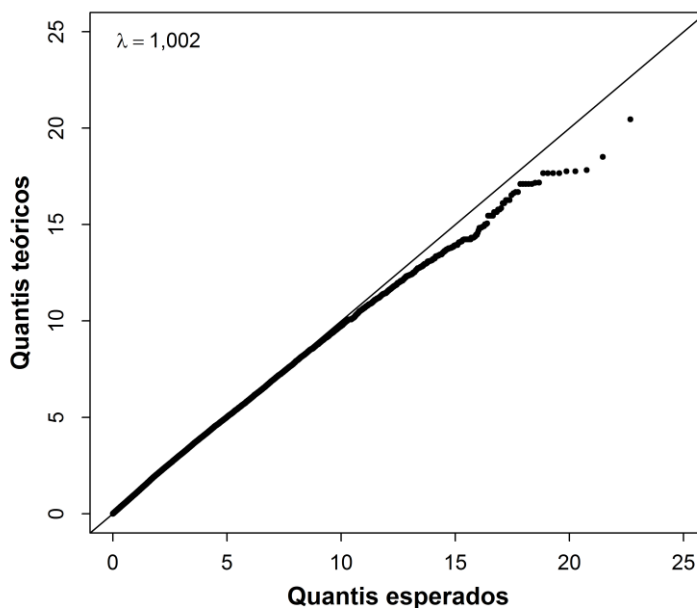
**Figura 4.** Comparativo da distribuição da heterozigosidade dos marcadores moleculares na população, anterior e posterior ao controle de qualidade (QC).

O aumento médio de 41% da distribuição da heterozigosidade após o controle de qualidade se explica devido ao grande número de marcadores homozigotos removidos no critério de MAF. A heterozigosidade é um fator de importância na obtenção da informação genética, pois na associação de polimorfismos com características fenotípicas ocorre a avaliação dos efeitos dos diferentes alelos, de modo que, a baixa frequência de um dos alelos não traz informação suficiente para inferir quaisquer correlações (BOTSTEIN et al., 1980).

#### Teste de associação

Sob a hipótese nula da inexistência de associação para qualquer *locus* SNP, assim espera-se que os pontos se distribuam sobre a linha de tendência central que relaciona os valores esperados e observados. A ocorrência de desvios nesta linha indicam *loci* que contrastam com a hipótese nula (associação significativa), sendo que um grande número é indicativo de inflação no teste estatístico devido à existência de algum viés nos dados (MATOS, 2012).

O gráfico quantil-quantil (Q-Q) (Figura 5) mostrou que o teste estatístico utilizado foi aceitável ( $\lambda = 1.002578$ ). Mesmo com valor de  $\lambda$  ótimo se pode observar uma suave deflação na distribuição dos pontos após o centro do gráfico, possivelmente devido ao número de indivíduos reduzido.



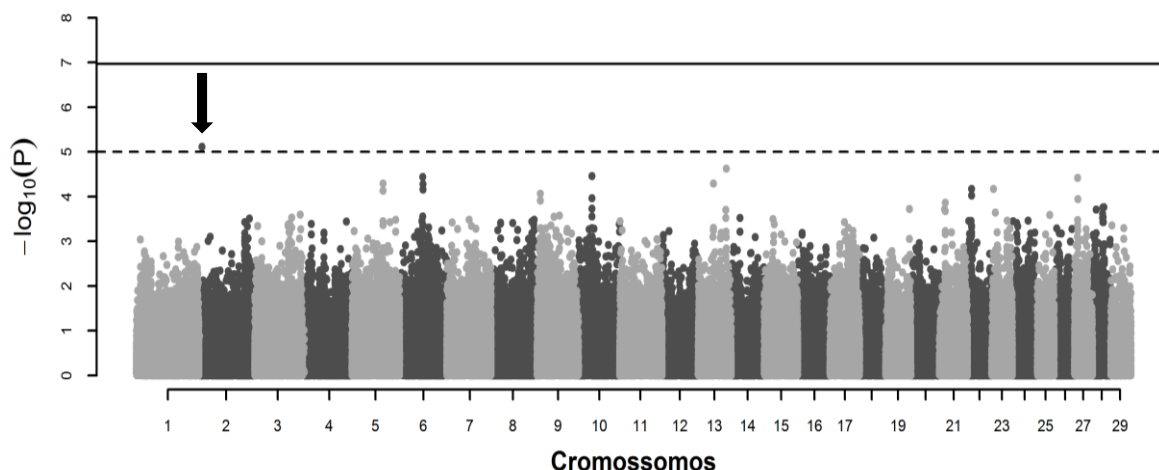
**Figura 5.** Gráfico quantil-quantil dos resultados do GWAS para IPP. Onde a linha preta representa a distribuição esperada e os pontos pretos à distribuição observada da estatística Qui-quadrado ( $\chi^2$ ).

A associação dos polimorfismos com a característica de IPP partiu do pressuposto de um modelo de herança aditiva, ou seja, que o valor fenotípico dependia da quantidade de réplicas do alelo menor. Desta forma, no GWAS realizado os SNP testados não conseguiram atingir o limiar de significância pela correção de Bonferroni. Entretanto, o SNP rs110169757 foi considerado o mais significativo e possivelmente correlacionado com o fenótipo de IPP, este se localiza no cromossomo autossômico bovino 2 (BTA2: 6151365) em região intergênica e se apresentou acima da linha de associação sugestiva (Figura 6).

Nos resultados do GWAS demonstrados na forma de Manhattan plot se pode observar a formação de picos difusos distribuídos nos diferentes cromossomos, remetendo que diversas regiões genômicas apresentaram algum tipo de relação com a característica em estudo. De fato, muitas das características reprodutivas têm padrão de herança poligênica, assim sendo, o fenótipo influenciado por múltiplos genes (MATOS, 2012).

Em estudo de associação genômica ampla, a interpretação dos resultados apresentam uma dificuldade fundamental conhecida como testes múltiplos, de modo que pesquisadores usam metodologias corretivas, por exemplo, o ajuste de Bonferroni, definindo, assim, um índice de significância (MOSKVINA e SCHIMIDT, 2008; FROMMLET et al., 2012). Este

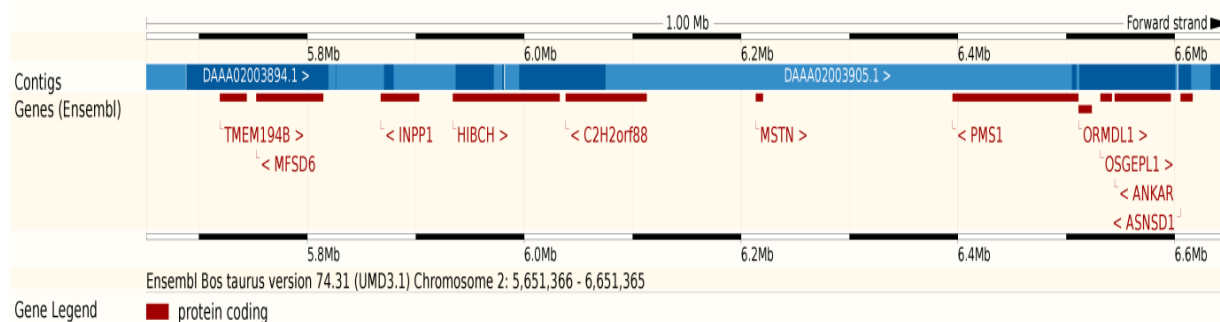
problema com múltiplos teste se intensificam, principalmente, quando a quantidade de marcadores aumenta e a população é reduzida (FINLAY et al., 2012). Neste contexto, é possível que neste estudo o método de correção tenha sido muito restrigente, assim para evitar que associações em potencial fossem eliminadas, assumimos um limiar de  $P = 10^{-5}$  para investigação.



**Figura 6.** Manhattan plot dos resultados de associação genômica para idade ao primeiro parto, onde a linha tracejada é a linha de associação sugestiva ( $\alpha = 10^{-5}$ ) e a linha contínua representa a correção ao ajustamento de Bonferroni ( $\alpha = 1.07 \times 10^{-7}$ ).

Estudos do tipo GWAS na espécie humana costumam apresentar número de indivíduos acima de 10,000, contrastando com a realidade ocorrente nos trabalhos com espécies domésticas (SUN, 2012). Partindo deste fato, pode-se sugerir que o grupo amostral analisado deste trabalho foi um fator limitante para obtenção de associações fortemente significativas.

A região de 500 Kb para ambos os lados da localização do polimorfismo (rs110169757) com base no alinhamento UMD v3.1 utilizando a ferramenta *BioMart* do Ensembl pode ser visualizada na Figura 7 e a lista com os 11 genes encontrados nesta mesma área na Tabela 3.



**Figura 7.** Imagem da região correspondente a 1 Mb em volta do SNP rs110169757. Em azul são representados os *contigs* do alinhamento UMD v3.1 e as barras vermelhas representam as regiões codificantes referentes ao genes.

Dentre os genes investigados houve evidências relacionando às características de desenvolvimento muscular, crescimento corporal, comportamental, reparação e manutenção do genoma nuclear (BARKER et al 1995; BERGER et al.,2012; JIMÉREZ et al., 2013; SINGH et al., 2014). Entretanto não foi encontrada correlação direta com fenótipo de idade ao primeiro parto.

Na análise do banco de QTLs em uma janela de 1 Mb ao redor do marcador mais sugestivo, localizou-se a presença de 19 QTLs (Tabela 4) relacionados com características de produção, carne e carcaça, leite e reprodução em diferentes raças, sendo três QTLs ligados a esta última. A região do BTA2 apontada neste estudo foi associada anteriormente com fenótipo reprodutivo, sendo o intervalo para o primeiro estro pós-parto relacionado à fertilidade da matriz e, fixação do úbere e *strenght*, correlacionados com as condições corporais de suporte a maternidade (DAWUDA et al, 1988; ASHWELL et al, 2005; SCHULMAN et al, 2008).

A compreensão da interação dos genes na influencia de um fenótipo é um enorme desafio, que muitas vezes é prejudicada por confundidores de origem ambiental sobre o fenótipo. Entretanto, as informações geradas neste estudo procurou estabelecer conexões funcionais do genótipo com o fenótipo de IPP, promovendo a investigação de processos biológicos que poderão contribuir em pesquisas posteriores destinadas à validação de genes candidatos e entendimento de mecanismos de interação gênica.

**Tabela 3.** Lista dos genes presentes em uma janela de 1 Mb ao redor do marcador mais significativo (rs110169757).

<b>Gene</b>	<b>Ensembl ID</b>	<b>BTA 2 coordenada</b>	<b>Distância do SNP (kb)</b>	<b>Tamanho do gene (pb)</b>	<b>Descrição</b>
C2H2orf88	ENSBTAG00000026994	6038403:6113008	38.3	580	<i>Bos taurus</i> small membrane A-kinase anchor protein
MSTN	ENSBTAG00000011808	6213566:6220196	62.2	2,768	<i>Bos taurus</i> myostatin
HIBCH	ENSBTAG00000007787	5934156:6032782	118.5	1,390	3-hydroxyisobutyryl-CoA hydrolase
PMS1	ENSBTAG00000018795	6395005:6511175	243.6	3,088	Postmeiotic segregation increased 1
INPP1	ENSBTAG00000007584	5867812:5903148	248.2	1,556	<i>Bos taurus</i> inositol polyphosphate-1-phosphatase
MFSD6	ENSBTAG00000007719	5752911:5814559	336.8	2,939	Major facilitator superfamily domain containing 6
ORMDL1	ENSBTAG00000012561	6511310:6523389	359.9	1,990	<i>Bos taurus</i> ORM1-like 1
OSGEPL1	ENSBTAG00000012557	6531525:6542021	380.1	1,394	<i>Bos taurus</i> O-sialoglycoprotein endopeptidase-like 1
ANKAR	ENSBTAG00000012554	6544589:6596238	383.2	4,302/4,329	Ankyrin and armadillo repeat containing
TMEM194B	ENSBTAG00000015678	5719428:5744015	407.3	1,634	Transmembrane protein 194B
ASNSD1	ENSBTAG00000000492	6605444:6616484	453.0	2,384	Asparagine synthetase domain containing 1

**Tabela 4.** QTLs do banco de dados *Cattle QTLdb* encontrados em uma janela de 1Mb nas proximidades do SNP mais significativo (rs110169757).

<b>Característica</b>	<b>Grupo</b>	<b>BTA 2 coordenada</b>	<b>QTLdb ID</b>	<b>PubMed ID</b>
<i>Strength</i>	Reprodução	5897389:10937591	1554	16230715
Fixação do úbere	Reprodução	5897389:7763197	1551	16230715
Intervalo para o primeiro estro pós-parto	Reprodução	6063867:6283723	5028	18298935
Produção de gordura no leite	Leite	5897389:32464114	1675	16167984
Percentual de proteína no leite	Leite	2134354:75456158	3531	14762091
Peso corporal (desmame)	Produção	6063867:6283723	15718	22303340
Largura da articulação do quadril	Produção	5897389:7763197	1555	16230715
Peso corporal (nascimento)	Produção	5897389:9662643	2751	9498354
Rendimento do produto no mercado	Produção	5897389:9662643	2752	9498354
Percentual de gordura	Carne e carcaça	6011260:6231116	6297	19422365
Teor de ácido linoleico conjugado	Carne e carcaça	5897389:9662643	4855	17894565
Área de olho de lombo	Carne e carcaça	5897389:9662643	2753	9498354
Grau de rendimento	Carne e carcaça	5897389:9662643	2754	9498354
Marmoreio	Carne e carcaça	5897389:9662643	2755	9498354
Espessura de gordura na 12 <sup>a</sup> costela	Carne e carcaça	5897389:9662643	2756	9498354
Percentual de gordura nos rins, pélvica e coração	Carne e carcaça	5897389:9662643	2757	9498354
Relação de ácido oleico/esteárico	Carne e carcaça	5897389:9662643	4862	17894565
Teor de ácidos graxos monoinsaturados	Carne e carcaça	5897389:9662643	4859	17894565
Área de olho de lombo	Carne e carcaça	2134354:9695621	4492	17596127

## Conclusão

Os resultados apresentados apontam que o uso da ferramenta genômica foi eficiente na avaliação e análise de milhares de SNP presentes no DNA de novilhas Nelore com fenótipo de idade ao primeiro parto, obtendo alta densidade de genótipos acurados. Sendo possível a verificação regiões de QTLs relacionados à características reprodutivas. O tamanho amostral possivelmente foi um fator limitante para encontrar associação entre SNP e característica de IPP.

**Agradecimentos** À FUNDECT - Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul, ao CNPq - Conselho Nacional de Desenvolvimento Científico pelo apoio financeiro e bolsa de iniciação científica concedida e à UFGD pelo apoio logístico. Ao aluno de doutorado Yuri Tani Utsunomiya (FCAV-Unesp) por ceder os *scripts* customizado utilizados para as análises e a Dr<sup>a</sup>. Marcia Cristina Matos ao apoiar bioinformático.

## Referências

- ABIEC - Associação Brasileira de Indústrias Exportadoras de Carne (2014) <http://www.abiec.com.br>. Accessed 25 Jan. 2014
- ACNB - Associação de Criadores de Nelore do Brasil (2014) <http://www.nelore.org.br/>. Accessed 25 Jan. 2014
- Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT (2010) Data quality control in genetic case-control association studies. *Nature protocols* 5 (9):1564-1573
- Ashwell MS, Heyen DW, Weller JI, Ron M, Sonstegard TS, Van Tassell CP, Lewin HA (2005) Detection of quantitative trait loci influencing conformation traits and calving ease in Holstein-Friesian cattle. *Journal of dairy science* 88 (11):4111-4119
- Aulchenko YS (2011) Abel Tutorial.243
- Aulchenko YS, de Koning DJ, Haley C (2007) Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics* 177 (1):577-585

- Baker SM, Bronner CE, Zhang L, Plug AW, Robatzek M, Warren G, Elliott EA, Yu J, Ashley T, Arnheim N, Flavell RA, Liskay RM (1995) Male mice defective in the DNA mismatch repair gene PMS2 exhibit abnormal chromosome synapsis in meiosis. *Cell* 82 (2):309-319
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*:289-300
- Berger JH, Charron MJ, Silver DL (2012) Major facilitator superfamily domain-containing protein 2a (MFSD2A) has roles in body growth, motor function, and lipid metabolism. *PLoS One* 7 (11):e50629
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics* 32 (3):314
- Bovine HapMap Consortium, Gibbs RA, Taylor JF, Van Tassell CP, Barendse W et al (2009) Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* 324 (5926):528-532
- CattleQTLdb (2013) Release 22. <http://www.animalgenome.org/cgi-bin/QTLdb/BT/index>. Accessed 14 Feb. 2014
- Dawuda PM, Eduvie LO, Esievo KA, Molokwu EC (1988) Interval between calving and first observable oestrus in post-partum Bunaji cows. *The British veterinary journal* 144 (3):258-261
- Dias LT, El Faro L, Albuquerque LGd (2004) Estimativas de herdabilidade para idade ao primeiro parto de novilhas da raça Nelore. *Revista Brasileira de Zootecnia* 33 (1):97-102
- Dobson H, Kamonpatana M (1986) A review of female cattle reproduction with special reference to a comparison between buffaloes, cows and zebu. *Journal of reproduction and fertility* 77 (1):1-36
- Finlay EK, Berry DP, Wickham B, Gormley EP, Bradley DG (2012) A genome wide association scan of bovine tuberculosis susceptibility in Holstein-Friesian dairy cattle. *PLoS One* 7 (2):e30545
- Frommlet F, Ruhaltinger F, Twaróg P, Bogdan M (2012) Modified versions of Bayesian Information Criterion for genome-wide association studies. *Computational Statistics & Data Analysis* 56 (5):1038-1051



- Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. *Nature reviews Genetics* 6 (2):95-108
- Hu ZL, Park CA, Wu XL, Reecy JM (2013) Animal QTLdb: an improved database tool for livestock animal QTL/association data dissemination in the post-genome era. *Nucleic acids research* 41:871-879
- Illumina (2012) BovineHD Genotyping BeadChip. [www.illumina.com](http://www.illumina.com). Accessed 26 Jan. 2014
- Illumina (2014) Product literature. <http://www.illumina.com/>. Accessed 26 Jan. 2014
- Jiang L, Jiang J, Yang J, Liu X, Wang J et al (2013) Genome-wide detection of copy number variations using high-density SNP genotyping platforms in Holsteins. *BMC genomics* 14:131
- Jiménez E, Arias B, Mitjans M, Goikolea JM, Roda E, Sáiz PA, García-Portilla MP, Burón P, Bobes J, Oquendo MA, Vieta E, Benabarre A (2013) Genetic variability at IMPA2, INPP1 and GSK3 $\beta$  increases the risk of suicidal behavior in bipolar patients. *European Neuropsychopharmacology* 23 (11):1452-1462
- Kinsella RJ, Kahari A, Haider S, Zamora J, Proctor G et al (2011) Ensembl BioMart: a hub for data retrieval across taxonomic space. *Database : the journal of biological databases and curation* 2011:30
- Matos MC (2012) Associação genômica ampla para características reprodutivas em bovinos da raça nelore. Universidade Estadual Paulista - UNESP, Joticabal
- Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF et al (2009) Development and characterization of a high density SNP genotyping assay for cattle. *PloS One* 4 (4):5350
- Moskvina V, Schmidt KM (2008) On multiple-testing correction in genome-wide association studies. *Genetic epidemiology* 32 (6):567-573
- Nishimura S, Watanabe T, Mizoshita K, Tatsuda K, Fujita T et al (2012) Genome-wide association study identified three major QTL for carcass weight including the PLAG1-CHCHD7 QTN for stature in Japanese Black cattle. *BMC genetics* 13:40
- Nogueira GP (2004) Puberty in South American *Bos indicus* (Zebu) cattle. *Animal reproduction science* 82-83:361-372
- Pereira E, Eler J, Costa F, Ferraz J (2001) Análise genética da idade ao primeiro parto e do perímetro escrotal em bovinos da raça Nelore. *Arq bras med vet zootec* 53 (1):116-121
- Pereira E, Eler JP, Ferraz JBS (2002) Análise genética de características reprodutivas na raça Nelore. *Pesquisa Agropecuária Brasileira* 37 (5):703-708

- Pirlo G, Miglior F, Speroni M (2000) Effect of age at first calving on production traits and on difference between milk yield returns and rearing costs in Italian Holsteins. *Journal of dairy science* 83 (3):603-608
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* 81 (3):559-575
- Resende MDVd, Aguiar AM, Abad JIM, Missiaggia AA, Sansaloni C et al (2010) Computação da seleção genômica ampla (GWS).
- Santana M, Eler J, Ferraz J, Mattos E (2012) Genetic relationship between growth and reproductive traits in Nellore cattle. *Animal : an international journal of animal bioscience* 6 (04):565-570
- Sasaki S, Ibi T, Ikeda S, Sugimoto Y (2013) A genome-wide association study reveals a quantitative trait locus for age at first calving in delta/notch-like EGF repeat containing on chromosome 2 in Japanese Black cattle. *Animal genetics*:1-3
- Schulman NF, Sahana G, Lund MS, Viitala SM, Vilkki JH (2008) Quantitative trait loci for fertility traits in Finnish Ayrshire cattle. *Genetics, selection, evolution : GSE* 40 (2):195-214
- Shin D-H, Cho K-H, Park K-D, Lee H-J, Kim H (2013) Accurate Estimation of Effective Population Size in the Korean Dairy Cattle Based on Linkage Disequilibrium Corrected by Genomic Relationship Matrix. *Asian-Australasian Journal of Animal Sciences* 26 (12):1672-1679
- Singh SP, Kumar R, Kumari P, Kumar S, Mitra A (2014) Characterization of 5' upstream region and investigation of TTTTA deletion in 5' UTR of myostatin (MSTN) gene in Indian goat breeds. *Animal biotechnology* 25 (1):55-68
- Sun YV (2012) Integration of biological networks and pathways with genetic association studies. *Human genetics* 131 (10):1677-1686
- Terasawa E, Fernandez DL (2001) Neurobiological mechanisms of the onset of puberty in primates. *Endocrine reviews* 22 (1):111-151
- Turner S, Armstrong LL, Bradford Y, Carlson CS, Crawford DC et al (2011) Quality control procedures for genome-wide association studies. *Current protocols in human genetics* 1:19
- Utsunomiya YT, Carmo AS, Neves HH, Carvalheiro R, Matos MC et al (2014) Genome-wide mapping of Loci explaining variance in scrotal circumference in nellore cattle. *PLoS One* 9 (2):e88561

- Utsunomiya YT, do Carmo AS, Carvalheiro R, Neves HH, Matos MC et al (2013) Genome-wide association study for birth weight in Nellore cattle points to previously described orthologous genes affecting human and bovine height. *BMC genetics* 14:52
- Vaiciunas A, Coutinho LL, Meirelles FV, Pires AV, Silva LFP (2008) Leptin and hypothalamic gene expression in early- and late-maturing *Bos indicus* Nellore heifers. *Genetics and Molecular Biology* 31 (3):657-664
- Valle ER, Andreotti R, Thiago LRLdS (2000) Técnicas de manejo reprodutivo em bovinos de corte. Embrapa Gado de Corte, Campo Grande
- Vargas CA, Elzo MA, Chase CC, Jr., Chenoweth PJ, Olson TA (1998) Estimation of genetic parameters for scrotal circumference, age at puberty in heifers, and hip height in Brahman cattle. *Journal of animal science* 76 (10):2536-2541
- Wiltbank J, Kasson C, Ingalls J (1969) Puberty in crossbred and straightbred beef heifers on two levels of feed. *Journal of animal science* 29 (4):602-605

## **Anexos**

### Normas da revista

#### **Molecular Biology Reports**

An International Journal on Molecular and Cellular Biology

Editor-in-Chief: Nilanjana Maulik

ISSN: 0301-4851 (print version)

ISSN: 1573-4978 (electronic version)

Journal no. 11033

- **Manuscript Submission**

Submission of a manuscript implies: that the work described has not been published before; that it is not under consideration for publication anywhere else; that its publication has been approved by all co-authors, if any, as well as by the responsible authorities – tacitly or explicitly – at the institute where the work has been carried out. The publisher will not be held legally responsible should there be any claims for compensation.

- **Title Page**

The title page should include:

1. The name(s) of the author(s)
2. A concise and informative title
3. The affiliation(s) and address(es) of the author(s)
4. The e-mail address, telephone and fax numbers of the corresponding author

- **Abstract**

Please provide an abstract of 150 to 250 words. The abstract should not contain any undefined abbreviations or unspecified references.

#### **Keywords**

Please provide 4 to 6 keywords which can be used for indexing purposes.

- **Text Formatting**

Manuscripts should be submitted in Word.

1. Use a normal, plain font (e.g., 10-point Times Roman) for text.
2. Use italics for emphasis.

3. Use the automatic page numbering function to number the pages.
4. Do not use field functions.
5. Use tab stops or other commands for indents, not the space bar.
6. Use the table function, not spreadsheets, to make tables.
7. Use the equation editor or MathType for equations.
8. Save your file in docx format (Word 2007 or higher) or doc format (older Word versions).

Manuscripts with mathematical content can also be submitted in LaTeX.

- Headings

Please use no more than three levels of displayed headings.

- Abbreviations

Abbreviations should be defined at first mention and used consistently thereafter.

- Citation

Reference citations in the text should be identified by numbers in square brackets. Some examples:

1. Negotiation research spans many disciplines [3].
2. This result was later contradicted by Becker and Seligman [5].
3. This effect has been widely studied [1-3, 7].

- Reference list

The list of references should only include works that are cited in the text and that have been published or accepted for publication. Personal communications and unpublished works should only be mentioned in the text. Do not use footnotes or endnotes as a substitute for a reference list.

The entries in the list should be numbered consecutively.

1. Journal article

Gamelin FX, Baquet G, Berthoin S, Thevenet D, Nourry C, Nottin S, Bosquet L (2009) Effect of high intensity intermittent training on heart rate variability in prepubescent children. *Eur J Appl Physiol* 105:731-738. doi: 10.1007/s00421-008-0955-8

Ideally, the names of all authors should be provided, but the usage of “et al” in long author lists will also be accepted:

Smith J, Jones M Jr, Houghton L et al (1999) Future of health insurance. *N Engl J Med* 341:325–329

2. Article by DOI

Slifka MK, Whitton JL (2000) Clinical implications of dysregulated cytokine production. *J Mol Med*. doi:10.1007/s001090000086

3. Book

South J, Blass B (2001) *The future of modern genomics*. Blackwell, London

4. Book chapter

Brown B, Aaron M (2001) The politics of nature. In: Smith J (ed) *The rise of modern genomics*, 3rd edn. Wiley, New York, pp 230-257

5. Online document

Cartwright J (2007) Big stars have weather too. IOP Publishing PhysicsWeb. <http://physicsweb.org/articles/news/11/6/16/1>. Accessed 26 June 2007

6. Dissertation

Trent JW (1975) *Experimental acute renal failure*. Dissertation, University of California

For authors using EndNote, Springer provides an output style that supports the formatting of in-text citations and reference list.

Authors preparing their manuscript in LaTeX can use the bibtex file `spbasic.bst` which is included in Springer’s LaTeX macro package.

- Footnotes

Footnotes can be used to give additional information, which may include the citation of a reference included in the reference list. They should not consist solely of a reference citation, and they should never include the bibliographic details of a reference. They should also not contain any figures or tables.

Footnotes to the text are numbered consecutively; those to tables should be indicated by superscript lower-case letters (or asterisks for significance values and other statistical data). Footnotes to the title or the authors of the article are not given reference symbols.

Always use footnotes instead of endnotes.

- Acknowledgments

Acknowledgments of people, grants, funds, etc. should be placed in a separate section before the reference list. The names of funding organizations should be written in full.

- Tables

1. All tables are to be numbered using Arabic numerals.
2. Tables should always be cited in text in consecutive numerical order.
3. For each table, please supply a table caption (title) explaining the components of the table.
4. Identify any previously published material by giving the original source in the form of a reference at the end of the table caption.
5. Footnotes to tables should be indicated by superscript lower-case letters (or asterisks for significance values and other statistical data) and included beneath the table body.

- Artwork and Illustrations Guidelines

For the best quality final product, it is highly recommended that you submit all of your artwork – photographs, line drawings, etc. – in an electronic format. Your art will then be produced to the highest standards with the greatest accuracy to detail. The published work will directly reflect the quality of the artwork provided.

- Electronic Figure Submission

1. Supply all figures electronically.
2. Indicate what graphics program was used to create the artwork.
3. For vector graphics, the preferred format is EPS; for halftones, please use TIFF format. MS Office files are also acceptable.
4. Vector graphics containing fonts must have the fonts embedded in the files.
5. Name your figure files with "Fig" and the figure number, e.g., Fig1.eps.

- Line Art

1. Definition: Black and white graphic with no shading.
2. Do not use faint lines and/or lettering and check that all lines and lettering within the figures are legible at final size.
3. All lines should be at least 0.1 mm (0.3 pt) wide.
4. Scanned line drawings and line drawings in bitmap format should have a minimum resolution of 1200 dpi.
5. Vector graphics containing fonts must have the fonts embedded in the files.

- Halftone Art

1. Definition: Photographs, drawings, or paintings with fine shading, etc.
2. If any magnification is used in the photographs, indicate this by using scale bars within the figures themselves.
3. Halftones should have a minimum resolution of 300 dpi.

- Combination Art

1. Definition: a combination of halftone and line art, e.g., halftones containing line drawing, extensive lettering, color diagrams, etc.
2. Combination artwork should have a minimum resolution of 600 dpi.

- Color Art

1. Color art is free of charge for online publication.
2. If black and white will be shown in the print version, make sure that the main information will still be visible. Many colors are not distinguishable from one another when converted to black and white. A simple way to check this is to make a xerographic copy to see if the necessary distinctions between the different colors are still apparent.
3. If the figures will be printed in black and white, do not refer to color in the captions.
4. Color illustrations should be submitted as RGB (8 bits per channel).

- Figure Lettering

1. To add lettering, it is best to use Helvetica or Arial (sans serif fonts).
2. Keep lettering consistently sized throughout your final-sized artwork, usually about 2–3 mm (8–12 pt).
3. Variance of type size within an illustration should be minimal, e.g., do not use 8-pt type on an axis and 20-pt type for the axis label.
4. Avoid effects such as shading, outline letters, etc.
5. Do not include titles or captions within your illustrations.

- Figure Numbering

1. All figures are to be numbered using Arabic numerals.
2. Figures should always be cited in text in consecutive numerical order.
3. Figure parts should be denoted by lowercase letters (a, b, c, etc.).
4. If an appendix appears in your article and it contains one or more figures, continue the consecutive numbering of the main text. Do not number the appendix figures, "A1, A2, A3, etc." Figures in online appendices (Electronic Supplementary Material) should, however, be numbered separately.

- Figure Captions

1. Each figure should have a concise caption describing accurately what the figure depicts. Include the captions in the text file of the manuscript, not in the figure file.
2. Figure captions begin with the term Fig. in bold type, followed by the figure number, also in bold type.
3. No punctuation is to be included after the number, nor is any punctuation to be placed at the end of the caption.
4. Identify all elements found in the figure in the figure caption; and use boxes, circles, etc., as coordinate points in graphs.
5. Identify previously published material by giving the original source in the form of a reference citation at the end of the figure caption.

- **Figure Placement and Size**

1. When preparing your figures, size figures to fit in the column width.
2. For most journals the figures should be 39 mm, 84 mm, 129 mm, or 174 mm wide and not higher than 234 mm.
3. For books and book-sized journals, the figures should be 80 mm or 122 mm wide and not higher than 198 mm.

- **Permissions**

If you include figures that have already been published elsewhere, you must obtain permission from the copyright owner(s) for both the print and online format. Please be aware that some publishers do not grant electronic rights for free and that Springer will not be able to refund any costs that may have occurred to receive these permissions. In such cases, material from other sources should be used.

- **Accessibility**

In order to give people of all abilities and disabilities access to the content of your figures, please make sure that

1. All figures have descriptive captions (blind users could then use a text-to-speech software or a text-to-Braille hardware)
2. Patterns are used instead of or in addition to colors for conveying information (color-blind users would then be able to distinguish the visual elements)
3. Any figure lettering has a contrast ratio of at least 4.5:1

- **Electronic Supplementary Material**

Springer accepts electronic multimedia files (animations, movies, audio, etc.) and other supplementary files to be published online along with an article or a book chapter. This feature can add dimension to the author's article, as certain information cannot be printed or is more convenient in electronic form.

- **Submission**

1. Supply all supplementary material in standard file formats.
2. Please include in each file the following information: article title, journal name, author names; affiliation and e-mail address of the corresponding author.
3. To accommodate user downloads, please keep in mind that larger-sized files may require very long download times and that some users may experience other problems during downloading.

- **Text and Presentations**

1. Submit your material in PDF format; .doc or .ppt files are not suitable for long-term viability.
2. A collection of figures may also be combined in a PDF file.

- **Spreadsheets**

1. Spreadsheets should be converted to PDF if no interaction with the data is intended.
2. If the readers should be encouraged to make their own calculations, spreadsheets should be submitted as .xls files (MS Excel).

- **Specialized Formats**

1. Specialized format such as .pdb (chemical), .wrl (VRML), .nb (Mathematica notebook), and .tex can also be supplied.

- **Collecting Multiple Files**

1. It is possible to collect multiple files in a .zip or .gz file.

#### Numbering

1. If supplying any supplementary material, the text must make specific mention of the material as a citation, similar to that of figures and tables.
2. Refer to the supplementary files as "Online Resource", e.g., "... as shown in the animation (Online Resource 3)", "... additional data are given in Online Resource 4".
3. Name the files consecutively, e.g. "ESM\_3.mpg", "ESM\_4.pdf".

- **Captions**

1. For each supplementary material, please supply a concise caption describing the content of the file.

- Processing of supplementary files

1. Electronic supplementary material will be published as received from the author without any conversion, editing, or reformatting.

- Accessibility

In order to give people of all abilities and disabilities access to the content of your supplementary files, please make sure that

1. The manuscript contains a descriptive caption for each supplementary material
2. Video files do not contain anything that flashes more than three times per second (so that users prone to seizures caused by such effects are not put at risk)